STUDY MATERIAL

FOUNDATION PROGRAMME

BUSINESS ECONOMICS

PAPER 3

TIMING OF HEADQUARTERS

Monday to Friday

Office Timings – 9.00 A.M. to 5.30 P.M.

Public Dealing Timings

Without financial transactions – 9.30 A.M. to 5.00 P.M.

With financial transactions – 9.30 A.M. to 4.00 P.M.

Phones

41504444, 45341000

Fax

011-24626727

Website

www.icsi.edu

E-mail

info@icsi.edu

FOUNDATION PROGRAMME – IMPORTANT NOTE

The study material has been written in lucid and simple language and conscious efforts have been made to explain the fundamental concepts of economics and to provide exposure to Indian economy and Statistical tools. This study material is divided into two main parts–

Part-A   Economics, and

Part-B   Elementary Statistics

There is computer based examinations for Foundation Programme of CS Course. Students are advised to go through instructions regarding computer based examinations available at ICSI website **www.icsi.edu.**

For supplementing the information contained in the study material, students may refer to the economic and financial dailies, commercial, legal and management journals, Economic Survey (latest), CS Foundation Course e-Bulletin, Suggested Readings and References mentioned in the study material and relevant websites.

The objective of the study material is to provide students with the learning material according to the syllabus of the subject of the Foundation Programme. In the event of any doubt, students may write to the Directorate of Studies in the Institute for clarification at **academics@icsi.edu**

Although due care has been taken in preparing and publishing this study material, yet the possibility of errors, omissions and/or discrepancies cannot be ruled out. This publication is released with an understanding that the Institute shall not be responsible for any errors, omissions and/or discrepancies or any action taken on the basis of contents of the study material.

Should there be any discrepancy, error or omission noted in the study material, the Institute shall be obliged if the same are brought to its notice for issue of corrigendum in the CS Foundation Course e-Bulletin.

SYLLABUS

PAPER 3: BUSINESS ECONOMICS

Level of Knowledge: *Basic Knowledge*

Objective: *To familiarize the basic concepts and theories of economics, elementary statistics and mathematics.*

PART A: ECONOMICS (80 MARKS)

1. **The Fundamentals of Economics**

   The Economic Problem-Scarcity and Choice; Nature and Scope- Positive and Normative Economics, Micro and Macro Economics; Central Problems of an Economy; Production Possibility Curve; Opportunity Cost; Working of Economic Systems; Economic Cycles

2. **Basic Elements of Demand and Supply**

   Demand- Meaning, Demand Schedule, Individual and Market Demand Curve, Determinants of Demand, Law of Demand, Changes in Demand; Supply- Meaning, Supply Schedule, Individual and Market Supply Curve, Determinants of Supply, Law of Supply, Changes in Supply; Equilibrium of Demand and Supply- Determination of Equilibrium Price and Quantity, Effect of a shift in Demand or Supply; Elasticity of Demand and Supply

3. **Theory of Consumer Behaviour**

   Cardinal Utility Approach-Law of Diminishing Marginal Utility, Law of Equi-Marginal Utility; Indifference Curve Approach- Indifference Curves, Properties of Indifference Curves, Budget Line, Consumer's Equilibrium

4. **Theory of Production and Costs**

   Theory of Production- Factors of Production, Basic Concepts, Production Function, Law of Variable Proportions, Returns to Scale; Producer's Equilibrium- Least-Cost Factor Combination and Output Maximisation for a given Level of Outlay; Theory of Costs- Basic Concepts, Short-run Total Cost Curves-Fixed and Variable, Short-run Average and Marginal Cost Curves, Relationship between Average and Marginal Cost Curve, Average and Marginal Cost Curves in the Long-run

5. **Analysis of Markets**

   Basic Concepts of Revenue, Revenue Curves, Relationship between Average and Marginal Revenue Curve; Concept of Market and Main Forms of Market; Equilibrium of the Firm- Meaning, Objectives of the Firm, Total Revenue-Total Cost Approach, Marginal Revenue-Marginal Cost Approach; Price and Output under Determination Perfect Competition, Monopoly, Monopolistic Competition and Oligopoly.

6. **Indian Economy- An Overview**

   Basic Characteristics of the Indian Economy; Major Issues of Development; Development Experience and Recent Trends in Indian Economy; Indian Economy in Comparison to Major Economies of the World

7. **Basic Elements of Money and Banking**

   Concept of Money-Its Functions, Quantity Theory of Money, Credit Creation; Central Bank (Reserve Bank of India)-Role and Functions; Commercial Banks-Role and Functions; Basic Elements of E-Banking; Monetary Policy in India

PART B: ELEMENTARY STATISTICS (20 MARKS)

8. Descriptive Statistics

Statistics- Definition, Functions, Scope, Application in Business, Law of Statistics, Limitations of Statistics; Collection and Presentation of Statistical Data-Primary and Secondary Data, Classification and Tabulation, Frequency Distribution, Cross Tabulation; Diagrams and Graphs; Measures of Central Tendency-Mean, Median, Mode; Measures of Dispersion-Mean Deviation, Standard Deviation, Range, Coefficient of Variation; Bi-variate Analysis-Covariance, Coefficient of Correlation.

9. Mathematics of Finance and Elementary Probability

Mathematics of Finance-Simple Interest, Compound Interest; Time Value of Money- Compounding & Discounting, Present Value & Future Value of an Annuity; Probability- Random Experiments, Sample Spaces, Events and Probability, Approaches to Probability- Classical & Empirical; Expected Value.

LIST OF RECOMMENDED BOOKS*

PAPER 3 : BUSINESS ECONOMICS

READINGS

1.	H. L. Bhatia	Micro Economic Theory; Modern Publisher, Gulab Bhawan, 6, Bahadurshah Zafar Marg, New Delhi.
2.	M.L. Jhingan	Micro Economic Theory; Konark Publishers Pvt. Ltd., A-149, Vikas Marg, Shakarpur, New Delhi-110 092.
3.	S. K. Agarwala	Economic System and Micro Economic Theory; Galgotia Publishing Company, 6A/4, WEA, Karol Bagh, New Delhi.
4.	D.D. Chaturvedi	Macro Economic - Analysis and Policy; Galgotia Publishing Company, 6A/4, WEA, Karol Bagh, New Delhi.
5.	D.M. Mithani	Macro Economics; Himalaya Publishing House.
6.	D.M. Mithani	Money Banking, International Trade and Public Finance; Himalaya Publishing House, Ramdoot, Dr. Bhalerao Marg, Girgaon, Bombay-400 004.
7.	Ruddar Datt & K.P.M. Sundaram	Indian Economy; S. Chand & Company Ltd., Ram Nagar, New Delhi.
8.	S. P. Gupta, P. K. Gupta & Man Mohan	Quantitative Techniques; Sultan Chand & Sons, 23, Daryaganj, New Delhi.
9.	R. S. Bhardwaj	Business Statistics, Excle Books, A-45, Narayana Phase-I, New Delhi
10.	A. K. Agarwal & Sahib Singh	Fundamental of Statistics; Frank Sons, Daryaganj, New Delhi.
11.	R. S. N. Pillai & V. Bagarathi	Statistics; S. Chand & Company Ltd. New Delhi.
12.	S. Saha	Business Statistics, New Central Book Agency, Kolkata
13.	B. M. Agarwal	Quantitative Techniques, Sultan Chand & Songs Pvt. Ltd., Daryaganj, New Delhi.
14.	Pindyck and Rubinfeld	Microeconomics; Pearson Education.
15.	Suraj B. Gupta	Monetary Economics; Sultan Chand & Sons Pvt. Ltd.

REFERENCES

1.	Samuelson & William D. Nordhans	Economics; McGraw Hills.
2.	P. N. Chopra	Advanced Economics Theory; Kalyani Publishers, New Delhi.
3.	M. Maria John Kennedy	Advanced Micro Economic Theory; Himalaya Publishing House, Delhi.

---

*This study material is sufficient from the point of view of syllabus. The students may refer these books for further knowledge and study of the subject.

4. A. C. L. Day   Outline of Monetary Economics

5. A. N. Agarwal   Indian Economy

6. I. C. Dhingra &   Economic Development & Planning in India.
    V. K. Garg

7. K.R. Choudhary   Basic Principles of Modern Micro Economics; Book Land Economic Series.

8. Mural R. Spiegal &   Theory and Problems of Statistics, Tata McGraw Hills, New Delhi
    Larry J. Stephans

9. S. P. Gupta   Statistical Methods; Sultan Chand & Sons, New Delhi.

10. Neva Goodwin,   Macro Economics in Context, PHI Learning Pvt. Ltd.
    Julie A Nelson,
    Jonathan Harris

11. Eleventh Five Year Plan Government of India Publication.
    (2007-2012)

12. Twelfth Five Year Plan Government of India Publication.
    (2012-2017)

13. Economic Survey   Government of India Publication.
    (Latest)

CONTENTS

PART A: ECONOMICS

LESSON 1

THE FUNDAMENTALS OF ECONOMICS

LESSON 2

BASIC ELEMENTS OF DEMAND AND SUPPLY

LESSON 3

THEORY OF CONSUMER BEHAVIOR

LESSON 4

THEORY OF PRODUCTION AND COST

LESSON 5

ANALYSIS OF MARKETS

LESSON 6

INDIAN ECONOMY – AN OVERVIEW

LESSON 7

BASIC ELEMENTS OF MONEY AND BANKING

PART B – ELEMENTARY STATISTICS

LESSON 8
DESCRIPTIVE STATISTICS

LESSON 9

MATHEMATICS OF FINANCE AND ELEMENTARY PROBABILITY

PART A
ECONOMICS

LESSONS

Economics has been recognized as a special area of study for over a century. Economics and economists are words that almost everyone has heard of and used. But, what exactly is economics? Very few people can give a good definition or description of what this field of study is all about. If ordinary citizens cannot give a good definition or description of economics, they can be excused because even economists struggled long to define their own field.

We have to perform various and varied activities throughout the day. After attaining adulthood, we have our own family and we are required to make arrangement for food, clothes, house and other necessaries of life for the members of our family. We have to activate ourselves to earn something, so that we may be able to meet the expenses. Our activities to generate income are termed as economic activities, which are responsible for the origin and development of economics as subject. Economics was originally introduced as a science of statecraft. It was concerned with the collection of revenue for the state i.e., government. Therefore, it becomes essential to acquaint students with basics of this subject so as to develop conceptual understanding.

Lesson 1
The Fundamentals of Economics

LEARNING OBJECTIVES

It is important that students get familiar with the basic purpose of studying economics. The basic concepts of economics have indispensable role in business and other activities of daily life. This lesson aims at providing answers to all the basic questions in economics before moving onto other lessons in this study material.

The Learning objectives of this Chapter are following:

(a) To get acquainted with the basic purpose of the study of Economics.

(b) To understand the problems which are being addressed in Economics.

(c) To establish the duality of scarcity and choice in every decision-making process.

(d) To be able to differentiate between Micro Economics and Macro Economics, Positive and Normative Economics.

(e) To go through formulation of central questions of the Economy.

WHY STUDY ECONOMICS?

Before getting ready to dwell upon the subject matter of Economics, its meaning, nature and scope a simple question needs to be answered – 'Why to study Economics?'. Let's approach this question with another question by asking- 'What are the things which will be missed out if one chooses not to study economics?' Until studying economics not many persons are aware of how industries, businesses and governments work in the interest of public. It is a natural scene in and around neighborhood and offices when you find people talking about price rise, rise in demand of a particular commodity, tax structure or tax exemption for a particular industry, without finding comprehensive explanation. Quite often we come across news briefings upon advantages and disadvantages of a certain tax structure, for example; Goods and Services Tax, or Pricing of 2G/3G/4G telecommunication spectrum, GDP growth rate and several others. Study of Economics provides a systematic understanding of economic events which are very significant for day to day functioning of a family, a company, a country or world at large.

While trying to answer this question, students can do one experiment in their neighborhood by asking basic questions to their neighbors as to – a) why does govt. impose taxes upon people and goods? b) why do prices for goods increase ever year? c) why $1 is exchanged with more than Sixty Indian Rupees? d) Why do American families are richer than Sri Lankan families? e) why Facebook is not charging us for its services? f) Why the call charges to New York are cheaper than call charges to Thailand? Make a table of observations comprising three columns namely, name, comfortable/ uncomfortable while answering question and have studied/not-studied economics. You would find that, irrespective of their profession, those who have studied economics are more comfortable answering your questions. In conclusion, the study of economics gives you wholesome understanding about events and activities happening around you.

A restaurant-bill provides us a lot of information about restaurant business in a locality. To interpret that information one must have a basic understanding about economics. The first visible information on the bill is the rate of tax charged on eatables from a restaurant. If due attention is paid, one must come across reason behind the specific prices charged for dishes listed in the bill. Prices are basically charged upon the cost of purchasing and processing for each item separately. Dairy and Bakery products are relatively less-costlier than poultry products, such are their costs in a restaurant. The cost involved in processing of an ice-cream is lesser than meat so the meat served in a restaurant is priced higher than the market price of the raw meat while the price of ice-cream served there will only be little bit higher than the price of ice-cream in market. This processing charge is the labour cost or wage payment for the workers in that restaurant. Labour wages in famous restaurants are higher than other restaurants and so are their processing charges. There can be several other insights from the bill like rents for the place occupied by restaurant, supply of unprocessed food materials, fresh vegetables, meat, dairy products etc. and others.

It shows that with or without understanding of economics we take part in various economic activities as an agent on a regular basis. The study of economics gives us an advantage to understand the dynamics behind economic activities and ability to interpret them. This sort of interpretation helps in understanding business for all possible changes, manipulation and further advancement in future. For students and professionals seeking career in business or doing business in any field of specialization, knowledge of economics becomes inevitable for realizing their acumen into some enterprising assets. In general, there are following reasons for studying economics;

a) Economics for citizenship – Solution to the most of the problems we confront in day to day life have political aspect. And majority of political problems are resolved through economic strategies and planning. Be it the issue of international trade agreement, tax restructuring or launching welfare program. The basic understanding of economics is beneficial for both – voters as well as elected officials.

b) Achieving Social Change – As a sensitive young student if one wants to make his locality, country and

this world a better place there is no dearth of social problems that required our attention. Be it poverty, inequality, hunger, crime, unemployment, terrorism, illiteracy, malnutrition or any other social problems, economics can certainly help us understand the origins of these problems, explain why previous efforts to solve them haven't succeeded, and help us to design new, more effective solutions.

c) Personal and Professional Development – Economics has been one of the most popular subject choice in leading colleges and universities in India at bachelor degree level for students intending to work in business. It has also been popular among students planning for career in law, politics, international law, industrial engineering and other profession only because it can help them in understanding their field better and may affect their business.

d) Tool for Rational Thinking – Besides discussing concepts, it helps in generating ideas. Economics has great deal of dependency over mathematics and statistics in pursuit of finding objective explanation for the problem and complex solutions to those problems. In this process, it adopts and sometimes develops some useful tools which are proven to be very helpful in rational thinking about issues. For example, the development of 'Human Development Index' or HDI which is a comprehensive tool for mapping not only economic but social and educational development of the country or region.

Studying economics is often more useful to understand social, political, technological, legal and environmental activities as the dynamics of scarcity and want plays out in decision making scenarios of these domains of knowledge. The detailed analysis on scarcity and want is presented in later part of the chapter. After discussing, 'why to study Economics?', the immediate step would be to discuss – 'what is economics?'.

THE ECONOMIC PROBLEM- SCARCITY AND CHOICE?

The word 'Economics' is derived from the Greek word '*Oikonomia*' which can be divided into two parts : (*a*) '*Oikos*', which means 'House', and (*b*) '*Nomos',* which means 'Management'. Together they mean managing household. The coverage of economics is broader than the meaning of the word 'Economics'.

Could you make an attempt to ask yourself – Is doing this Course the right choice for you? The expected outcome of doing the course may or may not promise what you want to do in life but, given the time and resources in possession, you found this course as a best suitable choice among all the available alternatives. One might have used this time in earning additional money by engaging oneself in some gainful employment rather than spending those times in doing this course. The purpose of investing some valuable time in pursuing this course may be to earn more money in future. In view of future benefits the cost incurred in taking this course appears worth investing which gives you a good enough reason to choose this course. Now again make an attempt to ask yourself – Is doing this Course the right choice for you? One may not be sure whether a decision taken will prove to be right or wrong eventually but, if it shows the benefits of decisions are higher than the cost of taking those decisions then the decision certainly is an economic decision. At this stage, it can be said that;

Economics is the study of choice under conditions of scarcity.

This definition has two parts; a) Choice and b) Scarcity. Scarcity is a situation in which the amount of something available is insufficient to satisfy the desire for it. As an individual, for instance, we face scarcity of time and spending power. Given more of either, we could have more of the good and services that we desire. At any given point of time, individuals face numerous scarcities. There are so many things one might like to have - sedan cars, branded clothes, stylish shoes, branded watches, sports equipments, studio apartments, independent bungalow, and so on…. i.e. individuals have unlimited wants but face the constraint of limited spending power.

Every individual confronts this kind of situation on day to day basis. Because of the scarcities of time and spending power, each of us is forced to make choices. We must allocate our scarce time to different activities: work, play, study, sleep, shopping and more. We must allocate our scarce spending power among different

goods and services; housing, food, furniture, travel, and many others. And each time we choose to buy something or do something, we also choose not to buy or do something else.

A little reflection suggests that just like individuals, at any given point of time, a society also faces unlimited wants and limited resources. The limited resources have alternative uses to which they can be put to use. Out of these alternative uses of scarce resources and unlimited wants, society has to make choice. The problem of scarcity and choice forms the core of Economics.

Given this, Economics would mean the study of ways in which mankind organizes itself to tackle the basic problem of scarcity of resources. Hence, economics is the study of alternate systems requisite to allocate these resources between competing ends. In view of allocation of resources, it can be said that;

Economics is the study of how we choose to use limited resources with alternative uses to obtain the maximum satisfaction of unlimited human wants.

HOW TO DEFINE ECONOMICS?

Steven D. Levitt, an award winning economist, said that "Since the science of economics is primarily a set of tools, as opposed to a subject matter, then no subject, however offbeat, need be beyond its reach." Steven Levitt had pointed in 'Freakonomics' that economics is primarily a set of tools before being a subject and the range of issues economics deals with is vast enough (ranging from individual consumer choices to buy a burger or not to buy a burger ….to enumerating causes of global financial crisis) to be captured in one definition. On this basis, it can be stated that it is difficult to have one universally acceptable definition of economics. However, systematic attempts were made in the past ever since the subject became a discipline of study.

There are four such notable definitions;

a) **The Wealth Definition of Economics**

Adam Smith (1723-1790) defined Economics as- Science of wealth with an objective to increase wealth and richness of a country.

Key Features of Wealth Definition:

  i) The objective of Economics is to increase the wealth of a country.

 ii) It considers production, distribution and consumption as core of economic activity.

 iii) It deals with causes of creation of wealth in an economy.

 iv) The term 'wealth' used in this definition considers material wealth.

b) **The Welfare Definition of Economics**

Alfred Marshall (1842 - 1924) defined Economics as – the study of mankind in the ordinary business of life. According to him, "Economics examines that part of individual and social action which is most closely connected with the attainment and with the use of material requisites of well being."

Key features of Welfare Definition:

  i) It defines Economics as the study of activities related with human being and their material welfare.

 ii) Marshall clarified that Economics is related with incomes of individuals and its uses for creating material welfare.

 iii) Collectively incomes of a group of individuals form the wealth of a nation and ultimate objective of Economics is to increase welfare of individual by their day to day activities.

c) **The Scarcity Definition of Economics**

Lionel Charles Robbins (1898 – 1984) defined Economics as – a science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.

Key features of Scarcity Definition:

i) It recognized that Economics is a science deals with the economic behaviours of human being.

ii) It also assessed that all the physical and non-physical resources have associated with scarcity.

iii) The resources have multiple utility and it can only be used either for one purpose or another. The use of most of the resources are mutually-exclusive.

iv) There is a need of efficient use of scarce resources, and primary objective of Economics is to ensure efficiency in use of resources with a purpose to satisfy human wants.

d) The Growth Definition of Economics:

Paul Samuelson (1915 – 2009) defined Economics as - the study of how man and society choose with or without the use of money to employ the scarce productive resources, which have alternative uses, to produce various commodities over time and distributing them for consumption, how or in the future among various person or groups in society.

Key features of Growth Definition:

i) It deals with allocation of scarce resource to be used in productive purposes.

ii) The selection of most efficient use of the resources from the alternative ways.

iii) The growth of economies will depend upon the consumption and production in the economy.

iv) This definition also points towards Economics as a study of economic system.

Economics has been defined by various Economists and social thinkers with different objectives and contexts. All these definitions are correct and none can be taken as universally acceptable. It points to the fact that Economics as a subject deals with divergent areas, issues and activities and has a wide scope.

This is why it is important for a student to understand the nature as well as scope of the subject along with its limitations, before delving deep into its content.

WHAT IS THE NATURE OF ECONOMICS?

Through the discussion in previous section, one can observe that, Economics has repeatedly been termed as science of wealth, or science of material welfare, which is good enough reason to take a notion whether or not Economics is really a science. The very first question to describe the nature of economics is to ask - whether economics is a science or an art or both.

Science uses empirical approach. Here are five key features of scientific study; a) Verifiability, b) Objectivity, c) Controllability d) Reliability and e) Predictability. Economics should have these features to qualify as a science subject. The definition of Economics clearly underlines it as a subject which is centered around economic behaviour of human beings. The study of human behaviour is different from the study of nature. The objectivity, controllability and other features of typical science subjects do not amicably fit for the study of human behavior in a social setting. Further, it is also mentioned that basic economics is a set of tools for measurement of economic process; be it GDP growth rate, inflation rate, elasticity of demand etc., and this feature adds significant limitation on Economics being an art. Therefore, Economics is being placed in the midst of being science or art. Let us explain this in detail.

Economics as a Science

A subject is considered science if:

- It is a systematised body of knowledge which studies the relationship between cause and effect.

- It is capable of measurement.

- It has its own methodological apparatus.

- It should have the ability to forecast.

If we analyse economics, we find that it has all the features of science. Like science, it studies cause and effect relationship between economic phenomena. To understand, let us take the law of demand. It explains the cause and effect relationship between price and quantity demanded for a commodity. It says, given that other things remain constant, as price rises, the demand for a commodity falls and vice versa. Here, the cause is price and the effect is change in quantity demanded. Similarly, the outcomes are measurable in terms of money. It has its own methodology of study (induction and deduction) and it forecasts the future market condition with the help of various statistical and non-statistical tools. Thus, a majority of economic laws are of this type and therefore, economics is science.

But it is to be noted that economics is not a perfect science like physical science. The fact is that we cannot rely upon the accuracy of the economic laws. The predictions made on the basis of economic laws can easily go wrong.

This is because economists do not have uniform opinion about a particular event. The problem of actual results differing from the predicted ones arises on account of the fact that in economics we cannot have controlled experiments.

This is so because economic processes involve human agency, their behaviours and adaptation to a particular economic situation. In other words, the subject matter of economics is the economic behaviour of man which is highly unpredictable. Money which is used to measure outcomes in economics is itself a dependent variable. It is not possible to make correct predictions about the behaviour of economic variables.

Economics as an Art

A discipline of study is termed as art if it tells us how to do a thing that is to achieve an end (objective). It is noteworthy that the final justification for studying economics lies in the possibility of our ability to use it for solving economic problems faced by us. Prof. J. M. Keynes says that "An art is a system of rules for the achievement of a given end."

We know that in practice, economics is used for achieving a variety of goals. Every individual economic unit whether acting as a consumer or a producer or an investor or a supplier of an input or in any other capacity has an economic goal to achieve. It decides its course of action by keeping in mind the end to be achieved and the situation faced by it. Even at national level the authorities formulate a variety of policies. In certain cases they attempt to plan and operate the entire economy so as to achieve a given set of ends. Therefore, economic laws are widely used and relied upon at all levels of our economic activities. And that makes economics an art.

Art is nothing but practice of knowledge. Whereas science teaches us to know, art teaches us to do. Unlike science which is theoretical, art is practical. If we analyse economics, we find that it has the features of being an art also. Its various branches provide practical solutions to various economic problems. It helps in solving various economic problems which we face in our day-to-day life.

From the above discussion, one can easily understand that Economics is not a pure science as other natural sciences. Economics does follow standard practices adopted by science subjects, since it has human element attached in its processes, it cannot completely satisfy the fundamental requirements of being a science subject. Economics is both a science and an art. It is science in its methodology and art in its application.

Positive and Normative Economics

The next question arises as to whether Economics is positive or normative in nature

The application of mathematical models, empirical testing of economic theories, scientifically obtainable results for business or policy decisions makes Economics as being positive in nature while moral, ethical and collective welfare approaches used in Economics make it normative in nature.

A positive or pure science analyses cause and effect relationship between variables but it does not pass value judgment. In other words, it states what is and not what ought to be. Positive statements are about facts. They state what the reality is. According to Robbins, economics is concerned only with the study of the economic decisions of individuals and the society as positive facts but not with the ethics of these decisions. Economics should be neutral between ends. It is not for economists to pass value judgments and make pronouncements on the goodness or otherwise of human decisions. An individual with a limited amount of money may use it for buying liquor and not milk, but that is entirely his business. A community may use its limited resources for making guns rather than butter, but it is no concern of the economists to condemn or appreciate this policy. Economics only studies facts and makes generalizations from them. It is a pure and positive science, which excludes from its scope the normative aspect of human behaviour.

Complete neutrality between ends is, however, neither feasible nor desirable. It is because in many matters the economist has to suggest measures for achieving certain socially desirable ends. For example, when he suggests the adoption of certain policies for increasing employment and raising the rates of wages, he is making value judgments; or that the exploitation of labour and the state of unemployment are bad and steps should be taken to remove them. Similarly, when he states that the limited resources of the economy should not be used in the way they are being used and should be used in a different way; that the choice between ends is wrong and should be altered, etc. he is making value judgments.

Normative economics is concerned with normative statements. Normative statements are concerned with what ought to be? In this case, economics is not concerned with facts rather it is concerned with how things should be. As normative science, economics involves value judgments. It is prescriptive in nature and describe 'what ought to be' or 'what should be the things'. For example, the questions like what should be the level of national income, what should be the wage rate, how much of national product be distributed among people - all fall within the scope of normative economics. Thus, normative economics is concerned with welfare propositions.

The above discussion shows that Economics is both positive as well as normative in nature.

SCOPE OF ECONOMICS

The scope of the subject of economics is vast and ever expanding. It is no more a branch of knowledge that deals only with the production and consumption. However, the basic thrust still remains on using the available resources efficiently while giving the maximum satisfaction or welfare to the people on a sustainable basis. One example gives us an understanding of how vast the scope of the subject of economics is. In December 2007, the IPCC (Intergovernmental Panel on Climate Change) was awarded the *Nobel Peace Prize* for their efforts to build up and disseminate greater knowledge about man-made climate change, and to lay the foundations for the measures that are needed to counteract such change. What did IPCC do to disseminate knowledge about climate change? IPCC basically presented economic analysis of impacts of climate change and estimates to mitigate the challenge of climate change. Economics is being extensively used in assessment of impact of climate change in almost every industrial zone, environmental projects, energy plants and investment in renewable energy resources like solar, wind, tidal and others. Economics is also used for assessment of economic efficiency of space mission, analysis of socio-economic problems like expenditure on health-care, eradication of poverty, management of government budget, taxation, investment in industrial production.

Given this, we can list some of the major branches of economics as under:

(i) ***Micro Economics:*** This is considered to be the basic economics. Microeconomics may be defined as that branch of economic analysis which studies the economic behaviour of the individual unit, may be a person, a particular household, or a particular firm. It is a study of one particular unit rather than all the units combined together. The microeconomics is also described as price and value theory, the theory of the household, the firm and the industry. Most production and welfare theories are of the microeconomics variety.

(ii) ***Macro Economics:*** Macroeconomics may be defined as that branch of economic analysis which studies behaviour of not one particular unit, but of all the units combined together. Macroeconomics is a study in aggregates. Hence, it is often called Aggregative Economics. It is, indeed, a realistic method of economic analysis, though it is complicated and involves the use of higher mathematics. In this method, we study how the equilibrium in the economy is reached consequent upon changes in the macro-variables and aggregates. The publication of Keynes' General Theory, in 1936, gave a strong impetus to the growth and development of modern macroeconomics.

(iii) ***International Economics:*** As the countries of the modern world are realising the significance of trade and commerce with other countries, the role of international economics is getting more and more significant nowadays.

(iv) ***Public Finance:*** The great depression of the 1930s led to the realization of the role of government instabilising the economic growth besides other objectives like growth, redistribution of income, etc. Therefore, a full branch of economics known as Public Finance or the fiscal economics has emerged to analyse the role of government in the economy. Earlier the classical economists believed in the *laissez faire* economy ruling out role of the government in economic issues.

(v) ***Development Economics:*** After the Second World War many countries got freedom from the colonial rule, their economics required different treatment for growth and development. This led to emergence of new branch of economics known as development economics.

(vi) ***Health Economics:*** A new realisation has emerged from human development for economic growth. Therefore, branches like health economics are gaining momentum. Similarly, educational economics is also coming up.

(vii) ***Environmental Economics:*** Unchecked emphasis on economic growth without caring for natural resources and ecological balance, now, economic growth is facing a new challenge from the environmental side. Therefore, Environmental Economics has emerged as one of the major branches of economics that is considered significant for sustainable development.

(viii) ***Urban and Rural Economics:*** Role of location is quite important for economic attainments. There is also much debate on urban-rural divide. Therefore, economists have realised that there should be specific focus on urban areas and rural areas. Therefore, there is expansion of branches like urban economics and rural economics. Similarly, regional economics is also being emphasised to meet the challenge of geographical inequalities. There are many other branches of economics that form the scope of economics. There are welfare economics, monetary economics, energy economics, transport economics, demography, labour economics, agricultural economics, gender economics, economic planning, economics of infrastructure, etc.

CENTRAL PROBLEMS OF AN ECONOMY

The problem of scarcity of resources which arises before an individual consumer also arises collectively before an economy. On account of this problem and economy has to choose between the following:

(i) Which goods should be produced and in how much quantity?

(ii) What technique should be adopted for production?

(iii) For whom goods should be produced?

These three problems are known as the central problems or the basic problems of an economy. This is so because all other economic problems cluster around these problems.

1. What to produce?

There are two aspects of this central problem – firstly, which goods should be produced, and secondly, what should be the quantities of the goods that are to be produced. The first problem relates to the goods which are to be produced. In other words, what goods should be produced? An economy wants many things but all these cannot be produced with the available resources. Therefore, an economy has to choose what goods should be produced and what goods should not be.

The priority has to be made between consumer or producer goods; or general or capital goods; or civil goods or defence goods.

The second problem is what should be the quantities of the goods that are to be produced. Production of goods depends upon the use of resources. Hence, this problem is the problem of allocation of resources. If we allocate more resources for the production of one commodity, the resources for the production of other commodities would be less.

2. How to produce?

The second central problem faced by any economy is which technique should be used for the production of given commodities. This problem arises because there are various techniques available for the production of a commodity such as, for the production of wheat, we may use either more of labour and less of capital or less of labour or more of capital. With the help of both these techniques, we can produce equal amount of wheat. Such possibilities exist relating to the production of other commodities also.

Therefore, every economy faces the problem as to how resources should be combined for the production of a given commodity. The goods would be produced employing those methods and techniques, whereby the output would be the maximum and cost of production would be the minimum.

3. For whom to produce?

The main objective of producing a commodity is its consumption in the economy. However, even after employing all the resources of an economy, it is not possible to produce all the commodities which are required. Therefore, an economy has to decide as to for whom goods should be produced. This problem is the problem of distribution of produced goods and services. Therefore, what goods should be consumed and by whom depends on distribution of National Product.

All the three central problems arise because resources are scarce. Had resources been unlimited, these problems would not have arisen. For example, in the event of resources being unlimited, we could have produced each and every thing we wanted, we could have used any technique and we could have produced for each and everybody.

Besides, what, how and for whom there are three more problems which are also regarded as basic problems.

PRODUCTION POSSIBILITY CURVE

Professor Paul Samuelson is a leading figure among those economists who have explained the working of the economic system through these three questions. According to Samuelson, the main functions of an economic system are to answer these three questions.

Explanation of these Problems using Production Possibility Curve

Professor Samuelson used the concept of the production possibility curve to explain the economic problem of a society. A production possibility curve is the locus of all such combinations of two commodities which can be produced in a country with its given resources and technology.

Fig.1.1 : Production Possibility Curve
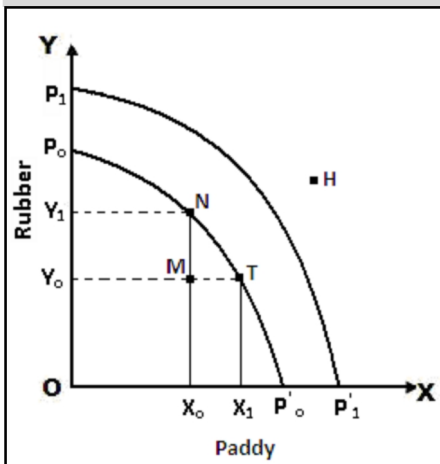
In Fig. 1.1, $P_0P'_0$ is the production possibility curve of a country. It shows different combinations of paddy (X) and natural rubber (Y) which of the country can produce with its available resources and technology. It can choose any such combination like N or T which lies on this curve.

1. Limited Resource: Here, the combination point N shows OY_1 amount of natural rubber and OX_0 amount of paddy. Again, the combination point T shows OY_0 amount of natural rubber and OX_1 amount of paddy. Thus, point N shows relatively higher amount of natural rubber as compared to point T. It implies that if the country wants to produce more of paddy, it has to reduce the production of natural rubber. This shows the limited availability of natural resources. Due to this reason, the country cannot choose any such combination like 'H' which lies beyond the production possibility curve.

2. Problem of 'What to Produce and in What Quantity': This curve also reflects the problem of 'what to produce'. If the country uses all of its resources for the production of only natural rubber, then the maximum possible production of natural rubber will be OP_0. In that case, there will be no production of paddy. Similarly, if the country uses all of its resources for the production of paddy then, the maximum possible production of paddy will be OP'_0. But in that case, the production of natural rubber will be zero.

3. Efficient Utilisation of Available Resources: If the country chooses the combination point M, i.e., if it produces OX_0 of paddy and OY_0 of natural rubber then it would indicate inefficient utilisation of resources. Here, the country can increase the production of paddy from OX_0 to OX_1 by keeping the production of natural rubber unchanged at OY0 (i.e., the country can move from point M to T). Similarly, in this situation, the country can also increase the production of natural rubber from OY_0 to OY_1 by keeping the production of paddy unchanged at OX_0 (i.e., the country can move from point M to N.). Thus, if the country chooses any combination of X and Y on the production possibility curve, it implies efficient utilisation of available resources. However, if it chooses any combination that lies below that curve, it would indicate inefficient utilisation or under utilisation of existing resources.

4. Improvement in Technology and Increase in the Amount of Resources: If new resources are available or if the level of technology is improved (e.g., application of high-yielding varieties of seeds, better methods of cultivation, better irrigational facilities, etc.) then the whole production possibility curve will shift outward. This is shown by $P_1P'_2$ curve in Fig. 1.1. In that case, the country can produce more of both X and Y commodities.

OPPORTUNITY COST

Opportunity cost is the value of alternative foregone in order to have something else. This value is unique for each individual. You may, for instance, forgo ice cream in order to have mashed potatoes. For you, the mashed potatoes have a greater value than dessert. But you can always change your mind in the future because there may be some instances when the mashed potatoes are just not as attractive as the ice cream. The opportunity cost of an individual's decisions, therefore, is determined by his or her needs, wants, time and resources (income).

This is important to the production possibility curve/ production frontier because a country will decide how to best allocate its resources according to its opportunity cost. Therefore, if the country chooses to produce more wine than cotton, the opportunity cost is equivalent to the cost of giving up the required cotton production.

Let's look at another example to demonstrate how opportunity cost ensures that an individual will buy the less expensive of two similar goods when given the choice. For example, assume that an individual has a choice between two telephone services. If he or she were to buy the most expensive service, that individual may have to reduce the number of times he or she goes to the movies each month. Giving up these opportunities to go to the movies may be a cost that is too high for this person, leading him or her to choose the less expensive service.

Remember that opportunity cost is different for each individual and nation. Thus, what is valued more than something else will vary among people and countries when decisions are made about how to allocate resources.

WORKING OF AN ECONOMIC SYSTEM

An economic system is an entire set of arrangements and institutions meant for meeting the two-fold objectives of a society:

– increasing the availability of resources

– ensuring the economic use

It is well known that economic systems, as created by different societies differ from each other, the economic system of even a given society keeps evolving and changing overtime, partly on account of ongoing efforts of the society to meet the problem of the scarcity of resources. Broadly speaking, types of economic systems are based upon per capita income, prioritization of individuals to spend their resources and scarcity of both income and resources. The best possible solution to these three potential problems is the basis of a successful economic system. Precedence set by society, its individuals and the government for the attainment of resource mobility and individual freedom is fundamental to the right choice of system for any society. There are three different types of economic system.

– Capitalist Economy

– Socialist Economy

– Mixed Economy

A Capitalist Economy

A Capitalist economic system is one which is characterized by free markets and the absence of government intervention in the economy. In practice a capitalist economy will need some government intervention, primarily to protect private property. In the real world, many economies which are viewed to have a capitalist economic system may have government spending taking up 35% of GDP. This is because the government pays for welfare, health, education and national defence. However, the economy is still viewed as capitalist because in the area of private enterprise firms are free to decide what to produce and for whom. Capitalist economic systems invariably lead to inequalities of wealth and income. However, it is argued that this inequality provides an incentive for wealth generation and economic growth. A Capitalist economic system is often contrasted to a Socialist or Communist economic system where economic decisions are made centrally by government agencies.

Features:

1. Capitalism derives its name from the fact that in this system, means of production are not owned by the government or by cooperatives. They are owned privately, that is by individuals and households. Business units (and therefore, the resources owned by them) also belong to individuals and households. The institution of private property also covers the right to inheritance. The institutions of property and inheritance have two strong implications.

2. People acquire a motive for earning more, because they are allowed to keep their earnings both for current and future use. For this reason, they are always on the look out for opportunities of increasing their income. In the process, if need be, they are also ready to work hard. The net result is that a capitalist system is characterised by a high production potential.

3. Private property and inheritance lead to ever-increasing inequalities of income and wealth. These inequalities, in their turn, result in unequal opportunities of earning an income. The market prices of various goods and services fail to correspond to their relative worth to the society. Therefore, a cumulative process develops in which the owners of capital are able to add to their incomes faster than the workers can add to their incomes, since they have to depend upon the income from their labour only.

4. Capitalism is also characterised by what is known as the policy of laissez-faire on the part of the authorities. The term laissez-faire means absence of state intervention in the working of the economy. The solution of the basic problems of the economy is left in the hands of market mechanism. In other words, the authorities do not try to regulate the prices, demand or supply. The market mechanism, through the interaction between demand and supply forces, brings about changes in prices. The prices, in turn, act as signals for individual economic units and guide them in their respective activities as consumers and producers, etc.

5. In theory, it is usually assumed that the market structure of a capitalist economy is competitive in nature. In practice, however, it need not be so. It is possible that while the authorities are following a policy of laissez-faire, the market itself is not competitive enough. It may have strong monopoly elements. It may be what we call a 'monopolistic competition', or there may be other forms of technical or institutional hurdles in the way of competition.

6. Another salient feature of capitalism is the use of money and credit. This is so because a capitalist system, by its very nature, tends to become quite complex with a large variety of goods, services and occupations. The producers undertake production mainly for sale in the market and not for self-consumption. Similarly, a capitalist economy tends to have production projects which have a long technical life. All these aspects of capitalism necessitate an elaborate system of financing its economic activities and therefore the use of money and credit.

7. In capitalism, all economic activities are guided by market forces. Producers produce only those goods

and services which are demanded by consumers in the market. The entire economy operates to meet the needs and preferences of the consumers. This characteristic of capitalism is known as that of 'consumer sovereignty'.

Merits:

1. Capitalist system is self- regulatory.

2. The process of economic growth is faster under capitalism. It has a tendency to register a huge rate of growth in national income and per capita income.

3. The capitalist system decides 'what to produce' and 'how to produce' in consonance with the forces of demand and supply.

4. Capitalistic system provides an incentive for efficient decision making and their implementation in the form of economic gains to the decision makers which ensure a high degree of operative efficiency in the system.

5. It provides flexibility to adapt to the changed circumstances.

6. It ensures individual freedom giving entrepreneurs incentive to work hard.

7. Capitalism ensures optimal allocation of resources in different uses.

Demerits:

1. Capitalism generates inequalities of income and wealth.

2. Wide differences in economic opportunities.

3. Distortion in the production pattern.

4. Capitalism wastes its productive resources.

5. Production of merit goods is not profitable.

6. Business units produce only those goods and services which are profitable.

7. Loss of human values and welfare.

8. Increases the wastage of resources as a result of competition.

A Socialist Economy

The concept of a socialist economy has its origin in the drawbacks of capitalism. There are no pre-determined details of a socialist system, but its main features are well recognized. This system tries to get rid of the drawbacks of capitalism and incorporate those features which are considered to be desirable. For example, it aims at removing the problems of inequalities of income and wealth, inequalities of economic opportunities, unemployment, cyclical fluctuations, and waste of productive resources. The advocates of socialism believe that most of these drawbacks come into being because of certain basic features of capitalism including the institutions of private property and inheritance, and the use of market mechanism.

In the words of H. D. Dickinson, "Socialism is an organisation of the society in which the material means of production are owned by the whole commodity and operated by organs, representative of and responsible to, all members of community according to a general plan, all members of community being entitled to get benefits from the results of such socialist planned production on the basis of equal rights."

According to Maurice Dobb, "The fundamental character of socialism consists in its abolition of class relations which form the basis of Capitalist production through expropriation of the propertied class and the socialisation of land and capital."

Accordingly, Socialism is basically designed to have the following salient features.

Features:

1. In a socialist economy, the institutions of private property and inheritance are abolished. The 'private sector' as we understand by this term, does not exist. It means that the ownership of means of production is not in the hands of individuals and households. Instead, they are owned by the government authorities and/or cooperatives or society. Individuals and households do not own any business concerns. And no one is an employee of a private business. Private ownership is allowed only in the case of consumption goods and personal belongings, and that too only to a limited extent. And to that extent, even inheritance of 'private property' may be allowed. An important implication of the abolition of private property and inheritance is that economic decision-making is no longer left in private hands.

2. A socialist economy is not guided by free working of a market mechanism. It is rendered ineffective. In a sense, its operation is "frozen". Consumers and producers are not allowed freedom in their decision-making. The consumers have to take decisions within the limits set by the authorities. They also lay down production schedules and decide what to produce, how much to produce, and what resources to be used as inputs. Thus, demand and supply forces are not to respond to changes in prices. Instead they are regulated with the objective of serving the national interests as a whole. Similarly, prices are not allowed to fluctuate in response to changes in demand and supply. They are also controlled and regulated by the authorities. Only in some cases, cooperatives may be permitted to change the prices within certain limits. Systematic operation of a complex economy necessitates a complex and vast set up of decision-making. In capitalism, this complex task is handled by market mechanism. But in socialism, its substitute has to be created. This is usually done in the form of centralised economic planning.

3. A socialist economy recognises the ill-effects of money and credit. In capitalist system, these create cyclical fluctuations and inequalities of income and wealth. Ideally, therefore, a socialist economy prefers not to have these institutions. But the hard reality is that it cannot do without them. In an economy, which produces a large number of goods and services, cannot have an efficient system of physical rationing. It has to create and operate a complex system of income distribution which is not possible without using money in some form. Accordingly, it is not able to completely discard the use of money and credit but restricts it to the minimum necessary extent.

4. Capitalism derives its name from the fact that in this system, means of production are not owned by the government or by cooperatives. They are owned privately, that is by individuals and households. Business units (and therefore, the resources owned by them) also belong to individuals and households. The institution of private property also covers the right to inheritance.

5. Another important feature of socialism is to have class less society. Since under socialism, no property is privately owned, so there is no question of existence of classes. Every person in the society gets the share in production according to his own merits.

Merits:

A socialist economy discards the use of market mechanism and replaces it with some form of regulatory authority, such as the planning commission. It also abolishes the institutions of private property and inheritance. Given these features of the socialist economy, it tries to remove the basic demerits of capitalism by pursuing the objectives of;

– distributive justice

– social security

– elimination of fluctuations of economy

  – coordinated development

  – elimination of social disputes.

Demerits:

While socialism is able to get rid of the problems of income inequalities, unemployment and cyclical fluctuations of national income and prices, it is not able to provide economic incentives and disincentives for hard work and initiative. As a result, it continues to suffer from slow growth rate, poor productivity of labour and low per capita income. On account of these weaknesses of the economy, it becomes difficult to raise the consumption standards of the masses. And this gives rise to the necessity of restructuring the economic system for better results.

A Mixed Economy

A mixed economy tries to– avoid the ill-effects of both capitalism and socialism and secure the benefits of both.

For this reason, it incorporates some elements of both capitalism and socialism. However, there is no pre-determined and standardised proportion in which their features could be selected and combined.

Features:

1. The selection of detailed features of a mixed economy is made with reference to the working of market mechanism, and its expected effects (both beneficial and harmful) on the society as a whole. In other words, we take up one segment of the economy at a time, and adopt the following procedure.

2. It is decided that working of the selected segment of the economy should be guided by free market mechanism if the net effect of this arrangement is expected to be beneficial for the society as a whole.

3. If the working of the segment under consideration can be made beneficial for the society by subjecting the working of market mechanism to some regulatory measures, then the said segment is subjected to be governed by a regulated market mechanism. In other words, in this case, the interaction between demand, supply and prices is regulated in a manner and to the extent found necessary.

4. Working of a segment under free market mechanism may be harmful for the society in certain respects. If it is possible to make its working beneficial by regulating the market mechanism, the said segment is again subjected to a regulated market mechanism. However, the extent and nature of regulatory measures are revised from time to time in light of the changing circumstances.

5. In some other cases, it may be found that market mechanism continues to have net ill-effects even after restrictions and regulations. In such cases, therefore, market mechanism is not allowed to operate at all. The authorities take over one or more functions of market mechanism, namely, demand decisions, supply decisions and prices. This is generally done through public sector undertakings, which need not be guided market forces.

6. Socialistic society offers equal opportunity to all to rise above level of standards. For this, state authority make available health, education, transport facilities to all, either subsidized or free of cost.

7. Since equalities of income and wealth in society is the major objectives in socialist economy, the authority tries to achieve it by social or state ownership of means of production.

Thus, in a mixed economy, the net result is that market mechanism is not totally abolished. It is allowed to operate with different degrees of freedom in different segments of the economy. Indian economy provides a very good example of a mixed economy as it operates in practice. In theory, a mixed economy is far superior to either capitalism or socialism since it tries to acquire beneficial features of both. In practice, however, it suffers from many drawbacks. Some of these arise on account of the fact that it is extremely difficult to work out the details of a mixed economy. The system has a tendency to suffer from several inner contradictions. Once the rules and procedures for its working have been formulated, it is not possible to revise them frequently or rapidly. The economy, therefore, fails to adjust itself to changing circumstances as rapidly as it should. The success of a

mixed economy also depends upon the integrity and expertise of the government administration, the expertise and freedom of the management, and the willingness of the workers to recognise their moral duty of increasing productivity.

ECONOMIC CYCLES

The term economic cycle (or business cycle) refers to economy-wide fluctuations in production or economic activity such as income employment, savings and investment over several months or years. These fluctuations occur around a long-term growth trend, and typically involve shifts over time between periods of relatively rapid economic growth (an expansion or boom), and periods of relative stagnation or decline (a contraction or recession or depression).

Business cycles are usually measured by considering the growth rate of real gross domestic product. Despite being termed cycles, these fluctuations in economic activity do not follow a mechanical or predictable periodic pattern.

STAGES OF THE ECONOMIC CYCLE

Economic Boom/ Inflation

A boom occurs when national output is rising strongly at a rate faster than the trend rate of growth (or long-term growth rate) of about 2.5% per year. In boom conditions, output and employment are both expanding and the level of aggregate demand for goods and services is very high. Typically, businesses use the opportunity of a boom to raise output and also widen their profit margins.

Characteristics of an Economic Boom

- Strong and rising level of aggregate demand - often driven by fast growth of consumption

- Rising employment and real wages

- High demand for imported goods & services

- Government tax revenues will be rising quickly

- Company profits and investment increase

- Increased utilization rate of existing resources

- Danger of demand-pull and cost-push inflation if the economy overheats.

Economic Slowdown

A slowdown occurs when the rate of growth decelerates - but national output is still rising. If the economy continues to grow (albeit at a slower rate) without falling into outright recession, this is known as a soft-landing.

Economic Recession

A recession means a fall in the level of real national output (i.e. a period when the rate of economic growth is negative). National output declines, leading to a contraction in employment, incomes and profits. The last recession in Britain lasted from the summer of 1990 through to the autumn of 1992. When real GDP reaches a low point at the end of the recession, the economy has reached the trough - economic recovery is imminent.

An economic slump is a prolonged and deep recession leading to a significant fall in output and average living standards.

Characteristics of an Economic Recession

– Declining aggregate demand for output

– Contracting employment / rising unemployment

– Sharp fall in business confidence and profits and a decrease in capital investment spending

– De-stocking and heavy price discounting

– Reduced inflationary pressure and falling demand for imports

– Increased government borrowing lower interest rates from central bank

– The last stage of economic recession is depression, at which the economic activities touches its low point in terms of production, employment, savings and investment.

Economic Recovery

A recovery occurs when real national output picks up from the trough reached at the low point of the recession.

The pace of recovery depends in part on how quickly aggregate demand starts to rise after the economic downturn. And, the extent to which producers raise output and rebuild their stock levels in anticipation of a rise in demand.

LESSON ROUND UP

– Study of Economics enable us to understand the logical background of the economic activities performed by us or happening in our surrounding.

– Economics is the study of choice under conditions of scarcity.

– Economics is the study of how we choose to use limited resources to obtain the maximum satisfaction of unlimited human wants.

– Adam Smith (1723 - 1790) defined Economics as- Science of wealth with an objective to increase wealth and richness of a country.

– Alfred Marshall (1842-1924) defined Economics as - "Economics examines that part of individual and social action which is most closely connected with the attainment and with the use of material requisites of well-being".

– Lionel Charles Robbins (1898 – 1984) defined Economics as – a science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.

– Paul Samuelson (1915 – 2009) defined Economics as - the study of how man and society choose with or without the use of money to employ the scarce productive resources, which have alternative uses, to produce various commodities over time and distributing them for consumption, how or in the future among various person or groups in society.

– The methodology of Economics is a scientific study while the subject matter of it is of a social science. Since its assumptions are based on moral philosophy, and its measurement tools follow scientific approach, Economics as a subject stand at the intersection point of Science and Art.

– A: Positive economics is objective and fact based, while normative economics is subjective and value based. Positive economic statements do not have to be correct, but they must be able to be tested and proved or disproved.

– Normative economic statements are opinion based, so they cannot be proved or disproved.

– Microeconomics is the study of economics at an individual, group or company level.

– Macroeconomics is the study of a national economy; its income, consumption, interest rate, unemployment rate, investments, government expenditure and external economy.

– Economics is not only the study of individual choices we make, as well as their consequences. When some of the consequences are harmful, economists study what – if anything – the government can or should do about them.

– The Central Problem of the Economy are; a) what to produce? b) how to produce? and c) for whom to produce?

– The production possibility frontier (PPF) is a curve depicting all maximum output possibilities for two goods, given a set of inputs consisting of resources and other factors. The PPF assumes that all inputs are used efficiently.

– Opportunity cost is the value of the next best choice that one gives up when making a decision.

– Capitalism is an economic system where private actors are allowed to own and control the use of property in accord with their own interests, and where the invisible hand of the pricing mechanism coordinates supply and demand in markets in a way that is automatically in the best interests of society. Government, in this perspective, is often described as responsible for peace, justice, and tolerable taxes.

– Socialism is an economic system characterized by social ownership, control of the means of production, and cooperative management of the economy. A socialist economic system would consist of an organization of production to directly satisfy economic demands and human needs, so that goods and services would be produced directly for use instead of for private profit driven by the accumulation of capital.

– Mixed Economy is an economic system in which both the private enterprise and a degree of state monopoly (usually in public services, defense, infrastructure, and basic industries) coexist. All modern economies are mixed where the means of production are shared between the private and public sectors.

GLOSSARY

Ends	The objectives pursued by human beings while engaged in economic activities.
Means	The instruments or resources used in attaining the perceived objectives.
Scarcity	The imbalance between ends and means.
Laissez Faire	Free market economics with a minimum of government intervention.
Economics	The Word Economics Has Greek Origin. OikosPlusNomos Meaning House Management. The word economics has something to do with economizing on the use of means to attain ends out of scarce resources.
Positive Economics	Deals with scientific issues and questions. Solves economics central problems without value judgment.
Normative Economics	Deals with ethical issues, questions and problems. Solves economic problems bringing in value judgment.
Micro Economics	Analysis of any single unit of economics.
Macro Economics	Analysis of all units of economics studied together.

Production Possibility Curve	The locus of output combinations which an economy can produce using technically most efficient methods of production and allocating resources in an economically efficient manner.

MULTIPLE CHOICE QUESTIONS

1. **Economics is the study of**

   A) Production technology

   B) Consumption decisions

   C) How society decides what, how, and for whom to produce

   D) The best way to run society

2. **The opportunity cost of a good is**

   A) The time lost in finding it

   B) The quantity of other goods sacrificed to get another unit of that good

   C) The expenditure on the good

   D) The loss of interest in using savings

3. **In a free market _____ _____**

   A) Governments intervene

   B) Governments plan production

   C) Governments interfere

   D) Prices adjust to reconcile scarcity and desires

4. **In the mixed economy**

   A) Economic problems are solved by the government and market

   B) Economic decisions are made by the private sector and free market

   C) Economic allocation is achieved by the invisible hand

   D) Economic questions are solved by government departments

5. **Normative economics forms _____ based on _____**

   A) Positive statements, facts

   B) Opinions, personal values

   C) Positive statements, values

   D) Opinions, facts

6. **Microeconomics is concerned with**

   A) The economy as a whole

   B) The electronics industry

   C) The study of individual economic units

D) The interactions within the entire economy

7. **Macroeconomics is the study of _____**

   A) Individual building blocks in the economy

   B) The relationship between different sectors of the economy

   C) Household purchase decisions

   D) The economy as a whole

8. **PPC stands for**

   A) Production possibility Curve

   B) Production preference Curve

   C) Price possibility curve

   D) Producer preferred curve

9. **Which of the following is not the basic economic problem?**

   A) What to produce

   B) How to produce

   C) For whom to produce

   D) Where will they be produced

10. **PPC shifts outward**

    A) If technology improves

    B) If technology deteriorates

    C) If quantity of productive resources decreases

    D) Both b and c

Answer Key : 1. (c), 2. (b), 3. (d), 4. (a), 5. (b), 6. (c), 7. (d), 8. (a), 9. (d), 10. (a)

Suggested Readings

1. Ahuja, H. L. (2009). Modern Micro Economics. S. Chand.

2. Gans, J., King, S., Stonecash, R., & Mankiw, N.G. (2011). *Principles of economics*. Cengage Learning.

3. Hall, R.E., & Lieberman, M. (2012). *Microeconomics: Principles and applications*. Cengage Learning.

4. Varian, H.R., Bergstrom, T.C., & West, J.E. (2003). Intermediate microeconomics (Vol. 7). New York: Norton.

Lesson 2
Basic Elements of Demand and Supply

LEARNING OBJECTIVES

The theory of demand and supply is a theory of price and output determination. In a market economy, consumers and producers take the changes in the market price as signals for making their consumption and production decisions. The following lesson explains how the market price is determined by the interaction of demand and supply and how market price is restored or equilibrium is restored if there is any disturbance in the market price. This lesson also discusses the concept of elasticity and various types of elasticities of demand and supply.

INTRODUCTION

The market system works through market forces of demand and supply. The market system functions in an orderly manner because it is governed by certain fundamental laws of market known as: (i) law of demand and law of supply. Analysis of the demand and supply facilitates in determining the prices of goods and services in the market. In a market economy, market mechanism determines the prices of a commodity where demand and supply intersect each other, that is, prices are the result of decisions taken by buyers and sellers in market place to buy and sell. Thus, the most important function of micro-economics is to explain the law of demand and supply, market mechanism and working of the price determination system.

The market of an economy comprises of two different groups of participants: Consumers and Producers. Demand analysis focuses on the behavior of consumers, while supply analysis examines the behavior of producers. The consumer indirectly tells the producer what she is willing to buy and how much she is willing to pay based on her actual spending patterns. The producer supplies the product if she can make a profit by doing so. The forces of demand and supply coordinate to arrive at an equilibrium price and quantity of output which best satisfies the consumers and reaps maximum profits to producers.

Given the understanding, it would be useful to have basic understanding about the law of demand and the law of supply, that is, how these laws make the market system work and how the equilibrium is determined.

CONCEPT OF DEMAND

In economics, demand has a particular meaning distinct from its ordinary usage. In common language demand and desire are treated as synonyms. In economics, demand comprises of three things:

 (i) Desire of commodity;

 (ii) Sufficient money to purchase the commodity; and

 (iii) Willingness to spend money to purchase that commodity.

This understanding makes it clear that a want or a desire does not become a demand unless an individual has ability to purchase and willingness to satisfy it. For instance, a poor man who wishes to have a car, his wish or desire for a car will not constitute the demand for car because he cannot afford to pay for it, that is, he has no purchasing power to make his wish or desire effective in the market.

- Important characteristic of demand is that demand for a commodity is always in reference to a particular price. Demand has no meaning unless it is related to price. Further, demand always means demand per unit of time. In other words, demand for a good at a particular price is the amount of it which will be brought at a particular point of time.

- Demand of a commodity is influenced by several factors such as desire of the consumer for a commodity, income of the consumer, the prices of substitute and complementary commodities etc.

- It is worth noticing that demand of an individual is different from the market demand.

- Individual demand refers to the quantity of a commodity that a particular person is willing to purchase at a given price over a period of time, say per day, per week, per month etc.

- Market demand is the total quantity that all the users of a commodity are willing to buy at a given price over a period of time. In other words, market demand is the sum of individual demands for a particular product.

Determinants of Demand

Demand for a good by a consumer can vary in response to several factors such as its own price, prices of other related goods, income of the consumer, tastes and preferences of the consumer etc. Symbolically,

$D_x = f(P_x, P_Y, Y, T,…)$

Where P_x is the price per unit of good X, P_Y, the prices of the related goods, Y is the income of the consumer, T represents the tastes and preferences of the consumer. Following are the leading determinants of demand:

(a) *Price of the Commodity:* The first determinant of the demand for a good is its own price. The consumer compares the marginal utility expected from a good with its price and decides whether it is worth buying or not. A fall in the price induces the consumer to buy more of the good and an increase in the price causes a fall in demand.

(b) *Prices of Related Commodities:* Prices of related commodities also affect the demand of the commodity (say X). There are two ways in which a good can be related to another good:

   – Substitute goods: If the price of a substitute good, Y increases, the demand for that good falls and the consumer wants to by more of X instead. In contrast, if the price of the substitute good falls the consumer increases the demand for that good and hence wants to buy less of X. It has positive cross price effect.

   – Complement goods: If the price of a complementary good, Y increase, the demand for that good falls so does the demand of its complement X. In the same way, a fall in the price of a complementary good causes an increase in the demand for X. It has negative cross price effect.

(c) *Level of Income:* The demand for a good is also affected by the levels of income of the consumer. With an increase in income the consumer wants to buy more of a good. However, if the good is considered an 'inferior' one, he is expected to reduce its demand when his income increases.

(d) *Expected Change in Price:* If price of a good is expected to increase, demand for that good also increases and vice-versa. A consumer wants to buy a good before its price goes up and will postpone its purchase if price is expected to fall.

(e) *Other Factors*: Other factors which affect the aggregate market demand for a good include the size of population, the marketing and sale campaigns by the suppliers, the 'selling expenses' incurred by the sellers, the tastes and preferences of the buyers, and distribution of income and wealth. For example, the richer sections are likely to spend a smaller proportion of their incomes on basic necessities and a larger proportion on luxuries and durable consumer goods.

Law of Demand

According to the law of demand, other things being equal, if price of a commodity falls, the quantity demanded of it will rise, and if price of the commodity rises, its quantity demanded will decline. It implies that there is an inverse relationship between the price and quantity demanded of a commodity, other things remaining constant. In other words, other things being equal, quantity demanded will be more at a lower price than at higher price.

The law of demand describes the functional relationship between price and quantity demanded. Among various factors affecting demand, price of a commodity is the most critical factor. Thus, demand of a commodity is mainly determined by the price of commodity.

$$Dx = f(Px)$$

Demand Schedule

A demand schedule is a tabulated statement that indicates the different quantities of a commodity that would be demanded at different prices. Demand schedule is of two types:

– Individual Demand Schedule

– Market Demand Schedule

Individual Demand Schedule: A hypothetical demand schedule is given in Table 2.1. A demand schedule has two columns, namely

– price per unit of the good (P_x)

– quantity demanded per period (D_x)

The demand schedule is a set of pairs of values of P_x and D_x.

Table: 2.1 Individual A's Demand Schedule for Commodity X

Price per unit of Commodity X P_x	Quantity Demanded of Commodity X (Units)
10	6000
20	5000
30	4000
40	3000
50	2000
60	1000

Table 2.1 presents the individual demand schedule. It is observed that as the price of commodity X increases the demand for commodity X starts declining. For instance, at price Rs. 10 per unit, 6000 units are demanded. When the price of commodity X increases to Rs. 60, the demand for commodity X declines to 1000 units. Similarly, one can read the table in reverse order and arrive at the conclusion that as price of good declines, its demand decreases. This inverse relationship between price and quantity gives the law of demand.

Demand Curve: A demand curve is a graphic representation of the demand schedule. It is a locus of pairs of price per unit (P_x) and the corresponding demand-quantities (D_x). (See. Figure 2.1)

Fig. 2.1: Demand Curve of individual A
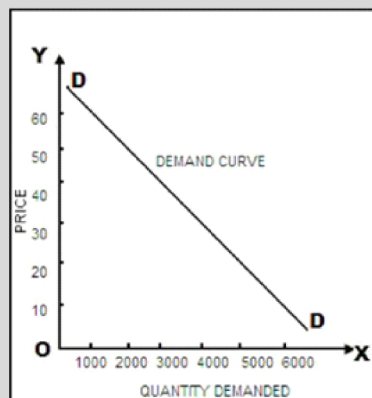
Fig 2.1 shows demand curve, where X-axis measures quantity demanded and Y-axis shows prices. As the price increases from 10 to 60 the quantity demanded declines from 6000 to 1000, establishing a negative relation among the two.

This also shows that the conventional demand curve, other things being constant, is downward sloping.

Market Demand Schedule: Market demand schedule is the horizontal summation of individual demand schedules. For instance, if the price of a bike is Rs.50,000 and at this price, Consumer A demands 2 bikes and Consumer B demands 1 bike (assuming there are only two consumers in this market) then market demand for the bike will be 3 (sum total of the demand of the two consumers). Likewise by adding up the individual demands for this commodity at different price levels, we can ascertain market demand for the commodity.

Table: 2.2 Market Demand Schedule for Commodity X

Per unit Price of Commodity X (Rs)	Quantity Demanded by Consumer A (Q_A)	Quantity Demanded by Consumer B (Q_B)	Market Demand $Q_A + Q_B$
10	6000	9000	15000
20	5000	8000	13000
30	4000	7000	11000
40	3000	6000	9000
50	2000	5000	7000
60	1000	4000	5000

Table 2.2 illustrates the market demand schedule. There are only 2 consumers (A and B) in the market. At price of Rs. 10 per unit, Consumer A demanded 6000 units of commodity X and consumer B demanded 9000 units of commodity X. Therefore, market demand of commodity X at this price is 15000 units. At price Rs. 60 per unit, consumer A demanded 1000 units of good X and consumer B demanded 4000 units of commodity X. Thus, at price Rs. 60 per unit, market demand of commodity X is Rs. 5000 units. It also follows the law of demand as market demand of commodity X reduces from 15000 units to 5000 when the price of commodity increases from Rs. 10 to Rs. 60 per units.

Shape of Market Demand Curve : Market demand curve can be derived graphically by horizontal summation of the individual demand curve at each price of commodity X.

Fig. 2.2: Market Demand Curve

Let us take another example to explain the horizontal summation of individual demand curves. Figure 2.2 presents the market demand curve. Market demand is simply the sum of demand of Consumer A and Consumer B. For instance, at price Rs. 3 per unit, consumer A demanded 10 units of good X and Consumer B demanded 15 units of good X. Market demand at price Rs. 3 is 25 units, that is, sum of Consumer A's and Consumer B's demand. Similarly, other values of market demand curve have been derived by adding up the demand of individual A and individual B at each price.

Reasons for Negative Slope of the Demand Curve

Some of the major reasons for the negative slope of the demand curve are listed below:

a) *The Law of Diminishing Marginal Utility*: This law provides that when a consumer buys additional units of a good, its marginal utility falls and consumer always compares the marginal utility of a good with the price to be paid for it; therefore, price which he is willing to pay for additional unit of a good falls.

b) *Increased Real Income*: A fall in the price of a good increases the real income of the consumer. He is able to buy more of the good under question, or buy more of other goods. Similarly, an increase in the price of a good reduces his real income. In this case, the income effect leads to a reduction in the demand of the good. This factor also contributes to the downward slope of demand curve. For instance, a person 'X' has got Rs. 10,000 as the nominal income. Now the amount of goods and services he can purchase with those Rs. 10,000 is nothing but real income. So when the price of a particular commodity falls, the real income, i.e. the purchasing power of the consumer increases and that is how at decreased price consumer can buy more of a particular commodity.

Change in Demand – Increase / Decrease versus Expansion/Contraction in Demand

There are various factors besides price of a commodity that affect the demand for that commodity. It should be noted that the 'location' of a demand curve (that is its distance from origin) is determined by factors other than its own price, while its slope is determined by its price. In other words, demand for a good changes when,

– a consumer moves from one point to another on the same demand curve (Movement along the demand curve)

– when the entire demand curve shifts its position (Movement from one demand curve to the other)

Movement along the Demand Curve: A demand curve is drawn on the assumption that all factors determining the demand behavior of a consumer, other than the price of the good itself, remain the same. When price of the good changes, the consumer moves along the given demand curve and changes the quantity demanded of the good.

Fig. 2.3: Movement along the Demand Curve
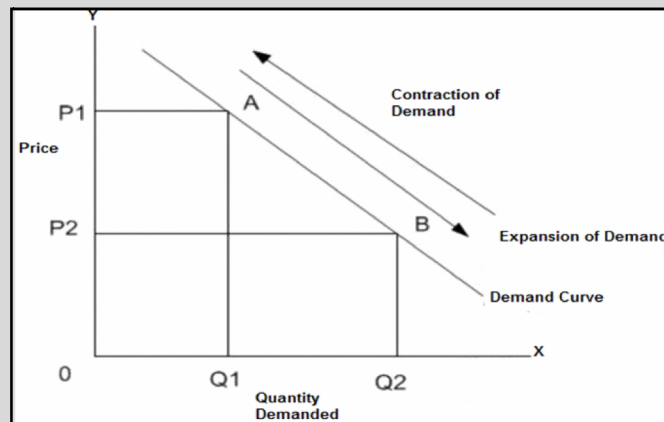
– A reduction in quantity of demand from Q_2 to Q_1 on account of an increase in price from P_2 to P_1 is termed as 'contraction' of demand (Figure 2.3). In this case, the consumer moves upward along the demand curve.

– In contrast, suppose the price of the good falls from P_1 to P_2, the consumer moves downward along the demand curve and increases the purchase of good from OQ_1 to OQ_2. This is termed as 'expansion' of demand.

Movement from One Demand Curve to the Other: If the demand for good changes without the change in price, the consumer shifts from one demand curve to the other. Such a movement is termed as 'increase' in demand if the demand curve shifts to the right. And it is termed as 'reduction' in demand when the movement is towards the left.

The shift in demand curve results from change in factors other than the price of the good. Other factors which influence the demand curve are change in price of related goods, change in income of the consumer, change in tastes and preferences of the consumer etc.

Fig. 2.4: Shift in Demand Curve
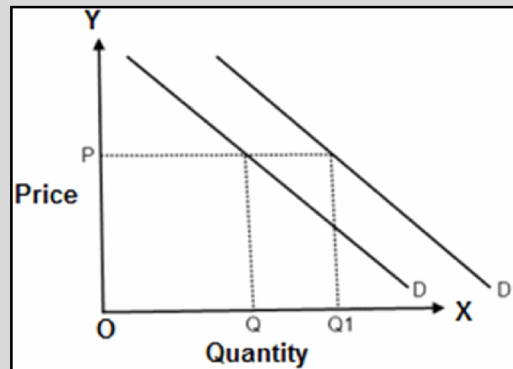
Figure 2.4 shows two demand curves D and D_1. With a given price OP per unit of good X, the consumer buys OQ quantity when he is on the demand curve D. Now the price of the commodity remains unchanged but the changes in other factors caused demand curve to shift from D to D_1. As a result of shift in demand, keeping price constant, the demand of good X increases from OQ to OQ_1.

Exceptions to the Law of Demand

The law of demand is widely applicable to a large number of goods. However, there are certain exceptions to it on account of which a change in the price of a good does not lead to a change in its quantity demanded in the opposite direction.

(a) *Expected Change in the Price of a Good:* While an actual change in the price of a good leads to a change in its demand in the opposite direction, expectations related to price changes the demand in the same direction. When the price of a good is expected to increase, consumers increase the demand so as to avoid paying a higher price later. Similarly, when the price of a good is expected to fall, the consumers postpone their purchases of it.

(b) The consumer may not consider a good as 'normal' or 'superior'. There are four types of such goods.

– **Inferior Goods:** Some goods are consumed generally by poorer sections of the society. It is believed

that with an increase in income such a consumer should move to a 'better' quality substitute good. For example, with an increase in income, a typical poor consumer shifts his demand from coarse grains to finer varieties of cereals. Therefore, with a fall in the price of a good (more so a necessity on which the consumer is spending a large part of his budget), the real income of the consumer goes up. If he considers the good under consideration as inferior, he reduces its demand and buys more of its substitute(s).

– **Giffen Goods:** Some special varieties of inferior goods are termed as Giffen goods. Cheaper varieties of food like bajra, cheaper vegetable like potatoes come under this category. Sir Robert Giffen of Ireland first observed that people used to spend more of their income on inferior goods like potato and less of their income on meat. But potatoes constitute their staple food. When the price of potato increased, after purchasing potato they did not have so many surpluses to buy meat. So the rise in price of potato compelled people to buy more of potato and thus raised the demand for potato. This is against the law of demand. This is also known as Giffen paradox. So giffen goods are products that people continue to buy even at high prices due to lack of substitute products.

– **Ignorance**: In some cases, the consumers suffer from the false notion that a higher priced good is of better quality. This happens mainly in the case of those goods where a typical consumer is not able to judge the quality easily. In such cases, the sellers may be able to sell more not by lowering the price but by raising it.

– **Conspicuous Consumption:** Certain goods are meant for adding to one's social prestige. These form the part of 'status symbol' for showing that their user is a wealthy or cultured person. The consumers consider it as a distinction to have these goods. In other words, a commodity may be purchased not because of its intrinsic value but because it is expected to add to the social prestige of the buyer. For example: Diamonds and expensive jewellery, expensive carpets. Their demand falls, if they are inexpensive.

(c) Change in Fashion: A change in fashion and tastes affects the market for a commodity. When a broad toeshoe replaces a narrow toe, no amount of reduction in the price of the latter is sufficient to clear the stocks. Broad toe on the other hand, will have more customers even though its price may be going up. The law of demand becomes ineffective.

(d) Complementary Goods: Law of demand may be violated in the case of complementary goods also. For example: if the price of the DVD player falls leading to increase in its demand, in spite of rise in price of DVDs, their demand will increase.

CONCEPT OF SUPPLY

Supply represents how much the market can offer. The quantity supplied refers to the amount of a good producers are willing to supply when receiving a certain price. The supply of a good or service refers to the quantities of that good or service that producers are prepared to offer for sale at a set of prices over a period of time. Supply means a schedule of possible prices and amounts that would be sold at each price. The supply is not the same concept as the stock of something in existence, for example, the stock of commodity X in Delhi means the total quantity of Commodity X in existence at a point of time; whereas, the supply of commodity X in Delhi means the quantity actually being offered for sale, in the market, over a specified period of time.

Determinants of Supply

At any point in time, the total quantity supplied of a good or service in the market is influenced by a number of factors. Some of the important factors include the following:

(a) *Costs of the Factors of Production:* The cost of factor inputs such as land, labour, capital etc. is one of the determinant factors which influence the market supply of a product. For instance, if the price of labour goes up, then the supply of the product will decline due to higher labor cost.

(b) *Change in Technology:* The change in technology as a result of constant research and development activities in terms of improved machinery, improved methods of organization and management helps the business units or firms to reduce the cost of production. All this contributes significantly to the increase in market supply at given prices.

(c) *Price of Related Goods (Substitutes):* Prices of related commodity also affect the supply of a commodity (say X). If the price of a substitute good, Y increases, the supply for that good increases and the producer would shift the allocation of resources to Y from X.

(d) *Change in the Number of Firms in the Industry (Market):* A change in the number of firms in the industry as a result of profitability also influences the market supply of a good. For example, an increase in number of firms in the industry attracted by higher profit would increase the quantity supplied of good at given prices.

(e) *Taxes and Subsidies:* A change in government fiscal policy in terms of change in tax rate or amount of subsidy may influence the supply of a good in the market. A decrease/increase in the amount of tax/subsidy on the good would allow firms to offer larger amount of a good at a given prices.

(f) *Goal of a Business Firm:* The goal of a business firm such as profit maximization, sales maximization or both is also responsible to influence the market supply of a good or service. In case the firm is interested to maximize profit, the same may be attained by decreasing the market supply of a good under certain conditions whereas goal of sales maximization will be attained by increasing the supply.

(g) *Natural Factors:* Natural Factors such as climatic changes, particularly in the case of agricultural products influence the supply.

Law of Supply

The law of supply states that a firm will produce and offer to sell greater quantities of a product or service as the price of that product or service rises, other things being equal. There is direct relationship between price and quantity supplied. In this statement, change in price is the cause and change in supply is the effect. Thus, the price rise leads to increase in supply and not otherwise. It may be noted that at higher prices, there is greater incentive to the producers or firms to produce and sell more. Other things include cost of production, change of technology, prices of inputs, level of competition, size of industry, government policy and non-economic factors.

Thus '*Ceteris Paribus*';

(a) With an increase in the price of a good, the producer is willing to offer more quantity in the market for sale.

(b) The quantity supplied is related to the specified time interval over which it is offered.

Three Alternative Ways of Expressing Supply

Supply of a good by an individual producer/firm or the market/industry as a whole is conventionally expressed in three alternative forms, namely;

– a supply function

– a supply schedule

– a supply curve

Supply Function: A supply function of an individual supplier is an algebraic form of expressing his behaviour

with regard to what he offers in market at the prevailing prices. In it, the quantity supplied per period of time is expressed as a function of several variables.

General form of the supply function is $S_x = f(P_x, C_x, T_x)$

Example of a supply function for good X is $S_x = 200 + 15P_x$

Where, S_x denotes quantity supplied of good X, P_x denotes the price of good X, C_x represents cost of production and T_x is technology of production.

Supply Schedule: A supply schedule is a tabular statement that shows different quantities or services that are offered by the firm or producer in the market for sale at different prices at a given time. It describes the relationship between quantities supplied of a good in response to its price per unit, while all non-price variables remain unchanged. A supply schedule has two columns, namely

- price per unit of the good (P_x)
- quantity supplied per period (S_x)

The supply schedule is a set of pairs of values of P_x and S_x. There are two types of supply schedule, namely

- individual supply schedule
- market supply schedule

Table: 2.3 Individual Supply Schedule for Commodity X

Price per unit of Commodity X P_x	Quantity of Commodity X (Units)
10	1000
20	2000
30	3000
40	4000
50	5000
60	6000

Individual Supply Schedule: It relates the supply of a good or service by one firm at different prices, other things remains constant or equal. Table 2.3 shows that as the price of good X increases from Rs. 10 to 60 the corresponding supply of the commodity increases from 1000 units to 6000 units.

Market Supply Schedule: The market supply schedule, on the other hand, like market demand schedule is the sum of the amounts of good supplied for sale by all the firms or producers in the market at different prices during a given time. Let us assume, there are two producers for a good (Table, 2.4). At price Rs. 10 per unit, producer A sells 1000 units and producer B offers 2000 units. Hence the total market supply at Rs. 10 per unit is 3000. As price increases from Rs 10 to Rs. 50, the market supply increases from 3000 units to 11000 units.

Table: 2.4 Market Supply Schedule for Commodity X

Per unit Price of Commodity X P_X	Quantity Supplied by Producer A (Q_A)	Quantity Supplied by Producer B (Q_B)	Market Supply $Q_A + Q_B$
10	1000	2000	3000
20	2000	3000	5000
30	3000	4000	7000
40	4000	5000	9000
50	5000	6000	11000
60	6000	7000	13000

Supply Curve: The supply curve is a graphical representation of the information given in supply schedule. The higher the price of the commodity or product, the greater will be the quantity of supply offered by the producer for sale and vice versa, other things remains constant.

Fig. 2.5: Supply Curve
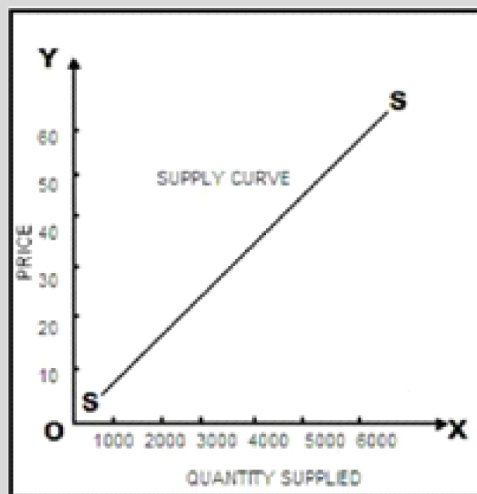
Fig 2.5 shows supply curve of producer A, where X-axis is quantity supplied and Y-axis shows prices. As the price increases from 10 to 60 the quantity supplied rises from 1000 to 6000, establishing a positive relation among the two.

This implies that the supply curve is upward sloping.

Reasons for positive slope of supply curve

Law of supply states that there exist positive relationship between the price of a product and its quantity supplied, *ceteris paribus*. The supply curve slopes upward from left to right. It means that the supply of a product increases with increase in its price and decreases with decrease in its price. Here, the question emerges that why law of supply behave in this fashion and not otherwise.

Some of the reasons or explanations given by the economists in this regard are stated as under:

Producers hire and use resources in order to make profits from the sale of the output produced, at least in the private enterprise sector of a mixed economy. In economics, it is assumed that a common objective of firms is not only to make profits, but to maximize profits. Hence, since most firms could supply other products apart from the good or service in question, it is unlikely that they would be prepared to supply large amounts of this good or service if the price were very low, because this implies low profits after production costs have been taken into

account. Indeed, there must be some price at which no firms would be prepared to supply a product because it is so low that it would not cover the minimum cost at which each unit of that product could be produced.

Conversely, the higher the price that firms can charge to sell a product, given the costs of the factors of production, the more profitable that product becomes and the larger the amount that will be supplied to the market. Hence, we can make the logical assumption that, other things being equal (including the number of firms in the market, the scale of production of each firm and the costs of factors), the higher the price of a good or service, the greater will be the quantity supplied of that good or service to the market, and vice versa.

Change in Supply – Increase / Decrease versus Expansion/Contraction in Supply

The 'location' of a supply curve (that is its distance from origin) of a product is determined by factors other than its own price, while its slope is determined by its own price. In other words, supply for a good changes when

– a producer moves from one point to another on the same supply curve (Movement along the supply curve)

– when the entire supply curve shifts its position (Movement from one supply curve to the other)

Movement along the Supply Curve: An increase in price will increase the quantity supplied, but a decrease in price will reduce the quantity supplied. The supply curve is positively sloped- upward and to the right, as against the demand curve which is negatively sloped.

A reduction in quantity of supply on account of a decrease in price is termed as 'contraction' in supply. In this case, the supplier moves downwards along the supply curve. In contrast, if the price of the good rises, the supplier moves upwards along the supply curve and offers to sell more of the good. This is termed as 'expansion' in supply.

Fig. 2.6: Change in Quantity Supplied
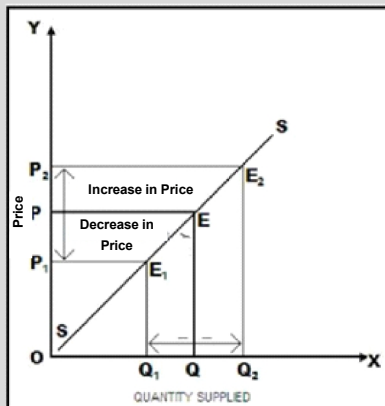
Fig 2.6 shows supply curve, where X axis is quantity supplied and Y axis shows prices.

Movement along the supply curve is shown in this fig, where the change in quantity supplied is only on account of the change in own price of the commodity, other things being equal.

Movement from One Supply Curve to the Other: If the supply changes without change in price, the supplier shifts from one supply curve to the other. Such a movement is termed 'increase' in supply when the producer moves to the outer supply curve to the right. And it is termed a 'reduction' in supply when the movement is to the inner supply curve to the left.

Fig. 2.7: Change in Supply
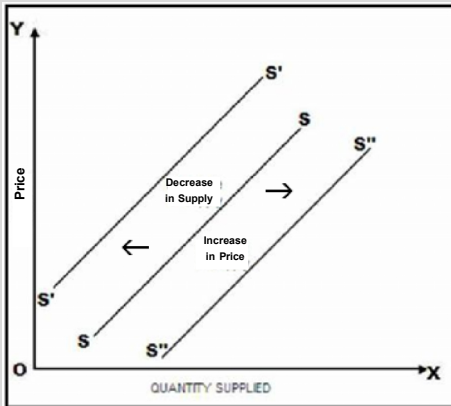
Thus, in Fig 2.7, we may consider supply curves S' and S" which shift to the left and right, respectively, from supply curve S implying decrease and increase in supply respectively.

Exceptions to the Law of Supply

The normal law of supply is widely applicable to a large number of goods. However, there are certain exceptions to it on account of which a change in the price of a good does not lead to a change in its supply in the same direction. The law of supply is not a universal principle that applies to all markets under all circumstances. There are, in fact, numerous important exceptions to the law of supply.

(a) *Expected Change in the Price of a Good:* While an actual change in the price of a good leads to a change in its supply in the same direction, an expected change in its price changes the supply in the opposite direction. When the price of a good is expected to increase, supplier decreases the supply-quantity so as to avoid selling at lower prices in the current period and sell more at even higher prices in the future period.

(b) *Market power:* If the supply side of the market is controlled by a small number of sellers then the law of supply might not operate. For example, in case of monopoly (single seller) may not necessarily offer a larger quantity even though the price is higher. Market control by the monopolist allows it to set the market price based on demand conditions and fix the quantity supplied without cost constraints being imposed from the supply side.

(c) *Competition: In* other market structures like oligopoly and monopolistic competition, sellers might be facing more competition, thereby offering to sell larger quantities at lower prices and negating the law of supply.

(d) *Perishable Goods:* In cases of perishable goods the supplier would offer to sell more quantities at lower prices to avoid running into losses due to damage of the product.

(e) *Legislation Restricting Quantity:* Suppliers cannot offer to sell more quantities at higher prices where the government has put regulations on the quantity of the good to be offered or the price ceiling at which the good is to be offered in market. Producers are unable to play with either of the factors on their own.

(f) *Agricultural Products:* Since the production of agricultural products can not be increased beyond a limit, the supply can also not be increased beyond this limit even if the prices are higher; the producer is unable to offer larger quantities.

(g) *Artistic and Auction Goods:* The supply of such goods can not be increased or decreased easily. Thus, it is difficult to offer larger quantities even if the prices shoot up.

DETERMINATION OF EQUILIBRIUM PRICE AND QUANTITY

In context of demand and supply, equilibrium is a situation in which quantity demanded equals quantity supplied and there is no incentive to buyers and sellers to change from this situation. The market clears itself and becomes stable (that is, at the market equilibrium, every consumer who wishes to purchase the product at the market price is able to do so, and the supplier is not left with any unwanted inventory). Equilibrium price is the price at which the demand is equal to supply.

Law of demand and law of supply explain separately the 'plans' of consumers as to how much they would buy at a given price and the 'plans' of producers as to how much they would offer for sale at the given price. The demand curve and the supply curve really show what consumers and producers would do if they were given the opportunity.

Although the demand would be very high at lower prices but in practice consumers may never get the opportunity to buy the product at that low price because suppliers are not willing to supply at that price. Similarly, although suppliers may be prepared to offer a large amount for sale at a high price, they may not be able to sell it at all because the consumers are not willing to buy at that price.

The demand for a product and the supply of a product are two sides of the market, and it is necessary to bring these together to establish equilibrium in the market which is the point where both the sides of the market are satisfied simultaneously.

This can be better understood with the help of the following illustration. (See Table 2.5). Let us take demand and supply schedule for good X and analyze equilibrium position. Equilibrium price is Rs. 40 and equilibrium quantity is 9000 units.

Table 2.5: Demand and Supply Schedule for Commodity or Good X

Per unit Price of Commodity X	Quantity Demanded (Units)	Quantity Supplied (Units)	Description
10	12000	3000	Excess Demand
20	11000	5000	Excess Demand
30	10000	7000	Excess Demand
40	9000	9000	Demand = Supply
50	7000	11000	Excess Supply
60	5000	13000	Excess Supply

Fig. 2.8: Demand and Supply Curves
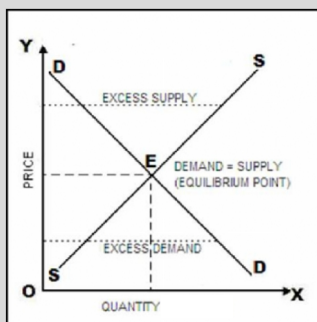
Fig 2.8 shows a demand and a supply curve, where X axis is quantity and Y axis shows prices.

In the figure the market equilibrium is established based on the data from Table 2.6

The equilibrium is the state when Demand equals Supply which is at point E.

Effect of Change in the Conditions of Demand and Supply on Market Price

The market price, or equilibrium price, is determined by the interaction of demand and supply curves. Remember that the demand curve and the supply curve for a commodity are drawn up on the assumption that all other factors which might affect the demand for or supply of the commodity remain constant. The equilibrium price will remain stable in the market as long as these other factors of demand and supply do not change. If any of these factors change, this will create excess demand or excess supply and so initial equilibrium price will also change.

For example, a condition of drawing a demand curve is that level of income does not change. If the level of income increases, there will be an increase in demand for a commodity X at the existing market price. Hence, if the price remains the same, supply will be the same as before, and with increased demand there will be a shortage, causing pressure on the existing price, which suppliers will then raise. On other hand, if consumer's level of income decreases, other things remaining constant, with decreased demand there will be deficient demand/excess supply, causing existing price to fall.

Fig. 2.9: Effect of Change in Demand on Market Price
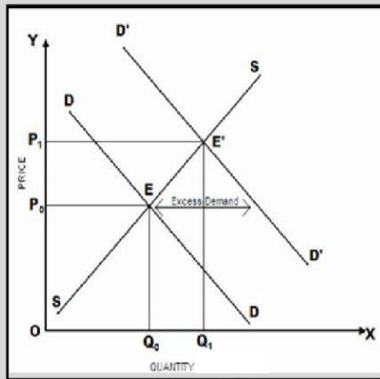
This effect can be represented diagrammatically as in Figure 2.9.

In Figure 2.9, the shift in demand curve from DD to D'D' shows the effect of an increase in demand as a result of an increase in consumers' income levels. Before the increase in demand, the equilibrium is shown by point E and equilibrium price was P_0 and the equilibrium output was Q_0. With the increase in demand, demand curve shifted to D'D'. As a result of the increase in demand, excess demand occurs at the price P_0, causing suppliers to expand output and raise the market price. As a result, excess demand got created resulting in increase in price and quantity supplied. A new equilibrium (shown by point E) is established at price P_1 and at output Q_1. Notice that a change in conditions of demand does not cause a movement of the supply curve – this could only result from changes in conditions of supply.

Fig. 2.10: Effect of Change in Supply on Market Price
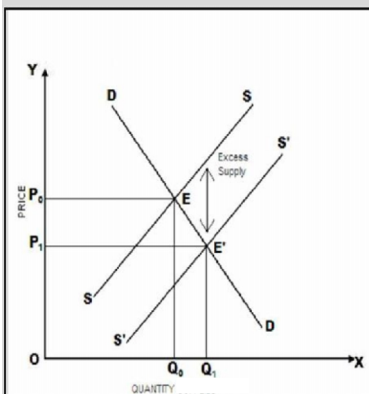
In Figure 2.10, the supply curves SS and S'S' show the effect of an increase in supply as a result of a favorable change in the conditions of supply (such as a reduction in the costs of production because of productivity improvements)— S'S' being the new supply curve. Before the increase in supply, the equilibrium price was P_0 and the equilibrium output was Q_0. As a result of the increase in supply, excess supply occurs at the price P_0, causing suppliers to lower the price in order to expand demand. A new equilibrium price is established at P_1 with a higher equilibrium output at Q_1. Notice again that a change in conditions of supply does not cause a shift in the demand curve.

It is useful to summarize the effects of changes in the conditions of demand and of supply on the equilibrium price and output, assuming that we are operating in a free market. It is as follows:

- Excess demand for a commodity will cause a rise in its price.

- Excess supply of a commodity will cause a fall in its price.

- Price will settle at one point where the quantity demanded equals the quantity supplied—the equilibrium price.

- An increase in demand (a movement of the demand curve to the right) will cause a rise in price and a rise in the quantity bought and sold.

- A decrease in demand (a movement of the demand curve to the left) will cause a fall in price and a fall in the quantity bought and sold.

- An increase in supply (a movement of the supply curve to the right) will cause a fall in price and a rise in the quantity bought and sold.

- A decrease in supply (a movement of the supply curve to the left) will cause a rise in price and a fall in the quantity bought and sold.

ELASTICITY OF DEMAND AND SUPPLY

Introduction

According to the law of demand, there is an inverse relationship between demand and price so that the demand curve has a negative slope. According to the law of supply, there is a positive relationship between supply and price so that the supply curve has a positive slope. This means that demand and supply respond to price. The law of demand and supply indicates the direction of change in demand in response to a change in price. It does not express the magnitude of change in demand or supply in response to a change in price. This information is provided by the tool of elasticity of demand.

Meaning of Elasticity

Elasticity refers to the ratio of the relative change in a dependent variable to the relative change in an independent variable i.e. elasticity is the relative change in the dependent variable divided by the relative change in the independent variable. For example, the ratio of percentage change in quantity demanded to percentage change in some other factor like price or income.

Elasticity of Demand

According to Marshall "the elasticity (or responsiveness) of demand in a market is great or small according as the amount demanded increases much or little for a given fall in price or diminishes much or little for a given rise in price".

The concept of elasticity of demand refers to the degree of responsiveness of demand of a good to a change in its determinants. Elasticity of demand differs in case of different commodities. For the same commodity, elasticity of demand differs from person to person. Analysis of elasticity of demand is not limited to price elasticity only, income elasticity of demand and cross elasticity of demand are also important to understand.

Elasticity of demand is mainly of three types:

- Price Elasticity of Demand

- Cross Price Elasticity of Demand

- Income Elasticity of Demand

Price Elasticity of Demand

Price elasticity of demand refers to the responsiveness of demand to a Change in price of a commodity. It may be noted that the price elasticity of demand has a negative sign because of the negative relationship between price and demand.

The formula for calculating price elasticity is:

$$E_d = \frac{\text{Change in Quantity Demanded}}{\text{Change in Price}}$$

There are five types of Price Elasticity of Demand depending upon the magnitude of response of demand to a change in price.:

– Perfectly elastic demand

– Perfectly inelastic demand

– Relatively elastic demand

– Relatively inelastic demand

– Unitary elastic demand

Perfectly elastic demand: The demand is said to be perfectly elastic when a very insignificant change in price leads to an infinite change in quantity demanded. A very small fall in price causes demand to rise infinitely. Likewise a very insignificant rise in price reduces the demand to zero. This case is theoretical which may not be found in real life. The demand curve in such a situation is parallel to X-axis. Numerically, elasticity of demand is said to be equal to infinity. ($E_d = \infty$)

Perfectly inelastic demand: The demand is said to be perfectly inelastic when a change in price produces no change in the quantity demanded of a commodity. In such a case quantity demanded remains constant regardless of change in price. The amount demanded is totally unresponsive to change in price. The demand curve in such a situation is parallel to Y-axis. Numerically, elasticity of demand is said to be equal to zero. ($E_d = 0$)

Relatively elastic demand: The demand is relatively more elastic when a smaller change in price causes a greater change in quantity demanded. In such a case a proportionate change in price of a commodity causes more than proportionate change in quantity demanded. For example: If price changes by 10% the quantity demanded of the commodity changes by more than 10%. The demand curve in such a situation is relatively flatter. Numerically, elasticity of demand is said to be greater than 1. ($E_d > 1$)

Relatively inelastic demand: It is a situation where a greater change in price leads to a smaller change in quantity demanded. The demand is said to be relatively inelastic when a proportionate change in price of a commodity causes less than proportionate change in quantity demanded. For example: If price changes by 20% quantity demanded changes by less than 20%. The demand curve in such a case is relatively steeper. Numerically, elasticity of demand is said to be less than 1. ($E_d < 1$)

Unitary elastic demand: The demand is said to be unitary elastic when a change in price results in exactly the same percentage change in the quantity demanded of a commodity. In such a situation the percentage change in both the price and the quantity demanded is the same. For example: If the price falls by 25%, the quantity demanded also rises by 25%. It takes the shape of a rectangular hyperbola. Numerically, elasticity of demand is said to be equal to 1. ($E_d = 1$).

Fig. 2.11 Types of Price Elasticity of Demand
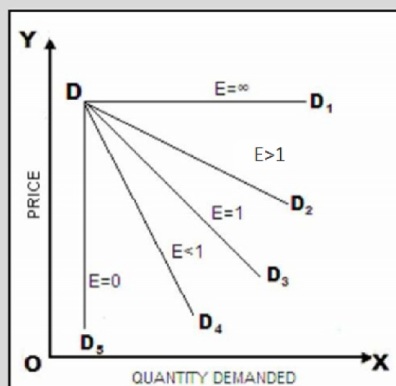
Fig 2.11 shows various demand curves, where X-axis shows quantity demanded and Y-axis shows prices.

Different types of price elasticities discussed above are shown in the diagram.

DD$_1$ – Perfectly Elastic Demand

DD$_2$ – Elastic Demand

DD$_3$ – Unitary Elastic Demand

DD$_4$ – Inelastic Demand

DD$_5$ – Perfectly Inelastic Demand

Price elasticity of demand can be measured through three popular methods. These methods are:

– Percentage Method

– Total Expenditure Method

– Graphic Method or Point Method

1. Percentage Method

According to this method, price elasticity is estimated by dividing the percentage change in quantity demanded by the percentage change in price of the commodity. Thus, given the percentage change of both quantity demanded and price; the elasticity of demand can be derived. If the percentage change in quantity demanded is greater that the percentage change in price, the elasticity will be greater than one.

If percentage change in quantity demanded is less than percentage change in price, the elasticity is said to be less than one. But if percentage change of both quantity demanded and price is same, elasticity of demand is said to be unit.

$E_d = (\triangle q/q) / (\triangle p/p) = (\triangle q/\triangle p) \times (p/q)$

Where, $\triangle$ q is the change in demand and $\triangle$ p is is the change in price, while original demand and price are q and p respectively.

2. Total Expenditure Method

Total expenditure method was formulated by Alfred Marshall. The elasticity of demand can be measured on the basis of change in total expenditure in response to a change in price. It is worth noting that unlike percentage method a precise mathematical coefficient cannot be determined in this case.

By the help of total expenditure method, it is determined whether the price elasticity is equal to one, greater than one or less than one. In this method, the initial expenditure before the change in price and the expenditure after the change in price are compared.

Total Outlay/ Expenditure = Price x Quantity Demanded

If, with a fall in price, it is found that the expenditure remains the same, elasticity of demand is said to be one ($E_d = 1$), if the total expenditure increases the elasticity of demand is said to be greater than one ($E_d > 1$), if the total expenditure diminishes with the fall in price, elasticity of demand is less than one ($E_d < 1$), and vice-versa.

Table: 2.6 Total Expenditure Method

Change in Prices	Elastic Demand ($e_d > 1$)	Unitary Elastic Demand ($e_d = 1$)	Inelastic Demand ($e_d < 1$)
Price Rise	Total expenditure falls	Total expenditure unchanged	Total expenditure rises
Price Fall	Total expenditure increases	Total expenditure unchanged	Total expenditure falls

The relationship presented in above table may also be shown graphically as under.

Fig. 2.12 : Price Elasticity of Demand using Total Expenditure Method
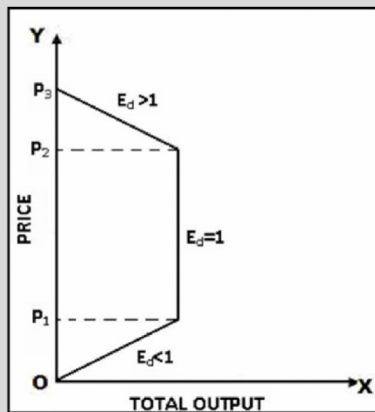
In Fig 2.12, demand for commodity X is unitary elastic over the price range OP_1 to OP_2 because total outlay or expenditure does not change with change in price. Demand is inelastic over the price range O to OP_1 because total expenditure increases or decreases with increase or decrease in prices respectively. Demand is elastic over the price range OP_2 to OP_3 because total outlay increases with decrease in prices and decreases with increase in prices.

Total Expenditure or Total Outlay method has certain drawbacks. Firstly, it is not able to give us an exact numerical measure of elasticity of demand. It only tells us whether the elasticity is equal to, less than, or more than one. Therefore, this method fails to compare demand elasticities of different goods. Secondly, it is not possible to use this method in measuring demand elasticity when demand changes in the same direction as the price, as in the case of giffen goods.

3. Graphic Method or Point Method

Graphic method is otherwise known as Point Method or Geometric Method. According to this method, elasticity of demand is measured on different points on a straight line demand curve. The price elasticity of demand at a point on a straight line is equal to the lower segment of the demand curve divided by upper segment of the demand curve at that point.

Thus, at mid point on a straight-line demand curve, elasticity will be equal to unity; at higher points (i.e., to the left of the mid-point) on the same demand curve elasticity will be greater than unity, at lower points (i.e. to the right of the mid-point) on the demand curve elasticity will be less than unity.

At a corner point on demand curve where there is no lower segment, elasticity of demand is equal to zero ($E_d = 0$) and where there is no upper segment on demand curve, elasticity of demand is equal to infinity, ($E_d = \infty$)

Fig. 2.13: Point Elasticity of Demand on a Straight Line Demand Curve
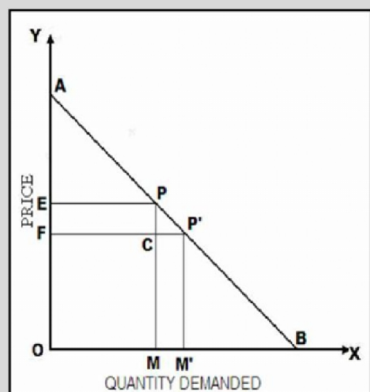
In Fig. 2.13, the demand curve is a straight line and meets X-axis and Y-axis at points B and A respectively, then elasticity of demand at any point such as P on the demand curve is given by the ratio:

$$\frac{\text{Segment of demand curve from point P to X- axis}}{\text{Segment of demand curve from point P to Y- axis}}$$

Fig. 2.14 Elasticity of Demand for Curvi-linear Demand Curve
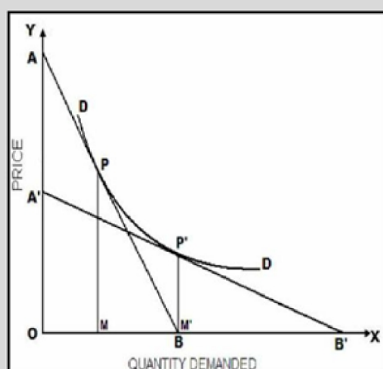
When demand curve is curvi-linear instead of a straight line, and we want to measure elasticity of demand at a point P on it, we should draw a tangent to it at point P, consider as if the tangent is the demand curve, and use the same approach as before in measuring elasticity of demand.

Thus, in Fig. 2.14, elasticity of demand at point P equals PB/AP = BM/OM, while at point P' it is P'B'/A'P' = B'M'/OM'.

Determinants of Price Elasticity of Demand

Price elasticity of demand is dependent upon a number of factors as follows:

(a) *Price Level*: The demand is generally elastic for moderately priced goods but, the demand for very costly and very cheap goods is inelastic. The rich do not bother about the prices of the goods that they buy. Very costly goods are demanded by the rich people and hence their demand is not affected much by change in prices. For example, increase in the price of maruti car from Rs. 3,00,000 to Rs. 3,20,000 will not make any noticeable difference in its demand. Similarly, the change in the price of very cheap goods (such as salt) will not have any effect on their demand, for their consumption which is very small and fixed.

