DIRECTORATE OF DISTANCE EDUCATION

UNIVERSITY OF JAMMU

JAMMU

SELF LEARNING MATERIAL

B.A. Semester-IV

Subject : ECONOMICS **UNIT : I - IV**

COURSE CODE : EC 401 **LESSON : 1 TO 16**

Dr. Hina S. Abrol

COURSE COORDINATOR

http:/www.distanceeducationju.in

Printed and published on behalf of the Directorate of Distance Education, University of Jammu, Jammu by the Director, DDE University of Jammu, Jammu.

ECONOMICS

Course Contributors
Dr. Reena Rani

Dr. Manohar Lal

Editing & Proof Reading by
Neelam Choudhary

Printed by Classic/19/ 150

CONTENTS

Detailed Syllabus

Semester IV

Quantitative Methods in Economics:

PREAMBLE : This course is to introduce students of Economics to the use of quantitative methods in Economics both as a tool for analysis as well as a tool for empirical validation and quantification. Thus statistical and mathematical tools are important to the study of Economics both to analyse data as well as to create mathematical models where complex arguments can be demonstrated in quantitative terms.

Unit I : Some Basic Concepts

Role of Mathematics in Economics, Number System - Integers, Rational Numbers; Irrational Numbers, Real Numbers and Imaginary Numbers; Gradient or Slope and Equation of Straight Line (Economic Examples - Budget line, linear demand curves), Rectangular Hyperbola; Equations-Different types of equations (non linear demand curves), Simple equations, simultaneous equations, quadratic equations & their solutions, (Economic Examples-Various Demand supply Equilibria)

Unit II : Functions and Differentiation

Functions–Different Functions in Economics; Limits and Continuity; Derivatives– Definition and Evaluation, Rules of Differentiation-differentiating polynomial functions; uses of Differentiation in Economics; Marginal Concepts and Elasticities. First and Second Order Conditions for optimisation, Examples from Economics - Utility Maximisation, Profit Maximisation

Unit III : Measures of Cenral Tendency and Dispersion

Measures of central tendency– Arthmetic Mean, Median, Mode, Geometric Mean and their merits and demerits; Measures of dispersion – Mean Deviation, Standard Deviation, Coefficient of Variation, Variance and their merits and demerits, Skewness-Definition; Coefficient of Skewness – Karl Pearson's measure, Kurtosis – Meaning: Measures of Kurtosis.

Unit IV : Bivariate Data

Covariance-definition & formula, Correlation Coefficient, Spearman's Rank correlation Coefficient, Curve fitting-fitting variable X against Y, Goodness of fit-Definition and formula, Uses in Economics-estimation of Demand Curves, supply curves, Consumption function.

Basic Reading List

1. Schaum's Series (2005), An introduction to Mathematical Economics, Tata McGraw Hill, New Delhi.

2. Chiang A.C & Wain Wight, Fundamentals of Mathematical Economics.

3. Chander, Romes (2007), Lectures on Elementary Mathematics for Economists, New Academic Publishing Co; New Delhi.

4. S.P. Gupta (2005), Statistical Methods, S. Chand & Sons New Delhi.

Additional Reading List

1. G.S. Manga (1972), Mathematics & Statistics for Economists, Vikas Publisting House, New Delhi.

2. C.B. Gupta & Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas Publishing House, New Delhi.

Unit I

Some Basic Concepts

ROLE OF MATHEMATICS IN ECONOMICS

LESSON NO. 1 **UNIT-I**

STRUCTURE

1.1. OBJECTIVES

After going through this lesson, you shall be able :

1. To understand the role of Mathematics in Economics.
2. To understand economic concepts.

3. To understand various views which have been given by different economists regarding importance of Mathematics in Economics.
4. To understand the importance of different tools of Mathematics in Economics.
5. To understand the number system.

1.2. INTRODUCTION

Mathematical modeling is used in economic analysis to study existing economic relationship and it helps economists to study how Mathematics is helpful in Economics by quantifying or providing measurement and meaning to economic concepts. Mathematics also plays a large role in the area of economic analysis. The concept of numbers and the structure of numbers that make up the number system are basic to the use of calculus— the most important branch of Mathematics useful for practical purposes especially for the students of Economics. In this lesson, we shall be discussing about the role of Mathematics in Economics and the number system.

1.3. ROLE OF MATHEMATICS IN ECONOMICS

Economics is the study of how resources are used as well as an analysis of the decisions made in allocating resources and distributing goods and services. Mathematics is the language of numbers and symbols that can be used to logically solve problems and precisely describe size, quantity & other concepts. Some complex problems could not be described and complicated problems could not be acted upon without the language of Mathematics and its support of logical processes to solve problems.

Mathematical modeling is used in economic analysis to study existing economic relationships and it helps economists study what might happen to the economy if a certain action is applied. Economic concepts and relationships can be measured in Mathematical indexes, formulas and graphs. Several areas of mathematics can be utilized in economic analysis including linear algebra, calculus, and geometry. Since economic concepts can be complex, it is important to use care in representing data & relationships.

Economics uses modeling to describe certain states of being and to analyze economic scenarios. Modeling suggests what will happen if certain actions are taken. Simulation of real world situation is possible with economic analysis and modeling which would not be possible without Mathematics.

Barnett, Ziegler & Byleen (2008, P – 210) described Mathematical modeling as 'the process of using Mathematics to solve real – world problems.'

1.3.1. Economic Literacy

Economics should be the concern of not only experts, but also laymen, as it influences all of us is our day to day lives. De Rooy introduces a concept called economic literacy where an individual understands the economic environment & how the environment affects the individuals. Some people may be frightened away from economic data because of mathematics. Economically literate people can interpret how economic metrics and news affect them and what economic relationships will create a successful environment for them as individuals. As a result, economically literate people will be interested in when conditions signal a recession or when prices rise along with unemployment e.g., individuals might care about inflation, because purchasing power declines with rising inflation. All these economic concepts are represented in mathematical terms.

1.3.2. Economic Concepts

Mathematics can help in visualizing and quantifying economic concepts. Formulas and graphs can be used to describe and display such concepts. Mathematics is also a way to deal with uncertainty in a problem. Many economic terms can be represented mathematically. The terms are used to describe values and behaviours concerning supply & demand, producer & consumer theory, imperfections in the market & strategic behaviour etc. The national economy is very complex so it is impossible for a single number or measure to accurately represent it whether as snapshot in time or over a period of time (De Rooy).

1.3.3. Mathematics in Price

Mathematics can be use to show relative size or whether something is high or low or large or small. A basic economic concept related to scarcity of resources is price. Price is "the exchange of goods & services for money" (Mc Afee, 2006 p-7). However, the true cost or opportunity cost is based on what must be given up or what can not be purchased because funds are allocated else where. McAfee (P-9) calls opportunity cost "the value of the best foregone alternative" Price – related economic measures include the consumer price index (CPI) and Personal Consumption Expenditure (PCE).

1.3.4. Role of Different Tools of Mathematics in Economics

The Following is the role of different tools of Mathematics in Economics.

1. **Matrices and Determinants:** The study of matrices and determinants is of immense significance in Economics. We find the application of matrices and determinants in various topics of Economics and Management such as linear Programming, theory of games, general equilibrium analysis, input output analysis etc.

2. **Maxima and Minima:** The Concept of maxima & minima plays a very useful utility role in almost all fields and especially in Economics. Every Consumer wants to maximize his utility. Every producer wants to maximise his revenue and minimise his costs which is possible through the Mathematical tools.

3. **Functions, limits and Continuity:** The concepts of function, limit and continuity are of fundamental importance in Economics. These concepts are very useful in understanding the derivatives of the function. Differential Calculus is the study of the changes that occurs in one variable when other variables on which it depends change. These are very useful in demand function, supply function and Consumption functions etc of Economics.

4. **Differential Equations and Difference equations:** Differential equations are used frequently in Economics. They are used to determine the conditions for dynamic stability in Microeconomic models of market equilibrium and to trace the time path of economic variables under various conditions:- They are used to estimate the total cost and revenue function from marginal cost and marginal revenue functions. Difference equations are widely used in Economics, to determine the conditions of dynamic stability in lagged economic models such as Cobweb Model, Harrod-Domer Model and Lagged Income Determination Model.

1.4. NUMBER SYSTEM

In Mathematical Economics we deal with numbers.

Let us, therefore, say a few words about number system before we learn different mathematical techniques.

The whole number system is divided into three main categories (systems) of numbers :

(a) Real numbers
(b) Imaginary numbers
(c) Complex numbers

We shall deal with the real number system in details.

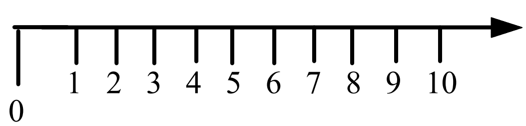
1.4.1. Natural Numbers, Whole Numbers & Integers

- **Natural Numbers :** The numbers used in counting, such as 1, 2, 3, 4... are called the natural numbers. The set of natural numbers is denoted by :

  N = {1, 2, 3, 4 ....}

- **Whole Numbers:** Whole numbers are simply the numbers

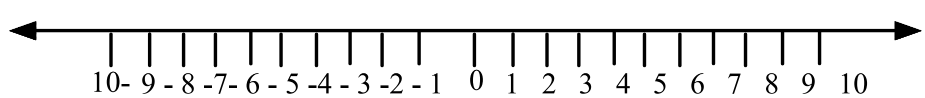
  0, 1, 2, 3, 4, 5, ...& (and so on)



No fractions

- **Counting Numbers:** "Natural Numbers" can mean either counting numbers { 1, 2, 3 ... } or whole number, { 0, 1, 2, 3 .... } depending on the subject.

- **Integers:** Integers are like whole numbers, but they also include negative natural numbers and zero, but still no fractions like ½. 1.1, 3.5 etc. are included. They are further divided into even and odd integers.



1.4.2. Rational Numbers

In Mathematics, a rational number is any number that can be expressed as the quotient or fraction P/q of two integers, P and q with the denominator $q \neq 0$, P is +ve, whereas q can be +ve or –ve integer.

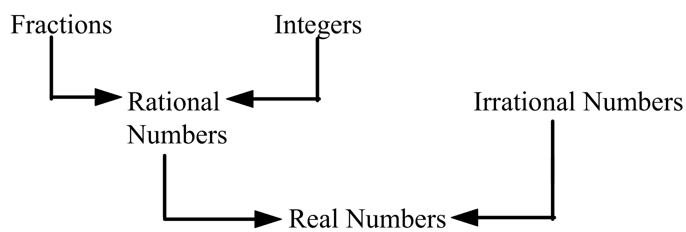
1.4.3. Irrational Numbers

Certain numbers can not be expressed in a form of fraction. Such numbers which can not be expressed as a ratio (m/n) of a pair of integers, such as $\sqrt{2}$, π, or e, are called irrational numbers.

$$\sqrt{2} = 1.4142, \ \pi = 3.1415, \ e = 2.718$$

1.4.4. Real Numbers

Integers, fractions, rational and irrational numbers all together form a system of numbers which is called Real number system. Real number system, therefore, can be explained in the following way:



We must remember two important things about the rational numbers :

1. There lie an infinite number of rational numbers between any two rational numbers. For example 1 & 2 are rational numbers; in between them $3/4$, $6/5$, $7/6$, $7/4$, $\frac{8}{7}$ and so on; are also the rational numbers.
2. There is at least one irrational number between the two rational numbers.

1.4.5. Imaginary Numbers

Just against the real numbers we have imaginary numbers. These numbers are concerned with the square root of negative numbers.

It is easy to find square root of positive numbers :

$$X^2 = 4, \; X = \pm 2$$

But if $X^2 = -4$, can we find the value of X?

When $X^2 = -4$, $X = \pm\sqrt{-4}$ which is not equal to ± 2 Thus, only a positive number can have a real valued square root.

To find the square root of a negative number, we make use of one (imaginary number) iota – symbolized as 'i'.

$$i = \sqrt{-1} \text{ so that } i^2 = -1$$

Suppose $X^2 = -4$

$\therefore X^2 = \pm\sqrt{4i^2} = \pm 2i$ and in this way we are able to get the square root of a negative number with the help of an imaginary quantity 'i'.

All the numbers which are expressed in terms of this imaginary quantity are called imaginary numbers.

$\sqrt{-4} = \pm 2i$, $\sqrt{-2} = \sqrt{2}i$ and $\sqrt{-8} = \sqrt{8}\,i$ are the example of imaginary numbers.

1.4.6. Complex numbers

A times we come across numbers which are combination of both imaginary and real numbers. Such numbers are called complex numbers.

For Example, $2+\sqrt{-9} = 2+3i$ contains real part (=2) and imaginary part (=3i) hence it is called complex number. Thus if a & b are two real numbers then $a + ib$ and $a - ib$ are two complex numbers:

Example 1: Evaluate: $\sqrt{-4}\times(1-\sqrt{-64})$

Sol. $\sqrt{-4}\times(1-\sqrt{-64}) = \sqrt{-1\times 4}\times(1-\sqrt{-1\times 64})$

$= 2i \times (1 - 8\ i) = 2\ i - 16\ i^2$

$= 2i - 16i^2 = 2i - 16\ (-i)$

$= 2i + 16$

$= 16 + 2\ i$

Example 2: Simplify $2i^{2} + 6i^3 + 3i^{16} - 6i^{19} + 4i^{25}$

Sol.

$= 2\ i^2 + 6i^2 \times i + 3(i^2)^8 - 6(i^2)^2 \times i + 4(-i^2)^{12} \times i$

$= 2(-1) + 6(-1) \times i + 3(-1)^8 - 6(-1)^9 \times i + 4(-1)^{12} \times i$

$= -2 - 6i + 3 + 6i + 4i$

$= 1 + 4i$

Example 3: Show that $1+i^{10}+i^{20}+i^{30}+i^{40}+i^{50}$ is a real number

Sol. $1+i^{10}+i^{20}+i^{30}+i^{40}+i^{50}$

$= 1+(i^2)^5 + (i^2)^{10} + (i^2)^{15} + (i^2)^{20} + (i^2)^{25}$

$= 1 + (-1)^5 + (-1)^{10} + (-1)^{15} + (-1)^{20} + (-1)^{25}$

$= 1 - 1 + 1 - 1 + 1 - 1 = 0$

Which is a real number.

Check Your Progress

Answer the questions in the space provided :

1. Write a note on natural numbers, whole numbers and integers.

____

____

____

____

____

2. Give examples of rational and irrational numbers.

3. What are real, imaginary and complex numbers ? Give examples.

1.5. PROPERTIES OF REAL NUMBERS

Important Properties of real numbers are:

(i) $a + b = b + a$

This called as Commutative law of addition

(ii) $(a + b) + c = a + (b + c)$

This is called as associative law of addition

(iii) $a \times b = b \times a$ or $ab = ba$

This is called as Commutative law of multiplication

(iv) $a \times (b \times c) = (a \times b) \times c$

Or

abc = bac   or   a(bc) = (ab)C

This is known as associative law of multiplication .

(v) a + b = c + b $\longrightarrow$ a = c

This is known as cancellation law of addition

(vi) a $\times$ b = c $\times$ b $\longrightarrow$ a = c

This is known cancellation law of multiplication

(vii) a + 0 = 0 + a = a

This is known as additive identity

(viii) a $\times$ 1 = 1 $\times$ a = a

This is known as existence of identities of multiplication.

(ix) a + (–a) = 0

This is known as existence of additive inverse.

(x) $a \times \frac{1}{a} = 1$

This is known as existence of multiplication inverse.

(xi) $\frac{0}{0}$ is called indeterminate.

(xii) 0 divided by any number is equal to 0.

(xiii) Any number divided by zero is equal to infinity (∞)

(xiv) a $\times$ (b+c) = ab+ac

This is know as distributive law.

1.6. LET US SUM UP

In this lesson, we discussed the role of Mathematics in Economics and the number system in detail.

1.7. EXAMINATION ORIENTED QUESTIONS

Q.1. Discuss the role of Mathematics in Economics.

Q.2. Discuss the role of different tools of Mathematics in Economics.

Q.3. Show that $1 + i^{10} + i^{20} + i^{30} + i^{40} + i^{50}$ is a real number.

Q.4. State the meaning of integers.

Q.5. What are irrational numbers?

1.8. SUGGESTED READINGS & REFERENCES

1. Aggarwal, C.S. & R.C. Joshi : Mathematics for students of Economics.
2. Allen, R.G.D.: Mathematical Analysis for Economists (Macmillan)
3. Kandoi B: Mathematics for Business and Economics with applications (Himalya Publishing House)

GRADIENT OR SLOPE OF A STRAIGHT LINE

LESSON NO. 2 **UNIT-I**

STRUCTURE

2.1. OBJECTIVES

After going through this lesson, you shall be able to understand:

1. Meaning of Gradient or slope of a straight line.
2. Find the slope of line.
3. Find the equation of straight line whose slope is given.

2.2. INTRODUCTION

The gradient also called slope of a straight line shows how steep a straight line is.

2.3. GRADIENT OR SLOPE OF A STRAIGHT LINE

The slope of a straight line is the tangent of the angle, which the line makes with the positive direction of X-axis. The slope of the line is generally denoted by 'm'. e.g. if there is a straight line AB, that makes an angle θ with the positive direction of X-axis, then $\tan\theta$ is the slope of line AB.

$\therefore\ m = \tan\theta$

Here few things need to be noted.

- If angle θ is acute i.e. $< 90°$ (fig. a), the straight line is positively sloped.

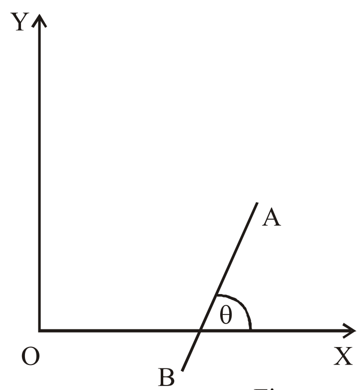


Fig. a

- If angle θ is obtuse i.e. $> 90°$ (fig. b), it gives rise to a negatively sloped straight line.

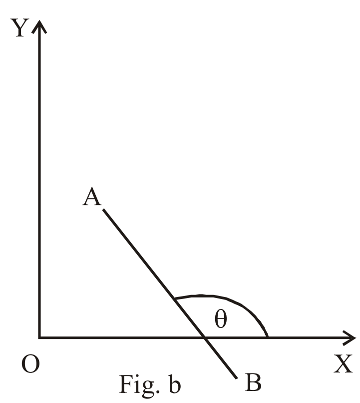


Fig. b

- If a line is parallel to X-axis, it means, it doesn't make any angle with X-axis

  $\therefore \quad \theta = 0° \;\&\; m$ i.e. slope $= \tan\theta = \tan 0° = 0$ (fig. c)

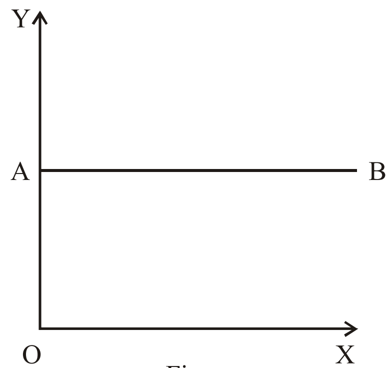


Fig. c

- If a line is perpendicular to the X-axis, $\theta = 90°$.

  $\therefore \quad$ Slope $= \tan\ \theta = \tan\ 90° = \infty$ (fig. d)

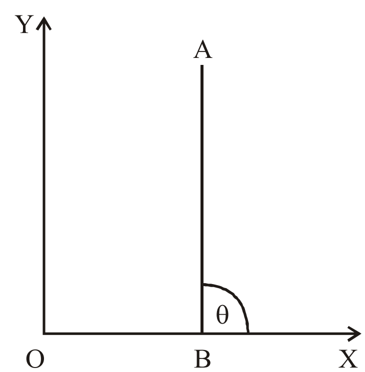
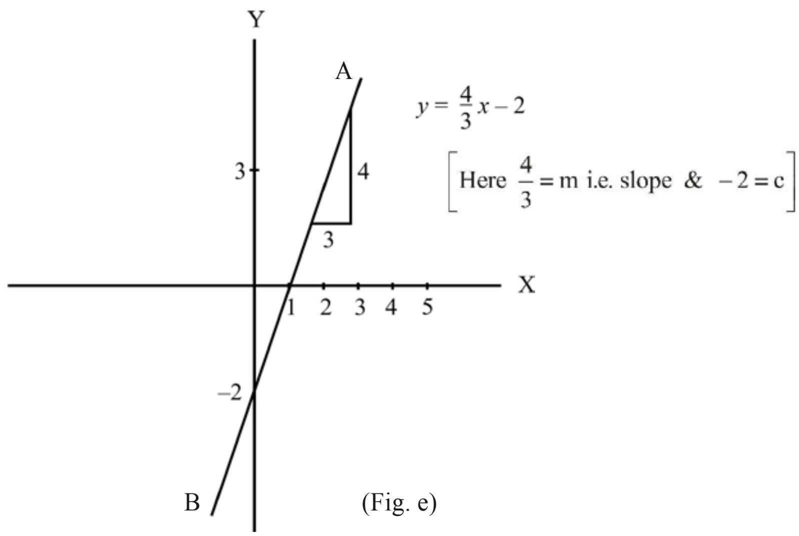


Fig. d

2.4. EQUATION OF STRAIGHT LINE

The equation of straight line is in the form $y = mx+C$ (m and C are numbers), where m is the gradient of the line and C is the y-intercept (where the graph crosses the y-axis).

Another way to calculate gradients is to divide vertical distance by horizontal distance.



(Fig. e)

In the above figure, line AB cuts Y axis at point –2, which means Y intercept is negative.

The above graph has equation $y = \frac{4}{3}x - 2$

Comparing it with y = mx + c, we find that

$m = \frac{4}{3}$ and $c = -2$

EXAMPLES

Example 1. Find the slope of $\sqrt{3}x + y = 12$

Solution: Given $\sqrt{3}x + y = 2$

$\Rightarrow \qquad y = -\sqrt{3}x + 12$

Compare it with $y = mx + C$

Here $\qquad m = -\sqrt{3}$

so slope $\qquad m = -\sqrt{3}$ **Ans.**

Example 2. Find the slope of line $y = 2x-7$

Solution: Given $y = 2x-7$

Compare it with $y = mx+C$

Hence $m = 2$

so slope of line $m = 2$

Example 3. Find the slope of line $3x+2y+6 = 0$

Solution: Given $3x+2y+6 = 0$

$$2y = -3x-6$$

$$y = \frac{-3}{2}x-3$$

Compare it with $y=mx + C$

Here $m = \frac{-3}{2}$ = slope

Example 4. Find the equation of straight line whose slope is 3 and cuts off an intercept –5 on y-axis.

Solution: Given slope $m = 3$

Intercept $C = -5$

so equation is $y = mx + C$

$y = 3x-5$ **Ans.**

Example 5. Prove that the lines $y = 7x$ and $y = 7x+\frac{1}{2}$ are parallel.

Solution: Line first $y = 7x$

Compare it with $y = mx+C$

so slope $m = 7$

Line second $y = 7x+\frac{1}{2}$

Compare it with $y = mx+C$

so slope $m = 7$

As slope of both lines are equal so lines are parallel.

Example 6. Find the equation of a straight line whose slope is $\frac{-2}{3}$ and intercept is 5

Solution: Given $m = \frac{-2}{3}$, $C = 5$

So equation of straight line is $y = mx + C$

Putting values in it, we get

$$y = \frac{-2}{3}x+5$$

Check Your Progress–I

1. What do you mean by the gradient of a straight line ?

2. Prove that the lines $y = 5x - 7$ and $2y = 10x + 5$ are parallel.

3. Find the equation of a straight line, where slope is 3 and passes through the origin.

4. Find the slope of $\sqrt{3}\,x + y = 12$

2.4.1. Equation of Straight Line in Point Slope Form

The equation of a straight line passing through a given point (x_1, y_1) and making an angle θ with X-axis is :

$$y - y_1 = m(x - x_1)$$

where $m = \tan\theta$ = Slope and x_1 and y_1 are the coordinates of the given point.

Example 1 : Find the equation of a straight line which passes through the point (–1, 3) and has slope 2.

Sol. The required equation is :

$$y - y_1 = m(x - x_1)$$

We are given that $x_1 = -1,\ y_1 = 3,\ m = 2$

$\therefore \qquad y-3 = 2[x-(-1)]$

i.e. $\qquad y-3 = 2x+2 \qquad$ or $\qquad 2x-y+5 = 0$

Example 2 : The gradient of a straight line is 12. It passes through a point whose coordinates are (5, 10). What is the function of the line ?

Sol. Here $x_1 = 5$, $y_1 = 10$ and $m = 12$

$\therefore \qquad y-y_1 = m(x-x_1)$

or $\qquad y-10 = 12(x-5)$

$\qquad y-10 = 12x-60$

By re-arranging, we get

$\qquad 12x-y-50 = 0$

Check Your Progress–II

1. Find the equation of a straight line having slope 0.50 and passing through the point (2, 2).

2. Find the equation of a straight line having slope of 5 and passing through the point (2, 3).

2.4.2. Equation of Straight Line in Intercept Form

If line AB cuts off a given intercept 'a' and 'b' on X and Y axis respectively, the equation is :

$$\frac{x}{a}+\frac{y}{b}=1$$

This is shown in fig. f, below :

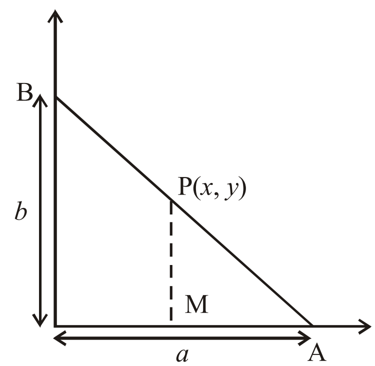


Fig. f

EXAMPLE

Example 1 : Find the equation of straight line which cuts off intercept $\frac{-1}{2}$ and 2 as the coordinate axis.

Sol. Equation of line in intercept form is

$$\frac{x}{a}+\frac{y}{b} = 1$$

Here $\quad a = \frac{-1}{2}, \quad b = 2$ given

$$\frac{x}{\frac{-1}{2}}+\frac{y}{2} = 1 \quad \Rightarrow \quad \frac{-2x}{1}+\frac{y}{2} =1$$

$\Rightarrow \qquad -4x + y = 2$

$\Rightarrow \qquad 4x-y+2 = 0$

Example 2: Find the equation of straight line passing through (2, 3) and cuts off equal intercepts on the axis.

Sol. Equation of line in intercept form is

$$\frac{x}{a}+\frac{y}{b} = 1$$

As it passes through (2, 3)

$$\therefore \qquad \frac{2}{a}+\frac{3}{b} = 1$$

Given $\qquad$ a = b = a

$$\Rightarrow \qquad \frac{2}{a}+\frac{3}{a} = 1$$

$$a = 5 = b$$

So equation is

$$\Rightarrow \qquad \frac{x}{a}+\frac{y}{b} = 1$$

$$\frac{x}{5}+\frac{y}{5} = 1$$

$$x + y = 5$$

$$x + y - 5 = 0 \quad \textbf{Ans.}$$

Example 3. Find equation of Straight line making intercepts of 3 and – 4 on *x*-axis and *y*-axis

Sol. Equation of straight line in intercept form is

$$\frac{x}{a}+\frac{y}{b} = 1$$

$$a = 3,\ b = -4 \quad \text{so}$$

$$\frac{x}{3}+\frac{y}{-4} = 1$$

$$\frac{x}{3}-\frac{y}{4} = 1$$

$$4x-3y = 12 \quad \textbf{Ans.}$$

Example 4. A line passes through (–3, 10) and sum of its intercepts on axes = 8, find the equation.

Sol. $a + b = 8$

$$\frac{x}{a}+\frac{y}{b} = 1$$

$\therefore$ $$\frac{-3}{a}+\frac{10}{b} = 1$$

But $a + b = 8$

$$a = 8 - b$$

So $$\frac{-3}{8-b}+\frac{10}{b} = 1$$

$$\frac{-3b+80-10b}{8b-b^2} = 1$$

or $$b^2-21b + 80 = 0$$

$$(b-5)(b-16) = 0$$

$$b = 5,\ 16$$

$\therefore$ $$a = 3,\ -8$$

So the given equations are :

$$\frac{x}{3}+\frac{y}{5} = 1$$

and $$\frac{x}{-8}+\frac{y}{16} = 1$$

2.4.3. Equation of Straight Line in Two Point Form

If a straight line AB passes through two points (x_1, y_1) and (x_2, y_2), its equation is :

$$y-y_1 = \frac{y_2-y_1}{x_2-x_1}(x-x_1)$$

Example 1. Find the equation of the line passing through the point (3, 5), (4, 7).

Sol. Equation in two points form is

$$y-y_1 = \frac{y_2-y_1}{x_2-x_1}(x-x_1)$$

Here $(x_1, y_1) = (3,5)$, $(x_2, y_2) = (4,7)$

So equation is $y-5 = \frac{7-5}{4-3}(x-3)$

$$y-5 = \frac{2(x-3)}{1}$$

$$y-5 = 2x-6$$

$$y = 2x-1$$

$$2x-y-1 = 0$$ **Ans.**

Examples 2. Prove that the points (0, 1) and (1, 2) and (–2, –1) will be on the same straight line.

Sol. Given $(x_1, y_1) = (0, 1)$, $(x_2, y_2) = (1, 2)$

$(x_3, y_3) = (-2, -1)$

Now equation of line in two form

$$(y-y_1) = \frac{y_2-y_1}{x_2-x_1}(x-x_1)$$

$$(y-1) = \frac{2-1}{1-0}(x-0)$$

$$y-1 = x$$

$$x-y+1 = 0$$

Now for points (–2, –1) to lie on this line, $x-y+1=0$ should hold true for Points (–2, –1).

$$\therefore \quad (-2)-(-1)+1 = 0$$

$$-2+1+1 = 0$$

$$-2+2 = 0$$

$$\Rightarrow \quad 0 = 0$$

Hence this point also lies on the same straight line, which means that these three points are collinear.

Examples 3 : Find the equation of straight line passing through points (3, 9) and (4, 8). Find the coordinates where it meets the axes.

Sol. Equation in two point form is

$$(y-y_1) = \frac{y_2 - y_1}{x_2 - x_1}(x-x_1)$$

$$(x_1, y_1) = (3, 9), (x_2, y_2) = (4,8)$$

$$y-9 = \frac{8-9}{4-3}(x-3)$$

$$y-9 = -(x-3)$$

$$\Rightarrow \quad y-9 = -x+3$$

$$x+y-12 = 0$$

Now when $x = 0$

$$0+y-12 = 0, \ y = 12$$

It meets y-axis at = (0, 12)

Further, when $y = 0$, $x-12 = 0$ or $x = 12$

$\therefore$ It meets x-axis at (12, 0).

Check Your Progress–III

1. Find out the equation of a straight line passing through points (1, 20) and (3, 10). Does the point (2, 15) lie on this line ?

2. Find the equation of a straight line passing through points (2, 0) and (4, –2). Also show that point (5, –3) lies on the same straight line.

3. Given points (0, 1), (1, 2) and (–2, –1) on a straight line, find the coordinates where it meets the axes.

2.5. APPLICATION OF STRAIGHT LINE IN ECONOMICS

The concept of straight line has wide applications in Economics. Let us discuss here some important applications :

In many cases, the relationship between economic variables may be linear. Demand curve is the linear relationship between the amount demanded and the price of a commodity. Therefore, the demand curve is a straight line. In general, the concept of straight line is widely used in economic analysis. We are giving certain facts regarding this application.

(*a*) The demand and supply curves are the linear functions of price and so represent the straight line. So the demand and supply curves are expressed by the equations $D = a - bp$ and $S = -a_1 + b_1p$ respectively (a, b,

a_1, b_1 being positive constants) where D stands for quantity demanded, S for quantity supplied and p for price per unit.

(*b*) The cost curve of a firm is a linear function of output. Therefore, the cost curve of a firm is a straight line expressed by the equation $C = a+bq$, where C = total cost, q = units of output, and a, b are positive constants. Slope of the line is marginal cost which remains constant at every level of output. When no output is produced *i.e.*, when $q = 0$, then total cost = a, which shows us that 'a' is fixed cost. It is also the y-intercept of the cost line.

2.5.1. Budget Line

In an earlier class, we have discussed what a budget line is. A budget line refers to a graphical specified depiction of the various combinations of two selected products that a consumer can afford at specified prices for the products given their particular income level.

Price Ratio and the Slope of the Budget line

Think of any point on the budget line. Such a point represents a bundle of goods which costs the consumer his/her entire budget. Now suppose the consumer wants to have one more unit of good. He can do it only if he gives up some amount of the other good. How much of good does he have to give up if he wants to have an extra unit of good ? It would depend on the prices of the two goods. A unit of good X costs P. Therefore, he will have to reduce his expenditure on good Y by P_x amount. With P_x, he could buy $\frac{P_x}{P_y}$ units of good Y. Therefore, if the consumer wants to have an extra unit of good X on which he is spending all his money, he will have to give up $\frac{P_x}{P_y}$ unit of good Y. In other words in the given market conditions the consumer can substitute good X for good Y at the rate $\frac{P_x}{P_y}$. The absolute value of the slope of the budget line measures the rate at which the consumer is able to substitute good X for good Y when he spends his entire budget.

The slope of price line represents the ratio of prices between two commodities

$$\therefore \quad \text{Slope of Price line} = \frac{P_x}{P_y}$$

It should be noted that the consumer can not have any point of combinations beyond the budget line.

The Budget line is also called as income line, because it represents the real income of the consumer.

It is also known as the "Consumption Possibility Line".

Equation of the Budget Line

Equation of budget line is $xP_x + yP_y = M$

where X, Y refer to $\rightarrow$ Two commodities

P_x, P_y $\rightarrow$ Prices of two commodities

M $\rightarrow$ Consumer's Income

The above equation can be expressed in intercept form as :

$$\frac{x}{M/P_x} + \frac{y}{M/P_y} = 1$$

Here

$\frac{M}{P_x}$ is the intercept on X-axis, which shows the number of units of X purchased if whole income is spent by the consumer on X-commodity.

Likewise

$\frac{M}{P_y}$ shows the Intercept on Y-axis or the amount of Y consumed, if nothing is spent on commodity X.

From this equation, we can again derive that

$$\text{Slope of budget line} = \frac{-\text{Coefficient of X}}{\text{Coefficient of Y}}$$

$$= \frac{-P_x}{P_y}$$

A budget line is thus negatively sloped.

EXAMPLES

Example 1: A consumer wants to consume two goods. The prices of the two goods are Rs. 4 and Rs. 5 respectively. The consumer's income is Rs 20.

(I) Write down the equation of the budget line.

(II) How much of good X can the consumer consume if she spends her entire income on that good?

(III) How much of good Y can consumer consume if she spends her entire income on that good?

(IV) What is the slope of the budget line?

(V) How does the budget line change if the consumer's income increases to Rs 40 but the prices remain unchanged?

(VI) How does the budget line change if the Price of good Y decreases by a rupee but the price of good X and the consumer's income remain unchanged?

(VII) What happens to the budget set if both the prices as well as the income double?

Solution: (I) Equation of the budget line is $xP_x + yP_y = M$

where $P_x = 4$

$P_y = 5$

$M = 20$

$\therefore \quad 4x + 5y = 20$

(II) Since $y = 0$

$4x + y = 20$

$4x = 20$

$$x = \frac{20}{4}$$

$x = 5$

The consumer can consume 5 units of good X if the consumer spends her entire income on good X.

(III) $$4x + 5y = 20$$

$$x = 0$$

$$\therefore \quad 5y = 20$$

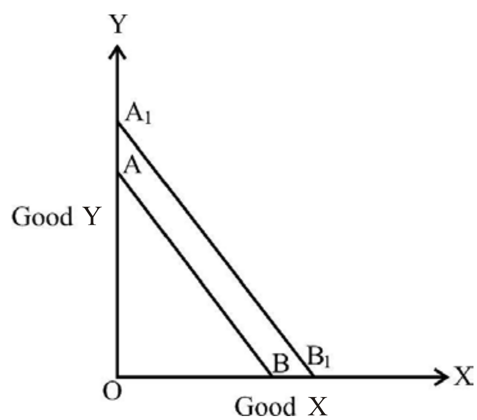
$$y = \frac{20}{5}$$

$$y = 4$$

The consumer can consume 4 units of good Y if the consumer spends her entire income on good Y.

(IV) Slope of the budget line $= \dfrac{-P_x}{P_y} = \dfrac{-4}{5} = -\ 0.8$

(V)



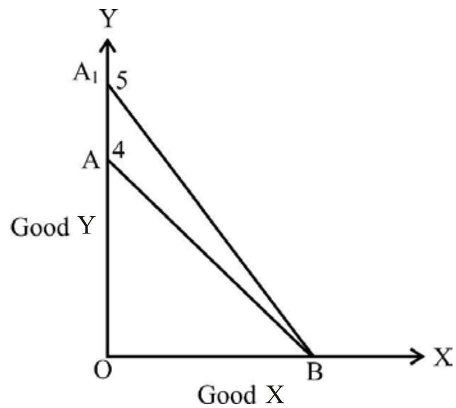
$$P_x = 4,\ P_y = 5,\ M = 40$$

$$x = \frac{M}{P_x} = \frac{40}{4} = 10 \text{ units}$$

$$y = \frac{M}{P_y} = \frac{40}{5} = 8 \text{ units}$$

The budget line will change as shown in the diagram (above). Here budget line AB has changed to A_1B_1. The new budget line is parallel to the original budget line, but to the right of it, as income has doubled.

(VI)



$$M = 20$$

New $P_y = 5 - 1 = 4$

New $y = \frac{20}{4} = 5$

The budget line will change as under:

The original budget line was AB.

The new budget line is A_1B. So there is a rightward expansion on the Y-axis.

(VIII) $M = 20 \times 2 = 40$

New $P_x = 4 \times 2 = 8$

$$x = \frac{40}{8} = 5$$

New $P_y = 5 \times 2 = 10$

$$y = \frac{40}{10} = 4$$

Since all the variables of budget set change in the same proportion, there will be no change in the budget set.

Example 2 : If a consumer spends his given income on A commodity he can purchase 40 units of A and if he spends his given income on B he can purchase 30 units of B. Find the budget equation. Also find the given level of income of the consumer and price per unit of the commodities.

Sol. We find from the given problem, that

$$\frac{M}{P_x} = \text{X-intercept} = 40, \quad \frac{M}{P_y} = \text{Y-intercept} = 30$$

$\therefore$ The equation of budget line is $\frac{x}{40}+\frac{y}{30}=1$ or $3x + 4y = 120$

Therefore, the required budget equation is $3x + 4y = 120$

Comparing it with the budget equation $xP_x + yP_y = M$

Here $M = 120,\ P_x = 3,\ P_y = 4$

$\therefore$ given income = 120

P_x = price of commodity A = 3

P_y = price of commodity B = 4.

Example 3 : A consumer gets an increase of Rs.100 per month and wishes to spend all of it on two goods X, Y whose prices are 4 and 5 respectively. Find the budget line and show that a doubling of all money prices has exactly the same effect on the budget line as halving money income.

Sol. New level of income = 100

P_x = price of X commodity = 4, $P_y = 5$

equation of budget line is $xP_x + yP_y$ = level of income

i.e., $x.4 + 5y = 100$ ....(1)

When prices are doubled, then from (1)

$8x + 10y = 100$ ....(2)

When money income is halved, then from (1),

$$x4 + 5y = 50 \qquad ....(3)$$

Since (2) and (3) is same, hence the doubling of all money prices has exactly the same effect on budget line as halving money income.

2.5.2. Linear Demand Curve

Both demand and supply curves are linear functions of price & hence represent a straight line. The demand curve is represented by the equation

$$D = a - bp$$

& Supply curve by :

$$S = -a_1 + b_1 p$$

where a, b, a_1 and b_1 are positive constants, D = Quantity, S = Quantity supplied and p is the price per unit.

Example 1. (*a*) When the price is Rs. 80 per watch, 10 watches are sold, 20 watches are sold when the price is Rs. 60. Find the linear demand function.

(*b*) When the price is Rs. 100 no watches are sold. When watches are free, 50 are demanded. Find the linear demand function.

(*c*) When the price is Rs. 50 there are 50 watches of brand XXX available for market. When the price is Rs. 70 there are 100 watches available for market. What is linear supply function ?

Sol. (a) Demand Curve

The demand curve passes through points whose co-ordinates are (10, 80) and (20,60), where x = demand in units, y = price in rupees.

The linear demand curve is $y - y_1 = \dfrac{y_2 - y_1}{x_2 - x_1}(x - x_1)$

Here $\qquad x_1 = 10,\ x_2 = 20,\ y_1 = 80,\ y_2 = 60$

$$y - 80 = \frac{60 - 80}{20 - 10}(x - 10)$$

i.e., $\qquad y - 80 = \dfrac{-20}{10}(x - 10) = -2x + 20$

or $\qquad 2x + y - 100 = 0$, is the required linear demand function.

(b) When $x = 0$, $y = 100$, where x = demand in units, y = price

When $x = 50$, $y = 0$

So linear demand curve through (0, 100) and (50, 0) is

$$y - 100 = \frac{0-100}{50-0}(x-0) = -2x$$

or $\qquad 2x + y - 100 = 0.$

(*c*) Linear supply function through (50, 50) and (100, 70) is

$$y - 50 = \frac{70-50}{100-50}(x-50)$$

where x = no. of watches, y = price in rupees.

i.e., $$y - 50 = \frac{20}{50}(x-50)$$

or $\qquad 5y - 250 = 2x - 100$

or $\qquad 2x - 5y + 150 = 0$ is the required supply function.

Example 2. From the following schedule :

Price per unit (in Rs.)	Supply (in units)	Demand (in units)
3	30	50
4	40	40
5	50	30

Find (*i*) linear supply and linear demand equations, (*ii*) equilibrium price and quantity.

Sol. (*a*) Demand Curve

The demand curve passes through the points whose co-ordinates are (50, 3), (40, 4), (30, 5). The linear demand curve passes through (50, 3) and (40, 4) is

$$P - P_1 = \frac{P_2 - P_1}{Q_2 - Q_1}(Q - Q_1) \qquad \text{....}(i)$$

where P = price, Q = Quantity demanded.

Here $Q_1 = 50$, $Q_2 = 40$, $P_1 = 3$, $P_2 = 4$.

$$P - 3 = \frac{4-3}{40-50}(Q - 50)$$

or $$P - 3 = -\frac{Q}{10} + 5$$

$$P = -\frac{Q}{10} + 8$$

The point (30, 5) also satisfies the equation

i.e. $$5 = -\frac{30}{10} + 8 \quad \textit{i.e.,}\ 5 = 5$$

∴ The equation of the demand curve passing through the given three point is

$$P = -\frac{Q}{10} + 8$$

(b) Supply Curve

The supply curve passes through the points whose co-ordinates are (30, 3), (40, 4), (50, 5). The linear supply curve passing through (30, 3) and (40, 4) is

$$P - P_1 = \frac{P_2 - P_1}{Q_2 - Q_1}(Q - Q_1)$$

where P = price, Q = Quantity supplied.

Here $P_1 = 3$, $P_2 = 4$, $Q_1 = 30$, $Q_2 = 40$.

$$P - 3 = \frac{4-3}{40-30}(Q - 30)$$

or $$P - 3 = \frac{Q}{10} - 3$$

$$P = \frac{Q}{10}.$$

The point (50, 5) also satisfies the equation.

Hence Demand Curve is $P = -\frac{Q}{10} + 8$ and Supply Curve is $P = \frac{Q}{10}$.

(c) At equilibrium, D = S

$$\frac{-Q}{10} + 8 = \frac{Q}{10}$$

or $$\frac{2Q}{10} = 8$$

or $$Q = 40$$

From supply curve $P = \frac{Q}{10}$

We get, $$P = \frac{40}{10} = 4$$

Equilibrium quantity = 40, and equilibrium price = 4.

This can also be verified from the given table which shows that at P = 4, quantity demanded is equal to quantity supplied.

2.6. LET US SUM UP

In this lesson, we discussed the concept of slope of a straight line and its wide applications is Economics like budget line, Linear Demand Curve etc.

2.7. EXAMINATION ORIENTED QUESTIONS

Q.1. If a consumer spends his given income on X-commodity he can purchase 50 units of X and if he spends his given income on Y he can purchase 20 units of Y. Find the budget equation. Also find the given level of income of the consumer and prices per unit of the commodities.

Q.2. When the price of a commodity is Rs.30 per unit, the demand and supply are 600 and 900 units respectively. A price of Rs.20 per unit changes the demand and supply to 1000 and 700 respectively. Assuming the demand and supply function to be linear, find

(i) demand function,

(ii) the supply function,

(iii) the equilibrium price and quantity.

2.8. SUGGESTED READINGS & REFERENCES

1. Aggarwal, C.S. & R.C. Joshi : Mathematics for Students of Economics.

2. Allen, R.G.D.: Mathematical Analysis for Economists (Macmillan).

RECTANGULAR HYPERBOLA

LESSON NO. 3 **UNIT-I**

STRUCTURE

3.1. OBJECTIVES

After going through this lesson, you shall be able to understand :

1. Meaning of Rectangular Hyperbola
2. Equation of Rectangular Hyperbola
3. Application of Rectangular Hyperbola in Economics

3.2. INTRODUCTION

In this lesson, we shall be discussing a very important concept i.e. Rectangular Hyperbola, which has many applications in Economics.

3.3. MEANING OF RECTANGULAR HYPERBOLA

A rectangular hyperbola is defined as the locus of a point which moves such that the product of its perpendicular distance from fixed line perpendicular to each is a positive constant say c^2. The fixed lines perpendicular to each other are called as asymptotes and their point of intersection is known as centre of rectangular hyperbola.

3.4. EQUATION OF RECTANGULAR HYPERBOLA

The shape of the rectangular hyperbola is as shown in the fig. below :

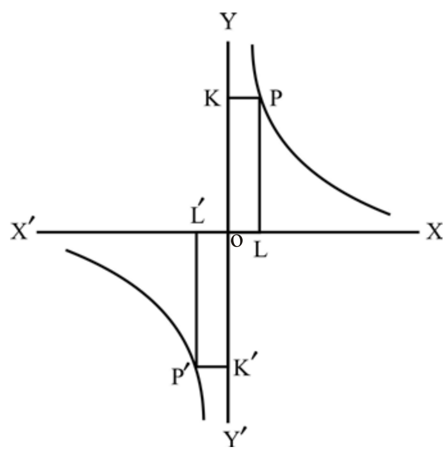


Fig. (*a*)

The asymptotes are horizontal and vertical lines respectively. The one portion of rectangular hyperbola lies in I Quadrant and similar second portion lies in III Quadrant.

Let P be any point on the curve in I Quad. Draw PL and PK $\perp$s on axes. As P moves to the right of the curve, the $\perp$ distance PL decreases and PK increases in such a way that the product PL × PK *i.e.* the area of the rectangle of OLPK remains constant (c^2). In Quadrant III, point P′ moves along the portion of the curve in such a way that product of $\perp$ distance from horizontal and vertical asymptotes (i.e. P′L′×P′K′) or area OL′P′K′ is equal to constant (c^2).

Let the horizontal and vertical asymptotes be respectively taken as x-axis and y-axis and centre of rectangular hyperbola be taken as origin. If the co-ordinates of point P be (x, y).

then $\because$ PL × PK = c^2, where c^2 is constant *i.e.* $xy = c^2$, which is the equation of rectangular hyperbola.

However, when the asympototes of the rectangular hyperbola are parallel to axes and centre be (a, b), then equation of rectangular hyperbola becomes

$$(x - a)(y - b) = c^2$$

3.5. APPLICATION OF RECTANGULAR HYPERBOLA IN ECONOMICS

A rectangular hyperbola plays an important role in economic theory. (e.g.) Average fixed cost which is defined as the ratio of total cost to output is represented by the rectangular hyperbola. In this case, the output axis and cost axis are the asymptotes and the product of the distance of any point on average fixed cost curve from the two axis is always equal to fixed cost and hence is a positive constant. This is shown in Fig. (b).

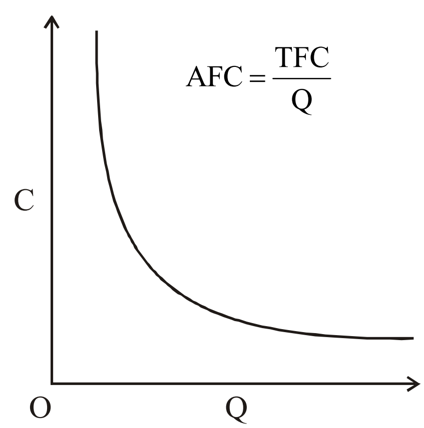


Fig. (b)

Second application is that the quantity theory of money says that a change in stock of money M implies an equiproportionate and direct change in the value of money $\frac{1}{p}$ on the opposite side, where p prepresents the price level.

$\therefore$ M = $C^2 \times P$ or $M \times \frac{1}{p} = C^2$ which is rectangular hyperbola where C^2 is constant.

Likewise, the demand curve may also be of the shape of rectangular hyperbola. Rectangular hyperbola demand curve shows that the total expenditure incurred by a consumer remains constant at all prices. Therefore, the elasticity of demand at any point on such a demand curve is constant and is equal to unity. This is one special type of demand curve.

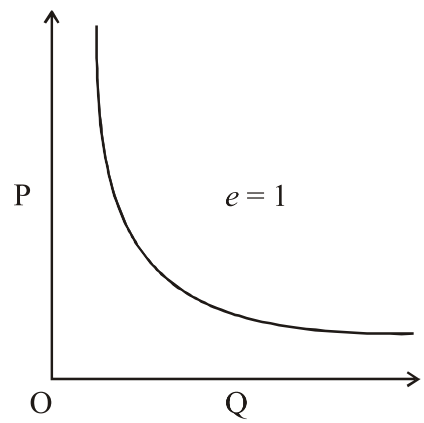


Fig. (c)

We can also express the average revenue curve in the shape of rectangular hyperbola. The marginal revenue at all levels of output is zero, and therefore marginal revenue curve Coincides with x-axis. So, the curve rectangular hyperbola has many applications in Economics.

EXAMPLES

Example 1 : If the total cost curve is $\pi = \alpha x \frac{x+\beta}{x+\gamma}$ where α, β, γ are Positve Constants, $(\beta > \gamma)$ Show that average cost curve is rectangular hyperbola.

Sol. Total Cost Curve

$$\pi = \alpha x \frac{x+\beta}{x+\gamma}$$

Average Cost Curve

$$\text{AVC} = \frac{\text{TVC}}{\text{Q}}$$

Here $\qquad$ $Q = x$

$$\text{AVC} = \frac{\pi}{x} = \frac{\alpha(x+\beta)}{(x+\gamma)}$$

$$y = \frac{\alpha(x+\beta)}{(x+\gamma)}$$

$$y(x+\gamma) = \alpha(x+\beta)$$

$$yx+y\gamma = \alpha x+\alpha\beta$$

$$xy+\gamma y = \alpha x+\alpha\beta$$

$$\alpha x-xy-\gamma y = -\ \alpha\beta$$

Let us add $\alpha\gamma$ to both sides so

$$\alpha x-y(x+\gamma)+\alpha\gamma = -\alpha\beta+\alpha\gamma$$

Or $$\alpha x+\alpha\gamma-y(x+\gamma) = -\alpha\beta+\alpha\gamma$$

$$\alpha(x+\gamma)-y(x+\gamma) = -\alpha\beta+\alpha\gamma$$

$$(\alpha-y)(x+\gamma) = -\alpha\beta+\alpha\gamma$$

which is of the form

$$(x-\text{a})\ (y-\text{b}) = \text{C}^2$$

So, it is a rectangular hyperbola.

Example 2 : Find the centre and the asymptotes of rectangular hyperbola $xy-2x-y-1 = 0$.

Sol. The given equation is :

$$xy-2x-y-1 = 0$$

Or $$xy-2x-y+2-2-1 = 0$$

Or $$x(y-2)-(y-2)-3 = 0$$

Or $$(x-1)(y-2) = 3$$

Compare it with

$$(x-a)(y-b) = c^2$$

Here $a = 1$, $b = 2$, $c^2 = 3$

$\therefore$ Centre of rectangular hyperbola is :

(a, b) *i.e.* $(1, 2)$ and asymptotes are $x - a = 0$ & $y - b = 0$

i.e. $x - 1 = 0$ & $y - 2 = 0$

3.6. LET US SUM UP

In this lesson, we discussed what a rectangular hyperbola is, how we can derive its equation and its application in Economics.

3.7. EXAMINATION ORIENTED QUESTIONS

Q.1. Define Rectangular Hyperbola. Write its equations and also its application in Economics.

Q.2. Show that the curve represented by $xy-3x-2y+1=0$ is a rectangular hyperbola.

Q.3. Find the equation of rectangular hyperbola whose asymptotes are parallel to the axis, centre is at $(2, 3)$ and $c^2=5$.

Q.4. Write down the equation of a rectangular hyperbola with asymptotes parallel to the axis, centre is $(-1, -2)$ and $c^2 = 10$.

3.8. SUGGESTED READINGS & REFERNCES

1. Chander Romesh (2007), Lectuer on Elementary Mathematics for Economists, New Academic Publishing Co, New Delhi.
2. G.S. Monga (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.
3. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas Publishing House, New Delhi.
4. Aggarwal, C.S. & R.C. Joshi : Mathematics for Students of Economics.

EQUATIONS

LESSON NO. 4 **UNIT-I**

STRUCTURE

4.1. OBJECTIVES

After going through this lesson this you, should be able to

1. Understand the meaning of different equations

2. Solve the Linear Equations

3. Solve Quadratic Equations

4. Know about the methods of Solving Simultaneous Equations

5. Know their applications in Economics.

4.2. INTRODUCTION

A mathematical equation is an expression containing at least one variable (= unkonwn value) and an "Equal Sign" (=) with a mathematical expression on each side of it. The equals sign says that both sides are of exactly the same value. In this lesson, we shall discuss about it different types of equations and their uses in Economics.

4.3. DEFININTION OF EQUATION

A polynomial of n^{th} degree in x equated to zero is termed as an equation of n^{th} degree in X. The unknown quantity in the equation is called variable. Thus, a + b = 0 ($a \neq 0$) is an equation of first degree in x. Equations of first degree in x are also called linear equations. Similarly, $ax^3 + bx^2 + cx + d = 0$ ($a \neq 0$) is an equation of 3rd degree in x or a cubic equation.

4.4. LINEAR EQUATIONS

4.4.1. Linear Equation in One Variable

$\frac{X + 3}{5} = 5$ is an equation of one unknown

(Whose value is required to be found out) and it is X Since only first power of the unknown appears in the given equation, the equation is called Linear Equation.

Equations in which maximum (or greatest) Power of the unknowns is only one are called Linear Equations

(i) $\frac{X - a}{X + a} + \frac{3b - X}{2b + X} = 0$ (a & b are constants)

(ii) $\frac{a}{bX} - a^2 = b^2 - \frac{b}{ax}$

The above two equations are linear equations. Equation $(X–3)^2 = 5$ is not a linear equation.

$\because \quad (X - 3)^2 = 5$

$\therefore X^2 - 6X + 9 = 5 \quad \text{or} \quad X^2 - 6X + 4 = 0$

The unknown X in the above equation possesses the power of 2, hence it is not a linear, but a quadratic equation or equation of second degree.

(i) $X^2 - 4 = 0$

(ii) $2X^2 + 5X = 8$

(iii) $3X^2 - 7X = 8$

are the examples of quadratic equations. In all these equations, unknown (X) possesses the power = 2

Similarly, we can define cubic equation as:

$$ax^3 + bx^2 + X + d = 0$$

and polynomial equation or multiple equation as

$$ax^n – bx^{n-1} + CX^{n-2} + ... + Px + q = 0$$

SOLUTION OF EQUATIONS

By solving a equation we mean the determination of the Particular value of the unknown which satisfies the given equation. It is called the root of the equations.

Suppose $5X + 8 = 0$

$\therefore \quad X = -\,{}^{8}\!/_{5}$

X can have only value $= -\frac{8}{5}$ to satisfy the equation $5X + 8 = 0$. No other value can satisfy this equation. So $-\frac{8}{5}$ is the root of this equation.

Note: In Linear equations we can obtain only one value of the unknown.

EXERCISES

Exercise 1 : Solve $5X + 6 - 7(X -2) = 5X$

Solution : The given equation can also be written as,

$$5X + 6 - 7X + 14 = 5X$$

$$5X - 7X - 5X = -14 - 6$$

$$-7X = -20 \qquad x = \frac{20}{7} \qquad \text{or} \qquad \frac{20}{7} = 2.857$$

Example 2 : Solve $\frac{a}{X-a} + \frac{b}{X-b} = \frac{a+b}{X-a-b}$

Solution: $\frac{a}{X-a} + \frac{b}{X-b} = \frac{a}{X-a-b} + \frac{b}{X-a-b}$

$$\therefore \quad \frac{a}{X-a} - \frac{a}{X-a-b} = \frac{b}{X-a-b} - \frac{b}{X-b}$$

$$\text{or} \quad \frac{a(X-a-b)-a(X-a)}{(X-a)\,(X-a-b)} = \frac{b(X-b)-b(X-a-b)}{(X-a-b)\,(X-b)}$$

$$\text{or} \quad \frac{ax-a^2-ab-ax+a^2}{(X-a)} = \frac{bx-b^2-bx+ab+b^2}{X-b}$$

$$= \frac{-ab}{X-a} = \frac{ab}{X-b}$$

$$= \frac{-1}{X-a} = \frac{1}{X-b}$$

$$= -1 \;(X - b) = 1 \;(X - a)$$

$$= -X + b = X - a$$

$$a + b = 2X$$

$$X = \frac{a+b}{2}$$

Check Your Progress–I

Solve the equations :

1. $\frac{3}{x-3}+\frac{5}{x-5}=\frac{8}{x-8}$

2. $\frac{1}{x+1}+\frac{1}{x+4}=\frac{1}{x+2}+\frac{1}{x+3}$

3. $\frac{x+5}{x+4}+\frac{x+9}{x+8}=\frac{x+8}{x+7}+\frac{x+6}{x+5}$

4.4.2. Linear Equations in case of Two Variables

The equation of the type of $a_1x + b_1y + C_1 = 0$ and $a_2x + b_2y + c_2 = 0$ are called simultaneous linear equations, if there is a particular pair of values of X and Y which solves both these equations.

The pair of values is called the solution of the two equations.

There are various methods of solving the simultaneous linear equations:

- **Method of Equalization of Coefficients**

Here we multiply the two equations by suitable numbers so that the coefficients of either of the two variables become equal in both the equations. Then by adding the two equation or by subtracting one from another, we get an equation in one variable only.

Let us explain the application of equalization of coefficients by taking the following example :

$$2X + y = 7 \quad \text{...(i)}$$

$$X + 3y = 6 \quad \text{...(ii)}$$

Multiply equation (ii) by 2, we get

$$2X + 6y = 12 \quad \text{...(iii)}$$

Subtracting eq (1) from equation (iii) we get

$$5y = 5$$

Or $\quad y = 1$

Substituting $y = 1$ in equation (i), we get

$$2X + 1 = 7$$

Or $\quad 2X = 6$

Or $\quad X = 3$

Therefore solution to the equation is (3, 1).

- **Method of Substitution**

We find the value of one of the variables in terms of the other from one equation and put this value in the other equation, in this way, we get an equation in one variable.

Let us now use second method

From equation (1) $\quad 2X = 7 - y$

Or $\quad X = \frac{7-y}{2}$

Let us substitute this value of X in terms of y in equation (ii), we get

$$\frac{7-y}{2} + 3y = 6$$

Or $\quad 7 - y + 6y = 12$

Or $\quad 7 + 5y = 12$

$5y = 12 - 7 = 5$

$y = 1$

$\therefore \quad X = \frac{7-1}{2} = 6/2 = 3$

Again solution to the equation is (3, 1).

- **Method of Comparison**

Here we find the value of one of the variables in terms of the other from one equation and put this value in the other equation. In this way, we get an equation in one variable.

Let us now use the method of comparison. From equation (i) we have already obtained $X = \frac{7-y}{2}$ from equation (ii) we have X = 6–3y. comparing X value from equation (i) with that of equation (ii), we have

$$\frac{7-y}{2} = 6 - 3y$$

Or $\quad 7 - y = 12 - 6y$

Or $\quad 6y - y = 12 - 7$

$5y = 5$

$y = 1$

substituting this value we get X = 3 so that again the solution to the equation is (3, 1).

Check Your Progress–II

Solve the following equations :

1. $4x + 3y = 7$

   $3x + 2y = 9$

2. $\frac{x}{2} + \frac{y}{5} = 11$

   $\frac{x}{3} + \frac{3y}{10} = 11.5$

3. $\frac{x-4}{3} = \frac{y-1}{4}, \; \frac{4x-5y}{7} = x-7$

4.4.3. Case of three variables

If we have a set of equations in three variables such as

$$x + y + z = 3 \quad \text{...(i)}$$

$$2x + y + 3z = 6 \quad \text{...(ii)}$$

$$4x - 2y - z = 1 \quad \text{...(iii)}$$

We solve these equations by eliminating one of the variables and reducing the system into two equations and two variables

Let us multiply eq (i) by 2 to get

$$2x + 2y + 2z = 6 \quad \text{...(iv)}$$

Sub equation (ii) from equation (iv) we get

$$Y - Z = 0$$

Now let us multiply equation (ii) by 2 we get

$$4X + 2y + 6z = 12 \quad \text{...(v)}$$

Subtract equation (iv) from equation (iii) we get

$$-4y - 7z = -11 \quad \text{...(vi)}$$

Now multiply equation (v) by 4 we obtain

$$4y - 4z = 0 \quad \text{...(vii)}$$

Add equation (vi) & equation (vii) we get

$$-11z = -11$$

Or $\quad Z = 11$

Substituting $Z = 1$ in equation (v) we get

$$Y = Z = 1$$

Now substitute value of y & z in equation (i) we get

$$X = 1$$

In a similar way, equation in 3 variables can be solved accordingly. But this procedure may not be conveniently used in equations is involving four variables as it will be very lengthy and cumbersome in such cases, we make use of matrix – algebra.

EXAMPLES

Example 1: $X + 2y + 3z = 1$

$3X + 4y + 5z = 3$

$2X + 2y + 4z = 2$

Solution: Given: $X + 2y + 3z = 1$ ...(1)

$3X + 4y = 5z = 3$ ...(2)

$2X + 2y + 4z = 2$ ...(3)

Multiply (1) by 3 and subtract from (2) we get

$2y + 4z = 0$ ...(4)

Multiply (2) by 2 and (3) by 3 and subtract, we get

$2y - 2z = 0$ ...(5)

Subtract (4) & (5) we get $6z = 0$, $Z = 0$

From (4), $y = 0$ from (1) $X = 1$ $\therefore$ $X = 1$, $y = 0$, $Z = 0$

Example 2: $\dfrac{XY}{X+y} = \dfrac{1}{9}, \dfrac{yz}{y+z} = \dfrac{1}{11}, \dfrac{ZX}{Z+X} = \dfrac{1}{10}$

Solution: The given equation can be written as

$$\frac{X+y}{XY} = 9, \ \frac{y+z}{yz} = 11, \ \frac{Z+X}{ZX} = 10$$

$\dfrac{1}{X} + \dfrac{1}{Y} = 9$ ...(1) , $\dfrac{1}{y} + \dfrac{1}{z} = 11$ ...(2) $\dfrac{1}{Z} + \dfrac{1}{X} = 10$ ...(3)

Adding, $2\left(\dfrac{1}{X} + \dfrac{1}{y} + \dfrac{1}{Z}\right) = 30$, $\dfrac{1}{X} + \dfrac{1}{y} + \dfrac{1}{Z} = 15$ ...(4)

From (1) & (4), $\dfrac{1}{Z} = 15 - 9 = 6$, $Z = \dfrac{1}{6}$

From (2) & (4), $\dfrac{1}{X} = 4$, $X = \dfrac{1}{4}$, Similarly $y = 1/5$

Example 3: $\frac{1}{X}+\frac{1}{y} = \frac{3}{2}, \qquad \frac{1}{y}+\frac{1}{Z}= \frac{5}{6}, \qquad \frac{1}{Z}+\frac{1}{X} = \frac{4}{3}$

Solution: $\frac{1}{X} + \frac{1}{y} = 3/2$ ...(1)

$$\frac{1}{y} + \frac{1}{2} = 5/6 \qquad ...(2)$$

$$\frac{1}{Z}+\frac{1}{X}=\frac{4}{3} \qquad ...(3)$$

Let us add all the equations, this will yield,

$$2\left(\frac{1}{X}+\frac{1}{y}+\frac{1}{Z}\right) = 3/2 + \frac{5}{6} + 4/3 = \frac{22}{6}$$

$$\frac{1}{X} + \frac{1}{y} + \frac{1}{Z} = \frac{22}{12} = \frac{11}{6} \qquad ...(4)$$

Now subtract equation (1) from equation (IV), we get

$$\frac{1}{Z} = \frac{11}{6} - 3/2 = \frac{11-9}{6} = \frac{2}{6} = \frac{1}{3}$$

Or $Z = 3$

Similarly by subtracting equation (II) & eq (III), from equation (IV) we get, X =1, y=2

So that complete solution is $X = 1,\ y = 2,\ Z = 3$

Check Your Progress–III

Solve the following equations :

1. $x-y-z=1$

   $y-z-x=1$

   $z-x-y=1$

2. $x + 2y + 3z = 1$

$2x + 2y + 4z = 2$

$3x + 4y + 3z = 3$

3. $x - y + z = 1$

$2x - 3y + z = 1$

$3x - y + 2z = 9$

4.4.4. Economic Applications of Linear Equations

We can apply linear equations in the determination of market equilibrium with laws of demand and supply. The Price of a commodity is determined by the intersection of demand and supply of the commodity in the market. When demand of a commodity is greater than its supply, price will surely rise and then more manufacturers will jump into the industry and supply will increase and with this the price will begin to fall.

