B.Com. II Year Commerce

PAPER VI

Business Mathematics and Statistics

Part A : Business Mathematics

Part B : Business Statistics

Unit : I – III & I – IV



SCHOOL OF OPEN LEARNING

(Campus of Open Learning)

University of Delhi

Department of Commerce

Prepared by : Dr. K.L. Dahiya

Graduate Course

Paper VI : Business Mathematics and Statistics

CONTENTS:

Part A: Business Mathematics

UNIT–I : Matrices and Determinants
Lesson : Matrices and Determinants

UNIT–II: Calculus
Lesson 1 : Calculus
Lesson 2 : Differentiation
Lesson 3 : Indefinite Integration

UNIT–III: Basic Mathematics of Finance
Lesson 4 : Simple and Compound Interest

Part B: Business Statistics

UNIT–I : Descriptive Statistics for Unviariate Data
Lesson 1 : Construction fo Frequency Distribution and Graphical Presentation
Lesson 2 : Measures of Central Tendency
Lesson 3 : Measures of Dispersion

UNIT–II : Correlation and Regression Analysis
Lesson 1 : Correlation
Lesson 2 : Repression Analysis

UNIT–III: Index Numbers
Lesson : Index Numbers

UNIT–IV: Time Series Analysis
Lesson : Analysis of Time Series

Editor:

Dr. K.L. Dahiya

SCHOOL OF OPEN LEARNING
UNIVERSITY OF DELHI
5, Cavalry Lane, Delhi - 110007

Academic Session 2013-14 (28000 copies)

© School of Open Learning

Published by : Executive Director, School of Open Learning, 5, Cavalry Lane, Delhi-110 007
Printed at : Nutan Printers, F-89/12, Okhla Industrial Area, Phase-I, New Delhi-110020.

UNIT - I

MATRICES AND DETERMINANTS

Matrix algebra is an important tool of decision making in the problems of business and economics under the assumption of linearity.

Matrix is a set of elements arranged in the form of a rectangular array, such as

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Each matrix is denoted by a single capital letter and by brackets like [ ] or ( ). Each number of the matrix is called an element. Elements written horizontally form a row and vertically written elements form a column. Each element has two subscripts, first one indicates its row and the second one indicates the column of the matrix. In the above example elements a_{11} means element of the first row and first column; similarly a_{23} indicates the element of second row and the third column. Rows are always read first and columns afterwards. The number of rows and the number of columns specify the dimension of a matrix. The above matrix is of 3 rows and 3 columns. Therefore, its dimension is 3×3 (read 3 by 3). The dimension of a matrix is written as a subscript to the matrix, such as

$$A=\begin{bmatrix} 1 & 4 & 3 \\ 2 & 1 & 8 \\ 0 & 9 & 5 \end{bmatrix}_{3\times3} \qquad B=\begin{bmatrix} p & 1 & q \\ c & t & r \end{bmatrix}_{2\times3}$$

$$C=\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots\cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots\cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots\cdots & a_{mn} \end{bmatrix}_{m\times n}$$

Total number of elements in a matrix are equal to the product of its rows and columns. For example, matrix A is of 3×3 order, therefore, the total number of elements in the above matrix will be 9. Similarly total number of elements in matrix C will be mn.

The following are some important facts about matrices :

a) If a matrix has only one row it is called a row matrix or row vector.

$X = [3 \;\; 2 \;\; 7]$

b) If a matrix has only one column it is called a column matrix or column vector.

$$Y=\begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix}$$

c) Two matrices are equal, if and only if, there corresponding elements are equal and their dimensions are the same. For example,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}_{2\times 2} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}_{2\times 2}$$

$$A = B \quad if \quad a_{11} = b_{11} \qquad a_{12} = b_{12}$$

$$a_{21} = b_{21} \qquad a_{22} = b_{22}$$

d) If all the elements of a matrix are zero it is called a Null matrix (or Zero matrix). For example,

$$O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{2\times 3}$$

Types of Matrices

1) Square Matrix

If the number of rows of a matrix are equal to the number of its columns i.e., m = n, it is called a square matrix. For example;

$$A = \begin{bmatrix} 3 & 7 \\ 4 & 2 \end{bmatrix}_{2\times 2} \qquad B = \begin{bmatrix} 1 & 3 & 4 \\ 5 & 7 & 6 \\ 4 & 0 & 2 \end{bmatrix}_{3\times 3}$$

Matrix A is a square matrix as there are only 2 rows and 2 columns. It is called a square matrix of 2×2. Similarly B is a square matrix of 3×3.

2) Diagonal Matrix

A square matrix is called a diagonal matrix when all its elements except the principal diagonal elements are zero. For example :

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}_{3\times 3}$$

A is a diagonal matrix.

3) Identify Matrix or Unit Matrix

A square matrix whose principal diagonal are 1 and all elements above and below the principal diagonal are zero is known as an Identify matrix. For example:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{3\times 3}$$

It is an Identity matrix of 3×3.

4) Scalar Matrix

If the diagonal elements of a matrix are equal it is called a Scalar matrtix. For example:

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \qquad X = \begin{bmatrix} x & 0 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \end{bmatrix}$$

Matrix A and X are Scalar matrices.

Algebra of Matrices

(1) Addition of Matrices :

Two matrices $A_{m\times n} = a_{ij}$ and $B_{m\times n} = b_{ij}$ can be added if their dimensions are same. If the addition matrix is called matrix C, then

$$C_{m\times n} = A_{m\times n} + B_{m\times n}$$

Dimensions of matrix C will be similar to the dimensions of matrices A and B. Matrix C can be obtained by adding corresponding elements of matrix A with the elements of matrix B. For example;

$$A = \begin{bmatrix} 5 & 4 & 0 \\ 7 & 1 & 2 \\ 1 & 3 & 6 \end{bmatrix}_{3\times 3} \qquad B = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 4 & 6 \\ 5 & 1 & 0 \end{bmatrix}_{3\times 3}$$

$$Then\ C = \begin{bmatrix} 5+1 & 4+3 & 0+2 \\ 7+2 & 1+4 & 2+6 \\ 1+5 & 3+1 & 6+0 \end{bmatrix}_{3\times 3} = \begin{bmatrix} 6 & 7 & 2 \\ 9 & 5 & 8 \\ 6 & 4 & 6 \end{bmatrix}$$

$$\therefore \quad C_{ij} = a_{ij} + b_{ij}$$

If the dimensions of two matrices are different, they cannot be added. For example

$$A = \begin{bmatrix} 0 & 2 \\ 7 & 6 \end{bmatrix}_{2\times 2} \qquad B = \begin{bmatrix} 5 & 1 & 2 \\ 7 & 3 & 6 \end{bmatrix}_{2\times 3}$$

$A_{2\times 2}$ and $B_{2\times 3}$ cannot be added.

(a) Matrix addition is commutative, i.e.,

$$A + B = B + A \quad \text{[provided their dimensions are same]}$$

(b) If A, B, C are three matrices of the same dimensions, then

$$A + (B + C) = (A + B) + C$$

It indicates that matrix addition is associative.

(c) If Null matrix $O_{m\times n}$ is added with a matrix $A_{m\times n}$ we get only matrix $A_{m\times n}$. For example :

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}_{2\times 2} \qquad A = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}_{2\times 2}$$

$$O_{2\times 2} + A_{2\times 2} = A_{2\times 2} = A_{2\times 2} + O_{2\times 2}$$

(2) Subtraction of Matrices

Subtraction of two matrices is possible if their dimensions are same. For example :

$(a_{ij}) - (b_{ij}) = (d_{ij})$ where $d_{ij} = a_{ij} - b_{ij}$

$$A = \begin{bmatrix} 7 & 0 & 1 \\ 2 & 7 & 5 \end{bmatrix}_{2\times 3} \qquad B = \begin{bmatrix} 2 & 8 & 4 \\ 1 & 3 & 6 \end{bmatrix}_{2\times 3}$$

$\therefore$ D = A − B

or $D = \begin{bmatrix} 7-2 & 0-8 & 1-4 \\ 2-1 & 7-3 & 5-6 \end{bmatrix} = \begin{bmatrix} 5 & -8 & -3 \\ 1 & 4 & -1 \end{bmatrix}_{2\times 3}$

Dimensions of matrix D will be same as that of A and B.

(3) Multiplication of a matrix by a Scalar

When a real number is multiplied with all the elements of a matrix, it is called a scalar multiplication and the number is known as a scalar (S). For example

S = Scalar = 2 $\qquad A = \begin{bmatrix} 1 & 3 & 9 \\ 0 & 7 & 5 \\ 6 & 2 & 8 \end{bmatrix}$

$$\therefore SA = \begin{bmatrix} 2\times 1 & 2\times 3 & 2\times 9 \\ 2\times 0 & 2\times 7 & 2\times 5 \\ 2\times 6 & 2\times 2 & 2\times 8 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 18 \\ 0 & 14 & 10 \\ 12 & 4 & 16 \end{bmatrix}$$

In general $SA = (Sa_{ij})_{m\times n}$

(4) Multiplication of two matrices

If the columns of the first matrix are equal to the rows of the second matrix (known as condition of conformity), the two matrices can be multiplied.

If $A_{m\times n} = (a_{ij})$ *and* $B_{n\times o} = (b_{jk})$ *then* $(AB)_{m\times o} = C = (C_{ik})$

For example:

$$A = \begin{bmatrix} 3 & 2 & 7 \\ 4 & 6 & 8 \end{bmatrix}_{2\times 3} \qquad B = \begin{bmatrix} 4 & 1 & 7 \\ 5 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix}_{3\times 3}$$

A and B can be multiplied becuase the columns of matrix A are 3 and rows of matrix B are 3. The dimensions of their product AB will be 2×3. Let the product matrix be known as matrix C. It will be arrived at by summing the products of the elements of the i^{th} row of matrix A and the corresponding elements of the k^{th} column of matrix B.

Each element of the Ist row of matrix A will be multiplied by each elements of the Ist column of matrix B, then its sum will give us the first element of the product matrix C. Then *each* element of the Ist row of A will be multiplied by all the elements of the second column of B. Their sum will give us the second element of Ist row of C. Similarly *each* element of Ist row of A will be multiplied by all the elements of 3rd column of matrix B. Their sum will denote the 3rd element of Ist row of C. This process will be repeated for all the subsequent rows of A.

$$C = \begin{bmatrix} (3\times 4)+(2\times 5)+(7\times 1) & (3\times 1)+(2\times 2)+(7\times 3) & (3\times 7)+(2\times 1)+(7\times 2) \\ (4\times 4)+(6\times 5)+(8\times 1) & (4\times 1)+(6\times 2)+(8\times 3) & (4\times 7)+(6\times 1)+(8\times 2) \end{bmatrix}$$

$$\text{or } C = \begin{bmatrix} (12+10+7) & (3+4+21) & (21+2+14) \\ (16+30+8) & (4+12+24) & (28+6+16) \end{bmatrix}$$

$$\text{or } C = \begin{bmatrix} 29 & 28 & 37 \\ 54 & 40 & 50 \end{bmatrix}_{2\times 3}$$

Illustraion 1. $A=\begin{bmatrix}5 & -1\\ 3 & 2\end{bmatrix}$ $B=\begin{bmatrix}-3 & 1\\ 4 & -2\end{bmatrix}$

find AB and BA.

Solution :

$$AB=\begin{bmatrix}(5\times-3)+(-1\times4) & (5\times1)+(-1\times-2)\\ (3\times-3)+(2\times4) & (3\times1)+(2\times-2)\end{bmatrix}$$

$$=\begin{bmatrix}(-15-4) & (5+2)\\ (-9+8) & (3-4)\end{bmatrix}=\begin{bmatrix}-19 & 7\\ -1 & -1\end{bmatrix}$$

$$BA=\begin{bmatrix}(-3\times5)+(1\times3) & (-3\times-1)+(1\times2)\\ (4\times5)+(-2\times3) & (4\times-1)+(-2\times2)\end{bmatrix}$$

$$=\begin{bmatrix}(-15+3) & (3+2)\\ (20-6) & (-4-4)\end{bmatrix}=\begin{bmatrix}-12 & 5\\ 14 & -8\end{bmatrix}$$

Illustration 2.

$$A=\begin{bmatrix}x & 5y\\ 2x & 0\end{bmatrix}\quad B=\begin{bmatrix}0 & \frac{y}{2}\\ \frac{x}{5} & \frac{-y}{10}\end{bmatrix}$$

find AB and BA.

Solution :

$$AB=\begin{bmatrix}(x\times0)+\left(5y\times\frac{x}{5}\right) & \left(x\times\frac{y}{2}\right)+\left(5y\times\frac{-y}{10}\right)\\ (2x\times0)+\left(0\times\frac{x}{5}\right) & \left(2x\times\frac{y}{2}\right)+\left(0\times\frac{-y}{10}\right)\end{bmatrix}$$

$$=\begin{bmatrix}\left(0+\frac{5xy}{5}\right) & \left(\frac{xy}{2}+\frac{-5y^2}{10}\right)\\ (0+0) & \left(\frac{2xy}{2}+0\right)\end{bmatrix}=\begin{bmatrix}xy & \frac{xy-y^2}{2}\\ 0 & -xy\end{bmatrix}$$

$$BA=\begin{bmatrix}(0\times x)+\left(\frac{y}{2}\times2x\right) & (0\times5y)+\left(\frac{y}{2}\times0\right)\\ \left(\frac{x}{5}\times x\right)+\left(\frac{-y}{10}\times2x\right) & \left(\frac{y}{2}\times5y\right)+\left(\frac{-y}{10}\times0\right)\end{bmatrix}$$

$$=\begin{bmatrix}\left(0+\frac{2xy}{2}\right) & (0+0)\\ \left(\frac{x^2}{5}-\frac{2xy}{10}\right) & \left(\frac{5xy}{5}+0\right)\end{bmatrix}=\begin{bmatrix}xy & 0\\ \frac{x^2-xy}{5} & xy\end{bmatrix}$$

We can see : In Illustration 1 and 2, AB $\neq$ BA.

Properties of Matrix Multiplication :

(a) Matrix multiplication is associative. Let $A_{m\times n}, B_{n\times o}$ and $C_{o\times p}$; then $[AB(C)]_{m\times p} = [A(BC)]_{m\times p}$

(b) Matrix multiplication is distributive.

Let $A_{m\times n}, B_{n\times p}$ and $C_{n\times p}$; then

$$A(B+C) = AB + AC$$

AB and AC will be of the dimensions m×p.

(c) Matrix multiplication is not necessarily commutative, i.e., $AB \neq BA$

(proved this property in Illustrations 1 and 2).

(d) Matrix multiplication by an Identity matrix. Let $A_{m\times n}$ and $I_{m\times n}$ then;

$$AI = A = IA$$

Example :

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 7 & -1 \\ 1 & 5 & 4 \end{bmatrix} \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 1 & 1 \\ 2 & 7 & -1 \\ 1 & 5 & 4 \end{bmatrix}$$

$AI = A$

Similarly IA will be equal to A.

(e) Multiplication of a matrix by itself.

If A is a square matrix, $A^2 = AA$.

Similarly $AA^2 = A.A.A = A^2A$ [By Associative Law]

Transpose of a Matrix

The transpose of a matrix is obtained by interchanging its rows into columns and columns into rows. It is denoted by A' or A^T. Let $A = (a_{ij})_{m\times n}$ then $A^T = (a_{ji})_{n\times m}$

For example

$$A = \begin{bmatrix} 4 & 2 & 6 \\ 1 & -1 & 3 \\ 1 & 2 & 0 \\ 5 & 4 & 1 \end{bmatrix} \qquad A' = \begin{bmatrix} 4 & 1 & 1 & 5 \\ 2 & -1 & 2 & 4 \\ 6 & 3 & 0 & 1 \end{bmatrix}$$

4×3 $\qquad$ 3×4

Properties of the Transpose of a matrix

(a) $(A')' = A$

(Transpose of a Transposed matrix is the original matrix.)

(b) $(AB)' = B'A'$

(The product of matrix AB when transposed is equal to the product of Transposed matrices taken in reverse order).

(c) $(A+B)' = A' + B'$

(The Transposed sum of two matrices is equal to the sum of their Transposes).

Symmetric and Skew Symmetric Matrices

(1) If $A' = A$, it is called a symmetric matrix (provided A is a square matrix).

Example :

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix} \quad A' = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}$$

Since $A = A'$, A is a symmetric Matrix.

(2) If $A' = -A$, A will be known as a skew symmetric matrix. For example

$$A = \begin{bmatrix} 0 & -4 \\ 4 & 0 \end{bmatrix} \quad A' = \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix} = -A$$

Determinants

A determinant of a square matrix is a unique number associated with it. To arrive at the determinant of a matrix it is written within vertical bars. e.g.

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

For example:

$$A = \begin{bmatrix} 3 & 5 \\ 2 & 7 \end{bmatrix}$$

Its determinant $|A| = \begin{vmatrix} 3 & 5 \\ 2 & 7 \end{vmatrix}$

$$|A| = (3\times 7) - (5\times 2) = 21 - 10 = 11$$

Determinant of a matrix of 3×3 is arrived at in the following way :

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$|A| = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$|A| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}\{a_{22} \times a_{33} - a_{23} \times a_{32}\} - a_{12}\{a_{21} \times a_{33} - a_{23} \times a_{31}\} + a_{13}\{a_{21} \times a_{32} - a_{22} \times a_{31}\}$$

For example $A = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 0 & 3 \\ 4 & 2 & 5 \end{bmatrix}$ than its determinant will be:

$$|A| = 2\begin{vmatrix} 0 & 3 \\ 2 & 5 \end{vmatrix} - 1\begin{vmatrix} 1 & 3 \\ 4 & 5 \end{vmatrix} + 4\begin{vmatrix} 1 & 0 \\ 4 & 2 \end{vmatrix}$$

$$= 2(0 \times 5 - 3 \times 2) - 1(1 \times 5 - 3 \times 4) + 4(1 \times 2 - 0 \times 4)$$

$$= 2(-6) - 1(-7) + 4(2)$$

$$= -12 + 7 + 8 = 3$$

(**Note** - If the matrix is of one element only its determinant will be the number itself and if the number is negative then the positive of that number will be its determinant).

Properties of Determinants :

(a) If the rows of a determinant are changed into columns or vice versa, the value of the determinant remains unchanged, For example,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}$$

(Note - Determinant of matrix A is equal to the determinant of A')

(b) In a determinant if two rows (or columns) are interchanged, its absolute value remains the same but the sign changes. For example,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = -\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

(c) The determinant of a matrix will be zero if two rows (or columns) are indentical.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{vmatrix} = 0 \qquad \text{Similarly} \qquad \begin{vmatrix} 3 & 2 & 7 \\ 4 & 2 & 7 \\ 6 & 2 & 7 \end{vmatrix} = 0$$

(d) If all the elements of a row (or column) of a determinant are zero, the value of the determinant will also be zero.

$$\begin{vmatrix} 0 & 5 & 2 \\ 0 & -1 & 4 \\ 0 & 1 & -3 \end{vmatrix} = 0$$

(e) If all the elements of a row (or column) of a determinant are multiplied by a number k, then its value is k times the original determinant, For example :

$$\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

(f) If all the elements of a row (or a column) of a determinant are written as sum (or difference) of two or more elements, then it can be written as sum (or difference) of two or more determinants. For example :

$$\begin{vmatrix} a_{11} \pm l_{11} & a_{12} \\ a_{21} \pm l_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \pm \begin{vmatrix} l_{11} & a_{12} \\ l_{21} & a_{22} \end{vmatrix}$$

(g) If k times of all the elements a row (or a column) are added (or subtracted) to all the corresponding elements of another row or column, the value of the determinant remains the same. For example :

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} \pm ka_{11} & a_{21} \pm ka_{12} & a_{23} \pm ka_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Illustration 3

Prove that $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (a-b)(b-c)(c-a)$

Solution : $(C_2 \to C_2 - C_1$ and $C_3 \to C_3 - C_1)$

$$\begin{vmatrix} 1 & 1-1 & 1-1 \\ a & b-a & c-a \\ a^2 & b^2-a^2 & c^2-a^2 \end{vmatrix} = (b-a)(c-a) \begin{vmatrix} 1 & 0 & 0 \\ a & 1 & 1 \\ a^2 & b+a & c+a \end{vmatrix}$$

$= (b–a)(c–a)\ (c+a–b–a) = (b–a)(c–a)(c–b)$

$= (a–b)(b–c)(c–a)$

Illustration 4

Prove that

$$\begin{vmatrix} a & b & c \\ a-b & b-c & c-a \\ b+c & c+a & a+b \end{vmatrix} = a^2 + b^2 + c^2 - 3abc$$

Solution : $(R_1 \to R_1 + R_3)$

$$\begin{vmatrix} a+b+c & a+b+c & a+b+c \\ a-b & b-c & c-a \\ b+c & c+a & a+b \end{vmatrix} = (a+b+c)\begin{vmatrix} 1 & 1 & 1 \\ a-b & b-c & c-a \\ b+c & c+a & a+b \end{vmatrix}$$

$(C_2 \to C_2 - C_1 \text{ and } C_3 \to C_3 - C_1)$

$$= (a+b+c)\begin{vmatrix} 1 & 1-1 & 1-1 \\ a-b & b-c-a+b & c-a-a+b \\ b+c & c+a-b-c & a+b-b-c \end{vmatrix}$$

$$= (a+b+c)\begin{vmatrix} 1 & 0 & 0 \\ a-b & 2b-c-a & -2a+b+c \\ b+c & a-b & a-c \end{vmatrix}$$

$$= (a+b+c)\begin{vmatrix} 2a-c-a & -2a+b+c \\ a-b & a-c \end{vmatrix}$$

$R_2 \to R_2 + R_1$

$$= (a+b+c)\begin{vmatrix} 2b-c-a+a-b & -2a+b+c+a-c \\ a-b & a-c \end{vmatrix}$$

$$= (a+b+c)\begin{vmatrix} b-c & b-a \\ a-b & a-c \end{vmatrix}$$

$$= (a+b+c)\ [(b-c)(a-c)-(b-a)(a-b)]$$

$$= (a+b+c)\ \{a^2+b^2+c^2-ab-bc-ac\}$$

$$= a^3+b^3+c^3-3abc$$

Product of two determinants

$|A|\ |B| = |AB|$ (Matrices A and B should be square matrices and their dimensions should be same. Dimensions of $|AB|$ will be equal to the dimensions of $|A|$ or $|B|$.)

Trace of a Matrix

Every square matix has its trace which is equal to the sum of its diagonal elements. For example :

$$A = \begin{bmatrix} 1 & 5 & 4 \\ 9 & -2 & 3 \\ 6 & 7 & 1 \end{bmatrix}$$

Trace (tr) of A = 1–2+1=0

Singular and Non Singular Matrices

If the determinant of a square matrix is zero it is called a singular matrix, i.e. $|A| = 0$

When $|A| \neq 0$ matrix A is called a non-singular matrix.

Minors and Cofactors of a Matrix.

A lower order determinant obtained after deleting the row and column is called a minor. For example :

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

for a_{11} the minor is $\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$ for a 12 the minor is $\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$ and similarly for

a_{13} the minor is $\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$

In a matrix total number of minors are equal to total number of elements.

When a minor is multiplied by $(-1)^{i+j}$ of the respective element, it is known as a cofactor. The co-factor of a_{11} in the above matrix will be

$$(-1)^{1+1} \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

Similarly the co-factor of element a_{12} will be

$$(-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

Again, in a matrix there will be as many co-factors as the number of elements. (When minors are alternatively multiplied by a minus sign, they form a co-factor.)

Adjoint of a Matrix

Transpose of a co-factor matrix is known as the adjoint matrix. For example, if the following is a co-factor matrix of matrix A.

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$ Its adjoint will be $$Adj = \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix}$$

Illustraion : 5

Find the adjoint of the following matrix

$$A = \begin{bmatrix} 2 & -1 & 1 \\ -10 & 4 & -5 \\ 5 & -2 & 2 \end{bmatrix}$$

Solution :

Co-factor matrix of $A = \begin{bmatrix} -2 & -5 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & -2 \end{bmatrix}$ Adj $A = \begin{bmatrix} -2 & 0 & 1 \\ -5 & -1 & 0 \\ 0 & -1 & -2 \end{bmatrix}$

Inverse of a Matrix

Every non-singular square matrix possesses an inverse which is denoted as A^{-1} (if the matrix is A). A^{-1} is arrived at by the following formula.

$$A^{-1} = \frac{1}{|A|}(\text{Adj } A)$$

(There are other methods also for ascertaining the inverse of a matrix. They will be discussed later. The above method of finding inverse of a matrix is known as adjoint method of finding inverse)

Note- 1. $AA^{-1} = I = A^{-1}A$.