(b) *Availability of Substitutes*: If a good has close substitutes, the price elasticity of demand for a commodity will be very elastic as some other commodities can be used for it. A small rise in the price of such a commodity will induce consumers to switch their consumption to its substitutes. For example gas, kerosene oil, coal etc. will be used more as fuel if the price of wood increases. On the other hand, the demand of such commodities which have no close substitutes is inelastic, such as salt.

(c) *Time Period*: A typical consumer finds it difficult to adjust his consumption of a good in the short run. He needs time to adjust to the changed situation. Therefore, demand elasticity of a good tends to increase in the long run.

(d) Proportion of Total Expenditure Spent on the Product: Elasticity of demand for a good is also dependent upon the proportion of a consumer's budget spent on it. On account of a price rise of a good, a consumer feels more concerned if he is spending a large proportion of his budget on it. The extent of change in demand by the consumer is not significant in the case of those goods on which the consumer spends a very small proportion of his monthly budget. In the former case elasticity of demand is higher, while in the latter case, it is low.

(e) Habits: Some products which are not essential for some individuals are essential for others. If individuals are habituated of some commodities the demand for such commodities will usually be inelastic, because they will use them even when their prices go up. A smoker generally does not smoke less when the price of cigarette goes up.

(f) Nature of the Commodities: The demand for necessities is inelastic and that for comforts and luxuries is elastic. This is so because certain goods which are essential will be demanded at any price, whereas goods meant for luxuries and comforts can be dispensed with easily if they appear to become costlier.

(g) Various Uses: A commodity which has several uses will have an elastic demand such as milk, wood etc. On the other hand, a commodity having only one or fewer uses will have inelastic demand. The consumer finds it easier to adjust the quantity demanded of a good when it is to be used for satisfying several wants than if it is confined to a single or few uses. For this reason, a multiple-use good tends to have more elastic demand.

(h) Postponing Consumption: Usually the demand for commodities, the consumption of which can be postponed, is elastic as the prices rise and are expected to fall again. For example, the demand for mp3 is elastic because its use can be postponed for some time if its price goes up, but the demand for rice and wheat is inelastic because their use cannot be postponed when their prices increase.

Cross Price Elasticity of Demand

The change in the demand of a good x in response to a change in the price of good y is called 'cross price elasticity of demand'. Its measure is

$$E_d = \frac{\text{Change in Quantity Demanded of Good X}}{\text{Change in Price of Good Y}}$$

Symbolically, $= (\triangle Q_x / \triangle P_y) / (Q_x / P_y)$

- Cross price elasticity may be infinite or zero.

- Cross price elasticity is positive infinity in case of perfect substitutes.

- Cross price elasticity is positive if the change in the price of good Y causes a change in the quantity demanded of good X in the same direction. It is always the case with goods which are substitutes.

- Cross price elasticity is negative if the change in the price of good Y causes a change in the quantity demanded of good X in the opposite direction. It is always the case with goods which are complements of each other.

- Cross price elasticity is zero, if a change in the price of good Y does not affect the quantity demanded of good X. In other words, in case of goods which are not related to each other, cross elasticity of demand is zero.

Income Elasticity of Demand

According to Stonier and Hague: "Income elasticity of demand shows the way in which a consumer's purchase of any good changes as a result of change in his income." It shows the responsiveness of a consumer's purchase of a particular commodity to a change in his income. Income elasticity of demand means the ratio of percentage change in the quantity demanded to the percentage change in income.

$$E_y = \frac{\text{Percentage Change in Quantity Demanded of Good X}}{\text{Percentage Change in Real Income of Consumer}}$$

Symbolically, $E_y = (\triangle Q_x / \triangle Y) / (Q_x / Y)$:

Where, Y denotes income of the consumer.

It is noteworthy that sign of income elasticity of demand is associated with the nature of the good in question.

Normal Goods: Normal goods have a positive income elasticity of demand so as consumers' income rises, demand also increases.

- Normal necessities have an income elasticity of demand between 0 and 1. For example, if income increases by 10% and the demand for fresh fruit increases by 4%, then the income elasticity is +0.4. Demand is rising less than proportionately to income.

- Luxuries have an income elasticity of demand greater than 1, $E_d > 1$.i The demand rises by more than the percentage change in income. For example, an 8% increase in income might lead to a 16% rise in the demand for restaurant meals. The income elasticity of demand in this example is +2. Demand is highly sensitive to income changes.

Inferior goods: Inferior goods have a negative income elasticity of demand. Demand falls as income rises. For example, as income increases, the demand for higher quality cereals goes up against the low-quality cheap cereals.

Price Elasticity of Supply

The law of supply indicates the direction of change in quantity supplied in response to a change in price. It does not express the magnitude of change in amount supplied in response to a change in price. This information is provided by the tool of elasticity of supply. Like the elasticity of demand, the elasticity of supply is the relative measure of the degree of responsiveness of quantity supplied of a commodity to a change in its price.

The greater the responsiveness of quantity supplied of a commodity to the change in its price, the greater is its elasticity of supply. To be more precise, the elasticity of supply is defined as a percentage change in the quantity supplied of a product divided by the percentage change in price. It may be noted that the elasticity of supply has a positive sign because of the positive relationship between price and supply.

The formula for calculating price elasticity of supply is:

$$E_S = \frac{\text{Percentage Change in Quantity Supplied}}{\text{Percentage Change in Price}}$$

There are five types of price elasticity of supply depending upon the magnitude of response of supply to a change in price: :

- Perfectly Elastic Supply
- Perfectly Inelastic Supply

– Relatively Elastic Supply

– Relatively Inelastic Supply

– Unitary Elastic Supply

Perfectly Elastic Supply: The supply is said to be perfectly elastic when a very insignificant change in price leads to an infinite change in quantity supplied. A very small rise in price causes supply to rise infinitely. Likewise a very insignificant fall in price reduces the supply to zero. The supply curve in such a situation is a horizontal line running parallel to x-axis. Numerically, elasticity of supply is said to be equal to infinity. ($E_s = \infty$)

Perfectly Inelastic Supply: The supply is said to be perfectly inelastic when a change in price produces no change in the quantity supplied of a commodity. In such a case, quantity supplied remains constant regardless of change in price. The amount supplied is totally unresponsive to change in price. The supply curve in such a situation is a vertical line, parallel to y-axis. Numerically, elasticity of supply is said to be equal to zero ($E_s = 0$).

Relatively Elastic Supply: The supply is relatively elastic when a small change in price causes a greater change in quantity supplied. In such a case a proportionate change in price of a commodity causes more than proportionate change in quantity supplied. For example, if price changes by 10% the quantity supplied of commodity changes by more than 10%. The supply curve in such a situation is relatively flatter. Numerically, elasticity of supply is said to be greater than 1 ($E_s > 1$).

Relatively Inelastic Supply: It is a situation where a greater change in price leads to smaller change in quantity supplied. The demand is said to be relatively inelastic when a proportionate change in price is greater than the proportionate change in quantity supplied. For example, if price rises by 20%, quantity supplied rises by less than 20%. The supply curve in such a case is relatively steeper. Numerically, elasticity of supply is said to be less than 1 ($E_s < 1$).

Unitary Elastic Supply: The supply is said to be unitary elastic when a change in price results in exactly the same percentage change in the quantity supplied of a commodity. In such a situation the percentage change in both the price and quantity supplied is the same. For example, if the price falls by 25%, the quantity supplied also falls by 25%. It is a straight line through the origin. Numerically, elasticity of supply is said to be equal to 1 ($E_s = 1$). (See Fig 2.18)

Fig. 2.15: Types of Price Elasticity of Supply
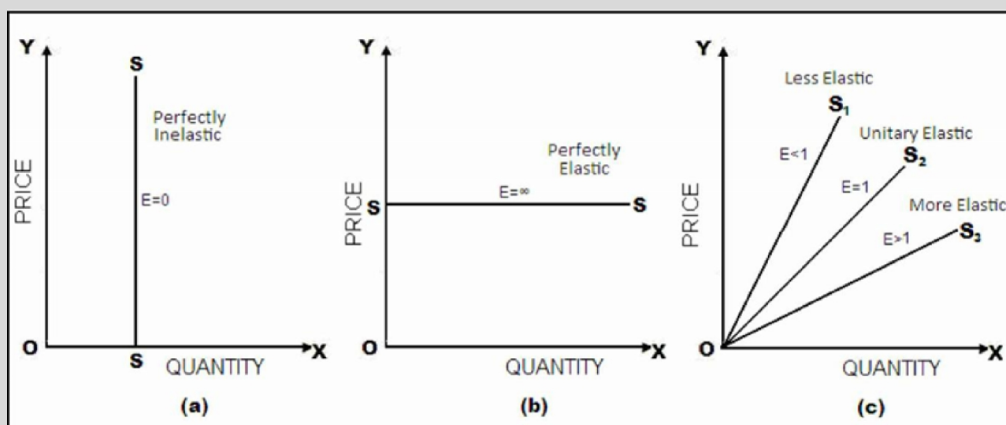
Determinants of price elasticity of supply

(a) **Time Period:** Time is the most significant factor which affects the elasticity of supply. If the price of acommodity rises and the producers have enough time to make adjustment in the level of output, the

elasticity of supply will be more elastic. If the time period is short and the supply cannot be expanded after a price increase, the supply is relatively inelastic.

(b) Ability to Store Output: The goods which can be safely stored have relatively elastic supply over the goods which are perishable and cannot be stored.

(c) Factor Mobility: If the factors of production can be easily moved from one use to another, it will affect elasticity of supply. The higher the mobility of factors, the greater is the elasticity of supply of the good and vice versa.

(d) Cost Relationships: If costs rise rapidly as output is increased, then any increase in profitability caused by a rise in the price of the good is balanced by increased costs as supply increases. If this is so, supply will be relatively inelastic. On the other hand, if costs rise slowly as output increases, supply is likely to be relatively elastic.

LESSON ROUND UP

– Demand means a desire or a wish to buy and consume a commodity or service provided consumer has adequate ability and is willing to buy.

– The consumer's decisions are guided by several elements, such as income, tastes and preferences etc. and an assumption is established that these factors remain constant.

– Law of demand states that "the amount demanded increases with a fall in price and diminishes with a rise in price".

– Demand for a good by an individual or the market as a whole is conventionally expressed in three alternative forms, namely;

  – a demand function

  – a demand schedule

  – a demand curve

– Demand for a good can change in two ways:

  – a consumer moves from one point to another on the same demand curve (Movement along demand curve)

  – when the entire demand curve shifts its position (Movement from one demand curve to the other)

– Demand Elasticity:

  – Types of demand elasticity: Price elasticity of demand; Cross price elasticity of demand and Income elasticity of demand.

  – Five cases of elasticity of demand: Perfectly elastic demand; Perfectly inelastic demand; Relatively elastic demand; Relatively inelastic demand and Unitary elastic demand

  – Methods of measuring price elasticity of demand: Percentage method or Arithmetic; Total expenditure method and Graphic method or Point method.

  – Factors affecting the elasticity of demand: Price level; Availability of substitutes; Time period; Proportion of total expenditure spent on the product; Habits; Nature of the commodities; Various uses and Postponing consumption.

– Supply means the quantity of goods offered for sale at pre determined price at a certain point of time.

– Law of Supply states that a firm will produce and offer to sell greater quantity of a product or service as the price of that product or services rises, other things being equal.

- Other things include cost of production, change of technology, price of related goods (substitutes and complements), prices of inputs, level of competition and size of industry, government policy and non economic factors.

- Supply of a good by an individual producer/firm or the market/industry as a whole is conventionally expressed in three alternative forms, namely;

    - a supply function

    - a supply schedule

    - a supply curve

- Supply for a good can change in two ways

    - a producer moves from one point to another on the same supply curve (Movement along supply curve)

    - when the entire supply curve shifts its position (Movement from one supply curve to the other)

- Supply Elasticity:

    - Five cases of elasticity of supply: Perfectly elastic supply; Perfectly inelastic supply; Relatively elastic supply; Relatively inelastic supply and Unitary elastic supply

    - Determinants of price elasticity of supply: Time period; Ability to store output; Factor mobility; Cost relationships and Excess supply

- The market price, or equilibrium price, is determined by the interaction of demand and supply at a given time with given conditions of demand and supply.

- Elasticity refers to the ratio of the relative change in a dependent to the relative change in an independent variable. There are two types of elasticities:

    - Elasticity of Demand

    - Elasticity of Supply

GLOSSARY

Ceteris Paribus	*Ceteris paribus* or *caeteris paribus* is a Latin phrase, literally translated as "with other things the same" or "all other things being equal or held constant".
Demand Curve	A demand curve is the curve showing the relationship between the quantities of a good which consumers would be willing to purchase at alternative prices.
Law of Demand	*Ceteris paribus,* the quantities demanded of a good and its own prices are inversely related.
Supply Curve	A supply curve is the curve showing relationship between the quantities supplied of a commodity by the producer at alternative prices.
Law of Supply	*Ceteris paribus,* the quantities supplied of a good and its own prices are positively related.
Price Elasticity of Demand	The relative response of a change in quantity demanded to a change in price. More specifically the price elasticity of demand is the percentage change in quantity demanded due to a percentage change in price.
Cross Elasticity	The relative response of a change in the demand for one good to a change in the

	price of another good. More specifically the cross elasticity of demand is percentage change in the demand for one good due to a percentage change in the price of another good.
Income Elasticity	The relative response of a change in demand to a change in income. More specifically the income elasticity of demand is the percentage change in demand due to a percentage change in buyers' income.
Price Elasticity of Supply	The relative response of a change in quantity supplied to a change in price. More specifically the price elasticity of supply is the percentage change in quantity supplied due to a percentage change in price.
Normal Goods	A normal good is a good that reacts positively to changes in buyers' income. If buyers have more income, then they purchase more of a normal good. If they have less income, then they reduce purchases of a normal good.
Inferior Goods	An inferior good is one that reacts negatively to changes in buyers' income. If buyers have more income, then they purchase less of an inferior good. If they have less income, then they increase purchases of an inferior good.
Giffen Goods	A consumer good for which demand rises when the price increases, and demand falls when the price decreases. Such goods are exceptions to the law of demand.
Superior Goods	Goods for which income elasticity of demand is greater than unity. For such goods the proportion of money spent on the goods tends to increase as the income increases.

MULTIPLE CHOICE QUESTIONS

1. **Supply curve represents the relationship between**

    a. Price and demand

    b. Price and supply

    c. Factors of production and price

    d. None of the above

2. **If percentage increase in demand is less than percentage increase in price, elasticity is said to be**

    a. Less than 1

    b. Equal to 1

    c. Greater than 1

    d. Equal to 0

3. **If demand curve is rectangular hyperbola, elasticity of demand is**

    a. Zero

    b. One

    c. Two

    d. Infinity

4. **Cross elasticity of demand between pen and ink is**

    a. Positive

    b. Negative

c. Zero

d. Infinity

5. **Change in demand due to increase in income is termed as**

a. Increase

b. Decrease

c. Extension

d. Contraction

6. **Market demand curve is _____ summation of individual demand curves**

a. Vertical

b. Horizontal

c. Parallel

d. None of the above

7. **A market can accurately be described as**

a. A place to buy things

b. A place to sell things

c. The process by which prices adjust to reconcile the allocation of resources

d. A place where buyers and sellers meet

8. **The equilibrium price clears the market; it is the price at which _____ _____**

a. Everything is sold

b. Buyers spend all their money

c. Quantity demanded equals quantity supplied

d. Excess demand is greater than zero

9. **If increase in price of good A increases the quantity demanded of good B, then A and B are said to be**

a. Substitute goods

b. Complementary goods

c. Neutral goods

d. None of the above

10. **A demand curve can shift because of change in**

a. Income

b. Prices of related goods

c. Taste and preferences

d. All of the above

Answer Key : 1. (b), 2. (a), 3. (b), 4. (b), 5. (a), 6. (b), 7. (d), 8. (c), 9. (a), 10. (d)

Suggested Readings

1. R.S. Pindyck, D.L. Rubinfeld and P.L. Mehta: Microeconomics

2. H. L. Ahuja :Modern Microeconomics

3. M. L. Seth:Principles of Economics

Lesson 3
Theory of Consumer Behavior

LESSON OUTLINE

– Introduction

– Utility

    – Meaning

– Cardinal Utility Approach

    – Assumptions

    – Basic Concepts

    – Law of Diminishing Marginal Utility

    – Law of Equi-Marginal Utility

    – Consumer Equilibrium

– Ordinal Utility Approach or Indifference Curve Analysis

    – Meaning of Indifference Curve

    – Indifference Map

    – Marginal Rate of Substitution

    – Assumptions

    – Properties of Indifference Curves

    – Budget line

    – Consumer Equilibrium

– Lesson Round Up

– Glossary

– Self-Test Questions

LEARNING OBJECTIVES

In this lesson, basic concepts of utility have been introduced. After developing an understanding of the basic concepts, cardinal and ordinal analysis of consumer behavior is done. The understanding of consumer behavior is incomplete without cardinal and ordinal utility analysis. Finally, it has been discussed as to what the conditions to consumer equilibrium are and how consumer attains equilibrium under the two types of utility analysis. The learning objectives of the lesson are as follows :

– To understand the concept of utility in economics.

– To learn about the cardinal and ordinal concepts of utility.

– To analyse the cardinal utility theory as well as ordinal utility theory of consumer behavior.

– To understand the concept and conditions of consumer equilibrium under the two theories.

INTRODUCTION

The theory of consumer behavior deals with the problem of constrained optimization that a consumer faces while making choices among various goods and services. In other words, the theory is concerned with the behavior of consumer, faced with resource constraints, in the market while demanding goods and services in order to maximize satisfaction. Consumer demands goods as they provide utility to him. Utility is nothing but the satisfaction derived from the consumption of goods and services. The concept of utility of a good or service is central to understand the consumer behavior because it influences the demand and price of a good or service. The different theories of consumer behaviour assume that consumer would always strive to maximize his utility. Being a rational individual, consumer wants to maximize his satisfaction given the income constraint. Therefore, it can be said that theories of consumer behavior are concerned with the question as to how a consumer should spend his limited income on different combinations of goods and services so that he may get maximum utility.

Economists have developed various theories of consumer behavior to understand the choices that people make in order to maximize their level of satisfaction. These include Cardinal Utility analysis, Indifference Curve analysis and Revealed Preference analysis. Cardinal Utility analysis is the oldest theory which was developed by neo classical economists in the 19th century. Marshall further developed the cardinal utility analysis in the 20th century. According to this analysis, utility can be measured in terms of numerical numbers. The technique of indifference curves was first originated by a classical economist Edge worth in 1881. He did not use it to explain the consumer demand analysis rather his main focus was to explain the possibilities of the exchange between two consumers. J.R. Hicks and R.G.D. Allen developed Indifference curve analysis in their well-known paper 'A Reconsideration of the Theory of Value'. They had criticized the Marshallian cardinal approach of utility and used the notion of ordinal utility approach to understand the consumer behavior. According to Indifference Curve Analysis, utility is ordinal as consumer can only rank or order the satisfaction he derives from different combinations of goods. In this lesson, the discussion is limited to cardinal and ordinal utility analysis.

UTILITY

MEANING OF UTILITY

Utility is want satisfying power of a good or service. It is also defined as the property of a good or service to satisfy the want of the consumer. Utility is subjective. It depends upon the mental assessment of the consumer and is determined by several factors which influence the consumer's judgment. These factors include, for example, the intensity of the want(s) to be satisfied. Utility of a good varies with the intensity of the want to be satisfied by its consumption. Thus, the satisfaction derived from same set of goods and services is different for different consumers. Alternatively, same set of goods give different satisfaction to different consumers. For instance, intake of a cup of tea may derive more utility to person A as compared to person B. This fact leads to a few important inferences:

– Utility of a good differs from consumer to consumer. This is because a given want can be felt in different intensities by different consumers.

– The utility of a good keeps changing even for the same consumer on account of changes in the intensity of the want(s) to be satisfied by its use. This change may be the result of a shift in the circumstances faced by the consumer, or it may take place in the process of the satisfaction of the want itself.

– The utility of a good is not to be equated with its usefulness. Satisfaction of a want need not add to the welfare of the consumer. For example, smoking, drug taking or consumption of similar other things are considered to be harmful to the health of the consumer. But the consumer may believe that they have

utility for him because he can use these to satisfy his wants.

In economics, we are not concerned with the 'normative' aspect of utility. It does not matter whether a good's consumption adds to the individual's well-being or not. So long as the consumers expect to derive some 'satisfaction' from a good (that is, so long the good has a 'utility' for them), they will be ready to buy it at some price and create a demand for it in the market.

A question here arises as to why it is important to define and measure utility? The measurement of utility helps us in understanding the demand behavior of individual consumers, and therefore, the market as a whole. The basis of the reasoning is that a consumer compares utility of a good with the price he has to pay for it. A consumer purchases additional units of same commodity so long as the utility from them is at least equal to the price to be paid for them. To understand the concept and its measurement further, we are discussing the following two economic approaches in relation to utility:

- – Cardinal Utility
- – Ordinal Utility

CARDINAL UTILITY APPROACH

Cardinal Utility approach to consumer behavior states that utility can be measured in cardinal numbers or definite numbers such as 1, 2, 3, 4, etc. Cardinal numbers are those definite numbers which can be subtracted or added. Fisher used the 'Util' as a measure of utility. In cardinal measurement, utility is expressed in absolute standard units, such as there being 20 units (utils) of utility from the first glass of water and 11 units from the second.

ASSUMPTIONS

Cardinal Utility analysis is based on certain assumptions. These are as follows:

- **Rational Consumer**: consumer is believed to be rational and his aim is to maximize utility subject to the income constraint.

- **Utility is Cardinally Measurable**: the utility is measurable and quantifiable in definite numbers. For example, a consumer can state that he obtains utility equal to 20 utils from the consumption of good X and 10 utils from the consumption of good Y.

- **Independent Utilities:** according to this school of thought, utility which a consumer derives from the consumption of a good completely depends upon the quantity of that commodity alone. It implies that the utility which a consumer gets from one commodity is independent of the consumption of other commodities.

- **Additive Utility:** the utility derived from the consumption of each commodity can be added to arrive at total utility derived by the consumer by consuming all the products purchased with his limited income.

- **Constant Marginal Utility of Money:** it is assumed that change in consumer's income do not lead to change in marginal utility of money for him. In other words, marginal utility of money does not vary with the amount of money an individual holds.

- **Diminishing Marginal Utility:** it is believed that the marginal utility gained by the consumer from additional units of commodity diminishes as quantity of consumption of the same commodity increases.

BASIC CONCEPTS

As discussed before, the consumer derives utility from the good he consumes and utility derived from the consumption of a good varies with the quantity of good consumed. It generates three concepts:

- Total Utility (TU)

- Average Utility (AU)

- Marginal Utility (MU)

Total Utility: it is the aggregate level of satisfaction obtained by the consumer from the consumption of all the different units of a commodity. If a consumer buys *n* units of good X then, for him, total utility from it is the summation of utilities derived from all the *n* units. Symbolically,

$TU_n = U_1 + U_2 + U_3 + \dots + U_n$

Or

$TU_n = \sum_{i=1}^{n} Ui$ and, i = 1, 2, 3, ... n

Where

TU_n = Total utility obtained from the consumption of *n* units of good X

U_i stands for the utility of ith unit of good X

Average Utility: it is derived by dividing the total utility obtained by a consumer with the number of units of commodity X consumed. Symbolically,

$$Average\ Utility = \frac{TUx}{Nx}$$

Where

TU_x = Total utility obtained from the consumption of good X

N_x = Number of units of good X consumed

Marginal Utility: Marginal utility is the additional utility derived from the consumption of an additional unit of commodity. Alternatively, it is the change that takes place in total utility by the consumption of an additional unit of a commodity. For example, the additional satisfaction that consumer gains from consuming n^{th} unit of 'X' is Marginal utility (MU) of that n^{th} unit of X. Marginal utility can be positive, zero or negative. Symbollically,

$$MU_n = TU_n - TU_{n-1}$$

Or

$$MU = \frac{\Delta TU}{\Delta Q}$$

Here

MU_n = Marginal utility obtained from n^{th} unit of good X

TU_n = Total Utility obtained from all the n units of good X

TU_{n-1} = Total Utility obtained from n-1 units of good X

ΔTU = Change in total Utility

ΔQ = Change in the quantity of commodity

Positive Marginal Utility: when the consumption of an additional unit of a commodity increases total utility then marginal utility of that unit will be positive.

Zero Marginal Utility: when the consumption of an additional unit of a good leads to no change in the total utility then marginal utility of that unit is zero.

Negative Marginal Utility: when the consumption of an additional unit of a commodity leads to decline in total utility, then the marginal utility of that unit is negative.

Laws of Cardinal Utility Analysis

Founders of cardinal utility analysis have developed two basic laws which are critical and have several economic implications. These two laws are as follows:

– Law of Diminishing Marginal Utility

– Law of Equi-Marginal Utility

LAW OF DIMINISHING MARGINAL UTILITY

The law of diminishing marginal utility is the foundation stone of utility analysis. Every individual experiences this phenomenon in his/her daily life. According to the law of diminishing marginal utility, other things being equal, marginal utility of a good starts falling as an individual consumes more units of it in a given period of time. Alternatively, as a consumer consumes more and more units of a commodity, the extra satisfaction that he derives from successive unit of a commodity goes on diminishing. Marginal utility eventually falls to zero and then becomes even negative. The law describes a familiar psychological tendency of the human beings. In the words of Marshal, "the additional benefit which a person derives from a given increase in his stock of a thing diminishes with every increase in the stock that he already has."

The law of diminishing marginal utility follows from the conventional (and realistic) assumption that the intensity of a given want keep on decreasing if the process of its satisfaction is continued without interruption, that is, a single want can be fully satisfied provided the consumer consumes a large enough quantity of the relevant good/service. Further, it is also assumed that the good to be consumed should be homogeneous. Its successive units should have the same technical specifications. Any change in them can cause a change in the intensity of the want being satisfied and thereby violate the law of diminishing marginal utility.

Explanation of the Law

The law of diminishing marginal utility is based on strict assumptions such as: (i) all units of the commodity are completely homogenous. (ii) The consumption of the successive units should be without any long time interval. (iii) Utility is psychological concept and it varies from person to person, but the law assumes that the utility can be measured cardinally such as in numbers 1, 2, 3 .... (iv). The consumer is rational human being and he strives for maximum satisfaction.

The law of diminishing marginal utility is violated if one or more of the assumptions upon which it is based get violated. Since utility of a good is related to the mental perception of the consumer regarding the intensity of the want to be satisfied and the capacity of the good to satisfy it, therefore, the law of diminishing marginal utility is violated if for some reason,

– the intensity of the want increases, or

– the consumer comes to think that the intensity of his want has increased.

It is for this reason that marginal utility of a good tends to increase if there is an unduly long interval between the consumption of two units of a good. Marginal utility of a good may also increase, if want of the consumer is intensified by consuming a very small quantity of it (such as, a very little quantity of water given to a very thirsty person).

Illustration

The verification of theoretical understanding is done through quantitative calculation as presented in Table 3.1.

Table 3.1: Diminishing Marginal Utility

No. of Units of Commodity X	Total Utility(TU)	Average Utility (AU)	Marginal Utility (MU)
1	14	14.0	14
2	22	11.0	8
3	28	9.3	6
4	32	8.0	4
5	34	6.8	2
6	34	5.7	0
7	32	4.6	-2
8	28	3.5	-4

To understand the relationship between diminishing marginal utility and other concepts of utility i.e. total utility and average utility, a hypothetical case of good X is presented in Table 3.1. It is important to note that the consumer is assumed to consume good X without any unreasonable time gap between the intakes of successive units of good X. This assumption is critical to understand that the intensity of hunger of the consumer diminishes as he consumes additional quantity of good X.

- As more and more units of good X are consumed, the marginal utility derived from each successive unit goes on diminishing, but total utility continues to increase at a diminishing rate as long as marginal utility is positive and greater than zero. For instance, as consumer increased the consumption of good X from unit 1 to unit 2, total utility increased from 14 to 22 whereas marginal utility declined from 14 to 8 (Table, 3.1). The total utility continued to increase until consumption of good X reached to 5 units. Correspondingly, upto the consumption of 5 units, marginal utility remained positive and greater than zero.

- When marginal utility from an additional unit declines to zero, total utility reaches its maximum. After this, marginal utility from the consumption of additional unit becomes negative. Correspondingly, total utility starts declining. For instance, marginal utility derived from the consumption of 6th unit of good X is zero. Correspondingly, total utility (34) reached at its maximum. Further, marginal utility derived from the consumption of 7th unit became (-) 2. As marginal utility turned negative total utility declined from 34 to 32.

- Average utility always remained positive and greater than zero. For instance, average utility declined from 14 to 3.5 as consumer increased its consumption from 1 unit to 8 units.

Graphical Presentation

Fig. 3.1: Relationship between Diminishing Marginal Utility and Total Utility
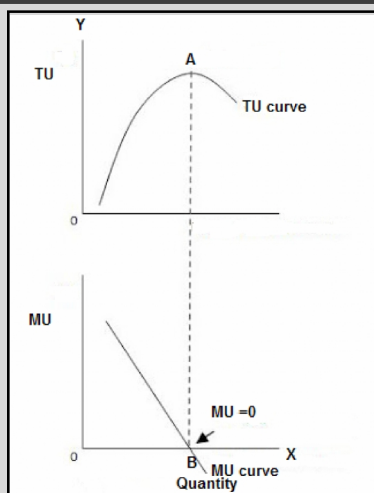
Figure 3.1 presents the relationship between diminishing marginal utility and total utility graphically. Units of the commodity are shown on OX-axis and utility on OY-axis.

– MU curve is downward sloping. The downward slope of MU curve implies that MU derived from each additional unit goes on declining.

– Between points O and B, marginal utility remains positive and greater than zero. Correspondingly, total utility is increasing until consumer consumed OB amount of good X.

– As MU becomes zero (corresponding to point B), total utility curve reached its maximum point which is shown as point A in figure 3.1. As total utility curve reaches its maximum point, it is a point of saturation. A is a point where total utility becomes constant. Alternatively, increase in the consumption of good X does not lead to any change in the total utility. After saturation, total utility curve starts declining.

– As MU derived from the consumption of additional unit becomes negative, TU curve starts declining.

In nutshell, MU curve slopes downwards. TU curve rises at a diminishing rate. It reaches its maximum distance from X-axis when MU is zero. Thereafter, it also slopes downwards, when MU is negative.

Exceptions to the law of diminishing marginal utility are as follows:

(i) Curious and rare things: the law of diminishing marginal utility does not apply in case of the hobby of an individual. For instance, the hobby of an individual is to collect old and rare coins, rare portraits etc. Marginal utility of such individual increases as stock of such curious and rare articles goes on increasing.

(ii) Misers: the law does not apply to misers who always wish to acquire more and more of wealth. Their desire for money seems to be insatiable.

LAW OF EQUI-MARGINAL UTILITY

Law of equi-marginal utility facilitates in explaining the demand behavior of a consumer and determination of his equilibrium when he faces the following situation. A consumer has a limited given income which he has to spend on various goods he wants. Now, how should a consumer spend his fixed income in purchasing between various commodities so as to maximize his total satisfaction? The law of equi--marginal utility tells us the way how a consumer maximizes his total utility while spending on various commodities.

According to the law of equi-marginal utility, a rational individual does not spend his entire income on one commodity. Consumer knows that if he purchases more units of the same commodity then marginal utility of each successive unit would decline. Hence, to get the maximum utility from his limited income, he spends his limited income in such a way that the last unit of money spent on various commodities yields equal marginal utility to him.

Assumptions

– Income of the consumer is given and remains constant.

– Each commodity holds the law of diminishing marginal utility, and its marginal utility schedule is known.

– Prices of the commodity remain constant.

– Commodity is divisible into small units. Thus, consumer can spend his income in small units of money.

– Tastes and preferences of the consumer remain constant.

It is assumed that the consumer decides to divide his total expenditure between different goods by taking into consideration not only their respective marginal utilities but also their per unit prices. A consumer is guided by marginal utility which he can derive by spending each additional rupee. It is on this basis that he decides to allocate his expenditure between alternative goods.

Case of two goods: Let us assume that a consumer purchases two goods X and Y on which he has to spend his entire income. He would be in equilibrium when he will spend his income among two goods (X and Y) in such a manner that the utility derived from the last unit of money spend on each good is equal. Symbolically,

$$\frac{MUx}{Px} = \frac{MUy}{Py}$$

MU_x= Marginal Utility of good X,

MU_y= Marginal Utility of good Y,

P_x= Price of good X,

P_y= Price of good Y.

- The marginal utility derived from a good should not be less than the price paid for it. That is, the ratio of

  $\frac{MUx}{Px}$ must be equal to 1 $(or\ MUx = Px)$. Similarly, the ratio of $\frac{MUy}{Py}$ must be equal to 1. Finally,

  $$\frac{MUx}{Px} = \frac{MUy}{Py}$$

- If MU_x/P_x and MU_y/P_y are not equal - Suppose, MU_x/P_x is greater than MU_y/P_y. In such a situation, consumer will substitute good x for good Y because good X is giving him more utility than good Y. As he substitutes good X for good Y, the quantity of good X will increase and quantity of good Y will decline. Accordingly, the marginal utility derived from good X will fall whereas marginal utility derived from good X will rise. The consumer will continue substituting good X for good Y until MU_x/P_x becomes equal to MU_y/P_y when the further reallocation will not increase his total utility. This will be an equilibrium situation as consumer will be deriving maximum satisfaction out of his fixed income.

Fig. 3.2: Equi-Marginal Principle in case of Two Goods
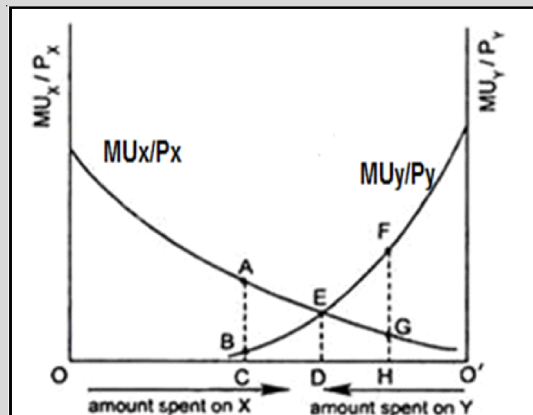
Figure 3.2 presents the operation of law of equi-marginal utility in case of two goods. As consumer moves right wards from O, amount spent on X will increase and, when he moves left wards from O', amount spent on Y will increase. How will consumer distribute his entire income to purchase combination of two goods (X and Y) so that he maximizes his utility form the given income?

Initially, consumer spent O'C on good Y and OC on good X. In that case, MU_x/P_x will be higher than MU_y/ P_y by the distance AB. In such a situation, consumer will increase spending on X and reduce on Y. Accordingly, quantity of X will increase and Y will reduce. The increase in quantity of X will reduce the MU derived from X and fall in quantity of Y will increase the marginal utility derived from Y. He will substitute X for Y until equality is

restored at point E. Similarly, when he spent OH on X and O'H on Y, then MU_X/P_X was lower than MU_Y/P_Y. In that case he will purchase more of Y and less of X unit he will be able to achieve equality, that is, MUx/Px = MUy/Py.

To sum, consumer maximizes his total utility by spending O'D amount on good Y and OD amount on good X. By purchasing this combination, the consumer equalizes marginal utilities of last rupee spent on X and Y at point E (i.e., $MU_X/P_X = MU_Y/P_Y = ED$). No other combination will give greater total utility.

Case of more than two goods: In the real world, a consumer may purchase more than two commodities. Suppose there are more than two goods on which the consumer is spending his income. In such a situation, he will maximize his utility when consumer distributes his expenditure between different goods in such a way that the marginal utility derived from the last rupee spent on each good is the same. In case of more than two goods (say n goods), the condition for a consumer to maximize utility is usually written in the following form:

By equating the ratios of marginal utilities to prices of goods, the consumer succeeds in deriving maximum possible utility from his expenditure. This is the best position which he can attain.

$$\frac{MUx}{Px} = \frac{MUy}{Py} = ----\frac{MUn}{Pn}$$

Illustration

The law of equi-marginal utility can be explained in terms of an arithmetical example. Table 3.2 presents an hypothetical example of the law of equi-marginal utility. It is assumed that our consumer is to spend 12 rupees and choose between four goods, A, B, C and D. Figures in the first row reveal that the first rupee spent on good A yields 40 units of utility for the consumer. If same rupee is spent on good B, the utility derived by the consumer is 38 units and so on.

Recalling that the consumer will spend each additional rupee on that good which brings him maximum marginal utility (that is having highest MU/P), we note that he will spend his 1st rupee on good D which brings him 45 units of utility. Similarly, the 2nd rupee is spent on good C (which brings 44 units of utility) ; the 3rd rupee is again spent on D (with marginal utility of 42); the 4th and 5th rupees are spent on goods A and C (not necessarily in this order); the 6th rupee goes to good D; 7th and 8th rupees are spent on goods A and B (not necessarily in this order); while the remaining four rupees are spent one each on A, B, C and D (not necessarily in this order). As a result, in all, he spends three rupees on A, two rupees on B, three rupees on C, and four rupees on D. The utility derived by him is 114 units from A, 74 units from B, 120 units from C, and 162 units from D, the total being 470 units. Any other division of his expenditure on these four goods would yield the consumer a smaller total utility. It should also be noted that when marginal utility from a rupee spent on two or more goods is the same, the consumer may spend it on either of them. Thus, in our example, we cannot say for certain whether the consumer will spend 4th rupee on A and 5th on C, or it will be the other way round. And if his total expenditure is only five rupees, the 5th rupee may be spent on either of the two goods with the same result.

Table 3.2: Illustration of the Law of Equi-Marginal Utility

Expenditure (in Rupees)	MU_a/P_a	MU_b/P_b	MU_c/P_c	MU_d/P_d
1	40	38	44	45
2	38	36	40	42
3	36	32	36	39
4	34	29	32	36
5	32	26	28	33

6	20	23	24	30
7	18	20	20	27
8	16	17	16	24
9	14	14	12	21

CONSUMER EQUILIBRIUM UNDER UTILITY ANALYSIS

Equilibrium is a situation which is characterized by the absence of any tendency to change. A consumer is in equilibrium when he has no tendency to reallocate his money expenditure. In other words consumer equilibrium refers to a situation wherein a consumer gets maximum satisfaction from the purchases of goods at given prices and given income. Any deviation from this point places the consumer in the sub-optimal situation.

Single Commodity case: a consumer with a single commodity attains equilibrium at a point when the marginal utility derived from that good is equated to its price.

– If marginal utility derived from an additional unit of the same good is higher than its price then consumer will purchase additional units of that commodity until he achieves equality, that is, MU=Price.

– On the other hand, if MU from the purchase of additional unit is less than the price then he will reduce the purchase of that commodity up to the point where MU derived from the purchase of additional unit of commodity is equal to its price.

Two or more commodity case: If there are two or more commodities, the condition for the equilibrium of the consumer is the equality of the ratios of the marginal utilities of the individual commodities to their prices. In case of n commodities, equilibrium condition can be written as follows :

$$\frac{MUx}{Px} = \frac{MUy}{Py} = ---- \frac{MUn}{Pn}$$

Application of the Cardinal Utility Approach

The law of cardinal utility analysis provide the basis to law of demand and the concept of consumer surplus.

– Law of Demand: According to the law of demand, there is an inverse relationship between quantity demanded of a commodity and its price. Consumer demands more quantity of same commodity as its price falls. It is so because more units yield diminishing marginal utility. Thus, consumer will purchase additional unit only when price of the unit is not higher than the MU he would get from additional purchase. Figure 3.3 shows the derivation of demand curve from marginal utility curve. Part A of figure 3.3 indicates that the marginal utility diminishes from the consumption of additional units. Correspondingly, the price of additional purchases also declines. From first unit of consumption of good X, MU_1 was the marginal utility. Consumer was ready to pay P_{-1} price. For the purchase of additional units of good X, he was not ready to pay P_{-1} price because his marginal utility from additional unit does not remain same. Accordingly, consumer purchased additional units at lower prices (Figure, 3.3). It shows that price of a commodity is determined by the MU of the commodity.

Fig. 3.3: Derivation of Demand curve from MU Curve
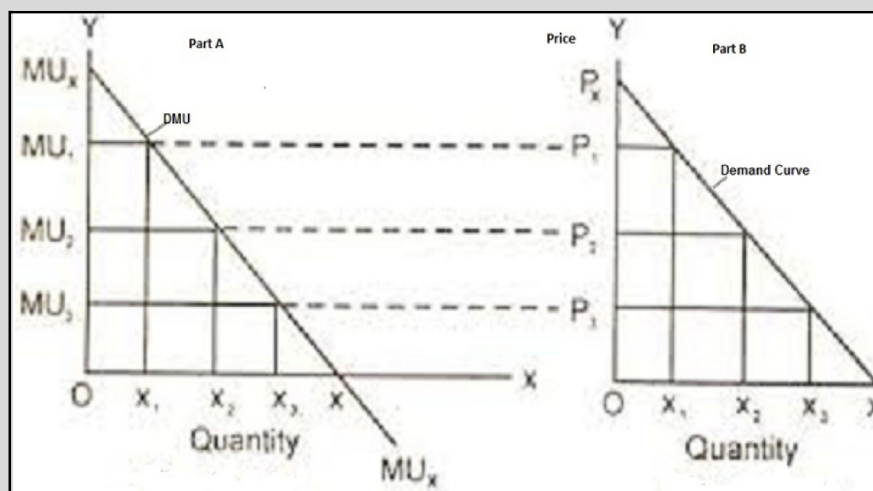
– The concept of consumer surplus is also based on law of diminishing marginal utility.

– Difference between the value-in use and value in exchange is also explained with the help of law of diminishing marginal utility.

– The law of diminishing utility provides basis for the progressive taxation.

Limitations

In reality, the cardinal utility approach suffers from several limitations which come in the way of its implementation by the consumer. Some of the limitations are as follows:

(a) The assumption that the goods on which the consumer spends his money are perfectly divisible, i.e., goods can be bought even in extremely small quantities does not hold true at times. The consumer is faced with lumpy goods. They are not divisible into very small quantities. He has to buy an entire quantity of a good or not at all. This is more so in the case of durable consumption goods. For example, he cannot buy half of a shirt, one-tenth of a bicycle. Consequently, he fails to apply the law of equi-marginal utility in practice.

(b) The law assumes independence of utility schedules of goods. It means that utility derived from one good is not affected by the quantity purchased of other goods. In reality, however, many goods are related to each other by being substitutes or complements to each other. In such cases, the marginal utility derived from a given good depends not only upon its own quantity, but also upon the quantity of the related good.

(c) Another shortcoming of the cardinal utility approach is related of constant utility of money which is completely unrealistic. As the stock of money increases to an individual, the marginal utility of money also changes.

(d) The law makes a questionable assumption that the consumer is able to accurately determine the marginal utility schedules of all the goods.

ORDINAL UTILITY APPROACH OR INDIFFERENCE CURVE ANALYSIS

Cardinal utility analysis suffers from various drawbacks, the major being the assumption of cardinal measurement of utility. The indifference curve analysis is advancement to the utility analysis as it tries to overcome the drawbacks of cardinal utility analysis and provides a technically superior analysis of demand or consumer behavior. The technique of indifference curve was first originated by classical economist Edgeworth in 1881. He did not use it to explain the consumer demand analysis rather his focus was to explain the possibilities of the exchange

between two consumers. J.R. Hicks and R.G.D. Allen. Developed Indifference curve analysis in their well-known paper 'A Reconsideration of the Theory of Value'. They had criticized the Marshallian cardinal approach of utility and used the notion of ordinal utility approach to understand the consumer behavior. According to the Ordinal Utility approach, consumer may not be able to indicate the exact amounts of utilities that he derives from commodities or any combination of them. But he is capable of judging whether the satisfaction obtained from a good or a combination of goods is equal to, lower than or higher than the other.

In indifference curve approach, the preferences are ordered or ranked in relation to one another. This approach, is, therefore an ordinal concept based on ordering of preferences compared with Marshall's approach of cardinality. To understand the indifference curve analysis, following components need to be discussed in detail. This approach to consumer behaviour is best understood in three distinct steps:

– Consumer Preferences (Indifference curve)

– Budget Constraints (Budget Line)

– Consumer Choices (Equilibrium)

MEANING OF INDIFFERENCE CURVE

An indifference curve is a locus of various combinations of two commodities that give equal satisfaction to the consumer. Since all the combinations on an indifference curve yield equal satisfaction to the consumer, the consumer is indifferent among them. Alternatively, a consumer is indifferent towards the different combinations located on indifference curve since each combination yields the same level of satisfaction, the total satisfaction or utility derived from any of these combinations remains constant. Symbolically,

$$U_0 = f(q_1, q_2 \ldots)$$

Where,

U_0 = total utility that remains constant;

q_1, q_2, and so on represent different combinations of two goods.

Graphical representation:

In the figure, there are two commodities X and Y, quantities of commodity X are measured on horizontal axis and commodity Y on vertical axis. IC curve represents an indifference curve showing various combinations of commodity X and Y, such as combination A and B that yield equal total satisfaction to the consumer. Combination A includes X1 units of X and Y1 units of Y. Similarly, combination B includes X2 units of X and Y2 units of Y. Combination A and B denote different quantities of X and Y. If consumer shifts from combination A to B, amount of commodity X increases while quantity of Y decreases. By increasing the amount of X while simultaneously decreasing the amount of Y, the consumer is able to keep the total utility constant.

Fig. 3.4: Indifference Curve
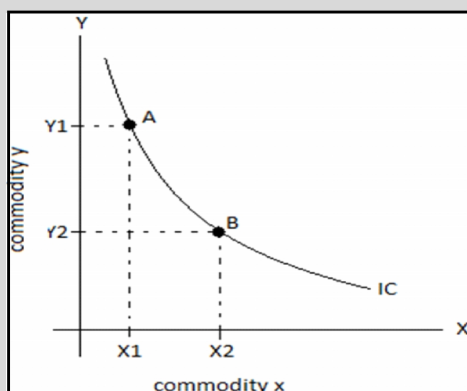
INDIFFERENCE MAP

An Indifference map refers to a collection of indifference curves corresponding to different levels of satisfaction. In other words, a set of indifference curves is known as an indifference map.

Fig. 3.5: Indifference Map
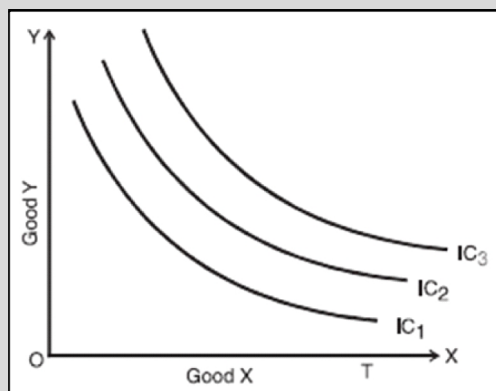
In the figure, an indifference map is given. There are three different indifference curves namely, IC_1, IC_2 and IC_3 in this map and each one correspond to different level of satisfaction. For instance, IC_1 represents lowest indifference curve corresponding to lowest level of satisfaction in this map. Similarly, IC_3 represents highest level of satisfaction in this map.

It is to be noted that on each indifference curve, the consumer would be indifferent towards various combinations but the indifference does not hold in case of different indifference curves. For example, a consumer will receive same satisfaction from all the combinations on IC_1 so he would be indifferent towards any of these combinations but if we compare IC_1 and IC_2 the consumer would not be indifferent as all the combinations on IC_2 yield higher satisfaction as compared to IC_1 and therefore he would prefer any combination on IC_2 as compared to IC_1. Further, all the combinations of IC_3 yield higher satisfaction as compared to any combination on IC_2 so the consumer would prefer to be on IC_3 than IC_2.

Thus, the consumer would always prefer a higher indifference curve since the higher indifference curve always represents greater quantities of both the commodities and hence higher level of satisfaction. Contrarily, a lower indifference curve always represents less quantity of both the goods and hence a lower level of satisfaction.

MARGINAL RATE OF SUBSTITUTION

In case of an indifference curve, a consumer has to always sacrifice some amount of one good, say good X to increase the consumption of the other good, say good Y, in order to keep his total utility constant. The rate at which the consumer is prepared to exchange one good for the other is known as Marginal Rate of Substitution (MRS). Alternatively, MRS measures the opportunity cost of obtaining one good in terms of giving up the other good. For example, MRS_{XY} (Marginal rate of Substitution of X for Y) refers to the amount of Y that the consumer is willing to give up per additional unit of X to maintain the same level of satisfaction i.e. to remain on the same indifference curve.

That is,

$MRS_{XY} = -\triangle Y/\triangle X$

MRS_{XY} is negative because of the reduction in Y that is required to maintain the same level of satisfaction while increasing the amount of X.

This is further illustrated with the help of the table below where a case of two goods, X and Y, has been taken, and the consumer is substituting X for Y.

Table 3.3: Marginal Rate of Substitution

Combination	Good X	Good Y	MRS_{xy}
A	1	40	-
B	2	32	1:8
C	3	26	1:6
D	4	22	1:4
E	5	20	1:2
F	6	19	1:1

Table 3.3 represents six combinations of good X and good Y. These combinations provide an idea on the amount which the consumer is ready to substitute of good Y to get an additional unit of good X. In combination A, the consumer includes 1 unit of good X and 40 units of good Y. When he increases the consumption of X from 1 to 2 units, he sacrifices 8 units of good Y, that is, marginal rate of substitution of X for Y is 1:8. As he increases the consumption of good X from 2 units to 3 units, the amount of good Y which consumer is willing to sacrifice to get an additional unit of good X is 6 units, that is, marginal rate of substitution of good X for Y is 1:6. Further, when consumer wants to increase consumption of good X from 3 units to 4 units, the consumer is willing to give-up only 4 units of good Y this time to get an additional unit of good X, that is, marginal rate of substitution of good X for Y is 1:4. The decline in the marginal rate of substitution of good X for good Y facilitates us to draw the following conclusion-As the amount of good X in consumption basket increases, the marginal utility derived by the consumer from an additional unit of it starts declining. On the other hand, as the amount of good Y in consumption basket declines, the marginal utility from additional sacrificed unit starts increasing.

Fig. 3.6: Marginal Rate of Substitution of Good X for Good Y
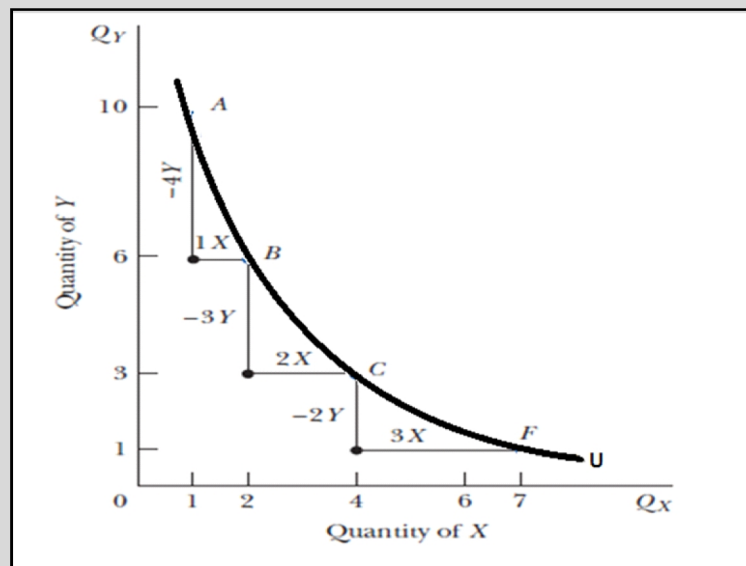
To further elaborate the point made above, let us discuss Figure 3.6. The figure shows that as consumer starts substituting X for Y, first he sacrifices 4 units of Y and get only 1 additional unit of X, then he sacrifices 3 units of Y and get 2 additional units of Y and further he sacrifices just 2 units of Y and get 3 additional units of X. It shows the decline in the marginal rate of substitution of good X for good Y. It is interesting to note that the slope of indifference curve declines as we move down the indifference curve. Contrarily, the slope of indifference curve increases as we move up. It follows from the property of convexity of the indifference curve. As consume move down, the amount of good Y reduces with him and amount of good X increases. Based on the principle of diminishing marginal utility, the increase in quantity of X reduces his utility from an additional unit of X. Fall in quantity of Y increases the utility from additional unit of good Y. Thus, consumer is willing to sacrifice less and less amount of Y to get an additional unit of X.

ASSUMPTIONS

1. **Rationality**: The consumer is assumed to be rational. Being a rational consumer, he aims to maximize his utility with his fixed limited income and at given market prices. Consumer is assumed to have full knowledge of all the relevant information.

2. **Ordinal Utility**: It is taken as axiomatically true that the consumer can rank his preferences (order the various 'baskets of goods') according to satisfaction of each basket. He need not know precisely the amount of satisfaction. It suffices that he expresses his preference for the various bundles of commodities.

3. **Diminishing marginal rate of substitution**: The slope of the indifference curve also called the marginal rate of substitution is declining. It implies that consumer sacrifice less and less amount of one good for each additional unit of another good while substituting between the two.

4. The total amount of utility of the consumer depends on the quantities of the commodities consumed:

   $U = f(q_1, q_2, ... q_n)$

5. **Consistency and transitivity of choice**: it is assumed that the consumer is consistent in his choice, that is, if in one period he chooses bundle A over B, he will not choose B over A in another period if both bundles are available to him. The consistency assumption may be symbolically written as follows:

If A>B, then B≯A

Similarly, it is assumed that consumer's choices are characterized by transitivity: if bundle A is preferred to B, and B is preferred to C, then bundle A, is preferred to C. Symbolically, it can be written as:

If A>B, and B>C, then A>C

PROPERTIES OF INDIFFERENCE CURVES

The following are the main characteristics or properties of indifference curves:

(a) **Indifference curves slope downward from left to the right:** by definition, indifference curve is a curve on which all the combinations of two goods yield same level of satisfaction. This property implies that if a consumer wants to have more quantity of good one, he has to give up some quantity of another good in order to maintain the same level of satisfaction.

(b) **Indifference curves are always convex to the origin:** it implies that the indifference curve is relatively flatter in its right-hand portion and relatively steeper in its left-hand portion. Only a convex indifference curve follows the principle of diminishing marginal rate of substitution. It implies that as more and more of one commodity is substituted for another, the consumer is willing to give up less and less of the commodity being substituted. This is called the diminishing marginal rate of substitution. If indifference curve be of any other shape such as concave to the origin, or straight line, then it does not follow the property of diminishing marginal rate of substitution.

• Straight line indifference curve shows prefect substitution between two goods. Thus, straight line indifference curve implies that MRS_{xy} remains constant as more units of X are acquired in place of Y.

Fig. 3.7: Indifference Curves – Case of Perfect Substitutes
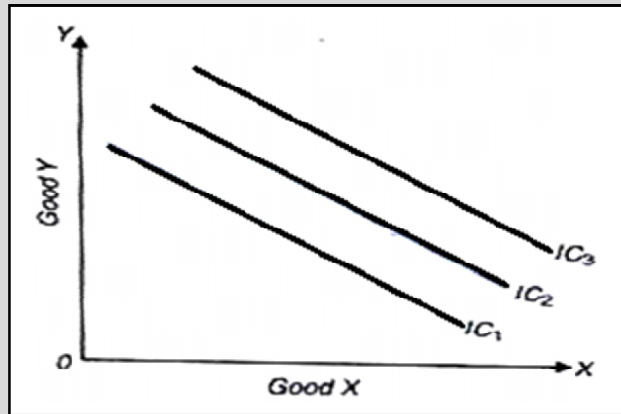
• **When indifference curve is concave to the origin:** it implies that the marginal rate of substitution of X for Y increases as the consumer obtains more of X in place of Y. Thus, it does not hold the assumption of diminishing marginal rate of substitution which is the basis premise of the utility analysis. Following figure present the case of concave indifference curve.

Fig. 3.8: Concave Indifference Curve
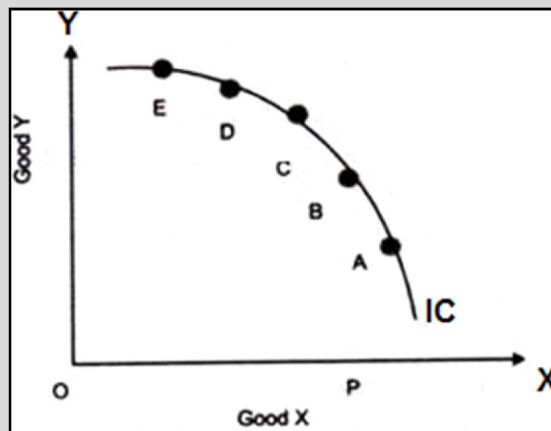
(c) A higher indifference curve represents a higher level of satisfaction than the lower indifference curve: This is because combinations lying on a higher indifference curve contain more of either one or both goods and more goods are preferred to less.

(d) Indifference curves can never intersect each other: No two indifference curves will intersect each other although it is not necessary that they are parallel to each other. In case of intersection the relationship becomes logically absurd because it would show that higher and lower levels are equal which is not possible. For example, in Fig. 3.5, IC_1 represent lower indifference curve and IC_2 represents higher indifference curve. By assumption, all the combinations on the higher indifference curve contain high quantity of both the goods. Accordingly, higher indifference curve yields higher level of satisfaction as

compared to lower indifference curve. Further, all the combinations on the same indifference curve yields same level of satisfactions.

- In figure below, two indifference curves are shown cutting each other at point A. By definition, all the combination on the same indifference curve yield equal level of satisfaction. Hence, points C and A yield equal satisfaction to the consumer as both are located on IC_2. Similarly, combinations B and A yield equal level of satisfaction to the consumer as both are located on IC_1. If combination C is equal to combination A in terms of level of satisfaction and combination B equals to combination A, then combination C would be equivalent to B in terms of level of satisfaction. This violates the basic premise of indifference curve since combination C contains more of commodity Y and commodity X than combination B. Being a rational consumer, he will prefer combination C to B as combination C will give more satisfaction to the consumer than B. But the intersection of two indifference curves lead us to an abrupt conclusion of A being equal to B in terms of level of satisfaction. Thus, it is clear that indifference curves cannot cut each other.

Fig. 3.9: Indifference Curve Intersection
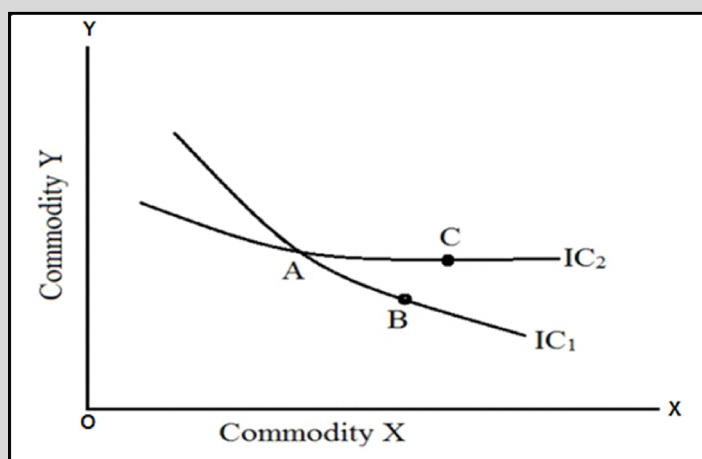
Budget Constraint

Budget constraint is the constraint that consumer faces as a result of limited incomes. Because of limited income, a consumer has to restrict his purchase and also prioritize his unlimited choices.

BUDGET LINE

It shows all those combinations of two goods which consumer can buy spending his entire money income on the two goods at their given prices. Symbolically, the budget line can be written algebraically as follow:

$$P_xX + P_yY = M$$

Where

Px = price of good X

Py = price of good Y

X = good X

Y = good Y

M = income of the consumer

Fig. 3.10: Budget line
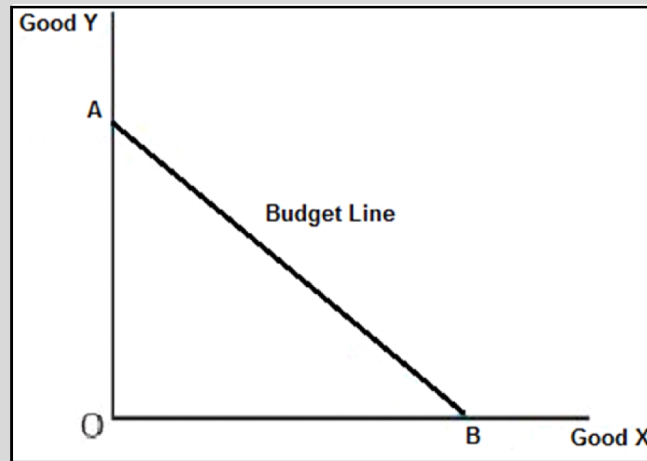
In this diagram, two-commodity case is discussed. The consumer spends his entire income on these two goods. With his entire income, the consumer can either purchase OA amount of Good Y or OB amount of Good X or he can purchase any combination of good X and Y that lies on Budget Line as shown by a linear downward sloping curve AB in the figure.

Using the budget line equation and budget line as shown in the figure there can be three cases.

– First, if consumer spends his entire income on good X and Y, he will select any combination lying on the budget line AB.

– Second, if he does not spend his entire income on these goods, he will select any combination below the budget line AB.

– Third, the consumer, with his given income, is unable to select any combination lying above the budget line AB. All the combinations lying above the budget line AB denote greater amount of either both the goods or atleast one good. These combinations can only be achieved if income of the consumer increases or prices of the goods fall which will cause increase in real income of the consumer.

Fig. 3.11: Budget Line with Increase in Income
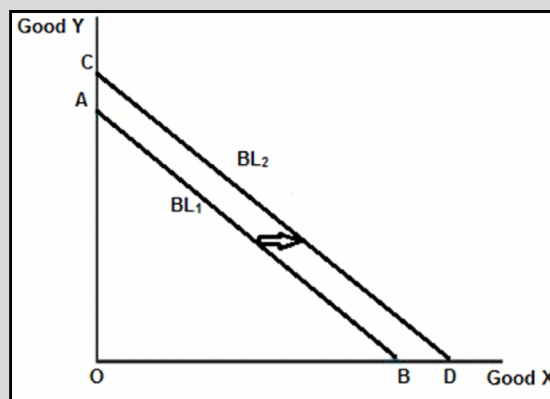
There are two budget lines in this figure, AB and CD. When budget line is AB, the consumer can select any combination in the area denoted by OAB. Now, suppose, consumer's income increases and as a result of this, budget line moves towards right from AB to CD. The consumption possibilities with this new budget line increases

since now the consumer can select any combination in the area denoted by OCD. The same result can also be achieved in the case when income of the consumer does not change but prices of both the goods fall proportionately causing a parallel outward shift in the budget line.

CONSUMER EQUILIBRIUM

A consumer is said to be in equilibrium when he is buying such a combination of goods that leaves him with no tendency to rearrange his purchases of goods. Following assumptions have been made to explain the equilibrium:

1. The consumer has a given indifference map exhibiting his scale of preferences for various combinations of two goods X and Y.

2. He has a fixed amount of money to spend on the two goods. He has to spend his given income on the two goods.

3. Prices of the two goods are given and constant for him. He cannot influence the prices of the goods by buying more or less of them.

The Conditions of Consumer's Equilibrium:

The following two conditions must be satisfied for consumer's equilibrium under the indifference curve analysis:

(i) MRS$_{XY}$ = P$_X$/P$_Y$ (Necessary Condition)

It indicates that slope of indifference curve must be equal to the slope of budget line at the point of equilibrium. It is so because if MRS$_{XY}$ > P$_X$/P$_Y$, it indicates that the consumer's willingness to pay for X is more than the price in the market. The consumer will continue to buy X till the ratio becomes equal for him. In the same way, if MRS$_{XY}$ < P$_X$/P$_Y$, it indicates that the consumer's willingness to pay for X is less than the price prevailing in the market. The consumer will reduce the amount of X till the ratio becomes equal for him.

(ii) Indifference Curve is convex to origin or MRS continuously falls (Sufficient Condition)

The second condition for consumer's equilibrium states that MRS must be diminishing at the point of equilibrium. In other words, indifference curve must be convex to origin at the point of equilibrium. This condition needs to be satisfied in order to achieve a stable equilibrium which is not possible to be achieved in case of MRS not falling continuously at the point of equilibrium.

Fig. 3.12: Consumer Equilibrium
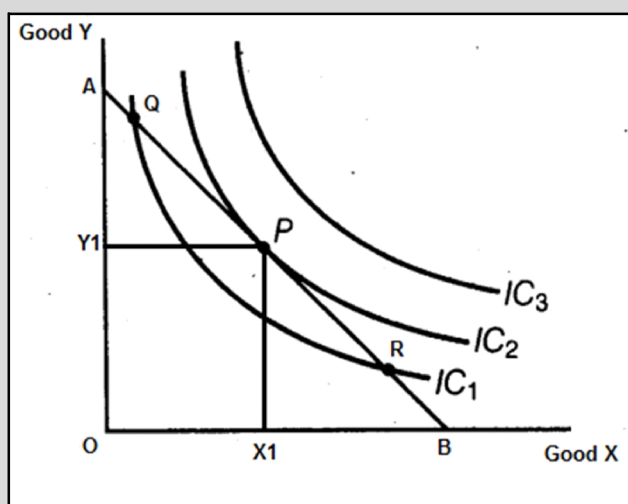
Figure 3.12 presents the consumer equilibrium. Being a rational consumer, consumer always tries to choose that combination of X and Y which brings him to the highest indifference curve which the market permits him to reach. In this context, we should note that the BL is a straight line and slopes downwards from left to right, and the indifference curves are convex to the origin. Therefore, if we take any particular indifference curve, we come

across three possibilities. The BL

- – does not touch it at all
- – intersects it twice
- – is tangent to it

It should also be noted that a BL can be tangent to one and only one indifference curve

Given the assumptions, highest possible indifference curve that can be reached would be one, to which BL is tangent. The point of tangency represents consumer's equilibrium.

- – In figure 3.12, BL touches to IC_2 curve at point P. Thus point P is the consumer equilibrium point as the slopes of BL and IC_2 are equal. Based on the analysis, consumer equilibrium is reached at point P as BL tangent to IC_2.
- – Point Q and R are not the equilibrium points as BL intersects IC but is not tangent. In other words, slopes of BL and IC_2 are not equal.

LESSON ROUND UP

- – The theories of consumer behavior deal with the problem of constrained optimization that a consumer faces while making choices among various goods and services.
- – Utility is want satisfying power of a good or service. It is also defined as property of the good and service which satisfies the want of the consumer.
- – Utility is subjective. It depends upon the mental assessment of the consumer and is determined by several factors which influence the consumer's judgment.
- – Cardinal Utility approach states that utility can be measured in cardinal numbers or definite numbers such as 1, 2, 3, 4, etc. Cardinal numbers are those definite numbers which can be subtracted or added. Fisher used the 'Util' as a measure of utility.
- – Concepts of Utility
  - – Total Utility - it is the aggregate level of satisfaction obtained by the consumer from the consumption of all the different units of a commodity.
  - – Average Utility -it is derived by dividing the total utility obtained by a consumer with the number of units of commodity X consumed.
  - – Marginal Utility -it is the additional utility derived from the consumption of an additional unit of commodity. Alternatively, it is the change that takes place in total utility by the consumption of an additional unit of a commodity.
- – According to the law of diminishing marginal utility, other things being equal, marginal utility of a good starts falling as an individual consumes more units of it in a given period of time. Alternatively, as a consumer consumes more and more units of a commodity, the extra satisfaction that he derives from successive unit of a commodity goes on diminishing.
- – According to the law of equi-marginal utility, a rational individual does not spend his entire income on one commodity. Consumer knows that if he purchases more units of the same commodity then marginal utility of each successive unit would decline. Hence, to get the maximum utility from his limited income, he spends his limited income in such a way that the last unit of money spent on different commodities yield equal marginal utility.
- – Equilibrium is a situation which is characterized by the absence of any tendency to change. Consumer equilibrium refers to a situation wherein a consumer gets maximum satisfaction from the purchases of goods at given prices and given income. Any deviation from this point places the consumer in the sub-optimal situation.
- – Single Commodity case: a consumer with a single commodity attains equilibrium at a point when the marginal utility derived from that good is equated to its price.

- Two or more commodity case: If there are two or more commodities, the condition for the equilibrium of the consumer is the equality of the ratios of the marginal utilities of the individual commodities to their prices.

- The indifference curve analysis is advancement to the utility analysis as it tries to overcome the drawbacks of cardinal utility analysis and provides a technically superior analysis of demand.

- According to the Ordinal Utility approach, consumer may not be able to indicate the exact amounts o futilities that he derives from commodities or any combination of them. But he is capable of judging whether the satisfaction obtained from a good or a combination of goods is equal to, lower than or higher than other.

- An indifference curve is a locus of various combinations of two commodities that give equal satisfaction to the consumer. Since all the combinations on an indifference curve yield equal satisfaction to the consumer, the consumer is indifferent among them.

- An Indifference map refers to a collection of indifference curves corresponding to different levels of satisfaction. In other words, a set of indifference curves is known as an indifference map.

- The rate at which the consumer is prepared to exchange one good for the other is known as Marginal Rate of Substitution (MRS). Alternatively, MRS measures the opportunity cost of obtaining one good in terms of giving up the other good.

- Budget constraint is the constraint that consumer faces as a result of limited incomes. Because of limited income, a consumer has to restrict his purchase and also prioritize his unlimited choices.

- A consumer is said to be in equilibrium when he is buying such a combination of goods that leaves him with no tendency to rearrange his purchases of goods.

GLOSSARY

Utility	It is defined as a want satisfying power of a commodity. It is the sensation which an individual derives from consuming a commodity. It can be measured on numerical scale as well as ordinally.
Cardinal Utility	Cardinal utility is a view of utility measurement based on the presumption that the satisfaction of wants and needs is a quantifiable characteristic of human activity. In other words, utility can be measured with numerical values (1, 2, 3, etc.) along a scale. If so, then the utility generated from consumption can be evaluated against an objective standard, which then makes it possible to compare utility among different goods and among different people.
Ordinal Utility	Ordinal utility is a view of utility measurement based on the presumption that the satisfaction of wants and needs is not a quantifiable characteristic of human activity and that preferences are subjective. Preferences among goods can be ranked (first, second, third, etc.) but not measured according to a scale. In this regard, consumers need only specify whether one good is more or less preferred than another. How much more or less a good is preferred is not important.
Total Utility	TU is the aggregate level of satisfaction obtained by the consumer from the consumption of all the different units of a commodity. If a consumer buys *n* units of good X then, for him, total utility from it is the summation of utilities derived from all the *n* units.
Marginal Utility	MU is the extra satisfaction generated from consuming one more unit of a good. Symbolically, Marginal Utility =Change in total utility/change in quantity
Average Utility	AU is derived by dividing the total utility obtained by a consumer with the number of units of X commodity consumed.

Law of Diminishing MU	The law of diminishing marginal utility means that the value of a good, the extra utility derived from good declines as more of the good is consumed. If the satisfaction obtained from a good declines, then buyers are willing to pay a lower price, hence demand price is inversely related to quantity demanded, which is the law of demand.
Law of Equi-Marginal Utility	According to the law of equi-marginal utility, a rational individual does not spend his entire income on one commodity. Consumer knows that if he purchases more units of the same commodity then marginal utility of each successive unit would decline. Hence, to get the maximum utility from his limited income, he spends his limited income in such a way that the last unit of money spent on different commodities yields equal marginal utility.
Consumers' Equilibrium	Consumer equilibrium exists when a consumer selects or buys the combination of goods that maximizes utility. This is achieved by equating the marginal utility-price ratio for each good consumed or by equating the ratio of prices and the ratio of marginal utilities. In other words, buyers are willing to pay relatively higher prices for goods that generate relatively more marginal utility.
Indifference Curve	An indifference curve is a curve which represents all those combinations of goods which give same satisfaction to the consumer. Since all the combinations on an indifference curve give equal satisfaction to the consumer, the consumer is indifferent among them.
Marginal Rate of Substitution	The rate at which an individual must give up "good A" in order to obtain one more unit of "good B", while keeping their overall utility (satisfaction) constant. The marginal rate of substitution is calculated between two goods placed on an indifference curve, which displays a frontier of equal utility for each combination of "good A" and "good B".
Budget Constraint	Budget constraint are the constraint that consumer faces as a result of limited incomes. Because of limited income, a consumer has to restrict his purchase and also prioritize his unlimited choices.
Budget line	Budget line shows all those combination of two goods which consumer can buy spending his entire money income on the two goods at their given prices. Symbolically, the budget line can be written algebraically as follow: $P_x X + P_y Y = M$

MULTIPLE CHOICE QUESTIONS

1. **The proponents of indifference curve approach are**

    a. Kuznets and Sweezy

    b. Hicks and Allen

    c. Marshal and Samuelson

    d. Samuelson and Sweezy

2. **The concept of cardinal utility was given by –**

    a. A. Marshall

    b. J. R. Hicks

    c. P. Samuelson

    d. J. M. Keynes

3. **Which of the following cannot be explained by Marshallian utility approach**

   a. Inferior good

   b. Normal goods

   c. Giffen goods

   d. None of the above

4. **If A > B and B> C, then A>C, this is known as assumption of**

   a. Intransitivity

   b. Transitivity

   c. Completeness

   d. None of the above

5. **Consumer reaches a point of saturation for a good when Marginal Utility is**

   a. Positive

   b. Negative

   c. Zero

   d. None of the above

6. **The slope of indifference curve is given by**

   a. MRS between two goods

   b. MRTS between two goods

   c. MRT between two goods

   d. None of the above

7. **The slope of budget line is given by**

   a. Px/Py

   b. MUx/MUy

   c. dX/dY

   d. None of the above

8. **One of the conditions of Consumer Equilibrium under ordinal utility analysis is achieved at a point where**

   a. Slope of indifference curve is equal to slope budget line

   b. Slope of isoquant is equal to slope of budget line

   c. Slope of budget line is equal to slope of iso-cost line

   d. None of the above

9. **Shape of indifference curve is a right angled triangle in case of**

   a. Substitute goods

   b. Complementary goods

c. Neutral goods

d. None of the above

10. **Sufficient condition for consumer equilibrium under indifference curve analysis is that Indifference curve should be**

a. convex to origin

b. Concave to origin

c. L shaped

c. None of the above

Answer Key : 1. (b), 2. (a), 3. (c), 4. (b), 5. (c), 6. (a), 7. (a), 8. (a), 9. (b), 10. (a)

Suggested Readings

1. R. Pindyck and D. Rubinfeld: Microeconomics

2. H. L. Ahuja: Modern Microeconomics

3. M. L. Seth: Principles of Economics

Lesson 4
Theory of Production and Cost

LESSON OUTLINE

LEARNING OBJECTIVES

The theory of production describes how a firm makes cost-minimizing and profit-maximising production decisions and how the firm's production cost varies with its output. It helps in understanding the functioning of supply-side constraints and decision making. The supply of a product is largely influenced by the costs of production. Therefore, it is important to understand the theory of cost. The theory of cost analyzes the relationship between cost and output, that is, how cost of production changes with change in production.

INTRODUCTION

The theory of production helps to understand producer behavior as it explains how producers make optimum production decisions subject to various cost constraints faced by a firm in the process of production. Alternatively, the theory of production describes how a firm makes cost-minimizing and profit-maximising production decisions and how the firm's production cost varies with its output. Thus, theory of production is a study of supply side constraints which firms face while taking production decisions. The production decisions of the firms depend upon relationship between the inputs and output as well as the prices of the inputs. The theory of production explains how firms can produce efficiently and how their costs of production change with the changes in costs of inputs and the level of output.

The various decisions a business enterprise makes about its productive activities can be classified into three layers of increasing complexity. The first layer includes decisions about methods of producing a given quantity of the output in a plant of given size and equipment. It involves the problem of what is called short-run cost minimization. The second layer, concerning the determination of the most profitable quantities of products to produce in any given plant, deals with what is called short-run profit maximization. The third layer, concerning the determination of the most profitable size and equipment of plant, relates to what is called long-run profit maximization.

FACTORS OF PRODUCTION

In order to produce goods, a firm needs various types of inputs which are known as 'Factors of Production', 'Agents of Production', 'Productive Resources' or sometimes even 'Productive Services'. In economics, the factors of production are divided into four categories: (i) Land, (ii) Labour, (iii) Capital and (iv) Entrepreneurship.

The level of production depends upon both the quantity of inputs and the efficiency with which they are employed in the process of production. It is also noteworthy that economic growth of a country, in a way, represents its productive capacity which, in turn, depends upon the technology and amount of productive resources.

1. **Land:** Land is not the creation of mankind as it is a gift of nature available to us free of cost. So, it is called as natural factor of production. It is also called as original or primary factor of production. Normally, land means surface of earth. But in economics, land has a wider meaning. According to Marshall, "By land is meant… materials and forces which nature gives freely for man's aid in land, water, air, light and heat." Thus, land is a stock of free gifts of nature as it is the sum total of those productive resources which are provided 'free of cost' by nature to us that is to say those resources on which no human effort has been expended to make them actually usable in a productive process. Land has certain distinguishing characteristics which include the following:

   – Supply of land is perfectly inelastic that is, it is fixed in quantity. Total quantity of land cannot be increased or decreased with human efforts.

   – Land is a Primary Factor of Production.

   – Land is a Passive Factor of Production: it cannot produce anything of its own. It needs help of Labour, Capital, Entrepreneur, etc.

   – Land is a perfectly immobile factor.

   – Economic reward for the use of land is rent.

2. **Labour:** The aggregate of all human physical and mental effort used in creation of goods and services.

3. **Capital:** Capital includes things like tools, machines, and other things that a business uses in order to produce their goods or services. At some level, all companies rely on their capital in order to run successfully.

Without these things, the company would be unable to carry out production.

4. **Entrepreneurship:** Entrepreneurship consists of three major functions, viz., coordination, management and supervision. Using these skills, the function of an entrepreneur is to bring the required factors together and making them work harmoniously.

Classification of factors of production: Inputs can be classified into two groups:

– Fixed factors (or inputs)

– Variable factors (or factors)

Fixed inputs: fixed inputs, in economic sense, are one whose supply is inelastic in the short-run and therefore they are used in a fixed quantity in the short-run.

Variable Factors: variable factors, in economic sense, are defined as one whose supply in the short-run is elastic. These includes labour, raw material and so on. Quantity of these factors can be changed, that is, increased or decreased, in the short-run according to the requirement of firm.

THEORY OF PRODUCTION

Theory of production studies the economic process of converting inputs into output, given the state of technology. It provides a basis for analysis of relationship between costs and amount of output. Further, it also provides a basis for the theory of firm's demand for factors of production.

Production Function: the production function is a technical relation that connects factor inputs and output. It describes the laws of proportion, that is, the transformation of factor inputs into product (output) at any particular time period. In real world, a firm can convert inputs into output in a number of ways, using various combinations of labour, raw material and capital goods. Thus, the production function describes the technical method which a firm adopts to convert inputs into output. For simplification, suppose there are only two inputs such as labour and capital. Firm uses these inputs to produce goods.

The production function can be expressed in form of an equation in which the output is the dependent variable and inputs are the independent variables. The equation is expressed as follows:

Q =f(K, L)

Where,

Q = Quantity of output

K = Capital

L = Labour

Method of Production: A method of production is a combination of factors of production (or factor inputs) required for the production of one unit of output. A unit of commodity can be produced by *n* number of methods of production. Suppose, a firm can produce a commodity x by using any of the following four methods:

Table 4.1: Methods of production

Factors/Methods	Capital units	Labour units
Method 1	3	2
Method 2	2	3
Method 3	4	1
Method 4	1	4

Which method would be technically efficient method of production?

A method of production A can be technically efficient relative to B, if method A uses less amount of one input and same amount of other input or less amount of both the inputs as compared to B.

Example 1: Suppose, method A uses 3 units of labour and 4 units of capital to produce one unit of commodity X and method B uses 4 units of labour and 4 units of capital to produce same unit of commodity X. In this case, method A is technically efficient method of production. In other words, method B is technically inefficient method of production. Theory of production deals with efficient methods of production as rational enterprises will not use inefficient methods of production.

Example 2: Suppose, there two methods of production, say A and B. Method A uses less quantity of some inputs and more quantity of other inputs as compared method B. Suppose, to produce good X, method A uses 3 units of labour and 4 units of capital and method B uses 2 units of labour and 5 units of capital. *Which method would be technically efficient method of production?* In this case, two methods (A and B) cannot be directly compared on the criteria of technical efficiency. Both methods may be considered technically efficient methods of production.

Which method will be chosen by manufacturer?

In this case, which one method would be chosen by the firm at any particular point of time will depend on the prices of factors of production. It is worth noticing that the choice of any particular method of production by manufacturer, among the set of technically efficient methods of production, is economic one, which is based on prices of factors of production.

– One should keep in mind that a technically efficient method of production is not necessarily economically efficient. There is a difference between technically and economically efficient method.

ISOQUANT

What is isoquant? An isoquant is the locus of all the combinations of factors of production which produces a given level of output. Alternatively, an isoquant includes all the technically efficient methods of production of a commodity which produces a given level of output. The isoquant can be of different shapes depending upon the degree of substitutability of factors of production. An isoquant can be of four different shapes:

1. **Linear isoquant:** linear isoquant assumes perfect substitutability of factors of production. If factors of production are perfectly substitutable then a unit of commodity can be produced by employing only labour or only capital or by any combination of capital and labour.

Fig. 4.1: Linear Isoquant
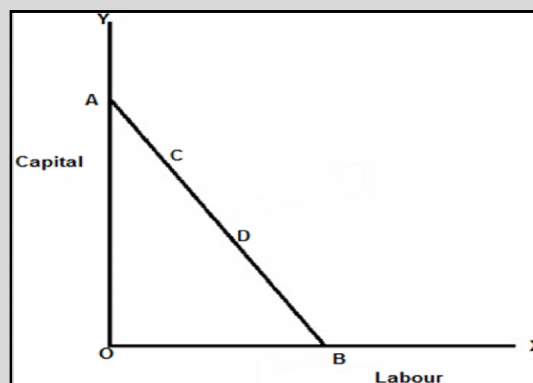
Figure 4.1 presents the shape of linear isoquant. Given the perfect substitutability of factors of production, the rate of substitution between two factors (labour and capital) of production on all the points remains constant. Producer can produce maximum output by using OA amount of capital or OB amount of labour or any combination lying on isoquant AB.

2. **Kinked Isoquant:** According to kinked isoquant curve, there is a limited substitutability between the factors of production. Substitution is limited because there are only few processes for producing any commodity which are represented by kinks on the isoquant. Thus, substitution is possible only at the kinks (See figure 4.2). This form of isoquant is called '**activity analysis-isoquant' or linear programming isoquant** because it is basically used in linear programming.

Fig. 4.2: Kinked Isoquant
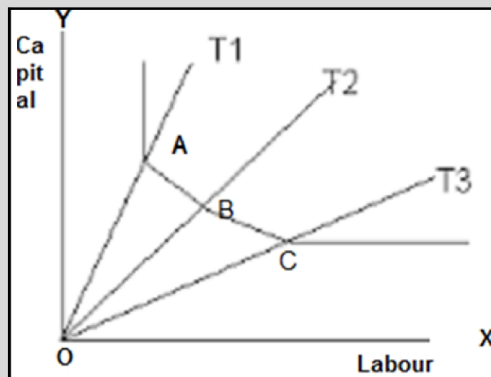
The shape of kinked isoquant is presented in figure 4.2. There are only three points on the isoquants such as A, B and C where substitution is possible between labour and capital. On the other hand, except on points A, B and C, it is not possible to change the composition of two factors of production.

3. **Input-output isoquant:** In case of input-output isoquant the rate of substitution between two factors of production is zero because it assumes strict complementarily between the factors of production. It means there is only one method of production to produce a commodity. This kind of isoquant is named after Leontief who invented the input-output analysis. As shown in Figure 4.3, point A on IQ_1 is the only combination of capital and labour which produces given amount of good efficiently.

Fig. 4.3: Input-output isoquant
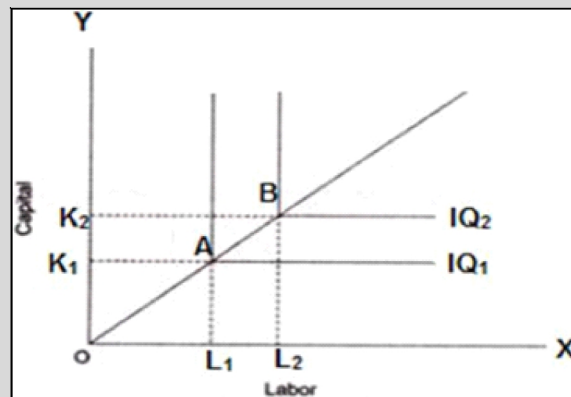
4. **Convex Isoquant:** this shape of isoquant assumes continuous substitutability between factors of productions only over a certain range, beyond which factors cannot be substituted for one another (Figure 4.4). It is so because, as we substitute capital for labour then the rate of substitution will decline. As we move down along isoquant the shape of isoquant gets flatter.

Fig. 4.4: Convex Isoquant
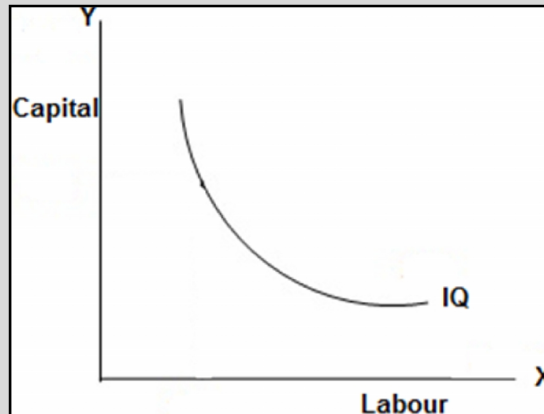
Concept of Production

To understand the production theory, various concepts need to be introduced. These include:

Total Product: it refers to the total amount of goods and services produced by a manufacturer with given amount of inputs.

Marginal Product: it refers to the addition to the total product resulting from the employment of additional unit of variable factor. In other words, it represents the successive increase in the total output due to the additional units of variable factor. Symbolically:

$MP_n = TP_n - TP_{n-1}$

MP_n = Marginal product of nth unit

TP_n = Total product produced by n units of factor

TP_{n-1} = Total output produced by n-1 units of factor.

OR

$$\text{Marginal Product} = \frac{\text{Change into total products}}{\text{Change in variable input}}$$

Average Product: Average product of a factor is the total output produced per unit of the factor employed. It can be obtained by dividing the total output by the total units of factor employed.

$$\text{Average Product} = \frac{\text{Total Product}}{\text{Inputs of variable factors}}$$

Types of Production Functions

There are two types of production functions which will be discussed under theory of production:

(i) Short Run Production Function

(ii) Long Run Production Function

Short run refers to a period of time in which quantities of one or more factors of production cannot be changed. In other words, in the short run there is at least one factor that cannot be varied. In contrast, long run is defined as that time period over which a firm can vary quantities of all the factors of production and therefore, can switch between different scales. In case of long run production function all inputs are variable. There are two alternative theories to these production functions i.e.

- – Law of Variable Proportions (or Law of Diminishing Returns): Short-Run Analysis of Production

- – Law of Returns to Scale (to analyze production in the long period)

Assumptions of Production Function

The analysis of production function rests on certain assumptions which are briefly listed below :

- – Perfect divisibility of both inputs and output.

- – Limited substitutability of one factor for the other as labour and capital are imperfect substitutes.

- – The level of technology remains constant.

- – The firm uses its inputs at maximum level of efficiency.

- – It relates to a particular unit of time.

- – A change in any of the variable factors produces a corresponding change in the level of output.

These are the general assumptions on the basis of which a production function is constructed.

LAW OF VARIABLE PROPORTIONS: SHORT-RUN ANALYSIS OF PRODUCTION

The law of variable proportions is the modern approach to the 'Law of Diminishing Returns (or The Laws of Returns). It is now usually called the Law of Variable Proportions. It can also be called the Law of Diminishing Marginal Product or Diminishing Marginal Returns or simply as Diminishing Returns. The law of variable proportions shows the production function with one input factor variable while keeping the other input factors constant.

Assumptions of Law of Variable Proportions

The law of variable proportions is valid under the following assumptions:

- – It is assumed that the state of technology remains unchanged. If there is an improvement in the technology, due to inventions, the average and marginal product may increase instead of decreasing.

- – It is assumed that there are some inputs whose quantity is fixed and some are variable. This is one of the ways by which we can alter the composition of the factor and know its effect on output. This law does not apply when all the factors are variable.

- – All units of the variable factor are homogenous.

The law of variable proportions explains the relationship between factor proportions of fixed and variable inputs on the one hand and output on the other. Alternatively, it refers to the input-output relation when output is increased by varying the quantity of one input. As quantity of variable factor increases and quantity of fixed factor remains constant, normally the marginal product and average product of the variable factor will increase upto a point. Thereafter, marginal production will start falling, the decline in marginal product will pull down the average product.

As the composition between variable factor and fixed factor changes, resultant change in output occurs in

varying proportions. The behavior of output with varying composition of fixed and variable factors can be divided into three distinct stages of production and returns to a factor.

Stages of Production

- First, Total product first increases at an increasing rate and then at a decreasing rate and this continues till the end of this stage. Average product is continuously increasing. MP first increases, becomes maximum and then starts falling. But the marginal product remains higher than the average product, that is, $MP_L >$ AP_L. The stage I ends where average product reaches its highest point, so here, the efficiency of variable factor (labour) is maximum. There are two important reasons for increasing returns: (i) indivisibility and (ii) specialization.

- In the second stage, total product continues to increase at a diminishing rate until it reaches its maximum point at the end of this stage. Both AP and MP diminish, but remain positive. At the end of the second stage, MP becomes zero. TP is maximum when MP is zero. AP shows a steady decline throughout this stage. Since both AP and MP decline, this stage is known as stage of diminishing returns. The main cause of application of the law of diminishing returns is the scarcity of fixed factor or the other factor of production.

- In the third stage, total product will start declining. The marginal product of the variable factor will become negative whereas the average product will remain greater than zero. The phenomenon of negative returns emerges as a result of application of excessive units of variable factor in relation to fixed factor, with the result that TP starts diminishing.

Graphic Presentation of Law of Variable Proportions

Three stages of law of variable proportions are graphically illustrated in figure 4.5.

Fig. 4.5: Three Stages of Production
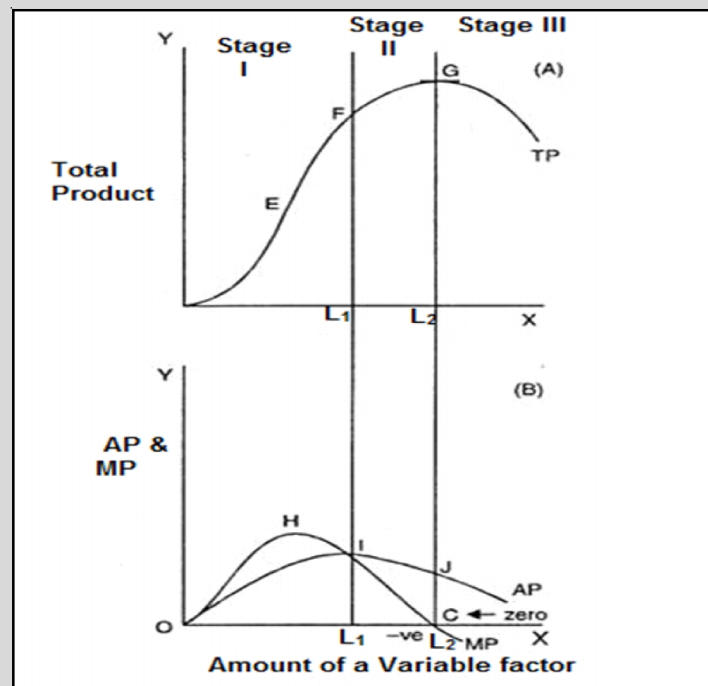
Figure 4.5 graphically presents the stages of production on account of change in variable factor when other factor remains constant. On x-axis, the quantity of the variable factor is measured whereas total product, average product and marginal product are measured on y-axis. Top panel of the figure 4.5 shows the total product curve and bottom panel shows average and marginal product curves.

Stage 1: It continues till the firm employs OL_1 units of variable factor.

- In this stage, MP first increases and reaches its maximum shown as point H in lower part of the figure. After reaching its maximum, it continues to decline in Stage 1.

- It is important to note that though MP starts declining in this stage after point H, it still remains higher than AP whereas AP continues to increase in Stage 1 and becomes maximum at the end of this stage.

- Corresponding to varying behavior of MP, TP increases at an increasing rate till the point of inflection as shown by point E in the upper part of the diagram.

- After point E, TP increases but at a decreasing rate upto point F in Stage 1.

- It should be noted that Stage 1 ends at the point where MP and AP are just equal to each other. In other words, this stage ends when MP curve coincides with AP curve at its maximum point.

Stage 2: It continues till the firm employs OL_2 units of variable factor.

- In this stage, MP continues to decline and cuts horizontal axis at point C at the end of the Stage 2.

- As against Stage 1, MP remains lower than AP at all points in Stage 2. AP starts falling and continues to decline in the positive zone.

- Corresponding to the behavior of MP, TP continues to increase at a decreasing rate till MP remains positive and reaches its maximum as shown by point G in the upper part of the diagram when MP turns zero i.e. at point C.

- It can be noted that point G and C lie directly over each other at OL_{-2} units of employment of the variable factor.

Stage 3: It starts after firms crosses OL_2 units of employment of the variable factor.

- In this stage, MP has turned negative from being zero at the start of this stage.

- AP still continues to decline as in Stage 2 but it remains positive.

- Corresponding to negative MP, TP starts declining in Stage 3 after reaching its maximum at the start of this stage.

- It can be noted that if firm still employs more units of the variable factor, MP will decline further in the negative zone and correspondingly TP will also continue to fall.

Production Decision of Business Firm

An important question is in which stage a rational producer will seek to produce. The objective of the rational producer is to minimize cost and maximize profits.

- A rational producer will not operate in stage three as marginal product of variable factor is negative in stage 3. As the marginal product of the variable unit becomes negative; a business firm can always increase its output by reducing the variable factor.

- A rational firm will not operate even in stage I. It is so because, firm is not making best use of its fixed factor. Thus, the increase in variable factor increases the average returns to the variable factor and total output is also increasing. Thus, a competitive firm has an incentive to expand all through stage 1.

– A rational firm will always operate in stage II. Both the marginal product and average product of the variable factor are diminishing but total product still increases. Total output reaches its maximum when marginal product becomes zero. It is this stage of operation that is suitable for rational producer.

Three Types of Returns to a Factor:

1. **Increasing Returns to Factor:** increasing returns to a factor refers to a situation in which output increases by a larger proportion when more and more units of a variable factor are combined with fixed factor of production. In such a situation, marginal product of the variable factor must be increasing.

   • In Table 4.2., when firm increased the number of variable units to 3, the corresponding increase in output is more than proportional increase in input. Not only this, marginal product of a variable factor is also increasing till the point firm employed three units of variable factor with the fixed factor.

   • A business firm experiences increasing returns when per unit cost continuously goes on diminishing.

2. **Diminishing Returns to Factor:** it occurs when the rise in marginal product from successive unit of a variable factor starts declining but remains positive. In Table 4.2, firm faces diminishing returns as it increases its variable units from 4 to 6.

Table 4.2: Stages of Law of Returns to Factor

Units of Fixed Factor (Capital)	Units of Variable Factor (Labour)	Total Product	Marginal Product	Average Product	Returns to variable Factor
5	1	4	4	4	
5	2	12	8	6	Increasing returns to
5	3	24	12	8	a factor
5	4	32	8	8	Diminishing returns
5	5	34	2	7	to a factor
5	6	34	0	6	
5	7	30	-4	4	
5	8	25	-5	3	Negative returns to
5	9	15	-10	2	a factor
5	10	4	-11	0	

3. **Negative Returns to Factor:** it refers to the stage when increase in variable factor causes fall in total output. As firm increases its variable input corresponding change in the output becomes negative (Table, 4.2).

Factors behind the Law of Returns

One of the important factors causing increasing returns to a variable factor is the indivisibility of fixed factor. Each unit of fixed factor, say capital, requires an optimum number of variable units. If the proportion of employed variable factor is less than the optimum requirement, it results in under utilization of the fixed factor and therefore lowers the productivity of the variable factor.

What causes diminishing returns to a variable factor? Once the optimum combination of fixed and variable factor is achieved, employment of additional amounts of variable units with fixed factor reduces the marginal product of variable factor because of un-optimal composition of fixed and variable factors.

Limitations of Law of Variable Proportions

– Homogeneous units of variable input (especially labour) is an unreal assumption. Every unit of labour or worker is different from each other.

– The assumption of constant input prices and technique of production also seems unrealistic in the present dynamic world.

RETURNS TO SCALE

The Law of Returns to Scale is a long run phenomenon. In the long run, all factors of production become variable. Alternatively, no factor of production is fixed in the long run. Thus, the firm is able to alter its scale of production in the long run by increasing all the factors of production which is not the case in short run. When all factors are changed in same proportion, the behaviour of output is analyzed with the help of laws of returns to scale. Thus, this law takes into consideration the change in the scale of production. The factors determining the scale of production of the firm cannot be changed in the short period. If the firm increases the units of both factors labour and capital, its scale of production increases. According to Koutsoyiannis "The term returns to scale refers to the changes in output as all factors change by the same proportion."

Explanation of Production Function

Suppose there are two factors of production such as labour (l) and capital (k). Output can be increased by increasing two factors in the same proportion. Such an increase in output would be called as returns to scale.

Initial production function is

$$Q = f(L, K)$$

Suppose, a business firm increases both capital and labour in the same proportion say z, then the production function can be written as:

$$Q_1 = f(zL, zK)$$

Q_1 refers to increase in output due to increase in factors of production in the same proportion.

– If Q_1 increases in the same proportion as the factors of production, that is, $Q_1/Q = z$, then it will be the case of constant returns to scale. For example, in Table 4.3, firm increased all the factors of production in same proportion from 4 to 7, and the corresponding change in output remained constant, i.e. 10, indicating constant returns to scale.

– If Q_1 increases by less than the proportionate increase in factors of production, that is, $Q_1/Q < z$, then it will be the case of diminishing returns to scale. For example, firm increased all the factors of production in the same proportion from 8 to 10, and the corresponding successive increase in output started declining indicating diminishing returns to scale (Table, 4.3)

– If Q_1 increases by more than the proportionate increase in factors of production, that is, $Q_1/Q > z$, then it will be the case of increasing returns to scale. For example, firm increased all the factors of production in the same proportion from 1 to 3, and the corresponding change in output was more than proportionate change in inputs indicating increasing returns to scale (Table, 4.3).

Table 4.3: The Returns to Scale

Units of Capital	Units of Labour	Total Output	Increase in output	Nature of returns to scale
1	1	4	4	Increasing returns to scale
2	2	10	6	
3	3	18	8	
4	4	28	10	Constant returns to scale
5	5	38	10	
6	6	48	10	
7	7	58	10	
8	8	65	7	Diminishing returns to scale
9	9	70	5	
10	10	74	4	

Isoquant and Returns to Scale

To understand and explain the laws of returns to scale, economists make use of isoquant curves. A brief description of the concept of isoquant has already been given in the beginning of the chapter. An isoquant is the locus of points representing different combinations of two factors of production (labour and capital) that yield the same level of output. The isoquant is also known as equal product curve and production indifference curve. To know about different shapes of isoquant curve and its usages, please refer figures 4.1, 4.2, 4.3 and 4.4. Although isoquant looks similar to indifference curves but these two are different, the following points should be noted:

  – An indifference curve represents different combinations of two consumer goods yielding same level of satisfaction whereas an isoquant represents whereas an isoquant represent different combinations of two producer goods producing same quantity of a commodity.

  – An indifference curve represents immeasurable utility (or satisfaction) whereas isoquant represents a measurable quantity of output of a product.

Returns to scale are of the following three types:

  1. Increasing Returns to Scale

  2. Constant Returns to Scale

  3. Diminishing Returns to Scale

Increasing Returns to Scale: increasing returns to scale occurs when a given percentage change in all the factors of production (say z) causes output to increase in agreater proportion. In other words, proportionate increase in all factors of production resulting-in a more than proportionate increase in output is a case of increasing returns to scale.

Fig. 4.6: Increasing Returns to Scale
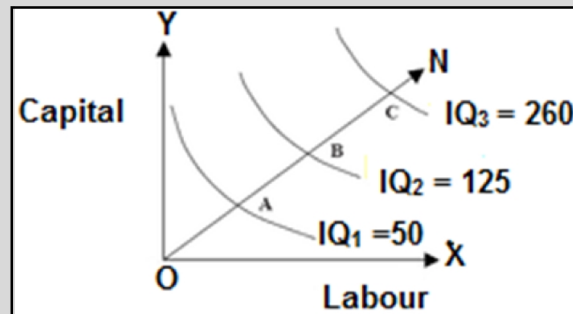
There are three isoquants - IQ_1, IQ_2 and IQ_3—represent three different levels of production – 50, 125 and 260 units, respectively. The product line ON shows the relationship between input and output. For instance, movement from point A to B denotes doubling the factors of production such as labour and capital. Input combination increases from IK +IL to 2K +2L (Figure, 4.6). The movement from A to B indicates increase in output by more than double, that is, 125 units. Similarly, movement from B to C indicates the increase in output by more than the corresponding change in factors of production (Figure, 4.6).

Causes of Increasing Returns to Scale: The law of increasing returns to scale comes into operation because of economies of scale. There are various factors which cause increasing returns to scale.