Thus, Price will be fixed at that point where the demand per unit and supply per unit of the quantity are same. In this way, market equilibrium will correspond to the point for which the demand equation and supply equation will have the same values of the variables involved.

For example

Market Demand function is $M_D = 150–2P$ and the market supply function is $M_S = –50+2P$

In this situation market equilibrium will be at that point

where

$$
\begin{aligned}
M_D &= M_S \\
150–2P &= –50 + 2P \\
– 4P &= –200 \\
P &= 50
\end{aligned}
$$

Thus when the price is Rs 50, the market demand will be $M_D=150–2(50)=50$

And Market supply will also we equal to 50. ($M_S = –50+2\times50 = 50$)

EXAMPLES

Example 1 : Henry Schultz estimates the demand curve for sugar in the U.S. during the period 1916 – 1929 to be D = 135-8 P, where D stands for quantity demanded and P stands for price:

(a) Find the Price if the quantity demanded is 93.

(b) How much sugar would be demanded if it were a free good?

(c) Find the amount demanded if the price is Rs.7.

(d) What is the highest price any one will pay?

Solution: (a) The estimated demand function is $D = 135 – 8P$ ...(1)

when $D = 93$ then from (1) $93 = 135 – 8P$, $8P = 42$, $P = 5.25$

(b) If sugar were a free good, $P = 0$

In this case we have $D = 135 - 8 \times 0 = 135$

(c) If $P = 7$, $D = 135 - 8 \times 7 = 135 - 56 = 79$

(d) Putting $D = 0$ in the equation $D = 135 - 8P$, we have $0 = 135 - 8P$ or $8P = 135$, $P = 16.875$

Thus, we see that if price is 16.875, the amount demanded is zero. Therefore, it is clear that the price must be something less than 16.875, if any amount of sugar is to be sold.

Example 2: The supply curve of a commodity is given to be $S = ap - b$

(i) Find the amount supplied if $P = 3 b/a$

(ii) What will be the price if the amount supplied is $7a - 3b$?

(iii) What is the lowest price at which the commodity will be supplied?

Solution: The given supply curve is $S = ap - b$

(i) If $P = 3 b/a$,

$$S = a\left(\frac{3b}{a}\right) - b = 2b$$

$$S = 2b$$

(ii)
$$S = a\ P - b \qquad \ldots(1)$$

$$\text{if } S = 7a - 3b$$

$$7a - 3b = a\ P - b$$

$$7a - 3b + b = aP$$

$$7a - 2b = aP$$

Or
$$P = \frac{7a}{a} \ \frac{-2b}{a} = 7 - 2 b/a$$

$$P = 7 - 2 b/a$$

(iii) Putting $S = 0$ in equation (1) we have $0 = aP - b$, $P = {}^{b}\!/_{a}$ it is evident that when $P = {}^{b}\!/_{a}$ nothing is supplied.

Therefore, price must be some what greater than ${}^{b}\!/_{a}$, if any unit of the commodity is to be supplied.

Examples 3: The supply curve for a commodity is given as

$$S = 1.1P - 0.05$$

(i)Find the amount supplied if price is 5.2

(ii)Find the price if the supply is 4.9

(iii)Determine the lowest price which will cause any supply of the commodity.

$$S = 1.1\ P - 0.05 \qquad ....(1)$$

Solution: (i) If $P = 5.2$ then $S = 1.1\ (5.2) - 0.05 = 5.72 - 0.05 = 5.67$

(ii) If $S = 4.9$ then $4.9 = 1.1\ P - 0.05$

$$4.9 + 0.05 = 1.1P$$

$$4.95 = 1.1P$$

Or $P = 4.5$

(iii) Putting $S = 0$ in (i), we have

$$0 = 1.1\ P - 0.05$$

$$0.05 = 1.1P$$

Or $P = 0.04 = {}^{1}\!/_{22}$

Therefore price must be some what greater than ${}^{1}\!/_{22}$ if any unit of commodity is to be supplied.

Example 4 : A market demand curve is $D = 120 - 5\ P$. Find the price if the quantity demanded is 20 units. Also find the quantity demanded if the price is Rs.18. What would be quantity demanded if it were a free good?

Solution: The demand curve is $D = 120 - 5\ P \qquad ...(1)$

(i)When $D = 20$ then from (1), $20 = 120 - 5P$, $5P = 100$

$$P = 20$$

(ii)When $P = 18$ then from (i), $D = 120 - 5 \times 18 = 30$

(iii)If it were a free good, then $P = 0$

$\therefore$ from (1), $D = 120$

Example 5: Given a demand curve $q_d = 12 - P/3$ where q_d stands for quantity demanded of a commodity what is the highest price one would pay for the commodity?

Solution: The demand curve is $qd = 12 - P/3$

Putting $q_d = 0$ in equation (1) we have $12 - P/3 = 0$, $P = 36$

Thus, we see that if price is 36, the amount demanded is zero

Therefore, it is clear that the price must be somewhat less than 36 if any amount of commodity is to be sold.

Example 6: Given the demand curve $D = 20 - 2P$ and the supply curve $S = -4 + 3P$

(a) Find the equilibrium price,

(b) Find the equilibrium quantity exchanged.

Solution: To find equilibrium price, we must have demand equal to supply, that is, $D = S$

Or $\quad 20 - 2P = -4 + 3P$

$$-5P = -24$$

4.8 units

Substitute value of P in demand & supply functions to obtain equilibrium quantity

$$D = 20 - 2P,\ D = 20 - 2(4.8) \quad \text{i.e., } D = 10.4 \text{ units}$$

$$S = -4 + 3P,\ S = -4 + 3(4.8) \quad \text{i.e., } S = 10.4 \text{ units}$$

Example 7: (a) The demand curve for accomodity is given as D = 20 – 5 P and the supply curve is S = 6P – 21. Find the equilibrium price and quantity.

(b) Find the equilibrium price and quantity if demand, D = 5 – 5 P and supply, S = – 5 + 5P

(c) The demand & supply functions be D = 234.115 – 0.086P and S = 1.832P – 37.465 respectively. Find the equilibrium Price & quantity exchanged.

(d) Find equilibrium price and quantity if demand is D = 80 – 20P and supply S = 24P – 84.

Solution: (a) The demand curve is $D = 20 - 5P$ ...(1)

and supply curve is $S = 6P - 21$ ...(2)

Or $20 - 5P = 6P - 21$, Or $11P = 41$, $P = {}^{41}\!/_{11}$

Substitute the value of P in equation (1) we get, $D = 20 - 5 \times \frac{41}{11} = \frac{15}{11}$

Equilibrium Price = ${}^{41}\!/_{11}$, Equilibrium quantity = ${}^{15}\!/_{11}$

(b) The demand curve is, $D = 5 - 5P$ ...(1)

Supply curve is , $S = -5 + 5P$ ...(2)

For equilibrium D = S

$$5 - 5P = -5 + 5P$$

$$5 - 5P = -5 + 5P$$

$$5 + 5 = 10P$$

$$P = \frac{10}{10} = 1$$

Substitute the value of P in equation (1)

$$D = 5 - 5$$

$$D = 0$$

(c) The demand curve is, $D = 234.115 - 0.086P$ ...(1)

Supply curve is, $S = 1.832P - 37.465$ ...(2)

For equilibrium, $D = S$

$$234.115 - 0.086P = 1.832P + 37.465$$

$$234.115 - 37.465 = 1.832P + 0.086P$$

$$196.65 = 1.918P$$

$$P = \frac{196.65}{1.918}$$

$$P = 102.52$$

Substitute in eq. (1)

$$D = 234.115 - 0.086\ (102.52)$$

$$D = 234.115 - 8.816$$

$$D = 225.299$$

(d) The demand curve is, $D = 80 - 20\ P$

supply curve is, $S = 24P - 84$

For equilibrium, $D = S$

$$80 - 20P = 24P - 84$$

$$80 + 84 = 24\ P + 20\ P$$

$$164 = 44P$$

$$= P = {}^{164}\!/_{44} = 3.72$$

$$D = 80 - 20P$$

Or $D = 80 - 20 \times 3.72$

$$D = 80 - 74.4 = 5.6$$

Example 8 : (a) Let the demand curve $q_d = 25 - 2P$ and the supply curve be $q_s = -2 + P$. Find the equilibrium price, amount sold, total amount of tax and the actual price per unit realized by the seller after imposition of a tax of Rs. 0.75

per unit.

(b) The demand & supply equations of a commodity are given by $q = \frac{5}{2} - \frac{1}{2}P$ & q = 2P – 3. What will be the equilibrium price & quantity? What will be the new equilibrium price if a tax of Rs. $6/5$ per unit is imposed on the commodity ?

(c) The demand for a commodity is given by 20–4P and supply curve is S = 10P – 8. What will be the equilibrium Price and quantity if subsidy of Rs.2 per unit is granted to the seller ?

Solution: (a) The demand curve is $q_d = 25 - 2P$ ...(i)

and supply curve is $q_s = -2 + P$ ...(ii)

When a tax of Rs. 0.75 *i.e.* Rs. $3/4$ per unit is imposed, then new supply curve is

$$Q_s = -2 + [P - (3/4)] = -2 + P - 3/4 = -11/4 + P \quad \text{...(iii)}$$

For equilibrium putting demand equal to new supply *i.e.* $q_d = q_s$

i.e. $$25 - 2P = -11/4 + P$$

or $$\frac{100 + 11}{4} = 3P$$

$$\frac{111}{4} = 3P = \frac{111}{12} \quad \text{or} \quad P = 9.25$$

Substituting P = 9.25 in (1) we get

$$q_d = 25 - 2 \times 9.25 = 6.50$$

So after tax imposition, equilibrium price = 9.25 and amount sold = 6.50

Total amount of tax = units of quantity exchanged × tax per unit

$$= 6.50 \times 0.75 = 4.875$$

Price actually realized by the seller = Price paid by buyer – tax per unit

$$= 9.25 - 0.75 = \text{Rs. } 8.50$$

(b) The demand function is $q = \frac{5}{2} - P/2$ ...(1)

and the supply function is $q = 2P - 3$ ...(2)

(i) The equilibrium price of a commodity is determined where the demand is equal to its supply, solving (1) and (2) simultaneously, we have

$$\frac{5}{2} - \frac{1}{2}P = 2P - 3$$

or $$2P + \frac{1}{2}P = \frac{5}{2} + 3$$

or $$\frac{5}{2}P = \frac{11}{2} \quad \text{or} \quad P = \frac{11}{2} \times \frac{2}{5} \quad \text{or} \quad P = \frac{11}{5}$$

Substituting $P = \frac{11}{5}$ in equation (1) we get

$$q = \frac{5}{2} - \frac{1}{2}\left(\frac{11}{5}\right) = 7/5$$

$\therefore$ The equilibrium Price is Rs. $11/5$ & equilibrium quantity is $7/5$

(ii) When a tax of Rs. $6/5$ per unit is imposed on the commodity, then the new supply equation becomes $q = 2[P - (6/5)] - 3 = 2P - 27/5$ and new demand equation is

$$q = \frac{5}{2} - P/2 \qquad \text{...(4)}$$

For equilibrium demand = Supply

So solving (3) & (4) simultaneously, we get

$$2P - 27/5 = \frac{5}{2} - \frac{P}{2}$$

$$2P + \frac{P}{2} = \frac{5}{2} + \frac{27}{5}$$

$$\frac{4P + P}{2} = \frac{25 + 27 \times 2}{10}$$

$$\frac{5P}{2} = \frac{25 + 54}{10}$$

$$\frac{5P}{2} = \frac{79}{10}$$

$$\frac{5P}{2} = \frac{79}{10}$$

$$P = \frac{79 \times \cancel{2}}{1\cancel{0} \times 5}$$

$$P = \frac{79}{25}$$

Putting $P = \frac{79}{25}$ in (4) we have $q = \frac{5}{2} \; \frac{-1}{2} \left(\frac{79}{25}\right) = \frac{23}{25}$

After tax imposition, equilibrium Price = Rs. $\frac{79}{25}$, equilibrium quantity = $\frac{23}{25}$.

(*c*) Given $D = 20 - 4P$

$S = 10P - 8$

In case of subsidy, the supply equation would change as :

$S = 10(P + 2) - 8$

Putting $D = S$

We get $20 - 4P = 10\,(P + 2) - 8$

$20 - 4P = 10P + 20 - 8$

$$14P = 8 \quad \Rightarrow \quad P = \frac{8}{14} = \frac{4}{7} = \text{Rs. } 0.57$$

Now we need to calculate the quantity demanded.

$$D = 20 - 4P$$

Or $$D = 20 - 4(0.57)$$

$$D = 17.71$$

Example 9: The marginal revenue & marginal cost curves of a monopolistic firm are given by :

$$MR = \beta - 2\alpha q \ \& \ MC = 2aq + b$$

Find the Profit maximizing level of out put .

Solution: The marginal revenue curve is MR = $\beta - 2\alpha q$ & marginal cost curve is MC = 2aq+b

For maximum profit MR = MC, $\beta - 2\alpha q = 2aq + b$

Or $$2aq + 2\alpha q = \beta - b$$

Or $$q(2a + 2\alpha) = \beta - b$$

$$q = \beta - b \big/ 2a + 2\alpha$$

4.5. QUADRATIC EQUATIONS

4.5.1. Solution of Quadratic Equations

As discussed earlier, in quadratic equations the unknown is of second degree or Power. The general form of such equations is given by $ax^2 + bx + c = 0$ (a, b, c are constant quantities).

There are two methods for solving such equations:

(I) By Factorization:

Example: Solve : $X^2 - 16X + 48 = 0$

We find the factors of the expression on L.H.S. of equality

$$\therefore \ X^2 - 16X + 48 = X^2 - 12X - 4X + 48$$

$$= X\ (X-12) - 4\ (X-12)$$

= (X–4) (X–12) are the factors

∴ the given equation is (X–4) (X–12) = 0

Since the product of two factors is zero,

∴ either (X–12) = 0 or (X– 4) = 0

If X – 12 = 0 then X = 12

If X – 4 = 0 then X = 4

Thus we get two values X = 12 & X = 4, both would satisfy the given equation $X^2 - 16X + 48 = 0$

Check: (1) Substitute X = 12 in the equation

$(12)^2 - 16(12) + 48 = 144 - 192 + 48 = 0 =$ R.H.S.

(2) Substitute X = 4 in the equation

$(4)^2 = 16(4) + 48 = 16 - 64 + 48 = 0 =$ R.H.S

The solution of the given equation therefore, is X = 12 and X = 4

Note: In case of quadratic equations, we get two values of unknown

(II) By use of quadratic formula:

The solution of the general form of quadratic equation $ax^2 + bx + C = 0$ is given by the formula

$$X = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

How do we get this solution?

$ax^2 + bx + C = 0$, Dividing through out by a $X^2 + \frac{b}{a}X + \frac{C}{a} = 0$

or $$X^2 + 2\frac{b}{2a}X + \frac{C}{a} = 0$$

or $$X^2 + 2\frac{b}{2a}X + \left(\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 + \frac{C}{a} = 0$$

$$\left(X+\frac{b}{2a}\right)^2 = \frac{b^2}{4a^2} - \frac{C}{a}$$

$$\left(X+\frac{b}{2a}\right)^2 = \frac{b^2-4ac}{4a^2}$$

Or $$X+\frac{b}{2a} = \pm\sqrt{\frac{b^2-4ac}{4a^2}} = \pm\frac{\sqrt{b^2-4ac}}{2a}$$

i.e $$X+\frac{b}{2a} = \pm\frac{\sqrt{b^2-4ac}}{2a}$$

or $$X=-\frac{b}{2a} = \pm\frac{\sqrt{b^2-4ac}}{2a} = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

4.5.2. Nature of the Roots of Quadratic Equations

(i) If $b^2 - 4ac > 0$, then the roots are real and distinct

(ii) If $b^2 - 4ac = 0$, the roots in that case are real and equal.

(iii) If $b^2 - 4ac < 0$, then the roots are imaginary.

Remember, a is the coefficient of X^2

b is the coefficient of X

and c is the constant quantity in the equation

$ax^2 + bx + c = 0$

Example 1 : Solve the equation:

$$3X^2 - 5X + 2 = 0$$

Comparing it with general quadratic form

$$a = 3,\ b = -5 \text{ and } c = 2$$

$$X=-\frac{(-5)\pm\sqrt{25-24}}{6} = \frac{5\pm1}{6}$$

$$X = \frac{5+1}{6} = \frac{6}{6} = 1, \quad X = \frac{5-1}{6} = 4/6 = 2/3$$

Example 2 : $X^2 - 5X + 6 = 0$

Again : $a = 1,\ b = -5,\ c = 6$

is $$X = \frac{5 \pm \sqrt{25-24}}{2} = \frac{5 \pm 1}{2}$$

$$X = \frac{5+1}{2},\ X = \frac{5-1}{2}$$

$$X = \frac{6}{2}, \qquad X = \frac{4}{2}$$

$$X = 3, \quad X = 2$$

TYPICAL EXAMPLES

Example 1 : $\frac{75}{y^2} + y^2 - 20 = 0$

or $75 + y^4 - 20y^2 = 0$

or $y^4 - 20y^2 + 75 = 0$

Let us put $y^2 = x$ then, we get

$x^2 - 20x + 75 = 0$

is $x = \frac{20 \pm \sqrt{400-300}}{2} = \frac{20 \pm 10}{2} = 15, 5$

Hence $y^2 = 15$ and also $y^2 = 5$

$\therefore \qquad y = \pm\sqrt{15}, \quad y = \pm\sqrt{5}$

Note: There are four roots to the equation because it is a bi-quadratic equation or an equation of order 4.

(ii) $$\sqrt{\frac{X}{1-X}} + \sqrt{\frac{1-X}{X}} = 13/6$$

Let again $\sqrt{\frac{x}{1-x}} = y$ then the equation will be $y + \frac{1}{y} = \frac{13}{6}$

or $\frac{y^2+1}{y} = \frac{13}{6}$ or $6y^2 + 6 = 13y$

or $6y^2 - 13y + 6 = 0$

i.e. $y = \frac{13 \pm \sqrt{169-144}}{12} = \frac{13 \pm \sqrt{25}}{12} = \frac{13 \pm 5}{12}$

$$y = \frac{13+5}{12}, \; y = \frac{13-5}{12}$$

$$y = \frac{18}{12} \qquad y = \frac{8}{12}$$

$$y = 3/2 \, , \; y = 2/3$$

Now $y = 3/2 = \sqrt{\frac{X}{1-X}} = 3/2$

Or $\frac{X}{1-X} = \frac{9}{4}$

Or $4X = 9 - 9X$

$4X = 9 - 9X$

$4X + 9X = 9$

Or $13X = 9$ or $X = 9/13$

Similarly $y = 2/3 = \sqrt{\frac{X}{1-X}} = 2/3$

Or $\frac{X}{1-X} = 4/9$

Or $\quad 9X = 4 - 4X$

Or $\quad 9X + 4X = 4$

$13X = 4$

$X = 4/13$

Example 2: $(X + 2)(X + 3)(X + 5)(X + 6) = 504$

or $\quad (X + 2)(X + 6)\ (X + 3)\ (X + 5) = 504$

or $\quad (X^2 + 8X + 12)\ (X^2 + 8X + 15) = 504$

Let $\quad X^2 + 8X = y$

Then $\quad (y + 12)(y + 15) = 504$

Or $\quad y^2 + 27y + 180 = 504$

$y^2 + 27y - 324 = 0$

Or $\quad y = \dfrac{-27 \pm \sqrt{729+1296}}{2} = \dfrac{-27 \pm \sqrt{2025}}{2} = \dfrac{-27 \pm 45}{2}$

$= \dfrac{-27 \pm 45}{2} = \dfrac{-27+45}{2}, \dfrac{-27-45}{2}$

$= 9, -36$

Now $y = 9$ implies that

$X^2 + 8X = 9 \qquad$ or $\qquad X^2 + 8X - 9 = 0$

Or $\quad X = \dfrac{-8 \pm \sqrt{64+36}}{2} = \dfrac{-8 \pm 10}{2}$

$= - \dfrac{-8+10}{2} = 2/2 = 1, \quad X = \dfrac{-8-10}{2} = \dfrac{-18}{2} = -9$

Similarly $y - 36$ will imply that

$X^2 + 8X = 36$

or $\quad X^2 + 8X + 36 = 0$

Or $$X = \frac{-8 \pm \sqrt{64 - 144}}{2}$$

Which yields imaginary roots.

Example 3: Solve the equation : $y^4 - 5y^2 + 6 = 0$

Solution: This equation in Power can be converted to quadratic equation easily. It has only even numbered Powers. Hence, suppose, $y^2 = X$ then the equation becomes:

$$X^2 - 5X + 6 = 0$$

$$X^2 - 3X - 2X + 6 = 0$$

$$X (X - 3) - 2(X - 3) = 0$$

$$(X - 3) \quad (X - 2) = 0, \; X = 3, \text{ or } 2$$

However we assume $y^2 = X$ hence $y^2 = 3$ or 2

This $y = \pm \sqrt{3}$ or $\pm \sqrt{2}$

Example 4: Solve: $\sqrt{X^2 - 2X + 3} = 2X + 3$

Solution: Here, the Portion under the root is quadratic but that outside is not. We square both sides

$$(X^2 - 2X + 3 \;) = (2X + 3)^2$$

$$(X^2 - 2X + 3) = 4X^2 + 6X + 9$$

$$-3X^2 - 8X - 6 = 0$$

$$3X^2 + 8X + 6 = 0$$

$$X = -\frac{8 \pm \sqrt{64 - 72}}{6}$$

$$X = \frac{-8 \pm \sqrt{-8}}{6}$$

$$X = \frac{-8 \pm i2\sqrt{2}}{6} = \frac{-4 \pm i\sqrt{2}}{3}$$

Hence, both the roots are imaginary.

Example 5: $\sqrt{\dfrac{X}{X+16}} + \sqrt{\dfrac{X+16}{X}} = \dfrac{25}{12}$

Solution: Let $\sqrt{\dfrac{X}{X+16}} = y$

Then the given equation becomes: $y + \dfrac{1}{y} = \dfrac{25}{12}$

$$12y^2 + 12 = 25y$$

$$12y^2 - 25y + 12 = 0$$

$$(3y - 4)(4y - 3) = 0$$

$\therefore \qquad y = 4/3 \quad \text{or} \quad 3/4$

On putting the value of y

$$\sqrt{\frac{X}{X+16}} = \frac{4}{3} \qquad \text{or} \qquad \sqrt{\frac{X}{X+16}} = 3/4$$

$$\frac{X}{X+16} = \frac{16}{9} \qquad \text{or} \qquad \frac{X}{X+16} = 9/16$$

$$9X = 16X + 256$$

$$16X = 9X + 144$$

Then $\qquad X = \dfrac{-256}{7} \qquad X = 144/7$

4.5.3. Economic Applications of Quadratic Equation

Let us now assume that demand and supply functions are of quadratic form.

Example 1: The demand curve for a commodity is $D = 10 - 5P^2$, the supply curve $S = -5 + 5P$

Find the equilibrium price & quantity exchanged in the market

Solution: We must have D = S to obtain equilibrium Price

$\therefore \quad 10 - 5P^2 = -5 + 5P$

Or $\quad -5P^2 - 5P + 15 = 0$

$P^2 + P - 3 = 0$

We solve this quadratic equation by formula

$$\therefore P = \frac{-1 \pm \sqrt{1+(3\times4)}}{2} = \frac{-1 \pm \sqrt{13}}{2}$$

i.e. either $\quad P = -\ \dfrac{-1+\sqrt{13}}{2}$ or $P = \dfrac{-1-\sqrt{13}}{2}$

Since P can not be negative quantity

$$\therefore \quad P = \frac{\sqrt{13-1}}{2}$$

Substitute this value in D or S curve

$$D = 10 - 5P^2 = 10 - 5\ \frac{\left(\sqrt{13}-1\right)^2}{2}$$

$$= \frac{10\sqrt{3} - 30}{4} = \frac{5}{2}[\sqrt{13}-3]$$

Example 2: Given $\quad 3D = 10.3P$

$2S = 8 + P^2$ find equilibrium price.

Solution: For equilibrium price D = S

$10 - 3P = 8 + P^2$

$3P^2 + 6P + 4 = 0$

$$P = \frac{-6 \pm \sqrt{36-(4\times3\times4)}}{6}$$

$$P = \frac{-6 \pm \sqrt{-12}}{6} = \frac{-6 \pm 2\sqrt{-3}}{6}$$

$$= \frac{-3 \pm \sqrt{-3}}{3} = \frac{-3 \pm i\sqrt{3}}{3}$$

Which is an imaginary quantity.

Price or quantity to be exchanged can never be imaginary quantities, hence given example appears to be indeterminate in price quantity.

Example 3: The demand for goods of an industry is given by the equation $pq = 100$, where p stands for price and q for quantity demanded. Supply is given by the equation $20 + 30p = q$. What is the equilibrium price and quantity demanded ?

Solution: The demand and supply equations, are given by

$$pq = 100 \quad ....(1) \qquad \text{and} \qquad 20 + 3p = q \qquad ....(2)$$

Put value of q from (2) in (1), we get

$$p(20 + 3p) = 100, \; 3p^2 + 20p = 100$$

Transposing, $3p^2 + 20p - 100 = 0$

Compare it with $ax^2 + bx + c = 0$. Here $a = 3$, $b = 20$, $c = -100$, 'x' = 'p'.

$$p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-20 \pm \sqrt{(20)^2 - 4 \times 3 \times (-100)}}{2 \times 3}$$

$$= \frac{-20 \pm 40}{6} = \frac{-20+40}{6}, \frac{-20-40}{6} = \frac{10}{3}, -10$$

But $p \neq -10$ [$\because$ price cannot be negative], $\therefore$ $p = 10/3$

Put in (1), $\frac{10}{3} \times q = 100, \; q = 30$

$\therefore$ Equilibrium price = 10/3, and Equilibrium quantity = 30.

Example 4: (i) Find S, given $p = 2$ and (ii) Find p, given S = 18 for the supply curve $S = 5p + 2p^2$.

Solution: (*i*) When $p = 2$, $S = 5 \times 2 + 2 \times 2^2 = 10 + 8 = 18$

(*ii*) When S = 18, then $18 = 5p + 2p^2$

$2p^2 + 5p - 18 = 0$, Here $a = 2$, $b = 5$, $c = -18$, 'x' = p

$$\therefore \qquad p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-5 \pm \sqrt{(5)^2 - 4 \times 2 \times (-18)}}{2 \times 2}$$

$$= \frac{-5 \pm \sqrt{169}}{4} = \frac{-5 \pm 13}{4} = \frac{8}{4}, \frac{-18}{4} = 2, -\frac{9}{2}$$

But $p \neq -\frac{9}{2}$ $\quad \therefore \; p = 2$ [$\because$ price cannot be negative]

Example 5: The marginal cost curve of a firm under perfect competition is given as $MC = q^2 - 8q - 1$.

If the market price is fixed at Rs. 19 per unit, find the equilibrium of output.

Solution: It is given that $MC = q^2 - 8q - 1$. Also price = 19.

We know that under perfect competition, a firm is in equilibrium when MC = price. [$\because$ MR = AR]

$\therefore \qquad q^2 - 8q - 1 = 19$

Or $\qquad q^2 - 8q - 20 = 0$

$$\therefore \qquad q = \frac{8 \pm \sqrt{(-8)^2 - 4(1)(-20)}}{2(1)} = \frac{8 \pm \sqrt{64 + 80}}{8}$$

$$= \frac{8 \pm \sqrt{144}}{2} = \frac{8 \pm 12}{2} = 10, -2$$

But $q \neq -2$ [$\because$ quantity cannot be –ve], $\therefore \; q = 10$

$\therefore$ equilibrium level of output = 10.

Check Your Progress–IV

Solve the following :

1. Given the cost function $C = 12x + 3x^2$. For what value of x, is the cost equal to Rs.139 ?

---

2. Given the demand law, $p = 85 - 4q - q^2$. Find the amount demanded (q), when $p = 40$. If the price rises to 64, how much will the demand contract?

3. Given the following demand laws and suppy laws, find the equilibrium price and the quantity :

(a) Demand curve : $D = 50 - \frac{1}{2}p^2$, supply curve : $S = \frac{1}{2}p + \frac{5}{2}p^2$.

(b) Demand curve : $D = 6 - p$, supply curve : $S = 4p + p^2$

(c) Demand curve : $D = 19 - 3p - p^2$, supply curve : $S = 5p - 1$.

4. The marginal cost curve of a firm under perfect competition is $MC = q^2 - 5q - 9$, where q represents units of output. If given price is 5, find equilibrium level of output.

5. The demand and supply equations are given by $p - q = 1$ and $p^2 + q^2 = 25$; when p and q stand for price and quantity respectively, find the equilibrium price and quantity.

6. The demand and supply equations are given by $p^2 + q^2 = 20$ and $2p + q = 8$; where p is the price and q is the quantity. Determine the equilibrium price and quantity.

7. Find the equilibrium price and quantity when the demand and supply functions are $Q_d = \dfrac{8p}{p-2}$ and $Q_s = p^2$.

4.6. SIMULTANEOUS EQUATIONS (2 UNKNOWNS)

The demand and supply of a commodity depend frequently not only on the price of a particular commodity, but also upon the prices of other related commodities (Substitutes or Complementary).

In such case, we can also determine the equilibrium quantities & prices of all the commodities concerned as follows :

Symbolically : $DA = f(P_A, P_B)$

$SA = f(P_A, P_B)$

i.e. Demand and supply of commodity A depends on the Price of A and also on the prices of some other commodity B

Similarly $D_B = f(P_A, P_B)$

$S_B = f(P_A, P_B)$

The condition of equilibrium is that for all commodities, the supply is equal to the demand.

Let us take the following example

Example 1: Supply & Demand functions of Ghee & Dalda are given as under

$$D_g = f(P_g, P_d) = 30\ P_g + P_d$$
$$D_d = f(P_g, P_d) = 20 + P_g - 2\ P_d$$
$$S_g = f(P_g, P_d) = 3 + 2\ P_g$$
$$S_d = f(P_g, P_d) = 17 + P_d$$

Determine equilibrium Prices and quantities of Dalda & Ghee

Solution:

By setting $D_d = S_d$ we get a linear equation

$$20 + P_g - 2P_d = 17 + P_d$$

$$3P_d - P_g = 3 \qquad ...(1)$$

By setting $D_g = S_g$ we get another equation

$$30 - P_g + P_d = 3 + 2P_g$$

$$3P_g - P_d = 27 \qquad ...(2)$$

Solve equations (1) and (2) we get $P_d = 4.5$ and $P_g = 10.5$

Substituting these values in original (given) functions,

$$D_g = 30 - P_g + P_d = 30 - 10.5 + 4.5 = 24$$

$$S_g = 3 + 2P_g = 3 + 2\ (10.5) = 24$$

$$D_g = 20 + P_g - 2P_d = 10 + 10.5 - 2\ (4.5) = 21.5$$

$$S_d = 17 + P_d = 17 + 4.5 = 21.5$$

and Equilibrium quantities are $D_1 = 6$, $D_2 = 2$

Simultaneous Equations (3 Unknowns)

Example 1: A unit of commodity A is produced by combining 1 unit of land, 2 units of labour and 5 units of capital. One unit of commodity B is produced by combining 2 units of land, 3 units of labour & one unit of capital. Similarly commodity C is produced by combining 3, 1 & 2 units of land, labour capital respectively. If the prices of commodities are given to be : $P_A = 27$, $P_B = 16$ and $P_C = 19$, find the wage (w) Rent (R) and Interest (I).

Solution: Under free competition we have free movement of the factors of production. There are a great number of independent firms. No firm can make profits and the compensation to all the factors of production is same in all the productions.

Under these very simplified and artificial conditions, our examples reduces to a system of linear equations :

We find the average cost (AC) of each commodity & equate it with its price (Since there should be no profits)

AC of commodity A = $1R + 2W + 5I$

AC of commodity B = $2R + 3W + 1I$

AC of commodity C = $3R + 1W + 2I$

$P_A = 27$, $P_B = 16$, & $P_C = 19$

We have 3 equations

$$R + 2W + 5I = 27$$
$$2R + 3W + I = 16$$
$$3R + W + 2I = 19$$

And

Solving these equations, we get R = 3, we = 2

& I = 4

Example 2: Demand & supply functions of two Commodities A & B are given to be

$$D_A = 10 - 2P_A = + P_B, \; D_B = 20 + P_A - 5P_B$$
$$S_A = 4P_A, \; S_B = -1 + 6P_B$$

Find the equilibrium prices and quantities.

Solution: By setting $D_A = S_A$

$$10 - 2P_A + P_B = 4P_A$$
$$6P_A - P_B = 10$$

By setting $D_B = S_B$

$$20 + P_A - 5P_B = -1 + 6P_B$$
$$P_A - 11P_B = -21$$

Solving equation (I) & (II) we get

$$P_B = \frac{136}{65} \;\&\; P_A = \frac{131}{65}$$

Equilibrium quantities can be obtained by substituting these values in D & S functions.

Example 3: Demand and supply curves for two commodities A & B are given as

$$D_1 = 12 - P_1 - 4P_2, \quad S_1 = -5 + 5P_1 + P_2$$
$$D_2 = 5 - P_1 - P_2, \quad S_2 = -3 + 5P_2$$

Find the equilibrium prices & quantities.

Solution: For equilibrium $D_1 = S_1$ and $D_2 = S_2$

Or $12 - P_1 - 4P_2 = -5 + 5P_1 + P_2$

and $5 - P_1 - P_2 = -3 + 5P_2$

Or $6P_1 + 5P_2 = 17$ ...(1)

and $P_1 + 6P_2 = 8$ ...(2)

Multiply (1) by 1 & (2) by 6 & subtract;

$$-31P_2 = -31$$

$$P_2 = 1$$

Substituting $P_2 = 1$ in equation (1) we have

$$6P_1 + 5 = 17,$$

$$P_1 = 2$$

when $P_1 = 2,$

$P_2 = 1$

then $D_1 = 12 - 2 - 4 = 6, \quad D_2 = 5 - 2 - 1 = 21$

$\therefore$ Equilibrium prices are $P_1 = 2,\ P_2 = 1$

EXERCISE

Q.1. Solve (1) $\frac{5X}{3} - \frac{X-2}{4}$

Ans. $X = 1$

Q.2. $(2X + 3)(3X + 5) = (6X + 1)(X+1) + 74$

Ans. $X = 5$

Q.3. Solve $\frac{X}{4} + \frac{Y}{5} + 1 = \frac{X}{5} + {}^{y}\!/_{4} = 23$

Ans. $X = 40, \quad y = 60$

Q.4. $\frac{4}{X} + \frac{10}{y} = 2, \ \frac{3}{X} + \frac{2}{y} = \frac{19}{20}$

Ans. $X = 4, \ y = 10$

Q.5. If the demand & supply curves of two commodities A & B are given below, find the equilibrium prices & quanties

$D_1 = 10 - P_1 - 2\ P_2, \quad S_1 = -3 + P_1\ P_2$

$D_2 = 6 - P_1 + P_2\ , \qquad S_2 = -2 + P_2$

Ans. $\therefore$ Equilibrium Prices are $P_1 = 2,\ P_2 = 3$

& Equilibrium quantities are $D_1 = 2,\ D_2 = 1$

Q.6. Solve (1) $2X^2 - 4X + 3 = 0$

Ans. $X = \dfrac{2 \pm \sqrt{2i}}{2}$

Q.7. Solve:

(I) $8X - 2 = {}^{3}\!/_{X}$

Ans. $X = \dfrac{3}{4}, -{}^{1}\!/_{2}$

(II) $\sqrt{X^2 - 5X + 31} = 2X + 1$

Hint: squaring $X^2 - 5X + 31 = 4X^2 + 4X + 1$

$X^2 + 3X - 10 = 0X = -5$, Or 2

Q.8. Solve $y^4 - 8y^2 - 9 = 0$

Hint: Put $y^2 = X$ $\therefore$ we have

$X^2 - 8X - 9 = 0$

Or $(X-9)\ (X+ 1) = 0, \quad X = 9,\ X = -1$

$Y^2\ -9,\ y^2 = -1$

$\therefore \qquad y = \pm\ 3,\ y = \pm\ i$

Q.9. Solve $\sqrt{\dfrac{X+1}{X-1}} + 8\sqrt{\dfrac{X-1}{X+1}} = 6$

Hint: Put $\sqrt{\dfrac{X+1}{X-1}} = y$ the given equation becomes

$$y + \frac{8}{y} = 6, \quad y^2 + 8 = 6y, \quad y^2 - 6y + 8 = 0$$

$$y = \frac{6 \pm \sqrt{36-4.18}}{2.1} = \frac{6 \pm \sqrt{4}}{2} = \frac{6 \pm 2}{2}, \ 4, 2$$

$y = 4$, but $y = \sqrt{\dfrac{X+1}{X-1}}$

$\therefore \sqrt{\dfrac{X+1}{X-1}} = 4$ squaring $\dfrac{X+1}{X-1} = \dfrac{16}{1}$

$16X - 16 = X + 1, \ 15X = 17$

$y = 2$ but $y = \sqrt{\dfrac{X+1}{X-1}}$

$\sqrt{\dfrac{X+1}{X-1}} = \dfrac{2}{1}$ squaring $\dfrac{X+1}{X-1} = 4$

$4X - 4 = X + 1, \ 3X = 5, \ X = 5/3$

$X = 17/15, \ 5/3$

Q.10. (i) Demand for goods of an industry is given by $2P + 3q = 8$

Where P is Price and q is quantity and the supply is given by the equation $P_2 + 2P_q + 3q^2 = 17$. What is the equilibrium Price & quantity?

(ii) The demand for goods of an industry is given by the equation $P_q = 100$, where P stands for price and q for quantity demanded. Supply is given by the equation $20 + 3P = q$ what is the equilibrium price & quantity demanded ?

Ans. (i) Equilibrium Price = 1, Equilibrium quantity = 2.

(ii) Equilibrium Price = $10/3$, Equilibrium quantity = 30.

4.7. LET US SUM UP

In this lesson, we discussed different types of equations and their applications in Economics.

4.8. EXAMINATION ORIENTED QUESTIONS

1. What is a Quadratic Equation?
2. What do you mean by an equation?

3. State the meaning of linear Equation.

4. State Quadratic formula and its derivation.

5. Explain the economic applications of the equations by giving an example.

6. $5x+6-7(x-2) = 5x$ Solve this

Sol.
$$
\begin{aligned}
5x+6-7x+14 &= 5x \\
5x-7x-5x &= -14 - 6 \\
-7x &= -20 \\
x &= \frac{20}{7}
\end{aligned}
$$

Q.7. Solve $\frac{3}{x-3}+\frac{5}{x-5} = \frac{8}{x-8}$

Sol.
$$
\begin{aligned}
\frac{3}{x-3}+\frac{5}{x-5} &= \frac{8}{x-8} \\
\frac{3x-15+5x-15}{(x-3)(x-5)} &= \frac{8}{x-8} \\
(8x - 30)(x - 8) &= 8(x^2 - 8x + 15) \\
8x^2 - 94x + 240 &= 8x^2 - 64x + 120 \\
-30x &= -120 \\
x &= 4
\end{aligned}
$$

Q.8.
$$
\begin{aligned}
\mathbf{3x + 2y} &\mathbf{= 9} \\
\mathbf{x + 3y} &\mathbf{= 10}
\end{aligned}
$$

Solve the equations.

Ans.
$$
\begin{aligned}
3x + 2y &= 9 \qquad \text{...(i)} \\
x + 3y &= 10 \qquad \text{...(ii)}
\end{aligned}
$$

Multiply equation (i) by 3 and equation (ii) by 2 and by substracting it, we get

$$9x + 6y = 27$$
$$2x + 6y = 20$$
$$- \quad - \quad -$$
$$7x = 7$$
$$x = 1$$

Put x = 1 in eqation (i), we get

$$3(1)+2y = 9$$
$$2y = 6$$
$$y = 3$$

Thus $x = 1$ and $y = 3$

Q.9. Find the roots of the equations

$$\boldsymbol{x^2 - 1 = 0}$$

Ans. $x = 1$, $x = -1$

Q.10. Find the roots of the equation

$$x^2 - 2x - 2 = 0$$

Ans. x = $1+\sqrt{3}$, $1-\sqrt{3}$

LONG ANSWER TYPE QUESTIONS

Q.1. Explain the economic applications of the equations by giving suitable examples.

Q.2. Find out the solution of $x^2 - 16x + 48 = 0$

Q.3. Solve the following equations

$$2x + 3y + 5z = -9$$
$$x + 10y + 6z = -13$$
$$-5x + y + 10z = 14$$

Ans. $x = -3$, $y = -1$ and $z = 0$

4.9. SUGGESTED READINGS & REFERENCES

1. Mehta & Madnani : Elementary Mathematics in Economics
2. Allen, R.G.D. : Mathematical Analysis for Econimists (Macmillan)
3. Chander Romesh (2007, Lecturer on Elmentary Mathematics for Econimists, New Delhi.
4. Aggarwal, C.S. & R.C. Joshi : Mathematics for Students of Economics.

Unit II

Functions and Differentiation

FUNCTIONS

LESSON NO. 5 **UNIT-II**

STRUCTURE

5.1. OBJECTIVES

After going through this lesson, you shall be able to understand :

• Meaning of Relation, Domain and Range of Relation

- Meaning of functions
- Different type of functions in Economics
- Meaning of Limits and Continuity.

5.2. INTRODUCTION

In mathematical operations, we come across two types of quantities– constants and variables. A quantity which does not change its value in mathematical operation is called a constant and a quantity which can assure different values is called a variable. When two variables x and y are connected in such a way that corresponding to each value of first variable, if there is a unique value of second variable y, then second variable y is said to be function of first variable x and is denoted by $y = f(x)$.The first variable x is called independent variable and second variable y is called dependent variable. In this lesson, we shall discuss about Cartesian Product, relation, range, functions and continuity.

5.3. CARTESIAN PRODUCT

Let A and B denote any two non empty sets. The Cartesian product of A and B, written as A×B, is the set of all ordered pairs (x, y) where $x \in A$ and $y \in B$

In symbols, $A \times B = \{(x,y)\ (x \in A\ \&\ y \in B)\}$

e.g. If $A = \{a, b, c, d\}$ & $B = \{1, 2, 3\}$ then

$A \times B = \{(a, 1)\ (a, 2)\ (a, 3), (b, 1), (b, 2), (b, 3), (c, 1)\ (c, 2)\ (c, 3), (d,1)\ (d, 2)\ (d, 3)\}$

Hence, $(a, 1) \in A \times B$ but $(1, a) \notin A \times B$.

5.4. RELATION

If A and B are any two sets, then a relation L is A×B in a subset of A×B

Examples. If A{1,2}, B{2,3}

then $A \times B = \{(x, y): x \in A, y \in B\} = \{(1, 2)\ (1, 3)\ (2, 2)\ \ (2, 3)\}$

$L = \{(x, y): x = y, (x, y) \in A \times B\} = \{(2, 2)\}$

M = {x, y): x is odd, y is even, (x, y) $\in$ A×B} = {(1,2)}

As L, M are subsets of A×B and hence represent relations in A×B

5.5. DOMAIN AND RANGE OF RELATION

Def. Given a relation L = {(x, y)} in A × B, the set of x's or first elements is called the domain of the relation and the set of y's or second elements is called the range of the relation.

Example: If A = {1, 2, 3} B = {2, 3, 4}

then L = {(1, 2,), (2, 3) (3, 4)} is a relation in A×B

In this case the domain L is the set D = {1, 2, 3} and range of L is the set R = {2, 3, 4}

5.6. DEFINITION OF FUNCTION

If A and B are two non-empty sets then a functions 'f' is defined as a subset of A×B such that no two distinct ordered pairs belonging to 'f' have the same first element.

Thus a function is a special kind of relation when one element of the domain is paired with one and only one element of the range, then this is called a function.

Example: If A = {1, 2, 3, 4, 5, 6}; B = {2, 4, 6} consider the following subsets of A×B.

f = {1, 2} (3, 4), (5, 6)}, L = {(1, 2), (1, 4), (1,6)}

The subset f is a function as it satisfies our definition. The domain of the function is Df = {1, 3, 5} and Range is Rf = {2, 4, 6}

But L is a relation in A×B, as in this case more than one element of range are paired with one element of domain.

5.6.1. Value of a Function

In $f(x)$, if we put $x = a$, we get $f(a)$ and $f(a)$ is called the value of the function at $x = a$. Thus, if $f(x) = 5x+2$, then $f(1) = 5(1)+2 = 7$ is he value of $f(x)$ at $x = 1$. In order to know the value of the function, given any value to x and then insert that value in place of x in $f(x)$. The following table illustrates this

Table – I

x	$f(x) = 5+4x$
0	5
1	9
5	25
10	45

Table – 2

x	$f(x)$ x^2-x+3
0	3
1	3
5	23
10	93

In the above two tables, we have illustrated two expressions.

In table 1, $f(x) = 5+4x$ and table 2, $f(x) = x^2–x+3$

We have explained how to calculate the value of $f(x)$ by inserting a given value. In general the value of $f(x)$ at $x = a$ is denoted by $f(a)$ and is determined by inserting 'a' in place of x in the expression for $f(x)$. Thus, in table 2, the value of $f(x)$ at $x = 5$ is $f(5)= 5^2–5+3 = 23$.

In certain cases, the function may not be defined at $x = a$ for example, Consider

$$f(x) = \frac{x-1}{x-2}$$

The determination of $f(x)$ at $x = 2$ leads to $f(2) = \frac{2-1}{2-2} = \frac{1}{0}$, which involves meaningless operation of division by zero.

Therefore, the function is not defined at $x = 2$, But function $f(x)$ is defined at other values of x.

Example: Find for what values of x the following functions are not defined.

(i) $f(x) = \dfrac{u-1}{x-2}$

(ii) $f(x) = \sqrt{25-x^2}$

(iii) $f(x) = \sqrt{x}$

Sol. (i) $f(x) = \dfrac{u-4}{x-2}$

Clearing for $x = 2$, $f(x)$ is not defined,

as at $x = 2$, $f(2) = \dfrac{1}{2-2} = \dfrac{0}{0}$ which is not defined

(ii) $f(x) = \sqrt{25-x^2}$

clearly for $x > 5$, $25-x^2<0$

$\therefore \quad f(x) = \sqrt{25-x^2}$ becomes imaginary.

$\therefore \quad f(x)$ is not defined for $x>5$

(iii) $f(x) = \sqrt{x}$

Clearly $f(x)$ is not defined for negative values of x

$\therefore$ $f(x)$ is not defined for $x < 0$

5.6.2. Types of Functions or Mapping

Constant Functions: A function whose range consists of only one element is called a constant function As an example

$$y = f(x) = 7$$

Which is alternatively expressible as $y = 7$ or $f(x) = 7$, whose value stays the same regardless of the value of x.

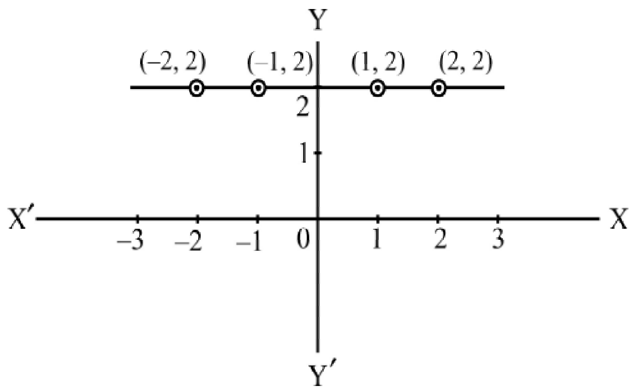
Example: Draw the graph of $f = \{(x, y); y = 2\}$

Solution: We have $f = \{(x, y); y=2\}$ .....(i)

equation (i) can be expressed as

$$f = \{[x, f(x)]: f(x) = 2\} \quad .....(ii)$$

Such a function is called Constant function because for any value of the independent variable x, the value of y is always to 2. The graph is drawn by joining the points like (-2, 2), (-1, 2), (1, 2) (2, 2) etc. as shown below.



Polynomial function: The Constant function is actually a "dengenerate" case of what are known as Polynomial functions. the 'word' Polynomial means "multiterm" and a Polymonial function has the general form

$$y = a_0 + a_1 x + a_2 x^2 + \ \ldots\ldots \ + a_n x^n$$

in which each term contains a co-efficient as well as a non-negative integral power of the variable x. Note that instead of the symbols a, b, c .... we have used the subscripted symbols a_1, a_2 ... ... a_n for the co-efficients. This is motivated by two considerations: (1) we can economize on symbols, since only the letter 'a' is "used up" in this way (2) subscripts helps to pinpoint the location of a particular co-efficient in the entire equation.

Depending on the value of the integer 'n', we have several subclasses of Polynomial functions.

Case of $n = 0$: $y = a_0$ [Constant functions]

Case of $n = 1$: $y = a_0 + a_1 x$ [linear function]

Case of $n = 2$: $y = a_0 + a_1 x + a_2 x^2$ [Quadratic function]

Case of $n = 3$: $y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ [cubic function]

and so on. The subscript indication of the powers of x are called exponents. The highest power involved, i.e. the value of n, is often called the degree of the polynomial function; a quadratic function for instance, is a second degree polynomial and a cubic function is a third degree polynomial.

One value function: When a function has only one value corresponding to each value of the independent variable the function is known as one-valued function. For example,

If $y = x^2$, y is a single valued function of x.

Multi-value or Many-valued function: When a function has several values corresponding to each value of the independent variables, it is known as multi-valued function. For example,

If $y^2 = x$, y is a two-valued function of x $(+\sqrt{x}$ *and* $-\sqrt{x})$

Linear functions: A function of the form $y = f(x)$, e.g. $y = mx$ +c; m and c being real constants (real numbers) is known as a linear function. The graph of linear function is always a straight line.

Cubic functions: A function of the form $y = f(x) = ax^3+bx^2+cx+d$ where a, b, c, d are real constants is called a cubic function.

Quadratic functions: A function of the form $y = f(x)\ ax^2+bx+c$ where a, b, and c are real constants [real numbers] is called a quadratic function. the graph of quadratic function is always a Parabala.

Even and odd functions: A function $y = f(x)$ is said to be an even function of x if $f(-x) = f(x)$ e.g. $f(x) = x^4+x^2$ is an even function of x because $f(-x) = (-x)^4 + (-x)^2 = x^4+x^2 - f(x)$.

A function $y = f(x)$ is said to be an odd function of x if $f(-x) = -f(x)$ *e.g.* $y = f(x) = x^3 - x$ is an odd function of x.

Rational functions: A rational function is defined as the ratio of two polynomials e.g. $y = \dfrac{f(x)}{g(x)}$ is a rational function where $f(x)$ and $g(x)$ are polynomial in x and $g(x) \neq 0$

For example $f(x) = \dfrac{x^3+2x+1}{x+5}$, $x+5 \neq 0$ is a rational function.

Non-Algebraic function: A function is said to be non-algebraic functions if the relation which involve the infinite terms and variables are not affected by the operations of additions, subtraction, multiplication, division, powers and roots. Exponential function, Logarithmic function, Trignometric function are called as non-algebraic or transcendental functions.

5.7. APPLICATIONS OF DIFFERENT FUNCTIONS IN ECONOMICS

A function is basically a relation between two or more than two variables i.e. dependence of one variable on one or more other variables. Thus, if the value of variable y depends on another variable x. We may write

$y = f(x)$ where f stands for function

A function is a technical term which is used to symbolise relationship between variables. Following are the different types of functions in Economics.

1. Demand function: This function expresses the relationship between the price of the commodity (Price is independent variable) and quantity of the commodity (quantity is dependent variable)

Demand fund $Q_x^d = f(P_x)$

the basic determinants of demand function are

$$Q_x^d = f(P_x, P_r, Y, T, W, E)$$

Q_x^d – Quantity demanded of a commodity x

P_x – Price of the good X

P_r – Price of the related good.

Y – Consumer's Income

T – Consumer's Taste and Preferences

W – Consumer's wealth

E – Consumer's expectations.

2. Supply function: It expresses the relationship between the Price of a good (independent variable) and quantity of the commodity supplied (dependent variable). It shows how much quantity of a good a seller offers at different prices. Hence S_x is represented as the quantity supplied of a good and P_x is the price of that good. Then supply function is

$$S_x = f(P_x)$$

The basic determinants of supply function are :

$$Q_x^s = f(G_f, P, I, T, P_r, E, G_p)$$

Here,Quantity supplied is Q_s and

G_f – Goal of the firm

P – Product Price

I – Input price

T –Technology

P_r – Price of the related good

E – Expectations of producers

G_p – Government policy

3. Utility function: Goods are demanded because they satisfy the wants of the people. Utility can be defined as the want satisfying power of a good. Utility is a subjective concept and resides in the minds of men. So, it varies from person to person *i.e.* different persons derive different utility from the different goods consumed. Thus, a utility function for a consumer consuming different goods may be represented as

$$U = f(x_1, x_2, ...x_n)$$

x_1, x_2, x_3 .... represent different goods.

4. Consumption Function: Also called propensity to consume, it shows the relationship between income and consumption. When there is an increase in income, the consumer spends a part of that income but not all of it, as it is also saved. So, with the increase in income, both consumption and savings will increase. Keynes believed that consumption expenditure is a function of current income i.e.

$$C = f(Y_t)$$

5. Production function: Production function is a functional relationship between physical input and physical output of a firm. Algebrically production function can be written as

$$Q = f(a, b,c, d)$$

Where Q – Quantity of output

a, b, c, d – quantitative factors

Q depends upon a, b, c, d factors respectively.

6. Cost Functions: Cost functions are derived functions from production function which describes the available efficient methods of production at anytime. Economic theory is divided between short run and long run cost function. If Q is the quantity produced by a firm at total cost C, we write cost functions as

$$C = f(Q)$$

which means that cost depends upon the quantity produced.

7. Revenue Functions: If R is the total revenue of a firm, X is the quantity demanded or sold and P is the price per unit of output, we write revenue function as the revenue earned as a function of price of the good and quantity of the goods sold.

$$R = f(P, S)$$

P – Price of the commodity

S – Total units sold.

8. Profit function

$$\pi = f(TR, TC)$$

π is the Profit which is the function of total revenue (TR) and total cost (TC).

9. Saving function: The relationship between the disposable income and saving describes the saving function. It can be written as

$$S = f(Y)$$

where S – saving, Y – Income

10. Investment function: This function shows the functional relation between investment and the rate of interest or income. So, the investment function.

$$I = f(i)$$

Where I – Investment

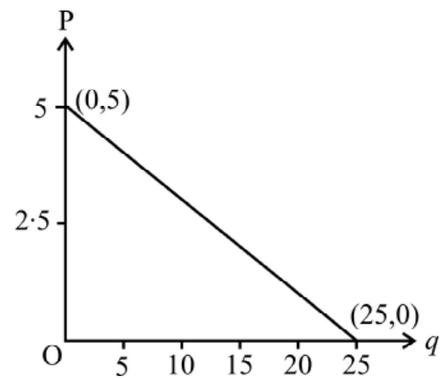
i – rate of interest

5.8. GRAPH OF LINEAR FUNCTION

The graph of a linear function can be drawn, by taking a simple example of linear equation.

Example: Draw the graph of a demand law given by $q = 25–5p$ where q is the demanded and p is the price.

Solution: Since the demand function is linear we need only two points corresponding to two ordered pairs of q and p which belong to the function q=25–5p or p =5. By plotting the points (25, 0) and (0, 5) and by joining them by a straight line, we get the required demand curve.



[**Hint :** For getting an explanation of how we plotted these points, revise lesson 2.]

5.9. GRAPHS OF QUADRATIC FUNCTIONS

The functions

$$q=\{[x,\ f(x)]:\ f(x)\ =\ ax^2+bx+c\}$$

is known as general quadratic function and its graph is called Parabola.

The following pairs (x, y) by using the rule $y = f(x) = ax^2+bx+c$ and construct a table. While constructing such a table we must take the following three pairs which become very helpful in drawing the parabola.

(a) Find the co-ordinates of the vertex, i.e the lowest or the highest point. This is done by completing the square on the R.H.S. of the equation $y = ax^2+bx+c$

(b) Find the points of intersection with x-axis by putting $y = 0$

(c) Find the points of intersection with y-axis by putting $x = 0$

The points so located are joined and Parabola is traced.

Example: Draw the graph of the function

$$f = \{(x, y): y = x^2–5x+6\}$$

Solution: We have $y = x^2–5x+6$

(1) By putting $y = 0$, we get

$$x^2+5x+6 = 0$$

$$(x–2)(x–3) = 0$$

(2) By putting $x = 0$, we get

$$y = (0)^2 – 5(0) + 6 = 6$$

Hence the curve cuts y-axis at (0, 6)

(3) Completing the square of R.H.S. of

$$y = x^2–5x+6$$

$$y = \left(x-\frac{5}{2}\right)^2 - \frac{1}{4}$$

$$y + \frac{1}{4} = \left(x-\frac{5}{2}\right)^2$$

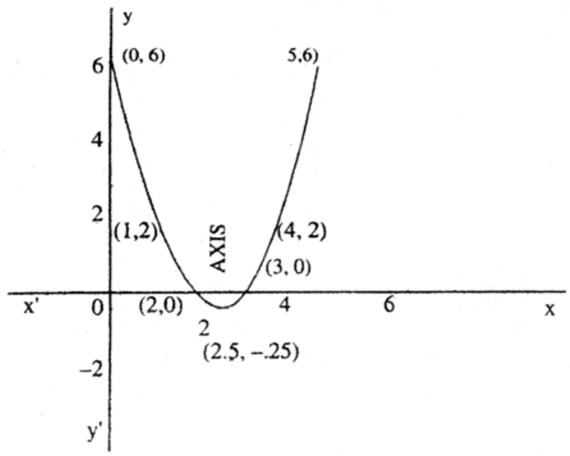
$\therefore$ The least value of $y = -\frac{1}{4}$ at $x = \frac{5}{2}$

(4) The axis of the Parabola is $x - \frac{5}{4} = 0$. The curve is symmerical about the line $x - \frac{5}{4} = 0$

(5) Besides the ordered pairs given by (1), (2) and (3) we obtaine few more and construct the following table:

x	–1	0	1	2	2.5	3	4	5	6
y	12	6	2	0	–0.25	0	2	6	12

By plotting these points in fig. and joining them by a smooth curve we get parabola.



5.10. GRAPH OF RECTANGULAR HYPERBOLA

The equation of rectangular hyperbola is given by $xy = C^2$ where C^2 is Constant. The technique of drawing a rectangular is explained with the help of following example

Example: Sketch the curve $xy = 32$ ...(1)

(1) Putting $x = 0$, we get

$$y = \frac{32}{0} = \infty \qquad ...(2)$$

(2) Implies that the curve meets y-axis at infinity

Putting $y = 0$, we get

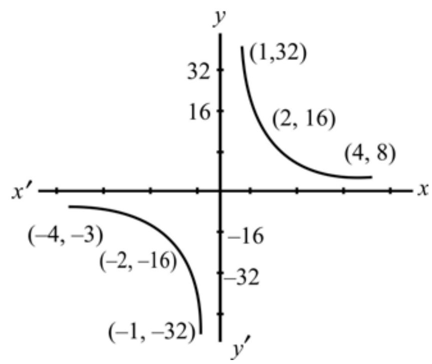
$$x = \frac{32}{0} = \infty \qquad ...(3)$$

(3) Implies that the curve meets x-axis at infinity.

Also the curve does not pass through origin because equation contains a constant term. Writing equation (1) as $y = \frac{32}{0}$ and giving different values, we can find corresponding values of y. Thus

x	–4	–2	–1	1	2	4
y	–8	–16	–32	32	16	3

The points are plotted in fig. and joined by a smooth curve.



Example 1: If $f(x) = x^2–5x+3$, then find out the following:

(i) $f(0)$(ii) $f(–2)$ (iii) $f(3)$ (iv) $f(\frac{1}{2})$

Sol. (i) $f(x) = x^2–5x+3$ ...(i)

From equation (i) by putting $x = 0$

$$f(0) = 0–0+3 = 3$$

$$f(0) = 3$$

(ii) By putting $x = (–2)$ in equation (i)

$$f(–2) = (–2)^2–5(–2)+3$$

$$\Rightarrow \quad 4+10+3 = 17$$

$$f(–2) = 17$$

(iii) By putting $x = 3$ in equation (i)

$$f(3) = (3)^2–5(3)+3$$

$$\Rightarrow \quad 9–15+3 = –3$$

$$f(3) = –3$$

(iv) By putting x = $\frac{1}{2}$ in equation (i)

$$f\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^2 –5\left(\frac{1}{2}\right)+3 = \frac{1}{4}-\frac{5}{2}+\frac{3}{1}$$

$$= \frac{1+10+12}{4} = \frac{3}{4}$$

$$f\left(\frac{1}{2}\right) = \frac{3}{4}$$

Example 2: If $f(x) = e^{-x}$, then find out the value of $\frac{f(-a)}{f(b)}$

Sol. $\quad f(x) = e^{-x} \quad$ ...(i)

By putting in equation (i)

$x = -a$ and $x = b$ we get

$$f(-a) = e^{-(-a)} = e^{a}$$

And $\quad f(b) = e^{-b}$

$$\frac{f(-a)}{f(b)} = \frac{e^{a}}{e^{-b}} = e^{a+b}$$

Example 3: If $f(x) = \frac{x^2-1}{3x+1}$, then find out

(i) $f(0)$(ii) $f(-1)$ (iii) $f(4)$

Ans. Function is given:

$$f(x) = \frac{x^2-1}{3x+1}$$

(i) By putting $x = 0$ we get

$$f(0) = \frac{0-1}{0+1} = -1$$

(ii) By putting $x = -1$ in equation (i) we get

$$f(-1) = \frac{(-1)^2-1}{3(-1)+1} = \frac{1-1}{-3+1} = 0$$

(iii) By putting $x = 4$ in equation (i) we get

$$f(4) = \frac{4^2 - 1}{3(4)+1} = \frac{16-1}{12+1} = \frac{15}{13}$$

$$f[f\{f(\mathrm{x})\}] = \frac{1}{1 - f\{f(x)\}} = \frac{1}{1 - \left(1 - \frac{1}{x}\right)}$$

$$= \frac{x}{x - x + 1} = \frac{x}{1} = x$$

$\Rightarrow \quad f[f\{f(\mathrm{x})\}] = x$

Example 4: If $f(x) = x$, $g(x) = |x|$ then find out $f(x)$ and $g(x)$

Sol. $f(x) = x$, $g(x) = |x|$

$$f(x)+g(x) = x+|x| = 2x, \; x \geq 0$$

$$= 0, \; x<0 \qquad [\because \text{for } x<0, \; |x| = -x]$$

Example 5: If $f(x) = \dfrac{x}{x-1}$ then find out the value of $\dfrac{f(a)}{f(a+1)}$

Sol. $f(x) = \dfrac{x}{x-1}$

By putting $x = a$ and $x = a + 1$ we get

$$f(\mathrm{a}) = \frac{(a)}{(a+1)}$$

And $\quad f(\mathrm{a}+1) = \dfrac{(a+1)}{(a+1-1)} = \dfrac{a+1}{a}$

$$\frac{f(a)}{f(a+1)} = \frac{\dfrac{a}{a-1}}{\dfrac{a+1}{a}} = \frac{a^2}{a^2-1}$$

Example 6: If function f: $R \to R$ is defined in this way

$$f(x) = 3x-1 \text{ if } x>3$$
$$= x^2-2 \text{ if } -2 \leq x \leq 3$$
$$= 2x+3 \text{ If } x<-2$$

then find out (i) $f(2)$ (ii) $f(4)$ (iii) $f(-1)$ (iv) $f(-3)$

Ans. (i) $f(2) = (2)^2-2 = 4-2 = 2$

(ii) $f(4) = 3(4)-1 = 12-1 = 11$

(iii) $f(-1) = (-1)^2-2 = 1-2 = -1$

(iv) $f(-3) = 2(-3)+3 = -6+3 = -3$

Example 7: If $f(x) = x^2-5x+3$ then findout $f[f(x)]$

Sol. $f(x) = x^2-5x+3$

$$f[f(x)] = [f(x)]^2-5[f(x)]+3$$
$$= (x^2-5x+3)^2-5(x^2-5x+3)+3$$
$$= x^4-25x^2-9-10x^3-30x+6x^2-5x^2+25x-15+3$$
$$= x^4-10x^3+24x^2-5x-3$$

Example 8: If $f(x) = \dfrac{1}{1-x}$, then verify $f[f\{f(x)\}] = x$

Sol. Given: $f(x) = \dfrac{1}{1-x}$

$$f\{f(x)\} = \frac{1}{1-f(x)} = \frac{1}{1-\frac{1}{1-x}} = \frac{-(1-x)}{x}$$

$$= \frac{x-1}{x} = 1-\frac{1}{x}$$

$\therefore$ $f(f(x) = x$

Example 9: If $f(x) = (a-x^n)^{\frac{1}{n}}$, then findout $f[f(x)]$

$$f(x) = (a-x^n)^{\frac{1}{n}}$$

$$f[f(x)] = [a-f(x)^n]^{\frac{1}{n}} = \left[a-\left\{\left(a-x^n\right)^{\frac{1}{n}}\right\}^n\right]^{\frac{1}{n}}$$

$$= \left[a-\left(a-x^n\right)\right]^{\frac{1}{n}} = \left(x^n\right)^{\frac{1}{n}} = x$$

EXAMPLES

Q.1. Trace the curve $y = x^2-5x+3$

Q.2. Explain the concept of functions.