2. The inverse of a matrix is unique i.e., there cannot be more than one inverse of a matrix.

3. $I = I^{-1}$

Illustration 6.

Find the inverse of the following matrix

$$A = \begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix}$$

Solution :

we know $A^{-1} = \frac{1}{|A|}(\text{Adj } A)$

$$|A| = \begin{vmatrix} 2 & 4 \\ 3 & 4 \end{vmatrix} = 8 - 12 = -4$$

Co-factor Matrix of $A = \begin{bmatrix} 4 & -3 \\ -4 & 2 \end{bmatrix}$

$$\text{Adj } A = \begin{bmatrix} 4 & -4 \\ -3 & 2 \end{bmatrix}$$

$$\therefore \quad A^{-1} = \frac{1}{|A|}.(\text{Adj } A) \qquad = -\frac{1}{4}\begin{bmatrix} 4 & -4 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ \frac{3}{4} & \frac{-1}{2} \end{bmatrix}$$

Illustration 7 :

Find the inverse of the following matrix

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 1 \\ 5 & 1 & 1 \end{bmatrix}$$

Solution :

$$A^{-1} = \frac{1}{|A|}.(\text{Adj } A)$$

$$|A| = 1(3-1) - 2)2-5) - 1(2-15) = 2+6+13 = 21$$

$$\text{Adj } A = \begin{bmatrix} 2 & -3 & 5 \\ 3 & 6 & -3 \\ -13 & 9 & -1 \end{bmatrix}$$

$$A^{-1} = \frac{1}{21}\begin{bmatrix} 2 & -3 & 5 \\ 3 & 6 & -3 \\ -13 & 9 & -1 \end{bmatrix} = \begin{bmatrix} \frac{2}{21} & \frac{-1}{7} & \frac{5}{21} \\ \frac{1}{7} & \frac{2}{7} & \frac{-1}{7} \\ \frac{-13}{21} & \frac{3}{7} & \frac{-1}{21} \end{bmatrix}$$

Elementary row operations and Inverse of a Matrix

Row operations method is another technique of finding the inverse of a matrix. For finding the inverse, following row (or column) operations may be undertaken :

(a) The rows (or columns) may be inter changed.

(b) All Elements of a row (or column) may be multiplied (or divided) by a non-zero constant and thus arrived row (or column) may be added or subtracted from another row (or column)

(c) A row (or column) may be replaced by such row (or column) which is calculated by multiplying the original row (or column) by a non-zero constant.

For finding the inverse, we should write the matrix and an Identify matrix in augmented form. By any of the above operations, we convert the matrix into Identity matrix and the adjacent Identity matrix into the inverse matrix.

For example :

$$\text{If } A = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$$

The augmented matrix will be

$$\left[\begin{array}{cc|cc} 4 & 2 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right]$$

$[R_1 \to R_2]$

$$\left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 4 & 2 & 1 & 0 \end{array}\right]$$

$[R_2 \to R_2 - 4R_1]$

$$\left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & -10 & 1 & -4 \end{array}\right]$$

$$\left[R_2 \to R_2 \times -\frac{1}{10}\right]$$

$$\left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & 1 & -\frac{1}{10} & \frac{2}{5} \end{array}\right]$$

$[R_1 \to R_1 - 3R_2]$

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{3}{10} & -\frac{1}{5} \\ 0 & 1 & -\frac{1}{10} & \frac{2}{5} \end{array}\right] \quad \therefore A^{-1} = \begin{bmatrix} \frac{3}{10} & \frac{-1}{5} \\ \frac{-1}{10} & \frac{2}{5} \end{bmatrix}$$

Illustration 8

Find the inverse of the following matrix by Row operation method.

$$A = \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 3 \\ 3 & 2 & 4 \end{bmatrix}$$

Solution :

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 7 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 3 & 2 & 4 & 0 & 0 & 1 \end{array}\right]$$

$R_3 \to R_3 - 3R_1$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 7 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & -4 & -17 & -3 & 0 & 1 \end{array}\right]$$

$R_1 \to R_1 - 2R_2$ and $R_3 \to R_3 + 4R_2$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & -2 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & -5 & -3 & 4 & 1 \end{array}\right]$$

$R_3 \to R_3 \div -5$, $R_2 \to R_2 - 3R_3$ and $R_1 \to R_1 - R_3$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2}{5} & \frac{-6}{5} & \frac{1}{5} \\ 0 & 1 & 0 & \frac{-9}{5} & \frac{-17}{5} & \frac{3}{5} \\ 0 & 0 & 1 & \frac{3}{5} & \frac{-4}{5} & \frac{-1}{5} \end{array}\right]$$

$$\therefore A^{-1} = \begin{bmatrix} \frac{2}{5} & \frac{-6}{5} & \frac{1}{5} \\ \frac{-9}{5} & \frac{-17}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{-4}{5} & \frac{-1}{5} \end{bmatrix}$$

EXERCISES

(1) If $A = \begin{bmatrix} 3 & 1 \\ 2 & 9 \end{bmatrix}, B = \begin{bmatrix} 7 & 3 \\ 6 & 0 \end{bmatrix}$, find$[A+B]$

Solution :

$$[A+B] = \begin{bmatrix} 3+7 & 1+3 \\ 2+6 & 9+0 \end{bmatrix} = \begin{bmatrix} 10 & 4 \\ 8 & 9 \end{bmatrix}$$

(2) If $A = \begin{bmatrix} 5 & -6 & 1 \\ 4 & 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 & 2 \\ 3 & 4 & 1 \end{bmatrix}$, find (A-B)

Solution :

$$(A-B) = \begin{bmatrix} 5-2 & -6-3 & 1-2 \\ 4-3 & 3-4 & 4-1 \end{bmatrix} = \begin{bmatrix} 3 & -9 & -1 \\ 1 & -1 & 3 \end{bmatrix}$$

(3) If $A = \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} a & b \\ 3 & 5 \end{pmatrix}$ find a and b

such that AB = BA. Also compute 3 A + 5B.

Solution :

$$AB = \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 3a+6 & 3b+10 \\ 4a+3 & 4b+5 \end{pmatrix}$$

$$BA = \begin{pmatrix} a & b \\ 3 & 5 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 3a+4b & 2a+b \\ 9+20 & 6+5 \end{pmatrix}$$

$\therefore AB = BA$

$$\text{or} \begin{pmatrix} 3a+6 & 3b+10 \\ 4a+3 & 4b+5 \end{pmatrix} = \begin{pmatrix} 3a+4b & 2a+b \\ 29 & 11 \end{pmatrix}$$

$3a + 6 = 3a + 4b$ $\qquad$ $3b + 10 = 2a + b$

$b = 1.5$ $\qquad$ $b - a = 5$

$4a + 3 = 29$ $\qquad$ $4b + 5 = 11$

$a = 6.5$ $\qquad$ $b = 1.5$

$\therefore a = 6.5$ and $b = 1.5$

$$3A = 3\begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 6 \\ 12 & 3 \end{pmatrix}$$

$$5B = 5\begin{pmatrix} a & b \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 5a & 5b \\ 15 & 25 \end{pmatrix}$$

$$\therefore \quad 3A + 5B = \begin{pmatrix} 9 & 6 \\ 12 & 3 \end{pmatrix} + \begin{pmatrix} 5a & 5b \\ 15 & 25 \end{pmatrix} = \begin{pmatrix} 5a+9 & 5b+6 \\ 27 & 28 \end{pmatrix}$$

(4) If $A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 0 & 2 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 2 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 & 2 \end{bmatrix}$

find AB, A'B, AC' and CB

Solution :

$$AB = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 0 & 2 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1\times2+0\times1+2\times2 & 1\times1+0\times-1+2\times2 \\ 1\times2+1\times1+3\times2 & 1\times1+1\times-1+3\times2 \\ 0\times2+2\times1+-1\times2 & 0\times1+2\times-1+-1\times2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 9 & 6 \\ 0 & -4 \end{bmatrix}$$

$$A'B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 2 & 3 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1\times2+1\times1+0\times2 & 1\times1+1\times-1+0\times2 \\ 0\times2+1\times1+2\times2 & 0\times1+1\times-1+2\times2 \\ 2\times2+3\times1+-1\times2 & 2\times1+3\times-1+-1\times2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 5 & 3 \\ 5 & -3 \end{bmatrix}$$

$$AC' = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 0 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1\times1+0\times0+2\times2 \\ 1\times1+1\times0+3\times2 \\ 0\times1+2\times0+-1\times2 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix}$$

$$CB = \begin{bmatrix} 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1\times2+0\times1+2\times2 & 1\times1+0\times-1+2\times2 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 5 \end{bmatrix}$$

(5) In a certain city there are 50 colleges and 400 schools. Each school and college has 18 peons, 5 clerks and one cashier. Each college in addition has one section officer and one librarian. The monthly salary of each of them is as follows :

Peon - Rs. 300, Clerks - Rs. 500, Cashier - Rs. 600

Section Officer - Rs. 700 and Librarian - Rs. 900.

Using matrix notation find (1) total number of posts of each kind in schools and colleges taken together, (ii) the total monthly salary bill of all the schools and colleges taken together.

Solution :

No. of colleges and schools; $A = [50 \quad 400]_{1\times2}$

No. of Employees $B = \begin{bmatrix} 18 & 5 & 1 & 1 & 1 \\ 18 & 5 & 1 & - & - \end{bmatrix}_{2\times5}$

Salary matrix $S = [300 \quad 500 \quad 600 \quad 700 \quad 900]_{1\times5}$

Total no. of posts $AB = [50 \quad 400]\begin{bmatrix} 18 & 5 & 1 & 1 & 1 \\ 18 & 5 & 1 & - & - \end{bmatrix}$

$= [8100 \quad 2250 \quad 450 \quad 50 \quad 50]_{1\times5}$

Total monthly salary of each college & schools is

$$BS' = \begin{bmatrix} 18 & 5 & 1 & 1 & 1 \\ 18 & 5 & 1 & - & - \end{bmatrix}_{2\times5} \begin{bmatrix} 300 \\ 500 \\ 600 \\ 700 \\ 900 \end{bmatrix}_{5\times1} = \begin{bmatrix} 10,100 \\ 8500 \end{bmatrix}_{2\times1}$$

Total monthly bill of all colleges & schools

$$A(BS') = [50 \quad 400]\begin{bmatrix} 10,100 \\ 8500 \end{bmatrix} = [3905000]$$

LESSON-1

CALCULUS

Functions, Limits and Continuity

Function

A function is a special type of relation that expresses how one quantity depends on another quantity. For example, when money is invested at some interest rate, the interest 'I' depends on the length of time, this dependence can be expressed by saying that "I is a function of t".

Let A and B be two non-empty sets. A function f from A to B is a rule which associates to each element x of A *a unique* y of B. The element y is called the image of x under f and we write $f(x) = y$ which is read "f of x equals y". A is called the *domain* of f and B is called the co-domain of f. The set of all images of the elements of the domain is called the *range* of f. The elements x and y are called the *independent* and *dependent* variables respectively.

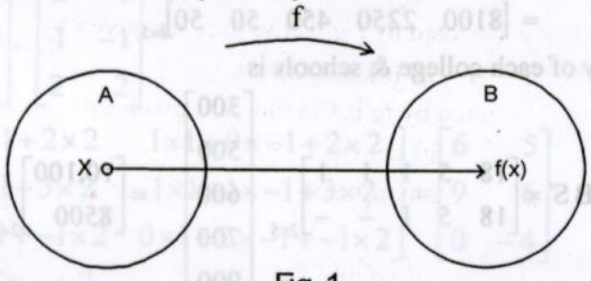


Fig. 1

Example : Let A = {a, b, c} and B = {1, 2, 3, 4}. The diagrams in Figure 2 show a correspondence by which elements of A are associated with elements of B. Which of these diagrams define a function from A to B?

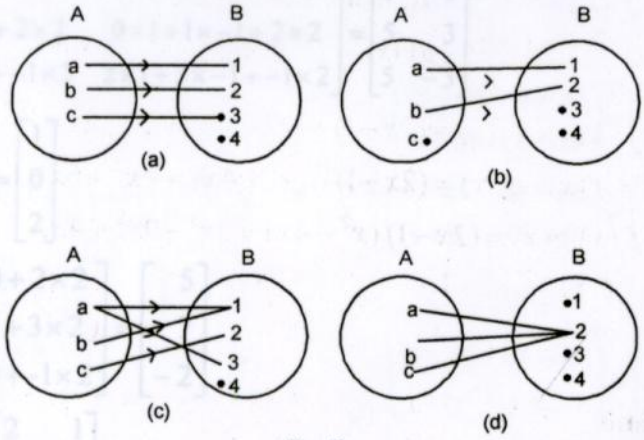


Fig. 2

Solution : In figure 2 (a), each element of A is associated with a unique element of B. Thus Figure 2 (a) defines a function.

In Figure 2 (b), not all elements in A are associated with elements in B. The element c in A has no image in B. Thus Figure 2 (b) does not define a function.

In Figure 2 (*c*), the element *a* of A is not mapped with exactly one element in B, but is mapped with 1 and 3. Thus Figure 2 (*c*) does not define a function.

In Figure 2 (*d*), each element in A is mapped with one element of B (the element happens to be the same for *a*, *b* and *c*). Thus Figure 2 (*d*) defines a function from A to B.

Example : On a given day, a factory can produce 0 to 4 units. The daily operating cost of the factory is given by :

x (number of units):	0	1	2	3	4
y (daily operating cost):	Rs.2000	5000	7000	9000	10000

Solution: To each x (number of units) there corresponds exactly one y (daily operating cost). Thus this defines a function. The domain of the function is the set A = {0, 1, 2, 3, 4} and the range is the set B = {2000, 5000, 7000, 9000, 10000}.

Algebra of Function. If f and g are two functions, then

(i) the sum of f and g, denoted by $f + g$, is the function given by

$$(f + g)(x) = f(x) + g(x),$$

(ii) the difference of f and g, denoted by f-g, is the function given by

$$(f - g)(x) = f(x) - g(x),$$

(iii) the product of f and g, denoted by fg is the function given by

$$(fg)(x) = f(x)\, g(x),$$

(iv) the quotient of f and g, denoted by f/g is the function given by

$$(f/g)(x) = \frac{f(x)}{g(x)}$$

In each of the above operations, x is assumed to lie in the domains of both f and g. However, in the quotient, any value of x for which $g(x) = 0$ is excluded.

Example : If $f(x) = 2x - 1$ and $g(x) = x^2 - 4x$, find the following :

(*a*) $(f + g)(x)$ (*b*) $(f - g)(x)$ (*c*) $(fg)(x)$ (*d*) $(f/g)x$.

Solution. (*a*) $(f+g)(x) = f(x)+g(x) = (2x-1)+(x^2-4x) = x^2-2x-1,$

(*b*) $(f-g)(x) = f(x)-g(x) = (2x-1)-(x^2-4x) = -x^2+6x-1,$

(*c*) $(fg)(x) = f(x)\, g(x) = (2x-1)(x^2-4x) = 2x^3-9x^2+4x,$

(*d*) $f(g)(x) = \dfrac{f(x)}{g(x)} = \dfrac{2x-1}{x^2-4x}$

Special Types of Functions

In this section we introduce some functions that have special forms and representations.

Constant Functions. A function of the form

$$f(x) = c$$

where c is a constant, is called a *constant* function. Thus to each x, f associates the same number c. For example, the function

$f(x) = 10$

is a constant function, all functional values are 10. Figure 3 shows that graph of the function $f(x) = 10$, a horizontal line 10 units above the X-axis.

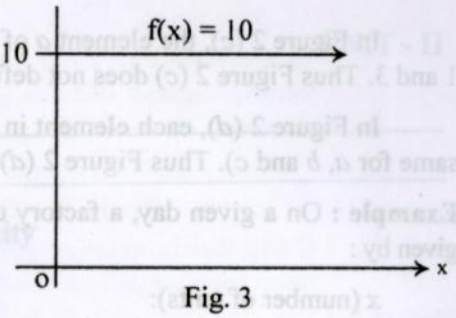


Fig. 3

Linear Function. A function of the form

$$f(x) = ax + b$$

where a and b are constants and $a \neq 0$, is called a *linear* function.

For example, the function $f(x) = 2x + 1$ is a linear function.

The graph of a linear function is a straight line. We say that the function $f(x) = ax + b$ has slope a.

Example : A manufacturer earns Rs. 4,500 in the first month and Rs. 6,000 in the second month. On plotting these points, the manufacturer observes a linear function may fit the data.

(i) Find the linear function that fits the data.

(ii) Using your model make a prediction of the earning for the third month.

Solution. (i) Let the linear function be $y = ax + b$, where y denotes earnings, x denotes the months and a and b are constants. From the given data, we obtain

$$4500 = a + b \quad \text{....(i)}$$
$$6000 = 2a + b \quad \text{....(ii)}$$

Subtracting *(i)* from *(ii)*, we get

$$a = 1500$$

Substituting this value in *(i)* gives $b = 3000$. Thus the required linear function is

$$y = 1500x + 3000$$

(ii) Earning for the third month is: $y = 1500\,(3) + 3000 =$ Rs. 7500.

Polynomial Function. A function of the form

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots\ldots + a_n$$

where n is a non-negative integer and $a_0, a_1, \ldots\ldots a_n$ are constants with $a_0 \neq 0$, is called a *polynomial* function in x of degree n, a_0 is called the leading co-efficient.

For example, the function $f(x) = 2x^3 - 5x^2 + 7x + 1$ is polynomial function in x of degree 3.

Note that a linear function is a polynomial of degree 1 and a quadratic function is a polynomial of degree 2. The domain of any polynomial is the set of all real numbers.

Rational Functions. A function of form

$$f(x) = \frac{p(x)}{q(x)}$$

where $p(x)$ and $q(x)$ are polynomials, is called a *rational* function.

Absolute Value Function. For any real number x, the function f defined by

$$f(x) = |x| = \begin{cases} x, \text{if } x \geq 0 \\ -x, \text{if } x < 0 \end{cases}$$

is called the *absolute value* function. The domain of f is the set of all real numbers and the range is the set of non-negative real numbers. The graph of $f(x) = |x|$ is shown in Figure 4

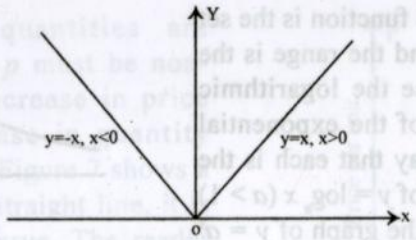


Fig. 4

Exponential Functions. A function defined by

$$f(x) = a^x$$

where a > 0 $\neq 1$ and the exponent x is any real number, is called an *exponential* function to the base a.

For example, $y = 2^x, y = 3^x$ and $y = \left(\frac{1}{2}\right)^x$ are exponential functions.

The domain of an exponential function is all real numbers and the range is all positive real numbers. If $a > 1$, the graph of $y = a^x$ rises from left to right (Figure 5). If $0 < a < 1$, then the graph of $y = a^x$ falls from left to right (Figure 5). Further, each exponential function has y intercept (0, 1). There is no x-intercept.

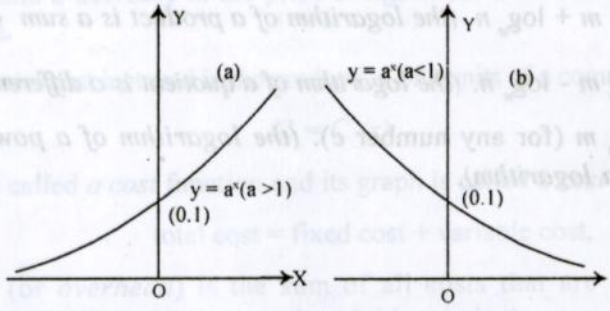


Fig. 5

Of particular interest and utility for business applications is the exponential function $y = e^x$, where e is the exponential number which is approximately 2. 7183.

If m and n are real numbers and a and b are positive, the following laws of exponents are satisfied:

$$(i)\ a^m a^n = a^{m+n} \quad (ii)\ \frac{a^m}{a^n} = a^{m-n} \quad (iii)\ (a^m)^n = a^{mn} \quad (iv)\ (ab)^n = a^n b^n \quad (v)\left(\frac{a}{b}\right)^n \qquad \text{....(i)}$$

$$= \frac{a^n}{b^n} \quad (vi)\ a^{-n} = \frac{1}{a^n} \quad (vii)\ a^0 = 1 \qquad (viii)\ a^1 = a. \qquad \text{....(iv)}$$

Logarithmic Function. Given a positive number a, where a $\neq 1$, and a positive number x, the logarithm to the base a of x, denoted by $\log_a x$, is defined by

$$y = \log_a x \text{ if and only if } a^y = x$$

Thus the logarithm of a number is the exponent to which we must raise the base to get the number. For example,

$\log_2 8 = 3$	because $2^3 = 8$,
$\log_3 81 = 4$	because $3^4 = 81$, and
$\log_{10} 1000 = 3$	because $10^3 = 1000$.

The domain of the logarithmic function is the set of all positive real numbers and the range is the set of all real numbers. Since the logarithmic function reverses the action of the exponential function and vice versa, we say that each is the inverse of the other. The graph of $y = \log_a x$ $(a > 1)$ can be obtained by reflecting the graph of $y = a^x$ $(a > 1)$ in the line $y = x$ (Figure 6).

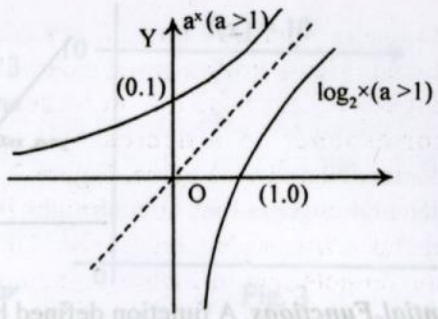


Fig. 6

The two most widely used bases for logarithms are "10" and "e", respectively.

Common Logarithms. Logarithms to the base 10 are called *common* logarithms.

Natural Logarithms. Logarithms to the base e are called *natural* logarithms. Natural logarithms are denoted by In or $\log_e$, or simply log. This base occurs in economic analysis and many situations involving growth or decay.

Some important properties of the logarithmic function, which are frequently used in calculus, are listed below :

Let m, n and a be positive numbers with $a \neq 1$. Then

(i) $\log_a (mn) = \log_a m + \log_a n$. *(the logarithm of a product is a sum of logarithms).*

(ii) $\log_a \left(\frac{m}{n}\right) = \log_a m - \log_a n$. *(the logarithm of a quotient is a difference of logarithms)*

(iii) $\log_a m^c = c \log_a m$ (for any number c). *(the logarithm of a power of a number is the exponent times a logarithm).*

(iv) $\log_a 1 = 0$

(v) $\log_a a = 1$

(vi) $\log_b m = \frac{\log_a m}{\log_a b}$

Property (vi) **is called the change of base formula.** It allows conversion of logarithms from one base to another.

Remark. It may be noted that $\log_a (m+n) \neq \log_a m + \log_a n$.

Functions related to Business and Economics. In this section we present some functions which are very useful in business and economics.

Demand Function. An equation that relates price per unit and quantity demanded at that price is called a *demand function*. Its graph is called a *demand curve*. If p is the price per unit of a certain product and x is the number of units of that product that consumers will demand during some time period at that price, then we can express the demand function as

$$x = f(p)$$

Here x is the dependent variable and p is the independent variable. Since

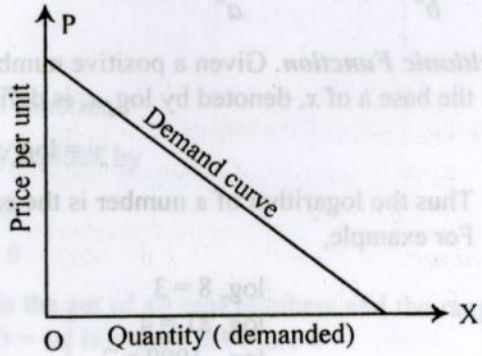


Fig . 7

negative prices and quantities are meaningless, both x and p must be non-negative. Usually, an increase in price corresponds to a decrease in quantity demanded and vice versa. Figure 7 shows a demand curve which is a straight line, it is called a *linear* demand curve. The reader should note that in Figure 7, the horizontal axis is used for the dependent variable x and the vertical axis for the independent variable p.