- **Technical and Managerial Indivisibility:** Certain inputs, such as machinery, managerial skills etc., are indivisible as these are available for a given production size. For instance, the capital and manager cannot be divided into parts to suit small size of production. Because of the indivisibility, such factors have to be employed in a minimum quantity even if the scale of production is small. In this case, as a business firm increases its production the per unit cost for successive unit starts declining which causes increasing returns to scale.

- **Specialization of factors:** as the scale of production increases the factors such as labour and managerial manpower become more specialized. The use of specialized manpower increases productivity per unit of inputs. This effect causes increasing returns to scale.

Constant Returns to Scale: Constant returns to scale occurs when the quantity of the factors of production is increased in such a manner that the ratio of the factors remains unchanged, output increases in the same proportion in which the factors of production are increased. In other words, when inputs are increased in a given proportion and output increases in the same proportion, the returns to scale are said to be constant. Thus, if quantity of all inputs is doubled then total output also doubles. Such a production function is often called linear homogenous production function or homogenous production function of the first degree. To make it more clear, the relationship between inputs and outputs are presented graphically in figure 4.7.

Fig. 4.7: Constant Returns to Scale
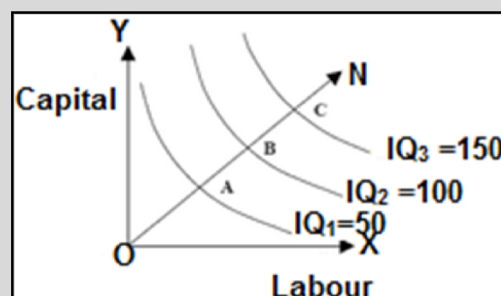
There are three isoquants - IQ_1, IQ_2 and IQ_3— which represent three different levels of production – 50, 100 and 150 units, respectively. The product line ON shows the relationship between input and output. For instance, movement from point A to B denotes doubling the factors of production such as labour and capital. Input combination increases from IK +IL to 2K +2L (Figure, 4.7). The movement from A to B indicates increase in output which is equal to proportionate increase in inputs, that is, output also doubles (from 50 to 100 units). Further, movement from B to C increases inputs from 2K +2L to 3K + 3L correspondingly output increases from 100 to 150 which is equal to proportionate change in factors of production.

Causes of Constant Returns to Scale: The constant returns to scale are attributed to the limits of economies of scale. When economies of scale disappear and diseconomies are yet to start, the returns to scale become constant. In other words, when inefficiency of production on a small scale are overcome and no problem regarding technical and managerial indivisibility remain, expansion in scale leads to a situation where returns increase in the same proportion as the factors of production.

 – Constant returns to scale are said to occur in those productive activities wherein factors of production are perfectly divisible. When factors of production are completely divisible, the production function is homogenous of degree one.

Diminishing Returns to Scale: it refers to a situation when the quantity of factors of production is increased in such a way that proportion of factors remains unchanged, and output increases in a smaller proportion as compared to increase in the quantity of factors of production. Alternatively, if the firm continues to expand beyond the stage of constant returns, the stage of diminishing returns to scale will start to operate. In this situation, the proportionate increase in all inputs results in less than proportionate increase in output, that is, if all inputs are doubled then total output is less than doubled. For example, if inputs are increased by 40%, but output increases by only 30%, it is a case of diminishing returns to scale. Diminishing returns to scale implies increasing costs.

Fig. 4.8: Diminishing Returns to Scale
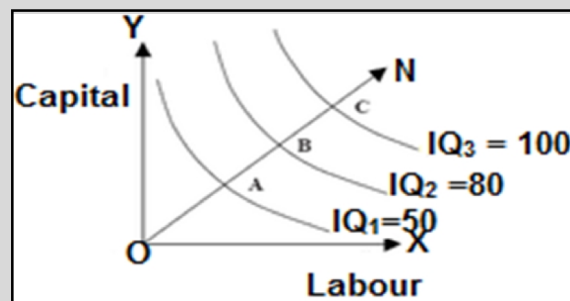
There are three isoquants - IQ1, IQ2 and IQ3—represent three different levels of production – 50, 100 and 150 units, respectively. ON product line represents the relationship between inputs and output. In figure 4.8, movement from point A to B indicates doubling the factors of production from IK +IL to 2K +2L but the corresponding increase in output is less than proportionate, that is, output increased from 50 to 80 units.

Causes of Diminishing Returns to Scale:

 – Diminishing returns to scale are caused by diseconomies of scale. It mainly happens as the size of production expands, managerial efficiency decreases which causing decrease in the rate of increase in output.

 – Another factor causing diminishing economies of scale is the exhaustibility of the natural resources. For instance, doubling the size of coal mining plant may not double the coal output because of limitedness of coal deposits.

PRODUCER'S EQUILIBRIUM OR LEAST COST COMBINATION

Producer equilibrium refers to a situation when producer is producing a given amount of output with least cost combination of factors. The least cost combination of factors is also known as optimal combination of factors. Least cost combination refers to that combination in which producer is able to achieve either of the following two conditions:

– production of maximum output from a given level of inputs.

– production of a given level of output at minimum cost.

To understand the least cost combination, it is important to know the prices of the inputs and the total expenditure to be incurred by the producer. Based on the budget and the prices, the isocost curve can be drawn.

Isocost Line

Isocost line shows various combinations of two factors of production that firm can buy with given outlays. In other words, isocost line is defined as the locus of various combinations of factors which a firm can buy with a given outlay. The isocost line is also known as the equal cost line or outlay line.

As it is assumed that there are only two factors of production. Thus, the total cost incurred on the factors of production for producing a good is the sum of the payments made to labour and capital. Thus, isocost line in equation form can be written as follow:

$C = r.K + w.L$

Where

C = total cost incurred by the firm on purchasing the quantities of factors used for production.

K = Capital

L = Labour

r= price of capital (or rate of interest on capital)

w= price of labour (or wage rate)

Shape of isocost curve: Isocostline is the straight line which touches both the axis, that is, x-axis and y-axis. Line AB is the isocostline. Line AB passes through all the combinations of labour and capital which the firm can buy at given outlay (say Rs. 1000) at given prices of the factors of production.

Fig. 4.9: Shape of IsoCostLine
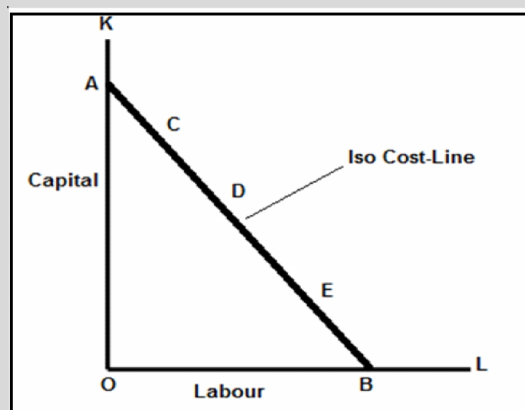
Slope of the IsoCostline: the slope of the isocostlien is the ratio of input prices. In figure 4.9, units of labour are shown on x-axis and units of capital are on y-axis. The slope of the isocostline would be the ratio of price of labour to price of capital.

Slope of the Isocostline = $\dfrac{Price\ of\ Labour}{Price\ of\ Capital}$

Shift in the isocost line: isocostline will shift under two circumstances:

- When prices of the factors remain unchanged but the total expenditure made by the firm will change. Suppose producer wants to increase total spending on the factors of production from Rs. 1000 to 1200 then isocostcurve will shift outward. The A'B' is the new iso cost-line. On the other hand if producer wants to reduce total expenditure on the factors of production from Rs. 1000 to 800. As a result, new isocostline A"B" shifted inwards.

Fig. 4.10: Shift in IsoCostLine
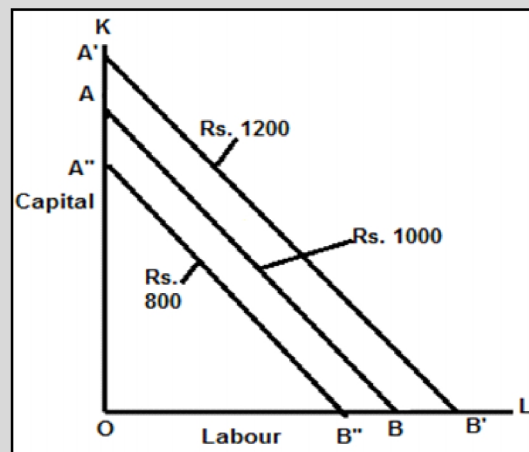
- When expenditure on the factors remains unchanged but the prices of the factor change. Suppose the prices of both the factors decrease proportionately then the real income of the producer will increase which will shift isocost curve outwards. On the other hand when prices of both the factors increases in the same proportion then the isocostline will shift inward.

Least Cost Combination

The least cost combination of factors refers to a firm producing the maximum volume of output from a given level of inputs on the one hand and on the other hand a firm produces a given level of output with the minimum cost when the factors are combined in an optimum manner.

Case 1: Maximization of output for a given cost: in this part, an attempt is made to explain the case when cost is given and objective of the producer is to maximize his output. Suppose the total given outlay is C and the prices of capital and labour are P_k and P_L.

Fig. 4.11: Maximizing Output for a Given Level of Inputs
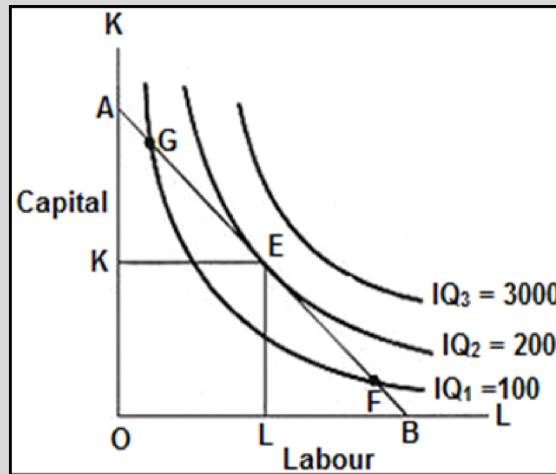
Figure 4.11 presents the case of maximizing output for a given level of inputs.

- AB is the isocostline represents the given cost outlay at given prices of capital (P_K) and labour (P_L).

- IQ_1, IQ_2 and IQ_3 are the three isoquants representing three different levels of output.

- It is important to note that with given amount of labour and capital, a business firm can produce any output level in the area OAB. It requires additional resources to produce any output level outside this area. Hence, IQ_3 level of output is not possible to attain as given level of labour and capital is not sufficient to produce that level of output.

- In this case, produce wants to reach the isoquant which provides the maximum level of output. In this case, the consumer can produce maximum output at point E on isoquant IQ_2. Thus, point E is the producer equilibrium point as it satisfied both the conditions for producer equilibrium.

  - At point E the slope of the isocostline (w/r) is equal to the slope of the isoquant (MP_L/MP_L). This is the first condition for equilibrium.

  - Second condition of equilibrium is that, the isoquant must be convex to the origin.

Case 2: Minimizing Cost for a Given Level of Output: in this case the level of output is given and the objective of the producer is to minimize the cost to produce given level of output. The conditions of equilibrium of the firm remain same, that is, slope of isoquant must be equal to slope of isocostline and the isoquant must be convex to the origin.

Fig. 4.12: Minimising Cost for a Given Level of Output
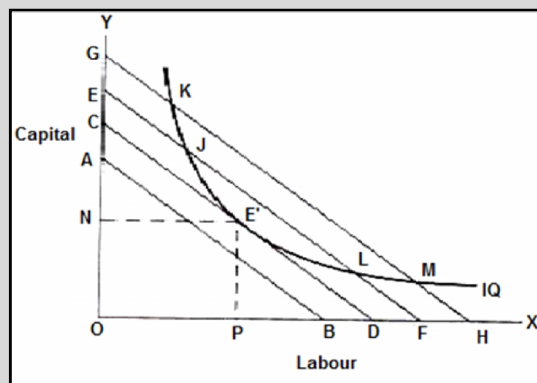
– The firm minimizes its cost by employing the combination of labour and capital determined by the point of tangency of the isoquant with the lowest isocostline. In figure 4.12, E' is the least cost combination point. At point E', both the conditions for equilibrium are satisfied.

– Point below E' is desirable because they show the least-cost combination but it is not attainable to produce output represent by isoquant IQ. Point above to E' show higher costs. Thus point E' is only optimal combination point of factor which produces given output at lowest possible cost.

THEORY OF COST

Theory of cost analyzes the relationship between cost and output, that is, how costs of production changes with changes in production. The relationship between cost and output is called cost function. The cost function of the firm depends upon the production conditions and the price of the factors used for production. How much costs a firm will incur on production depends upon the level of output. Thus, it is important to note that the supply of a product which a firm will offer to sell in the market depends largely upon the costs of a production incurred by a firm on various levels of output. In other words, it is the supply of a product is largely influenced by the costs of production. Costs of production are nothing but the prices of factors of production. There are four major inputs as discussed; land, labour, capital and entrepreneurship. The costs attached with each are; rent, wages, interest and profits respectively. Like production, costs of a firm may also be analyzed in the context of time period as follows:

– Short Run Costs

– Long Run Costs

Concepts of Cost

Before analyzing the cost concepts under short and long run, some basic concepts need to be discussed which do not change under short and long run.

– **Accounting Cost:** Accounting costs are the cash payments which a firm makes to other factor owners for purchasing and hiring for various factors of production. It is also called as explicit costs. When a business firm starts production of a good, firm has to pay prices for the factors used in the production process. For instance, a firm pays wages to the labourers employed, prices for the raw material, rent for place and building it hired for production and rate of interest on money borrowed for business. An accountant will take into account all the direct payments made by the business firm.

– **Economic Costs:** in addition to all these accounting costs, economic costs take in account the all the implicit costs which the entrepreneur could have earned if he would have invested his own money and sold his own services in the next best alternative uses. For instance, if an entrepreneur invests his certain amount of money in his own business would have invested somewhere else then he would have earned certain amount of rate of interest or dividends. Similarly, instead of working in his own business if he would have render his managerial and entrepreneurial services somewhere else he would have earned some amount of money. Thus, economist includes all these cost as implicit costs in the cost of production. Hence, economic costs take in account both the direct and indirect costs.

Economic costs = Accounting Costs + Implicit Costs

Opportunity Cost: Opportunity cost is the cost of any next best alternative use that is sacrificed. Suppose a piece of land can be used to grow both wheat and gram. If a farmer uses that piece of land only to grow wheat then he sacrifices the production of gram.

– Opportunity cost for a firm of using resources in the production of one good is the revenue foregone by

not using these resources in their next best alternative. Suppose the factors of production which a firm used for the production of good X may also be used for the production of good Y. Thus, the opportunity cost of production of good X is the foregone revenue which a firm would have earned from the production of good Y.

Private Cost: Private cost is the cost made by a firm for manufacturing a commodity. It is the sum of both explicit cost and implicit cost as a business firm takes into account all these cost while making both production decision and decision about prices of the product it has manufactured.

Social Cost: it is the cost to the society of an economic activity. In other words, social cost is the sum of the private cost and the net of negative externalities over positive externalities.

– The production of a commodity caused benefits and some loss to the society including the depletion of depletion of air quality due to high pollution, and other negative externalities generated by the manufacturing firm for a particular locality or for the society as a whole. A manufacturing firm does not take into account these costs while making decisions regarding the prices and output of a product. These externalities caused by the production of a firm to other constitute a part of social cost which a business firm completely ignores in any business decision but may be greater significance for the social point of view.

– For instance, a pesticide producing firm creates some negative externalities while producing pesticides which include foul smell, polluted air and effluents that are emitted as part of production of pesticides. These caused economic damage to the society in various forms including health problems to the people who live in the adjoining areas. These harmful effects are known as negative externalities.

– The beneficial effects caused by the production of a particular commodity are known as positive externalities.

Cost Function

A firm's cost function is the functional relationship between its output and its cost. Cost function is a determined by the production function of the firm and the prices it pays for different inputs. Hence, the theory of production forms the basis of the theory of cost. It is so because the rate of change in production with change in inputs determines the rate of change in production cost. It implies that the change in total, marginal and average costs depend on what law of production is in operation. It means the cost function is a derived function. Therefore, the theory of production will be examined in the framework of theory of production. Like production, costs of a firm will also be discussed in under two types of production functions:

– Short Run Costs

– Long Run Costs

Short-Run Costs

The short run is defined to be the period when at least some of the factors are fixed and some are variable. In the short-run firm's plant and equipment are the most difficult inputs to change quickly. Therefore, in short run there are two types of costs: fixed cost and variable cost.

– **Fixed Cost:** in the short run firm requires to incur few fixed costs initially in short period irrespective of the level of output. For ex., a firm would need land to build factory, an electricity connection to run machines, machinery to produce output, some mangers or employees to manage each and every function. All these expenditures are not independent to the level of output and are required to incur before production actually starts. Thus, fixed costs are those costs which are incurred in hiring the fixed factors of production whose amount cannot be changed in the short run.

- **Variable Cost:** the variable cost is the one which keeps changing with the changes in the level of output. It includes the expenditure incurred on the employment of variable factors including prices of the raw material, fuel and power used, payments to the workers etc. Variable cost remains zero until firm does not produce any output. Thereby, variable cost varies with the increase in the level of output.

- **Total Cost in the short run:** total cost is the sum of total fixed cost and total variable cost.

- Total Cost = Total Fixed Cost (TFC) + Total Variable Cost (TVC).

Fig. 4.13: Short-Run Total Cost, Total Variable Cost and Total Fixed Cost
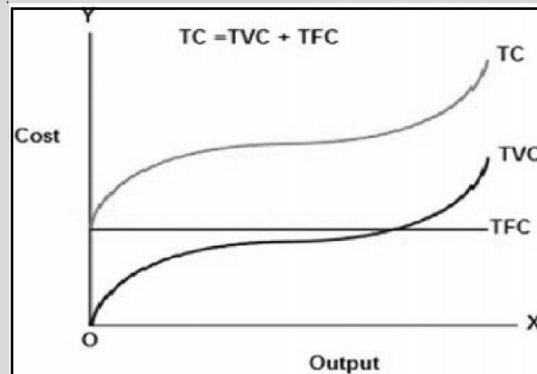
The shape of total cost, total variable cost and total fixed cost curves are presented in figure 4.13. Output is measured on the x-axis and cost on the y-axis.

- Total fixed cost is parallel to x-axis which indicates that it remains constant whatever be the level of output. Instead of starting from origin (or zero), it starts from y-axis which implies that it is fixed even when output is zero. Firm has to incur OA level of cost before it starts any production.

- On the other hand, total variable starts from origin and it varies with the level of output (Figure, 4.13). TVC was zero when firm was not producing any unit of output. As firm increases its level of output, TVC starts increasing and it decreases as firm reduces the quantity of output.

- Total cost also starts from point A which is located on y-axis. It starts from point A because it comprises both fixed and variable costs. It means when firm does not producing any unit of output, TFC remains equal to total fixed cost because firm does not incur any amount of variable cost. Same like TVC, TC also varies with the level of output.

Short-Run Average cost concepts: In the short run, there are three types of average costs: (i) Average fixed cost (AFC); (ii) Average variable cost (AVC) and (iii) Average total cost (ATC).

Average Fixed Cost: AFC is the per unit fixed cost. It can be calculated by dividing total fixed cost with the number of units of output produced. For instance, AFC was 200 for the first units of output. But as quantity of output increases, the AFC started declining. When firm produced 10th unit of output, AFC declined to Rs. 20.

$$AFC = \frac{TFC}{Units\ of Output}$$

Table 4.4: Calculation of Total Cost, Average Cost and Marginal Cost (in Rs.)

Units of Output	TFC	TVC	TC	AFC	AVC	ATC	MC
0	200	0	200	0	0.0	0.0	-
1	200	120	320	200.0	120.0	320.0	120
2	200	150	350	100.0	75.0	175.0	30
3	200	170	370	66.7	56.7	123.3	20
4	200	202	402	50.0	50.5	100.5	32
5	200	250	450	40.0	50.0	90.0	48
6	200	305	505	33.3	50.8	84.2	55
7	200	385	585	28.6	55.0	83.6	80
8	200	515	715	25.0	64.4	89.4	130
9	200	715	915	22.2	79.4	101.7	200
10	200	1015	1215	20.0	101.5	121.5	300

Average Variable Cost: Average variable is per unit variable cost. It is total variable cost divided by the units of output. For instance, AVC was 120 for the first units of output. But as quantity of output increases, the AVC started declining. AVC was nil when the quantity of output was zero. As production increases, AVC initially started declining and after certain level of output it started increasing (Table 4.4). It reached to its minimum when output increased upto five units and as firm continuously expanded its output AVC started increasing from 6th unit onward (Table 4.4).

$$AVC = \frac{TVC}{Units\ of Output}$$

Average Total Cost: The average total cost is the total cost of per unit of output. It is the sum o average fixed cost and average variable cost. As shown in Table 4.4, average total cost for the first unit of output was 320 and as firm increased the level of output ATC started declining and it continued to decline even when AVC starts increasing (Table 4.4). It continued to decline upto 7th unit of output. As firm continued the expansion of output after 7 units, average total cost started increasing.

$$ATC = AFC + AVC$$

Marginal Cost: Marginal cost refers to the increase in total cost caused by one more unit of output. In other words, MC is the addition to total cost of producing n units of output instead of n-1 units. For instance, MC for the first unit of output was 30 and as firm increased its production from 2 to 3 units, MC for the 3th unit reduced to 20 (Table 4.4). After reaching to its minimum point it started increasing. Symbolically, it can be written as:

$MC_n = TC_n - TC_{n-1}$

$$Or\ MC = \frac{Change\ in\ Total\ Cost}{Change\ in\ output}$$

Graphical Presentation of Average Costs Curves and Marginal Cost Curve

Understanding the relationship between AFC, AVC, ATC and MC are very critical. Except MC, other three are

very closely related to each other. It is so because, AFC is independent to the level of output as firm has to make investment in the form of plant and machinery as well as employ some managerial staff. These costs do not change with the level of output. On the other hand, other costs such as AVC, ATC and MC are the function of level of output.

Fig. 4.14: Short-Run Average and Marginal Cost Curves
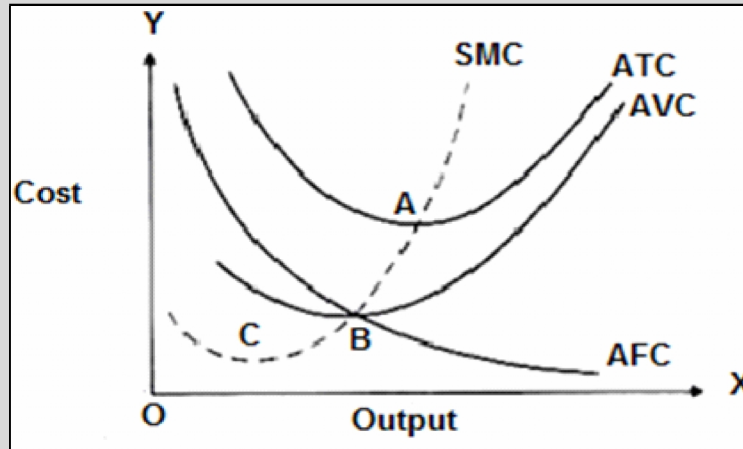
Figure 4.14 presents the relationship between short-run ATC, AVC, AFC and MC curves. It is noticed that AFC has been continuously declining. On the other hand, AVC, ATC and MC have initially declined and after certain level of output it started increasing.

 – AFC is declining continuously owing to the fact that it is a fixed cost and as units of output increase, per unit (average) fixed cost decreases.

 – AVC is U-shaped as units of output increases, initially it decreases (upto point B in the figure 4.14) owing to the economies of scale with the increase in level of output. At point B, per unit (average) variable cost reaches its minimum point as economies of scale vanish. After point B, diseconomies of scale with the increase in level of output pulls up the AVC curve as after this point, increase is output is accompanied by greater increase in variable cost.

 – ATC is also U-shaped, it initially declines upto point A because of economies of scale and after point A diseconomies of scale start operating which pulls up ATC upwards. But the interesting point is that ATC continues to decline even when AVC starts increasing i.e. minimum point of ATC lies to the right of minimum point of AVC. It is so because ATC is sum of AVC and AFC. Till decline in AFC is greater than increase in AVC, the net effect will pull down the ATC whereas when the magnitude of decline in AFC becomes less than magnitude of increase in AVC, the net effect will pull up the ATC i.e. ATC will start rising upwards.

 – MC curve is U-shaped. It is important to note that MC curve cuts both ATC and AVC at their respective minimum points. Further, MC curve lies below AVC and ATC curve till AVC and ATC decline. MC curve lies above these two curves when AVC and ATC curves start rising.

Long-Run Costs

In the long-run all costs become variable. It means that in the long-run, a firm can hire more quantity of both capital and labour, more of raw materials and other inputs, while technology remains constant. In other words,

there are no fixed factors in the long run and therefore there are no fixed costs. All factors are variable and as a result all costs are variable. For instance, if a firm closes down, its total cost (TC) also falls to zero. Similarly, TC increases with an increase in output, but its rate of increase may not be proportionate to the increase in output.

Average Cost Curve: Figure 4.15 presents the shape and derivation of long-run average cost curve. Suppose a firm has three options and corresponding to them, the short-run average cost curves are given. The short-run average cost curves corresponding to these plants are SAC_1, SAC_2, and SAC_3,

Fig. 4.15: Long run Average Cost Curve
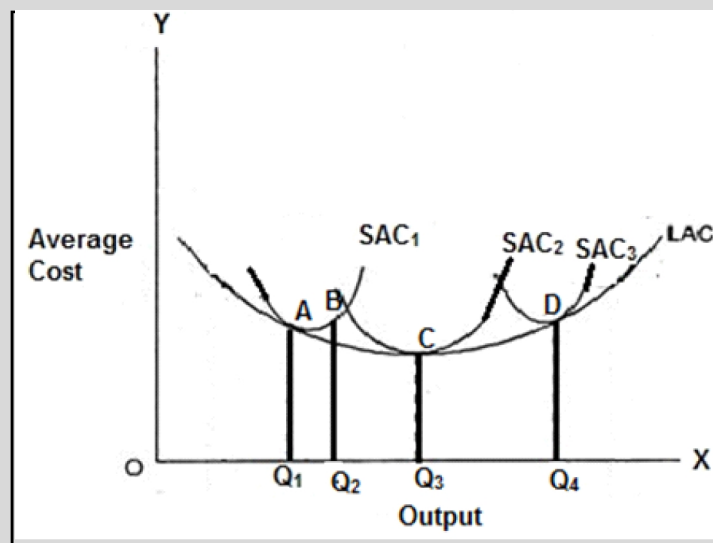
The firm decision about the size of plant depends upon the market size. If demand is small, the firm will use 1 for the purpose of production but doing so it will have to incur higher cost average cost. If the firm has to produce OQ_2 level of output, it has two options: (i) It can use plant 1. (ii) It can expand capacity and use plant 2.

It is important to note that SAC of plant 2 is lower than that of first plant, that is, per unit cost of production is lower in plant 2 as compared to plant 1 (Figure 4.15). This is due to the tendency of economies of scale. Economies of scale arise from several factors with the increase in production. Due to economies of scale, per unit cost of production declines provided inputs prices remain same.

Plant C is larger size than plant 2, but the SAC_3 is higher than SAC_2 curve. The average cost increases as firm when the plant third is added. This happens because of diseconomies of scale. Point C is the minimum average cost point, any increase after point C would produce output at higher cost.

The long-run average cost curve can be drawn by drawing a curve tangent to SAC_1, SAC_2 and SAC_3 as shown in figure 4.15. LAC curve is also known as 'envelop curve', 'planning curve'. It is a planning curve because it guides to the entrepreneurs in planning to expand the production in future. LAC also behaves like SAC curve. As shown in figure 4.15, SAC initially decreases with the expansion in production capacity until the optimum utilization of the second plant. Further increase in production capacity by adding third plant makes LAC move upward because SAC_3 lies above the level of SAC_2. This tendency of LAC shows that when the scale of production increase, average cost initially declines till it reaches to the optimal production point, then it increases due to diseconomies of scale.

Long-Run Marginal Cost Curve:

Long-run marginal cost (LMC) cost curve refers the change in total cost associated with an increase of one unit

of output when all inputs are variable. The LMC is derived from the short-run marginal cost curve. The process of derivation of LMC is shown in figure 4.16.

Fig. 4.16: Long-Run Marginal Cost Curve
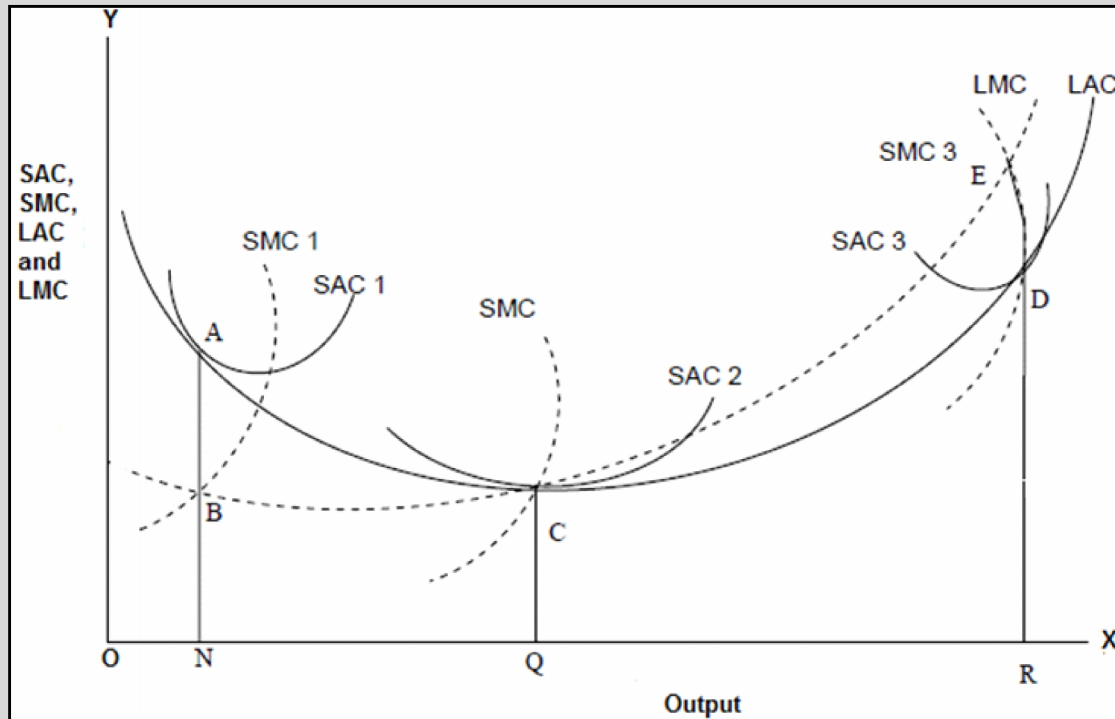
In figure 4.16, SMC curves are drawn corresponding to each SAC. To derive the LMC, consider the point of tendency between SAC curves and the LAC, that is, point A, C and D for the long-run planning curve. The tangential points determine the optimum output levels in the short-run with long-run production planning. Suppose, the perpendiculars are drawn from point A, C and D to the x-axis, the corresponding optimum output levels are ON, OQ and OR. The perpendicular AN intersects SMC_1 at point B. It implies that LMC is BN at output ON. In this manner, SMC can be obtained for different plants. For instance, at output level OQ, marginal cost is CQ determined by the intersection of LMC and SMC_2. Similarly, the perpendicular DR determines the optimum level of SMC_3 for output DR. When perpendicular DR is extended upward intersects SMC_3 at point E. Thus, the EQ measures the LMC at output OQ. If a curve is drawn through points B, C and E as shown by LMC, the curve represent the behavior of marginal cost in the long run. This is the LMC curve.

ECONOMIES AND DISECONOMIES OF SCALE

The term economies of scale denotes saving in cost of production with an increase in the scale of output or the size of the plant. It should be noted that the existence of economies does not mean a reduction in total cost in absolute terms. It only means a reduction in relative terms and manifests itself in a reduction in average cost of output. Similarly, the term diseconomies refer to an increase in average cost of output when the plant size is increased.

Economies and Diseconomies of Scale are broadly divided into two categories;

– Internal economies and diseconomies of scale

– External economies and diseconomies of scale

Internal economies and diseconomies of scale are those economies which occur due to endogenous (or internal) causes. For example, in case of a firm, internal economies and diseconomies arise on account of some actions of the firm itself.

In contrast, external economies and diseconomies of scale are the result of some exogenous or external causes which are not directly connected with the activities of the firm.

The sources of economies and diseconomies are obviously the causes on account of which they come into existence. These sources can be many and are most likely to vary from one situation to another as also from firm to firm and industry to industry. For this reason, it is possible to classify them in a large variety of groups.

Internal Economies of Scale

Internal economies of scale are associated with the expansion of the scale of output of the firm; they are derived indirectly as a result of expansion of the industry to which it belongs. Listed below are some of the leading sources of such economies.

Managerial Economies: These economies arise on account of the scope of employing better qualified and trained managers and other employees who are able to take quicker and more profitable decisions. In addition, management experts are exploring new methods of improving the management of the firm and reducing its cost of operations.

Financial Economies: It is a common knowledge that most firms have to depend upon borrowed funds. The lenders, while deciding the rate of interest to be charged on their loans, give due importance to the 'creditworthiness' of the borrower. And other things being equal, bigger firms enjoy greater creditworthiness than the smaller one. Accordingly, they are able to borrow funds at lower interest rates. For the same reason, they have also the option of raising additional sources through equity capital.

Technical Economies: With an increase in the scale of output, the choice of inputs and their varieties becomes wider for the firm. It can go in for those machines and equipment etc. which have a higher marginal productivity as compared with their cost. In other words, it is possible to get a larger output per unit of cost incurred on them.

Bye-Products: An increase in the scale of output also generates bigger flows of wastes. When the scale is small, the firm is not able to use these products for additional earnings. However, when the generation of waste crosses a critical limit, it often becomes possible for the firm to produce certain bye-products or sell off the waste to other firms and thus add to its income.

Better Utilization of Inputs: Various inputs, particularly machines and equipment are lumpy and indivisible. They also require time intervals for 'maintenance' and 'servicing' etc. Any one of them can go out of order and require repairs. If a machine goes out of order, or is otherwise not able to operate, then a firm with a small scale is not able to find its substitute and its production suffers. For example, if a transport company has only one truck and that needs some repair, its employees are left unemployed for the time being, though the firm has to pay them all the while. In contrast, a firm with a bigger scale is able to adjust the availability of its machinery, equipment and employees etc. in such a manner that the 'downtime' of various inputs is adequately taken care of.

Economies of Inventories: A bigger-size firm is in a better position to adjust its stocks of inputs and finished products etc. in such a manner that the normal discrepancy between flows of production and sales are ironed out.

Marketing Economies: A large firm also reaps the advantages of buying and selling in bulk. As a result, it is able to procure its inputs at concessional prices. Similarly, on account of bulk selling, its average selling costs come down. It can also have separate sales and marketing departments which can undertake the job of marketing its product in a professional manner. In addition, its sheer size imparts it better bargaining strength.

Advertising: When a firm is not operating under conditions of perfect competition, it is obliged to undertake various activities to promote its sales of which advertising happens to be an integral part. It is found that a small firm is not able to afford advertising because it has to be repetitive to be successful. Moreover, with an increase in the advertising budget, a firm is able to diversify its programme so as to cover more effective media and in an optimum proportion. As a result, its per unit advertising expenses come down.

Risk Economies: A large firm can diversify its product lines and thereby reduce the average risk faced by it since all product lines are not likely to generate losses simultaneously. The firm can compensate its losses from some lines with profits from the others. A large firm has also better command over resources compared with a small firm.

Internal Diseconomies of Scale

In the long run, a number of factors may cause an addition to average cost of a firm. These sources of diseconomies are directly or indirectly linked to the inability of the problems of management which not only crop up when a firm is started, but keep increasing with its size Chamberlin terms them 'Complexities of Management'. Briefly these are as follows:

Limits of Entrepreneurship: A firm, by definition, can have only one unit of entrepreneurship though it can have a large number of units of every other input. Moreover, entrepreneurship, in the ultimate analysis, manifests itself in the form of decision-making human beings who have to bear the consequences of their decisions. Therefore, the willingness to take risk varies between individual entrepreneurs and between different situations. Very frequently, wrong or sub-optimal decisions are taken which result in avoidable costs.

Similarly, with an increase in the size of the firm and its activities and departments, the entrepreneurship has to institute a system of delegating responsibility of assessing the ever-changing situation and taking on the spot decisions. It has also to set up an effective monitoring system. All this adds to the operating costs of the firm. In addition, it is found that on account of increase in the size of the firm, a number of additional departments may become necessary (such as a department which maintains leave accounts of the staff). As the administrative set up expands, delays set in and a number of avoidable expenses have to be undertaken.

Managerial Autonomy: With increasing size of the firm, various managerial wings (such as technical, financial, sales, marketing and others) have to be given sufficient autonomy of decision-making. The managers heading these departments, however, do not have a direct personal stake in the success of the firm. In addition, problems of coordination between various departments of the firm have to be tackled. The net result is often diseconomies of scale.

X- inefficiency: It is argued that the traditional theory is wrong in assuming that a firm always uses its resources in the best possible manner. This is because, in practice, most of the ground level decisions are taken by individual managers and other employees of the firm. And their motivation is advancement of their own careers rather than optimizing the performance of the firm. Quite frequently, they also lack adequate incentive to do their best. Unless there is a very sound system of 'reward and punishment' for each employee, their performance is likely to lack necessary motivation.

External Economies of Scale

These economies are those which are reaped by a firm not on account of its own efforts and increase in its scale but on account of the expansion and growth of the industry to which it belongs and also on account of overall development of the economy and markets.

Some of the major sources of these economies are as under:

Economies of Information: Availability of information is cheaper when we consider the industry as a whole. A firm requires continuous information regarding the prices of inputs and its product, as also likely changes in

them on account of shifts in government policies and other developments. If the industry as a whole sets up facilities for providing information through various means of communication, it is much cheaper, than if it is done by the firm alone. As a result, when the authorities or the industry as a whole provides means of information, it becomes economical for the firm to use the same sets up.

Research and Development: The results of research undertaken by the authorities or by the industry as a whole are economical for the firm to use. Also, it is generally commercially viable for a research organisation to undertake research on its own and sell the findings to individual firms on payment basis than for the firms to undertake the same individually.

Economies of Concentration: When an industry is concentrated in a certain locality or region, its firms get incidental saving in costs in the form of cheaper and more reliable services. These services cover, for example, repairs, consultancy, banking, credit, insurance, financial advice, packing, transport, housing, communication, , training, housing, health care, and so on. An individual firm is able to make use of these services at competitive and economical prices.

Economies of Specialization: When a number of associated and interlinked industries get located in the neighborhood, they all provide support to each other and their costs come down. As a result, the individual firms also benefit from this development.

External Diseconomies of Scale

As in the case of external economies, external diseconomies of scale are thrust upon a firm by extraneous forces. They are not directly related to the growth in the size of the firm under consideration.

There are several reasons for the prices of inputs to move up or their quality to deteriorate. For example, the authorities may impose a tax on one or more inputs. Or there may be a wage revision in the industry or imposed by the authorities on the economy as a whole.

Similarly, it is also possible that prices of some imported inputs may go up because of one or more reasons including for example, customs duties imposed by the domestic government or by the exporting country or increase in cost of production abroad, etc.

It is possible that on account of war, strikes, some natural calamity, or quantitative restrictions imposed by domestic country or a foreign country, the availability of an essential input may become costlier or insufficient.

LESSON ROUND UP

– Production is an important economic activity. Satisfaction of human wants is the objective of production using factors of production namely; Land, Labour, Capital and Entrepreneurship.

– Theory of production studies the economic process of converting inputs into outputs, given the state of technology. It provides a basis for analysis of relation between costs and amount of output.

– The production function is a technical relation connects factor inputs and output. It describes the laws of proportion, that is, the transformation of factor inputs into products (outputs) at any particular time period.

– Which one method would be chosen by the firm at any particular point of time will depend on the prices of factors of production.

– A technically efficient method of production is not necessarily economically efficient. There is a difference in technical and economic efficient method.

– An isoquant can be of four different shapes:

- – Linear isoquant

- – Kinked Isoquant

- – Input-output isoquant

- – Convex Isoquant

– There are two types of production function which we will discuss under theory of production:

  - – Short Run Production Function

  - – Long Run Production Function.

– The law of variable proportion shows the production function with one input factor variable while keeping the other input factors constant. It is a short run concept. The law of variable proportion is the modern approach to the 'Law of Diminishing Returns (or The Laws of Returns).

– Three types of returns to a factor

  - – Increasing returns to scale

  - – Diminishing Returns to Scale

  - – Negative Returns to Scale

– Law of Returns to Scale is a long run phenomenon. In the long run, all factors of production become variable. Alternatively, no factor of production is fixed in the long run. Thus, the firm is able to alter its stock of inputs in long run by increasing all the factors of production which is not the case with one variable factor. When all factors are changed in some proportion, the behaviour of output is analyzed with the help of laws of returns to scale.

– Least cost combination: Least cost combination refers to that combination in which consumer has to achieve either of the following two conditions:

  - – To produce maximum output from a given level of inputs.

  - – To produce a given level of output at minimum cost.

– Accounting Cost: Accounting costs are the cash payments which a firm makes to other factor owners for purchasing and hiring for various factors of production. It is also called as explicit costs.

– Economic Costs: in addition to all these accounting costs, economic costs take in account the all the implicit costs which the entrepreneur could have earned if he would have invested his own money and sold his own services in the next best alternative uses.

– Opportunity Cost: Opportunity cost is the cost of any next best alternative use that is sacrificed. Suppose a piece of land can be used to grow both wheat and gram. If a farmer uses that piece of land only to grow wheat then he sacrifices the production of gram.

– Social cost is the sum of the private cost and the net of negative externalities over positive externalities.

– Short-Run Costs:

  - – Fixed Cost

  - – Variable Cost

– In the short run, there are three types of average costs:

  - – Average fixed cost

– Average variable cost

– Average total cost

GLOSSARY

Factors of Production	A firm needs various types of inputs which are known as 'Factors of Production', 'Agents of Production', 'Productive Resources' or sometimes even 'Productive Services'. In economics, the factors of production are divided into four categories: (i) Land, (ii) Labour, (iii) Capital and (iv) Entrepreneurship.
Fixed inputs	Fixed inputs, in economic sense, are one whose supply is inelastic in the short-run and therefore they are used in a fixed quantity in the short-run.
Variable Factors	Variable factors, in economic sense, are defined as one whose supply in the short-run is elastic.
Isoquant	An isoquant is the locus of all the combinations of factors of production which produces a given level of output. Alternatively, an isoquant includes all the technical efficient method of production which produces a given level of output. The production isoquant can be of different shapes depending upon the degree of substitutability of factors of production.
Total Product	It refers to the total amount of goods and services produced by a manufacturer with given amount of inputs and keeping other factors remains constant.
Marginal Product	It represents the successive increase in the total output due to the additional units of variable factor
Average Product	Average product of a factor is the total output produced per unit of the factor employed. It can be obtained by dividing the total output by the total units of factor employed.
Short run	Short run refers to a period of time in which quantities of one or more factors of production cannot be changed.
Long run	Long run is defined as that time period over which a firm can vary quantities of all factors of production and therefore, can switch between different scales. In the long run production function all inputs are variable.
Returns to Scale	Law of Returns to Scale is a long run phenomenon. No factor of production is fixed in the long run. Thus, the firm is able to alter its stock of inputs in long run by increasing all the factors of production which is not the case with one variable factor.
Least cost combination	The least cost combination of factors refers to a firm producing the maximum volume of output from a given level of inputs on the one hand and on the other hand a firm produces a given level of output with the minimum cost when the factors are combined in an optimum manner.
Isocost line	Isocost line is defined as the locus of various combinations of factors which a firm can buy with a given outlay. The isocost line is also known as the equal cost line or outlay line.
Theory of Cost	Theory of cost analyzes the relationship between cost and output, that is, how costs of production changes with changes in production. The relationship between cost and output is called cost function.

Economic Costs	In addition to all these accounting costs, economic costs take in account the all the implicit costs which the entrepreneur could have earned if he would have invested his own money and sold his own services in the next best alternative uses.
Private Cost	Private cost is the cost made by a firm for manufacturing a commodity. It is the sum of both explicit cost and implicit cost as a business firm takes into account all these cost while making both production decision and decision about prices of the product it has manufactured.
Cost Function	A firm's cost function is the functional relationship between its output and its cost. Cost function is a determined by the production function of the firm and the prices it pays for different inputs. Hence, the theory of production forms the basis of the theory of cost.
Long-Run Costs	In the long-run all costs become variable. It means that in the long-run, a firm can hire more quantity of both capital and labour, more of raw materials and other inputs, while technology remains constant.
Internal Economies of Scale	Internal economies of scale are associated with the expansion of the scale of output of the firm; they are derived indirectly as a result of expansion of the industry to which it belongs.
Internal Diseconomies of Scale	In the long run, a number of factors may cause an addition to average cost of a firm. These sources of diseconomies are directly or indirectly linked to the inability of the problems of management.
External Economies of Scale	These economies are those which are reaped by a firm not on account of its own efforts and increase in its scale but on account of the expansion and growth of the industry to which it belongs and also on account of overall development of the economy and markets.

MULTIPLE CHOICE QUESTIONS

1. **Law of diminishing returns to factor is related to**

    a. Long-run

    b. Short-run

    c. Market period

    d. Both a and c

2. **Law of constant returns to scale is related to**

    a. Long-run

    b. Short-run

    c. Market period

    d. Both b and c

3. **Which of the following is not correct about land?**

    a. Unlimited supply

b. Immobility

c. Heterogeneity

d. None of the above

4. Cost incurred on plant and machinery is

a. Fixed cost

b. Variable cost

c. Sunk cost

d. None of the above

5. Which of the following denotes cost of sacrificed alternatives

a. Implicit cost

b. Accounting cost

c. Economic cost

d. Opportunity cost

6. Additional cost incurred on producing one more unit is known as

a. Average Cost

b. Marginal Cost

c. Implicit cost

d. Accounting cost

7. Which of the following factors are variable in the short-run

a. Labour

b. Capital

c. Both a and b

d. None of the above

8. Perfect substitutability of factors of production is assumed by

a. Input-output isoquant

b. Linear isoquant

c. Convex isoquant

d. Kinked isoquant

9. Producers would be incentivised to produce more as a result of increase in

a. Price

b. Cost of production

c. Competition

d. None of the above

10. Slope of an isoquant is given by

  a. dK/dL

  b. ds/dP

  c. dD/dP

  d. None of the above

Answer Key : 1. (b), 2. (a), 3. (a), 4. (a), 5. (d), 6. (b), 7. (a), 8. (b), 9. (a), 10. (a)

Suggested Readings

  1. Misra and Puri:Modern Microeconomics

  2. H. L. Ahuja:Modern Microeconomics

  3. M. L. Seth:Principles of Economics

Lesson 5
Analysis of Markets

LEARNING OBJECTIVES

There are a variety of markets ranging from perfect competition to monopoly to monopolistic competition. The equilibrium of these types of market is determined by different conditions which are based on the revenues earned and costs incurred. Broadly speaking, the market forces of demand and supply determines the equilibrium. It leads us to understand the determination of price and output level of the firm and industry separately.

Out of the various types of market, we study perfect competition, monopoly and monopolistic competition. There are various distinguishing features of each market and several common characteristics as well.

This lesson of the study examines a broad range of markets and explains how the pricing and output decisions of the firm depend upon the different types of market structure.

███

BASIC CONCEPTS OF REVENUE

The amount which a firm receives from the sale of the output is known as Revenue. According to Prof. Dooley," The Revenue of a firm is its sales receipts or income."

Revenue of a firm can be divided into three types,

1. Total Revenue

2. Average Revenue

3. Marginal Revenue

Total Revenue

Total Revenue is the amount receipt from the sale of the output. The total revenue depends on the total sale of the firm and the price of per unit output.

TR = Q X P

Where TR = Total Revenue, Q = Quantity or unit sales, P = Per unit price of the commodity

Average Revenue

We can obtain the average revenue by dividing the Total revenue by total unit sold.

AR = TR / Q

Where AR = Average Revenue, TR = Total Revenue, Q = Total unit sold

It should be noted that AR or average revenue is equal to price, in the coming topics we shall discuss why the average revenue is equal to price.

Marginal Revenue

Marginal Revenue is the amount of money received from the sale of additional unit or in other words, additional revenue received by the sale of an additional unit.

MR = TR_n – TR_{n-1}

Or

MR = $\triangle$TR / $\triangle$Q

Where MR = Marginal Revenue, $\triangle$TR = Change in total Revenue, $\triangle$Q = Change in the amount of sale, TR_n = Total Revenue of n units, TR_{n-1} = Total Revenue of n-1 units.

MR pertains to change in TR only on account of the last unit sold, while AR is based upon all the units sold by the firm. Therefore, any change in AR results in a much bigger change in MR. Reduction in MR is far bigger than that in AR; and similarly, an increase in MR is also much bigger than the corresponding increase in AR. The two are equal only when AR is constant.

The firm will not sell any quantity if TR or AR becomes zero or negative. However, MR can become negative if the fall in price is big enough.

Graphically, therefore, we have the following relationships.

Since TR, AR and MR are equal for the first unit sold, therefore, the three curves start from the same point. TR curve slopes upwards so long as MR is positive. If MR is falling with an increase in the quantity of sale, then TR curve will gain height at a decreasing rate. It reaches its maximum height when MR curve touches X-axis. TR curve slopes downwards when MR curve goes below X-axis to become a negative figure.

A change in AR causes a much bigger change in MR. Therefore, when AR curve has a negative slope; MR curve lies below it and has a greater slope. Similarly, when AR curve has a positive slope, MR curve lies above it and has a greater slope. When AR curve is parallel to X-axis, MR curve coincides with it.

In case AR is a straight line, MR curve will bisect each perpendicular distance of AR from Y-axis. However, if AR curve is parallel to X-axis, then MR curve coincides with it.

The above graphical relationships between AR and MR are shown in Fig. 5.1. In the first part of Fig. 5.1, AR has a constant value DD'. Therefore, AR curve starts from point D and runs parallel to X-axis. Since AR is a constant, MR is always equal to AR and the two curves coincide with each other.

In the second part of Fig. 5.1, AR curve starts from point D on Y-axis and is a straight line with a negative slope. It indicates that as quantity of good sold increases, it's per unit price falls at a given rate. Accordingly, MR curve also starts from point D and is a straight line. However, it is a locus of all those points which bisect the perpendicular distances between AR curve and Y-axis. For example, FM= MA.

Fig. 5.1 : Relationship between AR and MR
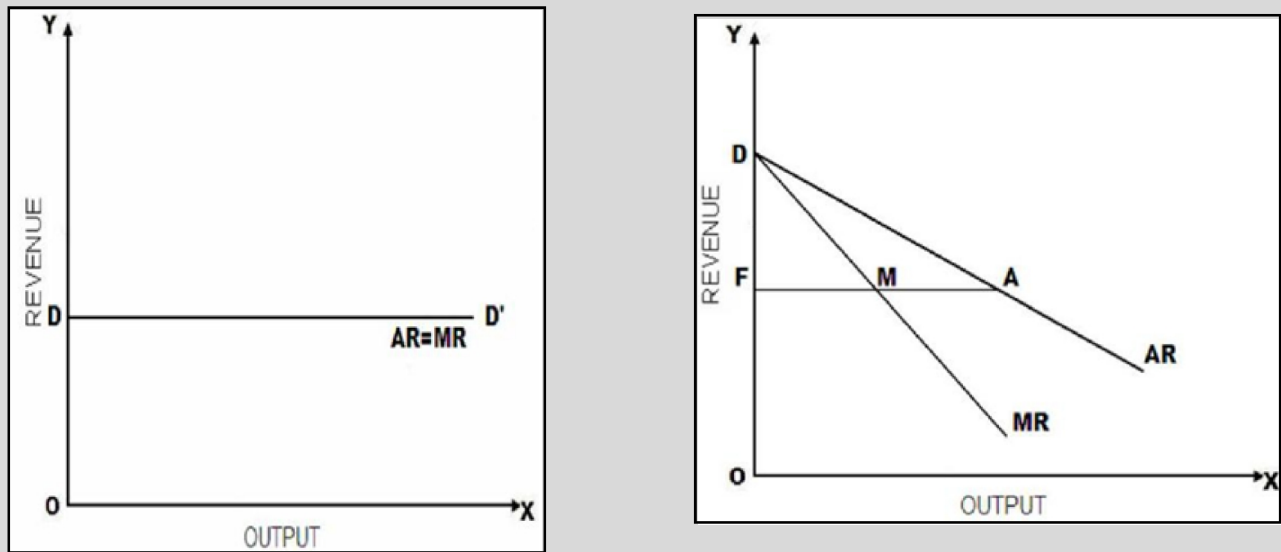
CONCEPT OF MARKET AND MAIN FORMS OF MARKET

The concept of market is central to the understanding of the determination of price and output of a commodity under consideration. In ordinary language, the term market refers to a public place in which goods and services are bought and sold. In economics, it has a different meaning. Different economists have tried to define market in different ways.

Cournot defines market as, "not any particular market place in which things are bought and sold, but the whole of any region in which buyers and sellers are in such free intercourse with each other that the prices of the same goods tend to equality easily and quickly".

To Ely, "Market means the general field within which, the force determining the price of particular product operate".

Stonier and Hague explain the term market as "any organisation whereby buyers and sellers of a good are kept in close touch with each other".

Thus a market has the following basic components :

– *BUYERS*: There should be buyers of the product. If a country consists of people who are very poor, there can hardly be market for luxuries like cars, LCDs etc.

– *SELLER*: A commodity should be offered for sale in the market. Otherwise there is no question of buying the commodity. Therefore, existence of sellers is a necessity for any market.

– *CONTACT*: Buyers and sellers should have close contact with each other.

– *PRICE*: There should be a price for the commodity. The exchange of commodities between buyers and sellers occurs at a particular price which is mutually agreeable to both the buyers and sellers.

This is because, in a modern economy, most of the production does not take place for self-consumption by the producers themselves. Thus, by the term market of a good, it should not be taken to mean a place where the buyers and sellers meet each other and conduct purchase and sale transactions. The market consists of two components;

– A Firm

– An Industry

A FIRM

A firm is a business unit engaged in the task of producing and selling of goods or services. It is identified by the fact that it is only one unit of entrepreneurship. The entrepreneurship may not be provided by a single individual. It may be exercised jointly by a board or a group of individuals in some defined manner. However, the firm has a unified and coordinated authority of decision making. In essence, these decisions relate to the objectives (such as, profit maximisation, or sales maximisation, etc.) and other policy decisions (such as, what to produce) of the business unit.

AN INDUSTRY

An industry is a set of firms which are conceptually closely associated in the sense of having some common type(s) of activities. A good example of an industry is a set of firms which are producing a certain type of a manufactured good or providing a certain type of service. The good supplied by the firms of the industry may be homogenous in the sense that the buyers believe that products of all firms are perfect substitutes of each other.

A consignment of the good, by itself, does not provide the identity of the supplier. The buyers, therefore, are indifferent between the sources of supply. Instead, they select between alternative supplier firms by comparing the prices being asked by them. In contrast, the industry may also comprise those firms which are producing differentiated products. This means that the product of one firm can be distinguished from that of the other. The buyers do not consider the products of different firms as perfect substitutes. However, if the firms are to belong to one industry, then their products must be close substitutes of each other.

There is no pre-determined number of firms which an industry must have. Their number can vary according to the structure of the market. At one extreme, it may have only one firm in which case it is called a monopoly or a one-firm industry. At the other extreme, an industry may have such a large number of firms that each of them accounts for an extremely small portion of the total supply of the industry and is not able to influence the price of the product. Between these two extremes, there can be several other possibilities.

There are various kinds of markets prevailing in the economy. The following will be discussed in the present lesson :

– PERFECT COMPETITION

– MONOPOLY

– MONOPOLISTIC COMPETITION

– OLIGOPOLY

EQUILIBRIUM OF THE FIRM

Meaning of Equilibrium

A firm is said to be in equilibrium when it has no tendency to change its level of output, that is, when it has no tendency either to increase or contract its level of output. The firm can earn maximum profit at the equilibrium point. The meaning of the word 'equilibrium' is state of balance. At that point where two opposite forces became equal, that point is known as equilibrium point. In case of a firm, in equilibrium position a firm will produce equilibrium quantity of a product and will charge as much price at which the equilibrium quantity of the product can be sold out in the market.

Objectives of the Firm

The main objectives can be explained as follows,

1. Achieving a target rate of return

2. Stabilization of price and profit margin

3. Realizing a target market share

4. Preventing price competition

5. Maximization of sales or sales revenue.

According to the classical theory of the firm, the sole aim of the business is profit maximization. Every entrepreneur wants to maximize his profit. But after some time critique of classical theory came into existence, according to that profit maximization can be a sole aim of business only in single owner entrepreneur. But now the picture has been changed, the salary and the perks of managers are more closely related with the sale of the firm rather than profit. Financial institutions keep keen eyes on the sale of the firms; they are more interested to finance those firms which have increasing sales. The realisation of a target market share is also the main objective of the firm.

Two Approaches to Producer's Equilibrium

Producer's equilibrium refers to the level of output of a commodity that gives the maximum profit to the producer of that commodity. Given the assumption of profit maximisation by the firm, it will be in equilibrium when there is no scope for either increasing its profit income or reducing its loss by changing the quantity of output i.e. it cannot improve its economic position by changing the output.

Profit(π) **= Total Revenue – Total Cost**

= TR – TC

Therefore, the output level at which 'total revenue less total cost' is maximum is called the equilibrium level of output. There are two approaches to arrive at producer's equilibrium.

– TR-TC Approach

– MR-MC Approach

TOTAL REVENUE – TOTAL COST APPROACH

According to this approach, there are two conditions of producer's equilibrium:

– The difference between TR and TC is maximum

– Profit falls if one more unit of output is produced (that is marginal cost becomes higher than marginal revenue if one more unit is produced)

Fig. 5.2 : Equilibrium of a Firm using TR and TC Curve
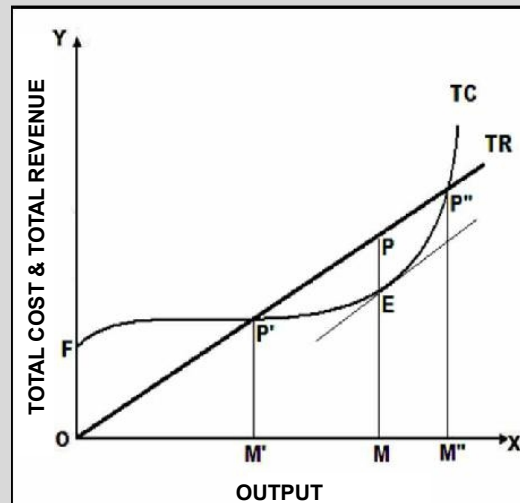
In the Fig. 5.2, X axis shows levels of output and Y axis shows costs and revenues.

TR is Total Revenue Curve and TC is Total Cost Curve. Further, P is equilibrium point, where the distance between TR and TC is maximum.

It can be seen that TC > TR before point P' and after P''. Therefore, producer will produce between P'P'' or M'M''. At the equilibrium point P, it can be seen that the tangent drawn to TC is parallel to TR i.e. at point P, slope of TC is equal to slope of TR. At any other point than P between P'P'' or M'M'', this equality is not achieved.

MARGINAL REVENUE – MARGINAL COST APPROACH

MR – MC approach is another way of identifying producer's equilibrium. It is derived from the TR – TC approach. The two conditions of equilibrium under MR-MC approch are as follows :

1. MR = MC

2. MC cuts the MR curve from below

MR = MC

When one more unit of output is produced, MR is the gain and MC is the cost to the producer. Clearly, so long as benefit is greater than the cost, or MR is greater than MC, it is profitable to produce more. Therefore, so long as MR is greater than MC, the firm has not achieved equilibrium level of output where the profit is maximum. The equilibrium is not achieved because it is possible for the firm to add to profits by producing more.

The producer is also not in equilibrium when MR is less than MC because benefit is less than the cost. By producing less, the producer or firm can add to his profits. When MC is equal to MR, the benefit is equal to cost, the producer is in equilibrium subject to condition that MC becomes greater than MR beyond this level of output.

When MC equals MR, the producer's profit would be less if he produces output more than or less than the 'MR = MC' level of output as explained above. Therefore, for attaining equilibrium 'MC equals MR' is a necessary condition (but not sufficient).

2. MC is greater than MR after MR = MC or MC cuts MR from below

MR = MC is a necessary condition but not sufficient to ensure equilibrium.

It is because the producer may face more than one MR = MC outputs. But out of these only that output beyond which MC becomes greater than MR is the equilibrium output. It is because if MC is greater than MR, producing beyond MR = MC output will reduce profits. And when it is no longer possible to add to profits the maximum profit level is reached.On the other hand, if MC is less than MR beyond the MR = MC output, it is possible to add to profits by producing more. Therefore this MR = MC level is not the equilibrium level. For a producer to be in equilibrium, MC must be equal to MR as well as MC must become greater than MR if more output is produced.

PERFECT COMPETITION

Perfect Competition in economic theory has a meaning diametrically opposite to the everyday use of the term. In practice, businessmen use the word competition as synonymous to rivalry. In theory, Perfect Competition implies no rivalry among firms. Perfect Competition, therefore, can be defined as a market structure characterized by complete absence of rivalry among the individual firms. That is to say perfect competition is a market structure where there is a perfect degree of competition and single price prevails.

According to Prof. Benham," A market is said to be perfect when all the potential sellers and buyers are promptly aware of the price at which transactions take place and all of the offers made by other sellers and buyers, and when any buyer can purchase from any seller, and vice versa." In the words of Prof. Marshall," The more nearly perfect a market is, the stronger is the tendency for the same price to be paid for the same thing at the same time in all parts of the market."

Perfect Competition Market is a hypothetical market structure where in every seller takes the market prices as the price of his own product, firms are incapable of influencing the market price either by acting singly or in a group.

Main Features

(i) *Homogeneous Product:* In a perfect competition, it is not possible to distinguish between the products of individual firms. There are no distinctive features of the product associated with any specific firm. The product, in that sense, is homogeneous and undifferentiated. To the buyer, product supplied by one firm is a perfect substitute of that supplied by another.

(ii) *Large Number of Sellers:* Perfect competition is characterized by a large number of firms. Here, the term large denotes the fact that no individual firm is in a position to significantly influence the total supply of the industry and thereby affects the price of the product. Every firm in the industry is thus, a price taker. It can sell any quantity of its own product at the going price. For it, the demand for its product is perfectly elastic. It, of course, must be remembered that the maximum quantity, which this firm can supply, is insignificantly small when viewed in relation to the aggregate supply of the industry as a whole.

(iii) *Large Number of Buyers:* Perfect competition is also characterized by a large number of buyers who are in competition with each other for the available supply. Their number is so large that any single buyer may change the quantity purchased without significantly affecting the total demand in the market and affecting the price of the product. Like an individual firm, an individual buyer is also a price taker. He can buy any quantity of the product he likes at going price. To him, the product has perfect elasticity of supply.

(iv) *Full Knowledge of Market:* It is assumed that in perfect competition, every buyer and seller has full knowledge of the prevailing price of the product, as also the prices being asked by the sellers and being offered by the buyers. This 'perfect knowledge' enables every buyer and seller to make use of any opportunity that may exist to strike a better bargain.

(v) *Economic Rationality:* Economic rationality is another feature of perfect competition. It means that every buyer and seller is motivated by his own economic interest in his decisions to buy or sell. This, coupled with the assumption of perfect knowledge, ensures that a uniform price prevails in the market.

(vi) *No Transportation Cost:* It is assumed that there is no transaction cost to be incurred by buyers and sellers in their activities. The price paid by a buyer is exactly equal to the price received by the seller. There is no resource cost in terms of time or other expenses to be incurred i.e. there are no transaction costs. In particular, a seller has no need to incur any selling expenses (say, in the form of advertisements) because his product is not differentiated from the products supplied by other sellers.

(vii) *Free Entry and Exit:* Perfect competition is also characterized by free entry and exit. Basically, the terms *entry* and *exist* apply to the suppliers, though their coverage can be extended to buyers also. It means that, given enough time, any existing firm can close down and leave the industry or any new firm can enter the industry. There is no legal, institutional, or technical hurdle in ding so. It is only estimated economic benefits or losses that guide the firms in these decisions. Similarly, any existing buyer of the product can increase his purchases, cut them or reduce them to zero. New buyers can also enter the market and offer to buy any quantity they like.

(viii) *Perfect Mobility of factors of production:* The one of the main features of perfect competition is perfect mobility of factors of production. The factors of production are free to move from one firm to another. It is also assumed that workers can also move between different firms, which imply that skills can be learned easily.

(ix) *No Government Regulation :* It is also assumed that there is no government regulation, government has no interference in terms of tariff, subsidies, etc.

EQUILIBRIUM OF THE FIRM UNDER PERFECT COMPETITION

Earlier we have discussed the various types of markets; we have also discussed the various objectives of the firm. Now we shall discuss the equilibrium of the firm under perfect competition, that what level of output an individual firm will decide to produce.

Under perfect competition, the firms are unable to alter the price of the product by changing the quantity of its own output. The prices of the input are given; therefore, cost conditions are also given. In other words, under perfect competition, it can only decide to alter the quantity of its output without changing price of the product.

A firm is said to be in equilibrium when its profits are maximum, which in lieu depends on the cost and revenue conditions of the firm. The concepts of cost and revenue vary in short run and long run. Thus a competitive firm has four equilibrium states differing on the basis of period of operation as follows:

- Short Run equilibrium of a Competitive Firm

- Long Run equilibrium of a Competitive Firm

- Short Run equilibrium of a Competitive Industry

- Long Run equilibrium of a Competitive Industry

The demand curve of a product under perfect competition

Now we shall discuss the derivation of firm's demand curve, with the help of market demand curve and market supply curve. In perfect competition the price given to the firm has been decided by the equilibrium of the market demand and market supply. It is shown in the Fig. 5.3. Price has been shown on Y axis and Quantity on X axis. The first panel of Fig. 5.3 represents inudstry and the second panel respresents the case of a firm.

Fig. 5.3 : Determination of Market Price under Perfect Competition
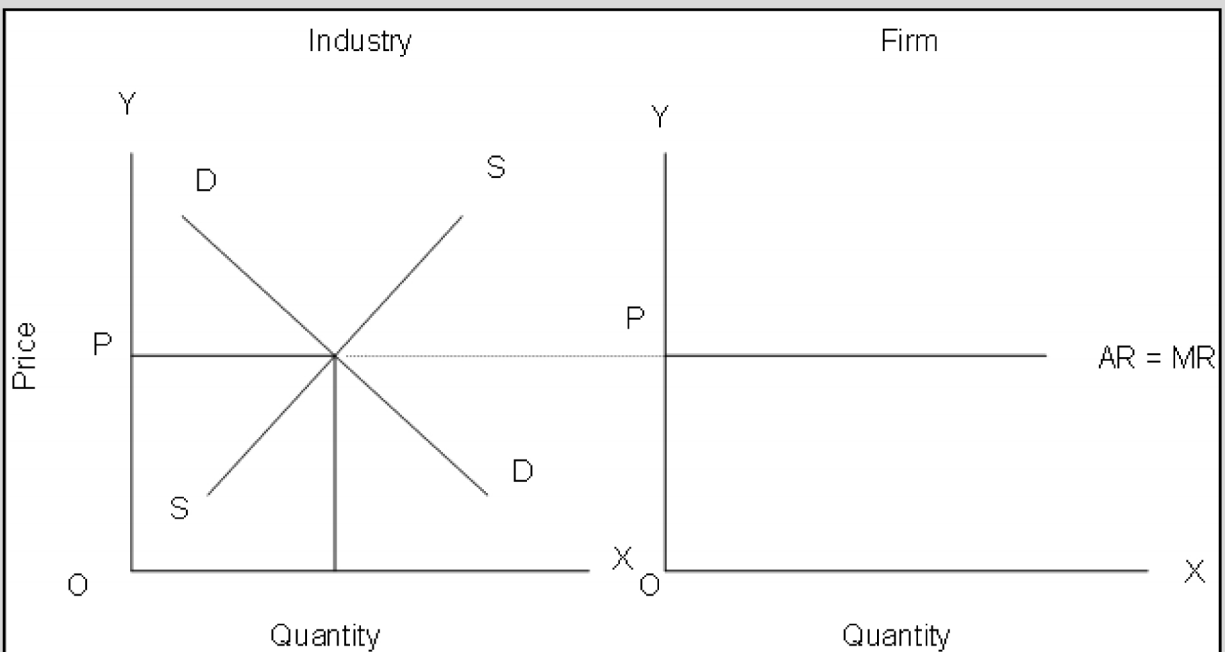
In the figure, market demand curve is DD, market supply curve is SS. On X axis we have assumed quantity and on Y axis we have assumed price of the commodity. The point at which market demand curve and market supply curve intersect each other, that point is known as equilibrium or market equilibrium. The price at this equilibrium is known as equilibrium price and the quantity corresponding to this equilibrium is known as equilibrium quantity.

The price decided in the market with the help of market demand and market supply is given to the existing firms under the perfect competition. Every firm in perfect competition is a price taker and in other words the industry is price maker. AR curve also becomes the demand curve of a firm as the firm can sell any amount it wants at the market price.

Short Run Equilibrium of a Competitive Firm

Under the short run period, the following are the major assumptions;

- Price of product is given in the market at which a firm can sell any quantity

- Plant size of firm is given(constant)

- Firm is facing given short run cost curves

The two conditions of equilibrium of a firm are : MR = MC & MC cuts MR from below. Graphically, this condition states that for equilibrium of the firm, MC curve should intersect MR curve from below and, after intersection, lie above MR curve. If we translate this condition in ordinary words, it means the following. The firm should keep on adding to its output as long as MR > MC because additional output adds more to its revenue than to its cost and

thus its profit income increases. Furthermore, if its MC is equal to MR but the firm finds that by adding to its output, MC becomes smaller than MR, then the firm should decide to increase its output.

Fig. 5.4 : Equilibrium of a Firm using MC and MR Curves
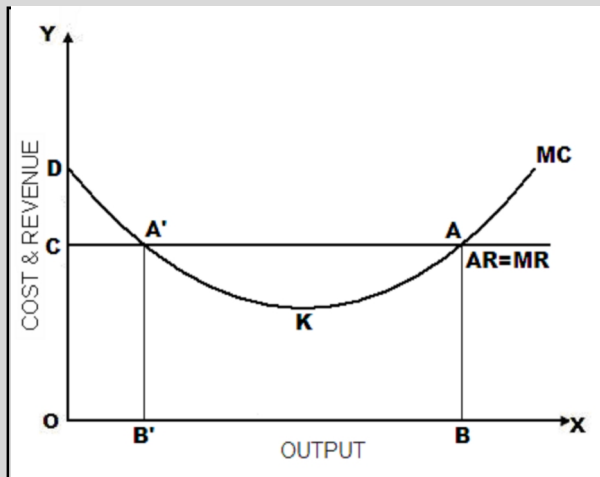
On account of perfect competition, the demand for the product of the firm is perfectly elastic. The firm can sell all its output at the going price in the market. Accordingly, its demand curve (AR curve) runs parallel to X-axis throughout its length and its MR curve coincides with AR curve.

As regards the supply side, we should recall the set of four cost curves of the firm, namely, the AFC, the AVC, the MC and the ATC. Out of these, the supply curve of the firm is that portion of the MC curve which lies above AVC curve and is upward sloping. The actual equilibrium of the firm is determined by the intersection of its supply and demand curves. An explanation of this phenomenon is provided below.

To begin with we note that, in the short run, the firm cannot avoid fixed costs. They have to be incurred even if production is reduced to zero. However, the variable costs are directly related to the quantity of output. The implication is that, in the short run, the firm cannot avoid losses by not producing. Therefore, it decides to continue production even at a loss, provided the loss does not exceed its fixed costs. It means that the firm would decide to produce if its average price (that is, per unit price of the good) equals or exceeds its AVC.

On account of the law of variable proportions, the average variable cost curve is U-shaped. Let us also recall that marginal cost represents a change in the total cost so that it is related only to the variable costs and not fixed costs. And since AVC curve is U-shaped, MC curve is also U-shaped. It lies below AVC curve when the latter is downward sloping. However, MC curve starts rising when the rate of fall in AVC curves slows down, intersects it at its lowest point, and rises above it.

We have also seen above that a firm attains its best possible position (that is, the position of maximum profit or minimum loss) when its MC curve cuts its MR curve from below. At the same time, price per unit of the product must be able to recover at least the average variable cost. When the price exceeds AVC, the firm is able to recover a part of its fixed costs also with a resultant reduction in its losses. In case the price equals the average of total cost, the firm is able to recover its full costs (including the component of 'normal profit'). And if the price is still higher, it earns an abnormal profit.

Three possibilities in short-run

There are three possibilities for a firm's equilibrium under perfect competition, a firm can earn normal profit

abnormal profit (supernormal profit) or it can bear the loss. At the equilibrium quantity, if the average cost is equal to average revenue, it means the firm is earning normal profit. In second case if the average cost is greater than the average revenue the firm is bearing the loss. In third case if the average cost is less than average revenue than the firm is earning the abnormal profit.

Normal Profit

Let's discuss the first possible situation of a firm under perfect competition; in this situation a firm is earning normal profit.

Fig. 5.5 : A Case of Normal Profits
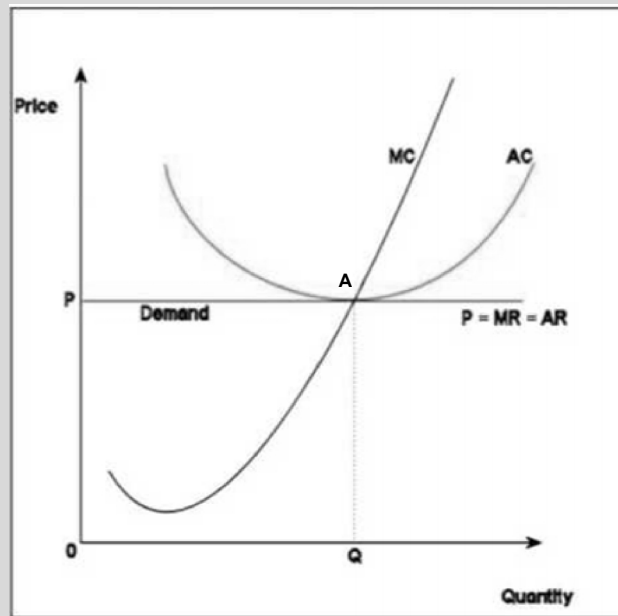
In the Fig. 5.5 costs and revenue have been assumed on Y axis and the quantity has been assumed on X axis, Marginal cost cuts the marginal revenue from below at point A. Corresponding to point A, P is equilibrium Price & Q is equililbrium quantity.

It should be noted that corresponding to equilibrium qauntity, average cost is also equal to average revenue. Or in other words per unit cost is just equal to per unit price of equilibrium quantity. It means the firm is earning normal profit.

Loss

The second possible situation of a firm's equilibrium under perfect competition is to incur losses.

Fig. 5.6 : A Case of Loss-making
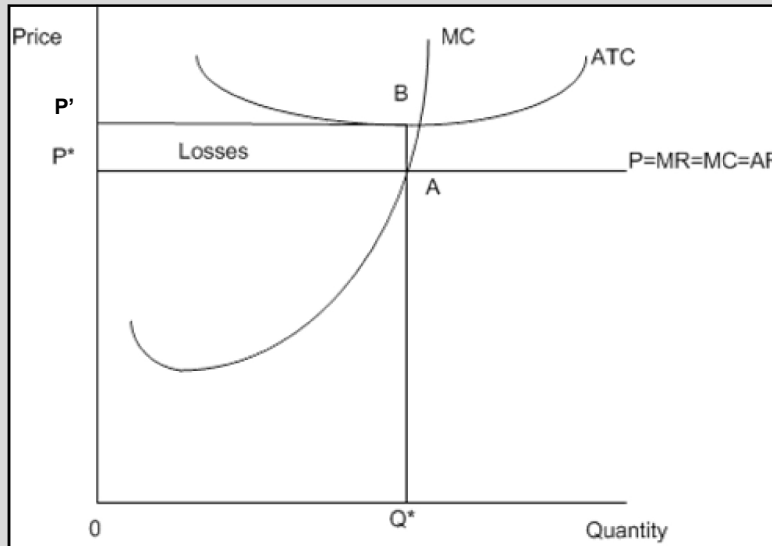
In the Fig. 5.6, again the cost and the revenue curves have been assumed on Y axis, and the quantity demanded has been assumed on X axis. As per the first equilibrium condition, marginal cost curve cuts the marginal revenue curve at point A, this point is known as equilibrium point. Corresponding to point A, P* and Q* are equilibrium Price & Quantity respectively. Also, corresponding to Q*, the average cost is more than the average revenue, or in other words average cost curve lies above the average revenue curve. In this case per unit cost of OQ* i.e. average cost is more than per unit revenue of OQ* i.e. average revenue. According to the figure per unit revenue is OP and per unit cost is OP', it means that per unit loss is PP', and the total loss on the quantity OQ*, is P*P'BA.

Supernormal Profit

Now we shall discuss the third possible situation of firm's equilibrium in short run, under perfect competition.

Fig. 5.7 : A Case of Super-normal Profits
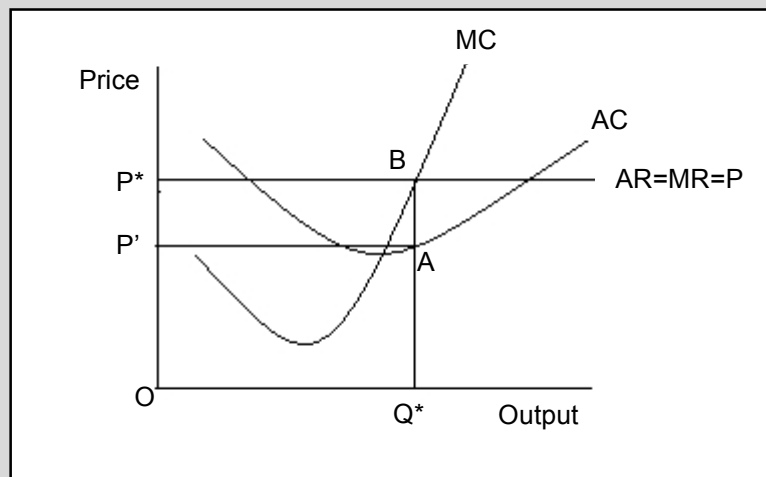
In short run all the three situations are possible, a firm can earn normal profit, abnormal profit (supernormal profit) or it can bear the loss. In the Fig. 5.7 per unit revenue or average revenue is OP* while per unit cost or average cost is OP', per unit receipts are high in comparison with per unit cost. That's why the average revenue curve lies above than average cost curve corresponding to Q*. In the above case the firm is earning abnormal profit, per unit profit is P'P* and the total profit for the quantity OQ* is P'P*BA.

These three possibilities are once again recapitulated in a single diagram [Fig. 5.8]

Fig. 5.8 : Equilibrium of a Competitive Firm in Short Run
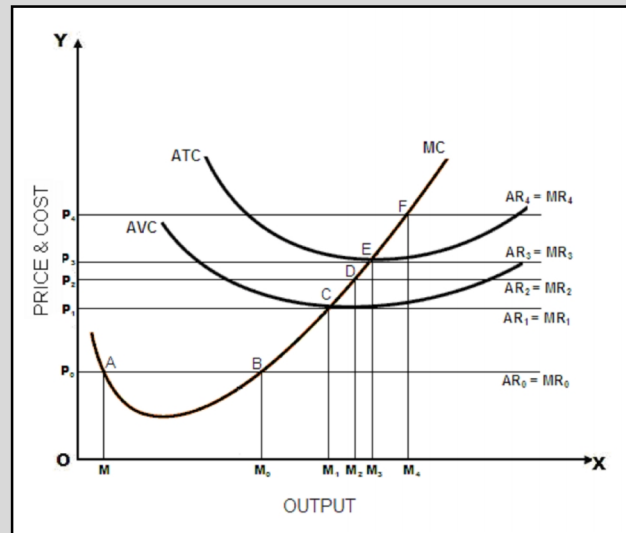
In Fig 5.8, we consider five different prices to illustrate the supply behaviour and associated equilibrium of the firm. There is an average revenue curve corresponding to each price. It runs parallel to X-axis and the MR curve also coincides with it.

1. When the price is OP_0, the corresponding MR_0 curve cuts MC curve at two points, A and B. At point A, none of the conditions of equilibrium is satisfied. At point B, MC curve cuts MR_0 curve from below but AR is still less than AVC. Therefore, the firm incurs a loss greater than its fixed cost if it decides to produce when the price is OP_0. The firm, therefore, decides to close down but it cannot leave the industry.

2. If the price happens to be higher and equal to OP_1 (that is, equal to the least possible average variable cost), the firm decides to produce. In this case, not only MC curve cuts MR_1 curve from below (a point C); AR_1 is also equal to AVC. Thus, we find that either the firm does not produce at all, or it produces at least equal to OM_1. Such a price indicates shut-down point of a firm.

3. In the third case, price (OP_2) exceeds AVC but is still less than ATC. MR_2 and MC curves intersect each other at point D. The firm produces OM_2. It still incurs a loss but less than its fixed costs because it is able to recover a portion of the latter.

4. In case the price rises to OP_3, the firm is able to recover its full cost including fixed costs. Its MC curve cuts MR_3 curve from below at point E and AR_3 = ATC. All the conditions of its equilibrium are satisfied. It produces OM_3.

5. If the price rises even further, say, P_4, the point of intersection of MR_4 and MC curves moves to F. The firm is able to recover not only its total cost but is able to earn an abnormal profit also. It produces OM_4.

It should be noted from above that in the short run, existing firms can close down but they cannot leave the industry and new ones cannot enter it. Therefore, when our firm is incurring a loss, it continues production so long, its' loses do not exceed fixed costs. Similarly, if it earns an abnormal profit, they are not wiped out by new firms entering the industry.

Long Run Equilibrium of a Competitive Firm

Long term is defined as that period in which the firm has the opportunity of varying all its inputs. There are no fixed costs and therefore average fixed cost curve vanishes. The average cost (AC) curve denotes average total cost (ATC) curve. More precisely, in the long run the firm can decide to go in for any of the alternative plants of different scales.

We have seen earlier that in the long run, the average cost (AC or LRAC) curve of the firm formed by its short run average cost curves (that is, plant curves) is also U-shaped. Up to a certain scale, there are increasing returns and LRAC curve slopes downwards. This is followed by the phase of constant returns in which LRAC curve is neither rising nor falling. And the third phase is that of diminishing returns to scale in which LRAC curve slopes upwards. We have also seen that corresponding to the U-shaped LRAC curve, long term marginal cost (LRMC) curve is also U-shaped and that it cuts LRAC from below at the lowest point of the latter.

Since the firm can vary all its inputs in the long run, it follows that it has the option to close down and leave the industry. Similarly, new firms can also enter the industry. This condition, termed 'free entry and exit of firms' has two implications.

– The firm is not compelled to operate when incurring a loss. It can leave the industry.

– No firm is able to earn an abnormal profit (that is, a profit in excess of the 'normal'one). It can only earn 'normal profit' which forms a part of its costs and is incorporated in its LRAC curve. This happens because abnormal profit earned by existing firms attracts new firms. And as they enter the industry, supply increases, price comes down and abnormal profit is wiped out.

Fig. 5.9: Long Run Equilibrium of a Competitive Firm
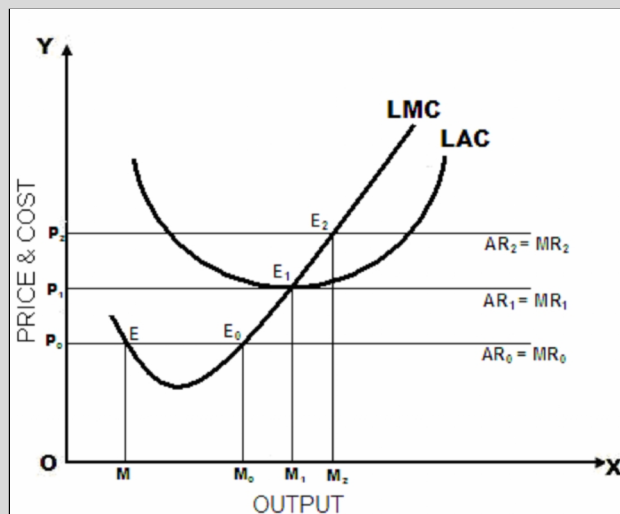
Determination of long run equilibrium of the firm under perfect competition is explained in Fig. 5.9 in which output is measured along X-axis and costs are measured along Y-axis. The firm is a price taker. For it the price of its product is given and fixed. It can sell any quantity it can produce at the going price. Its AR curve runs parallel to X-axis and MR curve coincides with it.

For the purpose of explaining the determination of firm's equilibrium, we are considering three alternative prices given to the firm by the industry.

1. Let us assume that the price in the market is below the optimum cost of the firm, say, OP_0 from which we get corresponding average revenue (AR_0) and marginal revenue (MR_0) curves. Now MR_0 curve cuts LMC curve at two points, E and E_0, but none of these is a long term equilibrium position of the firm. At point E,

LMC curve cuts MR_0 curve from above. At E_0, LMC curve cuts MR_0 curve from below, but $AR_0 < LAC$ and the firm incurs a loss.

2. Similarly, if the price of the firm's product is more than the optimum cost (least possible average cost) of the firm, the firm cannot be in stable equilibrium. Thus, let the price be OP_2 with average revenue curve (AR_2) and marginal revenue curve (MR_2). LMC curve is intersecting MR_2 curve from below at point E_2 and $AR_2 > LAC$. But the firm is not allowed to enjoy the abnormal profit. New firm enters the industry, supply increases and price falls till no firm can earn abnormal profit.

3. Long run stable equilibrium of the firm is attained at a point where when price of the product is equal to its optimum cost of production. In Fig. 5.9, this equilibrium price is OP_1. In this case, marginal revenue curve (MR_1) cuts LMC curve from below at lowest point of LAC (E_1). At the same time, AR_1 is equal to LAC, so that the firm does not incur a loss or earn an abnormal profit. At this stage, there is no incentive for existing firms to leave the industry or new ones to join it. The corresponding equilibrium output is OM_1.

In conclusion, we may also note that in long run equilibrium, the firm produces an 'optimum' output at the least possible average cost. It is this position where the firm is operating under 'constant returns' to scale. Consequently, its MC = AC. At the same time, MC = MR and AR = AC , so that we get AC = AR = MC = MR

EQUILIBRIUM OF INDUSTRY UNDER PERFECT COMPETITION

An industry comprises all the firms which are producing goods which the buyers consider substitutes of each other. As such the determination of price of such a product is the result of interaction between total demand for the output of all the firms taken together and their supply.

On the demand side, the important fact to be noted is that a change in its supply affects the price of the product also. The industry is not a price taker. Though the contribution of an individual firm in total supply is so insignificant that it cannot make any noticeable difference to the price of the product, this is not so with the industry. The change in supply made by the firms taken together alters the aggregate supply to such an extent that it cannot sell more without lowering the price. This results in a downward sloping demand curve for the industry.

The fact of a negatively sloped demand curve for the industry can also be understood as follows. A firm can sell more of its output by attracting customers from its competing firms. In the process, the total sales of the industry need not increase. But an industry can sell more when the existing buyers buy more of its product and/or new buyers enter the market and buy its product. Now it follows that existing buyers are already equating their marginal utility with the price. They would buy more only if price falls. Similarly, for the new buyers, the existing price is higher than the marginal utility of the product. And, therefore, they would also buy more of the good only if the price is reduced. Accordingly, the demand curve for the product of the firm must have a negative slope indicating that more of the product can be sold only by reducing its price.

The exact location and slope of the demand curve varies from product to product and also for the same product from one time interval to the other. There can also be several reasons on account of which its demand curve may shift in its slope and location. However, there is no theoretical basis for predicting these changes. Therefore, even while recognising that demand curve for a product can and often does shift over time, the economists assume that it retains its position when we move from short term to long term. In other words, demand curve for the industry is always drawn with a negative slope without specifically providing for a change in either its exact slope or its exact location.

Short Run Equilibrium of a Competitive Industry

We have noted above that when an industry changes the quantity of its supply, there is a corresponding change in the price of its product also. It follows, therefore, that when it is in equilibrium, there is no tendency on its part to change the quantity of its output. We also know that the price of the industry's product is determined by intersection

of its demand and supply curves. Having seen the nature of industry's demand curve, we may look into the nature of its supply curve which is obviously the summation of individual supply curves of the firms constituting it. That is to say, we get the industry's supply curve by adding the quantities which its firms are ready to sell at alternative prices. Since the supply curve of a firm, in the short run, is that upward sloping portion of its marginal cost curve which lies above its average variable cost curve, therefore, the supply curve of the industry also is upward sloping.

Fig. 5.10 : Short Run Equilibrium of a Competitive Industry
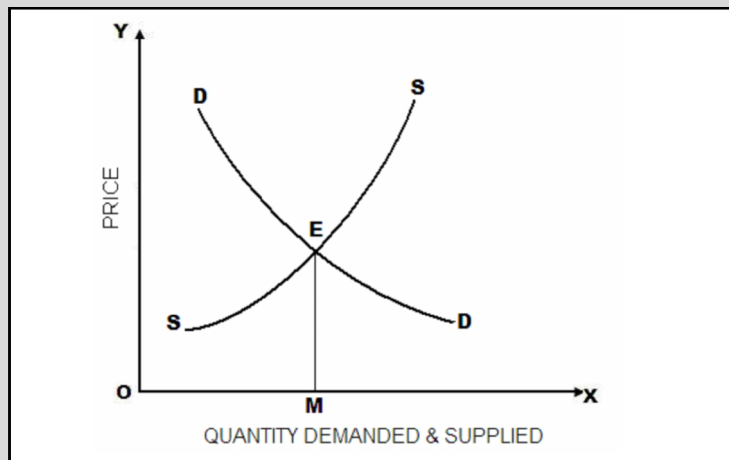
Short term equilibrium of the industry, which is determined by the intersection of its demand and supply curves, is illustrated in Fig. 5.10, in which quantities demanded and supplied are measured along X-axis and price per unit is measured along Y-axis. While the demand curve is downward sloping, the supply curve has a positive slope. In Fig. 5.10, short term demand and supply curves of the industry intersect at point E. The equilibrium price and output of the industry are, therefore, EM And OM respectively. It is noteworthy that in case the industry happens to be in a non-equilibrium position, it automatically gets adjusted to its equilibrium position This is because, in such a case, there is either an excess of supply or an excess of demand at the existing price. In case of excess supply, the firms are left with unsold stocks which they try to dispose off by reducing price. And in case of excess demand, some customers are not able to buy the quantities they wanted to. They, therefore, bid up price in competition with each other.

Long Run Equilibrium of a Competitive Industry

The slope of the demand curve of industry remains negative even in the long run, implying that it can sell more of the product only by reducing the price and vice versa.

However, the long run supply curve of the industry cannot be derived by horizontal summation of the supply curves of the individual firms. The reason is that in the long run, existing firms can leave the industry and new ones can join it. Moreover, it is possible that due to various reasons, the industry may suffer from some internal/ external diseconomies or enjoy some economies. These factors can shift the position of the long term supply curve of the industry. Another fact which complicates the derivation of long run supply curve of the industry is that while in the short run, individual firms may incur losses or enjoy abnormal profit, in the long run, these possibilities tend to be wiped out. When the industry is in equilibrium, its individual firms are also simultaneously in such equilibrium that they neither make an abnormal profit nor incur a loss.

Thus, long run supply curve of the industry is derived by taking into account all these determining variables. It is not derived by adding those portions of the MC curves of the individual firms which lie above the AC curves of the firms. Instead it is the locus of the pairs of those points which represent quantities of its output and the least average cost at which its firms can produce it.

Economists believe that the factors at work in the long run may result in one of the three situations, namely, that of diminishing, constant, and increasing returns, indicating the shift in the average cost of its firms. These three cases are explained below.

Constant Returns

In this case, the expansion of the industry does not lead to net economies or diseconomies. They are either not there or they get neutralised by each other. As a result, the average cost of the industry as a whole remains constant as it expands. Its demand curve is downward sloping and its supply curve is parallel to X axis. It means that while the quantity demanded responds, to a limited extent, to a change in price of the product, the elasticity of supply is perfect. The industry curtails its supply to zero if the price offered is reduced even by a small margin. On the other hand, at the going price, it is ready to sell as much as the buyers are ready to buy.

Fig. 5.11: Long Run Equilibrium of a Competitive Industry under Constant Returns
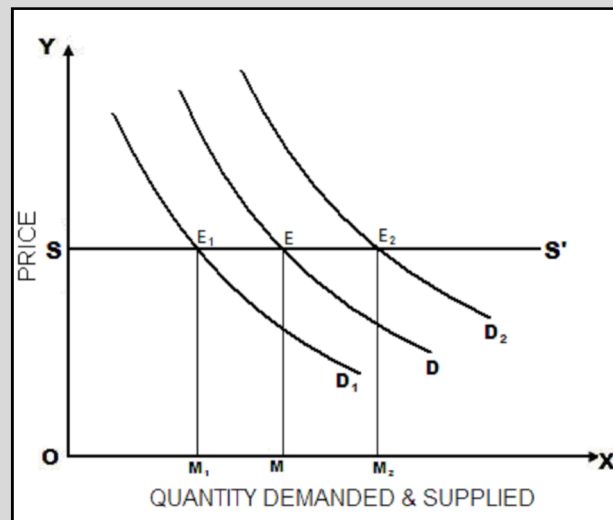
In Fig. 5.11, the supply curve of the industry is SS' and its demand curve is DD_1. The two intersect at point E_1 with OM_1 and E_1M_1 as equilibrium output and price respectively. At the same time, all the firms of the industry are also at equilibrium such that for each firm, its MC = MR = AC = AR and it operates at the least possible average cost. As a result, there is no incentive for any firm to leave the industry or enter it.

Fig. 5.11 also depicts the alternative equilibrium position if demand curve happens to be differently located. It is seen that shifting of the demand curve does not result in a change in price. It only results in a change in the quantity of supply.

Diminishing Returns

Let us assume that when existing firms of an industry expand their output, their average cost of production increases. Similarly, let us assume that if new firms enter the industry, they also face higher average product cost.

This implies that the expansion of the industry generates more diseconomies than economies and the net result is an increase in the average cost of production. The supply curve of the industry, therefore, slopes upwards and implies that the industry will be ready to sell more only if the price offered increases.

Fig. 5.12: Long Run Equilibrium of a Competitive Industry under Diminishing Returns
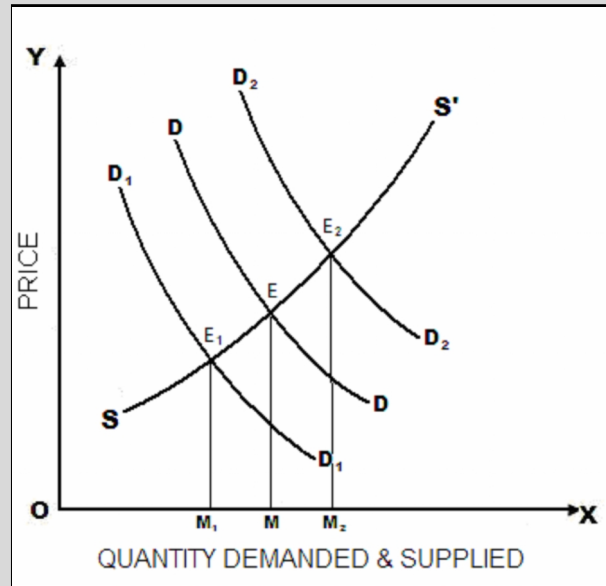
The demand curve of the industry is downward sloping in this case also. And its equilibrium is determined by the intersection of demand and supply curves. Diagrammatic illustration of determination of this equilibrium is given in Fig. 5.12, in which the demand and supply curves intersect at point E. The corresponding equilibrium price and supply happen to be EM and OM respectively. If there is an increase in demand so that the demand curve shifts to the position D_2D_2, there is an increase in both price and quantity of output. In contrast, if the demand decreases and the demand curve shifts to the position D_1D_1, there is a reduction in both price and quantity of supply.

Increasing Returns

In this case, the average cost of production of the industry declines with its expansion and its supply curve has a negative slope.

Some analysts claim that the very possibility of increasing returns is ruled out in a competitive industry. They say that it is logically contradictory to maintain that an industry can have increasing returns when its firms are facing diminishing returns. The critics point out that individual firms have U-shaped average cost curves even in the long run. As a result, when existing firms expand their output, their average cost is bound to increase, at least beyond a stage. It means that when an industry expands its output beyond a certain quantity, it is bound to face an increase in average cost of production.

Marshall's answer to the critics lies in his distinction between the effect on average cost of production when the expansion takes place in the industry as a whole instead of only in an individual firm. He maintained that it is possible that when a single firm expands it faces diminishing returns, but if all the firms expand simultaneously, there are increasing returns due to several economies that arise on account of their collective expansion. In other words, even when expansion of each individual firm results in an increase in its average cost of production, an expansion by the industry as a whole can still generate a variety of economies which are internal to the industry but external to the individual firms. For example, the expansion of the industry may cause an expansion in the industries supplying its inputs. This may lead to better quality and/or lower prices of certain inputs and other services etc. The industry may also find that, with an expansion in its size, adoption of certain technologies which were hitherto too expensive becomes economical. Several new specialised and technically efficient services (such as those of transport and repair works) may become available. As a result, while an individual firm continues facing diminishing returns, for the industry as a whole average cost may fall with expansion.

Fig. 5.13: Long Run Equilibrium of a Competitive Industry under Increasing Returns
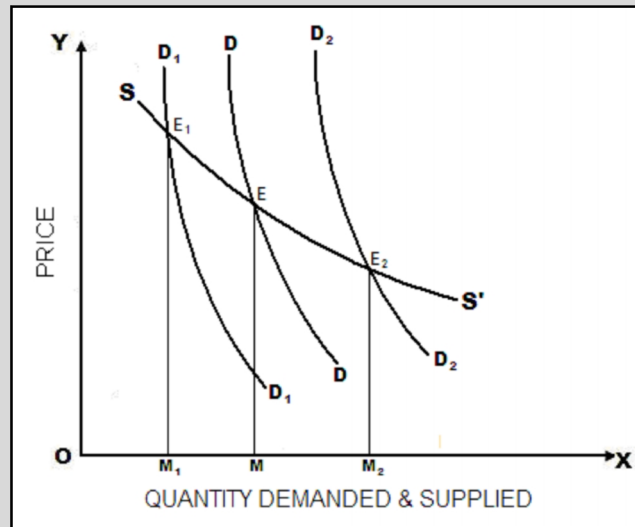
In Fig. 5.13, the supply curve of the industry SS' is negatively sloped depicting increasing returns, The demand curve, DD, is also negatively sloped, and the two intersect each other at E. It is noteworthy that for a stable equilibrium, it is necessary that after the point of intersection, demand curve lies below supply curve. It means that though the average cost of production falls with an increase in output, the price at which additional output can be sold falls still faster. In Fig. 5.13, this condition of stable equilibrium is satisfied. The industry produces an output of OM and sells it at a price of EM per unit.

It should be noted that when the industry is operating under increasing returns, an increase in the demand for its product will lead to an increase in quantity supplied but a fall in price. For example, in Fig. 5.13, with an increase in demand, the demand curve shifts from position DD to D_2D_2, This results in a new equilibrium position represented by point E_2 and a fall in price from EM to E_2M_2). In contrast, if there is a reduction in demand, the buyers will have to pay a higher price for a smaller output. For example, if the decrease in demand shifts the demand curve from DD to D_1D_1, output will fall from OM to OM_1 and the price will increase from EM to E_1M_1.

MONOPOLY

The term monopoly means a single seller. In economics, this term refers to a firm the product of which has no close substitute in the market. It is, in that sense, a single firm industry. Mono means single and poly means seller. It means a single seller. According to Prof. A.J.Braff, "under pure monopoly, there is a single seller in the market. The monopolist's demand is the market demand. The monopolist is a price maker. Pure monopoly suggests a no substitute situation."

Features of Monopoly

There are several features of monopoly; some of the main features of monopoly are as below :

1. Single seller and large number of buyers : The basic assumption of monopoly is single seller and large number of buyers. In monopoly firm is also an industry, there is no difference between firm and industry because there is only one seller or the producer. Firm's demand curve is also the demand curve for an industry. Because of the assumption of large number of buyers an individual buyer can not affect the price.

2. No close substitute : The second assumption of monopoly is that there is no close substitute of the product, produced by monopolist. If the close substitutes are available then the monopoly can not exist. Monopoly can only exist when the cross elasticity of the product, produced by the monopolist is equal to zero. In the absence

of a substitute product, the monopolist is free to fix a price of his choice. He can refuse to sell his product for a price below the one decided by him.

3. Strong barriers on entry of the new firms : Irrespective of the profit income of the existing producer firm, new firms cannot enter the industry. Hurdles to their entry may be on account of various reasons. There may be legal barriers, or the producer may own a technology or a naturally occurring substance which others cannot avail of. It is also possible that the size of the market may be too small and no new firm may find it economically worthwhile to enter it.

Revenue curves under Monopoly

As we know that the firm under monopoly is a price maker not a price taker, a Monopolist can increase or decrease the price. When there is a change in price, the average revenue and marginal revenue cannot be same. This is explained in Table 5.1 & Fig. 5.14.