Q.3. Find for what values of x the following functions are not defined:

(i) $f(x) = \dfrac{1}{\sqrt{(x-2)(x-4)}}$ (ii) $f(x) = \sqrt{x}$

Solution: (i) Let $f(x) = \dfrac{1}{\sqrt{(x-2)(x-4)}}$

When $2 < x < 4$ then $(x-2)(x-4)$ is -ve

$\therefore$ $\sqrt{(x-2)(x-4)}$ becomes imaginary

$\therefore$ $f(x)$ is not defined for $2 < x < 4$

Also at $x = 2$ and $x = 4$, $f(x)$ becomes infinite,

$\therefore$ $f(x)$ is not defined for $x = 2$ and $x = 4$

$\therefore$ $f(x)$ is not defined for $2 \le x \le 4$

(ii) $f(x) = \sqrt{x}$

clearly is not defined for negative values of x

$\therefore$ $f(x)$ is not defined for $x < 0$

Q.4. Find the values of x for which the following functions are not defined:

(i) $\frac{1}{5x-1}$ (ii) $\sqrt{x-6}$ (iii) $\frac{x^2-3x+2}{x-1}$

Sol. (i) Let $f(x) = \frac{1}{5x-1}$, $f(x)$ is not defined at $x = 3$

(ii) Let $f(x) = \sqrt{x-6}$, $f(x)$ is not a defined for $x < 6$ ($\because$ for $x <$ 6, if $f(x)$ is imaginary)

(iii) Here $f(x)$ is note defined for $x = 1$

Check Your Progress–I

Q.1. Discuss types of functions with examples.

Q.2. If $f(x) = \frac{x^2-1}{3x+1}$, then find out

(i) $f(0)$ (ii) $f(-1)$ (iii) $f(4)$

Q.3. Explain the Concept of a function.

Q.4. Define a constant function.

Q.5. What is a rational function?

Q.6. Define non-algebraic functions.

5.11. LIMITS AND CONTINUITY

5.11.1. Definition of Limits

The function $f(x)$ tends to a limit L as x tends to a, if the numerical between $f(x)$ and l can be made as small we like by making the positive difference between x and c small enough. In symbols.

$$\underset{x \to a}{\text{Lt}} f(x) = l$$

To be more rigorous, we state that $f(x)$ tends to limits l as x tends to a, if for each given E > 0, however small there exists a positive number of δ (depending upon E), such that $|f(x) - l| <$ E, for value of x for which $0 < |x - a| \leq \delta$

The function $|f(x) - l| <$ E is the same as l–E < $f(x)$ < l+E

LIMITS

Q.1. Evaluate (i) $\underset{x \to 2}{Lt} \dfrac{x^2 - 4}{x - 2} = \underset{x \to 2}{Lt} \dfrac{(x-2)(x+2)}{x-2} = \underset{x \to 2}{Lt} (x+2) = 4$

Second Method. Put x = 2+h, As $x \to 2$, $h \to 0$

$$\underset{x \to 2}{Lt} \frac{x^2 - 4}{x - 2} = \underset{h \to 0}{Lt} \frac{(2+h)^2 - 4}{2 + h - 2} = \underset{h \to 0}{Lt} \frac{h^2 - 4h}{h} = \underset{h \to 0}{Lt} (h + 4) = 4$$

(ii) $\underset{x \to 2}{Lt} \dfrac{x^2 - a^2}{x - a}$

Put $x = a + h$, as $x \to a$, $h \to 0$

$$\underset{x \to 2}{Lt} \frac{x^2 - a^2}{x - a} = \underset{h \to 0}{Lt} \frac{(a+h)^2 - a^2}{a + h - a} = \underset{h \to 0}{Lt} \frac{a^2 + h^2 + 2ah - a^2}{h}$$

$$\underset{h \to 0}{Lt} \frac{h(h + 2a)}{h} = \underset{h \to 0}{Lt} (h + 2a) = 2a$$

Q.2. Find (i) $\underset{x \to 2}{Lt} \dfrac{x^2 - 4}{x^2 - 2x}$ (ii) $\underset{x \to 1}{Lt} \dfrac{x^3 - 3x + 2}{x^3 + x^2 - 5x + 3}$

Sol. (i) $\underset{x\to 2}{Lt}\dfrac{x^2-4}{x^2-2x}$, Put $x = 2+h$, as $x\to 2,\ h\to 0$

$$= \underset{h\to 0}{Lt}\frac{(2+h)^2-4}{(2+h)^2-2(h+2)} = \underset{h\to 0}{Lt}\frac{h^2+4h}{h^2+2h} = \underset{h\to 0}{Lt}\frac{h(h+4)}{h(h+2)}$$

$$= \underset{h\to 0}{Lt}\frac{(h+4)}{(h+2)} = \frac{4}{2} = 2$$

(ii) Put $x = 1+h$, as $x\to 1,\ h\to 0$

$$\therefore \quad \underset{x\to 1}{Lt}\frac{x^3-3x+2}{x^3+x^2-5x+3} = \underset{h\to 0}{Lt}\frac{(1+h)^3-3(1+h)+2}{(1+h)^3+(1+h)^2-5(1+h)+3}$$

$$= \underset{h\to 0}{Lt}\frac{3h^2+h^3}{4h^2+h^3} = \underset{h\to 0}{Lt}\frac{3+h}{4+h} = \frac{3}{4}$$

Q.3. Evaluate $\underset{x\to 3}{Lt}\dfrac{x^2+9}{x-3}$

Solution: As long as $x\neq 3$, we have

$$\frac{x^2-9}{x-3} = \frac{(x-3)(x-3)}{x-3} = x+3$$

$$\therefore \quad \underset{x\to 3}{Lt}\frac{x^2+9}{x-3} = \underset{x\to 3}{Lt}\ (x+3) = 3+3 = 6$$

Another method:

Put $x = 3 + h$ so that as x tends to $3h$ tends to 0.

$$\therefore \quad \underset{x\to 3}{Lt}\frac{x^2+9}{x-3} = \underset{h\to 0}{Lt}\frac{(3+h)^2-9}{(3+h)-3} = \underset{h\to 0}{Lt}\frac{h(6+h)}{h}$$

$$= \underset{h\to 0}{Lt}(6+h) = 6+0 = 6$$

Q.4. Show that $\underset{x\to a}{Lt}\dfrac{x^n-a^n}{x-a}$

Sol. Put $x = a + h$, as $x \to a$, $h \to 0$

$$\underset{x\to a}{Lt}\frac{x^n - a^n}{x-a} = \underset{h\to 0}{Lt}\frac{(a+h)^n - a^n}{a+h-a} = \underset{h\to 0}{Lt}\frac{a^n(1+\frac{h}{a})^n - a^n}{h} = \underset{h\to 0}{Lt}\frac{a^n[(1+\frac{h}{a})^n - 1]}{h}$$

$$= \underset{h\to 0}{Lt}\frac{a^n[(1+n\frac{h}{a}+\frac{n(n-1)}{2!}\left(\frac{h}{a}\right)^2+.....-1]}{h} \quad \text{[By Binomial theorem]}$$

$$= \underset{h\to 0}{Lt}\frac{a^n[n\frac{h}{a}+\text{term with higher Power of } h]}{h}$$

$$= \underset{h\to 0}{Lt}\frac{a^n.\frac{h}{a}[n+\frac{n(n-1)}{2}\left(\frac{h}{a}\right)+......]}{h}$$

$$= \underset{h\to 0}{Lt}\, a^{n-1}.\frac{h}{a}[n+\frac{n(n-1)}{2!}\left(\frac{h}{a}\right)+......] = na^{n-1}$$

$$\therefore \quad \underset{x\to a}{Lt}\frac{x^n - a^n}{x-a} = na^{n-1}$$

Q.5. Evaluate $\underset{h\to 0}{Lt}\left[\frac{\sqrt{x+h}-\sqrt{x}}{h}\right]$

Solution: Rationalising the numerator, we get

$$= \underset{h\to 0}{Lt}\left[\frac{\sqrt{x+h}-\sqrt{x}}{h}\times\frac{\sqrt{x+h}+\sqrt{x}}{\sqrt{x+h}+\sqrt{x}}\right] = \underset{h\to 0}{Lt}\left[\frac{\left(\sqrt{x+h}\right)^2-\left(\sqrt{x}\right)^2}{h\sqrt{x+h}+\sqrt{x}}\right]$$

$$= \underset{h\to 0}{Lt}\frac{(x+h-x)}{h\left(\sqrt{x+h}+\sqrt{x}\right)} = \underset{h\to 0}{Lt}\frac{h}{h\left(\sqrt{x+h}+\sqrt{x}\right)}$$

$$= \underset{h\to 0}{Lt}\frac{1}{\sqrt{x+h}+\sqrt{x}} = \frac{1}{\sqrt{x}+\sqrt{x}} = \frac{1}{2\sqrt{x}}$$

Q.6. Evaluate

(i) $\underset{x\to 0}{Lt}\dfrac{\sqrt{x+2}-\sqrt{2}}{x}$,

Rationalizing $= \underset{x\to 0}{Lt}\dfrac{\sqrt{x+2}-\sqrt{2}}{x}\times\dfrac{\sqrt{x+2}+\sqrt{2}}{\sqrt{x+2}+\sqrt{2}}$

$$= \underset{x\to 0}{Lt}\frac{x+2-2}{x\left(\sqrt{x+2}+\sqrt{2}\right)} = \underset{x\to 0}{Lt}\frac{1}{\sqrt{x+2}+\sqrt{2}} = \frac{1}{\sqrt{2}+\sqrt{2}} = \frac{1}{2\sqrt{2}}$$

(ii) $\underset{x\to 0}{Lt}\dfrac{\sqrt{2-x}-\sqrt{2+x}}{x}$,

Rationalizing $= \underset{x\to 0}{Lt}\dfrac{\sqrt{2-x}-\sqrt{2+x}}{x}\times\dfrac{\sqrt{2-x}+\sqrt{2+x}}{\sqrt{2-x}+\sqrt{2+x}}$

$$= \underset{x\to 0}{Lt}\frac{2-x-(2+x)}{x\left(\sqrt{2-x}+\sqrt{2+x}\right)} = \underset{x\to 0}{Lt}\frac{-2}{\left(\sqrt{2-x}+\sqrt{2+x}\right)} = \frac{-2}{2\sqrt{2}} = \frac{-1}{2\sqrt{2}}$$

Q.7. Find the following limits

(i) $\underset{x\to 0}{Lt}\dfrac{\sqrt{1+x}-\sqrt{1-x}}{x}$ (ii) $\underset{h\to 0}{Lt}\dfrac{(x+h)^2-x^2}{h}$ (iii) $\underset{x\to 0}{Lt}\dfrac{\sqrt{1+x}-1}{x}$

Sol. $\underset{x\to 0}{Lt}\dfrac{\sqrt{1+x}-\sqrt{1-x}}{x}$, Rationalizing, we have

$$= \underset{x\to 0}{Lt}\frac{\sqrt{1+x}-\sqrt{1-x}}{x}\times\frac{\sqrt{1+x}+\sqrt{1-x}}{\sqrt{1+x}+\sqrt{1-x}}$$

$$= \underset{x\to 0}{Lt}\frac{1+x-(1-x)}{x(\sqrt{1+x}+\sqrt{1-x})} = \underset{x\to 0}{Lt}\frac{2}{\sqrt{1+x}-\sqrt{1-x}} = 1.$$

(ii) Again $\underset{h\to 0}{Lt}\dfrac{(x+h)^2-x^2}{h} = \underset{x\to 0}{Lt}\dfrac{x^2+2xh+h^2-x^2}{h}$

$$= \underset{x\to 0}{Lt}\frac{2xh+h^2}{h} = \underset{x\to 0}{Lt}(2x+h) = 2x$$

(iii) $$\underset{x\to 0}{Lt}\frac{\sqrt{1+x}-1}{x} = \underset{x\to 0}{Lt}\frac{\sqrt{1+x}-1}{x}\times\frac{\sqrt{1+x}+1}{\sqrt{1+x}+1}$$

$$= \underset{x\to 0}{Lt}\frac{1+x-1}{x(\sqrt{1+x}+1)} = \underset{x\to 0}{Lt}\frac{1}{\sqrt{1+x}+1} = \frac{1}{2}$$

5.11.2. Continuity

In order to understand the concept of continuity of a function, let us begin by considering the function $y = f(x)$ which may assume any form indicated in the figure below.

In diagram (a) we find a break in the curve at $x = c$. There are two values of $f(x)$ at the same point $x = c$. As x tends to c, $f(x)$ tends to k. But at point $x = c$, the curve jumps from to k to Q. The curve of the function is not continuous.

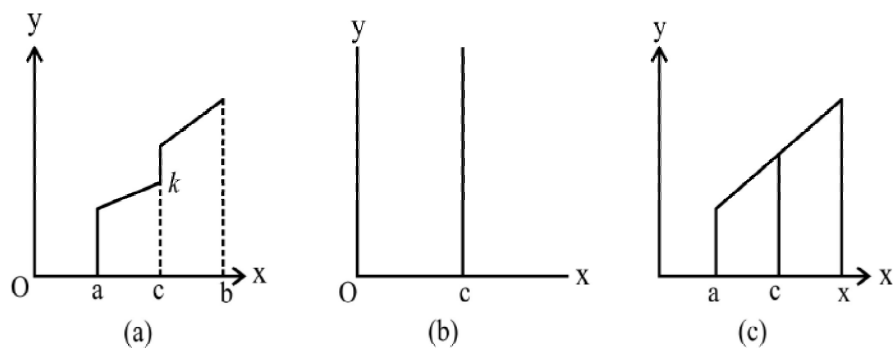


Fig.

Now consider fig. (b). It is evident that the value of $f(x)$ at c is infinity. thus the curve is constant continuous at $x = c$. In the diagram (c) there is no break or jump in the curve. The curve is continuous at $x = c$. By observing the above figures we can say that the function $f(x)$ will be continuous at $x = c$.

Definition: A function $f(x)$ is said to be continuous at $x = a$ if for any Positive number $\in$, however small, there exists a positive number $\delta(\in)$ such that

$| f(x) | - f(a) < \in$, for $|x - a| \leq \delta$

An alternative definition can be given as

we say that the function $f(x)$ is continuous x = a if (i) $f(a)$ exists

(ii) $\underset{h\to 0}{Lt} f(a+h) = \underset{h\to 0}{Lt} f(a-h) = f(a)$

Discontinuous functions: A function $f(x)$ is said to be discontinuous at x = a when either of the condition (i) and (ii) of continuity is not fulfilled, that is, when the function is not defined, at $x = a$

Or When the limits $\underset{h\to 0}{Lt} f(a+h)$, $\underset{h\to 0}{Lt} f(a-h)$ either do not exist or exist but not equal or equal but not equal to $f(a)$, in either case, $f(x)$ is said to be discontinuous at $x = a$.

Example 1: Show that $\dfrac{1}{x-a}$ is discontinuous at $x = a$.

Sol. The function is not defined at $x = a$. The reason is that, at $x = a$ the expression becomes meaningless. Therefore, the function is discontinuous at $x = a$.

Example 2. Show that $y = \dfrac{x^2-1}{x-1}$ is discontinuous at x =1 show also that the function becomes continuous, if the value of $y = 2$ at $x = 1$ is inserted.

Sol. (a) Now $f(x) = \dfrac{x^2-1}{x-1}$, $f(1) = \dfrac{1-1}{1-1} = \dfrac{0}{0}$, which is not defined

$\therefore$ $f(x)$ is not defined at $x = 1$

Hence the function is discontinuous at $x = 1$

But when we define our function as $y = \dfrac{x^2-1}{x-1}$, $x \neq 1$

$$y = \frac{x^2-1}{x-1}, \quad x \neq 1,\ y = 2 \text{ when } x = 1.$$

then $f(1)$ has a value equal to 2

$$\underset{h\to 0}{Lt} f(a+h) = \underset{h\to 0}{Lt} f(1+h) = \underset{h\to 0}{Lt} \frac{(1+h)^2-1}{1+h-1} = \underset{h\to 0}{Lt} (h+2) = 2$$

$$\underset{h\to 0}{Lt} (a-h) = \underset{h\to 0}{Lt} (1-h) = \underset{h\to 0}{Lt} \frac{(1-h)^2-1}{1-h-1} = \underset{h\to 0}{Lt} \frac{2-h}{1} = 2$$

$$\therefore \underset{h\to 0}{Lt} f(1+h) = \underset{h\to 0}{Lt} (1-h) = f(1)$$

So $f(x)$ is continuous at $x = 1$.

5.12. LET US SUM UP

In this lesson, we discussed about functions, their application in Economics and the concept of limits and continuity.

5.13. EXAMINATION ORIENTED QUESTIONS

(I) Show that $f(x) = 4x + 1$ is continuous at $x = 2$

(II) $f(x) = x^2 + 1$ is continuous at $x = 1$

(III) $f(x) = x^2 + 2$ is continuous at $x = 3$

5.14. SUGGESTED READINGS & REFERENCES

1. Allen R.G.D. (1976) : Mathematical Analysis for Economics, Macmillan.
2. Chander Romesh (2007) : Lectures on Elementary Mathematics for Economists, New Delhi.
3. Monga, G.S. (1972) : Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.
4. Leonard & Von Long (1978) : Introduction to Maths for Students of Economics, Cambridge.
5. Mehta & Madnani (1992) : Mathematics for Economists, S. Chand, New Delhi.

DERIVATIVE OR DIFFERENTIATION

LESSON NO. 6 **UNIT-II**

STRUCTURE

6.1. OBJECTIVES

After going through this lesson, you shall be able to understand :

- The Concept of Derivative.
- Rules of Differentiation.
- Parametric Equations and their Derivatives.

- Derivatives of Exponential Functions.
- Differentiation of Implicit Functions.
- Application of Differentiation in Economics.

6.2. INTRODUCTION

Differentiation is a method to compute the rate at which a dependent output y changes with respect to the change in the independent input x. This rate of change is called the derivative of y with respect to x. In more precise language, the dependence of y upon x means that y is a function of x. This functional relationship is often denoted by $y = f(x)$, where f denotes the function. If x and y are real numbers, and if the graph of y is plotted against x, the derivative measures the slope of this graph at each point.

The simplest case is when y is a linear function of x, meaning that the graph of y against x is a straight line. In this case, $y = f(x) = mx + c$, for real numbers m and c and the slope m is given by

$$m = \frac{\text{change in } y}{\text{change in } x} = \frac{\Delta y}{\Delta x}$$

In this lesson, we shall discuss in detail about the concepts of differentiation, types and uses in Economics.

6.3. DEFINITION OF DERIVATIVE

Let $y = f(x)$ be a finite and single valued function of x in an interval and be any point and $(x+h)$ be another point of the interval, then $\underset{h\to 0}{Lt}\frac{f(x+h)-f(x)}{h}$, if it exists, is the derivative of $f(x)$ at the point x. The derivative of $y = f(x)$ is denoted by

$$\frac{dy}{dx} = \frac{d}{dx} f(x) = f'(x) = y'$$

Differentiation 'ab-initio' or From First Principle or From Definition

Consider $y = f(x)$

If x changes by a small quantity (Δx), y also changes by Δy, Corresponding to change in x, then

$$y + \Delta y = f(x + \Delta x)$$

Or $\Delta y = f(x + \Delta x) - y$

$$\Delta y = f(x + \Delta x) - f(x)$$

$$\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x)-f(x)}{\Delta x}$$

$$\underset{\Delta x \to 0}{Lt} \frac{\Delta y}{\Delta x} = \underset{\Delta x \to 0}{Lt} \frac{f(x+\Delta x)-f(x)}{\Delta x}$$

$\frac{dy}{dx}$ or $\underset{\Delta x \to 0}{lim} \frac{\Delta y}{\Delta x}$ or $\frac{d(y)}{dx}$ is called the derivative or differential coefficient of y with respect to x.

6.4. RULES OF DIFFERENTIATION

Rule No. 1. Derivative of constant is *zero*

Let $y = f(x) = c$ is a constant function.

$$y + \Delta y = f(x + \Delta x) = C$$

$$y + \Delta y - y = f(x + \Delta x) - f(x) = 0$$

$$\Delta y = f(x + \Delta x) - f(x) = 0$$

$$\frac{\Delta y}{\Delta x} = f(x + \Delta x) - f(x) = 0$$

$$\underset{\Delta x \to 0}{Lt} \frac{\Delta y}{\Delta x} = \underset{\Delta x \to 0}{Lt} \frac{f(x+\Delta x)-f(x)}{\Delta x} = 0$$

$$\frac{dy}{dx} = 0$$

Hence $\underset{\Delta x \to 0}{\text{limit}} \frac{f(x+\Delta x)-f(x)}{\Delta x} = 0$

i.e. if $f(x) = C$

Then $\frac{d}{dx}(fx) = 0$

Rule No. 2. Power function rule:

The derivative of a power function: $y = x^n$ is nx^{n-1}

i.e. If $y = x^n$ is given

then, $\frac{dy}{dx} = nx^{n-1}$

Proof / Derivation

Let $y = x^n$

$$y+\Delta y = (x + \Delta x)^n$$

$$\Delta y = (x + \Delta x)^n - x^n$$

$$\Delta y = x^n\left[\left(1+\frac{\Delta x}{x}\right)^n - 1\right]$$ **Apply Binomial**

$$\Delta y = x^n\left[\left(1+n\frac{\Delta x}{x}+\frac{n(n-1)}{2!}\left(\frac{\Delta x}{x}\right)^2+......\infty\right)-1\right]$$

$$\Delta y = x^n\left[\left(1+n\frac{\Delta x}{x}+\frac{n(n-1)}{2!}\left(\frac{\Delta x}{x}\right)^2+......\right)\right]$$

$$\Delta y = x^n.\frac{\Delta x}{x}\left[n+\frac{n(n-1)}{2!}\left(\frac{\Delta x}{x}\right)+......\right]$$

Dividing both sides by Δx, we get

$$\frac{\Delta y}{\Delta x} = x^{n-1}\left[n+\frac{n(n-1)}{2!}\left(\frac{\Delta x}{x}\right)+......\right]$$

Taking limits as $\Delta x \to 0$, we get

$$\underset{\Delta x\to 0}{Limit}\frac{\Delta y}{\Delta x} = x^{n-1}\left[n+\frac{n(n-1)}{2!}\left(\frac{\Delta x}{x}\right)+......\right]$$

$$\frac{dy}{dx} = x^{n-1}(n) = nx^{n-1}$$

$$\frac{dy}{dx} = nx^{n-1}$$

For example, the derivative of $y = x^5$ is $\frac{dy}{dx} = \frac{d}{dx}(x)^5 = 5x^4$

the derivative of $y = x^8$ is $\frac{dy}{dx} = \frac{d}{dx}(x)^8 = 8x^7$

the derivative of $y = x$ is $\frac{dy}{dx} = \frac{d}{dx}(x) = 1(x)^0 = 1$

the derivative of $y = \frac{1}{x^3}$ is $\frac{dy}{dx} = \frac{d}{dx}(x)^{-3} = -3(x)^{-4} = \frac{-3}{x^4}$

and derivative of $y = 4x$ is $\frac{dy}{dx} = \frac{d}{dx}(4x) = 4$

Check Your Progress–I

Q.1. If $y = x^{14}$, find $\frac{dy}{dx}$.

Q.2. If $y = x^{9/2}$, find derivative of y with respect to x.

Q.3. Find the differential co-efficient of $x^{-3/2}$ with respect to x.

Q.4. If $y = x^6$, find rate of change of y with respect to small unit change of x.

Q.5. If $y = (3x + 4)^5$, find derivative of y with respect to x.

Q.6. Find differential coefficient of $(3x - 5)^{-7/2}$ w.r.t. to x.

Q.7. If $y = \sqrt{x} - \dfrac{1}{\sqrt{x}}$, find $\dfrac{dy}{dx}$.

Q.8. Find derivative of the following w.t. to x :

(i) $3x^2+4$, (ii) $3x^3+4x^2+5x-3$, (iii) $x(x^2+3)$

Q.9. Differentiate the following w.t. to x :

$$\frac{5x^4-7x^3+2x^2+x-3}{x^2}$$

Rule 3. Sum-difference rule Or Derivation of sum or difference of functions

Or

Prove that

$$\frac{d}{dx}(u-v) = \frac{d}{dx}(u) \pm \frac{d}{dx}(v)$$

Where u and v are derivable functions at x.

Proof: Let $y = u + v$

$$\therefore \quad y + \Delta y = [(u + \Delta u) + (v + \Delta v)]$$

$$\Delta y = \Delta u + \Delta v$$

Dividing both sides by Δx, we get

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}$$

Taking limits as $\Delta x \to 0$, we get

$$\frac{dy}{dx} = \frac{d}{dx}(u) + \frac{d}{dx}(v)$$

Hence $\frac{d}{dx}(u+v) = \frac{d}{dx}(u) + \frac{d}{dx}(v)$

Similarly, we can prove

$$\frac{d}{dx}(u-v) = \frac{d}{dx}(u) - \frac{d}{dx}(v)$$

Hence

$$\frac{d}{dx}(u \pm v) = \frac{d}{dx}(u) \pm \frac{d}{dx}(v)$$

Therefore, the theorem states that the derivative of the algebraic sum of two functions is equal to the corresponding algebraic sum of their derivtives, provided these derivatives exist.

For example, if $y = 5x^2+9x^5$, then $\frac{dy}{dx} = \frac{d}{dx}(5x^2) + \frac{d}{dx}(9x^5)$

or $\quad \frac{dy}{dx} = 10x + 45x^4$

If $\quad y = 2x^3-3x^2+8$, then $\frac{dy}{dx}(2x^3) - \frac{d}{dx}(3x^2) + \frac{d}{dx}(8)$

or $\quad \frac{dy}{dx} = 6x^2-6x+0 = 6x(x-1)$

EXAMPLES

Example 1. Diff. $x+\sqrt{x}+3$ w.r. to x.

Sol. $y = x+\sqrt{x}+3$, Diff. w.r. to x.

$$\frac{dy}{dx}=\frac{d}{dx}\ (x+\sqrt{x}+3)$$

$$=\frac{d}{dx}(x)+\frac{d}{dx}\sqrt{x}+\frac{d}{dx}3,$$

$$\left[\text{Use } \frac{d}{dx}x^n = nx^{n-1} \text{ and } \frac{d}{dx}(\text{c}) = 0\right]$$

$$=1+\frac{1}{2}x^{\frac{1}{2}}+0+1+\frac{1}{2\sqrt{x}}$$

Example 2. If $y=4x^5+\frac{1}{2}x^4-x^3+2x-8,$ find $\frac{dy}{dx}$

Sol. $\quad y = 4x^5+\frac{1}{2}x^4-x^3+2x-8$

$$\frac{dy}{dx} = \frac{d}{dx}\left(4x^5+\frac{1}{2}x^4-x^3+2x-8\right)$$

$$= 4\frac{d}{dx}(x^5)+\frac{1}{2}\frac{d}{dx}(x^4)-\frac{d}{dx}(x^3)+2\frac{d}{dx}(x)-\frac{d}{dx}(8)$$

$$= 4.5x^4+\frac{1}{2}.4x^3-3.x^2+2.1-0$$

$$\left[\text{Use } \frac{d}{dx}\left(x^n\right) = nx^{n-1}\right]$$

$$= 20x^4+2x^3-3x^2+2$$

Example 3. If $y=\sqrt{x}+\frac{1}{\sqrt{x}}$, then show that $2x\frac{dy}{dx}+y=2\sqrt{x}$

Sol. As $y=\sqrt{x}+\frac{1}{\sqrt{x}}$, Diff. w.r. to x

$$\frac{dy}{dx} = \frac{d}{dx}\left(\sqrt{x}+\frac{1}{\sqrt{x}}\right)$$

$$= \frac{d}{dx}(x^{1/2})+\frac{d}{dx}(x^{-1/2})$$

$$= \frac{1}{2}x^{-\frac{1}{2}}-\frac{1}{2}x^{-\frac{3}{2}}$$

$$\text{L.H.S.} = 2x\frac{dy}{dx}+y$$

$$= 2x\left[\frac{1}{2}x^{-\frac{1}{2}}-\frac{1}{2}x^{-\frac{3}{2}}\right]+\sqrt{x}+\frac{1}{\sqrt{x}}$$

$$= x^{\frac{1}{2}}-x^{-\frac{1}{2}}+x^{\frac{1}{2}}+x^{-\frac{1}{2}}$$

$$= 2\sqrt{x} = \text{R.H.S.}$$

Rule 4. Derivation of a product

Or

The derivative of the product of two functions is equal to the first function times the derivative of the second function plus the second function times the derivative of the first function.

Or

Prove that

$$\frac{d}{dx}(uv) = u\frac{d}{dx}(v) + v\frac{d}{dx}(u)$$

where u and v are two derivable function at x

Proof: Let $y = uv$

$$\therefore \quad y + \Delta y = (u+\Delta u)(v+\Delta v)$$

$$y + \Delta y = uv+u.\Delta v+v.\Delta u+\Delta u\ \Delta v$$

$$\Delta y = uv + u.\Delta v + v.\Delta u + \Delta u\ \Delta v - uv$$

$$\Delta y = u.\Delta v + v.\Delta u + \Delta u\ \Delta v$$

dividing both sides by Δx

$$\frac{\Delta y}{\Delta x} = \frac{u\Delta v}{\Delta x} + \frac{v\Delta u}{\Delta x} + \frac{\Delta u.\Delta v}{\Delta x}$$

As $\Delta x \to 0,\ \Delta u \to 0$

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$$

Thus, the theorem states that the derivative of the product of two functions = first function × derivative of the second function + second function × derivative of the first function.

For example; if $y = (4x+8)\ (4x^2)$

$$\frac{dy}{dx} = (4x + 8)\frac{d}{dx}(4x^2) + (4x^2)\ \frac{d}{dx}(4x + 8)$$

$$\frac{dy}{dx} = (4x+8)(8x) + (4x^2)(4)$$

$$= 32x^2 + 64x + 16x^2$$

$$= 48x^2 + 64x$$

Example 1. Differentiate $e^x \log x$ w.r. to x.

Sol. Let $f(x) = e^x \log x$, Diff. w.t. to x

$$f'(\mathrm{x}) = e^x\frac{d}{dx}\log x + \log x\frac{d}{dx}e^x, \text{ (by §3)}$$

$$= e^x\frac{1}{x} + \log x.e^x = e^x\left(\frac{1}{x} + \log x\right)$$

$$\left(\because \frac{d}{dx}\log x = \frac{1}{x}, \frac{d}{dx}e^x = e^x\right)$$

Example 2. Differentiate $x^5(2x^2+1)$ w.r. to x.

Sol. $y = x^5(2x^2+1)$, Diff. w.t. to x

$$\frac{dy}{dx} = x^5\frac{d}{dx}(2x^2+1) + (2x^2+1)\frac{d}{dx}(x^5)$$

$$= x^5\left[2\frac{d}{dx}x^2 + \frac{d}{dx}(1)\right] + (2x^2+1).5x^4$$

$$\left[\text{Use } \frac{d}{dx}x^n = nx^{n-1}\right]$$

$$= x^5(4x+0) + (2x^2+1).5x^4$$

$$= 4x^6 + 10x^6 + 5x^4 = 14x^6 + 5x^4$$

Example 3. If $y = \left(x+\frac{1}{x}\right)\left(\sqrt{x}+\frac{1}{\sqrt{x}}\right)$, find derivative of y w.r. to x.

Sol. Now $y = \left(x+\frac{1}{x}\right)\left(\sqrt{x}+\frac{1}{\sqrt{x}}\right)$, Diff. w.r. to x.

$$\frac{dy}{dx} = \left(x+\frac{1}{x}\right)\frac{d}{dx}\left(\sqrt{x}+\frac{1}{\sqrt{x}}\right) + \left(\sqrt{x}+\frac{1}{\sqrt{x}}\right)\frac{d}{dx}\left(x+\frac{1}{x}\right)$$

$$= \left(x+\frac{1}{x}\right)\left[\frac{d}{dx}(\mathrm{x}^{1/2}) + \frac{d}{dx}(\mathrm{x}^{-1/2})\right]$$
$$+\left(\sqrt{x}+\frac{1}{\sqrt{x}}\right)\left[\frac{d}{dx}(x) + \frac{d}{dx}(x)^{-1}\right]$$

$$\left[\text{Use } \frac{d}{dx}x^n = nx^{n-1}\right]$$

$$= \left(x+\frac{1}{x}\right)\left(\frac{1}{x}x^{-1/2} - \frac{1}{2}x^{-3/2}\right) + \left(\sqrt{x}+\frac{1}{\sqrt{x}}\right)\left(1-\frac{1}{x^2}\right)$$

$$= \left(\frac{x^2+1}{x}\right)\left(\frac{x-1}{2x^{3/2}}\right)+\left(\frac{x+1}{\sqrt{x}}\right)\left(\frac{(x-1)(\mathrm{x}+1)}{x^2}\right)$$

$$= \frac{(x^2+1)(x-1)}{2x^{5/2}}+\frac{(x+1)^2(x-1)}{x^{5/2}}$$

$$= \frac{x-1}{2x^{5/2}}[3x^2+4x+3]$$

Example 4. Differentiate $(7x-8)^4\,(5x-1)^3$ w.r. to x.

Sol. Now $\quad y = (7x-8)^4\,(5x-1)^3$. Diff. w.r. to x

$$\frac{dy}{dx} = (7x-8)^4\,\frac{d}{dx}(5x-1)^3+(5x-1)^3\frac{d}{dx}(7x-8)^4$$

$$\left[\text{Use } \frac{d}{dx}(uv)=u\frac{d}{dx}v+v\frac{d}{dx}u\right]$$

$$= (7x-8)^4.\,3\,(5x-1)^{3-1}.5 \;+\; (5x-1)^3.4(7x-8)^{4-1}.7$$

$$\left[\text{Use } \frac{d}{dx}(ax+b)^n=n(ax+b)^{n-1}.a\right]$$

$$= (7x-8)^3.\,(5x-1)^2\,[15\,(7x-8)+28(5x-1)]$$

$$= (7x-8)^3.\,(5x-1)^2\,[245x-148]$$

Check Your Progress–II

Differentiate the following w.r.t x :

Q.1. $(x+6)^2(x+2)^2(x+5)^3$

Q.2. $(3x+2)^{1/3}(x+1)$

Q.3. $(5-2x)(2x^3+3)$

Q.4. $(2x^2-3)(3x^2+x-1)$

Q.5. $(5x-7)^4(4x-3)^3$

Rule 5. Derivative of a Quotient of two functions

Prove that

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{d}{dx}(u)-u\frac{d}{dx}(v)}{v^2}$$

where u and v are both derivable functions of x and $v \neq 0$

Proof: Let $y = \frac{u}{v}$

$$y+\Delta y = \frac{u+\Delta u}{v+\Delta v}$$

$$\Delta y = \frac{u+\Delta u}{v+\Delta v}-\frac{u}{v}$$

$$= \frac{v(u+\Delta u)-u(v+\Delta v)}{v(v+\Delta v)}$$

$$\therefore \qquad \Delta y = \frac{v.\Delta u - u.\Delta v)}{v(v+\Delta v)}$$

Dividing both sides by Δx, we get

$$\frac{\Delta y}{\Delta x} = \frac{\frac{v.\Delta u - u.\Delta v)}{\Delta x}}{v(v+\Delta v)}$$

Proceeding to limits as $\Delta x \to 0$, also $\Delta v \to 0$, we get

$$\frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v(v+0)} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$$

The theorem states that the derivative of a quotient of two functions is equal to the denominator times the derivative of numerator minus the numerator times the derivative of the denominator, the whole divided by the square of the denominator provided that denominator is not equal to zero and their derivative exists.

The rule:

$$\frac{d}{dx}(\text{Quotient of two functions}) = \frac{D^r \frac{d}{dx} N^r - N^r \frac{d}{dx} D^r}{D^{r^2}}$$

Where D^r = Denominator

N^r = Numerator

Example 1. $\frac{d}{dx}\left(\frac{5x^2+1}{x}\right)$

Sol. $$\frac{d}{dx}\left(\frac{5x^2+1}{x}\right) = \frac{x\frac{d}{dx}(5x^2+1) - (5x^2+1)\frac{d}{dx}(x)}{x^2}$$

$$= \frac{x(10x) - (5x^2+1)(1)}{x^2} = \frac{10x^2 - (5x^2+1)}{x^2}$$

$$= \frac{(5x^2-1)}{x^2} = 5 - \frac{1}{x^2}$$

$$\frac{d}{dx}\frac{(5x^2-1)}{x^2} = 5 - \frac{1}{x^2}$$

Example 2. Find the derivative of $\dfrac{x+2}{3+\log x}$ w.r. to x.

Sol. Let $\quad y = \dfrac{x+2}{3+\log x}$. Diff. w.r. to x

$$\frac{dy}{dx} = \frac{(3+\log x)\dfrac{d}{dx}(x+2)-(x+2)\dfrac{d}{dx}(3+\log x)}{(3+\log x)^2}$$

$$\left[\text{Use } \frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\dfrac{du}{dx}u-u\dfrac{dv}{dx}v}{v^2}\right]$$

$$= \frac{(3+\log).1-(x+2)\left(\dfrac{1}{x}\right)}{(3+\log x)^2} = \frac{2+\log x-\dfrac{2}{x}}{(3+\log x)^2}$$

Example 3. Differentiate $\sqrt{\dfrac{1-x}{2+x}}$ w.r. to x.

Sol. Let $\quad y = \sqrt{\dfrac{1-x}{2+x}}$, Diff. w.r. to x

$$\frac{dy}{dx} = \frac{\sqrt{2+x}\dfrac{d}{dx}(1-x)^{1/2}-\sqrt{1-x}.\dfrac{d}{dx}(2+\mathrm{x})^{1/2}}{(\sqrt{2+x})^2}$$

$$= \frac{\sqrt{2+x}\left\{\dfrac{1}{2}.(1-x)^{1/2}(-1)\right\}^{\dagger}-\sqrt{1-x}.\dfrac{1}{2}(2+\mathrm{x})^{-1/2}}{2+x}$$

$$= \frac{-1}{2(2+x)}\left(\frac{\sqrt{2+x}}{\sqrt{1-x}}+\frac{\sqrt{1-x}}{\sqrt{2+x}}\right)$$

$$= \frac{-3}{2(2+x)^{3/2}\sqrt{1-x}}$$

Example 4. Diff. $\frac{3x}{7x^2+8}$ w.r.t. to x.

Sol. Let $y = \frac{3x}{7x^2+8}$, Differentiate w.r. to x.

$$\frac{dy}{dx} = \frac{d}{dx}\left(\frac{3\text{x}}{7x^2+8}\right)$$

$$= \frac{(7x^2+8)\frac{d}{dx}(3x) - 3x\frac{d}{dx}(7x^2+8)}{(7x^2+8)^2}$$

$$\left[\because \frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{d}{dx}(u) - u\frac{d}{dx}v}{v^2}\right]$$

$$= \frac{(7x^2+8)(.3.1) - 3x\left(\frac{d}{dx}7x^2 + \frac{d}{dx}(8)\right)}{(7x^2+8)^2}$$

$$= \frac{(7x^2+8)3 - 3x(14x+0)}{(7x^2+8)^2} \qquad \left[\text{Use } \frac{d}{dx}x^n = nx^{n-1}\right]$$

$$= \frac{21x^2 - 24 - 42x^2}{(7x^2+8)^2}$$

$$= \frac{24 - 21x^2}{(7x^2+8)^2}$$

Check Your Progress–III

Differentiate the following w.r. to x :

Q.1. $\dfrac{3x+4}{4x+5}$

Q.2. $\dfrac{x-b}{x+b}$

Q.3. $\sqrt{\dfrac{1-x}{1+x}}$

Q.4. $\dfrac{(2-5x)^2}{x^3-1}$

Q.5. $y=\frac{(x+2)(x+3)}{(x-2)(x-3)}$

Q.6. $\frac{x^2+1}{x^2+2}$

Q.7. $y=\frac{x^2+4}{x+2}$

Q.8. $\dfrac{(x-2)(2x+3)}{(x+7)(1-x)}$

Q.9. $\dfrac{(x+3)(x+4)}{(x-3)(x+4)}$

Q.10. $\dfrac{(2-5x)^3}{x^2-1}$

Q.11. $\dfrac{x^3-6x^2}{5x^2-1}$

Q.12. $\frac{(x+1)(2x-1)}{x-3}$

Q.13. $\frac{x^6}{ax+b}$

Q.14. $\frac{x^3+3x+4}{x+1}$

Q.15. $\dfrac{1+x}{\sqrt[3]{1-x}}$

Q.16. $\dfrac{2x^2-3x+1}{\sqrt{x-3}}$

Q.17. Find derivative of $\dfrac{6x^2}{2x+1}$ w.r.t. to x.

Q.18. (*i*) Differentiate $\frac{x+1}{\sqrt{x}}$ w.r. to x.

(*ii*) Differentiate the following w.r. to x, $y = \frac{x^2+3}{x}$.

Q.19. If $y = \frac{x}{x+5}$, prove that $x\frac{dy}{dx} = y(1-y)$.

Q.20. If $y = \frac{(x+2)(2x+1)}{x^3-1}$, find $\frac{dy}{dx}$.

Q.21. Find $\frac{dy}{dx}$, where $y = \sqrt{\frac{1-2x}{1+2x}}$

Q.22. (*a*) Diff. $\dfrac{x^3}{x^2+1}$ w.r. to x.

(*b*) Find $\dfrac{dy}{dx}$, for $y = \dfrac{x^2+2}{\sqrt{2x+1}}$

Q.23. Differentiate $\dfrac{3-4x}{5x-2}$ w.r. to x.

Rule No. 6. Chain rule (function of a function rule)

If we have a function $y = f(x)$ where x is in turn a function of another variable z, say $x = g(z)$, then the derivative of y with respect to z is equal to the derivative of y w.r.t. x times the derivative of x. w.r.t. to z

If $y = f(x)$ and $x = g(z)$

then $\dfrac{dy}{dz} = \dfrac{dy}{dx} \times \dfrac{dx}{dz}$

for example, if $y = (3x^2)$ where $x = 2z+5$ then

$$\frac{dy}{dx} = \frac{dy}{dx} \times \frac{dx}{dz} = \frac{d}{dx}(3x^2) \ \frac{d}{dz}(2z+5)$$

$$= (6x)(2) = 12x$$

The above rule also helps when we differentiate a function involving power. For instance.

$$y = (x^2+8x-8)^{16}$$

Let $(x^2+8x-8) = z$, so that $y = z^{16}$ and $z = (x^2+8x-8)$

we apply chain rule to find $\dfrac{dy}{dx}$ which is $\dfrac{dy}{dx}.\dfrac{dx}{dz}$

$$\therefore \quad \frac{dy}{dx} = \frac{d}{dz}(z^{16}).\frac{d}{dx}(x^2+8x-8)$$

$$= 16z^{15}(2x+8); \text{ substitute for } z$$

$$= 16(x^2+8x-8)^{15} \ (2x+8)$$

Example 1. If $y = 2w^2 + 1$, $w = 2z^2$, $z = 2x + 3x^2$, find derivative to y w.r. x.

Sol. We know $\dfrac{dy}{dx} = \dfrac{dy}{dw}.\dfrac{dw}{dz}\dfrac{dz}{dx}$ ...(1) [by §5(a)]

As $y = 2w^2 + 1$, $\dfrac{dy}{dw} = \dfrac{d}{dw}(2w^2+1) = 4w$

As $w = 2z^2$, $\therefore$ $\dfrac{dw}{dz} = \dfrac{d}{dz}(2z^2) = 4z$.

As $z = 2x + 3x^2$, $\dfrac{dz}{dx} = \dfrac{d}{dx}(2x + 3x^2) = 2 + 6x$

Substitute values of $\dfrac{dy}{dw}, \dfrac{dw}{dz}, \dfrac{dz}{dx}$ in (1)

$$\frac{dy}{dx} = (4w)(4z)(2 + 6x), \text{ but } w = 2z^2, \; z = 2x + 3x^2$$

$$= 4\,(2z^2)(4z)(2 + 6x)$$

$$= 32\,(2x+3x^2)^3(2 + 6x)$$

Example 2. If $z = y^3 + 7y^4$ and $y = 3x^2$, find derivative of z w.r., to x.

Sol. Now $\dfrac{dz}{dx} = \dfrac{dz}{dy}.\dfrac{dy}{dx}$ ...(1)

Since $z = y^3 + 7y^4, \; \dfrac{dz}{dy} = \dfrac{d}{dy}(y^3 + 7y^4) = 3y^2 + 28y^3$

Also as $y = 3x^3, \; \dfrac{dy}{dx} = \dfrac{d}{dy}(3x^2) = 6x$

Put values of $\dfrac{dz}{dy}$ and $\dfrac{dy}{dx}$ in (1),

$$\frac{dz}{dx} = (3y^2 + 28y^3).6x, \text{ but } y = 3x^2$$

$$= [3(3x^2)^2 + 28(3x^2)^3].6x = 162x^5 + 4536x^7$$

Q. Find the differential co-efficient of a^x, a being a constant.

Sol. Let $y = a^x$ ... (i)

Let δy be increment of y corresponding to an increment δx of x, so that

$$y + \delta y = a^{x+\delta x} = a^x.a^{\delta x} \quad \text{... (ii)}$$

Subtracting (i) from (ii), we get

$$\delta y = a^x.a^{\delta x} - a^x$$

or $\delta y = a^x(a^{\delta x} - 1)$

Diving both sides by δx, we get

$$\frac{\delta y}{\delta x} = a^x.\ \frac{a^{\delta x} - 1}{\delta x}$$

Proceeding to the limits as $\delta x \to 0$, we get

$$\frac{dy}{dx} = a^x . \log e^a \qquad \left[\because \underset{\delta x \to 0}{Lt} \frac{a^{\delta x} - 1}{\delta x} = \log e^a\right]$$

Some other important results.

Logarithmic function: $y = \log x, \dfrac{dy}{dx} = \dfrac{1}{x}$

Exponential function: $y = e^x$, $\dfrac{dy}{dx} = e^x$, if $y = e^{ax}$, $\dfrac{dy}{dx} = ae^{ax}$

Trignometric function:

$$y = \sin x, \frac{dy}{dx} = \cos x$$

$$y = \cos x, \frac{dy}{dx} = -\sin x$$

Polynomial function: The word Polynomial means multi term and a polynomial function of a single variable 'x' has the general form

$$y = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$$

Here a_0, a_1, a_2 ... a_n are constants and x_1, x_2.... are variables.

Example 1: Differentiate the following:

(i) $\cot x$ (ii) $\cos \log x$ (iii) $\log \sqrt{\sin x}$

(iv) $\sqrt{3x^2 + 4x + 3}$ (v) $\dfrac{1}{\sqrt{1 - x^2}}$ (vi) $e^{\frac{x^2}{2}}$

Sol. (i) $y = \cot x$

$$\therefore \quad y = \frac{\cos x}{\sin x}$$

$$= \frac{\sin x . \frac{d}{dx} \cos x - \cos x \frac{d}{dx} (\sin x)}{\sin^2 x} \quad \text{(Quotient rule)}$$

$$= \frac{-\sin x \sin x - \cos x \cos x}{\sin^2 x}$$

$$= \frac{-1}{\sin^2 x} \quad = - \ \operatorname{cosec}^2 x$$

$$\frac{dy}{dx} = -\operatorname{cosec}^2 x$$

(ii) $y = \cos \ \log \ x$

$$\therefore \quad \frac{dy}{dx} = (-\sin \ \log \ x) \ \frac{d}{dx} \log \ x \quad = \quad \frac{-1}{x} \sin \ \log \ x$$

(iii) $y = \log \sqrt{\sin x}$

$$\therefore \quad \frac{dy}{dx} = \frac{1}{\sqrt{\sin x}} \times \frac{1}{2} (\sin)^{-1/2} \cos \ x$$

$$= \frac{1}{2} \frac{\cos x}{\sin x} = \frac{1}{2} \cot \ x$$

(iv) $\sqrt{3x^2 + 4x + 3}$

Let $y = (3x^2 + 4x + 3)^{1/2}$

$$\frac{dy}{dx} = \frac{1}{2} (3x^2 + 4x + 3)^{-1/2} (6x + 4) \quad = \frac{6x+4}{2\sqrt{3x^2 + 4x + 3}}$$

(v) $y = \dfrac{1}{\sqrt{1 - x^2}} = (1-x^2)^{-1/2}$

$$\therefore \quad \frac{dy}{dx} = \frac{-1}{2} (1-x^2)^{-3/2} (-2x) = \quad \frac{2x}{2(1-x^2)^{3/2}} = \frac{x}{(1-x^2)^{3/2}}$$

(vi) $y = e^{\frac{x^2}{2}}$

$$\therefore \quad \frac{dy}{dx} = e^{\frac{x^2}{2}} \frac{d}{dx} \left(\frac{x^2}{2} \right) = \quad e^{\frac{x^2}{2}} \cdot \frac{1}{2} (2x)$$

$\therefore \quad \frac{dy}{dx} = x.\ e^{\frac{x^2}{2}}$

Examples 2: Differentiate the following:

(i) $7x^2–2x^3–8x–7$ (ii) $e^x.\log.x.$

(iii) $(3x^2+2)(2x–9)(x^3+7)$ (iv) $\frac{x}{4x+2}$

(v) $(x^3+2x)^2$ (vi) $y = \sqrt{x^3+4x^2+5x+6}$

(vii) $\log_a u$ find $\frac{dy}{dx}$

Sol. (i) Let $y = 7x^4–2x^3–8x–7$

$$\frac{dy}{dx} = 7\times4x^3–2\times3x^2–8\ –0$$

$$\frac{dy}{dx} = 28x^3–\ 6x^2–8$$

(ii)Let $y= e^x.\log.x.$

$$\frac{dy}{dx} = e^x.\ \frac{d}{dx}\log\ x\ +\ \log\ x\ \frac{d}{dx}e^x$$

$$= e^x\ \frac{1}{x} +\ \log\ x\ \frac{d}{dx}e^x\ =\ \frac{e^x}{x} +\ e^x\ \log\ x$$

$$= \frac{e^x}{x}[\ 1+x\ \log\ x]$$

(iii) $y = (3x^2+2)(2x–9)(x^3+7)$

$$\frac{d}{dx} = (3x^2+2)\frac{d}{dx}(2x–9)(x^3+7)+(2x–9)(x^3+7)\frac{d}{dx}(3x^2+2)$$

$$= (3x^2+2)[(2x–9)\frac{d}{dx}(x^3+7)+(x^3+7)\frac{d}{dx}(2x–9)]+(2x–9)(x^3+7)(3\times2x^1)$$

$= (3x^2+2)[(2x–9)(3x^2)+(x^3+7)(2)] \;+\; 6x(2x–9)(x^3+7)$

$= (3x^2+2)\;\;(2x–9)(3x^2)+2(3x^2+2)(x^3+7)+6x(2x–9)(x^3+7)$

(iv) $\quad y = \dfrac{x}{4x+2}$

$$\frac{dy}{dx} = \frac{(4x+2).1 - x\dfrac{d}{dx}(4x+2)}{(4x+2)^2}$$

$$\frac{dy}{dx} = \frac{4x+2-x(4)}{(4x+2)^2} = \frac{2}{(4x+2)^2}$$

(v) $\quad y = (x^3+2x)^2$

$$\frac{dy}{dx} = 2(x^3+2x)\frac{d}{dx}(x^3+2x) = 2[(x^3+2x)\;\;(3x^3+2)]$$

(vi) $\quad y = \sqrt{x^3+4x^2+5x+6}$

$$y = (x^3+4x^2+5x+6)^{½}$$

$$\frac{dy}{dx} = \frac{1}{2}(x^3+4x^2+5x+6)^{-½}\;(3x^2+6x+5)$$

$$= \frac{(3x^2+6x+5)}{2(x^3+4x^2+5x+6)^{½}} = \frac{(3x^2+6x+5)}{2\sqrt{x^3+4x^2+5x+6}}$$

(vii) If $y = \log_a x$ find $\dfrac{dy}{dx}$

$$y + \Delta y = \log_a (x+\Delta x)$$

$$\Delta y = \log_a (x+\Delta x) - y$$

$$= \log_a (x+\Delta x) - \log_a x$$

$$\Delta y = \log_a \frac{x+\Delta x}{x} \qquad \because \;\log m - \log n = \log\left(\frac{m}{n}\right)$$

$$= \log_a \left(1+\frac{\Delta x}{x}\right)$$

Dividing by Δx

$$\therefore \quad \frac{\Delta y}{\Delta x} = \frac{1}{\Delta x}\log_a\left(1+\frac{\Delta x}{x}\right)$$

$$= \frac{1}{x}\left(\frac{x}{\Delta x}\right)\log_a\left(1+\frac{1}{\frac{x}{\Delta x}}\right)$$

$$= \frac{1}{x}\log_a\left(1+\frac{1}{\frac{x}{\Delta x}}\right)^{x/\Delta x} \qquad \because \log m^n = n \log m$$

Let us set $\frac{x}{\Delta x} = n$ then $\frac{\Delta x}{x} = \frac{1}{n}$

And $\Delta x \to 0,\ n \to \infty$ Thus

$$\frac{\Delta y}{\Delta x} = \frac{1}{x}\log_a\left(1+\frac{1}{n}\right)^n$$

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \frac{1}{x}\lim_{\Delta x\to\infty}\log_a\left(1+\frac{1}{n}\right)^n$$

$$\therefore \quad \frac{dy}{dx} = \frac{1}{x}\log_a e \qquad [\because \log_a e = 1]$$

$$\frac{dy}{dx} = \frac{1}{x}$$

Example 3:

$$\frac{d}{dx}(x+2)^2(x+5)^3$$

$$= (x+2)^2 \frac{d}{dx}(x+5)^3 + (x+5)^3 \frac{d}{dx}(x+2)^2 \qquad \text{[Apply Product Rule]}$$

$$= (x+2)^2.3(x+5)^2(1)+(x+5)^3.2(x+2)(1)$$

$$= (x+2)(x+5)^2[3(x+2)+(x+5).2]$$

$$= (x+2)(x+5)^2[3(x+2)+2(x+5)]$$

$$= (x+2)(x+5)^2[3x+6+2(x+5)]$$

$$= (x+2)\ (x+5)^2[3x+6+2x+10]$$

$$= (x+2)(x+5)^2[5x+16]$$

$$= (x+2)(x+5)^2[5x+16]$$

Example 4: $\frac{d}{dx}\left(\frac{5x^2+1}{x}\right)$

Sol. Apply derivative of a quotient of two functions.

$$= \frac{x\frac{d}{dx}(5x^2+1)-(5x^2+1)\frac{d}{dx}(x)}{x^2} = \frac{x(10x)-(5x^2+1)(1)}{x^2}$$

$$= \frac{10x^2-(5x^2+1)}{x^2} = \frac{(5x^2-1)}{x^2} = 5-\frac{1}{x^2}$$

Example 5: Differentiate the following w.r.t. to x

$$y = \frac{1-\sqrt{x}}{1+\sqrt{x}}$$

Sol. $\frac{dy}{dx} = \frac{d}{dx}\left[\frac{1-\sqrt{x}}{1+\sqrt{x}}\right]$

$$= \frac{\left(1-\sqrt{x}\right)\frac{d}{dx}\left(1-\sqrt{x}\right)-\left(1-\sqrt{x}\right)\frac{d}{dx}\left(1+\sqrt{x}\right)}{\left(1+\sqrt{x}\right)^2}$$

$$= \frac{(1+x)\left(\frac{-1}{2}\right)x^{-1/2} - (1-\sqrt{x})\frac{1}{2}x^{-1/2}}{(1+\sqrt{x})^2} = \frac{\frac{-(1+\sqrt{x})}{2x^{1/2}} - \frac{(1-\sqrt{x})}{2x^{1/2}}}{(1+\sqrt{x})^2}$$

$$= \frac{\frac{-1-\sqrt{x}-1+\sqrt{x}}{2\sqrt{x}}}{(1+\sqrt{x})^2} = \frac{\frac{-2}{2\sqrt{x}}}{(1+\sqrt{x})^2} = \frac{-1}{\sqrt{x}(1+\sqrt{x})^2}$$

$$\frac{dy}{dx} = \frac{-1}{\sqrt{x}(1+\sqrt{x})^2}$$

6.5. DERIVATION OF AN INVERSE FUNCTION

Suppose $y = x^2$ is a single valued function of x. Then $x = \sqrt{y}$ is the inverse function of the first. Similarly assume

$$y = k+x \qquad \text{... (1)}$$

from (i)

$$x = y-k \qquad \text{... (2)}$$

Function (2) is the inverse function of (1) also function —(1) is the inverse function of (2). If both functions are single valued, then the derivative of an inverse function is the reciprocal of the derivative of the original function. In symbols, the theorem may be stated as

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

Suppose we have

$$y = x^2 \qquad \text{... (3)}$$

From (iii) $\frac{dy}{dx} = 2x$

The inverse of (3) is

$$x = \sqrt{y} \qquad \text{... (4)}$$

$$\frac{dx}{dy} = \frac{1}{2\sqrt{y}} = \frac{1}{2x}$$

This shows that the derivative of an inverse function is equal to the reciprocal of the derivation of the original function $2x$. Thus, if both the inverse functions are single valued, and if we know the derivative of the original function, then the derivative of the inverse function can be obtained at once.

The product of the derivatives of both the inverse function is unity.

$$\frac{dy}{dx} \cdot \frac{dx}{dy} = 1$$

Example 1: Find $\frac{dy}{dx}$ when $x^2 + y^2+2y = 20$

Sol. We have

$$x^2 + y^2+2y = 20$$

Diff. both sides w.r.t. to x, regarding y as a function of x, we get

$$2x+2y\frac{dy}{dx}+\frac{2dy}{dx} = 0$$

$$x + y\frac{dy}{dx}+\frac{dy}{dx} = 0$$

$$x+(1+y)\frac{dy}{dx} = 0$$

$$\therefore \qquad \frac{dy}{dx} = \frac{-x}{1+y}$$

If we consider x as a function of y, then $\frac{dx}{dy}$ can be obtained by differentiating the function w.r.t. y, Thus

$$2x.\ \frac{dx}{dy} + 2y\ + 2 = \ 0$$

$$x.\frac{dx}{dy} + y + 1 \ = \ 0$$

$$\frac{dx}{dy} \ = \ \frac{-(1+y)}{x}$$

This shows

$$\frac{dy}{dx} \cdot \frac{dx}{dy} \ = \ 1$$

Parametric Equations and their Derivatives

If x and y are both functions of a third variable, say t, then the equations are called Parametric equations. the variable t is called parameter

Let $\quad x = f(t) \quad$ ..... (1)

$\quad y = \Psi\,(t) \quad$ ..... (2)

Equation (1) and (2) are known as Parametric equations

If $x = f(\mathrm{t})$ and $y = \psi(t)$

$$\text{Then} \quad \frac{dy}{dx} = \frac{dy/dt}{dx/dt} \ = \ \frac{\psi'(t)}{f'(t)}$$

Example 1: Find $\frac{dy}{dx}$ when

$$x \ = \ \frac{3at}{1+t^2}\,,\ y \ = \ \frac{3at^2}{1+t^3}$$

$$\frac{dx}{dt} \ = \ \frac{(1+t)^2(3a) - 3at(2t)}{(1+t^2)^2}\,,$$

$$\frac{dy}{dt} = \frac{(1+t^3)(6at)-3at^2(3t^2)}{(1+t^3)^2}$$

$$= \frac{3a+3at^2-6at^2}{(1+t^2)^2}, \quad \frac{6at+6at^4-9at^4}{(1+t^3)^2}$$

$$= \frac{3a-3at^2}{(1+t^2)^2}, \quad \frac{6at-3at^4}{(1+t^3)^2}$$

$$= \frac{3a(1-t^2)}{(1+t^2)^2}, \quad \frac{3at(2-t^3)}{(1+t^3)^2}$$

$$\therefore \qquad \frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{3at(2-t^3)}{(1+t^3)^2} \times \frac{(1+t^2)^2}{3a(1-t^2)}$$

$$\frac{dy}{dx} = \frac{t(2-t^3)(1+t^2)^2}{(1+t^3)^2}$$

Example 2: Find $\frac{dy}{dx}$, when $x = at^3,\ y = 3at$.

Sol. As $\quad x = at^3$, *diff.* w.r.t. to t

$$\frac{dx}{dt} = \frac{d}{dt}(at^3) = 3at^2 \qquad \left[\because \frac{d}{dx}x^n = nx^{n-1}\right]$$

As $\quad y = 3at$, Diff. w.r.t. to t.

$$\frac{dy}{dt} = \frac{d}{dt}(3at) = 3a \qquad \left[\because \frac{d}{dx}x = 1\right]$$

Now $\quad \frac{dy}{dx} = \frac{dy}{dt} \Big/ \frac{dx}{dt} = \frac{3a}{3at^2} = \frac{1}{t^2}$

Example 3: If $x = \frac{a(1-t^2)}{1+t^2}$, $y = \frac{2bt}{1+t^2}$, find $\frac{dy}{dx}$.

Sol. As $x = \dfrac{a(1-t^2)}{1+t^2}$, Diff. w.r. to t.

$$\frac{dx}{dt} = \frac{(1+t^2)\frac{d}{dt}a(1-t^2) - a(1-t^2).\frac{d}{dt}(1+t^2)^{\dagger}}{(1+t^2)^2}$$

$$\left[{}^{\dagger}\text{Using } \frac{d}{dt}\frac{u}{v} = \frac{v\frac{d}{dt}u - u\frac{d}{dt}v}{v^2}\right]$$

$$= \frac{(1+t^2)a(-2t) - a(1-t^2)2t}{(1+t^2)^2} = \frac{-4at}{(1+t^2)^2}$$

Again $y = \dfrac{2bt}{1+t^2}$, Diff. w.r. to t

$$\frac{dy}{dt} = \frac{(1+t^2)\frac{d}{dt}(2bt) - 2bt\frac{d}{dt}(1+t^2)^{\dagger}}{(1+t^2)^2}$$

$$= \frac{(1+t^2)2b - 2bt(2t)}{(1+t^2)^2} = \frac{2b-2bt^2}{(1+t^2)^2}$$

As $$\frac{dy}{dx} = \frac{dy}{dt}\bigg/\frac{dx}{dt} = \frac{2b(1-t^2)}{(1+t^2)^2}\bigg/\frac{(-4at)}{(1+t^2)^2} = \frac{-b(1-t^2)}{2at}$$

Example 4. Differentiate w.r.to x

$y = \log(x^2+1)$

$$\frac{dy}{dx} = \frac{1^2}{(x^2+1)}\frac{d}{dx}(x^2+1) = \frac{2x}{x^2+1}$$

$$\frac{dy}{dx} = \frac{2x}{x^2+1}$$

$$\because \frac{d}{dx}\log x = \frac{1}{x}\frac{d}{dx}(x)$$

$$= \frac{1}{x}(1) = \frac{1}{x}$$

Characteristic of Logarithms

A characteristic of logarithms is that products become sums and quotients become difference. That is,

$$\log ab = \log a + \log b$$

$$\log\left(\frac{a}{b}\right) = \log a - \log b$$

Using this characteristics, we can differentiate complicated function of products and quotients by first taking logarithms to the base of the function.