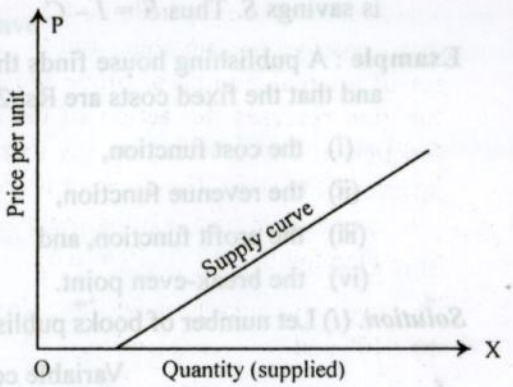


Fig. 8

Supply Function. An equation that relates price per unit and the quantity supplied at that price is called a *supply function* and its graph is called a *supply curve*. If p denotes the price per unit and x denotes the corresponding quantity supplied, then the supply function can be expressed as

$$x = g\ (p)$$

As before x and p are non-negative. Usually, an increase in price corresponds to an increase in the quantity supplied and a decrease in the price brings about a decrease in supply. Figure 8 shows a linear supply curve.

Cost Function. Let C the total cost incurred in the production of x units of a commodity. Then a function, say,

$$C = C\ (x)$$

relating C and x is called *a cost* function and its graph is called a *cost* curve. It may be noted that

total cost = fixed cost + variable cost,

where *fixed cost* (or *overhead*) is the sum of all costs that are independent of the level of production, such as rent, insurance etc. and *variable cost* is the sum of all costs that are dependent on the level of production, such as cost of material, labour, etc.

Revenue Function. Let R be the total revenue or income the company makes by selling x units of a product at price p per unit. Then R is given by the formula

$$R = px$$

R is called the total *revenue* function.

Profit Function. If $R(x)$ and $C(x)$ be the total revenue reveived and the total cost incurred, in the production of x units of a product, then the function P given by

$$P\ (x) = R\ (x) - C\ (x)$$

is called the *profict* function.

Break even point. The break-even point is the level of production where the revenue from the sales is equal to the cost of production. At the break-even point, the company is neither making a profit nor losing money.

Consumption Function. If I denotes the total national income and C denotes the total national consumption, then the function

$$C = f(I)$$

relating I and C is called the *consumption* function. The difference between I and consumption C

is savings S. Thus $S = I - C$.

Example : A publishing house finds that the cost of production directly attributed to each book is Rs. 40 and that the fixed costs are Rs. 25,000. If each book can be sold for Rs. 60, then determine:

(i) the cost function,

(ii) the revenue function,

(iii) the profit function, and

(iv) the break-even point.

Solution. (*i*) Let number of books published by the publishing house is x. From the given information,

$$\text{Variable cost} = 40x \quad \text{and} \quad \text{fixed cost} = \text{Rs. } 25{,}000$$

Hence the total cost function, $C(x)$, may be written as

$$C(x) = 40x + 25000$$

(*ii*) Since each book is sold for Rs. 60, the revenue function, $R(x)$, is given by

$$R(x) = 60x$$

(*iii*) Profit function, $P(x)$, is given by

$$P(x) = R(x) - C(x) = 60x - 40x - 25000 = 20x - 25000.$$

(iv) At the break-even point, $P(x) = 0$. That is,

$$20x - 25000 = 0 \quad i.e., \; x = 1250$$

Hence 1250 books should be produced and sold to achieve break-even point.

Example. A profit making company wants to launch a new product. It observes that the fixed cost of the new product is Rs. 35,000 and the variable cost per unit is Rs. 500. The revenue function for the sale of x units is given by $5000x - 100x^2$. Find (i) profit function, (ii) break-even values, and (iii) the values of x that result in a loss.

Solution. (*i*) If $R(x)$, $C(x)$ and $P(x)$ represents the revenue function, the cost function, and the profit function, then we have

$$R(x) = 5000x - 100x^2$$

$$C(x) = 35000 + 500x$$

$$P(x) = R(x) - C(x)$$

$$= 5000x - 100x^2 - 35000 - 500x$$

$$= -100x^2 + 4500x - 35000$$

Hence the profit function is $P(x) = -100x^2 + 4500x - 35000$

(*ii*) For the break-even values, we have

$$P(x) = 0$$

i.e., $$-100x^2 + 4500x - 35000 = 0$$

$\Rightarrow$ $$x^2 - 45x + 350 = 0$$

or $$(x - 10)(x - 35) = 0$$

$\Rightarrow$ $$x = 10 \text{ or } x = 35$$

Hence the break-even values are 10 and 35

(*iii*) To find the values of x that result in a loss, we have

$$P(x) < 0$$

i.e., $$-100x^2 + 4500x - 35000 < 0$$

$\Rightarrow$ $$x^2 - 45x + 350 > 0$$

or $$(x-10)(x-35) > 0$$

This is possible if either $x < 10$ or $x > 35$.

LIMIT OF A FUNCTION

Function denotes the relationship between dependent and independent variable as $y = f(x)$, y is a dependent variable and x is an independent variable. An independent variable may be assigned value in two ways viz. a fixed value as $x = a$ or tending towards certain constant as $x \to a$. Value of the function when independent variable is assigned, a fixed value is known as defined value of the function, but it is known as limit of the function if independent variable is tending towards certain constant. The implication is that x is not equal to a.

Mathematically limit of a function $f(x)$ when $x \to a$ is denoted by $\lim_{x \to a} f(x)$.

It can be defined as limit of a function $f(x)$ for a given value of the argument x is that constant l to which the function continuously approaches as x approaches the given value such that the difference between constant and the function may be made as small as we would like by making the argument x approaching sufficiently near to the assigned value a. Symbolically, $f(x) \to l$ as $x \to a$ but $|(x - a)| > 0$. In other words, the functin $f(x)$ is said to tend to the limit l as x tends to a if $\in > 0$ is chosen arbitrarily small and it is possible to find $\partial > 0$ such that whenever $|x - a| < \partial$, $|f(x) - l| < \in$ i.e., $f(x)$ can be made as close to l as we would like by taking x sufficiently close to a. It is denoted as $f(x) \to l$ as $x \to a$ which means $f(x)$ is l as x tends to a.

Although it is one of the most difficult concept in mathematics, it is generally used in practice. It is used to describe all those situations where target is such that it is possible to reach near the target, but it is not possible to achieve the target. In other words it is possible to reach near the point, but it is not possible to touch that point. For example, achieving optimum level of installed capacity or achieving optimum level of production are instances of limit.

The entire mathematical system of differential and integral calculus is based on the concept of limit of a function and related conditions of continuity of a function.

Evaluating limit of a function

On the basis of the definition of limit, limit of a function can be obtained by putting $x = a + h$ where $h \to 0$ or $x = a - h$ where $h \to 0$. h is a constant introduced to avoid the equality of x and a. This method can be applied in all conditions and is known as evaluating limit by first principle. It is a direct method of evaluating limit of a function.

Example. Evaluate $\lim_{x \to 1} f(x) = \dfrac{x^2 - 1}{x - 1}$

Solution : Applying direct method, let $x = \lim_{h \to o} (1 + h)$

$$\therefore \lim_{x \to 1} f(x) = \lim_{h \to o} \frac{(1+h)^2 - 1}{1 + h - 1}$$

$$= \lim_{h \to o} \frac{1 + h^2 + 2h - 1}{h}$$

$$= \lim_{h \to o} (h + 2) \qquad \text{(simplication)}$$

$$= \mathbf{2\ Ans.}$$

Example. Evalute $\lim_{h\to o} f(x) = \dfrac{x^n - a^n}{x-a}$

Solution : Let $x = \lim_{h\to o}(a+h)$

$$\lim_{h\to o} f(x) = \lim_{h\to o} \frac{(a+h)^n - a^n}{a+h-a}$$

$$= \lim_{h\to o} \frac{\left[a^n + na^{n-1}h + \dfrac{n(n-1)}{2}a^{n-2}h^2 ..... + h^2\right] - a^n}{h}$$

[Applying binomial expansion]

$$= \lim_{h\to o} \frac{na^{n-1}h + \dfrac{n(n-1)}{2}a^{n-2}h^2 + ... + h^n}{h}$$

$$= na^{n-1}$$

[as all the successive terms involve h which is tending towards zero]

Methods of Evaluating limit of a function

This limit can be evaluated under three different conditions, viz., when x tends towards a definite quantity as $x \to 2$ *or* $x \to -5$; when x tend toward zero and when x tends towards infinity. There are different methods which can be used for evaluating limit of a function under different conditions. (all these are over and above evaluating limit by direct method.)

(a) When x tends towards a definite quantity : When independent variable tends towards a definite quantity which may be positive or negative, the limit can be evaluated by any of the following methods.

(i) *By putting* $x = a$: It is the easiest metod of evaluating limit. It gives defined value as limit of the function. But it can be used only when given function is a polynomial function or a rational function and denominator is equal to zero for $x = a$. In addition to this the method can be used only if putting $x = a$ does not give any indeterminant form explained in the note below.

Example. Evaluate $\lim_{x\to 1} f(x) = \dfrac{4x^4 + 3x^2 - 1}{x^2 + 7}$

Solution : Since $f(x)$ is defined for $x = 1$, so putting $x = 1$,

$$\lim_{x\to 1} f(x) = \frac{4(1) + 3(1) - 1}{1+7}$$

$$= \frac{6}{8} = \frac{3}{4}.$$

Note : Although this method is easy, but it can not be used if limit comes in indeterminant form. It is taken as in indeterminant form if it comes in any of these form, viz., $\dfrac{0}{0}, \dfrac{\infty}{\infty}$ or $\dfrac{b}{0}$ where b is a constant. Limit is not determinable. If the defined value of a function takes any of these forms. These results are not defined in mathematics.

(ii) *By factorisation method* : This method is used when given function is a rational function. It involves factorising numerator and denominator to the extent possible, cancelling the common factor and putting $x = a$ in the simplified form.

Example. Evaluate $f(x) = \dfrac{2x^2 - 7x + 6}{5x^2 - 11x + 2}$ when $x \to 2$.

Solution : When $x = 2$, function is not defined, so applying factorisation method

$$\lim_{x \to 2} f(x) = \frac{(2x-3)(x-2)}{(5x-1)(x-1)} \qquad \text{[Factorisation]}$$

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

Putting $x = 2$

$$\lim_{x \to 2} f(x) = \frac{1}{9}$$ **Ans.**

Example. Evaluate $\lim_{x \to 1} f(x) = \dfrac{x^3 - 5x^2 + 2x + 2}{x^3 + 2x^2 - 6x + 3}$

Solution : When $x = 1$, $f(x)$ is not defined, so factorising the rational function.

$$f(x) = \frac{(x-1)(x^2 - 4x - 2)}{(x-1)(x^2 + 3x - 3)}$$

$$= \frac{x^2 - 4x - 2}{x^2 + 3x - 3}$$

Putting $x = 1$

$$\lim_{x \to 1} f(x) = \frac{1 - 4(1) - 2}{1 + 3(1) - 3} = -5$$

Example. Evaluate $f(x) = \dfrac{x^n - a^n}{x - a}$

Solution : Using factorisation method

$$f(x) = \frac{[x-a]\left[\dfrac{x^{n-1} + x^{n-2}a + x^{n-3}a^2 \ldots\ldots + a^{n-1}}{x-a}\right]}{x-a}$$

$$= x^{n-1} + x^{n-2}a + x^{n-3}a^2 \ldots\ldots + a^{n-1}$$

Putting $x = a$

$$\lim_{x \to a} f(x) = a^{n-1} + a^{n-1} + a^{n-1} + \ldots\ldots n \text{ times}$$

$$= na^{n-1}.$$

(iii) *By rationalisatin method* : This method is used when the given function is in the irrational form. The method involves rationalisation of the numerator or denominator as the case may be simplifying the function after rationalisation putting $x = a$ in the simplified form. The purpose of rationalisation is to simplify by rationalising the numerator or denominator as the case may be.

Example. Evaluate $\lim_{x\to 2} f(x) = \dfrac{x^2-4}{\sqrt{x+2}-\sqrt{3x-2}}$

Solution : When $x = 2, f(x)$ is not defined, so rationalising the denominator.

$$f(x) = \frac{x^2-4}{\sqrt{x+2}-\sqrt{3x-2}} \times \frac{\sqrt{x+2}+\sqrt{3x-2}}{\sqrt{x+2}+\sqrt{3x-2}}$$

$$= \frac{(x^2-4)\left[\sqrt{x+2}+\sqrt{3x-2}\right]}{x+2-3x+2}$$

$$= \frac{(x+2)(x-2)\left[\sqrt{x+2}+\sqrt{3x-2}\right]}{-2(x-2)} \qquad [(a-b)(a+B) = a^2-b^2]$$

$$= \frac{(x+2)\left[\sqrt{x+2}+\sqrt{3x-2}\right]}{-2}$$

Putting $x = 2$

$$\lim_{x\to 2} f(x) = \frac{4[2+2]}{-2}$$

$$= -8$$

Example. Evaluate $\lim_{x\to 2} f(x) = \dfrac{\sqrt{x^2-4}}{\sqrt{x-\sqrt{2}+\sqrt{x-2}}}$

Solution : Since $f(x)$ is not defined for $x = 2$,

Rationalising the denominator

$$f(x) = \frac{\sqrt{x^2-4}}{\sqrt{x-\left[\sqrt{2}+\sqrt{x-2}\right]}} \times \frac{\sqrt{x+}}{\sqrt{x+}}\left[\frac{\sqrt{2}-\sqrt{x-2}}{\sqrt{2}-\sqrt{x-2}}\right]$$

$$= \frac{\left(\sqrt{x^2-4}\right)\left[\sqrt{x+\left[\sqrt{2}-\sqrt{x-2}\right]}\right]}{x-2-x+2+2\sqrt{2}\cdot\sqrt{x-2}}$$

$$= \frac{\sqrt{x+2}\sqrt{x-2}\left[\sqrt{x+\left[\sqrt{2}-\sqrt{x-2}\right]}\right]}{2\sqrt{2}\cdot\sqrt{x-2}}$$

$$= \frac{\sqrt{x+2}\left[\sqrt{x+\left[\sqrt{2}-\sqrt{x-2}\right]}\right]}{2\sqrt{2}}$$

Putting $x = 2$

$$\lim_{x\to 2} f(x) = \frac{2\left[\sqrt{2}+\left[\sqrt{2}\right]\right]}{2\sqrt{2}} = 2$$

(b) **When x tends towards zero :** In case $x \to 0$, same methods can be applied for evaluating limit depending upon nature of function.

Example. Evaluate $\lim_{x\to 0} f(x) = \dfrac{x^2+2ax+a^2}{a^3}$

$$0+2a(0)+a^2 \qquad 1$$

Example. Evaluate $\lim_{x\to 0} f(x) = \dfrac{x}{\sqrt{1+x}-\sqrt{1-x}}$

Solution : Rationalising the denominator,

$$f(x) = \frac{x}{\sqrt{1+x}-\sqrt{1-x}} \times \frac{\sqrt{1+x}+\sqrt{1-x}}{\sqrt{1+x}-\sqrt{1-x}}$$

$$= \frac{x\left[\sqrt{1+x}+\sqrt{1-x}\right]}{1+x-(1-x)} = \frac{x\left[\sqrt{1+x}+\sqrt{1-x}\right]}{2x}$$

Putting $x = 0$

$$= \frac{2}{2} = 1.$$

(c) **When x tends towards ∞** : In this case, limit can be evaluated by dividing all the terms of numerator and denominator by the term having maximum power among all the terms of numerator and denominator taken together. It is in the case of rational functions. After this put $x = \infty$, now all the terms having x in denominator will be zero.

Example. Evaluate $\lim_{x\to\infty} f(x) = \dfrac{x^3+3x^2+6x+5}{x^3+2x+6}$

Solution : Term having maximum power is x^3; so dividing numerator as well as denominator by x^3

$$f(x) = \frac{\dfrac{x^3+3x^2+6x+5}{x^3}}{(x^3+2x+6)/x^3} = \frac{1+\dfrac{3}{x}+\dfrac{6}{x^2}+\dfrac{5}{x^3}}{1+\dfrac{2}{x^2}+\dfrac{6}{x^3}}$$

Putting $x = \infty$

$$\lim_{x\to\infty} f(x) = \frac{1+\dfrac{3}{\infty}+\dfrac{6}{\infty}+\dfrac{5}{\infty}}{1+\dfrac{2}{\infty}+\dfrac{6}{\infty}} = 1.$$

[Any number divided by infinity is equal to zero]

APPLICATIONS OF LIMIT IN BUSINESS AND ECONOMICS

The entire theory of differential and integral calculus is based on this concept which in turn have large applications in business and economics. Not only this, it is used for testing the continuity of functions. The concept limit is directly and indirectly applied in business and economics. Almost all applications are based either on the interpretation of e or geometrical progression with common ratio less than unit. The important applications related with e are continuous compounding and related with geometrical progressions are explained in other applications.

related with e are continuous compounding and related with geometrical progressions are explained in other applications.

1. Continuous compounding and discounting

As stated above, this application is based on e interpretation. It involves compounding or discounting.

In case of compounding involving interest on interest

$$\text{Amount} = P(1+r)^t$$

where P is principle, r is rate of interest in relation to 1 and t is time in years.

It gives amount due after t years if interest is compounded yearly.

But if interest is compounded half-yearly effective rate is half and time is doubled.

$$A(2) = P\left(1+\frac{r}{2}\right)^{2t}$$

where A(2) denotes amount after interest compounded twice a year.

If interest compounded quartely —

$$A(4) = P\left(1+\frac{r}{4}\right)^{4t}$$

If interest compounded monthly —

$$A(12) = P\left(1+\frac{r}{12}\right)^{12t}$$

If interest compounded m times —

$$A(m) = P\left[1+\frac{r}{m}\right]^{mt}$$

If interest compounded continuously then number of times interest compounded will tend to be infinity; and

$$A(m) = \lim_{m\to\infty}\left[P\left(1+\frac{r}{m}\right)^{mt}\right]$$

$$= P\left[\lim_{m\to\infty}\left[\left(1+\frac{r}{m}\right)^{\frac{m}{r}}\right]^{rt}\right] \qquad \left[\frac{m}{r}\cdot rt = mt\right]$$

$$= P\left[\lim_{m\to\infty}\left(1+\frac{r}{m}\right)^{\frac{m}{r}}\right]^{rt} \qquad \left[as \lim_{m\to\infty}\left[1+\frac{r}{m}\right]^{\frac{m}{r}} = e\right]$$

$$= P.e^{rt}$$

So if interest is compounded continuously $A = Pe^{rt}$ where e is a constant. The values of e^{rt} are provided by ready made tables known as exponential table values of e^{rt} corresponding to rt.

Example. Determine the amount at the end of 4 years of an investment of Rs. 1000 if the annual interest rate is 6 per cent being compounded

(i) annually

(ii) half-yearly

(iii) continuously.

Solution : (i) Amount $= P(1 + r)^t$

$= 1000\ (1 + 0.6)^4$

$= 1000\ (1{\cdot}2626)$ [From compound interest table]

$= \text{Rs. } 1262{\cdot}50$

(ii) Amount $= P\left(1+\frac{r}{2}\right)^{2t}$

$= 1000\ (1 + {\cdot}03)^8$

$= 1000\ (1{\cdot}2668)$ [From compound interest table]

$= \text{Rs. } 1266{\cdot}80$

(iii) Amount $= pe^{rt}$

$= 1000\ e^{0.06\times 4}$

$= 1000\ e^{0.24}$

$= 1000\ (1{\cdot}2712)$ [From exponential table]

$= \text{Rs. } 1271{\cdot}20.$

In case of discounting, it is just the reverse of compounding, r becomes rate of discount ; so $P = A(1 + r)^{-t}$ if discounting is annually, if it is half yearly $P = A\left(1+\frac{r}{2}\right)^{-2t}$

Therefore if discounting is m times

$$p(m) = A\left(1+\frac{r}{m}\right)^{-mt}$$

and when it is continously: the number if, m is tending towards ∞ .

So present value $= \lim_{m\to\infty}\left[A\left(1+\frac{r}{m}\right)^{-mt}\right]$

$$= A\left[\lim_{m\to\infty}\left(1+\frac{r}{m}\right)^{\frac{m}{r}}\right]^{-rt}$$

$$= Ae^{-rt}$$

Value of e^{-rt} are providing by exponential table according to rt.

Example. Mr. *X* has borrowed certain money for 3 years on the condition that he will pay Rs. 500 at the end of 3 years. Find the money borrowed if interest compounded (i) annually, (ii) continuously at 8 per cent.

Solution : Let he borrows Rs. P, then

$\therefore$

(i) $A = P(1+r)^t$

or, $P = A(1+r)^{-t}$

$= 500(1+\cdot 08)^{-3}$

$= 500(0\cdot 7938)$ (From present value table)

$= \text{Rs. } 396\cdot 90.$

(ii)

$P = Ae^{-rt}$

$= 500\ e^{-(0.8)(3)}$

$= 500\ e^{-.24}$

$= 500(\cdot 786628)$ (From exp. table)

$= \text{Rs. } 393\cdot 32 \text{ app.}$

2. Measurement of rate of growth

It is also based on e. It is used for the measurement of rate of growth of various economic phenomenon which changes continuously as population, prices etc. In such cases

$$A = Pe^r$$

If A, P and t are given r can be obtained which is the rate of growth.

$$\therefore \quad e^{rt} = \frac{A}{P}$$

It can be simplified in two different ways.

(i) A/P is a constant, and is equal to e^{rt}. Find value of rt corresponding to A/P in the exponential table. Say it is x

then $rt = x$

$$r = \frac{x}{t}.$$

(ii) Take natural log of both sides

$$rt\ In(e) = In\left[\frac{A}{P}\right]$$ [In represents natural log]

or $rt\ In = In\left[\frac{A}{P}\right]$ [as $Ine = 1$. log of any number with same base is always 1].

$$r = In\left[\frac{A}{P}\right]_t$$ [value of In $\left[\frac{A}{P}\right]$ can be obtained from natural log table]

Example. Population of Delhi had increased from 40 lacs during 1970 to 80 lakhs at the end of 1980. Find the rate of growth in population.

Solution : Population growth is continuous

$\therefore \quad A = Pe^{rt}$

or $\quad 80 = 40e^{10r}$

or $\quad e^{10r} = 2$

Simplifying it by both the methods

(i) In the exponential table e^{rt} is equal to 2 when $rt = .695$ app

$\therefore \quad 10\,r = \cdot 695$

$r = \cdot 695$

or $\quad 6 \cdot 95$ per cent

(ii) Taking natural log of both sides

$10\,r = In2$

$$r = \frac{\cdot 6931}{10} = 6 \cdot 93 \text{ per cent app.}$$

3. Other applications

These applications are based on the limit of sum of infinite terms of a geometrical progression being less than unit. These are explained below.

Example. What will be the present value of Rs. 1,000 flowing every year at 10% rate of discount for an infinite number of years.

Solution : Total present value $= \dfrac{1000}{(1+\cdot 10)} + \dfrac{1000}{(1+\cdot 10)^2} \ldots\ldots\ldots\ldots \infty$

$$= \frac{1000}{1 \cdot 10}\left[1 + \frac{1}{1 \cdot 10} + \frac{1}{(1 \cdot 10)^2} \ldots\ldots\ldots\ldots \infty\right]$$

It is forming a geometrical progression with a = 1 and r $= \dfrac{1}{1 \cdot 10}$ which is less than one.

$$\therefore \text{ Total present value} = \frac{1000}{1 \cdot 10}\left[\frac{1}{1 - \frac{1}{1 \cdot 10}}\right]$$

$$= \frac{1000}{\cdot 10} = \text{Rs.}10000$$

Example. Mr. X has made an initial investment of Rs. 1,000. What will be total investment after infinite number of years if $\dfrac{1}{10}$ of the previous is invested again and again.