Table 5.1 : Total Revenue, Average Revenue and Marginal Revenue

Quantity Sold	Price per Unit	(Total Revenue) TR	Average Revenue AR	Marginal Revenue MR
1	6	6	6	6
2	5	10	5	4
3	4	12	4	2
4	3	12	3	0
5	2	10	2	-2
6	1	6	1	-4

Fig. 5.14: Downward Sloping Revenue Curves
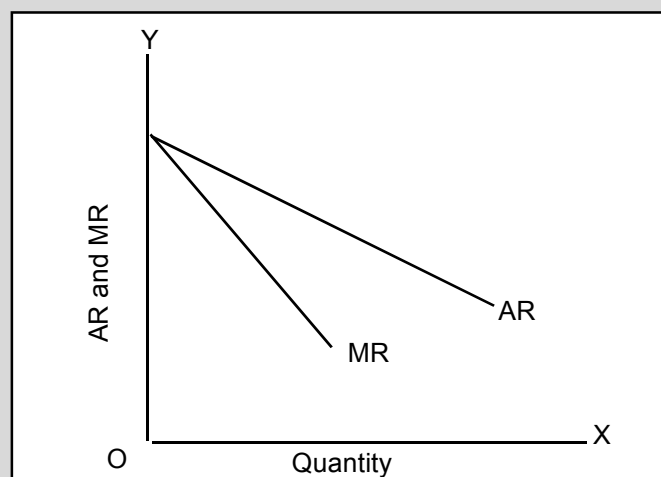
According to the 5.14, both the revenue curves, average revenue and marginal revenue are downward sloping. It is because of decrease in price.

If the monopolist wants to increase his sales, he has to reduce the price of his product so as to induce

– existing buyers to buy more

– new buyers to enter the market

Therefore, the demand conditions for his product are not the ones which are associated with a firm under competitive conditions. Instead, the demand conditions faced by him are similar to the ones which are faced by the industry as a whole. In other words, a monopoly firm faces a negatively sloped demand curve for his product. In the long run, the demand curve can shift both in its slope and location. However, there is no theoretical basis for determining the direction and extent of this shift.

As regards his cost of production, it may be assumed that a monopoly firm faces a given technology. Moreover, the monopolist faces conditions similar to those faced by a single firm under competitive conditions. He is not the sole buyer of the inputs used by his firm, but only one in the entire market. He has no control over the prices of the inputs used by him.

Equilibrium under Monopoly

The equilibrium conditions under monopoly are same as the equilibrium conditions of perfect competition. These are as follows :

1. Marginal Cost should be equal to Marginal Revenue.

2. Marginal Cost should cut the Marginal Revenue curve from below.

Short Run Equilibrium of a Firm under Monopoly

There are three possibilities for a firm's equilibrium under monopoly, a firm can earn normal profit, supernormal profit or it can bear the loss. If the average cost is equal to average revenue, it means the firm is earning normal profit, if the average cost is greater than the average revenue the firm is bearing the loss and if the average cost is less than average revenue than the firm is earning the abnormal profit.

Normal Profit

When the average cost is equal to average revenue corresponding to equilibrium output, the firm will earn normal profit. In the Fig. 5.15 MC cuts the MR curve at the point E, the equilibrium point. And AC touches the AR curve corresponding to that point. As we know that when per unit cost is equal to per unit revenue, the firm is earning normal profit.

Fig. 5.15: A Case of Normal Pfofits
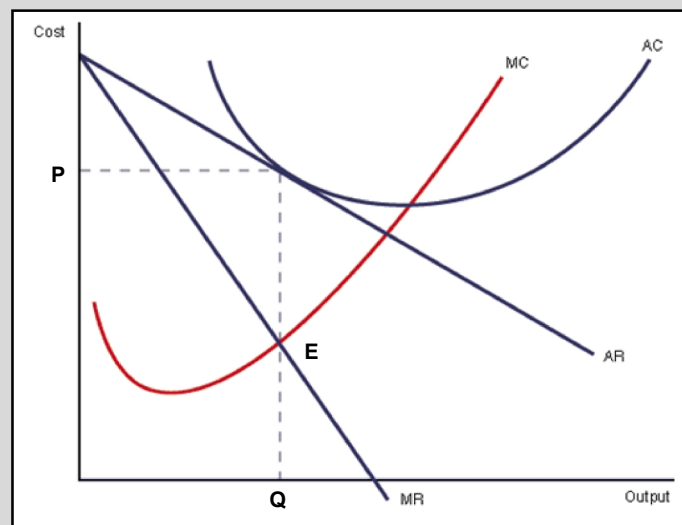
Supernormal profit

In the Fig. 5.16 price has been determined with the help of equilibrium, the price per unit is equal to OP or QA. According to the diagram per unit revenue or the average revenue is equal to OP, and per unit cost or average cost is equal to OP'. it means we are getting more and incurring less cost. In this situation, per unit profit is PP', that is OP – OP'. The total profit earned by the monopolist is equal to PP'BA.

Fig. 5.16: A Case of Supernormal profit
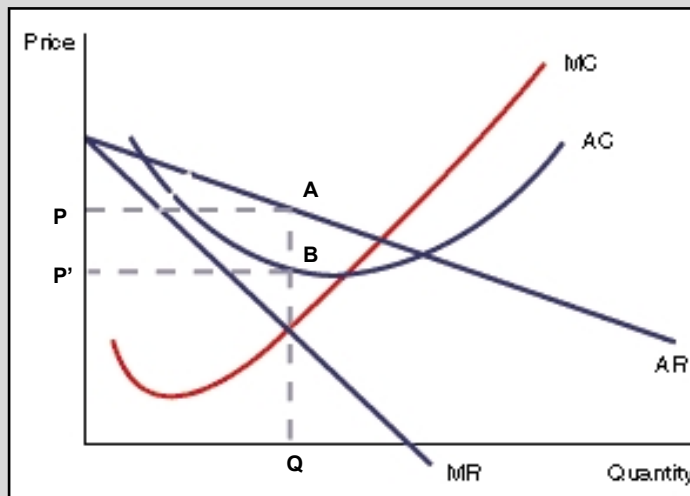
Loss

If the average cost is more than average revenue, in this case the firm has to incur more cost per unit in comparison with per unit revenue.

Fig. 5.17: A Case of Loss-making
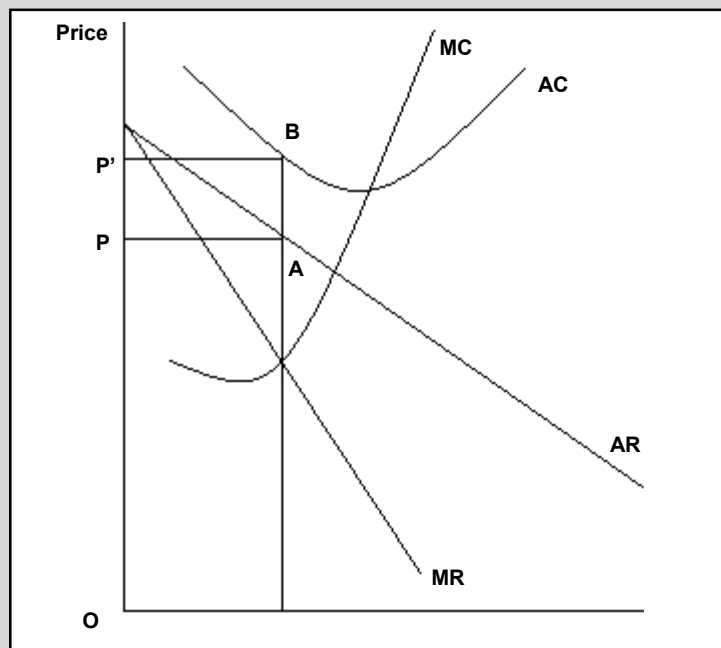
In the Fig. 5.17, we can see the average cost curve lies above the average revenue curve corresponding to equilibrium quantity. The average revenue or the per unit revenue is OP, and the per unit cost or the average cost is OP'. It means the firm is bearing the per unit loss PP'. The total loss incurred by the monopolist is equal to PP'BA. In short run, to make the entry unattractive for the new entrants, a monopolist can set a lower price and incur losses. But because of the monopoly power a monopolist earn only abnormal profit in the long run.

These three possibilities are once again recapitulated in Fig. 5.18

In the short run, by definition, the monopolist firm cannot vary all its factors of production. Its short run cost curves are similar to the ones faced by a firm operating in the short run and under perfect competition. It is also noteworthy that, in the short run, the monopolist may incur a loss but it will shut down the plant only if the loss exceeds its fixed costs. On the other hand, if the demand for its product is quite strong, it may make an extra profit.

Fig. 5.18 : Short Run Equilibrium of a Monopolist
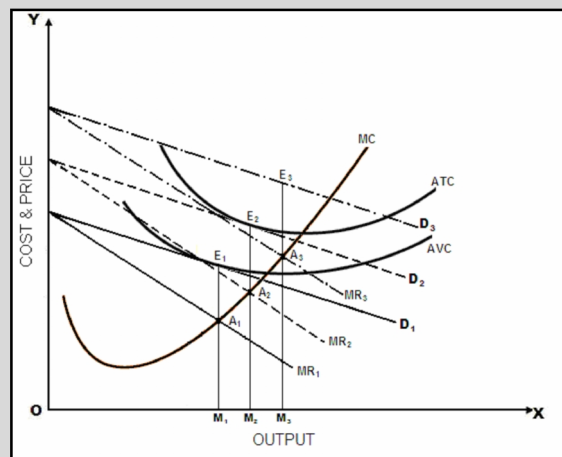
Determination of short run equilibrium of a monopoly is depicted in Fig. 5.18 with quantity of output measured along X-axis. Correspondingly, price and cost of production are measured along Y-axis. We have drawn three cost curves of the monopoly, namely, average variable cost curve, average total cost curve and marginal cost curve. Similarly, three alternative demand curves have been drawn to explain alternative possible positions of equilibrium.

1. Demand curve labeled D_1 is tangent to AVC curve at point E_1. Its corresponding marginal cost curve MC intersects MR_1 curve from below at point A_1. Thus, we note that while the first condition of equilibrium of the firm is satisfied, the monopolist is not able to recover his full cost of production. However, the loss which is equal to fixed costs cannot be reduced by closing down the plant. In this situation, therefore, the monopolist decides to produce OM_1 quantity of output, sell it at price E_1M_1 and suffer a loss equal to fixed costs. Note that there would be no production if the demand curve lies to the left of its position of D_1. In that case the monopolist would have added to his losses by operating his plant. Consequently, his best option would have been to close down the plant and minimize the loss to fixed costs.

2. If the demand curve lies to the right of D_1, the monopolist is able to recover a part of his fixed cost also. He is able to recover his full cost of production if the demand curve happens to be tangent to ATC curve (such as D_2). With D_2 as the demand curve, equilibrium position of the monopolist is given by the intersection of MC curve with MR_2 curve at A_2 which corresponds to the point of tangency (E_2) of D_2 with ATC curve. In this case not only MC curve cuts MR_2 curve from below but, at the same time, AR = ATC. The monopolist makes a normal profit by producing OM_2 and selling it at price E_2M_2.

3. The monopolist may earn an abnormal profit if the demand curve lies further to the right of D_2 (such as D_3). As before, the equilibrium position of the monopolist is determined with reference to the point of intersection between MC curve and MR_3 curve at A_3. The monopolist decides to produce OM_3 and is able to sell it at price E_3M_3.

Long Run Equilibrium of a Monopolist

In the long run, by definition, the monopolist can vary all the inputs. Therefore, the determination of equilibrium of the firm can be analyzed with the help of only two cost curves, namely, AC and MC. Moreover, the monopolist would not stay in the market if he is to operate at a loss. By implication, therefore, the demand curve must be tangent to the AC curve or must lie to the right and intersect it twice.

Fig. 5.19 : Long Run Equilibrium of a Monopolist
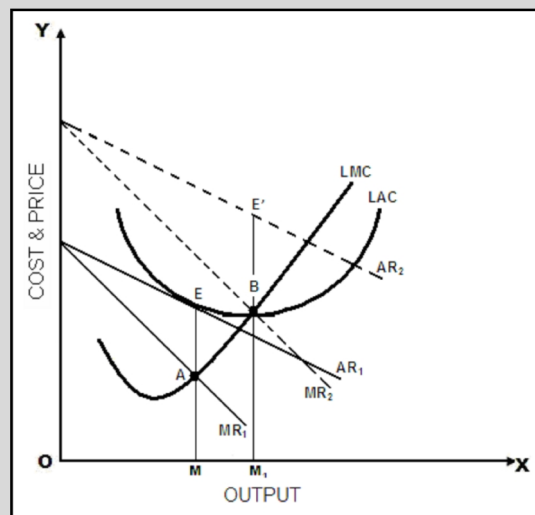
In Fig. 5.19, long term AC and MC curves of the monopolist are drawn U-shaped. The reasons for their being U-shaped have been discussed earlier. As before, equilibrium of the firm is that its MC curve should cut MR curve from below and its AR should be greater than or equal to AC. In Fig. 5.19, we depict two alternative cases of the determination of equilibrium of the monopolist, namely, (i) with normal profit, and (ii) with abnormal profit. Remember that the monopolist will not operate in the long run if there is a loss.

In the first case, demand curve (AR) of the monopolist is AR_1 and it is tangent to AC or LAC curve at point E. It should be noted that if demand curve lies to the left of AC curve, the monopoly is not able to recover its AC at any output and, therefore, it closes down and leaves the market. However, when AR curve is tangent to AC curve, the monopoly is able to recover its average cost (including normal profit) and is; therefore, ready to stay in the market and produce. Further, it should be noted that, by the very nature of relationship between MR and AR on the one hand and between MC and AC on the other, (i) the perpendicular drawn from point E to X-axis, (ii) the MC curve and (iii) MR curve are concurrent (at point A). As a result, all the conditions of equilibrium of monopoly are satisfied. Thus, at point A, MC curve of the firm cuts MR curve from below. At the same time, AR is equal to AC at point E. Accordingly; the monopoly produces OM and sells it a price EM per unit which covers its average cost including normal profit.

In the second case, the monopolist is able to make an abnormal profit. Its marginal revenue curve is MR_2 which cuts its MC curve from below at point B. The corresponding height of its average revenue curve (AR_2) is $E'M_1$. Accordingly, the monopolist produces OM_1 and is able to sell it at $E'M_1$ per unit with an extra profit of $E'B$ per unit. Moreover, this extra profit is not competed away because there is no substitute good in the market and no new firm can enter the market and produce it.

Discriminating Monopoly

Since a monopolist can fix the price of his product, it is possible that he may find it profitable not to charge a uniform price for his entire output but sell it at varying prices. The term used to denote this phenomenon is 'price discrimination' and the monopoly practicing it known as 'discriminating monopoly'. Thus, price discrimination is the absence of a policy of uniform price for the entire output. Instead, the monopolist sells portions of it at two or more prices. And this happens even when the product is homogeneous and one unit of it cannot be distinguished from the other. Professor Pigou mentions three types of price discrimination.

- The monopolist may charge a separate price for each unit sold by him. This is known as the price discrimination of first degree.

- The monopolist may sell his output in 'batches' or 'lots', charging a separate price for each batch or lot. This is known as price discrimination of second degree.

- The monopolist may split up the market for his product on the basis of the buyers, He may divide buyers into two or more categories and charge from each category a different price. This is known as price discrimination of third degree.

Reasons for Price Discrimination

There are three main reasons for price discrimination viz.

- Monopoly awareness of consumer ignorance for the cost of product

- There are two different markets

- Charge of discriminating prices for the same product or services from the different customers

The monopolist is aware of consumer ignorance for the cost of product due to lack of knowledge and communication of proper information. In rendering professional services or personal services discriminating price can be charged by a monopolist from different customers. A doctor specialist having monopoly in his professional specialisation can charge higher fee from rich and lesser fee from poor clients. Such discrimination is possible when the service rendered or commodities cannot be resold.

Price discrimination may be practiced under the circumstances when cost difference might exist due to distance between one market or another, lower price in poorer market and higher prices in sophisticated market could be charged. Such price discrimination occurs when firm's different markets are separated by distance or by national frontiers; cheaper commodity in one market may not be resold at dearer markets because of excessive cost of transportation.

Equilibrium under Price Discrimination

Profit maximisation is the basic objective before a discriminating monopolist. For this purpose he would like to avail of every opportunity what is available to any ordinary monopolist. The basic rule according to W.J. Baumol which he has to follow is that marginal revenue in all the markets where he is practicing discrimination should be similar. Monopolist can allocate his product in different markets on the basis of available MR from the best combination. The total market shows the aggregate AR and MR and the profit maximizing output of OQ. The optimum level of output for the firm is determined at OQ since at this level of output MR = MC. The whole output OQ cannot be profitably sold in one market because of the limited size of market. Therefore, the monopolist must allocate output OQ between two markets i.e. X and Y in such proportion that the necessary condition of profit maximisation is satisfied i.e. in both markets i.e. $MC = MR_x = MR_y$

The price and output in each market is determined by the intersection of the aggregate marginal cost curve and the specific marginal revenue curve. In market X the price charged is P_x which is higher than P_y in market Y. P_x

is higher because the demand in market X is more inelastic than in market Y. Assuming that the costs are similar for each market, the supplier will make more profit in market X than in market Y. The total output of this monopolist is OQ which equals $OQ_x + OQ_y$. (See figure 5.20)

Fig.5.20: Equilibrium under Price Discrimination
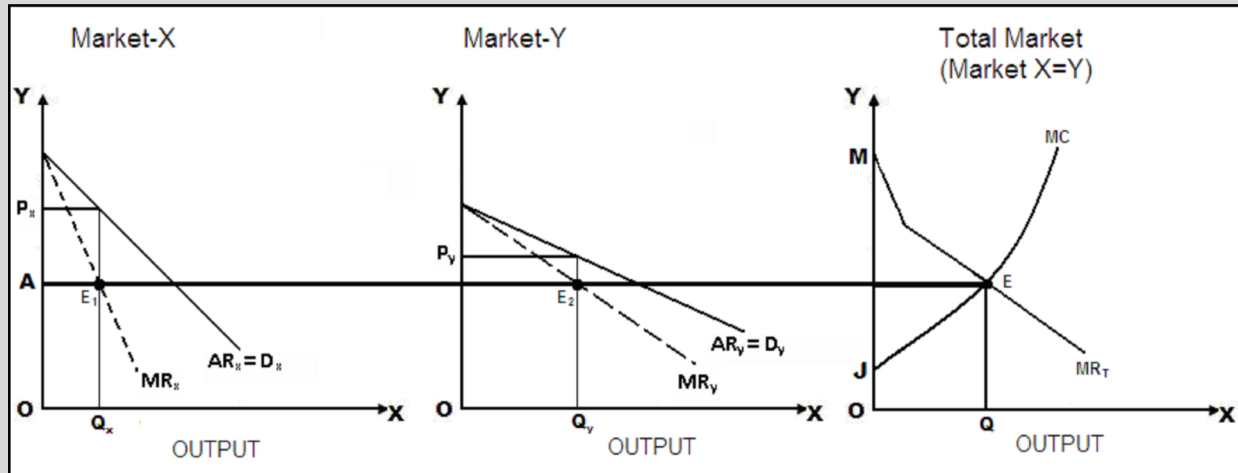
MONOPOLISTIC COMPETITION

A monopolistic competition is defined as that market structure in which each seller produces a 'differentiated product'. The concept of product differentiation means that the product marketed by one seller can be distinguished from the products marketed by other sellers in some form or other. Some of the important methods of product differentiation include: trade marks, brand names size packing or color etc. of the item and technical specifications etc.

According to Prof. Leftwitch," Monopolistic competition is a market situation in which there are many sellers of a particular products but the product of each seller is in some way differentiated in the minds of consumers from the product of every other seller. According to Prof. H.H.Liebhafsky," Monopolistic competition has today come to mean a state of affairs in which there is a large number of sellers selling non-homogenous or slightly differentiated products and in which freedom of entry exists."

Thus, in this market structure, each seller is a monopolist of his differentiated product. The buyers can get it only from him and from none else. At the same time, however, the products offered by different sellers are close substitutes of each other. The buyers are always comparing the prices of their products together with the perceived 'quality' of each. In other words, there is also an intense competition between suppliers for a share in the market. For this reason, it is a market structure in which there is a competition between a group of firms while each firm is a monopolist of its own product. It is, therefore, termed as monopolistic competition.

However, defining a monopolistic competition in this manner, though very realistic, poses certain problems of its own.

– Since the products supplied by the competing firms are not homogeneous, therefore, we cannot define the concept of a 'market demand for the product' precisely, that is, it is not possible to determine the average revenue curve of the 'industry' as a whole.

– Not only that, it is very difficult to even define an industry in a precise manner for the reason that its constituent firms are not supplying the same product. At the most, we may think of a 'group' of firms selling close substitutes of each other.

– It is not possible to have a satisfactory definition of even a 'group'. This is because the 'product group' (such as scooters, or motor cycles) under consideration is itself in competition with other 'product groups'.

Given these limitations, let us mention some of the salient features of the monopolistic competition.

Features of Monopolistic Competition

1. The first feature of monopolistic competition, as mentioned above, is product differentiation. A buyer can get a specific type of the 'product' only from one final source (may be, through the dealers and sub-dealers, etc.).

2. Product differentiation necessitates incurring of selling expenses on the part of firms under market structure of monopolistic competition.

3. Monopolistic competition is characterized by a large number of sellers. The demand and supply conditions of these sellers are inter-dependent. However, in spite of their large number, no individual seller becomes a price taker. He has the authority to demand a price of his choice, though he also considers the demand conditions for his product while exercising this authority. In other words, in spite of there being a large number of sellers, the demand curve for the product of an individual seller is downward sloping. Its demand is not perfectly elastic. It also has large number of buyers also.

4. The fact that each firm produces a 'differentiated product' implies that it can distinguish it further by varying its 'quality'. An improvement in the 'quality' implies an increase in its average cost of production while deterioration in quality implies a reduction in average cost of production. Also an improvement in quality is expected to 'increase' the demand for the product so that, for each given quantity, the buyers are ready to pay a higher price.

5. The firms under the monopolistic competition face a competitive market as regards the inputs used by them. They also have to operate within a given technological range. The result is that no firm is able to compete out its rival by producing a 'better quality' product at a lower average cost.

6. It is assumed that each firm has an accurate knowledge of its demand and cost conditions. This feature implies that the firm is able to estimate the impact of any change in the quantity and/or quality of its product on both its cost of production and average revenue. This knowledge, therefore, enables the firm to maximize its expected profit income.

7. Every existing firm can leave the 'group' of firms belonging to the 'product group'. Similarly, new firms can enter the group and produce close substitutes of the existing products in the group. This free entry and exit of firms ensures that, in the long run, no firm incurs a loss and no firm is able to earn abnormal profit.

8. At the same time, every firm in monopolistic competition is assumed to pursue the goal of profit maximisation. Its aim is not to maximize sales proceeds, or an increase in the market share, etc.

9. It is also assumed that in monopolistic competition all firms have identical cost and demand conditions. This simplifying assumption helps us in analyzing the determination of group equilibrium. It enables us to analyze the working of an individual firm and use it as the representative of the working of the entire group. In the absence of this assumption, we have to separately work out the determination of output, product quality, and price of each firm within the group.

Since product differentiation and selling expenses form the foundation of monopolistic competition, let us explain them a little further.

Product Differentiation

The concept of product differentiation covers all those criteria on the basis of which the product of one firm can be distinguished from that of the other.

The differentiation may be or may not be real. A real differentiation refers to the technical features of the product, including its technical life and performance, durability, cost of operation and maintenance, and the like. In contrast, a

non technical or imaginary differentiation may also be there. It may take the form of brand names, trade marks, packing, size, shape, and so on. The non technical differentiation of a product aims at adding to its subjective appeal to the buyers so that they are induced to 'increase' its demand and are ready to pay more for it. In reality, however, the two forms of differentiation are so mixed up that it becomes nearly impossible to demarcate between the two.

However, irrespective of the nature of product differentiation, it is expected that its adoption by a firm would 'increase' its demand. It provides an opportunity to the firm to tell the buyers that the combination of 'product quality' and price offered by it is better than similar combinations offered by its competing firms. On this account, therefore, a firm under the monopolistic competition is not a price taker. The demand curve for its product develops an inverse relationship with the product price. The firm can raise the price of its product without losing all its customers. And it has to lower the price if it wants to sell a larger quantity of its product. In other words, the demand curve of each firm in monopolistic competition corresponds to that for the 'industry' as a whole, with the difference that it is comparatively more elastic. It is downward sloping and not parallel to X-axis.

It must also be recalled that, in monopolistic competition, the products of firms are close substitutes of each other. As a result, they have high positive cross elasticities. The market for the product of one firm is not clearly separated from the market segments occupied by the products of its rival firms. The firm under consideration can gain or lose a market share of its product both by its own price decisions and those of its rivals. In addition to this, the demand conditions for the product of a firm are affected by its own 'selling expenses', as also by the expenses incurred by its rivals.

Selling Expenses

This leads us to the consideration of the other basic feature of monopolistic competition, namely selling expenses. They are to be distinguished from the production expenses in the sense that the latter are incurred so as to 'produce' the product.

'Selling expenses' are all those outlays which are made in order to create and/or increase its demand. They are aimed at shifting the demand curve of the advertised product to the right so that the buyers should agree to pay more for a given quantity (or are ready to buy more at a given price). Selling expenses may be incurred in several ways though in each case the objective of incurring them is to 'increase' the demand for the product under consideration. They not only include advertisement through different media, show rooms and selling campaigns but also 'offer discounts' and 'incentives' to the dealers and customers. Furthermore, selling expenses may be informative and educative in the sense that they try to convince the customers through 'informing' them of the 'benefits' of using the items being advertised that they should buy this item in preference to something else. Similarly, some selling expenses are incurred to neutralize the perceived impact being created by the activities of the rival firms (or rival product groups). It is also possible that a firm may decide to adopt an aggressive posture and incur selling expenses to create a market for its product or capture a bigger share of the existing market of the 'product group'.

It should be noted that, under monopolistic competition, a major portion of selling expenses incurred in the economy is likely to be neutralizing in its effect or it may only encourage consumption expenditure at the cost of saving. Since, in the ultimate analysis, all selling expenses imply a resource cost for the economy, it is worth asking whether 'non-educative' competitive selling expenses are desirable.

As noted above, the forms of selling expenses are undergoing continuous changes over time. The options available to the advertisers have rapidly expanded with new media. 'Selling' or 'marketing' has acquired a professional dimension. Repetitive and expensive sales campaigns have become necessary for most firms for survival in the market.

It is generally agreed that the impact of selling expenses on demand for an item follows the law of variable proportions. That is to say, the average selling expenses per unit of sales exhibit the three phases of (i) increasing, (ii) constant, and (iii) diminishing returns in terms of addition of the quantity demanded on account of successive additions to selling expenses. This phenomenon is explained by pointing out that advertisement and other forms

of selling expenses gain effectiveness only if they are incurred heavily and repeatedly. Moreover, an addition to the selling budget opens up more effective and wider options to the firm. However, the sources of these benefits tend to get exhausted as the size of selling budget is increased. The average selling cost curve (ASC curve) assumes a U-shape. However, given the selling budget of the firm, the average selling cost curve will be a rectangular hyperbola. It will be similar to the average fixed cost curve of the firm.

EQUILIBRIUM OF THE FIRM UNDER MONOPOLISTIC COMPETITION

While analyzing the determination of equilibrium of a firm under monopolistic competition, it should be noted that we can have two types of its demand curve. Both will be downward sloping.

– It is assumed that a price revision by our firm is not associated by a retaliatory action by its rival firms. In this case, the demand curve of the firm under consideration will have smaller slope. The elasticity of demand for its product will be more elastic.

– In case, the rival firms follow the price revision by our firm, then the demand for its product will become less elastic. Our firm will have to resort to a bigger reduction in price to get a given addition in the demand for its product. Similarly, for a given fall in demand, it will be possible to raise the price by a higher amount. Its demand curve will be steeper in this case.

Short Run Equilibrium under Monopolistic Competition

At this stage, we need not repeat the conditions of equilibrium of a firm. Under Monopolistic Competition, like in case of Monopoly, revenue curves are downward sloping because in order to sell more, producer will have to decrease the price. Further, as in case of monopoly, in short-run a firm can either earn normal profit, super normal profit or losses under Monopolistic Competition. We may also recall that in the short run, the firm may be able to earn an abnormal profit if the demand for its product is quite strong. The abnormal profit is not competed away because in short run new firms cannot enter the 'group' and enhance the supply of the 'product group'. In other words, it is not possible for additional supplies of close substitutes to enter the market and reduce the absolute share of our firm. Similarly, it is also possible for it to operate at a loss subject to the condition that the loss should not exceed its fixed costs.

As in the case of other market structures, in monopolistic competition also a firm is faced with certain fixed costs in the short run. In addition to some production costs, they may include some components of firm's selling expenses as well. Similarly, some other components of selling expenses can fall in the category of variable expenses.

Fig. 5.21: Short Run Equilibrium under Monopolistic Competition
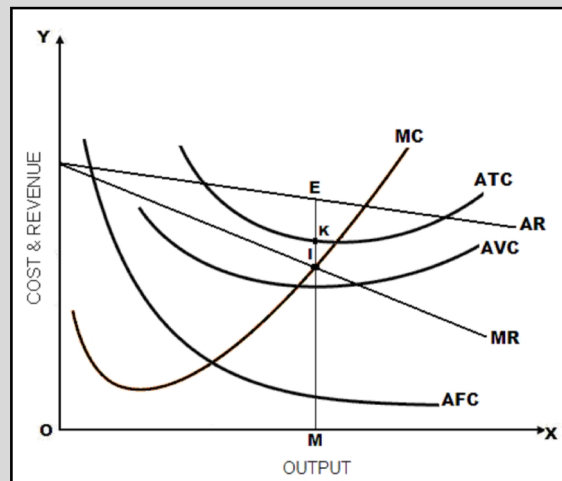
In Fig. 5.21, both AR and MR curve of the firm have negative slopes. AVC curve comprises not only variable production costs but also variable components of selling expenses. Both MC and AVC curves are expectedly U-shaped and the former cuts the latter at its lowest point. ATC curve represents average of the total cost of the firm, including fixed selling expenses.

Equilibrium of the firm is determined with reference to the point I where MC curve of the firm intersects its MR curve from below. The firm decides to produce a quantity of OM while it is able to charge a price of EM per unit. It should be noted that the firm is making an extra profit of EK per unit which is not competed out by the entry of rival firms.

However, depending upon the relative location of cost and revenue curves, the firm could also be in equilibrium with

  – only normal profit

  – covering a part of fixed costs and therefore, incurring a loss less than its fixed costs,

  – a loss equal to total fixed costs

  – In the last case, the AR curve would be tangent to AVC curve

Let us now consider the question of 'group equilibrium', that is, simultaneous equilibrium of all the firms in the group. We notice that the cost and demand conditions of individual firms differ from each other. Moreover, they produce differentiated products so that it is not possible to derive the supply and demand curves for the group as a whole. Chamberlin overcomes this difficulty by making the 'uniformity assumption'. He assumes that all firms in the group have identical cost and demand conditions. As a result, when in equilibrium, all firms produce same quantities of their respective products and sell them at the same price. It is obvious that this is a highly unrealistic assumption. It was probably better to admit that we cannot talk of the concept of 'group equilibrium' and that we have to consider the determination of equilibrium of each firm individually.

Long Run Equilibrium under Monopolistic Competition

In discussing long run equilibrium of a firm under monopolistic competition we should specifically note the following facts.

  – In the long run, there are no fixed costs. The firm can not only vary all its inputs, it can also vary its selling costs. Moreover, it is possible for the firm to make a choice between alternative product qualities.

  – No firm can be compelled to operate at a loss. It can always leave the industry. Therefore, when loss making firms leave the group, the absolute market shares of the remaining firms increase. Their demand curves shift rightwards and upwards. This process continues till it becomes possible for the remaining firms to produce without incurring a loss.

  – Similarly, if the demand is strong so as to bring abnormal profit income to the existing firms, new firms can enter the 'product group', and produce close substitutes of the existing products. This increases the total supply of the 'product group' and with given demand, reduces the demand shares of the existing firms. As a result, the demand curve of an individual firm cannot stay above its long term average cost curve.

  – Each firm operating under monopolistic competition is thus able to make a choice between alternative combinations of

    – product quality

    – product differentiation

    – selling costs

  – In addition, it must also take note of the fact that any variation of price on its part can lead to a price reaction by its rivals also. Conceptually, therefore, it faces a much steeper demand curve than would be case without price reaction by its rivals.

Fig. 5.22 : Long Run Equilibrium of Firm under Monopolistic Competition	**Fig. 5.23: Long Run Equilibrium of Firm under Monopolistic Competition**
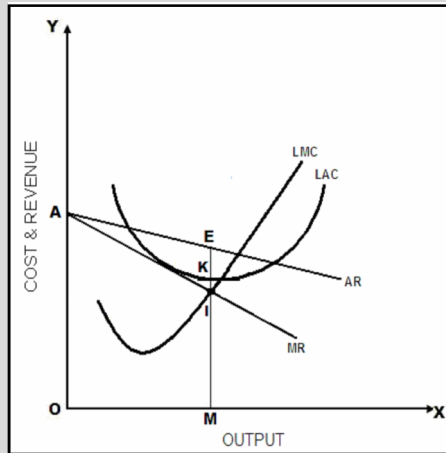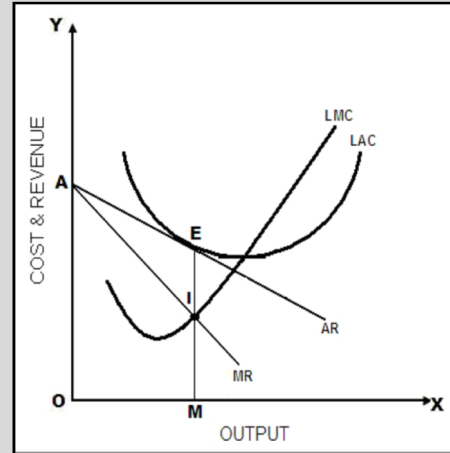
It follows from the above that a firm under monopolistic competition is exposed to a continuous interaction with rest of the firms in the group. Its decisions are not independent of what the other forms are doing. It also notes the fact that its demand curve not only depends upon the actions of its rivals but also upon the quality of its own product and the selling expenses incurred by it. Therefore, it considers alternatives combinations of its cost components relating to product quality and selling expenses etc. and estimates the corresponding slope and position of the demand curve.

Let us assume that our firm has selected the product quality and the amount of selling expenses as represented by LAC curve in Fig. 5.22 which has its corresponding long term MC curve (LMC). The latter curve intersects MR curve from below at I. In this case, the firm would decide to produce a quantity OM of its product and sell it at a per unit price of EM which brings it a per unit profit of EK. However, as noted above, in the long run, the firm is not allowed to earn abnormal profit. New firms enter the market and produce close substitutes. This results in a reduction in the market shares of the existing firms including ours. Its AR curve shifts leftwards till it becomes tangent to LAC curve, as at point E in Fig. 5.23. Once this happens, our firm earns only a normal profit. There is no incentive for new firms to enter the market. Similarly, the process of entry of new firms and reduction in market shares of existing firms cannot continue indefinitely. Once the profit income of the existing firms is reduced to normal level, there is no incentive left for new firms to enter the market.

As a result, long term equilibrium of a firm under monopolistic competition is given by the point of tangency between its AR curve with its LAC curve which is at point E in Fig. 5.23 in the case of our firm.

It should be noted that in monopolistic competition, firms earn only normal profit, in the long run as they do under perfect competition. Presence of monopolistic competition does not improve their profitability. Further, on account of

- – negatively sloping demand curves
- – U-shaped long term average cost curves
- – free exit and entry of firms
- – competition among firms

They all produce below their 'optimum capacity' as the equilibrium is to the left of minimum point of the average cost curve. Some productive capacity of theirs remains unutilized. This increases their average cost of production which they charge from the consumers by adding them to the supply price. The firms operating under monopolistic competition also incur selling expenses which they charge from the consumers.

Similarly, the consumers also suffer in several ways.

- They get a smaller supply compared with what the economy can produce by effectively utilizing its productive capacity.

- The average cost of production is more than the minimum average cost of optimum level of output and the consumers have to bear it in the form of higher prices.

- They also pay for the selling expenses incurred by the sellers.

- In the net, therefore, 'for a given quality of product they pay a higher price, and for a given price they get an inferior product'.

The only possible gain to the consumers is in the form of the 'variety' of products. However, some thinkers believe that the variety does not bring in any addition of utility to the body of consumers taken together because every consumer cannot afford to buy 'better' quality of every product. And those who have a sense of deprivation suffer a loss of utility on that account.

The economy also suffers in the sense that its productive capacity is not utilized fully. Also the firms use up a large amount of productive resources in selling expenses most of which are incurred only to counteract the impact of rivals and have no real benefit for the economy.

OLIGOPOLY

The term oligopoly is derived from two Greek words: 'oligi' means few and 'polein' means to sell. An oligopoly is an industry dominated by a few firms, e.g. supermarkets, petrol, car industry etc. Oligopoly is a market situation in which there are a few firms selling homogeneous or differenti-ated products. The former is called pure or perfect oligopoly and the latter is called imperfect or differentiated oligopoly. Oligopoly is, sometimes, also known as 'competition among the few' as there are few sellers in the market and every seller influences and is influenced by the behaviour of other firms.

In India, markets for automobiles, cement, steel, aluminium, etc, are the examples of oligopolistic market. In all these markets, there are few firms for each particular product.

The main characteristics of oligopoly:

1. **Few firms:** Under oligopoly, there are few large firms although the exact number of firms is not defined. The market is characterised by severe competition as each firm produces a significant portion of the total output.

2. **Barriers to Entry:** Under oligopoly, firms can earn abnormal profits in the long-run as the market is characterised by barriers to entry (for example, patents, licenses, economies of scale requirement of large capital, control over crucial raw materials, etc) which prevent entry of new firms into the industry.

3. **Non-Price Competition:** Under oligopoly, firms try to avoid price competition due to the fear of price war and hence depend upon non-price methods like advertising, after sale services, warranties etc. in order to influence demand and build brand recognition.

4. **Interdependence:** As under oligopoly, there are few firms that hold a significant share in total output of the industry, each firm is affected by the price and output decisions of rival firms. It is in this sense that there is huge interdependence among firms in oligopoly.Therefore, a firm takes into account the action and reaction of its competing firms while determining its price and output levels.

5. **Nature of the Product:** The product of the firms under oligopoly may be either homogeneous or differentiated.

6. **Selling Costs:** Under oligopoly, firms try to avoid price competition and indulge in non-price competition.

As there is severe competition and huge interdependence in oligopoly markets, selling cost are important for competing against the rival firms for the larger market share.

7. ***No unique pattern of pricing behaviour:*** Under oligopoly, firms want to act independently and earn maximum profit on the one hand and they want to cooperate with rivals to remove uncertainty on the other hand. Depending upon their different motives, there may be various situations emerging in real life and hence it is not possible to predict anyunique pattern of pricing behaviour among the firms in oligopoly markets. The firms may either compete with others or collude with others leading to different pricing situations.

8. ***ndeterminateness of Demand Curve:*** Unlike other market structures, it is not possible to determine demand curve ofa firmunder oligopoly. It is so because on the one hand, there is huge interdependence among rival firms and on the other hand, there is uncertainty regarding the reaction of rivals. The rivals may react in different ways when a firm changes its price and that renders the demand curve of a firm indeterminate.

Firms' Behavior under Oligopoly

Depending upon the objectives of the firms, magnitude of barriers to entry and nature of government regulation, there may be different possible outcomes in relation to firms' behaviour under oligopoly. These are as follows:

1. Stable prices (e.g. through kinked demand curve)

2. Price wars (competitive oligopoly)

3. Collusion for higher prices.

Broadly, we can say that oligopoly may either be collusive or it may be non-collusive. A collusive oligopoly is a market situation wherein the firms cooperate with each other in determining price or output or both. A non-collusive oligopoly refers to a market situation wherein thefirms in an oligopoly market compete with each other rather than co-operating.

In view of the fact there are various models of firms' behavior under oligopoly, we shall restrict our discussion to one model namely, kinked demand curve model (non-collusive oligopoly).

1. Non-Collusive Oligopoly-Sweezy's Kinked Demand Curve Model (Price-Rigidity):

Price rigidities are often observed in oligopolistic markets.Paul Sweezy in 1939 used unconventional demand curve called as kinked demand curve to explain these rigidities in oligopoly markets.

Reason for the kink in demand curve:

Under the present model, it is assumed that firms behave in two-fold manner (asymmetric behavioural pattern) in reaction to a price change by a rival firm. In simple words, the asymmetric behavioural pattern is that firms follow price cuts by a rival but firms do not follow price increases. In other words, if a seller increases its price, its rivals will not follow the price increase and the market share of this firm will reduce significantly as a result of higher price compared to rivals (the demand curve corresponding to price increase would be relatively elastic as small increase increase in price will reduce demand of this firm significantly). Contrarily, if a seller reduces its price, its rivals will also reduce their price atleast equal to the price reduction by this firm in order to prevent their market share from falling. The firm lowering the price initially will not be able to gain from price reduction due to reaction of its rivals(the demand curve corresponding to price decrease would be relatively inelastic as this firm will not be able to increase its demand much by reducing its price). Two different types of elasticities corresponding to different types of reaction by rival firms result in a kind in demand curve of a firm in oligopoly markets.

Why price rigidity?

As is clear from the above, any firm cannot gain or it can even lose by changing its price from the prevailing price

in the market. Neither increase nor decrease in rpice would result in increase in demand for the firm changing its price. It leads to firms sticking to the same price over time and hence results in price rigidity under oligopoly.

Explanation of the Model:

Given these assumptions, the price-output relationship in the oligopolist market is explained in Fig. 5.24 where KPD is the kinked demand curve and OP_0 the prevailing price in the oligopoly market for the OR product of one seller. Starting from point P, corresponding to the current price OP_1, any increase in price above it will considerably reduce his sales, for his rivals are not expected to follow his price increase. This is so because the KP portion of the kinked demand curve is elastic, and the corresponding portion KA of the MR curve is positive. Therefore, any price-increase will not only reduce his total sale but also his total revenue and profit.

On the other hand, if the seller reduces the price of the product below OP_Q (or P), his rivals will also reduce their prices. Though he will increase his sales, his profit would be less than before. The reason is that the PD portion of the kinked demand curve below P is less elastic and the corresponding part of marginal revenue curve below R is negative. Thus in both the price-raising and price-reducing situations, the seller will be a loser. He would stick to the prevailing market price OP_0 which remains rigid.

In order to study the working of the kinked de-mand curve, let us analyse the effect of changes in cost and demand conditions on price stability in the oligopolistic market.

Let us suppose, the prevailing price in the market is OP_0. As per assumption, if one seller increases the price above OP_0 the rival sellers do not change their price and keep the prices of their products at OP. It will lead to product of price changing firm to becoming costlier than others and consequently, demand for the costlier product will fall significantly. It can also be seen that demand curve of firm for any price above OP_0 i.e. KP segment of demand curve, is relatively elastic. High elasticity will reduce the demand significantly as a result of price increase.

On the other hand, if a seller reduces the price below OP_0, the rivals will also follow the price cut so as to prevent their demand from falling. It can also be seen that demand curve of firm for any price below OP_0 i.e. PD segment of demand curve, is relatively inelastic. Low elasticity will not increase the demand significantly as a result of price cut.

This is called asymmetrical behavioural pattern which results in kink in the demand curve and hence price rigidity in oligopoly markets. Prices remain rigid at the kink denoted by point P i.e. price will remain sticky at OP_0 and output will be equal to OR at this price. Due to the difference in elasticities of two segments of demand curve, MR curve becomes discontinuous corresponding to the point of change in elasticity of the demand curve as represented by the kink. At output less than OR, demand curve is KP and corresponding MR curve is KA. For output larger than OR, demand curve is PD and corresponding MR curve is BMR.

Fig. 5.24: Kinked – Demand Curve Model
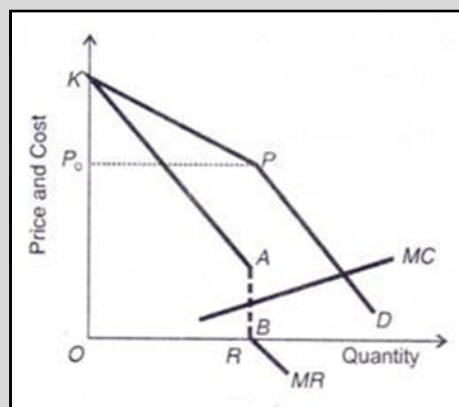
Another point to note is that as per equilibrium condition 'MC=MR', equilibrium establishes at price OP_0 and quantity OR as MC cuts MR in the discontinuous portion of MR below the kink. It may also be noted that even if cost increases, i.e. MC curve move upwards, or cost decreases, i.e. MC curve move downwards, in its discontinuous portion i.e. segment AB, the price will remain sticky at OP_0 so long as movement of MC curve does not go beyond the AB.

2. Collusive Oligopoly

Kinked demand curve model is based on the assumption that oligopoly firms though interdependent in the market decide to act independently. But there is another possibility that firms may try to remove uncertainty related to acting independently by entering into pricing agreements with each other. Such an agreement is called collusion. A collusion may be formal or informal. Generally, a collusion may take the form of cartel or price leadership. Cartel refers to an association of independent firms, within the same industry, that follow the common policies relating to price, output, sale, profit maximisation and distribution of products. Price leadership is based on tacit or informal collusion. Under price leadership, one firm that may be a large or dominant firm acts as the price leader and fixes the price for the product while other firms follow it. Price leadership is of various types. But the discussion is restricted to only dominant firm price leadership.

Comparison between Perfect Competition, Monopoly, Monopolistic Competition and Oligopoly

Features	Perfect Competition	Monopoly	Monopolistic Competition	Oligopoly
Number of selling firms	Large	Single	Large	Few large sellers
Number of buyers	Large	Large	Large	Large
Size of the market for each firm	Very small	Large	Small	Large
Nature of product	Homeogenous	Unique product, no close substitute	Differentiated products	Homeogenous or Differentiated
Entry and exit condition	Free	Strong Barriers to entry	Free	Barriers to entry
Price policy of firm	Price taker	Price Maker	Price Maker (some control over price depending on consumer brand loyalty)	Varies
Market knowledge	Complete Information	Incomplete	Incomplete	Incomplete
Price Elasticity of Demand for individual firm	Perfectly elastic (Infinite)	Less	High	–
AR and MR Curves	Equal	Different	Different	Different
Selling Cost	Nil	Optional	Large	Large
Equilibrium Conditions	MR=MC MC cuts MR from below	MR=MC MC cuts MR from below	MR=MC MC cuts MR from below	–
Nature of decision variable	Only output	Both price and output are within his control but at one time decision can be taken only about one of them. It faces a trade off between price and output	Nature and extent of product differentiation and hence the level of selling expenses.	–

Concept of Externalities

In economics, an externality is said to exist when the actions of a specified group of economic agents have significant economic repercussions on agents outside the group. In other words, the term externality is used to describe the cost or benefit incurred by the third party who did not choose to receive that cost or benefit. It is the consequence of economic activities endured by an unrelated third party due to lack of control over the factors that creates the cost or benefit. Externalities may be positive or negative.

*Positive Externalities***:** Positive externalities are defined as economic activities that have positive effects on unrelated third party. In case of production, positive externalities are positive effects that originate during the production process of a good or service. For example, a farmer, who has an orchard near to bee farms, in the process of production of fruits also helps bee farmers around the area. The bees can collect ample amount of nectar to prepare honey and it increases benefit of bee farmers. In this case, bee farmers are incurring positive externalities as a result of the production activity of the farmer in the nearby orchard. On the other hand, bees help in pollination at fruit farm which is the positive externality incurred by orchard owner as a result of production activity of bee farmers. In this case, both the parties are benefitting from the production activity of each other, even though both are making production decisions independently.

*Negative Externalities***:** Negative externalities are defined as economic activities that have negative effects on unrelated third parties. In case of production, negative externalities are negative effects that originate during the production process of a good or service.For example, pollution caused by a chemical manufacturing industry that dumps its waste chemicals into a nearby water stream. The industry is degrading the natural state of water resource by mixing harmful chemicals in it. The pollution by this firm will constitute negative externality for the people living in and around that environment and using the water for drinking and other purposes.

LESSON ROUND UP

- The concept of a market is central to the understanding of the determination of price and quantity of output of a commodity under consideration. Market means the general field within which, the force determining the price of particular product operate

- The market consists of two components:

  - A Firm

  - An Industry

- There are various kinds of markets prevailing in the economy.

  - Perfect Competition has following characteristics:

  - Large Number of Sellers and Buyers

  - Homogeneous Product

  - Free Entry and Exit

  - Firm is a price taker

- Full Knowledge of Market

- Economic Rationality

- No Transportation Cost

- The perfect competition firm is in equilibrium when MC= MR and MC cuts the Mr curve from below.

- Monopoly

- Monopoly means a single seller and large number of buyers

- No close substitute

- A single firm industry

- Monopolist is free to fix a price of his choice. Firm is price maker.

- Irrespective of the profit income of the existing producer firm, new firms cannot enter the industry

- A buyer will buy it only if its price does not exceed its marginal utility to him

- The equilibrium for monopolist is at a point where MR= MC.

- One of the important features of the monopolist is price discrimination, where he charges different prices for homogeneous product to different consumers.

- Monopolistic Competition

  - Monopolistic Competition means large number of sellers and buyers

  - Close substitutes

  - All firms are monopolists of their differentiated product

- The main feature is product differentiation and selling expenses comprising of marketing costs.

- Oligopoly

- A few sellers and large number of buyers

- Homogeneous and differentiated Products

GLOSSARY	
Profit Maximizing Conditions	The necessary condition is marginal revenue equal to marginal cost. The sufficient condition is marginal cost curve is intersecting the marginal revenue curve from below.
Economic Profits	Difference between total revenue and total cost incurred on inputs.
Normal Profits	When total revenue and total cost equals. Normal Profits = Zero Economic Profits
Homogeneous products	Two or more products that are identical in every possible respect.
Short Run Equilibrium	A competitive firm attains short run equilibrium when MR= SMC with SMC rising.
Long Run Equilibrium	A competitive firm attains long run equilibrium when MR= LMC with LMC rising and P=AR=LAC.
Break Even Condition	When total sale proceeds covers total costs of production.
Shut Down Condition	When price of the good falls below the average variable costs. If the price of the firm is such that it is not even able to cover its variable costs, the firm should shut down.
Degree of Monopoly	Monopoly power is the degree of power held by the monopolist to set the price for a good.
Price Discrimination	When the similar product is sold to different consumers at different prices.

| Product Differentiation | Slight differences that exist between two or more goods that are essentially the same and which satisfy the same basic want or need. This is generally pursued in monopolistic competition. |
| Selling Costs | All expenditures pertaining to selling activities after the product has been produced. An important component is advertising and other sales promotion expenditures including giving free gifts and other promotional activities. |

MULTIPLE CHOICE QUESTIONS

1. **Which one is not a feature of Perfect Competition?**

    (a) A Large Number of Buyers

    (b) A Large Number of Sellers

    (c) Firm is a Price Maker

    (d) Homogenous Product

2. **In long run a firm under Perfect Competition can?**

    (a) Earn Normal Profit

    (b) Earn Abnormal or Supernormal Profit

    (c) Bear the Loss

    (d) Shut down the Production process.

3. **Which one is not a feature of Monopoly?**

    (a) Single Seller

    (b) Firm is a Price taker

    (c) Unique Product

    (d) Strong Barriers to entry

4. **What is known as Shut -down Price?**

    (a) When price of the good falls below the average variable costs.

    (b) When price of the good is equals to the average variable cost.

    (c) When price of the good is more than the average variable cost.

    (d) When MC = MR

5. **A market will be in equilibrium, when**

    (a) Quantity demanded equals quantity supplied

    (b) Excess demand and excess supply are zero

    (c) The market is cleared by the equilibrium price

    (d) All of the above

6. **When a firm's average revenue is equal to its average cost, it gets —**

    (a) Supernormal Profit

    (b) Normal profit

(c) Loss

(d) None of the above

7. Which of the following statements about firms in different types of market is false?

(a) A Perfect competitor has no influence over the price of its product.

(b) A monopolistic competitor may engage in non price competition

(c) An oligopolist may monitor the prices and products of all the other firms in its market

(d) A monopolist must be a large firm

8. Compared to the case of perfect competition, a monopolist is more likely to -

(a) Charge a higher price

(b) Produce a lower quantity of the product

(c) Make a greater amount of economic profit

(d) All of the above

9. For which of the following market structure, it is assumed that there are barriers to entry?

(a) Perfect Competition

(b) Monopolistic Competition

(c) Monopoly

(d) All of the above

10. In order to practice Price discrimination, which of the following is needed?

(a) Some degree of monopoly power

(b) An ability to separate the market

(c) An ability to prevent reselling

(d) All of the above

Answer key : 1. (c), 2.(a), 3.(b), 4. (a), 5. (d), 6.(b), 7.(d), 8.(d), 9.(c), 10. (d)

Suggested Readings

1. H. L. Ahuja, Modern: Microeconomics

2. A. Koutsoyiannis: Modern Microeconomics

Lesson 6
Indian Economy – An Overview

LESSON OUTLINE

LEARNING OBJECTIVES

Indian economy is one of the fastest growing economies in the world. It has experienced with five-year planning system for nearly seven decades and now following a new approach to planning and growth. In its journey of economic growth and development, some of the issues still remained unresolved and require continuous efforts on the part of policymakers.

It is important to review basic characteristics of Indian economy as well as various issues still faced by the economy in order to have better understanding from an economic perspective. The five-year planning system and recent NITI Aayog require a special mention in this case. From the policy perspective, it is also important to compare Indian economy with other major economies in the world to have a comprehensive understanding of the areas where our economy is performing better than other major economies of the world and where it is lagging behind.

BASIC CHARACTERISTICS OF THE INDIAN ECONOMY:

Despite global headwinds and a truant monsoon, India registered robust growth of 7.2 per cent in 2014-15 and 7.6 per cent in 2015-16, thus becoming fastest growing major economy in the world. As per the estimates of the International Monetary Fund (IMF), global growth averaged 3.1 per cent in 2015, declining from 3.4 per cent registered in 2014. While growth in advanced economies has improved modestly since 2013, the emerging economies have witnessed a consistently declining trend in growth rate since 2010. It is against this background that the recent Indian growth story appears particularly bright and attracts the attention of policy makers around the world.

For an overview of the Indian economy, the first step is to discuss the basic characteristics i.e., key macro-economic indicators of the economy. These are as follows:

(a) National Income:

Income is probably the most frequently used term in economics, used by experts and lay men. Income level is the most commonly used tool to determine the well-being and happiness of nations and their citizens. This remains true even today, even if we know that 'income' is not an exhaustive idea to know about the well-being of the society. Hence, a brief study, of the trend of the national income in India is very much essential for attaining a clear understanding about the impact of various plans and economic reforms on the Indian economy.

In general, income of an economy is measured through Gross Domestic Product (GDP). The following broad trends in the changing composition of the domestic production are revealed:

Table 6.1 : Share of Gross Domestic Product by industry of Origin (1999-00 series)

(Percentage distribution)

		1950-51	1980-81	2013-14*	2014-15*	2015-16*
I.	**Agriculture & Allied Services**	**55.4**	**38.0**	**13.9**	**16.1**	**18.4**
II.	**Industry**	**15.0**	**24.0**	**26.2**	**31.4**	**28.3**
	(a) Mining & Quarrying	1.4	2.0	1.9	2.9	
	(b) Manufacturing	8.9	13.8	14.9	18.1	
	(c) Electricity, Gas & Water supply	0.3	1.6	1.9	2.3	
	(d) Construction	4.4	6.6	7.4	8.1	
III.	**Services**	**29.6**	**38.0**	**59.9**	**52.5**	**53.3**
	(e) Trade, Transport and Communications	11.3	17.4	26.4	19.4	
	(f) Finance, Insurance, Real Estate etc.	7.7	7.5	20.6	20.5	
	(g) Community, Social and Personal services	10.6	13.1	12.9	12.6	
		100.0	100.0	100.0	100.0	100.0

Source: Handbook of Statistics of the Indian Economy, RBI and CSO
Note: *2004-05 survey

The share of the primary sector which includes, agriculture, forestry and fishery has gone down from 55.4 per cent of GDP in 1950-51 to 38 per cent in 1980-81 and further declined to 13.9 per cent in 2013-14. The main cause of the decline is a rapid fall in the share of agriculture alone. There is also contraction in the share of forestry from about 6 per cent in 1950-51 to nearly 0.7 per cent in 2013-14. The share of fishery has remained more or less constant around 1 per cent throughout the period. However, in the recent years (2014-16), there is a marginal improvement in the agriculture.

The share of industry which includes mining, manufacturing, electricity, gas and water supply and construction has shown a steady increase from 15 per cent in 1950-51 to 24 per cent in 1980-81 and 28.3 per cent in 2015-16. Among the major constituents of the secondary sector, the share of manufacturing industries to GDP also increased from 8.9 per cent in 1950-51 to 18.1 per cent in 2013-14. The share of construction to GDP improved from 4.4 per cent in 1950-51 to 8.1 per cent in 2014-15.

The share of service sector has three components: (a) Trade, Transport, Storage and Communication, (b) Finance, Insurance, Real estate and Business services and (c) Community, social and personal services. The share of service sector indicated a sharp improvement from 29.6 per cent in 1950-51 to about 59.9 per cent in 2013-14. There was a significant increase in the share of trade, transport and communications from just 11.3 per cent in 1950-51 to 31.8 per cent in 1970-71 and then finally to 53.3 per cent in 2015-16. The expansion of transport, especially road transport and communications, during the last decade of mobile revolution has been the major contributor to this increase.

The process of economic development involves a rapid expansion of public administration especially economic and welfare services such as education, health and family welfare. Taking community and personal services as a group, there was an improvement in its share from 10.6 per cent in 1950-51 to 12.6 per cent in 2014-15.

The structural change in the composition of national income by industrial origin is the consequence of the process of economic growth initiated during the plans. However, agriculture did not indicate a fast and consistent rate of growth. As is evident from Table 6.2 the rate of growth of agriculture showed a decline from 3 per cent during 1950-51 and 1960-61 to 1.5 per cent during 1970-71 and 1980-81 and thereafter, it picked up to 3.4 per cent during 1980-81 and 1990-91. However, it again declined to 2.6 per cent during 1990-91 and 2000-01 and then improved to 3.8 per cent between 2004-05 and 2013-14.

Table 6.2: Compound Annual Growth Rate of Various Economic Activities

	1950-51 and 1960-61	1960-61 and 1970-71	1970-71 and 1980-81	1980-81 and 1990-91	1990-91 and 2000-01	1999-00 and 2007-08	2004-05 and 2013-14
Agricultural, forestry & fishing	3.0	2.3	1.5	3.4	2.6	2.8	3.8
Mining and quarrying	5.6	3.9	4.9	8.4	3.9	5.1	2.8
Manufacturing	6.0	5.2	4.0	7.6	5.9	7.6	6.9
Electricity, gas and water supply	10.3	11.1	6.8	9.1	6.5	4.6	6.3
Construction	6.3	5.5	2.9	4.6	5.1	10.5	7.5
Trade, hotels and restaurants	5.3	4.8	4.2	5.9	7.2	8.8	9.0
Transport, storage and communications	5.7	5.5	6.1	5.6	8.0	13.8	9.0
Financing, insurance & real estate	3.0	3.4	4.0	9.9	8.4	8.8	11.5
Community and personal services	4.1	5.3	4.4	6.1	6.7	5.6	7.0
Total services	4.3	4.8	4.4	6.7	7.5	7.9	9.2
GDP at factor cost	3.9	3.7	3.1	5.6	5.6	7.3	7.6

Source: Computed from the National Accounts Statistics of India, EPWRF and CSO and Economic Advisory Council Estimates.

The growth of the service sector which was around 4.5 per cent during 1950-51 to 1980-81, showed a sharp increase to about 9.0 per cent during 2004-05 and 2013-14. The growth rate of the manufacturing sector which was 6.4 per cent during 1950-51 and 1960-61 improved to 7.6 per cent during 1980-81 and 1990-91, and marginally decreased to 6.9 between 2004-05 and 2013-14. The GDP of the economy grew at a rate of 3.9 per

cent during 1950-51 to 1980-81 and it improved to 7.6 per cent during 2004-05 and 2013-14. The slow growth of the agricultural and manufacturing sector has been a major road block for the overall growth. It is only the fast growth of service sector – which remained the principal driver of the growth of economy.

Thus the above analysis clearly, shows a '*service-led*' pattern of economic growth followed by India in the later part of its economic transformation. A peculiar feature of the economic development in India is that the country, without going through industrialization, has jumped to the post-industrial stage.

(**b**) **Sectoral Employment in India:** The question arises: Did this structural shift in GDP share composition have an impact on the employment pattern or distribution of workforce in India? Data provided in Table 6.3 based on the various rounds of the National Sample Survey reveals that the share of workforce deployed in agriculture declined from 68 per cent in 1983 to 53.2 per cent in 2009-10. Along with this share of employment in industry increased from 13.9 per cent in 1983 to 21.5 per cent 2009-10. During the same period the share of services in total employment increased from 18.2 per cent to 25.4 per cent. This analysis shows that, in India, the sector which has observed sharp decline in its GDP contribution (agriculture sector) has not observed the corresponding decline in its contribution to employment. On the other hand, the sector which has observed a sharp increase in its share of GDP failed to register a sharp increase in its contribution to employment, i.e. services sector.

Table 6.3: Sectoral Distribution of Work Force in India

Sector of Employment	1983	1993-94	2009-10
Agriculture	68.1	63.9	53.2
Industry	13.9	14.9	21.5
Services	18.2	21.2	25.4
Total	100.0	100.0	100.0

Note : Figure relate to usual status of individuals
Workforce covers those involved in gainful activity regularly + those involved in gainful activity occasionally.

Source: Various rounds of National Sample Survey.

(**c**) **Capital Formation:** Capital formation is an important step towards economic development of a country. Insufficiency of capital is considered as an important limiting factor in case of the underdeveloped or developing economies. Hence, both the amount of production and its composition largely depend on the amount of capital available in the country. Capital formation includes physical (machines, tools, instruments, etc.) as well as human (knowledge, skill, health etc.) capital.

Since capital formation is an essential determinant of economic growth, it is quite necessary to ask how can a community achieve a high rate of investment or capital formation? The foremost important process of physical capital accumulation is an increase in the volume of real domestic savings.

Table 6.4: Savings and Investment in India (as percent of GDP)

Year	1960s	1970s	1980s	1991-92 to 1996-97	1997-98 to 2003-04	2004-05 to 2011-12	2013-14
					(2004-05) series		
Gross Domestic Capital Formation (CDCF)	14.3	17.3	20.8	23.9	24.9	35.3	34.8
Gross Domestic Savings (GDS)	12.3	17.2	19.0	22.7	24.9	33.5	30.1
Saving Investment Gap	-2.0	- 0.1	- 1.0	- 1.2	0.0	- 1.8	-4.7

Indian economy has outlined a path of capital accumulation in which a major part of the resources needed for

economic development were supplied by real domestic savings and a part by inflow of capital from abroad. In India, household sector is the major contributor to Gross Domestic Savings (GDS) accounting for 29.8 percent of GDP in 2015 (World Bank). Data provided in Table 6.4 presents saving-investment gap (difference between GDCF and GDS) which is indicative of foreign savings. During 1960s this gap was the highest, to the extent of 2 per cent of GDP, which later showed a decline and reached to zero during 1997-98 to 2003-04. Later during 2004-05 and 2011-12 it increased to 1.8 per cent.

(**d**) **Inflation:** Inflation management is one of the hardest tasks an economic policy maker has to undertake. Inflation is a sustained rise in the general level of prices. Price stability is thus considered as an essential condition for economic stability and growth. In a country like India, which is subject to various structural problems, fluctuations in the price level due to natural as well as economic factors are very much common. Accordingly, an atmosphere of uncertainty is created by fluctuation in prices which always goes against the spirit of economic development of the country. The rate of inflation is measured on the basis of price indices which are of two kinds – Wholesale Price Index (WPI) and Consumer Price Index (CPI).

Table 6.5: Inflation based on Different Indices (in per cent)

	2010-11	**2011-12**	**2012-13**	**2013-14**	**2014-15**	**2015-16* (Apr-Dec)**
WPI	9.6	8.9	7.4	6	2	-3
CPI	-	-	10.2	9.5	5.9	4.8

Source: Department of Industrial Policy and Promotion (DIPP) for WPI, Central Statistics Office (CSO) for CPI .
Notes: * WPI and CPI figures are provisional.

Persistent and elevated levels of inflation, in particular food inflation, were a major concern of the government during 2010-11 to 2013-14. During this high-inflation phase, average WPI inflation was 8.0 per cent and average inflation based on the CPI was way above comfort level at 9.7 per cent. CPI– based food inflation touched double digits and likewise WPI-based food inflation rose to a high of 9.3 per cent during the same period. The rising trend reversed from 2014-15 onwards and the economy has been experiencing sustained moderation in general inflation ever since. A comparative picture of inflation based on the major series of the price indices for the last five years is given in Table 6.5.

(**e**) **Foreign Capital/Investment:** The need for foreign capital arises because of the following reasons. In most developing countries like India, domestic capital is inadequate for the purpose of economic growth. Foreign capital is typically seen as a way of filling gap between the domestically available supply of savings, plus foreign exchange plus, government revenue and the planned investment necessary to achieve developmental targets. These countries are not only 'capital poor' but they are also backward in technology required for rapid economic development. They require trained personnel, technical know-how and advanced management skills. They also need modern machines and equipments. Foreign capital in addition to providing much needed resources is expected to also help to solve the problem of technological backwardness of these countries.

Table 6.6: Foreign Investment Flows (Gross) by Category (US $ million)

	Foreign investment			*Portfolio Investment*			*Grand Total*
	Foreigners (1)	*NRI's (2)*	*3 = 1+2*	*FII's (4)*	*Others (5)*	*6 = 4+5*	*7 = 3 + 6*
1991-92	66	63	129	4	0	4	133
2001-02	5,214	916	6,130	1,505	516	2,021	8,151
2013-14	21,177	14,869	36,046	5,001	20	5,021	41,067
2014-15	21,761	22,530	44,291	40,923	-	40,923	85,214
Total (1991-92 to 2014-15)	1,92,421 (33.0)	4,58,111 (27.1)	3,50,502 (60.1)	2,00,427 (34.4)	32,188 (5.5)	2,32,615 (39.9)	5,83,147 (100.0)

Note: Foreigners include investment flows under RBI automatic route and SIA/FIPB route.
　　　　Others include Euro-equities (GDR amounts raised by Indian Corporates) and Offshore funds and others.
　　　　Figures in brackets are percentages of total foreign investment.
Source: Compiled from Economic Survey, Handbook of Statistics on Indian Economy, RBI Bulletin.

After the announcement of New Industrial Policy (1991), there has been acceleration in the flow of foreign capital in India. As per data provided by the Government of India (GOI), during 1991-92 to 2012-13, total foreign investment flows were of the order of $456.9 billion, out of which about $270.2 billion (59.1 per cent) were in the form of Foreign Direct Investment (FDI) and the remaining $ 186.7 billion (40.9 per cent) were in the form of portfolio investment. It clearly shows the preference of foreign firms was more towards incurring portfolio investment. Moreover, out of the total direct foreign investment of the order of $ 270.2 billion, nearly 45 per cent ($ 120.7 billion) was contributed by Non-Resident Indians. Thus, the net contribution of foreign firms in direct investment was about 55 per cent of total foreign investment flows.

(**f) Foreign Trade:** During the pre-British period, India was famous for her textile and handicrafts exports. During the British period, gradually pattern of India's trade changed and it became an exporter of raw materials and importer of British machine made goods. But during the post-independence period, India experienced a thorough change in the composition and direction of foreign trade. In the post-economic reforms period, a good number of non-traditional commodities have got included in the export list of the country. The direction of trade has also changed remarkably.

Since the introduction of five-year planning system in the country, India had to depend heavily on the imports of various types of machineries and equipment for the development of various types of industries in the country. Such imports were known as developmental imports. At that stage, India had to import capital goods extensively in order to set up various industries, plants, locomotives, hydro-electric projects etc. In the next stage, India had to import a huge quantity of intermediate goods and raw materials in order to utilize the productive capacity developed in the initial period of planning. These types of imports were known as maintenance imports. Thus foreign trade has been helping the country in different stages of its economic development.

Table 7 below gives information on India's exports and imports for a decade. As may be seen from the Table 7 that for the four years from 2000-01 to 2003-04, India's dependency on other countries was much less and its export performance was steadily improving. During this period, the trade deficit was much less and confined to thousand crore rupee figures although there was more than 56 per cent annual growth in balance of trade in 2003-04. If those were indications, in the year 2004-05, there was increase in both exports and imports with annual growth of almost 40 per cent, a double rate of growth compared to previous year; exports also performed well at Rs. 3.75 lakh crore with 28 per cent growth in 2004-05 in comparison to 15 per cent growth in 2003-04. However, the impact was noticed more in balance of trade figures as it crossed one lakh crores (Rs. 125725 crore) and grew by 91 per cent in 2004-05 compared to 56 per cent in 2003-04. From this point, the country could not recover and the balance of trade figures were ever increasing since then except for 2009-10 when the growth was negative due to global slowdown impact.

The global slowdown during 2008 had its impact on the economy of almost all the countries, including India. The impact was such that during 2009-10, in case of both exports and imports, there was decrease in the rate of growth; for exports it was merely 0.6 per cent, whereas for imports, it was negative growth of – 0.8 per cent for the first time in 2009-10. The resulting impact was that the balance of trade also showed negative growth of 2.9 per cent. However, in 2010-11 and 2011-12, the growth in exports and imports recovered. In 2010-11, the export growth was 35 per cent but import growth was less at 23 per cent, resulting in only 4 per cent growth in balance of trade. However, in 2011-12, the imports grew significantly by 39 per cent and exports grew by 28 per cent, resulting in 63 per cent growth in balance of trade.

Table 6.7: India's Exports, Imports, and Balance of Trade from 2000-01 to 2011-12

Years	Value in Rs. crores			Percentage growth		
	Exports	*Imports*	*Balance of Trade*	*Exports*	*Imports*	*Trade deficit*
2000-01	203571	230873	-27302			
2001-02	209018	245200	-36182	2.68	6.21	32.53
2002-03	255137	297206	-42069	22.06	21.21	16.27
2003-04	293367	359108	-65741	14.98	20.83	56.27
2004-05	375340	501065	-125725	27.94	39.53	91.24
2005-06	456418	660409	-203991	21.6	31.8	62.25
2006-07	571779	840506	-268727	25.28	27.27	31.73
2007-08	655864	1012312	-356448	14.71	20.44	32.64
2008-09	840755	1374436	-533681	28.19	35.77	49.72
2009-10	845534	1363736	-518202	0.57	-0.78	-2.9
2010-11	1142922	1683467	-540545	35.17	23.45	4.31
2011-12	1465959	2345463	-879504	28.26	39.32	62.71
2012-13	1634319	2669162	-1034843	11.48	13.8	17.66
2013-14	1905011	2715434	-810423	16.56	1.73	-21.69
2014-15	1897026	2734049	-837023	-0.42	0.69	3.28

Source: Compiled and computed from the data provided by Handbook of Statistics on Indian Economy, Economic Survey, RBI Bulletin

The top ten destination countries for Indian exports, viz. UAE, USA, Singapore, China, Hong Kong, Netherlands, Saudi Arabia, UK, Germany and Japan, remained more or less the same during the last four years except for some minor change in relative positions here and there. These ten countries together accounted for a share of more than 50% of our total exports. The major items in India's export basket are Petroleum products, Engineering goods and Gems & Jewellery. These three products together contribute 50 per cent of India's exports.

India's imports from the People's Republic of China remained consistently on top with the percentage share of imports from China ranging between 11 and 12 per cent during 2008-09 and 2012-13. This indicates the importance of Chinese goods in Indian markets. After China UAE, Saudi Arab, Switzerland and USA have remained the next top four major import partners for India during 2009-10 to 2012-13. Further, petroleum & crude oil products and pharmaceutical products have remained major components of India's import basket.

DEVELOPMENT ISSUES IN INDIAN ECONOMY:

Despite the fact that Indian economy has been attaining praise worthy progress in terms of growth rate, India is

still considered as a low income developing economy. A quarter of the country's population lives below poverty line. Poverty is not only acute but a chronic problem in India. At the same time, there exist un-utilised or under-utilised human and natural resources. It is, therefore, quite important to understand the basic issues of development of an economy which is one of the poor and at the same time one of the fastest growing economies of the world.

1. Low per capita income: Developing economies are noted for existence of low per capita income. The per capita income of India in 2014 was $ 1560. It may be noted from Table 6.8 that in the same year the average per capita Gross National Income (GNI) of USA, highly developed country, was 35 times that of India. It shows that difference between the level of living of an average American and Indian was quite large and significant. Even the neighbouring country China, a major developing country, had an average 5 times higher per capita GNI than India. In case of India not only the per capita income is low but the distribution of income is also unequal. The unequal distribution of income and wealth makes the problem of poverty one of the critical and acute obstacles in the path towards economic progress.

Table 6.8: Per Capita Gross National Income at Market Prices -2014 (in US Dollars)

Country	Atlas method
USA	55,320
India	1,560
China	7,520

Source: The World Bank, World Development Indicators

2. Huge dependence of population on agriculture: Occupational distribution of population in India clearly reflects the backwardness of the economy. Indian agriculture has come a long way since independence, with chronic food scarcity given way to grain self-sufficiency despite a two-and-a-half fold increase in population. In India, even in 2014, about 47 per cent of the working population was engaged in agriculture. However, its contribution to national income was only 17 per cent, according to data published by the World Bank. It shows that there is very low productivity per person in agriculture. Further the expansion of industries was not sufficient to attract the labour force from the primary activities. Hence, India is still lagging on this front.

3. Heavy population pressure: The Indian economy is facing the problem of population explosion. It is clearly evident from the total population of India which was 121.08 crores as per 2011 census. It is the second highest populated country, China being the first. The main problem in India is the high level of birth rates coupled with a falling level of death rates. The fast rate of growth of population necessitates a higher rate of economic growth in order to maintain the same standard of living of the population. To maintain a rapidly growing population, the requirements of food, clothing, shelter, medicine, schooling, etc. all rise. This, a rising population imposes greater economic burden upon the country. The failure to sustain the living standard keeps the poor and underdeveloped countries poor and under developed.

4. Existence of chronic unemployment and under-employment: The huge unemployed working population is another important feature of Indian economy. In India labour is an abundant factor. As a result, it is very difficult to provide gainful employment to the entire working population. Rapid growth of population coupled with inadequate growth of secondary and tertiary occupations are responsible for the occurrence of chronic unemployment and under-employment in our country. In India, unemployment is the result of deficiency of capital.

Moreover, larger number of labour force is engaged in the agricultural sector of the Indian economy than what is really justified in terms of GDP contribution of this sector. This has reduced the marginal product of agricultural labourer either to a negligible amount or to zero. This led to the existence of disguised unemployment in Indian agricultural sector. Adding to this, in urban areas, the problem of educated unemployment has also taken a serious turn. The Planning Commission has estimated that during 2004-05, the rate of unemployment has risen to 8.36 per cent from 7.32 per cent in 1999-2000. Though there is no doubt that unemployment exists in a greater degree in the urban areas, the rural areas also suffer from the problem of unemployment. This is visible from Table 6.9.

Table 6.9: All India Rural & Urban Unemployment Rates (2015-16) (%)

Sector	Male	Female	Transgender	Person
Rural	4.2	7.8	2.1	5.1
Urban	3.3	12.1	10.3	4.9
Rural + Urban	4.0	8.7	4.3	5.0

Source: Labour Bureau, Ministry of Labour & Employment

Note: Unemployment figures are based on Usual Principal Status (UPS) approach.

5. Slow improvement in Rate of Capital Formation: The basic characteristic of the Indian economy was the existence of capital deficiency. However, in recent years, there is a slow and steady improvement in India's rate of capital formation, which is reflected in Table 6.10. In a country like India where the rate of population growth is 1.6 per cent during 2000-05, about 6.4 per cent investment is required to offset the additional burden imposed by increasing population. Thus, India requires as high as 14 per cent level of gross capital formation so that it may cover depreciation and maintain the same level of living. A higher rate of gross capital formation alone can pave the way for economic growth to improve standard of living of the population. It is indeed a pleasure to note that Gross Domestic Saving in 2014 has reached a high level of 29.0 per cent and Gross Capital Formation was also high at 31.4 per cent.

**Table 6.10: Gross Capital Formation and Gross Domestic Saving
as per cent of Gross Domestic Product**

Country	Gross Capital Formation		Gross Domestic Saving	
	1990	2014	1990	2014
USA	18	19.3*	16	16.3*
India	24	31.4	23	29.0
China	35	47.7*	38	50.4*

Source: The World Bank, World Development Indicators (2015)

Notes: * 2013 figure

6. Inequality in wealth distribution: Oxfam's report, 'An economy for the 99 percent', 2017, shows that the gap between rich and poor is far greater than had been feared. As per this report, eight men own the same wealth as the 3.6 billion people who make up the poorest half of humanity. Data on the distribution of global wealth – particularly in India and China – indicates that the poorest half of the world has less wealth than had been previously thought. In India, the richest 1 per cent of Indians have 58 per cent of total Indian wealth. Fifty seven billionaires have the same amount of wealth as the bottom 70 per cent of India. The CEO of India's top information firm earns 416 times the salary of a typical employee in his company.

7. Poor Quality of Human Capital: Indian economy is suffering from the problem of poor quality of capital. If we enlarge the definition capital formation to include the use of any resource that enhances productive capacity, then besides physical capital the knowledge and training of the population will also form a part of capital. Hence, the expenditure on education, skill formation, research and improvements in health are included in human capital. Under United Nations Development Programme (UNDP), countries have been ranked on the basis of Human Development Index (HDI). This index is based on life expectancy, educational attainment and per capita income. It is an alternative indicator of socio-economic development of the country. As per the Human Development Report, India ranks 130 out of 188 countries. India's HDI value for 2014 is 0.609. India has improved its ranking by 6 places between 2009 and 2014 Table 6.11. In comparison to other nations in the BRICS grouping, India has the lowest rank with Russia at 50, Brazil at 75, China at 90, and South Africa at 116.

Table 6.11: India's Position in the Global Human Development Index, 2014

Country	HDI 2014		Change in rank	GNI per capita ($)	LEB (years)	Expected years of schooling (years)	Mean years of schooling (years)
	Value	Rank	2009-14	2014	2014	2014[a]	2014[a]
Brazil	0.755	75	3	15175	74.5	15.2	7.7
Russian Fed.	0.798	50	8	22352	70.1	14.7	12.0
India	**0.609**	**130**	**6**	**5497**	**68.0**	**11.7**	**5.4**
China	0.727	90	13	12547	75.8	13.1	7.5
South Africa	0.666	116	4	12122	57.4	13.6	9.9
World	**0.711**			**14301**	**71.5**	**12.2**	**7.9**

Source: HDR 2015.

Notes: [a]Data refers to 2014 or the most recent year available; Gross National Income (GNI) per capita is based on 2011 dollar purchasing power parity (PPP); LEB is Life Expectancy at Birth.

8. Low level of technology: Prevalence of Low level of technology is one of the important characteristics of developing country like India. Since new techniques are expensive and require a considerable degree of skill for their application in production, the twin requirements for the absorption of new technology are the availability of capital and training of an adequate number of skilled personnel. Hence, deficiency of physical and human capital, in terms of illiteracy and the absence of skilled labour force are the major hurdles in the spread of technology in the economy.

The Indian economy also suffers from this basic weakness that majority of the farmers are too poor to buy even the essential inputs, such as improved seeds, fertilisers and insecticides and machineries like harvesters, tractors, sowing machines etc. Further, majority of the enterprises are also micro and small enterprises, and it is beyond the means for them to employ modern and more productive techniques.

9. Lack of access to basic amenities: According to the Census of India 2011, around 7 per cent of India's population (650 million) lives in rural and slum areas. Only 46.6 per cent of households in India have access to drinking water within their premises. A far lower, 43.5 per cent of households have access to tap water. Similarly, only 46.9 per cent households have toilet facilities within the household premises. This results in developing less strength to fight diseases and is also partly responsible for the low level of efficiency of the Indian workers.

Further, dedicated skilled health personnel are a pre-requisite for efficient and effective delivery of health services. However, the availability of such personnel to meet various needs of the health sector is a huge challenge in India. An evaluation study of the National Rural Health Mission (2011) conducted in seven selected states – UP, MP, Jharkhand, Odisha, Assam, J&K, and Tamil Nadu - assessed the levels of shortfalls faced by some of these states in health personnel. There was a 95 per cent shortage of skilled health personnel in Jharkhand, around 80 per cent in MP and 70 per cent in UP (Table 6.12). Assam reported only 11 per cent shortage, while Tamil Nadu did not report any shortage at all.

Table 6.12: Shortfalls in Health Personnel under the NRHM

States	Number of specialists		Shortfall of specialists
	Required	**Available**	**(per cent)**
UP	2060	618	70
MP	1080	220	80
Jharkhand	776	40	95
Assam	412	365	11

Source: NITI Aayog

10. Demographic characteristics: The demographic characteristics of India are not at all satisfactory rather these are associated with high density of population, a smaller proportion of the population in working age group of 50-60 years and a comparatively larger proportion of population in the minor age group of 0-15 years. As per 2011 census, the density of population in India was 382 per sq. km. as compared with world density of population of 41 per sq km. Even in China, the density is nearly 123 per sq km. Again, as per 2011 census, 29.5 per cent of the total population is in the age group of 0-14 years, 62.5 per cent is in the working age group of 15-59 years and about 8.0 per cent in the age group of 60 and above. All these show that dependency burden of our population is very high.

11. Under-utilisation of natural resources: In respect of natural endowments India is considered as a very rich country. Various types of natural resources, such as land, water, minerals, forest and power resources are available in sufficient quantity in the various parts of the country. However, due to its very inherent problems like inaccessible regions, primitive techniques and shortage of capital such huge resources remained largely under-utilised. Until recently, India was not in a position to develop even 5 per cent of total hydro power potential of the country.

12. Lack of infrastructure: Lack of infrastructural facilities is one of the serious problems from which the Indian economy is suffering till today. These infrastructural facilities include transportation and communication facilities, electricity generation and distribution, banking and credit facilities, health and educational institutes etc. The two most vital sectors i.e. agriculture and industry could not make much headway in the absence of proper infrastructural facilities in the country. Moreover, due to the absence of proper infrastructural facilities, development potential of different regions of the country largely remains under-utilised.

These are some of the major development issues of Indian economy. These issues need to be resolved in order to unlock the full potential of the economy. Addressing these issues would surely result in faster growth in the economy both on the economic front as well as the social front.

DEVELOPMENT EXPERIENCE AND RECENT TRENDS IN INDIAN ECONOMY

For the smooth functioning of any economy, planning plays an important role. The Planning Commission was set up in 1950 and was entrusted with the responsibility of the creation, development and execution of India's five year plans. The guiding principles of India's Five-Year Plans are provided by the basic objectives of growth, employment, self-reliance and social justice. Apart from these basic objectives, each five-year plan took into account the new constraints and possibilities faced during the period and attempted to make the necessary directional changes and emphasis.

India's five year plans used to be supervised by erstwhile Planning Commission. However, in 2014, India has announced the dissolution of the Planning Commission, and its replacement by a think tank called the NITI Aayog (National Institution for Transforming India), which will be discussed later. Here is an overview of India's Five Year Plans

(a) Approach to Each Plan

First Plan (1951-56): The first Prime Minister of India, Jawaharlal Nehru presented the first five-year plan to the

Parliament of India on 8 December, 1951. This plan was based on "Harrod-Domar" model. At the time of this plan India faced three problems – influx of refugees, severe food shortage and mounting inflation. The agricultural sector was hit hardest by the partition of India and needed urgent attention. India had also to correct the disequilibrium in the economy caused by the Second World War. Accordingly, the First Plan emphasized, as its immediate objectives, the rehabilitation of refugees, rapid agricultural development so as to achieve food self-sufficiency in the shortest possible time and control of inflation.

Second Plan (1956-61): The strategy of growth laid emphasis on rapid industrialization with a focus on the development of heavy industries and capital goods, such as iron and steel, chemicals, including nitrogenous fertilizers, heavy engineering and machine building industry. The plan was developed by Professor Mahalanobis.

Third Plan (1961-66): Third Plan set as its goal the establishment of a self-reliant and self-generating economy. However, working of Second Plan had also shown that the rate of growth agricultural production was the main limiting factor in India's economic development. Accordingly, the Third Plan specifically incorporated the development of agriculture as one of the objectives of planning in India besides, for the first times, considering the aim of balanced, regional development.

Enough misfortunes awaited this plan-two wars, one with China in 1961-62 and the other with Pakistan in 1965-66 along the Gujarat border and a severe drought-led famine in 1965-66 had to be faced. Due to heavy drain and diversion of funds, this plan could not meet its targets.

Three Annual Plans: The period of the three consecutive Annual Plans was 1966-69. The original draft outline of Fourth Plan prepared in 1966 under the stewardship of Ashok Mehta had to be abandoned on account of the pressure exerted on the economy by two years of drought, devaluation of the rupee and inflationary recession. Hence, the government decided to go for an Annual Plan for 1966-67. Due to the same reasons the government went for another two such plans in the forthcoming years. It is popularly known as the period of "Plan Holiday".

Fourth Plan (1969-74): The Plan was based on the Gadgil strategy with the two principal objectives of "growth with stability and "progressive achievement of self-reliance". The Fourth Plan aimed at 5.5 per cent average rates of growth in the national income and the provision of national minimum for the weaker sections of the community – this latter came to be known as the objectives of 'growth with justice' and "*Garibi Hatao*" (removal of poverty). Droughts and the Indo-Pak War of 1971-72 led the economy to capital diversions creating financial crunch for the Plan.

Fifth Plan (1974-79): The Plan proposed to achieve the two main objectives, viz., removal of poverty and attainment of self-reliance, through promotion of higher rate of growth, better distribution of income and a very significant increase in the domestic rate of saving. This policy also focused on import substitution and export promotion. Further National program on Minimum needs in which primary education, drinking water, rural roads, housing etc. were included.

Annual Plan (1978-80): This plan was termed as Rolling Plan. In this plan targets of the previous years were to be achieved in next year. Rolling Plan was first advocated by Gunnar Myrdal.

Sixth Plan (1980- 85): This Plan emphasised on socio-economic infrastructure in rural areas; eliminating rural poverty and reducing regional disparities through the Integrated Rural Development Programme (IRDP -1979); 'target group'[1] approach initiated[2].

Seventh Plan (1985 - 90): This Plan was introduced in April 1985, after the country had enjoyed a reasonable

---

1. 'Target group' approach of planning is selecting the group of people where a particular problem is and attacking the problem directly.

2. Followed by this, a number of national level programmes and schemes were launched during the plan, which tried to attend to the specific areas and the specific concerns of socio-economic development (this is 'target group' approach): National Rural Employment Programme (NREP)- 1980, Restructured Twenty-Point Programme- 1982, Biogas Programme- 1982, Development of Women and Children in Rural Areas (DWERA)- 1983, Rural Landless Employment Guarantee Programme (RLEGP)- 1983, Self-Employment to Educated Unemployed Youth Programme (SEEUP)- 1983, Dairy Development Programme (DDP) –1983, Village and Small Industries Development Programme (VSIDP)- 1983, Tribal Development Agency (TDA) –1983, National Seeds Programme (NSP) –1983, Intensive Pulses Development Programme (IPDP)- 1983, Intensive Cotton Development Programme (ICDP)- 1983, Khadi and Village Industries Programme (KVIP)- 1983, Programme for Depressed Areas (PDA) –1983, Special Programme for Women and Children (SPWC)- 1983.

rate of economic growth of the order of 5.4 per cent during the Sixth Plan. The Seventh plan emphasised on rapid foodgrain production, increased employment creation and productivity in general. The basic tenets of planning, i.e., growth, modernisation, self-reliance and social justice remained as the guiding principles.

The approach to the Eighth Five Year Plan (1990-95) was approved in September 1989 and the Eighth Plan was to be introduced in April 1990. However, there were a series of changes in Government at the Centre, necessitating constant reconstitution of the Planning Commission and preparation of a series of versions of the approach to the Eighth Plan (1992-97). Finally, the fourth version of the **Eighth Plan (1992-97)** was approved at a time the country was going through a severe economic crisis caused by a balance of payments crisis, a rising debt burden, ever-widening budget deficits, mounting inflation and recession in industry. The Government initiated the process of fiscal reforms with a view to provide a new dynamism to the economy. The Eighth Plan (1992-97) reflected these changes in its attempt to accelerate economic growth and improve the quality of life of the common man.

The Ninth Plan (1997-2002) was launched when there was an all round 'slowdown' in the economy led by the South East Asian Financial Crisis (1996-97). Though the liberalisation process was still criticised, the economy was very muchout of the fiscal mess of the early 1990s. With a general nature of 'indicative planning' the Plan not only did target an ambitious high growth rate (7 per cent) but also tried to direct itself towards time-bound 'social' objectives. There was an emphasis on the seven identified Basic Minimum Services (BMS) with additional Central Assistance for these services with a view to achieving complete coverage of the population in a time-bound manner. The BMS included

  (i) Safe drinking water;

 (ii) Primary health service;

 (iii) Universalisation of primary education;

 (iv) Public housing assistance to the shelter-less poor families;

  (v) Nutritional support to children;

 (vi) Connectivity of all villages and habitations; and

(vii) Streamlining of the public distribution system.

Tenth Plan: The Plan (2002-07) commenced with the objectives of greater participation of the National Development Council (NDC) in its formulation. Some highly important steps were taken during the plan, which undoubtedly points out to a change in the planning policy mindset of the policy makers, major ones being:

  (i) Doubling per capita income in 10 years;

 (ii) Accepting that the higher growth rates are not the only objective – it should be translated into improving the quality of life of the people;

 (iii) For the first time the Plan went to set the 'monitorable targets' for eleven select indicators of development for the centre as well as for the states;

 (iv) 'Governance' was considered a factor of development;

  (v) States' role in planning to be increased with the greater involvement of the PRIs;

 (vi) Policy and institutional reforms in each sector, i.e., reforms in the PSUs, legal reforms, administrative reforms, labour reforms, etc.;

(vii) Agriculture sector declared as the prime moving force (PMF) of the economy;

(viii) Increased emphasis on the social sector (i.e., education, health, etc.);

 (ix) Relevance between the processes of economic reforms and planning emphasised; etc.

Eleventh Plan (2007-12): The NDC in December 2006 approved the Approach to the 11th Plan document titled "*Towards Faster and more Inclusive Growth"* and directed the Planning Commission to prepare a detailed plan to assess the resources required to meet the broad objectives in Approach Paper. The detailed version of the Eleventh Five Year Plan (2007-02) was approved by the NDC in December 2007.

Objectives of the Eleventh Plan: The Plan envisaged a high growth of GDP of the order of 9 per cent for the country as a whole. This implies that per capita GDP would grow at about 7.5 per cent per year to double in 10 years. Further, the Plan documents add that the target is not just faster growth but also inclusive growth which ensures broad based improvement in the quality of life of the people, especially of the poor SCs/STs, OBCs and the minorities.

Vision for the Eleventh Plan: The broad vision of the 11th Plan includes several inter-related components.

1. Rapid growth that reduces poverty and creates employment opportunities;

2. Access to essential services in health and education especially for the poor;

3. Empowerment through education and skill development;

4. Extension of employment opportunities using National Rural Employment Guarantee Programme;

5. Environmental sustainability;

6. Reduction of gender inequality; and

7. Improvement in governance.

Twelfth Plan (2012 – 17): In December 2012, the Planning Commission published the near final draft of 12th Five Year Plan – Faster, More Inclusive and Sustainable Growth. Simultaneous achievement of these elements is seen as crucial for the success of the Plan. The 12th Plan says that '*[it] must be guided by a vision of India moving forward in a way that would ensure a broad-based improvement in living standards of all sections of the people through a growth process which is faster than in the past, more inclusive and also more environmentally sustainable'*.

Thus the 12th Plan called for more attention to be given to problem of sustainability. It states that *'No development process can afford to neglect the environmental consequences of economic activity, or allow unsustainable depletion and deterioration of natural resources'* and several chapters are devoted to the issues of water, land use, environment, forestry and wildlife.

The Draft Approach Paper lays down the major targets of the Plan, the key challenges in meeting them, and the broad approach that must be followed to achieve the stated objectives which are summed-up as follows:

1. Growth rate of 9 per cent is targeted for the Plan.

2. It emphasizes the need to intensify efforts to have 4 per cent average growth in the agriculture sector during the Plan period.

3. Restrain inflationary pressure.

4. It proposes that the major flagship programmes which were instrumental for promoting inclusiveness in the Eleventh Plan should continue in the Twelfth Plan.

5. For the GDP to grow at 9 per cent, commercial energy supplies will have to grow at a rate between 6.5 and 7 per cent per year.

6. It suggests the need to take steps to reduce energy intensity of production processes, increase domestic energy supply as quickly as possible and ensure rational energy pricing.

7. It draws attention to evolving a holistic water management policy.

8. It argues for a new legislation for land acquisition.

9. It maintains that health, education and skill development will continue to be the focus areas in the Twelfth Plan.

10. Requirement of large investments in infrastructure sector development—Public Private Partnerships (PPPs) in infrastructure sector needs to be imparted so that present infrastructure shortages can be addressed early.

11. It has emphasised the importance of the process of fiscal correction.

12. It also emphasizes the need for focusing more on efficient use of available resources in view of the resource constraints.

(b) Introduction of NITI Aayog in India

NITI (National Institution for Transforming India) Aayog is basically a policy think tank of Government of India and State Governments that replaces 65-year old Planning Commission. Union Government of India had announced the formation of NITI Aayog on 1st January, 2015.

The NITI Aayog has a governing council comprising of all State Chief Ministers and Lt. Governors of Union Territories and it works towards fostering a 'Co-operative federalism' for providing a "national agenda" to the Centre and States.

The body is comprised of a CEO and a Vice Chairperson, to be appointed by the Prime Minister, in addition to some full-time members and two part-time members, while four Union Ministers would serve as ex-officio members. Besides, there would be specific regional councils, while experts and specialists from various fields would be called as special invitees nominated by the Prime Minister.

NITI Aayog serves as a "think tank" of the government as a "directional and policy dynamo" and would provide both to the governments at the centre and in the states with strategic and technical advice on key policy matters including economic issues of national and international importance.

Thus NITI Aayog will never plan, rather it will formulate policy. By following these policies, various Ministries of the Central Government will prepare developmental projects considering the need of long term development. NITI is in favour of cooperative federal structure where both the Centre and States jointly prepare developmental policies. But NITI, at the same time, wants to promote healthy competition among the developing states.

NITI Aayog is to have regional councils to focus on developmental activities on specific areas and is patterned on the National Reforms Development Commission of China.

While the Planning Commission had the power to allocate funds to states for attaining regional development, the NITI Aayog does not have such powers. Rather, the task of allocating funds to states is now being vested with the Finance Ministry's Department of Expenditure. Its primary job would be to undertake long term policy and design frameworks and take necessary initiatives for attaining faster development and finally to monitor these activities sincerely. Thus, NITI Aayog is responsible for actively monitoring and evaluating implementation of the Government programmes and initiatives.

It was time to consider developing such an institution that would provide direction to Monitoring and Evaluation (M & E) activities in the country, laying stress upon quality standards and sound ethical procedures and provide for appropriate institutional mechanisms.

NITI Aayog would therefore mean:

   (a) A group of people with authority entrusted by the government to formulate/regulate policies concerning transforming India.

   (b) It is a commission to assist government in both social and economic issues.

   (c) It is an institute of think tank with experts in it.

   (d) It is an body to actively monitor and evaluate implementation of government programmes and initiatives.

The following are some of the important aims and objectives of NITI Aayog:

   1. NITI Aayog sets its aims to provide a critical directional and strategic input into the development process of the country.

   2. NITI Aayog aims to serve as a "think tank" of the government both at central and state levels with relevant strategic and technical advice on key policy matters including economic issues of national and international importance.

   3. NITI Aayog now seeks to replace the centre-to-state one way flow of policy framed by the Planning Commission by an amicable settled policy framed by a genuine and continuing partnership of states.

   4. The NITI Aayog will also seek to put an end to slow and tardy implementation of policy by fostering better Inter-Ministry co-ordination and better centre-state co-ordination. It will help evolve a shared vision of

national development priorities, and foster co-operative federalism, in order to focus on the view that strong states make a strong nation.

5. The NITI Aayog has set it objectives to develop mechanisms to formulate credible plans to the village level and aggregate these progressively at higher levels of government. This Aayog will ensure special attention to the sections of society that may be at risk of not benefitting adequately from economic progress.

6. The NITI Aayog, will create a knowledge, innovation and entrepreneurial support system through a collaborative community of national and international experts, practitioners and partners. The Aayog will offer a platform for resolution of inter-sectoral and inter-departmental issues in order to accelerate the implementation of the development agenda.

7. The NITI Aayog will monitor and evaluate the implementation of programmes, and focus on technology upgradation and capacity building.

Undertaking the above activities, the NITI Aayog will aim to accomplish the following objectives and opportunities:

(i) An effective administration paradigm in which the Government is an "enabler" rather than a "provider of first and last resort".

(ii) Attaining progress from "food security" to focus on a mix of agricultural production as well as attain actual returns that farmers get from their produce.

(iii) To ensure that India is an active player in the debates and deliberations on the global commons.

(iv) To ensure that the economically vibrant middle-class remains actively engaged, and its potential is fully utilized.

(v) Leverage India's pool of entrepreneurial, scientific and intellectual human capital.

(vi) Incorporate the geo-economic and geo-political strength of the non-resident Indian Community.

(vii) Use urbanization as an opportunity to create a wholesome and secure habitat through the use of modern technology.

(viii) Use technology to reduce opacity and potential for misadventures in governance.

Moreover, the NITI Aayog aims to enable India to face complex challenges in a better way through the following measures:

(i) Leveraging of India's demographic dividend and realization of the potential of youth, men and women through imparting education, skill development, elimination of gender bias and also by providing employment.

(ii) Elimination of poverty, and the enhance the chance for every Indian to live a life of dignity and self-respect.

(iii) Redressal of inequalities based on gender bias, caste and economic disparities.

(iv) To integrate villages institutionally into the development process of the country.

(v) To provide policy support to more than 50 million small business which are considered as a major source of employment generation.

(vi) To safeguard our environmental and ecological assets.

The positive side of this institution is to establish a dynamic institutional mechanism where eminent individuals outside the government system could contribute towards policy making. Moreover, one of the major tasks of the NITI Aayog is to actively monitor and evaluate implementation of programmes and initiatives which is something new under the present setup.

MAJOR SECTORAL POLICIES

Apart from the above mentioned development experiences of India through planning, India also adopted various sector-oriented policies, which are summarised as below:

(c) AGRICULTURE

GREEN REVOLUTION

After independence in 1947, India was not in a condition to fulfill the requirement of food grains. It was partly due to the natural calamities like famines, droughts, which struck India from time to time and created the problem of food shortage. As a result, India had to depend on other countries to overcome this problem. Our farmer suffered a lot. Hence, to improve agricultural production India adopted the new agricultural strategy during 1964-65. Green Revolution aimed to improve agricultural production within a limited period and maintaining a high level agricultural production over a long period of time.

To improve agricultural production, green revolution envisages several steps, such as, use of high yielding varieties of seeds, use of chemicals, fertilizers, pesticides, use of improved technology, multiple cropping, irrigation facilities, providing agricultural credit to farmers, suitable price mechanism for agricultural production and land reforms. As a result of green revolution, the production of food grains increased a lot which ultimately made India a self reliant country in food grains.

Before the Green Revolution, farmers were struggling to make both ends meet as they did not have any agricultural facilities. For irrigation, they depend on monsoon which always played a gamble with them. If monsoon failed their crops failed which ultimately increased their problem.

In such circumstances, the Green Revolution proved to be a boon. It not only made our country self-sufficient in food-grains but also improved the condition of our farmers. They started living a comfortable life. Thus, Green Revolution helped in satisfying the appetite of the millions as well as relieving the economy of the problem of food scarcity.

(d) Industrial Policy

Any government action aimed at affecting industry may be considered to be part of industrial policy. A country must formulate industrial policy as an instrument of industrialisation. In a country like India, where private sector is allowed to co-exist in business, its control and regulation is necessary. Industrial policy is a necessary step in this direction.

Industrial Policy Resolution (IPR) of 1948:

In a mixed economy of our sort, the government should declare its industrial policy clearly indicating what should be the sphere of the State and of the private enterprise. A mixed economy means co-existence of the two sectors public and private. This the Government of India did by a policy resolution on 30 April 1948 called the first Industrial Policy Resolution of 1948, which made it clear that India was going to have a mixed economy.

The Industrial Policy Resolution, 1948 divided the industrial structure into four groups:

1. Basic and strategic industries such as arms and ammunition, atomic energy, railways, etc., shall be the exclusive monopoly of the State.

2. The second group consisted of key industries like coal, iron and steel, ship-building, manufacture of telegraph, telephone, wireless apparatus, mineral oils, etc. In such cases the State took over the exclusive responsibility of all future development and the existing industries were allowed to function for ten years after which the State would review the situation and explore the necessity of nationalisation.

3. In the third group, 18 industries including automobiles, tractors, machine tools, etc., were allowed to be in the private sector subject to government regulation and supervision.

4. All other industries were left open to the private sector. However, the State might participate and/or intervene if circumstances so demanded.

The IPR1948 also emphasised the importance of small-scale and cottage industries in the Indian economy. The Industries (Development and Regulation) Act was passed in 1951 to implement the Industrial Policy Resolution, 1948.