Example: Find $\frac{dy}{dx}$, when $y = (x + 2)(x+4)$

Solution: $y = (x+2)(x+4)$

Taking log of both sides,

$$\log y = \log (x+2)(x+4) = \log (x + 2) + \log (x + 4)$$

$$\frac{d}{dx}\log y = \frac{d}{dx}\log(x+2) + \frac{d}{dx}\log(x+4)$$

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{x+2} + \frac{1}{x+4} = \frac{x+4+x+2}{(x+2)(x+4)}$$

$$\frac{dy}{dx} = y\frac{2x+6}{(x+2)(x+4)} = (x+2)(x+4)\frac{2x+6}{(x+2)(x+4)} = 2x + 6$$

Derivative of Exponential Functions

Case I: Let an exponential function be

$$y = a^x$$

To find the derivative of y with respect to x, we apply the technique of taking logarithms to the base and then differentiating we have

$$\log y = \log a^x$$

$$= x \log a \quad (\because \log m^n = n \log m)$$

$$\frac{dy}{dx}(\log y) = \log a$$

$$\frac{1}{y}\frac{dy}{dx} = \log a$$

$$\frac{dy}{dx} = y \log a$$

Example 1: $y = a^{x^2}$

Taking logarithms of both sides, we get

$$\log y = \log a^{x^2}$$

$$\log y = x^2 \log a$$

Diff. w.r. to x

$$\frac{1}{y}\frac{dy}{dx} = 2x \log a$$

$$\frac{dy}{dx} = y.2x \log a$$

$$= 2x.\ a^{x^2} \log a$$

Example 2: $y = a^{x^2+x}$

$$\log y = \log a^{x^2+x}$$

$$\log y = (x^2+x)\log a$$

$$\frac{1}{y}\frac{dy}{dx} = (2x+1)\log a$$

$$\frac{dy}{dx} = y(2x+1)\log a$$

$$\frac{dy}{dx} = a^{x^2+x}(2x+1)\log a$$

Example 3 : $y = x^x$

Sol. Taking logarithms of both sides

$$\log y = \log x^x$$

$$\log y = x.\log x$$

$$\frac{dy}{dx}\log y = x\log x + \log x.\frac{d}{dx}(x)$$

$$\frac{1}{y}\frac{dy}{dx} = x\frac{1}{x} + \log x.1$$

$$= 1+\log x$$

$$\frac{dy}{dx} = y(1+\log x) \quad = x^x (1 + \log x)$$

Since $\quad 1 = \log e$

$$\frac{dy}{dx} = x^x (\log e + \log x)$$

$$\frac{dy}{dx} = x^x \log ex (\log m + \log n = \log mn)$$

Example 4: $y = e^{x^2}$

$$\frac{dy}{dx} = e^{x^2}\frac{d}{dx}(x^2)$$

$$\frac{dy}{dx} = 2xe^{x^2}$$

Typical Illustrations

Example 1: $y = \log\left(x+\sqrt{x^2+c^2}\right)$

$$\frac{dy}{dx} = \frac{1}{x+\sqrt{x^2+c^2}}\frac{d}{dx}\left(x+(x^2+c^2)^{½}\right)$$

$$= \frac{1}{x+\sqrt{x^2+c^2}}\left[1+\frac{1}{2}\left(x^2+c^2\right)^{-½}(2x)\right]$$

$$= \frac{1}{x+\sqrt{x^2+c^2}}\left[1+\frac{x}{\sqrt{(x^2+c^2)}}\right]$$

$$= \frac{1}{x+\sqrt{x^2+c^2}}\left[\frac{\sqrt{x^2+c^2}+x}{\sqrt{(x^2+c^2)}}\right]$$

$$\frac{dy}{dx} = \frac{1}{\sqrt{(x^2+c^2)}}$$

Example 2: If $y\sqrt{x^2+1} = \log\left(x+\sqrt{x^2+1}\right)$

Prove that $(x^2+1)\frac{dy}{dx}+xy-1=0$

Sol. Given $y\sqrt{x^2+1} = \log\left(x+\sqrt{x^2+1}\right)$

Differentiating both sides w.r.to x we have

$$y\frac{1}{2}(x^2+1)^{½-1}(2x)+\sqrt{x^2+1}\frac{dy}{dx} = \frac{1}{x+\sqrt{x^2+1}} - \frac{d}{dx}\left(x+\sqrt{x^2+1}\right)$$

$$\frac{y}{2}(x^2+1)^{-½}(2x)+\sqrt{x^2+1}\frac{dy}{dx} = \frac{1}{x+\sqrt{x^2+1}} - \left[1+\frac{1}{2}\left(x^2+1\right)^{½-1}.2x\right]$$

$$\frac{dy}{dx} = \frac{(1-xy)}{x^2+1)}$$

Taking L.H.S.

$$(x^2+1)\frac{(1-xy)}{x^2+1)} + xy - 1 = 0$$

$$(1–xy) + xy –1 = 0$$

L.H.S. = R.H.S.

$$(x^2+1)\frac{dy}{dx} + xy - 1 = 0$$

Check Your Progress–IV

Q.1. Find derivative of y w.r. to x, when

(i) $2y = 1 - z^2$, $z = x^2+3$

(ii) $y = \dfrac{3-u}{2+u}$, $u = \dfrac{4x}{1-x^2}$

Q.2. If $y = \dfrac{2(x+1)}{x^2+2x-3}$, $x = 2t^2 + 3t$, obtain $\dfrac{dy}{dt}$.

Q.3. Find $\dfrac{dy}{dx}$, when :

(i) $x = 4t^2 + 3t + 1, y = 7t - 1$

(ii) $x = \dfrac{2t+3}{t^2-1}$, $y = \dfrac{3t-2}{t^2-1}$

(iii) $x = t^4 + 3t^2 + t, \;\; y = 7t^2 + 6t - 9$

(iv) $x = \dfrac{2at}{1+t^2}, \; y = \dfrac{2b(1-t^2)}{1+t^2}$

(v) $x = \dfrac{at^2}{1+t^2}, \;\; y = \dfrac{2bt}{1+t^2}$

6.6. DIFFERENTIATION OF IMPLICIT FUNCTION

EXERCISE

Q.1. (a) Find $\dfrac{dy}{dx}$, when $x^2y+xy^{2} = 25$

(b) Find $\dfrac{dy}{dx}$ of the following:

(i) $x^3+y^3+\log\, y - e^x = 0$

(ii) $x^a+a^x+x^x+a^a-y = 0$

(iii) $f(x,y) = x^3-2x^2y+3xy^2-25 = 0$

(iv) $x^2+y^2+2y = 15$

(v) $x^3+y^3 = 3xy$

(vi) $x^2+y^2 = 25$

(vii) $x^2+3xy+2y^2 = 6$

Ans. (a) $\dfrac{dy}{dx} = \dfrac{-(y^2+2xy)}{x^2+2xy}$

(b) (i) $x^3+y^3+\log y - e^x = 0$

Differentiate w.r.t. to x

$$3x^2+3y^2\frac{dy}{dx}+\frac{1}{y}\frac{dy}{dx}-e^x=0,\ \frac{dy}{dx}\left(3y^2+\frac{1}{y}\right)=e^x-3x^2$$

$$\therefore \qquad \frac{dy}{dx}=\frac{e^x-3x^2}{3y^2+1/y}$$

(ii) and (iii) can be solved similarly

Ans. (ii) $ax^{n-1}+a^x \log a + x^x(1+\log x)$

(ii) $\dfrac{4xy-3y^2-3x^2}{6xy-2x^2}$

(v) $\dfrac{y-x^2}{y^2-x}$

Q.2. (a) Find $\dfrac{dy}{dx}$, when $y = \sqrt{x+\sqrt{x+\sqrt{x+\ldots\ldots\infty}}}$

(b) Find $\dfrac{dy}{dx}$, when (i) $ax^2+2hxy+by^2+2gx+2fx+x = 0$

(ii) $x^2y = 75$ (ii) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=2$

(iv) Find $\dfrac{dy}{dx}$, when $y = x+\cfrac{1}{x+\cfrac{1}{x+\cfrac{1}{x+\ldots\ldots\infty}}}$

Sol. (a) $y = \sqrt{x+\sqrt{x+\sqrt{x+\ldots\ldots\infty}}}$ squaring

$y^2 = x + \sqrt{x + \sqrt{x + \ldots \propto}} = x + y$ Differentiate w.r.to x

$$2y\frac{dy}{dx} = 1 + \frac{dy}{dx} \; or \; (2y-1)\frac{dy}{dx} = 1, \frac{dy}{dx} = \frac{1}{2y-1}$$

(b) (i) $ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$, Differentiate w.r. to x

$$a\frac{dy}{dx}(x)^2 + 2h\frac{d}{dx}(xy) + b\frac{d}{dx}(y)^2 + 2g\frac{d}{dx}(x) + 2f\frac{d}{dx}(y) + \frac{d}{dx}(1) = 0$$

$$2ax + 2h(x\frac{d}{dx} + y) + 2by\frac{dy}{dx} + 2.g.1. + 2f\frac{dy}{dx} = 0$$

$\therefore$ $$\frac{dy}{dx}(2hx + 2by + 2f) = -(2ax + 2hy + 2g)$$

$\therefore$ $$\frac{dy}{dx} = \frac{-ax + hy + g}{hx + by + f}$$

(ii) Can be solved similarly

Ans. (ii) $\frac{-2y}{x}$

(iii) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $b^2x^2 + a^2y^2 = a^2b^2$

Differentiate both sides w.r. to x

$$b^2(2x) + a^2.2\frac{dy}{dx} = \frac{-b^2x}{a^2y}$$

(iv) Can be solved similarly.

6.7. DERIVATIVES OF HIGHER ORDER

If we want to find the rate of change in y due to a small change in x, we find $\frac{dy}{dx}$.

For example, for the functional relationship,

$$y = 2 + 4x^4, \; \frac{dy}{dx} = 16x^3$$

Now, the expression $\frac{dy}{dx} = 16x^3$ can be differentiated again which will give,

$$\frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dx}(16x^3) = 48x^2$$

This result is said to be second differential of y, is written as $\frac{d^2y}{dx^2}$ or $f''(x)$ where as the first differential is written as $f'(x)$.

Now, since $\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right)$, $\frac{d^2y}{dx^2}$ is the rate of change of $\frac{dy}{dx}$ and is therefore, rate of change of rate of change.

For example, if y is income and x is time then $\frac{dy}{dx}$ is rate of change of income over time, and $\frac{d^2y}{dx^2}$ tells whether this rate of change (or slope) is increasing or decreasing or stationary.

By successive differentiation, we can find third $\frac{d^3y}{dx^3}$, fourth $\frac{d^4y}{dx^4}$ differentials and so on.

Hence given: $y = 2 + 4x^4$

$$\frac{dy}{dx} = \frac{d}{dx}(2 + 4x^4) = 16x^3$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx}(16x^3) = 48x^2$$

$$\frac{d^3y}{dx^3} = \frac{d}{dx}(48x^2) = 96x$$

$$\frac{d^4y}{dx^4} = \frac{d}{dx}(96x) = 96$$

$$\frac{d^5y}{dx^5} = \frac{d}{dx}(96) = 0$$

6.8. SIGN OF THE DIFFERENTIAL COEFFICIENT

If at any point,

(1) $\frac{dy}{dx} > 0$, y is increasing as x rises; and y is decreasing as x declines. If this is so for all points then y is said to be monotonically increasing function of x. When plotted on a graph paper, the curve rises from left to right around the point.

(2) $\frac{dy}{dx} = 0$, y does not change at all with x. The curve is either a straight line parallel to axis of x or the slope of the curve at that point is zero. Thereore this point is either a point of minima or maxima or inflextion.

(3) $\frac{dy}{dx} < 0$, y is increasing as x falls and vice-versa. If this is so, for all points for which the function has been defined, then y is said to be monotonically decreasing function of x. The curve falls from left to right.

Now look at the second derivative

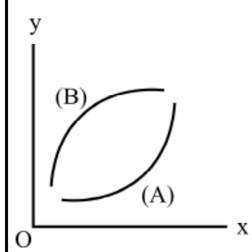
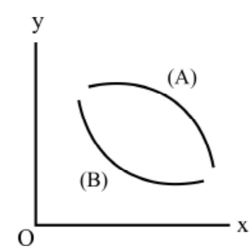
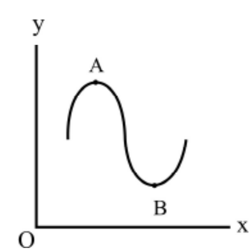
If $\frac{d^2y}{dx^2} > 0$; The slope of the curve at the point increases.

If $\frac{d^2y}{dx^2} < 0$, The slope of the curve at the point decreases.

6.9. NATURE OF THE CURVE AND ORDER CONDITIONS

With help of the first and second derivatives (called first and second order conditions) of the given function, we are able to examine the nature of the curve which that function gives rise to. Without going in to proofs, We present the order conditions along with the nature of the curve in the tabular form below:

Suppose the given function is $y = f(x)$

First Order Curve **Condition**	**Second Order** **Condition**	**Nature of** **the Curve**	**Shape of the**
$\frac{dy}{dx} > 0$	$\frac{d^2y}{dx^2} > 0$ (A) $\frac{d^2y}{dx^2} < 0$ (B)	Concave upward Concave downward (Rising in both cases)	
$\frac{dy}{dx} < 0$	$\frac{d^2y}{dx^2} > 0$ (A) $\frac{d^2y}{dx^2} < 0$ (B)	Concave upwards Concave downwards (Falling in both cases)	
$\frac{dy}{dx} = 0$	$\frac{d^2y}{dx^2} > 0$ (A) $\frac{d^2y}{dx^2} < 0$ (B)	Minimum point on the curve Maximum point on the curve	

6.10. SUCCESSIVE DERIVATIVE

Q.1. (a) Find third derivative of $y = 4x^5-2x^4+x^3+7x^2-4x+5$

(b) Find the second derivative of $y = x^2e^{-x}$

Sol. (a) $y = 4x^5-2x^4+x^3+7x^2-4x+5$, differentiate w.r. to x:

$$\frac{dy}{dx} = 20x^4-8x^3+3x^2-14x-4, \text{ Differentiate w.r. to } x$$

$$\frac{d^2y}{dx^2} = 80x^3-24x^2-6x+14, \text{ Differentiate w.r. to } x.$$

$$\frac{d^3y}{dx^3} = 240x^2-48x+6$$

(b) $y = x^2e^{-x}$, Differentiate w.r. to x

$$\frac{dy}{dx} = x^2\frac{d}{dx}e^{-x}+e^{-x}\frac{d}{dx}(x^2)$$

$= x^2e^{-x}(-1)+e^{-x}2x = x^2e^{-x} + e^{-x}\ 2x$

$= xe^{-x}(2-x)$, Differentiate again w.r. to x

$$\frac{d^2y}{dx^2} = xe^{-x}\frac{d}{dx}(2-x)+(2-x)\frac{d}{dx}xe^{-x}$$

$= xe^{-x}(-1) + (2-x)\ [xe^{-x}(-1)+e^{-x}\cdot 1]$

$= -xe^{-x} + (2-x)\ (-xe^{-x}+e^{-x})$

$= -\ 4xe^{-x} + x^2e^{-x} + 2e^{-x}$

$= e^{-x}(x^2 - 4x + 2)$

Q.2. Find third derivative of following w.r.to x

(i) $x \log x$ (ii) $x^3 \log x$

(iii) x^2e^{2x} (iv) $(3-2x)^5$ at $x = 0$

Sol. (i) Let $y = x \log x$ Differentiate w.r. to x

$$\frac{dy}{dx} = x.\frac{1}{x}+\log x.\ 1 = 1 + \log x, \text{ differentiate again w.r.to } x$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx}(1+\log x) = \frac{1}{x} = x^{-1}$$

$$\frac{d^3y}{dx^3} = \frac{d}{dx}(x^{-1}) = -x^{-2}$$

(ii) Let $y = x^3 \log x$, Differentiate w.r.to to x

$$\frac{dy}{dx} = x^3 \frac{d}{dx}\log x + \log x \frac{d}{dx}x^3$$

$$= x^3.\frac{1}{x} + \log x(3x^2)$$

$$= x^2 + 3x^2 \log x$$

$$\frac{d^2y}{dx^2} = 2x + 6x\log x + 3x^2.\frac{1}{x}$$

$$= 2x+6x\log x +3x$$

$$= 5x+6x \log x$$

$$\frac{d^3y}{dx^3} = 5 + 6x\frac{1}{x} + 6\log x.1 = 11+6\log x$$

(iii) Let $y = x^2e^{2x}$ Differenetiate w.r.to x

$$\frac{dy}{dx} = x^2 \frac{d}{dx}e^{2x} + e^{2x}\frac{d}{dx}x^2$$

$$= x^2e^{2x}.2x = \mathrm{e}^{2x}(2x^2+2x)$$

Differentiate again w.r. to x.

$$\frac{d^2y}{dx^2} = e^{2x}\frac{d}{dx}(2x^2+2x) + (2x^2+2x)\frac{d}{dx}e^{2x}$$

$$= e^{2x}(4x+2)+(2x^2+2x)\mathrm{e}^{2x}.2$$

$$\frac{d^3y}{dx^3} = e^{2x}\frac{d}{dx}(4x^2+8x+2)+(4x^2+8x+2)\frac{d}{dx}e^{2x}$$

$$= e^{2x}(8x+8) + (4x^2 + 8x + 2)e^{2x}\cdot 2$$

(iv) Let $y = (3–2x)^5$ at $x = 0$, Differentiate w.r. to x

$$\frac{dy}{dx} = 5(3\text{-}2x)^3(–2), \text{ Differentiate again w.r. to } x.$$

$$\frac{dy}{dx} = 20 \ (3–2x)^3(–2)^2 \text{ Differentiate again w.r. to } x$$

$$= 480 \ (3 - 2x)^2$$

at $x = 0$, $\frac{dy}{dx} = 5\times 3^3(–2)$, $\frac{d^2y}{dx^2} = 20\times 3^3(4)$

$$\frac{d^3y}{dx^3} = 60\times 3^2(–8)$$

6.11. PARTIAL AND TOTAL DEFFERENTIATION

Till now we have considered functions of only one variable. However, in Economics, in most of the relations, more than one variable are involved. For example, utility (U) derived from consumption depends on amounts of several commodities (x_1, x_2 ....) one has. Thus,

$$U = f\ (x_1,\ x_2\ \dots\ x_n)$$

Following aspects need to be studied in such functions:

1. If the quantity of only one commodity x_1 change while those of all other commodities ($x_2, x_3 \dots x_n$) remain uncharged then how will utility change, that is what will be the marginal utility (MU) of x_1?

2. Similarly, we may like to find out the MU of other commodities.

3. If the quantities of all the commodities are changed then what will

happen to total utility?

In (1) and (2) we shall have to calculate partial differential and in (3) total differential.

The marginal utility of x_1 will be $\frac{\partial y}{\partial x_1}$ and that of x_2 $\frac{\partial y}{\partial x_2}$. These are termed as partial differential coefficients with respect to x_1 and x_2 respectively.

6.12. THE METHOD OF CALCULATING PARTIAL DIFFERENTIAL COEFFICIENTS

The method is same as that of differentiation in the case of one variable. Hence we assume that all variables other than one are considered to be constants.

For example, let us partially differentiate

$$z = 5x^3 - 2x^2y + 4xy^2$$

with respect to x and y

(i) Assuming y to be constant

$$\frac{\partial z}{\partial x} = 5(3x^2) - 2y\ (2x) + 4y^2\ (1)$$

$$= 15x^2 - 4xy + 4y$$

(ii) Assuming x to be constant:

$$\frac{\partial z}{\partial y} = 0 - 2x^2(1) + 4x\ (2y)$$

$$= -2x^2 + 8xy$$

Example 1: Partially differentiate $z = \frac{e^x}{1-y}$

$$\frac{\partial z}{\partial x} = \frac{1}{(1-y)}\frac{\partial}{\partial x}\left(e^x\right) = \left(\frac{1}{1-y}\right)e^x$$

$$\frac{\partial z}{\partial x} = e^x\frac{\partial}{\partial y}\left(\frac{1}{1-y}\right)$$

$$= \quad e^x \frac{1}{(1-y)^2} \quad = \quad \frac{e^x}{(1-y)^2}$$

Example 2: Find partial differential coefficients of first and second order :

If $z \ = \ x^2 \ . \ e^{2y}$

$$\frac{\partial z}{\partial x} \quad = \quad e^{2y}\frac{\partial}{\partial x}(x^2) = e^{2y} \cdot 2x = 2xe^{2y}$$

$$\frac{\partial z}{\partial x} \quad = \quad x^2 \frac{\partial}{\partial y}(e^{2y}) = x^2 \cdot 2e^{2y} = 2x^2 e^{2y}$$

and $$\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) \quad = \quad \frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}(2xe^{2y}) = e^{2y}(2) = 2e^{2y}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) \quad = \quad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial y}(2xe^{2y}) = 2x(e^{2y} \times 2) = 4xe^{2y}$$

Similarly,

$$\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) \quad = \quad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}(2x^2.e^{2y}) = e^{2y}(4x) = 4xe^{2y}$$

and $$\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) \quad = \quad \frac{\partial^2 z}{\partial y^2} \quad = \quad \frac{\partial}{\partial y}(2x^2.e^{2y}) = 2x^2(2e^{2y}) = 4x^2 e^{2y}$$

It is to be noted that $\dfrac{\partial z}{\partial x \partial y} \quad = \quad \dfrac{\partial z}{\partial y \partial x}$

Example 3: $z \ = \ 2x^3 + 5x^2 y + xy^2 + y^2$. Find the second order partial differential coefficients.

Symbol	*Alternative symbol*	*Partial differential coefficient*
$\frac{\partial z}{\partial x}$	$= fx$	$= 6x^2 + 10xy + y^2$
$\frac{\partial^2 z}{\partial x^2}$	$= fx.x$	$= 12x + 10y$
$\frac{\partial^2 z}{\partial x \partial y}$	$= fxy$	$= 10x + 2y$

Similarly, $z = 2x^3 + 5x^2y + xy^2 + y^2$.

$\frac{\partial z}{\partial y}$	$= fy$	$= 5x^2 + 2xy + 2y$
$\frac{\partial^2 z}{\partial y \partial x}$	$= fyx$	$= 10x + 2y$
$\frac{\partial^2 z}{\partial y^2}$	$= fyy$	$= 2x + 2$
$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$	$= fxy = fyx$	$= 10x + 2y$

Total Derivative

Definition: If $z = f(x, y)$ then the total differential dz of z is defined as:

$$dz = fy.dx + fy.dx = \frac{\partial z}{\partial x}dx + \frac{\partial x}{\partial y}dy$$

Similarly if $z = f(x_1, x_2 \ldots \ldots x_n)$

$$dx = f_1 dx_2 + f_2 dx_2 + \ldots \ldots + f_n dx_n$$

Example 1: If $z = 3x^2 + xy - 2y^3$ then

(1) $fx = 6x + y$

(2) $fy = x – 6y^2$

$\therefore \quad dz = (6x + y)dx + (x - 6y^2)dy$

Example 2: If $z = \sqrt{x+y}$ find dz

Sol. $$dz = d(x+y)^{\frac{1}{2}} = \frac{1}{2}(x+y)^{\frac{-1}{2}} d(x+y) = \frac{1}{2\sqrt{x+y}}(dx+dy)$$

Example 3: If $z = \sqrt{3x^2 - y^2}$ find dz

$$dz = d(3x^2 - y^2)^{\frac{1}{2}} = \frac{1}{2} \times \frac{1}{\sqrt{3x^2 - y^2}} d(3x^2 - y^2)$$

$$= \frac{1}{2} \times \sqrt{3x^2 - y^2}\ (3xdx - 2ydy) = \frac{3xdx - 2ydy}{2\sqrt{3x^2 - y^2}}$$

Example 4: If U = $q_1^{\frac{1}{2}} q_3^{\frac{1}{2}}$ is the utility function, find the Marginal Substitution rate between the two commodities.

Sol. Now, total utility remains unchanged on an indifference curve, that is dU = 0. Hence, we have to find the total differential of U.

$$dU = d\left[q_1^{\frac{1}{2}} q_2^{\frac{1}{3}}\right] = 0$$

$$= q_2^{\frac{1}{3}}\left[\frac{1}{2} q_1^{\frac{-1}{2}} dq_1\right] + q_1^{\frac{1}{2}}\left[\frac{1}{3} q_3^{\frac{2}{3}} dq_2\right] = 0$$

$$= \frac{1}{2} q_2^{\frac{1}{3}} q_1^{\frac{-1}{2}} dq_1 - \frac{1}{3} q_1^{\frac{1}{2}} q_3^{\frac{-2}{3}} dq_2 = 0$$

Then MRS $$\frac{dq_2}{dq_1} = \frac{\frac{-1}{2} q_2^{\frac{1}{3}}\ q_1^{\frac{-1}{2}}}{\frac{1}{3} q_1^{\frac{1}{2}}\ q_2^{\frac{-2}{3}}}$$

$$= \frac{-3}{2} q_1^{-1} q_2^{1} = \frac{-3}{2}\left(\frac{q_2}{q_1}\right)$$

6.13. APPLICATION OF DIFFERENTIATION IN ECONOMICS

In this section our aim is to express some important concepts in economic theory as derivatives i.e. to interpret the derivatives with reference to some economic relation. Let us take the concept of marginal utility, marginal cost etc. This is an important and fundamental concept in economic theory. This concept is involved whenever we have to determine the rate of change in one variable with respect to some other variable: for example the rate of change in income with respect to investment or rate of change in demand with respect to price or the rate of change in output with respect to some input such as labour or capital etc. However these concepts are discussed in more detail in the next two lessons.

Cost Curves

Let us assume π is total cost and q is the quantity of output produced and also that the cost function is given by

$$\pi = a+bq+cq^2$$

where $a,$ b and c are constants.

(a) Marginal Cost (MC):

Marginal cost is defined as $\frac{d\pi}{dq}$

Thus we have

$$\frac{d\pi}{dq} = b+2cq$$

(b) Average Cost (AC):

Average cost is shown by

$$\frac{\pi}{q} = \frac{a}{q}+b+cq$$

Relation between average and marginal cost

The slope of the average cost curve is obtained by finding the derivation of $\frac{\pi}{q}$. Thus we can use the formula for the quotient :

$$\frac{d}{dq}\left(\frac{\pi}{q}\right) = \frac{q\frac{d}{dq}(\pi)-\pi\frac{d}{dq}(q)}{q^2} = \frac{q\frac{d\pi}{dq}-\pi}{q^2} = \frac{\frac{d\pi}{dq}}{q}-\frac{\pi}{q^2}$$

$$= \frac{1}{q}\left[\frac{d\pi}{dq} - \frac{\pi}{q}\right] = \frac{1}{q}[\text{MC} - \text{AC}]$$

From the above relationship/result of the slope of AC, we have the following relationship:

(a) When average cost curve slopes downwards, its slope will be negative i.e.

$$\frac{d}{dq}\left(\frac{\pi}{q}\right) < 0 \quad \text{or} \quad \text{MC} - \text{AC} < 0 \quad \text{or} \quad \text{MC} < \text{AC}$$

It implies that when AC curve slopes downwards MC will lie below it.
(b) When AC curve reaches a minimum point, its slope become zero i.e.

$$\frac{d}{dq}\left(\frac{\pi}{q}\right) = 0 \quad \text{or} \quad \text{MC} - \text{AC} = 0 \quad \text{or} \quad \text{MC} = \text{AC}$$

It implies that MC cuts AC curve at the minimum point of AC curve.
(c) When AC curve rises upwards, its slope is positive i.e.

$$\frac{d}{dq}\left(\frac{\pi}{q}\right) > 0 \quad \text{or MC} > \text{AC}$$

It implies that when AC curve slopes upwards, MC will lie above AC curve.

Uses of Differentiation in Economics

The uses of differentiation in Economics can be discussed under following points :

1. A derivative at a point in a curve can be viewed as the slope of the line tangent to that curve at that point.
2. Derivative can be used to calculate instantaneous rate of change.
3. Derivative can also be used to understand the shape of the function (concave upwards or downward convex).
4. It can also help to solve optimisation problems.
5. Derivatives or change in one variable relative to the change in another are identical to the economic concept of Marginalism which examines the

change in one variable due to change in another variable. Marginal means small incremental changes such as incremental changes in work hours or factor output.

6. By calculating the Marginal Revenue and Marginal Costs, business managers can maximise their profits and measure the rate of increase in profit which results from increase in production.
7. Marginal concepts are very important concept as all decisions are taken at the margin. With more and more production, Marginal Revenue falls and Marginal Cost rises, so for maximising profits, the producer should produce where MR = MC.
8. With derivatives we can calculate the rate of change of anything with respect to time which gives the knowledge of upcoming events and different behaviour events can present.

6.14 LET US SUM UP

In this lesson, we have discussed the concept of derivative, rules of differentiation and the application of differentiation in Economics.

6.15 SUGGESTED READINGS & REFERENCES

1. Leonard & Uon Long (1978) : Introduction to Maths for Students of Economics, Cambridge.
2. Mehta & Madnani (1992) : Mathematics for Economists, S. Chand, New Delhi
3. Monga, G.S. (1972) : Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.
4. Aggarwal, C.S. & R.C. Joshi : Mathematics for Students of Economics.

MARGINAL CONCEPTS

LESSON NO. 7 **UNIT-II**

STRUCTURE

7.1. OBJECTIVES

After going through this lesson, you shall be able to understand :

1. The concept of Marginal Utility, Marginal Propensity to Consume, Marginal Propensity to Save (MPS), Marginal Cost (MC) & Marginal Revenue (MR).
2. Price Elasticity of Demand
3. Relationshing between Average Revenue, Marginal Revenue and Elasticity of Demand
4. Increasing and Decreasing function
5. Maximum value, minimum value and Point of inflexion of a function at a point.

7.2. INTRODUCTION

Marginal is the variation in Y corresponding to a very small variation in X (X is independent and Y is the dependent variable). Mathematically, in order to convert a function into marginal function, just differentiate that function in first order. In this lesson, we shall discuss this concept in detail.

7.3. MARGINAL UTILITY

The change that takes place in total utility by consuming an additional unit of a good is known as marginal utility. In other words an addition made to total utility by consuming one more unit of a good.

i.e. $$MU = \frac{d}{dQ}(T.U)$$

For example

Total utility (TU) is $U = 9x^2+5x+4$

$$\therefore \text{Marginal utility} = \frac{d}{dx}(TU) = \frac{d}{dx}(9x^2+5x+4)$$

Marginal Utility $= 18x+5$

7.4. MARGINAL PROPENSITY TO CONSUME (MPC)

It can be defined as change in consumption due to change in income of the consumer. In other words, MPC refers to the relationship between additional income and additional consumption. It may be the ratio of change in consumption to the change in income.

$$\text{MPC} = \frac{\Delta c}{\Delta y} = \frac{dc}{dy}$$

e.g. If TC = 24 + 0.84 , find MPC

$$\text{MPC} = \frac{d}{dy}(24 + 0.84)$$

$$= 0 + .8 = 0.8$$

7.5. MARGINAL PROPENSITY TO SAVE (MPS)

It is the amount by which saving changes in response to an incremental change in disposable income. Simply, it is a relationship between marginal income and marginal saving. It is a ratio of change in saving to change in income.

$$\therefore \quad \text{MPC} = \frac{dS}{dy}$$

For example if S = 50 + 0.64

$$\therefore \quad \text{MPS} = \frac{dS}{dy} = \frac{d}{dy}(50 + 0.64)$$

$$\text{MPS} = 0 + 0.6 = 0.6$$

7.6. MARGINAL COST (MC)

MC is an additional cost which is caused by producing an extra unit of output. In other words, it is ratio of change in total cost to the change in output.

$$\text{MPC} = \frac{dQ}{dy}(\text{TC})$$

If TC = $x^2 - 3x + 10$

$$\therefore \qquad \text{MPC} = \frac{d}{dQ}(x^2 - 3x + 10)$$

$$\therefore \qquad \text{MPC} = 2x - 3$$

7.7. MARGINAL REVENUE (MR)

It is the net revenue earned by selling an additional unit of the product. In other words, marginal revenue is the ratio of the change in total revenue to the change in selling one more unit of the good.

$$\text{MR} = \frac{d}{dQ}(\text{TR})$$

If TR = $50Q–4Q^2$ find MR

$$\text{MR} = \frac{d}{dQ}(\text{TR}) = \frac{d}{dQ}(50Q - 4Q^2)$$

$$\text{MR} = 50 - 8Q$$

Check Your Progress–I

Q.1. (*a*) The demand curve of a monopolist is given by $p = \dfrac{50-x}{5}$.

(*i*) Find the marginal revenue for any output x.

(*ii*) What is MR, when $x = 0$ and $x = 25$?

(*b*) Let the revenue function be given by $R=14x - x^2$ and cost function $T=x(x^2–2)$. Find AC, MC, AR and MR.

____

____

____

____

____

Q.2. (*a*) Determine price elasticity of demand and marginal revenue if $q = 30–4p^2$ where q is price and $p = 3$.

(*b*) Given the demand curve $p=400–2q–3q^2$, where p is price and

q is quantity, demonstrate the relation between marginal revenue and elasticity of demand.

Q.3. (*a*) If the total cost is represented by $y = 500+1000x-24x^2+4x^3+x^4$, where y is total cost and x is output, find average and marginal cost.

(*b*) Consider the total cost function $y = 20+20x+0.5x^2$, where y denotes total cost and x quantity produced, find the Average and Marginal Costs.

Q.4. Given the total cost function $C=Q^3-3Q^2+15Q$, find the Average Cost and Marginal Cost.

[**Hint :** $AC=\frac{C}{Q}=Q^2-3Q+15$, $MC=\frac{dC}{dQ}=3Q^2-6Q+15$]

Q.5. Given AR=$3x^2+4x+3$, find Marginal Revenue.

____

____

____

____

____

7.8. PRICE ELASTICITY OF DEMAND

Q. Derive an expression for price elasticity of demand for the demand function $q = f(P)$.

Sol. Price elasticity of demand is defined as the value of the ratio of the relative change in the demand to the relative change in price.

Precisely, If we suppose that the demand changes from q to $(q+\delta q)$ when the price changes from P to P+δP, then elasticity of demand is :

$$\eta_d = \frac{\text{Proportionate change in quantity demanded}}{\text{Proportionate change in price}}$$

$$= \frac{\frac{\delta q}{q}}{\frac{\delta P}{P}} = \frac{P}{q} \cdot \frac{\delta q}{\delta P}$$

In terms of calculus, Elasticity of demand at Price P is

$$\eta_d = \underset{\delta P \to 0}{Lt} \frac{P}{q} \frac{\Delta q}{\Delta P} = \frac{P}{q} \underset{\delta P \to 0}{Lt} \frac{\delta q}{\delta P}$$

Since according to the law of demand, price and demand move in the opposite direction thus

$$\eta_d = \frac{-P}{q} \underset{\delta P \to 0}{Lt} \frac{\delta q}{\delta P} = \frac{-P}{q} \frac{dq}{dP}$$

The absolute value of η_d is $|\eta_d| = \frac{P}{q} \frac{dq}{dP}$

Elasticity of Demand

Example: If $y = 1+2x-x^2$. Find the elasticity of y with respect to x for $x = 5$.

Sol. $$e = \left(\frac{x}{y}\right)\frac{dy}{dx}$$

Given: $$y = 1 + 2x - x^2$$

$$\frac{dy}{dx} = 2 - 2x$$

$$\therefore \quad e = \left(\frac{x}{y}\right)(2-2x)$$

By substituting for y and $x = 5$, we have

$$e = \frac{x}{(1+2x-x^2)}(2-2x)$$

$$e = \frac{5}{(1+10-25)}(2-10) = \frac{20}{7}$$

Example: Given the demand curve as P = 76 – 73D, find elasticity of demand for D = 0·75.

Sol. By definition, elasticity of demand e_d is given by

$$= \frac{P}{D}\cdot\frac{dD}{dP}$$

But $$\frac{dP}{dD} = -73$$

$$\therefore \quad \frac{d\text{D}}{d\text{P}} = \frac{1}{\frac{d\text{p}}{d\text{D}}}$$

$$e_d = \frac{P}{D}\left(\frac{1}{-73}\right)$$

To find the elasticity for D = 0·75, substitute it in P = 76 – 73D, so that P = 21·25

$$\therefore \qquad e_d = -\left(\frac{1}{73}\right)\left(\frac{21{\cdot}25}{0{\cdot}75}\right) = -\ 0{\cdot}388$$

7.9. MARGINAL REVENUE AND ELASTICITY OF DEMAND

Total revenue in defined as R = P × q (P is price per unit and q is quantity sold)

$$\therefore \qquad \text{MR} = \frac{d}{dq}(\text{R}) = \frac{d}{dq}(\text{P} \times q) = \left(\frac{dP}{dq}\right)q + P$$

This $\quad \text{MR} = \text{P} + q\left[\frac{d\text{P}}{dq}\right]$ holds true for any demand function.

7.10. RELATIONSHIP BETWEEN AR, MR AND η

If AR is the Average Revenue and MR is the Marginal Revenue at any level of output then

$$\eta = \frac{\text{AR}}{\text{AR–MR}}$$

Proof: $\quad \text{MR} = \frac{d}{dq}(R) = \frac{d}{dq}(P \times q)$

Applying product rule, we obtain

$$\text{MR} = \text{P}{\cdot}1 + q\frac{dP}{dq} = P\left(1 + \frac{q}{P}.\frac{dP}{dq}\right) \qquad \text{...(i)}$$

Since $\quad \eta = \frac{-P}{q} \cdot \frac{dq}{dP}$

$$\frac{-1}{\eta} = \frac{q}{P}.\frac{dP}{dq}$$

Hence (i), can be written as

$$\text{MR} = P\left(1 - \frac{1}{\eta}\right) = \text{AR}\left(1 - \frac{1}{\eta}\right) \qquad (\because \text{ P = AR})$$

$$1-\frac{1}{\eta} = \frac{\text{MR}}{\text{AR}}$$

$$\frac{-1}{\eta} = \frac{\text{MR}}{\text{AR}} - 1$$

$$\frac{1}{\eta} = 1 - \frac{\text{MR}}{\text{AR}}$$

$$= \frac{\text{AR–MR}}{\text{AR}}, \qquad \eta = \frac{\text{AR}}{\text{AR–MR}}$$

(a) MR > 0, $\eta > 1$

(b) MR = 0, $\eta = 1$

(c) MR < 0, $\eta < 1$

Example 1: Using P = 100 – 5 q show $\eta = \frac{\text{AR}}{\text{AR–MR}}$ at P = 20

Sol. We have, P = 100 – 5q

When P = 20 then q = 16

From P = 100 – 5q

Differentiate both side w.r. to P, we have

$$1 = -5\frac{dq}{dP}$$

$$\frac{5dq}{dP} = -1$$

$$\frac{dq}{dP} = \frac{-1}{5}$$

Now $$\eta = \frac{-P}{q}\cdot\frac{dq}{dP} = \left(\frac{-20}{16}\right)\times\left(\frac{-1}{5}\right) = \frac{1}{4}$$

Calculation of AR and MR at q = 16

$$\text{MR} = \frac{dR}{dq} = \frac{d}{dq}(100–5q)q$$

$$= \frac{d}{dq}\ (100q - 5q^2)$$

$[\because\ P = 100 - 5q;\ PQ = 100q - 5q^2;\ R = 100q - 5q^2]$

$$= 100 - 10q$$

When $q = 16$

$$MR = -60$$

$$AR = 100 - 5q = 100 - 80 = 20$$

Now putting values of AR, MR and η at

$P = 20$, $q = 16$, it can be verified that

$$\eta = \frac{AR}{AR - MR}$$

$$\frac{1}{4} = \frac{20}{20-(-60)} = \frac{20}{80} = \frac{1}{4}$$

Example 2: Calculate elasticities of demand in terms of q for the demand function $q = a - bP$.

Sol. We know the elasticity of demand is given by

$$\eta = \frac{-P}{q}\frac{dq}{dP}$$

$\therefore$ $\frac{dq}{dP} = -b$ from the given function

$$\therefore \quad \eta = \frac{-P}{q}(-b) = \frac{Pb}{q}$$

Example 3: Consider the demand law

$$P = a - bq^2$$

Find MR and show that

$$MR = P + q\frac{dP}{dq}$$

Sol. Total Revenue = Price × quantity = $(a - bq^2)q = aq - bq^3$

$$\frac{d(R)}{dq} = MR = a - 3bq^2$$

$$R = Pq$$

$$\frac{dR}{dq} = P \cdot 1 + q\frac{dP}{dq}$$

Example 4: (i) A demand function is given by $q = aP^{-b}$. Calculate price elesticity of demand.

(ii) Determine the price elasticity of demand for the demand function P = qe^q

(iii) For the demand function Q = BP^{-n}. Calculate Price elasticity of demand, comment when $n = 1$.

(iv) Find the elasticity of demand of the demand function P = 12 – 4x at P = 2

Sol. (i) $q = aP^{-b}$, Differentiate w.r. to P $\frac{dq}{dP} = a(b)P^{-b-1}$

$$\eta_d = \frac{-P}{q}\frac{dq}{dP} = \frac{-P}{aP^{-b}}[-abP^{-b-1}] = b$$

(ii) P = qe^q differentiate w.r. to q, $\frac{dP}{dq} = qe^q + e^q \cdot 1 = e^q(q+1)$

$$\eta_d = \frac{-P}{q}\frac{dq}{dP} = \frac{-P}{q} \times \frac{1}{e^q(q+1)} = \frac{-qe^q}{q} \times \frac{1}{e^q(q+1)} = \frac{-1}{(q+1)}$$

(iii) $$\eta_d = \frac{-P}{Q}\frac{dQ}{dP}$$

$$\frac{dQ}{dP} = B(-n)P^{-n-1} = -BnP^{-n-1}$$

$$\eta_d = \frac{-P}{Q}(-BnP^{-n-1}) = \frac{P}{Q}BnP^{-n-1}$$

$\eta = 1$ $$\eta_d = \frac{P}{Q}B \cdot 1P^{-1-1} = \frac{P}{Q}BP^{-2} = \frac{P}{Q} \times \frac{B}{P^2} = \frac{B}{QP}$$

(iv) $P = 12 - 4^x$

$$\frac{dP}{dx} = 4^x \log 4$$

$$\eta = \frac{-P}{q}\frac{dx}{dP} = \frac{-P}{q}\left|\frac{1}{4x \log 4}\right|$$

at $P = 2$

$$\eta = \frac{-2}{q}\left|\frac{1}{4x \log 4}\right|$$

Example 5: Determine the price elasticities of demand of the following:

(i) $P = qe^{-q}$ (ii) $P = qe^{\frac{1}{q^2}}$

Sol. (i) $P = qe^{-q}$

Differentiate w.r. to q $\frac{dP}{dq} = q[-e^{-q}] + e^{-q}(1) = e^{-q}[1-q]$

$$\eta_d = \frac{-P}{q}\frac{dq}{dP} = \frac{-P}{q} \times \frac{1}{e^{-q}(1-q)}$$

$$= \frac{qe^{-q}}{q} \times \frac{1}{e^{-q}(1-q)} = \frac{1}{(q-1)}$$

$$\eta_d = \frac{1}{(q-1)}$$

(ii) $P = qe^{\frac{1}{q^2}}$ differentiate w.r. to q

$$\frac{dP}{dq} = qe^{\frac{1}{q^2}}\left[\frac{-2}{q^3}\right] + e^{\frac{1}{q^2}} = \frac{e^{\frac{1}{q^2}(q^2-2)}}{q^2}$$

$$\eta_d = \frac{-P}{q}\frac{dq}{dP} = \frac{-qe^{\frac{1}{q^2}}}{q} \times \frac{q^2}{e^{\frac{1}{q^2}}(q^2-2)} = \frac{-q^2}{2-q^2}$$

$$\eta_d = \frac{-q^2}{2-q^2}$$

Example 6: Show that elasticity of demand at all points on the curve $xy = \alpha^2$ will be numerically equal to one where x is price and 4 is quantity.

Sol. $\frac{xdy}{dx} + y \cdot 1 = 0, \quad \frac{dy}{dx} = \frac{-y}{x}$

Price elasticity of demand $= \frac{-x}{y}\frac{dy}{dx}$

$$= \frac{-x}{y}\left(\frac{-y}{x}\right) = 1, \text{ where } x = \text{price}; \; y = \text{quantity}$$

Example 7: The demand function is Q = 25 – 4P + P^2; where Q is quantity demand for the commodity and P is its price. What is elasticity of demand when price is (a) Rs. 8 (b) Rs. 5?

Sol. $Q = 25 - 4P + P^2, \quad \frac{dQ}{dP} = -4 + 2P$

$$\eta_d = \frac{-P}{Q}\frac{dQ}{dP} = \frac{-P(-4+2P)}{25-4P+P^2}$$

At $P = 8, \; \eta_d = \frac{8(4-16)}{25-32+64} = \frac{-8\times 12}{57} = \frac{-96}{57}$

at $P = 5, \; \eta_d = \frac{5(4-10)}{25-20+25} = -1$

Example 8 : Determine Price elasticity of demand and marginal revenue if $q = 30–4p–p^2$ where q is the quantity demanded and p is price and p = 3.

Sol. $q = 30–4P–P^2$

$$\frac{dq}{dP} = -4-2P$$

When $P = 3, \; q = 30–12–9 = 9$

Price $\quad E_d = \dfrac{-P}{q}\dfrac{dq}{dP}$

$$= \frac{-3}{9}(-4-2\times 3)$$

$$= \frac{-1}{3}(-10) = \frac{10}{3}$$

$$MR = P+q\frac{dq}{dP}$$

$$= P+q\left(\frac{1}{-4-2P}\right)$$

$$= 3+9\left(\frac{1}{-4-2\times 3}\right)$$

$$= 3-\frac{9}{3} = \frac{21}{10}$$

7.11. ELASTICITY OF TOTAL COST AND AVERAGE COST

Q.1. Explain the concept of elasticity of total cost and elasticity of average cost.

Sol. Suppose a firm produces q units of output at a total cost $\pi = f(q)$, then the elasticity of total cost is measured by

$$E = \frac{q}{\pi}\cdot\frac{d\pi}{dq} = \frac{\frac{d\pi}{dq}}{\frac{\pi}{q}} = \frac{MC}{AC}$$

Now elasticity of average cost is measured by

$$E = \frac{q}{t}\cdot\frac{dt}{dq} = \frac{q}{t}\frac{d}{dq}\left(\frac{\pi}{q}\right) = \frac{\frac{q}{t}\left[q\frac{q\pi}{dq}-\pi 1\right]}{q^2}$$

$$= \frac{1}{t}\left[\frac{d\pi}{dq} - \frac{\pi}{q}\right] \quad = \frac{\text{M–AC}}{t}[t = \text{AC}]$$

$$= \frac{\text{MC–AC}}{\text{AC}} = \frac{\text{MC}}{\text{AC}} - 1$$

$$= \text{E} - 1$$

Example 1: The income elasticity of the consumer expenditure is defined to be E = $\frac{dC}{dy} \cdot \frac{y}{C}$, where C is consumer expenditure y is income if $C = a+by$, where a and b are fixed constants, show that E=$\frac{1}{1+\frac{a}{by}}$. Show that E is independent of both C and y if a = 0.

Sol. Give C = a + by differentiate w.r. to y, $\frac{dC}{dy} = b$.

Now elasticity of commer expenditure is

$$\text{E} = \frac{dC}{dy} \cdot \frac{y}{C} \quad = b\frac{y}{a+by} \quad = \frac{by}{by\left(1+\frac{a}{by}\right)}$$

$$\therefore \qquad \text{E} = \frac{1}{1+\frac{a}{by}}$$

When $a = 0$, E = $\frac{1}{1+0} = 1$

$\therefore$ E is independent of both C and y if a = 0

Check Your Progress–II

Q.1. (*a*) Given the demand function $q = 100 - 2p - 2p^2$. Calculate the elasticity of demand when $p = 10$.

(*b*) Find elasticity of demand, if the demand function is

$x = 250–5p+p^2$. Also find elasticity of demand at $p = 8$.

$$\left[\begin{array}{l} Hint.\ \dfrac{dx}{dp} = -5+2p,\ \eta_d = \dfrac{-p}{x}\dfrac{dx}{dp} = \dfrac{p(5-2p)}{250-5p+p^2} \\ \qquad\qquad when\ p=8,\ \eta_d = \dfrac{8(5-2\times 8)}{250-5\times 8+8^2} = \dfrac{-44}{137} \end{array}\right]$$

Q.2. (*a*) Find η, for (*i*) $x = 50–p^2$, (ii) $x = a–bp$.

(*b*) Prove that for the demand curve $qp^{-a} = b$, where a, b are constants, $\eta_d = –a$.

Q.3. (*a*) If $p = 20 - q$, find value of p for which $\eta=1$.

(b) If the demand function is $p = 4 = x^2$, for what value of x, will the elasticity of demand be unitary ?

Q.4. If the demand law is $p = (4 - 5x)^2$, on what value of x, is the elasticity of demand unitary ?

Q.5. If the demand law is given by $q = \dfrac{20}{p+1}$, find η_d w.r. to price at the point when $p = 3$.

Q.6. (*a*) Given the demand function $p = 50 - 3q$. Find the elasticity of demand when $p = 5$.

(b) Given the demand function $p = 20 - 2x$, find price elasticity of demand at $p = 4$ & $p = 8$.

7.12. MAXIMA AND MINIMA

7.12.1. Increasing and Decreasing Function

Q. Explain increasing and decreasing function. If $y = f(x)$ is increasing function, then $\dfrac{dy}{dx} > 0$ and if $y = f(x)$ is decreasing function then $\dfrac{dy}{dx} < 0$.

Sol. Increasing and Decreasing function: A function $y=f(x)$ is said to be an increasing function if as x increases y also increases and the function is said to be decreasing if as x increases y decreases.

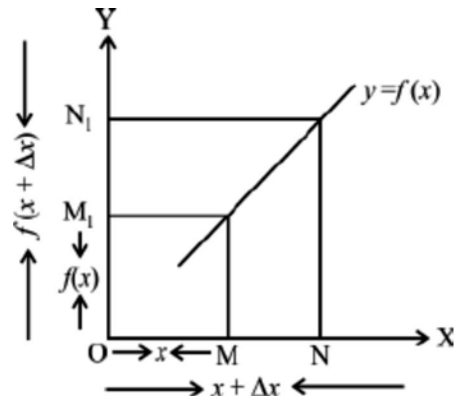


Fig.1

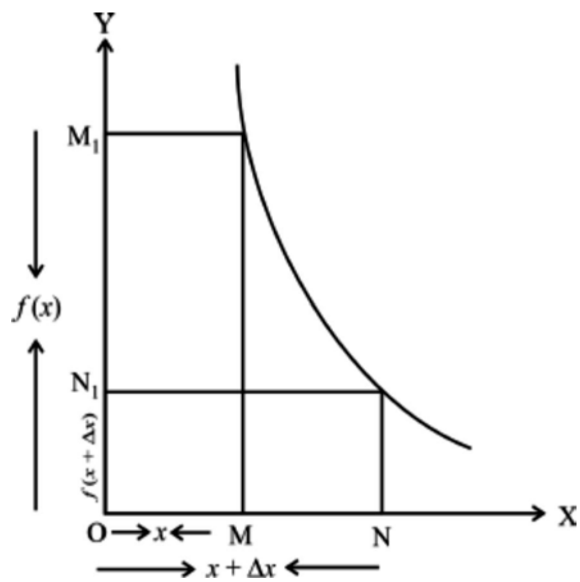


Fig.2

If $y = f(x)$ be an increasing function of x, then as x increases from OM to ON, y increases from OM_1 to ON_1. (Fig. 1)

$$\therefore \qquad ON_1 > OM_1$$

$\therefore \qquad f(x+\Delta x) > f(x)$

or $f(x+\Delta x) - f(x) > 0$

or $\qquad \dfrac{f(x+\Delta x)-f(x)}{\Delta x} > 0 (\Delta x > 0)$

Take limits as $\Delta x \to 0$

or $\qquad \underset{\Delta x \to 0}{Lt} \dfrac{f(x+\Delta x)}{\Delta x} > 0$

or $\qquad \dfrac{d}{dx} f(x) > 0, \quad$ i.e. $\quad \dfrac{dy}{dx} > 0$

If $y = f(x)$ be a decreasing function of x, then as x increases OM to ON, y decreases from OM_1 to ON_1. (Fig. 2)

$\therefore \qquad ON_1 < OM_1$

or $\qquad f(x + \Delta x) < f(x)$

or $\quad f(x+\Delta x) - f(x) < 0, \qquad (\Delta x > 0)$

or $\qquad \dfrac{f(x+x)-f(x)}{\Delta x} < 0$

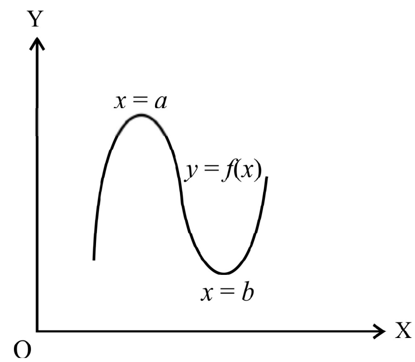
Take limits as $\Delta x \to 0$

or $\qquad \underset{\Delta x \to 0}{Lt} \dfrac{f(x+\Delta x)-f(x)}{\Delta x} < \qquad 0$

or $\qquad \dfrac{d}{dx} f(x) < 0 \qquad$ i.e. $\dfrac{dy}{dx} < 0$

7.12.2. Maximum Value, Minimum Value and Point of Inflexion of a Function at a Point

Maximum and Minimum: A function $y = f(x)$ is said to be maximum at a point if it ceases to increase and begins to decrease at that point and the function is said to be minimum at a point of it ceases to decrease and begins to increase at that point.



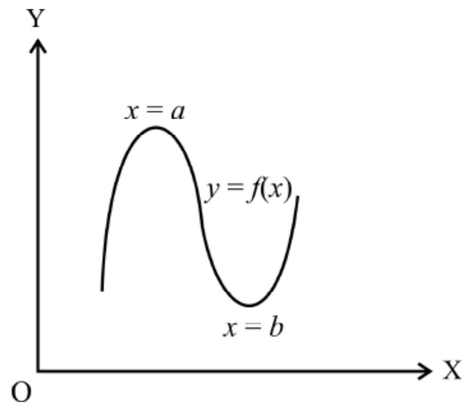
Point of inflexion: If $f'(x)$ does not changes sign as x passes through the value a, where $f'(a) = 0$ then $f(a)$ is neither a maxmum nor a minimum value of $f(x)$ at $x = a$. Then $x = a$ is called the point of inflexion on the curve $y = f(x)$.

Example 1: Prove that at the maximum point $\frac{dy}{dx} = 0$ and $\frac{dy}{dx}$ changes from positive to negative and at the minimum point $\frac{dy}{dx} = 0$ and $\frac{dy}{dx}$ changes from negative to positive.

Sol. Let $y\ (x)$ be the given function. Let $x = a$ be a maximum point, then as the function ceases to increase and begins to decrease at the maximum point.

$\therefore\ \frac{dy}{dx}$ changes from positive to negative at the maximum point.

So at the maximum point $\frac{dy}{dx}=0$ and $\frac{dy}{dx}$ changes from positive to negative.



Similarly at the minimum point, $\frac{dy}{dx}$ = 0 and $\frac{dy}{dx}$ changes from negative to positive.

(i) If $f'(a) = 0$ and $f''(a) < 0$ then $f(a)$ is the maximum value of $f(x)$ at $x = a$

(ii) If $f'(a) = 0$ and $f''(a) > 0$, then $f(a)$ is the minimum value of $f(x)$ at $x = a$

PRACTICAL QUESTIONS

Q.1. Examine whether the following functions have maxima or minima:

(i) $y = x^3 + x^2 + x + 1$

(ii) $y = x + 10$

Sol. (i) Let $y = x^3 + x^2 + x + 1$

$$\frac{dy}{dx} = 3x^2 + 2x + 1 = 0$$

$$x = \frac{-2 \pm \sqrt{-8}}{6}$$

Thus there is no real value of x satisfying $\frac{dy}{dx} = 0$

Hence the function does not have a maximum or a minimum value.

(ii) Hence $\frac{dy}{dx} = 10$

For maximum or minimum $\frac{dy}{dx} = 0$, which is not possible. Hence, the function does not have a maximum or a minimum value.

Q.2. The total cost function of ABC Co. Ltd. is given by

$$C = 4x^3 - 9x^2 + 10x + 10$$

Find (i) Average Cost (AC) (ii) Slope of AC (iii) Marginal Cost (MC) (iv) Slope of MC.

Sol. Given $C = 4x^3 - 9x^2 + 10x + 10$

(i) Average Cost (AC) $= \frac{C}{x} = 4x^2 - 9x - 10 + \frac{10}{x}$

(ii) Slope of AC $= \frac{d}{dx}$ (AC) $= 8x - 9 - \frac{10}{x^2}$

(iii) Marginal cost $= \frac{dC}{dx} = 12x^2 - 18x + 10$

(iv) Slope of MC $= \frac{d}{dx}$ (MC) $= 24x - 18$

Q.3. (a) Find the minimum value of $y = 20 - 4x + x^2$.

(b) State whether $x^2 - 270x + 18400$ has maximum or minimum value. Also find that value.

(c) Find maximum and minimum values of

(i) $x^2 - 6x + 13$

(ii) $4x^2 - 2x + 1$

Sol. (a) $y = 20 - 4x + x^2$, differentiate w.r. to x

$$\frac{dy}{dx} = -4 + 2x, \text{ differentiate w.r. to } x$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx}(-4 + 2x) = 2$$

For extreme values, $\frac{dy}{dx} = 0$, i.e. $-4 + 2x = 0$ or $x = 2$

Since $\frac{d^2y}{dx^2}$ in positive.

$\therefore \quad x = 2$ gives minimum value and minimum value of

$$y = 20 - 4 \times 2 + 2^2 = 16$$

(b) Given $y = x^2 - 270x + 18400$, differentiate w.r.t. to x

$$\frac{dy}{dx} = 2x - 270, \text{ differentiate again w.r. to } x$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx}(2x - 270) = 2$$

For extreme value $\frac{dy}{dx} = 0$

$\therefore \quad 2x - 270 = 0, \quad x = 135$

Since $\frac{d^2y}{dx^2}$ is positive, therefore $x = 135$ gives minimum value and minimum value of

$$y = (135)^2 - 270 \times 135 + 18400 = 175$$

(c) (i) $y = x^2 - 6x + 13$, differentiate w.r. to x

$$\frac{dy}{dx} = 2x - 6; \quad \frac{d^2y}{dx^2} = 2 \frac{dy}{dx} = 0 \text{ gives } 2x - 6, \ x = 3$$

as $\frac{d^2y}{dx^2} = 2 > 0 \quad \therefore \quad x = 3$ gives minimum value.

(ii) $y = 4x^2 - 2x + 1$, differentiate w.r. to x

$$\frac{dy}{dx} = 8x - 2 \qquad \therefore \quad \frac{d^2y}{dx^2} = 8$$

$$\frac{dy}{dx} = 0 \text{ gives } \quad 8x - 2 = 0 \qquad \therefore \quad x = \frac{1}{4}$$

As $\frac{d^2y}{dx^2} = 8 > 0 \qquad \therefore \quad x = \frac{1}{4}$ gives minimum value.

y minimum = $4\left(\frac{1}{2}\right)^2 - 2\frac{1}{4} + 1 = \frac{1}{4} + \frac{1}{2} + 1 = \frac{3}{4}$

Q.4. Find the maximum and minimum value of

$$2x^3 - 15x^2 + 36x + 20$$

Sol. $y = 2x^3 - 15x^2 + 36x + 20$

$$\frac{dy}{dx} = 6x^2 - 30x + 36 = 6(x^2 - 5x + 6)$$

For stationary value, $\frac{dy}{dx} = 0$

$$6(x^2 - 5x + 6) = 0$$

$$6(x - 3)(x - 2) = 0$$

Thus, stationary values occur at $x = 2$ and $x = 3$

Now, $\frac{d^2y}{dx^2} = 12x - 30$

For $x = 2$

$$\frac{d^2y}{dx^2} = 24 - 30 = -6 < 0$$

at $x = 2$

$$\frac{d^2y}{dx^2} < 0$$

Hence y is maximum and maximum value of y is

$$y = 2(2)^3 - 15(2)^2 + 36(2) + 20$$

$$= 2 \times 8 - 15 \times 4 + 72 + 20$$

$$= 16 - 60 + 72 + 20 = 48$$

at $x = 3$ $\frac{dy}{dx} = 0$ and $\frac{d^2y}{dx^2} > 0.$

Hence y is minimum and minimum value is

$$y = 2(3)^3 - 15(3)^2 + 36(3) + 20$$

$$= 54 - 135 + 108 + 20 = 47$$

Q. 5. Show that the maximum value of the function $y = x^3 - 27x + 108$ is 108 more than the minimum value.

Sol. Let $y = x^3 - 27x + 108$

$$\frac{dy}{dx} = 3x^2 - 27$$

$$\therefore \quad \frac{dy}{dx} = 0$$

$$3x^2 - 27 = 0$$

$$x^2 = 9$$

$$\therefore \quad x = \pm 3$$

Now, $\dfrac{d^2y}{dx^2} = 6x$

For $x = 3$

$$\frac{d^2y}{dx^2} = 18 > 0$$

at $x = 3$, $\dfrac{dy}{dx} = 0$, $\dfrac{d^2y}{dx^2} > 0$.

Hence at $x = 3$, y is minimum and value is

$$y = x^3 - 27x + 108 = 27 - 27(3) + 108$$

$$= 27 - 81 + 108 = 54$$

at $x = -3$, $\dfrac{dy}{dx} = 0$, $\dfrac{d^2y}{dx^2} < 0$

Hence at $x = -3$, y is maximum and max. value is

$$y = x^2 - 27x + 108 = (-3)^3 - 27(-3) + 108 = 162$$

$\therefore$ Maximum value (162) is more than minimum value (54).

Check Your Progress–III

Q.1. Find the maximum and minimum values of following functions :

(*a*) $x^3 + 2x^2 - 4x - 8$

(*b*) $2x^3 - 3x^2 - 12x + 4$

(*c*) $x^4 - 14x^2 + 24x + 9$

(*d*) $y = 2x^3 - 21x^2 + 36x - 20$

Q.2. (*a*) Show that the curve $y = x + \frac{1}{x}$, has one maximum and one minimum value. Show that the latter is larger than the former.

(*b*) Show that the maximum value of the function $f(x) = x^3 - 27x - 108$ is 108 more than the minimum value.

(*c*) Show that $x^5 - 5x^4 + 5x^3 - 1$ has maximum value when $x = 1$, a minimum value when $x = 3$ and neither maximum nor minimum value when $x = 0$.

Q.3. Examine the function for maxima and minima :

(*i*) $y = x^2 - 3x$

(*ii*) $y = \dfrac{6x}{x^2 + 4}$

7.13. MAXIMUM PROFIT UNDER PERFECT COMPETITION AND MONOPOLY

Example: Derive the first order and second order conditions of profit maximization by a firm. Interpret in case of perfect competition and monopoly.

Sol. Let us denote revenue by R, cost by π and profit by M. Profit may be defined as

$$M = R - \pi$$

In order to maximize M, following two conditions must be satisfied.

(1) **First Order Conditiion:** The first order condition for the necessary condition is satisfied by setting first derivative of M w.r. to *q* equal to zero.

i.e. $$\frac{dM}{dq} = \frac{dR}{dq} - \frac{d\pi}{dq} = 0$$

$$\therefore \quad \frac{d\pi}{dq} = \frac{dR}{dq}$$

i.e. Marginal Cost = Marginal Revenue

(2) **Second Order Condition:** The second order condition requires

$$\frac{d^2M}{dq^2} = \frac{d^2R}{dq^2} - \frac{d^2\pi}{dq^2} < 0$$

or $\dfrac{d^2\pi}{dq^2} > \dfrac{d^2R}{dq^2}$

i.e. Marginal cost exceeds marginal revenue to the right of the equilibrium point. Thus, necessary and sufficient condition for a firm to be in equilibrium are

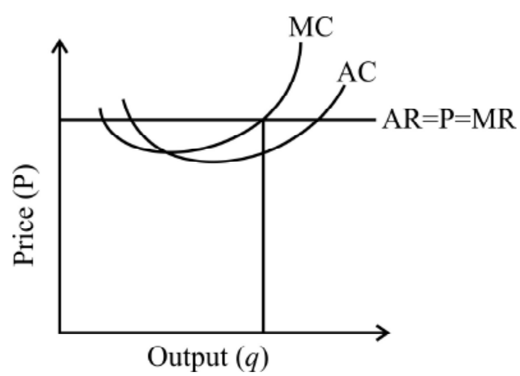
(i) MC = MR

(ii) MC curve must cut MR curve from below.

Under perfect competition P remain constant, for an individual firm is a price taker not a price maker. A firm can sell as such quantity as its wants, but at the given price. Therefore, P is not a function of *q*. But under imperfect competition, P is a function of *q*,

i.e. $P = f(q)$

Since $R = Pq,\ MR = P + q\dfrac{dP}{dq}$



In this expression $\dfrac{dP}{dq}$ shows the rate of change of P w.r. to *q*

Under Perfect competition $\dfrac{dP}{dq} = 0$

Hence $MR = P + q \times 0 = P$

That is why MR curve is parallel to quantity axis under perfect competition and it concides with the average revenue curve.

Therefore, first order condition under perfect competition becomes MC = P

Under Monopoly: The profit M of the monopolist will be maximum when MC = MR.

Example 1: A monopolist firm has the following total cost and demand function $C = aQ^2+bQ+C$; $p = \beta - \alpha Q$

What is the profit maximizing output level when the firm is assumed to fix the output? Verify that this is the same result as when firm fixes the price.

Sol. (i) When firm fixes the output level

$$TR = OP = Q(\beta - aQ),\ MR = \frac{dR}{dQ} = \beta - 2\alpha Q$$

$$TC = aQ^2 + bQ + C,\ MC = \frac{dC}{dQ} = 2aQ + b$$

Condition for profit maximizing output level is MR = MC.

$$\beta - 2\alpha Q = 2aQ + b$$

or $$\beta - b = 2aQ + 2\alpha Q$$

or $$\beta - b = 2Q\ (a + \alpha)$$

$$\frac{\beta - b}{2(a+\alpha)} = Q$$

Hence, profit maximizing level of output is $\left[\dfrac{\beta - b}{2(a+\alpha)}\right]$

(ii) When firm fixes the price. Now TR and TC should be put in ferms of price P.

$$\text{Since } P = \beta - \alpha Q, \qquad Q = (\beta - P)\alpha$$

$$\beta - P = \alpha Q$$

$$\frac{\beta - P}{\alpha} = Q$$

$$TR = QP = \left(\frac{\beta - P}{\alpha}\right)P = \frac{\beta P - P^2}{\alpha},\ \frac{dR}{dP} = \frac{1}{\alpha}(\beta - 2P)$$

$$\text{TC} = a\left(\frac{\beta - P}{\alpha}\right)^2 + b\left(\frac{\beta - P}{\alpha}\right) + C, \frac{dC}{dP} = \frac{-2a\beta + 2aP}{\alpha^2} - \frac{b}{\alpha}$$

For profit maximization, $\frac{d}{dp}(\text{TR–TC}) = 0$

Hence MR–MC = 0

MR = MC

$$\frac{1}{\alpha}(\beta - 2P) = \frac{-2a\beta + 2aP}{\alpha} - \frac{b}{\alpha}$$

i.e. $$P = \frac{2a\beta + \alpha\beta + \beta b}{2(\beta + a)}$$

i.e. $$P = \frac{2a\beta + \alpha\beta + \alpha b}{2(\alpha + a)}$$

Since the demand function is P = β–αQ

$$\therefore \quad \frac{2a\beta + \alpha\beta + \alpha b}{2(\alpha + a)} = \beta - \alpha Q$$

Or $$\alpha Q = \beta - \frac{2a\beta + \alpha\beta + \alpha b}{2(\alpha + a)}$$

Or $$Q = \frac{\beta - b}{2(\alpha + a)}$$

Which is the same level as when firm fixes the output.