Solution : Total investment $= 1000 + 1000\left[\frac{1}{10}\right] + 1000\left[\frac{1}{10}\right]^2 \ldots\ldots t$

where $t \to \infty$

$$= \lim_{t\to\infty}\left[\sum_{i=1}^{t} 1000\left[\frac{1}{10}\right]^t\right]$$

$$= 1000\left[1 + \frac{1}{10} + \left(\frac{1}{10}\right)^2 + \left(\frac{1}{10}\right)^3 \ldots\ldots\infty\right]$$

It is forming a geometrical progression with $a = 1$ and $r = \frac{1}{10}$

$$= 1000\left[\frac{1}{1-\frac{1}{10}}\right]$$

$$= \text{Rs. } \frac{10000}{9}$$

Example. A firm has purchased a machinary for Rs.c being depreciated at r per cent per annum on diminishing balance method. Using limit, show that sum of depreciation of all the years is equal to initial cost. (r may be used as in relation to 1)

Solution : Total depreciation $= cr + (c - cr)r + [(c - cr) - (c - cr)r]^r$

$+ \ldots\ldots$ t year

$= cr\,[1 + (1 - r) + [(1 - r) - (1 - r)r]\ r \ldots\ldots t$ years]

$= cr\,[1 + (1 - r) + (1 - r)^2 \ldots\ldots t$ year]

Total depreciation $= \lim_{t\to\infty}[cr[1 + (1 - r) + (1 - r)^2 \ldots\ldots t]$

$[1 + (1 - r) + [(1 - r)]^2 \ldots\ldots t$ years is forming a geometrical progression with $a = 1$ and $r = (1 - r)$ which is less than one]

So total depreciation $= cr\left[\frac{1}{(1-(1-r)}\right]$.

CONTINUITY OF FUNCTIONS

Functions can also be classified as continuous and discontinuous. A function is taken as continuous if its graphical presentation gives a smooth curve. The interval for which it gives a smooth curve. Otherwise, function is known as discontinuous.

The continuity of a function can be tested with the help of limit concept. There are two possible cases in this situtation, viz, continuity at a point and continuity in an interval. Both the situations are explained below.

Continuity at a point

It involves testing the continuity of a function at a particular value of x. It is taken as continuous at a point if the following conditions are fulfilled :

(a) Defined value exists at that point, i.e., $f(a)$ exists.

(b) Limit at that point exists *i.e.*, $\lim_{x\to a} f(x)$ exists.

(c) Both defined value and limit are equal, *i.e.*, $f(a) = \lim_{x\to a} f(x)$

If any of these conditions are not fulfilled, the function is taken as discontinuous. There are different types of discontinuities based on the unfulfilled condition. It is taken as missing point discontinuity if defined value does not exist. If limit does not exist at that point, the discontinuity is taken as finite discontinuity. But if both does not exist, the discontinuity is known as infinite discontinuity.

Example. Given $f(x) = \begin{cases} \dfrac{x^3-1}{x-1} & \text{if } x<1 \\ 3x & \text{if } x>1 \end{cases}$

Determine whether function is continuous at $x = 1$.

Solution : For continuity of the function,

(1) Defind value, $f(1) = 3(1) = 3$. [as $f(x) = 3x$ when $x = 1$]

(2) Left hand limit of $f(x)$, $\lim\limits_{x\to 1} f(x) = \dfrac{(x-1)(x^2+x+1)}{x-1} = 1+1+1 = 3$

Right hand limit of $f(x)$, $\lim\limits_{x\to 1} f(x) = 3(1) = 3$

So both left hand and right hand limit are existing and equal, $\lim\limits_{x\to 1} f(x) = 3$

(3) Both defined value and limit are existing and equal, therefore, function is continuous at $x = 1$.

Example. Given $f(x) = \dfrac{x^2+3x-10}{x-2}$ for $x \neq 2$

What value should be asigned to $f(x)$ to make $f(x)$ continous at $x = 2$

Solution : $\lim\limits_{x\to 2} f(x) = \dfrac{(x-2)(x+5)}{(x-2)}$ [By Factorising]

$= (2+5) = 7$

So $f(x)$ must be equal to 7 for making $f(x)$ continuous at $x = 2$. Only then both defined value and limit will be existing and equal.

2. Continuity in an interval

The limit concept can also be used for testing the continuity of a function in an interval. This interval may be open or close. Hence continuity in an interval may be

(a) continuity in an open interval

(b) continuity in a close interval

(a) *Continuity in an open interval* : An open interval is defined as that interval in which end points are not included. All the values in between end points are included. It is denoted by (a, b).

So continuity in the open interval of (a, b) amounts to testing continuity at all points in between a and b. If the function is continuous at all the points in between a and b, it is taken as continuous in the given interval, otherwise not.

Testing of continuity in the open interval is same as testing at a point becuase at each point, function should fulfill all the three conditions for continuity. It may be applied in the form of taking some points in the given interval and testing the continuity at each point separately.

(b) *Continuity in a close interval* : A close interval is that interval in which end points are also included. It is denoted by $[a, b]$ read as close interval of a and b. The continuity of the function in this close interval requires three conditions :

(i) Right hand continuity at the point a

(ii) Continuity in the open interval of (a, b)

(iii) Left hand continuity at the point b.

(i) A function is taken as having right hand continuity at the point a if the following three conditions are fulfilled.

(a) Defined value of $f(x)$, i.e., $f(a)$ is existing.

(b) right hand limit of $f(x)$ *i.e.*, $\lim_{x \to a+} f(x)$ is existing

(c) $f(a) = \lim_{x \to a+} f(x)$

If any of these conditions is not fulfilled, the function is taken as discontinuous at $x = a$.

(ii) Continuity in the open interval of (a, b) is same as explained earlier.

(iii) Left hand continuity at this point b will be existing only if :

(a) Defind value of $f(x)$, *i.e.*, $f(b)$ is existing

(b) left hand limit of $f(x)$ *i.e.*, $\lim_{x \to b-} f(x)$ is existing

(c) both are equal *i.e.*, $f(b)$ $\lim_{x \to b-} f(x)$

It implies that if a function is continuous in the close interval, it is always continuous in the same open interval. (as it is one of the conditions for continuity in the the close interval). But if a function is continuous in the open interval, it may not necessarily be continuous in the close interval. It may be that function which have only one side continuity either left hand side or right hand continuity.

Continuous Functions

There are some functions which are always in its domain. If a function appears in any of these forms, it can be directly stated that function is continuous without applying tests. These are as follows

(a) *Polynomial function* : A function appearing in the form $f(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n x^0$ where n is a non-negative integer, is known as polynomial function of n^{th} degree. Such functions are always continuous for all the interval because defined value and limit at all points exist and are equal. So if the given function is a polynomial function, it can be taken as continuous in the given interval.

(b) *Rational function* : A rational function appearing in the form of division of two functions is taken as continuous for all the interval if it does not include any point where denominator is equal to zero. If it includes any point where denominator is equal to zero, neither defined value not limit will exist, so function will be discontinuous.

It can be applied in the form as to put the denominator of rational function equal to zero and solve for x. If the given interval consists of that value of x function is taken as discontinuous, otherwise continuous.

Example. Given $f(x) = \dfrac{x^3 + 1}{x^2 - 9}$, determine all values of x for which $f(x)$ is continuous.

Solution : The given $f(x)$ is a rational function, so putting denomination $x^2 - 9 = 0, x = \pm 3$ so $f(x)$ is continuous for all intervals excluding –3 and +3.

Example. Given $f(x) = \sqrt{x^2 - 9} X \in R$

Determine its continuity in the interval $(-\infty, -3), (-\infty, -3), (3, +\infty), (3, +\infty), (-3, 3)$.

Solution : The given function $f(x)$ is defined for all values except for less than 3 and greater than –3 excluding both. For all values in between –3 and + 3, it becomes root of negative which is not defined in real number system. So limit will also not exist and function is discontinuous for value between –3 and 3.

According to the given intervals.

(i) $(-\infty,-3)$, function is continuous

(ii) $(-\infty,-3)$, function is continuous

(iii) $(3,+\infty)$, function is continuous

(vi) $(3,+\infty)$, function is continuous

(v) $(-3,3)$, function is discontinuous

Example. Let $G(x) = \frac{1}{\sqrt{x-2}}, \neq 2, \ x \in R$

Tell for what interval the function $G(x)$ is continuous.

Solution. If x is a real number denominator it is defined only when $x > 2$.

$\therefore$ $G(x)$ is Continuous for the interval $(2, \infty)$

APPLICATIONS IN BUSINESS AND ECONOMICS

The same continuity of function is also relevant in business and economics. It can be directly applied for testing whether a function is continuous at a point or in given interval. In addition to this, it is necessary that a function must be continuous for differentiation which in turn has number of applications.

Example. An electric company charges from its customers the following amount for services; Rs. 3.00 for the first 20 kilowatts hours or less, Rs. 0.10 per kilowatt hour for the next 80 kilowatt hours; and 0.08 per kilowatt hour for any hours above 100 kilowatt hours. Determine the total cost function of x kilowatt hours, $TC(x)$ and test its continuity at $x = 20$ kilowatt hour and x = 100 kilowatt hours.

Solution : Accordingly,

$$TX(x) = \begin{cases} 3 & \text{if } 0 < x \le 20 \\ 3 + \cdot 10(x-20) & \text{if } 20 < x \le 100 \\ 11 + \cdot 05(x-100) & \text{if } x > 100 \end{cases}$$

At $x = 20$

Defined value, TC (20) = 3

$\lim_{x \to 20} TC(x) = 3$ (comparing left hand and right hand limit)

So both are existing and equal, TC(x) is continuous.

At $x = 100$

Defined value, $TC(100) = 11$

$\lim_{x \to 100} TC(x) = 11$ (comparing left hand and right hand limit)

$\therefore$ Function is continuous at $x = 100$.

Example. The total cost function of a firm stating total cost at different levels of output is defined as

$$TX(x) = \begin{cases} 20x & \text{if } 0 < x \leq 100 \\ 18x & \text{if } 100 < x \leq 200 \\ 16x & \text{if } x > 200 \end{cases}$$

Test where the function is continuous in the intervals (100, 200), (100, 200), (100, 200).

Solution : For open interval (100, 200), $TC(x)$ is defined as $18x$ which is polynomial function so it is continuous.

(ii) For the close interval (100, 2000), $TC(x)$ is defined as $20x$ for $x = 100$ but its right hand limit is given $18x$, so it is not continuous at $x = 100$ and function is discontinuous for this interval.

(iii) For the interval (100, 200), $TC(x)$ is defined as $18x$ which is continuous.

LESSON 2

DIFFERENTIATION

Meaning of differentiation

The term 'calculus' means method of calculation. This meaning does not have much relevance in the practical use of the term in relation to differentiation. Calculus is classified as differential and integral. Differential calculus is related with average rate of change.

The method involved is known as differentiation and the result obtained is known as derivative or differential co-efficient.

Practically derivative can be defined as the ratio of change in the value of function to change in the value of independent variable where the change in the latter is tending towards zero. Technically, if $y = f(x)$, then $\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$ is the derivative of y with respect to x where

Δy denotes change in y

Δx denotes change in x

This derivative is often denoted by $\frac{dy}{dx}$ or $f(x)$ or y.

Therefore, derivation is the rate of change in the value of function when limit of change in the value of independent variable is approaching zero.

The steps involved in differentiation are given below :

(i) introduces certain change Δx (delta x) in the dependent variable x and find the corresponding change in the value of function. Technically, if $y = f(x)$, then introducing Δx change in x and Δy change in y.

$$y + \Delta y = f(x + \Delta x)$$
$$\Delta y = (x + \Delta x) - y$$
$$= f(x + \Delta x) - f(x)$$

(ii) divide both sides by Δx giving the ratio of change in value of a function to change in the value of an independent variable. The same is also called as average rate of change.

$$\text{Therefore } \frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

(iii) Lastly, take its limit when $\Delta x \to 0$. This limit must exist for the derivative.

$$\text{Therefore } \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

$$\text{or } \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

The same is the derivative of y with respect to x and it is known as derivative by first principle.

Example. If $y = x^n$, find its derivative by first principle.

Solution : Give $y = x^n$.

then $y + \Delta y = (x + \Delta x)^n$ (introducing change)

$$\Delta y = (x + \Delta x)^n - y^y$$

$$\Delta y = (x + \Delta x)^n - x^n \qquad \left[\because y = x^4\right]$$

$$\Delta y = \left[x^n + nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}\Delta x^2 ..... + \Delta x^n\right] - x^n$$

(apply bionomial distribution)

$$\Delta y = nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}\Delta x^3 ..... + \Delta x^n$$ (simplifying)

Dividing both sides by Δx

$$\frac{\Delta y}{\Delta x} = \frac{nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}\Delta x^2 ..... + \Delta x^n}{\Delta x}$$

$$= nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}\Delta x ..... + \Delta x^{n-1}$$

Taking its limit when $\Delta \to 0$.

$$\lim_{\Delta x \to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0}\left[nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}\Delta x ..... + \Delta x^{n-1}\right]$$

or $$\frac{\Delta y}{\Delta x} = nx^{n-1}$$

Derivative as Instantaneous Rate of Change

This derivative is defined both mathematically and geometrically. According to mathematical interpretation, derivative is used both as velocity of the function and instantaneous rate of change. In physical sciences, it is used as velocity of the function giving rate of movement. In commerce and economics, derivative is used as instantaneous rate of change. Instantaneous rate of change denotes the rate of change in the value of a function at a particular value of independent variable. Derivative gives the same as it is the rate of change in y when limit of rate of change in x is tending towards zero which amounts to a particular value of x.

Example. The demand function of a commodity is $x = 20 - \frac{P}{2}$. Find the rate of change in quantity demanded.

Solution : Given $x = 20 - \frac{P}{2}$

then $x + \Delta x = 20 - \frac{P + \Delta P}{2}$

$$\Delta x = -\frac{\Delta P}{2}$$

Dividing both sides by ΔP and taking its limit when $\Delta P \to 0$

$$\lim_{P \to 0} \frac{\Delta x}{\Delta P} = \frac{-1}{2}$$

or $$\frac{dx}{dP} = -\frac{1}{2}$$ which is constant at all prices.

Example. The total cost of producing x units of a commdity is $TC(x) = x^3 + 100$. Show that rate of change in total cost is varying at all levels.

Solution : Given $TC(x) = x^2 + 100$

Introducing Δx change in x

$$TC(x) + \Delta TC(x) = (x + \Delta x)^2 + 100$$

$$\Delta TC(x) = (x + \Delta x)^2 + 100 - (x^2 + 100)$$

$$= x^2 + \Delta x^2 + 2x\Delta x + 100 - x^2 - 100$$

$$= \Delta x^2 + 2x\Delta x$$

Dividing both sides by Δx

$$\frac{\Delta TC(x)}{\Delta x} = \Delta x + 2x$$

Taking its limit when $\Delta x \to 0$

$$\lim_{\Delta x \to 0} \frac{\Delta TC(x)}{\Delta x} = 2x$$

or $$\frac{dTC(x)}{\Delta x} = 2x$$

which implies it is changing at all values of x.

Standard Forms of Differentiation

There are some standard forms of differentiation which can be used for obtaining the derivatives of the functions appearing in different forms. All these standard forms are based on the first principle of differentiation. (Proofs have been omitted). These are applicable to all forms of functions.

1. If $y = c$ (a constant) $\quad \frac{dy}{dx} = 0$

2. If $y = c.f(x)$ $\quad \frac{dy}{dx} = \frac{c.d\ f(x)}{dx}$

3. If $y = f(x) \pm g(x)$ $\quad \frac{dy}{dx} = \frac{df(x)}{dx} \pm \frac{d\ g(x)}{dx}$ (Summation form)

4. If $y = f(x).g(x)$ $\quad \frac{dy}{dx} = f(x)\frac{d\,g(x)}{dx} + g(x)\frac{df(x)}{dx}$ (Product form)

5. If $y = \frac{f(x)}{g(x)}$ $\quad \frac{dy}{dx} = \frac{g(x)\frac{df(x)}{dx} - f(x)\frac{dg(x)}{dx}}{g^2(x)}$ (Quotient form)

6. If $y = f(u)$ and $u = g(x)$.

$$\frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx} \quad \text{(Chain rule)}$$

7. ***Implicit Functions*** : Functions are also classified as explicit and implicit function. Explicit functions are characterised as those where one variable can be expressed as a function of another variable only as $y = f(x)$. Otherwise it is taken as implicit function as $f(x, y) = 0$. For example, $y = x^2 + 2x + 4$ is an explicit function whereas $x^2 + 2xy + y^2 = 0$ is an implicit function. The derivative of implicit function is obtained using chain rule. For example derivative of y^2 with respect to x can be obtained assuming $u = y^2$, then

$$\frac{du}{dx} = \frac{du}{dy}\cdot\frac{dy}{dx}$$

On this basis, the rule is that all terms involving y will be first differentiated with respect to y and multiply it by $\frac{dy}{dx}$. After this solve it for $\frac{dy}{dx}$.

8. ***Inverse Functions*** : Two functions are taken as inverse if the independent variable of one function becomes the dependent variable of the other and vice-versa. For example, if $y = f(x)$, then $x = f'(y)$.

Then $\quad \frac{dy}{dx}\cdot\frac{dx}{dy} = 1 \quad$ or $\quad \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$

Derivative of Algebraic Functions

Algebraic functions are defined as those functions where quantities are related by mathematic operations such as +, -, ×, ÷, power and square root only. For example, if

$$y = \sqrt{x^2 + 100y} \quad or \quad = \frac{x^2 - 4}{x + 2}$$

The rule of differentiation of algebraic functions is multiply the power of x in the base and reduce the power by one. It is based on the first principle.

$$\therefore y = x^n, \quad then \frac{dy}{dx} = nx^{n-1}$$

Example. If $y = \frac{\sqrt{x+1} + \sqrt{x-1}}{\sqrt{x+1} - \sqrt{x-1}}$, prove that $\frac{dy}{dx} = 1 + \frac{x}{\sqrt{x^2 - 1}}$

Solution : Given $y = \frac{\sqrt{x+1} + \sqrt{x-1}}{\sqrt{x+1} - \sqrt{x-1}}$

$$= \frac{\sqrt{x+1}+\sqrt{x-1}}{\sqrt{x+1}-\sqrt{x-1}} \times \frac{\sqrt{x+1}+\sqrt{x-1}}{\sqrt{x+1}+\sqrt{x-1}} \quad \text{(Rationalising the denominator)}$$

$$= \frac{x+1+x-1+2\cdot\sqrt{x^2-1}}{x+1-x+1} = \frac{2x+2\cdot\sqrt{x^2-1}}{2}$$

$$y = x + \sqrt{x^2-1}$$

$$\therefore \quad \frac{dy}{dx} = 1 + \frac{1}{2}(x^2-1)^{1/2}\cdot 2x = 1 + \frac{x}{\sqrt{x^2-1}}$$

Example. If $y = \dfrac{x}{a-\sqrt{a^2-x^2}}$, prove that $\dfrac{dy}{dx} = \dfrac{-a\left(a+\sqrt{a^2-x^2}\right)}{x^2\left(\sqrt{a^2-x^2}\right)}$

Solution : Given $y = \dfrac{x}{a-\sqrt{a^2-x^2}}$

$$= \frac{x}{a-\sqrt{a^2-x^2}} \cdot \frac{a+\sqrt{a^2-x^2}}{a+\sqrt{a^2-x^2}} \quad \text{(Rationalisation)}$$

$$= \frac{x\left[a+\sqrt{a^2-x^2}\right]}{a^2-a^2+x^2}$$

$$y = \frac{a+\sqrt{a^2-x^2}}{x}$$

$$\therefore \quad \frac{dy}{dx} = \frac{x.a\left(a+\sqrt{a^2-x^2}\right) - \left(a+\sqrt{a^2-x^2}\right)\dfrac{dx}{dx}}{x^2} \quad \text{(Quotient form)}$$

$$= \frac{x.\left|\frac{1}{2}\left(a^2-x^2\right)^{1/2}-2x\right| - \left(a+\sqrt{a^2-x^2}\right)}{x^2}$$

$$= \frac{\dfrac{-x^2}{\sqrt{a^2-x^2}} - a - \sqrt{a^2-x^2}}{x^2} = \frac{-x^2 - a\cdot\sqrt{a^2-x^2} - a^2 + x^2}{x^2\cdot\sqrt{a^2-x^2}}$$

$$= \frac{-a\left[a+\sqrt{a^2-x^2}\right]}{x^2.\sqrt{a^2-x^2}}$$

Derivative of Exponential Function

Another type of non-algebraic function used in business and economics is exponential functions. These functions are defined as those where independent variable appears as power of some constant base. For example $y = 10^x$ or $y = e^x$. These functions are classified as natural and other exponential functions. If the base is e (basis of natural log system having an approximate value of 2.7183), the function is known as natural exponential, otherwise other exponential function. Such functions are often used to denote continuous growth.

Derivative of natural exponential functions is based on the rule if

$$y = e^x, \quad \text{then} \quad \frac{dy}{dx} = e^x.$$

It can be proved by first principle (proof ignored). Derivative of the other exponential function can be obtained easily with the help of logarithmic differentiation which is explained in the next section.

Example. Differentiate $y = x^n\, e^{ax+b}$ with respect to x.

Solution : Given $y = x^n\, e^{ax+b}$

$$\therefore \quad \frac{dy}{dx} = x^n \cdot e^{ax+b} \cdot a + e^{ax+b}\, nx^{n-1}$$

$$= x^{n-1}\, e^{ax+b}\,[ax+n]$$

Example. Differentiate the following with respect to x

(i) $y = x\,e^x$, (iii) $y = \dfrac{e^{x^2}}{x}$, (iii) $y = e^{\sqrt{2x}}\, e^{-x}$,

(iv) $y = 10xe^{-x/50}$, (v) $y = (x+2)\, e^{\frac{1}{x+2}}$

Solution : (i) $y = x \cdot e^x$

$$\therefore \quad \frac{dy}{dx} = x \cdot \frac{de^x}{dx} + e^x \cdot \frac{dx}{dx}$$

$$= x \cdot e^x + e^x \qquad = e^x\,[x+1]$$

(ii) $y = \dfrac{e^{x^2}}{x}$

$$\frac{dy}{dx} = x. \frac{\dfrac{de^{x^2}}{dx} - e^{x^2} \cdot \dfrac{dx}{dx}}{x^2}$$

$$= \frac{x \cdot e^{x^2} \cdot 2x - e^{x^2}}{x^2} \qquad = \frac{e^{x^2}}{x^2}\,[2x^2 - 1]$$

(iii) $y = e^{\sqrt{x^2}} + e^{-x}$

$$\frac{dy}{dx} = e^{\sqrt{2x}} \cdot \frac{d\sqrt{2x}}{dx} + e^{-x} \cdot \frac{d(-x)}{dx}$$

$$= e^{\sqrt{2x}} \cdot \frac{1}{2}(2x)^{-\frac{1}{2}} \cdot 2 - e^{-x} \qquad = \frac{e^{\sqrt{2x}}}{\sqrt{2x}} - e^{-x}$$

(iv) $y = 10xe^{-x/50}$

$$\frac{dy}{dx} = 10\left[x \cdot \frac{d\,e^{-x/50}}{dx} + e^{-x/50} \cdot \frac{dx}{dx}\right]$$

$$= 10\left[xe^{-x/50} \; -\frac{1}{50} + e^{-x/50}\right]$$

$$= 10e^{-x/50}\left[-\frac{x}{50} + 1\right]$$

(v) $y = (x+2)e^{\frac{1}{x+2}}$

$$\frac{dy}{dx} = (x+2)\frac{de^{\frac{1}{x+2}}}{dx} + e^{\frac{1}{x+2}} \cdot \frac{d(x+2)}{dx}$$

$$= (x+2)\, e^{\frac{1}{x+2}} \cdot \frac{-1}{(x+2)^2} + e^{\frac{1}{x+2}}$$

$$= e^{\frac{1}{x+2}}\left[-\frac{1}{x+2} + 1\right]$$

$$= e^{\frac{1}{x+2}}\left[\frac{x+1}{x+2}\right]$$

Derivative of a Logarithmic Function

Logarithmic Functions are also used for denoting functions used in business and economics. Such functions are defined as those whose independent variable appears with log ratio as $y = \log ax$ or $y = \log_{10} x$. Such functions are also classified as natural logarithmic functions, common logarithmic functions and other logarithmic functions based on the base of log. In case of natural log, base is e [a constant with approximate value of 2.7183] as $y = \log_a x$ also denoted by $y = lnx$. But if base of log is 10, it is known as common log as $y = \log_{10} x$ also denoted by $\log x$. If base of log is other than 10 and e, it is known as other logarithmic function as $y = \log_a x$ and $a \neq 10 + e$. In such cases base must be specified.

It is also possible to change the base of log as if $y = lnx$ with e, then with base 10 it becomes $y = \frac{\log x}{\log e}$ whose log e is a constant.