Industrial Policy Statement of 1956:

On 30 April 1956, the Government revised its first Industrial Policy (i.e., the policy of 1948), and announced the Industrial Policy of 1956. The reasons for the revision were: (i) introduction of the Constitution of India, (ii) adoption of a planning system, and (iii) declaration by the Parliament that India was going to have a socialist pattern of society.

All these principles were incorporated in the revised industrial policy as its most avowed objectives.

The Resolution classified industries into three categories having regard to the role which the State would play in each of them:

1. Schedule A consisting of 17 industries would be the exclusive responsibility of the State. Out of these 17 industries, four industries, namely arms and ammunition, atomic energy, railways and air transport would be Central Government monopolies; new units in the remaining industries would be developed by the State Governments.

2. Schedule B, consisting of 12 industries, would be open to both the private and public sectors; however, such industries would be progressively State-owned.

3. Schedule C contained the remaining industries. In their case the private sector was to have the main initiative of development. However, they were to fit within the economic and social priorities and policies of the government and were to be subject to the provisions of the Industries (Development and Regulation) Act, 1951.

The IPR 1956, stressed the importance of cottage and small-scale industries for expanding employment opportunities and for wider decentralisation of economic power and activity. Regional disparities in industrialisation were to be reduced.

Industrial Policy Statement, 1977

The industrial policy was modified in December 1977 by a Statement in the Parliament. The main thrust of the new industrial policy was a shift in favour of small-scale sector, which was classified, into three sectors.

1. Cottage and household industries, which were a source of providing large scale self-employment.

2. Tiny sector which was defined on the basis of a small amount of investment below a specified limit.

3. Small-scale industries which were larger than the first two categories but were still having investment within specified limits.

Industrial Policy of 1980

IPR 1980 re-emphasised the socio-economic goals of the earlier policy with certain clarifications and extensions, namely,

– optimum utilisation of installed capacity

– achieving higher productivity and more employment

– removal of regional disparities through preferential treatment for industrially backward districts

– enlarging the agricultural base of the economy be extending a preferential treatment to agro-related industries

– faster promotion of export oriented and import substitution industries.

The Policy statement of 1980 also recognised the need for improving the management of the public sector undertakings.

Industrial Policy of 1991:

The new liberalised industrial policy was announced by the Government of India on 24 July 1991. The New Industrial Policy scrapped the asset limit for Monopolies and Restrictive Trade Practices (MRTP) companies and abolished industrial licensing of all projects, except few groups. It has raised the limit for foreign participation or foreign capital in the country's industrial landscape.

The highlights of the new policy are:

1. Industrial licensing would be abolished for all projects except selected sectors. The exemption from licensing would apply to all substantial expansion of existing and new units.

2. The policy provided for automatic clearance for import of capital goods.

3. As for the MRTP Act, the policy states that the pre-entry scrutiny of investment decisions by the so-called MRTP companies would no longer be required.

5. The policy intended to scrap the asset limit of the MRTP companies.

6. The policy envisaged disinvestment of government equity in public sector to mutual funds, financial institutions, general public and workers.

   Reservation for the public sector, as on 2008, was very limited (just 2)—covering only manufacturing involving certain substances relevant for atomic energy (as well as production of atomic energy) and provision of railway transport.

7. In order to invite foreign investment in high priority industries, it was decided to provide approval for direct foreign investment up to 51 per cent forewing equity in such industries.

8. In a departure from the prevailing locational policy for industries, the policy provides that in locations other than cities of population of more than one million, there would be no requirement for obtaining industrial approvals except for industries subject to compulsory licensing.

(e) Economic Reforms

The year 1991 will remain a year of notable and widespread changes in the policy for economic development in India. Soon after resumption of office on June 21, 1991, the Government adopted a number of stabilization measures that were designed to restore external and internal confidence. The government has made some radical changes in its policies regarding foreign investment, trade, exchange rate, industry, banking and fiscal affairs, etc. These various elements taken together constitute an economic policy. Since July, 1991 Government of India has announced several new policies under the name of 'New Economic Reforms'. The new economic reforms gave a new direction and dimension to the Indian economy.

Nature of economic reforms :

The nature of new economic reforms / policy can be stated as below :

(i) Liberalisation: The fundamental feature of the new economic reforms was that it provided freedom to the entrepreneurs to establish any industry / trade / business venture. The meaning of the economic liberalization is freedom for economic decision. It means producers, owners and consumers of factors of production all are free to take their decisions to promote their self-interest. Government announced liberalization policy in the industrial sector, external sector (foreign trade exchange rate), banking and financial sector, fiscal sector etc. The capital markets had also been freed and opened to the private enterprises. Since 1991 foreign equity participation upto 51 percent or more was allowed. Industrial sector was de-licensed. Monopolies and Restrictive Trade Practices (MRTP) Act was abolished. Infrastructure sector was opened to the foreign investments. Foreign Exchange Regulation Act (FERA) was amended and Foreign Exchange Management Act (FEMA) was enacted.

(ii) Extension of privatization: Another feature of the new economic reforms was the extension in the scope of privatization. Privatization is a process which reduces the role of state or public sector in country's economic activity. The basic objective of privatization everywhere is to improve the performance of public sector undertakings so as to lessen the financial burden on tax payers. In the wave of privatization, out of 17 industries reserved for public sector, 11 industries had been given to the private sector. The scope of privatization had further been extended by offering greater opportunities of investment to the foreign private investors.

(iii) Globalisation of economy: The new economic reforms/policy have made the economy outwardly oriented. Globalisation refers to a process of increasing economic integration and growing economic interdependence between countries in the world economy. Economic liberalization and privatization of public sector enterprises ultimately lead to globalization of the economy. In other words, globalization is nothing but flow of capital, commodities, technology and labour across the national boundaries. Now economic activities were to be governed both by domestic market and also the world market.

Paradigm of economic reforms in India since 1991:

Government of India accepted and adopted the new economic reforms. The new economic reforms gave a new direction and dimension to the Indian economy. Following table summarises paradigm of economic reforms in India since 1991.

Table 6.13: Paradigm of economic reforms in India since 1991

Pre-Reform Period	*Post-Reform Period*
1. Quantitative licensing on trade and industry.	1. Abolition of industrial and trade licensing.
2. State regulated monopolies of utilities and trade.	2. Removal of state monopolies, privatization and disinvestment.
3. Restriction of foreign investment and technology.	3. Liberal regime for FDI, portfolio investment, foreign technology.
4. Government control on finance and capital markets	4. Liberalisation of financial and capital markets.
5. Foreign exchange control, no convertibility of rupee.	5. Abolition of exchange control, full convertibility on current account.
6. Import substitution and export of primary goods.	6. Export promotion and export diversification, no import bias.
7. Closed economy.	7. Open economy.
8. High duties and taxes with multiple rates and large dispersion.	8. Reduction and rationalization of taxes and duties.
9. Self-reliance.	9. Integrate with world markets.
10. Explicit subsidies on food, fertilizers and some essential items.	10. No change budget subsidies on LPG, Kerosene.
11. Frequent state interventions.	11. Selective and effective state interventions.
12. Sector-specific monetary, fiscal and tariff policies.	12. Sector-neutral monetary, fiscal and tariff policies.
13. Not much concern for deficit.	13. Contain all kinds of deficit.
14. Hidden subsidies on power, urban transport, public goods.	14. No change, but user charges are being rationalized and subsidies targeted.
15. PSUs as engine of growth	15. Private investment as engine of growth.
16. End-use and sector-specific multiple and controlled interest rates.	16. Flexible interest rates without any end-use or sector specifications.
17. Restrictions on Foreign Direct Investment (FDI) and Multi-national corporations (MNCs)	17. Inducement to FDI and MNC's
18. Tax concessions on exports and savings.	18. Rationalised and being phased out.
19. Underdeveloped capital market.	19. Reforms in capital market.
20. State controlled credit.	20. Credit policy reforms.

Recent Trends in Indian Economy

(f) Goods and Services Tax (GST)

On 3rd August, 2016 entire India was glued to Rajya Sabha when the Rajya Sabha successfully passed the Constitutional Amendment Bill for GST (GST Bill) which eventually paved the way for GST in India. As the Government is likely to introduce GST soon in 2017, it is important for us to understand the GST from an economic perspective.

Introduction:

GST is abbreviation for Goods and Service Tax. GST is also known as Value Added Tax (VAT) in few countries. GST / VAT is a consumption based tax wherein the basic principle is to tax the value addition at each business stage. To achieve this, tax paid on purchases is allowed as a set off/ credit against liability on output/income.

GST is levied on all transactions of goods and services. Thus, in principle, GST should not differentiate between 'goods' and 'services'.

Internationally, GST was first introduced in France and now more than 150 countries have introduced GST. Most of the countries, depending on their own socio-economic formation, have introduced National level GST or Dual GST.

GST is different from the current tax structure in many ways. Currently, taxes treat goods and services differently. However, 'goods' attract Excise at manufacturing level and VAT at the time of sale. In contrast, services attract only one levy i.e. Services tax on provision of taxable services.

Table 6.14: Chronology on GST

1974	Report of L K Jha committee suggested moving to Value Added Tax (VAT) regime
1986	Introduction of Modified Value Added Tax (MODVAT)
1991	Chelliah Committee recommends VAT/Goods and Service Tax (GST)
1 July 1994	Service tax introduced in India
1999	'Empowered Committee' comprising of representatives of 29 States is formed for the purpose of introduction of State VAT
2000	Implementation of uniform State sales tax rates (1%, 4%, 8%, 12%)
2002	Introduction of input credit against services of same category
1 April 2003	Haryana was the first State to introduce VAT. Subsequently, other States introduced VAT (20 States in 2005, 5 States in 2006, Tamil Nadu in 2007 and lastly, Uttar Pradesh in 2008).
1 September 2004	Central level taxes integrated by introduction of 'CENVAT'
January 2005	White Paper, a policy document indicating basic policies of 'State VAT' was released by the 'Empowered Committee'
1 April 2005	Value Added Tax introduced in 20 States
February 2006	Finance Minister comments in the Budget Speech that there is a large consensus that the country must move towards a national level GST that must be shared between the Center and States. He proposes 1 April 2010 as the date for introducing GST.
1 April 2006	VAT implemented in 5 more States
1 January 2007	VAT implemented in Tamil Nadu
April 2007	Central Sales Tax phase out initiated (4% -2%)
May 2007	Empowered Committee (EC) of State Finance Ministers in consultation with Central Government, constituted a Joint Working Group (JWG), consisting of officers of Central and State Governments

	to examine various models and options for GST and to give their assessment of the same to the EC.
November 2007	Joint Working Group (JWG) presented its report on the GST to the EC. The EC accepted the report on GST submitted by the JWG
1 January 2008	VAT implemented in Uttar Pradesh
April 2008	Empowered Committee finalises the overall strategy for GST introduction in India
November 2009	'First Discussion Paper' on GST released by EC
December 2009	Task Force submits its report on GST to 13th Finance Commission
January 2009	Department of Revenue releases its comments on 'First Discussion Paper' on GST
February 2011	IT strategy (by Mr. Nandan Nilekani) for GST released
March 2011	115th Constitution Amendment Bill (CAB) introduced in Parliament. However, this 115th CAB was lapsed (in May 2014) with change of Government at Center.
1 April 2011	Point of Taxation Rules, 2011 introduced
1 July 2012	- Negative List of Services Regime introduced- Place of Provision of Services Rules, 2012 introduced
July 2014	Union Finance Minister states in the Budget Speech 2014 that "I do hope we are able to find a solution in the course of this year and approve the legislative scheme which enables the introduction of GST"
19 December 2014	The 122nd Constitutional Amendment Bill was introduced in Lok Sabha
6 May 2015	The 122nd CAB passed in Lok Sabha
6 May 2015	122nd CAB introduced in Rajya Sabha. 122nd CAB was referred to Select Committee of Rajya Sabha
22 July 2015	Rajya Sabha Select Committee tabled its report
October 2015	Government placed in the public domain four reports on key business processes i.e. registration, payment, refunds and returns in GST regime.
14th June 2016	Draft GST law made available
3 August 2016	Rajya Sabha cleared the Constitutional Amendment Bill
8 August 2016	Lok Sabha cleared the Constitutional Amendment Bill
September/ October 2016	- More than half of the State Assemblies need to ratified the Constitutional Amendment Bill - Presidential Assent received- GST Council formed
November 2016	Formation of GST Council and recommendation on GST Law by GST Council
January 2017	- GST Council consensus for GST to be rolled out from 1st July, 2017 instead of 1st April, 2017 - Notification of GST Rules
February/March 2017	- Draft Laws approved by GST Council - Cabinet approves for GST Supplementary Bill - Lok Sabha passes the Four Bills
April 2017	- Rajya Sabha passes the Four Bills - President gives assent to Four GST Bills

Structure: India is proposing to implement 'dual GST'. In 'dual GST' regime, all the transactions of goods and services made for a consideration would attract two levies i.e. CGST (Central GST) and SGST (State GST).

Taxes that will be subsumed in GST: GST would be levied on all the transactions of goods and services made

for a consideration. This new levy would replace almost all of the indirect taxes. In particular, it would replace the following indirect taxes:

At Central level

– Central Excise Duty (including Additional Duties of Excise)

– Service Tax

– Additional Customs Duty, commonly known as Countervailing Duty (CVD) (levied on imports in lieu of Excise duty)

– Special Additional Duty of Customs (SAD) (levied on imports in lieu of VAT)

– Surcharges, and

– Cesses.

At State level

– VAT/Sales tax

– Entertainment tax (unless it is levied by the local bodies)

– Luxury Tax

– Taxes on lottery, betting and gambling

– Entry tax not in lieu of Octroi

– State Cesses and Surcharges in so far as they relate to supply of goods and services.

However, certain items / sectors would be outside the GST regime.

There was a burden of "tax on tax" in the pre-existing Central excise duty of the Government of India and sales tax system of the State Governments. The introduction of Central VAT (CENVAT) has removed the cascading burden of "tax on tax" to a good extent by providing a mechanism of "set off" for tax paid on inputs and services upto the stage of production, and has been an improvement over the pre-existing Central excise duty. Similarly, the introduction of VAT in the States has removed the cascading effect by giving set-off for tax paid on inputs as well as tax paid on previous purchases and has again been an improvement over the previous sales tax regime.

But both the CENVAT and the State VAT have certain incompleteness. The incompleteness in CENVAT is that it has yet not been extended to include chain of value addition in the distributive trade below the stage of production. It has also not included several Central taxes, such as Additional Excise Duties, Additional Customs Duty, Surcharges etc. in the overall framework of CENVAT, and thus kept the benefits of comprehensive input tax and service tax set-off out of the reach of manufacturers/dealers. The introduction of GST will not only include comprehensively more indirect Central taxes and integrate goods and services taxes for set-off relief, but also capture certain value addition in the distributive trade.

Similarly, in the present State-level VAT scheme, CENVAT load on the goods has not yet been removed and the cascading effect of that part of tax burden has remained unrelieved. Moreover, there are several taxes in the States, such as, Luxury Tax, Entertainment Tax, etc. which have still not been subsumed in the VAT. Further, there has also not been any integration of VAT on goods with tax on services at the State level with removal of cascading effect of service tax. In addition, although the burden of Central Sales Tax (CST) on inter-state movement of goods has been lessened with reduction of CST rate from 4% to 2%, this burden has also not been fully phased out. With the introduction of GST at the State level, the additional burden of CENVAT and services tax would be comprehensively removed, and a continuous chain of set-off from the original producer's point and service provider's point upto the retailer's level would be established which would eliminate the burden of all cascading effects, including the burden of CENVAT and service tax. This is the essence of GST. Also, major

Central and State taxes will get subsumed into GST which will reduce the multiplicity of taxes, and thus bring down the compliance cost. With GST, the burden of CST will also be phased out.

Thus GST is not simply VAT plus service tax, but a major improvement over the previous system of VAT and disjointed services tax.

GST is a tax on goods and services with comprehensive and continuous chain of set-off benefits from the producer's point and service provider's point upto the retailer's level. It is essentially a tax only on value addition at each stage, and a supplier at each stage is permitted to set-off, through a tax credit mechanism, the GST paid on the purchase of goods and services as available for set-off on the GST to be paid on the supply of goods and services. The final consumer will thus bear only the GST charged by the last dealer in the supply chain, with set-off benefits at all the previous stages.

(g) Demonetization

Demonetization is the act of stripping a currency unit of its status as legal tender. The old unit of currency must be retired and replaced with a new currency unit.

On 8 November 2016, Prime Minister Narendra Modi announced the "demonetisation" of Rs. 500 and Rs. 1,000 notes with effect from the next day. Through the days after 8 November, the government has made a set of claims with regard to the objectives and outcomes of the demonetisation scheme.

The first claim is that demonetisation would plug terror financing. The second claim is that demonetisation would help unearth "black money." The third claim is that the unearthed black money would expand the fiscal space of the government. When unaccounted cash is not returned to the banking system, the Reserve Bank of India (RBI) can use the savings to pay the government a dividend. The fourth claim is that demonetisation would help reduce interest rates in the banking system. The fifth claim is that demonetisation would help formalise India's informal economy, reduce the extent of transactions in cash and help create a "less-cash economy." A number of incentives have been offered to induce people to use digital transactions.

Demonetization as a cleaning exercise may produce several good things in the economy. The immeasurable benefits of having more transparency and reduced volume of black money activities can be expected as long term benefits.

INDIAN ECONOMY IN COMPARISON TO MAJOR ECONOMIES OF THE WORLD

Per Capita Income

The per capita income of an Indian was $ 1,560 in 2014 and it has increased to $ 1,600 in 2015. However, in both years compare to other major countries (USA, UK, Japan, Germany and China) India's Per Capita GNI was low. It may be noted that in 2015 the average Per Capita GNI of USA was around 35 times that of India, while at the purchasing power parity rates, it was around 10 times only. In other words, per capita income at official exchange rates exaggerated the disparities, while the purchasing power parity figures corrected the position. Even after this adjustment, though the per capita income differences got narrowed down, still the difference between the level of living of an average American and an Indian was quite large and significant.

The comparison could be extended to another competitive developing country like China. In the year 2015, the average Per Capita GNI of China was around 5 times higher than that of India, while at the purchasing power parity rates, it was around 2 times only. Still, the difference between the people level of living of in China and in India is notable.

Table 6.15: GNI Per Capita (in US Dollars)

Countries	Atlas method		Purchasing Power Parity Basis	
	2014	2015	2014	2015
USA	55,320	55,980	56,130	57,540
UK	43,760	43,390	39,810	40,610
Japan	41,920	36,680	38,060	38,870
Germany	47,680	45,940	47,560	48,410
India	1,560	1,600	5,610	6,030
China	7,520	7,930	13,460	14,390

Source: World Development Indicators.

Population growth in different countries

Population growth rates are different in different parts of the world. Taking the world as a whole, the total population was 4,426 million in 1980 and stood at 6,893 million in 2010. The annual compound growth rate during 1980-90 was 1.7 per cent, but the growth rate started declining thereafter, and was of the order 1.15 per cent in 2009-10.

Table 6.16: World Population Growth by Groups

Country Group	Population (millions)		Per cent of total		Average annual growth rate (%)	
	1980	2010	1980	2009	1980-90	2000-09
Low income economies	1,384	796	31.3	11.5	2.4	2.10
Middle income economies	2,217	4,970	50.0	72.1	1.6	1.12
High income economies	825	1,127	18.6	16.3	0.7	0.63
Total	4,426	6,893	100.0	100.0	1.7	1.15
India	687	1,210	15.5	17.6	2.1	1.38
China	981	1,338	22.2	19.4	1.5	0.51

From Table 6.16, it is obvious that the low income economies and the middle income economies accounted for about 84 per cent of the world population. During 2000-09, the growth rate in low income economies was relatively higher at 2.10 per cent and that in middle income economies was slightly lower at 1.12 per cent.

China and India, which account for nearly 37 per cent of the world population show a population growth of about 2 per cent per annum during 1980-90. It may, however, be noted that in China, population growth rates declined to 0.5 per cent during 2000-09, whereas in India, it is still high at 1.38 per cent. Indeed, India ranked first in terms of annual population increase for the period 2010-15 (World Population Prospects, 2015).

Sex Composition of Population

From Table 6.17, we find that aggregate sex ratio has declined in the world from 986 in 2001 to 982 in 2015. Indeed, almost all the countries trend has been similar, except Brazil. In Brazil, during the same period, the sex ratio has increased from 1,025 to 1,034. In India, a disturbing revelation of 1991 census is the decline in the ratio

of females per 1,000 males. The sex ratio declined from 934 in 1981 to 929 in 1991. In 2011, there is a slight improvement in the proportion of females to 940. However, it declined in 2015 to 929 females per 1,000 males, while in USA it stood at 1,017.

Table 6.17: Trends in Sex Ratio in Five Most Populous Countries (2001-2015)
(females per 1,000 males)

Country	2001	2011	2015
World	986	984	982
China	944	926	941
India	933	940	929
USA	1,029	1,025	1,017
Indonesia	1,004	988	986
Brazil	1,025	1,042	1,034

Source: World Population Prospects

Human Development: International Comparisons

Under United Nations Development Programme (UNDP), countries have been ranked on the basis of Human Development Index (HDI). This index is based on life expectancy, educational attainment and per capita income. It is an alternative indicator of socio-economic development of a country. As per the Human Development Report (HDR) 2015, India ranks 130 out of 188 countries. India's HDI of 0.609 is below the average countries in the medium human development group (0.630) but marginally higher than the HDI average of South Asian countries (0.607). Between 1980 and 2014, India's GNI per capita increased by about 338 per cent. Over the same period, the Life Expectancy at Birth (LEB) increased by 14.1 years, mean years of schooling by 3.5 years and expected years of schooling by 5.3 years. As compared to other BRICS nations, India reports the least mean years of schooling, and LEB that is lower than that of Brazil, China and Russia, but higher than that of South Africa. Bangladesh, with a lower GNI per capita than India, has a much higher LEB. The indices of education as reflected by the HDI show that the progress made in the education sector needs to be faster, with a focus on greater coverage.

Table 6.18 : HDI Component Indices of Selected Countries 2014 and 1980

Country	HDI 2014					HDI 1980				
	LEB (years)	Expected years of schooling (years)[a]	Mean years of schooling (years)[a]	GNI Per Capita ($)	HDI Value	LEB (years)	Expected years of schooling (years)	Mean years of schooling (years)	GNI Per Capita ($)	HDI Value
Russian Fed.	70.1	14.7	12	22352	0.798	67.3	12.2	7.1		
Sri Lanka	74.9	13.7	10.8	9779	0.757	68.2	10	7.1	2562	0.571
Brazil	74.5	15.2	7.7	15175	0.755	62	9.9	2.5	10457	0.547
China	75.8	13.1	7.5	12547	0.727	66.5	8.4	3.9	758	0.43
South Africa^	57.4	13.6	9.9	12122	0.666	56.9	11.1	4.8	9756	0.569
India	68	11.7	5.4	5,497	0.609	53.9	6.4	1.9	1255	0.362
Bangladesh	71.6	10	5.1	3191	0.57	53.5	4.9	2	1148	0.338
Pakistan	66.2	7.8	4.7	4866	0.538	57	3.7	1.8	2437	0.353

Source: HDR 2015.

Notes:[a] Data refers to 2014 or the most recent year available; $: Gross National Income (GNI) per capita is based on 2011 dollar purchasing power parity (PPP); ^ In respect of South Africa, data for 1980 is based on HDR 2014.

Along with HDI, HDR 2015 also gives information on the Gender Development Index (GDI) for all the 188 countries. The HDI value for females in India is 0.525 in 2014, which remains unchanged in comparison to that in 2013. As can be seen in Table 6.19, expect Pakistan, all the other four South Asian countries have reported higher HDI values for females in comparison to India. The mean years of schooling for girls in India at 3.6 years is substantially lower than the figure for males and shows the extent of educational deprivation of girl children in India's cultural context.

Table 6.19: GDI of Selected Countries 2014

Country	GDI		HDI Value		LEB (years)		Expected years of schooling		Mean years of schooling		GNI Per Capita ($)	
	Value	Group										
	2014		2014		2014[a]		2014[a]		2014		2014	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Sri Lanka	0.948	3	0.73	0.769	78.2	71.5	14.2	13.3	10.7	10.9	5452	14307
China	0.943	3	0.705	0.747	77.3	74.3	13.2	12.9	6.9	8.2	10128	14795
India	0.795	5	0.525	0.66	69.5	66.6	11.3	11.8	3.6	7.2	2116	8656
Bangladesh	0.917	4	0.541	0.59	72.9	70.4	10.3	9.7	4.5	5.5	2278	4083
Pakistan	0.726	5	0.436	0.601	67.2	65.3	7	8.5	3.1	6.2	1450	8100

Source: HDR 2015.

Notes: Countries are categorized into five Groups based on their absolute deviations of HDI values between men and women. Group 5 represents low equality of HDI values between men and women with above 10 per cent absolute deviations. [a]Data refers to 2014 or the most recent year available; $: Gross National Income (GNI) per capita is based on 2011 dollar purchasing power parity (PPP).

Research and Development

In India, the R&D sector grew by 20.8 per cent in 2012-13 and contributed 1.4 per cent of GDP (Central Statistical Organisation's old method). As per the CSO's new method, there is no separate head for R&D. It is a part of the professional scientific & technical activities including R&D classification which grew at 3.8 per cent and 25.5 per cent respectively in 2013-14 and 2014-15. India's R&D expenditure has been low and the science, technology and innovation (STI) Policy 2013 envisages raising it to 2 per cent of GDP with enhanced participation of the private sector.

However, according to the Global Competitiveness Report 2015-16, India's capacity for innovation has been lower than that of many countries like the USA, the UK, South Korea, and even South Africa. Even in quality of scientific research institutions, India scores lower than China and South Africa. This is also exhibited through its poor score on university–industry collaboration in R&D as compared to some other BRICS (Brazil, Russia, India, China and South Africa) nations like China and South Africa. In terms of patents granted per million population, India fares badly compared to other BRICS countries. In terms of company spending on R&D also, India ranks below China. Only in terms of availability of scientists and engineers, does India score better or is equal to other BRICS countries.

Table 6.20: Global Competitiveness Index: R&D and Innovation

Country	Capacity for Innovation		Quality of scientific research institutions		Company spending on R&D		University-Industry collaboration on R&D		Availability of scientists and engineers		PCT patents granted/million population	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
USA	5.9	2	6.1	4	5.6	3	5.8	2	5.4	4	160.3	11
UK	5.4	10	6.3	2	4.9	17	5.7	4	4.9	18	89.9	18
South Korea	4.8	24	4.8	27	4.6	21	4.6	26	4.4	40	220.7	7
South Africa	4.6	32	4.7	33	3.8	32	4.5	31	3.4	106	6.9	46
China	4.2	49	4.2	42	4.2	23	4.4	32	4.5	36	13.4	32
Brazil	3.8	80	3.6	80	3.3	60	3.8	54	3.3	115	3.5	51
India	4.2	50	4.1	45	3.9	31	3.9	50	4.2	49	1.6	61
Russia	3.8	84	4	58	3.2	75	3.6	67	4.1	64	7.7	41

Global Links

The world economy is bound together by trade in goods and services, financial flows, and movements of people. As national economies develop, their links expand and grow more complex. The indicators in *Global links* measure the size and direction of these flows and document the effects of policy interventions, such as tariffs, trade facilitation, and aid flows, on the development of the world economy.

Equity flows

Equity flows comprise foreign direct investment (FDI) and portfolio equity. It has long been recognized that FDI flows can carry the benefits of knowledge and technology transfer to domestic firms and the labor force, productivity spillover, enhanced competition, and improved access for exports abroad. Moreover, they are the preferred source of capital for financing a current account deficit because FDI is non-debt creating. Investors continue to be attracted by improved business regulatory environment, growth prospects, and buoyant and expanding domestic markets. World Development Indicators (2015) noted that although many economies receive FDI, flows remain highly concentrated: Brazil, China and India account for more than half. In 2014, India's foreign direct investment inflow was $ 33,871 millions. In the same year, China's and USA's FDI inflow was around 9 times and 4 times that of India respectively.

Global portfolio equity flows rebounded substantially, with an overall annual increase of 41 percent at the end of 2014. Equity flows to high-income economies increased 42 percent, and equity flows to middle-income economies increased 27 percent. Investors sought after emerging markets being perceived as offering high returns, leading to some diversification in the destination of portfolio equity flows, but in general portfolio equity flows remained highly concentrated in only a handful of middle-income countries. China recorded a 58 percent increase in net portfolio equity flows, to $52 billion; India recorded a 40 percent decline, to $12 billion; and Brazil's net inflows remained unchanged, at $12 billion.

Merchandise Trade

In 2014, India's merchandise trade was 38.3 per cent of GDP. Indeed, it is higher than USA (23.2 per cent of GDP). Still, India's merchandise trade was lower than that of Germany (70.4 per cent of GDP) and China (41.5 per cent of GDP). In India major contribution in exports is not from merchandise but from services. The services

sector in India remained the most vibrant sector in terms of contribution to national and state incomes, trade flows, FDI inflows, and employment. India's services sector covers a wide variety of activities such as trade, hotel and restaurants, transport, storage and communication, financing, insurance, real estate, business services, community, social and personal services, and services associated with construction.

Table 6.21: Major Indicators for Global Links (2014)

Countries	Merchandise trade (% of GDP)	Net migration (thousands)	Foreign direct investment Net inflow $ millions	Portfolio equity Net inflow $ millions
USA	23.2	5008	131,829	155,077
UK	39.8	900	45,457	49,402
Japan	32.7	350	9,070	32,997
Germany	70.4	1250	8,390	3,764
India	38.3	–2,598	33,871	12,369
China	41.5	-1800	289,097	51,916

Source: World Development Indicators

GDP Shares in selected countries: Table 6.22 provides information regarding shares in GDP of agriculture, industry, manufacturing and services. The data reveal wide variation in the output pattern across different countries. USA, which is a developed country, shows that the share of agriculture in GDP is just 1 per cent in both the years i.e. 2000 and 2015. Further, during the same years, the share of services to GDP was 76 and 78 per cent respectively. If we closely look at the trend, it could be observed that in case of all the developed countries GDP is highly contributed by services, followed by industry and manufacturing and least contributing sector is agriculture. In China, for the year 2015, the share of services sector in GDP was 50 per cent. Further, for the same year the share of industry, manufacturing and agriculture in GDP was 41, 30 and 9 per cent respectively. India was unique that it skipped the industrial phase and shifted directly to services and that this transition is an economic achievement. This could be observed from the trend that, in India, between 2000 and 2015, the share of services sector in GDP has increased from 51 to 53 per cent. For the same period, agriculture contribution has decreased from 23 to 17 per cent. Though industry's contribution has increased from 26 to 30 percent, manufacturing sector's contribution to GDP was almost stagnant (around 16 per cent).

Table 6.22: Percentage Share of Sector-wise Contribution of GDP

Country	Gross domestic product $ billions		Agriculture % of GDP		Industry % of GDP		Manufacturing % of GDP		Services % of GDP	
	2000	2015	2000	2015	2000	2015	2000	2015	2000	2015
China	1,211.30	11,007.70	15	9	46	41	32	30	40	50
India	**476.6**	**2,095.40**	**23**	**17**	**26**	**30**	**15**	**16**	**51**	**53**
Germany	1,950.00	3,363.40	1	1	31	30	23	23	68	69
Japan	4,731.20	4,123.30	2	1	31	27	21	19	67	72
UK	1,635.40	2,858.00	1	1	25	19	15	10	74	80
USA	10,284.80	18,036.60	1	1	23	21	16	12	76	78
World	**33,391.00**	**73,891.90**	**5**	**4**	**31**	**28**	**19**	**15**	**64**	**68**

It can be concluded that Indian economy is on the right track of growth and development. It is one of the major economies in the world today which is also growing the fastest in the world.

KEY FEATURES OF BUDGET 2017-2018

INTRODUCTION

Budget 2017-18 contains 3 major reforms. First, presentation of Budget advanced to 1st February to enable the Ministries to operationalise all activities from the commencement of the financial year. Second, merger of Railways Budget with General Budget to bring Railways to the centre stage of Government's Fiscal Policy and Third, removal of plan and nonplan classification of expenditure to facilitate a holistic view of allocations for sectors and ministries

ROADMAP & PRIORITIES

- Agenda for 2017-18 is : "Transform, Energise and Clean India" – TEC India
- TEC India seeks to
  - Transform the quality of governance and quality of life of our people;
  - Energise various sections of society, especially the youth and the vulnerable, and enable them to unleash their true potential; and
  - Clean the country from the evils of corruption, black money and non-transparent political funding
- Ten distinct themes to foster this broad agenda:
  - **Farmers** : committed to double the income in 5 years;
  - **Rural Population** : providing employment & basic infrastructure;
  - **Youth** : energising them through education, skills and jobs;
  - **The Poor and the Underprivileged** : strengthening the systems of social security, health care and affordable housing;
  - **Infrastructure**: for efficiency, productivity and quality of life;
  - **Financial Sector** : growth & stability by stronger institutions;
  - **Digital Economy** : for speed, accountability and transparency;
  - **Public Service** : effective governance and efficient service delivery through people's participation;
  - **Prudent Fiscal Management**: to ensure optimal deployment of resources and preserve fiscal stability;
  - **Tax Administration**: honouring the honest.

FARMERS

- Target for agricultural credit in 2017-18 has been fixed at a record level of Rs. 10 lakh crores
- Farmers will also benefit from 60 days' interest waiver announced on 31 December 2016
- Coverage under Fasal Bima Yojana scheme will be increased from 30% of cropped area in 2016-17 to 40% in 2017-18 and 50% in 2018-19 for which a budget provision of Rs. 9000 crore has been made
- New mini labs in *Krishi Vigyan Kendras* (KVKs) and ensure 100% coverage of all 648 *KVKs* in the country for soil sample testing
- Dedicated Micro Irrigation Fund in NABARD to achieve 'per drop more crop' with an initial corpus of Rs. 5,000 crores
- Coverage of National Agricultural Market (e-NAM) to be expanded from 250 markets to 585 APMCs. Assistance up to Rs. 75 lakhs will be provided to every e-NAM

• A model law on contract farming to be prepared and circulated among the States for adoption

RURAL POPULATION

• Against target of 5 lakh farm ponds under MGNREGA, 10 lakh farm ponds would be completed by March 2017. During 2017-18, another 5 lakh farm ponds will be taken up

• Women participation in MGNREGA has increased to 55% from less than 48%

• MGNREGA allocation to be the highest ever at Rs. 48,000 crores in 2017-18.

• Allocation for Pradhan Mantri Awaas Yojana – Gramin increased from Rs. 15,000 crores in BE 2016-17 to Rs. 23,000 crores in 2017-18 with a target to complete 1 crore houses by 2019 for the houseless and those living in kutcha houses.

• For imparting new skills to people in rural areas, mason training will be provided to 5 lakh persons by 2022

• Total allocation for Rural, Agriculture and Allied sectors is Rs. 187223 crores

YOUTH

• To introduce a system of measuring annual learning outcomes in our schools

• Innovation Fund for Secondary Education proposed to encourage local innovation for ensuring universal access, gender parity and quality improvement to be introduced in 3479 educationally backward districts.

• SWAYAM platform, leveraging IT, to be launched with at least 350 online courses. This would enable students to virtually attend courses taught by the best faculty

• National Testing Agency to be set-up as an autonomous and self-sustained premier testing organisation to conduct all entrance examinations for higher education institutions

• *Pradhan Mantri Kaushal Kendras* to be extended to more than 600 districts across the country. 100 India International Skills Centres will be established across the country.

• Skill Acquisition and Knowledge Awareness for Livelihood Promotion programme (SANKALP) to be launched at a cost of Rs. 4000 crores. SANKALP will provide market relevant training to 3.5 crore youth

THE POOR AND THE UNDERPRIVILEGED

• *Mahila Shakti Kendra* will be set up with an allocation of Rs. 500 crores in 14 lakh ICDS *Anganwadi* Centres. This will provide one stop convergent support services for empowering rural women with opportunities for skill development, employment, digital literacy, health and nutrition

• National Housing Bank will refinance individual housing loans of about Rs. 20,000 crore in 2017-18

• To create additional 5,000 Post Graduate seats per annum to ensure adequate availability of specialist doctors to strengthen Secondary and Tertiary levels of health care

• Two new All India Institutes of Medical Sciences to be set up in Jharkhand and Gujarat

• To foster a conducive labour environment, legislative reforms will be undertaken to simplify, rationalise and amalgamate the existing labour laws into 4 Codes on (i) wages; (ii) industrial relations; (iii) social security and welfare; and (iv) safety and working conditions.

INFRASTRUCTURE

• For transportation sector as a whole, including rail, roads, shipping, provision of Rs. 2,41,387 crores has been made in 2017-18.

• For 2017-18, the total capital and development expenditure of Railways has been pegged at Rs. 1,31,000 crores. This includes Rs. 55,000 crores provided by the Government

• Railway lines of 3,500 kms will be commissioned in 2017-18. During 2017-18, at least 25 stations are expected to be awarded for station redevelopment.

- 500 stations will be made differently abled friendly by providing lifts and escalators.

- It is proposed to feed about 7,000 stations with solar power in the medium term

- A new Metro Rail Act will be enacted by rationalising the existing laws. This will facilitate greater private participation and investment in construction and operation.

- In the road sector, Budget allocation for highways increased from Rs. 57,976 crores in BE 2016-17 to Rs. 64,900 crores in 2017-18

- 2,000 kms of coastal connectivity roads have been identified for construction and development

- Select airports in Tier 2 cities will be taken up for operation and maintenance in the PPP mode

- A new and restructured Central scheme with a focus on export infrastructure, namely, Trade Infrastructure for Export Scheme (TIES) will be launched in 2017-18

FINANCIAL SECTOR

- Foreign Investment Promotion Board to be abolished in 2017-18 and further liberalisation of FDI policy is under consideration

- An expert committee will be constituted to study and promote creation of an operational and legal framework to integrate spot market and derivatives market in the agricultural sector, for commodities trading. e-NAM to be an integral part of the framework.

- Bill relating to curtail the menace of illicit deposit schemes will be introduced. A bill relating to resolution of financial firms will be introduced in the current Budget Session of Parliament. This will contribute to stability and resilience of our financial system

- A mechanism to streamline institutional arrangements for resolution of disputes in infrastructure related construction contracts, PPP and public utility contracts will be introduced as an amendment to the Arbitration and Conciliation Act 1996.

- Government will put in place a revised mechanism and procedure to ensure time bound listing of identified CPSEs on stock exchanges. The shares of Railway PSEs like IRCTC, IRFC and IRCON will be listed in stock exchanges.

- In line with the *'Indradhanush'* roadmap, Rs. 10,000 crores for recapitalisation of Banks provided in 2017-18

DIGITAL ECONOMY

- 125 lakh people have adopted the BHIM app so far. The Government will launch two new schemes to promote the usage of BHIM; these are, Referral Bonus Scheme for individuals and a Cashback Scheme for merchants

- Aadhar Pay, a merchant version of Aadhar Enabled Payment System, will be launched shortly

- A proposal to mandate all Government receipts through digital means,beyond a prescribed limit, is under consideration

- Proposed to create a Payments Regulatory Board in the Reserve Bank of India by replacing the existing Board for Regulation and Supervision of Payment and Settlement Systems

PUBLIC SERVICE

- To utilise the Head Post Offices as front offices for rendering passport services

- Web based interactive Pension Disbursement System for Defence Pensioners will be established

- To rationalise the number of tribunals and merge tribunals wherever appropriate

PRUDENT FISCAL MANAGEMENT

- Stepped up allocation for Capital expenditure by 25.4% over the previous year

- For the first time, a consolidated Outcome Budget, covering all Ministries and Departments, is being laid along with the other Budget documents

- FRBM Committee has recommended 3% fiscal deficit for the next three years, keeping in mind the sustainable debt target and need for public investment, fiscal deficit for 2017-18 is targeted at 3.2% of GDP and Government remains committed to achieve 3% in the following year

- Revenue Deficit of 2.3% in BE 2016-17 stands reduced to 2.1% in the Revised Estimates. The Revenue Deficit for next year is pegged at 1.9% , against 2% mandated by the FRBM Act

PERSONAL INCOME-TAX

- Existing rate of taxation for individual assesses between income of Rs. 2.5 lakhs to 5 lakhs reduced to 5% from the present rate of 10%

- Surcharge of 10% of tax payable on categories of individuals whose annual taxable income is between Rs.50 lakhs and Rs. 1 crore

- Simple one-page form to be filed as Income Tax Return for the category of individuals having taxable income upto Rs. 5 lakhs other than business income

RAPID (Revenue, Accountability, Probity, Information and Digitisation)

- Maximise efforts for e-assessment in the coming year

- Enforcing greater accountability of officers of Tax Department for specific act of commission and omission

LESSON ROUND UP

- A peculiar feature of the economic development in India is that the country,without going through industrialization, has jumped to the post-industrial stage.

- Inflation management is one of the hardest tasks an economic policy maker has to undertake.Inflation is a sustained rise in the general level of prices. Price stability is thus considered as an essentialcondition for economic stability and growth.

- Foreigncapital is typically seen as a way of filling gap between the domestically available supply of savings, plus foreignexchange plus, government revenue and the planned investment necessary to achieve developmental targets.After the announcement of New Industrial Policy (1991), there has been acceleration in the flow of foreigncapital in India.

- The major items in India's export basket are Petroleum products, Engineeringgoods and Gems & Jewellery. These three products together contribute 50 per cent of India's exports.

- The Planning Commission wasset up in 1950 and was entrusted with the responsibility of the creation, development and execution of India'sfive year plans.

- India's five year plans used to be supervised by erstwhile Planning Commission. However, in 2014, India hasannounced the dissolution of the Planning Commission, and its replacement by a think tank called the NITIAayog (National Institution for Transforming India).

- NITI Aayog serves as a "think tank" of the government as a "directional and policy dynamo" and would provideboth to the governments at the centre and in the states with strategic and technical advice on key policy mattersincluding economic issues of national and international importance.

- While the Planning Commission had the power to allocate funds to states for attaining regional development,the NITI Aayog does not have such powers. Rather, the task of allocating funds to states is now being vestedwith the Finance Ministry's Department of Expenditure. Its primary job would be to undertake long term policyand design frameworks and take necessary initiatives for attaining faster development and finally to monitorthese activities sincerely. Thus, NITI Aayog is responsible for actively monitoring and evaluating implementationof the Government programmes and initiatives.

- The year 1991 will remain a year of notable and widespread changes in the policy for economic

development inIndia. In 1991, the Government adopted a number of stabilizationmeasures that were designed to restore external and internal confidence. The government has made someradical changes in its policies regarding foreign investment, trade, exchange rate, industry, banking and fiscalaffairs, etc. These various elements taken together constitute an economic policy. Since July, 1991 Governmentof India has announced several new policies under the name of 'New Economic Reforms'. The new economicreforms gave a new direction and dimension to the Indian economy.

– GST / VAT is a consumption based tax wherein the basic principle is to tax the value addition at each businessstage. To achieve this, tax paid on purchases is allowed as a set off/ credit against liability on output/income.

– Demonetization is the act of stripping a currency unit of its status as legal tender.

GLOSSARY	
NITI Aayog:	It is the successor in interest to the Planning Commission, superseding the erstwhile Planning Commission with effect from 1st January, 2015.
Planning Commission:	A commission delegated to propose plans for future activities and developments of the country for its effective growth.
Green Revolution:	The term Green Revolution refers to the renovation of agricultural practices beginning in Mexico in the 1940s. Because of its success in producing more agricultural products there, Green Revolution technologies spread worldwide in the 1950s and 1960s, significantly increasing the produced per acre of agriculture
HYV:	High- yielding varieties (HYVs) are any of a group of genetically enhanced crops such as rice, maize and wheat that have an increased growth rate, an increased percentage of usable plant parts or an increased resistance against crop diseases. High Yielding Variety Seeds are developed to get higher yielding. The main advantages of High Yielding Variety seeds are yielding of farms with the use of HYV seeds in high compared to the use of traditional seeds. These crops formed the basis for the Green Revolution.
Public Investment:	It is another name for Government investment.
GDP:	Gross Domestic Product that is the market value of all final goods and services produced within a country in a given period.
New Industrial Policy:	In the light of worldwide recession, a policy was announced in July 1991, which; besides liberalization of economy and globalisation, also aimed at building upon the gains achieved, to correct the distortions, maintain a sustained growth in productivity and gainful employment and attain international competitiveness.
Globalization:	A generalized historical process through which more economic activity takes place across national borders. Forms of globalization include international trade (exports and imports), foreign direct investment, international financial flows, and international migration.
GDP Per Capita:	The level of GDP divided by the population of a country or region.
Inflation:	A process whereby the average price level in an economy increases over time.
Consumer Price Index:	The consumer price index (CPI) is a measure of the overall price level paid by consumers for the various goods and services they purchase. Monthly or annual changes in the CPI provide a good measure of the rate of consumer price inflation.
Development:	Economic development is the process through which a country's economy expands and improves in both quantitative and qualitative terms. Economic development requires the coming together of several different processes and conditions: the accumulation of real capital; the development of education, skills, and human capacities; improvements in governance, democracy, and stability; and changes in the sectoral make-up of the economy.
Employment:	Employment is a specific form of work, in which the worker performs their labour for

someone else in return for a money wage or salary.

Exports: An export is the sale of a product from one country (either a good or a service) to a purchaser in another country.

Imports: Goods or services which are produced in a foreign country and purchased domestically. Imports include money spent on vacations or purchases in foreign countries.

Industrial Policy: Government policies aimed at fostering the domestic development of particular desirable or productive industries, in order to boost productivity, create higher-paid jobs, and enhance international trade performance. Tools of industrial policy can include measures to stimulate investment in targeted industries; trade policies (such as tariffs, export incentives, or limits on imports); and technology policies.

Unemployment: Individuals who would like to be employed, and are actively seeking work, but cannot find a job, are considered "officially" unemployed.

MULTIPLE CHOICE QUESTIONS

1. **NITI Aayog stands for**
   a. National Institution for Transforming India
   b. National Institution for Technology and Innovation
   c. Newspaper Integration and Teaching Institution
   d. None of the above

2. **NITI Aayog is a successor of:**
   a. Planning Commission
   b. Pay Commission
   c. Productivity Council
   d. Finance Commission

3. **New Industrial Policy was announced in**
   a. March 1956
   b. April 1990
   c. July 1991
   d. January 2001

4. **Following is the last Five Year Plan in India**
   a. 10th Five Year Plan
   b. 12th Five Year Plan
   c. 13th Five Year Plan
   d. 11th Five Year Plan

5. **The largest employment generating sector in India is**
   a. Primary Secondary
   b. Secondary Sector
   c. Manufacturing Sector
   d. Service Sector

6. **Recently, Indian experienced demonetisation on**
   a. 7th November, 2016
   b. 8th November, 2016
   c. 9th November, 2016

d. 8th December, 2016

7. **Poverty is measured in terms of:**

    a. Calorie

    b. Education

    c. Wealth

    d. None of the above

8. **Parallel economy is also referred to as**

    a. Grey economy

    b. Black Economy

    c. Black market

    d. None of the above

9. **In year 2014, HDI rank of India among BRICS countries was**

    a. Lowest

    b. Highest

    c. In the Middle

    d. None of the above

10. **India's per capital income at market prices in 2014 was US$**

    a. 15060

    b. 1560

    c. 1005

    d. 2000

Answer key : 1. (a), 2. (a), 3. (c), 4. (b), 5. (a), 6. (b), 7. (a), 8. (b), 9. (a), 10. (b)

Selected Readings

1. Datt & Sundharam: Indian Economy

2. P.K. Dhar: Indian Economy

3. Ramesh Singh: Indian Economy

4. Economic Survey, 2015-16 and 2016-17

Lesson 7
Basic Elements of Money and Banking

LESSON OUTLINE

LEARNING OBJECTIVES

With the advent of banking sector and emerging financial markets in the economy, one requires a precise understanding concerning money; what exactly would term money imply; any object (not necessarily rupee) that circulates widely as a means of payment or does it involve something more to it. Besides this, what is the connection between money, banks, and credit system?

The purpose of this study is to get the students acquainted with the concept of money and thereby relating it to different approaches which helps in defining the functions of money.

The chapter further deals with the process of credit creation and how different theories of money works.

Since the entire financial structure revolves around banks and financial institutions it becomes significant to deliberate upon the role and functions of central and commercial banks.

Further, economy of any nation depends or say is run by effective monetary policy. At this stage, it becomes essential to acquaint students with the tools of monetary process and policy in India.

CONCEPT AND NATURE OF MONEY

Economists and thinkers have found that it is not easy to define money though easily recognisable. This is because, over recorded history, a large variety of items have performed the role of money. For this reason, it cannot be defined by referring to its tangible features, such as, the stuff it is made of, its weight, size, shape, colour, chemical composition, and so on. The exact form of money has undergone innumerable changes. These changes have been influenced by the dynamism of the society as a whole which includes its political, social, and economic institutions and the policies of the government.

The evolution of money has proved to be an unending and continuous process which can be seen from the fact that apart from the commodity money and metallic money, lately there is an emergence of paper money and a variety of other financial instruments.

It is noteworthy that money is not something which, having come into existence continues to be in its original form. Various things have served as money at different times and places. These have varied from shells, goats, cows, and rice to silver and gold pieces and to coins, paper currency, notes and to demand deposits of bank.

Money was brought into existence to tide over the difficulties of barter, but in the process it has helped economy in such a manner that its use has become indispensable. It has responded to the ever changing nature of the economy and its growing complexity. In turn, it has helped the economy in acquiring those complex features without which the latter could not develop.

Money and Near-money

Money consists of (a) legal tender money[1] i.e., coins and currency notes and (b) bank money (i.e., demand deposits). Money is the perfect liquid asset and can be directly used for making purchases of goods and services. However, the coins and currency are used only for small transactions. In case of large transactions where large payments are involved (for example, purchasing a land or a machine), bank money in the form of bank cheques or bank drafts is used. In modern economies, people mostly use bank money in their transactions and thus bank money forms a major proportion of money supply. In this way, bank money is considered as liquid as the legal tender money.

There are other assets also which cannot be technically regarded as money, but are claims to money and perform some functions of money. Such assets are called near-money. Near-money refers to all those assets which possess many of the characteristics of money, have high degree of liquidity. However, near-money cannot be directly used for making transactions. They must first be converted into money proper before spending. Hence, near-money assets are highly liquid, but are not as liquid as the money is. They are close substitutes of money, but not the perfect substitutes. Some examples of near-money are bills of exchange, bonds, debentures, shares, etc.

Definition of Money

Crowther defines money as "anything that is generally acceptable as a means of exchange and at the same time acts as a measure and store of value".

Approaches to the Definition of Money

On the basis of constituents of money, there are following approaches:

– Traditional Approach

---

[1] Legal tender money is enforced by law. That means, no one can refuse to accept it as a means of payment.

  – Monetarist Approach

  – Liquidity or Generalized Purchasing Power Approach

  – Central Bank Approach

1. Traditional Approach: In the traditional approach, money is regarded only as a medium of exchange. The definition of money emphasises its characteristics like spendability, liquidity, etc. According to this approach money (M) includes currency (C) and demand deposits (DD).

$$M = C + DD$$

All other assets can be considered as money if they are first converted into currency or demand deposits.

2. Monetarist Approach: This approach is associated with Milton Friedman and other Quantity Theorists or monetarists. Friedman defined money as "a temporary abode of purchasing power". According to this view, money can act as a temporary abode of purchasing power, if it is kept in the form of cash, demand deposits or any other asset which is close to currency, i.e., near-money asset.

Thus, money (M), according to the monetarist approach, includes currency (C), demand deposits (DD) and time deposits (TD):

$$M = C + DD + TD$$

Time deposits have been included in money on the basis of two monetarist assumptions: (a) Money has been regarded as having the highest correlation with income, (b) Money includes all those assets which are perfect substitutes and time deposits are very close substitutes for currency and deposits.

3. Liquidity Approach:

Gurley and Radcliffe Committee has given liquidity approach to the definition of money. In this approach, the scope of the constituents of money has been further widened to include in money the monetarist definition plus the liabilities of non-banking intermediaries.

Thus, money (M) includes currency, (C), demand deposits (DD), time deposits (TD), saving bank deposits (SB), shares (S), bonds (B) etc.

$$M = C + DD + TD + SB + S + B$$

This approach adopts substitutability criterion to define money as a weighted sum of currency, demand deposits and their substitutes.

Currency and demand deposits have been given unit weight; the assets which are completely unrelated to currency and demand deposits have been given zero weight; the assets which are imperfect substitutes of currency and demand deposits have been given weights between zero and one.

4. The Central Bank Approach:

The Central Bank approach takes a step further in extending the scope of the constituents of money by regarding money as the total amount of credit extended by a wide variety of sources.

Thus, according to this approach, money (M) includes currency (C), demand deposits (DD), time deposits (TD), credit from non-bank financial institutions (NBFI) and credit from unorganised agencies (CUA):

$$M = C + DD + TD + NBFI + CUA$$

FUNCTIONS OF MONEY

The functions of money have been well summed up in a couplet:

<div style="text-align:center">

"Money is a matter of functions four,

A Medium, a Measure, a Standard and a Store."

</div>

In recent days, the economists have added some new functions to these four functions.

The functions of money can be classified under three main heads:

- Primary Functions or Original Function
- Secondary Functions
- Contingent Functions

I. Primary Functions

The main functions performed by money are called primary or original function. The primary functions of money are as follows:

1. **Medium of Exchange:** The most important function of money is to serve as a means of payment. To be a successful medium of exchange, money must be commonly accepted by people in exchange for goods and services. While functioning as a medium of exchange, money benefits the society in a number of ways: (a) It overcomes the inconvenience of barter system (i.e., the need for double coincidence of wants) by splitting the act of barter into two acts of exchange, i.e., sales and purchases through money. (b) It promotes transactional efficiency in exchange by facilitating the multiple exchanges of goods and services with minimum effort and time. (c) It promotes allocative efficiency by facilitating specialisation in production and trade (d) It allows freedom of choice in the sense that a person can use his money to buy the things he wants most, from the people who offer the best bargain and at a time he considers the most advantageous.

2. **Measure of Value:** Money serves as a common measure of value in terms of which the value of all goods and services is measured and expressed. By acting as a common denominator or numeraire, money has provided a language of economic communication. It has made transactions easy and simplified the problem of measuring and comparing the prices of goods and services in the market. Prices are but values expressed in terms of money.

   Money also acts as a unit of account. In India, the unit of account is the 'Rupee', in USA the 'Dollar', in Japan the 'Yen' etc. As a unit of account, it helps in developing an efficient accounting system because the values of a variety of goods and services which are physically measured in different units (e.g. quintals, metres, litres, etc.) can be added up. This makes possible the comparisons of various kinds, both over time and across regions. It provides a basis for keeping accounts, estimating national income, cost of a project, sale proceeds, profit and loss of a firm, etc.

   To be satisfactory measure of value, the monetary units must be invariable. In other words, it must maintain a stable value. A fluctuating monetary unit creates a number of socio-economic problems. Normally, the value of money, i.e., its purchasing power, does not remain constant; it rises during periods of falling prices and falls during periods of rising prices.

II. Secondary Functions

The relatively less important functions of money are called secondary functions. Since, these functions originates from primary functions, these are also called derived functions. The secondary functions of money are as follows: -

3. **Standard of Deferred Payment:** Money serves as the standard of deferred payment or units in which future or deferred payments are made. This function applies to interests, rents, salaries, pensions, insurance

premium etc. The lending and borrowing acts are easily expressed in money. Due to the qualities of stability in value, general acceptability and durability, money is regarded best for these transactions.

Thus, money not only helps current transactions by functioning as a medium of exchange, but facilitates credit transaction (i.e., exchanging present goods on credit) through its function of standard of deferred payments. But, to become a satisfactory standard of deferred payments, money must maintain a stable value over time.

4. **Store of Value:** The fourth function of money is to serve as a store of value because it is easy to spend and easy to store. Money serves as a store of value in the short run as well as long run. By performing the function of store of value, the money provides security to individuals to meet unpredictable expenditures and to pay debt that are in terms of money. In barter exchange system, all commodities could not be stored for a longer period of time. Money has a unique nature of durability and stability in value; thus it is can be stored for a long time. This has introduced in people, a trend of making savings from the incomes for future purposes.

   This function of money i.e. a store of value is necessary but not sufficient condition to term anything as money. Although, money functions as a store of value, but all things functioning as store of value cannot be termed as money. For instance, things like diamond, jewellery work as a store of value. But, these do not serve the primary functions of money, hence are not termed as money; therefore not used as money.

5. **Transfer of Value:** Money has general acceptability as a means of exchange. So it is easier to transfer value from one place to another. At present, money is stored in the form of bank deposits. Depositor can transfer the amount of money deposited in his bank account to the account of another person. It means it is easier to transfer value in the form of money. Money is a means through which transfer of value from one place to another has become easier and quicker. So transfer of value in the form of money through space continues to be important. For example, a businessman of Orissa who sells his property and goes to Delhi and settles down there is a case of transfer of value through space. Further, money serves as the function of transfer of value or purchasing power. People transfer value by selling commodities or property to others and by buying commodities and property from others. Money has facilitated the transaction of goods in distant places. For example, it is much easier to transfer one lakh rupees through bank draft from person A in Amritsar to person B in Mumbai than remitting the same value in commodity terms, say wheat.

III. Contingent Functions

Prof. Kinley has mentioned four contingent functions of money. These functions are as follows:

6. **Basis of Credit:** In present days, the use of credit money like Cheque, Draft, bill of exchange, Promissory Notes is expanding widely. The credit instruments are issued on the basis of cash reserve. The credit instrument like Cheque is issued on the basis of money deposits. Hence, money is a basis of credit.

7. **Distribution of National Income:** Money facilitates the distribution of national income between people. Total output of the country is jointly produced by a number of people as workers, land owners, capitalists, and entrepreneurs, and, in turn, will have to be distributed among them. Money helps in the distribution of national product through the system of wage, rent, interest and profit. The distribution of production among factors would be difficult in the age of modern specialized labour in the absence of money.

8. **General Form of Capital:** Money works as a general form of capital. This increases the liquidity and mobility of capital. In fact, all forms of wealth (e.g., land, machinery, stocks, stores, etc.) can be converted into money.

9. **Maximum Benefit:** People derive maximum satisfaction from own income by the help of money. According to law of equi-marginal utility, people derive maximum satisfaction when they spend by making marginal utility equal in all goods. People spend money to make marginal utilities of all commodities equal and

derive maximum satisfaction by the help of money. Likewise, the producers also spend money in different factors so as to make marginal productivity of all factors equal. This increases total output and yields maximum benefit to the producers.

FIG. 7.1: FUNCTIONS OF MONEY
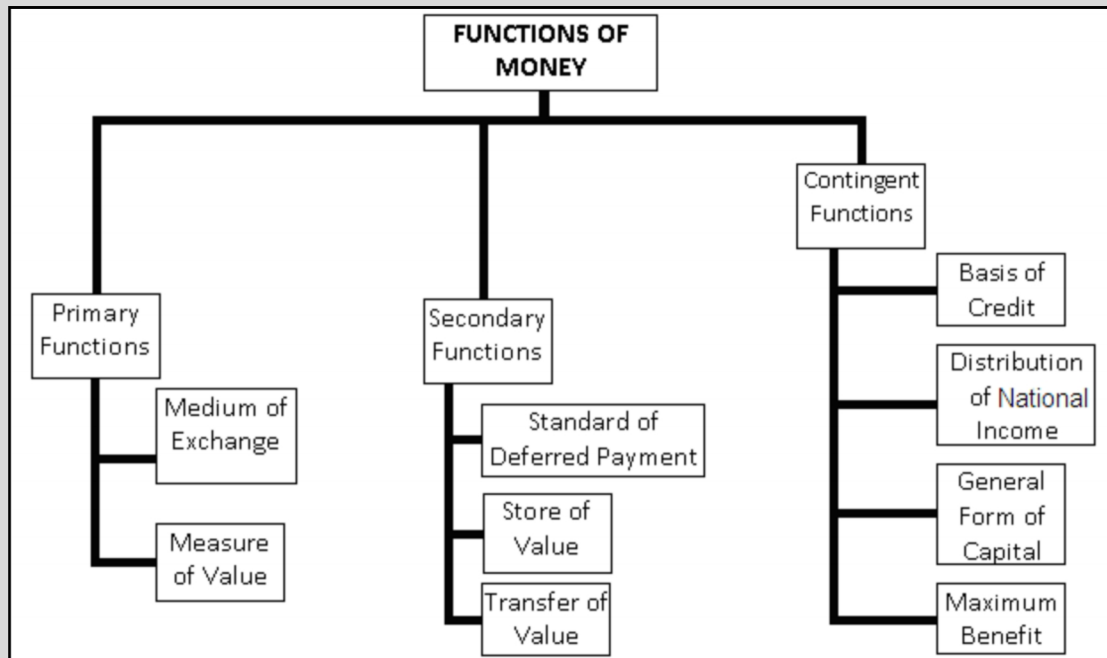
QUALITIES OF GOOD MONEY

To be able to perform the above mentioned functions of money, it must possess the following qualities:

1. **General Acceptability:** Money should be acceptable to all without any hesitation. As mentioned above, since it is declared by the law, money as the legal tender has the quality of general acceptability.

2. **Portability:** Money should be easy to carry or transfer from one place to another.

3. **Durability:** Money as a material must last for a long time without losing its value. For instance, ice and fruits cannot become good money because they lose their value quickly with the passage of time as ice melts and fruits perish.

4. **Divisibility:** Money material must be easily divisible for the purchase of smaller units of the commodities. Cows, for example, cannot function as good money because a cow cannot be divided without losing its value.

5. **Homogeneity:** Money should be homogeneous. Its units should be identical; they should be of equal quality and physically indistinguishable. If money is not homogeneous, the individuals will not be certain of what they are receiving when they make transactions.

6. **Cognisability:** Money should be easy to recognise. If it is not easily recognisable, it would be difficult for the individuals to determine whether they are dealing with money or some inferior asset.

7. **Stability:** The value of money should remain stable and should not change for a long period of time. If the value of money is not stable, it will not be able to function as a measure of value, as a store of value and as a standard of deferred payment.

NEW MONETARY AND LIQUIDITY AGGREGATES

Money can, therefore, be defined for policy purposes as the set of liquid financial assets, the variation in the stock of which could have impact on aggregate economic activity. As a statistical concept, money could include certain liquid liabilities of a particular set of financial intermediaries or other issuers. Thus, like other countries, a range of monetary and liquidity measures are compiled in India. Various monetary and liquidity aggregates compiled in India and their definitions are set out in Table 7.1.

Table 7.1: Monetary Aggregates in India

Monetary Aggregates	New Monetary Aggregates	Liquidity Aggregates
M_1 = Currency with the public +Demand deposits with the banking system + 'Other' deposits with RBI	NM_1 = Currency with the public +Demand deposits with the banking system + 'Other' deposits with RBI (unchanged)	L_1 = NM_3 + All deposits with the post office savings banks (excluding National Savings Certificates)
M_2 = M_1 + savings deposits of post office savings banks	NM_2 = NM_1 + time liability portion of savings deposits with banks + Certificates of Deposits (CDs) issued by banks + Term deposits maturing within one year	L_2 = L_1 + Term deposits with term lending institutions + Term borrowing of Financial Institutions (FIs) + CDs issued by FIs
M_3 = M_1 + Time deposits of banks	NM_3 = NM_2 + term deposits over one year maturity + call/term borrowings of banks	L_3 = L_2 + Public deposits of non-banking financial companies (NBFCs)
M_4 = M_3 + All deposits with post office savings banks (excluding National Savings Certificates)	NM_4 = Abolished	

CREDIT CREATION

Meaning

A bank differs from other financial institutions because it can create credit. Banks have the ability to expand their demand deposits as a multiple of their cash reserves. This is because of the fact that demand deposits of the banks serve as the principal medium of exchange, and, in this way, the banks manage the payments system of the country. In short, expansion of deposits is called credit creation and the ability of the banks to expand deposits makes them unique and distinguish them from other non-bank financial institutions. Demand deposits are an important constituent of money supply and the expansion of demand deposits means expansion of money supply.

The whole structure of banking is based on credit. Credit means getting the purchasing power (i.e., money) now by a promise to pay at some time in future. In other words of Kent, "Credit may be defined as the right to receive payment of the obligation to make payment on demand or at some future time on account of an immediate transfer of goods." In a sense, the words credit, debt and loan are synonymous; credit or loan is the liability of the debtor and the asset of the bank. The creditor believes that the debtor will return the loan and so decides to give the loan. Advancing credit or loan essentially depends upon the (a) confidence, (b) character, (c) capacity, (d) capital, and (e) collateral of the debtor.

Bank credit means bank loans and advances. A bank keeps a certain proportion of its deposits as minimum reserve for meeting the demands of the depositors and lends out the remaining excess reserve to earn income. The bank loan is not paid directly to the borrower but is only credited in his account. Every bank loan creates an equivalent deposit in the bank. Thus, credit creation means expansion of bank deposits. The word 'creation' refers to the ability of the bank to expand deposits as a multiple of its reserves.

During the process of credit creation, there arise two important issues which hold utmost consideration, namely profitability and liquidity.

 – **Profitability**: Banks are guided by the profit earning motives in its operations. So the loans granted must be in a manner which earns an interest higher than what banks pay on its deposits.

 – **Liquidity**: Banks on the other hand, subsequently should be able to meet its commitment to pay cash to its depositors as and when the depositors decide to exercise their right to demand cash against their deposits.

The bank credit creating process is based on the assumption that during any given time interval, only a fraction of its customers will be in genuine need of cash and that all customers would not turn up for demanding cash against their deposits at one point of time. At any given time, only a fraction of total deposits would be used to pay cash. Therefore, banks can lend the remaining proportion of money into market thereby, creating credit.

The net result is that a bank is able to meet the demand for encashment by its depositors by maintaining cash reserves which are only a fraction of its deposit liabilities. This phenomenon is known as that of "fractional cash reserves", or "cash deposit ratio" being less than one.

The experience of the 1920s, 1990s and 2000s keep renewing economists' interest in the role of credit in macroeconomic fluctuations as they are of the view that not just money but credit also matters for macroeconomic and financial stability. The Great Depression of the 1920s is also viewed as a credit boom gone wrong. Further, sub-prime lending crisis occurring between 2007 and 2010 also reminded of importance of proper regulation of lending and credit agencies. It is generally argued that credit creation may fuel bubbles, make the system unstable and lead economies into crises. The world has observed various Minsky moments and all these form interesting case studies for understanding the financial crises and role of credit. Therefore, central banks all over the world not only set monetary policy to manage or control inflation but they also take care of credit market conditions for the stability of the economic systems.

Basic Concepts

In order to understand the process of credit creation, the proper knowledge of some basic concepts is necessary:

1. **Bank as a Business Institution:** Bank is a business institution which aims at maximising profits through loans and advances from the deposits.

2. **Bank Deposits:** Bank deposits form the basis for credit creation. Prof. Halm has classified deposits into two types: (a) primary deposits; and (b) secondary or derivative deposits.

   (a) **Primary deposits:** When a bank accepts cash from the customer and opens a deposit account in his name it is called primary or passive account. The creation of the primary deposit does not mean creation of credit (or money). These deposits simply convert currency money into deposit money. However, the primary deposits form the basis for the creation of credit (or money). It is out of these deposits that the bank grants loans and advances.

   (b) **Secondary or Derivative Deposits:** When a bank grants loans and advances, it, instead of giving cash to the borrower, opens deposit account in his name. This is secondary or derivative deposit. Every loan creates a deposit. It is called derivative deposit because it has been derived from the loan transaction of the bank. Creation of derivative deposits means creation of credit (or money). With the creation of these deposits, money supply in the form of bank deposits is increased.

3. **Cash-Reserve Ratio (CRR):** From the general experience, banks know that all depositors will not withdraw all deposits at the same time. Therefore, they keep a fraction of the total deposits for meeting the cash demand of the depositors and lend out the remaining excess deposits. The percentage of total deposits which the banks are required to hold in cash reserves for meeting the depositors' demand for cash is called CRR.

4. **Excess Reserves:** The reserves that a bank holds above the required cash reserves are called excess reserves. It is the excess reserves out of which the loans are granted and thereby the credit is created.

5. **Credit multiplier:** The banks can create multiple times credit from a given amount of cash. In the process of multiple credit creation, the total amount of derivative deposits created by the banks will be a multiple

of the initial excess reserves.

$$Credit\ Multiplier\ (K) = \frac{Total\ derivative\ deposits\ (\Delta D)}{Initial\ excess\ reserves\ (\Delta R)}$$

$$(Or)$$

$$\frac{1}{Cash\ reserve\ ratio}$$

For instance, if the initial excess reserves of Rs. 1000 produce total derivative deposits of Rs. 5,000 then the credit multiplier is 5.

Credit Creation by Single Bank: The process of credit creation can be analysed in two ways: (a) credit creation by a single bank; and (b) credit creation by the banking system as a whole. In the single bank system, only one bank operates and all the cash deposits and cheques are to be made with this bank alone. The process of credit creation by a single bank can be illustrated with the help of a hypothetical example given in Table 7.2.

Table 7.2: Credit Creation by Single Bank

Rounds	Primary Deposits	Cash Reserves (r = 20%)	Credit Creation or Derivative Deposits (ΔD)
1. (Person A)	Rs. 1000 (Initial primary deposits)	Rs. 200	Rs. 800 (Initial excess reserves ΔR)
2. (Person B)	800	160	640
3. (Person C)	640	128	512
4. (Person D)	512	102	410
-	-	-	-
-	-	-	-
Total	5000	1000	4000

Suppose the customary CRR maintained by the bank is 20%. Now if person A deposits Rs. 1000 with the bank, the bank does not keep the entire cash in reserve but only the 20% of it to meet the day-to-day cash demand. Thus, after keeping Rs. 200 (i.e., 20% of Rs. 1000), the bank lends the remaining Rs. 800 to person B by opening a credit account in his name. Again, keeping 20% of Rs. 800 (i.e., Rs. 160), the bank advances the remaining Rs. 640 (i.e., Rs. 800 – 160) to person C. Similarly, keeping 20% of Rs. 640 (i.e., Rs. 128), the bank advances the remaining Rs. 512 (i.e., Rs. 640 – 128) to person D and so on. This process will continue till the initial primary deposit of Rs. 1000 and the initial excess reserves of Rs. 800 lead to additional (derivative) deposits of Rs. Rs. 800 + 640 + 512 + ... = Rs. 4000. By adding up all the deposits (i.e., primary plus derivative), we get total deposits of Rs. 5000. Here credit multiplier (which is the reciprocal of the CRR, i.e., 20%) is 5 and the credit creation (or the total derivative deposits, i.e., Rs. 4000) is five times the initial excess reserves (i.e., Rs.800).

Multiple Credit Creation by Banking System: In the real world, there are many banks in existence in the banking system. The banking system can grant loans many times the excess reserves of cash created by it. When an individual bank creates derivative deposits, it loses cash to other banks; the loss of deposit of one bank is the gain of deposit by some other bank. This transfer of cash within the banking system creates, in turn, primary deposits and increases the possibility for a further creation of derivative deposits by the banks receiving cash. The process of the banking system to increase credit many times more than the initial excess reserves is called multiple credit creation.

Table 7.3: Multiple Credit Creation by Banking System

Banks	Primary Deposits	Cash Reserves (r = 20%)	Credit Creation or Derivative Deposits (ΔD)
A	Rs. 1000 (Initial primary deposits)	Rs. 200	Rs. 800 (Initial excess reserves ΔR)
B	800	160	640
C	640	128	512
D	512	102	410
-	-	-	-
-	-	-	-
Total	5000	1000	4000

The process of multiple credit creation can be explained with the help of an example given in Table 7.3.

Suppose, with the initial primary deposit of Rs. 1000 and the CRR of 20%, bank A has initial excess reserves of Rs. 800 (i.e., Rs. 1000 minus 20% of Rs. 1000 = Rs.800). The bank creates derivative deposits equal to its initial excess reserves of Rs.800 by granting loans to the borrowers. The borrowers make payments of Rs. 800 by cheques to other people who are the customers of bank B. The CRR being 20%, the excess reserves of bank B are Rs. 640 (i.e., Rs. 800 minus 20% of Rs. 800 = 640) which it converts into derivative deposits by giving loans to borrowers. Further, the borrowers from bank B make payment of Rs. 640 by cheques to some people who are the customers of Bank C. This creates the primary deposits of Rs.640 in bank C, which, in turn, leads to the creation of excess reserves and derivative deposits of Rs. 512 (i.e., Rs. 640 minus 20% of Rs.640 = Rs.512) in bank C.

This process will continue until the initial primary deposit of Rs. 1000 with bank A lead to the creation of total deposits (primary plus derivative) of Rs. 5000 and the initial reserve of Rs. 800 in bank A leads to the multiple expansion of total derivative deposits of Rs. 4000 in the entire banking system. Thus, credit creation or the creation of total derivative deposits (ΔD) by the banking system will be Rs. 800 + 640 + 512 + ... = Rs. 4000, i.e., 5 times of the initial excess reserves (ΔR) of Rs.800.

$$Credit\ multiplier = \frac{Total\ derivative\ deposits\ (\Delta D)}{Initial\ excess\ reserves\ (\Delta R)} = \frac{4000}{800} = 5$$

or

$$= \frac{1}{Cash\ reserve\ ratio} = \frac{1}{0.2} = 5$$

Limitations of Credit Creation

Though banks would prefer to hold an unlimited capacity of creating credit with an objective of profitability, in practice banks are faced with numerous limitations in doing so. These limitations result in making the credit creation process by banks as a non-profitable activity. Thus, a bank would keep creating additional credit so long as:

1. loans extended by the banks to its customers have a negligible chance of turning into bad debts

2. the difference between the interest rates that the banks charge on loans and advances lent is greater than that the banks give to depositors on the money deposited with the bank

In other words, the limitations of credit creation by banks operate through shifts in their balance between liquidity and profitability.

The limitations on credit creation by banks may be considered in respect to the following factors:

– the capacity of banks to create credit,

– the willingness of banks to create credit

– demand for credit in the market

Capacity to create credit is a matter of

– the availability of cash deposits with the bank

– the factors that determine their cash deposit ratio

It may be recalled that a bank is willing to create credit so long as

– it estimates that it will be profitable to do so

– it is able to maintain the liquidity requirements

As regards the demand for credit

– demand for credit must exist in market

– borrowers should be creditworthy (to avoid bad debts against their loans)

– amount of loan granted should not exceed the paying capacity of the borrower

Leakages

– banks may not willing to utilise their surplus funds for granting loans, probably the economy is heading towards a recession.

– if public withdraw some cash to hold it with themselves, that will reduce the power of the banks to create credit.

In many situations, due to a policy pursued by the central bank, it is difficult to get fresh loans from the banks and we get the impression that demand for credit is unlimited. But factually it is not so. A persistent excess demand for credit exists only under inflationary conditions. But these conditions cannot last forever because continuous inflationary pressure can disrupt the financial system itself. In other words, the concept of a limitless demand for bank credit is only an illusion and not a reality.

QUANTITY THEORY OF MONEY

Economists have always been interested in answering a two-fold question, namely, what determines the level of prices in general and what causes a variation in it. The Quantity Theory of Money seeks to explain the factors that determine the general price level in an economy. The theory states that the price level is directly determined by the supply of money. There are two versions of the Quantity Theory of Money:

– Transaction Approach

– Cash Balance Approach

Transaction Approach

The transaction approach to the Quantity Theory of Money is given by Fisher which may be explained with the help of following equation of exchange.

MV= PT

where,   M is the total supply of money

V is the velocity of circulation of money

P is the general price level

T is the total transactions in physical goods

This equation means, in an economy the total value of all goods sold during any period (PT) must be equal to the total quantity of money spent during that period (MV).

Fisher Version/Transaction Approach

Assumptions:

(i) Price level is to be measured over a period of time

(ii) There are no credit sales in the market

(iii) Money is only a medium of exchange

(iv) Each unit of money can change hands several times during the said time interval. Thus, total cash payment= MV, where: M= Money supply and V= Velocity.

(v) All cash payments received during the year must be equal to the volume of goods and services sold multiplied their respective prices. Thus, total cash payment=PT, where: P=Price and T=Time period.

When these two assumptions are made the Equation of Exchange becomes the Quantity Theory of Money which shows that there is an exact, proportional relationship between money supply and the price level. In other words, the level of prices in the economy is directly proportional to the quantity of money in circulation. That is, doubling the total supply of money would double the price level.

This version proceeds with the idea that price level is determined by the demand for and supply of money.

Fisher also extended the equation of exchange so as to include demand (bank) deposits (M') and their Velocity (V') in the total supply of money. Thus, the equation of exchange becomes:

$$MV + M'V' = PT$$

Or

$$P = \frac{MV + M'V'}{T}$$

M = Currency

M' = Bank's money

V & V' = Respective Velocities

Thus, the level of general price level depends on all the five variables of equation.

In the words of Irving Fisher, "Other things remaining unchanged, as the quantity of money in circulation increases the price level also increases in direct proportion and the value of money decreases and vice versa".

If the quantity of money is doubled the price level will also double and the value of money will be one half. On the other hand if the quantity of money is reduced by one half the price level will also be reduced by one half and value of money will be twice.

Fisher's Transaction Approach can explain the causes of hyperinflation that occurs during war or emergency. It can also explain certain long term trend in prices. But it cannot explain normal peace time inflation. This shortcoming has been modified by the Cambridge version or the Cash-Balance Approach.

Cash Balance Approach

The Cash-Balance Approach to the Quantity Theory of Money may be expressed as:

$$\pi = kR/M$$

where, π = the purchasing power of money

k = the proportion of income that people like to hold in the form of money

R = the volume of real income

M = the stock of supply of money in the country at a given time

This equation shows that the purchasing power of money or the value of money (π) varies directly with k or R, and inversely with M.

Since π is the reciprocal of the general price level i.e.

$$\pi = 1/P$$

$$\Rightarrow \frac{1}{P} = \frac{kR}{M}$$

$$\Rightarrow M = kRP$$

If we multiply the volume of real income (R) by the general price level (P), we have the money national income(Y).

$$M = kY$$

where, Y is the country's total money income

In the Cash Balance approach k was more significant than M for explaining changes in the purchasing power (or value) of money. This means that the value of money depends upon the demand of the people to hold money.

CENTRAL BANKS

A central bank is an apex institution of a country's monetary and financial system. Since the monetary system (which includes commercial banks) is a dominant part of the financial system of a country; the central bank is the apex system to the country's financial system also. As such it plays a leading role in organizing, running, supervising, regulating and developing the monetary- financial system. The best way to define a central bank is to say that it is the apex financial institution of the country and subject to certain legal boundaries is vested with the authority to regulate, guide and help the financial system. It is not guided by profit motive and gives precedence to the economic interests of the country over its gains. Till the end of 19th century, central banking came into existence out of slow evaluation of some existing commercial banks. These banks were mostly in private ownership but through their strength and dominating size, came to acquire certain powers which were considered there of central banking. However, overtime, their working was bound by codes and ethics of rules and practices. Consequently, central banking developed into a distinct entity of its own. Central banks came into existence in the early 20th century through official legislation. The practice of creating a full-fledged central bank so as to take charge of the existing financial system got an impetus by the recommendations made by the International Financial Conference held at Brussels in 1920s. A large number of Central Banks were established thereafter. A central bank in a country takes into purview the design and conduct of the monetary and credit policy as its special responsibilities. The central bank is the central regulatory authority in any economy on the banking and financial structure of the country. RBI, the central bank of our country came into existence in 1935.

Functions of a Central Bank

Central banking functions have evolved gradually over decades. Their evolution has been guided by ever-changing need to find new methods of regulating, guiding and helping the financial system (particularly, the

banks). In other words, the evolution of central banking functions has tended to coincide with the evolution of the financial systems of the world economies. There are two types of functions that a central bank performs, namely;

– Leading Functions

– Other Functions

Leading Functions

(a) **Issue of Notes:** It is one of the primary functions of a central bank. The entire financial system of a country, with ever increasing volume and variety of the financial instruments, institutions and markets, needs a stable supply of legal tender money. This legal tender tends to vary, both in volume and composition, to the changing requirements of the economy. Accordingly, the central bank of the country is granted the sole right to issue currency (including that of the government of the country). The central bank exercises a monopoly over issuing bank notes in the economy.

The central bank is given monopoly of note issue for two reasons:

– Currency notes issued by a central bank are its liabilities corresponding to which it acquires certain assets which can be a source of income to it.

– Through appropriate legislation or otherwise, it can be ensured that the central bank does not over-issue currency notes.

The main advantages of granting the monopoly right of note issue to the central bank are given below:

1. It brings uniformity in the monetary system of note issue and note circulation.

2. The central bank can exercise better control over the money supply in the country. It increases public confidence in the monetary system of the country.

3. Monetary management of the paper currency becomes easier. Being the supreme bank of the country, the central bank has full information about the monetary requirements of the economy and, therefore, can change the quantity of currency accordingly.

4. It enables the central bank to exercise control over the creation of credit by the commercial banks.

5. The central bank also earns profit from the issue of paper currency.

6. Granting of monopoly right of note issue to the central bank avoids the political interference in the matter of note issue

In the initial stages, central banks were privately owned and were competing with other banks for business. During those days, therefore, they were tempted to over issue notes so as to earn more income. Consequently, the authorities felt that steps should be taken against this misuse of the privilege of monopoly of note issue. The solution of this problem was sought in removing, partially or fully, the temptation to over issue notes, that is, in ensuring that the assets acquired by the central bank against its note issue are not income yielding. The non-income yielding assets for backing the note issue were to be gold bullion and coins.

A maximum limit of note issue may be prescribed without any reference to its backing in terms of gold. Clearly, this method is highly restrictive. The supply of legal tender fails to respond to increasing needs of an expanding economy. Revising the limit through legislation can be quite cumbersome and disruptive.

It may be prescribed that the gold backing of note issue must not fall below a prescribed percentage.

(b) **Bank to Bankers:** The second main function of a central bank is that of being a bank to the banks. This signifies that the central bank has the same relationship with the commercial banks in the country that the

latter share with their customers. It provides security to their cash reserves, gives them loan at the time of need, gives them advice on financial and economic matter and work as clearing house among various member banks. The central bank acts as the bankers' bank in three capacities:

– as custodian of the cash preserves of the commercial banks

– as the lender of the last resort

– as clearing agent

In this way, the central bank acts as a friend, philosopher and guide to the commercial banks.

As a custodian of the cash reserves of the commercial banks the central bank maintains the cash reserves of the commercial banks. Every commercial bank has to keep a certain percentage of its cash balances as deposits with the central banks. These cash reserves can be utilized by the commercial banks in times of emergency.

The centralization of cash reserves in the central bank has the following advantages:

1. It inspires confidence of the public in the banking system of the country.

2. It provides the basis of a larger and more elastic credit structure than if these amounts were scattered among the individual banks.

3. Centralised reserves can be used to the fullest possible extent and in the most effective manner during the periods of seasonal strains and financial emergencies.

4. It enables the central bank to provide financial accommodation to the commercial banks which are in temporary difficulties. In fact the central bank functions as the lender of the last resort on the basis of the centralised cash reserves.

5. The system of centralized cash reserves enables the central bank to influence the creation of credit by the commercial banks by increasing or decreasing the cash reserves through the technique of variable cash-reserve ratio.

6. The cash reserves with the central bank can be used to promote national welfare.

As the supreme bank of the country and the bankers' bank, the central bank acts as the lender of the last resort. In other words, in case the commercial banks are not able to meet their financial requirements from other sources, they can, as a last resort, approach the central bank for financial accommodation. The central bank provides financial accommodation to the commercial banks by rediscounting their eligible securities and exchange bills.

The main advantages of the central bank's functioning as the lender of the last resort are:

1. It increases the elasticity and liquidity of the whole credit structure of the economy.

2. It enables the commercial banks to carry on their activities even with their limited cash reserves.

3. It provides financial help to the commercial banks in times of emergency.

4. It enables the central bank to exercise its control over banking system of the country.

As the custodian of the cash reserves of the commercial banks, the central bank acts as the clearing house for these banks. Since all banks have their accounts with the central bank, the central bank can easily settle the claims of various banks against each other with least use of cash. The clearing house function of the central bank has the following advantages:

1. It economizes the use of cash resources by banks while settling their claims and counter-claims.

2. It reduces the withdrawals of cash which enables the commercial banks to create credit on a large scale.

3. It keeps the central bank fully informed about the liquidity position of the commercial banks.

(c) Banker to the Government: The central bank of the country also acts as a banker to the government. This function normally involves three things;

– providing ordinary banking services to the government

– being a public debt agent and underwriter to the government

– being an advisor

As a Banker to Government, the central bank performs the same functions for the government as a commercial bank performs for its customers. It maintains the accounts of the central as well as state government; it receives deposits from government; it makes short- term advances to the government; it collects Cheques and drafts deposited in the government accounts; it provides foreign exchange resources to the government for repaying external debt or purchasing foreign goods or making other payments.

As an Agent to the Government, the central bank collects taxes and other payments on behalf of the government. It raises loans from the public and thus manages the public debt. It also represents the government in the international financial institutions and conferences.

As a Financial Adviser to the Government, the central bank gives advise to the government on economic, monetary, financial, banking and fiscal matters such as deficit financing, devaluation, trade policy, foreign exchange policy, etc.

(d) Custodian of Foreign Exchange Reserves: Central bank is the custodian of the foreign currency obtained from various countries. This has become an important function of central bank. These days, because with its help it can stabilize the external value of the currency. This arrangement helps the authorities in managing and co-ordinating the monetary matters of the country more effectively. This is because there is a direct association between foreign exchange reserves and quantity of money in the market. The foreign exchange reserves are influenced by international capital movements, international trade credits etc. Because of the interaction between the domestic money supply, price level, and exchange reserves, the central bank frequently faces several contradictory tendencies which have to be reconciled.

(e) Management and Regulation of Exchange Rate: A related function which is assigned to the central bank is the management, regulation and stabilisation of the exchange rate. This task is facilitated when the central bank is also the custodian of official foreign exchange reserves. The need for a stable exchange rate has emerged with the rapidly growing global connections. In this context, it was essential that this function be handled by an expert agency; for which central bank of the country is considered the best agency. The central bank being the apex institution of the entire financial system of the country possesses maximum data and has the expertise of estimating the financial trends and the type of corrective measures needed. Moreover, it has several regulatory powers over the financial system. It can contemplate and take the complementary measures needed for ensuring the success of the steps taken in the area of exchange rate.

(f) Credit Control: These days, the most important function of a central bank is to control the volume of credit for bringing about stability in the general price level and accomplishing various other socio economic objectives. The central bank has acquired the rights and powers of controlling the entire banking. A central bank can adopt various quantitative and qualitative methods for credit control such as bank rate, open market operation, changes in reserve ratio selective controls, etc. Over the years, credit control has become a leading function of a modern central bank. In earlier days, the term credit control referred to the regulation of only the volume of money and credit. Currently, the term is used in a wider meaning and covers not only the volume of money and credit, but also its components, its flows, its allocation between alternative uses and borrowers, terms and conditions attached to credit. The need for credit control

arises because it is observed that "money cannot manage itself". Left to unregulated market forces, flows of money and credit have the tendency to accentuate cyclical fluctuations.

Other Functions

(a) Collection of Data: Central banks in almost all the countries collects statistical data regularly relating to economic aspects of money, credit, foreign exchange, banking etc. from time to time. The committees and commission are appointed for studying various aspects relating to the aforesaid problem.

(b) Central Banking in Developing Countries: The basic problem of less developed countries is the problem of lack of capital formation whose main causes are lack of saving and investment. Therefore, central bank can play an important role by promoting capital formation through mobilizing savings and encouraging investment.

It is believed that an underdeveloped country requires an all- frontal approach in solving its problems of poverty and growth. Though regulation of the volume of money and credit and its other dimensions, the central bank plays a key role in its growth policy, much more is needed to make it really effective. Viewed in this manner, the functions of a central bank come to cover a much wider field than is conventionally considered in the case of central banks of developed countries.

Each country has its own Central Bank

For instance,	For India	Reserve Bank of India
	For USA	Federal Reserve
	For Europe	European Central Bank
	For China	People's Bank of China

Reserve Bank of India (RBI)

The RBI was established on April 1, 1935, under the Reserve Bank of India Act, 1934 as a private shareholders' bank with fully paid up share capital of Rs.5 crores on the basis of the recommendations of the Hilton Young Commission. The Government held shares of nominal value of Rs. 2,20,000. It was nationalised on January 1, 1949.

The executive head of the bank is called Governor, who is assisted by Deputy Governors and other Executive officers in the administration of the bank. For general direction, the bank has a central board of directors, supplemented by four local boards at Delhi, Calcutta, Madras and Bombay for four respective regional areas; northern, eastern, southern and western.

The head office of the central bank is in Mumbai. The bank has 22 offices across the country.

COMMERCIAL BANKS

A commercial bank is a financial institution authorized to provide a variety of financial services, including consumer and business loans, savings accounts etc. Earlier commercial banks were limited to accepting deposits of money or valuables for safekeeping and verifying coinage or exchanging one jurisdiction's coins for another's. By the 17th century, most of the essentials of modern banking, including foreign exchange, the payment of interest, and the granting of loans, were in place.

Broadly speaking, the functions of a commercial bank can be classified in the following two categories:

– Primary Functions
– Secondary Functions

Primary Functions

(a) **Accepting Deposits:** The commercial banks accept deposits from public, businessmen and others in form of

- Saving Deposits
- Time Deposits
- Current Deposits

Under the saving deposits banks accept small deposits from households or persons in order to encourage savings in the economy.

Fixed deposits are accepted for a fixed time period specified in advance. It carries higher rate of interest as compared to saving deposits.

In case of current account, the banks undertake the obligation of paying all cheques against the deposits of the customers subject to adequate fund in the account. In many cases the current account balance runs into overdrafts to a certain pre-specified limit. No interest is paid on such deposit types. These are mostly availed by business houses.

(b) **Lending of Fund:** The other important activity covered in the array of functions of the commercial banks is lending of the fund to the users in form of Loans and Advances; Cash Credit, Overdraft and Discounting of Bills etc.

Loans are like advances extended by the bank to his customers with or without security for a specified period of time at an agreed rate of interest. In this case, banks credit the loan amount in the customer's account who may withdraw the same as per his needs.

Under the Cash credit facility, banks offer his customers to borrow cash up to certain specified limit against the security of goods.

Overdraft is like an arrangement offered by the banks where customers are permitted temporarily to overdraw from his current account without security upto a certain pre specified limit. Banks also deals in discounting and purchasing bills. In both the situations, banks after charging discounts and commissions to credit the amount of bills in customer's accounts which shall be recovered from customers debtors on maturity of the instrument..

Secondary Functions

(a) **Agency Service:** Banks acts as an agents to their customers by rendering various services on behalf of customers like:

- Collection of bills, draft, cheques dividends etc**.**
- Payments of insurance premium, loan instalments, rent, bills etc.
- Acting as representatives of customers for stock exchange operations such as purchase and sale of securities etc.
- Acting as an executors, administrators, trustee of an estate of customers
- Other services such as preparation income tax returns, claiming of tax refunds etc.

(b) **General Utility Services:** Commercial banks also offers variety of general utility services such as issuing travellers cheques, locker facilities for keeping valuables in safe custody, issue of debit and credit cards, etc. to their customers.

Thus the functions of commercial banks can be summarized as below:

Primary Functions	Secondary Functions
Accepting Deposits – The commercial banks accept deposits from public, businessmen and others in the form of – Saving Deposits – Time Deposits – Current Deposits	Agency Service – Collection of bills, draft, cheques, dividends, etc. – Payments of insurance premium, loan installments, etc. – Acting as an executor, administrator or trustee – Other services such as preparation of income tax returns, claiming of tax refunds etc.
Lending funds	**General Utility Services**
– Loans and Advances – Cash Credit – Overdraft – Discounting of Bills	– Issuing travelers cheques – Locker facilities for keeping valuables in safe custody – Issue of debit and credit cards

BASIC ELEMENTS OF E-BANKING

Introduction

E-banking refers to electronic banking. It is also called as "Virtual Banking" or "Online Banking" or "Internet Banking". With the growth of internet and wireless communication technologies, telecommunications etc. in recent years, the structure and nature of banking and financial services have gone for an immense change. E-banking is the latest in this series of technological wonders in the recent past which involves use of internet for delivery of banking products and services.

Banks have traditionally been in the forefront of harnessing technology to improve their products, services and efficiency. They have, over a long time, been using electronic and telecommunication networks for delivering a wide range of value added products and services. The delivery channels include direct dial – up connections, private networks, public networks etc. and the devices include telephone, Personal Computers including the Automated Teller Machines, etc. With the popularity of Personal Computers, easy access to Internet and World Wide Web (WWW), internet is increasingly used by banks as a channel for receiving instructions and delivering their products and services to their customers. Nowadays, it is also possible to use internet banking on mobile phones using a Wi-Fior 3G connection. This form of banking is generally referred to as E-banking, although the range of products and services offered by different banks vary widely both in their content and sophistication. Hence, banking activity is now no longer confined to the branches where a customer has to approach the branch in person, for withdrawing cash or deposit a cheque or request for a statement of accounts. Under this system, online banking is possible where every bank customer is provided with a personal identification number (PIN) for making online transactions with the bank through internet connections.

Types of E-Banking

Broadly, the levels of banking services offered through E-Banking can be categorized in to three types:

(i) The Basic Level Service is the banks' websites which disseminate information on different products and services offered to customers and members of public in general. It may receive and reply to customers' queries through e-mail.

(ii) In the next level are Simple Transactional Websites which allow customers to submit their instructions, applications for different services, queries on their account balances, etc., but do not permit any fund-based transactions on their accounts,

(iii) The third level of Internet banking services are offered by Fully Transactional Websites which allow the customers to operate on their accounts for transfer of funds, payment of different bills, subscribing to other products of the bank and to transact purchase and sale of securities, etc.

The above forms of e-banking services are offered by traditional banks, as an additional method of serving the customer or by new banks, who deliver banking services primarily through Internet or other electronic delivery channels as the value added services. Some of these banks are known as 'virtual' banks or 'Internet-only' banks and may not have any physical presence in a country despite offering different banking services.

On the basis of above discussion we can divide the e-banking websites into the following ways :

1. Informational Websites

Informational websites provide customers access to general information about the financial institution and its products or services.

2. Transactional Websites

Transactional websites provide customers with the ability to conduct transactions through the financial institution's website by initiating banking transactions or buying products and services. Banking transactions can range from something as basic as a retail account balance inquiry to a large business-to-business funds transfer. E-banking services, like those delivered through other delivery channels, are typically classified based on the type of customer they support. The following table lists some of the common retail and wholesale e-banking services offered by financial institutions.

Table 7.4: Common E-Banking Services

Retail Services	Wholesale Services
Account management	Account management
Bill payment	Cash management
New account opening	Small business loan applications, approvals, or advances
Consumer wire transfers	Commercial wire transfers
Investment / Brokerage services	Business-to-business payments
Loan application and approval	Employee benefits / pension administration
Account aggregation	

Importance of E-banking

(a) For Banks:

1. **Reduced Transaction Cost:** E-banking helps the banks by reducing their transactional costs. As the transaction can be converted online, it is considered to be the cheapest of all the modes of transaction.

2. **Perfect Information:** E-banking makes perfect information available to all the participants by bringing efficiencies in the search process.

3. **Reduction in Paper Work:** E-banking eliminates paper waste, which is beneficial not only for those who handle the work but also well for the environment.

4. **Loyalty of Customers:** E-banking customers are found to be more loyal than non-E-banking customers as e-banking satisfies the customers so it helps in increasing the strength of banks.

5. **Minimisation of Fixed Cost:** As the banks need not to make more branches for the customers so it reduces the fixed cost of making the buildings and overhead expenses. As the cost to the customer is minimal the cost to the bank is also minimal. Though there is a fee charged for the services of e-banking but that is also low.

(b) For Customers:

1. **Convenience to Customers:** E-banking offers the services like anywhere, any time banking means 24 hours a day customers can transact their business transaction which has increased the convenience to customers.

2. **New products and services:** E-banking offers different products and services to customers. Now the customers can make their account inquiry, fund transfer, bill presentment and Demat holdings, etc.

3. **Low Cost Per Transaction:** Customers need not to visit the branch personally so it saves times, money of the customer, provides convenience, easy accessibility and has a positive impact on the customer satisfaction. It has claimed that e-banking offers the customers more benefits at lower cost.

4. **Increase in Customer Loyalty:** As the services provided through E-banking are safe, ensures security, brings transparency in the work so it increases the loyalty of the customers towards banking transactions.

5. **Removes geographical barriers:** E-banking reduces the traditional geographical barriers as it could reach to customers located at different countries. Hence, it provides the convenience of anywhere, anytime banking.

(c) For Business:

1. **Activity Review:** Business owners, accounting staff and other approved employees can access routine banking activity such as deposits, cleared checks and wired funds quickly through an online banking interface. This ease of review helps ensure the smooth processing of all banking transactions on a daily basis, rather than waiting for monthly statements. Errors or delays can be noted and resolved quicker, potentially before any business impact is felt.

2. **Productivity:** E-banking leads to productivity gains. Automating routine bill payments, minimizing the need to physically visit the bank and the ability to work as needed rather than on banking hours may decrease the time involved in performing routine banking activities. Additionally, online search tools, banking actions and other programs can allow staff members to research transactions and resolve banking problems on their own, without interacting with bank employees. In some cases, month-end settlement for credit card transactions and bank accounts can be automated by using e-banking files.

3. **Lower Banking Costs:** Banking relationships and costs are often based on resource requirements. Businesses that place more demands on banking employees and need more physical assistance with wire transfers, deposits, research requests and other banking activities often incur higher banking fees. Opting for e-banking minimizes business overhead and banking expenses.

4. **Reduced Errors:** Utilizing e-banking reduces banking errors. Automation of payments, wires or other consistent financial activities ensures payments are made on time and may prevent errors caused by keyboard slips or user error. Additionally, opting for electronic banking eliminates errors due to poor

handwriting or mistaken information. In many cases, electronic files and daily reviews of banking data can be used to double or triple check vital accounting data, which increases the accuracy of financial statements.

5. **Reduced Fraud:** Increased scrutiny of corporate finances through audits and anti-fraud measures requires a high level of visibility for all financial transactions. Relying on e-banking provides an electronic footprint for all accounting personnel, managers and business owners who modify banking activities. E-banking offers visibility into banking activities, which makes it harder for under-the-table or fraudulent activities to occur.

E-Banking in Indian Scenario: The banking industry in India is facing competition from non-traditional banking institutions offering banking and financial services over the internet with the deregulation of banking industry and with the emergence of new technologies. Banking-related internet sites in India offer only the basic services and the ICICI bank first offered the internet banking services in 1997 under the brand name 'Infinity'. Later on other new generation banks like Indusind bank, Citi bank, and federal bank, HDFC bank, and ABN Amro, etc. have also introduced to attract high profile customers. At present, all major banks in India are becoming proactive in providing e-banking services to its customers.

Popular Services Covered Under E-Banking Indian banks offer to their customers following E-Banking products and services:

1. Automated Teller Machines

2. Telephone Banking

3. Electronic Clearing Cards

4. Smart Cards

5. Electronic Funds Transfer (EFT) System

6. Electronic Clearing Services

7. Mobile Banking

8. Internet Banking

9. Telebanking

10. Door Step Banking

Facilities Available Under Internet Banking in India:

(i) Bill Payment Service:

Bill payment service is a utility service of e-banking. Accordingly, each bank has tie-ups with various utility companies, service providers, and insurance companies across the country. Such tie-ups can facilitate online payment of bills of electricity, telephone, mobile phone, credit card, insurance premium bills etc.

In order to make online payment of bills, a simple one-time registration for each bill has to be made and a standing instruction has to be made to make online payment of recurring bills automatically. Most interestingly, the bank usually does not charge customers for such online bill payment.

(ii) Fund Transfer:

Internet banking has made provision for transfer of any amount of fund from one account to another of the same or any other bank. Accordingly, customers can send money anywhere in India. Once a customer logs in his

account, he needs to mention the payee's account number, his bank and the branch. The transfer will take place in a day or so, whereas in a traditional method it takes about three to four working days.

(iii) Credit Card to Customers:

Internet banking provides the facility of credit card to its customers. With internet banking, customers can not only pay their credit card bills online but also gets a loan on their cards. Not just this, they can also apply for an additional card, request a credit line increase and in case the card is lost, one can report lost card online.

(iv) Railway Pass and Online Booking:

Through e-banking facility to issue Railway pass is also available. Indian Railways has tied up with ICICI bank for this purpose and one can now make railway pass for local trains online. The pass can be delivered to the customer at his doorstep. Initially, the facility was limited to Mumbai, Thane, Nashik, Surat and Pune. The bank would just charge Rs 10 + certain per cent of service tax. Moreover, online booking of e-tickets of Railways, Airlines etc. can also be made with some arrangement with banks through e-banking.

(v) Investing through E-Banking:

Through e-banking, opening a fixed deposit account has become easier. A customer can now open an FD account online through funds transfer. Online banking can also be a great friend for investors. Moreover, investors with interlinked de-mat[2] account and bank account can easily trade in the stock market and the amount will be automatically debited from their respective bank accounts and the shares will be credited in their de-mat account.

Besides, some banks provide its customers the facility to purchase mutual funds directly from the online banking system. Nowadays, most leading banks offer both online banking and de-mat account facilities.

(vi) Recharging Prepaid Phone:

Through Internet banking, recharging of prepaid phone has also become possible. It is no longer needed to rush to the vendor to recharge prepaid phones as and when talk time runs out. Here the customer just tops-up his prepaid mobile cards by logging in to e-banking. By just selecting operator's name, entering mobile number and the amount of recharge, the prepared phone of the customer is again back in action within few minutes.