Example 2: A radio manufacturer produces 'x' sets per weeks at a total costs of Rs. $\left(\frac{x^2}{25} + 3x + 100\right)$. He is a monopolist and the demand for this market is x = 75–3p, where p is the price in rupees per set. Show that maximum net revenue is obtained when about 30 sets are produced per week. What is the monopoly price?

Sol. Let x be the number of sets which maximizes the net revenue of the monopolist.

$\therefore$ TC for x set is given C = $\frac{x^2}{25}$ + $3x$ + 100, MC = $\frac{2x}{25}$ + 3

Demand function is x = 74–3p, $\therefore$ P = $\left(\frac{75-x}{3}\right)$ per set

$\therefore$ TR for x sets = $\left(\frac{75-x}{3}\right)x$ = $25x - \frac{x^2}{3}$ $\therefore$ MR = $\frac{dR}{dx}$ = $25 - \frac{2x}{3}$

Net revenue will be maximum at the level of output when MR = MC

$$\therefore \quad 25-\frac{2x}{3} = \frac{2x}{25}+3$$

$$\text{Or} \quad \frac{56}{75}x = 22$$

$$\text{Or} \quad x = 30 \text{ (approx)}$$

Since P = $\frac{75-x}{x}$ $\therefore$ When x = 30, P = $\left(\frac{75-30}{3}\right)$ = Rs 15 per set.

Example 3: If the total cost function of a firm is C = $\frac{1}{3}x^3-5x^2+3x+10$, where x is output and price under perfect competition is given as 6, find for what value of x, the Profit M is given as M = Total Sales – Total Cost, will be maximized. Examine both first and second order conditions.

Sol. Total cost for x units is C = $\frac{1}{3}x^3-5x^2+30x+10$

Since Price is Rs, 6 so MR = 6, TR = $6x$

$$\text{Profit M} = \text{TR–TC} = 6x-\left(\frac{1}{3}x^3-5x^2+30x+10\right)$$

$$\frac{dM}{dx} = 6 - (x^2 - 10x + 24)$$

$$\frac{d^2M}{dx^2} = -(2x-10)$$

First order conditions for max. profit is $\frac{dM}{dx} = 0$

i.e. $x^2-10x+24 = 0 \therefore x = 4, 6$

At $\quad x = 4, \frac{d^2M}{dx^2} = -(8-10) = 2>0$

at $\quad x = 6, \frac{d^2M}{dx^2} = -(12-10) = -2<0$

$\therefore$ At $x = 6$, Profit is maximum.

Example 4: The demand function of a firm is P = 500 – 0.2x and its cost function is C = 25x+10,000 (P = Price, x = output and C = Cost). Find the output at which the profits of the firm are maximum. Also find the price it will change.

Sol. $\quad R = Pq = 200q-0.2q^2$

$$\text{Profit} = M = R-\pi = 200q-0.2q^2-100q-30{,}000$$

$$= 100q-0.2q^2-30{,}000$$

Differentiate w.r. to q

$$\frac{dM}{dq} = 100-0.04q, \quad \frac{d^2M}{dq^2} = -0.04$$

First order condition $\frac{dM}{dq} = 100-0.04q$, gives $q = 2500$

Second order condition, $\frac{d^2M}{dq^2} = -\ 0.04<0$,

thus M is maximum at $q = 2500$ substituting $q = 2500$ in .... (i)

we get maximum profit = 95000

Price at which profit is maximum

$$P = 200 - 0.02q = 200 - 0.02\ (2500) = 150$$

Example 5: If the demand function of the monopolist is given by P = 200–0.5 and its cost function is given by C = 100 + 5q + 7q^2

Solution: $R = Pq = 200q - 0.5q^2$

$$\text{Profit} = M = R - C$$
$$= 200q - 0.5q^2 - 100 - 5q - 7q^2$$
$$= 195q - 100 - 7.5q^2$$

$$\frac{dM}{dq} = 195–15q, \quad \frac{d^2M}{dq^2} = –15$$

First order condition: $\frac{dM}{dq} = 0$, i.e. $195 - 15q = 0$, $q = 13$

Second order Condition: $\frac{d^2M}{dq^2} = –15 < 0$

Thus, M is maximum action $q = 13$

when $q = 13$, $P = 200 - 0.5\times13 = 200 - 6.5 = 193.5$

Maximum Profit= $195 \times 13 - 100 - 7(13)^2 = 1167.5$

Example 6: A monopolist has the following demand and cost function

$$P = 100 - 5q, \; C = q^2 + 4q + 5.$$

Find the level of Profit maximizing output and the maximum profit.

Sol. Demand function is $P = 100 - 5q$

$$R = Pq = (100 - 5q)q = 100q - 5q^2$$
$$\text{Profit} = M = R - C$$
$$= 100 - 5q^2 - (q^2 + 4q + 5)$$
$$= 96q - 6q^2 - 5$$

$$\frac{dM}{dq} = 96 - 12q, \quad \frac{d^2M}{dq^2} = –12$$

First order condition, $\frac{dM}{dq} = 0$, i.e. $96q - 12q = 0$, $q = 6$

Second order condition, $\frac{d^2M}{dq^2} = 0 < 0 \quad \therefore$ at $q = 6$ Profit is maximum.

$$\text{Max. Profit} = 96 \times 6 - 6 \times 6^2 - 5$$
$$= 576 - 216 - 5 = 355$$

7.14. LET US SUM UP

In this lesson, we discussed in detail the marginal concept and its wide applications in Economics.

7.15. EXAMINATION ORIENTED QUESTIONS

Q.1. Find the relationship between marginal revenue and elasticity of demand.

Q.2. Define MC, MR and marginal utility.

Q.3. Derive an expression for price elasticity of demand for the demand function $q = f(P)$.

7.16. SUGGESTED READINGS & REFERENCES

1. Monga, G.S. (1972) : Mathematics and Statistics for Economists, Vikas Publishing House.
2. Allen, R.G.D. (1976) : Mathematical Analysis for Economic Macmillan.
3. Chiang, A.C. (1974) : Fundamental Method of Mathematical Economics.
4. Aggarwal, C.S. & R.C. Joshi : Mathematics for Students of Economics.

UTILITY MAXIMISATION

LESSON NO. 8 **UNIT-II**

STRUCTURE

8.2. OBJECTIVES

After going through this lesson, you shall be able to understand :

1. Utility Maximisastion
2. Optimal Behaviour of a Consumer
3. Conditions for Utility Maximisation

8.1. INTRODUCTION

Utility maximisation means that how a consumer gets maximum satisfaction with limited income. Utility (u) derived by the consumer depends on the quantity consumed of various commodities (X_1, X_2 .....) i.e. U = f (x_1, x_2....). In this lesson, we shall discuss this concept in detail.

8.3. UTILITY MAXIMISATION

As consumers, we make choices every day about what and how much to buy and use. In order to know how consumers make these decisions, economists (reasonably) assume that people make choices that maximize their levels of happiness (i.e. people are "Economically rational"). Economists even have their own world for happiness:

1. Utility: The amount of happiness gained from consuming a good or service. This concept of economic utility has some specific properties that are important to keep in mind.

2. Sign matters: Positive utility numbers (i.e. numbers greater than zero) indicate that consuming a good makes the consumer happier.

3. Bigger is better: The greater the utility number, the more happiness the consumer receives from consuming an item. (Note that this is consistent with the first point since large negative numbers are smaller, i.e. less than small negative numbers).

4. Ordinal but not Cardinal Properties: Utility is not measuried by cardinal numbers such as 1, 2, 3, etc but it is ranked I, II, III ...... etc. Thus utility is ordinal and not cardinal.

Example 1: Given the utility function $u = q_1, q_2$ where q_1 and q_2 are commodities and u is utility, hold u constant, find $\frac{dq_2}{dq_1}$.

Sol. We know du $= fq^1 \mathrm{d}q_1 + fq^2 dq_2$,

Here $fq^1 = q_2$

$fq^2 = q_1$

du $= q_2 dq_1 + q_1 dq_2$ when u is constant, d$u = 0$

$0 = q_2 dq_1 + q_1 dq_2$

$$\therefore \qquad \frac{dq_2}{dq_1} = \frac{-q_2}{q_1}$$

Example 2: (a) Explain the price Line or the Budget Constraint line and Derive its equation.

(b) Find the condition for optimal Behaviour of a consumer.

Sol. (a) Price line or the budget line is defined as the locus of a point representing the combinations of two commodities (say Q_1 and Q_2) that can be purchased with the given income and at the given price of the commodities.

Let P_1 and P_2 be the Price of commodities Q_1 and Q_2 respectively and let the consumer possess y_0 amount of money or income to be spent on Q_1 and Q_2, then the budget equation will be

$$y_0 = P_1q_1+P_2q_2 \qquad .... (1)$$

where P_1q_1 is the portion of y_0 spent on Q_1 and the balance P_2q_2 spent on Q_2.

The equation (1) can be written as $q_2 = \frac{1}{P_2}y_0 - \frac{P_1}{P_2}q_1$, Slope of Budget line = $\left(\frac{-P_1}{P_2}\right)$. Thus the slope of budget line is the negative of price ratio.

(b) **Optimal Behaviour of a Consumer:** We know that at given income and the prices, the consumer behaviour is such that his utility is maximised. Let the preference function and budget constraint be

$$u = f(q_1, q_2)$$

And $\qquad y_0 = P_1q_1 + P_2q_2$

Now to find the condition that the utility to be maximum, we require that two conditions must be satisfied.

First order Condition: From the given preference function and the constraint function we have.

$$Z = f(q_1, q_2) + \lambda(y_0 - P_1q_1 - P_2q_2) \qquad ...(2)$$

$$\frac{\partial Z}{\partial q_1} = fq^1 - \lambda P_1 = 0 \qquad ...(3)$$

$$\frac{\partial Z}{\partial q^2} = fq^2 - \lambda P_2 = 0 \qquad ...(4)$$

$$\frac{\partial Z}{\partial \lambda} = y_0 - P_1q_1 - P_2q_2 = 0 \qquad ...(5)$$

From (3) and (4) we have $fq^1 = \lambda P_1$, $fq^2 = \lambda P_2$

Dividing , we get

$$\frac{fq^1}{fq^2} = \frac{P_1}{P_2}$$

Thus, the utility of the consumer will be maximum if marginal rate of substitution is equal to the price ratio of the two commodities Q_1 and Q_2.

Second Order Condition: the second condition is the diminishing marginal substitutability.

Now Marginal rate substitution = $\frac{dq_2}{dq_1} = \frac{-fq_1^0}{fq^2}$

The second condition is that $\frac{d^2q_2}{dq_1} > 0$

i.e. rate at which MRS changes with respect to q_1 is positive.

Example 3: Given the utility function $u = x^2y^3$ and the budget constraint $10 = x + 4y$. find values of x and y that maximize utility.

Sol. Let $Z = x^2y^3 + \lambda(10 - x - 4y)$ ...(1)

$$\frac{\partial Z}{\partial x} = 2xy^3 - \lambda = 0 \quad ...(2)$$

$$\frac{\partial Z}{\partial y} = 3x^2y^2 - 4\lambda = 0 \quad ...(3)$$

$$\frac{\partial Z}{\partial \lambda} = 10 - x - 4y = 0 \quad ...(4)$$

From (2) and (3), $\lambda = 2xy^3 = \frac{3}{4}x^2y^2$ or $x = \frac{8y}{3}$ from (4),

$$10 - \frac{8y}{3} - 4y = 0,\ 20y = 30,\ y = \frac{3}{2} \qquad \therefore\ x = 4$$

Slope of Price line $= \frac{-P_1}{P_2} = \frac{-1}{4}$ (Here $P_1 = 1$, $P_2 = 4$)

Slope of indifference curve = $\frac{-fx}{fy} = \frac{-2xy^3}{3x^2y^2} = \frac{-2y}{3x}$

At $x = 4$, $y = \frac{3}{2}$, slope of indifference curve $\frac{-2\left(\frac{3}{2}\right)}{3(4)} = \frac{-1}{4}$

So first order condition of utility maximization is met.

From second order condition: Let $r = \frac{dy}{dx} = \frac{-fx}{fy} = \frac{-2yx^{-1}}{3}$

$$\frac{dr}{dx} = \frac{\partial r}{\partial x} + \frac{\partial r}{\partial y}\frac{dy}{dx} = \frac{2}{3}yx^{-2} - \frac{2}{3}x^{-1}\left(\frac{2}{3}yx^{-1}\right) = \frac{10}{9}yx^{-2} > 0$$

$\therefore$ Second order condition is also satisfied.

When $x = 4, y = \frac{3}{2}$, utility is maximum.

Example 4: Find the optimum commodity purchases for a consumer whose utility function and budget constraint are respectively given to be

$$U = f(q_1, q_2) = q_1^{15}q_2, \quad 3q_1 + 4q_2 = 100$$

Solution: Utility function and budget constraint are

$$U = q_1^{15}q_2 \qquad \text{...(I)}$$

$$\text{And } 100 - 3q_1 - 4q_2 = 0 \qquad \text{...(II)}$$

$$\text{Let } Z = q_1^{15}q_2 + \lambda(100 - 3q_1 - 4q_2) \qquad \text{...(III)}$$

Setting partial derivative of Z w.r.to q_1, q_2 and λ equal to zero.

i.e. $$\frac{\partial Z}{\partial q_1} = 15q_1^{14}q_2 - 3\lambda = 0 \qquad \text{...(IV)}$$

$$\frac{\partial Z}{\partial q_2} = q_1^{15} - 4\lambda = 0 \qquad \text{...(V)}$$

$$\frac{\partial Z}{\partial \lambda} = 100–3q_1–4q_2 = 0 \qquad ...(VI)$$

From (V), $\lambda = \frac{1}{4} q_1^{15}$, Put in (IV)

$$15q_1 15q_2 – 3.\frac{1}{4} q_1^{15} = 0$$

$$15q_2 – \frac{3}{4} q_1 = 0, \qquad 5q_2 = \frac{1}{4} q_1$$

$q_2 = \frac{1}{20} q_1$ put on (VI), $100 – 3q_1 – 4.\frac{1}{20} q_1 = 0$

$$100 = \frac{16}{5} q_1, \quad q_1 = \frac{500}{16} = \frac{125}{4}$$

Slope of price line $= \frac{-P_1}{P_2} = \frac{-3}{4}$

[$\because$ Here from budget constraint $P_1 = 3$, $P_2 = 4$] ...(VI)

Slope of indifference $= \frac{-fq_1}{fq_2} = \frac{-15q_1^{\frac{1}{4}}q_2}{q_1^{\frac{1}{5}}} = \frac{-15q_2}{q_1}$ ...(VII)

$$q_1 = \frac{125}{4}, \; q_2 = \frac{1}{20}, \; q_1 = \frac{1}{20}\cdot\frac{125}{4} = \frac{25}{16}$$

Slope of indifference curve $= \frac{15}{20} \cdot \frac{q_1}{q_2} = \frac{-3}{4}$

From (VI) and (VII) we see that first order condition of utility maximization is met.

Second order condition:

Let $$r = \frac{dq_2}{dq_1} = \frac{-fq_1}{fq_2} = \frac{15q_2}{q_1} = -15q_2q_1^{-1}$$

$$\frac{dr}{dq_1} = \frac{\partial r}{\partial q_1} + \frac{\partial r}{\partial q_2} \cdot \frac{dq_2}{dq_1}$$

$$= 15q_2q_1^{-2} + (-15q_2^{-1})\ (-15q_2q_1^{-1})$$

$$= \frac{15q_2}{q_1^2} + \frac{225q_2}{q_1^2} > 0$$

Second order condition is also met

Thus when $q_1 = \frac{125}{4}$, $q_2 = \frac{25}{16}$, (total utility is maximum)

8.4. LET US SUM UP

In this lesson, we discussed in detail the concept of utilitity maximisation behaviour of a consumer.

8.5. EXAMINATION ORIENTED QUESTIONS

Q.1. If $U = -x^2 + 2x + 9$, then find maximize utility.

Ans. 10

Q.2. If $U = -100 + 160Q - 10Q^2$

then find maximum utility.

Ans. 540

Q.3. If $u = 10 + 6x - 2x^2$ find maximum utility.

Ans. 14·5

8.6. SUGGESTED READINGS & REFERENCES

1. Allen, R. G.D.; Mathematical Analysis for economists (Macmillan)
2. Yamane Taro : Mathematics for Economics— A Elementary Survey (Prentice Hall of India Pvt. Ltd.)
3. Dowling, Edward T: Introduction to Mathematical Economics (Tata Macgraw)
4. Aggarwal, C.S. & R.C. Joshi : Mathematics for Students of Economics.

Unit III

Measures of Central Tendency and Dispersion

MEASURES OF CENTRAL TENDENCY

LESSON NO. 9 **UNIT-III**

STRUCTURE

9.1. OBJECTIVES

After going through this lesson, you should be able to understand :

1. Measures of Central tendency
2. Characteristics of a good averages
3. Purpose and Functions of Averages
4. Mathematical Averages and Positional Averages

9.2. INTRODUCTION

Statistics deals with a large number of big figures which are not only confusing to mind but also difficult to analyse. Therefore, condensation of data is necessary in statistical analysis to reduce the size of the data and enable us to make comparative studies of related variables. To reduce the complexity of data and to make them comparable, it is essential that the various phenomena which are being compared are reduced to one figure each. A figure which is used to represent a whole series should neither have the lowest value in the series nor the highest value but a value some where between these two limits, possible in the centre where most of the items of the series cluster. Such figures are called measures of central tendency or averages.

9.3. CHARACTERISTICS OF A GOOD AVERAGE OR MEASURE OF CENTRAL TENDENCY

According to **G.U. Yule**, the desirable characteristics of a satisfactory average are:

1. **It should be rigidly defined** i.e. the definion should be clear and unambiguous so that it leads to one and only one interpretation by different Persons. In other words, the definition should not leave anything to the discretion of the investigator or the observer.
2. **It should not be affected by the extreme values:** By extreme observations we mean very small or very large observations. Thus a few very small or vary large observations should not unduly affect the value of a good average.
3. **It should be affected as little as possible by fluctuation of sampling:** By this we mean that if we take independent random samples of the same size from a given population and compute the average for each of these samples then for as ideal average, the values so obtained from different samples should not vary much from one another. The difference in the values of the average for different samples is attributed to the so-called fluctuations of sampling. This property is also explained by saying that an ideal average should possess sampling stability.
4. **It should be suitable for further mathematical treatment:** In other words, the average should possess some important and interesting mathematical properties so that its use in further statistical theory is enhanced. For example, if we are given the averages and sizes (frequencies) of a number of different groups then for an ideal average, we should be in a position to compute the average of the combined group. If an average is not amenable to further algebraic manipulation then obviously its use will be very much limited for further applications in statistical theory.
5. **It should be based on all the observations:** Thus in the computation of an ideal average the entire set of data at our disposal should be used and there should not be any loss of information resulting from not using the available data. Obviously, if the whole data is not used in computing the average, it will be unrepresentative of the distribution.

6. It should be easy to understand and calculate even for a non-mathematical person. In other words, it should be readily comprehensible and should be computed with sufficient ease and rapidly and should not involve heavy arithmetical calculations.

Measures of various orders

Statistical series may differ from each other in the following three ways and accordingly three measures are designed to study the differences:

(a) They may differ in the values of the variable around which most of the items cluster-measures of the first order or measures of central tendency or average are used.

(b) They may differ in the extent to which items are dispersed round the central value - measures of the second order, or measures of dispersion are used.

(c) They may differ in the extent of departure from a normal distribution - measures of the third order or skewness kurtosis etc., are used.

9.3.1. PURPOSE AND FUNCTIONS OF AVERAGES

Some of the main purposes and function of Averages are as under:

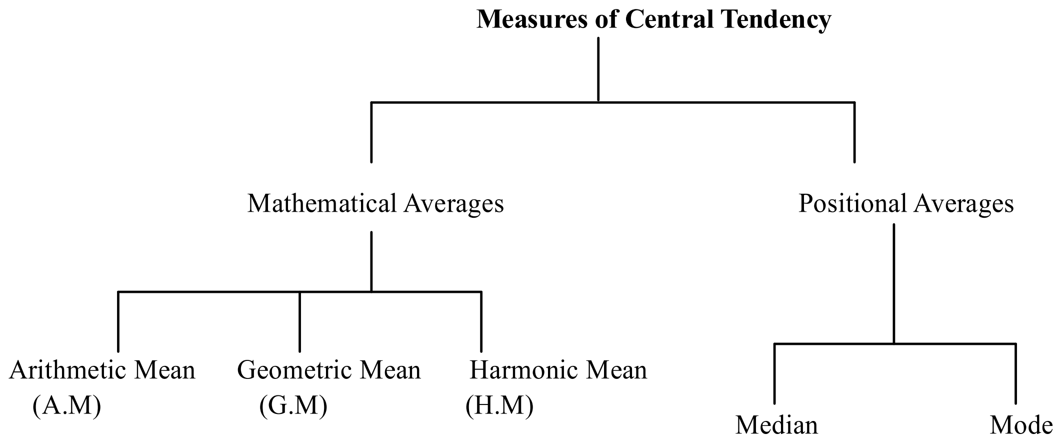
1. **Brief Description:** The main purpose of average is to present a simple and systematic description of the principal features of the raw data. As a result of it, data can be easily understood.

2. **Comparison:** Averages help in making comparison of different sets of data. For example, a comparison of the per capita income of India and U.S.A. shows that per capita income of India is much less than the per capita income of U.S.A. Accordingly, it is concluded that India is a poor country.

3. **One Value for all:** Averages represent the universe or mass of statistical data. One value represents all values of the series. Accordingly conclusion can be drawn in respect of the universe as a whole.

4. **Statistical Analysis:** Averages constitute the basis of statistical analysis. For example, if one knows the average marks secured by the students of a class in their different subjects, one can easily analyse the subjects in which the students are weak.

5. **Formulation of Policies:** Averages help in formulation of Policies.



Average are broadly classified in two categories:

(1) Mathematical Averages

(2) Positional Averages

There are three types of mean which are suitable for a particular type of data. They are—

(a) Arithmetic mean or Average

(b) Geometric mean

(c) Harmonic mean

Arithmetic Mean (A.M.) in Individual observations: It is also popularly known as average. If mean is mentioned, it implies arithmetic mean, as the other means are identified by their full names. It is the most commonly used measures of Central tendency.

Definition: Sum of the observed values of a set divided by the number of observations in the set is called a mean or an average.

If X_1, X_2..... X_N are N observed values, the mean or average is given as

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n} = \frac{\Sigma X}{n}$$

Where is the sum of the observations.

For examples, the arithmetic mean of 5, 8, 10, 15, 24 and 28 is

$$\frac{5+8+10+15+24+28}{6} = \frac{90}{6} = 15$$

Calculation of Arithmetic Mean in Discrete Series

When the data are arranged or given in the form of frequency distribution i.e. there are k variate values such that a value X_i has a frequency f_i (i = 1, 2, ... k)

The formula for the mean is

$$\overline{X} = \frac{f_1X_1 f_2X_2 + ... + f_KX_K}{f_1 + f_2 + ... + f_K} = \frac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i}, i = 1, 2 \ldots k$$

Where N $= f_1 + f_2 + ... + f_K = \sum_{i=1}^{K} f_i$

N = is the total frequency.

Calculation of A.M. in Continuous series

In case of continuous or grouped frequency distribution, the value of X is taken as the mid-value of the corresponding class.

If the data are given with K class interval i.e. the data are in the form as follows:

Class Interval	*Frequency*
X_1 - X_2	f_1
X_2 - X_3	f_2
X_3 - X_4	f_3
...	...
...	...
...	...
...	...
$X_K – X_{K+1}$	f_K

The arithmetic mean

$$\overline{X} = \frac{f_1Y_1f_2Y_2 + ...+ f_KY_K}{f_1+f_2 + ...+ f_K} = \frac{1}{N}\sum_{i=1}^{K} f_iY_i$$

Where Yi is the mid-point of the class interval $X_i - X_i + 1$ and is given as

$$Y_i = \frac{X_i + X_{i+1}}{2} \quad , \ i = 1, 2 \ ... \ ... \ K$$

In this situation the values in the interval are considered to be centered at the mid point of the interval

Remark. The symbol Σ is the letter capital sigma of the Greek alphabet and is used in Mathematics to denote the sum of values.

Steps for the computation of Arithmetic Mean

1. Multiply each value of X or the mid-value of the class (in case of grouped or continuous frequency distribution) by the corresponding frequency.
2. Obtain the total of the products obtained in Step – I above to get ΣfX
3. Divide the total obtained in step 2 by N = Σf, the total frequency.

The resulting value gives the arithmetic mean.

Example 1: Daily cash earnings of 15 workers working in different industries are as follows:

Daily earning (Rs.)

11.63, 8.22, 12.56, 12.14, 29.23, 18.23, 11.49, 11:30, 17.00, 9.16, 8.64, 27.56, 8.23, 19.77, 12.81

Average daily earning of a worker can be calculated by the formula:

$$\overline{X} = A = \frac{\sum fx}{\sum f}$$

$$\overline{X} = \frac{1(11.63 + 8.22 + --+12.81)}{15} = \frac{217.97}{15} = 14.53$$

The average daily earning of a worker in Rs. 14.53

Example 1: The intelligence quotents (IQ's) of 10 boys in a class are given below:

70, 120, 110, 101, 88, 83, 95, 98, 107, 100

Find the mean I.Q

Solution: Mean I.Q $\left(\overline{X}\right)$ of the boys are given by

$$\overline{X} = \frac{\sum X}{n}$$

$$= \frac{1}{10}(70 + 120 + 110 + 101 + 88 + 83 + 95 + 98 + 107 + 100)$$

$$= \frac{972}{10} = 97.2$$

Example 2: The distribution of age at first marriage of 130 males was given below

Age in years (X) : 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29

No. of males (y) : 2, 1, 4, 8, 10, 12, 17, 19, 18, 14, 13, 12

The average age can be computed by the formula:

$$A = \frac{18 \times 2 + 19 \times 1 + 20 \times 4 + \ldots. + 29 \times 12}{2+1+8+\ldots.+12}$$

$$= \frac{3240}{130}$$

$$= 24.92 \text{ years}$$

Example 3: The distribution of size of holdings of cultivated land in an area was as follows:

Size of holding:	0–12	2–4	4–6	6–8	8–10	10–20	20–40	40–60
No. of holdings (f):	48	19	10	14	11	9	2	1

Solution:

Size of holdings (hectares)	Mid Points (4)	No. of holdings (f)	fx
0 - 2	1	48	48
2 - 4	3	19	19 × 3 = 57
4 - 6	5	10	10 × 5 = 50
6 - 8	7	14	14 × 7 = 70
8 - 10	9	11	11 × 9 = 99
10 - 20	15	9	9 × 15 = 135
20 - 40	30	2	2 × 30 = 60
40 - 60	50	1	1 × 50 = 50
			Σfx = 569

$$\text{A.M.} = \frac{\sum fx}{\sum f} = \frac{569}{114} = 4.99$$

The average size of holdings is 4.99 hectares.

Step Deviation Method for Computing Arithmetic Mean

It may be pointed out that the formula $\frac{\sum fx}{N}$ can be used conveniently if the values of X or / and f are small. However, if the values of X or / and f are large the calculation of mean by the formula $\frac{\sum fx}{N}$ is quite tedious and time consuming. In such a case the calculations can be reduced to a great extent by using the step deviation method which consists in taking the deviations (differences) of the given observations from any arbitrary value A.

Let $d = X - A$

Then $fd = f\,(X–A)$

Taking the sum over various values of X we get

$$\sum fd = \sum fx - A\sum f$$

$$\sum fd = \sum fx - A.N \qquad [\because N = \Sigma f]$$

A being constant can be taken outside the summation sign i.e. ΣCX = CΣX where C is constant

$\therefore$ Dividing both sides by N

We get

$$\frac{\sum fd}{N} = \frac{\sum fx}{N} - \frac{A}{N} N = \frac{\sum fx}{N} - A$$

$$\frac{\sum fd}{N} = \overline{X} - A$$

$$\boxed{\overline{X} = A + \frac{\sum fd}{N}}$$

This formula is much more convenient to use for numerical problems than the formula.

In case of grouped or continous frequency distribution, with class intervals or continuous frequency distribution, with class intervals of equal magnitude, the calculations are further simplified by taking

$$d = \frac{X - A}{h}$$

Where X is the mid- value of the class and h is the common magnitude of the class intervals

$$hd = (X - A)$$

Multiplying both sides by f we get

$$hfd = f|X - A| = fx - fA$$

Summing both sides over the values of X we get

$$h\sum fd = \sum fx - A\sum f = \sum fx - N.A$$

Dividing both sides by N we get

$$h\frac{\sum fd}{N} = \frac{\sum fx}{N} - A = \overline{X} - A$$

$$\boxed{\overline{X} = A + h\frac{\sum fd}{N}}$$

Steps for computation of Mean by Step Deviation Method

Step 1: Compute d = (X -A) h, A being any arbitrary number and h is the common magnitude of the classes, Algebraic signs + or - are to be taken with the deviations.

Step 2: Multiply d by the corresponding frequency to get Σfd

Step 3: Find the sum of the Products obtained in Step 2.

Step 4: Divide the sum obtained in Step 3 by N, the total frequency.

Step 5: Multiply the value obtained in step 4 by h.

Step 6: Add A to the value obtained in step 5.

The resulting value gives the arithmetic mean of the given distribution.

Example 1: Calculute the mean for the following frequency distribution:

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Number of students	6	5	8	15	7	6	3

(i) By the direct formula

(ii) By the step deviation method.

Solution:

Marks	Mid-value (X)	No. of students (f)	fx	$d = \frac{X-35}{10}$	fd
0 - 10	5	6	30	–3	–18
10 - 20	15	5	75	–2	–10
20 - 30	25	8	200	–1	–8
30 - 40	35	15	525	0	0
40 - 50	45	7	315	1	7
50 - 60	55	6	330	2	12
60 - 70	65	3	195	3	9
		N=Σf=50	Σfx=1670		Σfd=8

(i) Direct Method:

$$\text{Mean } \left(\overline{X}\right) = \frac{\Sigma fx}{\Sigma f} = \frac{1670}{50} = 33.4 \text{ marks}$$

(ii) Step Deviation Method:

We have A = 35 and h = 10

$$\overline{X} = A + \frac{h.\Sigma d}{N} = 35 + \frac{10 \times (-8)}{50} = 35 - 1.6 = 33.4 \text{ marks}$$

Correcting Incorrect values: It sometimes happens that due to an oversight or mistake in copying, certain wrong items are taken while calculating mean. The problem is how to find out the correct mean. The process is very simple. From incorrect ΣX deduct wrong items and add correct items and then divide the correct ΣX by the number of observations. The results, so obtained, will give the value of correct mean.

Example 1: The mean marks of 100 students were found to be 40. Later on it was discovered that a score of 53 was misread as 83. Find the correct mean corresponding to correct score.

Solution: We are given N = 100, $\overline{X}$ = 40

$$\overline{X} = \frac{\Sigma X}{N}$$

$$\Sigma X = N\overline{X} = 100 \times 40 = 4000$$

But this is not correct ΣX

Correct ΣX = Incorrect ΣX – wrong item + correct item

= 4000 – 83 + 53 = 3970

$$\therefore \text{Correct } \overline{X} = \frac{\text{Correct } \Sigma X}{N} = \frac{3970}{100} = 39.7$$

Hence the correct average 39.7

Example 2: Mean of 100 observations is found to be 40. If at the time of computation two items are wrongly taken as 30 and 27 istead of 3 and 72. Find correct mean.

Solution: $\overline{X} = \dfrac{\Sigma X}{N}$ or $\Sigma X = N\overline{X}$

Here $\overline{X} = 40$ $N = 100$

$\therefore$ $\Sigma X = 100 \times 40 = 4000$

Correct ΣX = Incorrect ΣX – wrong item + correct item

$= 4000 - 57 + 75$

Correct $\Sigma X = 4018$

Correct Mean $= \dfrac{4018}{100} = 40.18$

Mathematical Properties of Arithmetic Mean

Arithmetic mean possess somevery interesting and important mathematical properties as given below.

Property I: The algebraic sum of the deviations of the given set of observations from their arithmetic mean is zero.

Mathematically, $\Sigma\left(X - \overline{X}\right) = 0$

Or for a frequency distribution: $\Sigma f\left(X - \overline{X}\right) = 0$

Proof: $\Sigma f\left(X - \overline{X}\right) = \Sigma f\left(X - f\overline{X}\right) = \Sigma fx - \Sigma f\overline{X}$

$= \Sigma fx - \overline{X}\Sigma f$ $\left(\because \ \overline{X} \text{ is constant}\right)$

$= \Sigma fX - \overline{X}N$ $\left(\because \Sigma f = N\right)$

$= \quad \overline{X} = \dfrac{1}{N}\Sigma fx = \Sigma fX = N\overline{X}$

$\therefore \ \Sigma f\left(X - \overline{X}\right) = \ N\overline{X} - \overline{X}N = 0$

This would more clear with the help of example:

X	$(X - \overline{X})$
10	- 20
20	- 10
30	0
40	+ 10
50	+ 20
$\Sigma X = 150$	$\Sigma(X - \overline{X}) = 0$

$$\text{Here } \overline{X} = \frac{\Sigma X}{N} = \frac{150}{5} = 30$$

When the sum of the deviations from the actual mean i.e. 30 is taken it comes out to be zero. It is because of this property that the mean is characterized as a point of balance i.e. the sum of the positive deviations from it is equal to the sum of the negative deviations from it.

Property 2: Mean of the combined Series

If we know the sizes and means of two component series, then we can find the mean of the resultant series obtained on combining the given series.

If n_1 & n_2 are the sizes and $\overline{X_1}, \overline{X_2}$ are the respective means of two groups then the mean $\overline{X}$ of the combined group of size $n_1 + n_2$ given by

$$\overline{X} = \frac{n_1 \overline{X_1} + n_2 \overline{X_2}}{n_1 + n_2}$$

Proof: We know that if $\overline{X}$ is the mean of n observations then $\overline{X} = \frac{\Sigma X}{n} = \Sigma X = n\overline{X}$

i.e. Sum of n observations = n × Arithmetic Mean

If $\overline{X}_1$ is the mean of n_1 observations of the first group and $\overline{X}_2$ is the mean of n_2 observations of the second group then we get.

The sum of n_1 observations of the first group = $n_1 \times \overline{X}_1 = n_1 \overline{X}_1$

The sum of other n_2 observations of the second group = $n_2 \times \overline{X}_2 = n_2 \overline{X}_2$

$\therefore$ The sum of $(n_1 + n_2)$ observations of the

Combined group = $n_1 \times \overline{X}_1 = n_2 \overline{X}_2$

Hence, the mean $\overline{X}$ of the combined group $n_1 + n_2$

Observations is given by $\overline{X} = \dfrac{\text{Sum of}\left(n_1 + n_2\right)\text{ observations}}{n_1 + n_2} = \dfrac{n_1 \overline{X}_1 + n_2 \overline{X}_2}{n_1 + n_2}$

Remarks: Some writers use the notation $\overline{X}_{12}$ for the combined mean of the two groups thus we may write

$$\overline{X}_{12} = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2}{n_1 + n_2}$$

<u>**Generalisation**</u> In general, if $\overline{X}_1, \overline{X}_2 -- \overline{X}_K$ are the arithmetic means of k groups with n_1, n_2 ... n_K observations respectively the can similarly prove that the mean $\overline{X}$ of the combined group of size $n_1 + n_2 + ... + n_K$ given by

$$\overline{X} = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2 + --+ n_K \overline{X}_K}{n_1 + n_2 + --+ n_K}$$

The following example shall illustrate the application of the above formula

Example 1: The Mean height of 25 male workers in a factory is 61cms & the mean height of 35 female workers in the same factory is 58 cms. Find the combined mean height of 60 workers in the factory

Solution:

$$\overline{X}_{12} = \frac{N_1 \overline{X}_1 + N_2 \overline{X}_2}{N_1 + N_2}$$

$N_1 = 25, \overline{X}_1 = 61,\ N_2 = 35,\ \overline{X}_2 = 58$

$$\overline{X}_{12} = \frac{(25 \times 61)+(35\times58)}{25 + 35} = \frac{1525+2030}{60} = \frac{3555}{60} = 59.25$$

If we have to find out the combined mean of three sub-groups the above formula can be extended as follows:

$$\overline{X}_{123} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + N_3\overline{X}_3}{N_1 + N_2 + N_3}$$

Property 3: The sum of the squared deviations of the items from arithmetic mean is minimum, that is less than the sum of the squared deviations of the items from any other value. The following example would clarify the point:

X	$(X - \overline{X})$ $\overline{X} = 4$	$(X - 4)^2$
2	–2	4
3	–1	1
4	0	0
5	1	1
6	2	4
$\Sigma X = 20$	$\Sigma(X-\overline{X}) = 0$	$\Sigma(X-\overline{X})^2 = 10$

The sum of the squared deviations is equal to 10 in the above case. If the deviations are taken from any other value the sum of the squared deviations would be greater than 10. For example let us calculate the squares of the deviations of items from a value less than the arithmetic mean say, 3.

X	(X–3)	$(X–3)^2$
2	–1	1
3	0	0
4	1	1
5	2	4
6	3	9
		$\Sigma(X–3)^2 = 15$

It is clean that $\Sigma(X-3)^2$ is greater. This property state that the sum of the squares of items is least from the mean is of immense use in regression analysis which shall be discussed later.

Merits and Demerits of Arithmetic Mean

Merits: Arithmetic mean is most widely used in Practice because of the following reasons.

1. It is defined by a rigid mathematical formula with the result that every one who computes the average gets the same answer.
2. It is affected by the value of every item in the series.
3. It is the simplest average to understand and easiest to compute. Neither the arranging of data is required for calculating median nor grouping of data as required for calculating mode is required while calculating mean.
4. Being determined by a rigid formula, it lends itself to subsequent algebraic treatment better than the median or mode.
5. It is relatively reliable in the sense that it does not vary too much when repeated samples are taken from the same population, at least not as much as some other kind of statistical descriptions.
6. The mean is typical in the sense that it is the centre of gravity, balancing the values on either side of it.
7. It is a calculated value and not based on Position in the series.

Demerits:

1. Arithmetic mean cannot be used in the case of open end classes such as less than 10, more than 70, etc., since for such classes we cannot determine the mid-value X of the of the class intervals unless (i) we estimate the end intervals or (ii) we are given the total value of the variable in the open end classes. In such cases mode or median (discussed latter) may be used.
2. Arithmetic mean cannot be obtained if a single observation is missing or lost or is illegible unless we drop it out and compute arithmetic mean of the remaining values.
3. Arithmetic mean can not be used if we are dealing with qualitative characteristics which can not be measured quantitatively such as intelligence, honesty, beauty etc. In such cases median (discussed later) is the only average to be used.

4. Since the value of mean depends upon each and every item of the series, extreme items i.e. very small and very large items, unduly affect the value of the average. For example, if in a tutorial group there are 4 students and their marks in a test are 60, 70, 10 and 80, the average marks would be 60 + 70 + 10 + 80 = $\frac{220}{4} = 53$ on single item, i.e. 10 has reduced the average marks considerably. The smaller the number of observations, the greater is likely to be the impact of extreme values.
5. It can not be determined by inspection nor can it be located graphically.

MEDIAN

In the words of **L.R. Connor**, "The median is that value of the variable which divides the group in two equal parts, one part comprising all the values greater and the other, all the values less than median".

The median by definition refers to the middle value in a distribution. In case of median one-half of the items in the distribution have a value the size of the median value or smaller and one-half have a value the size of the median value or larger. The median is just the 50th percentile value below which 50% of the values in the sample fall. It splits the observations into two halves.

Calculation of Median — Individual observations

Steps:

1. Arrange the data in ascending or descending order of magnitude (Both arrangements would give the same answer).
2. In a group composed of an odd number of values such as 7, add 1 to the total number of values and divide by 2. Thus 7+1 would be 8 which divided by 2 gives 4 - the number of the value starting at either end of the numerically arranged groups which will be the median value. In a large group, same method would be followed. In a group of 199 items the middle value would be 100th value.

This would be determined by $\frac{199+1}{2} = 100$

In the form of formula

$$\text{Med} = \text{Size of } \frac{N+1}{2} \text{th item}.$$

Example 1: From the following data of the weekly wages of 7 workers. Compute the median wage:

Wages (in Rs.) 100 150 80 90 160 200 140

Solution:

S.No.	Wages arranged in ascending order
1	80
2	90
3	100
4	140
5	150
6	160
7	200

$$\text{Median} = \text{Size of } \frac{N+1}{2}\text{th item}$$

$$= \frac{7+1}{2}\text{th item} = 4\text{th item}$$

Size of 4th item = 140

Hence median wage = Rs. 140

We thus find that median is the middle most item – 3 Persons get a wage less than Rs. 140 and equal number i.e. 3 get more than Rs. 140.

The procedure for determining the median of an even-numbered group of items is not as obvious as above. If there were, for instance different values in a group, the median is really not determinable since both the 5^{th} & 6^{th} values are in the centre. In practice, the median value for a group composed of an even number of items is estimated by finding the arithmetic mean of the two middle values – that is adding the two values in the middle and dividing by two expressed in the form of formula, it amount to

$$\text{Median} = \text{size of } \frac{N+1}{2}\text{th item}$$

Thus we find that it is both when N is odd as well as even than 1 has to be added to determine median value.

Example 2: Obtain the value of median from the following data:

391, 384, 591, 407, 672, 522, 777, 753, 2448 , 1,490

Solution: Calculation of Median

S.No.	Data arranged in ascending order X
1	384
2	391
3	407
4	522
5	591
6	672
7	753
8	777
9	1,490
10	2,488

$$\text{Median} = \text{Size of } \frac{N+1}{2}\text{th item}$$

$$= \frac{11}{2} = 5.5^{th} \text{ item}$$

$$\text{Size of } 5.5^{th} \text{ item} = \frac{\text{5th item + 6th item}}{2}$$

$$= \frac{591+672}{2} = \frac{1263}{2} = 631.5$$

Computation of Median— Discrete Series

1. Arrange the data in ascending or descending order of magnitude.
2. Find out the cumulative frequencies
3. Apply the formula : Median = Size of $\frac{N+1}{2}$
4. Now look at the cumulative frequency column and find that total which is either equal to $\frac{N+1}{2}$ or next higher than that and determine the value of variable corresponding to this. That gives the value of median.

Example 1: From the following data find the value of median:

Income (Rs.)	100	150	80	200	250	180
No. of Persons	24	26	16	20	6	30

Solution: Calculation of Median

Income arranged in ascending order	No. of Persons (f)	*C.f.*
80	16	16
100	24	40
150	26	66
180	30	96
200	20	116
250	6	122

Median = Size of $\frac{N+1}{2}$th item = $\frac{122+1}{2} = 61.5^{th}$ item

Size of 61.5th item = 150

Calculation of Median— Continuous Series

Steps involved for its computation are:

1. Prepare 'less than' cumulative frequency (C.f.) distribution
2. Find $\frac{N}{2}$
3. See c.f. is just greater than $\frac{N}{2}$
4. The corresponding class contains the median value and is called the median class.

The value of median is now obtained by using the interpolation formula:

$$\text{Median} = 1+\frac{h}{f}\left(\frac{N}{2}-C\right) \qquad ...(1)$$

Where

l : is the lower limit of the median class,

f : is the frequency of the median class,

h : is the magnitude or width of the median class,

n : Σf is the total frequency and C is the cumulative frequency of the class preceding the median class.

Remarks: The interpolation formula is based on the following assumptions:

1. The distribution of the variable under consideration is continuous with exclusive type classes without any gaps.
2. There is an orderly and even distribution of observations with in each class.

However, if the data are given as a grouped frequency distribution where classes are not continuous, then it must be converted into continuous frequency distribution before applying the formula: This adjustment will affect only the value of 1 in 1.

Example 1: The frequency distribution of weights in grams of mangoes of a given variety is given calculate the Median.

Solution: Since the interpolation formula for median is based on continuous frequency distribution we shall first convert the given inclusive class interval series in to exclusive class interval series.

Weights in grams	No. of Mangoes	(Less than) *C.f.*
409.5 - 419.5	14	14
419.5 - 429.5	20	34
429.5 - 439.5	42	76
439.5 - 449.5	54	130
449.5 - 459.5	45	175
459.5 - 469.5	18	193
469.5 - 479.5	7	200
	$\Sigma f = 200 = N$	

$\frac{N}{2} = 100$. The C.f. just greater than 100 is 130

Hence the corresponding class 439.5 – 449.5 is the median class

$$M\& = 1 + \frac{h}{f}\left(\frac{N}{2} - C\right)$$

$$= 439.5 + \frac{10}{54}(100 - 76)$$

$$= 439.5 + \frac{10 \times 24}{54}$$

$$= 439.50 + 4.44 = 443.94 \text{ gms.}$$

Example 2: Find the median from the following data:

Wages/Week (Rs.)	No. of workers
50 - 59	15
60 - 69	40
70 - 79	50
80 - 89	60
90 - 99	45
100 - 109	40
110 - 119	15

Solution: Calculation of Median

Wages/ Week (Rs.)	No. of workers	*C.f.*
50 - 59	15	15
60 - 69	40	55
70 - 79	50	105
80 - 89	60	165
90 - 99	45	210
100 - 109	40	250
110 - 119	15	265

Median = Size of $N/2$ th item

$$= \frac{265}{2} = 132.5^{th} \text{ item}$$

Median lies in the class 80–89. But the real limit of this class is 79.5 – 89.5

$$\text{Median} = l + h/f \left(N/2 - C\right)$$

$$l = 79.5$$

$$\frac{N}{2} = 132.5$$

$$C.f = 105$$

$$f = 60$$

$$h = 10$$

$$\text{Median} = 79.5 + \left(\frac{132.5 - 105}{60}\right) \times 10$$

$$= 79.5 + 4.58$$

$$= 83.98$$

Example 3: Compute Median from the following data:

Mid value	Frequency
115	6
125	25
135	48
145	72
155	116
165	165
175	175
185	185
195	195

Solution: Since we are given mid-values, we should find out the upper and lower limits of the various classes.

Calculation of Median:

Class Intervals	f	$C.f$
110 - 120	6	6
120 - 130	25	31
130 - 140	48	79
140 - 150	72	151
150 - 160	116	267
160 - 170	60	327
170 - 180	38	365
180 - 190	22	387
190 - 200	3	390

$$\text{Med} = \text{Size of } \frac{N}{2}\text{th item} = \text{Size of } \frac{390}{2} = 195^{th} \text{ item}$$

Median lies in class 150 – 160

l = 150, $N/2$ = 195, C.f. = 151, 6 = 116, i = 10

$$\text{Median} = l + \frac{h}{6}\left(\frac{N}{2} - C\right)$$

$$= 150 + \left(\frac{195-151}{116}\right) \times 10$$

$$= 150 + 3.79 = 153.79$$

Example 4: Find the missing frequency in the following distribution if N is 100 and median is 30.

Marks	No. of students
0 - 10	10
10 - 20	?
20 - 30	25
30 - 40	30
40 - 50	?
50 - 60	10

Solution: Let the frequencies of the class 10 – 20 be X Frequency of the class 40–50 = 100 – (75 + X) or 25–X [75 is the total of 10 + 25 + 30 + 10].

Calculation of Median:

Marks	No. of students	*C.f*
0 - 10	10	10
10 - 20	X	10 + X
20 - 30	25	35 + X
30 - 40	30	65 + X
40 - 50	(25 -X)	90
50 - 60	10	100

$$\text{Median} = l + \frac{h}{f}\left(\frac{N}{2} - C\right)$$

Since median is 30, it lies in the class 20 – 30

$$l = 20, \ \frac{N}{2} = 50, \ \text{C.f are} = (\ 35\ – X), \ f = 30$$

$$30 = 20 + \frac{10}{30}\left(50 - (35 - \times)\right)$$

$$= 20 + \frac{50 - (35 - \times)}{30} \times 10$$

$$= 20 + \frac{50 - 35 + X}{30} \times 10$$

$$= 20 + \left(\frac{15 + X}{30}\right) \times 10$$

$$30 - 20 = \frac{150 + 10X}{30}$$

$$10 = \frac{150 + 10\ X}{30}$$

$$300 - 150 = 10\ X$$

$$150 = 10\ X$$

$$\frac{150}{10} = X$$

$$X = 15$$

Hence the frequency of the class corresponding to 10 – 20 = 15 and the frequency of the class corresponding to 40 - 50 = 25 - X = 25 - 15 = 10

Calculation of Median When class Intervals are unequal

When class intervals are unequal, the frequencies need not be adjusted to make the class intervals equal and the same formula can be applied as discussed above

Example 1: Calculate median from the following data:

Marks	0-10	10-30	30-60	60-80	80-90
No. of students	5	15	30	8	2

Solution: Calculation of Median

Marks	*f*	*C.f*
0 - 10	5	5
10 - 30	15	20
30 - 60	30	50
60 - 80	58	58
80 - 90	60	60

$$\text{Med} = \text{Size of } \frac{N}{2} \text{th item}$$

$$= \text{Size of } \frac{60}{2}$$

$$= 30^{th} \text{ item}$$

Median lies in class 30 – 60

$$l = 30, \quad N/2 = 30, \; C.f. = 20, \; f = 30, \; h = 30$$

$$\text{Med} = 30 + \frac{30 - 20}{30} \times 10$$

$$= 30 + 10 = 40$$

If we make the class intervals equal, the same answer will be obtained.

Marks	f	$C.f$
0 - 10	5	5
10 - 20	7.5	12.5
20 - 30	7.5	20
30 - 40	10	30
40 - 50	10	40
50 - 60	10	50
60 - 70	4	54
70 - 80	4	58
80 - 90	2	68

Med = Size of $\frac{N}{2}$th item

$$= \frac{60}{2} = 30^{th} \text{ item}$$

Median lies in the class 30 – 40

$$\text{Med} = l + \frac{h}{f}\left(\frac{N}{2} - C\right)$$

$$l = 30, \quad \frac{N}{2} = 30, \quad C.f. = 20, \quad f = 10, \quad h = 10$$

$$\text{Med} = 30 + \frac{10}{10}\ (30 - 20)$$

$$= 30 + \frac{10}{10} \times 10 = 30 + 10 = 40$$

Properties of Median

1. Median is a positional average and hence it is not influenced by the extreme values.
2. Median can be calculated even in the case of open end intervals.
3. Median can be located even if the data are incomplete.
4. It is not good representative of data if the number of items is small.
5. It is not amenable to further algebraic treatment.

Merits and Limitations of Median

Merits:

1. It is especially useful in case of open-end classes since only the position and not the values of items must be known. The median is also recommended if the distribution has unequal classes, since it is easier to compute than the mean.
2. It is not influenced by the magnitude of extreme deviations from it. For example, the median of 10, 20, 30, 40 and 150 would be 30 where as the mean is 50. Hence very often when extreme values are present in a set of observations, the median is a more satisfactory measure of the central tendency than the mean.
3. Median is the only average to be used while dealing with qualitative characteristics which can not be measured quantitatively but can still be arranged in ascending or decending order of magnitude e.g. to find the average intelligence, average beauty, average honesty etc., among a group of people.
4. The value of median can be determined graphically whereas the value of mean can not be graphically ascertained.
5. It is rigidly defined.
6. Median is easy to understand and easy to calculate for a non-mathematical person.

7. Perhaps the greatest advantage of median is, however, the fact that the median actually does indicate what many people incorrectly believe the arithmetic mean indicates. The median indicates the value of the middle item in the distribution. This is a clear-cut meaning and makes the median a measure that can be easily explained.

Demerits:

1. For calculating median it is necessary to arrange the data, whereas other averages do not need any such arrangement.
2. Median, being a positional average, is not based on each and every item of the distribution. It depends on all the observations only to the extent whether they are smaller than or greater than it; the exact magnitude of the observations being immaterial. Let us consider a simple example. The median value of 35, 12, 8, 40 and 60 i.e. 8, 12, 35, 40, 60 is 35, now if we replace the value 8 and 12 by any two values which are less than 35 and the values 40 and 60 by any two values greater than 35 the median is unaffected. This property is sometimes described by saying that median is sensitive.
3. It is erratic if the numbers of items is small.
4. The median in some cases can not be computed exactly as can be mean. When the number of items included in a series of data is even, the median is determined approximately as the mid-point of the two middle items.

Check Your Progress–I

Q.1. Pocket allowance of 5 students respectively are:

125, 75, 150, 175, 200

Find out arithmetic mean. (**Ans.** Average Pocket Allowance: Rs.145)

____

____

Q.2. Weight of 15 Persons is as follows:

Weights (kgs) 20, 28, 34, 39, 42, 50, 53, 54, 59, 64, 72, 74, 74, 78, 79

Find out mean weight. (**Ans.** 54.7 kg)

Q.3. Calculate average of the following discrete series.

Size	30	29	28	27	26	25	24	23	22	21
Frequency (f)	2	4	5	3	2	7	1	4	5	7

(**Ans.** Average = 24.925)

Q.4. Calculate arithmetic mean from the given data:

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of students (f)	6	5	8	15	7	6	3

(**Ans.** $\overline{X}$ = 33.4 marks)

Q.5. The value of median for the following data is 46. Find the missing frequencies.

x	10-20	20-30	30-40	40-50	50-60	60-70	70-80
f	12	30	?	65	—	25	18

The total of frequencies is 229. (**Ans**. $x = 34$, $y = 45$)

Q.6. Find out Median value of the following distribution:

Wage Rate	0-10	10-20	20-30	30-40	40-50
No. of workers	22	38	46	35	20

(**Ans.** Median = 24.46)

MODE

It is another measure of central tendency. Mode is a value of a particular type of items which occurs most frequently. For instance, if shoe size no. 7 has maximum demand, size no. 7 is the modal value of shoe sizes.

Definition: Mode is the value which occurs most frequently in a set of observations and around which the other items *of the set* cluster densely. In other words, mode is the value of a series which is predominant in it. In the words of Croxton and Cowden, "*The mode of a distribution is value at the point around which the items tend to be most heavily concentrated. It may be regarded as the most typical of a series of values.*"

According to A.M. Tuttle, '*Mode is the value which has the greatest frequency density in its immediate neighbourhood*'. Accordingly mode may also be termed as the *fashionable value* (a derivation of the French word 'la Mode') of the distribution.

Calculation of mode— Individual Observations

For determining mode count the number of times the various values repeat themselves and the value occurring maximum number of times in the modal value.

Example 1: Calculate the mode from the following data of the marks obtained by 10 students:

X : 10, 27, 24, 12, 27, 27, 20, 18, 15, 30

Solution: Since the item 27 occurs the maximum number of times i.e. 3 hence the modal marks are 27. When there are two or more values having the same maximum frequency, one cannot say which is the modal value and hence mode is said to be ill-defined. Such a series is also known as bimodal or multimodal.

Calculation of Mode in Discrete Series:

Variate value (X)	3	4	7	8	9	11	12
Frequency (f)	2	6	5	14	10	6	3

Clearly $x = 8$ has maximum frequency 14. Hence 8 is the modal value.

If there is an irregular distribution that is, if the trend of frequency changes all of a sudden for certain value(s), the variate value or class corresponding to the maximum frequency should not be accepted as mode.

Example 1: The distribution of marks of 174 students out of 25 marks is

Marks (X)	3	4	6	7	9	10	13	15	18	20
Frequency (f)	4	8	15	20	32	16	14	35	10	6

The given distribution is an irregular distribution because the frequencies are increasing up to the value of X = 9 and then gradually decreasing except for X=15 corresponding to which f = 35. This frequency is not consistent with the trend of data. Hence to consider X = 15 as mode is not proper.

Therefore, a better modal value can be worked out by the method of grouping. Various steps involved in the method of grouping are :

(i) Write the variate values, in order, in column (1) and the frequencies corresponding to them in column (2).

(ii) Add frequencies in Pairs starting from the first and place them in a position between the two frequencies in column (3).

(iii) Omit the first frequency and repeat the step (ii) and place the added frequencies in column (4).

(iv) Again group the frequencies in three's starting from the first and place sum of each group against the mid-frequency in column (5).

(v) Leave first frequency and repeat step (iv) creating column (6).

(vi) Again leave first two frequencies and repeat step (iv) placing the added values in column (7). Draw brackets in each column against added frequencies for two's or three's

(vii) Parenthesise the maximum frequency of each column.

The end frequencies which is not used in grouping are left out. For any distribution, the above mentioned seven columns are to be created.

Marks	*Frequency*	*Add frequencies*				
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
3	4	12		27		
4	8		23		43	
6	15	35				67
7	20		(52)	(68)		
9	32	48			(62)	
10	16		30			
13	14	(49)		59		65
15	(35)		45		51	
18	10	16				
20	6					

Once the frequency table is prepared, another table known as analysis table has to be prepared. In this table, the variate values are written in the caption (column heads) and column numbers along the sub-head. In the body of the table, give value 1 to each of the variate values which has been summed up in constituting the maximum frequency. Total 1's of each column of the analysis table.

The variate value having the maximum column total is the mode.

Analysis table

Columns	Marks (X)									
	3	4	6	7	9	10	13	15	18	20
(2)								I		
(3)							I	I		
(4)				I	I					
(5)				I	I	I				
(6)					I	I	I			
(7)			I	I	I					
Total			I	3	4	2	2	2		

In the above table for X = 9, the maximum sum of I's is 4, hence the modal value is 9

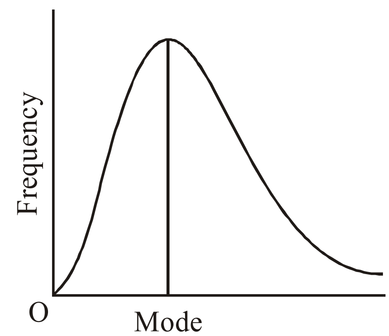
Remarks:

(1) It is worth pointing out that by inspection one would have concluded that the mode is 15 as it has maximum frequency 35, however this not correct, as revealed by the analysis table.

(2) The given distribution is unimodal as it has only one modal value. There may be two or more columns having equal maximum frequency in the analysis table. In such case, each corresponding variate value would have been taken as bi modal and with more than two modes is known as multimodal.

Mode of a Continuous Distribution

If the distribution is with continuous class intervals, mode can be easily calculated in the manner described here. One must take care that the distribution is continuous and in order (ascending or descending). The class intervals for all the classes are equal. If they are unequal, they should be made equal presuming that the frequencies are uniformly distributed throughout the class interval.



Let the grouped frequency distribution be as follows:

Classes	Frequency
X_1 - X_2	f_1
X_2 - X_3	f_2
X_3 - X_4	f_3
!	
X_{P-1} - X_P	f_{P-1}
X_P - X_{P+1}	f_P
X_{P+1} - X_{P+2}	f_{P+1}
!	!
!	!
X_K - X_{K+1}	f_K

Assume that the distribution is in order and the maximum frequency is fP. Then, the modal class is $X_P - X_{P+1}$.

The exact value of mode can be found out by the interpolation formula

$$M_o = X_P + \frac{f_P - f_{P-1}}{(f_P - f_{P-1}) + (f_P - f_{P+1})} (X_{P+1} - X_P) \qquad ...(1)$$

If we denote the lower limit of the modal class by L_O, the maximum frequency by f, the frequency preceding f by f_{-1} and following f by f_{+1} and the class interval by I, then

The formula (1) can be written as

$$M_o = L_o + \frac{f - f_{-1}}{(f - f_{-1}) + (f - f_{+1})} \times I$$

Putting $f - f_{-1} = \Delta_1$ $f - f_{+1} = \Delta_2$, the formula for mode is

$$M_o = L_o + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times I$$

Example 1: We find the mode of the frequency distribution given below:

Classes (wt in kg.)	Number of children
2.0 - 2.4	5
2.4 - 2.8	5
2.8 - 3.2	9
3.2 - 3.6	4
3.6 - 4.0	4
4.0 - 4.4	3

Obviously by inspection the modal class is (2.8 - 3.2) we calculate mode by the formula:

$$M_o = L_o + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times I$$

In the above case

$$L_o = 2.8 \quad \Delta_1 = 9 - 5 = 4, \; \Delta_2 = 9 - 4 = 5$$

$$I = 3.2 - 2.8 = 0.4$$

$$\text{Hence } M_o = 2.8 + \frac{4}{4+5} \times 0.4$$

$$= 2.8 + 0.178 = 2.978 \text{ kg}$$

In case the frequency distribution is such that the modal class can not be ascertained merely by inspection, the method of grouping should be adopted.

Many frequency distributions have more than one mode. But we are interested in a single central value and hence for such distributions mode is considered as an ill-defined measure of central tendency.

For a moderately skewed or asymmetrical frequency distribution, mode can be calculated by Karl Pearson's empirical formula

Mean – Mode = 3 (Mean – Median)

[Mode = 3 Median – 2 Mean] ... (2)

In case where mode is ill-defined, formula (2), can be used to determine the modal value.

Merits of mode

(1) It is easy to understand.

(2) It is simple to calculate and locate.

(3) If quantitative data in ranking is possible, mode is very useful.

(4) It is not affected by extreme items. Hence it is more useful as compared with mean when there are extreme observations.

(5) It can be calculated even in case of open ended classes.

Demerits

(1) It is not rigidly defined. It is ill-defined, when maximum frequency is repeated or maximum frequency occurs either in the beginning or in the end. Many a time, we come across an irregular distribution. We have bimodal or multimodal distributions in this case. Here the mode is not a representative measure of location.

(2) It is not based on all the observations of the series.

(3) It is not suitable for further mathematical treatment e.g., it is not possible to find the combined mode of two or more series.

Which is the best average?

There are many measures of central tendency like arithmetic mean, median, mode etc. Which of these is the best average? There is no clear-cut answer to it. Different averages are suitable for different situations.

However, while selecting the relevant average, the following points must be kept in mind:

1. **Objective:** Selection of average must conform to the objective of study. For instance, if all values are to be given equal importance, arithmetic mean will be the most appropriate. However, if the value occurring most frequently in a series in to be identified, mode will be the most relevant.
2. **Number of variables:** In case the number of variables in a series is very small, arithmetic mean is the best measure of the central tendency of the series.
3. **Distribution of Items and Frequency:** If the value of large number of items in a series is small, but that of one or two items is large, then arithmetic mean may not be useful. If most of the values are located at the middle of the series or related to qualitative facts, then the use of median should be the best option.
4. **Importance to the Highest and the lowest Items:** If no importance is to be attached to the highest and the lowest items of a series, then the use of median or mode should be the most suitable.
5. **Types of Series:** If in a series large number of items are similar to each other, then the use of mode is not suitable.

Conclusion: Median and Mode are not based on all items of the series. Only arithmetic mean is based on all items of the series. Also only arithmetic mean is capable of further algebraic treatment; mode and median are not.

Geometric Mean

In algebra, geometric mean is calculated in case of geometric Progression, but in Statistics we need not bother about the progression. Here, it is the particular type of data for which the geometric mean is of importance because it gives a good mean value. If the variate values are measured as ratios, proportions or percentages, geometric mean gives a better measure of central tendency than other means.

Definition: Geometric mean of 'N' variate values is the Nth root of their product. Like arithmetic mean it also depends on all observations. It is affected by the extreme values but not to the extent of average. However there is one great drawback in it, that it cannot be calculated if any one or more values are zero or negative. In case an even number of observations are negative, an absurd value of geometric mean will be available from a practical point of view. Hence, if there is a zero or negative value in the set of variate values, it should not be used.

Suppose X_1, X_2 ... X_N are N variate values, then the geometric mean is given as

$$G = n\sqrt{(X_1)(X_2)...(X_N)}$$

Where X_1, X .... X_n are various value of the series 'n' is the number of items. If we have only two observations, e.g. 49 & 25, their

$$GM = \sqrt{49 \times 25} = \sqrt{1225} = 35$$

But if n>2, n^{th} root can't be calculated easily. To make the calculation of finding out n^{th} root, simpler logarithms are used. G.M. by using logs is found and thus

$$G.M = \text{Antilog of } \frac{\sum \log X}{N}$$

Merits:

1. It is based on all the values in the series.
2. It is rigidly defined.
3. It is capable of being applicable to study social and economic phenomena.
4. It is suitable for further mathematical treatment.
5. It is not affected much by extreme values.
6. It is not affected much by fluctuations of sampling.
7. It is very useful for calculation of average rate of growth of population, rate of profits, sales, construction of index numbers etc.

Demerits:

1. If any value is zero, G.M. can not be calculated because (X_1) (X_2) .... (X_n) is to be found. If any of these is zero, the multiplication result will be zero and interpretation would because impossible.
2. Knowledge of logarithms is essential. Therefore it is found difficult to compute by a person with no background in Mathematics.
3. The value of G.M. obtained may not be there in the series, therefore it cannot be termed as the true representative of the data.

Properties of Geometric Mean

The following are two important mathematical properties of Geometric Mean.

1. The product of the value of series will remain unchanged, when the value of geometric mean is substituted for each individual value. For example,

the geometric mean for series 2, 4, 8 is 4: therefore we have

$$2 \times 4 \times 8 = 64 = 4 \times 4 \times 4$$

2. The sum of the deviations of the logarithms of the original observations above or below of the logarithm of the geometric mean is equal. This also means that the value of the geometric mean is such as to balance the ratio deviations of the observations from it. Thus using the same previous numbers, we find that

$$\left(\frac{4}{2}\right)\left(\frac{4}{4}\right) = 2 = \left(\frac{8}{4}\right)$$

Because of this property, this measure of central value is especially adapted to average ratios, rates of change and logarithmically distributed series.