Derivative can be obtained directly for only natural logarithmic functions as if $x = lnx$, then $\frac{dy}{dx} = \frac{1}{x}$ (proof has been omitted). So if other than natural log function is given, first change its base to e and then differentiate. The standard forms are applicable in the same way.

Example. Differentiate the following with respect to x.

(i) $y = ln\,(x^2+1)$, (ii) $y = ln\left(\sqrt{1+x^2}\right)$, (iii) $y = ln^2\,(x+1)$

(iv) $y = x \cdot ln\, x$ (v) $y = ln \cdot \sqrt{\frac{x+1}{x-1}}$

Solution : (i) $y = ln\,(x^2 + 1)$

$$\frac{dy}{dx} = \frac{1}{x^2+1} \cdot \frac{d\left(x^2+1\right)}{dx}$$

$$= \frac{2x}{x^2+1}$$

(ii) $y = ln\left(\sqrt{1+x^2}\right)$

$$\frac{dy}{dx} = \frac{1}{\sqrt{x^2+1}} \cdot \frac{d\sqrt{1+x^2}}{dx}$$

$$= \frac{1}{\sqrt{1+x^2}} \cdot \frac{1}{2}\left(1+x^2\right)^{-1/2} \cdot 2x$$

$$= \frac{x}{1+x^2}$$

(iii) $y = ln^2(x+1)$

or $y = [ln(x+1)]^2$

$$\frac{dy}{dx} = 2ln(x+1) \cdot \frac{d[ln(x+1)]}{dx}$$

$$= \frac{2ln(x+1)}{x+1}$$

(iv) $y = x \cdot lnx$

$$\frac{dy}{dx} = x \cdot \frac{1}{x} + lnx$$

$$= 1 + lnx.$$

(v) $y = ln \cdot \sqrt{\frac{x+1}{x-1}}$

$$\frac{dy}{dx} = \sqrt{\frac{x-1}{x+1}} \cdot d\frac{\sqrt{\frac{x+1}{x-1}}}{dx}$$

$$= \sqrt{\frac{x-1}{x+1}} \frac{1}{2}\left(\frac{x+1}{x-1}\right)^{-1/2} \cdot \frac{(x-1)-(x+1)}{(x-1)^2}$$

$$= \sqrt{\frac{x-1}{x+1}} - \left(\frac{x-1}{x+1}\right)^{-1/2} \cdot \frac{1}{(x-1)^2}$$

$$\therefore = -\left[\frac{1}{(x+1)(x-1)}\right]$$

$$= -\frac{1}{x^2-1}$$

In addition to this, logarithmic differetiation can be used in different ways. It is based on the derivative of logarithmic function. It basically involves three steps (i) taking natural log of both the sides, (ii) Differentiating implicitly both sides with respect to x, and (iii) solving for $\frac{dy}{dx}$. Some forms based on these are given below :

(a) Derivative of natural exponential function :

as $y = e^x$. Taking natural log of both sides

$lny = x \cdot lne$ $[lne = 1]$

$\therefore$ $lny = x$

Differentiating both sides with respect to x

$$\therefore \quad \frac{1}{y}\cdot\frac{dy}{dx}=1$$

$$\therefore \quad \frac{dy}{dx}=y$$

$$=e^x$$

(b) Derivative of other exponential functions as $y = a^x$. Taking natural log of both sides

$lny = x.lna$ [lna is a constant]

Differentiating both sides with respect to x.

$$\therefore \quad \frac{1}{y}\cdot\frac{dy}{dx}=lna$$

or $$\frac{dy}{dx}=lna\cdot y$$

$$=lna\cdot a^x$$

(c) Derivative of a function raised to the power another function as $y = [f(x)]^{g(x)}$. It can be differentiated only with the help of logarithmic differentiation. Taking natural log of both sides.

$lny = g(x) \cdot ln\, f(x)$.

Differentiating both sides with respect to x

$$\frac{1}{y}\cdot\frac{dy}{dx}=g(x)\cdot\frac{1}{f(x)}\cdot f(x)+ln\, f(x)\cdot g'(x)$$

$$\frac{dy}{dx}=\left[\frac{g(x)}{f(x)}\cdot f(x)+ln\, f(x)\cdot g'(x)\right]\cdot y$$

$$=\left[\frac{g(x)}{f(x)}\cdot f(x)+ln\, f(x)\cdot g'(x)\right][f(x)]^{g(x)}$$

Logarithmic differentiation can also be applied for differentiating algebraic and exponential functions if these appears in combined forms.

Example. Differentiate the following with respect to x :

(i) $y = 5^{x+2}$ (ii) $y = a^x \,.\, x^a$ (iii) $y = x^{e^x}$

Solution: (i) $y = 5^{x+2}$. Taking natural log of both sides $lny = (x + 2)$ lns. Differentiating both sides with respect x.

$$\frac{1}{y}\cdot\frac{dy}{dx}=lns.$$

$$\frac{dy}{dx}=lnx\cdot y$$

$$=lns\cdot 5^{x+2}$$

(ii) $y = a^x \cdot x^a$

$$\frac{dy}{dx} = a^x \cdot \frac{d(x^n)}{dx} + x^a \cdot \frac{d(a^x)}{dx}$$

$$= a^x \cdot ax^{a-1} + x^a \cdot a^x \cdot lna$$

$$= a^x \cdot x^{a-1}\left[a + x \cdot lna\right] \left[\frac{da^x}{dx} = a^x \cdot lna\right]$$

(iii) $y = x^{e^x}$. Taking natural log of both sides and differentiating.

$$lny = e^x \cdot lnx$$

$$\therefore \quad \frac{1}{y} \cdot \frac{dy}{dx} = e^x \cdot \frac{1}{x} + lnx \cdot e^x$$

$$\frac{dy}{dx} = e^x\left[\frac{1}{x} lnx\right] \cdot y$$

$$= e^x\left[\frac{1}{x} + lnx\right] x^{e^x}$$

Application based on derivation as instantaneous rate of change :

Whenever we want to find out the rate of change in the function at a particular level then find the derivative of the function and evaluate it for that level. Some examples are given below to explain this :

Example : A firm has a revenue function given by TR(x) = 10x where x is the units demanded and TR(x) is gross revenue with a production cost function given by

$$TC(x) = 1{,}00{,}000 + 50\left[\frac{x}{1000}\right]^2$$

Find the rate of change in profit if x = 1,00,000 units.

Solution :

Given $TR(x) = 10x$, $\quad TC\ (x) = 1{,}00{,}000 + 50\left[\frac{x}{1000}\right]^2$

$\therefore$ Profit function, $\quad n(x) = TR\ (x) - TC\ (x)$

$$= 10x - 1{,}00{,}000 - 50\left[\frac{x}{1000}\right]^2$$

Rate of change in profit at x will be:

$$\frac{dx(x)}{dx} = 10 - 50\left[\frac{2x}{1000000}\right]$$

$$= 10 - \frac{x}{10{,}000}$$

Putting $\quad x = 1{,}00{,}000$ units

$$= 10 - \frac{1{,}00{,}000}{10{,}000} = 0$$

So rate of change in profit at x = 1,00,000 is zero indicating that profit is constant at this level.

Example. The demand function of a commodity at $x = \sqrt{200} - p^2$ where x units are demanded when price is Rs. p per unit. Determine the level of output where rate of change in total revenue is zero.

Solution : Given $x = \sqrt{200 - p^2}$

or $P = \sqrt{200 - x^2}$

$\therefore \; TR\ (x) = P \cdot x = \left(\sqrt{200 - x^2}\right)x$

Rate of change in TR(x) will be:

$$\frac{dTR\,(x)}{dx} = \sqrt{200 - x^2} + x \cdot \frac{1}{2}\left(200 - x^2\right)^{-1/2} - 2x$$

$$= \sqrt{200 - x^2} - \frac{x^2}{\sqrt{200 - x^2}}$$

$$= \frac{200 - 2x^2}{\sqrt{200 - x^2}}$$

Putting it equal to zero and solving for x

$$= \frac{200 - 2x^2}{\sqrt{200 - x^2}} = 0$$

$$200 - 2x^2 = 0$$

$$or \quad x^2 = 100$$

$$or \quad x = 10 \text{ units}$$

Therefore rate of change in total revenue is zero at $x = 10$ units.

APPLICATIONS

The applications based on simple differentiation (derivative concept) are explained below :

1. Measurement of price elasticity of demand

Price elasticity of demand is an important economic concept. It is based on law of demand. Law of demand states the quantitative relationship between price and quantity demanded and if price is high, quantity demanded is low and vice-versa. The same quantitative relationships are quantified by price elasticity of demand. It is always negative because of negative relationships between price and quantity stated by law of demand. In economics, it is defined as

$$E_p = \frac{\text{Relative change in quantity demanded}}{\text{Relative change in price}}$$

Symbolically, if demand function is $x = f(p)$.

$$then \quad E_p = \frac{\frac{\Delta x}{x}}{\frac{\Delta p}{p}} = \frac{\Delta x}{\Delta p} \cdot \frac{p}{x} \; where\ \Delta p \text{ change in price and } \Delta x \text{ change in quantity demanded.}$$

This gives the price elasticity of demand for certain Δp change in price. But if price elasticity of ired at given price p instead of certain change then limit of Δp will tend towards zero and

$$E_p = \lim_{\Delta p \to 0} \frac{\Delta x}{\Delta p} \cdot \frac{p}{x} \text{ at a price } p$$

$$= \frac{dx}{dp} \cdot \frac{p}{x} \left[\text{as } \lim_{\Delta p \to 0} \frac{\Delta x}{\Delta p} = \frac{dx}{dp} \right]$$

If a demand function is given and price elasticity of demand is required at a price, it is given by

$$E_p = \frac{dx}{dp} \cdot \frac{p}{x}$$

where $\frac{dx}{dp}$ is the derivative of demand function with respect to p, p is the price at which elasticity is required, and x is the quantity demanded at that price.

Example : The demand function of a commodity is $x = \sqrt{100 - p^2}$. Determine the price elasticity of demand (i) when price decreases from Rs. 8 to Rs. 6, (ii) at price Rs. 8.

Solution : (i) Given $\sqrt{100 - p^2}$

$\therefore$ when p = Rs. 8 $x = 6$ units

and when p = Rs. 6 $x = 8$ units

Therefore Δp = Rs. 2 $\Delta x = -2$

$$\therefore E_p = \frac{\Delta x}{\Delta p} \cdot \frac{p}{x} = \frac{-2}{2} \cdot \frac{8}{6} = \frac{-4}{3}$$

(ii) At p = Rs. 8, $x = 6$ units.

$$E_p = \frac{dx}{dp} \cdot \frac{p}{x} = \frac{d\sqrt{100 - p^2}}{dp} \cdot \frac{p}{x}$$

$$= \frac{-p}{\sqrt{100 - p^2}} \cdot \frac{p}{x}$$

$$= \frac{-p^2}{x^3} \left[\sqrt{100 - p^2} = x \text{ Given} \right]$$

$$= -\frac{64}{36} = -\frac{16}{9}$$

Example. The linear demand function of a commodity shows that qunatity demanded is nil when price is Rs. 50 per unit and it is just 1000 units when price is zero. Determine the price and quantity where price elasticity and demand is unit.

Solution : Linear demand function can be obtained using intercept form as intercept on x-axis is 1000 units and on y-axis is Rs. 50 [given].

$$\frac{x}{1000} + \frac{P}{50} = 1$$

or $x + 20P = 1000$

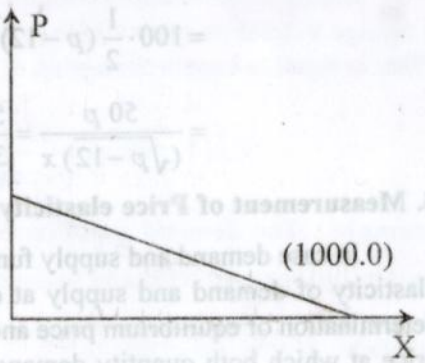
or $x = 1000 - 20P$. .....(i)

Corresponding $E_p = \frac{dx}{dp}\frac{P}{x} = \frac{-20P}{x}$

$$\frac{-20P}{x} = -1$$

[unit means -1 for E_p]

or $x = 20P$ ...(*ii*)



Solving (i) and (ii) simultaneously for P

$$1000 - 20P = 20P$$

or $40P = 1000$

$\therefore$ P = Rs. 25

Corresponding $x = 20(25) = 500$ unit.

So it has unit elasticity when P = Rs. 25 and $x = 500$ units.

2. Measurement of Price elasticity of supply

Price elastictity of supply at a price can be measured with the help of simple differentiation in the same way as price elasticity of demand.

This concept is based on the law of supply which states the positive relationship between price and quantity supplied (from firm's point of view). It is defined for a supply function : $x = g(P)$ as

$$E_p = \frac{\text{Relative change in quantities supplied}}{\text{Relative change in price}}$$

$$= \frac{\frac{\Delta x}{x}}{\frac{\Delta p}{p}} = \frac{\Delta x}{\Delta p}\cdot\frac{p}{x}$$ where Δp is change in price.

At given price, price elasticity of supply is given by :

$$E_p = \lim_{\Delta p \to 0} \frac{\Delta x}{\Delta p}\cdot\frac{p}{x} = \frac{dx}{dp}\cdot\frac{p}{x}$$

where $\frac{dx}{dp}$ is the derivative of supply function with respect to p. It always takes positive value because of positive relationship between price and quantity supplied.

Example. The weekly supply of butter (x kg) depends on the price (Rs. p per kg) according to

$$x = 100\sqrt{p-12} + 150$$

Determine the price elasticity of supply at p = Rs. 21.

Solution : Given $x = 100\sqrt{p-12} + 150$ $\therefore$ when p = Rs. 21, $x = 450$ units.

$$\therefore \quad E_p = \frac{dx}{dp}\cdot\frac{p}{x}$$

$$= 100 \cdot \frac{1}{2}(p-12)^{-\frac{1}{2}} \cdot \frac{p}{x}$$

$$= \frac{50\,p}{(\sqrt{p-12})\,x} = \frac{50\,(21)}{3\,(450)} = \frac{7}{9}$$

3. Measurement of Price elasticity of demand and supply at equilibrium price

In case demand and supply functions for a commodity are given, it is possible to measure both price elasticity of demand and supply at equilibrium price using simple differentiation. The first step is the determination of equilibrium price and quantity using demand and supply function. Equilibrium price is the price at which both quantity demanded and supplied are equal and corresponding quantity is known as equilibrium quantity, so

$X = f(p)$ (demand function)

and $X = g(p)$ (Supply function)

Then put $f(p) = g(p)$ and some for p, this gives equilibrium price putting this in either demand function or supply function and solving for x gives us equilibrium quantity.

At this equilibrium price, price elasticity of demand is obtained by using demand function and price elasticity of supply is obtained by using supply function. The only difference in this application and the previous two is that these elasticities are calculated at equilibrium price only whereas the earlier two were calculated at any arbitrary price (given).

Example : A firm has the following demand and supply functions

$$P = \sqrt{100 - X^2} \quad \text{(Demand)}$$

$$X = -10 + 2P \quad \text{(Supply)}$$

Determine the price elasticity of demand and supply at equilibrium price.

Solution : Given $P = \sqrt{100 - X^2}$ or $X = \sqrt{100 - P^2}$

or $P^2 = 100 - X^2$

For equilibrium condition , Put $X = -10 + 2P$ in the demand function

$$\therefore \quad P^2 = 100 - [-10 + 2P]^2$$

$$P^2 = 100 - 100 + 40P - 4P^2$$

or $5P^2 - 40P = 0$

$P - 8 = 0$ (equilibrium price) $\therefore$ P = 8

Corresponding equilibrium quantity

$X = -10 + 2\,(8) = 6$ units

So price elasticity of demand at P = Rs. 8 with $X = 6$ units will be :

$$E_p = \frac{dX}{dP} \cdot \frac{P}{X} = \frac{d\sqrt{100 - P^2}}{dP} \cdot \frac{P}{X}$$

$$= \frac{1}{2}\left(100 - P^2\right)^{-1/2} \cdot -2P \cdot \frac{P}{X}$$

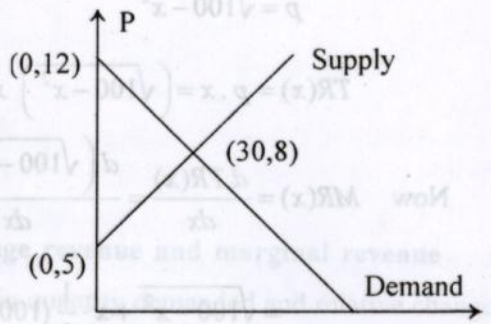
$$= \frac{-P^2}{\left(\sqrt{100-P^2}\right)X} = \frac{-P^2}{X}$$

$$= \frac{-64}{36} = \frac{-16}{9}$$

$$E_p = \frac{dX}{dP} \cdot \frac{P}{X} = \frac{d(-10+2P)}{dP} \cdot \frac{P}{X}$$

$$= \frac{2P}{X} = \frac{2(8)}{6} = \frac{8}{3}$$

Example : A firm finds that both quantity demanded and supplied are 30 units when market price is Rs. 8 per unit. It is known that demand is nil if price is increased to Rs. 12 per unit and firm is not willing to supply if price is reduced to Rs. 5 per unit. Determine the firm's price elasticity of demand and supply at equilibrium price assuming linear relationship between price and quantity in both cases.



Solution : The demand equation being straight line, so its equation can be obtained as :

$$\therefore \frac{P-12}{12-8} = \frac{X-0}{0-30}$$

or $$-30\,(P-12) = 4X$$

or $$x = 90 - \frac{15}{2}P \quad \text{(Demand)}$$

Similarly supply function will be

$$\frac{P-5}{5-8} = \frac{X-0}{0-30}$$

or $$-30\,(P-5) = -3x$$

or $$X = 10P + 50 \quad \text{(supply)}$$

Price elasticity of demand and supply at equilibrium price, P = Rs. 8 will be :

$$E_p = \frac{dx}{dP} \cdot \frac{P}{X} = \frac{-15}{2} \cdot \frac{P}{X} = \frac{-15\,(8)}{2\,(30)} = -2$$

$$E_s = \frac{dx}{dP} \cdot \frac{P}{X} = \frac{10P}{X} = \frac{10\,(8)}{30} = \frac{8}{3}$$

4. Derivation of marginal revenue function [*MR(x)*] from total revenue function *TR(x)*

Total revenue function, *TR(x)*, gives the total revenue of x units sold. It can be obtained from the demand function as $TR(x)$=P.x where p is the price when x units are demanded. On the other hand, marginal revenue function, *MR(x)* gives the additional revenue of xth unit. In economics, marginal revenue is defined as change in total revenue caused by change in number of units sold as additional revenue of xth unit = $TR(x) - TR(x-1)$. It has an important relationship with total revenue as total revenue will increase till marginal is positive, it will be zero if marginal is zero and decrease if marginal is negative.

*MR(x)*can be obtained with the help of simple differentiation if *TR(x)* is given. *MR(x)* gives the rate of change in total revenue which in turn is also given by derivative of total revenue function. Therefore, *MR(x)* is equal to simple derivation of *TR(x)* :

$$\therefore \quad MR(x) = \frac{dTR(x)}{dx}$$

Example : The demand function of a commodity is $x = \sqrt{100 - p^2}$. Determine the corresponding marginal revenue function and find marginal revenue when $x = 6$ units.

Solution : Given $x = \sqrt{100 - p^2}$

or $p = \sqrt{100 - x^2}$ (corresponding inverse function)

$$TR(x) = p \,.\, x = \left(\sqrt{100 - x^2}\right) x$$

Now $$MR(x) = \frac{d\,TR(x)}{dx} = \frac{d\left(\sqrt{100 - x^2}\right) x}{dx}$$

$$= \sqrt{100 - x^2} + x \cdot \frac{1}{2}(100 - x^2)^{-1/2} \cdot -2x$$

$$= \sqrt{100 - x^2} - \frac{x^2}{\sqrt{100 - x^2}}$$

$$MR(x) = \frac{100 - 2x^2}{\sqrt{100 - x^2}}$$

$$MR(6) = \frac{100 - 2(6)^2}{\sqrt{100 - (6)^2}} = \frac{28}{8} = \frac{7}{2} = Rs.\,3.50.$$

Example. A firm can sell 100 units if it charges Rs. 5 per unit. It knows that for each Rs. 0.10 reduction in price, it can sell 20 more units. Determine the firm's marginal revenue function. Use it to determine marginal revenue when $x = 100$ units.

Solution : Given P = Rs. 5,

Quantity demanded, $x = 100$ units

$\therefore$ when P = Rs. 4,

quantity demanded, $x = 300$ units.

(as for each Rs. 0.10 reduction, quantity demanded increases by 20 units, so if price declines 10 times to Rs. 4, quantity demanded also increases 10 times to 300 units).

Using two point form :

$$\frac{P - 5}{5 - 4} = \frac{X - 100}{100 - 300}$$

or $$-200\,(P - 5) = X - 100$$

or $x = -200P + 1100$

or $P = \frac{11}{2} - \frac{X}{200}$

So corresponding, $TR(x) = P \,.\, X = \left(\frac{11}{2} - \frac{X}{200}\right) X$

$$= \frac{11}{2} X - \frac{X^2}{200}$$

$$\therefore \quad MR(X) = \frac{dTR(X)}{dx} = \frac{d\left[\frac{11}{2} X - \frac{X^2}{200}\right]}{dx}$$

$$= \frac{11}{2} - \frac{X}{100}$$

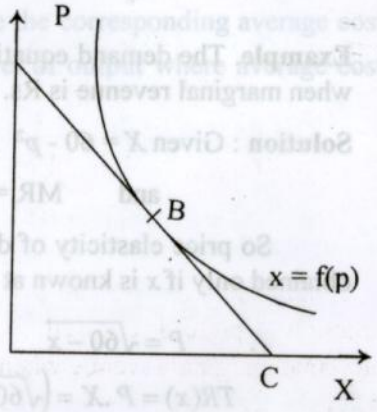
When $X = 100$ units

$$MR(100) = \frac{11}{2} - \frac{100}{100} = \frac{9}{2} = Rs.\ 4.50$$

5. Measurement of Price elasticity in terms of average revenue and marginal revenue

Price elasticity of demand is the ratio of relative change in quantity demanded and relative change in price. In economics there are three important methods of its measurement, viz. (i) total expenditure method, (ii) point method, (iii) arc method.

According to point method, price elasticity of demand at a point on the demand curve is equal to lower half of the tangent at that point divided by upper half of his tangent, price elasticity of demand at B is equal to $\frac{BC}{AB}$.



The same result, if further analysed using relationship between average revenue and marginal revenue gives elasticity of demand :

$$E_p = \frac{AR}{AR - MR}$$

$$or \quad MR = AR \left[1 - \frac{1}{E_p}\right]$$

The same result can be provided using simple differentiation.

$$TR(X) = P \,.\, X$$

$$\frac{dTR(X)}{dX} = P \,.\, \frac{dX}{dX} + X \,.\, \frac{dP}{dX} \qquad \text{[Using product form]}$$

$$MR(x) = P + X \cdot \frac{dP}{dX}$$

$$= P + X \cdot \frac{p}{X \cdot E_p} \qquad \left[E_p = \frac{dX}{dp} \cdot \frac{P}{X} \; or \; \frac{X \cdot E_p}{P} = \frac{dX}{dP}\right.$$

$$MR = P\left[1+\frac{1}{E_p}\right] \qquad \left[\therefore \frac{dP}{dX} = \frac{p}{X \cdot E_p}\right]$$

$$= AR\left[1+\frac{1}{E_p}\right] \qquad [P = AR]$$

$$MR = AR\left[1-\frac{1}{E_p}\right] \qquad [\because E_p \text{ is always negative}]$$

$$or \quad E_p = \frac{AR}{AR-MR}$$

Alternatively, $MR = P + X = \frac{dP}{dP}$

$$= P + P \cdot \frac{X}{P} \cdot \frac{dP}{dX} \quad \text{[dividing and multiply second form by } p\text{]}$$

$$= \left[P + P \cdot \frac{1}{E_p}\right]$$

$$MR = AR\left[1-\frac{1}{E_p}\right] \qquad \left[as \frac{X}{P} \cdot \frac{dP}{dX} = \frac{1}{E_p}\right]$$

Example. The demand equation of a commodity is $x = 60 - P^2$. Determine the price elasticity of demand when marginal revenue is Rs. 4.00

Solution : Given $X = 60 - p^2$

and MR = Rs. 4.00

So price elasticity of demand requires average revenue when marginal revenue is Rs. 4. It can be obtained only if x is known at which MR = Rs. 4.00

$$\therefore \quad P = \sqrt{60-x}$$

$$TR(x) = P.X = \left(\sqrt{60-x}\right) x$$

$$MR(x) = \frac{dTR(x)}{dx}$$

$$= \sqrt{60-x} + \frac{x(-1)}{2\sqrt{60-x}}$$

$$= \frac{2(60-X)-X}{2.\sqrt{60-X}}$$

$$MR(x) = \frac{2P^2 - X}{2P} = 4 \quad \text{(Given)}$$

or $\quad 8P = 2P^2 - X$

or $\quad x = 2P^2 - 8P$

Using (i) and (ii) equations

$$60 - P^2 = 2P^2 - 8P$$

or $$3P^2 - 8P - 60 = 0$$

$$3P^2 - 18P + 10P - 60 = 0$$

$$3P[P-6] + 10[P-6] = 0$$

$$(3P + 10)(P - 6) = 0$$

So either $P = \frac{-10}{3}$ (not possible, therefore it is ignored)

or P = Rs. 6

or AR = Rs. 6 when MR = Rs. 4

$$\therefore \quad E_p = \frac{AR}{AR - MR} = \frac{6}{6-4} = 3$$

6. Derivation of marginal cost function from total cost function :

Total cost function gives the functional relationship between total cost and number of units produced. Marginal cost function gives the additional cost of last unit. It can be obtained from total cost function with the help of simple differentiation.

$$MC(X) = \frac{dTC(X)}{dX}$$

Example. A firm's total cost function is $TC(X) = X^3 + 2000$. Determine the corresponding average cost function, A(X) marginal cost function, M(X) and also determine the level of output where average cost equals marginal cost.