(vii) Shopping at Fingertips:

Internet banking provides facility of shopping at fingertips. Leading banks have tied-up with various shopping websites. With a range of all kind of products, one can shop online and the payment is also made conveniently through his account.

Risks involved in E-Banking

1. Operational risk:

Operational risk also referred as the transactional risk, which is the most common form of risk associated with e-banking. It takes the form of inaccurate processing of transactions, non-enforceability of contracts, compromises in data integrity, data privacy and confidentiality, unauthorized access to bank's systems and transactions etc. Besides inadequacies in technology, human factors like negligence by customers and employees, fraudulent activity of employees and crackers / hackers etc. can become potential source of operational risk.

2. Security risk:

Security of transaction is the most important area of concern. Customers always want secrecy of their transaction. As all information is on-line so sometime customers confidential information can retrieve and also implant virus.

---

2. Demat (Dematerialisation) of Shares means an investor surrenders his physical shares and in turn gets electronic shares, i.e., paperless trading.

Further threats from hackers, unauthorized access to banks information proves harmful not only for banks but also loose customer's confidence.

3. System architecture and design

Appropriate system architecture and control is an important factor in managing various kinds of operational and security risks. Banks face the risk of wrong choice of technology, improper system design and inadequate control processes. Technology which is outdated, not scalable or not proven could land the bank in investment loss, a vulnerable system and inefficient service. Not updating bank's system in keeping with the rapidly changing technology, increases operational risk because it leaves holes in the security system of the bank. Also, staff may fail to understand fully the nature of new technology employed.

4. Reputational risk

Reputational risk is the risk of getting significant negative public opinion, which may result in a critical loss of funding or customers. Such risks arise from actions which cause major loss of the public confidence in the banks' ability to perform critical functions or impair bank-customer relationship. The main reasons for this risk may be system or product not working to the expectations of the customers, significant system deficiencies, significant security breach (both due to internal and external attack), inadequate information to customers about product use and problem resolution procedures, significant problems with communication networks that impair customers' access to their funds or account information especially if there are no alternative means of account access.

5. Legal risk

Legal risk arises from violation of laws, rules, regulations, or prescribed practices, or when the legal rights and obligations of parties to a transaction are not well established. Given the relatively new nature of e-banking, rights and obligations in some cases are uncertain and applicability of laws and rules is uncertain or ambiguous, thus causing legal risk.

If banks are allowed to play a role in authentication of systems such as acting as a Certification Authority, it will bring additional risks. A digital certificate is intended to ensure that a given signature is, in fact, generated by a given signer. Because of this, the certifying bank may become liable for the financial losses incurred by the party relying on the digital certificate.

6. Money laundering risk

As e-banking transactions are conducted remotely banks may find it difficult to apply traditional method for detecting and preventing undesirable criminal activities. Application of money laundering rules may also be inappropriate for some forms of electronic payments. Thus banks expose themselves to the money laundering risk.

7. Cross border risks

Internet banking is based on technology that is designed to extend the geographic reach of banks and customers. Such market expansion can extend beyond national borders. This causes various risks.

- It includes legal and regulatory risks, as there may be uncertainty about legal requirements in some countries and jurisdiction ambiguities with respect to the responsibilities of different national authorities.

- If a bank uses a service provider located in another country, it will be more difficult to monitor it thus, causing operational risk.

- Cross border transaction accentuates credit risk, since it is difficult to appraise an application for a loan from a customer in another country compared to a customer from a familiar customer base.

8. Strategic Risk

This risk is associated with the various issues related to development of a business plan, availability of sufficient resources to support this plan, credibility of the vendor (if outsourced), change in the work environment for employees, and level of the technology used in comparison to the available technology etc.

9. Other risks

Traditional banking risks such as credit risk, liquidity risk, interest rate risk and market risk are also present in e-banking. These risks get intensified due to the very nature of e-banking on account of use of electronic channels as well as absence of geographical limits.

All the above mentioned risks arise due to flaws in design, insufficient technology, and employee's negligence, unauthorized work environment either intentional or unintentional. To reduce such kinds of risks there is a need to adopt appropriate technology, appropriate system proper access control so that security of banking transaction can be enhanced and is helpful in building customer satisfaction.

RBI's Recommendation on E-Banking: RBI has constituted a working group to examine different issues of e-banking and suggested different measures on technology, security, legal practices. Some of these recommendations are as under:

1. The Indian banks association should design a standard which should be adopted by all the banks in India keeping in mind the security concern.

2. Banks should adopt adequate security measures to maintain the secrecy and confidentiality of data and to implement it, logical access control should be used.

3. Anti-money laundering (ALM) technology system should be developed for reporting and query purpose.

4. To adopt a fraud free culture, internal grievance redressal system should be installed.

5. An explicit security plan with documentation should be ensured and all the physical access control should be strictly ensured.

6. To adopt an extensive e-banking network which also prove to be help for rural and remotest areas of the country.

MONETARY POLICY IN INDIA

INTRODUCTION

Monetary policy is a regulatory policy by which the central bank or monetary authority of a country controls the supply of money, availability of bank credit and cost of money, that is, the rate of Interest.

Monetary policy / monetary management is regarded as an important tool of economic management in India. RBI, which is the central bank of India or monetary authority of India, controls the supply of money and bank credit. The Central bank has the duty to see that legitimate credit requirements are met and at the same credit is not used for unproductive and speculative purposes. RBI rightly calls its credit policy as one of controlled expansion.

OBJECTIVES OF MONETARY POLICY OF INDIA:

The main objective of monetary policy in India is 'growth with stability'. Monetary management regulates availability, cost and use of money and credit. Following are the main objectives of monetary policy in India:

1. Growth with Stability

Traditionally, RBI's monetary policy was focused on controlling inflation through contraction of money supply

and credit. This resulted in poor growth performance. Thus, RBI have now adopted the policy of 'Growth with Stability'. This means sufficient credit will be available for growing needs of different sectors of economy and at the same time, inflation will be controlled with in a certain limit.

2. Regulation, Supervision and Development of Financial Stability

Financial stability means the ability of the economy to absorb shocks and maintain confidence in financial system. Threats to financial stability can come from internal and external shocks. Such shocks can destabilize the country's financial system. Thus, greater importance is being given to RBI's role in maintaining confidence in financial system through proper regulation and controls, without sacrificing the objective of growth. Therefore, RBI is focusing on regulation, supervision and development of financial system.

3. Promoting Priority Sector

Priority sector includes agriculture, export and small-scale enterprises and weaker section of population. RBI with the help of bank provides timely and adequately credit at affordable cost of weaker sections and low income groups.

4. Generation of Employment

Monetary policy helps in employment generation by influencing the rate of investment and allocation of investment among various economic activities of different labour Intensities.

5. External Stability

With the growth of imports and exports India's linkages with global economy are getting stronger. Earlier, RBI controlled foreign exchange market by determining exchange rate. Now, RBI has only indirect control over external stability through the mechanism of 'Managed Flexibility', where it influences exchange rate by buying and selling foreign currencies in open market.

6. Encouraging Savings and Investments

RBI by offering attractive interest rates encourages savings in the economy. A high rate of saving promotes investment. Thus the monetary management by influencing rates of interest can influence saving mobilization in the country.

7. Redistribution of Income and Wealth

By control of inflation and deployment of credit to weaker sectors of society the monetary policy may redistribute income and wealth favouring to weaker sections.

8. Regulation of NBFIs

Non-Banking Financial Institutions (NBFIs), like UTI, IDBI, IFCI and others plays an important role in deployment of credit and mobilization of savings. RBI does not have any direct control on the functioning of such institutions. However it can indirectly affects the policies and functions of NBFIs through its monetary policy.

MONETARY POLICY OF RBI

The Monetary Policy of RBI is not merely one of credit restriction, but it has also the duty to see that legitimate credit requirements are met and at the same time credit is not used for unproductive and speculative purposes. The monetary policy is implemented by RBI through instruments of credit control. These instruments are classified into two types:

– Quantitative or General Measures

– Qualitative or Selective Measures

(A) General or Quantitative Credit Control Methods

In India, the legal framework of RBI's control over the credit structure has been provided Under Reserve Bank of India Act, 1934 and the Banking Regulation Act, 1949. The Quantitative Instruments are also known as the General Tools of monetary policy. These tools are related to the Quantity or Volume of the money. They are designed to regulate or control the total volume of bank credit in the economy. The general tool of credit control comprises of following instruments.

1. Bank Rate Policy (BRP)

The Bank Rate Policy (BRP) is a very important technique used in the monetary policy for influencing the volume or the quantity of the credit in a country. The bank rate refers to rate at which the central bank (i.e. RBI) rediscounts bills and prepares of commercial banks or provides advance to commercial banks against approved securities. It is "the standard rate at which the bank is prepared to buy or rediscount bills of exchange or other commercial paper eligible for purchase under the RBI Act". The Bank Rate affects the actual availability and the cost of the credit, i.e., interest rate. Any change in the bank rate necessarily brings out a resultant change in the cost of credit available to commercial banks. If the RBI increases the bank rate then it reduces the volume of commercial banks borrowing from the RBI. It deters banks from further credit expansion as it becomes a more costly affair. On the other hand, if the RBI reduces the bank rate, borrowing for commercial banks will be easy and cheaper. This will boost the credit creation. Thus any change in the bank rate is normally associated with the resulting changes in the lending rate and in the market rate of interest. However, the efficiency of the bank rate as a tool of monetary policy depends on existing banking network, interest elasticity of investment demand, size and strength of the money market, international flow of funds, etc.

2. Open Market Operation (OMO)

The open market operation refers to the purchase and/or sale of short term and long term securities by the RBI in the open market. This is very effective and popular instrument of the monetary policy. The OMO is used to wipe out shortage of money in the money market, to influence the term and structure of the interest rate and to stabilize the market for government securities, etc. It is important to understand the working of the OMO. If the RBI sells securities in an open market, commercial banks and private individuals buy it. This reduces the existing money supply as money gets transferred from commercial banks to the RBI. Contrary to this when the RBI buys the securities from commercial banks in the open market, commercial banks and individuals sell it and get back the money they had invested in them. Obviously the stock of money in the economy increases. This way when the RBI enters in the OMO transactions, the actual stock of money gets changed. Normally during the inflation period in order to reduce the purchasing power, the RBI sells securities and during the recession or depression phase it buys securities and makes more money available in the economy through the banking system. Thus under OMO there is continuous buying and selling of securities taking place leading to changes in the availability of credit in an economy.

3. Variation in the Reserve Ratios (VRR)

The Commercial Banks have to keep a certain proportion of their total assets in the form of Cash Reserves. Some part of these cash reserves are their total assets in the form of cash. Apart of these cash reserves are also to be kept with the RBI for the purpose of maintaining liquidity and controlling credit in an economy. These reserve ratios are named as Cash Reserve Ratio (CRR) and a Statutory Liquidity Ratio (SLR). The CRR refers to some percentage of commercial bank's net demand and time liabilities which commercial banks have to maintain with the central bank and SLR refers to some percent of reserves to be maintained in the form of gold or foreign securities. In India the CRR by law remains in between 3-15 percent while the SLR remains in between 25-40 percent of bank reserves. Any change in the VRR (i.e. CRR + SLR) brings out a change in commercial banks reserves positions. Thus by varying VRR commercial banks lending capacity can be affected. Changes in the VRR helps in bringing changes in the cash reserves of commercial banks and thus it can affect the banks credit creation. RBI increases VRR during the inflation to reduce the purchasing power and credit

creation. But during the recession or depression it lowers the VRR making more cash reserves available for credit expansion.

4. Repo Rate and Reverse Repo Rate

In addition to the above instrument, another tool used by central bank is Repo and Reverse Repo Rate. Repo rate is the rate at which commercial banks borrow money from the central bank in case of any shortage of funds they face. RBI lends money to bankers against approved securities for meeting their day to day requirements or to fill short term gap. A reduction in the repo rate will result in increased borrowing powers of the banks hence increased amounts with the banks to lend out. Consequently, this will increase the supply of credit. Reverse Repo rate is the opposite concept. It is the rate at which RBI borrows money from commercial banks. An increase in reverse repo rate can cause banks to transfer more funds to RBI due to attractive interests hence, reducing lending capacity.

(B) Qualitative Instruments or Selective Tools

The Qualitative Instruments are also known as the Selective Tools of monetary policy. These tools are not directed towards the quality of credit or the use of the credit. They are used for discriminating between different uses of credit. It can be discrimination favoring export over import or essential over non-essential credit supply. This method can have influence over the lender and borrower of the credit. The Selective Tools of credit control comprises of following instruments.

1. Fixing Margin Requirements

The margin refers to the part of a loan which a borrower has to raise in order to get finance for his purpose. In simple words, marginal requirement is the difference between the market value of the security and its maximum loan value. If a security has a market value of Rs. 100 and if the marginal requirement is 60% the maximum loan that can be advanced for the purchase of security is Rs. 40. A change in a margin implies a change in the loan size. This method is used to encourage credit supply for the needy sector and discourage it for other non-necessary sectors. This can be done by increasing margin for the non-necessary sectors and by reducing it for other needy sectors. Example, if the RBI feels that more credit supply should be allocated to agriculture sector, then it will reduce the margin and even 85-90 percent loan can be given.

2. Consumer Credit Regulation

Under this method, consumer credit supply is regulated through hire-purchase and installment sale of consumer goods. Under this method the down payment, installment amount, loan duration, etc. is fixed in advance. This can help in checking the credit use and then inflation in a country.

3. Publicity

This is yet another method of selective credit control. Through it Central Bank (RBI) publishes various reports stating what is good and what is bad in the system. This published information can help commercial banks to direct credit supply in the desired sectors. Through its weekly and monthly bulletins, the information is made public and banks can use it for attaining goals of monetary policy.

4. Credit Rationing

Central Bank fixes credit amount to be granted. Credit is rationed by limiting the amount available for each commercial bank. This method controls even bill rediscounting. For certain purpose, upper limit of credit can be fixed and banks are told to stick to this limit. This can help in lowering banks credit exposure to unwanted sectors.

5. Moral Suasion

It implies to pressure exerted by the RBI on the Indian banking system without any strict action for compliance

of the rules. It is a suggestion to banks. It helps in restraining credit during inflationary periods. Commercial banks are informed about the expectations of the central bank through a monetary policy. Under moral suasion central banks can issue directives, guidelines and suggestions for commercial banks regarding reducing credit supply for speculative purposes.

6. Control through Directives

Under this method the central bank issue frequent directives to commercial banks. These directives guide commercial banks in framing their lending policy. Through a directive the central bank can influence credit structures, supply of credit to certain limit for a specific purpose. The RBI issues directives to commercial banks for not lending loans to speculative sector such as securities, etc. beyond a certain limit.

7. Direct Action

Under this method the RBI can impose an action against a bank. If certain banks are not adhering to the RBI's directives, the RBI may refuse to rediscount their bills and securities. Secondly, RBI may refuse credit supply to those banks whose borrowings are in excess to their capital. Central bank can penalize a bank by charging some rates. At last it can even put a ban on a particular bank if it does not follow its directives and work against the objectives of the monetary policy.

LESSON ROUND UP

- Money is an important component of the economy. Money was brought into existence to tide over the difficulties of barter system that was the traditional method to trade.

- There are two approaches adopted by economists to define money
  - Functional Approach
  - Liquidity or Generalized Purchasing Power Approach

- The functional approach defines money based on following functions:
  - Medium of Exchange
  - Measure of Value
  - Standard of Deferred Payment
  - Store of Value
  - Transfer of Value
  - Basis of Credit
  - Distribution of Social Income
  - General Form of Capital
  - Maximum Benefit

- The liquidity approach provides acceptability of money by creditors in discharge of their claims, and by sellers as sales proceeds, means that money has a generalized purchasing power. The liquidity approach emphasizes the function of money as a store of money.

- Money facilitates the process of credit creation which is the process of inducing and extraction money supply in the economy.

- A bank differs from other financial institutions because it can create credit. The process of 'Credit Creation' begins with banks lending money out of primary deposits.

– During the process of credit creation, there arise two important issues which hold utmost consideration, namely profitability and liquidity.

– Quantity Theory of Money seeks to explain the factors that determine the general price level in an economy

– There are two versions of the Quantity Theory of Money:

  – Transaction Approach

  – Cash Balance Approach

– The overall control of an economy's financial and monetary structure is controlled by the apex institution known as central bank which is Reserve bank of India.

– A commercial bank is a financial institution authorized to provide a variety of financial services, including consumer and business loans, savings accounts etc.

– E-banking is defined as the automated delivery of new and traditional banking products and services directly to customers through electronic, interactive communication channels.

– E-banking includes the systems that enable financial institution customers, individuals or businesses, to access accounts, transact business, or obtain information on financial products and services through a public or private network, including the Internet.

– Customers access e-banking services using an intelligent electronic device, such as a personal computer (PC), personal digital assistant (PDA), automated teller machine (ATM), kiosk, or Touch Tone telephone.

– While the risks and controls are similar for the various e-banking access channels, this booklet focuses specifically on Internet-based services due to the Internet's widely accessible public network.

– By monetary policy, is meant the policy which results in changes in the supply of money. Monetary Policy of a country is the Central Bank's policy for controlling the availability, cost and use of money and credit with the help of monetary measures in order to achieve certain definite goals.

– The monetary policy is implemented by RBI through instruments of credit control. These instruments are classified into two types:

  – Quantitative or General Measures

  – Qualitative or Specific Measures

– The leading instruments associated with general credit control measures are:

  – Bank Rate Policy

  – Open Market Operations

  – Variable Reserve Requirements

  – Repo Rate and Reverse Repo Rate

– Selective credit control measures are exercised through issuing specific instructions to the banks.

GLOSSARY

Barter System	Barter System is that system in which goods are exchanged for goods. In ancient times when money was not invented trade as a whole was on barter system.
Functional Approach	The approach to define money based on the functions of money.

Liquidity Approach	The approach to define money based on its characteristic of liquidity.
Deferred Payment	A debt which has been incurred and will be paid back at some point in the future.
Purchasing Power	The ability to purchase which is generally measured by income.
	The value of money, as measured by the quantity and quality of products and services it can buy. It is also called buying power.
Financial Institutions	An establishment that focuses on dealing with financial transactions, such as investments, loans and deposits. Conventionally, financial institutions are composed of organizations such as banks, trust companies, insurance companies and investment dealers.
Digital certificate	The electronic equivalent of an ID card that authenticates the originator of a digital signature.
E-Banking	The remote delivery of new and traditional banking products and services through electronic delivery channels.
Firewall	A hardware or software link in a network that relays only data packets clearly intended and authorized to reach the other side.
Hacker	An individual who attempts to break into a computer without authorization.
Internet	The global system of interconnected computer networks that use the Internet protocol suite (TCP/IP) to link billions of devices worldwide.
Kiosk	A publicly accessible computer terminal that permits customers to directly communicate with the financial institution via a network.
Legacy systems	A term commonly used to refer to existing computers systems and applications with which new systems or applications must exchange information.
Smart cards	A card with an embedded computer chip on which information can be stored and processed.
Wire transfer	Often used to refer to any electronic transfer of money from one person to another
Monetary Policy	The actions of a central bank, currency board or other regulatory committee (monetary authorities) that determine the size and rate of growth of the money supply, which in turn affects interest rates.
Minsky moment	A sudden major collapse of asset values which is part of the credit cycle or business cycle.

MULTIPLE CHOICE QUESTIONS

1. **Which one is not a secondary function of Money?**

   (a) Standard of deferred payment

   (b) Medium of Exchange

   (c) Store of value

   (d) Transfer of value

2. **According to the Quantity theory of Money**

   (a) MV = PT

(b) MP = VT

(c) MT = VP

(d) T = MPV

3. **Reserve Bank of India (RBI) was established in ?**

(a) July 1, 1937

(b) July 1, 1935

(c) April 1, 1935

(d) April 1, 1937

4. The head office of the RBI is situated in?

(a) Delhi

(b) Mumbai

(c) Kolkata

(d) Chennai

5. **"Anything that is generally acceptable as a means of exchange and at the same time acts as a measure and store of value." This definition of money is given by?**

(a) Crowther

(b) Adam Smith

(c) Ricardo

(d) Keynes

6. **According to Central Bank approach, Money is?**

(a) M = C+DD+TD+SB+S+B

(b) M = C+DD+TD+NBFI+CUA

(c) M = C+DD+TD

(d) M = C+DD

7. **Which one is not a function of commercial bank?**

(a) Accepting Deposis

(b) Lending of Funds

(c) Issuing Travelers Cheques

(d) Custodian of Foreign Exchange Reserves

8. **The formula to calculate the value of Credit Multiplier(k) is?**

(a) $\Delta D/\Delta R$

(b) $\Delta R/\Delta D$

(c) $\Delta D \times \Delta R$

(d) $\Delta D + \Delta R$

9. **Which one is known as "The Lander of the Last Resort."**

(a) World Bank

(b) IMF

(c) RBI

(d) Commercial Banks

10. **Who Implements the Monetary Policy in India?**

(a) Central Government

(b) State Government

(c) RBI

(d) World Bank

Answer Key : 1.(b), 2. (a), 3. (c), 4. (b), 5.(a), 6. (b), 7.(d), 8. (a), 9.(c), 10.(c)

Suggested Readings:

1. B.S. Bhatia: Management of Service Sector

2. E-Banking: Federal Financial Institutions Examination Council, Handbook

3. R.R. Paul: Money and Banking

4. Suraj B. Gupta: Monetary Economics Institutions, Theory and Policy

5. G. Crowther: An Outline of Money

6. R.S. Sayers: Modern Banking

7. RBI Governor speeches on Monetary Policy

PART B
ELEMENTARY STATISTICS

Due to advanced communication network, rapid changes in behaviour, varied expectations and new market openings, it sometimes become difficult task of making quick and appropriate decisions. Therefore, there is a need to depend more upon quantitative techniques like statistics and basic business mathematics. On the other hand, it is always required to take proper care in interpreting the values in statistics, lest it results in misleading conclusions.

For instance, a classical story provides that a village teacher was to cross river along with his family. He calculated the average height of family members as 5 feet and average depth of river is as 3 feet.

On the other side of the river, his children were missing. He again made calculation and arrived at the same result, so he wrote a couplet in urdu;

'Araba jyon ka tyon, toh kunba dooba kyon'.

Statistics is broadly divided into two main categories. The two categories of statistics are descriptive statistics and inferential statistics. Descriptive Statistics: It is used to present the general description of data which is summarised quantitatively. Inferential statistics: It is used to make valid inferences from the data which are helpful in effective decision making for managers or professionals.

This section introduces students to the basic tools of statistics and their applicability to give them confidence in drawing relevant conclusions from provided data.

Statistics

Lesson 8
Descriptive Statistics

LESSON OUTLINE

LEARNING OBJECTIVES

"Statistics", that a word is often used, has been derived from the Latin word 'Status' that means a group of numbers or figures; those represent some information of our human interest. We find statistics in everyday life, such as in books or other information papers or TV or newspapers. Collection and presentation of data is an integral part of any research or any economic decision making activity of an organization.

In a statistical enquiry, after collection, classification and presentation of data, the important step is to draw relevant and meaningful conclusions out of the information so organized. This is where the need to understand basic statistical tools & techniques emerges. The important are averages, dispersion and Bivariate analysis. This section deals with three most common measures of central tendency which are, the mean, the median and, the mode and we also discuss some of the measures of dispersion like, mean deviation, standard deviation, range and coefficient of variation. This lesson also deals with bi-variate analysis, in which we'll discuss the covariance and coefficient of correlation.

This lesson deals with these aforementioned techniques in order to bring understanding of basic statistical enquiry in students.

STATISTICS : AN OVERVIEW

The word 'Statistics' is as old as human society itself. It is regarded in general as groups of numerical statements of facts relating to certain human experience, put in systematic and logical manner. Perhaps this word of English language had its origin from the Latin word 'Status meaning political state'. Words 'Stato' and 'Statistik' were also used in similar sense in Italian and German languages respectively. In ancient days', the word statistics was used only for facts and figures required by the state for its official use and was known as "the science of king", "the science of state craft" and 'political arithmetic'. At that time, statistics was used to collect information relating to crimes, military strength and wealth for devising military and fiscal policies. It was also used to satisfy the needs of the ruling kings, the needs like number of soldiers, number of objects etc. which helped the state in administration. But today, statistics is not merely a by product of the administrative set up of the state but it embraces all sciences – social, physical and natural and is finding numerical applications in various diversified fields such as agriculture, industry, sociology, psychology, planning, economics, business management and so on. It is rather impossible to think of any sphere of human activity where statistics has not crept in.

The word statistics has acquired a wider meaning over the years. The set of quantitative data relating to phenomena whether concerning state or otherwise are called statistics. The word statistics is used in different senses by different persons. To a layman, statistics has become 'synonymous with figures'. For others, statistics is a discipline. It has grown today as a separate discipline or subject of study like mathematics or economics. Thus, the word 'statistics' is basically used in two senses i.e. as:

– Statistics Data

– Statistical Method

Statistics in the sense of numerical data basically refers to quantitative aspects and is in the form of numerical data. Statistics as statistical method refers to a body of knowledge which contains statistical principles and methods used for collection, organisation, presentation, analysis and interpretation of data. It is a branch of scientific methods used for dealing with those phenomena which can be described numerically either by measurement or by counting.

To resolve the controversy regarding the meaning of statistics, it is essential to go through various definitions given by experts on the subject from time to time.

DEFINITIONS Of STATISTICS

I. Statistics as Numerical Data

Some of the important and popular definitions of statistics as numerical data are presented as under:

"Statistics refers to the classified facts representing the conditions of the people in the State ... especially those facts which can be stated in numbers or in tables of numbers or in any tabular classified arrangement".

Webster

"Statistics are collection of noteworthy facts concerning state both historical and descriptive".

Gotfriedand Achenwall

"Statistics is an aggregate knowledge brought together for a practical end namely the knowledge of concrete states".

Wappans

"Statistics is numerical statement of facts in any development of enquiry placed in relation to each other".

Bowley

"By statistics, we mean quantitative data affected to a marked extent by multiplicity of causes".

Kendal

All the above mentioned definitions are narrow in scope. These definitions restrict their scope of statistics to facts relating to state only. According to modern thinkers, the scope of the science of statistics is considerably wide today and is applicable to almost anything worth mentioning such as biology, demoghraphy, economics, commerce etc. In a sense, above definitions are incomplete because of their poor coverage.

Prof. Horace Secrist has given most acceptable, popular and comprehensive definition of statistics in the sense of numerical data. According to him, "By statistics we mean aggregate of facts affected to marked extent by multiplicity of causes, numerically expressed, enumerated or estimated according to a reasonable standard of accuracy, collected in a systematic manner for a pre-determined purpose and placed in relation to each other".

According to this definition, the numerical data should possess the following major characteristics in order to be known as statistics.

Characteristics of Statistics

(i) *Aggregate of Facts:* Statistics must relate to aggregate of facts. A single figure even though numerically expressed cannot be called statistics because such figures are unrelated and cannot be compared. Hence, we cannot draw any conclusions out of it. For instance, a single figure relating to the height of a student or marks of a student in the class in a particular subject is meaningless and irrelevant unless we are providing the complete data about all the students of the class. It makes the comparison among students possible. Thus, we can say, that it is not the datum but the data which is the subject matter of statistics.

(ii) *Affected by Multiplicity of Causes:* Generally speaking, statistical facts are not traceable to single factor. Since statistics are commonly used in social science, it is affected by multiple factors at the same time. For example, the prices of particular commodity are affected by number of factors such as supply, demand, import, export, income distribution, money in circulation, taste, etc. In natural sciences, like physics and chemistry, it is possible to isolate the effect of various factors on a particular item but it is very difficult to do so in social science. However, the science of statistics has provided the novel technique such as multiple correlation to study the joint effects of a number of factors on a single item.

(iii) *Numerically Expressed:* It is only the numerical data which constitute statistics. Any qualitative expression such as rich, poor, good, bad, intelligent, young, old, honesty etc do not constitute statistics, and this is outside the scope of study of statistics unless a numerical equivalent value is assigned to each expression. For example, intelligence is not statistics but intelligence quotient designed by Dr. Fraud could be accepted as statistics of attributes. Thus, in statistics facts are expressed in numbers and not words but it does not mean that all numerical statements are statistics.

(iv) *Enumerated or Estimated:* Statistical data relating to any subject could be collected either by actual counting or measurement or by estimation. In case, where the scope of enquiry is very large, enumeration is not advisable because it involves high costs and is also time consuming too. In such cases, data are estimated by using the powerful techniques of sampling and estimation. On the other hand, if number of units is less, they are actually counted or measured. For example to find out the number of students appearing in an examination in a particular class, we have to actually count, but to find out the number of persons attending a particular public meeting, we rely on estimates.

(v) *Reasonable Degree of Accuracy:* Statistical data are collected only with a reasonable degree of accuracy. The data which are collected by enumeration are supposed to be accurate and precise but it is not so in the case of estimation. The degree of accuracy of estimated value largely depends on the nature and purpose of the enquiry. There cannot be a uniform standard of accuracy for all type of enquiries. For example, if the height of the students of a class are being measured, it is all right if measurement is correct to a centimeter but if we are measuring distance between two places say Delhi and Calcutta, the difference of a few kilometers may be ignored. Thus, in many statistical studies mathematical accuracy cannot be maintained but it is important that reasonable standards of accuracy must be maintained for drawing meaningful conclusions.

(vi) *Systematic Manner:* The collection of data must be in a careful and systematic manner because data collected in haphazard and unsystematic manner may lead to erroneous, misleading fallacious conclusions. It requires careful planning of data collection approach in order to attain reasonable degree of accuracy.

(vii) *Pre-determined Purpose:* It is essential that one must be clear in advance about the objectives and purpose of enquiry or data to be collected before actually going for it. If data are not collected with some objective, the whole effort in term of money and result would be waste. For example, if the, purpose of enquiry is to measure level of education in labour strata, it would not serve any purpose unless data are collected from the area dominated by labour strata.

(viii) *Placed in Relation to Each Other:* The comparability is one of the important attributes for the numerical facts to be known as statistics. The numerical data may be compared period-wise or region-wise. For example, per capita income of India may be compared at a point of time to that of Pakistan, China, Brazil, Mexico, provided the data are homogeneous. They are related to same phenomena.

II. Statistics as Statistical Method

Statistics has developed as a full fledged discipline over a period of time. Basically, it is the science which provides tools for analysis and interpretation of huge amount of data collected for the purpose of decision-making in various fields of scientific enquiry.

Some of the popular definitions of statistics as statistical methods are given below for examination:

"Statistics may be called as the science of counting." **A.L. Bowley**

"Statistics may be rightly called the science of averages". **A.L. Bowley**

"The science of statistics is the method of judging collective, natural or social phenomenon from the results obtained by the analysis of an enumeration or collection of estimates." **King**

"Statistics is the science which deals with classification and tabulation of numerical facts as the basis for explanation, description and comparison of phenomenon." **Lovin**

"Statistics is the science which deals with the method of collecting, classifying, presenting comparing and interpreting numerical data collected to throw some light on any sphere of enquiry." **Selligman**

"Modern statistics refers to a body of methods and principles that have been developed to handle the collection, description, summarisation and analysis of numerical data. Its primary objective is to assist the researcher in making decision or generalisation about the nature and characteristics of all the potential observations under consideration of which the collected data form only a small part." **Lincon L. Chao**

"Statistics may be defined as the science of collection, presentation, analysis and interpretation of numerical data." **Croxton & Cowden**

Earlier definitions of statistics given by Prof. A.L. Bowley are inadequate and defective because they restrict the scope of statistics to counting and average. Prof. King's definition is comprehensive but it focuses on collection

of data, treating other methods of secondary importance. This is not a correct approach. Lovin's definition is fairly satisfactory because it indicates simple and scientific exposition of statistical methods.

Croxton and Cowdon have given more satisfactory and acceptable definition. This definition brings into its fold various stages of statistical enquiry i.e., collection, presentation, analysis and interpretation of numerical data.

Thus, on examining above mentioned definitions, statistics may be defined as a discipline which is concerned with collection, organisation, presentation, analysis and interpretation of numerical data.

According to this definition, a brief detail of various stages of statistical enquiry are given below:

Stages of Statistical Enquiry

(i) *Collection of Data:* Collection of statistical data constitutes one of the important aspects of statistical inquiry. During this stage, the relevant data are collected from various sources. The sources of data may be primary or secondary. In the case of primary data, data are originally collected by an investigator or agency, on the other hand, the data published or unpublished which have already been collected and processed by some agency or person and taken over from there and used by any other agency for their statistical work are termed as secondary. It may be observed that the distinction between primary and secondary data is matter of degree only.

(ii) *Organisation and Presentation of Numerical Data:* Organisation and presentation of data is the second important stage of statistical enquiry. Generally, secondary data are available in organized form but primary data collected through survey in raw form needs their organisation. Some of the important components of organisation of data are editing, classification and tabulation of data. Editing of data removes omission, inaccuracies and inconsistencies whereas in classification, general items which have common characteristics are brought together before the tabulation. Subsequently, they are presented in table and finally the data is given the shape of a diagram, chart and graph wherever necessary.

(iii) *Analysis of Numerical Data:* In this phase of statistical enquiry, the numerical data after collection, organisation and presentation is analyzed, in order to have deep understanding of the subject matter. In the process of analyzing numerical data, some of the popular measures, such as averages or measures of central tendency, dispersion, skewness, correlation, regression, association and attributes, interpolation and extrapolations are used. Probability and distribution, sampling, analysis of variance, index number and time series are also used if required to simplify the data.

(iv) *Interpretation of Numerical Data:* The interpretation of data refers to drawing inferences and conclusions from the given data. The interpretation of data is most sensitive and difficult task which requires higher degree of skill, experience and common sense and balance judgment on the part of the investigator. If the data is misinterpreted, it may lead to misleading conclusions which results in wastage of time and resources and thus the object of enquiry may be defeated.

Importance and Functions of Statistics

Statistics is a useful science. There is hardly any area in social science or natural science where statistical methods are not effectively used. In fact, to a very striking degree, the modern culture has become a statistical culture and there is no ground for misgivings regarding the practical realization of the dream of H.G. Wells "statistical thinking one day be as necessary for effective citizenship as the ability to read and write." The science of statistics has gained an enormous importance and popularity because of various functions performed by it.

(i) *Preciseness and Definiteness:* Statistics presents the facts in a precise and definite form. It is generally agreed that facts which are presented in numerical form are more precise and convincing. Statistics, in a sense, leaves, no scope for vagueness. For example, the statement that total population of India has

increased during 1981 to 1991. This statement does not give any idea about increase in population. But if we say that total population of India has increased to 856 millions during 1981-1991, this statement is more precise and conclusive.

(ii) *Simplification in Understandable Form:* Statistics presents the large volume of complex data in meaningful and understandable form. In the words of W. King's, it is for the purpose of simplifying these unwieldy masses of facts that statistical science is useful. It reduces them to numerical totals or average which may be abstractly handled like any other number. It draws pictures and diagrams to illustrate general tendencies and thus in many ways adopts this group of ideas to the capacity of our intellects.

(iii) *Facilitates Comparison:* Comparison is also one of the important functions of statistics. It enables us to understand the behaviour of data over a time period or at a point of time and has certain reasons for it. Averages, standard deviations, coefficient of variation, time series, Index number, correlation, graphs, diagrams are some of the statistical techniques which make the comparison of particular phenomenon possible.

(iv) *Enrich Individual Knowledge and Experience:* Statistics is a science which provides opportunities to individuals to enrich their knowledge and experience. It is well established fact that statistics with the help of noble techniques and sound methodology has opened new vistas of knowledge which was closed to mankind otherwise. In a sense, statistics has provided such a master key to mankind that we can use it anywhere and can study any problem and on right terms.

(v) *Formulation of Policies:* Statistics also helps in the formulation of various economic, business and other policies at state, national or global level. For example, framing of government policies on education, taxation, pollution, law and order need a sound statistical support. Poor statistical base and network in the economy may prove a handicap in designing sound policy and achieving expected results thereafter. Business organizations also make use of statistics to design their policies in areas of finance, marketing and personnel. They collect relevant information on various parameters, analyze it and draw inferences and then in the light of conclusions derived from or gathered information design their policies.

(vi) *Business Forecasting:* Forecasting means formal process of predicting future events that will significantly affect the functioning of enterprise. Government and business units needs forecasting for designing better plans and policies and sound decision-making in order to use existing man and material resources optimally. To achieve this end, statistics offers various forecasting techniques such as time series, extrapolation in order to make reliable predictions on future. About statistical forecast, Prof. Bowley rightly observed that "A statistical estimate may be good or bad, accounts for the reverse, but almost in all cases it is likely to be more accurate than a casual observers' impression".

(vii) *Framing and Testing Hypothesis:* Statistics also helps to formulate hypothesis and to test its validity and subsequently for developing sound theories. The hypothesis may be about any phenomena on political, social or economic aspects. In the words of Lord J.M. Keynes, it is rightly pointed out that the function of statistics is first to suggest empirical laws which may or may not be capable of subsequent deductive reasoning by checking its results and submitting them to the test of experience".

Functions of Statistics

Some of the important functions of statistics are as under:-

(1) It helps in comparison

(2) Its facilitates in the formulation of appropriate policies

(3) Its facilitates in the formation and also in the testing of hypothesis

(4) Its helps in business predictions

(5) Its helps in the presentation of data in an organised manner

Comparison

When the data is not compared with the same kind of data, we are not able to find the population of the real meaning. For example if we say that, the population of 2016-17 is going to be increased as compared to 2010-11. In this example we are not able fo get the real meaning. Now look at the second example, that the population in 2016-17 will cross 125 crores in caparison with 120 crores in 2010-11. It gives us the real meaning.

Formulation of Appropriate Policies

According to Robert W. Burgess," The fundamental gospel of statistics is to push back the domain of ignorance, rule of thumb, arbitrary or premature decision, traditions and dogmatism and to increase the domain in which decisions are made and principles are formulated on the basis of analysed quantitative facts." Statistics provide us the framework to formulate appropriate policies. Like, if we have the data of National income, its distribution, contribution of different sectors, growth of it, etc. It helps the government in investment policy, and also in public expenditure policy.

Formation and Testing of Hypothesis

No doubt, statistics is very useful in formulation and testing of hypothesis, and it also helps in developing new theories. For example, paracetamol helps in reducing the fever, increase in bank rate controls the inflation, can be tested by statistics.

Business Prediction

For the prediction of future sale, we should analyse the sale of previous years, instead of making guess. A knowledge of future trend is very fruitful in making suitable plans for good business.

Presentation of Data

With the help of statistics we can present the data with figures and charts. It would be impossible to any one to recall the increase in per capita income in each decade, but with the help of figures the per capita income growth per decade can be easily remembered.

NATURE OF STATISTICS- A SCIENCE OR AN ART

Statistics as a Science

Statistics is a science or an art is a debatable issue. Science is a systematic body of knowledge which studies cause and effect relationship and tries to find out generalization. It basically explains the facts. The main features of science are:

– It is a systematic study of a subject

– It establishes the cause and effect relationship of a fact

– Its laws are universal in nature.

Keeping various statistical methods in view, statistics may be defined as a science in which we study various methods of collection, editing, classification, tabulation, presentation of facts through diagrams and graphs and condensation of data regarding various social, political and economic problems so that relationship may be established between various facts and the problems may be analysed, interpreted and forecasting may be done.

However, statistics is not an exact science or pure science like physics, chemistry etc. It may be called a science of scientific method. Statistics basically help other sciences to derive their own laws. Statistical knowledge is not for its own sake but for the sake of other knowledges. According to Tippet, "as a science the statistical method is a part of the general scientific methods and is based on the same fundamental ideas and processes."

Statistics as an Art

If science is the knowledge, art is action; an art is an actual application of science. A science teaches us to know and art teaches us to do. An art basically has the following characteristics:

– It is a group of actions to solve a problem.

– It does not describe facts but examines merits and demerits and tells us the way to achieve the objective.

On the basis of above characteristics of an art, statistics may be regarded as an art of applying the science of scientific methods. As an art, statistic facilitates better understanding and solution to the real life problems by way of offering quantitative information. The successful application of the methods of statistics usually depends upon the considerable degree of skill and vast exposure of statistician, who is using the statistical data for solving the real life problem. According to Tippet, "Statistic is both a science and an art. It is a science in that its methods are basically systematic and have general application and an art in that there successful application depends, to a considerable degree, on the skill and special experience of the statistician, and on his knowledge of the field of application."

Scope of Statistics

It is very difficult to mark the scope of statistics because there is hardly any sphere of human activity where statistics does not creep-up. All disciplines whether social sciences or natural sciences or management, are using statistics in one form or another.

Broadly, the scope of statistics can be confined to the following two main aspects i.e.,

– Divisions or classifications of statistics

– Applications of statistics

Divisions or Classification of Statistics

Broadly speaking, statistics may be classified into a few major categories which are discussed below:

1. **Descriptive Statistics**

    Descriptive statistics deals with the numerical data. These data represents observations obtained either through counting or through some measurement process. They may define an entire population or universe and be derived through some sampling procedure. In other event, they represent the raw material of any subsequent statistical processing and may in some cases be sufficient by themselves to answer particular question or problem.

    Descriptive statistics include the techniques such as collection and tabulation of data, averages or measures of central tendency, index number, dispersion, skewness, which help in summarizing and describing the main features of data. It also focuses its attention on characteristics of data which on account of its huge size would have escaped the attention of users.

2. **Applied Statistics**

    This division of statistics is concerned with the application of statistical methods and techniques to the

specific problems and facts of real life situations as they exist. Some of the important techniques which constitute the applied statistics are sample survey, quality control, quantitative analysis for business decisions, index numbers etc. For example, for the purpose of business forecasting such as demand of a particular product in the market, we have to make use of methods like extrapolation or least square which form the contents of applied statistics.

3. **Inferential Statistics**

It covers those methods and techniques of statistics which are used to draw conclusions and statistical inferences about the parameters of population on the basis of estimates derived from a sample. Similarly, the hypothesis in regard to the characteristics of sample is tested on the basis of population parameters. There are several significance tests which are designed for this purpose. These tests lay down the possibility of being true in individual cases. The statistical inferences or conclusions drawn from given data are not merely guesses but they are backed by prior exercise of logical thought and are based on fundamental theories of science of mathematics.

Applications of Statistics

Statistics has developed wide areas of its applications in the modern age of information technology. Some of the important areas of statistical applications are discussed as under:

1. **State**

Statistics is indispensable for functioning of the State. State authorities need different type of facts and figures on various matters to frame policies and guidelines in order to perform its functions smoothly. In earlier days, when statistics was science of statecraft, the object was to collect data relating to manpower, crimes, income, wealth etc. for formulating suitable military and fiscal policies. Presently, with the change in the nature of functions of the State i.e. from maintenance of law and order to promoting human welfare, the scope of application of statistics to State affairs has changed drastically. Today, State authorities collect statistics through their own agencies on multiple aspects i.e., population, AGRICULTURE, national income, defence, science, mineral resources, oceanography, space research, etc. Almost all ministries at Central or State level, depend heavily on statistical information for their smooth functioning. The availability of statistical information helps the government to frame suitable policies and guidelines to improve the overall working of the system.

2. **Economics**

Economics is concerned with the allocation of limited resources among unlimited ends in the most optimal manners. Statistics, provides statistical information to answer the basic questions in the economics i.e. what to produce, how to produce and for whom to produce. Statistical information basically helps to understand the economic problems and the formulation of economic policies. Earlier, in olden days, the applications of statistics to economics was limited because (i) at that time economic theories and principles were based on deductive logic; (ii) the statistical techniques were not developed for its applications in other disciplines. But in the present era of computers and information technology, statistical data and advanced techniques of statistical analysis have proved immensely useful. In economics, now the scholars have shifted their stand from deductive logic to inductive logic to explain any economic proposition. The inductive logic of reasoning requires observation of economic behaviour of large number of units. Thus, it requires a sound statistical support in the form of data as well as techniques.

The use of statistics and statistical techniques can be better examined in the following term:

–   **Test and Verification of Economic Theories/Principles/Hypothesis:** Over the years economists have developed multiple theories and principles based on deductive reasoning in the areas of consumption, production, exchange, distribution, investment, business cycles and taxation etc. All these theories are merely for academic interests only unless they are put to empirical test or verification. Statistics serves this purpose and enables us to compare the theories in real life situations.

–   **Understanding and Study of Economic Problems:** Statistics helps to understand the various economic problems with clarity and precision and also enables to frame suitable policies in relevant areas for better results. For example, income and wealth statistics are helpful to frame policies for reducing disparities of income whereas price statistics helps to understand the problem of inflation and cost of living in the economy.

–   **Economic Planning:** Economic planning is an important aspect of economic policies of a country. An effective economic planning requires sound information base for different components of economy so that planner should be able to face future smoothly. Here, statistics helps to provide data and also tools to analyze for optimal results. Time series analysis, index numbers, forecasting techniques are some of the powerful techniques, which are immensely useful for the analysis of data in economic planning. Statistical techniques are also helping in framing planning models. In India economic models of different five year plans have extensively used the tools of statistics.

–   **Measurement of National Income and Components:** Statistics also enables to study and measure various national income components and their compilations. It collects information on income, saving, investment, expenditure etc. and also establishes relations among them.

3. **Business Management and Industry**

    In modern world, management of business organisation has become a complicated exercise as a result of change in size, technical know-how, quantum of production, number of personnel's employed, and capital employed, workers and increasing level of competition. Management while planning, organizing, controlling and communicating is confronted with alternative courses of action. In the face of uncertainty, management cannot adopt trial and error method. It is here that statistical data and powerful statistical techniques of probability, expectations, sampling, test of significance, estimation theory, forecasting and so on plays an indispensable role. In the words of Chao, "Statistics is a method of decision-making in the face of uncertainty on the basis of numerical data and calculated risks". Statistics thus, provides information to the business units which help in deciding location and size of business, demand forecasting, production planning, quality control, marketing decisions and personnel administration. In Industry, statistics is extensively used in 'Quality Control'.

4. **Social Sciences and Natural Science**

    Statistics has its vast applications in social sciences. In social science, particularly sociology, the most important application is in the field of demography for studying mortality, fertility, marriage, population and growth. In education and psychology, it is used to determine the intelligent quotient (IQ).

5. **Biology and Medicine:** Statistical tools for collection, presentation and analysis of observed data relating to the causes and incidence of diseases are of paramount importance in biology and medical sciences. For instance, the statistical pulse rate, body temperature, blood pressure, heart beats, weight, etc. of the patient greatly help the physician for proper diagnosis of disease. The most important application of statistics in medical science lies in using the test of significance for testing the efficacy of manufacturing drugs or injections or medicines for controlling/curbing specific diseases.

6. **Research:** Statistics is also extensively used to conduct research in new areas and to open new vistas of knowledge to mankind. Experiments include impact of fertilizer on crops, soil etc.

REVIEW QUESTIONS

1. The word 'Statistic' is basically used in two senses, i.e., as : _____ and _____

2. _____ has given most acceptable definition of statistics.

3. Arrange in order in context of statistical enquiry

(a) Interpretation (b) Analysis (c) Collection (d) Organisation and Presentation

Answers: 1. Statistic Data, Statistic Method 2. Prof. Horace Secrist
3. (c)-(d) - (b)-(a)

Laws of Statistics

There are two important laws of statistics on which whole theory of sampling is based. These two laws are:

– Law of Statistical Regularity

– Law of Inertia of a Large Numbers

1. **Law of Statistical Regularity:** This law provides the basis of theory of probability in statistics. According to this law, if a large sample is taken from a population in a random way, it is bound to be fairly representative of the population. Prof. W.I. King has rightly observed in this regard. "The law of statistical regularity formulated in a mathematical theory of probability lays down that a moderately large number of items chosen at random from a very large group are almost sure on the average to have the characteristics of the large group".

This law holds good if it meets the following two conditions.

– random sample is taken

– the size of sample is large

2. **Law of Inertia of Large Numbers**: This law is corollary of the law of statistical regularity. This law states that the larger the size of sample, the more accurate the results are likely to be. It is because large numbers have greater stability, steadiness, consistency as compared to small numbers.

Limitations of Statistics

Statistics in spite of its immense use has certain limitations. Some of the important limitations are listed as under:

(i) *Deals in Quantitative Data:* Statistics deals with quantitative data only and not the qualitative and descriptive facts such as honesty, efficiency, intelligence, blindness, etc.

(ii) *Study of Groups and not Individuals:* Statistics deals with groups and not with individuals. This is a serious limitation of statistics. For example, income of an individual or profit of a particular business unit does not constitute the statistics because those figures are unrelated and incomparable whereas aggregate of figures relating to prices and consumption of various commodities, over different period of time will be statistics.

(iii) *In-exact Science:* Statistical laws are not exact. Their results are true only on averages. They are valid

only under certain set of assumptions. It is because of this, the science of statistics is less exact than natural sciences such as physics, chemistry, etc.

(iv) *Misuse:* Statistics deal with figures which are innocent by themselves and do not bear on their face the level of their quality and can be easily distorted, manipulated by dishonest politicians or unscrupulous people for their selfish motives. It is a delicate subject and dangerous tool in the hands of non-experts. The user of statistical methods should not only have sound knowledge of the subject but also have self-control of an artist. In the words of W.I. King, "Statistics are like clay from which you can make a God or a devil as you please". At another place he opined "science of statistics is the useful servant but only of great values to those who understand its proper use".

DISTRUST OF STATISTICS

Distrust literally means lack of confidence or belief. The statistical statement or science of statistics is always subject to doubt and suspicion to the public because of its misuse by unscrupulous elements for their selfish motive. The common beliefs about statistics are as follows:

– An ounce of truth will produce tonnes of statistics.

– It can prove anything.

– It can prove nothing.

– Figures do not lie; liars make figure.

– Figures are innocent and easily believable.

– There are three type of lies—lies, damn lies and statistics.

– Figure though accurate, might be manipulated by selfish persons to conceal the truth and present a misleading and distorted picture of the facts to general public.

Here, it may be clarified that statistics is only a tool which if rightly used may prove extremely useful and if misused, might be disastrous. Statistics neither approves nor disapproves anything. There is a need for utmost care and precaution in the interpretation of statistical data in all of its manifestations. "Statistics should not be used as a blind man uses a lamp post for support instead of illumination".

COLLECTION AND PRESENTATION OF STATISTICAL DATA

Introduction and Meaning

Collection of facts and figures about a phenomenon is one of the most important functions of statistics. For any statistical enquiry whether it is related to business, management, economics or natural sciences, the basic issue is to collect the facts and figures relating to a particular phenomenon.

Collection of data refers to systematic recording of results either by counting or by enumeration. The collected data should be quantitative i.e. numerical in nature so that we can apply statistical methods in presentation, analysis and interpretation of the same. The entire structure of statistical analysis for any enquiry is based on systematic collection of data.

Once statistician is clear about the objectives and scope of enquiry, statistical units to be used, the degree of accuracy aimed in the final results, the next step in data collection is to decide about the sources of data and type of data to be used.

Broadly speaking, there are two types of data:

– Primary

– Secondary

Primary Data

It is the data, which is collected for the first time by investigators or enumerators to serve a particular purpose. Such a data is original in nature. The sources from where these data can be collected are known as primary sources.

In India, the organizations such as Central Statistical Organisation (CSO), Census of India, National Sample Survey (NSSO), and Reserve Bank of India (RBI) are the organizations, which collect and publish the primary data and so they are primary sources of data. However any researches can collect primary data based on the requirements.

Secondary Data

Secondary data refers to the data which is originally collected and published by the authorities other than who require it. Such data is already available in some government publications, research study, journals or newspapers.

Main sources of secondary data may be classified in the following two categories:

1. **Published Sources:** Published sources of secondary data basically refers to various national/international organization/agencies which collect and publish the statistical data relating to business, trade, labour, price, consumption, production, investment, savings, population, unemployment, banks and financial institutions, corporations etc. These publications offer very useful sources of secondary data. Some of the important sources of secondary data in this category include the following:

   – Publications of the Central and State Governments, of Foreign Governments and international bodies like IBRD, IMF, ADB, ILO, UNO, WTO, WHO etc.

   – Publications of Semi-Government organisation e.g. Reserve Bank of India Bulletin, Reports on Currency and Finance.

   – Publications of CSO and NSSO and other research bodies (ICAR & IARI, Delhi).

    – Publications of various Chambers of Commerce, Trade Associations and Co-operative Societies.

    – Reports of various Committees and Commissions appointed by the Government.

2. **Unpublished Sources:** It covers all those sources of secondary data where records are maintained by private agencies or business firms for their own use and are restrictedly available for use of general public. Data collected by research institutions are also included in the category of unpublished sources of secondary data.

Distinction between Primary and Secondary Data

It is very difficult to distinguish between primary and secondary data. The difference is merely of degree of detachment of original source. The data which is primary for one may be secondary to other. Some of the major points of distinction between the two types of data are listed as under:

Primary Data	Secondary Data
Primary data is first hand information and original in nature.	Secondary data is in the form of compilation of existing data or already published data.
The collection of primary data involves huge resources in terms of money and time, finance and energy.	Secondary data is relatively less costly.
Primary data is usually collected by keeping in mind the purpose for which it is collected so its suitability will be more.	Secondary data may or may not suit the purpose.
Primary data may be used as it is in its original form.	The use of secondary data requires lot of care and precaution.
Primary data are more reliable, accurate and adequate.	Secondary data are not always, reliable, accurate and adequate.

Choice of Data

A proper choice of type of data needed for any statistical investigation basically depends on consideration of various factors such as nature, objective and scope of enquiry, availability of financial resources, time, accuracy expected and the status of agency.

It is observed that presently in large number of statistical investigations, secondary data are generally used because of the availability of large amount of reliable published data from government sources (Central or State), Government agencies, international institutions and organizations, private organizations and research institutions, journals and magazines, etc. But primary data are usually collected under `those situations where secondary data suitable for enquiry under study is either not available or it is not directly suitable for the purpose.

Precautions in the Use of Secondary Data

Secondary data needs utmost precautions in their use because they are collected for different purpose and by different agencies. Prof. Bowley rightly stated in this regard — "It is never safe to take the published statistics at their face value without knowing their meaning and limitations." So, it is essential for the users of secondary data to ensure the reliability, adequacy and suitability of data to the given problem under investigation. Some of the major precautions in this regard include the following:

1. **Suitability of Data:** It is expected that the user of the secondary data should ensure that the data are suitable for the purpose of enquiry. The suitability of data can be judged in the light of the nature and scope of investigation, terms and units defined, etc.

2. **Reliability of Data:** The reliability of secondary data is another precaution, which needs special attention of the users. The users should take note of following points regarding reliability of secondary data:

    – the reliability, integrity and experience of the collecting organisation

    – the reliability of the source of information

    – the methods used for collection and analysis of data

    – the degree of accuracy desired by the company.

In case the data were collected on the basis of sample, the users should satisfy themselves regarding:

    – adequacy of sample

    – use of proper sampling technique

    – collection of data by trained, experienced and unbiased investigator under the supervision of competent officer

    – use of proper estimation techniques to estimate the population parameters.

3. **Adequacy of Data:** Even if the available secondary data are suitable and reliable from the users point of view, it may not be adequate enough for the purpose of the given enquiry. Adequacy of data is to be judged in the light of the requirements of the survey and the geographical area covered by the available data. Adequacy of the available data is also considered in the light of time period. Hence, in order to arrive at conclusions free from limitations and inaccuracies, the secondary data must be subjected to thorough scrutiny and editing before they are accepted for use.

Methods of Collecting Primary Data

Primary data may be collected by any one of the following method:

    – Direct personal interview

    – Indirect oral interview

    – Information received from local agencies or correspondents

    – Mailed questionnaire method

    – Schedule sent through enumerators

    – Telephonic survey

1. **Direct Personal Interview:** Under this method, investigators personally visit the respondents, ask them questions pertaining to the enquiry and collect the desired information. The information or data so collected from such type of interviews is an original one in nature.

Merits:

    – It proves highly accurate and reliable data.

    – It makes available to investigators supplementary information which helps them in the interpretation of results.

Demerits:

    – It is a costly and time consuming method.

    – The data collected under this method may be influenced by subjective attitude of investigators.

    – The success of this method depends on the capability of the interviewer and the manner in which he

conducts the interview to gather the required information or data.

This type of method is suitable only in those cases where scope of enquiry is confined to single locality, area or region and there is no constraint on the utilization of resources like money and time.

2. **Indirect Oral Investigation:** Under this method, investigators collect data by contacting third-party or informants who are supposed to possess the information on the problem under investigation. For example, the information on the gambling, drinking, or smoking habits of an individual can best be obtained by interviewing his personal friends, relatives, neighbors who know him thoroughly well. Usually under this method, a small list of questions is prepared and these questions are put to different informants in the environments and then their answers are recorded. This type of procedure to collect actual data may be adopted by enquiry committee or commissions of enquiry.

Merits:

– It allows to extract information from the parties about the person who is reluctant to supply any information about the subject matter of enquiry if approached directly;

Demerits:

– It is time consuming and costly.

– The information collected under this method is sometimes not reliable.

– The success of this method is directly dependent on several factors, such as knowledge and representative character of informants, biasedness of witness about the matter and ability and sincerity of interviewer to extract information.

3. **Information Received from Local Agencies or Correspondents:** According to this method, the correspondents or local agencies in different regions collect the required information in their own way and subsequently submit their reports or transmit the information to the Central Office where data are processed and finally analyzed. This method of collecting information is usually adopted by newspaper agencies, which require periodical information in areas such as sports, economic trends, share markets, law and order, etc.

Merits:

– This method is economical and expeditious.

– It is ideal method for the situation where data is required on a continuous basis.

Demerits:

– The quality of data collected through this method is highly unreliable because of biasedness of local agents and the different techniques of data collection adopted by them.

– It is not useful for comprehensive and extensive study of a problem.

4. **Mailed Questionnaire Method:** Under this method, a well designed questionnaire is mailed to the informants with a request to fill it up and return the same within the specific time schedule. It is pointed out in the covering letter to the respondents that information supplied by them in the questionnaire will be kept strictly confidential. The investigators send questionnaire along with self-addressed envelope to respondents for quick and better response. In the prevalent knowledge based environment, an investigator may also e-mail the questionnaire to target group of respondent and can mobilize the desired information subject to certain conditions. This type of method is very popular among research workers, private agencies, etc.

Merits:

– It is most economical method in terms of time, money and manpower.

– It may be used more effectively in cases where scope of enquiry is very wide and extensive.

– Direct information from respondents, leaves less scope for personal bias.

Demerits:

– This method does not work where majority of respondents are illiterate.

– There is low degree of reliability of the information supplied by informants.

– This method is not flexible.

The success of this method is based on the ability to design the questionnaire, knowledge of the respondents and their response to questionnaires.

5. **Schedule or Questionnaire sent through Enumerators:** Under this method of enquiry, the trained enumerators personally visits the informants and explain the objectives of enquiry, ask questions and record their replies. This method is commonly used by big business houses, research institutions and large public undertakings.

Merits:

– It enables to extract information from all type of respondents literate and illiterate.

– The information collected by this method is more accurate and reliable.

– This method is ideal for extensive surveys.

– In the case of poor response, the enumerators can personally persuade the informants to supply information.

Demerit:

– It is the most expensive method because the enumerators who collect information are paid workers.

– It is time consuming.

– It is hectic and tiring.

– The success of this method depends on skill of enumerators, object of enquiry, etc.

6. **Telephonic Survey:** Under this method investigator instead of presenting himself before the informants contact them on telephone and collect the desired information. This method is more convenient and less time consuming but subjective in nature.

Merit:

– It is convenient and less time consuming.

Demerits:

– It is subjective in nature.

CENSUS AND SAMPLE INVESTIGATION

Census and Sample Investigation are the two popular approaches used by the statistician to collect the primary data.

Under the census or complete enumeration method, data are collected for each and every unit of the population or universe which is a complete set of items which are of interest in any particular situation. For example if the marks obtained by all the students of B.A. (Hons.) Economics of Delhi University are recorded for analysis it will be called a census investigation. Another example is the case of population Census.

This method is generally recommended in those cases where area of investigation is limited and requires intensive examination or study of the population.

Merits of Census Investigation

The technique of census investigation is advocated by its users because of its various merits including the following:

1. **Intensive Study:** Under census investigation, data are obtained form each and every unit of population. It enables to study more than one aspect of each item of population. For example, in the census of India study which is conducted after every 10 years, data are collected regarding population size, males and females, education levels, handicap member in the family, source of income, caste, religion etc.

2. **Reliable Data:** The data collected under this method is likely to be more reliable, representative and accurate because in census every item is observed personally.

3. **Suitable Choice:** The census method is a suitable choice in the situations when different items of population are not homogeneous.

4. **Basis of Various Surveys:** Data of complete enumeration or census can also be used as a basis of various surveys.

Demerits of Census Investigation

Despite various merits discussed above, the census technique has certain demerits also. Some of these demerits include the following:

1. **Costly Method:** Census is very costly method of investigation since the data are collected by observing each and every item of population. This method is usually adopted by the government organization to collect detailed data such as population census or agricultural census, census of Industrial protection etc.

2. **Time Consuming:** This method is time consuming and requires more manpower to collect the original data.

3. **Possibilities of Errors:** There is large number of possibilities of errors in this method due to non-response, measurement, lack of preciseness of the definition of statistical units or personal bias of the investigators.

Sample Investigation

On the other hand, sample investigation refers to the process of learning about population on the basis of sample drawn from it.

To be more precise, in sampling technique instead of every items of the universe only a part of the universe is studied and conclusions are drawn on the same basis for the entire universe. The basic objectives of collection and analysis of sample investigation is to reveal certain characteristics of population:

– To estimate parameters of population, i.e. means, median and mode etc.;

– To test certain validity of statement about a population;

– To investigate the changes in population over time.

Basically, the sampling process involves mainly following three elements:

– selecting the sample

– collecting the information

– making an inferences about population.

Essentials of Sampling or Sample Investigation

Some of the important essentials of sample investigation include the following:

1. **Representativeness:** A sample is selected in such a way that it should in true sense represent the universe otherwise the results obtained from it may be misleading.

2. **Adequacy:** The size of the sample should be adequate enough to represent the parametric characteristics of the population.

3. **Independence:** There should be independent and at random selection of all the items covered under sample.

4. **Homogeneity:** Homogeneity is another essential element of sample investigation. According to this, there is no basic difference in the nature of units of universe and that of sample.

Merits of Sample Investigation

Some of the important merits of sample investigation include the following:

1. **Less Costly:** Data collection through sample investigation is less costly because in this case the task of data collection is confined only to a fraction of population.

2. **Less Time Consuming:** Sample investigation is less time consuming because it takes less time in collection, analysis and interpretation of data.

3. **Reliability:** Data collected under sample investigation are more accurate and reliable because of use of services of well trained investigators/experts.

4. **Detailed Information:** Since the sample investigation saves both time and money, it is always possible to collect detailed information in sample survey.

5. **Flexibility:** The data collected thorough sample investigation has greater scope of flexibility.

Demerits of Sample Investigation

Despite all the above mentioned merits, sample investigation has various demerits also including the following:

1. It is impossible to attain cent percent accuracy because the conclusions drawn about characteristics of population are based on the results obtained from the selected sample.

2. It creates the new source of error, i.e. 'random error' or 'sampling error'.

3. The results of sampling investigation are not satisfactory unless it uses the services of experts.

4. The sample may not be the representative of population because it depends on the mind set and attitude of the investigators.

5. It is not possible to use sample investigation method when the population possesses the heterogeneous character.

Distinction between Census and Sample

Some of the major points of distinction between the census and sample method of investigation are presented as under:

Attributes	Census Investigation	Sample Investigation
Type and nature of Enquiry	Comprehensive	Selective
Time required	More	Less
Accuracy & Reliability	As high as 100%	Less than 100%
Types of errors	Non sampling	Sampling & Non-sampling
Suitability	For small surveys where the coverage of every item is necessary	For large coverage
Organizational work	Difficult	Easier
Cost	High	Low

PRESENTATION OF DATA

The presentation of statistical data generated from various sources is highly cumbersome and complicated exercise. It is an essential step after data collection to make the data really useful to its users. It broadly covers following aspects:

– Classification of data

– Tabulation of data

– Frequency distribution of data

– Diagrammatic presentations of data

– Graphic representation of data.

I. Classification of Data

Introduction

The data collected is usually very voluminous and large in quantity and as such they are not fit for analysis and interpretation. Hence the first step after collecting data is to classify it.

Classification of data means grouping of related facts into classes. According to Secrist, "classification of data is the process of arranging data into sequences according to their common characteristics or separating them into different but related parts". In the words of AM Tuttle, "a classification is a scheme for breaking a category into a set of parts called classes according to some precisely defined differing characteristics possessed by all the elements of the category". Both the definitions focus on arrangement of data on the basis of certain characteristics or criteria. In simplified words, it is similar to sorting of letters received in the post office into different lots on the basis of geographical areas.

For example, the students registered for Company Secretaryship Course during a financial year may be classified on the basis of any of the following criteria:

– Gender

– Level of Qualifications

– Regions

– Legal/Non-Legal Qualifications

– Professional Qualifications

– Others.

Thus, the same data can be classified in different ways based on different criteria, depending upon the requirements of the users. This process can be compared to sorting of letters in the post office

Advantages of Classification

Some of the advantages of classification of data are listed as under:

– It arranges and presents huge volume of unwieldy raw data in condensed and meaningful form so that it is easily understandable to the users.

– It facilitates meaningful comparison of data.

– It enables us to understand the relationship among variables which are classified on the basis of different criteria.

– It highlights significant features of the data at a glance.

– It makes the statistical treatment of the data easier because of its arrangement in relatively homogeneous groups based on certain criteria.

Principles of Classification

However, there is no hard and fast rule to classify raw data. It depends on the nature of data and the objectives of the study for which data are needed. The following principles may prove useful for sound technical classification of data.

1. **Ambiguity:** There should not be any ambiguity and confusion regarding placing any observation/variable in the given arrangement or classification. It means that each and every item in the data must belong to one of the classes. A good and ideal classification is one which is free from residue class i.e. 'all others' or the rest because such class do not reveal the characteristics of data completely.

2. **Mutually Exclusiveness:** Each item of the data should be mutually exclusive or non-overlapping that is an observed value belongs to one and only one of the classes.

3. **Stability:** The classification of data should follow the principle of stability throughout the period of analysis in order to enable the user to properly carry out the analysis and comparison of data. For example, if the sales of refrigerators are classified by regions then advertising expenditure incurred for promoting the sale should also be classified regions wise.

4. **Conformity to Objectives:** The data should be classified in such a way that it conforms to the objectives of the enquiry. For example, if we are interested in studying the relationship between company secretary qualification and gender then it is meaningless to classify the students registered for Company Secretaryship course according to age and region because it will not serve any purpose.

5. **Flexibility:** Flexibility is an essential attribute for ideal classification so that it should be adjustable to new and changing conditions/circumstances. No classification is permanent to be used forever

Basis of Classification

Classification of data depends upon characteristics of data, and objectives of the enquiry.

Broadly speaking, the data may be classified on the following four bases:

– Geographical i.e. in relation to place, area or region

- Chronological i.e. on the basis of time

- Qualitative i.e. on the basis of some character or attribute

- Quantitative i.e. in terms of numerical values or magnitudes

Geographical Classification: Under this classification, data are classified according to geographical or location wise i.e., area wise, zone wise, region wise, state wise or country wise. For example, registration of students for Company Secretaryship course during the month of May 2014 and 2015 are given in the following tables.

Students Registration for CS Course		
Region	May 2014	May 2015
Eastern	355	319
Northern	418	438
Southern	522	403
Western	318	222
Foreign	2	1
TOTAL	**1615**	**1383**

This type of classification is generally listed in alphabetical order or in some other form depending upon the size or values as the case may be.

Chronological Classification: Under this type, data are classified on the basis of differences in time. For example, we may present the growth of population in India over different years as under:

Year	Population (Crores)
1891	23.6
1901	23.6
1911	25.2
1921	25.1
1931	27.9
1941	31.9
1951	36.1
1961	43.9
1971	54.8
1981	68.3
1991	84.4
2001	110.0

This type of classification is widely used in understanding the economic and business problems and to analyze their trend values over a period of time.

Qualitative Classification: In this type of classification, data are classified on the basis of certain qualitative/ descriptive characteristics or presence and absence of attributes, such as intelligence, beauty, honesty, occupation and gender, etc. This type of classification may be simple or multiple. In simple classification, the data are classified

into two categories i.e., presence and absence of attributes. This form of simple classification would be as follows:

Population

Blind   Non-blind
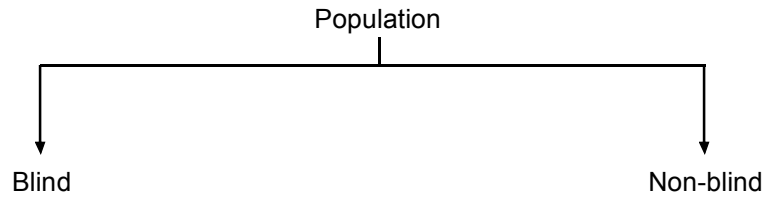
However, in multifold classification the given data are classified into more than two categories.

For example, the problem of blindness among students is being studied on the basis of gender and their smoking habits. The classification of data in such a case would be as under:

Example of Classification of Data
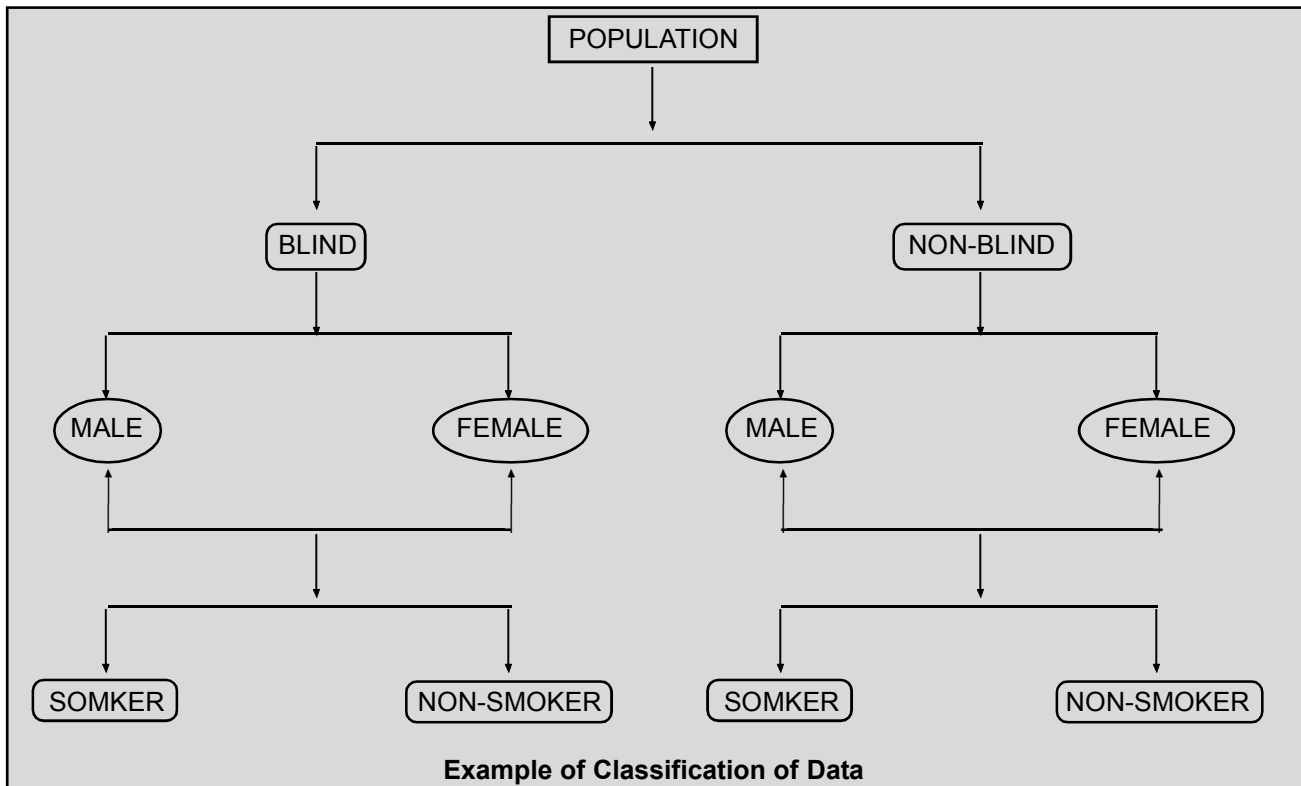
Quantitative Classification: It refers to the classification where data are classified on the basis of figures or criteria which are capable of quantitative measurement such as height, age, weight, income, expenditure, number of marks obtained by students etc. For example, marks obtained by 100 students of B.Com. (Hons.) in the statistics paper may be classified as under:

Marks obtained by B.Com Students	
Marks in Statistics	**Number of Students**
0-20	15
20-40	15
40-60	45
60-80	20
80-100	5

In the above mentioned table, the marks of the students is the variable and number of students is the frequency.

Variable means a characteristic which varies. A variable may be discrete or continuous. A discrete variable is one which does not assume fractional values within the range of numbers whereas continuous variable is one which is capable of manifesting every conceivable fractional value within the range of fractional possibilities.

II. Tabulation of Data

Introduction

Tabulation refers to systematic and logical arrangement of data in columns and rows in accordance with some salient features and characteristics. According to A.M. Tulte, 'A statistical table is the logical listing of related quantitative data in vertical columns and horizontal rows of numbers with sufficient explanatory and qualifying words, phrases and statements in the forms of titles, heading and explanatory notes to make clear the full meaning of data and their origin. In the words of Secrist, "Table are means of recording in permanent form the analysis that is made through classification and by placing in just opposition things that are similar and should be compared".

Thus, it is clear from above two definitions that tabulation is a systematic and logical presentation of data such that it is easily understandable and comparable. It attempts to furnish the maximum information in the minimum possible space without sacrificing the quality and usefulness of data. The basic objectives of tabulation are to summarize the information in such a fashion that it can be easily analyzed compared and interpreted.

Significance of Tabulation of Data

Some of the major points of significance of Tabulation of data are discussed as under:

1. Presentation of Complex Data in Systematic and Simplified Form: Tabulation presents the complex raw data in systematic and simplified form so that it is easily understandable to a common man without wasting much time.

2. Comparison: Tabulation facilitates meaningful comparison of data because of division of table into parts, sub-parts, columns and rows.

3. Identification of Data: Tabulation enables easy identification of data as source of reference because data are arranged in tables with specific titles and numbers.

4. Economy in Space and Time: Tabulation also eliminates unnecessary and irrelevant details and presents the meaningful information only. In this way it saves time and space both.

5. Discloses Trend and Pattern of Data: Tabulation discloses the trend and pattern of data which cannot be otherwise understood if the presentation is descriptive.

6. Facilitates Statistical Analysis: Tabulation facilitates statistical analysis easier because of systematic presentation of data. It enables to make use of various statistical techniques such as averages, dispersion, correlation or time series, when data are tabulated in systematic fashion.

Essential Parts of a Table

Before going into details of the techniques of tabulation, it is useful to discuss about the essentials of a table. A table consists of statistical data systematically arranged in rows and columns. The actual form and structure of a table depends upon the nature of data tabulated and the purpose of study under investigation.

However, a table should contain the following parts:

 – Table Number

 – Title of the Table

 – Sub-title or Head Note

 – Captions and stub

 – Body

 – Footnotes

 – Source Note

1. Table Number: Each and every table should be serially numbered so that it becomes easier for reference purposes in future. The number of a table may be either in the centre at the top of title or in the side of title at the top or in the bottom of the table at left hand side.

2. Title of the Table: It is a must for every table to have a suitable title which usually appears at the top of table. The title should be brief, precise, clear and self-explanatory so that the reader or user gets clear idea about the contents of table at a glance. Title should be written in phrases and in bold letters.

3. Sub-tile or Head Note: A Sub-title or head notes refer to brief or short statement for explaining any point or points not included in the heading. It is placed below the title centered and is enclosed in brackets. The head notes are applicable for entire tables. For example, unit of measurements are always expressed as head notes such as 'rupees in thousands' or 'area in hectares' or 'quantity in liters' or in 'tonnes' etc.

4. Captions and Stubs: Captions and stubs are heading for the vertical columns and horizontal rows in a table respectively. They should be brief, concise and self-explanatory. Captions are written in the centre of the columns in small letters to economize the space whereas stubs are written at the extreme left of the row. The stubs are kept as narrow as possible without loosing precision and clarity of statements.

5. Body of the Table: The body constitutes an important part of the table. It contains the numerical information which is presented to the users. The information is shown in columns and rows arranged according to captions and stubs.

6. Footnotes: Footnotes are used to clarify anything which is not otherwise clear by the heading, title, captions, stub etc. Footnote is given at the bottom of a table. The footnotes basically supply information on any of the following points:

 – to point out any specific detail of data

 – the sources in the case of secondary data

 – limitation of data if any

 Footnotes may be identified by members say 1, 2, 3 ... or letters i.e. a, b ..... or symbols *, **/! etc.

7. Source Note: Below the footnotes or below the table, the source note of the data may be mentioned for the verification of the reader. The source note is required if the secondary data is used. In case data are taken from a research journal or periodical, then source note should contains the details such as name of the journal/periodical, date of its publication, volume number, table number if any, page number etc. It may help the users to satisfy himself about the accuracy of data.

A specimen table, containing all parts are given as under, in order to understand it clearly.

Table		
TITLE		
		HEAD NOTE IF ANY
STUB	CAPTIONS	TOTAL
STUB		
ENTRIES		ROW TOTAL
TOTAL	COLOUMN TOTAL	GRAND TOTAL
FOOTNOTES/ SOURCE NOTES		

Classification of Table

Tabulation of data may be classified on different basis depending upon the objectives, nature and the coverage of enquiry. Broadly speaking, it may be classified into following ways:

1. *Simple and Complex Table:* Simple Tabulation of data refers to the table where data are classified on the basis of single characteristics. It is also known as one way table whereas in the case of complex table manifold table data are classified and presented with respect to two or more characteristics simultaneously. The Table 1 & 2 respectively will illustrate both the types of tables.

TABLE 1	
India's Gross National Product at Factor Cost	
	(Rupees in thousand Crores)
YEAR	**GROSS NATIONA PRODUCT** **(at Current Prices)**
1990-91	503.5
1994-95	903.9
1995-96	1059.8
1996-97	1230.5
1997-98	1376.8
1998-99	1601.1
1999-2000@	1771.0
@ Quick Estimates	
Source: Economic Survey 2000-2001	

The above table is a simple one, because it furnishes the data about one single characteristics of Indian economy i.e. Gross National Product. Table 2 is an example of the complex table which reflects two or more than two characteristics. In this case, it shows, India's export, import and trade balance over the years, i.e. from 1990-91 to 2000-2001.

Here, it may be clarified that as the order of table increases, the table becomes more and more confusing and complicated.

TABLE 2			
India's Export, Import and Trade Balance			
			(Rupees in Crores)
YEAR	Exports	Imports	Trade Balance
1990-91	32553	43198	-10645
1991-92	44041	47851	-3810
1992-93	53688	63375	-9687
1993-94	69751	73101	-3350
1994-95	82674	89971	-7297
1995-96	106353	122678	-16325
1996-97	118817	138920	-20103
1997-98	130101	154176	-24075
1998-99	139753	178332	-38580
1999-2000	162925	204583	-41658
2000-01(P)	129460	157256	-27796
(P) Provisional Estimates			
Source: Ministry of Finance, Government of India, Economic Survey 2000-2001			

2. *General Purpose or Reference Table and Specific Table:* General purpose tables are those tables, which serve general purpose and provide reference to the users. The basic purpose of this type of table is to present the data in such a fashion that individual item may be found readily by a reader. In this type of tables, we include the tables published by government or government agencies such as tables included in Economic Survey, Statistical Abstract of India, RBI Bulletin etc. On the other hand, specific purpose tables are those tables which are of analytical nature and prepared with the idea of making comparative studies and studying the relationship and significance of the figures provided by the data. Such type of tables uses interpretative figures such as ratio, percentage, in order to facilitate comparison. These tables are called derivative tables since they are derived from the general tables.

General Guidelines and Rules of Tabulation

There is no hard and fast rule to prepare a table. Basically, it depends on the nature of data and the objectives of the enquiry, and the skill and common sense on the part of investigator. In construction and tabulation common sense is the chief requisite and experience is the chief teacher. Tabulation of data is basically an art of presenting the data in a meaningful manner. Some of the general rules which may be followed in the construction of tables are given as follows:

– The table should be compact, easily understandable and free from confusions.

– The table should be designed keeping in view the size of paper to be used.

– There should be logical placement of items in the table to avoid confusion.

– All items should be clearly stated.

– The table should be complete and self-explanatory.

– Avoid the use of Ditto marks (").

– Proper care should be taken for ruling and spacing between different rows and columns.

– The unit of measurement should be clearly stated in the form of head note.

Requisite of a Good Statistical Table

A good statistical table should possess the following requisites:

– A statistical table should aim at simplicity. Too much of detail and complexities should be avoided. But at the same time, it should be complete in itself and give explanation wherever necessary.

– It should have a clear, comprehensive and explanatory title. If necessary, there may be sub-title also.

– The stub and caption should be clear and concise and should be arranged in a way so that the objectives, whether a general purpose table, special purpose table, are fully achieved. The rows and columns can be lettered or numbered to facilitate reference.

– The head note should be clear and should indicate the units.

– Totals and sub-totals should be shown in the table to indicate the overall total and sub-totals of various columns as also to indicate that all items have been accounted for.

– The reference if any should be noted clearly at the bottom of the table.

– Whenever necessary the derived statistics should be shown in the table.

– A table should have balanced length and breadth.

– Use of abbreviation should be avoided.

– Greater clarity can be achieved by the use of variations in column width, width of lines, and faces of types (bold, italics etc.).

– There should be consistency in ruling.

– Zero should not be used to indicate the information that is not available. This can be well indicated by not available.

The following illustrations explain the above points.

Illustration 8.1

Present the following information regarding the employees of a factory in a suitable tabular form:

In 1985 out of a total of 3,500 workers of a factory 2,400 were members of a trade union.

The number of women employed was 400 of which 350 did not belong to a trade union. In 1990 the number of union workers increased to 3,160 of which 2,580 were men. On the other hand, the number of non-union workers fell down to 416 of which 360 were men.

In 1995 there were 3,600 employees who belonged to a trade union and 100 who did not belong to a trade union. Of all the employees in 1995, 600 were women, of whom only 16 did not belong to trade union.

Solution

Table showing the Gender-wise Distribution of Union and Non-Union Members in a Factory									
Category	1985			1990			1995		
	M	F	Total	M	F	Total	M	F	Total
Members	2350	50	2400	2580	580	3160	3016	584	3600
Non Members	750	350	1100	360	56	416	84	16	100
Total	3100	400	3500	2940	636	3576	3100	600	3700

M - Males; F - Females

Illustration 8.2

In a trip organised by a Local College of Delhi University, there were 80 persons, each of whom paid ₹150 on an average. There were 60 students each of whom paid ₹160. Members of the teaching staff were charged at a higher rate. The number of servants was 8 (all males) and they were not charged anything. The number of ladies was 30% of the total of which two was lady staff member.

Tabulate the above information.

Solution:

Table showing the Type of Participants, according to Gender and Contribution Made					
Type of Participants	Gender			Contribution per member (₹)	Total Contribution (₹)
	Males	Females	Total		
Students	38	22	60	160	9,600
Teaching Staff	10	2	12	200	2,400
Servants	8	—	8	—	—
Total	56	24	80	—	12,000

Notes :

Total contribution = Average contribution X No. of persons who joined the trip

$$= 150 \times 80 = ₹12,000$$

Contribution of the staff per head has been obtained by deducting the contribution of students from the total and dividing the difference by the number of teaching staff, i.e.

$$= \frac{(12,000) - (60 \times 160)}{12} = \frac{12,000 - 9,600}{12} = \frac{2,400}{12} = ₹\ 200$$

Illustration 8.3

In 1990 out of total 2400 students in a college in a metropolitan city, 1800 were for graduation and the rest for post graduation. Out of 1800 graduate students 600 were girls. However, in all there were 800 girls in the college. In 1995, the number of graduate students increased to 2000 out of which 700 were girls but the number of PG students falls to 500 of which only 400 were boys. In 2005, out of 1000 girls, 800 were for graduation where as total number of graduates were 2500. The number of boys and girls in PG classes were equal.

You are required to present the above information in tabular form and calculate the percentage increase in number of graduate students in 2005 as compared to 1990.

Solution:

Year	Graduate Students		PG Students		Total
	Boys	**Girls**	**Boys**	**Girls**	
1990	1200	600	400	200	2400
1995	1300	700	400	100	2500
2005	1700	800	200	200	2900

Percentage increase in graduate students in 2005 as compared to 1990

= 700/1800 x 100 = 38.8%

REVIEW QUESTIONS
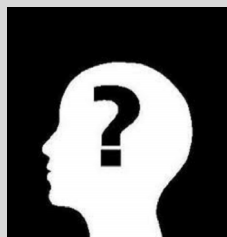
1. _____ refers to systematic recording of result either by counting or by enumeration.

2. Simple tabulation of data is based on _____ characterstics of data.

3. _____ tables are derived from general tables.

Answers: 1. Collection of data 2. Single 3. Derivative

III. Frequency Distribution of Data

Introduction

A tabular arrangement of data with corresponding frequency is known as frequency distribution. In other words, it is a specification of the ways in which the frequencies of the different elements of the population are distributed according to the size of variables under consideration.

To be more precise in the words of Croxton and Cowdon "Frequency distribution is a statistical table which shows the set of all distinct values of the variables arranged in the order of their magnitude either individually or in groups with their corresponding frequencies side by side". Thus ,a frequency distribution consists of two parts one part shows magnitude of values whereas other part shows number of times a value or group of values has repeated.

Classification of Frequency Distribution

Broadly speaking, a frequency distribution may be classified into the following two categories:

– Ungrouped frequency distribution of discrete frequency distribution.

– Grouped continuous frequency distribution or continuous frequency distribution.

1. **Ungrouped or Discrete Frequency Distribution:** Under discrete frequency distribution, we count the number of times each value of a variable is repeated and the number (technically known as tally) is placed alongside the value to which it corresponds. (See Illustration)

Illustration 8.4

From the given information of number of rooms in house in a locality, construct the discrete frequency distribution.

1	3	2	2	2	2	1	2	1	1	1	3
2	2	3	3	3	3	3	3	2	3	4	1
3	2	6	2	1	6	2	1	5	1	5	2
1	5	2	3	4	2	7	4	2	3	3	4
4	3	4	4	2	3	1	4	3	2	2	1

Solution:

In this case, the variable is the number of rooms in each house and since it assumes only integer value so it is a discrete variable. The above information can be condensed in the form of discrete frequency distribution as under:

FREQUENCY TABLE		
No. of Rooms in the house(X)	**Tally Bars**	**Frequency (f)**
1	~~HHT~~ ~~HHT~~ //	12
2	~~HHT~~ ~~HHT~~ ~~HHT~~ ///	18
3	~~HHT~~ ~~HHT~~ ~~HHT~~ /	16
4	~~HHT~~ ///	8
5	///	3
6	//	2
7	/	1
TOTAL		**60**

In the above table X shows the number of rooms and (f) the frequency i.e. availability of rooms in the house. Literally bar is used to count the frequency.

The above distribution is discrete frequency distribution because every X value is associated with corresponding (f) values i.e. $X_1 = 1$, $f_1 = 12$; $X_2 = 2$, $f_2 = 18$, $X_3 = 3$ and $f_3 = 16$, $X_4 = 4$, $f_4 = 8$, $X_5 = 5$, $f_5 = 3$, $X_6 = 6$, $f_6 = 2$, $X_7 = 7$, $f_7 = 1$

2. **Grouped Frequency Distribution:** It is the distribution where the frequency refers to group of values. It is a tabular arrangement of data by classes together with the corresponding frequencies. This type of distribution is prepared when some variable assumes any fractional values particularly height and weight of the students.

The following table on frequency distribution of weight of 100 male students at Delhi University will help to understand the above point more clearly:

FREQUENCY TABLE		
Weights (X)	**Tally Bars**	**Frequency (f)**
40-50	ЖЖ	5
50-60	ЖЖ ЖЖ ЖЖ ///	18
60-70	ЖЖ ЖЖ ЖЖ ЖЖ ЖЖ ЖЖ ЖЖ ЖЖ ЖЖ	42
70-80	ЖЖ ЖЖ ЖЖ ЖЖ ЖЖ //	27
80-90	ЖЖ ///	8
Total		**100**

Basic Terminology Used in the Construction of Grouped Frequency Distribution

The basic terminology which is frequently used in the construction of grouped frequency distribution includes the following aspects.

Class Limit: The end numbers or the highest and lowest values that can be included in a class interval are known as the class limits of that class. For example, in above table 40-50 and 80-100 are the lower and upper class limits.

Class Interval: It is the difference between the upper limit and lower limit of the same class. The lower limit of a class is usually represented by symbol l_1 and upper limit by l_2.

$$\text{Class Mid Point} = \frac{\text{Upper Limit} + \text{Lower Limit}}{2}$$

Class Frequency: The number of observations included in a particulars class is known as the frequency of that class.

Exclusive Classification: Exclusive classification is the one whose class intervals are so determined that upper limit of one class is the lower limit of the next class. Example of this type of classification is as below:

Wages of workers (₹)	Workers(f)
500-600	35
600-700	30
700-800	20
800-900	10
900-1000	5
Total	**100**

The basic point of confusion in this type of classification is on placing the variable having upper unit value. For example, a worker getting ₹600 would be placed in the class of ₹500-600 or ₹600-700. The normal convention in this case is to include the value equal to upper limit in the next class. A more precise way to present the data is as under:

Wages of workers (₹)	Workers(f)
500 and above but below 600	35
600 and above but below 700	30
700 and above but below 800	20
800 and above but below 900	10
900 and above but below 1000	5
Total	100

Inclusive Classification: It refers to that classification where both the class limits are included in the class itself while determining the class intervals. The following example will illustrate this type of classification:

Marks	Students
1-10	4
11-20	5
21-30	6
31-40	10
41-50	5
Total	30

In this case, both the class limits lower as well as upper are included in the class interval. For example, marks 1 and 10 both are included in the class interval 1-10.

Class Boundaries: Under a situation where group frequency distribution is having a gap between upper limit of one class and lower limit of the next class, there arises a need of obtaining class boundaries. These class boundaries are obtained as under:

Lower Class Boundary = Lower class limit - 1/2.d

Upper Class Boundary = Upper class limit + 1/2.d

Where, d is known as the correction factor and it is the difference between upper class limit of any class interval and lower class limit of next class interval. The following example helps to understand the above point.

Class Limit (Wages)	Class Boundaries	Workers
6-10	5.5 - 10.5	4
11-20	10.5 - 20.5	5
21-30	20.5 - 30.5	6
31-40	30.5 - 40.5	10
41-50	40.5 - 50.5	5
Total		30

It may be noticed in the above table that upper boundary of one class is the same as the lower boundary of the next class. Basically, it is the process of making the distribution continuous.

Un-equal Class Intervals: In this type of cases, Class intervals for different classes are different. Following example illustrates this type of case:

Marks	Students
1-10	4
10-15	4
15-25	12
25-40	5
40-50	5
Total	**30**

Open End Classes: An open end class refers to those situations where the last class has no upper limit and the first class has no lower limit. The necessity of open end classes arises when there are few very high values or low values which are far apart from the majority of observation. The following example illustrates the case of open end classes.

Marks	Students
Less than 50	9
50-100	4
100-150	3
150-200	2
200 and above	2
Total	**20**

However, the use of open end classes creates difficulties in finding mid-values while analysing grouped frequency distribution. The best way to overcome this type of problem is to avoid the use of open end classes as far as possible.

Class Width or Size of Class Interval: It is the difference between the lower and upper class boundaries. If all class interval or frequency distribution has equal width then the width is denoted by C. It is generally preferable to have classes of equal width. Unequal width is resorted when some classes are over crowded and some have practically no frequency.

Frequency Density - It refers to the frequency of a class per unit of width and indicates concentration of frequency in the class.

$$\text{Frequency Density (FD)} = \frac{\text{Class Frequency}}{\text{Width of Class}}$$

General Rules for Constructing a Frequency Distribution

1. Determine the largest and smallest numbers in raw data and thus find the range (difference between largest and smallest numbers).

2. Divide the range into a convenient number of class interval having the different sizes or open end classes. The number of class interval depends on the number of items to be classified, magnitude of data and the level of accuracy. Class intervals are chosen in such a way that class marks or mid-points coincide with actually observed data. This tends to reduce the so called grouping error involved in further mathematical analysis. However, class boundaries should not coincide with actually observed data.

3. Determine the number of observations into each class intervals i.e. find the class frequencies. This is best done by using tally marks.

Determination of Number of Classes and Class Intervals

1. Determination of Number of Classes: In a group frequency distribution, the determination of number of classes/class interval is very tedious task. Generally, there is no hard and fast rule to decide the number of classes. Practically, the number of classes should be neither large nor too small. If the number of classes are large, some of the groups may have very small or no frequencies. This might give irregular pattern of frequencies in different class thus making the frequency distribution irregular. On the other hand, if the number of classes is small, large number of frequencies may get concentrated in small number of groups. This obscures some of the important features and characteristics of data, and thus results in loss of information. The number of class basically depends upon the total frequency, the nature of data, accuracy desired and case of computation. However, there is need to keep a balance between these two factors i.e. irregularity of frequency distribution and loss of information in order to arrive at the optimum number of classes. The number of classes is inversely proportional to magnitude of class interval.

   Prof. Sturges has suggested a rule (formula) in order to decide the approximate number of classes. The rule is as under:

   $K = 1 + 3.322 \log 10 \ N$

   Where, K = Number of class intervals

   N = Total Number of observations in data

   The value obtained by the above formula is rounded to the next higher figure.

   For example, If $N = 10$

   then, $K = 1 + 3.322 \log 10$

   $= 1 + 3.322 \times 1$

   $= 1 + 3.322 = 4.322 \sim 4$

   If $N = 100$,

   then $K = 1 + 3.322 \log_{10} 100$

   $= 1 + 3.322 \times 2$

   $= 1 + 6.644 = 7.644 \sim 8$

   Struges Rules very ingeniously restrict the number of classes 4 and 8 which is fairly reasonable number from practical point of a view. This rule does not work if the number of observations is very large or small.

2. Determination of Size of class: Like the number of classes, the magnitude of class intervals should be optimal and reasonable. Since the number of classes are inversely proportional to the magnitude of class intervals in a given distribution, the approximate value of the magnitude (width) of the class interval can be obtained by using Struges rules as udner:

   $$\text{Magnitude of class (i)} = \frac{\text{Range}}{\text{Number Of Class}}$$

   Where, $\text{Range} = X_{max} - X_{min}$

   = Highest value in distribution— Lowest value in distribution

   $$i = \frac{H - L}{1 + 3.322 \ \text{Log} 10 \ N}$$

It may be pointed out, from practical point of view that it is desirable to take class interval of equal or uniform magnitude, throughout the frequency distribution. This will facilitate the computation of various statistical measures and also result in meaningful comparison between different classes and different frequency distribution.

Cross Tabulations

Cross tabulation is a useful data analytic technique. Data referring to a particular variable can be broken down by reference to one or more other variables through successive cross tabulations. For example in an opinion poll, instead of simply providing a table for the entire sample of who votes for a particular party, crosstabulation would permit the breakdown of voting preferences by other variables such as gender, class, trade union affiliation, etc.

In Statistics, a contingency table (also known as a cross tabulation or crosstab) is a type of table in a matrix format that displays the (multivariate) frequency distribution of the variables. They are heavily used in survey research, business intelligence, engineering and scientific research. They provide a basic picture of the interrelation between two variables and can help find interactions between them. Suppose that we have two variables, sex (male or female) and handedness (right or left handed). Further suppose that 100 individuals are randomly sampled from a very large population as part of a study of sex differences in handedness. A contingency table can be created to display the numbers of individuals who are male and right handed, male and left handed, female and right handed, and female and left handed. Such a contingency table is shown below.

Handedness/Gender	Right Handed	Left Handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

The above table allows us to see at a glance that the proportion of men who are right handed is about the same as the proportion of women who are right handed although the proportions are not identical. If the proportions of individuals in the different columns vary significantly between rows (or vice versa), we say that there is a *contingency* between the two variables. In other words, the two variables are *not* independent. If there is no contingency, we say that the two variables are *independent*.

IV. Diagrammatic Presentation of Data

Diagrams are another important method to represent the statistical data. They are nothing but geometrical figures such as lines, bars, squares, rectangles, circles, cubes, etc.

Diagrammatic presentation of data provides an easier method of understanding.

Advantages of Diagrammatic Presentation

Some of the advantages of diagrammatic presentation include the following:

1. *Easy to Understand:* Data presented in the form of diagram can be understood even by a common man. Due to their attractive and impressive character, the diagrams are very frequently used by various newspapers and magazines for the 'explanation of certain phenomena'. Diagrams are also useful in modern advertising campaign.

2. *Simplified Presentation:* Diagrams are used to represent a huge mass of complex data in a simplified and intelligible form.

3. *Reveals Hidden Facts:* A diagrammatic presentation may bring out certain hidden facts and relations which may not be revealed from classification and tabulation of data.

4. *Quick to Grasp:* The data, when presented in the form of diagrams, require less time to understand it.

5. *Easy to Compare:* The data, when presented in the form of diagrams, are easy to compare.

6. *Universal Acceptability:* The diagrams have universal applicability. These are used almost in every field of study like economics, business, administration, social institutions etc.

Limitations of Diagrammatic Presentation

In spite of the above advantages of diagrams, their usefulness is some what limited. One has to be very careful while drawing conclusions from diagrams. The main limitations are as under :

1. *Provides Vague Ideas:* Diagrams give only a vague idea of the problem which may be useful for a common man but not for an expert who wishes to have an exact idea of the problem.

2. *Limited Information:* The information given by diagrams vis-à-vis classification and tabulation is limited.

3. *Low Precision:* The level of precision of values indicated by diagrams is very low.

4. *Restricts Further Analysis of Data:* Diagrams do not lead to any further analysis of data.

5. *Portray Limited Characteristics:* Diagrams can portray only limited number of characteristics. Larger the number of characteristics, the more difficult is to understand those using diagrams.

6. *Possibility of Misuse:* Diagrams are liable to be misused for presenting an illusory picture of the problem.

7. *Fail to Present Meaningful Look in Certain Situations:* Diagrams do not give a meaningful look when various measurements have wide variation.

8. *Careful in Use:* Diagrams drawn on a false base line should be analyzed very carefully.

General Principles of a Diagrammatic Presentation

A diagrammatic presentation is a simple and effective method of presenting the information contained in statistical data. The construction of a diagram is an art, which can be acquired only through practice. However, the following rules should be observed in their construction, to make them more effective and useful tool of understanding a given problem:

1. Every diagram must have a suitable title written at its top to convey the subject matter in brief and unambiguous manner. The details about the titles, if necessary, should provided below the diagram in the form of a footnote.

2. A diagram should be constructed in a manner that it has an immediate impact on the viewer. It should be neatly drawn and an appropriate balance should be maintained between its length and breadth. A diagram should neither be too small or too large. Various aspects of the problem may be emphasized by using different shades or colors.

3. Diagrams should be drawn accurately by using proper scales of measurements. Accuracy should not be compromised to attractiveness.

4. The selection of an appropriate diagram should be carefully done keeping in view the nature of data and the objective of investigation.

5. When a diagram depicts various characteristics distinguished by various shades and colors, in index explaining these should be given for clear identification and understanding.

6. The source of data must also be indicated particularly when it has been from a secondary source.

7. As far as possible, the constructed diagram should be simple so that even a layman can understand them without any difficulty.

Types of Diagrams

There are a number of diagrams which can be used for the presentation of data. The selection of a particular diagram depends upon the nature of data, objective of presentation and the ability and experience of the person doing this task. Some popular types of diagrams are discussed below.

1. **Line Diagram:** In case of a line diagram, different values are represented by the length of the lines, drawn vertically or horizontally. The gap between successive lines is kept uniform. Although the comparison is easy, the diagram is not very attractive. This diagram is used when the number of items is very large.

The following example enables to understand.

Illustration 8.5

The income of 10 workers in a particular week was recorded as given below. Represent the data by a line diagram.

S.No. of workers	1	2	3	4	5	6	7	8	9	10
Income (in ₹)	240	350	290	400	420	450	200	300	250	200

Solution :
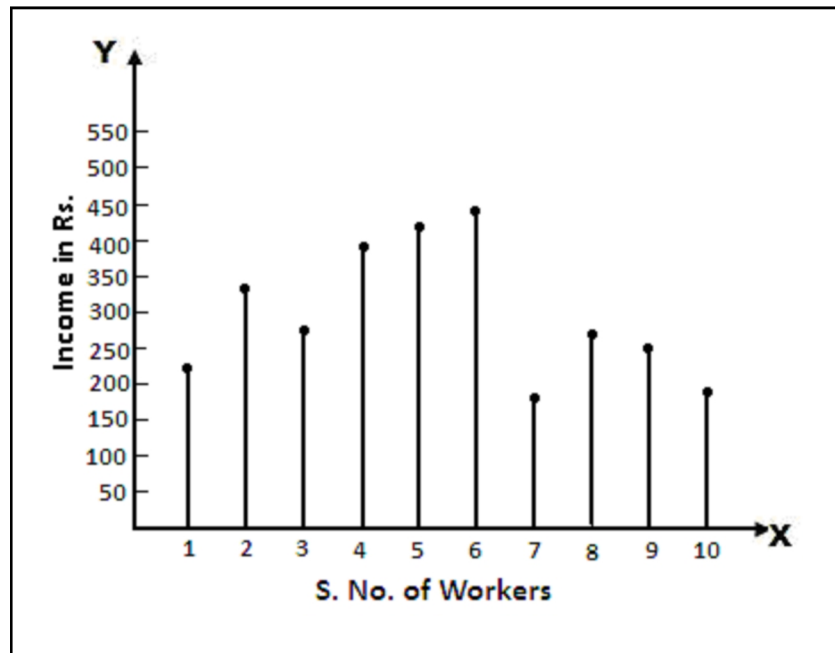
Line Diagram of Income of 10 workers

2. **Simple Bar Diagram:** In case of a simple bar diagram, the vertical or horizontal bars, with height proportional to the value of the item, are constructed. The width of a bar is chosen arbitrarily and is kept constant for every bar. The gaps between succesive bars are also kept same. Bar diagrams are particularly suitable for representing individual time series or spatial services.