Calculation of Geometric Mean— Individual Observations

$$\text{G.M.} = \text{Antilog}\left(\frac{\Sigma \log X}{N}\right)$$

Steps:

1. Take the logarithmas of the variable X and obtain the total Σ log X.
2. Divide Σ log X by N and take the antilog of the value so obtained. This gives the value of Geometric Mean.

Example: Calculate Geometric Mean from the following data:

125, 1462, 38, 7, 0.22, 0.08, 12.75, 0.5

Solution: Calculation of G.M.

X	log X
125	2.0969
1462	3.1650
38	1.5798
7	0.8451
0.22	1.3424
0.08	2.9031
12.75	1.1055
0.5	1.6990

$$\text{G.M.} = \left(\frac{\Sigma \log X}{N}\right) = \text{A.L.} \left(\frac{6.7368}{8}\right)$$

$$= \text{A.L. } (0.8421) = 6.952$$

Calculation of Geometric Mean-Discrete Series

$$\text{G.M.} = \text{Antilog} \left(\frac{\Sigma f \log X}{N}\right)$$

Steps:

(i) Find the logarithms of the variable X.

(ii) Multiply these logarithms with the respective frequenciest obtain the total Σf log X.

(iii) Divide Σf log X by the total frequency and take the antilog of the value so obtained

$$\text{G.M.} = \frac{A.L. \Sigma f \log X}{N}$$

Calculation of Geometric Mean— Continuous Series

$$\text{G.M.} = \text{Antilog} \left(\frac{\Sigma f \log X}{N}\right)$$

Steps:

(i) Find out the mid-points of the classes and take their logarithms.

(ii) Multiply these logarithms with the respective frequencies of each class and obtain the total $\Sigma f \log m$.

(iii) Divide the total obtained in step (ii) by the total frequency and take the antilog of the value so obtained.

Example: Find the Geometric mean for the data given below:

Marks	Frequency
4–8	6
8–12	10
12–16	18
16–20	30
20–24	15
24–28	12
28–32	10
32–36	6
36–40	2

Solution: Calculation of Geometric Mean

Marks	M.P.	f	log m	$f \times$ log m
4–8	6	6	0.7782	4.6692
8–12	10	10	1.0000	10.0000
12–16	14	18	1.1461	20.6298
16–20	18	30	1.2553	37.6590
20–24	22	15	1.3424	20.1360
24–28	26	12	1.4150	16.9800
28–32	30	10	1.4771	14.7710
32–36	34	6	1.5315	9.1890
36–40	38	2	1.5798	3.1596
		N = 109		$\Sigma f \times$ log m=137.1936

$$\text{G.M.} = \text{A.L.} \left(\frac{\Sigma f \log m}{N}\right) = \left(\frac{137.1936}{109}\right)$$

$$= \text{A.L.} \ (1.2587) = 18.14$$

Harmonic Mean: The Harmonic Mean is based on the reciprocals of the numbers averaged. It is defined as the reciprocal of the arithmetic mean of the reciprocal of the individual observation. Thus by definition

$$\text{H.M.} = \frac{N}{\left(\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} + --+\frac{1}{X_n}\right)}$$

When the number of items is large the computation of harmonic mean in the above manner becomes tedious. To simplify calculations, we obtain reciprocal of the various items from the table and apply the following formula :

In individual observations, $\text{H.M} = \frac{N}{\Sigma\left(1/X\right)}$

In discrete series, $\text{H.M} = \frac{N}{\Sigma\left(f \times \frac{1}{X}\right)}$

In continuous series $\text{H.M.} = \frac{N}{\Sigma\left(f \times \frac{1}{m}\right)}$

m – mid point.

Example: From the following data compute the value of Harmonic mean:

Marks	10	20	25	40	50
No. of students	20	30	50	15	5

Solution: Calculation of Harmonic mean:

Marks (X)	f	f/X
10	20	2.000
20	30	1.5000
20	50	2.000
40	15	0.375
50	5	0.100
	N = 120	$\Sigma\left(\frac{f}{X}\right) = 5.975$

$$\text{H.M.} = \frac{N}{\Sigma\left(\frac{f}{X}\right)} = \frac{120}{5.975} = 20.08$$

Merits and Limitations of Harmonic Mean

Merits:

1. Its value is based on every item of the series.
2. It lends itself to algebraic manipulation.
3. In problems relating to time and rates it gives better results than other averages.
4. It is rigidly defined.
5. It is not affected much by fluctuations of sampling.

Limitations:

1. It is not easily understood
2. It is difficult to compute.
3. It gives largest weight to smallest items. This is generally not a desirable feature and as such this average is not very useful for the analysis of economic data.
4. Its value can not be computed when there are both positive and negative items in a series or when one or more items are zero.

Because of these limitations the harmonic mean has very little practical application and is not a good representation of a statistical series, unless the phenomenon is such where small items need to be given a very high weightage.

Check Your Progress-II

Q.1. Given below is the data of age of 9 children of street. Find the median: 5, 8, 7, 3, 4, 6, 2, 9, 1 (**Ans.** Median = 5)

____

____

____

____

Q.2. Find the median of the following values:
30, 20, 15, 10, 25, 35, 18, 21, 28, 40, 36 (**Ans.** Median = 25)

Q.3. Find median of the series of the following table:

Item	3	4	5	6	7	8
Frequency	6	9	11	14	23	10

(**Ans.** Median = 6)

Q.4. Find out Median of the following series:

Size	15	20	25	30	35	40
Frequency	10	15	25	5	5	20

(**Ans.** 25)

Q.5. Find out Median wage rate from the following data-set.

Wage Rate	5-15	15-25	25-35	35-45	45-55	55-65
No. of workers	4	6	10	5	3	2

(**Ans.** Rs. 30)

Q.6. Calculate Mode of the following data:

Wage	25	50	75	80	85	90
No. of workers	4	6	9	3	2	1

(**Ans.** Mode (Z) = 75)

Q.7. Find out mode of the following data:

Class Interval	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No. of children	4	5	3	2	6	7	3

(**Ans.** Mode (Z) = 31)

Q.8. Calculate the Mean, Median and Mode of the number of persons per house in a village with the help of the following information:

No. of persons per house	1	2	3	4	5	6	7	8	9	10
No. of Houses	26	113	120	95	60	42	21	14	5	6

(**Ans.** $\overline{X} = 3.78$, $M = 3$, $Z = 1.44$)

Q.9. Calculate the Median and Mode from the following data:

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of workers	2	18	30	45	35	20	6	3

(**Ans.** $Z = 36$, $M = 36.55$)

Q.10. Calculate mode when arithmetic mean is 146 and median is 130.

(**Ans.** $Z = 98$)

Q.11. If mode is 63 and median is 77, calculate arithmetic mean. (**Ans.** $\overline{X}$ = 84)

9.4. LET US SUM UP

In this lesson, we discussed the different types of averages and their merits and demerits.

9.5 EXAMINATION ORIENTED QUESTIONS

Short Answer Type Questions

Q.1. What do you mean by 'Average'? State merits and Demerits of Arithmetic mean.

Q.2. What are the different kinds of statistical averages?

Q.3. Pocket expenditure of 6 students is respectively, (Rs.) 6, 12, 18, 24, 30 and 36. Find out the arithmetic mean.

Q.4. Give formula of calculating arithmetic average of a continuous series and discuss merits of A.M.

Q.5. What do you mean by central tendency of a series?

Q.6. Define and explain Arithmetic mean.

Q.7. Show that the sum of deviations of the values of the variable from their arithmetic mean is equal to zero.

Q.8. State four merits of arithmetic average.

Q.9. State four demerits of arithmetic average.

Q.10. What are the four objectives of a statistical average ?

Q.11. Define Median. State four merits of Median.

Q.12. Define Mode. State five demerits of Mode.

Q.13. Give formula for finding out Median and State four demerits of median.

Q.14. Give formula for estimating mode and State three merits of Mode.

Q.15. Give formula for finding out median of a continuous series.

Q.16. What is a Positional average?

Q.17. What are the uses of Mode?

Q.18. Mention four merits of Median.

Q.19. State four demerits of Median.

Q.20. State four merits of Mode.

Q.21. State four demerits of Mode.

Q.22. What is meant by central Tendency? Discuss the various methods of measuring it and point out merits and demerits of Mean.

Q.23. Why is the Arithmetic Mean the most commonly used measure of central tendency?

Q.24. Define Median. What are its merits and demerits?

Q.25. Explain the concepts of mode and Median. Also state their relative merits and demerits.

Q.26. What is Mode? Discuss the merits, demerits and uses of the mode.

Q.27. Define Geometric Mean and discuss its merits and demerits. Where is its use recommended?

Q.28. Calculate GM of the following data :

1, 7, 18, 65, 91 and 103 (Hint : Ans. 20.62)

Q.29. Define Harmonic Mean and discuss its merits and demerits. Under what situations, would you recommend its use?

Q.30. Find out Harmonic Mean from the following :

2574, 465, 75, 5, 0.8, 0.08, 0.005, 0.0009 (Hint Ans. 0.00604)

9.6. SUGGESTED READINGS & REFERENCES

1. S.P. Gupta (2005), Statistical Methods, S. Chand and Sons, New Delhi
2. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas publishing House, New Delhi.
3. Monga, G.S. (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.
4. Gupta, S.C. and Indra Gupta, Business Statistics.

MEASURES OF DISPERSION

LESSON NO. 10 **UNIT-III**

STRUCTURE

10.1. OBJECTIVES

After going through this lesson, you would be able to understand.

1. Meaning of Dispersion
2. Purpose of Measuring Disperion
3. Properties of Good measure of Dispersion
4. Concepts of Range, Interquartile Range, Quartile Deviation (Q.D.), Mean Deviation, Standard Deviation Coefficient of Variation

10.2. INTRODUCTION

The term dispersion refers to the variability of the size of items. Dispersion explains that the size of various items in a series are not uniform. Father, they vary. For example, if in a series the lowest and highest values vary only a little, the dispersion is said to be low. But if this variation is very high, dispersion is said to be considerable. In a series of ten students, the marks obtained are

10, 6, 8, 5, 10, 10, 8, 10, 5, 8 (the average = 8). In another class, 10 students obtained the following marks, 10, 10, 5, 2, 10, 10, 3, 10, 10 (the average = 8). The dispersion in the second case is more because the size of items in this series vary considerably, inspite of the fact that the averages of the two have come out to be 8. Where as the measures of central tendency provide a central value of the distribution, the measures of dispersion are required to measure the amount of variation (dispersion or Scatter) of values about the central value. For example suppose the monthly income (in rupees) of five households are—

Monthly income in Rupees

4500	6000	5500	3750	4700

The arithmetic mean of income is Rs. 4,890 and the median income Rs.4,700

The amount of variation in incomes is shown by deviations from the central value.

Deviations from the Arithmetic Mean

–390	1110	610	-1140	-190

Deviations from the Median

–200	1300	800	-950	0

We observe that some deviations are positive and others are negative. Also some deviations are large and others are small. We require an over all summary measure of variation in all values about the central value. This summary measure is called the measure of dispersion.

10.3. DEFINITIONS

A.L. Bowley defines dispersion, as - "Dispersion is the measure of variations of the item".

Prof. L.R. Connor defines dispersion as "Dispersion is a measure of the extent to which the individual items vary".

According to Spriegel, "The degree to which numerical data tend to spread about an average value is called the variation or dispersion of data".

All the above definitions suggest that the term dispersion refers to the variability in the size of items. This variability is measured with respect to the average of the series. Therefore measures of dispersion are also termed as averages of the second order.

Commonly used measures of dispersion are—

1. Range
2. Inter quartile range and Quartile deviation
3. Mean deviation
4. Variance
5. Standard deviation
6. Coefficient of variation

Before giving the details of various measure of dispersion, it is worthwhile to discuss the purposes of these measures and enunciate the properties of a good measure of dispersion.

10.3.1. PURPOSES / OBJECTIVES OF MEASURING DISPERSION

1. **Gateway to other statistical measures:** Measures of dispersion, especially variance and standard deviation lead to many statistical techniques like correlation, regression, analysis of variance, etc.
2. **To have an idea about the reliability of central value:** In a way, it is the measure of degree of scatteredness. If Scatter is large, an average is less reliable. If the value of dispersion is small, it indicates that a central value is a good representative of all the values in the set.
3. **To compare two or more sets of values with regard to their variability:** Two or more sets can be compared by calculating the same measure of dispersion having the same unit of measure. A set with smaller value possesses lesser variability.
4. **To Provide information about the structure of a series:** A value of measure of dispersion gives an idea about the spread of the observations.

10.3.2. PROPERTIES OF A GOOD MEASURE OF DISPERSION

1. It should be simple to understand.
2. It should be least affected by extreme values.
3. It should be calculated with reasonable ease, i.e., the formula should be such that if does not complicate the computation of a measure of dispersion.

4. It should be capable of further algebraic treatment.

5. It should be based on all the items of the series.

Range: The range is defined as the difference between the largest and the smallest value of the variable in the given set of values. Suppose, the values of X are arranged in ascending order as

$$X_1 < X_2 < \dots < X_n$$

So that X_n is the largest and X_1 is the smallest value, then the range is defined as

$$R = X_n - X_1$$

All values, of X lie with in the range.

If the range is large, the spread of values is large and hence the variation of values of X is large. On the other hand, if the range is small, the spread of values of X is small and the variation of X is small.

Definition

It is the difference between the largest and the smallest observation in a set.

If we denote the largest observation by L and the smallest observation by S the formula is

Range $R = L - S$

Coefficient of range: A relative measure known as coefficient of range is given as:

$$\text{Coefficient of range} = \frac{L-S}{L+S}$$

Hence range or coefficient of range, better the result.

Example 1: Calculate Range and Coefficient of Range of the following data

4, 7, 8, 46, 53, 77, 8, 1, 5, 13

Solution: Range $L - S = 77 - 1 = 76$

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{77-1}{77+1} = \frac{76}{78} = 0.97$$

Example 2: Calculate coefficient of Range from the following data:

Marks	No. of students
10 - 20	8
20 - 30	10
30 - 40	12
40 - 50	8
50 - 60	4

Solution: Coefficient of Range = Range = L – S = 60 – 10 = 50

$$\frac{L-S}{L+S} = \frac{60-10}{60+10} = \frac{50}{70} = 0.714$$

Characteristics of the Range

Let us discuss the following characteristics of the range:

1. It is rigidly defined.
2. It is easy to calculate and simple to interpret.
3. It does not depend on all values of variable.
4. The range depends on the units of measurement of variable. It has the same units as of the variable considered. The range of income is in terms of rupees, the range of distances may be in kilometers etc.

Inter quartile Range (I.R)

Definition: The difference between the third quartile and first quartile is called interquartile range. Symbolically

$$I.R. = Q_3 - Q_1$$

Quartile Deviation (Q.D.)

This is half of the inter quartile range i.e.

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

Also the coefficient of quartile deviation is given by the formula:

Coeff. of Q.D.= $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

Computation of Quartile Deviation

The process of computing quartile deviation is very simple since we have to just compute the values of the upper and lower quartiles. The following illustrations would clarify calculations.

Individual observations

Example 1: Find out the value of quartile deviation and its coefficient from the following data:

Roll No.	1	2	3	4	5	6	7
Marks	20	28	40	12	30	15	50

Solution: Calculation of Quartile Deviation:

Marks arranged in ascending order: 12, 15, 20, 28, 30, 40, 50

$$Q_1 = \text{Size of } \frac{N+1}{4}\text{th item}$$

$$= \text{Size of } \frac{7+1}{4} = 2^{nd} \text{ item}$$

Size of 2^{nd} item is 15 . Thus $Q_1 = 15$

$$Q_3 = \text{Size of } 3\left(\frac{N+1}{4}\right)\text{th item} = \text{Size of } \left(\frac{3 \times 8}{4}\right)\text{th item} = 6^{th} \text{ item}$$

Size of 6^{th} item is 40 thus $Q_3 = 40$

$$\text{Q.D.} = \frac{Q_3 - Q_1}{2} = \frac{40-15}{2} = 12.5$$

$$\text{Coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{40-15}{40+15} = \frac{25}{55} = 0.45$$

Discrete Series

Example 1: Compute coefficient of quartile deviation from the following data:

Marks	10	20	30	40	50	80
No. of students	4	7	15	8	7	2

Solution: Calculation of coefficient of Quartile Deviation

Marks	Frequency	*C.f.*
10	4	4
20	7	11
30	15	26
40	8	34
50	7	41
80	2	43

$$Q_1 = \text{Size of } \frac{N+1}{4}\text{th item}$$

$$= \frac{43+1}{4}\text{th item} = 11^{th} \text{ item}$$

Size of 11^{th} item is 20. Thus $Q_1 = 20$

$$Q_3 = \text{Size of } 3\left(\frac{N+1}{4}\right)\text{th item} = \frac{3\times 44}{4} = 33^{rd} \text{ item}$$

Size of 33^{rd} item is 40 Thus $Q_3 = 40$

$$\text{Q.D.} = \frac{Q_3 - Q_1}{2} = \frac{40-20}{2} = 10$$

$$\text{Coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{40-20}{40+20} = 0.33$$

Continuous Series

Example 1: In the following table the figures relating to the wages of 60 workers in a factory are given. Calculate the Quartile Deviation (Q.D.) and its coefficient.

Wages (in Rs.)	No. of workers
20 - 25	2
25 - 30	10
30 - 35	25
35 - 40	16
40 - 45	7

Solution: Q.D. $= \dfrac{Q_3 - Q_1}{2}$

Calculation of Quartile Deviation and its coefficient

Wages (in Rs.)	No. of workers *(f)*	*C.f*
20 - 25	2	2
25 - 30	10	12
30 - 35	25	37
35 - 40	16	53
40 - 45	7	60
	N = 60	

Calculation of Q_1:

$$Q_1 = \text{Size of } \frac{N}{4}\text{th item} = {}^{60}\!/_{4} = 15\text{th item}$$

Q_1 lies in the class 30 – 35

$$Q_1 = L + \frac{{}^{N}\!/_{4} - C.f}{f} \times i$$

$L = 30, \ N/4 = 15, \ C.f = 12 \ f = 25, \ i = 5$

$$Q_1 = 30 + \frac{15-12}{25} \times 5$$

$$= 30 + 0.6 = 30.6$$

Calculation of Q_3:

$$Q_3 = \text{Size of } \frac{3N}{4} \text{th item} = \frac{3 \times 60}{4} = 45\text{th item}$$

Q_3 lies in the class 35 – 40

$$Q_3 = L + \frac{3N/4 - C.f}{f} \times i$$

$$L = 35, \ \frac{3N}{4} = 45, \ C.f = 37, \ f = 16, \ I = 5$$

$$Q_3 = 35 + \frac{45-37}{16} \times 5 = 35 + 2.5 = 37.5$$

$$Q.D. = \frac{Q_3 - Q_1}{2} = \frac{37.5 - 30.6}{2} = 3.45$$

$$\text{Coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{37.5 - 30.6}{37.5 + 30.6} = 0.101$$

The Mean Deviation

Suppose $X_1, X_2 - X_n$ are n values of variable X and $\overline{X} = \frac{1}{n}\sum X_i$ is their arithmetic mean the deviations from $\overline{X}$ are given by

$$X_1 - X_1, \quad X_2 - \overline{X}, \ .... \ X_n - \overline{X}$$

Some of these deviations will be Positive and others negative. Ignoring the sign of deviations,

$$\left|X_1 - \overline{X}\right|, \left|X_2 - \overline{X}\right|, --\left|X_n - \overline{X}\right|$$

are the absolute values of the deviations where the two Parallel bars | | indicate that the absolute value is taken (This is also called the modulus value).

The arithmetic mean of the absolute deviations is called the mean deviation or mean absolute deviation. Thus

$$\frac{1}{n}\Sigma\left|X_1 - \overline{X}\right|$$

Is the mean deviation of X about the arithmetic mean.

Properties

1. Mean deviation removes one main objection of the earlier measures, that is it involves each value of the set.
2. It is not affected much by extreme values.
3. Its main drawback is that algebraic negative signs of the deviation are ignored which is mathematically unsound.
4. Mean deviation is minimum when the deviation are taken from median.

Note: If the deviations are taken from mean and the signs of the deviations are taken into consideration, the sum of the deviation is zero i.e., $\Sigma\left(X_i - \overline{X}\right) = 0$

Individual Series

Calculate Mean Deviation from Mean and Mean Deviation from Median.

Example 1: Calculation of Mean Deviation from Arithmetic Mean and Median also find coefficient of Mean Deviation:

S.No.	1	2	3	4	5	6	7	8	9
Wage	40	42	45	47	50	51	54	55	57

Solution: Calculate of Mean Deviation from Arithmetic Mean and Coefficient of Mean Deviation.

S.No.	Wage (Rs.)	Deviation from Arithmetic Mean
1	40	| 40 - 49 | = | - 9 | = 9
2	42	| 42 - 49 | = | - 7 | = 7
3	45	| 45 - 49 | = | - 4 | = 4
4	47	| 47 - 49 | = | - 2 | = 2
5	50	| 50 - 49 | = | - 1 | = 1
6	51	| 51 - 49 | = | - 2 | = 2
7	54	| 54 - 49 | = | - 5 | = 5
8	55	| 55 - 46 | = | - 6 | = 6
9	57	| 57 - 49 | = | - 8 | = 8

$$N = 9, \ \Sigma X = 441$$

$$\overline{X} = \frac{\sum X}{N} = \frac{441}{9} = 49 \text{ Rupees}$$

$$\text{Mean Deviation} = \frac{1}{N}\Sigma\left(X_i - \overline{X}\right) = \frac{1}{9}(44) = 4.89 \text{ Rupees}$$

$$\text{Coefficient of Mean Deviation} = \frac{\frac{1}{N}\Sigma(X_1 - \overline{X})}{\overline{X}} = \frac{4.89}{49} = 0.09$$

(ii) Calculation of Mean Deviation and Coefficient of Mean Deviation from Median

S.No.	1	2	3	4	5	6	7	8	9
Wage	40	42	45	47	50	51	54	55	57
|(X-M)|	10	8	5	3	0	1	4	5	7

$$\text{Median} = \text{Size of } \left(\frac{9+1}{2}\right)\text{th item} = 5^{th} \text{ item} = 50 \text{ Rupees}$$

Mean Deviation from Median = $\frac{\Sigma|X - \text{Median}|}{N}$ = $\frac{43}{9}$ = 4.78 Rupees

Coefficient of Mean Deviation from Median = $\frac{\frac{\Sigma|X - \text{Median}|}{N}}{\text{Median}}$

$= \frac{4 \cdot 78}{50} = 0 \cdot 096$

Calculation of Mean Deviation - Discrete Series

In discrete series the formula for calculating mean deviation is

$$\text{M.D.} = \frac{\Sigma f|D|}{N}$$

|D| denotes deviation from median ignoring signs

Steps:

(i) Calculate the median of the series.

(ii) Take the deviations of the items from median ignoring signs and denote them |D|

(iii) Multiply these deviation by the respective frequencies and obtain the total $\Sigma f|D|$

(iv) Divide the total obtained in step (ii) by the number of observations. This gives us the value of mean deviation.

Example 2: Calculate mean deviation from the following series:

X	10	11	12	13	14
f	3	12	18	12	3

Solution: Calculate the Mean Deviation

X	*f*	|D|	*f* | D⁻ |	*C.f.*
10	3	2	6	3
11	12	1	12	15
12	18	0	0	33
13	12	1	12	45
14	3	2	6	48
	N = 48		$\Sigma f\|D\|=36$	

$$\text{M.D} = \frac{\Sigma f|D|}{N}$$

$$\text{Median} = \text{Size of } \left(\frac{N+1}{2}\right)^{th} \text{ item}$$

$$= \frac{48+1}{2} = 24.5^{th} \text{ item}$$

Size of 24.5^{th} item is 12

Hence Median = 12

$$\text{M.D.} = \frac{36}{48} = 0.75$$

Calculation of Mean Deviation - Continuous Series

For calculating mean deviation in continuous series the procedure remains the same as in discrete Series. The only difference is that here we have to obtain the mid point of the various classes and take deviations of these points from median. The formula is the same i.e.

$$\text{M.D.} = \frac{\Sigma f|D|}{N}$$

Example 3: Calculate Mean deviation from mean from the following data:

Marks	No. of students
0 - 10	4
10 - 20	6
20 - 30	10
30 - 40	20
40 - 50	10
50 - 60	6
60 - 70	4

Solution: Calculation of Mean deviation from Mean

Marks	(X) Mid-Point	f	fx	|0| $\|X-\overline{X}\|$	$f\|D\|$
0 - 10	5	4	20	30	120
10 - 20	15	6	90	20	120
20 - 30	25	10	250	10	100
30 - 40	35	20	700	0	0
40 - 50	45	10	450	10	100
50 - 60	55	6	330	20	120
60 - 70	65	4	260	30	120
		$\Sigma f = 60$	Σfx=2100		$\Sigma f\|D\|$=680

$$\overline{X} = \frac{\sum fx}{\sum f} = \frac{2100}{60} = 35$$

$$\text{M.D.} = \frac{\sum f|D|}{N} = \frac{680}{60} = 11.33$$

Merits and Demerits of Mean Deviation

Merits:

1. **Simple:** It is very simple and easy measure of dispersion.
2. **Based on all values:** Mean deviation is based on all the items of the series. It is therefore more representative than the range or quartile.
3. **Less effect of Extreme Values:** Mean deviation is less affected by extreme values than the range.

Demerits:

1. **In accuracy:** Calculation of mean deviation suffers from lack of accuracy, because the + or - signs are ignored.
2. **Not capable of Algebraic Treatment:** Mean deviation is not capable of any further algebraic treatment.
3. **Unreliable:** In case deviations are taken from mode and mode being uncertain, mean deviation also become uncertain and therefore unreliable.

Standard Deviation

Standard Deviation is a most satisfactory scientific method of dispersion. Standard Deviation concept was introduced by Karl Pearson in 1893. This is sometimes called as "Root Mean Square Deviation'. This is generally denoted by (σ) of the Greek language. Standard Deviation is the square root of the arithmetic mean of the squares of deviation of the items from their mean value. Standard deviation has two main features:

(I) The value of its deviations is taken from arithmetic mean.

(II) Plus and minus signs of the deviations become redundant once the deviations are squared. Finally, square root of the arithmetic mean of the squares of the deviation is calculated. It is this square root which is called Standard Deviation. This is always in positive value.

In the words of Spiegel, "The standard deviation is the square root of the arithmetic mean of the squares of all deviations, deviations being measured from arithmetic mean of the items".

Coefficient of Standard Deviation

This is a relative measure of the dispersion of series. It is generally used whenever variation in different series is compared. Coefficient of Standard Deviation is estimated by dividing the value of Standard Deviation by the mean of the series thus

$$\text{Coefficient of Standard Deviation} = \frac{\sigma}{\overline{X}}$$

Individual Series and Standard Deviation

In case of individual observations standard deviation may be computed by applying any of the following methods:

1. By taking deviation of the items from the actual mean.
2. By taking deviations of the items from an assumed mean.

1. Deviations takes from Actual Mean: When deviations are taken from actual mean the following formula is applied

$$\sigma = \sqrt{\frac{\Sigma X^2}{N}}$$

$$X = \left(X - \overline{X}\right)$$

Step:

(1) First of all mean value of the concerned series is determined. That is we find out $\overline{X}$ as $\overline{X} = \frac{\Sigma X}{N}$

(2) Deviation of each item $\overline{X}$ is determined. That is, we find the values of X as X = $X - \overline{X}$

(3) Each value of the deviation is squared. The sum total of the square of the deviation is obtained that is we find out ΣX^2

(4) ΣX^2 is divided by the number of items (N) in the series. Square root of $\frac{\Sigma X^2}{N}$ will be the standard deviation. That is, we calculate the value of $\sqrt{\frac{\Sigma X^2}{N}}$. Thus, following formula for the calculation of standard deviation.

$$\text{S.D. or } \sigma = \sqrt{\frac{\Sigma X^2}{N}} \text{ or } \sqrt{\frac{\Sigma\left(X - \overline{X}\right)^2}{N}}$$

(Here σ = Standard deviation, ΣX^2 – Sum total of the squares of deviations, $\overline{X}$ – Mean value, $X - \overline{X}$ – deviation from mean value, N – Number of items.

2. Deviation taken from Assumed Mean: When the actual mean is in fractions, say it is 123.674, it would be too cumbersome to take deviation from it and then obtain squares of these deviations. In such a case either the mean may be approximated or else the deviation be taken from an assumed mean.

When deviation are taken from assumed mean the following formula is applied

$$\sigma = \sqrt{\frac{\Sigma d^2}{N} - \left(\frac{\Sigma d}{N}\right)^2}$$

Steps:

(I) Take the deviations of the items from an assumed mean i.e. obtain (X–A). Denote these deviations by *d*. Take the total of these deviations, i.e. obtain Σd.

(II) Square there deviations & obtain the total Σd^2.

(III) Substitute the values of Σd^2, Σd and N in the above formula.

Example 1: Following are the marks obtained by 10 students of a class. Calculate standard deviation and coefficient of standard deviation.

Marks: 12 8 17 13 15 9 18 11 6 1

Solution:

S.No.	Marks (X)	Deviation ($X = X - \overline{X}$) $\overline{X} = 11$	$X^2 = (X - \overline{X})^2$ Square of the Deviation
1	12	12 – 11 = 1	1
2	8	8 – 11 = -3	9
3	17	17 – 11 = 6	36
4	13	13 – 11 = 2	4
5	15	15 – 11 = 4	16
6	9	9 – 11 = 2	4
7	18	18 – 11 = 7	49
8	11	11 – 11 = 0	0
9	6	6 – 11 = 5	25
10	1	1 – 11 = 10	100
N=10	$\Sigma X = 110$	$\Sigma X = 0$	$\Sigma X^2 = 244$

$$\overline{X} = \frac{\sum X}{N} = \frac{110}{10} = 11$$

$$\sigma = \sqrt{\frac{\sum X^2}{N}} = \sqrt{\frac{244}{10}} = 4.93 \text{ Marks}$$

$$\text{Coefficient of S.D.} = \frac{\sigma}{\overline{X}} = \frac{4.93}{11} = 0.45$$

Deviation taken from assumed Mean

Example 1: Find out standard deviation given the following data:

8, 10, 12, 14, 16, 18, 20, 22, 24, 26

Solution:

S.No.	Size (X)	Deviations from Assume and Average (D=X-A) A= 20	Square of the Deviation d^2
1	8	8 – 20 = –12	144
2	10	10 – 20 = –10	100
3	12	12 – 20 = –8	64
4	14	14 – 20 = –6	36
5	16	16 – 20 = –4	16
6	18	18 – 20 = –2	4
7	20 (A)	20 – 20 = 0	0
8	22	22 – 20 = +2	4
9	24	24 – 20 = +4	16
10	26	26 – 20 = +6	36
N = 10		Σd = –30	Σd^2 = 420

$$\sigma = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2} = \sqrt{\frac{420}{10} - \left(\frac{-30}{10}\right)^2}$$

$$= \sqrt{42 - (-3)^2} = \sqrt{33} = 5.74$$

Standard Deviation = 5.74

Calculation of Standard Deviation - Discrete Series

For calculating standard deviation in discrete series any of the following methods may be applied.

1.Actual mean method

2.Assumed mean method

1. Actual Mean Method: When this method is applied deviations are taken from the actual mean i.e. we find $\left(X - \overline{X}\right)$ & denote these deviation by X. These deviations are then squared and multiplied by the respective frequencies. The following formula is applied

$$\sigma = \sqrt{\frac{\sum fX_2}{N}} \quad \text{where } X = \left(X - \overline{X}\right)$$

However, in Practice this method is rarely used because if the actual mean is in fractions the calculations take a lot of time.

2. Assumed Mean Method – When this method is used the following formula is applied

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \quad \text{where } d = (X–A)$$

Step:

(I) Take the deviations of the items from an assumed mean & denote these deviations by d

(II) Multiply these deviations by the respective frequencies & obtain the total, Σfd.

(III) Obtain the squares of the deviations, i.e. calculate d^2.

(IV) Multiply the squared deviations by the respective frequencies & obtain the total Σfd^2.

Example 1: Calculate the standard deviation from the data given below:

Size of item	Frequency
3.5	3
4.5	7
5.5	22
6.5	60
7.5	85
8.5	32
9.5	8

Solution: Calculation of Standard Deviation

X (Size of item)	*f*	D=(X–A)=(X–6.5)	*fd*	*fd*^{2}
3.5	3	-3	-9	27
4.5	7	-2	-14	28
5.5	22	-1	-22	22
6.5	60	0	0	0
7.5	85	1	85	85
8.5	32	2	64	128
9.5	8	3	24	72
	N = 217		Σfd=128	Σfd^2=362

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

$\Sigma fd^2 = 362$, $\Sigma fd = 128$ $N = 217$

$$\sigma = \sqrt{\frac{362}{217} - \left(\frac{128}{217}\right)^2} = \sqrt{1.668 - 0.348} = \sqrt{1.32} = 1.148$$

Example 2: Find the Standard deviation of the following data:

Size	1	2	3	4	5	6	7	8
Frequency	5	10	15	20	15	10	10	15

Solution: Calculation of Standard Deviation (σ) in Discrete Series (Assumed mean method)

Size X	Frequency (f)	Deviation d=X-A	D^2	fd	fd^2
1	5	-4	16	-20	80
2	10	-3	9	-30	90
3	15	-2	4	-30	60
4	20	-1	1	-20	20
5	15	0	0	0	0
6	10	1	1	10	10
7	10	2	4	20	40
8	15	3	9	45	135
	$\Sigma f=100$			$\Sigma fd=25$	$\Sigma fd^2=435$

$$\text{S.D. or } \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} = \sqrt{\frac{435}{100} - \left(\frac{-25}{100}\right)^2}$$

$$\sigma = \sqrt{\frac{435}{100} - \left(\frac{-1}{4}\right)^2} = \sqrt{\frac{435}{100} - \frac{1}{16}} = 2.07$$

Thus Standard deviation, $\sigma = 2.07$

Calculation of Standard Deviation - Continuous Series

In case of continuous series standard deviation may be computed by applying any of the following two methods

1. By taking deviation of the items from actual mean
2. By taking deviation of the items from actual mean

Deviation of the item from actual mean

Example: Given the following series calculate standard deviation from actual mean

Size	0-2	2-4	4-6	6-8	8-10	10-12
f	2	4	6	4	2	6

Solution:

Size X	Mid value (m)	Frequency (f)	fm	$(X-\overline{X})^2$	$f(X-\overline{X})^2$
0 - 2	1	2	2	30.25	60.50
2 - 4	3	4	12	12.25	49.00
4 - 6	5	6	30	2.25	13.50
6 - 8	7	4	28	0.25	1.00
8 - 10	9	2	18	6.25	12.50
10 - 12	11	6	60	20.25	121.50
		Σf=24	ΣfX=156		ΣfX^2=258

$$\overline{X} = \frac{\Sigma fm}{\Sigma f}$$

$$\overline{X} = \frac{156}{24} = 6.5$$

$$\sigma = \sqrt{\frac{\sum fX^2}{N}} \quad \text{or} \quad \sigma = \sqrt{\frac{\sum f\left(X-\overline{X}\right)^2}{N}} = \sqrt{\frac{258}{24}} = 3.28$$

Deviation of the item from assumed mean

This method is the same as used in case of Discrete series. The only difference is that whereas in Discrete Series, deviations are obtained of the actual values of the series, in case of frequency distribution series deviation are obtained of the mid-values of the class intervals.

$$\text{S.D.} \quad \text{or} \quad \sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$$

Example 1: Calculate standard deviation of the following series by using assumed mean method:

Size	0-2	2-4	4-6	6-8	8-10	10-12
Frequency	2	4	6	4	2	6

Solution:

Size X	Mod value (m)	Frequency (f)	Deviation from assumed averaged = (m-A) A=5	fd	fd^2
0-2	1	2	− 4	− 8	32
2-4	3	4	− 2	− 8	16
4-6	5	6	0	0	0
6-8	7	4	2	8	16
8-10	9	2	4	8	32
10-12	11	6	6	36	216
		Σ=24		Σfd=36	Σfd^2=312

S.D. or $\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} = \sqrt{\frac{312}{24} - \left(\frac{36}{24}\right)^2} = 3.28$

S.D. or $\sigma = 3.28$

3. Step Deviation Method: When this method is used we take a common factor from the given data. The formula for computing standard deviation is:

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

Where $d' = \frac{(X - A)}{C}$ and C = common factor

The use of the above formula simplies calculation.

Example 1: Find the standard deviation for the following distribution.

X	4.5	14.5	24.5	34.5	44.5	54.5	64.5
f	1	5	12	22	17	9	4

Solution: Calculation of Standard Deviation

X	f	D= (X-34.5)/10	fd′	fd'^2
4.5	1	-3	-3	9
14.5	5	-2	-10	20
24.5	12	-1	-12	12
34.5	22	0	0	0
44.5	17	17	17	17
54.5	9	18	18	36
64.5	4	12	12	36
	N = 70		Σfd′=22	$\Sigma fd'^2$=130

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

$\Sigma fd'^2 = 130, \quad \Sigma fd' = 22 \qquad C = 10, \; N = 70$

Substituting the values

$$\sigma = \sqrt{\frac{130}{70} - \left(\frac{22}{70}\right)^2} \times 10 = \sqrt{1.857 - 0.097} \times 10$$

$$= \sqrt{1.857 - 0.097} \times 10$$

$$= \sqrt{1.76} \times 10$$

$$= 1.326 \times 10$$

$$= 13.26$$

Calculation of Standard Deviation - continuous Series

In continuous series any of the methods discussed above for discrete frequency distribution can be used. However, in practice it is the step deviation method that is mostly used. The formula is

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

$$d' = \frac{(m-A)}{C} \quad \text{where } C = \text{Common factor}$$

Steps:

(I) Find the mid points of various classes

(II) Take the deviations of these mid-points from an assumed mean and denote these deviation by d.

(III) Whenever possible take a common factor and denote this column by d′.

(IV) Multiply the frequencies of each class with these deviations.

(V) Square the deviations and multiply them with the respective frequencies of each class and obtain $\Sigma fd'^2$.

Thus the only difference in Procedure in case of continuous series is to find mid-points of the various classes.

Example: Calculate the standard deviation for the following distribution giving 300 telephone calls according to their duration in seconds.

Duration (in seconds)	No. of calls
0 - 30	9
30 - 60	17
60 - 90	43
90 - 120	82
120 - 150	81
150 - 180	44
180 - 210	24

Solution: Calculation of Standard Deviation

Duration (in seconds)	No. of calls (f)	Mid point (m)	$d' = \frac{(m-A)}{C}$	fd′	fd′²
0-30	9	15	-3	-27	81
30-60	17	45	-2	-34	68
60-90	43	75	-1	-43	43
90-120	82	105	0	0	0
120-150	81	135	1	81	81
150-180	44	165	2	88	176
180-210	24	195	3	72	216
				$\Sigma fd'=137$	$\Sigma fd'^2=665$

$$\sigma = \sqrt{\frac{\Sigma fd'^2}{N} - \left(\frac{\Sigma fd'}{N}\right)^2} \times C$$

$\Sigma fd'^2 = 665$, $\Sigma fd' = 137$ $N = 300$, $C = 30$

$$\sigma = \sqrt{\frac{665}{300} - \left(\frac{137}{300}\right)^2} \times 30 = \sqrt{2.217 - 0.209} \times 30$$

$$= \sqrt{2.008} \times 30 = 1.417 \times 30 = 42.51$$

Merits and Demerits of Standard Deviation

Merits

1. **Based on all values:** The calculation of standard deviation is based on all the values of a series. It does not ignore any value. Accordingly it is a comprehensive measure of dispersion.
2. **Certain Measure:** Standard deviation is a clear & certain measure of dispersion. Therefore, it can be used in all situations.

3. **Little effect of a change in sample**: Change in sample cause little effect on standard deviation. This is because deviation is based on all the values of a sample.
4. **Algebraic Treatment:** Standard Deviation is capable of further algebraic treatment.

Demerits:

1. **Difficult:** It is difficult to calculate and make use of standard deviation as a measure of dispersion.
2. **More Importance of Extreme Value:** In the calculation of standard deviation, extreme values tend to get greater importance.

Coefficient of Variation (C.V)

All the measures of dispersion discussed so far have units. If two series differ in their units of measurement, their variability can not be compared by any measure given so far. Also, the size of measures of dispersion depends upon the size of values. Hence in situations where either the two series have different units of measurements or their means differ sufficiently in size, the coefficient of variation should be used as measure of dispersion. It is unitless measure of dispersion & also takes into account the size of the means of the two series. It is the best measure to compare the variability of two series or sets of observations. A series with less coefficient of variation is considered more consistent or stable. It was first used by the famous statistician Karl Pearson. That is why it is called Karl Pearson's coefficient of variation. In the words of Karl Pearson, Coefficient of variation of a series of variate values is the ratio of the standard deviation to the mean multiplied by 100

If σ is the standard deviation & $\overline{X}$ is the mean of the set of values, the coefficient of variation is

$$\text{C.V.} = \frac{\sigma}{\overline{X}} \times 100 = \text{Coefficient of Standard Deviation}$$

Coefficient of variation is 100 times the coefficient of dispersion based on standard deviation of a statistical series.

C.V. – Coefficient of variation

$\overline{X}$ – Mean

σ – Standard deviation

Properties:

1. It is one of the most widely used measures of dispersion because of its virtues.
2. Smaller the value of C.V., more consistent are the data & vice versa. Hence, a series with smaller C.V. is more consistent, i.e., it possesses less variability.
3. For field experiments, C.V. is generally reported. If C.V. is low, it indicates more reliability of experimental findings.
4. It is a relative measure of variability.

Calculation of Coefficient of variation

Individual Series and Coefficient of variation

Example 1: Calculate coefficient of variation of the following series:

S.No.	1	2	3	4	5	6	7	8	9	10
Marks	53	58	25	30	54	42	32	48	46	52

Solution: Calculation of CV: Individual Series:

S.No.	Marks X	Deviation X = $(X-\overline{X})$, $\overline{X}$ =44	Square of the deviation $(X-\overline{X})$ or X^2
1	53	–9	81
2	58	14	196
3	25	–19	361
4	30	–14	196
5	54	10	100
6	42	–2	4
7	32	–12	144
8	48	4	16
9	46	2	4
10	52	8	62
N=10	ΣX=440		$\Sigma X^2 = \Sigma(X-\overline{X})^2$ =1,166

$$\overline{X} = \frac{\sum X}{N} = \frac{440}{10} = 44 \text{Marks}$$

$$\sigma = \sqrt{\frac{\sum X^2}{N}} = \sqrt{\frac{\sum(X - \overline{X})^2}{N}} = \sqrt{\frac{1,166}{10}} = \sqrt{116.6} = 10.8 \text{ Marks}$$

$$\text{C.V.} = \frac{\sigma}{\overline{X}} \times 100 = \frac{10.8}{48} \times 100 = 24.54$$

Thus C.V. = 24.54

Discrete Series and Coefficient of variation

Example 1: Calculate coefficient of variation of the following data

Item	10	12	14	16	18	20	22
Frequency	4	6	10	15	9	4	2

Solution: Calculation of C.V.

Items X	Frequency f	Deviation d d=X–A, A=16	fd	fd^2
10	4	–6	–24	144
12	6	–4	–24	96
14	10	–2	–20	40
16	15	0	0	0
18	9	2	18	36
20	4	4	16	64
22	2	6	12	72
	N = 50		$\Sigma fd = -22$	$\Sigma fd^2=452$

$$\overline{X} = A + \frac{\Sigma fd}{N} = 16 + \left(\frac{-22}{50}\right) = 15.56$$

$$\text{S.D. or } \sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} = \sqrt{\frac{452}{50} - \left(\frac{-22}{50}\right)^2} = \sqrt{9.04 - 0.19}$$

$$\sigma = \sqrt{8.85} = 2.97$$

$$\text{Coefficient of variation} = \frac{\sigma}{\overline{\overline{X}}} \times 100 = \frac{2.97}{15.56} \times 100 = 19$$

Thus coefficient of variation = 19

Continuous Series and Coefficient of variation

Example 1: From the data given below state which series is more consistent:

Variable	Series A	Series B
10 - 20	10	18
20 - 30	18	22
30 - 40	32	40
40 - 50	40	32
50 - 60	22	18
60 - 70	18	10

Solution: In order to find out which series is more consistent we shall compare the coefficient of variation

$$\text{C.V.} = \frac{\sigma}{\overline{\overline{X}}} \times 100$$

Calculation of coefficient of variation (Series A)

Variable	f	Mixed value (m)	d = X–A A=35	fd	fd^2
10 - 20	10	15	–20	–200	4000
20 - 30	18	25	–10	–180	1800
30 - 40	32	35	0	0	0
40 - 50	40	45	10	400	4000
50 - 60	22	55	20	440	8800
60 - 70	18	65	30	540	16200
	N=140			$\Sigma fd=1000$	$\Sigma fd^2=34800$

Calculation of Mean:

$$\overline{X} = A + \frac{\Sigma fd}{N} = 35 + \frac{1000}{140}$$

$$= 35 + 7.14 = 42.14$$

Calculation of Standard Deviation

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

$$= \sqrt{\frac{34800}{140} - \left(\frac{1000}{140}\right)^2} = \sqrt{248.571 - (7.14)^2}$$

$$= \sqrt{248.571 - 50.94} = \sqrt{197.5} = 14.05$$

$$\text{C.V.} = \frac{\sigma}{\overline{X}} \times 100 = \frac{14.05}{42.14} \times 100 = 33.34 \text{ percent}$$

Calculation of coefficient of variation (Series B)

Variable	f	Mixed value (m)	d=X–A A=35	fd	fd^2
10 - 20	18	15	–20	–360	7200
20 - 30	22	25	–10	–220	2200
30 - 40	40	35	0	0	0
40 - 50	32	45	10	320	3200
50 - 60	18	55	20	360	7200
60 - 70	10	65	30	300	9000
	N=140			Σfd=400	Σfd^2=28800

Calculation of Mean

$$\overline{X} = A + \frac{\Sigma fd}{N} = 35 + \frac{400}{140}$$

$$= 35 + 2.85 = 3785$$

Calculation of Standard Deviation

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} = \sqrt{\frac{28800}{400} - \left(\frac{400}{140}\right)^2}$$

$$= \sqrt{72 - (2.857)^2} = \sqrt{72 - 8.1628}$$

$$= \sqrt{63.8372} = 7.98$$

$$\text{C.V.} = \frac{\sigma}{\overline{X}} \times 100 = \frac{7.98}{37.85} \times 100 = 21.08 \text{ percent}$$

C.V. Series (A) = 33.34 percent

C.V. Series (B) = 21.08 percent

Since coefficient of variation is less for Series B

Hence series B is more consistent or stable.

Example 1: Two factories A and B are located in some industrial estate. Average wage and its standard deviation are given below separately for A and B Find out coefficient of variation

Factory	Average weekly wage	S.D.	No. of workers
A	35	5	476
B	30	10	524

Solution: Coefficient of variation of factory A:

$$CV_A = \frac{\sigma}{X} \times 100 = \frac{5}{35} \times 100 = 14.28$$

Coefficient of variation of factory B:

$$CV_B = \frac{\sigma}{X} \times 100 = \frac{10}{30} \times 100 = 33.33$$

Variance

The term variance was used to describe the square of the Standard deviation by R.A. Fisher in 1913. The concept of variance is highly important in advanced work where it is possible to split the total into several parts, each attributable to one of the factors causing variation in their original series. Variance is defined as follows:

$$\text{Variance} = \frac{\Sigma\left(X - \overline{X}\right)^2}{N}$$

Thus variance is nothing but the square of the Standard deviation

$$\text{Variance} = \sigma^2$$

$$\sigma = \sqrt{\text{Variance}}$$

Illustration 1: The following table gives the marks obtained by a group of 80 students in an examination. Calculate the variance.

Marks obtained	No. of students
10 - 14	2
14 - 18	4
18 - 22	4
22 - 26	8
26 - 30	12
30 - 34	16
34 - 38	10
38 - 42	8
42 - 46	4
46 - 50	6
50 - 54	2
54 - 58	4

Calculation of variance

Marks obtained	Mid-value (m)	f	d=X–A	fd	d^2	fd^2
10 - 14	12	2	-20	-40	400	800
14 - 18	16	4	-16	-64	256	1024
18 - 22	20	4	-12	-48	144	257
22 - 26	24	8	-8	-64	64	512
26 - 30	28	12	-4	-48	16	192
30 - 34	32	16	0	0	0	0
34 - 38	36	10	4	40	16	160
38 - 42	40	8	8	64	64	512
42 - 46	44	4	12	48	144	576
46 - 50	48	6	16	96	256	1536
50 - 54	52	2	20	40	400	800
54 - 58	56	4	24	96	576	2304
		N=80		Σfd=120		Σfd^2=8992

$$\sigma^2 = \frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2 = \frac{8992}{80} - \left(\frac{120}{80}\right)^2$$

$$= 112.4 - (1.5)^2 = 112.4 - 2.25$$

$$= 110.15$$

10.4. LET US SUM UP

In this lesson, we discussed the concept of dispersion and its different measures.

10.5. EXAMINATION ORIENTED QUESTIONS

Short Answer Type Questions

Q.1. Define Dispersion. Why is the study of dispersion required?

Q.2. Define Range. How it is measured?

Q.3. Define Quartile Deviation. How is it measured?

Q.4. Define Quartile Deviation. Give formula of Quartile Deviation.

Q.5. Define Mean Deviation. Give merits of mean Deviation.

Q.6. What is coefficient of Quartile Deviation?

Q.7. Define Standard Deviation. Give merits and demerits of Standard Deviation.

Q.8. Give formula of Standard Deviation for discrete series also state merits of S.D.

Q.9. Give formula of Mean Deviation for individual series also discuss merits of mean deviation.

Q.10. Define coefficient of Range. How is it measured?

Q.11. Define coefficient of variation. How is it measured?

Q.12. Illustrate the meaning of the term dispersion with examples.

Q.13. Explain Quartile deviation with help of formula.

Q.14. What are the merits of standard deviation?

Long Answer Type Questions

Q.1. What is meant by Mean Deviation? What are the methods to calculate it?

Q.2. What is meant by coefficient of variation? How will you calculate it in case of a discrete series?

Q.3. What is standard deviation? What are its advantages and disadvantages?

Some More Questions

Q.1. Calculate range and coefficient of Range of the following data:

4, 7, 8, 46, 53, 77, 8, 1, 5, 13

(**Ans.** R = 76, Coefficient of Range = 0.97)

Q.2. Given the following data - set, calculate Range and the coefficient of Range.

Size	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5
Frequency	4	5	6	3	2	1	3	5

(**Ans.** R = 7, CR = 0.44)

Q.3. Find the Quartile Deviation and the coefficient of Quartile Deviation of the following series wages of 9 workers in Rupees:

170, 82, 110, 100, 150, 150, 200, 116, 250

(**Ans.** Q.D = 40, coefficient of Q.D.= 0.276)

Q.4. Given the following data, estimate the coefficient of Q.D.

15, 20, 23, 23, 25, 25, 27, 40

(**Ans.** Coefficient of Q.D. = 0.12)

Q.5. Find the Mean Deviation and Coefficient of Mean Deviation

Class interval	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	15	19	14	20	18	14

(**Ans.** MD = 14.46, coefficient of MD = 0.466)

Q.6. Calculate standard deviation, given the following data:

10, 12, 14, 16, 18, 22, 24, 26, 28 (**Ans.** σ = 6.05)

Q.7. Calculate standard deviation and coefficient of standard deviation, given the following data:

Income (Rs.)	5	10	15	20	25	30	35	40
No. of workers	26	29	40	35	26	18	14	12

(**Ans.** S.D = 10.02, coefficient of SD = 0.516)

Q.8. Find out the standard deviation of the marks secured by 10 students

S.No.	1	2	3	4	5	6	7	8	9	10
Marks	43	48	65	57	31	60	37	48	78	59

(**Ans.** SD = 13.26)

Q.10. Calculate range, standard deviation and coefficient of variation of marks secured by students.

50, 55, 57, 49, 54, 61, 64, 59, 58, 56

(**Ans.** Range = 15, S.D = 4.38 and Coefficient of V = 7.78)

Q.11. Calculate coefficient of variation from the following data:

Variables	10	20	30	40	50	60	70
Frequencies	6	8	16	15	32	11	12

(**Ans.** CV= 37.34)

Q.12. Estimate coefficient of variation of the following data:

Weight /kg	0-20	20-40	40-60	60-80	80-100
No. of Persons	81	40	66	49	14

(**Ans.** C.V = 63.75)

Q.13. What are the characteristics of a good measure of dispresion? Write various methods of dispersion.

Q.14. Find the Standard deviation of the following distribution:

Age	20-25	25-30	30-35	35-40	40-45	45-50
No. of Persons	170	110	80	45	40	35

(**Ans.** σ = 7.936)

Q.15. Find out the range and its coefficient of the following items:

Size: 145, 367, 268, 73, 185, 619, 280, 115, 870, 316

(**Sol.** Range = 797; Coefficient of Range = 0.85)

Q.16. Calculate Range and its coefficient in Discrete series:

Size of the items	10	20	30	40	50	60	70
Frequency	5	4	2	1	2	3	5

Sol. $R = L - S$

$R = 70 - 10 = 60$

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{70-10}{70+10} = \frac{60}{80}$$

Coefficient of Range = 0.75

Q.17. Find out the range and its coefficient from the following series:

Size	5-10	10-15	15-20	20-25	25-30	30-35	35-40
f	3	8	7	12	9	8	6

Sol. Here, L = 40 and S = 5

Range = L – S = 40 – 5 = 35

$$\text{Coefficient of Range} = \frac{40-5}{40+5} = \frac{35}{45} = 0.78$$

10.6. SUGGESTED READINGS

1. S.P. Gupta (2005), Statistical Methods, S. Chand and Sons, New Delhi
2. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas publishing House, New Delhi.
3. Monga, G.S. (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.

SKEWNESS

LESSON NO. 11 **UNIT-III**

11.1. OBJECTIVES

After going through this lesson, you would be able to understand.

1. Meaning of Skewness
2. Difference between Dispersion and Skewness
3. Measures of Skewness
4. Various relative measures of Skewness such as Karl Pearson's Coefficient of Skewness, Bowley's coefficient of Skewness, Kelly's Coefficient of Skewness, Measures of Skewness based on moment etc.

11.2. INTRODUCTION

Meaning of central tendencies reveal the concentration of frequencies towards the central value of the series and methods of dispersion reveal the dispersal of values in relation to the central value. But the nature of dispersal of values on either sides of an average is not known by measuring dispersion. The measures of

‘Skewness’ tell about the pattern of dispersal of items from an average, whether it is symmetrical or not. The nature of distribution is further studied deeply by calculating 'Moments' which reveals whether the symmetrical curve is normal, more flat than normal or more peaked than normal. Similarly, kurtosis is yet another measure which tells us about the form of a distribution. Thus, it can be said that the central tendencies and dispersion measure should be supplemented by measures of Skewness and Kurtosis so that a more elaborate picture about the distribution given can be obtained. The study becomes more important in subjects of Economics, Sociology and other Social Sciences where usually normal distribution in a series does not occur. However, these hold importance in Biological Sciences and other physical sciences as well. In this lesson, we shall discuss these concepts in detail.

11.3. MEANING OF SKEWNESS

The word 'Skewness' is the opposite of symmetry and its presence tells us that a particular distribution is not symmetrical or in other words it is skewed. The word ‘Skewness’ can be understood by the following definitions given by eminent statisticians, economists and mathematicians.

1. “When series is not symmetrical it is said to be asymmetrical or skewed” – Croxten and Cowden.
2. “Measures of skewness tell us the direction and the extent of skewness. In symmetrical distribution the mean, median and mode are identical. The more the mean moves away from the mode, the larger the asymmetry of skewness.” – Simpson and Pafka

The concept of Skewness will be clear from the following three diagrams showing a symmetrical distribution, a positively skewed distribution and a negative skewed distribution.

(a) Symmetrical Distribution:

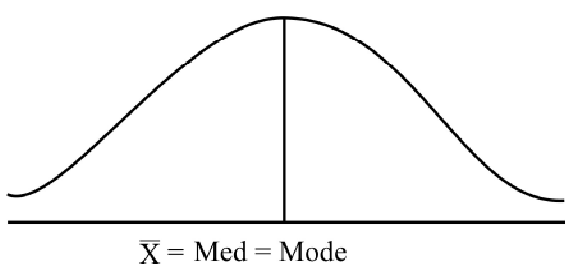


Fig.

It is clear from the above diagram that in a symmetrical distribution, the value of mean, median and mode coincide. The spread of the frequencies is the same on both sides of the centre point of the curve.

A distribution which is not symmetrical or bell shaped is called a skewed distribution and such a distribution could either be positively skewed or negatively skewed as would be clear from the following diagram:

(b) Positively skewed Distribution:

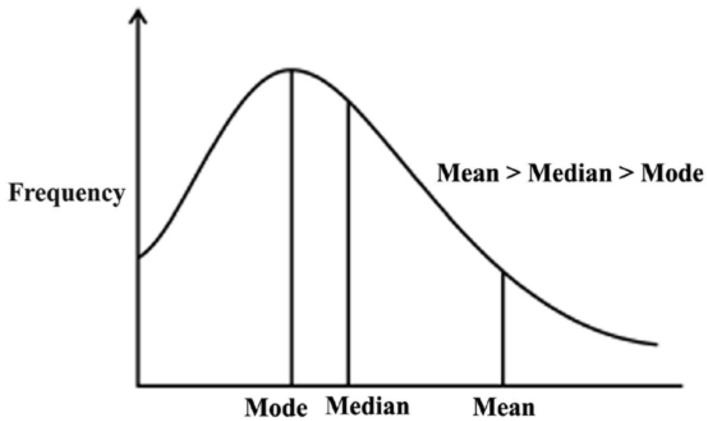


Fig.

(c) Negatively skewed Distribution:

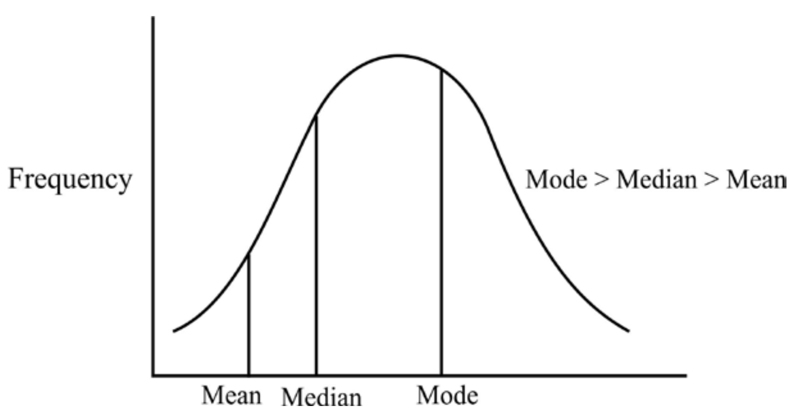


Fig.

In a positively skewed distribution, the curve has a longer tail to the right. Here the value of mean is maximum and that of mode least - the median lies in between the two. In a negatively skewed distribution, the longer tail lies towards the left. Here, the value of mode is maximum and that of mean the least. The median lies in between the two. In the positively skewed distribution the frequencies are spread out over a greater range of values on the high- value end of the curve (the right-hand side) than they are on the low- value end. In the negatively skewed distribution the position is reversed i.e. the excess tail is one on the left hand side. A positively skewed distribution will have greater variation towards the higher values of the variable and a negatively skewed distribution has greater variation towards the lower values of the variable.

11.3.1. Difference between Dispersion and Skewness

Dispersion is concerned with the amount of variation rather than with its direction. Skewness tells us about the direction of the variation or the departure from symmetry. In fact measures of skewness are dependent upon the amount of dispersion.

11.3.2. Measures of Skewness

Measures of skwness tell us the direction and extent of asymmetry in a series, and permit us to compare two or more series with regard to these. They are either absolute or relative.

Absolute measures of Skewness: They are based on the assumption that in a skewed distribution the values of mean, median and mode do not coincide and so the difference between any two of these values indicate the extent of skwness. As per these methods, skewness can be measured by anyone of these:

(a) Mean - Mode (b) Mean - Median (c) Median - Mode

Drawbacks: This method suffers from the following drawbacks:

1. It would be expressed in the unit of value of the distribution and could, therefore, not be compared with another comparable series expressed in different units.
2. Similar looking curves may show large differences in the results.

Relative measures of distribution: Relative measures of skewness facilitate comparisons. They are obtained by dividing the absolute measures by anyone of the measures of dispersion. They are also called coefficient of skewness. Various relative measures of skewness are -

(a) Karl Pearson's Coefficient of Skewness: Karl Pearson's coefficient, also popularly known as Pearsonian coefficient of skewness, is based upon the difference between mean and mode. This difference is divided by standard deviation to give a relative measure. The formula thus becomes.

$$S_{Kp} = \frac{\text{Mean} - \text{Mode}}{\text{Standard Deviation}}$$

S_{Kp} — Karl Pearson's coefficient of skewness

But, as we know, in case of an asymmetrical (a skewed) distribution,

Mode $\rightarrow$ $M_0 = 3Md - 2M$

$$\therefore \qquad S_{Kp} = \frac{M - (3Md - 2M)}{\sigma} = \frac{3(M - Md)}{\sigma}$$

There is no limit to this measure in theory and this is slight drawback. But in practice the value given by this formula is rarely very high and usually lies between ± 3.

When a distribution is symmetrical, the values of mean, median and mode coincide and, therefore, the coefficient of skewness will be zero. When a distribution is positively skewed, the coefficient of skewness shall have plus sign (>0) and when it is negatively skewed, the coefficient of skewness shall have minus sign (0<). The degree of skewness shall be obtained by the numerical value say 0.8 or 0.2 etc. Thus this formula gives both the direction as well as the extent of skewness.

(b) Bowley's Coefficient of Skewness: This measure of skewness is also known as quartile coefficient of skewness. It is based upon the values of quartiles and median. When $Q_3 - M = M - Q_1$ the distribution is said to symmetrical.

Bowley's absolute measure of skewness = $(Q_3 - M) - (M - Q_1)$

Where M=Median & Q_1 & Q_3 first and third quarter respectively.

Bowley's relative measure of skewness $(S_B) = \dfrac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)}$

According to this formula, Sk=0, If $Q_3 - Md = Md - Q_1$

i.e. If Median is equidistant from the upper and lower quartiles.

It is positive if

$Q_3 - Md > \text{Median} - Q_1$

or $Q_3 + Q_1 > 2Md$

and negative If $Q_3 - Md < Md - Q_1$

or $Q_3 + Q_1 < 2Md$

Further, $Sk = +1$, if $Md - Q_1 = 0$

i.e. $Md = Q_1$

and $= -1$, If $Q_3 - Md = 0$

or $Q_3 = Md$

So, the theoretical limit of the formulae are ± 1. If the distribution is symmetrical, the value comes out to be zero.

Drawback: Since quartiles are not based on all the observations of a series, the coefficient by this method may come out to be zero even when the data is not symmetrical. Therefore, for the purposes of comparison, Karl Pearson's method is more useful.

Use: However, when in a distribution, extreme values are prominent or where a distribution is open-ended and positional measures are called for, Bowley's coefficient gives better results.

(3) Kelly's coefficient of Skewness: Bowley's measure discussed above neglects the two extreme quarters of the data. It would be better for a measure to cover the entire data especially because in measuring skewness, we are often interested in the more extreme items. Bowley's measure can be extended by taking any two deciles equidistant from median or any two percentiles equidistant from the median. Kelly has suggested the following formula for measuring Skewness based upon the 10th and the 90th percentiles (or the first and ninth deciles).

Absolute measures

$$S_K = (P_{90} - P_{50}) - (P_{50} - P_{10})$$

$$= P_{90} + P_{10} - 2P_{50}$$

But $P_{90} = D_9$

& $P_{10} = D_1$

$$S_K = (D_9 - D_5) - (D_5 - D_1)$$

$$= D_9 + D_1 - 2D_5$$

It should be noted that

$$D_5 = P_{50} = \text{Median}$$

$\therefore$ $S_K = P_{90} + P_{10} - 2Md$

Or $D_9 + D_1 - 2Md$

$$S_{Kk} = \frac{P_{90} + P_{10} - 2Md}{P_{90} - P_{10}}$$

Also $$Sk = \frac{D_9 + D_1 - 2Med}{D_9 - D_1}$$

Where S_{Kk} = Kelly's coefficient of Skewness

This measure of Skewness has some theoretical attraction if skewness is to be based on Percentiles. However, this method is not popular in Practice and generally Karl Pearson's method is used.

4. Measures of Skewness based on Moments. This method is explained in a later section.

Check Your Progress–I

1. What is skewness? Draw a figure to show a positively skewed and a negatively skewed distribution. What are the positions of Mean, Median and Mode in these?

____

____

____

____

____

2. What are the different measures of skewness? Discuss their relative importance.

____

____

____

____

____

Conclusion: From the above, it may be concluded that Skewness measures the lack of symmetry in the distribution. The skewness may be positive or negative. Karl Pearson's method of measuring skewness is the most accepted one which is based upon, mean, mode and standard deviation as it takes into account all the items of the series.

Example 1: If Mean = 120, Mode = 123 and Karl Pearson's coefficient of skewness is 0.3, find the coefficient of variation.