Solution : Given $TC(X) = X^3 + 2000$

So $$AC(X) = \frac{TC(X)}{X} = X^2 + \frac{2000}{X}$$

and $$MC(X) = \frac{dTC(X)}{dX} = 3X^2$$

(Putting both equal and solving for X

$$\frac{X^2 + 2000}{X} = 3X^2$$

or $$2X^3 = 2000$$

$$X^3 = 1000$$

$$X = 10 \text{ units.}$$

MAXIMA AND MINIMA

In calculas, extreme values of a function are of two types, viz., relatively extreme and absolutely extreme. Relatively extreme values are also known as local extreme values. They are extreme in relation to surounding values of a function. Some are described as maxima and minima. Absolutely extreme values are extreme in the entire range of the function. These values are also known as global extreme values of maximum and minimum. These may be more than one relatively extreme values of a function but there is only one absolutely extreme value of a function in the figure.

f(x) has relatively maximum values at *B* and *D* and relatively minimum at point *A* and *C*. But it has absolutely maximum at *B* and absolutely minimum at *C*.

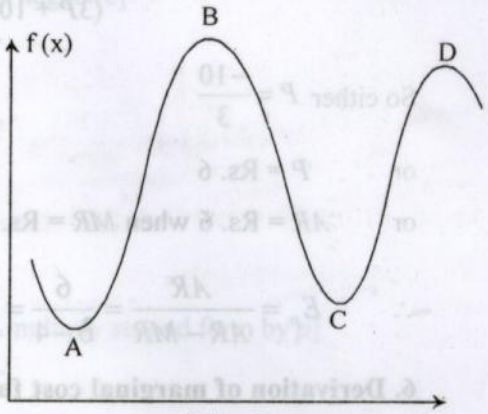


Fig. 1

Calculas gives only relatively extreme values.If a profit function is given or average cost function is given, objective is to determine the level of output where cost per unit is minimum.

Absolutely extreme values can be obtained from relatively extreme values by simple comparison. If a function has two relatively maximum values, then absolutely maximum can be obtained by comparing both relatively maximum values of a function. But if the function has only one relatively extreme value it becomes absolutely extreme also.

The relatively extreme values can be obtained using calculas in the same way as explained in the determination of relatively extreme values under Successive Differentiation.

(*i*) *If* $y = f(x)$, then put $\frac{dy}{dx} = 0$ and solve it for x.

(*ii*) At this values of x, *if* $\frac{d^2y}{dx^2} > 0$, y takes relatively minimum value,

if $\frac{d^2y}{dx^2} = 0$, y becomes point of inflection, and

if $\frac{d^2y}{dx^2} < 0$, y takes relatively maximum value.

Example. If $f(x) = ax^2 + bx + c$, $a \neq 0$, prove that the point

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

is a relatively maximum if a < 0 and is a relatively minimum if $a > 0$.

Solution : Given $f(x) = ax^2 + b + c$

$$\therefore \quad \frac{df(x)}{dx} = 2ax + b$$

Putting it equal to zero and solving it for x for relatively extreme value

$$\therefore \quad 2ax + b = 0$$

$$x = -b/2a$$

At $x = \frac{-b}{2a}$

$\frac{d^2 f(x)}{dx^2} = 2a$

It will be positive if a > 0, so taking relatively minimum value of

$$f\left(\frac{-b}{2b}\right) = a\left[\frac{-b}{2a}\right]^2 + b\left[\frac{-b}{2a}\right] + c = \frac{4ac - b^2}{2a}$$

It will be negative if $a < 0$, so taking relatively minimum value of

$$\frac{4ac - b^2}{2a}$$

Applications in business and economics

There are a number of important applications in business and economics based on maximum and minima. Some of these involves maximisation, others minimisation depending on the nature of the function. Some of the important applications are explained below :

1. Maximisation of total revenue function :

If a firm's demand function is given, it is possible to determine the corresponding level of output whose total revenue is maximum using maxima and minima. But it will be possible only if firm is operating under imperfect competition (where price decreases as firm sells more units). In case of perfect competition, price is constant for a firm resulting in total revenue as an increasing function. The total revenue is maximum when marginal revenue is zero.

(i) Put $\frac{dTR(x)}{dx} = 0$ and solve for x.

(ii) The value of x, where $\frac{d^2 TR(x)}{dx^2} < 0$,

$TR(x)$ is maximum

Example. The demand function of a commodity is $p = 15e^{-x/3}$ for $0 \le x \le 8$ where Rs. p is the price per unit when x units are demanded. Determine the price and quantity for which total revenue is maximum.

Solution : Given $p = 15e^{-x/3}$

$$\therefore TR(x) = p \,.\, x = \left(15e^{-x/3}\right)x$$

$$= 15xe^{-x/3}$$

and $$\frac{dTR(x)}{dx} = 15\left[x \,.\, e^{-x/3} \,.\, \frac{-1}{3} + e^{-x/3} + 1\right]$$

$$= 15e^{-x/3}\left[\frac{-x}{3} + 1\right]$$

For maximisation, put it equal to zero and solve for x

$$\therefore 15e^{x/3}\left[\frac{-x}{3} + 1\right] = 0$$

$\therefore$ either $15e^{-x/3} = 0$ or $\left[\frac{-x}{3}+1\right]=0$

$\frac{-x}{3}+1=0$ $\quad [15e^{-x/3} = 0$, not possible as it is equal to $p]$

$\therefore \quad x = 3$ units

$$\text{At } x=3, \frac{d^2TR(x)}{dx^2} = 15\left[e^{-x/3}.\frac{-1}{3}+\left(\frac{-x}{3}+1\right)e^{-x/3}.\frac{-1}{3}\right]$$

$$= -5e^{x/3}\left[1+\frac{x}{3}+1\right]$$

$$= -5e^{-1} < 0.$$

So $TR(x)$ is maximum when $x = 3$ units

At this level, $p = 15e^{-1}$

Example. A shoe manufacturer produces both high grade and low grade shoes in x units (hundred pairs) and y unit (hundred pairs) per day respectively. Assume that the equipment is capable of producing the shoes in the following relationship :

$$y = \frac{42-5x}{10-x} \quad (x<10)$$

if the high grade shoes are sold at a price twice the low grade shoes. Determine the pairs of both the shoes he should produce to maximize his total revenue.

Solution : Let the price of low grade shoes be Rs. p per pair, then price of high grade shoes will be Rs. 2 p per pair. Accordingly total revenue will be :

$$TR = 2px + py$$

or $$TR(x) = 2px + p\left[\frac{42-5x}{10-x}\right] \qquad \left[\text{Putting } y = \frac{42-5x}{10-x}\right]$$

For maximizing $TR(x)$, take its first derivative w.r.t. x, put it equal to zero and solve for x

$$\frac{dTR(x)}{dx} = 2p + p\left[\frac{(10-x).-5-(42-5x).-1}{(10-x)^2}\right]$$

$$= 2p + p\left[\frac{-8}{(10-x)^2}\right]$$

Putting it equal to zero and solving for x

$$2p + p\left[\frac{-8}{(10-x)^2}\right] = 0$$

$$\frac{8}{(10-x)^2} = 2 \qquad \text{[Cancelling } p \text{ from both sides]}$$

or $$(10 - x)^2 = 4$$

$\therefore$ $10 - x = +2$ or $10 - x = -2$

or $x = 8$ or 12

$\therefore$ $x = 8$ [$x < 10$, and $x = 12$ not possible]

At $x = 8$ $\frac{d^2TR(x)}{dx^2} = \frac{-16}{(10-x)^3}$ which is negative,

$RT(x)$ is maximum when $x = 8$, putting this in the function $y = \frac{42 - 5(8)}{10 - 8} = 1$

Therefore, total revenue is maximum when the manufacturer produces 800 pairs of high grade and 100 pairs of low grade.

2. Minimisation of Average cost function

Maxima and minima can also be used for determining the level of output where per unit cost is minimum corresponding to given total or average cost function. It is often used in economic analysis. It gives the optimum level of output.

(i) Put $\frac{dA(x)}{dx} = 0$ and solve it for x (ii) the value of x for which $\frac{d^2AC(x)}{dx^2} > 0$, per unit cost is minimum corresponding average cost can be obtained by puting this x in $AC(x)$.

Example. A machine initially costs Rs. 6,400 with no scrap value. The cost of operating is Rs. 500 in the first year and increases to Rs. 800 in each successive year. Determine (i) the number of years it should be operated for minimising total operating cost per year, and (ii) corresponding cost per year.

Solution. Capital cost = Rs. 6400

Operating cost = Rs.500 + Rs. 1300 + Rs. 2100.....upto x years.

$= \frac{x}{2}[2(500) + (x - 1)\,800]$ [Second x term of an A.P.]

$= \frac{x}{2}\,[800 + 200]$

$= 400x^2 + 100x$

$\therefore$ $TC(x)$ = operating cost + capital cost

$= 400x^2 + 100x + 6{,}400.$

Corresponding per year cost function will be :

$$AC(x) = \frac{TC(x)}{x} = 400x + 100 + \frac{6400}{x}$$

For minimising it, take its derivative w.r.t. x

$\therefore$ $\frac{dAC(x)}{dx} = 400 - \frac{6400}{x^2}$

Putting it equal to cost and solving it for x.

$$400 - \frac{6400}{x^2} = 0$$

or $x = 4$ years

At $x = 4$ years, $\frac{d^2AC(x)}{dx^2} = \frac{12800}{x^3} > 0$, therefore cost per year is minimum if machine is operated for

4 years. Then $AC(x) = 400(4) + 100 + \frac{6400}{4}$ = Rs. 3300.

3. Maximixation of profit function

If a firm's demand function with total cost function is given, it is possible to determine the firm's equilibrium by maximizing the corresponding profit function. It is taken as profit, function $\pi(x) = TR(x) - TC(x)$.

For maximizing (i) put $\frac{d\pi(x)}{dx} = 0$ and solve for x, (ii) the value of x for which $\frac{d^2\pi(x)}{dx^2} < 0$, profit is maximum corresponding profit can be obtained by putting in $\pi(x)$.

Example. The total cost function of a firm is

$$TC(x) = \frac{1}{3}x^2 - 5x^2 + 28x + 10$$

where $TC(x)$ is the total cost of x units. A tax at the rate of Rs. 2 per unit of output is imposed and the producer adds it to his cost. If the market demand function is given by

$$P = 2530 - 5x$$

where P is the price per unit of output, determine the profit maximising price and output.

Solution : Given $P = 2530 - 5x$

$$TC(x) = \frac{1}{3}x^3 - 5x^2 + 28x + 10$$

corresponding $TR(x) = P.\ x,$

After tax $TC(x) = \frac{1}{3}x^2 - 5x^2 + 28x + 10 + 2x = \frac{1}{3}x^2 - 5x^2 + 30x + 10$

$= (2530 - 5x)\ x$

$= 2530x - 5x^2$

$\pi(x) = TR(x) - TC(x)$

$$= 2530x - 5x^2 - \frac{1}{3}x^3 + 5x^2 - 30x - 10$$

$$= -\frac{1}{3}x^3 + 2500x - 10$$

For maximising $\pi(x)$, take its derivative w.r.t. x

$$\frac{d\pi(x)}{dx} = -x^2 + 2500$$

Putting it equal to zero and solving for x

$$-x^2 + 2500 = 0$$

or $x = 50$ unit.

At $x = 50$ units, $\frac{d^2\pi(x)}{dx^2} = -2x < 0$

$\therefore$ profit is maximum when $x = 50$ units,

Corresponding $P = 2530 - 5\,(50)$

$= \text{Rs. } 2280.$

Example. The relationship between sales, *TR(x)* and advertising cost, Rs. x is given by

$$TR(x) = \frac{32000x}{500+x}$$

It is known that gross profit is 25 per cent of the sales. Determine (i) the corresponding net profit as a function of x and $\pi(x)$ (ii) the value of x which maximises $\pi(x)$ and (iii) maximized $\pi(x)$.

Solution : (i) Given $$TR(x) = \frac{32000x}{500+x}$$

$\therefore$ gross profit $$= \frac{1}{4}\left[\frac{32000x}{500+x}\right]$$

$$= \frac{8000x}{500+x}$$

$\therefore \pi(x) = \text{Gross profit} - x \text{ (advertisement cost)}$

$$= \frac{8000x}{500+x} - x$$

(ii) Form maximizing it, take its first derivative w.r.t. x,

$$\frac{d\pi(x)}{dx} = \frac{(500+x)\,8000 - 8000x}{(500+x)^2} - 1$$

$$= \frac{8000\ (500)}{(500+x)^2} - 1$$

Putting it equal to zero and solve for x

$$= \frac{8000\ (500)}{(500+x)^2} - 1 = 0$$

$$(500 + x)^2 = 8000\ (500)$$

$$500 + x = 2000$$

$$x = \text{Rs. } 1500$$

At $x = Rs.1500$, $\dfrac{d^2\pi(x)}{dx^2} = \dfrac{-8000000}{(500+x)^2} < 0$

So $\pi(x)$ is maximum when $x = \text{Rs. } 1500.$

(iii) Corresponding profit, $\pi\,(1500) = \dfrac{8000\ (1500)}{500+1500} - 1500$

$= \text{Rs. } 4500.$

4. Minimisation of Total Cost

Usually, total cost function gives the functional relationship between total cost and the number of units produced. It means it is an increasing function, so minimisation of this is not possible. But sometimes, the total cost is defined in such a way that it decreases upto certain level and then increases. Only in such cases, this application is relevant. The steps involved for minimisation are as follows :

(i) Put $\frac{dTC(x)}{dx} = 0$ and solve for x

(ii) The value of x for which $\frac{d^2TC(x)}{dx^2} > 0$

$TC(x)$ is minimum.

Example. A cylinder with a capacity of $\frac{440}{7}$ cubic ft. is made of two types of material. The cost of material of the top and bottom is Rs. 6 per sq. ft and the sides material cost Rs. 4.80 per sq. ft. Determine the dimensions of the cylinder which minimizes total cost and minimized total cost.

Solution : Given Volume $= \frac{440}{7}$ cubic ft.

$$\text{So } \pi r^2 h = \frac{440}{7} \text{ cubic ft.}$$

$$h = \frac{440}{7\pi r^2}$$

$$= \frac{440}{7} \times \frac{7}{22r^2}$$

$$= \frac{20}{r^2}$$

Corresponding cost of material of top and bottom

= [Total area of top and bottom] x Rate

$= (2\pi r^2)6 = \text{Rs.}\, 12\pi r^2$

Cost of sides = [Total area of sides] × Rate

$= (2\pi rh)$. Rs.4.80

$= \text{Rs.}\, 9.60\pi rh.$

$\therefore$ Total cost of material $= 12\pi r^2 + 9.60\,\pi rh$

$$TC(r) = 12\pi r^2 + 9.60\pi \cdot \left[\frac{20}{r^2}\right]$$

$$= 12\pi r^2 + \frac{192\pi}{r}.$$

For minimising, $TC(r)$, take its first derivative w.r.t. r

$$\frac{dTC(r)}{dr} = 24\pi r - \frac{192\pi}{r^2}$$

Putting it equal to zero and solving it for r

$$24\pi r - \frac{192\pi}{r^2} = 0$$

$$r^2 = \frac{192\pi}{24\pi} = 8$$

$$r = 2 \text{ ft.}$$

At $r = 2$ft.

$$\frac{d^2TC(r)}{dr^2} = 24\pi + \frac{384\pi}{r^3}$$

So at $r = 2$ ft, $TC(x)$ is minimum.

Corresponding $h = \frac{20}{(2)^2} = 5 ft.$

and minimised $TC\ (2) = 48\pi + 96\pi$

$= Rs.144\pi$

5. Maximisation of tax Revenue

The theory of maxima and minima can also be used for maximizing tax revenue.The method used is same as in case of optimizing other functions as total revenue function, total cost function etc. The only difference is that the level of output sold is a function of the rate of tax. Therefore total tax revenue is equal to rate of tax multiplied by units sold.

The units sold is obtained considering the production tax imposed at the rate of Rs. t per unit. It will be different in case of perfect competition and monopoly. In the case of perfect competition, output sold is given by equilibrium price whereas in the second case, it is same as profit maximizing output given by

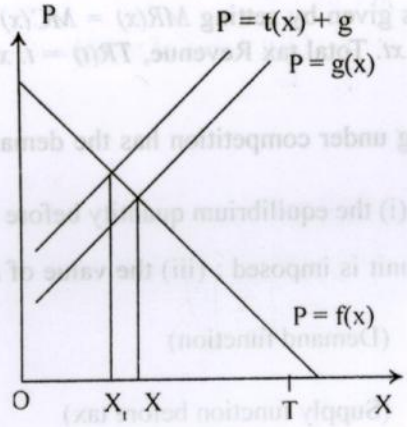


Figure 1

equality of marginal revenue and marginal cost.

Therefore in the case of perfect competition, the tax imposed is added in the supply function which will shift the supply function upward and increase the equilibrium price to some output (depending on the elasticity of demand) and reduces the equilibrium quantity as shown in figure 1.

$\therefore$ if $P = f(x)$ (Demand function)
and $P = g(x)$ (Supply function before tax)
then $P = g(x) + t$ (Supply function after tax)
$\therefore$ Equilibrium output is given by putting
$f(x) = g(x) + t$ and solving it for x. It will give x as a function of tax rate t.
Total tax revenue = $t \,.\, xt$.

Therefore for maximizing tax take its derivative w.r.t. t, put it equal to zero and solve for t. (ii) the value of t for which $\frac{d^2TR(t)}{dt^2} < 0$; tax revenue is maximum.

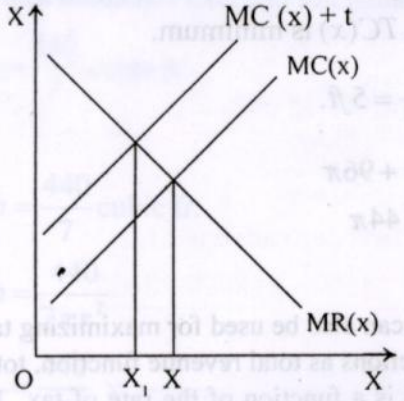


Fig . 2

In the case of monopoly, the firm adds it to the marginal cost function as shown in Figure 2.

The profit maximizing output is given by setting $MR(x) = MC(x) + t$ and solving it for x. At this profit must be minimum. It is same as xt. Total tax Revenue, $TR(t) = t.\, x$ which can be maximized in the same way as explained above.

Example. An industry operating under competition has the demand function $P = 10 - x$ with the supply function $P = 4 + \frac{x}{2}$. Determine (i) the equilibrium quantity before tax; (ii) the equilibrium quantity after tax if tax at the rate of Rs. t per unit is imposed ; (iii) the value of t which maximizes tax revenue; (iv) maximized tax revenue.

Solution : (i) Given $P = 10 - x$ (Demand function)

and $P = 4 + \frac{x}{2}$ (Supply function before tax)

$\therefore$ Equilibrium quantity will be

$$10 - x = 4 + \frac{x}{2}$$

or $x = 4$ units

Corresponding equilibrium price will be, P = 10 - 4 = Rs. 6.

(ii) After tax supply function becomes $P = 4 + \frac{x}{2} + t$. So new equilibrium after tax will be

$$10 - x = 4 + \frac{x}{2} + t$$

or $xt = 4 - \frac{2}{3}t.$

(iii) if $xt = 4 - \frac{2}{3}t$, total tax revenue will be

Tax Revenue $(t) = t\left(4 - \frac{2}{3}t\right)$

For maximizing it, take its first derivative w.r.t. t.

$$\therefore \quad \frac{d \text{ tax Revenue}(t)}{dt} = 4 - \frac{4}{3}t$$

Putting it equal to zero and solving for t.

$$4 - \frac{4}{3}t = 0$$

or t = Rs. 3 per unit.

At t = Rs. 3 per unit.

$$\frac{d^2 \text{ tax Revenue}(t)}{dt^2} = -\frac{4}{3} < 0$$

So tax revenue is maximum when t = Rs. 5 per unit.

(iv) Maximized tax revenue = $t \cdot x_t$

$$= 3\left(4 - \frac{2}{3}t\right)$$

= Rs. 6

LESSON 3

INDEFINITE INTEGRATION

Calculus is an important branch of mathematics. It is broadly classified as differential and integral calculus. Differential calculus gives the derivative of a function used as an instantaneous rate of change. Integral calculas, on the other hand, deals with the total rate of change.

Integral calculus is reverse of the process of differentiation (also known as antidifferentiation) had its own applications in business and economics. Two important applications are; obtaining the original function from its derivative function and the measurement of area under curve.

The integral of a function is denoted by sign $\int$ (known as integral). Integration is also related to a variable just like differentiation. So function is always followed by *dx* or *du* or any variable in relation to which function is integrated. Integration of a function can be technically defined as reverse of differentiation.

Let $y = f(x)$ and its derivative, $\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x}$

Putting it equal to $f(x)$,

$$\frac{dy}{dx} = f(x)$$

Then $\int f(x).dx = f(x) + c.$

Left hand side denotes the integral of $f'(x)$ with respect to x and right hand side states it as equal to $f(x) + c$, where $f(x)$ is the original function and c is an arbitrary constant. The same is known as integration and the result obtained is described as indefinite integral of $f'(x)$ with respect to x. Indefinite integral is used for obtaining the original function from its derivative.

Integration is reverse of the process of differentiation. In case of differentiation, the power of variable is multiplied to the base and the power is reduced by one for algebraic function. So integral process involves increasing the power of the variable by one and dividing by the increased power. For example,

if $y = x^2 + 2x - 5,$

$$\frac{dy}{dx} = 2x + 2 \qquad \text{or} \qquad dy = (2x+2)dx$$

Then $\int (2x+2)dx = \frac{2x^2}{2} + 2x + c$

$$= x^2 + 2x + c.$$

where c is an arbitrary constant. This provides for that constant which has been ommitted by differentiation. The value of c can be obtained under conditions as in the above example if y value is given for certain value of x.

Different Methods of Integration

There are different methods of integration used in different conditions. Some of these are as discussed below :

(a) **Simple integration.** It just involves the reversing process of differentiation for obtaining integral.

(i) *Algebraic function* : If the function to be integrated is an algebraic function, the integral can be obtained by increasing the power of variable by one and dividing the varible by increased power. Some standard forms of simple integration are as follows :

1. If $dy = f'(x)\,dx$, $\int f(x).dx = f(x) + c$

2. If $dy = a\,[f'(x)dx]$ where a is a constant,

$$\int a\,xf'(x)\,dx = a\int f'(x)dx = af(x) + c.$$

3. If $dy = f_1'(x) + f_2'(x)]\,dx$

$$\int\left(f_1'(x) + f_2'(x)\right)dx = \int(f_1'(x)\,dx + \int(f_2'(x)\,dx$$
$$= f_1(x) + f_2(x) + c.$$

4. If $dy = x^n\,.dx$, $\int x^n\,.dx = \dfrac{x^{n+1}}{n+1} + c.$

Example. Find the integral of the following :

(i) $\int x dx$ (ii) $\int dx$ (iii) $\int \sqrt[3]{x^2}\,.dx.$

(iv) $\int\left(\dfrac{1}{x^4} + \dfrac{1}{\sqrt[4]{x}}\right)dx$ (v) $\int (x-2)^2\,dx$

Solution.