Illustration 8.6

Represent the following data by a bar diagram :

Years :	1993-94	1994-95	1995-96	1996-97
Coffee Exports ('0000 tons) :	13.67	13.73	17.06	18.12

Solution :
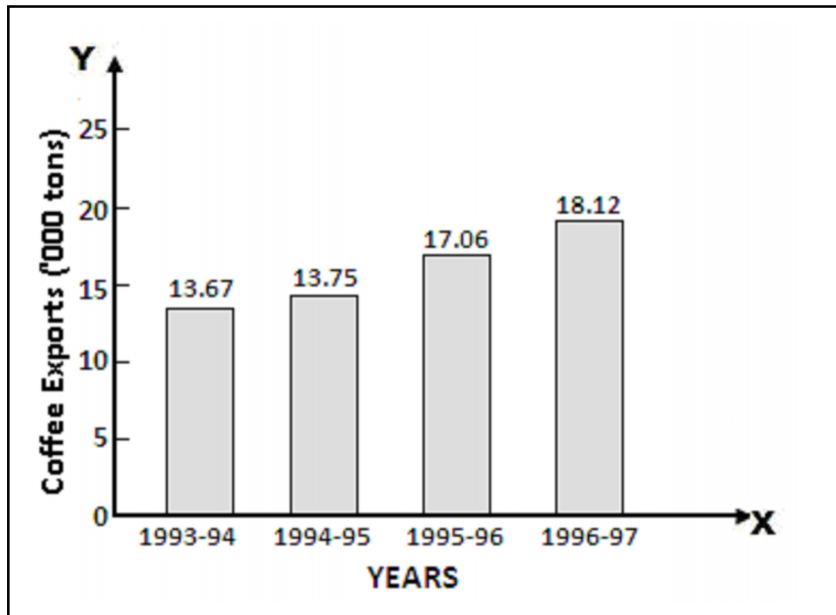
Bar Diagram of Coffee Exports

3. **Multiple Bar Diagram:** This diagram is also known as compound bar diagram. It is used when we want to show comparisons between two or more sets of data. A set of bars for a period or related phenomena are drawn side by side without gaps while various sets of bars are separated by some arbitrarily chosen constant gap. Different bars are distinguished by different shades or colors.

Illustration 8.7

Represent the following data on faculty-wise distribution of students by a multiple bar diagram :

College	Students		
	Arts	Science	Commerce
A	1200	600	500
B	1000	800	650
C	1400	700	800
D	750	900	300

Solution :
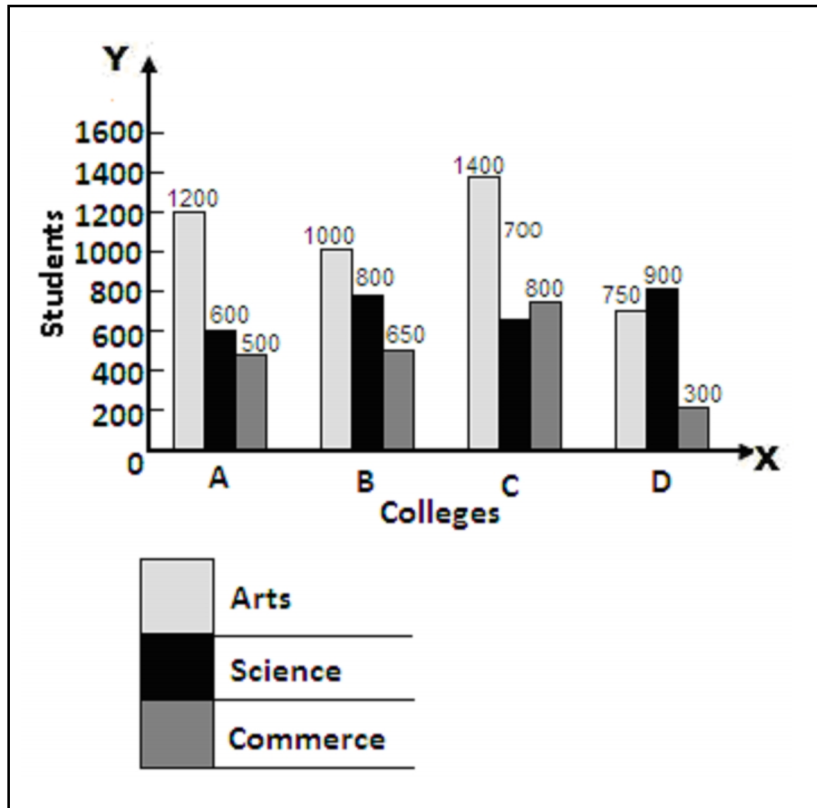
Multiple Bar Diagram

4. **Sub-divided or Component Bar Diagram:** In this diagram, the bar corresponding to each phenomenon is divided into various components. The portion of the bar occupied by each component denotes its share in the total. The sub-divisions of different bars should always be done in the same order and these should be distinguished from each other by using different shades or colors. A sub-divided bar diagram is useful when it is desired to represent the comparative values of different components of a phenomenon.

Illustration 8.8

The following table gives the value of (₹ in Crores) of contracts secured from abroad, in respect of Civil Construction, industrial turnkey projects and software consultancy in three financial years. Construct a component bar diagram to denote the share of an activity in total export earnings from the three projects.

Years	1994-95	1995-96	1996-97
Civil construction	260	312	338
Turnkey Projects	442	712	861
Consultancy services	1740	1800	2000
Total	2442	2824	3199
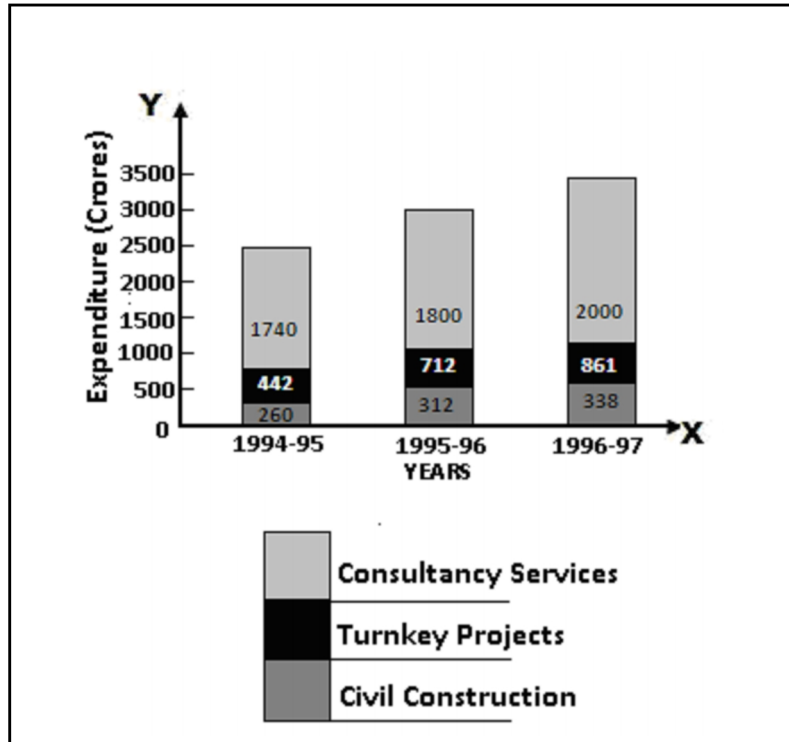
Component Bar Diagram

5. **Circular or Pie Chart:** Pie chart is an alternative to a component bar diagrams. A pie chart basically consists of a circle divided into sector by radii in such a way that areas of the sectors are proportional to the values of the component items under investigation. The whole circle, of course, represents the whole of data under investigation. It is also known as circular chart or percentage chart diagram.

Steps to draw Pie Chart

– Express the different components of the given data in percentages of the whole

– Multiply the each percentage components to 3.6 (since the total angle at the centre being 360 degree, it will represent the whole i.e. 100%)

– Draw a circle with convenient radius

– Divide the circle into different sectors with those central angles

– Shade different sector differently

Use of Pie Chart

The use of pie chart is quite popular as the circle provides a visual concept of the whole (100%). Pie charts are also one of the most commonly used charts because they are simple to use. Despite its popularity, pie charts should be used sparingly for two reasons:

– They are best used for displaying statistical information when there are no more than six components only, otherwise, the resulting picture will be too complex to understand.

– Pie charts are not useful when the values of each components are similar because it is difficult to see the differences between slice sizes.

(See Illustration 8.9)

Illustration 8.9

Represent the following data, on India's exports (₹ in Crores) by regions during April to February, 1997.

Region:	Europe	Asia	America	Africa
Exports:	32699	42516	23495	5133

Solution:

Total exports = 32699 + 42516 + 23495 + 5133 = ₹10,3843 crores

The angles for different regions are:

Europe $= \dfrac{32699 \times 360}{103843} = 113°$

Asia $= \dfrac{42516 \times 360}{103843} = 147°$

America $= \dfrac{23495 \times 360}{103843} = 82°$

Africa $= \dfrac{5133 \times 360}{103843} = 18°$
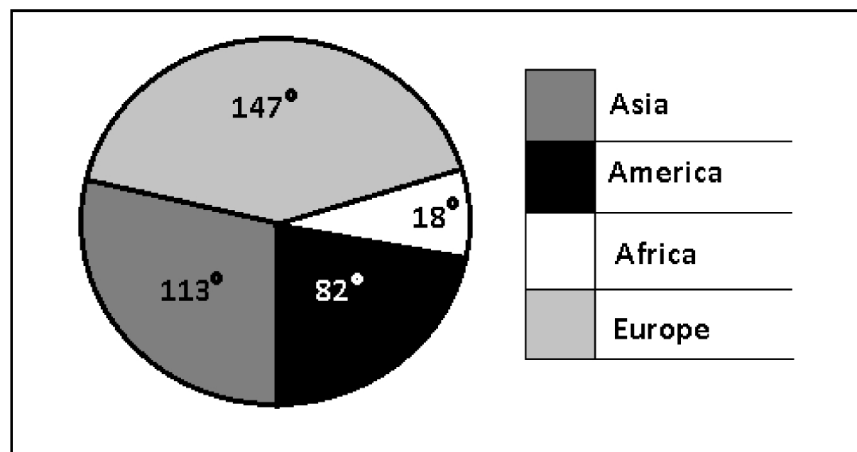
Pie-Chart

V. Graphic Presentation

Graphic presentation is another way of presentation of data. Graphs are commonly used for the presentation of time series and frequency distributions. Graphic presentation of data, such as diagrammatic presentation, also provides a quick and easier way of understanding broad features and to facilitate the process of comparison of two or more situations. In addition to this, it can also be used as a tool of analysis e.g. we can find median, mode etc. by means of graph.

Construction of a Graph

A point in a plane can be located with reference to two mutually perpendicular lines. The horizontal line is called the X-axis and the vertical line the Y-axis. Their point of intersection is termed as origin. The position of a point in a place is located in terms of its distances from the two axes. If a point P is 3 units away from Y-axis and 5 units away from X-axis, its location will be as shown in the figure:
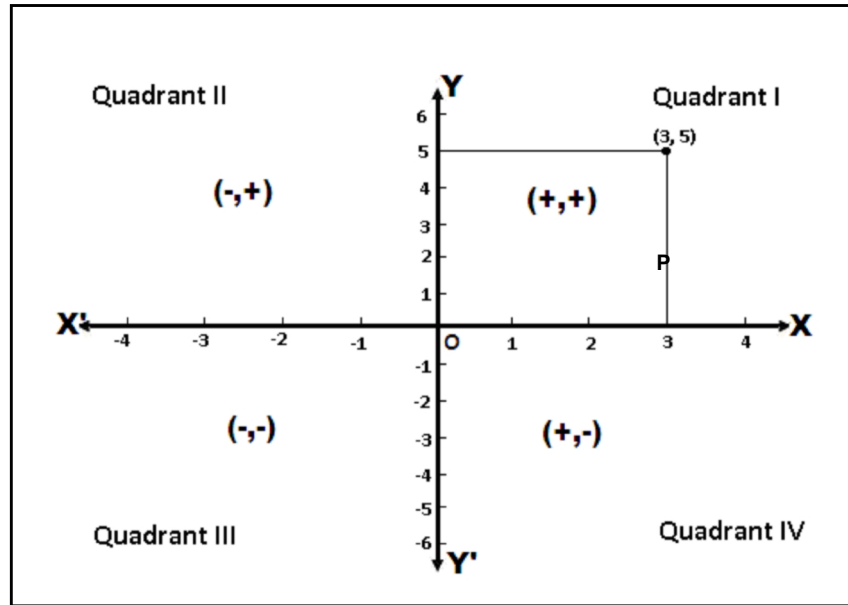
Graphical Plane

It should be noted here that the distance of the point from Y-axis is measured along X-axis and its distance from X-axis is measured along Y-axis. To measure 3 units from Y-axis, we move 3 units along X-axis and erect a perpendicular at this point. Similarly, to measure 5 units from X-axis, we move 5 units along Y-axis and erect a perpendicular. The point of intersection of these two perpendiculars is the required point. The position of the point is denoted by the pair of numbers (3, 5). These numbers are respectively termed as abscissa and ordinates of the point. Jointly they are termed as the coordinates of a point. The coordinates of a point, in general form, are written as (X, Y).

The four parts of the plane are called quadrants. It may be noted that X and Y are both positive in first, X is negative and Y is positive in second, X and Y are both negative in third and X is positive and Y is negative in fourth quadrant.

Different points can be plotted for a different pair of values e.g. for data on demand of a commodity at different prices, we can locate a point for each pair of quantity and price. These points are then joined by a line or a curve to get the required graph.

General Rules for a Graphic Presentation

For an attractive and effective graphic presentation, statistical data requires certain rules, principles and guidelines. Some of the major principles in this regard are discussed as under :

  (i) *Suitable Title:* It is essential for a graph to have a suitable title. It indicates clearly the subject for which the facts or data are presented in the graph.

 (ii) *Unit of Measurement:* The unit of measurement of a variable should be clearly stated below the title.

(iii) *Suitable Scale:* There should be suitable scale so that the whole data may be represented in accurate manner. The respective sizes of the scale of X axis and Y axis cannot be rigidly laid down. About choice of scale, Bowley has rightly pointed out, "It is difficult to lay down rules for the proper choice of scales by which the figures should be plotted out. It is only the ratio between the horizontal and vertical scales that need to be considered. The figure must be sufficiently small for the whole for it to be visible at once: if the figure is complicated, related to long series of years and varying numbers, minute accuracy must be sacrificed for this consideration. Supposing the horizontal scale is decided, the vertical scale must be

chosen so that the part of the line which shows the greatest rate of increase is well inclined to the vertical which can be managed by making the scale sufficiently small and on the other hand all important fluctuations must be clearly visible for which the scale may need to be decreased. Any scale which satisfies both these conditions will fulfill its purpose".

(iv) *Index:* A brief index explaining various types of shades, colours, lines and designs used in the construction of the graph along with scale of interpretation showed may be given for clear understanding.

 (v) *Sources of Data:* The sources of information wherever possible should be given at the bottom of the graph.

(vi) *Simplicity:* The Graph should be as simple as possible so that they are easily understood even by a ordinary person who does not have any exposure in the areas of statistics or mathematics.

(vii) *Neatness:* Since the graphs is visual aids for presentation of statistical data, it is therefore, imperative that they are made neat, clean and very attractive by proper size and lettering, and use of appropriate devices like different colours, dashes, dark lines broken lines etc.

Merits of a Graph

The technique of graph offers numerous advantages to the users. Some of the major advantages are listed as under:

– The graph presents the data in such a fashion, that it is easier for the user to understand.

– It gives attractive presentation to statistical data as compared to tables. By looking at the graph, it is easier to notice the main features of data at a glance in trends and fluctuations.

– It is time saving device.

– It facilitates comparison of data relating to two different periods of time or two different regions.

– It does not require knowledge of mathematics on the part of the users to understand the message of graph.

– It is helpful to locate the median, mode and mean values of the data.

– It is helpful in forecasting, interpolation and extrapolation of data.

Limitations of a Graph

Although the technique of graph is a useful one, but it is not free from limitations. Some of its major limitations are as under :

– It lacks complete accuracy about facts.

– It depicts only few characteristics of data.

– The graph cannot be used in support of some statement.

– The graph is not a substitute for tabulation in all circumstances and for all purposes.

– It is not easy to a layman to understand and interpret the graph.

– It shows only the unreasonable tendency of the data and actual values are not always clear from the graph.

Types of Graphs

Basically, the graphs are broadly divided into two categories:

– Graph of time series or time series graph

– Graph of frequency distribution

Time Series Graph

A time series graph also known as historigram refers to a graph which depicts the value of a variable over a different point of time. In time series graph, time is the most important factor and the variable is related to time which may be either year, month, week, day, hour or even minutes and second. It is the most helpful device to understand and analyze the changes in a variable at different point of time. These types of graphs are widely used by economists, businessmen and statistician because:

– it is easy to understand;

– it enables to present more information of complicated nature in very simple and precise way.

The construction of this type of graph is very simple and does not require any technical skill on the part of users. Some of the major steps involved in the construction of time series graph are given as under:

– Represent time on horizontal or X axis and value of variable on vertical or Y axis. Ensure that the unit of time and value of variable should be clearly stated.

– Start Y value with zero and devise a suitable scale to present the value of a variable so that the whole data can be accommodated with in given space. For instance, 1 centimeter on the vertical scale may be equal to 100 units, 500 units or 1000 units. Like wise, it may be equal to year; month or days on horizontal axis.

– Plot the values of variable corresponding to the true factor and join different points by drawing a straight line.

– If there are more than one variable, than these value may be plotted on the graph by different types of lines.

A brief details of various types of graph are given as under:

 1. **Line Graph:** A line graph is a way to summarize how two pieces of information are related and how they vary depending on one another.

In graph, the successive plotted points are joined with the help of a straight line.

Advantages:

– It can compare multiple continuous data sets easily

– Interim data can be inferred from graph line

Disadvantages:

– It used only with continuous data.

Illustration 8.10

Represent the following population of India data by a suitable graph:

Year	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991
Population (in Crores)	23.8	25.2	25.1	27.9	31.8	36.1	43.9	54.8	68.5	83.7

Solution:
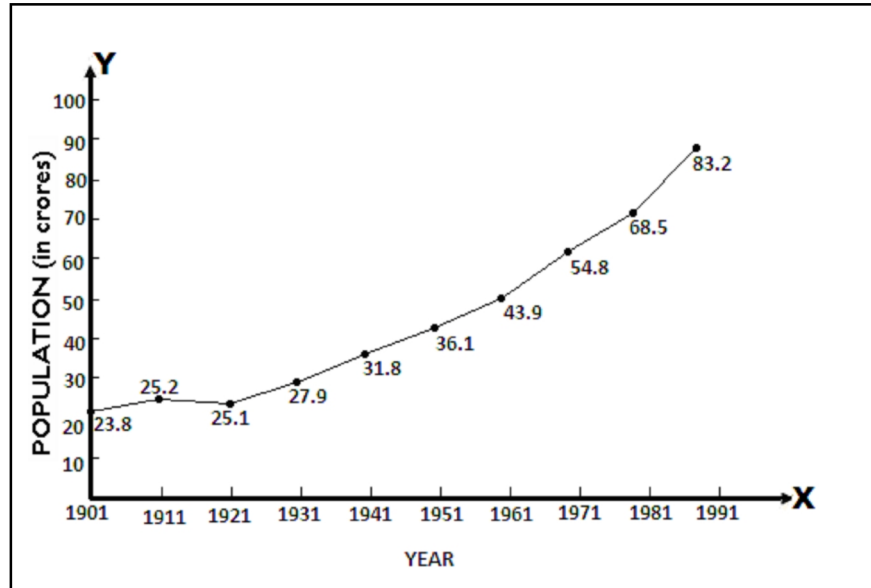
Line Graph

Use of False Base Line: The basic principle of drawing a time series graph is that vertical line must start from zero. If the fluctuations in the values of the dependent variable (to be presented on Y axis) are very small relative to their magnitude and if the minimum of these values is very distant (far greater) from zero, the point of origin, then for an effective portrayal of these fluctuations, the vertical scale is stretched by using false base line. In such a situation the vertical line is broken and the space between the origin 0 and minimum value (or some convenient value near that) of dependent variable is omitted by drawing two zigzag horizontal line above the base line. The scale along Y axis is framed accordingly. The false base line graph technique is useful from two points of views.

- to magnify the minor fluctuation in time series data

- to economize the space.

For better understanding of use of false base line see illustration 9.11.

Illustration 8.11

The yearly exports (₹ in Crores) of coffee are given below. Represent the data by a line graph.

Year	1984-85	1985-86	1986-87	1987-88	1988-89
Exports (₹ in Crores)	210	265	297	262	294

Solution :
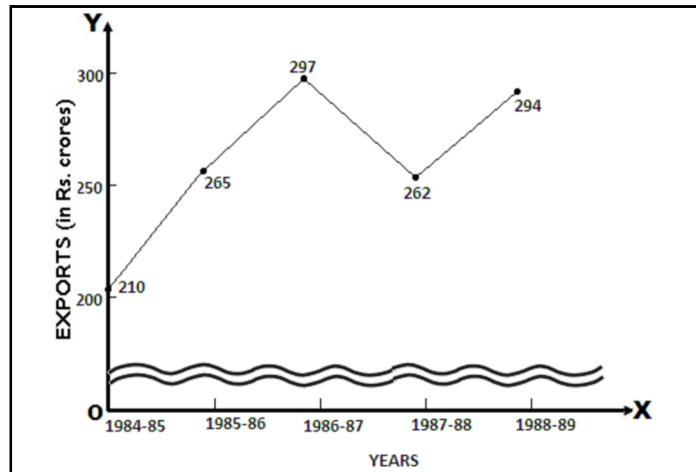
Line Graph with false base line

Since the fluctuations in values are small in relation to their size, the graph will be drawn by the use of false base line, as shown below:

2. Net Balance Graph

When the net balance of, say, income and expenditure or revenue and cost or exports and imports etc. are to be shown, and then net balance graph is used. Different shades are used for positive and negative differences. See the following Illustration.

Illustration 8.12

Represent the following data on revenue and cost, of a company during July 1997 to December 1997, by a net balance graph.

Months (1997)	July	Aug.	Sep.	Oct.	Nov.	Dec.
Revenue (₹ in Lakhs)	30	35	28	30	33	32
Cost (₹ in Lakhs)	28	32	30	31	29	30

Solution :

We draw line graphs of revenue and cost separately. The gap between them will be either profit (when revenue is greater than cost) or loss. The distinction between profit and loss is made by shading the respective gaps in different styles, as shown in the following graph
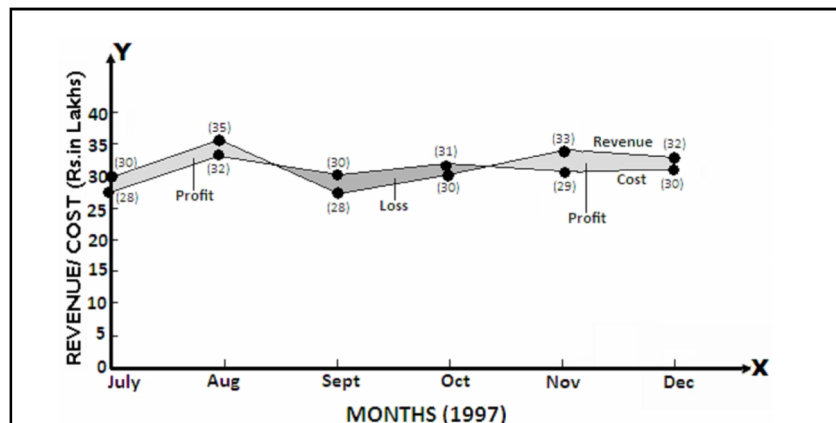
Net Balance Graph

GRAPHS OF A FREQUENCY DISTRIBUTION

A frequency distribution can also be represented by means of a graph. The most common forms of graphs are discussed below.

1. Histogram: A histogram is a graph of a grouped frequency distribution in which class intervals are plotted on X axis and their respective frequencies on Y axis. On each class interval, a rectangle is created with height proportional to the frequency density of the class.

The histogram is a popular graphing tool. It is used to summarize discrete or continuous data which are measured on an interval scale. It is often used to illustrate the major features of the distribution of the data in a convenient form. A histogram divides up the range of possible values in a data set into classes or groups. For each group, a rectangle is constructed with a base length equal to the range of values in that specific group, and an area proportional to the number of observations falling into that group. This means that the rectangles will be drawn of non-uniform height. A histogram has an appearance similar to a vertical bar graph, but when the variables are continuous, there are no gaps between the bars. When the variables are discrete, however, gaps should be left between the bars.

Generally, a histogram will have bars of equal width, although this is not the case when class intervals vary in size. Choosing the appropriate width of the bars for a histogram is very important. Thus, the histogram consists simply of a set of vertical bars. Values of the variable being studied are measured on an arithmetic scale along the horizontal x-axis. The bars are of equal width and correspond to the equal class intervals, while the height of each bar corresponds to the frequency of the class it represents.

– Construction of a Histogram when Class Intervals are equal: When class intervals are equal, then there is no need of computing frequency density. We can take the height of each rectangle to be equal to the frequency of the class. (See Illustration 9.13).

Illustration 8.13

Draw a histogram for the following frequency distribution :

Class Intervals	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	15	25	30	35	50	40	10

Solution:
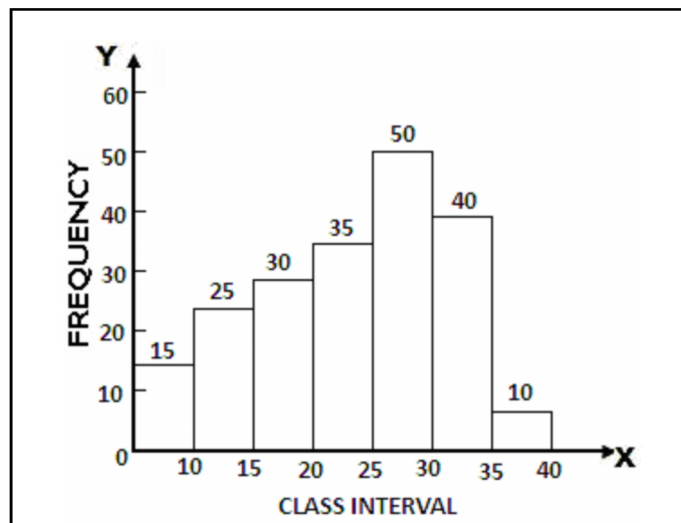
Histogram

– Construction of a Histogram when Class Intervals are not equal: When different class intervals of a frequency distribution are not of equal width, we compute frequency density (= frequency ÷ class interval) of each class. The product of shortest class interval and the frequency density of a class is taken as the height of the corresponding rectangle. Thus, we can write the height of the rectangle of a class as under:

$$= \frac{\text{Frequency of the class} \times \text{Width of the Smallest Class Interval}}{\text{Width of the class Interval}}$$

OR
$$= \frac{\text{Frequency of the class}}{\text{Adjustment factor}}$$

Where Adjustment Factor (AF) $= \dfrac{\text{Width of the Class Interval}}{\text{Width of the Smallest Class Interval}}$

The following illustration helps to understand the steps to construct a histogram when class intervals are not uniform.

Illustration 8.14

Represent the following frequency distribution by a histogram :

Class Intervals	0-10	10-15	15-30	30-40	40-60
Frequency	8	10	36	40	32

Solution:

Class Intervals	Frequency (f)	Adjustment Factor (AF)	Adjusted Frequency = f/AF
0-10	8	2	4
10-15	10	1	10
15-30	36	3	12
30-40	40	2	20
40-60	32	4	8
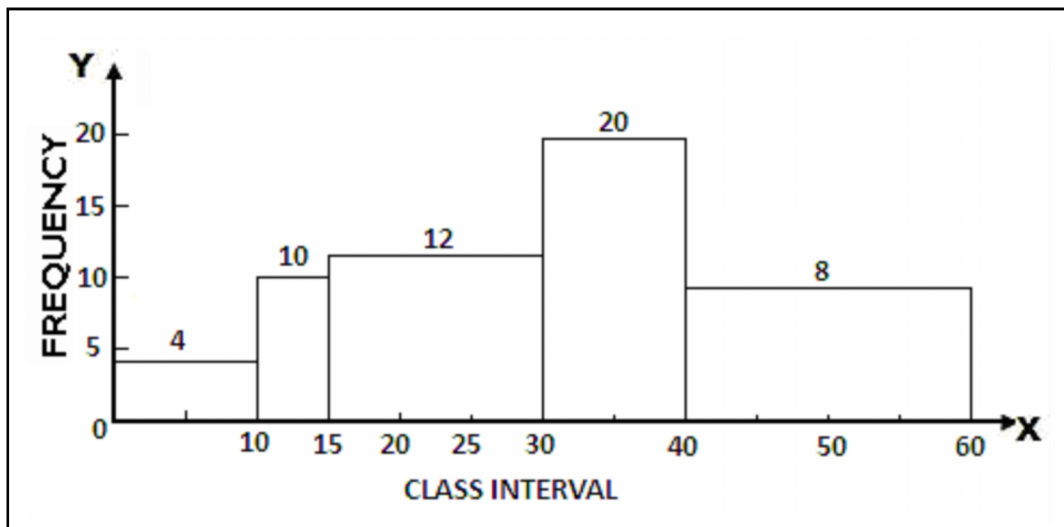
Use of Histogram: The histogram is used for variable whose values are numerically expressed and measured on an interval scale. It is generally used in a situation when dealing with large data sets (greater than 100 observations). A histogram can also help to detect any unusual observations (outliers) or any gaps in the data.

2. *Frequency Polygon or Histograph:* A frequency polygon or Histograph is another method of representing a frequency distribution on a graph.

A frequency polygon is drawn by joining the midpoints of the upper widths of adjacent rectangles of the histogram, with straight lines.

Two hypothetical class intervals, in the beginning and at the end of the distribution, are created. The ends of the polygon are extended up to base line by joining them with the mid-points of the hypothetical classes. This step is required to make area under the polygon equal to the area under histogram. Frequency polygon can also be constructed without making rectangles. The points of frequency polygon are obtained by plotting mid-points of classes against heights of various rectangles, which are equal to the frequencies if all the classes are of equal width. A Histograph smoothes out the abrupt changes that may appear in histogram and is useful for demonstrating continuity of the variable being studied. (See Illustration 9.15)

Illustration 8.15

The daily profits (in ₹) of 100 shops are distributed as follows :

Profit/Shop:	0-100	100-200	200-300	300-400	400-500	500-600
No. of Shops:	12	18	27	20	17	6

Construct a frequency polygon of the above distribution.
Solution:
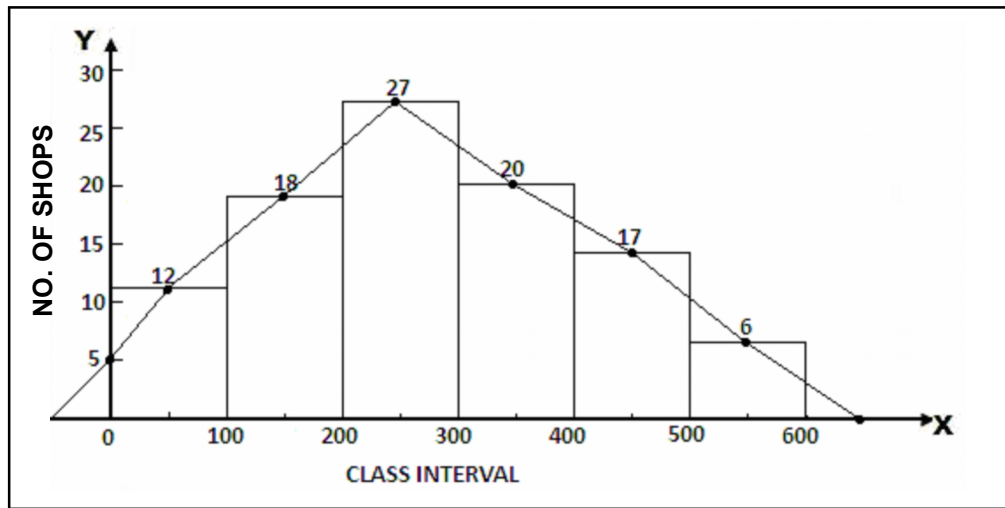
FREQUENCY POLYGON

Illustration 8.16

Represent the following data by a frequency polygon.

Class Intervals:	5-15	15-25	25-35	35-45	45-55	55-65
Frequency:	12	15	20	18	10	5

Solution:

We shall construct the frequency polygon without making the histogram.
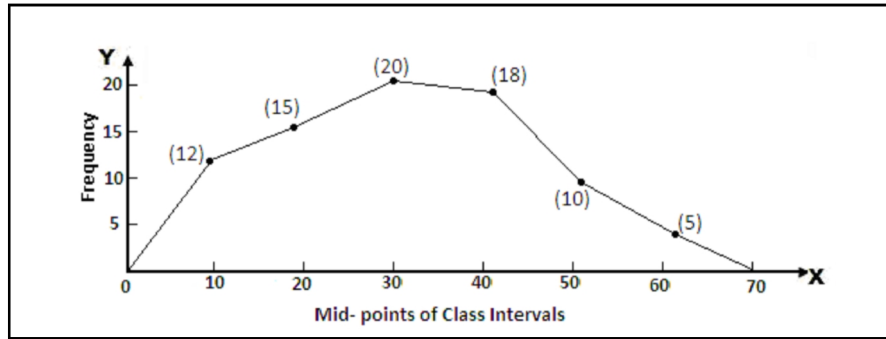
FREQUENCY POLYGON

3. **Frequency curve:** When the verticals of a frequency polygon are joined by a smooth curve the resulting figure is known as a frequency curve. As the number of observations increases, there is need of accommodating more and more classes to accommodate them and hence the width of each class becomes smaller and smaller. In such a situation, the variable under consideration tends to become continuous and the frequency polygon of the data tends to acquire the shape of a frequency curve. Thus, a frequency curve may be regarded as a limiting form of frequency polygon as the number of observations tends to become large. The construction of a frequency curve should be done very carefully by avoiding, as far as possible, sharp and sudden turns. Smoothing should be done so that the area under the curve is approximately equal to the area under histogram or frequency polygon. The dotted curve represents the frequency curve of the data. See Illustration 15.

4. **Cumulative Frequency Curve or Ogive:** The graphical representation of a cumulative frequency distribution is termed as a cumulative frequency curve or ogive. Since, a cumulative frequency distribution can be of 'less than' or more than type, accordingly, there are two types of ogives; 'less than ogive' and 'more than ogives'. The ogives can be better understood with the help of following illustration 9.17:

Illustration 8.17

Draw 'less than' and 'more than' ogives for the following distribution of weekly wages of 100 workers.

Weekly Wages (₹)	0-100	100-200	200-300	300-400	400-500
No. of Workers	8	30	35	20	7

Solution:

First we contract the 'less than' and 'more than' type cumulative frequency distributions.

Wages Cumulative less than (₹)	Cumulative Frequency (Workers)	Wages Cumulative More than (₹)	Cumulative Frequency (Workers)
100	8	0	100
200	38	100	92
300	73	200	62
400	93	300	27
500	100	400	7

Note that the two ogives intersect at median.

An Ogive can be used to determine positional averages like median, quartiles, deciles, percentiles etc. We can also determine the percentage of cases lying between certain limits. Various frequency distributions can be compared on the basis of their Ogives.
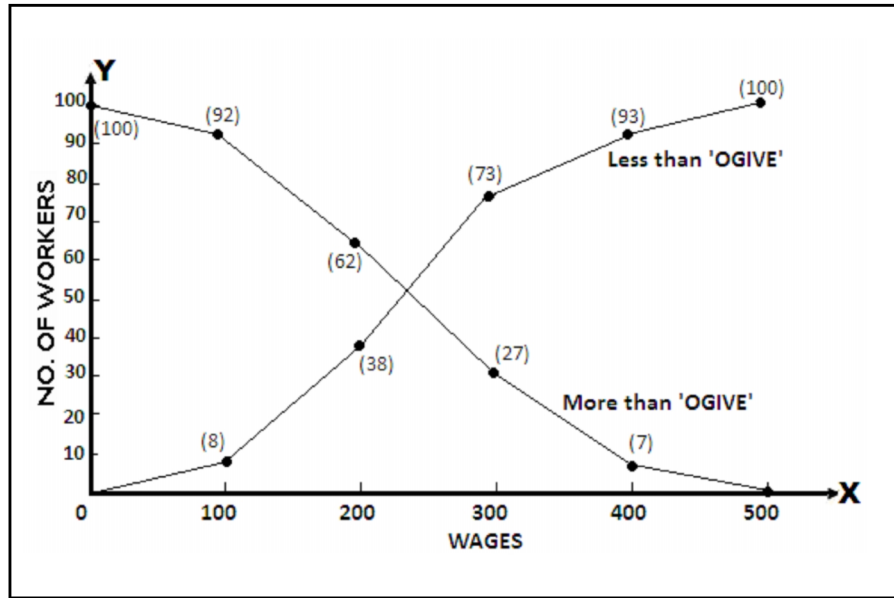
Scatter Diagram

To find out the nature of relationship we prepare a dot chart, this dot chart is known as scatter diagram.

When the given data is plotted on a graph paper in the form of the values of X and Y, we got many points as the number of observations. By looking at various points on graph paper we can get the idea whether the variables are related or not. The greater the scatter of the plotted points on graph paper the lesser is the relationship between the two variables.

Positive Correlation   **Negative Correlation**
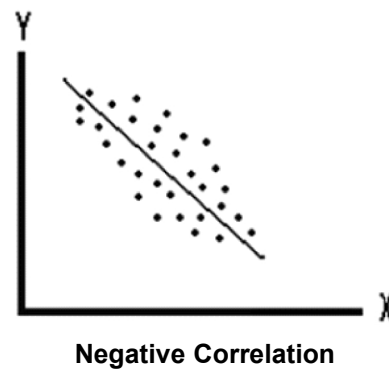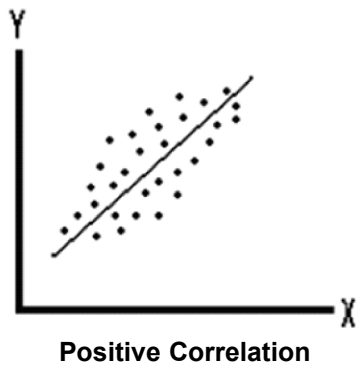
MEASURES OF CENTRAL TENDENCY

Introduction

A measure of Central Tendency or an average of a distribution is a single figure which can be regarded as a representative of the whole distribution. Different authors have defined the average of a distribution in different ways.

Some of the important definitions of an average are given as under:

"An average is an attempt to find one single figure to describe the whole of figures".

Clark and Sekkade

"An average is a value which is typical or representative of a set of data".

Murry R. Spiegal

"An average is a single value within the range of data that is used to represent all values in the series. Since an average is somewhere within the range of the data, it is also called a measure of Central Value".

Croxton and Cowden

"A measure of central tendency is a typical value around which other figures congregate".

Simpson and Kafka

Thus, it is clear from the above definitions that an average is a single figure that represents the frequency distribution of the large number of items.

Objectives and Significance

Some of the important objectives and significance of statistical averages are discussed as under:

(i) *To Present Huge Mass of Data in a Summarized Form:* It is very difficult for human mind to grasp a large body of numerical figures. A measure of average is used to summarize such data into a single figure which is easier to understand and remember.

(ii) *To Facilitate Comparison:* Different sets of data can be compared by comparing their averages. For example, the level of wages of workers in two firms can be compared by mean (or average) wages of workers in each of them.

(iii) *To Help in Decision Making:* Most of the decisions to be taken in research planning etc. are based on the average values of certain variables. For example, if the average monthly sales of a company are falling, the sales manager may have to take certain decisions to improve it.

(iv) *To Establish Precise Relationship:* It also focuses and establishes precise relationships among variables.

Requisites of a Good Statistical Average

An average is a single value representing a group of values and hence it should possess the following characteristics:

(i) *Firmly Defined:* An average should not be flexible or that there would be no confusion with regard to its meaning and connotation. If an average is left to the estimation of an observer and if it is not rigid, it cannot be a representative of a series. The bias of the investigator in such cases would affect the value of the average considerably.

(ii) *Easy to Understand and Simple to Compute:* An average should be readily understood even by a layman

otherwise its use is bound to be very limited. The properties of the average should be such that they can be easily understood by persons of ordinary intelligence. It should also be simple to compute so that it can be used widely. But ease of computation should not be sought at the expense of other advantages. For instance, a difficult average may be desirable so as to ensure greater accuracy.

(iii) *Based on all the Items:* The average should depend upon each and every item of the series. If some of the items are not taken into account the average cannot be said to be a representative one. Further, the average itself is altered if any of the items is dropped. For example, the arithmetic mean of 35, 43, 25, 70, 60, 85 is 53. If one item is dropped say 43, the arithmetic mean would become 55.

(iv) *Capable of Further Algebraic Treatment:* An average could be used for further statistical computations so that its utility gets enhanced. For example, if we are given the data about the average heights of boys and girls in a class and their numbers separately, then we should be able to work out the combined average.

(v) *Not be Affected by Extreme Observations:* The average obtained in a particular field from two independent sample studies should not materially differ from each other. If one or two items unduly affect the averages i.e. either increases its value or reduces its value the average cannot be really typical of the entire series.

(vi) *Sampling Stability:* An average should be least affected by sampling fluctuations, i.e., if we take independent random samples of the same size say 35 different groups from a given population and compute the average of each group, it is expected to get approximately the same value. The value so obtained from different samples should not vary much from one another. There may be some difference but these samples in which this difference is less are considered better than those in which this difference is more.

(vii) *Capable of Calculation in case of Distributions Containing Open and Class-interval:* Sometimes, we are bound to use open end class interval for classification. Even in such situations, an average can be calculated very easily without making any assumptions regarding such open end class interval.

Types of Average

Various measures of average can be classified into two broad categories.

– Mathematical Averages: It covers Arithmetic Mean or Mean, Geometric Mean and Harmonic Mean. Further, each of these averages can be simple or weighted. (GM and HM are out of the scope of the syllabus.)

– Positional Averages: Median, Mode, Quartiles, Deciles and Percentiles.

Arithmetic Mean

Arithmetic mean is defined as the sum of observations divided by the number of observations.

If there are N individual observations X_1, X_2 ... X_n, and then their arithmetic mean, denoted by X, is given by

$$\overline{X} = \frac{X_1 + X_2 \ldots\ldots + X_n}{N} = \frac{\Sigma X_i}{N}$$

Where

$\overline{X}$ = Arithmetic Mean

ΣX_i = Sum of the Values of the Variable X

N = Number of items.

Further, if the frequency of observations X_i is f_i (i = 1, 2 .... n),

$$\overline{X} = \frac{f_1 X_1 + f_2 X_2 \ldots\ldots + f_n X_n}{f_1 + f_2 + \ldots\ldots + f_n} = \frac{\Sigma f_i X_i}{\Sigma f}$$

The following illustration will help us to calculate Arithmetic mean in different situations.

If the frequency distribution is grouped, and then mid-points of the class intervals are taken as X_i values to calculate Arithmetic Mean.

Illustration 8.18

Compute the arithmetic mean of the following frequency distribution:

Class Intervals:	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency:	8	10	12	15	18	14	11	5

Solution:

Class Intervals	Mid-values (X)	Frequency(f)	fX
0-10	5	8	40
10-20	15	10	150
20-30	25	12	300
30-40	35	15	525
40-50	45	18	810
50-60	55	14	770
60-70	65	11	715
70-80	75	5	375
Total		**Σf = 93**	**ΣfX = 3685**

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

or $\quad \overline{X} = \dfrac{3685}{93} = 39.6$

Short-cut Method

This method is used when the magnitude of X_i values is large. The use of short-cut method is helpful in the simplification of calculation work.

Let A be the assumed mean, We can write

$d_i = X_i - A$

or $\quad \sum f_i d_i = \sum f_i(X_i - A) = \sum f_i X_i - A \sum f_i$

Dividing both sides by N ($=\sum f_i$). we get

$$\frac{\sum f_i d_i}{N} = \frac{\sum f_i X_i}{N} - A$$

or $\quad \overline{X} = \dfrac{\sum f_i X_i}{N} = A + \dfrac{\sum f_i d_i}{N}$

Illustration 8.19

The following is the distribution of weekly wages of workers in a factory. Calculate the arithmetic mean of the distribution.

Weekly Wages (₹)	240-269	270-299	300-329	330-359	360-389	390-419	420-449
No. of Workers	7	19	27	15	12	12	8

Solution:

Note that the given class intervals are inclusive. However, for the computation of mean, they need not be converted into exclusive intervals.

Weekly Wages (₹)	Mid-value (X)	Frequency (f)	$d_i = X - A$ A=X-344.5	$f_i d_i$
240-269	254.5	7	−90	−630
270-299	284.5	19	−60	−1140
300-329	314.5	27	−30	−810
330-359	**A= 344.5**	15	0	0
360-389	374.5	12	30	360
390-419	404.5	12	60	720
420-449	434.5	8	90	720
Total		$\Sigma f_i = 100$		$\Sigma f_i d_i = -780$

$$\overline{X} = A + \frac{\Sigma f_i d_i}{\Sigma f_i}$$

$$= 344.5 + \frac{(-780)}{100}$$

$$= 336.7$$

Step Deviation or Coding Method

In a frequency distribution, if the X_i values are equally spaced, the computation work can be further eased by using step-deviation method.

We can define $U_i = \dfrac{X_i - A}{h}$ where, h is the difference of any two successive observations.

Multiplying both the sides by f_i and taking sum of all observations, we get —

$$\Sigma f_i U_i = \frac{\Sigma f_i (X_i - A)}{h} = \frac{1}{h} \{\Sigma f_i X_i - NA\}$$

$$\text{or }, \Sigma f_i X_i - NA = h \, \Sigma f_i U_i$$

Dividing both the sides by N, we get

$$\overline{X} = \frac{\Sigma f_i X_i}{N} = A + h \frac{\Sigma f_i U_i}{N}$$

Weekly Wages (₹)	Mid-value (X)	Frequency (f)	$d_i = X - 344.5$	$f_i\, d_i$	$U_i = \dfrac{X-344.5}{30}$	$f_i\, U_i$
240-269	254.5	7	–90	–630	–3	–21
270-299	284.5	19	–60	–1140	–2	–38
300-329	314.5	27	–30	–810	–1	–27
330-359	**344.5**	15	0	0	0	0
360-389	374.5	12	30	360	1	12
390-419	404.5	12	60	720	2	24
420-449	434.5	8	90	720	3	24
Total		$\Sigma f_i = 100$		$\Sigma f_i\, d_i = -780$		$\Sigma f_i U_i = -26$

Here h= 30, A = 344.5

Using this formula, we can find the mean of the above illustration as given below:

$$7 \quad = \quad 344.5 + \left(\frac{-26}{100} \times 30 \right)$$

$$= \quad 344.5 - 7.8$$

$$= \quad 336.7$$

Weighted Arithmetic Mean

In the computation of simple arithmetic mean, it is assumed that all the items (or observations) of the distribution are of equal importance. However, if all the items are not of equal importance, then simple arithmetic mean will not be a good representative of the distribution. Hence, the weighing of different items become necessary and weight are assigned in proportion to the importance of the item in the group i.e. more important item is assigned more weight.

Let $W_1 W_2 \ldots W_n$ be the respective weights of $X_1, X_2 \ldots X_n$

Then,

Weighted Arithmetic Mean, $\overline{X}_w = \dfrac{\Sigma W_i X_i}{\Sigma W_i} \qquad \dfrac{W_1 X_1 + W_2 X_2 + \ldots\ldots + W_n X_n}{W_1 + W_2 + \ldots\ldots\ldots + W_n}$

Illustration 8.20

Ram purchases 20 kgs. of Wheat, 10 kgs. of rice and 2 kgs. of ghee every month. If the price of wheat is ₹10 per kg., price of rice is ₹14 per kg. and price of ghee is ₹120 per kg. Find the average price per kg. per month.

Solution:

The simple average of the prices, $\quad \overline{P} = \dfrac{10 + 14 + 20}{3} = \dfrac{144}{3} = ₹48$

Thus, the simple average is not representative of all observations because it is highly affected by the price of ghee, which is not very important in the group, since its consumption in a month is only 2 kg. In such a situation, the weighted arithmetic mean will be appropriate. The importance to various items can be based on its quantity consumed.

$$\text{Thus,}\ \overline{P}_w = \frac{20 \times 10 + 10 \times 14 + 2 \times 120}{20 + 10 + 2} = \ ₹18.125$$

Properties of Arithmetic Mean

Arithmetic mean possesses the following important properties:

1. The sum of deviations of a set of observations is zero when these deviations are taken from their arithmetic mean. The deviation of a set of observation X_i from its mean is written as $X_i - \overline{X}$. Then, sum of deviations is written as $S(X_i - \overline{X})$. The above property implies that

   $$\Sigma(X_i - \overline{X}) = 0$$

2. The sum of square of deviations a set of observations is minimum, when the deviations are taken from their arithmetic mean.

   This property implies that,

   $$\Sigma f_i (X_i - \overline{X})^2 \leq \Sigma f_i (X_i - A)^2 \text{ where A is any arbitrary origin.}$$

3. Arithmetic mean is capable of being treated algebraically. This property implies that if any two of the three values, X, $\Sigma f_i X_i$ and N, are known, the third can be computed easily.

4. If N_1 and $\overline{X}_1$ are the number of observations and mean of a series and N_2 and $\overline{X}_2$ are the respective values of another series, then mean $\overline{X}_{12}$ of the series may be obtained by combining the two series i.e.

   $$\left(\overline{X}_{12} = \frac{N_1 \overline{X}_1 + N_2 \overline{X}_2}{N_1 + N_2} \right)$$

5. If a constant A is added (or subtracted) from every value of the series, then the mean of the series also gets added (or substracted) by it.

   According to this property if we write $X_1 + A, X_2 + A,....$ then the mean of the changed series = $\overline{X} + A$, where $\overline{X}$ is mean of the $X_1, X_2 ...... X_n$ series.

6. If every value of a series is divided (or multiplied) by a constant h, then the mean of the series also gets divided (or multiplied) by it.

   According to this property, if we write $\dfrac{X_1}{h}, \dfrac{X_2}{h}$ .... then the mean of the changed series = $\overline{X}/h$

Illustration 8.21

The average rainfall for a week, excluding Sunday, was 10 cms. Due to heavy rainfall on Sunday, the average for the week rose to 15 cms. How much rainfall was on Sunday?

Solution:

A week can be treated as composed of two groups. First group consisting of 6 days, excluding Sunday for which $N_1 = 6$ and $X_1 = 10$. The second group consisting of only Sunday for which $N_2 = 1$. Note that the mean of this group is equal to the observation itself that we want to find. Let this be X_2. We are also given X_{12} (Combined mean) = 15.

$$\therefore 15 = \frac{6 \times 10 + \overline{X}_2}{7} \text{ or } 60 + \overline{X}_2 = 105 \text{ or } \overline{X}_2 = 45 \text{ cms}$$

Illustration 8.22

The mean weight of 20 employees in a company is 65 kgs. And the mean weight of 8 employees is 70 kgs. And another 7 is 55 kgs. You are required to find out the mean weight of remainder.

Solution:

Mean weight of 20 employees in a company = 65 kgs

Total weight of 20 employees = 65 x 20 = 1300 kgs

Total weight of 8 employees whose mean weight is 70 kgs. = 560 kgs

Total weight of 7 employees whose mean weight is 55 kgs. = 385 kgs

Total weight of remaining 5 employees = 1300 – (560 + 385) = 355 kgs

Mean weight of these 5 employees = 355/5 = 71 kgs

Illustration 8.23

Two groups of students reported mean weights of 160 kg. and 150 kg. respectively. Find out, when the mean weight of both the groups together be 155 kg?

Solution

Let N_1 be the number of students reported in the first group and N_2 be the number of students reported in second group, then:

	Group-I	Group-II
Number of students	N_1	N_2
Mean Weights (kg.)	$\overline{X}_1 = 160$	$\overline{X}_2 = 150$

Combined Mean Weight $\overline{X}_{12}$ = 155 kg

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}$$

$$155 = \frac{160N_1 + 150N_2}{N_1 + N_2}$$

$$155N_1 + 155N_2 = 160N_1 + 150N_2$$

Thus, the combined mean weight of both the groups together will be 155 kg when $N_1 = N_2$

Illustration 8.24

The mean weight of 150 students in certain class in a co-educational college in Delhi University is 60 kg. The mean weight of boys in the Class is 70 kg. and that of girls is 50 kg. Find the number of boys and the number of girls in the Class.

Solution:

Let N_1 and N_2 stand for number of boys and girls in the class respectively.

$\overline{X}_1$ = Mean weight of boys in the class

$\overline{X}_2$ = Mean weight of girls in the class

$\overline{X}_{12}$ = Combined mean weight of boys and girls in the class

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}$$

$$60\ (N_1 + N_2) = \ N_1 + 50\ N_2$$

$$60 \times 150 = \ 70\ N_1 + 50\ N_2$$

$$9000 = 70 N_1 + 50 N_2 \quad\quad \text{…(i)}$$

$$150 = N_1 + N_2 \quad\quad \text{…(ii)}$$

Solving (i) and (ii) we may get,

$$9000 = 70 N_1 + 50 N_2$$

$$10500 = 70 N_1 + 70 N_2$$

$$\pm$$

$$\overline{-1500 = -20 N_2}$$

$$20 N_2 = 1500$$

$$N_2 = 1500 / 20 = 75$$

$$N_1 + N_2 = 150$$

$$N_1 + 75 = 150$$

$$\Rightarrow \quad N_1 = 75$$

Hence, $\quad N_1 = 75 ; N_2 = 75$

Illustration 8.25

The mean of a certain number of observations is 40. If two or more items with values 50 and 64 are added to this data, the mean rises to 42. Find the number of items in the original data.

Solution:

Let 'n' be the number of observations whose mean $\overline{X} = 40$

Total of n Values, $\quad \overline{x} = \dfrac{\Sigma x_i}{n} \quad \therefore \Sigma x_i = n(40) = 40n$

Two more items of values 50 and 64 are added therefore, total of (n+2) values:

$$= \Sigma X_i + 50 + 64$$

$$= 40 n + 50 + 64$$

$$= 40 n + 114$$

Now, New Mean is 42

$\therefore \quad\quad \overline{X} = \dfrac{\text{New Total of (n+2) values}}{n+2}$

$\therefore \quad\quad 42 = \dfrac{40n + 114}{n + 2}$

$\therefore \quad 42n + 84 = 40n + 114$

$\therefore \quad\quad 2n = 30$

$\therefore \quad\quad n = 15$

Therefore, the number of items in the original data = 15

Illustration 8.26

The mean weight of 98 students is found to be 50 lbs. It is later discovered that the frequency of the class interval (30-40) was wrongly taken as 8 instead of 10. Calculate the correct mean.

Solution:

Incorrect mean, $\overline{X}$ = 50 kg

Σf_i = 98

∴ Incorrect $\overline{X}$ = $\dfrac{\text{Incorrect } \Sigma f_i X_i}{\Sigma f_i}$

50 = $\dfrac{\text{Incorrect } \Sigma f_i X_i}{98}$

∴ Incorrect $\Sigma f_i X_i$ = 98 x 50 = 4900

Now, Correct $\Sigma f_i X_i$ = Incorrect $\Sigma f_i X_i$ – (8 x 35) + (10 x 35)

Note, the class-mark of class interval (30 – 40) is 35 and for the calculation of the mean, we consider class marks.

Correct $\Sigma f_i X_i$ = 4900 – 280 + 350

= 4970

Also, Correct Σf_i = 98 + 2 = 100

∴ Correct Mean $\dfrac{\text{Correct } \Sigma f_i X_i}{\text{Correct } \Sigma f_i}$

= $\dfrac{4970}{100}$

$\overline{X}$ = 49.70 lbs

Illustration 8.27

The average marks of three batches of students having 70, 50 and 30 students respectively are 50, 55 and 45. Find the average marks of all the 150 students, taken together.

Solution:

Let $\overline{X}$ be the average marks of all 150 students taken together

	Batch-I	**Batch-II**	**Batch-III**
Marks	$\overline{X}_1$=50	$\overline{X}_2$=55	$\overline{X}_3$=45
No. of Students	n_1=70	n_2=50	n_3=30

$$\overline{X} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2 + n_3\overline{x}_3}{n_1 + n_2 + n_3} = \frac{70 \times 50 + 50 \times 55 + 30 \times 45}{70 + 50 + 30} = \frac{7600}{150}$$

= 50.67 marks

Illustration 8.28

The mean salary paid per week to 1000 employees of an establishment was found to be ₹900. Later on, it was

discovered that the salaries of two employees were wrongly recorded as ₹750 and ₹365 instead of ₹570 and ₹635. Find the corrected mean salary.

Solution:

We first find the corrected sum of observation.

Corrected sum of observations = Uncorrected sum of observation – Sum of incorrect observations + Sum of correct observations

= 900 × 1,000 – 1,115 + 1,205 = 9,00,090

∴ Correct Mean = 9,00,090/1,000 = ₹900.09

Merits and Demerits of Arithmetic Mean

Merits:

Out of all averages, arithmetic mean is the most popular average. It is widely used by experts and policy makers because of its following merits:

– It is easy to understand and easy to calculate.

– It is a rigidly defined average.

– It is based on all the observations and hence can be regarded as representative of the data.

– It is capable of being treated mathematically and hence is widely used in statistical analysis.

– It can be computed even if detailed distribution is not known but sum of observations and number of observations is known.

– It provides a good basis for the comparison of two or more distributions.

– It is widely used in business and commerce.

Demerits**:**

Although arithmetic mean satisfies most of the characteristics of a good statistical average, yet it has certain limitations also. Some of these limitations are given below:

– It can neither be determined by inspection nor by graphical location.

– It cannot be computed for a qualitative data because qualitative data are not measurable.

– It is unduly affected by extreme observations and hence is not a good representative of data having some extreme observations.

– The value of arithmetic mean of a data may not be an observation of the data and as such it is called a fictitious average.

– It cannot be computed for a distribution with open ends.

– It assigns equal weight to high and small observations which distort the averages.

– Arithmetic mean cannot be calculated in case even a single observation in a series is missing.

– In the absence of complete distribution of observations, the arithmetic mean may lead to fallacious conclusions. For example, there may be two different distributions with same arithmetic means.

MEDIAN

Median of a distribution is that value of the variant which divides it into two equal parts. In terms of the frequency curve, the ordinate drawn at median divides the area under the curve into two equal parts.

Median is a positional average because its value depends upon the position and not on magnitude of the items.

Determination of Median

1. When individual observations are given

The given observations are arranged in either ascending or descending order of magnitude. If n is the number of observations, then median is given by

(i) the size of $\frac{(N+1)^{th}}{2}$ observation if n is odd or

(ii) the mean of $(N/2)^{th}$ and $(N/2 + 1)^{th}$ observation if N is even.

Illustration 8.29

Find median of the observations in each of the following two cases :

   (i) 25, 14, 28, 30, 25, 15, 32

   (ii) 35, 20, 55, 27, 15, 40

Solution:

   (i) Arranging the given observations in ascending order of magnitude, we have

      14, 15, 25, 25, 28, 30, 32

      Here, N is 7, therefore median is the size of $\left(\frac{7+1}{2}\right)^{th}$ i.e. 4th item. Hence, median is 25.

   (ii) Arranging the observations in ascending order, we may get

      15, 20, 27, 35, 40, 55

Since N = 6, therefore median in average of $\left(\frac{6}{2}\right)^{th}$ and $\left(\frac{6}{2}+1\right)^{th}$ term i.e., 3rd & 4th terms

Thus, the median is mean of 3rd and 4th observation = $\left(\frac{27+35}{2}\right)$ = 31.

Note: the same value of median is obtained if the observations are arranged in the descending order.

Alternative Definition of Median

We have defined median as a value so that half (i.e. 50%) of the observations lie below it and half of the observations lie above it. However, the median = 25, obtained in illustration 12(i) is such that only two (i.e. 28.6%) observations are below it and three (i.e. 42.9%) observations are above it. This indicates that the definition of median, given above, is somewhat ambiguous. To avoid this ambiguity, the median of a distribution is often defined as that value of the variant *such that at least half of the observations are less than or equal to it and at least half of the observations are greater than or equal to it.*

2.When grouped frequency distribution is given (Interpolation Formula):

For a continuous or grouped frequency distribution, median is that value of the variant such that the ordinate at it divides the area under the histogram into two equal parts.

The following steps are involved in the computation of median.

   – Locate the median class i.e. the class in which median lies. If N ($= \Sigma f_i$) is the number of observation, then median class is the class in which N/2th observation lies.

   – The median is given using the following interpolation formula:

$$M_d = L_m + \frac{\frac{N}{2} - C}{f_m} \times h,$$

Where,

M_d = Median

L_m = lower limit of the median class

f_m = frequency of the median class

h = width of the median class

C = Cumulative frequency of class preceding the median class.

Important Mathematical Property of Median

It may be noted that the sum of deviations of the variables from their median, ignoring signs, is always the minimum. For example, the median of 2, 4, 6, 8, 10 is 6. The deviations from 6 ignoring signs are 4, 2, 0, 2, 4, is 12. This total is smaller than the one obtained if deviations are calculated from any other value. Thus, if deviations are taken from 8, the deviation ignoring signs would be 6, 4, 2, 0, 2 and the total is 14.

Illustration 8.30

The following table gives the marks obtained by 300 students of a management course. Find the median of the distribution.

Marks obtained	0-15	15-30	30-45	45-60	60-75	75-90	90-100
No. of students	26	34	64	76	60	30	10

Solution:

Marks obtained	Frequency (f)	Cumulative Frequency (C)
0-15	26	26
15-30	34	60
30-45	64	124
45-60	76	200
60-75	60	260
75-90	30	290
90-100	10	300

Since N/2 = 150, therefore, 45-60 is the median class.

Further, L_m = 45, f_m = 76, h = 60-45 = 15 and C = 124

$$M_d = \frac{45 + (150 - 124) \times 15}{76}$$

$$= 50.13 \text{ marks}$$

Remarks:

- Since the variable in a grouped frequency distribution is continuous, we always take exact value of N/2, including figures after decimals, when N is odd.
- The above formula is also applicable when classes are of unequal width.
- When class intervals are of inclusive type, then these should be changed to exclusive type before the computation of median.
- Median can be computed even if class intervals are with open ends.

Illustration 8.31

Find median of the following data

Age greater than (in yrs.)	0	20	30	40	50	60	70
No. of Persons	230	218	200	165	123	73	288

Solution

Note that it is a greater than type cumulative frequency distribution. First we convert it into a less than type form.

Class Intervals	Greater than Cumulative Frequency	Frequency	Less than Cumulative Frequency
0-10	230	12	12
10-20	218	18	30
20-30	200	35	65
30-40	165	42	107
40-50	123	50	157
50-60	73	45	202
60-70	28	20	222
70 and above	8	8	230

N/2 = 230/2 = 115, therefore median class is 40-50

Also L_m = 40, f_m = 50, h = 10, C = 107

$$\therefore M_d = 40 + \frac{115 - 107}{50} \times 10 = 41.6 \text{ years}$$

Illustration 8.32

The following table gives the distribution of daily wages of 900 workers. However, the frequencies of the two classes 40-50 and 60-70 are missing. If the median of the distribution is ₹59.25, find the missing frequencies.

Wages (₹)	30-40	40-50	50-60	60-70	70-80
No. of Workers	120	?	200	?	185

Solution:

Let f_1 and f_2 be the missing frequencies of the classes 40-45 and 60-70 respectively. We can write the given distribution in the following tabular form :

Class Intervals Wages (₹)	Frequency (f)	Cumulative Frequency (C)
30-40	120	120
40-50	f_1	$120 + f_1$
50-60	200	$320 + f_1$

60-70	f_2	$320+f_1+f_2$
70-80	185	900

Since median is given to be 59.25, the median class is 50-60. Thus, we can write,

$$59.25 = 50 + \frac{450 - (120 + f_1)}{200} \times 10 = 50 + \frac{330 - f_1}{20}$$

Solving this for f_1, we get, $\qquad f_1 = 145$

Further, $\qquad\qquad\qquad f_2 = 900 — (120 + 145 + 200 + 185) = 250$

Merits and Demerits of Median

Merits:

- It is easy to understand and easy to compute. In case of individual observations or ungrouped data, it can be located even by inspection.

- It is rigidly defined.

- It is not much affected by extreme values.

- It can be determined even if the given distribution has open ends.

- It can also be located graphically.

- It can be computed for a qualitative data where ranking of items, according to the intensity of characteristics, is possible.

- It can be used in a situation when data is qualitative i.e., beauty, honesty etc.

Demerits:

- When the number of observations is large, the location of median requires their arrangement in ascending or descending order which is a very cumbersome process.

- It is a positional average and hence is not capable of being treated mathematically.

- For a grouped frequency distribution, the median is only estimated but not located.

- It is not based on all the observations and therefore cannot be regarded their good representative.

- Since weighted median, like weighted arithmetic mean, is not defined, it is not a suitable average when all the items are not of equal importance.

MODE

It is that value of the variant which repeats maximum number of times in a distribution and around which other observations are densely distributed. In the words of A.M. Tuttle, "Mode is the value which has the greatest frequency density in its immediate neighborhood".

If the frequency distribution is fairly regular, then mode is given by the value corresponding to maximum frequency. A frequency distribution is said to be regular in which the frequencies first increase in a gradual manner, reach a peak value and finally decline gradually in, approximately, the same manner as the values of the variable increase (or decrease).

The concept of mode as a measure of average is preferable to mean and median when it is desired to know the most typical size e.g. the most common size of shoes, the most common size of a garment etc.

Determination of Mode

1. By Inspection

When a frequency distribution is regular then mode can be determined just by inspection.

Illustration 8.33

Compute mode of the following distribution

X	1	2	3	4	5	6	7	8	9	10
f	4	6	8	12	17	11	7	6	5	3

Solution:

Since the distribution is fairly regular, therefore, mode of the above distribution is = 5.

2. By Method of Grouping

This method is used when the distribution is not regular. This method is explained through the following illustration.

Illustration 8.34

Determine mode of the following distribution

X	10	11	12	13	14	15	16	17	18	19
f	8	15	20	100	98	95	90	75	50	30

Solution:

The above distribution is not regular because there is a sudden increase in frequency from 20 to 100. Therefore, mode cannot be located by inspection. To obtain mode, we write the following grouping table.

Determination of Mode

X	f (1)	(2)	(3)	(4)	(5)	(6)
10	8	23		43		
11	15		35		135	
12	20	120				218
13	**100**		**198**	**293**		
14	98	**193**			**283**	
15	95		185			**260**
16	90	165		215		
17	75		125		155	
18	50	80				
19	30					

The highest frequency total in each of the six columns of the above table is identified and analyzed in the following table.

Variables

Columns	10	11	12	13	14	15	16	17	18	19
1				1						
2					1	1				
3				1	1					
4				1	1	1				
5					1	1	1			
6						1	1	1		
Total	0	0	0	3	4	4	2	1	0	0

Since the values 14 and 15 are both repeated maximum number of times, therefore the mode is ill defined. In this case, we can find the value of mode by the use of the formula. Mode = 3 Median-2 Mean, to be discussed later in this Chapter.

We note that for the given distribution Mean = 15.09 and Median = 15. Therefore, Mode = 3 × 15-2 × 15.09 = 14.82

Note: If the most repeated values in the above analysis table were not adjacent, the distribution would have been bi-modal.

3. By Interpolation Formula

For a grouped frequency distribution, we can determine modal class either by inspection or by method of grouping.

The exact location of mode is then determined by the use of the following interpolation formula :

$$Z = L_m + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times h$$

Where,

Z	=	mode
L_m	=	lower limit of the modal class
f_m	=	frequency of modal class
f_1	=	frequency preceding modal class frequency
f_2	=	frequencies succeeding modal class frequency
h	=	width of the modal class.

Illustration 8.35

The frequency distribution of marks obtained by 60 students of a class is given below:

Marks	30-34	40-44	45-49	50-54	55-59	60-64
Frequency	3	5	12	18	14	62

Find mode of the distribution.

Solution:

The given class intervals should be converted into class boundaries. Since the distribution is regular, the modal class, by inspection, is 49.5 – 54.5

Further L_m = 49.5, f_m = 18, f_1 = 12, f_2 = 12, h = 5

$$\therefore \text{Mode}, Z = 49.5 + \frac{18-12}{(2 \times 18)-12-14} \times 5$$

$$= 49.5 + \frac{6}{36-26} \times 5$$

$$= 49.5 + \frac{6}{10} \times 5$$

$$= 52.5 \text{ marks}$$

Merits and Demerits of Mode

Merits:

- It is easy to understand and easy to calculate. Very often, it is located by just inspection.
- It can be located even when the variable is not measurable but categorization or ranking of observations is possible.
- It is not affected by extreme observations.
- It can be determined even for a distribution with open ends.
- Model value is a value around which there is a concentration of observations and therefore regarded as best representative of the observations.
- It can also be located graphically.

Demerits:

- It is not based on all the observations.
- It is not capable of further mathematical treatment.
- In certain cases, mode may not be rigidly defined and hence the important requisite of a good measure of central tendency is not satisfied.
- It is not easy to calculate unless the number of observations is sufficiently large and reveal a marked tendency of concentration around a particular value
- It is not suitable when different items of the data are of unequal importance.
- The mode of a distribution depends upon the choice of class width and hence is an unstable average.
- It is not suitable in the situation where weights are given to extreme values.
- Its value is most affected by fluctuations of sampling.

REVIEW QUESTIONS

1. The sum of square of derivatives from _____ is minimum.

2. Histogram is useful to determine the value of _____.

3. The relationship between mean, median and mode is _____.

4. State True or False

– Median can be computed in an open-end frequency distribution.

– Mean is not affected by the choice of assumed average.

Answers: 1. Mean 2. Mode 3. $M_o = 3M_d - 2\overline{X}$ 4. True, True

Comparative Merits and Demerits of Mean, Median and Mode

– All the three measures of averages are easy to understand and easy to compute. In certain situations, median and mode can be located just by inspection.

– Mean is highly affected while median is slightly affected by extreme observations. Mode is not affected by extreme observations.

– Mean is based on the magnitude of all the observations as opposed to median and mode.

– Only Mean is capable of algebraic treatment.

– If a distribution has open ends, we can compute median and mode but not mean.

– If a distribution has unequal class intervals, we can compute mean and median but not mode.

– All the three averages can be computed if the characteristics is measurable; median and mode can be computed if the characteristics is not measurable but items can be arranged in the order of its intensity, only mode can be computed if the characteristics is neither measurable nor can be arranged in order of intensity.

– Mean and median are rigidly defined while mode may not be rigidly defined in all situation.

– Like weighted arithmetic mean, weighted median or weighted mode is not defined.

Relation Between Mean, Median and Mode

A frequency distribution is said to be symmetrical about an ordinate (AB) if its frequency curve is divided into two parts, by this ordinate, such that one part is mirror image of the other. For a symmetrical distribution, we have Mean = Median = Mode, as shown in figure:

If a distribution is not symmetrical, it is termed as skewed. Empirically, it has been found that, for a moderately skewed distribution, we have

$$\overline{X} - M_o = 3\,(\overline{X} - M_d) \text{ or } M_o = 3M_d - 2\overline{X}$$
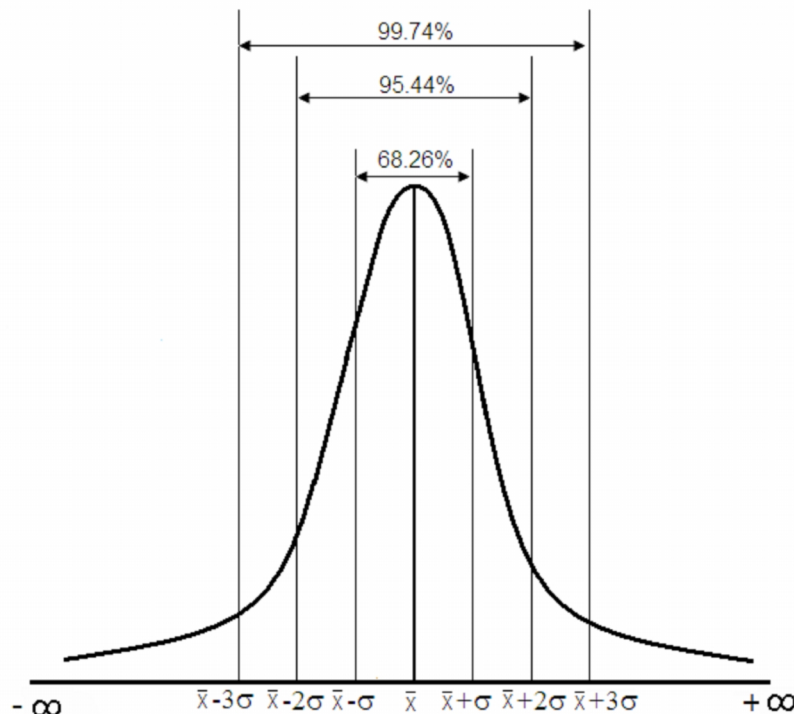
Thus, it may be noted that $Z = 3M_d - 2\overline{X}$

Illustration 8.36

In a moderately a symmetrical distribution–

    (a) The mode and median are 75 and 60 respectively, find mean.

    (b) The mean and mode are 50 and 37.5 respectively, find median.

Solution:

    (a) Using the empirical relation between mean, median and mode, we can write

$$\overline{X} - M_o = 3\,(\overline{X} - M_d)$$

or $2\overline{X} = 3\,M_d - M_o$

or $\overline{X} = \dfrac{3M_d - M_o}{2} = \dfrac{3 \times 60 - 75}{2} = 52.5$

    (b) Also, $M_d = \dfrac{2\overline{X} + Mo}{3} = \dfrac{2 \times 50 + 37.5}{3} = 45.83$

Choice of a Suitable Average

The choice of a suitable average depends upon the following three broad considerations :

1. **Considerations based on the Suitability of Data for an Average:** The nature of the given data may itself indicate the type of average that could be used. For example, if the characteristics are not measurable, it is not possible to compute mean. Similarly, if the distribution is skewed, median may represent the data more appropriately because mean is highly affected by extreme observations and mode is not affected at all. Further, if the class intervals are of unequal width, mean and median can be satisfactorily calculated. However, an approximate value of mode can be computed by making class intervals of equal width under the assumption that observations in a class are uniformly distributed. The accuracy of the mode will depend upon the validity of this assumption. If a distribution has open ends, we can compute mode and median. An approximate value of mean can be computed only under the assumption that the width of first class is equal to the width of second class and the width of last class is equal to the width of last but one class.

2. **Considerations based on the Purpose of Investigation:** If the purpose of investigation is to compute an average value (e.g. average income), then computation of mean is more appropriate. However, if we want to study the pattern of the distribution, e.g. number of persons having income between two values, the computation of median, quartiles, deciles and percentiles etc. might be more appropriate. In contrast to this, mode is computed when we want to determine the most common or modal size of the distribution.

3. **Considerations based on the Merits of an Average:** The presence or absence of various characteristics of an average may also affect its selection in a given situation. If the requirement is that the average should be rigidly defined, then mean or median should be chosen in preference to mode. Similarly, if it is desired to compute an average based on all the observations, we select mean. Further, if the requirement is that an average should not be affected by extreme observations, then mode is most appropriate etc.

MEASURES OF DISPERSION

Statistical average or the measures of central tendency gives us an idea of the concentration of the observations about the Central parts of the distribution but it does not tell us the full story. It is hardly fully representative of a mass unless we know the manner in which the individual items scatters around it. For instance, if we are given only the averages of a series of observations, we cannot form complete idea about the distribution since there

may exist a number of distributions whose averages are same but which may differ widely from each other in number of ways. Thus, it is necessary that the measures of statistical averages must be supported and supplemented by some additional summary measures to describe a distribution. One such measure is dispersion or a measure of variability of observations.

Meaning and Definition of Dispersion

Dispersion means the variation or scattering or deviation of the different values of a variable from their average. In other words, it is the extent of scatteredness of items around a measure of central tendency. It indicates the extent to which the values vary among themselves. Some of the important definitions of dispersion given by different experts are as under:

"Dispersion is the measure of variations of items". **A.L. Bowley**

"Dispersion is the measure of extent to which individual items vary". **L.R. Connor**

"The measure of the scatterdness of the mass of figures in a series about an average is called the measure of variation or dispersion". **Simpson and Kafka**

"The degree to which numerical data tend to spread about an average value is called the variation or dispersion of the data". **Spiegel**

Objectives and Significance of Dispersion

Some of the important objectives of dispersion along with their significance are:

1. **To test the reliability of an average:** The reliability of a average depends upon the extent of dispersion in the data. A low value of dispersion implies greater degree of homogeneity among the observations and hence more reliable average.

2. **To compare the extent of variability in two or more distributions**: Using a measure of dispersion, we can compare the degree of homogeneity of observations in different distributions. A distribution having lower value of dispersion is said to be more uniform or consistent.

3. **To facilitate the computation of other statistical measures:** Various measures of dispersions are used in the computations of important statistical measures like correlation, regression, test statistics, confidence intervals, control limits etc.

4. **To serve as a basis for control of variations:** The knowledge about the dispersion may be utilized to control variations in various areas of medicines, industry, incomes etc.

Characteristics of Good Measure of Dispersion

A good measure of dispersion should possess the following characteristics:

– It should easy to calculate and easy to understood.

– It should be rigidly defined.

– It should be based on all the observations.

– It should be capable of further mathematical treatment.

– It should not be in duly affected by extreme observations.

– It should not be much affected by the fluctuations of sampling.

Types of Dispersion

Various measures of dispersion can be classified into two broad categories:

1. The measures which express the variations in observations in terms of differences between certain selected observations e.g. range, inter quartile range, inter percentile range etc.

2. The measures which express the variations in observations in terms of average of deviations from a central value e.g. mean deviation, standard deviation etc.

Both categories of measure can be further divided into absolute and relative measures of dispersion. An absolute measure is expressed in terms of the unit of measurement of the variable. As opposed to this, a relative measure of dispersion, popularly known as co efficient of dispersion is expressed as a pure number, independent of units of measurement of the variable. Relative measures are often used for comparison of dispersion of various distributions.

Methods of Dispersion

The following are some of the important methods measures of dispersions :

– Mean Deviation

– Standard Deviation

– Range

– Coefficient of Variation

Mean Deviation

Mean deviation is also known as average deviation. It is the average difference between the items in a distribution and the mean of that series. There is an advantage in taking the deviations from median, however, in practice, the arithmetic mean is more frequently used in calculating the mean deviation.

Computation of Mean Deviation

If X1, X2, X3 and Xn are N given observations, then the Mean Deviation will be

$$\sum \left| X - X \right| \ / \ N = \frac{\sum \left| X - X \right|}{N}$$

= Mean Devation from Median

$$\sum \left| X - X \right| \ / \ N = \frac{\sum \left| X - M \right|}{N}$$

Coefficient of Mean Deviation

= Mean Deviation / Median

Illustration 8.37

Find the Mean Deviation of 3,6,6,7,8,11,15,16

Step 1

Find the Mean = $\sum$X/N

= (3+6+6+7+8+11+15+16) / 8

= 72/8

=9

Step 2

Find the distance of each value from that mean

Value	Distance from Mean
3	**6**
6	3
6	3
7	2
8	1
11	2
15	6
16	7

Mean Deviation = 6+3+3+2+1+2+6+7 / 8

= 30/8

=3.75

In the above example Mean = 9 and Mean Deviation = 3.75

STANDARD DEVIATION

Since mean deviation is the arithmetic mean of the absolute values of deviations, it is not very convenient to be algebraically manipulated. This necessitates a search for a measure of dispersion which is capable of being subjected to further mathematical treatment. One such measure is standard deviation or root-mean square deviation. Standard deviation may be defined as the positive square root of arithmetic mean of the squares of all the deviations of the values from their arithmetic mean. In brief, it refers to the square root of the mean of the squares of deviation from mean.

The standard deviation is usually denoted by Greek letter   , is given as

$$\text{For N individual observations } = \sqrt{1/N \ \left(X_i - \overline{X}\right)^2}$$

$$\text{For frequency distribution } = \sqrt{1/N \ f_i\left(X_i - \overline{X}\right)^2}$$

Note: We consider only positive square root of the above expression as standard deviation. Square of the standard deviation is known as variance of the distribution.

The above formulae are appropriate if X is a whole number. If X is not a whole number, the standard deviation is conveniently computed by using transformed form of the above formula, given below:

$$\sigma \quad = \quad 1/N \ \Sigma(X_i - \overline{X})^2$$

$$= \quad 1/N \ \Sigma(X_i - \overline{X})(X_i - \overline{X})$$

$$= \quad 1/N \ \Sigma(X_i^2 - X_i\overline{X}) - \overline{X}/N \ \Sigma(X_i - \overline{X})$$

Here in above equation, the second term is sum of deviations from $\overline{X}$ which is equal to zero.

$$= \quad 1/N \ (\Sigma X_i^2 - \overline{X}^2)$$

$$= \quad 1/N \ \Sigma X_i^2 - (\Sigma X_i/N)^2$$

$$= \text{Mean of Squares - Square of the mean}$$

Similarly, for a frequency distribution, we can write —

$$\sigma^2 = \left[\frac{\Sigma f_i X_i^2}{\Sigma f_i} - \left(\frac{\Sigma f_i X_i}{\Sigma f_i} \right)^2 \right]$$

The following illustration will help us understand the concept of standard deviation better.

Illustration 8.38

Calculate standard deviation of the following data

X	10	11	12	13	14	15	16	17	18
f	2	7	10	12	15	11	10	6	3

Solution:

X	f	fX	X - Mean	$f(X-\overline{X})^2$
10	2	20	-4	32
11	7	77	-3	63
12	10	120	-2	40
13	12	156	-11	2
14	15	210	0	0
15	11	165	1	11
16	10	160	2	40
17	6	102	3	54
18	3	54	4	48
Total	**76**	**Σ fX= 1064**		**$\Sigma f(X-\overline{X})^2 = 300$**

$$\overline{X} = 1064/76 = 14$$
$$\sigma_x = \sqrt{300/76}$$
$$= \sqrt{3.95}$$
$$= 1.99$$

Short-cut Method

This method is based on an important property of variance (or standard deviation), given below:

The variance of a distribution is independent of the change of origin but not of change of scale.

Change of Origin

If a constant value, say, A is subtracted from each observation $X_1, X_2 \ldots \ldots X_n$, then resulting observation X_1-A, $X_2-A, \ldots \ldots X-A$ are said to be measured from A. This operation is termed as shifting of origin at A because $X_i - A = 0$ when $X_i = A$ (i = 1, 2 . . . . . n).

Note: That the origin of X_i values is at 0

In this case $\sigma_X^2 = \sigma_d^2$, when d = (X – A)

Change of Scale

To make change of scale, we can divide (or multiply) every observation by a suitable constant.

For example, if X_i are measured in inches, then $Y_i = X_i/12$ (i = 1, 2 . . . . . n) will be in feet. It is possible to make simultaneous change of origin and scale by making the transformation

$$dx_i \quad = \quad \frac{(X_i\text{-}A)}{h} \quad (i = 1, 2 . . . . . n)$$

where, A refers to change of origin and h refers to change of scale. It can be shown that:

$$\sigma_X^2 = h^2\sigma^2 dx_i$$

$$\sigma_X^2 = \left(\frac{(f_i dx_i)^2}{N}\right) - \left(\frac{(f_i dx_i)}{N}\right)^2 \times h^2$$

$$x = \sqrt{\frac{\{(f_i dx_i)^2\}}{N} - \left(\frac{(f_i dx_i)}{N = \Sigma f_i}\right)^2} \times h$$

Illustration 8.39

Calculate standard deviation of the following data:

Marks:	0-10	10-20	20-30	30-40	40-50
Frequency:	10	8	15	8	4

Solution:

Calculation of Standard Deviation

Marks	Mid-values (X)	f	dx_i = X- 25/10	f dx_if	dx_i^2
0-10	5	10	-2	-20	40
10-20	15	8	-1	-8	8
20-30	25	15	0	0	0
30-40	35	8	1	8	8
40-50	45	4	2	8	16
Total		45		-12	72

$$\sigma_x = \sqrt{\frac{72}{45} - \frac{144}{45 \times 45}} \times 10 = 12.36$$

Illustration 8.40

Find the mean and standard deviation of the following observations:

X = 2, 5, 7, 8, 13

Transform the above observations such that:

(i) The mean of the transformed series become 10, while the standard deviation remaining unchanged.

(ii) The mean of the transformed series becomes 10 and the standard deviation becomes double the standard deviation of the given series.

Solution:

$$\overline{X} = \frac{2+5+7+8+13}{5} = 7$$

$$\sigma = \sqrt{\frac{4+25+49+64+169}{5} - 49} = 3.63$$

(i) Since the standard deviation is not affected by change of origin, the transformed series with mean 10 and s.d. 3.63 is obtained by adding 3 to each observation. Thus, the required series is 5, 8, 10, 11, 16

(ii) If every observation of the series is multiplied by 2, the mean and s.d. of the new series would be 14 and 7.26. Then, the required series is obtained by subtracting 4 from each observation. Thus, the required series is 0, 6, 10, 12, 22

Illustration 8.41

The mean and standard deviation of 10 observations are 35 and 2 respectively. Find the changed mean and standard deviation if –

(i) each observation is increased by 4, and

(ii) each observation is multiplied by 5.

Solution:

(i) When each observation is increased by 4, the mean of the changed observations will also increase and it will become equal to 35 + 4 = 39.

Since the increasing of all observations by a constant is change of origin, the standard deviation shall not change.

Thus, the mean and standard deviation when each observation is increased by 4 would be 39 and 2 respectively.

(ii) When each observation is multiplied by 5, the mean of the new observations will be 35 x 5 = 175 and their standard deviation will be 2 x 5 = 10.

Merits and Demerits of standard Deviation

Merits :

- It is rigidly defined measure of dispersion.
- It is based on all the observations.
- It is capable of being treated mathematically.

Demerits :

- As compared to Range and Quartile Deviation etc., it is difficult to understand.
- It is difficult to compute.
- It is highly affected by extreme observations.
- Since the unit of standard deviation is same as the unit of observations, it cannot be used to compare the dispersions of the distributions expressed in indifferent units.

Range

Range is the simplest method of dispersion. Symbolically

Range = L –S

Where L = Largest item, and

      S = Smallest item.

It is defined as the difference between the value of the largest item and the value of the smallest item.

Coefficient of Range

$= (L - S) / (L + S)$

The relative measure corresponding to range is known as Coefficient of Range.

Illustration 8.42

The following are the prices of X commodity from Monday to Saturday; calculate the range and coefficient of range.

Day	Price of X
Monday	200
Tuesday	210
Wednesday	208
Thursday	160
Friday	220
Saturday	250

Range = L – S

L = 250 and S = 160,

Range = 250 – 160 = Rs. 90.

Coefficient of Range = L –S / L + S

= 250 – 160 / 250 + 160

= 90 / 410

= 0.22

In the above example range is 90 and coefficient of range is .22.

Coefficient of Range in Continuous Series

In case of continuous series, first we have to find the difference between the upper limit of the highest wage class and the lower limit of the lowest wage class.

Illustration 8.43

Marks	No. of Students
10-20	8
20-30	10
30-40	12
40-50	8

Coefficient of Range= L – S / L + S

= 50 – 10 / 50 + 10

= 40 / 60

= .667

Merits and Limitations of Range

The merits and limitations of range are as under:

Merits

- Amongst all the methods of studying dispersion range is the simplest method, and also the easiest to compute.
- It takes minimum time to compute the value of range.

Limitations

- Range is not based on each and every item of the distribution.
- Range cannot tell us anything about the character of the distribution within the two extreme observations.

Inter – Quartile Range

The range which includes the middle 50 percent of the distribution, is known as interquartile range. In other word one quarter of the observations at the lower end, another quarter of the observations at the upper end of the distribution are excluded in computing the interquartile range.

Inter – Quartile Range = $Q_3 - Q_1$

Quartile Deviation = $\left(\dfrac{Q_3 - Q_1}{2} \right)$

Quartile deviation is an absolute measure of dispersion. The relative measure corresponding to this measure, is known as the coefficient of quartile deviation.

Coefficient of Q.D = $\dfrac{\left(\dfrac{Q_3 - Q_1}{2} \right)}{\left(\dfrac{Q_3 + Q_1}{2} \right)}$

$= \left(\dfrac{Q_3 - Q_1}{Q_3 + Q_1} \right)$

Illustration 8.44

Find out the value of Q.D and Coefficient of Q.D.

Roll No.	1	2	3	4	5	6	7
Marks	20	28	40	12	30	15	50
Marks arranged in ascending order:	12	15	20	28	30	40	50

Q_1 = Size of (N +1)/4 th item

= Size of 7+1/4 = 2nd item

Size of 2nd item is 15. Thus Q$_1$ = 15

Q$_3$ = Size of 3(7+1)/4 = 6th item

Size of 6th item is 40.

Q.D = 40 -15 / 2

= 12.5

Coefficient of Q.D = Q3-Q1 / Q3 + Q1

= 25/55 = .455

Co-efficient of Variation (CV)

The main purpose of finding coefficient of variation (often abbreviated as CV) is used to study of quality assurance by measuring the dispersion of the population data of a probability of frequency distribution, or by determining the content or quality of the sample data of substances. The method of measuring the ratio of standard deviation to mean is also known as relative standard deviation often abbreviated as RSD. It only uses positive numbers in the calculation and expressed in percentage values. Therefore, the resultant value of this formula CV [Standard Deviation (σ)/Mean (X) will be multiplied by 100. CV is important in the field of probability & statistics to measure the relative variability of the data sets on a ratio scale. In probability theory and statistics, it is also known as unitized risk or the variance coefficient.

The standard deviation is an absolute measure of dispersion. A relative measure, known as co-efficient of disperson or variation, is given by

$$C.V = \frac{\sigma}{\overline{X}} \times 100$$

This measure is used to compare variability or homogeneity or stability or uniformity or consistency of two or more sets of data. The data having a higher value of the co-efficient of variation is said to be more dispersed or less uniform etc.

Note: The dispersion of two or more distributions can be compared on the basis of their standard deviations only if their means are equal and the units of measurements are same.

Illustration 8.45

The mean and standard deviations of two brands of light bulbs are given below:

	Brand I	**Brand II**
Mean	800 hours	770 hours
Standard deviation	100 hours	60 hours

Which brand of bulbs is more reliable ?

Solution:

Since means of the two brands are different, their reliability can be compared by computing the co-efficient of variations.

Coefficient of Variation = $\frac{\sigma}{\overline{X}} \times 100$

Putting the available information in the above formula, we get as under:

C.V of brand I = (100/800) x 100 = 12.5%

C.V of brand II = (100/700) x 60 = 7.79%

Since C.V of brand II is less, therefore, mean life of brand II is more reliable than mean life of brand I.

BI – VARIATE ANALYSIS

Bivariate analysis is one of the simplest form of statistical analysis. It is related to the analysis of two variables (X and Y) for the purpose of determining the empirical relationship between them. Bivariate analysis is useful in testing the relationship (correlation and regression analysis). Bivariate analysis help us in determining the value for one variable, if we know the value of the other variable. There are mainly three types of bivariate analysis like, scatter plots, regression analysis and correlation analysis.

Covariance

Covariance indicates how two variables are related. A positive covariance means the variables are positively related, while a negative covariance means the variables are inversely related. The formula for calculating covariance of data is shown below :

$$COV(X_1 Y) = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{N}$$

x = the independent variable

y = the dependent variable

N = number of data points

X= the mean of the independent variable x

Y= the mean of the dependent variable y

To understand how covariance is used, consider the table below, which describes the rate of economic growth (x_i) and the rate of return on the S&P 500 (y_i).

Economic Growth % (x_i)	S & P 500 Returns % (Y_i)
2.1	8
2.5	12
4.0	14
3.6	10

Using the covariance formula, you can determine whether economic growth and S&P 500 returns have a positive or inverse relationship. Before you compute the covariance, calculate the mean of x and y.

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{N}$$

$$\overline{X} = \frac{2.1 + 2.5 + 4.0 + 3.6}{4}$$

$$= \frac{12.2}{4}$$

$$= 3.1$$

$$\overline{y} = \frac{\sum\limits_{i=1}^{n} Y_i}{N}$$

$$\overline{y} = \frac{8+12+14+10}{4}$$

$$= \frac{44}{4}$$

$$= 11$$

Now you can identify the variables for the covariance formula as follows :

x = 2.1, 2.5, 4.0, and 3.6 (economic growth)

y = 8, 12, 14, and 10 (S&P 500 returns)

X= 3.1

Y= 1

Substitute these values into the covariance formula to determine the relationship between economic growth and S&P 500 returns.

$$COV(X_1Y) = \frac{\sum\limits_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{n}$$

$$= \frac{(2.1-3.1)(8-11)+....}{4}$$

$$= \frac{(-1)(-3)+(-0.6)(1)+(0.9)(3)+....}{4}$$

$$= \frac{3+(-0.6)+2.7+(-0.5)}{4}$$

$$= \frac{4.6}{4}$$

$$= 1.15$$

The covariance between the returns of the S&P 500 and economic growth is 1.15. Since the covariance is positive, the variables are positively related – they move together in the same direction.

Coefficient of Correlation

In practical world we have a large number of problems which deals with two or more than two variables. If change in one variable is accompanied by change in other variable, both the variables are known as correlated. There are number of examples, which shows such kind of relationship like, age and playing habit, rainfall and production of crop and so on. The degree of relationship is measured though the correlation analysis. The measure of correlation is known as correlation coefficient. Some of the important definitions of correlation are as under:-

1. According to A. M. Tuttle," Correlation is an analysis of the covariation between two or more variables."

2. According to YaLun Chou, "Correlation analysis attempts to determine the ' degree of relationship' between variables."

3. Prof. Simpson and Kafka says, "Correlation analysis deals with the association between two or more variables."

Significance of Correlation Analysis

Correlation analysis has a vital role to play in our practical life because of the following reasons:

With the help of correlation analysis we can measure the degree of relationship between the variables.

With the help of one variable, we can estimate the value of another variable. For this we use another tool of statistics, that is known as regression analysis.

Correlation analysis also helps in business decisions, it helps in the estimation of costs, prices and sales on the basis of functionally related variables.

Types of Correlation

Correlation is described in several ways, three of the most important ways are as under:-

1. Positive and Negative Correlation

2. Simple and Multiple Correlation

3. Linear and Non- Linear Correlation

Positive and Negative Correlation

Whether the correlation is positive or negative, it depends on the change in the direction of variables. If both the variables are increasing or decreasing or in other words if the direction of cxhange in both the variables is same, correlation is said to be positive. On the other hand if the variables are varying in opposite directions, correlation is said to be negative.

Positive Correlation

X	Y
10	16
15	25
20	35
25	48

Negative Correlation

X	Y
10	30
20	25
30	18
40	10

Simple and Multiple Correlation

If the problem is related to only two variables it comes under simple correlation. But if there are three or more than three variables are studied it is a problem of multiple correlation. For example, if we study the relationship between rainfall and the production of wheat, it is related to simple correlation. But if we include one more variable, for example the amount of fertilizer used, now it becomes the problem of multiple correlation. Here our discussion is restricted to the simple correlation only from the point of view of syllabus.

Linear and Non-Linear Correlation

If the amount of change in one variable tends to bear constant ratio to the amount of change in the other variable then the correlation is said to be linear. On the other hand, if the amount of change in one variable does not bear a constant ratio of the amount of change in the other variable then the correlation is said to be Non-Linear.

POSITIVE LINEAR CORRELATION   **CURVI-LINEAR CORRELATION**
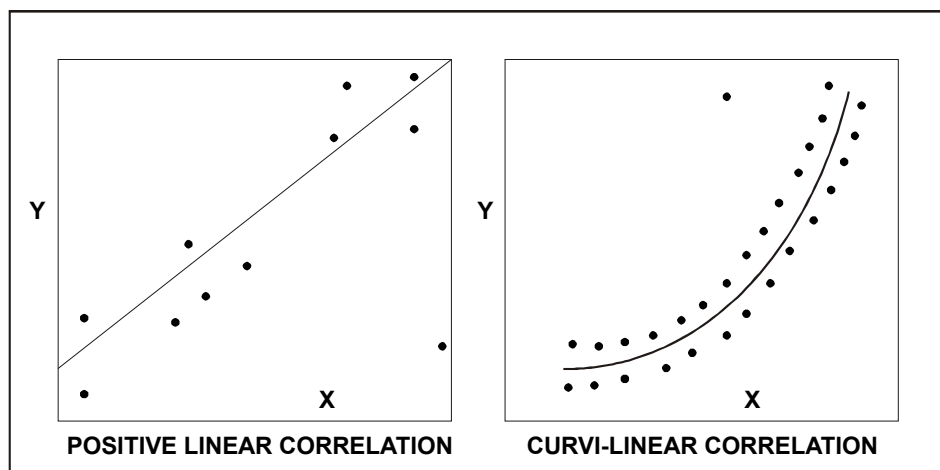
Methods of Studying Correlation

There are number of methods to ascertaining whether two variables are correlated or not. Some of the important methods are as under:-

1. Scatter Diagram Method

2. Karl Pearson's Coefficient of Correlation

3. Spearman's Rank Correlation Coefficient

Above three methods, first method is based on the graphs and on the knowledge of diagrams. But rest of the two methods are mathematical methods. Let's discuss in detail,

Scatter Diagram Method

This is the simplest method to determine whether two variables are Positively related, Negatively related or Unrelated. To find out the nature of relationship we prepare a dot chart, this dot chart is known as scatter diagram.

When the given data is plotted on a graph paper in the form of the values of X and Y, we got many points as the number of observations. By looking at various points on graph paper we can get the idea whether the variables are related or not. The greater the scatter of the plotted points on graph paper the lesser is the relationship between the two variables.
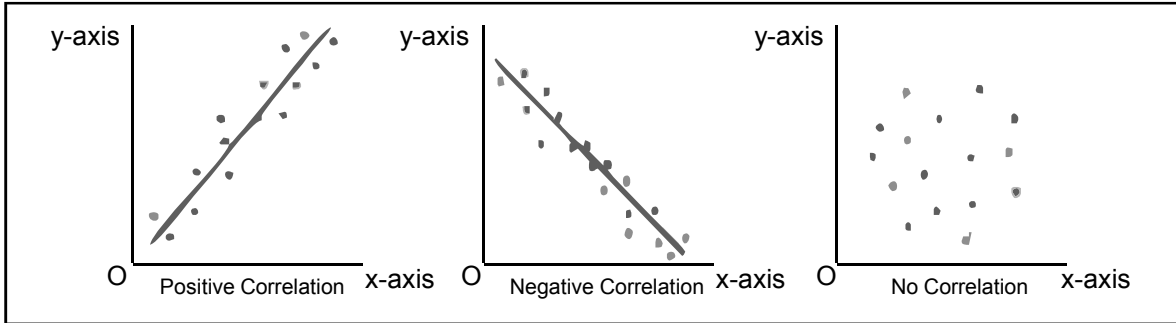
The more closely the points on graph paper come to a straight line falling from the lower left hand corner to the upper right hand corner, correlation is said to be positive. If all the points are lying on a straight line rising from the upper left hand to the lower right hand corner, correlation is said to be negative. Or the other hand if the plotted points on a graph paper are in a haphazard manner, it shows no correlation between variables.

Illustration 8.46

By putting the individual value of the two variables, we obtain two curves. One is for variable X and second for variable Y. If the direction of both the curves are moving in the same direction, correlation is said to be positive. On the other hand, if the curves are moving in the opposite direction correlation is said to be negative.

Year	Average Income(in Rs)	Average Expenditure(in Rs)
2005	3100	3000
2006	3320	3200
2007	3500	3350
2008	3700	3500
2009	4200	4000
2010	4300	4100
2011	4500	4200
2012	4650	4400
2013	4800	4650
2014	**5000**	**4500**
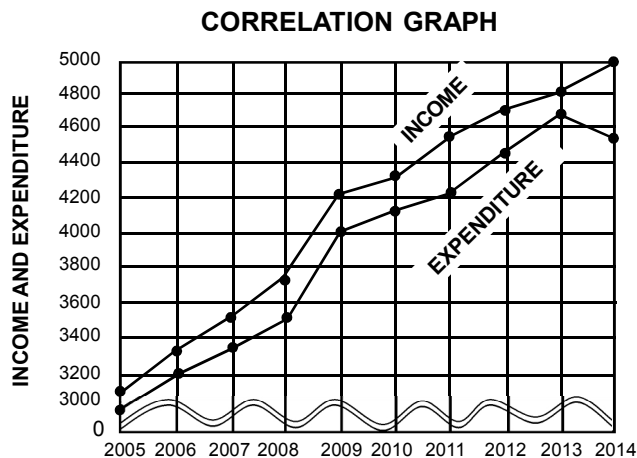
CORRELATION GRAPH

This method of correlation is normally used in case of time series, where the data is given over a period of time. The above graph shows that both the variables, income and expenditure are correlated.