Solution: Karl Pearson's coefficient of Skewness $= \frac{\overline{X} - M_O}{S.D} = 0.3$

$$S.D. = \frac{\overline{X} - M_O}{0.3} = \frac{-3}{0.3} = -10$$

Or We take it to be 10 $(\because$ S.D is always positive$)$

$$\text{C.V.} = \frac{\text{S.D.}}{\text{X}} \times 100 = \frac{10}{120} \times 100 = 8.33\%$$

Example 2: In a negatively skewed data, mean, median and mode are 25, 28 and 22 respectively. Can you say whether the given data is correct or not?

Solution: In a negatively skewed distribution

Mode > Median > Mean

$$22 \not> 28 > 22$$

$\therefore$ The information is correct.

Example 3: Calculate coefficient of variation if Pearson's measure of skewness is 0.4, arithmetic mean is 56 and Median = 50

Sol. $$S_K = \frac{3\left(\overline{X} - M\right)}{\sigma}$$

$$\overline{X} = 56,\ M = 50,\ S_K = 0.4$$

$$0.4 = \frac{3\left(56 - 50\right)}{\sigma}$$

$$\sigma = \frac{18}{0.4},\ \sigma = 45$$

Coefficient of valation C.V $= \dfrac{\sigma}{\overline{X}} \times 100$

$$= \frac{45}{56} \times 100 = 90\%$$

Example 4: If mean = 40, Standard deviation = 10, Karl Pearson's coefficient of Skewness = 0.5, Find median and mode

Solution: Karl Pearson's coefficient of Skewness

$$S_k = \frac{\overline{X} - M_O}{\sigma}$$

$$\overline{X} = 40,\ M_O = ?\ \ \sigma = 10,\ S_K = 0.5$$

$$0.5 = \frac{40 - M_O}{10}$$

$$0.5 \times 10 = 40 - M_O$$

$$M_O = 35$$

Now $\quad M_O = 3\ M - 2\ \overline{X}$

$$35 = 3XM - 2 \times 40$$

$$35 = 3\ M - 80$$

$$35 + 80 = 3M$$

$$\frac{115}{3} = M$$

$$38.33 = M$$

Median = 38.33 and Mode = 35

Example 5: Find the coefficient of skewness from the following data:

Value	6	12	18	24	30	36	42
Frequency	4	7	9	18	15	10	5

Solution: $\quad S_K = \dfrac{\text{Mean} - \text{Mod}}{\sigma}$

Calculation of coefficient of Skewness

X Value	f	d′=(X-24)6	fd′	fd′²
6	4	-3	-12	36
12	7	-2	-14	28
18	9	-1	-9	9
24	18	0	0	0
30	15	1	15	15
36	10	2	20	40
42	5	3	15	45
	N =68		$\Sigma fd'=15$	$\Sigma fd'^2=173$

Calculation of Mean: $\overline{X} + \frac{\sum fd`}{N} \times C$

$A = 24, \ \Sigma fd' = 15, \ N = 68, \ C = 6$

$$\overline{X} = 24 + \frac{15}{68} \times 6 = 24 + 1.32 = 25.32$$

Calculation of Mode: By inspection modal value is 24.

Calculation of Standard Deviation

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

$\Sigma fd'^2 = 173, \ \Sigma fd' = 15, \ N = 68, \ C = 6$

$$\sigma = \sqrt{\frac{173}{68} - \left(\frac{15}{68}\right)^2} \times 6 = \sqrt{2.544 - 0.046 \times 6}$$

$$= 1.58 \times 6 = 9.48$$

$\overline{X} = 25.32, \ M_O = 24, \ \sigma = 9.48$

$$S_K = \frac{25.32 - 24}{9.48} = \frac{1.32}{9.48} = 0.139$$

Example 6: Calculate Karl Pearson's coefficient of skewness from the following data:

Class	Frequency
70 - 80	18
60 - 70	22
50 - 60	30
40 - 50	35
30 - 40	21
20 - 30	11
10 - 20	6
0 - 10	5

Solution: $S_K = \dfrac{\text{Mean} - \text{Mode}}{\sigma}$

Calculation of coefficient of Skewness

Class	f	m.P. (m)	D=(m-45)/10	fd′	fd′²
0 - 10	5	5	-4	-20	80
10 - 20	6	15	-3	-18	54
20 - 30	11	25	-2	-22	44
30 - 40	21	35	-1	-21	21
40 - 50	35	45	0	0	0
50 - 60	30	55	1	30	30
60 - 70	22	65	2	44	88
70 - 80	18	75	3	54	162
	N=148			$\Sigma fd'=47$	$\Sigma fd'^2=479$

Calculation of Mean:

$$\overline{X} = A + \frac{\Sigma fd`}{N} \times C$$

$A = 45, \ \Sigma fd' = 47, \quad N = 148, \quad C = 10$

$$\overline{X} = 45 + \frac{47}{148} \times 10 = 45 + 3.176 = 48.176$$

Calculation of Mode:

By inspection mode lies in the class 40 – 50

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

$L = 40, \ \Delta_1 = 35 - 21 = 14, \quad \Delta_2 = 35–30 = 5, \ i = 10$

$$\text{Mode} = 40 + \frac{14}{14+5} \times 10 = 40 + 7.368 = 47.368$$

Calculation of Standard Deviation:

$$\sigma = \sqrt{\frac{\Sigma fd'}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times \text{C}$$

$$= \sqrt{\frac{479}{148} - \left(\frac{47}{148}\right)^2} \times 10$$

$$= \sqrt{3.236 - 0.101} \times 10$$

$$= \sqrt{3.135} \times 10 = 1.771 \times 10 = 17.71$$

$$\overline{\text{X}} = 48.176, \text{Mode} = 47.368 \ \ \sigma = 17.71$$

$$\text{S}_{\text{Kp}} = \frac{48.176 - 47.368}{17.71} = \frac{0.808}{17.71} = 0.046$$

Example 7: Calculate Bowley's coefficient of skewness of the following data :

Weight (lbs)	*No. of persons*	*Weight (lbs)*	*No. of persons*
Under 100	1	150-159	65
110-109	14	160-169	31
110-119	66	170-179	12
120-129	122	180-189	5
130-139	145	190-199	2
140-149	121	200 and over	2

Solution. Here we are given the frequency distribution with inclusive type classes. Since the formulae for median and quartiles are based on continuous frequency distribution with exclusive type classes without any gaps, we obtain the class boundaries which are given in the last column of the following table:

COMPUTATION OF QUARTILES

Weight (lbs)	*No. of persons (f)*	*'Less than' c.f.*	*Class boudaries*
0-99	1	1	0-99.5
100-109	14	15	99.5-109.5
110-119	66	81	109.5-119.5
120-129	122	203	119.5-129.5
130-139	145	348	129.5-139.5
140-149	121	469	139.5-149.5
150-159	65	534	149.5-159.5
160-169	31	565	159.5-169.5
170-179	12	577	169.5-179.5
180-189	5	582	179.5-189.5
190-199	2	584	189.5-199.5
200-209	2	N=586	199.5-209.5

Here $N=586, \quad \frac{N}{4}=146.5, \quad \frac{N}{2}=293, \quad \frac{3N}{4}=439.5$

The *c.f.* just greater than N/2 i.e., 293 is 348. Hence the corresponding class 129.5-139.5 is the median class. Using the Median Formula, we get :

$$\therefore \quad Md(Q_2) = 129.5 + \frac{10}{145}(293-203) = 129.5 + \frac{10 \times 90}{145}$$

$$= 129.5 + 6.2069 = 135.7069 \simeq 135.71$$

The *c.f.* just greater than N/4 *i.e.* 146.5 is 203. Hence the corresponding class 119.5–129.5 contains Q_1.

$$\therefore \quad Q_1 = 119.5 + \frac{10}{122}(146.5-81) = 119.5 + \frac{10 \times 65.5}{122}$$

$$= 119.5 + 5.3688 = 124.8688 \simeq 124.87$$

The *c.f.* just great than 3N/4 i.e. 439.5 is 469. Hence the corresponding class 139.5–149.5 contains Q_3.

$$\therefore \qquad Q_1 = 139.5 + \frac{10}{121}(439.5–348) = 139.5 + \frac{10 \times 91.5}{121}$$

$$= 139.5 + 7.5620 = 147.0620 \simeq 147.06$$

Bowley's co-efficient of skewness is given by :

$$\text{Sk (Bowley)}= \frac{Q_3 + Q_1 - 2Md}{Q_3 - Q_1} = \frac{147.06 + 124.87 - 2 \times 135.71}{147.06 - 124.87}$$

$$= \frac{271.93 - 271.42}{22.19} = \frac{0.51}{22.19} = 0.0298$$

Example 8: Calculate the coefficient of skewness from the following data by using quartiles.

Marks	*No. of persons*	*Marks*	*No. of persons*
Above 0	180	Above 60	65
Above 15	160	Above 75	20
Above 30	130	Above 90	5
Above 45	100		

Solution. We are given 'more than' cumulative frequency distribution. To compute quartiles, we first express it as a continuous frequency distribution without any gaps as given in the following table :

COMPUTATION OF QUARTILES

Marks	*No. of Students (f)*	*'Less than' c.f.*
0-15	180-160=20	20
15-30	160-130=30	50
30-45	130-100=30	80
45-60	100-65=35	115
60-75	65-20=45	160
75-90	20-5=15	175
above 90	5	N=180
Total	N=Σf=180	

$$\frac{N}{2} = \frac{180}{2} = 90, \quad \frac{N}{4} = \frac{180}{4} = 45 \quad \text{and} \quad \frac{3N}{4} = 135$$

The *c.f.* just greater than N/2 i.e., 90 is 115. Hence the corresponding class 45-60 is the median class.

$$\therefore \qquad \text{Md} = l + \frac{h}{f}\left(\frac{N}{2} - C\right) = 45 + \frac{15}{35}(90 - 80)$$

$$= 45 + \frac{15 \times 10}{35} = 45 + 4.29 = 49.29$$

The *c.f.* just greater than N/4 *i.e.*, 45 is 50. Hence the corresponding class 15-30 contains Q_1.

$$\therefore \qquad Q_1 = l + \frac{h}{f}\left(\frac{N}{4} - C\right) = 15 + \frac{15}{30}(45 - 20)$$

$$= 15 + \frac{15 \times 25}{30} = 15 + 12.5 = 27.5$$

The *c.f.* just greater than 3N/4 *i.e.*, 135 is 160. Hence the corresponding class 60-75 contains Q_3.

$$\therefore \qquad Q_3 = l + \frac{h}{f}\left(\frac{3N}{4} - C\right) = 60 + \frac{15}{45}(135 - 115)$$

$$= 60 + \frac{20}{3} = 60 + 6.67 = 66.67$$

Hence Bowley's coefficient of skewness based on quartiles is given by :

$$\text{Sk (Bowley)} = \frac{Q_3 + Q_1 - 2Md}{Q_3 - Q_1} = \frac{66.67 + 27.50 - 2 \times 49.29}{66.67 - 27.50}$$

$$= \frac{94.17 - 98.58}{39.17} = -\frac{4.41}{39.17} = -0.1126$$

Example 9. In a frequency distribution the coefficient of skewness based on quartiles is 0.6. If the sum of the upper and lower quartiles is 100 and median is 38, find the value of upper quartiles.

Solution. We are given :

$$\text{Sk} = \frac{Q_3 + Q_1 - 2Md}{Q_3 - Q_1} = 0.6 \qquad \text{....(i)}$$

Also $Q_3 + Q_1 = 100$ and Median = 38 ....(ii)

Substituting in (i) we get

$$\frac{100 - 2 \times 38}{Q_3 - Q_1} = 0.6$$

$$\Rightarrow \quad Q_3 - Q_1 = \frac{100 - 76}{0.6} = \frac{24}{0.6} = 40$$

Thus we have :

$$Q_3 + Q_1 = 100$$

and $Q_3 - Q_1 = 40$

Adding, we get :

$$2Q_3 = 140$$

$$Q_3 = \frac{140}{2} = 70$$

Hence the value of the upper quartile is 70.

Note : Examples 7-9 have been taken from the Book 'Business Statistics' by S C Gupta and Indra Gupta.

11.4 LET US SUM UP

In this lesson, we discussed in detail about the different measures of skewness.

11.5. EXAMINATION ORIENTED QUESTIONS

Q.1. Find the coefficient of skewness for the following distribution:

Variable	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	2	5	7	13	21	16	8	3

Ans. – 0.147

Q.2. Calculate Karl Pearson's coefficient of skewness from the following data:

Mark above	0	10	20	30	40	50	60	70	80
No. of students	150	140	100	80	80	70	30	14	0

Hint: Convert the data in to continuous series calculate frequencies (150 - 140), (140 - 100), (100 - 80), (80 - 70) .... and so on use formula

$$S = \frac{3\left(\overline{X} - M\right)}{\sigma}$$

Ans. S = – 0.75

11.6. SUGGESTED READINGS & REFERENCES

1. S.P. Gupta (2005), Statistical Methods, S. Chand and Sons, New Delhi
2. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas publishing House, New Delhi.
3. Monga, G.S. (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.
4. Gupta, S.C. and Indra Gupta, Business statistics.

MOMENTS & KURTOSIS

LESSON NO. 12 **UNIT-III**

STRUCTURE

12.1. OBJECTIVES

After going through this lesson, you would be able to understand :

1. Meaning of Moments & Purpose of Moments
2. Meaning of Kurtosis & various types of Kurtosis

12.2. INTRODUCTION

In this unit, we will discuss about two very important concepts i.e., Moments and Kurtosis.

12.3. MOMENTS

According to **Trederic Mills**, "Moment is a familiar mechanical term for the measure of a force with reference to its tendency to produce rotation. The strength of this tendency depends, obviously, on the amount of force and the distance from the origin of the point at which the force is exerted". The above definition reveals that the term 'moment' in Mechanics refers to the measure of a force with reference to its tendency to produce rotation.

Moreover, the strength of this force depends upon (a) the amount of force and (b) the distance from the origin at which the force is applied. If on both sides of the origin, the amount of force X distance are equal, there will be a balance, otherwise not.

In statistics, the use of this term is almost analogous. In a frequency distribution, the class frequencies are looked upon as the forces and the deviations of the different values from the mean (or any other central value) are considered to be distances. This means if forces (frequencies) are represented by f_1, f_2, f_3 - and distance as x_1, x_2, x_3 - the f_1x_1 is the moment of the first force, $f_2\, x_2$ of the second, $f_3\, x_3$ of third and so on.

On adding these moments we get Σfx and when it is divided by total force (frequency) Σf, we get the value of moment from $\frac{\Sigma fx}{\Sigma f}$.

Let the symbol x be used to represent the deviation of any item in a distribution from the arithmetic mean of that distribution. The arithmetic mean of the various powers of these deviations in any distribution are called the moments of the distribution. If we take the mean of the first power of the deviations, we get the first moment about the mean; the mean of the squares of the deviation gives us the second moment about the mean; the mean of the cubes of the deviations gives us the third moment about the mean; and so on. The moments about mean (also called central moments) are denoted by Greek latter μ (read as mu); thus μ_1 stands for first moment about mean, μ_2 stands for second moment about mean etc.

Symbolically

$$\mu_1 = \frac{\Sigma\left(X - \overline{X}\right)}{N} \quad \text{or} \quad \frac{\Sigma X}{n}$$

[Since sum of deviations from arithmetic mean is always zero, μ_1 would be zero.]

$$\mu_2 = \frac{\Sigma(X - \overline{X})^2}{N} \quad \text{or} \quad \frac{\Sigma X^2}{N}$$

$$\mu_3 = \frac{\Sigma(X - \overline{X})^3}{N} \quad \text{or} \quad \frac{\Sigma X^3}{N}$$

For a frequency distribution

$$\mu_1 = \frac{\Sigma f(X-\overline{X})}{N} \quad \text{or} \quad \frac{\sum fX}{N}; \quad \mu_2 \frac{\sum f\left(X-\overline{X}\right)^2}{N} = \frac{\sum fX^2}{N}$$

$$\mu_3 = \frac{\sum f(X-\overline{X})^3}{N} \quad \text{or} \quad \frac{\sum fX^3}{N}$$

$$\mu_4 = \frac{\sum f(X-\overline{X})^4}{N} \quad \text{or} \quad \frac{\sum fX^4}{N}$$

Assumed Mean: When the arithmetic mean of a series is not a whole number but a fraction, deviations are to be taken from assumed mean to avoid tedious calculations. In that case be, the formulae will be :

$$\mu_1' = \frac{1}{N}\Sigma(X-A) \qquad \mu_2' = \frac{1}{N}\Sigma(X-A)^2$$

$$\mu_3' = \frac{1}{N}\Sigma(X-A)^3 \qquad \mu_4' = \frac{1}{N}\Sigma(X-A)^4$$

(Where $N = \Sigma f$)

These moments about assumed mean (also called raw moments) are then converted into moments about actual mean by using the following relationship:

$$\mu_1 = \mu_1' - \mu_1'^2 = 0$$

$$\mu_2 = \mu_2' - (\mu')_1^2$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

Things to be noted :

The first four moments about an arbitrary point 'A' are of great importance for the calculation of measures of central tendency and dispersion

e.g. $\overline{X} = \mu_1'$ (about origin)

$\mu_2 = \sigma^2$ (Dispersion)

$\mu_3 = \beta_1$ (Skewness)

and $\beta_2 = \frac{\mu_4}{\mu_2^2}$ (Kurtosis)

The last concept shall be discussed in a later section.

Moments about Zero

The moments about Zero are often denoted by V_1, V_2, V_3 etc. and are obtained as follows:

$$V_1 = \frac{\sum fX}{N} \qquad V_2 = \frac{\sum fX^2}{N}$$

$$V_3 = \frac{\sum fX^3}{N} \qquad V_4 = \frac{\sum fX^4}{N}$$

Also

The first moments about Zero or $V_1 = A + \mu_1'$

The second moment about Zero or $V_2 = \mu_2 + (V_1)^2$

Third moment about Zero or $V_3 = \mu_3 + 3V_1V_2 - 2V_1^3$

The fourth moment about Zero or $V_4 = \mu_4 + 4V_1V_3 - 6V_1^2V_2 + 3V_1^4$

12.3.1. Purpose of Moments

Following are the basic objectives for which moments for a series are calculated.

They are :

1. Moments are calculated to study the nature of the distribution.
2. The calculation of moments helps us to know us about the nature of symmetry.
3. The moments tell us whether the distribution is flatter than the normal curve or is more Peaked than the normal curve.
4. They also tell us the type of symmetrical curve whether it is normal or U-shaped or a triangle etc.
5. Two important constants of a distribution are calculated from μ_2, μ_3 and μ_4. They are :

   (i) β_1 (read as beta one) $= \dfrac{\mu_3^{\,2}}{\mu_2^{\,3}}$

   (ii) β_2 (read as beta two) $= \dfrac{\mu_4}{\mu_2^{\,2}}$

   β_1 measures skewness & β_2 measures kurtosis

Conclusion: From the above, it can be concluded that 'moment' in Statistics describes the balance of values about the mean. If the two sides about mean are balanced, the series is in symmetry otherwise not. The nature of symmetry & kurtosis are also obtained through moments.

12.4. KURTOSIS

Kurtosis is another measure which gives a description about the nature and form of a distribution. The word 'kurtosis' has been derived from. Greek language which means 'bulginess'. Kurtosis is measured by coefficient β_2, $\beta_2 = \frac{\mu_4}{\mu_2^{\,2}}$ where μ_4 is the fourth moment of the distribution and μ_2 is the second moment of the distribution. The standard value of β_2 is taken to be 3. When β_2 is less than 3, it is called a Platykurtic, if the value comes out to be greater than 3, then it is leptokurtic and if the value is equal to 3, the curve is mesokurtic or normal curve.

Some times γ_2, the derivative of β_2, is used as a measure of kurtosis. γ_2 is defined as

$$\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3$$

For a normal distribution, $\gamma_2 = 0$. If γ_2 is positive, the curve is leptokurtic and if γ_2 is negative, the curve is platykurtic.

The following diagram illustrates the shape of three different curves mentioned above :

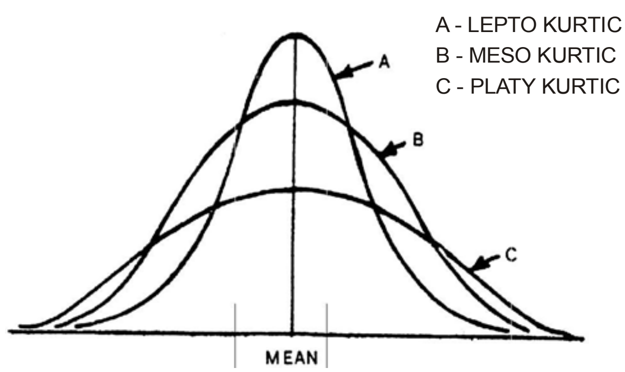


Fig.

The above diagram clearly shows that these curves differ widely with regard to convexity, an attribute which Karl Pearson referred to as 'Kurtosis'. While skewness shows the direction of dispersion, Kurtosis shows the shape and hump (middle part) of a frequency distribution or it deals with the flatness and peakedness of the frequency curve. Curve of type B is known as Normal Curve, as it is neither flat nor peaked. It is called mesokurtic. Curves more peaked than B or normal curve are said to lack Kurtosis or having negative Kurtosis or are leptokurtic. Here curve A is leptokurtic. Curve C, flatter than curve B, is called platykurtic, as it shows positive Kurtosis.

Here we can add the remark and sketch given by a British Statistican W.S. Gosset in the following way "Platykurtic curves like the platypus, are squat with short tails; leptokurtic curves are high with long tails like the kangaroos noted for leaping."

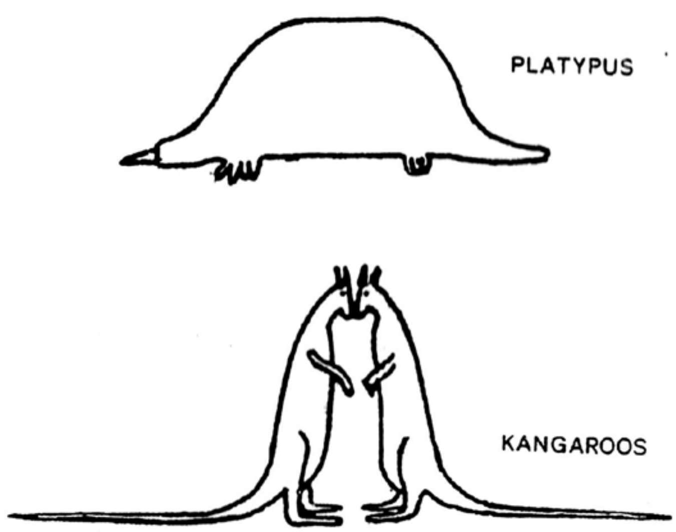


Fig. (Source-Business Statistics by SC Gupta & Indra Gupta)

Definitions

Simpson and Kafka: "The degree of kurtosis of a distribution is measured relative to the Peak of a normal curve."

Croxton and Cowden: "A measure of Kurtosis indicates the degree to which a curve of a frequency distribution, is Peaked or flat topped."

Example 1: Analyse the frequency distribution by the method of moment

X	2	3	4	5	6
f	1	3	7	3	1

Solution: Calculation of first four moments

X	f	$(x)=(X-\overline{X})$, $\overline{X}=4$	fx	fx^2	fx^3	fx^4
2	1	-2	-2	4	-8	16
3	3	-1	-3	3	-3	3
4	7	0	0	0	0	0
5	3	1	3	3	3	3
6	1	2	2	4	8	16
	N=15		$\Sigma fx=0$	$\Sigma fx^2=14$	$\Sigma fx^3=0$	$\Sigma fx^4=38$

$$\mu_1 = \frac{\sum f(X-\overline{X})}{N} = \frac{\sum fX}{N} = \frac{0}{15} = 0$$

$$\mu_2 = \frac{\sum f(X-\overline{X})^2}{N} = \frac{\sum fX^2}{N} = \frac{14}{15} = 0.933$$

$$\mu_3 = \frac{\sum f(X-\overline{X})}{N} = \frac{\sum fX^3}{N} = 0$$

$$\mu_4 = \frac{\sum f(X-\overline{X})^4}{N} = \frac{\sum X^4}{N} = \frac{38}{15} = 2.533$$

$$\sigma = \sqrt{\text{Variance}} \quad \text{or} \quad = \sqrt{\mu_2}$$

$$\sigma = \sqrt{0.933} = 0.966$$

$$\beta_1 = \frac{{\mu_3}^2}{{\mu_2}^3} = \frac{0^2}{(0.933)^3} = 0$$

In a symmetrical distribution β_1 is zero. Hence this distribution is symmetrical

$$\beta_2 = \frac{\mu_4}{\mu_2{}^2} = \frac{2.533}{(0.933)^2} = \frac{2.533}{0.87} = 2.91$$

Since the value of β_2 is less than three, the distribution is platykurtic.

Example 2: The first four moments of a distribution about X=2 are 1, 2:5, 5.5 and 16. Calculate the four moments about $\overline{X}$ and about zero.

Solution: We are given μ_1' =1, μ_2' = 2.5, μ_3' = 5.5 and μ_4' = 16. From these moments about arbitrary origin we can find out moments about mean with the help of the following relationships:

$\mu_2 = \mu_2' - (\mu_2')^2$

$\mu_3 = \mu_3' - 3\mu_1' \mu_2' + 2(\mu_1')^3$

$\mu_4 = \mu_4' - 4\mu_1' \mu_3' + 6(\mu_2')(\mu_1') - 3(\mu_1')^4$

Substituting the values

$\therefore \mu_2 = 2.5 - (1)^2 = 1.5$

$\mu_3 = 5.5 - (1)(2.5) + 2(1)^3 = 5.5 - 7.5 + 2 = 0$

$\mu_4 = 16 - 4(1)(5.5) + 6(2.5)(1)^2 - 3(1)^4 = 16 - 22 + 15 - 3 = 6$

Thu moments about mean are $\mu_1 = 0, \mu_2 = 1.5, \mu_3 = 0$

$\mu_4 = 6$

Moments about zero

Let moments about zero be denoted by V_1, V_2, V_3 etc.

The first moment about zero i.e. $V_1 = A + \mu_1'$

The second moment about zero i.e. $V_2 = \mu_3 + V_1{}^2$

The third moment about zero i.e. $V_3 = \mu_3 + V_1V_2 - 2V_1{}^3$

The fourth moment about zero is $V_4 = \mu_4 + 4V_1V_3 - 6V_1{}^2V_2 + 3V_1{}^4$

$V_1 = 2 + 1 = 3$

$V_2 = 1.5 + (3)^2 = 10.5$

$V_3 = 0 + (3 \times 3 \times 10.5) - 2(3)^3 = 94.5 - 54 = 40.5$

$V_4 = 6 + 4(3)(40.5) - 6(3)^2(10.5) + 3(3)^4$

$= 6 + 486 - 567 + 243 = 168$

Example 3: The first four central moments of a distribution are, 0, 2.5, 0.7 and 18.75. Comment on the skewness and kurtosis of the distribution

Solution: Testing Skewness:

We are given $\mu_1 = 0, \mu_2 = 2.5, \mu_3 = 0.7$ and $\mu_4 = 18.75$

Skewness is measured by the coefficient β_1

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

Here $\mu_3 = 0.7,\ \mu_2 = 2.5$

Substituting the values $\beta_1 = \dfrac{(0.7)^2}{(2.5)^3} = +0.031$

Since $\beta_1 = +0.031$, the distribution is slightly skewed i.e. it is not perfectly symmetrical.

Testing Kurtosis

For testing kurtosis we compute the value β_2. When a distribution is normal or symmetrical, $\beta_2 = 3$. When a distribution is more peaked than the normal, β_2 is more than 3 and when it is less peaked than the normal, β_2 is less than 3.

$$\beta_2 = \frac{\mu_4}{\mu_2^2},\ \mu_4 = 18.75,\ \mu_2 = 2.5$$

$$\beta_2 = \frac{18.75}{(2.5)^2} = \frac{18.75}{6.25} = 3$$

Since β_2 is exactly three, the distribution is mesokurtic.

12.5. LET US SUM UP

In this lesson, we discussed the concepts of moments and kurtosis.

12.6. EXAMINATION ORIENTED QUESTIONS

Q.1. From the following information calculate the first four moments about mean and analyse the distribution.

X	0	1	2	3	4	5	6	7
f	1	8	28	56	70	56	28	8

Hint: $\mu_1 = \frac{\Sigma f(X-\overline{X})}{N}$, $\mu_2 = \Sigma f(X-\overline{X})^2$, $\mu_3 = \frac{\Sigma f(X-\overline{X})^3}{N}$

$$\mu_4 = \frac{\Sigma f(X-\overline{X})^4}{N}, \ \sigma = \sqrt{\mu_2}, \ \beta = \frac{\mu_4}{\mu_2^{\ 2}}$$

Ans. $\mu_1 = 0, \mu_2 = 2, \mu_3 = 0, \mu_4 = 11, \sigma = 1.4142, \beta_2 = 2.75$

Q.2. From the following data calculate the first four moments (1) about 25 (II) about mean

Weekly income (in Rs)	0-10	10-20	20-30	30-40
No. of workers	1	3	4	2

Hint: moment about A = 25, $\mu_1' = \frac{\Sigma fd}{N} \times i$, $\mu_2' = \frac{\Sigma fd}{N} i^2$

$$\mu_3' = \frac{\Sigma fd^3}{N} \times i^3, \quad \mu_4' = \frac{\Sigma fd^4}{N} i^2$$

Ans. $\mu_1' = -3$, $\mu_2' = 90$, $\mu_3' = -900$, $\mu_4' = -21{,}000$

$\mu_1 = 0$, $\mu_2 = 81$, $\mu_3 = -144$, $\mu_4 = 14{,}817$

Q.3 The first for moments of a distribution are 0, 2.5, 0.7 and 18.75. Find kurtosis and comment

Ans. $\beta_2 = 3$ distribution is normal / mesokurtic curve

Q.4 N = 10, $\Sigma fdx = 50$, $\Sigma fdx^2 = 1967.2$, $\Sigma fdx^3 = 2925.8$, $\Sigma fdx^4 = 86650.2$

Find if the curve is platykurtic or not

Ans. $\beta_2 = 2.22$ (i.e <3) $\therefore$ the curve is platykurtic.

Q.5. Find the first four central moments of the following number:

1, 2, 8, 9, 10

Sol. Moment $\overline{X} = 6$

X	$(X-\overline{X})$	$(X-\overline{X})^2$	$(X-\overline{X})^3$	$(X-\overline{X})^4$
1	–5	25	–125	625
2	–4	16	–64	256
8	2	4	8	64
9	3	9	27	81
10	4	10	64	256
$\Sigma X=30$	$\Sigma(X-\overline{X})=0$	$\Sigma(X-\overline{X})^2=70$	$\Sigma(X-\overline{X})^3=-90$	$\Sigma(X-\overline{X})^4=1282$

$$\overline{X} = \frac{\Sigma X}{N} = \frac{30}{5} = 6$$

$$\mu_1 = \frac{\Sigma(X-\overline{X})}{N} = \frac{0}{5} = 0$$

$$\mu_2 = \frac{\Sigma(X-\overline{X})^2}{N} = \frac{70}{5} = 14$$

$$\mu_3 = \frac{\Sigma(X-\overline{X})^3}{N} = \frac{-90}{5} = -18$$

$$\mu_4 = \frac{\Sigma(X-\overline{X})^4}{N} = \frac{1282}{5} = 256.4$$

Q.6. What are the types of Kurtosis?

Q.7. What are moments? Write their utility and purpose.

Ans. (I) Moment is a familiar mechanical term which refers to the measures of a force with respect to its tendency to provides rotation. The strength of the tendency depends on the amount of force and the distance from the origin of the point at which the force is exerted.

(II) The following is the summary of how moments helps in analysing a frequency distribution.

	Moment	*What is measures*
(i)	First moment about origin or zero	Mean
(ii)	Second moment about the mean	Variance
(iii)	Third moment about the mean	Skewners
(iv)	Fourth moments about the mean	Kurtyosis

(III) The concept of moment is of great significance in statistical work. With the help of moments, we can measure the central tendency of a set of observations, their variability, asymmetry and the height of the peak of their curve.

12.7. SUGGESTED READINGS & REFERENCES

1. S.P. Gupta (2005), Statistical Methods, S. Chand and Sons, New Delhi
2. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas publishing House, New Delhi.
3. Monga, G.S. (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.
4. Gupta, S.C. and Indra Gupta, Business statistics.

Unit IV

Bivariate Data

COVARIANCE

LESSON NO. 13 **UNIT-IV**

STRUCTURE

13.1. OBJECTIVES

After going through this lesson, you would be able to understand :

1. Bivariate distribution of X & Y
2. Covariance and Correlation
3. Correlation Coefficient
4. Methods of Studying Correlation

5. Coefficient of Correlation and Probable Error
6. Coefficient of Determination
7. Properties of the Coefficient of Correlation

13.2. INTRODUCTION

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.

13.3. BIVARIATE DISTRIBUTION OF X AND Y

In the earlier chapter, we studied univariate distributions, where observations were given on a single variable. We obtained a measure of central tendency (like the arithmetic mean, median, etc. and a measure of dispersion (like the standard deviation, mean deviation etc) of the set of values of the variable.

Now, suppose we have observations on two variables, X and Y for several individuals. Here, we have a bivariate distribution of X and Y.

We can still calculate the central value (arithematic mean, median, etc) and dispersion (standard deviation, mean deviation etc.) of each variable, X and Y separately. However, we would also like to know, if there is any association between the values of the two variables. We would like to know for example, how does the value of one variable change if the value of the other increases or decreases by a certain amount ? Is the change in the same direction and in the same proportion or the change is more or less than proportionate and in the reverse direction?

A numerical measure of association between two variables is given by Karl Pearson's Coefficient of Correlation.

13.4. COVARIANCE AND CORRELATION

Covariance and Correlation describe how two variables are related.

Variables are positively related if they move in the same direction. Variables are inversely related if they move in opposite directions. Both Covariance and Correlation indicate whether variables are positively or inversely related. Correlation also tells us the degree to which the variables tend to move together.

Covariance can also be explained with help of an example. For example, if the economic growth increases, stock market returns tend to increase as well.

These variables are said to be positively related because they move in the same direction. You may also hear that as world oil production increases, gasoline prices fall. These variables are said to be negatively or inversely related because they move in opposite directions.

The relationship between two variables can be illustrated in a graph. In the examples below, the graph on the left illustrates how the positive relationship between economic growth and market returns might appear. The graph indicates that as economic growth increases, stock market returns also increase. The graph on the right is an example of how the inverse relationship between oil production and gasloline prices might appear. It illustrates that as oil production increases, gas prices fall.

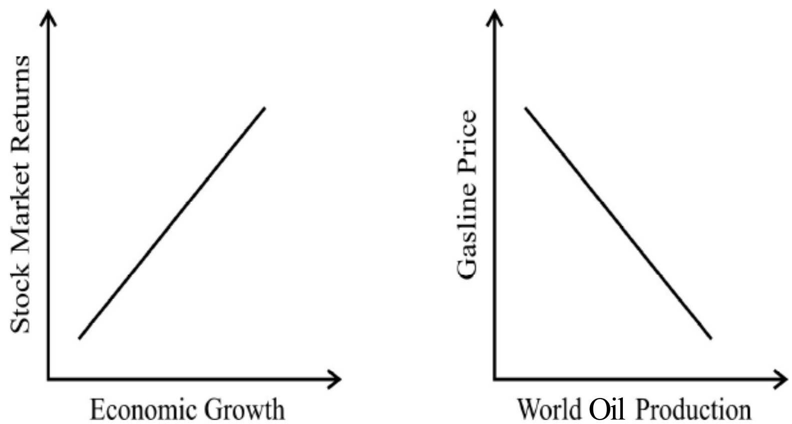


Fig.

To determine the actual relationships of these variables, you would use the formulas for covariance and Correlation.

13.5. COVARIANCE

Covariance indicates how the 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 sample data is shown below.

$$\text{Cov }(X,Y) = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{n-1}$$

X = the independent variable

Y = the dependent variable

n = number of data points in the sample

$\overline{X}$ = the mean of the independent variable X

$\overline{Y}$ = 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 stock market (Y_i)

Economic Growth % (X_i)	Stock Market return % (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 stock market returns have a positive or inverse relationship. Before you compute the Covariance, Calculate the mean of *x* and *y*.

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 (Stock Market return)

$\bar{x} = 3.1$

$\bar{y} = 11$

Substitute these values in to the Covariance formula to determine the relationship between economic growth and share market return.

The Covariance between the return of share market and economic growth is 1.53. Since the Covariance is positive, the variables are positively related — they move together in the same direction.

13.6. CORRELATION COEFFICIENT

Correlation: Correlation is another way to determine how the two variables are related. In addition to telling you whether variables are positively or inversely related, correlation also tells you the degree to which variables tend to move together.

As stated above, Covariance measures variables that have different units of measurement. Using Covariance, you could determine whether units were increasing or decreasing, but it was impossible to measure the degree to which the variables moved together because Covariance does not use one standard unit of measurement. To measure the degree to which variables move together, you must use correlation.

Correlation standardizes the measure of interdependence between two variables and Consequently, tells you how closely the two variables move. The correlation measurement, called a Correlation Coefficient, will always take on a value between 1 and –1.

If the Correlation Coefficient is one, the variables have a perfect positive Correlation. this means that if one variable moves a given amount, the second moves proportionally in the same direction.

If the Correlation Coefficient is Zero, no relationship exists between the variables. If one variable moves, you can make no predictions about the movement of the other variable; they are uncorrelated.

If Correlation Coefficient is –1; the variables are perfectly negatively correlated (or inversely correlated) and move in opposite direction to each other. If one variable increases, the other variable decreases proportionally. To calculate the Correlation Coefficient for two variables, you would use the correlation formula, shown below :

$$r_{(x,y)} = \frac{Cov(x, y)}{Sx.Sy}$$

$r(x,y)$= Correlation of the variables x and y

Cov(x,y) = Covariance of the variables x and y

Sx = Sample standard deviation of the random variables x

Sy = Sample standard deviation of the random variable y

Importance of Correlation

(1) For the formation and testing of economic laws: There are so many economic laws and assumptions like law of demand, supply etc. which tells that when price of commodity increases, demand contracts and if price decreases, demand extension takes place. This law can also be tested by scatter diagram. We can study the correlation between demand and Price by collecting data at different rates. If correlation is negative, then the law holds true. Rest of the laws can also be tested.

(2) For research purpose: From correlation we can discover new laws also. In this way correlation can also be used in research.

(3) Policy formulation: From the result of the correlation we can formulate the policies for the solution of the problems.

13.7. METHODS OF STUDYING CORRELATION

The various methods of ascertaining whether the two variables are correlated or not are:

I. Scatter Diagram Method
II. Graphic Method
III. Karl Pearson's Coefficient of Correlation
IV. Rank Method

Of these, the first two are based on the knowledge of diagram and graphs whereas the others are the mathematical methods. Each of these methods shall be discussed in detail.

I. Scatter Diagram Method

In a scatter, we plot the values of the two variables, as a set of points, on a graph paper. The cluster of points, so obtained is called the scatter diagram. Let us illustrate this with following example.

Example 1: Suppose we have observations on

(i) the monthly income (X) and
(ii) the total monthly expenditure on food (Y), in rupees, of five rural househoulds, as shown inTable.

Monthly Income and Expenditure on Food of Five rural Households:

Variables	Households				
	1	2	3	4	5
Income (X) in Rs	550	600	800	700	650
Expenditure on food (Y) in Rs.	400	450	550	550	400

We note that the income of the first household is Rs 550 per month and its expenditure on food per month is Rs. 400. We may plot this as a point (X,Y) on the graph paper, where X = 550 and Y = 400. We measure 550 along the X-axis and 400 along the Y-axis. Similarly, for the second household X = 600 and Y = 450 so that the co-ordinates of the second point are (600, 450). We measure 600 along the X-axis and 450 along Y-axis, and so on.

The cluster of points is shown in fig. this is called Scatter diagram.

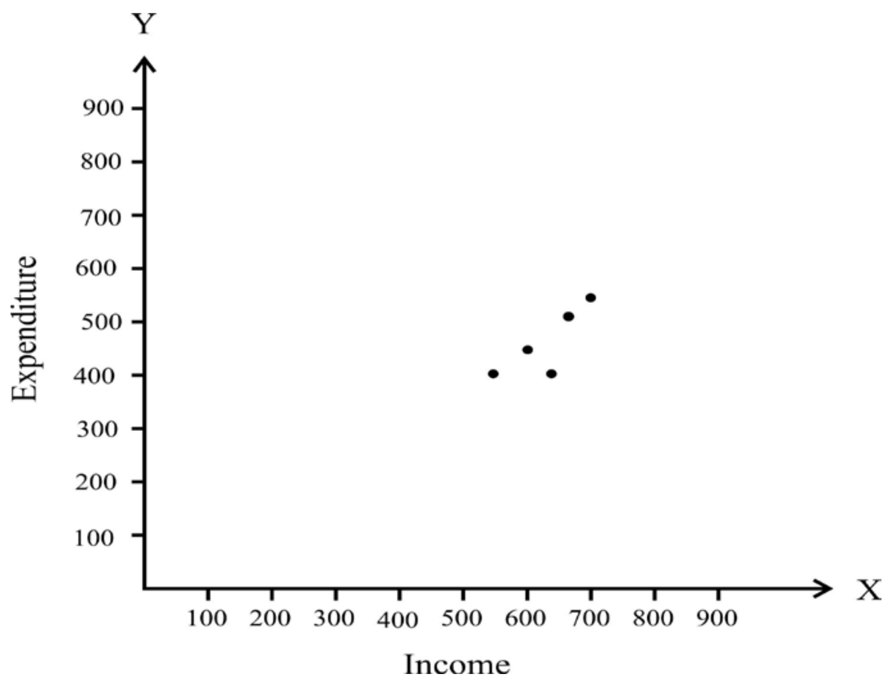


Fig. Scatter Diagram.

It is clear from the scatter diagram that the points tend to cluster about an upward sloping straight line. In other words, the expenditure on food (Y) increases as the income (X) of the household increases.

In general, if the straight line is sloping upward, the values of X and Y change in the same direction (i.e an increase in the value of X is accompanied with an increase in the value of Y). Otherwise, if the straight line is sloping downwards, an increase in the value of X is accompanied with a decrease in the value of Y.

It is the slope of the straight line (which depends on the angle that the straight line makes with the X-axis and is equal to $\frac{Y}{X}$) that determines the rate of change in the value of Y for a unit changes in the value of X.

MERITS AND DEMERITS OF A SCATTER DIAGRAM

Merits

(a) It is easy to draw a scatter diagram

(b) It is an easy first step in determining the form of relationship (linear or non-linear) between two variables.

(c) In case of linear relationship between Y and X, it gives a clear visual picture of the proportionate change in the value of Y for a change in the value of X.

Demerits

(a) The strength, or the degree of association between two variables cannot be determined in numerical terms from a scatter diagram.

(b) Scatter diagram does not indicate the direction of causation. It does not tell, if Y causes X or X cause Y.

II. Graphic Method

When this method is used, the individual values of the two variables are plotted on the graph paper. We thus obtain two curves, one for X variable and another for Y variable. By examining the direction and closeness of the two curves so drawn, we can infer whether or not the variables are related. If both the curves drawn on the graph are moving in the same direction (either upward or downward), Correlation is said to be positive.

On the other hand, if the curves are moving in the opposite directions, correlation is said to be negative.

The following example shall illustrate the method.

Illustration: From the following data, ascertain whether the income and expenditure of the workers of a factory are correlated:

Year	Average income (in Rs.)	Average expenditure (in Rs.)
1970	100	90
1971	102	91
1972	105	93
1973	105	95
1974	101	92
1975	112	94
1976	118	100
1977	120	105
1978	125	108
1979	130	110

Solution: The graph shows that the variables, income and expenditure are closely related.

Fig.

This method is noramlly used where we are given data over a period of time i.e, in case of time series. However, as with the scatter diagram method, in this method also we cannot get a numerical value describing the extent to which the variables are related.

III. Karl Pearson's Coefficient of Correlation

Karl Pearson's Coefficient of Correlation Of the several mathematical methods of measuring correlation, the Karl Pearson's method, popularly known as Pearsonian Coefficient of Correlation is denoted by the symbol *r*. It is one of the very few sybmols that are used universally for describing the degree of correlation between two series.The formula for computing Pearsonian *r* is:

$$r = \frac{\sum xy}{N\sigma_x\sigma_y} \qquad ...(1)$$

$x = (X - \overline{X}),\ y = (Y - \overline{Y})$

σ_x = Standard deviation of series X

σ_y = Standard deviation of series Y.

N = Number of pairs of observations.

r = the (product moment) correlation coefficient.

This method is to be applied only where the deviations of items are taken from actual mean and not from assumed mean.

The value of the Coefficient of Correlation as obtained by the above formula shall always lie between ± 1. When $r = +1$, it means there is perfect positive correlation between the variables. when $r = -1$, it means there is perfect negative correlation between the variables. when $r = 0$, it means there is no relationship between the two variables. However, in practice, values of *r* as +1, –1 and 0 are rare. We normaly get value which lie between +1 and –1 such as 0.8, –0.4 etc. The coefficient of correlation describes not only the magnitude of correlation but also its direction. Thus, +0.8 would mean that correlation is positive because the sign of *r* is + and the magnitude of Correlation is 0.8. Similarly –0.4 means correlation is negative.

The above formula for computing Perasonian coefficient of correlation can be transformed to the following form which is easier to apply

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}} \qquad ...(2)$$

where $x = X - \overline{X}$ and $y = Y - \overline{Y}$

The coefficient of correlation is said to be a measure of Covariance between two series. The Covariance of two series X and Y is written as

$$\text{Covariance} = \frac{\Sigma xy}{N}$$

Where x and y stands for derivations of X and Y series from their respective means.

In order to find out the value of correlation Coefficient, first we calculate Covariance and then in order to convert it to a relative measure we divide the Covariance by the Standard deviation of two series. The ratio so obtained is called Karl Pearson's Coefficient.

$$r = \frac{\Sigma xy}{N\sigma_x \sigma_y}, \quad \sigma_x = \sqrt{\frac{\Sigma x^2}{N}}, \quad \sigma_y = \sqrt{\frac{\Sigma y^2}{N}}$$

$$r = \frac{\Sigma xy}{N\sqrt{\frac{\Sigma x^2}{N} \cdot \frac{\Sigma y^2}{N}}} = \frac{\Sigma xy}{\sqrt{\Sigma x^2 . \Sigma y^2}}$$

It is obvious that while applying this formula we have not to calculate separately the standard deviations of X and Y series as is required by formula (1).

This simplifies greatly the task of calculating correlation coefficient.

Steps

(i) Take the deviation of X-series from the mean of X and denote these deviations by x.

(ii) Square these deviations and obtain the total, i.e. Σx^2.

(iii) Take the deviations of Y-series from the mean of Y and denote these deviations by Y.

(iv) Square these deviations and obtain the total i.e. Σy^2.

(v) Multiply the deviations of X and Y series and obtain the total i.e. Σxy.

(vi) Substitute the values of Σxy, Σx^2, and Σy^2 in the above formula.

The following examples will illustrate the procedure.

Example 1. Calculate the Coefficient of Correlation from the following data.

X	9	8	7	6	5	4	3	2	1
Y	15	16	14	13	11	12	10	8	9

Solution: Calculation of Coefficient of Correlation

X	$(X-\overline{X})=x$	x^2	Y	$(Y-\overline{Y})=y$	y^2	xy
9	+4	16	15	+3	9	12
8	+3	9	16	+4	16	12
7	+2	4	14	+2	4	4
6	+1	1	13	+1	1	1
5	0	0	11	–1	1	0
4	–1	1	12	0	0	0
3	–2	4	10	–2	4	4
2	–3	9	8	–4	16	12
1	–4	16	9	–3	9	12
$\Sigma X=45$	$\Sigma x=0$	$\Sigma x^2=60$	$\Sigma Y=108$	$\Sigma y = 0$	$\Sigma y^2= 60$	$\Sigma xy=57$

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$$

$$x = (X-\overline{X}) \text{ and } \quad y = (Y-\overline{Y})$$

$$\overline{X} = \frac{45}{9} = 5, \quad \overline{Y} = \frac{108}{9} = 12$$

$\Sigma xy = 57,\ \Sigma x^2 = 60,\ \Sigma y^2 = 60$

$$r = \frac{57}{\sqrt{60 \times 60}} = \frac{57}{60} = 0.95$$

Example 2. Making use of the data summarized below, calculate the coefficient of correlation r_{12}.

Case	X_1	X_2
A	10	9
B	6	4
C	9	6
D	10	9
E	12	11
F	13	13
G	11	8
H	9	4

Solution: Calculation of Coefficient of Correlation:

Case	X_1	$(X_1 - \overline{X})$	x_1^2	X_2	$(X_2 - \overline{X_2})$	x_2^2	$x_1.x_2$
A	10	0	0	9	+1	1	0
B	6	–4	16	4	–4	16	16
C	9	–1	1	6	–2	4	2
D	10	0	0	9	+1	1	0
E	12	2	4	11	+3	9	6
F	13	3	9	13	+5	25	15
G	11	1	1	8	0	0	0
H	9	–1	1	4	–4	16	4
N=8	$\Sigma X_1=80$	$\Sigma x_1=0$	$\Sigma x_1^2=32$	$\Sigma X_2=64$	$\Sigma x_2=0$	$\Sigma x_2^2=72$	$\Sigma x_1 x_2=43$

$$\overline{X} = \frac{\Sigma X_1}{N} = \frac{80}{8} = 10,\quad \overline{X_2} = \frac{\Sigma X_2}{N} = \frac{64}{8} = 8$$

$$r_{12} = \frac{\Sigma x_1.x_2}{\sqrt{\Sigma x_1^{\,2}.\Sigma y_2^{\,2}}}$$

$\Sigma x_1.x_2 = 43.\ \Sigma x_1^{\,2} = 32,\ \Sigma x_2^{\,2} = 72$

Substituting the values

$$r_{12} = \frac{43}{\sqrt{32 \times 72}} = \frac{43}{\sqrt{2304}}$$

$$= \frac{43}{48} = +0.896$$

Note: It may be noted that the above formula is the same as given earlier i.e.

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2.\Sigma y^2}}$$

The only difference is that of the symbols. Since in this example, we are given series X_1 and X_2, we changed the symbols in the formula accordingly.

Direct Method of Finding out Correlation

Correlation Coefficient can also be calculated without taking deviations of items either from actual mean or assumed mean, i.e. actual X and Y values. The formula in such a case is

$$r = \frac{N\Sigma XY - \Sigma X\Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

This formula would give the same answer as we get when deviation of items are taken from actual or assumed mean. The following example shall illustrate the point.

Example : Calculate the coefficient of Correlation by using direct method i.e, without taking the deviations of items from actual or assumed mean.

X	9	8	7	6	5	4	3	2	1
Y	15	16	14	13	11	12	10	8	9

Solution: Calculation of Correlation coefficient by direct method.

$$r = \frac{N\Sigma XY - \Sigma X\Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

X	X^2	Y	Y^2	XY
9	81	15	225	135
8	64	16	256	128
7	49	14	196	98
6	36	13	169	78
5	25	11	121	55
4	16	12	144	48
3	9	10	100	30
2	4	8	64	16
1	1	9	81	9
$\Sigma X=45$	$\Sigma X^2=285$	$\Sigma Y=108$	$\Sigma Y^2=1356$	$\Sigma XY=597$

$N = 9$, $\Sigma XY = 597$, $\Sigma X = 45$, $\Sigma Y = 108$, $\Sigma X^2 = 285$, $\Sigma Y^2 = 1356$

$$r = \frac{N\Sigma XY - \Sigma X\Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{9 \times 597 - 45 \times 108}{\sqrt{9 \times 285 - (45)^2}\sqrt{9 \times 1356 - (108)^2}}$$

$$r = \frac{5373 - 4860}{\sqrt{2565 - 2025}\sqrt{12204 - 11664}}$$

$$= \frac{513}{\sqrt{540 \times 540}} = \frac{513}{540} = +0.95$$

Assumptions of the Pearsonian Coefficient

The Karl Pearson's Coefficient of Correlation is based on the following assumptions :

(1) There is linear relationship between the variables i.e. when the two variables are plotted on a scatter diagram, a straight line will be formed by the points so plotted.

(2) The two variables under study are affected by a large number of independent causes so as to form a normal distribution. Variables like height, weight, price, demand, supply, etc, are affected by such forces that a normal distribution is formed.

(3) There is a cause and effect relationship between the forces affecting the distribution of the items in the two series. In such a case, relationship is not formed between the variables i.e. if the variables are independent there cannot be any correlation. For example, there is no relationship between income and height because the forces that affect these variables are not common.

Merits and Limitations of the Pearsonian Coefficient

Merits

(1) It is ideal measure of correlation coefficient because it is based on arithmatic mean and standard deviation.

(2) From this, the direction of correlation is determined + shows positive and –shows negative correlations.

(3) From this method, numeric measure of correlation is found. It always remains between +1 and –1.

Amongst the mathematical methods used for measuring the degree of relationship, Karl Pearson's method is the most popular. The correlation coefficient summarizes in one figure not only the degree of Correlation but also the direction i.e. whether correlation is positive or negative.

The chief limitations of the method are :

(1) The Correlation Coefficient always assumes linear relationship regardless of the fact whether that assumption is correct or not.

(2) Great care must be exercised in interpreting the value of this coefficient, as very often the coefficient is misinterpreted.

(3) The value of the Coefficient is unduly affected by the extreme items.

(4) As compared with other methods, this method takes more time to compute the value of correlation coefficient.

13.8. COEFFICIENT OF CORRELATION AND PROBABLE ERROR

The Probable error of the coefficient of correlation helps in interpreting its value.With the help of probable error, it is possible to determine the reliability of the value of the coefficient in so far as it depends on the conditions of random sampling. The probable error of the Coefficient is obtained as follows.

$$\text{P.E}_r = 0.6745\frac{1-r^2}{\sqrt{N}}$$

where r is the Coefficient of Correlation and N is the number of pairs of observation.

(1) If the value of r is less than the probable error, there is no evidence of correlation i.e. the value of r is not at all significant.

(2) If the value of r is more than six times the Probable error, the Coefficient of Correlation is Practically certain, i.e. the value of r is significant.

(3) By adding and substracting the value of probable error from the Coefficient of Correlation, we get respectively the upper and lower limits within which coefficient of correlation in the population can be expected to lie symbolically,

$$\rho = r \pm \text{P.E}$$

Where ρ(rho) is correlation in the Population.

Let us compute probable error, assuming a Coefficient of correlation of 0.80 and a sample of 16 pairs of items. We will have

$$\text{P.E}_\text{r} = 0.6745\frac{1-0.8^2}{\sqrt{16}} = 0.06$$

The limits of the correlation in the Population would be $r \pm$ P.E, i.e. 0.8 ± 0.06 or 0.74 – 0.86.

Instances are quite common wherein a correlation coefficient of 0.5 or even 0.4 is obviously considered to be a fairly high degree of correlation by a writer or research worker. Yet a correlation coefficient of 0.5 means that only 25% of the variation is explained. A correlation Coefficient of 0.4 means that only 16% of the variation is explained.

Conditions for the use of Probable error

The measure of probable error can be properly used only when the following three conditions exist :

(1) The data must approximate a normal frequency curve (bell-shaped curve).

(2) The statistical measure for which the P.E is computed must have been calculated from a sample.

(3) The sample must have been selected in an unbiased manner and the individual items must be independent.

However, these conditions are generally not satisfied and as such reliability of the correlation coefficient is determined largely on the basis of exterior tests of reasonableness which are often of the statistical character.

Example: If r = 0.6 and N = 64, find out the Probable error of the Coefficient of correlation and determine the limits for population r.

Sol: $$P.E_r = 0.6745\frac{1-r^2}{\sqrt{N}}$$

$r = 0.6$ and $N = 64$

$$P.E. = 0.6745\frac{1-(0.6)^2}{\sqrt{64}} = \frac{0.6745 \times .64}{8} = 0.054$$

Limits of Population Correlation = 0.6 ± 0.054 = 0.0546 – 0.654

13.9. COEFFICIENT OF DETERMINATION

The Coefficient of determination, which is given as r^2, explains to what extent the variation of dependent variable Y is being expressed by the independent variable X.

$$r^2 = \frac{\text{variance explained}}{\text{total variance}}$$

More the value of r^2, better it is. A high value clearly shows that a good linear relation exists between the two variables. Obviously, if $r = 1$, then $r^2 = 1$ which is an indicator of Perfect relationship between the two variables. Also the quantity $(1-r^2)$ is called the coefficient of non-determination or coefficient of alienation $(1-r^2)$ is thus a measure of deviation from perfect linear relationship. For eg. if $r = 0.8$ it means high degree of correlation. But $r^2 = 0.64$ which means that only 64% of variation in the dependent variable are due to the independent variable, remaining 36% variations are due to other independent variables yet to be located.

Spurious Correlation: If we have paired observations on two variables and the assumption holds good, it is always possible to calculate the correlation coefficient between the two variables. But correlation coefficient is not always meaningful unless the two variables are properly chosen. For example, if we find out the correlation between the total imports and the number of shoes produced per year, it would not lead to any conclusion, because there is no cause and effect relationship between the total imports and number of shoes produced per year. Such a correlation is known as spurious or non-sense correlation between the two variables. Another example of spurious correlation is the correlation between the number of tourists who visited Taj Mahal and the number of TV sets produced. From this discussion,we come to the conclusion that it is not enough to find out the correlation coefficient only, but it is more important to see the nature of the variables. If they have a cause and effect relationship, and if one variable can be explained with the help of the other, then the correlation has sense, otherwise it is spurious.

13.10. PROPERTIES OF THE COEFFICIENT OF CORRELATION

The following are the important properties of the Correlation Coefficient r:

(1) The Coefficient of correlation lies between −1 and +1

Symbolically $-1 \leq r \leq +1$ or $|r| \leq |$

Proof: Let x and y be deviations of X and Y series from their mean, σ_x and σ_y be their standard deviation.

Expand the function:

$$\left(\frac{x}{\sigma x}+\frac{y}{\sigma y}\right)^2 = \Sigma\left[\frac{x^2}{\sigma x^2}+\frac{y^2}{\sigma y^2}+\frac{2xy}{\sigma x.\sigma y}\right]$$

$$= \frac{\Sigma x^2}{\sigma x^2} + \frac{\Sigma y^2}{\sigma y^2} + \frac{2\Sigma xy}{\sigma x.\sigma y}$$

But $\frac{\Sigma x^2}{\sigma x^2}$ = N $[\because \sigma x^2 = \frac{\Sigma x^2}{N} \therefore \frac{\Sigma x^2}{\sigma x^2} = \frac{\Sigma x^2}{\Sigma x^2} \times N = N]$

Similarly $\frac{y^2}{\sigma y^2}$ = N

$$\frac{2\Sigma xy}{\sigma x.\sigma y} = 2Nr \quad [\because r = \frac{\Sigma xy}{\sigma x.\sigma y}]$$

Hence $\left(\frac{x}{\sigma x} + \frac{y}{\sigma y}\right)^2 = N+N+2Nr = 2N+2Nr = 2N(1+r)$

But $\Sigma\left(\frac{x}{\sigma x} + \frac{y}{\sigma y}\right)$ is the sum of squares of real quantities and as such it cannot be negative, at the most it can be zero.

$\because \quad 2N(1+r) \geq 0$

Hence r cannot be less than –1, at the most it can be –1.

Similarly be expanding $\Sigma\left[\frac{x}{\sigma x} + \frac{y}{\sigma y^2}\right]^2$ it can be shown that this is equal to 2N(1–r)

This again cannot be negative; at the most it can be zero.

$\therefore$ r cannot be greater than +1, at the most it can be +1.

Hence $-1 \leq r \leq +1$

(2) The Co-efficient of Correlation is independent of change of scale and origin of the variable X and Y.

Proof: By change of origin we mean substracting some Constant from every given value of X and Y and by change of scale we mean dividing or multiplying every value of X and Y some Constant.

We know that $r = \dfrac{\Sigma\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)}{\sqrt{\Sigma\left(X-\overline{X}\right)^2 \Sigma\left(Y-\overline{Y}\right)^2}}$

Where $\overline{X}$ and $\overline{Y}$ refers to the actual means of X and Y series.

Let us now change the scale and origin. Deduct of fixed quatity a from X and b from Y. Also divide X and Y series by a fixed value i and c. After these changes are introduced, new values of x and y obtained from original X and Y shall be

$$x = \frac{X-a}{i} \text{ and } y = \frac{Y-b}{c}$$

$$\text{Mean of } x = \frac{\Sigma\left(\dfrac{X-a}{i}\right)}{N}$$

$$= \frac{\dfrac{\Sigma X - Na}{i}}{N} = \frac{\Sigma X - Na}{Ni}$$

But $\dfrac{\Sigma X - Na}{Ni} = \dfrac{\overline{X}-a}{i}$

Thus mean of $x = \dfrac{\overline{X}-a}{i}$

Similarly it can be shown that mean of $y = \dfrac{Y-b}{c}$

The value of the coefficient of correlation r for new set of values will be

$$r = \frac{\Sigma\left(\dfrac{X-a}{i}-\dfrac{\overline{X}-a}{i}\right)\left(\dfrac{Y-b}{c}-\dfrac{\overline{Y}-b}{c}\right)}{\sqrt{\Sigma\left(\dfrac{X-a}{i}-\dfrac{\overline{X}-a}{i}\right)^2 \times \Sigma\left(\dfrac{Y-b}{c}-\dfrac{\overline{Y}-b}{c}\right)^2}}$$

$$= \frac{\Sigma\left(\frac{X-a-\overline{X}+a}{i}\right)\left(\frac{Y-b-\overline{Y}+b}{c}\right)}{\sqrt{\Sigma\left(\frac{X-a-\overline{X}+a}{i}\right)^2 \times \Sigma\left(\frac{Y-b-\overline{Y}+b}{c}\right)^2}}$$

$$= \frac{\Sigma\frac{\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)}{ic}}{\sqrt{\Sigma\frac{\left(X-\overline{X}\right)^2}{i^2} \times \Sigma\frac{\left(Y-\overline{Y}\right)^2}{c^2}}}$$

$$= \frac{\Sigma\frac{\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)}{ic}}{\sqrt{\Sigma\frac{\left(X-\overline{X}\right)^2\left(Y-\overline{Y}\right)^2}{i^2c^2}}}$$

$$= \frac{\Sigma\left(X-\overline{X}\right)\Sigma\left(Y-\overline{Y}\right)}{\sqrt{\Sigma\left(X-\overline{X}\right)^2\Sigma\left(Y-\overline{Y}\right)^2}}$$

Thus the Coefficient of Correlation is independent of change of scale and origin.

13.11. RANK CORRELATION COEFFICIENT

It is not always possible to take measurements on units or objects. Many characters are expressed in Comparative terms such as beauty, smartness, temperament, etc. In such cases, the subjects are ranked pertaining to that particular character instead of taking measurements on them. Sometimes, the units are also ranked according to their quantitative measure. In these types of studies, two situations arise :

(i) the same set of units is ranked according to two characters A and B.

(ii) two judges give ranks to the same set of units, independently pertaining to one character only. In both these situations, we get paired ranks for a set of units. For example,

(a) two judges rank the girls independently in a beauty competition.

(b) the students are ranked according to their marks in Mathematics and Statistics.

In all these situations, the usual Pearsonian Correlation Coefficient cannot be obtained. Hence, the Psychologist, Charles Edward Spearman (1906) developed a formula for Correlation Coefficient, which is known as rank correlation or Spearman's Correlation. It is denoted by r_s. Suffix s to r is a Connotation for Spearman, the name of the inventor.

The formula for r_s is derived as the ratio of Covariance to the product of Standard deviation of two series of ranks. Here the derivation is omitted and thus the formula for rank correlation is

$$r_s = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)} \quad \text{or} \quad 1 - \frac{6\Sigma D^2}{N^3 - N}$$

Where r_s denotes rank coefficient of Correlation and D refers to the difference of ranks between paired items in two series.

The value of this Coefficient is interpreted in the same way as Karl Pearson's Correlation Coefficient and ranges between +1 and –1. When r_s is +1, there is complete agreement in the order of the rank and the ranks are in the same direction. When $r_s = -1$, there is complete agreement in the order of the ranks and they are in opposite directions.

In ranks correlation, we may have two types of problems:

(A) Where ranks are given

(B) Where ranks are not given

(A) Where Ranks are given: Where actual ranks are given to us, the steps required for computing rank correlation are:

(i) Take the differences of the two ranks, i.e. $(R_1 - R_2)$ and denote these differences by D.

(ii) Square these differences and obtained the total ΣD^2

(iii) Apply the formula $r_s = 1 - \frac{6\Sigma D^2}{N^3 - N}$

Example 1: Two judges gave the following ranks (from highest to the lowest) 1 to eleven girls who contested in a beauty competition. Whether or not, there is an agreement between the independent rankings of the two judges can be ascertained only by finding out the ranks Correlation between the ranks awarded by two judges.