(i) $\int x \cdot dx = \dfrac{x^2}{2} + c$

(ii) $\int dx = \int 1 \cdot dx = x + c$

(iii) $\int \sqrt[3]{x^2} \cdot dx = \int x^{2/3} dx = \dfrac{x^{\frac{2}{3}+1}}{\frac{2}{3}+1} + c.$

$$= \frac{3}{5}x^{5/3} + c$$

(iv) $\int\left(\dfrac{1}{x^4} + \dfrac{1}{\sqrt[4]{x}}\right)dx = \int\left(x^{-4} + x^{-1/4}\right)dx$

$$= \frac{x^{-3}}{-3} + \frac{x^{3/4}}{3/4} + c$$
$$= \frac{x^{-3}}{-3} + \frac{4}{3}x^{3/4} + c.$$

v) $\int (x-2)^2\,dx = \int (x^2 - 4x + 4)dx$

$$= \frac{x^3}{3} - 2x^2 + 4x + c.$$

(ii) *Logarithmic function :* If a logarithmic function where independent variable appears with a log ratio is given as $y = lnx$ (ln denotes natural log with the base e). Its derivative, $\frac{dy}{dx} = \frac{1}{x}$, therefore the integral of $\frac{1}{x}$ with respect to x is lnx or $\int \frac{1}{x} dx = lnx + c$.

Therefore, it can be generalised that whenever in a rational function, numerator is the derivative of the denominator or is equal to some constant mulitplied by numerator, the integral of the function is the natural log of denominator. For example ln $\int \frac{1}{x} dx$, 1 the numerator is the derivative of the x appearing in denominator; therefore, its integral is $lnx + c$.

Example. Evaluate the following :

(i) $\int \frac{4x+6}{x^3+3x+1} dx$ (ii) $\int \frac{x^2+2x}{x^3+3x^2+5} dx$ (iii) $\int \frac{x^3-x}{x^4-2x^2+2} dx$

Solution : (i) Let $y = x^2 + 3x + 1$.

$$\therefore \quad \frac{dy}{dx} = 2x + 3.$$

It shows that numerator is $2(2x + 3)$; so applying log form :

$$\int \frac{4x+6}{x^2+3x+1} \cdot dx = 2 \int \frac{(2x+3)}{x^2+3x+1} \cdot dx = 2ln\,(x^2+3x+1) + c.$$

(ii) Let $y = x^3 + 3x^2 + 5$, $\therefore \quad \frac{dy}{dx} = 3x^2 + 6x$

It shows that numerator is $\frac{1}{3}[3x^2 + 6x]$

Applying log form.

$$\int \frac{x^2+2x}{x^3+3x^2+5} \cdot dx = \frac{1}{3} \int \frac{3x^2+6x}{x^3+3x^2+5} \cdot dx$$

$$= \frac{1}{3} ln\,(x^3+3x^2+5) + c.$$

(iii) Let $y = x^4 - 2x^2 + 2$, $\therefore \quad \frac{dy}{dx} = 4x^3 - 4x$

It shows that numerator is $\frac{1}{4}[4x^3 - 4x]$

So applying log form :

$$\int \frac{x^3-x}{x^4-2x^2+2} \cdot dx = \frac{1}{4} \int \frac{4x^3-4x}{x^4-2x^2+2} dx$$

$$= \frac{1}{4} ln\,[x^4-2x^2+2] + c.$$

(iii) *Exponential function :* All those functions where independent variable appears as a power of constant base, are known as exponential function. If the base is equal to *e*, the function is known as natural exponential function as $y = e^x$. If derivative, $\frac{dy}{dx} = e^x$. So its integral, $\int e^x \, . \, dx = e^x + c.$

Similarly if $y = e^{bx}; \frac{dy}{dx} = be^{bx}$ and $\int be^{bx} \, dx = e^{bx} + c.$

Example. Evaluate the following :

(i) $\int abe^{bx} \cdot dx$ (ii) $\int 4e^{-x} \cdot dx$ (iii) $\int (2e^{3x} + e^{-x}) \, dx$

Solution : (i) Put $u = bx, \frac{du}{dx} = b$ or $dx = \frac{du}{b}$

Making this substitution:

$$\int abe^u \cdot \frac{du}{b} = \int ae^u du = a \int e^u \cdot du$$

$$= ae^u + c$$

$$= ae^{bx} + c.$$

(ii) Putting $u = -x, \frac{du}{dx} = -1$ or $dx = -du.$

Making this substitution :

$$\int 4e^{-x} \cdot dx = \int 4e^u (-du) = -4 \int e^u \cdot du$$

$$= -4e^u + c$$

$$= -4e^{-x} + c.$$

(iii) $\int (2e^{3x} + e^{-x}) dx = \int 2e^{3x} \cdot dx + \int e^{-x} \cdot dx$

$$= \frac{2e^{3x}}{3} - e^x + c.$$

(b) **Integration by substitution :** In case of composite functions, the integral can be obtained by making substitution. For example. $\int [f(x)]^n dx$ can be obtained by substituting $u = f(x)$, and dx $\frac{du}{f'(x)}$ $\left[\text{because if } u = f(x), \frac{du}{dx} = f'(x), \text{ or } dx = \frac{du}{f'(x)}\right]$

Therefore

$\int u^n \cdot \frac{du}{f'(x)} = \frac{1}{f'(x)} \left[\frac{u^{n+1}}{n+1}\right] + c \cdot f(x)$ must be a constant, otherwise it is not possible. The substitution is made in such a way that the integral can be simplified to the possible extent.

Example. Find the integral of the following :

(i) $\int \sqrt{3x+4} \cdot dx$ (ii) $\int x(5+3x)^{2^8} dx$

(iii) $\int \frac{x^2 + 2x}{\sqrt{x^3 + 3x^2 + 1}} \cdot dx$

Solution : (i) Let $u = 3x + 4$, $\therefore \frac{du}{dx} = 3$ or $dx = \frac{du}{3}$

Making this substitution,

$$\int \sqrt{3x+4} \cdot dx = \int \sqrt{n} \cdot \frac{du}{3} = \frac{1}{3} \int \sqrt{u} \cdot du$$

$$= \frac{1}{3} \frac{u^{3/2}}{3/2} + c$$

$$= \frac{2}{9} (3x+4)^{3/2} + c.$$

(ii) Let $u = 5 + 3x^2$, $\therefore \frac{du}{dx} = 6x$ or $dx = \frac{du}{6x}$

Making this substiution

$$\int x(5+3x^2)^8 \, dx = \int x.(u)^8 \cdot \frac{du}{6x}$$

$$= \frac{1}{6} \int (u)^8 \cdot du$$

$$= \frac{1}{6} \frac{u^9}{9} + c$$

$$= \frac{(5+3x^2)^2}{54} + c.$$

(iii) Let $u = x^3 + 3x^2 + 1$, $\therefore \quad \frac{du}{dx} = 3x^2 + 6x$

or $$dx = \frac{du}{3(x^2+2x)}$$

Making this substitution :

$$\int \frac{x^2+2x}{\sqrt{x^3+3x^2+1}} \cdot dx = \int \frac{x^2+2x}{\sqrt{u}} \cdot \frac{du}{3(x^2+2x)}$$

$$= \frac{1}{3} \int \frac{1}{\sqrt{u}} \cdot du$$

$$= \frac{1}{3} \frac{u^{1/2}}{1/2} + c$$

$$= \frac{2}{3} \sqrt{(x^3+3x^2+1)} + c.$$

(c) **Integration by parts** : Integration can also be done by parts. This method is used when function involves product of two or more functions. The integration of such functions is based on reversing the process of differentiation of product form.

Let $y = uv$. (both u and v are taken as functions of x)

$$\therefore \quad \frac{dy}{dx} = u \cdot \frac{dv}{dx} + v \cdot \frac{du}{dx}$$

Integrating both sides

$$\int \frac{dy}{dx} = \int u \cdot \frac{dv}{dx} + \int v \cdot \frac{du}{dx}$$

$$uv = \int u.dv + \int v.du$$

$$\therefore \int u \cdot dv = uv - \int v - du.$$

If a function is a product of two functions, one is taken as u and other equal to dv.Normally the function which can be differentiated easily is taken as u and which can not be integrated easily is taken as dv.

Example. Evaluate the following :

(i) $\int x^3 \cdot e^{x^2}\, dx$ (ii) $\int \ln x.dx$ (iii) $\int x^n \cdot \ln x\, dx$

Solution : (i) $u = x^2 \quad \therefore \quad \frac{du}{dx} = 2x$ or $dx = \frac{du}{2x}$

Making this substitution

$$\int x^3\, e^{x^2}\, dx = \int x \cdot u \cdot e^u \cdot \frac{du}{2x}$$

$$= \frac{1}{2} \int u \cdot e^u \cdot dx$$

Now integrating by parts :

$$u = u,\ du = 1$$

$$dv = e^u,\ v = e^u$$

$$\frac{1}{2} \int u \cdot e^u \cdot du = \frac{1}{2} \left[ue^u - \int e^u \cdot du\right]$$

$$= \frac{1}{2}\left[ue^u - e^u\right] + c$$

$$= \frac{1}{2}\left[x^2 e^{x^2} - e^{x^2}\right] + c$$

$$= \frac{1}{2} e^{x^2}\left[x^2 - 1\right] + c.$$

(ii) $\int \ln x \; dx = \int 1 \cdot \ln x \, dx$

Using integration by parts :

Let $u = Inx, du = \dfrac{1}{x}$

$dv = 1 \quad v = x$

$$\int 1 \cdot \ln x dx = x \cdot \ln x - \int x \cdot \frac{1}{x} \cdot dx$$

$$= x \cdot \ln x - x + c$$

$$= x[\ln x - 1] + c.$$

(iii) For integration by parts

$$\int x^n \cdot \ln x \, dx$$

Let $u = Inx, \quad du = \dfrac{1}{x}$

$dv = x^n, \qquad v = \dfrac{x^{n+1}}{n+1}$

$$\therefore \int x^n \cdot Inx \, dx = \ln x \cdot \frac{x^{n+1}}{n+1} - \int \frac{x^{n+1}}{n+1} \cdot \frac{1}{x} \cdot dx$$

$$= \ln x \cdot \frac{x^{n+1}}{n+1} - \frac{1}{n+1} \int x^n \cdot dx$$

$$= \ln x \cdot \frac{x^{n+1}}{n+1} - \frac{x^{n+1}}{(n+1)^2} + c$$

$$= \frac{x^{n+1}}{n+1}\left[\ln x - \frac{1}{n+1}\right] + c$$

Applications of Indefinite Integral in Business and Economics

Integration is broadly classified as definite and indefinite integration. Both these have their own applications. The most important application of integration, i.e., measurement of area under a curve, is based on definite integration, but indefinite integration has its own applications. These are in the form of obtaining original function from its derivative function. For example, finding total cost function from the marginal cost function. Some of these are explained below :

(a) Derivation of total cost function and average cost function from the marginal cost function : If marginal cost function, $MC(x)$ is integrated with respect to x, it will give total cost function, $TC(x)$ [since $MC(x)$ is the derivative of $TC(x)$]. Therefore,

$$\int MC\,(x)\,dx = TC\,(x) + C$$

where $TC'(x)$ is denoting the integral of $MC(x)$ and C is arbitrary constant which was eliminated from $TC(x)$ in the process of differentiation. $TC'(x)$ is same as total variable cost function.

The value of C can be obtained if the total cost of some units of x is given. Put the value of x in $TC(x) = TC'(x) + C$ = given cost and solve for C. So $TC(x) = TC(x) + C$ (value obtained). This C is the fixed cost part of total cost and is same as total cost at zero level of output. Therefore it can be obtained by setting $x = 0$ in such cases.

If the total cost function is obtained, average cost function can be obtained just be dividing the $TC(x)$ by x (as average cost is the total cost divided by number of units produced).

Therefore $$AC(x) = \frac{TC(x)}{x}.$$

Example. The marginal cost function of manufacturing x pairs of shoes is $MC(x) = 6 + 10x - 6x^2$. The total cost of producing one pair of shoes is Rs. 12.00. Determine the total and average cost function.

Solution : Given $$MC(x) = 6 + 10x - 6x^2$$

$$\therefore\ TC(x) = \int MC(x)\,.\,dx$$

$$= \int (6 + 10x - 6x^2)\,dx$$

$$= 6x + 5x^2 - 2x^3 + C.$$

Given $TC(1) = 6(1) + 5(1) - 2(1) + C = 12$

$\therefore\ C = 3.$

Putting $C = 3$ in $TC(x)$

$$TC(x) = 6x + 5x^2 - 2x^3 + 3$$

and $$AC(x) = \frac{TC(x)}{x} = \frac{6x + 5x^2 - 2x^3 + 3}{x}$$

$$= 6 + 5x - 2x^2 + \frac{3}{x}$$

Example. The marginal cost of producing an item is given by $MC(x) = \frac{x}{\sqrt{x^2 + 400}}$ where x represents the unit produced. Determine its total cost function and average cost function if cost at zero level of output is Rs. 900.

Solution : $$MC(x) = \frac{x}{\sqrt{x^2 + 400}}$$

$$\therefore\ TC(x) = \int \frac{x}{\sqrt{x^2 + 400}}\,.\,dx$$

Using substitution method, let $u = x^2 + 400$

$$\frac{du}{dx} = 2x \quad \text{or} \quad dx = \frac{du}{2x}$$

Making this substitution

$$TC(x) = \int \frac{x}{\sqrt{u}} \cdot \frac{du}{2x}$$

$$= \frac{1}{2} \int \frac{1}{\sqrt{u}}\,.\,du$$

$$= \frac{1}{2} \cdot \left(\frac{u^{1/2}}{1/2} \right) + C$$

$$= \sqrt{u} + C$$

$$= \sqrt{x^2 + 400} + C$$

Given total cost at zero level is Rs. 900.

$\therefore \quad TC(0) = \sqrt{0 + 400} + C = 900$

$C = 880$

Putting C = Rs.880 in $TC(x)$,

$$TC(x) = \sqrt{x^2 + 400} + 880$$

$$AC(x) = \frac{\sqrt{x^2 + 400} + 880}{x}$$

Example. The cost of producing xth unit is given by $MC(x) = 20\ e^{0.5x}$. Determine the corresponding total cost function and average cost function if total cost at zero level is Rs. 80.

Solution : Given $MC(x) = 20\ e^{0.5x}$

$$\therefore \quad TC(x) = \int 20e^{0.5x} \, . \, dx$$

$$= \frac{20e^{0.5x}}{0.5} + C.$$

$$= 40e^{0.5x} + C.$$

Using $TC(0) =$ Rs. 80

$$TC(0) = 40\ e^{0.5(0)} + C = 80$$

$$C = \text{Rs. } 40$$

Putting C = Rs. 40

$$TC(x) = 40\ e^{0.5x} + 40$$

$$AC(x) = \frac{40e^{0.5x} + 40}{x}$$

(b) Derivation of total revenue function and average revenue function from marginal revenue function : This indefinite integration can also be applied for obtaining total revenue function, $TR(x)$ from marginal revenue function. $MR(x)$, which in turn can be used for deriving average revenue function, $AR(x)$ also known as demand function. It is as follow.

$$TR(x) = \int MR\,(x)\, dx = TR'(x) + C$$

where $TR'(x)$ is simply the integral of $MR(x)$ and C is the arbitrary constant. The value of C is evaluated assuming total revenue at zero level is zero. So setting $TR(0) = 0$, value of C is obtained.

Corresponding to the $TR(x)$,

$$AR(x) = \frac{TR(x)}{x}$$

Example. The additional revenue of xth unit is given by $MR(x) = x^2 - 12x + 27$. Determine its total revenue function and average revenue function.

Solution : Given $MR(x) = x^2 - 12x + 27$

$$\therefore \quad TR(x) = \int (x^2 - 12x + 27)\, dx$$

$$= \frac{x^3}{3} - 6x^2 + 27x + C$$

Assuming $TR(0) = 0$

$$TR(0) = 0 + C = 0$$

$$C = 0.$$

So, $$TR(x) = \frac{x^3}{3} - 6x^2 + 27x$$

$$AR(x) = \frac{\frac{x^3}{3} - 6x^2 + 27x}{x}$$

or

$$P = \frac{x^2}{3} - 6x + 27.$$

Example. Find the total revenue function $TR(x)$ and demand function of the marginal revenue function.

$$MR(x) = \frac{ab}{(x+b)^2} - C.$$

Solution : Given $MR(x) = \frac{ab}{(x+b)^2} - C$

$$TR(x) = \int \left(\frac{ab}{(x+b)^2} - C. \right) dx$$

$$= \frac{-ab}{-x+b} - Cx + C_1 \qquad [C_1 \text{ is arbitrary constant}]$$

Putting $TR(0) = 0$

$$TR(o) = -\frac{ab}{0+b} - C(0) + C_1 = 0$$

$$C_1 = a$$

$$TR(x) = -\frac{ab}{(x+b)} - CX + a = +\frac{ax}{s+b} - Cx \qquad \text{(simplified form)}$$

and $$AR(x) = +\frac{\frac{ax}{x+b} - Cx}{x} - C$$

or, $$P = +\frac{a}{x+b} - C$$

Example. The additional revenue of xth unit of a commodity is given by $MR(x) = \frac{100}{x+2}$. Determine the corresponding total revenue function and demand function.

Solution : Given $$MR(x) = \frac{100}{x+2}$$

$$TR(x) = \int \frac{100}{x+2} \cdot dx$$

$$= 100\, ln(x+2) + C$$

Assuming $$TR(0) = 0$$

$$TR(0) = 100\,[ln 2] + C = 0$$

$$C = -100\, ln\, 2$$

So, $$TR(x) = 100\, ln(x+2) - 100\, In\, 2$$

$$= 100\left[ln\left(\frac{x+2}{2}\right)\right]$$

$$= 100\left[ln\left(\frac{x}{2}+1\right)\right]$$

Corresponding $$AR(x) = \frac{100\left[ln\left(\frac{x}{2}+1\right)\right]}{x}$$

$$P = \frac{100\left[ln\left(\frac{x}{2}+1\right)\right]}{x}$$

DEFINITE INTEGRATION

Integration is broadly classified as indefinite and definite integration. Indefinite integration involves antiderivative of a function. It is basically used for obtaining original function from its derivative function. It is known as indefinite because the constant of integration, c is arbitrary.

Definite integral is one step ahead of indefinite integral. It is mainly used for the measurement of area under a curve. If a function $f(x)$ is continuous in the close interval $[a, b]$, then the symbol $\int_a^b f(x)\, dx$ denotes the value of definite integral of F over the interval $[a, b]$.

According to this theorem, if $f(x)$ is a continuous function in the close interval $[a, b]$, then the area under the curve between a and b is given by evaluating the integral of $f(x)$ with respect to x with the limits a and b. (These are known as lower bound and upper bound respectively). It is shown in figure 1. Its value can be obtained in this way.

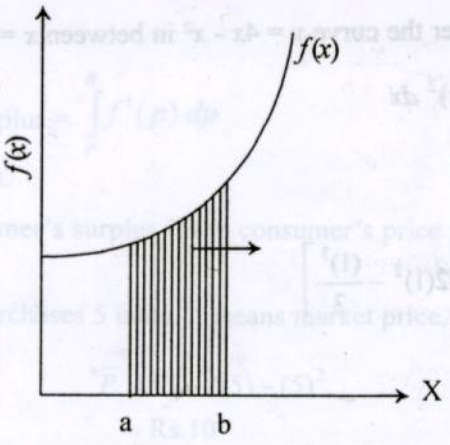


Fig. 1

Fig. 1

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b$$

$$= F(b) - F(a)$$

This gives a numeric value denoting area.

[Note : Value of the aribitrary constant c is zero in this case].

Example. Evaluate the following :

(i) $\int_{-1}^{3}(3x^2+5x-1)\,dx$ (ii) $\int_1^5 e^x\,dx$

(iii) $\int_2^8 lnx\,dx$

Solution : (i) $\int_{-1}^{3}(3x^2+5x-1)\,dx = \left(\frac{3x^2}{3}+\frac{5x^2}{2}-x\right)\Big|_{-1}^{3}$

$$= \left(x^3+\frac{5}{2}x^2-x\right)\Big|_{-1}^{3}$$

$$= \left(3^3+\frac{5}{2}.3^2-3\right)-\left(-1^3+\frac{5}{2}.-1^2+1\right) = \frac{83}{2}$$

(ii) $\int_1^5 e^x\,dx = e^x\Big|_1^5 = e^5 - e^1 = 145.69$ app.

(Value of e^5 and e^1 are given by exponential table)

(iii) $\int_2^8 lnx\,dx = (x\,lnx - x)\Big|_2^8$

$$= (8ln\,8-8)-(2ln\,2-2)$$

$$= 8ln\,8 - 2ln2 - 6$$

$$= 9.249$$

[Value of $ln\,8$ and $ln\,2$ are given by natural log table]

Example. Find the area under the curve $y = 4x - x^2$ in between $x = 1$ and $x = 3$ on x-axis.

Solution : Area $= \int_{-1}^{3} (4x - x)^2 \, dx$

$$= 2x^2 - \frac{x^3}{3}\Big|_1^3$$

$$= \left[2(3)^2 - \frac{(3)^2}{3}\right] - \left[2(1)^2 - \frac{(1)^3}{3}\right]$$

$$= \frac{22}{3} \text{ square units.}$$

Applications in Business and Economics

The most important application of definite integration is the measurement of area under a curve. Some of these are as follows :

(a) **Measurement of consumer's surplus** : In economics, consumer's surplus is defined as the difference between the price consumer is willing to pay and the price which he actually pays. If the price consumer willing to pay, is more than the price he actually pays, the difference is consumer's surplus. Otherwise there will be no consumer's surplus. The price which the consumer willing to pay is given by demand function and the price which he pays is given by market demand and supply.

Diagrammatically it is shown in figure 2. Let $P = f(x)$ is a price function stating the price consumer is willing to pay for different values of x and P is the price fixed by market demand and supply at which consumer is purchasing x units. Then total price consumer is willing to pay, is given by the integral of $f(x)$ with the limits $[0, \bar{x}]$ whereas price actually paid is $\bar{p}\,\bar{x}$. So the difference, i.e., area of $AB\bar{P}$ is the measurement of consumer's surplus.

Therefore, Consumer's surplus $= \int_0^{\bar{x}} f(x), dx - \overline{px}$

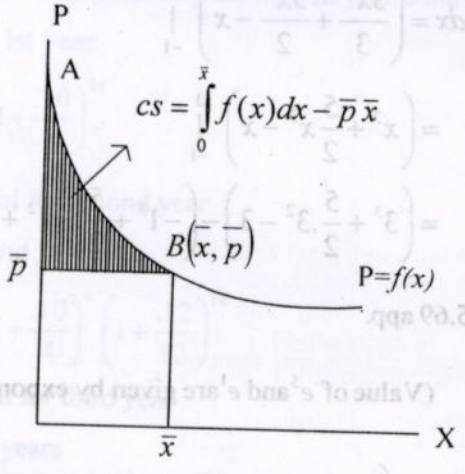


Fig. 2

An alternative method is to measure consumer's surplus with the help of integrating demand function in relation to price (*i.e.*, $x = f'(p)$ which is inverse of price function). It involves obtaining the demand function corresponding to price function and integrating it with respect to P with the limits $[\bar{P}, B]$ where B is the price at which quantity demanded is zero. It can be obtained by setting $x = 0$ in

the demand function.

Therefore consumer's surplus $= \int_{\bar{P}}^{B} f^{1}(p)\,dp$

Both gives the same result.

Example. Find the consumer's surplus if the consumer's price function is $p = 45 - 2x - x^2$ and he purchase 5 units.