Karl Pearson's Coefficient of Correlation

The most widely used method in coefficient of correlation is Karl Pearson's Coefficient of Correlation. The Pearson coefficient of correlation is denoted by the symbol r*. There are two methods to calculate the value of Pearson coefficient of correlation.

Direct Method

Short Cut Method

A. Direct Method

Type I : This method is used when given variables arc small in magnitude,

Formula : $r = \dfrac{N\sum XY}{\sqrt{N\sum X^2 - (\sum X)^2} \quad \sqrt{N\sum Y^2 - (\sum Y)^2}} \dfrac{\sum X \sum Y}{}$

Illustration 8.47

Calculate Karl Pearson's coefficient of correlation between

Age (years) : 1   2   3   4   5

Weight (kg.) : 3   4   6   7   12

Solution : $\sum X = 15 : \sum Y = 32 : \sum X^2 = 55 : \sum Y^2 = 254 : \sum XY = 117$

Age (X)	Weight	X²	Y²	XY
1	3	1	9	3
2	4	4	16	8
3	6	9	36	18
4	7	16	49	28
5	12	29	144	60
15	32	55	254	117

$Asr = \dfrac{N\sum XY}{\sqrt{N\sum X^2 - (\sum X)^2} \quad \sqrt{N\sum Y^2 - (\sum Y)^2}} \dfrac{\sum X \sum Y}{}$

$\therefore r = \dfrac{5 \times 117}{\sqrt{5 \times 55 - (15)^2}} \dfrac{15 \times 32}{\sqrt{5 \times 254 - (32)^2}}$

$= \dfrac{585 \times 480}{\sqrt{275 = 225}\sqrt{1270 = 1024}} = \dfrac{105}{\sqrt{50 \times 246}} = \dfrac{105}{\sqrt{12300}} = \dfrac{105}{\sqrt{11.90}} = 0.9467\,Ans.$

Type II : It is direct formula to find r. This formula can effectively be used where $\overline{X}$ and $\overline{Y}$ is not infractions. The formula is

$r = \dfrac{\sum XY}{\sqrt{\sum x^2 . \sum y^2}}$: where dx is the deviation of X variable from its $\overline{X}$

y is the deviation of Y variable from its $\overline{Y}$.; xy is the product of the two above dx² is the square of x²;y² is the square of dy.

Illustration 8.48

Calculate coefficient of correlation between death and birth rate for the following data.

Birth Rate	24	26	32	33	35	30
Death Rate	15	20	22	24	27	24

Solution :

Birth Rate X	Death Rate Y	$(X-\bar{X})$ = x	$(Y-\bar{Y})$ = y	$(X-\bar{X})^2$ = x²	$(Y-\bar{Y})^2$ = y²	$(X-\bar{X})(Y-\bar{Y})$ = xy
24	15	-6	-7	36	49	42
26	20	-4	-2	16	4	8
32	22	2	0	4	0	0
33	24	3	2	9	4	6
35	27	5	5	25	25	25
30	24	0	2	0	4	0
$\Sigma X = 180$	$\Sigma Y = 132$	$\Sigma x = 0$	$\Sigma y = 0$	$\Sigma x^2 = 90$	$\Sigma y^2 = 86$	$\Sigma xy = 81$
$\bar{X} = \dfrac{180}{6} = 30$	$\bar{Y} = \dfrac{132}{6} = 22$					

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 . \Sigma y^2}} = \frac{(81)}{\sqrt{90 \times 86}} \frac{81}{\sqrt{7740}} \frac{81}{87398} = .92$$

B. Short Cut Method

In case the mean is a whole number above method is simple. But when the mean is in fractions, short-cut method is used. In this method, the deviations are calculated from assumed mean and the following formula is applied.

$$r = \frac{N\Sigma dxdy}{\sqrt{N\Sigma dx^2 - (\Sigma dx)^2}} \frac{\Sigma dx \Sigma dy}{\sqrt{N\Sigma dy^2 - (\Sigma y)^2}}$$

Where

Σdx = Sum of deviations of X series from its Assumed Mean i.e. Σ (X-Ax)

Σdy = Sum of deviations of Y series from its assumed Mean i.e. Σ (Y-Ay)

Σdx^2 = Sum of squared deviations of X Series from its Assumed Mean i.e. Σ $(X-A_x)^2$

Σdy^2 = Sum of squared deviations of Y Series from its Assumed Mean i.e. Σ $(Y—A_y)^2$

$\Sigma dxdy$ = Sum of products of deviations of X and Y series from their respective assumed means. $\Sigma dxdy = \Sigma (X - A_x) (Y - A_y)$

N = Number of pairs

$\bar{Y} \bar{X}$

Illustration 8.49

Calculate coefficient of correlation between X series and Y series using Karl Pearson's Method.

X :	14	12	14	16	16	17	16	15
Y :	13	11	10	15	15	9	14	17

Solution : $\sum dx = 0 : \sum dx^2 = 18 : \sum dy = 8 : \sum dy^2 = 62 : \sum dxdy = 6$

X	Y	dx =X-A$_x$	dx²	dy =Y-A$_Y$	dy²	dxdy
14	13	-1	1	-1	I	1
12	10	-3	9	-3	9	9
14	10	-1	1	-4	16	4
16	15	1	1	1	1	1
16	15	1	1	1	1	1
17	9	2	4	-5	25	-10
16	14	I	1	0	0	0
15	17	0	0	3	9	0
	1	0	18	-8	62	6

Let A$_x$ = 15 and A$_y$ = 14

$$r = \frac{N\sum dxdy}{\sqrt{N\sum dx^2 - (\sum dx)^2}} - \frac{\sum dx \sum dy}{\sqrt{N\sum dy^2 - (\sum dy)^2}}$$

$$= \frac{8 \times 6 - (0) \times (-8)}{\sqrt{8 \times 18 - (0)^2} \times \sqrt{8 \times 62(-8)^2}}$$

$$= \frac{40 - 0}{\sqrt{(144 - 0)}\sqrt{(496 - 64)}} = \frac{48}{\sqrt{144 \times 432}} = \frac{48}{\sqrt{62208}} = \frac{48}{\sqrt{249.41}} = .192 = .19$$

Imp. Note ; Above formula can also be given the shape as following

$$r = \frac{\sum dxdy - \frac{\sum dx . \sum dy}{N}}{\sqrt{\sum dx^2 - \frac{(\sum dx)^2}{N}}\sqrt{\sum dy^2 - \frac{(\sum dy)^2}{N}}}$$

Spearman's Rank Correlation Coefficient

When the population is not normal or the shape of the distribution is not known, to avoid any assumption about the population being studied by ranking the observations according to size and basing the calculations on the ranks and not on the original observations. In such cases where data are ranked data we use speciman's rank correlation.

Spearman's rank correlation coefficient is defined as :

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

The value of coefficient of correlation is interpreted in the same manner as in pearson's coefficient of correlation, ranges between -1 to +1. let's discuss the Spearman's coefficient of correlation in detail with the help of an example:

Illustration 8.50

English (Marks)	Maths (Marks)	Rank (English)	Rank (Maths)	D (R_1 - R_2)	D^2
56	66	9	4	5	25
75	70	3	2	1	1
45	40	10	10	0	0
71	60	4	7	3	9
62	65	6	5	1	1
64	56	5	9	4	16
58	59	8	8	0	0
80	77	1	1	0	0
76	67	2	3	1	1
61	63	7	6	1	1

Where D= difference between ranks and D2 = difference squared.

After putting the value of n, D and D^2 in the formula we get,

$$r_s = 1 - \frac{6\sum D^2}{n(n^2 - 1)}$$

$$r_s = 1 - \frac{6 \times 54}{10(10^2 - 1)}$$

$$r_s = 1 - \frac{324}{990}$$

$$r_s = 1 - 0.33$$

$$r_s = 0.67$$

LESSON ROUND UP

– Statistics in the sense of numerical data basically refers to quantitative aspects and is in the form of numerical data. Statistics as statistical method refers to a body of knowledge which contains statistical principles and methods used for collection, analysis and interpretation of data.

– Statistics as statistical methods defined as a discipline which is concerned with collection, organization, presentation, analysis and interpretation of numerical data.

– There are two important laws of statistics i.e the law of statistical regularity; and the law of a large numbers.

– Collection of data refers to systematic recording of results either by counting or by enumeration. Primary

data is the data, which is collected for the first time by investigators or enumerators working under his supervision to serve a particular purpose. Such a data is of original nature. Secondary data refers to the data which is originally collected and published by the authorities other than who require it.

– Classification of data means grouping of related facts into classes. Classification of data depends upon characteristics of data, and objectives of the inquiry.

– Tabulation refers to systematic and logical arrangement of data in columns and rows in accordance with some salient features and characteristics. A tabular arrangement of data with corresponding frequency is known as frequency distribution.

– Presentation of data is to graphically represent the classified and tabulated data. There are number of diagrams which can be used for the presentation of data. The selection of a particular diagram depends upon the nature of data, objective of presentation and the ability and experience of the person doing this task. Some of the important types of diagram include line diagram, Simple Bar Diagram, Multiple Bar Diagram, Sub-divided or Component Bar Diagram, Circular or Pie Chart and scatter diagram.

– A measure of Central Tendency or an average of a distribution is a single figure which can be regarded as a representative of the whole distribution. An average is a single value within the range of data that is used to represent all values in the series. Since an average is somewhere within the range of the data, it is also called a measure of Central Value.

– The objectives and significance of statistical averages include presentation of huge mass of data in summarized form, facilitation of comparison, help indecision makingand establishing the precise relationships.

– There are two types of averages i.e. mathematical and positional. Mathematical averages include the measures such as arithmetic mean, geometric mean and harmonic mean whereas positional averages include measures like median and mode.

– Bivariate analysis is one of the simplest form of statistical analysis. It is related to the analysis of two variables for the purpose of determining the empirical relationship between them.

– Covariance indicates how two variables are related. A positive covariance means the variables are positively related, while a negative covariance means the variables are inversely related.

– If change in one variable is accompanied by change in other variable, both the variables are known as correlated. There are number of examples, which shows such kind of relationship like, age and playing habit, rainfall and production of crop and so on. The degree of relationship is measured though the correlation analysis.

GLOSSARY

Statistics A type of mathematical analysisinvolving the use of quantified representations, models and summaries for a given set of empirical data or real world observations. Statistical analysis involves the process of collecting and analyzing data and then summarizing the data into a numerical form.

Statistical enquiry Statistical enquiry refers to a process of collection of data and presentation of the collected data in meaningful ways.

Analytical Statistics The branch of statistics that is used in analyzing the data in terms of drawing meaningful conclusions, comparing two sets of data etc.

Inductive Statistics The branch of statistics that deals with generalizations, predictions, estimations, and decisions from data initially presented.

Inferential Statistics	The branch of statistics that draws conclusions that extend beyond the immediate data alone. Applied Statistics
	The branch of statistics that extends to practical application of statistics, as opposed to the study of its theoretical basis.
Descriptive Statistics	Any calculation which attempts to provide a concise summary of the information content of data that is collection,
Hypothesis	A tentative explanation for an observation, phenomenon, or scientific problem that can be tested by further investigation.
Data	Data collected in original form for statistical enquiry.
Census	The collection of data from every element in a population or universe or arena of statistical enquiry.
Sample	The collection of data from subgroup or subset of the population.
Discrete Variable	Variables (data) which assume a finite or countable number of possible values, usually obtained by counting.
Frequency	The number of times a certain value or class of values occurs.
Cumulative Frequency	The running total of the frequencies at each class interval level.
Frequency Distribution	The organization of raw data in table form with classes and frequencies.
Class Limits	Separate one class in a grouped frequency distribution from another. The limits could actually appear in the data and have gaps between the upper limit of one class and the lower limit of the next.
Class Width	The difference between the upper and lower boundaries of any class.
Class Boundary	It is the process of making the distribution continuous. The lower class boundary is found by subtracting 0.5 units from the lower class limit and the upper class boundary is found by adding 0.5 units to the upper class limit.
Class Mark	The middle point of the class. It is found by adding the upper and lower limits and dividing by two. It can also be found by adding the upper and lower boundaries and dividing by two.
Histogram	A graph which displays the data by using vertical bars of various heights to represent frequencies. The horizontal axis can either be the class boundaries, the class marks, or the class limits.
Frequency Polygon	A line graph. The frequency is placed along the vertical axis and the class midpoints are placed along the horizontal axis. These points are connected with straight lines.
Frequency Curve	A smooth frequency polygon.
Cross Tabulation	Cross tabulation is a useful data analysis technique. Data referring to a particular can be broken down by reference to one or more other variables through successive cross tabulations.
Ogive	A frequency polygon of the cumulative frequency or the relative cumulative frequency.
Pie Chart	Graphical depiction of data as slices of a pie. The frequency determines the size of the

	slice. The number of degrees in any slice is the relative frequency times 360 degrees.
Scatter Diagram	To find out the nature of relationship we prepare a dot chart, this dot chart is known as scatter diagram.
Mean	Sum of all the values divided by the number of values.
Median	The midpoint of the data after being ranked (sorted in ascending order). There are as many numbers below the median as above the median.
Mode	The most frequent number. There may not be a mode; there may be one mode; or there may be many modes.
Weighted Mean	The mean when each value is multiplied by its weight and summed. This sum is divided by the total of the weights.
Mean Deviation	It is the average difference between the items in a distribution and the mean of that series.
Standard Deviation	A measure of dispersion calculated as the square root of the variance.
Range	It is defined as the difference between the value of the largest item and the value of the smallest item.
Coefficient of Variation	This measure is used to compare variability or homogeneity or stability or uniformity or consistency of two or more sets of data.
Bi-variate Analysis	It is related to the analysis of two variables for the purpose of determining the empirical relationship between them.
Covariance	Covariance indicates how two variables are related.
Coefficient of Correlation	Correlation analysis deals with the association between two or more variables.

MULTIPLE CHOICE QUESTIONS

1. **Statistics can be defined as –**

   (a) An art

   (b) Science

   (c) Both art as well as science

   (d) Neither art nor science

2. **Information collected for a specific purpose is known as-**

   (a) Experimental information

   (b) Primary data

   (c) Secondary data

   (d) External information

3. **The information that already exist somewhere is known as -**

   (a) Secondary data

   (b) External information

(c) Primary data

(d) Experimental data

4. **The relationship between mean, median and mode is -**

(a) Mode = 3 Median – 2 Mean

(b) Median = 2 Mean – 3 Mode

(c) Mode = 2 Mean – 3 Median

(d) Mean = 3 Median – 2 Mode

5. **Given Mean 30, Mode 30, the Median would be -**

(a) 40

(b) 25

(c) 35

(d) 30

6. **The formula to calculate the coefficient of quartile deviation is -**

(a) Q3 – Q1/Q3 + Q1

(b) Q3 +Q1/Q3-Q1

(c) Q3+Q1/Q3+Q1

(d) Q3-Q1/Q3-Q1

7. **A coefficient of correlation is computed to be +0.91 means that-**

(a) Coefficient of correlation can not have this value

(b) The relationship between two variables is strong but negative

(c) The relationship between two variables is strong and positive

(d) The relationship between two variables is weak.

8. **The formula to calculate the coefficient of correlation of Range is-**

(a) L+S / L-S

(b) L-S / L+S

(c) L+S / L+S

(d) L-S / L-S

9. **The value of coefficient of correlation lies between -**

(a) -1 and 0

(b) 0 and 1

(c) -1 and 1

(d) None of the above

10. **Suppose the variable x and y, opposes each other then the value of the coefficient of correlation will be -**

    (a) Negative

    (b) Positive

    (c) Zero

    (d) None of the above

Answer Kys : 1.(c), 2. (b), 3. (a), 4. (a), 5.(d), 6. (a), 7.(c), 8. (b), 9.(c), 10.(a)

Suggested Readings

1. D.C. Sancheti& V.K. Kapoor : Statistics, Theory, Methods & Applications

2. Fundamentals of Statistics - S.C. Gupta

3. Statistics, Theory, Method and Applicants - D.C. Sancheti & V.K. Kapoor

4. Business Statistics – Naval Bajpai

Lesson 9
Mathematics of Finance and Elementary Probability

LEARNING OBJECTIVES

Finance, one of the important fields for individuals, businesses, and organizations; addresses the ways in which monetary resources are to be raised, allocated, and used over time, taking into account the risks entailed in the projects. With the emergence of fields like actuarial science and insurance; scope of financial management in the organization have increased manifold. It includes a number of interrelating disciplines, including Statistics and Economics.

Introduction to theory of Probability is useful in understanding, studying, and analyzing complex real world systems. Probability theory can be used to model and develop complex real world systems. A good understanding of Probability theory is needed to develop simulations.

In this context, this lesson deals with the practical aspect of the interest mechanism, annuity and the basic introduction to Probability to acquaint students with terminologies of the theory and provides contextual understanding of the matter to the extent it may be useful in our day to day lives for decision making. Apart, it also has a stand in risk management in our professions, business and managerial aspects for smooth running of operations and decision making.

██

INTEREST

Interest is the consideration paid by the borrower to the money lender for the use of his money.

The concept of interest involves two persons:

– Money lender

– Borrower

Borrower: Person who takes the money for his use.

Money Lender: Person who gives the money in the hope of getting some consideration i.e. the money lender charges from the borrower an amount which is more than what he gives to him. The difference is called interest. The rate of interest is decided mutually by the two parties i.e., the lender and the borrower.

Money lender can be an individual person or a bank or some other financial institution whereas borrower can be a single person or a company.

The money lender gives his money to the borrower and at the time of return, charges more than what he initially lent. This extra amount charged by the money lender for the use of his money is called **interest**. The rate of interest is decided mutually by the lender and the borrower. The sum borrowed is called the **Principal** or Capital and the sum paid by the borrower at the end of the period is called **amount**. The period for which the money is used is called **time**.

Interest can be reckoned in two ways:

– Simple Interest

– Compound Interest

SIMPLE INTEREST

The principal remains the same for the entire period, the money remains with the borrower.

If ₹ P is the principal i.e., amount borrowed and r is the rate of interest per year and n, the number of years for which the money remains with the borrower, then

Simple Interest = P x r x n

= Principle x Rate x Time

If rate of interest is 'i' per rupee, per year.

Then,

Simple interest for n years = P x i x n

If the rate of interest is r% per year

Then,

Simple Interest for n years = $\dfrac{P \times r \times n}{100}$

Amount due after n years = $\left(P + \dfrac{\mathrm{Pr}\, n}{100} \right) = P\left(1 + \dfrac{r\, n}{100} \right)$ = Principal+Interest for n years   (i)

COMPOUND INTEREST

In case of compound interest, interest also earns interest. At the end of each conversion period, interest is added to the principal and the amount thus formed becomes the principal for the next conversion period.

Amount due after n years = $P\left(1+\dfrac{r}{100}\right)^n$ (ii)

Compound Interest = $P\left(1+\dfrac{r}{100}\right)^n - P$

Note: Please note the difference between Result (i) and (ii).

Conversion Period

The period of time for which the interest is calculated is called the conversion period.

The conversion period is decided by the borrower and the lender. The conversion period may be one year, 6 months, 3 months or any other period.

If the conversion period is one year, the amount due at the end of n years = $P\left(1+\dfrac{r}{100}\right)^n$

Compound Interest = $P\left(1+\dfrac{r}{100}\right)^n - P$

If the conversion period is 6 months, then for n years the number of conversion periods = 2n and if r is the annual rate of interest, then rate of interest for 6 months will be r/2. The amount due after n years or 2n conversion periods will be equal to;

$$= P\left(1+\dfrac{r/2}{100}\right)^{2n}$$

Similarly,

If the conversion period is three months, the amount due at the end of n years

$$= P\left(1+\dfrac{r/4}{100}\right)^{4n}$$

If there are m conversion periods in one year then,

Amount due at the end of n years = $P\left(1+\dfrac{r/m}{100}\right)^{mn}$

Illustration: 9.1

A sum of ₹1200 becomes ₹1323 in two years at compound interest compounded annually. Find the rate percent.

Solution:

Let r % be the rate per annum.

$$1323 = 1200\left(1+\dfrac{r}{100}\right)^2$$

$$\left(1+\frac{r}{100}\right)^2 = \frac{1323}{1200}$$

$$\Rightarrow 1+\frac{r}{100} = \sqrt{\frac{1323}{1200}} = 1.05$$

$$\Rightarrow \frac{r}{100} = 1.05 - 1$$

r = 0.05 x 100

= 5%

Illustration: 9.2

Find the compound interest for ₹ 1,000 for 4 years at 5% per annum when:

   (i) The interest is compounded annually

  (ii) It is compounded half-yearly

 (iii) It is compounded quarterly

Solution:

(i) $CI = 1000\left(1+\frac{5}{100}\right)^4 - 1000$

$$= 1000(1.05)^4 - 1000$$
$$= 1215 - 1000 = ₹215$$

(ii) $CI = 1000\left(1+\frac{5/2}{100}\right)^{4 \times 2} - 1000$

$$= 1000(1.025)^8 - 1000$$
$$= 1000(1.218) - 1000$$
$$= 1218 - 1000$$
$$= ₹218$$

(iii) $CI = 1000\left(1+\frac{5/4}{100}\right)^{4 \times 4} - 1000$

$$= 1000\left(1+\frac{1.25}{100}\right)^{16} - 1000$$

$$= 1000(1.0125)^{16} - 1000$$
$$= 1000(1.21989) - 1000$$
$$= 1219.89 - 1000$$
$$= ₹219.89 \cong ₹ 220$$

Illustration: 9.3

The difference between simple and compound interest on a sum for three years at 5% per annum is ₹76.30. Find the sum.

Solution:

Let the sum = ₹ P

Simple Interest for three years $= P \times \dfrac{5}{100} \times 3 = \dfrac{15P}{100} = 0.15P$

Compound Interest $= P\left(1 + \dfrac{5}{100}\right)^{3} - P$

$= P(1.05)^3 - P$

$= 1.15763P - P = 0.15763P$

Difference between compound interest and simple interest;

$\Rightarrow$ CI – SI $= 0.15763P - 0.15P$

$= 0.00763P$

$\Rightarrow 0.00763P = 76.30$

$\Rightarrow \quad P = \dfrac{76.30}{0.00763} = ₹10,000$

REVIEW QUESTIONS

(1) If the interest on a certain sum for first year at 5% per annum compounded is ₹25 then, the total interest at the end of second year will be:

  (a) ₹51.25

  (b) ₹50

  (c) ₹26.25

  (d) None.

(2) The difference in compound interest and simple interest for one year for ₹1000 at 10% per annum is :

  (a) ₹10

  (b) ₹5

  (c) ₹0

  (d) None.

Answer: 1.(a) 2. (c)

Time Value of Money

The money which is receivable at present has more value than the money receivable in the future. Hence, the relationship that exists between the value of money receivable at present and the value of money receivable at future is referred as time value of the money.

Compounding and Discounting

Let's discuss compounding and discounting in detail, first we'll discuss compounding.

Compounding

The method used to determine the future value of present investment is known as compounding. Or in simple words, if we invest some money today, what will be the amount we get at a future date.

Discounting

The method used to determine the present value of future cash flows is known as discounting. Or in simple words, what should be the amount we need to invest today, to get a specific amount in future.

Before discussing the compounding and discounting in detail, first we should discuss the concept of annuity.

ANNUITY

An annuity is a fixed sum paid at regular intervals under certain stated conditions. The interval may be either one year or half-year or quarter year or one month, etc., If nothing is mentioned about the period of interval, it is taken to be one year.

Types of Annuity

(i) *Annuity Certain*: Annuity payable for a fixed number of intervals is called annuity certain.

(ii) *Contingent Annuity*: An annuity whose payments continue till the happening of an event, the date of which cannot be foretold is called contingent annuity, i.e., the life insurance premium is paid periodically to the life insurance company by the insured person for certain number of years or till he dies whichever is earlier. In case, the insured person dies before the maturity of the policy, the periodic payments become contingent annuity and in this case, the contingent event is the death of the insured person. The monthly pension of the retired person continues till his death.

(iii) *Perpetual Annuity (or Perpetuity)*: An annuity which continues forever (infinite number of years) is called perpetual annuity.

(iv) *Immediate Annuity (or ordinary annuity)*: If the periodic payments are made at the end of each period; the annuity is called an immediate annuity or ordinary annuity.

(v) *Annuity Due*: If the periodic payments are made in advance at the beginning of each period, the annuity is called annuity due.

(vi) *Deferred Annuity*: If the period payments start only after a certain specified period (called deferred period), it is called a deferred annuity. A deferred annuity can be either a deferred immediate annuity or a deferred annuity due.

If the annuity is deferred for n years, the first payment will become due at the end of nth year or $(n+1)^{th}$ year depending on whether the deferred annuity is a deferred annuity due or a deferred immediate annuity.

(vii) *Forborne Annuity*: An annuity left unpaid for a certain number of years is called forborne annuity for that number of years.

(viii) *Periodic Payment/Uniform Annuity/Variable Unity*: The amount of each payment of annuity is called the periodic payment or periodic rent of the annuity. If the periodic payments are all equal, the annuity is called level of uniform annuity. If the payments change in every period, it is called variable annuity.

Payment Period

The time between two successive payment dates of an annuity is called its payment period or payment interval.

Term of an Annuity

The total time from the beginning of the first payment period to the end of the last payment period is called the term of the annuity.

Amount of an Annuity (Future Value of an Annuity)

The amount of an annuity is defined as the total worth of all the payments of the annuity at the time of conclusion (i.e., at the end of the term) of the annuity. The amount of an annuity can also be defined as the sum of the future values of each of the Periodic Payments.

Amount of Annuity = Sum of the Periodic Payments + Sum of the Interests on the Periodic Payments

Case I: Immediate Annuity or Ordinary Annuity

In this case, the payments are made at the end of each period i.e., the payment made at the end of the first period will earn interest for (n-1) periods at the end of n periods.

If r% per annum is the rate of interest and n the number of periods, A is the annuity payment (fixed sum) then ₹A

paid at the end of first period will become $₹A\left(1+\dfrac{r}{100}\right)^{n-1}$ at the end of n periods.

Total amount of the annuity at the end of n periods;

$$= A\left(1+\frac{r}{100}\right)^{n-1} + A\left(1+\frac{r}{100}\right)^{n-2} + A\left(1+\frac{r}{100}\right)^{n-3} + ----- + A$$

$$= A\left[\frac{\left(1+\dfrac{r}{100}\right)^{n}-1}{\dfrac{r}{100}}\right] = A\left[\frac{(1+i)^{n}-1}{i}\right] \text{ where; } \frac{r}{100} = i$$

Case II: Annuity Due

Annuity due i.e., payments are made in advance at the beginning of each period i.e., the payment made at the beginning of first period will earn interest for n periods and the total amount of the annuity:

$$= A\left(1+\frac{r}{100}\right)^{n} + A\left(1+\frac{r}{100}\right)^{n-1} ...... A\left(1+\frac{r}{100}\right)$$

$$= A\left(1+\frac{r}{100}\right)\left[\frac{\left(1+\dfrac{r}{100}\right)^{n}-1}{\dfrac{r}{100}}\right]$$

$$= A(1+i)\left[\frac{(1+i)^{n}-1}{i}\right] \text{ where, } \frac{r}{100} = i$$

PRESENT VALUE

The present value of a given sum of money due at the end of a certain period of time is that sum which if invested now at the given rate of interest accumulates to the given sum at the end of the period.

If P is the Principal [Present Value] and r% per period is the rate of interest, then by compounding the interest period wise, the amount 'A' due after n periods is given by:

$$A = P\left(1 + \frac{r}{100}\right)^n \Rightarrow P = \frac{A}{\left(1 + \frac{r}{100}\right)^n}$$

= Present Value of 'A' due at the end of n years

We say that the present value of amount A due at the end of n period at r% per annum is $\dfrac{A}{\left(1 + \dfrac{r}{100}\right)^n}$

The present value or capital value of an annuity is the sum of the present values of all the payments.

Present value of ₹ A paid in the begining of second period $= \dfrac{A}{1 + \dfrac{r}{100}} = \dfrac{A}{1+i}$

Present value of ₹ A paid in the begining of nth period $= \dfrac{A}{\left(1 + \dfrac{r}{100}\right)^{n-1}}$

Present Value of the Annuity

It can be defined as Sum of the present values of all the payments received in annuity

$$= A + \frac{A}{\left(1 + \dfrac{r}{100}\right)} + \frac{A}{\left(1 + \dfrac{r}{100}\right)^2} + \dots + \frac{A}{\left(1 + \dfrac{r}{100}\right)^{n-1}}$$

$$= A\left[1 + \frac{1}{1 + \dfrac{r}{100}} + \frac{1}{\left(1 + \dfrac{r}{100}\right)^2} + \dots + \frac{1}{\left(1 + \dfrac{r}{100}\right)^{n-1}}\right]$$

$$= A\left[\frac{1 - \dfrac{1}{\left(1 + \dfrac{r}{100}\right)^n}}{1 - \dfrac{1}{\left(1 + \dfrac{r}{100}\right)}}\right] = A\left[\frac{1 - \left(1 + \dfrac{r}{100}\right)^{-n}}{1 - \left(1 + \dfrac{r}{100}\right)^{-1}}\right]$$

When the annuity becomes a perpetuity, i.e., $n \to \infty$ $\left(1 + \dfrac{r}{100}\right)^{-n} = 0$

$\therefore$ Present value of perpetuity $= \dfrac{A}{1 - \left(1 + \dfrac{r}{100}\right)^{-1}}$

If $\dfrac{r}{100}$ = i; the rate of interest per rupee per period, then

Present value of annuity due for n periods $= A\left[\dfrac{1 - (1+i)^{-n}}{1 - (1+i)^{-1}}\right]$

And present value of perpetuity $= \dfrac{A}{1 - (1+i)^{-1}}$

Find the amount of annuity of ₹4000 per annum for 10 years reckoning compound interest at 10% per annum.

Solution :

Considering immediate annuity, the required amount

$$= 4000\left(1 + \dfrac{10}{100}\right)^{9} + 4000\left(1 + \dfrac{10}{100}\right)^{8} + \ldots + 4000\left(1 + \dfrac{10}{100}\right)^{1} + 4000$$

$$= 4000 + 4000\left(1 + \dfrac{10}{100}\right) + 4000\left(1 + \dfrac{10}{100}\right)^{2} + \ldots + 4000\left(1 + \dfrac{10}{100}\right)^{9} \text{ [writing in the reverse order]}$$

$$= 4000\left[1 + 1.1 + (1.1)^{2} + \ldots + (1.1)^{9}\right] \left[\because 1 + \dfrac{10}{100} = 1.1\right]$$

$$= 4000\left[\dfrac{(1.1)^{10} - 1}{1.1 - 1}\right] = 4000\left[\dfrac{(2.594 - 1)}{0.1}\right] = \dfrac{4000}{0.1} \times 1.594$$

= ₹63,760.

PROBABILITY

The theory of probability involves the study of a random phenomena.

In the vocabulary of statistics, "Random" does not mean haphazard but rather refers to a kind of order that emerges only in the long run. The long-run regularity of random phenomena can be described mathematically just as the fall of an apple or the motion of an asteroid can be predicted. The mathematical study of randomness is called *probability theory*.

Probability theory originated in the middle of the 17th Century. It began with Blaise Pascal (1623-1662). His interest and that of his contemporaries, As interest in the natural sciences proliferated, so did demands for new laws of probability. Among prominent early contributions were the 19th Century mathematicians. Laplace De Moivre, Gauss, and Poisson. Today, probability theory finds application in diverse disciplines such as biology, economics, operations research, astronomy, etc. A biologist might be interested in the distribution of bacteria in a culture, an economist in economic forecasts, a production engineer in the inventory of a particular commodity, an astronomer in the distribution of stars in different galaxies and so on. Considerable interest is also shown in this field as a mathematical discipline in its own right. The name of the brilliant Russian mathematician, A. Kolmogorov (1903-87) should be mentioned as a pioneer in this area.

Probability refers to chance, i.e., chance of happening of some event. In every walk of life, we make statements that are probabilistic and that carry overtones of chance. For example, we might talk about the probability that a bus will arrive on time, or that a child to be born will be a son, or that the stock market will go up, and so on. What is the characteristic feature in all the phenomena above? It is that they are not deterministic. Past Information, no matter how voluminous, will not allow us to formulate a rule to determine precisely, what will happen in the future. The theory of probability involves the study of this type of phenomena called *random phenomena*.

Probability theory has a central place in the theory and applications of statistics, since we are analyzing and interpreting data that involve an element of chance or uncertainty. The reliability of our conclusions is supported by accompanying probability statements.

To understand the 'Theory of Probability', the knowledge of some basic concepts of set theory is very essential and is given below:

SOME BASIC CONCEPTS OF SET THEORY

Definition of a Set

A collection of distinct and well-defined objects is called a **set**. The objects forming the set are called **elements** or members of the set. 'Distinct' means no two elements in a set are the same. By 'well-defined', we mean that we are given a rule by which we can determine whether a particular element is a member of the set under consideration.

Sub-set

A set 'A' is called a sub-set of a set 'B' if each element of set 'A' belongs to the set 'B'. Set 'A' can be smaller than or equal to set 'B'. If set 'A' is smaller than set 'B', then set 'A' is called proper sub-set of set 'B' and is denoted symbolically as $A \subset B$. If A = B, then also A is called sub-set of B because every set is its own sub-set also. Hence, $A \subseteq B$ denotes set 'A' is a sub-set of 'B'. (The symbol includes the possibility A =B).

Examples:

1. Set of even integers is a proper sub-set of the set of positive integers.

2. Set (1, 2, 3) is a proper sub-set of the set (1, 2, 3, 4, 5).

3. Members of family form a set because each member is different and by seeing any person, we can say whether he/she belongs to a particular family or not.

Note: Elements of a set are enclosed inside the curly brackets { }.

If $A \subseteq B$ i.e., if A is sub-set of B, then B is called **Super set** of A and is denoted by $B \supseteq A$.

Equal Sets

Two sets A and B are said to be equal if all the elements of set A belongs to set B and all the elements of set B belongs to set A.

Null or Empty or Void Set

It is a set having no element and is usually denoted by f or { }.

Disjoint Sets

Two sets A and B are said to be disjoint, if there is no common element among them, i.e., if there is no element which belongs to both A and B together.

For example: If A = {1,2,3} and B = {4,5,6}, then the sets A and B are disjoint.
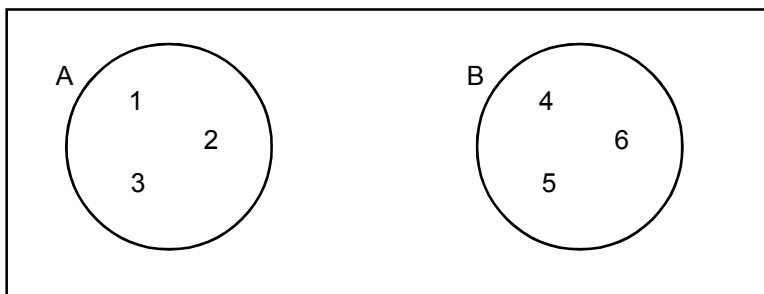
VENN DIAGRAM OF DISJOINT SETS

Union of Two Sets

If A and B are two given sets, then their union is the set of those elements which belong either to A or B or to both, i.e., those elements which belong to at least one of the two sets, A and B.

The Union of A and B is denoted symbolically as A U B.

$\qquad$ AUB = {X $\in$ A or X $\in$ B or X $\in$ both A and B}

$\qquad$ Let A = {1, 2, 3, 4, 5} ; B = {2, 3, 4, 6,7}

Then AUB = {1, 2, 3, 4, 5, 6, 7}

$\qquad$ = Elements of set A + Elements of Set B – Elements common to set A and B

In case of disjoint sets,

$\qquad$ AUB = Elements of set A + Elements of set B

Any elements belonging to (A∪B) will belong to **at least one** of the two sets.

Intersection of Sets

If A and B are two sets, then their intersection is the set of those elements which are common to both the sets A and B, i.e., which belong jointly to A and B, or which belong simultaneously to A and B. For example, if A = {1, 2, 3, 4,5,}; B = {2, 3, 4, 6,7}, then the intersection of A and B (denoted symbolically as A ∩ B is the set containing the elements 2,3,4, i.e., A ∩ B = {2, 3, 4}.

VENN DIAGRAM FOR UNION AND INTERSECTION
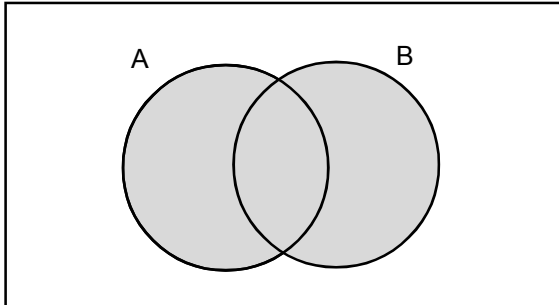
Shaded area represents A∪B

Union of two sets
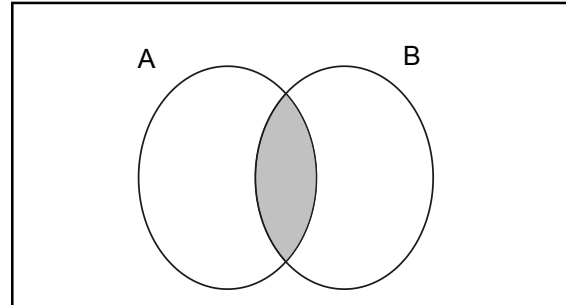
Shaded area represents A∩B

Intersection of two sets

Note : Set A is a sub-set of A∪B, B is a sub-set of AUB but (A∩B) is a sub-set of A as well as B.

Complement of a Set

Let U be any set (generally called Universal set) and A be its sub-set. Then the complement of set A in relation to U is that set whose elements belong to U but not to A. Complement of A is denoted by A' or A^c.

If U = {1, 2, 3, 4,5} and A = {1, 2, 5} then A^c = {3, 4}

i.e. A^c = {x : x ∈ U and x ∉ A} = U − A, (A ∪ A^c) ∪ (A ∩ A^c)= ϕ
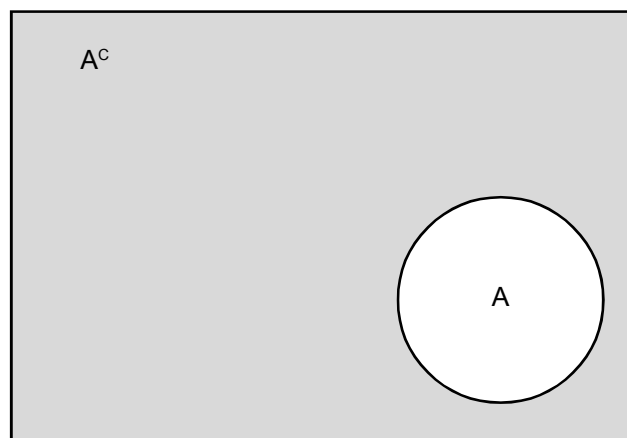
Shaded area represents the complement of set A in relation to U.

De Morgan's Law

  (i) (A ∪ B)^c = (A^c ∩ B^c)

  (ii) (A ∩ B)^c = (A^c ∪ B^c)

Note : Complement of a set is always in relation to a universal set.

Difference between Two sets

If A and B are two sets, then their difference (A –B) is the set of all those elements of A which do not belong to B.

Element of (A – B) Elements of A – Elements common to A and B

$$= \text{Elements of A} - \text{Elements of (A } \c \text{ B)}$$

Elements of (B – A) = Elements of B – Elements of (A $\c$ B)

Elements of (A – B) belong exclusively to A, i.e., exactly to A or only to A and elements of (B – A) belong exclusively to B.

If A = {3, 4, 5, 6, 7}, B = {6, 7, 8, 9},

then, (A – B) = {3, 4, 5,}, (B – A) = {8, 9)

Some Important Results

If A and B are any two sets, then –

   (i) A – B = A ∩ B'

   (ii) B – A = B ∩ A'

   (iii) (A – B) ∪ (B – A) = (A ∪ B) – (A ∩ B)

Factorial

Factorial n is the continued product of first n natural numbers. Factorial n is symbolically written as n!.

   n! = 1 x 2 x 3x 4x ……………… x (n – 1) x n

   = n x (n – 1) x (n – 2) x …………. X 4 x 3 x 2 x 1

   = n [(n – 1)]! = n(n – 1)[(n – 2)]!

Permutation and Combination

Permutation: Permutation is an ordered arrangement of a group of objects.

The permutation of n objects taken r at a time is represented as $_nP_r$

Where, $_nP_r = n!/(n-r)!$

For example, the permutation or arrangement of 9 different balls in 3 different rows can be done in $_9P_3$ ways.

= 9!/(9-3)! = 9x 8 x 7 = 504 ways

Combination

A combination is a selection of all or part of a set of objects, without regard to the order in which objects are selected.

For example, suppose we have a set of three letters: A, B, and C. We might ask how many ways we can select 2 letters from that set. Kach possible selection would be an example of a combination. The complete list of possible selections would be: AB, AC, and BC.

Mathematically, the number of Combinations ot *n* objects taken *r* at a time is given as $_nC_r$ $_nC_r$ = n! / r!(n - r)!

For example picking a team of **3** people from a group of 10 can be done in $_{10}C_3$ **ways.**

$_{10}C_3$ = 10!/(7! · 3!)= 10 x 9 x 8/(3 x 2 x 1) = 120.

Number of combinations of n things taken r at a time (r ≤ n). [Symbolically written as ^nC_r] is given by:

$$^nC_r = \frac{n!}{r!(n-r)!}$$

$$^nC_r = {}^nC_{n-r}$$

$0! = 1$, (By definition)

1!=1

Random experiment

If an experiment is conducted repeatedly under essential homogeneous conditions, the results are not unique but may be one of the various possible outcomes. An experiment for which we know the set of all different results but it is not possible to predict which one of the set will occur at any particular execution of the experiment is called a **random experiment.**

For example: tossing a fair coin, casting an unbiased die and drawing a card from a pack of 52 cards.

Let us take the experiment of **tossing a coin.** If we toss a coin then we get two possible outcomes either a 'head' **(H)** or a 'tail' **(T),** and it is impossible to predict whether the result of a toss will be a 'head' or 'tail'.

Let us consider a similar experiment **rolling a die** from a box. It we through a die then there are only six possible outcomes. The faces of the die are marked as 1,2, 3, 4, 5 and 6 and these are the only possible outcomes. But here also the outcome of a particular throw is completely unpredictable.

In mathematical theory we will only consider the experiments or observations, for which we know a priori the set of all different possible outcomes, such that it is impossible to predict which particular outcome will occur at any particular performance of the experiment. As such, if a random experiment is repeated under identical conditions the outcomes or results may vary or fluctuate at random.

Trial : Performing of an Experiment is called a Trial.

Outcomes : Whenever we do an experiment like flipping a coin or rolling a die, we get an **outcome.** For example, if we flip a coin we get an outcome of heads or tails, and if we roll a die we get an outcome of 1, 2, 3, 4, 5, or 6.

Event : An Event is a collection of basic outcomes that have some desirable property. Let's say that, as an example, E is the event where our roll of a six-sided die has a value less than or equal to 3 I his makes E the collection of basic outcomes where the result is value of 3 or less symbolically, here E - {O_1,O_2,O_3}. Note that, depending on the event, there could only be one - or possibly even zero - basic outcomes that qualify as being part of a particular event.

Experiment

↓ ← *Outcomes*

Events

| Simple Events | Compound events | Equally likely events | Mutually exclusive events | Exhaustive events | Indepen-dent events | Dependent events |
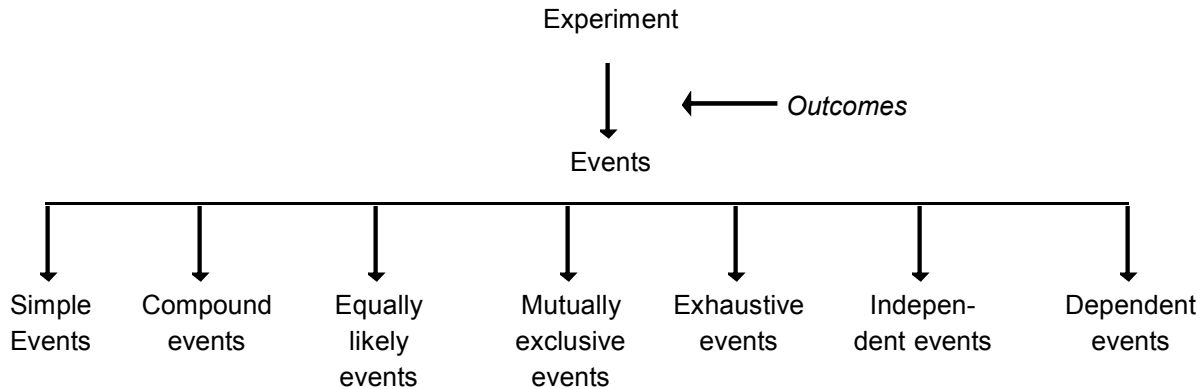
(i) Elementary (i.e. simple) **Events** are those events which cannot be decomposed further.

(ii) Composite Events (or Compound Events) are those events which can be decomposed further into elementary events. In other words, an elementary event corresponds to a single possible outcome of an experiment and a compound event is an aggregate of some elementary events and can be decomposed into simple events.

Examples :

(i) Tossing of a coin is an experiment and the two possible outcomes:

– Turning up of head and

– Turning up of tail,

are simple events. The event 'turning up of a head or tail' is compound event because it can be decomposed into two simple events : (a) 'turning up of head' and (b) 'turning up of tail'.

(ii) Throwing of a die is an experiment and the six possible outcomes are simple events [a die is a six faced cube and the faces are numbered from 1 to 6. When a die is thrown (or rolled), any of the six faces may come up].

The event that 3 comes up is a simple event but the event 'an odd number comes up' is a composite event because it can be decomposed further into three simple events: 1 comes up, 3 come sup and 5 comes up.

Similarly, an event '4 comes up' is a simple event 'an even number comes up' is a composite event because it can be decomposed further into three simple events : 2 comes up, 4 comes up and 6 comes up.

Equally Likely Events

Events are said to be equally likely, if after taking into account all the conditions, no event can be expected to occur in preference to any other event in the same experiment.

For example: If a fair coin is tossed, there is equal chance of turning up of head or tail because in a fair (**unbiased or perfect**) coin, we cannot expect head to turn up more frequently than tail or vice versa, when a fair dice is thrown, there are six equally likely events.

Mutually Exclusive Events

Events are said to be mutually exclusive, if the occurrence of any one of them prevents the occurrence of any other event in the same experiment, i.e., at the same time.

Examples:

(i) If a fair coin is tossed, turning up of 'head' and turning up of 'tail' are two mutually exclusive events because if head turns up, then tail cannot turn up in the same experiment and vice versa.

(ii) If a fair dice is thrown (or rolled or flipped), there are 6 mutually exclusive events because only one face can come up at a time in any experiment.

Mutually exclusive events cannot occur simultaneously in the same experiment. Mutually exclusive events may or may not be equally likely. For example: if a coin is biased, the two events head and tail are mutually exclusive but not equally likely.

Independent Events

Two or more events are said to be independent when the outcome of one does not affect, and is not affected by the outcome of other. For example, if a coin is tossed twice, the result of the second throw would in no way be affected by the result of the first throw.

Dependent Events

Dependent events are those events, in which the occurrence or non-occurrence of one event in any one trial affects the probability of other events in other trials. For example, the probability of drawing a queen from a pack of 52 cards is 4/52 or 1/13. But if the card drawn is not replaced in the pack, the probability of drawing again a queen is 3/51.

Sample Space

The set of all possible outcomes of an experiment is called the sample space of the experiment and is denoted by S. Each outcome is called an element or a sample point of the sample space. Sample space is also called universal set or event space or possibility set. For example, if a die is thrown, there are six possibilities and we say that the sample space associated with the experiment of throwing a die consists of six sample points and is denoted as S = {1, 2, 3, 4, 5, 6}.

The sample space associated with the experiment of throwing a fair coin consists of two sample points, i.e., head and tail and is denoted as S = {H, T} where H denotes head and T denotes tail.

An event is a sub-set of the sample space. An event may consists of one or more sample points. For example when a dice is thrown, the event that an even number appears consists of three sample points but the event that 1 appears consists of only one sample point.

If you roll a six-sided die, there are six possible outcomes, and each of these outcomes is equally likely. A six is as likely to come up as a three, and likewise for the other four sides of the die.

The probability of a given event is an expression of chance of occurrence of an event. A probability is a number which ranges from 0 to 1. 0 for an event which cannot occur and 1 for an event which certain to occur.

Exhaustive Events

The total number of all possible outcomes of a random experiment constitute an exhaustive set of events, i.e., the events are said to be exhaustive, if all possible cases {i.e., outcomes} are considered.

APPROACHES OF PROBABILITY

Mathematical (or Classical) Definition of Probability

If an experiment can result in (m + n), equally likely mutually exclusive and exhaustive cases and m of them are favourable to the occurrence of an event 'A' and the remaining, i.e., n are against the occurrence of the event 'A' then the probability of the occurrence of event 'A' denoted by P(A), is defined as the ratio:

$$\frac{m}{m+n} = \frac{\text{Number of cases favourable to the occurrence of the event 'A'}}{\text{Total number of equally likely, mutually exclusive and exhaustive cases}}$$

Probability of occurrence of event 'A' is denoted as = P(A)

The probability of the occurrence of any event always lies between 0 and 1 and these are the two limits of probability

Instead of saying that the probability of occurrence of event A is $\dfrac{m}{m+n}$;

it is sometimes stated that;
Odds are m to n in favour of event A or n to m against the event A, i.e.,
Odds in favour of the event A =

$$\frac{\text{No. of cases favourable to the occurrence of the event A}}{\text{No. of cases against the occurrence of the event A}} = \frac{\dfrac{m}{m+n}}{\dfrac{n}{m+n}} = \frac{m}{n}$$

Odds against event A = $\dfrac{\text{Number of cases against the occurrence of the event A}}{\text{Number of cases in favour of the occurrence of the event A}}$

$$= \frac{n}{m} = \frac{P(A)}{P(A)}$$

Note : The ratio $\dfrac{m}{n}$ or $\dfrac{n}{m}$ is always expressed in its lowest form; i.e., as the ratio of two integers having no common factor.

If m = 0 i.e., if the number of cases favourable to the occurrence of event A = 0 then, P(A) = 0 i.e., event A is an impossible event. If n = m, $P(A) = \dfrac{m}{m} = 1$ i.e., event A is sure to happen and is called a **Certain (or sure) event**.

If neither m=0, nor n=0, then the probability of the occurrence of any event A is always less than 1. Therefore, the probability P of the occurrence of any event satisfies the relation $0 \le P \le 1$.

Important Notes:

– If the events are mutually exclusive and exhaustive, the sum of their individual probabilities of occurrence is equal to one. For example: If A, B and C are three mutually exhaustive events then; and P(A) + P(B) + P(C) = 1.

– The probability of the occurrence of one particular event is also called **Marginal Probability** of that event.

To choose an object at random from N objects means that each object has the same probability $\dfrac{1}{N}$ of being chosen.

Relative Frequency theory of Probability or Empirical Probability

In 1800s, British statisticians started collecting data of births and deaths for calculating risks in life insurance and commercial insurance. Today this concept is known as relative frequency of probability. The probability of an

event can be defined as the relative frequency with which it occurs in an indefinitely large number of trials. If an event occurs 'a' times out of n, its relative frequency is a/n; the value which is approached by a/n when n becomes infinity is called the limit of the relative frequency.

Symbolically P(A) = Lim $\dfrac{a}{n}$

In practice the estimate of P(A) can be written as if it were actually P(A) and the relative frequency definition of probability may be expressed as:

P(A) = $\dfrac{a}{n}$

Classical probability is normally encountered in problems dealing with games of chance, for example, dice and card. On the other hand, probability derived from past experience is called empirical probability and is used in many practical problems, these two concepts of probability are widely used.

ADDITION RULE OF PROBABILITY (THEOREM OF TOTAL PROBABILITY)

If A and B are two events then the probability that at least one of them occurs is denoted by $(P \cup B)$ and is given by $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

When the event are mutually exclusive

$P(A \cap B) = 0$

And in that case $PA(A \cup B) = P(A) + P(B)$

```
                    ┌─────────────────────────┐
                    │ P (at least one of A and B) │
                    │     Additional Rule     │
                    └─────────────────────────┘
                                │
                                ▼
                          ╱───────────╲
                         ╱  Are A and   ╲   Yes    ┌─────────────────────────┐
                        ⟨   B mutually    ⟩ ──────▶ │ P(at least one of A and B) │
                         ╲  exclusive?   ╱          │    =P(A) + P(B)         │
                          ╲───────────╱             └─────────────────────────┘
                                │ No
                                ▼
                    ┌──────────────────────────────┐
                    │ P (at least one of A and B)   │
                    │ =P(A) + P(B) − P(A and B)     │
                    └──────────────────────────────┘
```
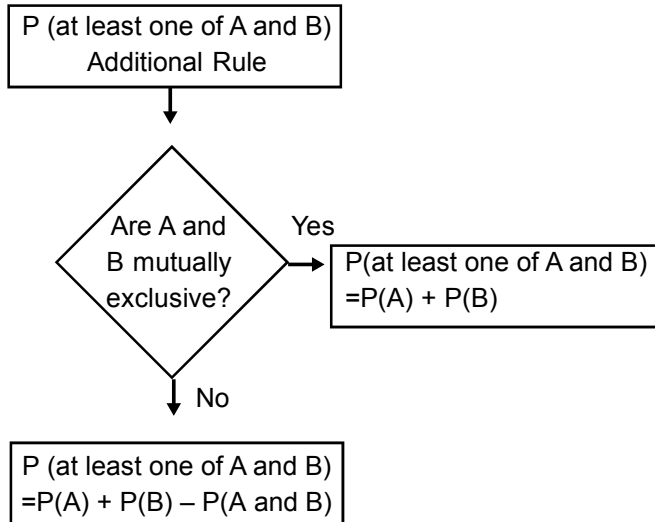
Multiplication Rule

If A and B are two events, the probability of their joint occurrence or simultaneous occurrence is given by:

$P(A \cap B) = P(A) \cdot P(A/B)$

When the events are independent;

P(A/B) = P(A)

P(B/A) = P(B)

```
┌─────────────────────────┐
│    Multiplication Rule   │
│       P (A and B)        │
└─────────────────────────┘
              │
              ▼
┌────────────────────────┐  Yes   ┌──────────────────────────┐
│ Are A and B independent │───────▶│  P(A and B) = P(A) x P(B) │
└────────────────────────┘        └──────────────────────────┘
              │
              ▼    NO
┌────────────────────────────┐
│ P(A and B)  =  P(A) P(A/B)  │
│             =  (P(B) P(A/B) │
└────────────────────────────┘
```

∴ When the events are independent, P(A ∪ B) = P(A) + P(B) − P(A ∩ B).

P(A/B) is the conditional probability of the occurrence of event A, when it is known that event B has already occurred.

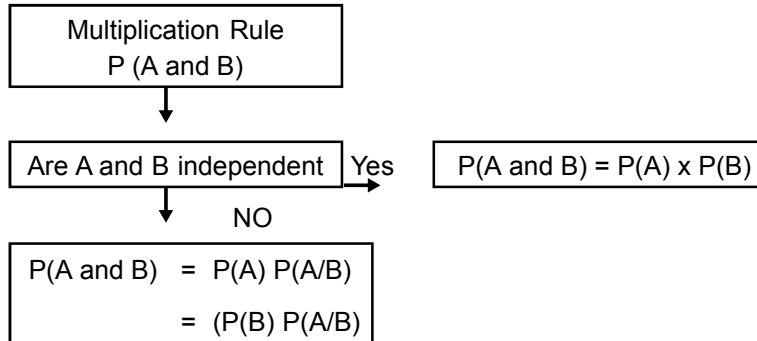Similarly, P(B/A) is the conditional probability of the occurrence of event B, when it is known that event A has already occurred.

If the events are independent, the occurrence of A do not affect the occurrence of B.

∴ P(B/A) = P(B)

Also, P(A/B) = P(A)

For example, in a family the gender of the second child is not affected by the gender of the first child or if a coin is thrown two times, the outcome of the second throw is not affected by the outcome of the first throw.

SOLVED ILLUSTRATIONS

Illustration 9.5

If two coins are thrown, there are four possibilities, viz.,

{HH, HT, TH TT}

Let E_1 be the event of getting one head and one tail,, then

$$P(E_1) = \frac{2}{4} = \frac{1}{2}$$

If E_2 is the event of getting two tails, then

$$P(E_2) = \frac{1}{4}$$

If E_3 is the event of getting two heads, then

$$P(E_3) = \frac{1}{4}$$

If E_4 is the event of getting head or tail, then

$P(E_4) = \dfrac{4}{4} = 1$ **[Certain or Sure Event]**

If E_5 is the event of getting three head (or three tails or two heads and one tail or one head and two tails), then

$P(E_5) = \dfrac{0}{4} = 0$ **[Impossible Event]**

Illustration 9.6 If two coins are thrown, what is the probability that head occurs on both the coins?

Solution : If two coins are thrown, the following cases are possible:

Coin I	Coin II
Head	Tail
Tail	Head
Tail	Tail
Head	Head

In total, there are four possibilities out of which there is only one possibility where both the coins throw head.

Therefore, required probability = $\dfrac{1}{4}$

Note:

(i) Similarly, the probability that both the coins throw tail = $\dfrac{1}{4}$

(ii) The probability that one coin throw head and the other coin throw tail $= \dfrac{2}{4} = \dfrac{1}{4}$

Illustrations 9.7: The percentages of 100 teenage offenders are in the following six categories :

Categories	Group A	Group B	Total
Non-offenders	28	42	70
First time offenders	05	15	20
Repeat offenders	07	03	10
Total	40	60	100

Suppose a teenager is chosen at random and the following events are defined.

(i) Teenager is from Group A

(ii) Teenager is from Group B

(iii) Teenager is not an offender

Find the following probabilities :

(i) P(A)   (ii) P(B)   (iii) P(C)

Solution :

P(A)=Probability that the Teenagers is from Group A = $\dfrac{\text{Number of teenagers in Group A}}{\text{Total Number of teenagers}}$ $\dfrac{28+5+7}{100}$ $\dfrac{40}{100}$ = 0.4

P(B)=Probability that the Teenagers is from Group B = $\dfrac{\text{Number of teenagers in Group B}}{\text{Total Number of teenagers}}$ $\dfrac{45+15+3}{100}$ $\dfrac{60}{100}$ = 0.6

P(C)=Probability that the Teenagers is not an offender = $\dfrac{\text{Number of teenagers who are not offenders}}{\text{Total Number of teenagers}}$ $\dfrac{28 + 42}{100}$ $\dfrac{70}{100}$ = 0.7

Illustrations 9.8: An analyst develops the following table of joint probabilities relating the size of firm (measured in terms of number of employees) and type of firm.

Number of Employees	Industry		
	Construction	Manufacturing	Retail
Under 20	0.2307	0.0993	0.5009
20 – 99	0.0189	0.0347	0.0876
100 or more	0.0019	0.0147	0.0113

If one is selected at random, find the probability of the following events :

  (i) P(A) = The firm employs fewer than 20 employees

 (ii) P(B) = The firm is in the retail industry

(iii) P(C) = A firm in the construction industry employs between 20 to 99 workers

(iv) P(D) = A firm in the retail industry employed more than 20 workers.

Solution:

  (i) Required Probability, P(A) = 0.2307 + 0.0993 + 0.5009 = 0.8303

 (ii) Required Probability, P(B) = 0.5009 + 0.0876 + 0.0113 = 0.5998

(iii) Required Probability P(C) = $\dfrac{0.0189}{0.2307 + 0.0189 + 0.0019}$ = $\dfrac{0.0189}{0.2515}$ = 0.075

(ivi) Required Probability P(D) = $\dfrac{0.876 + 0.0113}{0.5009 + 0.876 + 0.0113}$ = 0.296

Illustration 9.9: Find the probability that: (i) a leap year has 53 Sundays and (ii) a non-leap year has 53 Sundays.

Solution :

A leap year contains 366 days which is equal to 366/7 = 52 weeks + 2 days.

These two days will always be consecutive days and the following possibilities will exist:

    Sunday   -   Monday

    Monday   -   Tuesday

Tuesday	-	Wednesday
Wednesday	-	Thursday
Thursday	-	Friday
Friday	-	Saturday
Saturday	-	Sunday

In the above seven possibilities, each day is repeated twice. Therefore, the probability of Sunday to be repeated 53 times in a leap year = 2/7

A non-leap year contain 365 days i.e., 365/7 = 52 weeks and 1 day. This one day can be any of the seven days of week.

∴ The probability that a non-leap year contains 53 Sundays = 1/7

Illustrations 9.10: Three unbiased coins are tossed. What is the probability of obtaining:

    (a) all heads

    (b) two heads

    (c) one head

    (d) at least one head

    (e) at least two heads

    (f) all tails.

Solution :

When an unbiased coin is thrown, the probability of getting a head, P(H) = Probability of getting a tail, P(T) = ½

When three coins are thrown, the following sample space is obtained:

S = (HHH); (HTT); (HTH); (HHT); (THH); (TTH); (THT); (TTT)

    (a) Probability of all heads = $\dfrac{1}{2} \times \dfrac{1}{2} \times \dfrac{1}{2} = \dfrac{1}{8}$

       (Because out of 8 possibilities, there is only one possibility which contains three heads)

    (b) Probability of two heads i.e., Probability of (H T H), (H H T) or (T H H)

$$= \dfrac{1}{8} + \dfrac{1}{8} + \dfrac{1}{8} = \dfrac{3}{8}$$

    (c) Probability of one head, i.e., Probability of (H T T), (T T H) or (T H T)

$$= \dfrac{1}{8} + \dfrac{1}{8} + \dfrac{1}{8} = \dfrac{3}{8}$$

    (d) Probability of at least one head = 1 – Probability of getting all tails

$$1 - \dfrac{1}{8} = \dfrac{7}{8}$$

(e) Probability of at least two heads, i.e. Probability (H H T), (T H H), (H T H), or (H H H)

$$= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$$

(f) Probability of all tails, i.e., Probability (T T T) $= \frac{1}{8}$

Illustrations 9.11: A card is drawn from a well-shuffled deck of 52 cards. Find the Probability of drawing a card which is neither a heart nor a king.

Solution:

In 52 cards, there are 13 cards of heart including one king. In the remaining 39 cards, there are three kings. Therefore, there are 36 cards which do not contain either card of heart or a card of king.

Required Probability $= \frac{36}{52} = \frac{9}{13}$

Illustrations 9.12: Five men in a company of 20 are graduates. If 3 men are picked out of the 20 at random, what is the probability that they are all graduates? What is the probability of at least one graduates?

Solution:

Total sample size = 20

Number of Graduates = 5

Number of Non-graduates = 20 – 5 = 15

Three graduates can be taken out of 5 Graduates in $^5C_3$ ways and 3 persons out of 20 can be taken in $^{20}C_3$ ways.

(i) Required probability $= \dfrac{^5C_3}{^{20}C_3} = \dfrac{\dfrac{5!}{3!(5-2)!}}{\dfrac{20!}{3!(20-3)!}} = \dfrac{\dfrac{5!}{3! \times 2!}}{\dfrac{20!}{3! \times 17!}} = \dfrac{5!}{3! \times 2!} \times \dfrac{3! \times 17!}{20!}$

$$= \frac{5 \times 4}{2 \times 1} \times \frac{3 \times 2 \times 1}{20 \times 19 \times 18}$$

$$= \frac{1}{114}$$

(ii) Probability of at least one graduate=

$$= \frac{\left(^5C_1 \times {}^{15}C_2\right) + \left(^5C_2 \times {}^{15}C_1\right) + \left(^5C_3 \times {}^{15}C_0\right)}{^{20}C_3} = \frac{(5 \times 105) + (10 \times 15) + 10}{1140}$$

$$= \frac{525 + 150 + 10}{1140} = \frac{685}{1140} = \frac{137}{228}$$

Illustrations 9.13: In a single throw of two dice, find the probability of getting: (a) P(A) = a total of 11, (b) P(B) = a total of 8 or 11, and (c) P(C) = same number on both the dice.

Solution:

When two dice are thrown together, total number of possibilities are 6 x 6 = 36.

(a) Total of 11 will be obtained in the following ways:

First Dice	Second Dice
5	6
6	5

i.e., 2 cases

∴ Required Probability, P(A) = $\dfrac{2}{36} = \dfrac{1}{18}$

(b) Total of eight will be obtained in the following cases :

First Dice	Second Dice
2	6
6	2
5	3
3	5
4	4

i.e., 5 cases

∴ Required Probability, P(B) = $\dfrac{5}{36}$

(c) There are six cases in which same number will appear on both the dice i.e., (1,1), (2,2), (3,3), (4,4), (5,5), (6,6)

∴ Required Probability, P(C) = $\dfrac{6}{36} = \dfrac{1}{6}$

Illustration 9.14:

(a) A pair of dice is rolled, find the probability:

  (i) P(A)= that the sum is neither 8 nor 10 and

  (ii) P(B)= that the sum is greater than 12.

Solution:

When a pair of die is rolled, total number of possibilities is 6x6=36

(a) Let X denote the event of getting the sum 8 and Y denote the event of getting the sum 10. There are 5 possibilities for getting the sum 8, i.e., (2.6), (3,5), (4,4), (5,3), (6,2).

There are 3 possibilities for achieving the sum 10 i.e., (4,6), (5,5), (6,4)

Probability of getting the sum 8 = P(X) = 5/36

Probability of getting the sum 10 = P(Y) = 3/36 = 1/12

Probability that the sum is either 8 or 10 = $\frac{5}{36} = \frac{3}{36} = \frac{8}{36} = \frac{2}{9}$ {events are mutually exclusive}

Probability that the sum is neither 8 nor 10, P(A)=1 – Probability that sum is either 8 or 10

= 1 – P(A)

$= 1 - \frac{2}{9} = \frac{7}{9}$

(b) When two dice are thrown, a total sum greater than 12 is impossible and hence the probability of such an event is Zero.

RANDOM VARIABLES

By *random* we mean *unpredictable*; when the same is applied in case of a random variable, means, we cannot with certainty predict its future value. Even through the entire past history of the variable is known, it is still unpredictable with certainty. If the variable is of the deterministic type, no such uncertainty exists.

However, quite a few random variables do exhibit statistical regularity. Consider a simple experiment of tossing an unbiased coin. We do not know in advance whether the outcome of a particular toss would be a head or tail. But, we know for sure that in a long sequence of tosses, about half of the outcomes would be heads. If this does not happen, we suspect either the coin or the person tossing it is biased. Statistical regularity of averages is an experimentally verifiable phenomenon in many cases involving random quantities. Hence, we are tempted to develop mathematical tools for the analysis and quantitative characterization of random variables.

Random Variable Definition

Although it may look simple at first sight to give a definition of what a random variable is, it proves to be quite difficult in practice. A random variable, usually written *X*, is a variable whose possible values are numerical outcomes of a random experiment. It therefore is a function that associates a unique numerical value with every outcome of an experiment. The value of the random variable will vary from trial to trial as the experiment is repeated.

The following paragraphs will cover the details.

Random Experiment: As Random variables are outcomes of a random experiment, it is essential to understand a random experiment as well. Where Random variables are outcomes, a random experiment is a process leading to an uncertain out-come, before the experiment is run. It is usually assumed that the experiment can be repeated indefinitely under essentially homogeneous conditions. Result of a random experiment is not unique but it can be one of the various possible outcomes. Simple example is tossing an unbiased coin, where outcomes can be head or tail. You keep on tossing the coin a number of times under essential homogeneous conditions, the outcomes would keep on flipping between Head & Tail, without exactly knowing which toss would result in to what.

The outcome of an experiment need not be a number, for example, the outcome in a coin toss experiment can be 'heads' or 'tails'. However, we often want to represent outcomes as numbers. A random variable is a function that associates a unique numerical value with every outcome of an experiment. In the given example, if there are three trails (say); the number of times "Head" appears can be a random variable, which can assume values as, 0, 1, 2, & 3. Because in three trials, you can have minimum zero Heads and maximum three Heads.

Types of Random Variables

Classification of random variables is done based on their probability distribution. A random variable has either

an associated probability distribution (discrete random variable) or probability density function (continuous random variable). Based on that, there are two types of random variable – Discrete and Continuous.

Discrete Random Variables

A discrete random variable is one which may take on only a countable number of distinct values such as 0,1,2,3,4,........ Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, the number of people in an ATM queue, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten etc.

The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or the probability mass function.

Suppose a random variable X may take k different values, with the probability that $X = x_i$ defined to be $P(X = x_i) = p_i$. The probabilities p_i must satisfy the following:

   1: $0 < p_i < 1$ for each i

   2: $p_1 + p_2 + ... + p_k = 1$.

Example

   1. A coin is tossed ten times. The random variable X is the number of "Tails" that are noted. X can only take the values 0, 1, ..., 10, so X is a discrete random variable. The above two properties hold in this case. For e.g. probability of 8 Tails, p_8 will definitely fall in the range 0 to 1. And also, the sum of probabilities for all possible values of Tails, $p_0 + p_1 + p_2 + ... + p_{10} = 1$.

      Note in case of ten trials, number of tails can be 0 to 10.

Continuous Random Variables

A continuous random variable is one which takes an infinite number of possible values (usually in a given range). Continuous random variables are usually measurements like, height, weight, the amount of sugar in an orange, time required to finish a task, interest earn etc. For e.g.: life of an individual in a community. A person may die immediately on his birth (life equals to zero years) or after attaining an age of 110 years (say). Within this range, he may die at any age. Therefore the variable "age" can take any value in the range 0 to 110, in this case.

A continuous random variable is not defined at specific values, since the values are infinite and therefore probability at a specific value is almost zero. Instead, it is defined over an interval of values, and is represented by the area under a curve.

Suppose a random variable *X* may take all values over an interval of real numbers. Then the probability that *X* is in the set of outcomes *A, P(A)*, is defined to be the area above *A* and under a curve. The curve, which represents a function *p(x)*, must satisfy the following:

   1: The curve has no negative values ($p(x) \geq 0$ for all x)

   2: The total area under the curve is equal to 1.

A curve meeting these requirements is known as a density curve.

A light bulb is burned until it burns out. Suppose the life of bulb ranges between zero hours (minimum) to 100 hours (maximum). The random variable Y is its lifetime in hours. Y can take any positive real value in the range 0 to 100, so Y is a continuous random variable. It is immaterial to calculate probability of Y at a specific point in the specified range; instead we wish to calculate probability between any two end points in the range, like 0-10,

50-70, less than 20, more than 90 etc. At any point in the complete range (0-100), $p(x) \geq 0$, and the total area in the probability curve from $p(x=0)$ to $p(x=100)$ would be equal to one.

EXPECTED VALUE

The expected value (or population mean) of a random variable indicates its average or central value. It is a useful summary value (a number) of the variable's distribution.

Stating the expected value gives a general impression of the behaviour of some random variable without giving full details of its probability distribution (if it is discrete) or its probability density function (if it is continuous).

Two random variables with the same expected value can have very different distributions. There are other useful descriptive measures which affect the shape of the distribution, for example standard deviation.

The expected value of a random variable X is symbolised by E(X).

If X is a discrete random variable with possible values x1, x2, x3, ..., xn, and p(xi) denotes P(X = xi), then the expected value of X is defined by:

$$\mu = E(X) = \sum P(x_i) - x_i$$

where the elements are summed over all values of the random variable X.

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

$$\mu = E(X) = \int x \cdot f(X) dx$$

Example

Discrete case: When a die is thrown, each of the possible faces 1, 2, 3, 4, 5, 6 (the xi's) has a probability of 1/6 (the p(xi)'s) of showing. The expected value of the face showing is therefore:

$$\mu = E(X) = (1 \times 1/6) + (2 \times 1/6) + (3 \times 1/6) + (4 \times 1/6) + (5 \times 1/6) + (6 \times 1/6) = 3.5$$

Notice that, in this case, E(X) is 3.5, which is not a possible value of X.

Expected Values of Random Variables

We already looked at finding the mean in the section on averages. Random variables also have means but their means are not calculated by simply adding up the different variables.

The mean of a random variable is more commonly referred to as its Expected Value, i.e. the value you expect to obtain should you carry out some experiment whose outcomes are represented by the random variable.

In Probability theory, the **expected value** (or **expectation**, **mathematical expectation**, **EV**, **mean**,) refers, intuitively, to the value of a random variable one would "expect" to find if one could repeat the random variable process an infinite number of times and take the average of the values obtained. More formally, the expected value is a weighted average of all possible values. In other words, each possible value that the random variable can assume is multiplied by its assigned weight, and the resulting products are then added together to find the expected value. The weights used in computing this average are the probabilities in the case of a discrete random variable, or the values of a probability density function in the case of a continuous random variable.

Example:

A local club plans to invest $10000 to host a baseball game. They expect to sell tickets worth $15000. But if it rains on the day of game, they won't sell any tickets and the club will lose all the money invested. If the weather forecast for the day of game is 20% possibility of rain, is this a good investment?

Make a table of probability distribution.

Outcome	+$3000	– $12000
Probability	0.80	0.20

Use the weighted average formula.

Expected value = 5000(0.8) – 12000(0.2)

= 4000–2400

= 1600

The club can expect a return of $1600. So, it's a good investment, though a bit risky.

In other cases, we are asked to find the values of one or more variables involved in the model for which the experiment has a given expected value.

Illustration 9.15

A company makes electronic gadgets. One out of every 50 gadgets is faulty, but the company doesn't know which ones are faulty until a buyer complains. Suppose the company makes a $3 profit on the sale of any working gadget, but suffers a loss of $80 for every faulty gadget because they have to repair the unit. Check whether the company can expect a profit in the long term.

Solution

Write the probability distribution.

$$E(X) = \frac{49}{50}3 + \frac{1}{50}(-80)$$

$$= \frac{147}{50} - \frac{80}{50}$$

$$= \frac{67}{50}$$

$$= 1.34$$

Since the expected value is positive, the company can expect to make a profit. On average, they make a profit of $1.34 per gadget produced.

The intuitive explanation of the expected value above is a consequence of the law of average numbers: the expected value, when it exists, is almost surely the limit of the sample mean as the sample size grows to infinity. More informally, it can be interpreted as the long-run average of the results of many independent repetitions of an experiment (e.g. a dice roll).

Suppose random variable X can take value x_1 with probability p_1, value x_2 with probability p_2, and so on, up to value x_k with probability p_k. Then the **expectation** of this random variable X is defined as

$$E(X) = x_1 p_1 + x_2 p_2 + ... + x_k p_k$$

Since all probabilities p_i add up to one ($p_1 + p_2 + ... + p_k = 1$), the expected value can be viewed as the weighted average, with p_i's being the weights:

$$E[X] = \frac{x_1p_1 + x_2p_2 + \dots + x_kp_k}{1} = \frac{x_1p_1 + x_2p_2 + \dots + x_kp_k}{p_1+p_2+\dots+p_k}$$

If all outcomes x_i are equally likely (that is, $p_1 = p_2 = \dots = p_k$), then the weighted average turns into the simple average. This is intuitive: the expected value of a random variable is the average of all values it can take; thus the expected value is what one expects to happen *on average*. If the outcomes x_i are not equally probable, then the simple average must be replaced with the weighted average, which takes into account the fact that some outcomes are more likely than the others. The intuition however remains the same: the expected value of X is what one expects to happen *on average*.

Example 1. Let X represent the outcome of a roll of a six-sided die. More specifically, X will be the number of pips showing on the top face of the die after the toss. The possible values for X are 1, 2, 3, 4, 5, and 6, all equally likely (each having the probability of 1/6). The expectation of X is

$$E[X] = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6} = 3.5$$

Let X be a discrete random variable taking values $x1$, $x2$, ... with probabilities $p1$, $p2$, ... respectively. Then the expected value of this random variable is the infinite sum

$$E[X] = \sum_{i=1}^{\infty} x_i p_i$$

Given that the random variable X is continuous and has a probability distribution f(x), the expected value of the random variable is given by:

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

Uses and applications

The expected value plays important roles in a variety of contexts. In regression analysis, one desires a formula in terms of observed data that will give a "good" estimate of the parameter giving the effect of some explanatory variables upon a dependent variable. The formula will give different estimates using different samples of data, so the estimate it gives is itself a random variable. A formula is typically considered good in this context if it is an unbiased estimator – that is, if the expected value of the estimate (the average value it would give over an arbitrarily large number of separate samples) can be shown to equal the true value of the desired parameter.

In decision theory, and in particular in choice under uncertainty, an agent is described as making an optimal choice in the context of incomplete information. For risk neutral agents, the choice involves using the expected values of uncertain quantities, while for risk averse agents it involves maximizing the expected value of some objective function.

LESSION ROUND UP

– *Borrower*: Person who takes the money for his use.

– *Money Lender*: Person who gives the money in the hope of getting some consideration i.e. The money lender charges from the borrower an amount which is more than what he gives to him. The difference is called interest. The rate of interest is decided mutually by the two parties i.e., the lender and the borrower.

– *Simple Interest*: If Rs. P is the principal i.e., amount borrowed and r of the rate of interest per year and n, the number of years for which the money remains with the borrower, then

Simple Interest   = Prn

   = P × R × n

– *Compound Interest*: In case of compound interest, interest also earns interest. At the end of each conversion period, interest is added to the principle and the amount thus formed becomes the principle for the next conversion period.

Amount due after n years =

Compound Interest =

– *Annuity*: An annuity is a fixed sum paid at regular intervals under certain stated conditions. The interval may be either one year or half-year or quarter year or one month, etc., If nothing is mentioned about the period of interval, it is taken to be one year.

– *Types of Annuity*: There are eight type of annuity- Annuity Certain, Contingent Annuity, Perpetual Annuity (or Perpetually), Immediate Annuity (or ordinary annuity), Annuity Due, Deferred Annuity, Forborne Annuity, Periodic Payment/Uniform Annuity/Variable Unity.

– Probability refers to chance, i.e., chance of happening of some event. Probability theory has a central place in the theory and applications of statistics, since we are analyzing and interpreting data that involve an element of chance or uncertainty. The reliability of our conclusions is supported by accompanying probability statements.

– *Sample Space*: The set of all possible outcomes of an experiment is called the sample space of the experiment and is denoted by S. Each outcome is called an element or a sample point of the sample space. Sample space is also universal set or event space or possibility set.

– The probability distribution is a table or an equation that links each outcome of a statistical experiment with its probability of occurrence.

  – Results on Probability:

  – P(S) = 1, S = complete sample space

  – $0 \leq P(E) \leq 1$

  – P(j) = 0

  – For any events A and B we have : P(A*"B) = P(A)+P(B)-P(A)"B)

  – If A denotes (not-A), then P(A)=1-P(A)

GLOSSARY

Simple Interest   Interest computed only on the principal and not on principal plus interest earned or incurred in the previous period(s).

Compound Interest	Interest computed on the principal amount to which interest earned to-date has been added. Where compound interest is applied, the investment grows exponentially and not linearly as in the case of simple interest. In compound interest, interest also earned interest.
Annuity	It is series of fixed payments at fixed intervals, guaranteed for a fixed number of years or the lifetime under stated conditions.
Experiment	An operation which can produce some we-defined outcomes is called an experiment.
Random Experiment	An experiment in which all possible outcomes are known and the exact output cannot be predicted in advance, is called a random experiment.
Sample Space	When we perform an experiment, then the set S of all possible outcomes is called the sample space.
Event	Any subset of a sample space is called an event (simple or compound).
Discrete Random	A discrete random variable is one which may take on only a countable number of distinct values such as 0, 1, 2, 3, 4, ... Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete.
Continuous Random Variables	A continuous random variable is one which takes an infinite takes an infinite number of possible values. Continuous random variables are usually measurements.

MULTIPLE CHOICE QUESTIONS

1. **Two or more events are said to be____when the outcome of one does not affect, and is not affected by other.**

   (a) Dependent

   (b) Independent

   (c) Simple

   (d) Compond

2. **What time will be taken by an amount of Rs. 1000 to yield Rs. 150 as interest at 5% per annum of simple interest?**

   (a) 2 Years

   (b) 1 Year

   (c) 3 Years

   (d) None of the above

3. **Rahul took a loan for 3 years at the rate of 5% per annum on Simple Interest, If the total interest paid was Rs. 615, the principal amount was**

   (a) 4200

   (b) 4300

   (c) 4400

   (d) 4100

4. **Find compound interest on Rs. 4000 at 4% per annum for 2 years, compounded annually**

(a) 328.40

(b) 326.40

(c) 324.40

(d) 328.40

5. **A sum of Rs. 1100 becomes Rs. 1300 in two years at compound interest compounded annually. Find the rate percent.**

(a) 8.7

(b) 8.5

(c) 8.6

(d) 8.8

6. **Time value of money indicates that**

(a) A unit of money obtained today is worth less than a unit of money obtained in future

(b) There is no difference in the value of money obtained today and tomorrow

(c) A unit of money obtained today is worth more than a unit of money obtained in future

(d) None of the above

7. **Three unbiased coins are tossed, what is the probability of getting at least 2 heads?**

(a) 1/2

(b) 1/3

(c) 1/6

(d) 1/8

8. **A bag contains 10 green and 20 blue balls, One ball is drawn at random. What is the probability that ball is blue**

(a) 1

(b) 1/3

(c) 2/3

(d) 4/3

9. **What is the probability of getting a sum 7 from two throws of dice.**

(a) 4/6

(b) 1/3

(c) 2/3

(d) 1/6

10. **If the occurrence of one event means that another cannot happen, then the events are**

    (a) Independent

    (b) Mutually exclusive event

    (c) Empirical

    (d) Equally likely event

Answer Kys : 1.(b), 2. (c), 3. (d), 4. (b), 5.(a), 6. (c), 7.(a), 8. (c), 9.(d), 10.(a)

FOUNDATION PROGRAMME

BUSINESS ECONOMICS

PRACTICE TEST PAPER

(This Test Paper is for practice and self study only and not to be sent to the Institiute)

1. Macroeconomics deals with _____.

    (a) economic aggregates

    (b) behaviour of firms

    (c) activities of individual units

    (d) behaviour of a particular market

2. The meaning of Greek word '*oikonomikus*' is

    (a) Business Management

    (b) Business Economics

    (c) House Management

    (d) Wealth Management

3. Who is known as father of Economics?

    (a) Adam Smith

    (b) Joan Robinson

    (c) Alfred Marshall

    (d) John Maynard Keynes

4. In economics the central problem is:

    (a) Money

    (b) Consumption

    (c) Allocation

    (d) Scarcity

5. A recession is:

    (a) a period during which aggregate output declines.

    (b) a period of increasing unemployment.

    (c) a period of increasing excess capacity.

    (d) All of the above

Students are advised to go through instruction regarding computer based examination available on ICSI website www.icsi.edu

6. The quantity of a product that buyers wish to purchase at each possible price, is known as

   (a) Demand

   (b) Supply

   (c) Excess demand

   (d) Excess Supply

7. _____ and _____ do not directly affect the demand curve.

   (a) the price of related goods, consumer incomes

   (b) consumer incomes, tastes

   (c) the costs of production, bank opening hours

   (d) the price of related goods, preferences

8. An increase in consumer income will increase demand for a _____ but decrease demand for a _____.

   (a) substitute good, inferior good

   (b) normal good, inferior good

   (c) inferior good, normal good

   (d) normal good, complementary good

9. Education is

   (a) Merit Good

   (b) Public Good

   (c) Private Good

   (d) None of the above

10. If the demand curve for a good is horizontal and the price is positive, then a leftward shift of the supply curve results in –

    (a) A price of zero

    (b) an increase in price

    (c) A decrease in price

    (d) no change in price

11. Which of the following is an example of a normative statement?

    (a) A higher price for a good causes people to want to buy less of that good.

    (b) A lower price for a good causes people to want to buy more of that good.

    (c) To make the good available to more people, a lower price should be set.

    (d) Government should provide basic healthcare to all citizens.

12. Which of the following are assumptions underlying the PPC?

    (a) Only two goods are produced

    (b) Technology, population, and capital are variable

(c) Prices determine the position on the PPC

(d) All the above

13. If the price elasticity of demand for a good is .75, the demand for the good can be described as:

(a) Giffen

(b) Elastic

(c) Inferior

(d) Inelastic

14. Which of the following is not a property of indifference curve?

(a) Two curves do not intersect.

(b) Higher indifference curve shows higher level of satisfaction.

(c) Indifference curve slopes upward

(d) None of the above

15. Equilibrium is defined as a situation in which –

(a) no government regulations exist

(b) neither buyers nor sellers want to change their behavior

(c) demand curves are perfectly horizontal

(d) suppliers will supply any amount that buyers wish to buy

16. The extra utility from consuming one more unit of a good is called.

(a) Marginal utility

(b) Total utility

(c) Average utility

(d) Bonus utility

17. Marginal revenue is the _____ when output is _____.

(a) change in average revenue, increased

(b) change in total revenue, increased by one unit

(c) change in average revenue, increased by one unit

(d) change in total revenue, constant

18. Opportunity cost is –

(a) the additional benefit of buying an additional unit of a product.

(b) that which we forgo, or give up, when we make a choice or a decision.

(c) A cost that cannot be avoided, regardless of what is done in the future.

(d) the cost incurred in the past before we make a decision about what to do in the future.

19. The law of diminishing returns states that:

(a) as a firm uses more of a variable resource, given the quantity of fixed resources, the average product

of the firm will increase

(b) as a firm uses more of a variable resource, given the quantity of fixed resources, marginal product of the firm will eventually decrease

(c) in the short run, the average total costs of the firm will eventually diminish

(d) in the long run, the average total costs of the firm will eventually diminish

20. An upward shift in marginal cost _____output and an upward shift in marginal revenue _____ output

   (a) reduces, reduces

   (b) reduces, increases

   (c) increases, increases

   (d) increases, reduces

21. A profit maximizing monopolist produce at a point where:

   (a) MC = MR and MC cuts MR from below

   (b) MC> MR

   (c) MC < MR

   (d) MC = AR and MC cuts AR from below

22. In the short-run, a firm should close down if profits are zero.

   (a) true

   (b) false

   (c) may be

   (d) depends on determinants of demand

23. Other things being equal, a fall in the price of milk will cause _____.

   (a) A decrease in the demand for milk and an increase in the supply of milk

   (b) An increase in the quantity demanded of milk and a decrease in the quantity supplied of milk.

   (c) A decrease in the quantity demanded of milk and an increase in the quantity supplied of milk

   (d) An increase in the demand for milk and a decrease in the supply of milk

24. If the price of iPods increases and as a consequence, the demand for MP3s increases, then iPods and MP3s are –

   (a) Complementary products

   (b) Independent products.

   (c) Substitute products.

   (d) Ceteris paribus products

25. Refer to the below supply and demand diagram. If the market price is at P_1, then
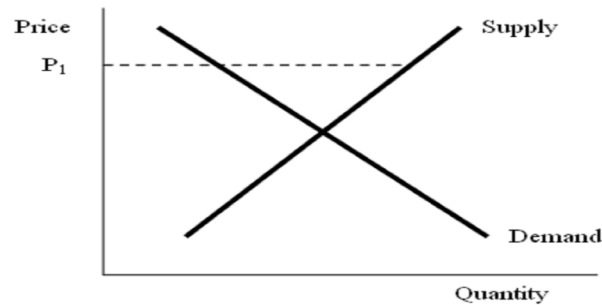
(a) a shortage will occur and the market price will fall.

(b) a shortage will occur and the market price will rise.

(c) a surplus will occur and the market price will fall.

(d) a surplus will occur and the market price will rise.

26. If marginal utility is zero:

(a) Total utility is zero

(b) An additional unit of consumption will decrease total utility

(c) An additional unit of consumption will increase total utility

(d) Total utility is maximized

27. The price of an item drops 10% in such a way that the Price Elasticity of Demand of that item is unit-elastic. We would expect the quantity of the item demanded to

(a) Drop by 5%

(b) stay the same

(c) increase by 5%

(d) increase by 10%

28. The reason that the marginal cost curve eventually increases as output increases for the typical firm is because –

(a) of diseconomies of scale

(b) of minimum efficient scale

(c) of the law of diminishing returns

(d) normal profit exceeds economic profit

29. The phrase '*ceteris paribus*' is best expressed as

(a) 'all else equal.'

(b) 'everything affects everything else.'

(c) 'scarcity is a fact of life.'

(d) 'there is no such thing as a free lunch.'

30. Following is the necessary condition for a firm to be in equilibrium.

    (a) MC=MR

    (b) MC cuts MR from below

    (c) MC=AR

    (d) None of the above

31. The central banking functions in India are performed by the

    (a) Central Bank of India

    (b) Reserve Bank of India

    (c) State Bank of India

    (d) Punjab National Bank

32. Banks create money by

    (a) printing it

    (b) issuing debit cards

    (c) accepting cheques

    (d) lending out part of their deposits

33. The money supply is _____.

    (a) Currency in circulation only

    (b) Demand deposits only

    (c) Cheques + money + bank cards + credit cards

    (d) Currency in circulation plus demand deposits

34. There are no barriers to entry under

    (a) perfect competition

    (b) oligopoly

    (c) monopoly

    (d) all of the above

35. An indifference curve represents

    (a) Combinations of two inputs.

    (b) Combinations of two goods and services yielding same level of satisfaction.

    (c) Both a and b.

    (d) None of the above.

36. 12th Five Year Plan ended in the year.

    (a) 2017

    (b) 2018

    (c) 2020

(d) None of the above

37. A profit-maximizing firm in the short run will expand output _____.

(a) until marginal cost begins to rise

(b) until total revenue equals total cost

(c) until marginal cost equals average variable cost

(d) as long as marginal revenue is greater than marginal cost

38. A non discriminating monopolist will find that marginal revenue _____.

(a) exceeds average revenue or price

(b) is identical to price

(c) is sometimes greater and sometimes less than price

(d) is less than average revenue or price

39. Monopolistic competition is similar to perfect competition because of –

(a) A large number of firms in the industry

(b) Standardized Product

(c) Difficult entry and exit in the industry by individual firms

(d) Price Discrimination

40. As a consumer move upward along an indifference curve, his total utility –

(a) First increases and then decreases

(b) First decreases and then increases

(c) Remains constant

(d) Increases

41. A distinguishing characteristic of monopolistic competition is –

(a) Large number of firms

(b) Low entry barriers

(c) Product standardization

(d) Product differentiation

42. The price elasticity of demand for a textbook is estimated to be 1, no matter what the price or quantity demanded. In this case,:

(a) a 10 percent increase in price will result in a 10 percent increase in the quantity demanded.

(b) a 10 percent increase in price will result in a 10 percent decrease in the quantity demanded.

(c) an increase in price will decrease the total revenue of sellers.

(d) A decrease in price will increase the total revenue of sellers.

43. The book entitled '*An Inquiry into the Nature and Causes of the Wealth of Nations*' was written by

(a) J. M. Keynes

    (b) J. S. Mill

    (c) J. B. Say

    (d) Adam Smith

44. Firms are assumed to _____ costs and to _____ profits.

    (a) incur, desire

    (b) pay, make

    (c) charge, earn

    (d) minimize, maximize

45. Which is not a determinant of demand?

    (a) income

    (b) the cost of inputs in production

    (c) the prices of related goods

    (d) future price expectations

46. If demand is _____ then price cut will _____ spending.

    (a) perfectly inelastic, increase

    (b) elastic, increase

    (c) elastic, decrease

    (d) None of the above

47. A positive cross elasticity of demand coefficient indicates that:

    (a) a product is an inferior good

    (b) a product is a normal good

    (c) two products are substitutes

    (d) two products are complementary goods

48. In perfect competition, the marginal revenue curve

    (a) and the demand curve facing the firm are identical

    (b) is always above the demand curve facing the firm

    (c) is always below the demand curve facing the firm

    (d) intersects the demand curve when marginal revenue is minimized

49. Which of the following is an example of economies of scale

    (a) As the computer industry has expanded, the number of professionally trained computer programmers has also increased, which has caused the salaries of computer programmers to increase

    (b) To attract firms to locate in its state, the state government reduced the tax rate that businesses must pay on its profits, thus lowering the costs to firms who locate in the state

    (c) A firm increases in size and is therefore able to lower its health insurance costs because as the size of the group insured increases the premium per person decreases substantially

(d) As the demand for calculators increased, the price of calculators actually fell.

50. From society's point of view, society would be better off if a monopolist –

(a) produced more and charged a higher price

(b) produced less and charged a higher price

(c) produced more and charged a lower price

(d) produced less and charged a lower price

51. The problem with barter economy remained that it required –

(a) less time and trouble to trade as compared with a money economy.

(b) a banking system from trade to occur.

(c) double coincidence of wants.

(d) that there be a single coincidence of wants.

52. In case of Perfect Competition, number of sellers would be –

(a) Large

(b) Single

(c) Varied but not too many

(d) None of the above

53. Cardinal utility approach is associated with the name of –

(a) A. Marshal

(b) J. R. Hicks

(c) R. G. D. Allen

(d) None of the above

54. If firms can neither enter nor leave an industry, the relevant time period is the –

(a) Short run

(b) Intermediate run

(c) Long run

(d) Immediate run

55. The short run is a time period in which –

(a) all resources are fixed

(b) the level of output is fixed

(c) the size of the production plant is variable

(d) some resources are fixed and others are variable

56. Two of the major factors contributing to growth are –

(a) money and efficiency

(b) money and luck

(c) resources and luck

(d) resources and a good climate

57. Which of the following is not correctly marked with respect to the year?

(a) Reserve bank of India 1977

(b) Industries(Development and Regulation)Act 1951

(c) Bureau of Industrial and Financial Reconstruction 1987

(d) Third five year plan 1961

58. Homi J. Bhabha was the first chairman of –

(a) Finance Commission

(b) Planning Commission

(c) Election Commission

(d) Atomic Energy Commission

59. The three annual plans from 1966-69 are also known as –

(a) Plan Break

(b) Plan Holiday

(c) All the above

(d) None of the above

60. TRAI stands for –

(a) Telecommunication Regulatory Authority of India

(b) Telecom Regulatory Authority of India

(c) Telecom Regulatory Area of India

(d) Telecom Regulatory Authority Institute

61. In the production function, equation expressed as Q_x = f(L,K,T…………n), L stands for –

(a) land

(b) labour

(c) longlife

(d) luxury goods

62. Reasons for increasing returns in Stage I of law of variable proportions is _____.

(a) Indivisibility

(b) Specialisation

(c) Both (a) and (b)

(d) None of the above

63. Homogeneous products are

(a) identical products

(b) differentiated products

(c) Both (a) and (b)

(d) None of the above

64. Profit is denoted by which of the following symbols?

(a) Sigma (σ)

(b) Summation (Σ)

(c) Pie (π)

(d) alpha (α)

65. An oligopoly market is characterized by

(a) A few dominating sellers.

(b) Strong barriers to entry

(c) Interdependence among sellers

(d) All of the above.

66. Which of the following has NOT been an objective of planning in India?

(a) creation of capitalist economy

(b) Self-Reliance

(c) Increase in National Income

(d) None of the above

67. Following was the last Five Year Plan in India

(a) 11th Five Year Plan

(b) 12th Five Year Plan

(c) 13th Five Year Plan

(d) 14th Five Year Plan

68. Agricultural Price Commission is now named as –

(a) Commission for Agricultural Costs and Prices (CACP)

(b) Costs and Prices Commission for Agricultural (CPCA)

(c) Agricultural Commission for Costs and Prices (ACCP)

(d) None of the above

69. IMR stands for

(a) Infant Mortality Rate

(b) Indian Mortality Rate

(c) Innocent Mortality Rate

(d) None of the above

70. HDI stands for

    (a) Human Development Index

    (b) Happiness and Development Index

    (c) Human Demographic Index

    (d) Human Development in India

71. Which of the following is not the basic economic problem?

    (a) What to produce

    (b) How to produce

    (c) For whom to produce

    (d) Where to produce

72. PPC shifts outward

    (a) If technology improves

    (b) If technology deteriorates

    (c) If quantity of productive resources decreases

    (d) Both b and c

73. Which of these is not a method of data collection?

    (a) Questionnaires

    (b) Interviews

    (c) Experiments

    (d) Observation

74. As consumption of a good increases utility derived from successive unit

    (a) Decreases

    (b) Increases

    (c) Remains constant

    (d) Either increases or remains constant

75. A consumer is said to be in equilibrium if utility from _____ of money spent on each commodity is equal

    (a) Last unit

    (b) First Unit

    (c) Some units

    (d) All units

76. Under ordinal utility analysis, we cannot _____ utility but utility derived can be _____

    (a) Measure, ranked

    (b) Rank, measured

(c) Increased, measured

(d) Rank, increased

77. Which of the following cost curves is a horizontally parallel line?

(a) Total Fixed Cost

(b) Total Variable Cost

(c) Total Cost

(d) Total Revenue Cost

78. Which of the following is not the part of statistical enquiry?

(a) Collection of data

(b) Presentation of data

(c) Manipulation of data

(d) Interpretation of data

79. Mean of 0.3, 5, 6, 7, 9,12, 0.2 is –

(a) 4.9

(b) 5.6

(c) 8

(d) None of the above

80. "Statistics is both, a science and an art" is given by –

(a) Keynes

(b) Tippets

(c) Bowley

(d) Spiegel

81. Which of the following terms best describes data that were originally collected at an earlier time by a different person for a different purpose?

(a) Primary data

(b) Secondary data

(c) Experimental data

(d) Field note

82. The graphical representation of a cumulative frequency distribution is called –

(a) Histogram

(b) Pie chart

(c) Ogive

(d) Historigram

83. In order to compare two or more related series, we consider –

    (a) Multiple bar chart

    (b) Grouped bar chart

    (c) Line graph

    (d) Both (a) or (b)

84. Which of the following statements is false?

    (a) Statistics is derived from the Latin word '*Status*'

    (b) Statistics is derived from the Italian word '*Statista*'

    (c) Statistics is derived from the French word 'Stato'

    (d) Statistics is derived from the German word '*Statistik*'

85. Consecutive rectangles in a Histogram have no space in between –

    (a) True

    (b) False

    (c) May or may not have space

    (d) Must have space in between

86. For a normal curve, the mean, median and mode are –

    (a) All equal

    (b) All zero

    (c) All different

    (d) The mean is always the largest and the mode always the smallest

87. Standard deviation is denoted by –

    (a) Square of sigma

    (b) Sigma

    (c) Square root of sigma

    (d) None of these

88. Statistics can best be considered as –

    (a) an art

    (b) a science

    (c) both (a) and (b)

    (d) neither (a) nor (b)

89. Which of the following is NOT a continuous variable?

    (a) height

    (b) weight

    (c) time

(d) number of persons in a family

90. In the second stage of the Law of Variable Proportions

(a) TP decreases at a decreasing rate

(b) TP increases at a decreasing rate

(c) TP increases at an increasing rate

(d) TP decreases

91. Census investigation is not preferred over simple investigation because –

(a) It is costlier

(b) It is time consuming

(c) Both a & b

(d) None of the above

92. The point of intersection of the 'less than' and the 'more than' ogives corelates to

(a) mean

(b) mode

(c) median

(d) none of the above

93. Mean is affected by _____

(a) Change of scale

(b) Origin

(c) both (a) and (b)

(d) none of the above

94. Law of returns to Scale is a _____concept

(a) Long-run

(b) Short-run

(c) Very short-run

(d) None of the above

95. Standard deviation can be calculated from _____ average.

(a) any

(b) mean only

(c) mean and median only

(d) none of the above

96. When LAC rises upwards, it indicates:

(a) Falling average cost

(b) Diseconomies > Economies

(c) Economies > Diseconomies

(d) Economies = diseconomies

97. The standard deviation of 10, 16, 10, 16, 10, 10, 16, 16 is –

(a) 4

(b) 6

(c) 3

(d) 0

98. If a card is drawn at random from a pack of 52 cards, what is the chance of getting a spade?

(a) 4/13

(b) 5/13

(c) 0.25

(d) 0.25

99. The sum required for to earn monthly interest of ₹ 1200 @ 18% p.a. is –

(a) 50000

(b) 60000

(c) 80000

(d) 100000

100. If the price of all commodities in a place has increased 125 times in comparison to the base period prices, the index number of prices for the place is now –

(a) 100

(b) 125

(c) 225

(d) 440

ANSWERS

1. (a)	34. (a)	68. (a)
2. (c)	35. (b)	69. (a)
3. (a)	36. (a)	70. (a)
4. (d)	37. (d)	71. (d)
5. (d)	38. (d)	72. (a)
6. (a)	39. (a)	73. (c)
7. (c)	40. (c)	74. (a)
8. (b)	41. (d)	75. (a)
9. (a)	42. (b)	76. (a)
10. (d)	43. (d)	77. (a)
11. (d)	44. (d)	78. (c)
12. (a)	45. (b)	79. (b)
13. (d)	46. (b)	80. (b)
14. (c)	47. (c)	81. (b)
15. (b)	48. (a)	82. (c)
16. (a)	49. (c)	83. (d)
17. (b)	50. (c)	84. (c)
18. (b)	51. (c)	85. (a)
19. (b)	52. (a)	86. (a)
20. (b)	53. (a)	87. (b)
21. (a)	54. (a)	88. (c)
22. (b)	55. (d)	89. (d)
23. (b)	56. (a)	90. (b)
24. (c)	57. (a)	91. (c)
25. (c)	58. (d)	92. (c)
26. (d)	59. (b)	93. (c)
27. (d)	60. (b)	94. (a)
28. (c)	61. (b)	95. (b)
29. (a)	62. (c)	96. (b)
30. (a)	63. (a)	97. (c)
31. (b)	64. (c)	98. (a)
32. (d)	65. (d)	99. (c)
33. (d)	66. (a)	100. (c)
	67. (b)	

NOTES