Girls No.	1	2	3	4	5	6	7	8	9	10	11
	Rank										
Judge A	3	4	1	2	5	10	11	7	9	8	6
Judge B	2	4	3	1	7	9	6	11	10	5	8

Sol. Following the usual procedure, the rank correlation is calculated:

												Total
$D=R-R_2$	1	0	–2	1	–2	1	5	–4	–1	3	–2	0
D^2	1	0	4	1	4	1	25	16	1	9	4	66

$D = R_1 - R_2$, $D^2 = (R_1 - R_2)^2$

Here $\Sigma D^2 = 66$ and $N = 11$

$$r_s = 1 - \frac{6\Sigma D^2}{N^3 - N} = 1 - \frac{6 \times 66}{11^3 - 11} = 1 - \frac{396}{1331 - 11} = 1 - \frac{396}{1320} = 1 - 0.3 = 0.7$$

The value of rank correlation $r_s = 0.70$, which is quite high. Hence it can be concluded that there is an agreement between judges with regard to the beauty of the girls.

Examples 2: The rates of 12 students according to their marks in Mathematics and Statistics were as follows:

Student No.	1	2	3	4	5	6	7	8	9	10	11	12
Mathematics	5	2	1	6	8	11	12	4	3	9	7	10
Statistics	4	3	2	7	6	9	10	5	1	11	8	12

Sol.

												Total
D	1	–1	–1	–1	2	2	2	–1	2	–2	–1	–2	0
D^2	1	1	1	1	4	4	4	1	4	4	1	4	30

Here $\Sigma D^2 = 30$ and $N = 12$

$$r_s = 1 - \frac{6\Sigma D^2}{N^3 - N} = 1 - \frac{6 \times 30}{12^3 - 12} = 1 - \frac{180}{1728 - 12}$$

$$= 1 - \frac{180}{1716} = 1 - 0.1048 = 0.895$$

The Correlation between the ranks of marks in the two subjects is very high. From this, it is inferred that the students who are good in Mathematics are also good in statistics.

Examples 3: Two judges in a beauty Competition rank the 12 entries as follows:

X	1	2	3	4	5	6	7	8	9	10	11	12
Y	12	9	6	10	3	5	4	7	8	2	11	1

What degree of agreement in there between the judgement of the two judges?

Sol. Calculation of Rank Correlation Coefficient

X (R_1)	Y (R_2)	$R_1 - R_2$	D^2
1	12	–11	121
2	9	–7	49
3	6	–3	9
4	10	–6	36
5	3	+2	4
6	5	+1	1
7	4	+3	9
8	7	+1	1
9	8	+1	1
10	2	+8	64
11	11	0	0
12	1	+11	121
			$\Sigma D^2 = 416$

$r_s = 1 - \frac{6\Sigma D^2}{N^3 - N}$ Here $\Sigma D^2 = 416$ and $N = 12$

$$r_s = 1 - \frac{6 \times 416}{12^3 - 12} = 1 - \frac{6 \times 416}{12^3 - 12} = 1 - \frac{2496}{1716} = 1 – 1.454 = – 0.45$$

B. Where Ranks are not given

When we are given the actual data and not the ranks, it will be necessary to assign the ranks. Ranks can be assigned by taking either the highest values as 1 or the lowest value as 1. But whether we start with the lowest value or the highest value we must follow the same method in case of both the variables.

Example 1: Calculate the rank Correlation Coefficient for the following data.

X	92	89	87	86	83	77	71	63	53	50
Y	86	83	91	77	68	85	52	82	37	57

Sol. Calculation of Rank Correlation Coefficient:

X	R_1	Y	R_2	$(R_1–R_2)^2=D^2$
92	10	86	9	1
89	9	83	7	4
87	8	91	10	4
86	7	77	5	4
83	6	68	4	4
77	5	85	8	9
71	4	52	2	4
63	3	82	6	9
53	2	37	1	1
50	1	57	3	4
				$\Sigma D^2 = 44$

$$r_s = 1 - \frac{6\Sigma D^2}{N^3 - N}$$ Here $\Sigma D^2 = 44$ and $N = 10$

$$r_s = 1 - \frac{6 \times 44}{10^3 - 10} = 1 - \frac{264}{990} = 1 – 0.267 = 0.733$$

Example 2: Calculate the Coefficient of Rank correlation from the following data:

X	87	22	33	75	37
Y	29	63	52	46	48

Sol. Calculation of Rank Correlation Coefficient

X	R_1	Y	R_2	$(R_1-R_2)^2=D^2$
87	5	29	1	16
22	1	63	5	16
33	2	52	4	4
75	4	46	2	4
37	3	48	3	0
				$\Sigma D^2 = 40$

$$r_s = 1-\frac{6\Sigma D^2}{N^3-N} \qquad \text{Here } \Sigma D^2 = 40 \text{ and } N = 5$$

$$= 1-\frac{6\times 40}{5^3-5} = 1-2 = -1$$

Example 3: Calculate the coefficient of correlation from the following data by the Spearman's Rank Difference method.

Price of Tea (Rs.)	Price of Coffee (Rs.)
75	120
88	134
95	150
70	115
60	110
80	140
81	142
50	100

Sol. Calculation of Spearman's Rank Correlation Coefficient

Price of Tea (Rs) X	R_1	Price of Coffee (Rs) Y	R_2	$D^2=(R_1-R_2)^2$
75	4	120	4	0
88	7	134	5	4
95	8	150	8	0
70	3	115	8	0
60	2	110	2	0
80	5	140	6	1
81	6	142	7	1
50	1	100	1	0

$$r_s = 1-\frac{6\Sigma D^2}{N^3-N} = 1-\frac{6\times 6}{8^3-8}$$

$$= 1-0.071 = +\ 0.929$$

Equal Ranks: In some cases it may be found necessary to rank two or more individuals or entries as equal. In such a case it is customary to give each individual an average rank.Thus if two individuals are ranked equal to fifth place they are each given the rank $\frac{5+6}{2}$, that is 5.5. If three are ranked equal at fifth place they are given the ranks $\frac{5+6+7}{3}$ = 6. In other words, where two or more items are to be ranked equal, the rank assigned for purposes of calculating coefficient of correlation is the average of the ranks which these individuals would have got had they differed slightly form each other.

Where equal ranks are assigned to some entries, an adjustment in the above formula for calculating the rank coefficient of correlation is made. The adjustment consists of adding $\frac{1}{12}(m^3-m)$ to the value of ΣD^2, where m stands for the number of items whose ranks are common. If there are more than one such group of items with common rank, this value is added as many times the number of such groups. the formula can thus be written.

$$r_s = 1 - \frac{6\left\{\Sigma D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + ...\right\}}{N^3 - N}$$

Example: From the following data of the marks obtained by 8 students in the Accountancy and statistics papers compute Rank Coefficient of correlation.

Marks in Accountancy	15	20	28	12	40	60	20	80
Marks in Statistics	40	30	50	30	20	10	30	60

Sol. Computation of Rank Correlation

Marks in Accountancy (X)	Ranks assigned (R_1)	Marks in Statistics (Y)	Ranks assigned (R_2)	R_1–R =D	D^2
15	2	40	6	–4	16.00
20	3.5	30	4	–0.5	0.25
28	5	50	7	–2	4.00
12	1	30	4	–3	9.00
40	6	20	2	+4	16.00
60	7	10	1	+6	36.00
20	3.5	30	4	–0.5	0.25
80	8	60	8	0	0
					ΣD^2=81.50

$$r_s = 1 - \frac{6\left\{\Sigma D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + ...\right\}}{N^3 - N}$$

here = ΣD^2 = 81.5, N = 8

The item 20 is repeated 2 times in series X. so m = 2

In series Y th item 30 occurs 3 times and so $m = 3$

Substituting these values in the above formula

$$r_s = 1 - \frac{6\left\{81.5 + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(3^3 - 3) + ...\right\}}{8^3 - 8}$$

$$= \quad 1 - \frac{6(81.5 + 0.5 + 2)}{5.4}$$

$$= 1 - \frac{6 \times 84}{504} \;=\; 1 \;-\; \frac{504}{504} \;=\; 0$$

Merits and Limitations of the Rank method

Merits

(1) This method is simpler to understand and easier to apply, as compared to the Karl Pearson's method. The answer obtained by this method and the Karl Pearson's method will be the same provided no value is repeated i.e. all the items are different.

(2) Where the data are of qualitiative nature like honesty, efficiency, intelligence, etc., this method can be used with great advantage. For example, the workers of two factories can be ranked in order of efficiency and the degree of correlation established by applying this method.

(3) This is the only method that can be used where we are given actual data.

Limitations

(1) This method cannot be used for finding out correlation in a grouped frequency distribution.

(2) Where the number of items exceeds 30, the calculations become quite tedious and require a lot of time. Therefore, this method should not be applied where N exceeds 30 unless we are given the ranks and not the actual values of the variable.

When to use Correlation Coefficient?

The rank method has two principal uses.

(1) The initial data are in the form of ranks.

(2) If N is fairly small (say, no more than 25 or30) rank method is some times applied to interval daa as an approximation to the more time consuming r. This requires that the interval data be transferred to rank orders for both variables. If N is much in excess of 30, the labour required in ranking the scores becomes greater than is justified by the anticipated saving of time through the rank formula.

13.12. LET US SUM UP

In this lesson, we discussed in detail the concept of Covariance and Correlation.

13.13. EXAMINATION ORIENTED QUESTIONS

Q.1. Define Co-variance. Explain Co-variance with the help of example.

Q.2. Discuss Bivariate distribution of X & Y.

Q.3. Find Rank coefficient of correlation from the following data.

Mathematics	29	32	53	47	45	32	70	45	70	53
Hindi	56	60	72	48	72	35	67	67	75	31

Ans. Hint 0.4

Q.4. Find correlation by rank method:

X	50	55	65	50	55	60	50	65	70	75
Y	110	110	115	125	140	115	130	120	119	160

Ans. Hint 0.115

Q.5. Write in detail about the situation of Spearman rank correlation.

Q.6. Write in detail about degree of correlation.

Q.7. Explain the importance of correlation.

Q.8. Write merits and demerits of Karl Peason's correlation coefficient.

Q.9. Discuss Pal Pearson's coefficient of correlation with the help of formula.

13.14. SUGGESTED READINGS & REFERENCES

1. S.P. Gupta (2005), Statistical Methods, S. Chand and Sons, New Delhi
2. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas publishing House, New Delhi.
3. Monga, G.S. (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.

CURVE FITTING

LESSON NO. 14 **UNIT-IV**

STRUCTURE

14.1. OBJECTIVES

After going through this lesson, you would be able to understand.

1. Curve fitting and method of least squares
2. Regresion Analysis

14.2. INTRODUCTION

Curve fitting and Method of Least squares

Very often there exists a relation between available observations in respect of two variables. For instance, expenditure depends on income; weight and height of a person are interdependent; agricultural production depends on the amount of rainfall; price of commodity depends on demand of that commodity, etc. These relationships between variables can be expressed by means of some mathematical equation, representing a certain geometrical curve. The process of finding such a curve or its equation on the basis of a given set of observations is called curve fitting.

In other words, suppose (x_1,y_1) (x_2,y_2) (x_3,y_3) .... (x_n,y_n) are n paired values as two variables x and y. Again, suppose that x is independent variable and y is dependent variable; and y is function x. Then, the device of finding a functional relationship of the form $y = f(x)$ between two variables x and y on the basis of a finite set of pairs of observations (x_1,y_1), (x_2,y_2), (x_3,y_3) .... (x_n,y_n) on these variables is called curve fitting. Following are equations of some common types of curves :

Straight line: $y = a + bx$

Parabola: $y = a + bx + cx^2$

Cubic Curve: $y = a + bx + cx^2 + dx^3$

Polyomial of nth degree: $y = a + bx + cx^2 + ... + px^n$

Exponential Curve: $y = ab^x$

Modified expontential Curve $y = a+bc^x$

Logistic curve: $\frac{1}{y} = a+bc^x$

Gompertz Curve $\log y = a+bc^x$

Geometric values $y = ax^b$

The following rules help to select the appropriate equation to be fitted to a given set of observations:

(a) Plot the corresponding observations of x and y as points on a graph paper. If the pattern of points shows approximately a linear path, use the straight line curve.

(b) If $\log y$ when plotted against x shows a linear path, the exponential curve is used.

(c) If $\log y$ when plotted against $\log x$ shows a linear path, the geometric curve is used.

Curve fitting is very important tool both from the point of view of theoretical and practical Statistics. In theoretical Statistics, the study of regression and correlation can be regarded as fitting of linear curves (such as straight lines, planes, hyper-planes) to be given bivariate or multivariate frequency or probablity distributions. In practical Statistics, it enables us to get a close functional relation between the two variables x and y. In general, the relation is expressed by a polynomial, but other types of relationship, algebraic, exponential or logarithmic can also be fitted by using the principle of least squares.

Methods of least squares: It is a device for finding the equation of a specified type of curve, which best fits a given set of observations. In other words, to find the values for a set of unknown quantities so that a set of given equations may be satisfied as nearly as possible, when the number of equations is greater than the number of unknown and the equations are not strictly compatible with each other, Legendre suggested a method, known as Legendre principle or the method of least squares with the theory of probability and placed the method on a logical foundation.

Let there be a number of independent linear equations in n unknown x_1, x_2, x_3 ......x_n:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+a_{13}x_3+........+a_{1n}x_n \ = \ b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3 \ +......+a_{2n}x_n \ = \ b_2 \\ .... \\ a_{m1}x_1+ax_{m2}+a_{m3}x_3+.......+a_{mn}x_n \ = \ b_m \end{array}\right\} \quad \begin{array}{l} i.e. \ \sum_{i=1}^{n} a_i x_j = b_{i,} \\ \\ i = 1, 2,3, 4, ...m \end{array}$$

Where a's and b's are constants

If m = n, we have a unique set of value satisfying the given system of equation. If m > n i.e, if the number of equations is greater than the number of unknowns,. no such solution exists. Therefore, we try find those value of x_1, x_2, x_3, ...x_n which satisfy the given system of equations as nearly as possible.

The principle of least squares states that these values are those which make sum of squares (S) a minimum, where

$$S = \sum_{i=1}^{m}\left(a_{i1}x_1 + a_{12}x_2 + a_{i2}x_3 + ...a_{in}x_n - b_i\right)^2 = \sum_{i=1}^{m} E_i^2$$

$$E_i = a_{i1}x_1 + a_{12}x_2 + a_{i2}x_3 + ...a_{in}x_n - b_i$$

Where

The extreme values of the function $f = f(x_1, x_2, x_3 ......x_n)$ are given by $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} .... \frac{\partial f}{\partial x_n} = 0$, provided that the partial derivatives exist. Applying this concept, S will have amaximum or minimum, for those values of the unknown $x_i(j = 1, 2, 3, ...n)$ which satisfy the following n equations.

$$\frac{\partial f}{\partial x_j} = 0 \qquad j = 1, 2, 3, ... n$$

$$\text{i.e.} \quad \sum_{i=1}^{m} a_{i1}E_i = 0, \ \sum_{i=1}^{m} a_{i2}E_i = 0, \ \sum_{i=1}^{m} a_{i3}E_i = 0, \ .... \ \sum_{i=1}^{m} a_{in}E_i = 0.$$

These n equations are called the normal equations and can be solved like simultaneous equations for the n unknowns x_1, x_2, x_3 ......x_n . These solutions obtained from the normal equations make the second order Partial derivatives of S positive and so S is minimum.

FITTING STRAIGHT LINE

Let us consider an equation of straight line

$$y = a + bx$$

to be fitted to a given set of n pairs of observatios (x_1,y_1), (x_2,y_2), (x_3,y_3) .... (x_n,y_n). applying the method of least squares, the values of a and b are so determined as to minimise $\sum_{i=1}^{n}\left(Y_i - a - bx_i\right)^2$ This leads to two normal equations:

$$\Sigma y = a_n + b\Sigma x \text{ and } \Sigma xy = a\Sigma x + b\Sigma x^2$$

The values of n, Σx, Σy, Σx^2 and Σxy are substituted on the basis of the given data. We have then two equations involving a and b, solving which the values of a and b are obtained.

Fitting Parabola

Let us consider an equation of Parabola.

$$y = a + bx + cx^2$$

to be fitted to a given set of n pairs of observations (x_1y_1), (x_2,y_2), (x_3,y_3), ..., (x_n,y_n). Applying the method of least squares, the values of a, b, and c are so determined as to minimise $\sum_{i=1}^{n}\left(Y_i - a - bx_i - cx_i^2\right)^2$

This leads to three normal equations:

$$\Sigma y = an + b\Sigma x + c\Sigma x^2;$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$$

and $$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$$

The values of n, Σx, Σy, Σx^2, Σx^3, Σx^4, Σxy, and Σx^2y are substituted on the basis of the given data.

We have then three equations involving a, b, and c solving which the values of a, b and c are obtained.

FITTING EXPONENTIAL AND GEOMETRIC CURVE

To fit curves with equations of the form $y = ab^x$ and $y = ax^b$, the procedure is to take logarithms of both sides and then form normal equations.

Exponential curve: To fit exponential curve of the form $y = ab^x$, we take logarithms of both sides and get

$$\log y = \log a + x \log b$$

which can be written as: $Y = A + Bx$,

where, $Y = \log y$, $A = \log a$, and $B = \log b$.

The normal equations are

$$\Sigma Y = nA + B\Sigma x$$

and $$\Sigma \underline{x} Y = A\Sigma x + B\Sigma x^2$$

These equations are solved for A and B, and then taking antilog, we find the values of a and b.

Geometric curve: Similarly, the geometric curve of the form $y = ax^b$ may be written as

$$\log y = \log a + b \log x$$

$$\Rightarrow \qquad Y = A + bX$$

where, $Y = \log y$, $A = \log a$, and $X = \log x$

The normal equations are

$$\Sigma Y = nAb\Sigma X + \Sigma XY$$

$$= A\Sigma X + b\Sigma X^2$$

These equations are solved for A and b, and then a = anti log A is obtained.

Fitting a curve of the Form

$$y = ax + \frac{b}{x}$$

Let the set of n paired values be (x_1, y_1) (x_2, y_2),(x_3, y_3).....(x_n, y_n)

or (x_i, y_i), $i = 1, 2, 3, ...n$. Then, the sum of the squares of resudeals is

$$R = \sum_{i=1}^{n}\left(y_i - ax_i - \frac{b}{x_i}\right)^2$$

Differetiating R w.r.to a and be separately and setting these partial derivatives equal to zero, we get the normal equations.

$$\frac{\partial R}{\partial a} = 0 = -2\sum_{i=1}^{n}\left(y_i - ax_i - \frac{b}{x_i}\right)$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} b = a\sum_{i=1}^{n} x_i^2 + nb$$

$$\frac{\partial R}{\partial b} = 0 = -2\sum_{i=1}^{n}\left(\frac{1}{x_i}\right)\left(y_i - ax_i - \frac{b}{x_i}\right)$$

$$\sum_{i=1}^{n}\left(\frac{y_i}{x_i}\right) = \sum_{i=1}^{n} a + b\sum_{i=1}^{n}\left(\frac{1}{x_i^2}\right) = na + b\sum_{i=1}^{n}\left(\frac{1}{x_i^2}\right)$$

Substituting the values and solving the above equations, we get the values of a and b.

Fitting A curve of the form

$$y = ax^2 + \frac{b}{x}$$

Let the set of paired values be $(x_1 y_1)$ (x_2,y_2),(x_3,y_3) .... (x_n,y_n)

or (x_i, y_i), $i = 1, 2, 3, ... n$. Then the sum of the squares of residuals is

$$\mathrm{R} = \sum_{i=1}^{n}\left(Y_i - ax_i - \frac{b}{x_i}\right)^2$$

Differentiating R with respect to a and b respectively and setting these partial derivatives equals to zero, we get the normal equations.

$$\frac{\partial R}{\partial a} = 0 = -2\sum_{i=1}^{n} x_i^2\left(y_i - ax_i - \frac{b}{x_i}\right)$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i^4 + b\sum_{i=1}^{n} x_i$$

$$\frac{\partial R}{\partial b} = 0 = -2\sum_{i=1}^{n}\left(1/x_i\right)\left(Y_i - ax_i^2 - \frac{b}{x_i}\right)$$

$$\Rightarrow \qquad \sum_{i=1}^{n}\left(\frac{y_i}{x_i}\right) = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n}\left(\frac{1}{x_i^2}\right)$$

Substituting the values and solving the above equations, we get the values of a and b.

14.3. REGRESSION ANALYSIS

The word 'regression' is used to denote estimation or prediction of the average value of one variable for a specified value of the other variable and the regression analysis is mathematical measures of the average relationship between two or more variables, in terms of the original units of the data. The estimation is done by means of suitable equations, derived on the basis of available bivariate data. Such an equation is known as a regression equation and its geometrical representation is called a regression curve.

In regression analysis, there are two types of variables.

(a) 'Predicted' or 'dependent' or 'regressed' or 'explained' variable–the value of which can be predicted and

(b) 'Independent' or 'predictor' or 'regressor' or explanator' variable–on the basis of which the values of this variable are predicted.

In simple regression (or linear regression) the relationship between the variables is assumed to be linear.

The estimate of x (say, x') is obtained from an equation of the form.

$$x' - \bar{x} = bxy(y - \bar{y}) \qquad \text{...... (i)}$$

and the estimate of y (say y') from another equation (usually different from the former) of the form

$$y' - \bar{y} = b_{yx}(x - \bar{x}) \qquad \text{...... (ii)}$$

Equation (i) is known as regression equation of x on y, and equation (ii) as regression equation of y on x. The coefficient bxy in equation (i) is called regression Coefficient of x and y, and byx in equation (ii) is called regression coefficient of y on x.

The geometrical representation of linear regression equations (i) and (ii) are known as regression lines. These lines are ‘best fitting straight lines obtained by the method of least squares.

The regression Coefficients *bxy* and *byx* show some important properties, which can be summarised as follows :

(a) The product of the two regression coefficients, *bxy* and *byx*, is equal to the square of Correlation coefficient i.e. $bxy.byx = r^2$. In other words, the correlation coefficient is the geometric mean between the regression coefficients.

(b) r, bxy, byx, all have the same sign. If the correlation coefficient r is zero, the regression coefficients *bxy* and *byx* are also zero.

(c) One of the two regression coefficients must be numerically less than unity and the other must be numerically greater than unity or both will be unity or both will be numerically less than unity. Both coefficient cannot be greater than unity.

(d) The regression lines always intersect at the point $(\bar{x}, \bar{y})$.

(e) The regression Coefficients are independent of the change of origin, but not of scale.

(f) Arithmatic mean of the regression coefficient is greater than the correlation coefficient, provided the correlation coefficient is positive.

(g) The two regression equations are usually different. However, when $r = \pm 1$ they become identical, and in this case, there is an exact linear relationship between the variables. When $r = 0$, the regression equations reduce to $y - \bar{y}$ and $x = \bar{x}$, and neither y not x can be estimated from linear regression equations.

(h) The angle between the two regression lines depends on the correlation coefficient r. When $r = 0$, the two lines are perpendicular to each other; when $r = -1$ or $r = -1$ they coincide. As r increase numerically from 0 to1, the angle between the regression lines diminishes from 90º to 0º.

	x on y	y on x
Regression Equation	$x = a+bx$ $x-\bar{x} = bxy(y-\bar{y})$	$y = a+bx$ $y'-\bar{y} = \text{b}yx(x-\bar{x})$
Normal Equations	$\Sigma x = na+b\Sigma y$ $\Sigma xy = a\Sigma y+b\Sigma y^2$	$\Sigma y = na+b\Sigma x$ $\Sigma xy = \text{a}\Sigma x+\text{b}\Sigma x^2$
Regression coefficients	$bxy = r\frac{\sigma_x}{\sigma_y}$	$byx = r\frac{\sigma_y}{\sigma_x}$
When deviation are taken from mean	$r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma xy}{\Sigma y^2}$ $x = (x'-\bar{x})$, $y = y'-\bar{y}$	$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{\Sigma x^2}$
When deviations are taken from assumed Mean	$r\frac{\sigma_x}{\sigma_y} = \Sigma dx\Sigma dy - \frac{(\Sigma dx)(\Sigma dy)}{N}$ $\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}$ dx = (X–A); dy = (Y–A)	$r\frac{\sigma_y}{\sigma_x} = \Sigma dx\Sigma dy - \frac{(\Sigma dx)(\Sigma dy)}{N}$ $\Sigma dx^2 - \frac{(\Sigma dx)^2}{\text{N}}$
When the original values of X and Y series are taken standard Error or estimate coefficient of correlation.	$r\frac{\sigma_x}{\sigma_y} = \frac{XY - N\bar{X}\bar{Y}}{\Sigma y^2 - N(\bar{Y})^2}$ $\text{S}x = \sigma x\sqrt{1-r^2}$ $r = \sqrt{bxy \times byx}$	$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma XY - N\bar{X}\bar{Y}}{\Sigma X^2 - N(\bar{X})^2}$ $\text{S}y = \sigma_y\sqrt{(1-r^2)}$

Example 1. From the following data obtained the two regression equations.

X	6	2	10	4	8
Y	9	11	5	8	7

Sol. Obtaining Regression Equations.

X	Y	XY	X^2	Y^2
6	9	54	36	81
2	11	22	4	121
10	5	50	100	25
4	8	32	16	64
8	7	56	64	49
$\Sigma X=30$	$\Sigma Y=40$	$\Sigma XY=214$	$\Sigma X^2=220$	$\Sigma Y^2 = 340$

Regression equation Y on X: $Y = a+bX$

To determine the values of a and b the following two normal equations are to be solved.

$$\Sigma Y = Na + b\Sigma X$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

Substituting the valuess

$$40 = 5a + 30b \qquad \text{...(i)}$$

$$214 = 30a + 220b \qquad \text{...(ii)}$$

Multiplying equation (i) by 6

$$240 = 30a + 180b \qquad \text{...(iii)}$$

$$214 = 30a + 220b \qquad \text{...(iv)}$$

Deducting (iv) from (iii)

$$-40b = +26$$

$$b = -\frac{26}{40} = -0.65$$

Substituting the value of b in equation (i)

$$40 = 5a + 30(-0.65)$$

or $$5a = 40 + 19.5$$

$$= 59.5$$

or $$a = 11.9$$

Putting the values of a and b in the equation, the regression of Y on X is

$$Y = 11.9 - 0.65X$$

Regression line of X and Y: $X = a + bY$

and the two normal equations are

$$\Sigma X = Na + b\Sigma Y$$

$$\Sigma XY = a\Sigma Y + \Sigma Y^2$$

$$30 = 5a + 40b \qquad ...(i)$$

$$214 = 40a + 340b \qquad ...(ii)$$

Multiplying equation (i) by 8:

$$240 = 40a + 320b \qquad ...(iii)$$

$$214 = 40a + 340b \qquad ...(iv)$$

Deducting (iv) from (iii)

$$26 = -20b$$

$$b = -1.3$$

Substituting the value of b in eq(i):

$$30 = 5a + 40(-1.3)$$

$$5a = 30 + 52 = 82$$

$$\therefore \quad a = 16.4$$

Putting the values of a and b in the equation, the regression line of X on Y is

$$X = 16.4 - 1.34$$

Deviation taken from Arithmetic Means of X and Y

The above method of finding out regression equations is tedious. The calculations can very much be simplified if instead of dealing with the actual values of X and

Y, we take the deviation of X and Y series from their respective means. In such a case the two regression equations are written as follows.

(i) Regression Equation of X and Y

$$X-\overline{X} = r\frac{\sigma_x}{\sigma_y}(Y-\overline{Y})$$

$\overline{X}$ is the mean of X series

$\overline{Y}$ is the mean of Y series.

$r\frac{\sigma_x}{\sigma_y}$ is known as the regression coefficient of X on Y. The regression of X and Y. The regresion coefficient of X on Y is denoted by the symbol *bxy* or b_1. It measures the change in X corresponding to a unit change in Y. When deviations are taken from the means of X and Y, the regression coefficient of X on Y is obtained as follows:

$$bxy \quad \text{or} \quad r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma xy}{\Sigma y^2}$$

Instead of finding out the value of correlation coefficient, σx, σy etc. we can find the value of regression coefficient by calculating Σxy and Σy^2 and dividing the former by the latter.

(ii) Regression Equation of Y on X

$$Y-\overline{Y} = r\frac{\sigma_y}{\sigma_x}(X-\overline{X})$$

$r\frac{\sigma_y}{\sigma_x}$ is the regression coefficient of Y on X. It is denoted by *byx* or b_2. It measures the change in Y corresponding to a unit change in X. When deviations are taken from actual means the regression coefficient of Y on X can be obtained as follows.

$$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{\Sigma x^2}$$

It should be noted that the under root of the product of two regression coefficients gives us the value of correlation coefficient.

Symbolically:

$$r = \sqrt{bxy \times byx}$$

Proof: $bxy = r\frac{\sigma_x}{\sigma_y}$ and $byx = r\frac{\sigma_y}{\sigma_x}$

$$bxy \times byx = r\frac{\sigma_x}{\sigma_y} \times r\frac{\sigma_y}{\sigma_x} = r^2 \qquad \therefore r = \sqrt{bxy \times byx}$$

Example 2: From the data of illustration 1, calculate the regression equations taking deviation of items from the mean of X and Y series.

Sol. Calculation of Regression Equations.

X	$(X-\overline{X})=x$	x^2	Y	$(Y-\overline{Y})=y$	y^2	xy
6	0	0	9	1	1	0
2	–4	16	11	3	9	–12
10	4	16	5	–3	9	–12
4	–2	4	8	0	0	0
8	2	4	7	–1	1	–2
ΣX=30	$\Sigma x=0$	$\Sigma x^2=40$	ΣY=40	$\Sigma y=0$	$\Sigma y^2=20$	$\Sigma xy=-26$

Regression Equation of X on Y: $X-\overline{X} = r\frac{\sigma_x}{\sigma_y}(Y-\overline{Y})$

$$r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma xy}{\Sigma y^2} = \frac{-26}{20} = -1.3$$

$$\overline{X} = \frac{30}{5} = 6; \quad \overline{Y} = \frac{40}{5} = 8$$

Hence (X–6) = (–1.3) (Y – 8)

$$= -1.3Y + 10.4$$

$$X = -1.3Y + 16.4$$

or $\quad X = 16.4 - 1.3Y$

Regression Equation of Y on X: $Y - \overline{Y} = r\frac{\sigma_y}{\sigma_x}(X - \overline{X})$

$$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{\Sigma x^2} = \frac{-26}{40} = 0.65$$

$$Y - 8 = -\ 0.65(X - 6)$$

$$= -\ 0.65X + 3.9$$

$$Y = -\ 0.65X + 11.9$$

or $\quad Y = 11.9 - 0.65X$

Thus we find that the answer is the same as obtained earlier. However, the calculations are very much simplified without the use of normal equations.

Deviations taken from Assumed Means

When actual means of X and Y variables are in fractions, the calculations can be simplified by taking the deviations from the assumed means. When deviations are taken from assumed means, the entire procedure of finding regression equations remains the same. The only difference is that instead of taking deviations from actual means, we take the deviations from assumed means. The two regression equations are

$$X - \overline{X} = r\frac{\sigma_x}{\sigma_y}(Y - \overline{Y})$$

The value of $r\frac{\sigma_x}{\sigma_y}$ will now be obtained as follow:

$$r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma dxdy - \frac{\Sigma dx \times \Sigma dy}{N}}{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}}$$

dx = (X–A) and dy = (Y–A)

Similarly the regression equation of Y on X is

$$Y - \overline{Y} = r\frac{\sigma_y}{\sigma_x}(X - \overline{X})$$

$$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma dxdy - \frac{\Sigma dx \times \Sigma dy}{N}}{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}}$$

It should be noted that in both the cases, the numerator is the same; the only difference is in the denominator.

Example: From the data of illustration 1 obtain regression equations taking deviations from 5 in case of X and 7 in case of Y.

Solution: Calculation of Regression Equations

X	(X–5)=dx	dx^2	Y	(Y–7)=dy	dy^2	dxdy
6	1	1	9	2	4	2
2	–3	9	11	4	16	–12
10	+5	25	5	–2	4	–10
4	–1	1	8	1	1	–1
8	+3	9	7	0	0	0
ΣX=30	Σdx=+5	Σdx^2=45	ΣY=40	Σdy=5	Σdy^2=25	$\Sigma dxdy$=–21

Regression Equation of X and Y: $(X-\overline{X}) = bxy(Y-\overline{Y})$

$$bxy = \frac{\Sigma dxdy - \frac{\Sigma dx \times \Sigma dy}{N}}{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}}$$

$$= \frac{-21 - \frac{(5)(5)}{5}}{25 - \frac{(5)^2}{5}}$$

$$= \frac{-21-5}{20}$$

$$= -1.3$$

$$\overline{X} = \frac{30}{5} = 6, \quad \overline{Y} = \frac{40}{5} = 8$$

So the regression equation is

$$(X-6) = -1.3(Y-8)$$

$$= (1.3Y+10.4)$$

or $\quad X = 16.4 - 1.3Y$

Regression Equation of Y on X: $Y-\overline{Y} = byx(X-\overline{X})$

$$byx = \frac{\Sigma dxdy - \frac{\Sigma dx \times \Sigma dy}{N}}{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}} = \frac{-21 - \frac{5\times 5}{5}}{45 - \frac{5^2}{5}} = \frac{-26}{40} = 0.65$$

So the regression equation is

$$Y-8 = 0.65\ (X - 6)$$

$$= -\ 0.65X + 3.9$$

or $\quad Y = 11.9 - 0.65X$

14.4. LET US SUM UP

In this lesson, we discussed the concept of curve fitting in detail.

14.5. SUGGESTED READINGS

1. S.P. Gupta (2005), Statistical Methods, S. Chand and Sons, New Delhi
2. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas publishing House, New Delhi.
3. Monga, G.S. (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.

USES IN ECONOMICS [χ^2 TEST AND GOODNESS OF FIT]

LESSON NO. 15 **UNIT-IV**

15.1. OBJECTIVES

After going through this lesson you would be able to understand

— Goodness of Fit

— χ^2 Test and Goodness of Fit

— Uses of χ^2 Test

15.2. INTRODUCTION OF GOODNESS OF FIT

The goodness of fit of a statistical model describes how well it fits a set of observations. Measures of goodness of fit summarize the discrepancy between observed values and the values expected under the model in question. Such measures can be used in statistical hypothesis testing, e.g. to test for normality of residuals, to test whether two samples are drawn from identical distributions (see kolmogorov-Smirnov test), or whether outcome frequencies follow a specified distribution (see Pearson's chisquared test). In the Analysis of Variance, one of the components into which the variance is partitioned may be a lack-of-fit sum of squares.

Fit of Distributions: In assessing whether a given distribution is suited to a data-set, the following tests and their underlying measures of fit can be used:

* Kolmogorov – Smirnov test;
* Cramer – VonMiser Criterion;
* Anderson – Darling test;
* Chi square test;
* Akaike information criterion;
* Hosmer– Lemeshow test.

15.2.1. χ^2 Test and Goodness of Fit

In all these tests, e.g., Z-test, t-test and F-test, we have to make assumptions about the population values or parameters, therefore such tests are called parametric tests.

Parametric tests are not applicable in situations where it is not possible to make any dependable assumptions about the form of the parent distribution from which samples have been drawn. To overcome such situations, various tests have been evolved which do not require any assumptions about the parameters. Such tests are distribution free tests and are called Non-Parametric Tests.

Chi-square test is a non-Parametric test which was first developed by Hamlet and later Karl Pearson derived it independently and applied it as a Test of Goodness of fit. It described the magnitude of difference between observed frequencies and the frequencies expected under certain assumptions, and therefore, with the help of such tests, it is possible to find out whether such differences are significant or are insignificant and could have arisen due to fluctiations of sampling.

The chi-square test involves the following steps:

Step 1. Calculate the expected frequencies, let us denote them by E.

Step 2. Find out the difference between observed frequencies (denoted by O) and the expected frequencies i.e. (O–E)

Step 3. Find out $\frac{(O-E)^2}{E}$

Step 4. Do a summation of all the values of $\frac{(O-E)^2}{E}$ and this will be the value of χ^2, i.e. $\chi^2 = \Sigma\left[\frac{(O-E)^2}{E}\right]$

Step 5. Compare the calculated value of χ^2 with the tabulated value, χ^2 for the given degrees of freedom and at the desired level of significance.

Step 6. If the calculated value of χ^2 is greater than the tabulated value, the difference between observed and expected values is significant. If the calculated value is less than the tabulated value, the differences between the observed and expected frequencies is not significant and could have arisen due to fluctuations of sampling.

Degrees of Freedom: The term degrees of freedom refers to the number of independent constraints in a set of data. In case of association table, the degrees of freedom are calculated by the formula:

$$v = (c-1)\ (r-1)$$

Where v stands for the degrees of freedom, 'c' for the number of columns and 'r' for the number of rows.

Chi-Square Test for Association of Attributes

The word association is used to indicate the degree of relationship between attributes (the corresponding word for variables is correlation). Let us consider the following (2×2) Contigency table.

Attributes A

		A	α	Total
Attribute B	B	(AB) = a	(αB) = b	(B) = a+b
	β	(Aβ) = c	(αβ) = d	(β) = c+d
	Total	(A) = a+c	(α) = b+d	N

For a (2×2) contingency table, a measure of association is given by Yule's Coefficient of association as

$$Q = \frac{(AB)(\alpha\beta)-(A\beta)(B\alpha)}{(AB)(\alpha\beta)+(A\beta)(B\alpha)}$$

$$= \frac{ad-bc}{ad+bc}$$

Simplified Formula for (2×2) Table

Let us consider a contingency table having 2 rows and 2 columns with the four cell frequencies *a*, *b*, *c*, *d* as follows.

			Total
	a	b	R_1
	c	d	R_1
Total	C_1	C_2	N

$R_1 = a + b$, $R_2 = c + d$, $C_1 = a + c$, $C_2 = b + d$

$$N = a + b + c + d$$

$$= R_1 + R_2 = C_1 + C_2$$

In this case, Chi-square statistic is given by

$$X^2 = \frac{N(ad - bc)^2}{R_1 R_2 C_1 C_2}$$

with degrees of freedom = (2–1) × (2–1) = 1

15.2.2. USES OF χ^2 TEST

The χ^2 test is one of the most popular statistical inference procedures today. It is applicable to a very large number of problems in practice which can be summed up under the following heads :

(1) χ^2 test as a test of independence – With the help of χ^2, we can find out whether two or more attributes are associated or not. Suppose we have N observations classified according to some attributes.

We may ask whether the attributes are related or independent. Thus, we can find out whether quinine is effective in Controlling fever or not. In order to test whether or not the attributes are associated, we take the null hypothesis that there is no association in the attributes under study. In other words, the two attributes are independent. If the calculated value of χ^2 is less than the table value at a certain level of significance (generally 5% level), we say that the results of the experiment provide no evidence for doubting the hypothesis or, in other words, the hypothesis that the attributes are not associated holds good. On the other hand, if the calculated value of χ^2 is greater than the table value at a certain level of significance, we say that the results of the experiment do not support the hypothesis or, in other words, the attributes are associated. It should be noted that χ^2 is not measure of the degree or form of relationship, it only tells us whether two principles of classification are or are not significantly related, without reference to any assumptions concerning the form of relationship.

(2) χ^2 test as a test of goodness of fit – χ^2 test is very popularly known as test of goodness of fit for the reason that it enables us to ascertain how well the theoretical distributions such as Binomial, Poisson, Normal etc. fit empirical distributions; i.e. those obtained from sample data. When an ideal frequency curve whether normal or some other type is fitted to the data, we are interested to find out how well this curve fits with the observed facts. A test of the concordance (goodness of fit) of the two can be made just by inspection, but such a test is obviously inadequate.

(3) **χ^2 test as a test of homogeneity:** The χ^2 test of homogeneity is an extension of the chi-square test of independence. Tests of homogeneity are designed to determine whether two or more independent random samples are drawn from the same population or from different populations. Instead of one sample as we use with independence problem, we shall now have two or more samples.

Example 1. In an anti - malarial compaign in a certain area, quinine was administered to 812 persons out of a total population of 3,248. The number of fever cases is shown below.

Treatment	Fever	No Fever	Total
Quinine	20	792	812
No Quinine	220	2,216	2,436
Total	240	3,008	3,248

Discuss the usefulness of quinine in checking malaria.

Sol. Let us take the hypothesis that quinine is not effective in checking malaria. Applying χ^2 test.

$$\text{Expectation of (AB)} = \frac{(A)\times(B)}{N}$$

$$= \frac{240\times 812}{3248} = 60$$

E i.e. expected frequency corresponding to first row and first column is 60.

The table of expected frequencies shall be:

60	752	812
180	2,256	2,436
240	3,008	3,248

O	E	$(O–E)^2$	$\frac{(O-E)^2}{E}$
20	60	1,600	26.667
220	180	1,600	8.889
792	752	1,600	2.128
2,216	2,256	1,600	0.709
			$\Sigma\frac{(O-E)^2}{E} = 38.393$

$$\chi^2 = \Sigma\frac{(O-E)^2}{E} = 38.393$$

$$v = (r–1)\ (c–1)$$

$$= (2–1)\ (2–1) = 1$$

For $v = 1, \quad \chi^2_{0.05} = 3.84$

The calculated value of χ^2 is greater than the table value. The hypothesis is rejected. Hence quinine is useful in checking malaria.

Example 2. In an experiment of immunization of cattle from tuberculosis, the following results were obtained.

	Affected	Not affected
Inoculated	12	26
Not inoculated	16	6

Calculate χ^2 and discuss the effect of vaccine in controlling susceptibility to tuberculosis (5%) value of χ^2 for one degreee of freedom 3.84.

Sol. Let us take the hypothesis that the vaccine is not effective in controlling susceptibility to tuberculosis. Applying χ^2 test.

$$\text{Expectation of (AB)} = \frac{(A)\times(B)}{N} = \frac{38\times 28}{60} = 17.7$$

The table of expected frequencies will be as follows.

17.7	20.3	38
10.3	11.7	22
28	32	60

Since one of the observed frequencies is less than 10, we will use Yates' Correction and then apply χ^2 test:

O	E	$(O–E)^2$	$\frac{(O-E)^2}{E}$
12.5	17.7	27.04	1.528
15.5	10.3	27.04	2.625
25.5	20.3	27.04	1.332
6.5	11.7	27.04	2.311
			$\Sigma\frac{(O-E)^2}{E} = 7.792$

$$\chi^2 = \Sigma\frac{(O-E)^2}{E} = 7.792$$

$$v = (r–1)(c–1) = (2–1)(2–1) = 1$$

$$v = 1, \ X^2_{0.05} = 3.84$$

Since the calculated value of χ^2 is greater than the table value, the hypothesis is not true. We therefore conclude that vaccine is effective in controlling susceptibility to tuberculosis.

Example 3. From the data given below about the treatment of 250 patients suffering from a disease, state whether the new treatment is superior to the conventional treatment:

Treatment	No. of Patients		
	Favourable	Not favourable	Total
New	140	30	170
Conventional	60	20	80
Total	200	50	250

(Given for degree of freedom = 1, chi-square 5 percent = 3.84)

Sol. Let us take the hypothesis that there is no significant difference between the new and conventional treatment. Applying χ^2 test.

$$\text{Expectation of (AB)} = \frac{200 \times 170}{250} = 136$$

The table of expected frequencies shall be as follows :

136	34	170
64	16	80
200	50	250

O	E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
140	136	16	0.118
60	64	16	0.250
30	34	16	0.471
20	16	16	1.000
			$\Sigma \frac{(O-E)^2}{E} =$ 1.839

$$\chi^2 = \Sigma\frac{(O-E)^2}{E} = 1.839$$

$$v = (r-1)(c-1) = (2-1)(2-1) = 1$$

$$v = 1,\ \chi^2_{0.05} = 3.84$$

The calculated value of χ^2 is less than the table value. The hypothesis is accepted. Hence there is no significant difference between the new and conventional treatment.

15.3. LET US SUM UP

In this lesson, we discussed χ^2 as a goodness of fit measure.

15.4. EXAMINATION ORIENTED QUESTIONS

Q.1. Explain Goodness of fit with the help of an example.

Q.2. State Chi-square test with help of an example.

Q.3. Discuss uses of χ^2 Test.

15.5. SUGGESTED READINGS

1. S.P. Gupta (2005), Statistical Methods, S. Chand and Sons, New Delhi
2. C.B. Gupta and Vijay Gupta (1995), An Introduction to Statistical Methods, Vikas publishing House, New Delhi.
3. Monga, G.S. (1972), Mathematics and Statistics for Economists, Vikas Publishing House, New Delhi.

SUPPLY CURVES

LESSON NO. 16 **UNIT-IV**

STRUCTURE

16.1. OBJECTIVES

After going through this lesson, you shall be able to understand:

— Supply and supply function

— Law of supply and supply curve

— Shift in supply: Increase and Decrease in supply

16.2. INTRODUCTION

Supply means the amount offered for sale at a given price. Supply should be carefully distinguised from stock. Stock is the total volume of a commodity which can be brought into the market for sale at a short notice and supply means the quantity which is actually brought in the market. In this lesson, we shall discuss about this concept in detail.

16.3. THE MEANING OF SUPPLY

We have discussed in earlier class that demand is a schedule of the quantities of good that will be purchased at various prices. Similarly the supply refers to the schedule of the quantities of a good that the firms are able and willing to offer for sale at various prices. How much of a commodity the firms are able to produce depends on the resources available to them and the technology they employ for producing a commodity. How much of a commodity the firms will be willing to offer for sale depends on the profits they expect to make by producing and selling the commodity. Profits in turn depends on the price of the commodity on the one hand and unit cost of production on the other.

SUPPLY FUNCTION

The quantity of a commodity that firms will be able and willing to offer for sale in the market depends on several factors. The important factors determining supply of a commodity are :

1. The price of the commodity.
2. The prices of inputs (i.e. resources) used for the production of the commodity.
3. The state of technology.
4. The number of firms producing and selling the commodity.
5. The prices of related goods produced.
6. Future expectations regarding prices.

We will explain these factors determining supply of a commodity in detail in a later section. However, it may be noted that out of the above determinants of supply, the own price of the commodity, the prices of inputs (i.e. resources) used to produce the commodity, and the technology are three important factors and therefore the supply function of a commodity is often written taking these factors as independent variables. Thus supply of a commodity is written as

$$Q^S_X = S\ (P_X,\ F_1,\ F_2 \ldots\ldots\ldots,\ F_m)$$

Where Q^S_X is the quantity supplied of the commodity X, P_X is it own price, F_1, $F_2 \ldots$ F_m are the prices of inputs used to produce the commodity X and the state of technology determines the form of supply function S. It must be noted that the form of the function refers to the precise quantitative relation between the independent variables such as own price of the commodity X and prices of factors such as F_1, F_2 etc.

16.4. THE RELATION BETWEEN PRICE AND QUANTITY SUPPLIED: LAW OF SUPPLY

Supply of a commodity is functionally related to its price. The law of supply relates to this functional relationship between price of a commodity and its quantity supplied. In contrast to the inverse relationship between the quantity demanded and the changes in price, the quantity supplied of commodity generally varies directly with price. That is, the higher the price, the larger is the quantity supplied of a commodity.

The supply schedule and the upward-sloping supply curve reflect the law of supply. According to the law of supply, when the price of a commodity rises, the quantity supplied of it in the market increases, and when the price of a commodity falls, its quantity demanded decreases, other factors determining supply remaining the same. Thus, according to the law of supply, the quantity supplied of a commodity is directly or positively related to price. It is due to this positive relationship between price of a commodity and its quantity supplied that the supply curve of a commodity slopes upward to right as can be seen from supply curve SS in figure (a).

Table: Supply Schedule of Wheat

Price Per Quintal (Rs.)	Quantity Supplied (in quintals)
500	100
510	150
520	200
530	225
540	250
550	275

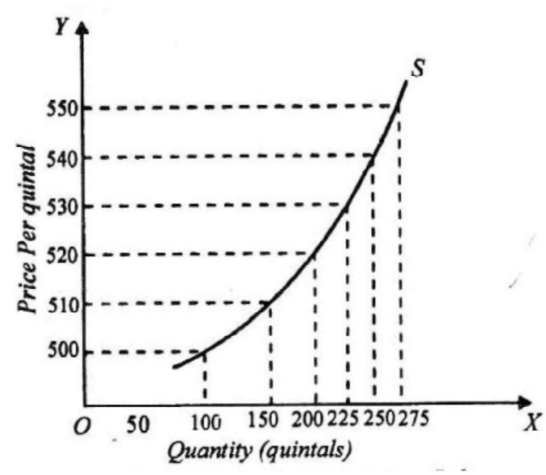


Fig. (a): Supply Curve Showing Direct Relationship Between Price and Quantity Supplied.

When price of wheat rises from Rs. 520 to Rs. 530 per quintal, the quantity supplied of wheat in the market increases from 200 quintals to 225 quintals per period.

16.5. SHIFTS IN SUPPLY : INCREASE AND DECREASE IN SUPPLY

As stated above, the supply of a commodity in Economics means the entire schedule or curve depicting the relationship between price and quantity supplied of the commodity, given the other factors influencing supply.

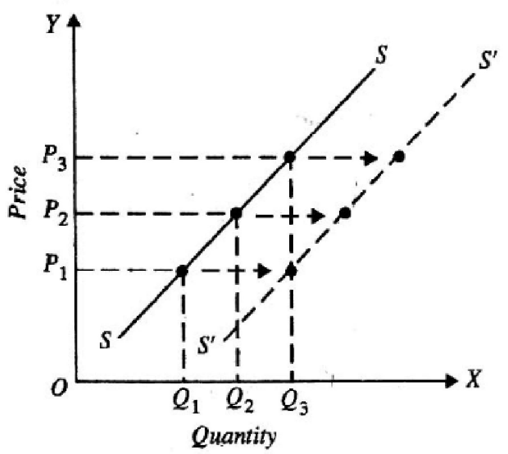


Fig. (b): Increase in supply causing a rightward shift in the supply curve.

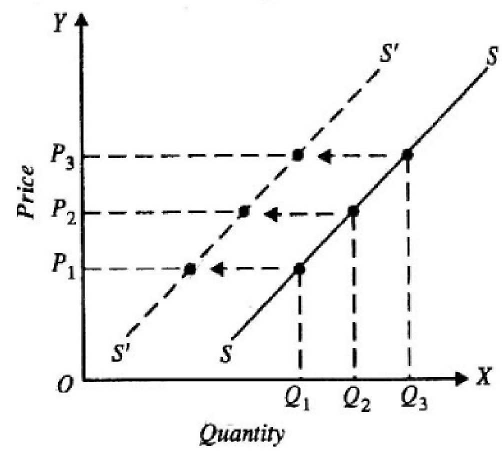


Fig.(c): Decrease in Supply Causing a leftward shift in the supply curve.

These other factors are the state of technology, prices of inputs (resources), prices of other related commodities, etc., which are assumed constant when the relationship between price and quantity supplied of a commodity is examined. It is the change in the factors other than price that cause a shift in the supply curve. For example, when prices of inputs such as labour and raw materials used for the production of a commodity decline, this will result in lowering the cost of production which will induce the producers to produce and make available a greater quantity of the commodity in the market at each price. This increase in supply of a commodity due to the reduction in prices of inputs will cause the entire supply curve to shift to the right as shown in Figure (b) where the supply curve shifts from SS to S″S″. As shown by arrow marks, at price P_1, P_2 and P_3 quantity supplied increases when supply increases, causing a rightward shift in the supply curve. Similarly, progress in technology used for production of a commodity which increases productivity and reduces cost per unit will also cause the supply curve to shift to the right.

On the other hand, decrease in supply means the reduction in quantity supplied at each price of the commodity as shown in Figure (c) where as a result of decrease in supply the supply curve shifts to the left from SS to S″S″. As shown by the arrow marks, at each price such as P_1, P_2, P_3, the quantity supplied on the supply curve S″S″ has declined as compared to the supply curve SS. The decrease in supply occurs when the rise in prices of factors (inputs) used for the production of a commodity produced leads to higher cost per unit of output which causes a reduction in quantity supplied at each price. Similarly, the imposition of an excise duty or sales tax on a commodity means that each quantity will now be supplied at a higher price than before so as to cover the excise duty or sales tax per unit. This implies that quantity supplied of the commodity at each price will decrease as shown by the shift of the supply curve to the left.

Another important factor causing a decrease in supply of a commodity is the rise in prices of other commodities using the same factors. For example if the price of wheat rises sharply, it will become more profitable for the farmers to grow it. This will induce the farmers to reduce the cultivated area under other crops, say sugarcane, and devote it to the production of wheat. This will lead to the decrease in supply of sugarcane whose supply curve will shift to the left.

Further, agricultural production in India greatly depends on the rainfall due to monsoons. If monsoons come in time and rainfall is adequate, there are bumper crops, the supply of agricultural products increases. However, in a year when monsoons are untimely or inadequate, there is a sharp drop in agricultural output which causes a shift in the supply curve of agricultural output to the left.

We thus see that there are several factors other than prices which determine the supply of a commodity and any change in these other factors will cause a shift in the entire supply curve.

16.6. FACTORS DETERMINING SUPPLY

It is clear from the supply schedule and the supply curve given above that the quantity supplied varies directly with price of the product. A supply schedule and supply curve shows that the supply of a product is function of its price. However, the supply depends not only on the price of a product but on several other factors too. It should be remembered that in economic theory whereas the effect of changes in price of a product on the quantity supplied of it is depicted and explained by movement along a given supply schedule or curve, the effect of other factors is represented by the changes or shifts of the entire supply schedule or supply curve. While making a supply schedule or drawing a supply curve we assume that these factors remain the same. Thus when these other factors change, they cause a shift in the entire supply curve. The factors other than price which determine supply are the following:

(a) **Production Technology:** The change in technology significantly affects the supply function by altering the cost of production. If there occurs an improvement in production technology used by the firm, its production efficiency increases which reduces the unit cost of production and consequently the firms would supply more than before at the given price. That is, the supply would increase implying thereby that the entire supply curve would shift to the right.

(b) **Price of Inputs:** Changes in prices of factors or inputs used in production also cause a change in cost of production and consequently bring about a change in supply. For example, if either wages of labour increase or prices of raw materials and fuel go up, the unit cost of production will rise. With higher unit cost of production, it will be profitable to produce less and therefore less would be supplied or offered for sale than before at various given prices. This implies that supply curve would shift to the left.

(c) **Prices of related Products:** When we draw a supply curve we assume that the prices of other products remain unchanged.

(d) **Number of Producers (or firms):** If the number of firms producing a product increases, the market supply of the product will increase causing a rightward shift in the supply curve.

(e) **Future Price Expectations:** The supply of a commodity in the market at any time is also determined by seller's expectations of future prices.

(f) **Taxes and Subsidies:** Taxes and subsidies also influence the supply of a product. If an excise duty or sales tax is levied on a product, the firms will supply the same amount of it at a higher price or less quantity of it at the same price.

It follows from above that technology, prices of factors and products, expectations regarding future prices and taxes and subsidies are the important determinants of supply which cause rightward or leftward shift in the whole supply curve.

16.7. LET US SUM UP

In this lesson, we discussed the concept of supply curve in detail.

16.8. EXAMINATION ORIENTED QUESTIONS

Q.1. What in supply? Explain Law of Supply.

Q.2. What are the factors which determine supply of a commodity?

Q.3. Distinguish between movement along a supply curve and shift in the supply curve. What are the factors which causes shift in the suppply curve?

16.9. SUGGESTED READINGS

1. Mithani, D.M.: Micro Economics, Himalaya Publishing House, Mumbai.
2. Koutsoyiannis, A: Modern Micro Economics, Macmillan Publisher Ltd. New Delhi.
3. Ahuja, H.L.: Advanced Economic Theory— Micro Economic Analysis, S. Chand & Co. New Delhi.

Glossary of Statistical Terms

Assumed Mean: An approximate value in order to simplify calculation.

Attribute: A characteristic that is qualitative in nature. It cannot be measured.

Average is a measure of central tendency of the distribution of the values of a certain variable, e.g. the Arithmetic mean, median and mode are averages which measure the central value in different ways.

Bimodal Distribution: A distribution which has two mode values.

Bivariate Distribution: Frequency distribution of two variables.

Census Method: A method of data Collection, which requires that observations are taken on all the individuals in a populaion.

Class Frequency: Number of observations in a class.

Class Interval: Difference between the upper and the lower class limits.

Class Marks: Class mid point.

Class Mid-point: Middle value of a class. It is the representative value of different observations in a class. It is equal to

$$\frac{\text{Upper class limit+lower class limit}}{2}$$

Classification: Arranging or organising similar things into groups or classes.

Constant: A constant is also a quantity used to describe an attribute, but it will not change during calculation or investigation.

Dispersion: Dispersion is a measure of scatter or dispersion of values about the central value.

Continuous variable: A quantitative variable that can take any numerical value.

Data: A (often large) set of numbers systematically arranged for conveying specific information on a subject for better understanidng or decision making.

Decile: A partition value that divides the data into ten equal parts.

Discrete variable: A quantitative variable that takes only certain values. It changes from one value to another by finite "jumps". The intermediate values between two adjacent values are not taken by the variable.

Dispersion

Enumerator: A person who collects the data.

Estimator is the method of obtaining an estimate of the parameter values from sample data. For example, the sample arithemetic mean is an estimator of the population arithematic mean; and the sample variance is an estimator of the population variance etc.

Estimate is the numerical value of the estimator that we obtain from a given sample.

Frequency: The number of times an observation occurs in raw data. In a frequency distribution it means the number of observation in a class.

Frequency curve: The graph of a frequency distribution in which class frequencies on Y-axis are plotted against the values of class marks on X-axis.

Frequency Distribution: A classification of a quantitative variable that shows how different values of the variable are distributed in different classes along with their corresponding class frequencies.

Multimodal Distribution. The distribution that has more than two modes.

Observation: A unit of raw data.

Partition value: The median, quartiles, deciles and percentiles are called partition values. The median is that value of the variable, which divides the entire set of values in two equal halves; the quartiles are the values (Q_1, Q_2, Q_3) which divide them in four equal parts, deciles divide then in ten equal parts and percentiles in hundred equal parts.

Percentiles: A value which divides the data into hundred equal parts. So there are 99 percentiles in the data.

Policy: The measure to solve an economic problem.

Population: Population means all the individual units for whom the information has to be sought.

Qualitative classification: Classification based on quality. For example classification of people according to gender, marital status etc.

Qualitative Facts: Information or data expressed in terms of qualitites.

Quantitative Facts: Information or data expressed in number.

Random sampling: It is a method of sampling in which the representative set of informants is selected in such a way that every individual is given equal chance of being selected as an informant.

Range Difference between the maximum and the minimum values of a variable.

Relative Frequency of a class as proportion or percentage of total frequency.

Sampling Error: It is the numerical difference between the estimate and the true value of the parameter.

Time Series: Data arranged in chronological order or two variable data where one of the variables is time.

Univariate Distribution: The frequency distribution of one variable.

Variable: A variable is a quantity used to measure an "Attribute" (such as height, weight, number etc.) of some thing or some persons, which can take different values in different situations.

Weighted Average: The weighted average is calculated by providing the different data points with different weights.

Model Test Paper - 1
ECONOMICS (Semester - IV)

Quantitative Methods in Economics

Time Allowed: 3 Hrs. ***Maximum Marks: 80***

Section-A

Note: Attempt any one question from each unit. Each question carries 6 marks.
(6×4 = 24)

Unit - I

Q.1. Discuss the Role of Mathematics in Economics.

Q.2. Define real number. State its properties.

Unit -II

Q.3. Write a short note on types of functions with examples.

Q.4. Prove that

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{d}{dx}(u) - u\frac{d}{dx}(v)}{v^2}$$

where u and v are both derivable function at x and $v \neq 0$.

Unit - III

Q.5. Define Arithematic Mean. State one notable property of Arithematic Mean.

Q.6. Define dispersion. What are the main measures of dispersion

Unit - VI

Q.7. State χ^2 Test and Goodness of fit.

Q.8. Discuss Rank Correction Coefficient.

Section B

Note: Attempt any one question from each unit. Each question carries 14 marks. (4×14 = 56)

Unit - I

Q.9. A consumer wants to consume two goods. The prices of the two goods are Rs. 4 and Rs. 5 respectively. The consumer's income is Rs. 20.

(i) Write down the equation of the budget line.

(ii) How much of good 1 can the consumer consume if she spends her entire income on that good?

(iii) How much of good 2 can consumer consume if she spends her entire income on that good?

(iv) What is the slope of the budget line?
(2+5+5+2=14)

Q.10. Define Meaning of Equation. Discuss the various types of Equations.

Unit - II

Q.11. Derive expression for Price elasticity of demand for the demand function $q = f(P)$

Q.12. The demand for a certain product is represented by the equation.

$P = 100 - 5q$

(i) Find the marginal revenue for any output

(ii) What is MR, when $q = 0$ and $q = 4$?

Unit - III

Q.13. Calculate the median of x variable given below:

Item	3	4	5	6	7	8
Frequency	6	9	11	14	23	10

[Hint: **Ans.** Median = 6]

Q.14. Calculate Standard deviation of the following Series.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of students	5	10	20	40	30	20	10	4

[Hint: **Ans.** S.D = 15.69]

Unit - IV

Q.15. Explain the Concept of curve fitting.

Q.16. Calculation the coefficient of correlation from the following data.

X	9	8	7	6	5	4	3	2	1
Y	15	16	14	13	11	12	10	8	9

[Hint: $r = 0.95$]

Model Test Paper - 2

ECONOMICS (Semester - IV)

(Quantitative Methods in Economics)

Time Allowed: 3 Hrs. ***Maximum Marks: 80***

Section-A

Note: Attempt any one question from each unit. Each question carries 6 marks. (6×4 = 24)

Unit - I

Q.1. Define Linear equation. Find the slope of the line $3x + 2y + 6 = 0$

Q.2. Define Budget line. Write down Equation of the budget line. (3, 3)

Unit - II

Q.3. Define Imaginary Numbers, Integers, Whole numbers.

Q.4. What are the types of Equations? (2, 2, 2)

Unit - III

Q.5. The mean height of 25 male workers in a factory is 61 cms and the mean height of 35 female workers in the same factory is 58 cms. Find the combined mean height of 60 workers in the factory.

Q.6. Define Median and Obtain the value of Median from the following data:
391, 384, 591, 407, 672, 522, 777, 753, 2488.

Unit - IV

Q.7. State the meaning of Correlation. What are the types of correlation?

Q.8. Write down the importance of Karl Pearson's Correlation Coefficient.

SECTION B

Note: Attempt any one question from each unit each question carries 14 marks. (4×14 = 56)

Unit - I

Q.1. Define slope of Equation of straight line and solve $x^2 - 16x + 48 = 0$

Q.2. A market demand Curve is D = 120 – 5p. Find the price if the quantity demanded is 20 units. Also find the quantity demanded if the price is 18. What would be quantity demanded if it were a free good?

Unit - II

Q.3. If $f(x) = x^2 - 5x + 3$ then find out the following:

(i) $f(0)$

(ii) $f(-2)$

(iii) $f(3)$

(iv) $f(½)$

Q.4. (a) If $y = x^n$, then Prove that

$$\frac{dy}{dx} = nx^{n-1}$$

(b) Find $\frac{dy}{dx}$ if $y = \frac{2}{(x+2)^2}$

Unit - III

Q.5. Discuss the various measures of central tendency.

Q.6. Following are the marks obtained by 10 students of a class. Calculate Standard deviation and coefficient of standard deviation.

Marks: 12 8 17 13 15 9 18 16 6 1

Unit - IV

Q.7. Find Correlation by Rank method.

X	50	55	65	50	55	60	50	65	70	75
Y	110	110	115	125	140	115	130	120	119	160

Q.8. Discuss method of Least squares.

Model Test Paper - 3

ECONOMICS (Semester - IV)

(Quantitative Methods in Economics)

Time Allowed: 3 Hrs. ***Maximum Marks: 80***

Section-A

Note: Attempt any one question from each unit. Each question carries 6 marks.

Unit - I

Q.1. (a) Solve $\frac{1}{x+1}+\frac{3}{x+4} = \frac{4}{x+3}$

(b) Solve $(x+2)\ (x+3)\ (x+5)\ (x+6) = 504$

Q.2. Discuss the Role of Mathematics in Economics.

Unit - II

Q.3. Differentiate the following:

(i) $y = \frac{x}{4x+2}$

(ii) $y = e^x \log x$

Q.4. Find the relationship between marginal revenue and elasticity of demand i.e.

Prove that $MR = AR\left(1+\frac{1}{e_d}\right)$

Unit - III

Q.5. Define Standard Deviation. Give formula of Standard Deviation for discrete series.

Q.6. What is meant by Mean Deviation? What are the methods to calculate it?

Unit - IV

Q.7. Define Covariance with help of example.

Q.8. State Chi-Squared test with help of example.

Section - B

Note: Attempt any one question from each unit. Each question carries 14 marks. $(4\times14 = 56)$

Unit - I

Q.1. Explain the economic application of the equations by giving an example.

Q.2. Given the demand curve $D = 20 - 2p$ and the supply curve $S = -4 + 3p$.

(a) Find the equilibrium price.

(b) Find the equilibrium quantity exchanged.

Unit - II

Q.3. Show that the maximum value of the function $y = x^3 - 27x + 108$ is 108 more than the minimum value.

Q.4. Determine Price elasiticity of demand and marginal revenue if $q = 30 - 4p - p^2$ where q is quantity demanded and p is price and $p = 3$.

Unit - III

Q.5. Find out Median value of the following distribution:

Wage Rate	0-10	10-20	20-30	30-40	40-50
No. of workers	22	38	46	35	20

Q.6. Calculate Mean Deviation from mean from the following:

Marks	No. of students
0-10	4
10-20	6
20-30	10
30-40	20
40-50	10
50-60	6
60-70	4

Unit - IV

Q.7. Calculate the coefficient of Rank correlation from the following data:

X	87	22	33	75	37
Y	29	63	52	46	48

Q.8. Write merits and Demerits of Karl Pearson's correlation coefficient.