Solution : If consumer purchases 5 units, it means market price,

$$\bar{P} = 45 - 2(5) - (5)^2$$
$$= \text{Rs.}10$$

$$\therefore \text{ Consumer's surplus } = \int_0^5 (45 - 2x - x^2)\,dx - 10(5)$$
$$= \left(45x - x^2 - \frac{x^3}{3}\right)\Big|_0^5 - 50 = \frac{475}{3} - 50 = \frac{\text{Rs. } 325}{3} \text{ or Rs.}108.33 \text{ app.}$$

Alternatively, inverse function of the given price function, i.e., demand function will be

$$45 - 2x - x^2 - P = 0$$

or
$$x^2 + 2x - 45 + P = 0$$

$$\therefore \qquad x = \frac{-2 \pm \sqrt{4 + 4(45 - P)}}{2} = \sqrt{46 - P} - 1$$

[negative value not possible, so ignored]

$$\therefore \text{ Consumer's surplus } = \int_{10}^{45} \sqrt{46 - P} - 1 \qquad [P = \text{Rs.}45 \text{ when } x = 0 \text{ units}]$$

$$= -\frac{2}{3}(46 - P)^{3/2}\Big|_{10}^{45} - P\Big|_{10}^{45}$$

$$= \left(-\frac{2}{3} + 144\right) - 35 \qquad = \frac{\text{Rs.}325}{3} = \text{Rs.}108.33 \text{ app.}$$

Example. A monopolist's demand and total cost function are P = 25 - x and TC(x) = 10 + 9x respectively. Determine the consumer's surplus (assuming monopolist maximises profit).

Solution : In case of monopoly, the firm supplies units where it makes maximum profit. It is given by setting marginal revenue equal to marginal cost. It is shown below.

Given $\qquad P = 25 - x \qquad$ and $\qquad TC(x) = 10 + 9x$

$\therefore \qquad TR(x) = P \,.\, x = 25x - x^2 \qquad$ and $\qquad MC(x) = 9$

$$MR(x) = \frac{dTR(x)}{dx} = 25 - 2x$$

Setting $MR(x) = MC(x)$ and solving for x

$$25 - 2x = 9 \quad \text{or}$$
$$x = 8 \text{ unit.}$$

Profit is maximum at $x = 8$ unit as marginal cost is constant

$\therefore \quad P = 25 - 8 = \text{Rs.}17$

$$\text{Consumer's surplus} = \int_0^8 (25 - x)\,dx - 17(8)$$

$$= \left(25x - \frac{x^2}{2}\right)\Big|_0^8 - 136$$

$$= 168 - 136 \quad = \text{Rs. } 32.$$

(b) Measurement of Producer's Surplus. As in the case of consumer's surplus, producer's surplus is the difference between the price received by the producer and the price at which producer is willing to supply. For example, if producer sells product at Rs. 10.00 per unit whereas he is willing to sell at Rs. 8.00, per unit, the producer's surplus is Rs. 2.00. The price producer is willing to accept is given by his supply curve; and the price he actually gets is determined by demand and supply. Diagrammatically it is shown by figure 3.

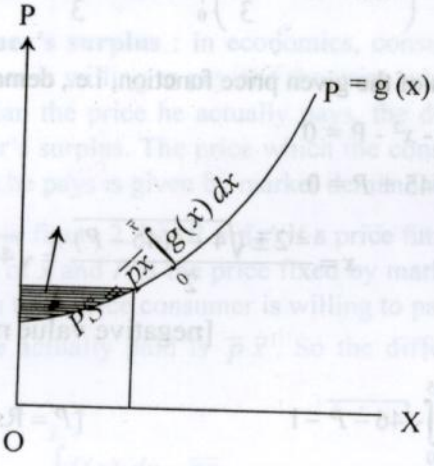


Fig.3

$\therefore$ Therefore Producer's Surplus $= \overline{px} - \int_0^{\bar{x}} g(x)\,dx$

where $\bar{p}$ and $\bar{x}$ are the market price and quantity supplied. $g(x)$ is the supply function of the producer. $\bar{p}$ $\bar{x}$ denotes the price producer gets by supplying $\bar{x}$ units and $\int_0^{\bar{x}} g(x)dx$ gives the price producer is willing to accept and the difference represents producers surplus.

An alternative method is to take the inverse of given supply function $P = g(x)$ which will be $x = g'(P)$ denoting x as a function of P. The integral of g'(P) with the limit $(P_0, \bar{P})$ gives the producer's surplus. P_0 is denoted the price at which producer is willing to supply any unit it can be obtained by setting $x = 0$ in the either function and solving for P.

$$\text{Therefore Producer's Surplus} \quad = \int_{P_0}^{\bar{P}} g'(P)dp$$

Example. The supply equation of a commodity is $10P - 2\sqrt{x + 300} = 0$ and market price is Rs. 8.00. Find the producer's surplus.

Solution : Given supply equation, $10P - 2\sqrt{x+300} = 0$

or, $$10P - 2\sqrt{x+300} = 0$$

$$P = \frac{1}{5}\sqrt{x+300}$$

when $P = 8$, quantity supply will be $8 = \frac{1}{5}\sqrt{x+300}$

or, $$1600 = x + 300$$

$\therefore$ $$x = 1300 \text{ units.}$$

Producer's surplus $= 8(1300) - \int_0^{1300} \frac{1}{5}\sqrt{x+300}\, dx$

$$= 10{,}400 - \frac{1}{5}\ (x+300)^{3/2} \cdot \frac{2}{3} \Big|_0^{1300}$$

$$= 10{,}400 - \frac{2}{15}\left[(x+300)^{3/2} \cdot \Big|_0^{1300}\right]$$

$= 10{,}400 - 7,\ \ 840 = Rs.\ 2560.$ (simplification)

Example. Under pure competition, the demand and supply functions of a commodity are $P = \frac{8}{x+1} - 2$ and $P = 1/2\ (x + 3)$ respectively. Determine the consumer's and producer's surplus at equilibrium price.

Solution : $$P = \frac{8}{x+1} - 2$$ (Demand)

$$p = \frac{1}{2}(x+3)$$ (Supply)

So equilibrium quantity is given by

$$\frac{8}{x+1} - 2 = \frac{1}{2}(x+3)$$

or, $16 - 2(x+1)2 = (x+3)(x+1)$

$x^2 + 8x - 9 = 0$

$(x + 9)(x - 1) = 0$

So $x = -9$ or 1

$x = 1$ [as $x = -9$ is not possible]

Putting $x = 1$, in demand or supply function, equilibrium price, P = Rs. 2.

So consumer's surplus $= \int_0^1 \left(\frac{8}{x+1} - 2\right) dx - 2(1)$

$$= (8ln\,(x+1) - 2x)\Big|_0^1 - 2$$

$$= [(8ln\,2 - 2) - (8ln\,1 - 0)] - 2$$

$$= 8ln\,2 - 4$$

$$= 8(.6932) - 4 = \text{Rs. } 1.55 \text{ app.} \qquad [\text{as } ln\,1 = 0]$$

Producer's surplus $= 2(1) - \int_0^1 \frac{1}{2}(x+3)\,dx$

$$= 2 - \frac{1}{2}\left[\left(\frac{x^2}{2} + 3x\right)\Big|_0^1\right]$$

$$= 2 - \frac{1}{2}\left[\frac{7}{2}\right] = 2 - \frac{7}{4}$$

$$= \text{Rs. } \frac{1}{4} \quad \text{or} \quad 0.25 \text{ app.}$$

(c) Determination of maximized Profit : Definite integration can also be applied for determining the maximized profit if the marginal revenue and marginal cost functions are given. Marginal revenue gives the rate of change in total revenue and marginal cost gives the rate of change in total cost. The firm makes maximum profit when the rate of change in both are equal and afterwards rate of change in total cost is more than the rate of change in total revenue (*i.e.*, marginal cost is increasing). Diagrammatically it is shown in figure 4.

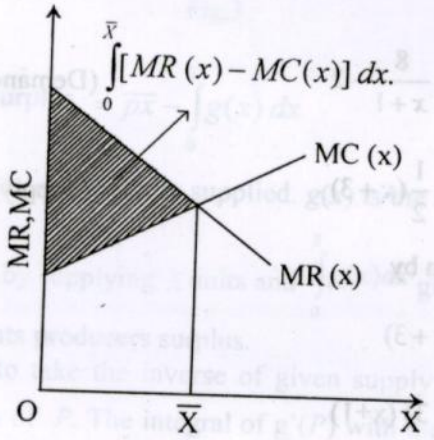


Fig. 4

The area shaded represents the maximized profit. It is the area bounded by marginal revenue function MR (x) and marginal cost function MC (x) with the limit $[0, \bar{X}]$ $\bar{X}$ is denoting the profit maximizing output obtained by setting MR (x) =MC (x) and solving X. At this MC (x) must be increasing for maximum profit.

So maximized profit $= \int_0^{\bar{X}} [MR(x) - MC(x)]\, dx.$

This gives the additional profit only. So loss at zero level of output must be subtracted from this for obtaining actual profit. It is same as fixed cost in the short-run.

So maximized profit $= \int_0^{\bar{X}} (MR(x) - MC(x))\, dx. - C$

where C is the fixed cost or loss at zero level of output.

If is happen that marginal revenue is equal to marginal cost at two level of output as shown in figure 5.

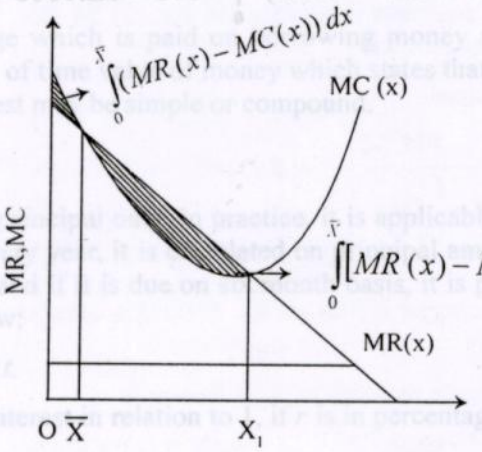


Fig. 5

Upto OX, marginal cost lies above marginal revenue representing loss. From $\bar{X}$ to X_1 marginal revenue lies above marginal cost represent profit. So maximized profit is the difference between profit area and loss area.

$$\text{Therefore Maximized profit} = \int_{\bar{X}}^{\bar{X}_1} [MR(x) - MC(x)]\, dx - \int_0^{\bar{X}} [MC(x) - MR(x)]\, dx$$

$$= \int_{\bar{X}}^{\bar{X}_1} [MR(x) - MC(x)]\, dx$$

Example. A firm's marginal revenue function, $MR(x) = 1000 - 4x$ and marginal cost function, $MC(x) = 3x^2 - 118x + 1315$. Determine the maximized profit using definite integration assuming fixed cost is Rs. 595.

Solution : Given $MR(x) = 1000 - 4x$, $MC(x)\ 3x^2 - 118x + 1315$.

For profit maximizing output, put $MR(x) = MC(x)$ and solve for x.

$\therefore \quad 1000 - 4x = 3x^2 - 118x + 1315$

or, $\quad 3x^2 - 114x + 315 = 0$

$$x = \frac{114 \pm \sqrt{(114)^2 - 4(3)\ (315)}}{6} = 3 \text{ or } 35$$

But it is making maximum profit when $X = 35$ only as marginal cost is increasing at this level (It can be verified by the derivative of $MC(x)$ at $x = 35$. It is positive).

$$\text{So Maximized profit} = \int_0^{35} [(1000 - 4x) - (3x^2 - 118x + 1315)]\, dx - 595$$

$$= \int_0^{35} (-3x^2 + 114x - 315)\, dx - 595$$

$$= (-x^3 + 57x^2 - 315x) \Big|_0^{35} - 595 = \text{Rs.}1000$$

BASIC MATHEMATICS OF FINANCE

[SIMPLE AND COMPOUND INTEREST]

Mathematics of finance deals with simple interest, compound interest, trade discount, present and the future values of annuities, nominal and the effective rate of interest, yield to maturity etc. It is very important as all these concepts are commonly used in the financial decision making.

Simple and Compound Interest

Interest denotes the charge which is paid on borrowing money and is received on lending the money. It is based on the concept of time value of money which states that Rs.100 of today is not equal to Rs. 100 one year after. This interest may be simple or compound.

Simple Interest

It refers to the interest on principal only. In practice, it is applicable only when interests is paid as and when it is due. Therefore, every year, it is calculated on principal amount. For example, if interest is due annually, it is paid annually and if it is due on six month basis, it is paid on six month basis. Simple interest can be computed as bellow:

Simple Interest (S.I.) = $P.r.t.$

where P = principal; r = rate of interest in relation to 1, if r is in percentages it is divided by 100: t = time period in years.

Corresponding Amount = $P+S.I. = P+P.r.t. = P(1+rt)$

Note 1. - Principal also denotes the present value or value in the beginning or value in the zero period whereas Amount denotes the future value at the t^{th} period.

Note 2. - Whenever interest is paid as and when it is due, it is always assumed to be a case of simple interest.

Example 1: A sum of Rs. 5000 is deposited by Mr. X in a finance company under quarterly interest scheme. Find the total interest earned by him in 5 years at 12% per annum.

Solution: Interest of each quarter = Prt

$$= 5000 \times \frac{12}{100} \times \frac{1}{4} = \text{Rs.}150$$

Total interest earned = Rs.150 × 5 × 4 = Rs.3000

The main feature of the simple interest is that the total interest for the given period is always same whether paid annually, six monthly, quarterly or monthly.

Example 2: What time will be required for a sum of money to double itself if simple rate of interest is 10% per annuam.

Solution: Let the sum or Principal = Rs. P

$$\therefore \quad \text{Amount} = 2P = P(1+rt)$$

$$\Rightarrow 2P = P(1+.10t)$$

$$\Rightarrow .10t = 1$$

$$\Rightarrow t = 10 \text{ years}$$

Example 3 : If in four years Rs. 2520 amount to Rs. 3024, in what time at the same rate will Rs. 520 amounts to Rs. 611 at simple interest.

Solution: Let the rate of interest = $100r\%$

$\therefore$ Amount $= P(1+rt)$

$$3024 = 2520(1+4r)$$

or $r = 0.05$ or $r = 5\%$

Let the time period = t years

$\therefore$ $611 = 520\,(1+.05t)$

$\Rightarrow t = 3.5$ years

Example 4: What is the cash payment equivalent to Rs.297.54 due in 3 years and 6 months, simple interest being reckoned at 4% per annum?

Solution: Cash payment = Principal

$$\therefore \text{ Principal} = \frac{A}{1+rt} = \frac{295.54}{1+\left(.04\times\frac{7}{2}\right)} = \text{Rs. } 261$$

Example 5: What sum will discharge a debt of Rs. 5300 due after one and half year at 5% per annum simple interest doday?

Solution: Sum required to discharge the debt today = Principle

$$\Rightarrow \text{ Principal} = \frac{A}{1+rt} = \frac{5300}{1+.05\left(\frac{3}{2}\right)} = \text{Rs. } 4930.24$$

Example 6: In how many years at 4% per annum simple interest will the interest be one quarter of the principal.

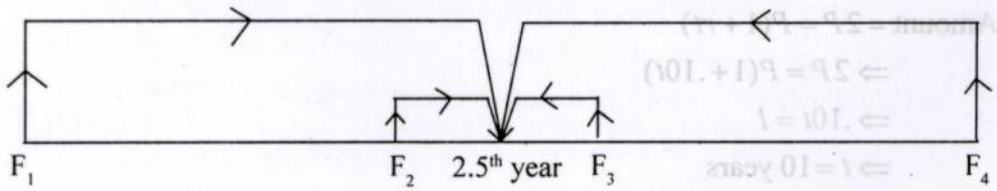
Solution: Let the principal = Rs. P.

$$\text{Simple Interest} = \frac{1}{4}\text{ P } = P(.04)^t$$

$$\Rightarrow t = \frac{1}{\cdot 16} = 6\frac{1}{4} \text{ years.}$$

Equating the different cash flows involving simple interest

The equivalent value of the cash flows at different points of time involving simple interest can also be obtained at a point of time formally known as focal date. The equivalent value denotes a lump-sum equal to all the flows. In such cases it is preferable to take the help of time scale. All the flows may be denoted on the time scale (exact scale is not required). For example:- Four different flows- F_1, F_2, F_3 and F_4 are there at the end of each year and their equivalent value is required at 2.5th year lying in between second and third year, then it is given by the sum total of the value of all these four flows at 2.5th year.



So the flows which are on the left hand side of the focal date their future value is required and the flows which are on right hand side, their present value is required. (This rough time scale enables us in deciding whether future value of present value of the particular flow is required to arrive at the equivalent value).

Example 7: At 5% simple interest, find the value today of the following obligations: Rs. 1000 due today, Rs. 2000 due after 6 months with interest at 6% and Rs. 4000 due in one year with interest at 8%.

Solution: Since value is required in the beginning or in zero period and the flows are at the end of 6 months and 1 year (which are on the right hand side of focal date), present value of both the flows due at the end of 6 months and one year is required.

Amount due at the end of 6 months = $P(1+rt)$

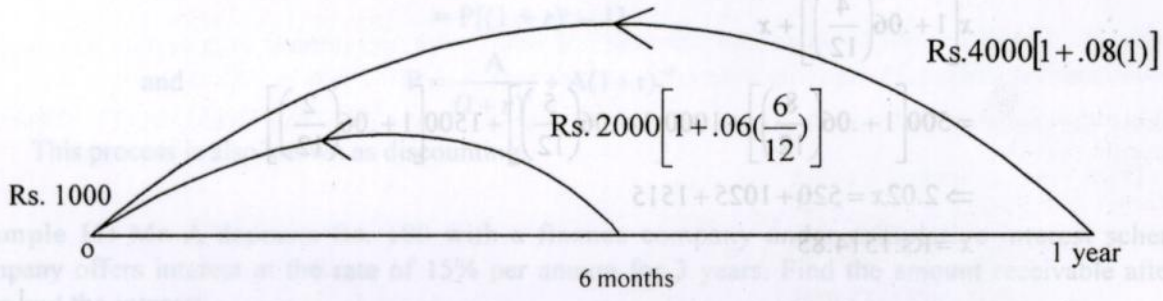
$$= 2000\left[1+.06\left(\frac{6}{12}\right)\right] = \text{Rs. } 2060$$

Amount due at the end of 1 year = 4000 [1+.08(1)] = Rs. 4320

$\therefore$ The value of all the flows in the beginning at 5%

$$= 1000 + \frac{2060}{\left[1+.05\left(\frac{1}{2}\right)\right]} + \frac{4320}{[1+.05(1)]} = \text{Rs. } 7124.04$$

Same can be analysed using time scale in this way



Example 8: A man owes Rs. 500 due in 3 months with interest at 5% and Rs.1500 due in 9 months with interest at 4% if money is worth 6%, what single payment made at the end of 6 months will discharge his debts?

Solution: Let the single payment at the end of six months = Rs.X

$\therefore$ X = furture value of the amount due at the end of 3 months + present value of the amount due at the end of 9 months

Amount due at the end of 3 months at 5%

$$= 500\left[1+(.05)\left(\frac{3}{12}\right)\right] = \text{Rs. } 506.25$$

$\therefore$ Value of Rs. 506.25 at the end of 6 month at 6% $= 506.25\left[1+.06\left(\frac{3}{12}\right)\right]$

$$= \text{Rs. } 513.85 \qquad \text{.....(1)}$$

Amount due at the end of 9 months at 4%

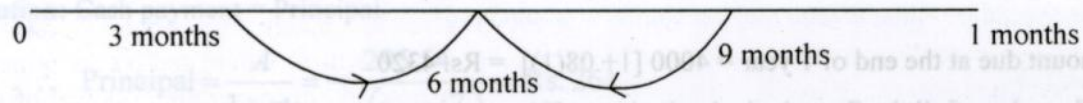
$$= 1500\left[1 + .04\left(\frac{9}{12}\right)\right] = \text{Rs. } 1545$$

$\therefore$ The value of Rs. 1545 at the end of 6 month at 6% $= 1545\left[1 + .06\left(\frac{3}{12}\right)\right]$

$= \text{Rs. } 1522.16$ ..... (ii)

$\therefore$ Total amount due at the end of 6 months

$= \text{Rs.}1513.85 + \text{Rs. } 1522.16 = \text{Rs. } 2036.01$



Example 9: X owes Rs. 500 due in 2 months, Rs. 1000 due in 5 months and Rs. 1500 due in 8 months. He wishes to discharge his ibligation in two equal payments; one in six months and other due in 10 months. Find the payment if money is worth 6% simple interest and the 10th month is taken as focal date.

Solution: Let each payment is of Rs. x. Therefore the value of both the payments at the end of 10 months should be equal to value of all the obligations at the end of 10 months.

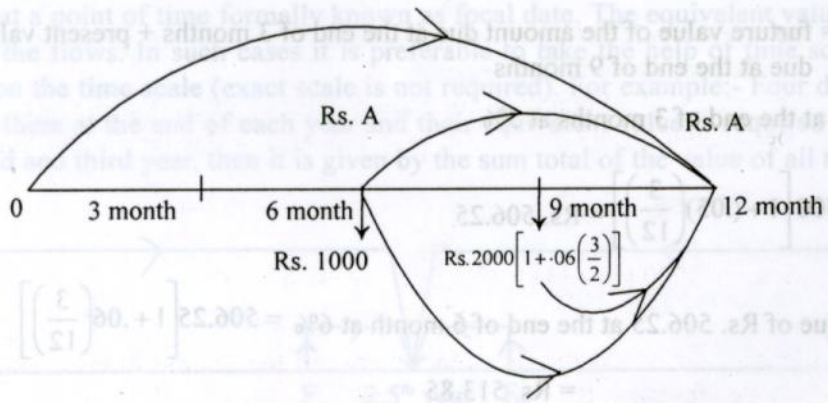
$\therefore$ $x\left[1 + .06\left(\frac{4}{12}\right)\right] + x$

$$= 500\left[1 + .06\left(\frac{8}{12}\right)\right] + 1000\left[1 + .06\left(\frac{5}{12}\right)\right] + 1500\left[1 + .06\left(\frac{2}{12}\right)\right]$$

$$\Rightarrow 2.02x = 520 + 1025 + 1515$$

$$x = \text{Rs.}1514.85$$

Example 10: X owes Y Rs. 1000 due in 6 months without interest and Rs.2000 with interest for $1\frac{1}{2}$ year at 6% due in 9 months. Y agrees to accept 3 equal payments; first due today, second at the end of 6 months and the third at the end of 1 year. Find the equal payments using one year from today as the focal date if money is worth 8% to Y.



Solution: Let each payment is of Rs. A

$$\therefore \quad A\left[1+.08(1)\right]+A\left[1+.08\left(\frac{6}{12}\right)\right]+A$$

$$=1000\left[1+.08\left(\frac{6}{12}\right)\right]+2000\left[1+.06\left(\frac{3}{2}\right)\right]\left[1+.08\left(\frac{3}{2}\right)\right]$$

$$\Rightarrow 1.08A+1.04A+A=1040+2223.6$$

$$\Rightarrow 3.12A=1046$$

$$A = \text{Rs. } 1046.03$$

Compound Interest

It carries interest on the principal and on the interest due also. It is mainly used when interest is not paid as and when due. Therefore, in each successive periods interest will be calculated on the principal plus the interest due. In other words, amount of each successive period becomes the principal of next period. Thus interest of each successive period will go on increasing. For the first period, both the simple interest and the compound interest are equal. In this case-

$$\text{Amount (A)} = P(1+r)^t$$

Where A is the amount also known as future value, P is the principal or present value, r = rate of interest per annum in relation to 1 and t = time period in years.

$$\text{Compound interest (CI)} = A-P = P[1+r]^t - P$$

$$= P[(1+r)^t - 1]$$

and $$P = \frac{A}{(1+r)^t} + A(1+r)^{-t}$$

This process is also known as discounting.

Example 11: Mr. A deposits Rs. 100 with a finance company under cumulative interest scheme. Company offers interest at the rate of 15% per annum for 3 years. Find the amount receivable after 3 years and the interest.

Solution:

$$\text{Amount receivable} = P(1+r)^t$$

$$= 10{,}000\,(1+.15)^3 = 15208.75$$

$$\text{Interest receivied} = 15208.75 - 10{,}000 = \text{Rs. } 5208.75$$

Example 12: Mr. A deposits Rs. 5000 in a bank for 3 years offering interest at the rate of 9% per annum in the first year, 10% per annum during the second year and 11% per annum in the third year. Find his balance after 3 years.

Solution:

$$\text{Balance after 3 years} = 5000\,(1+.09)\,(1+.11) = \text{Rs. } 6654.45$$

Example 13: Find the sum on which the difference between the simple interest and the compound interest for 3 years at 5% per annum is Rs. 73.20.

Solution: Let the sum or Principal = Rs. P.

$\therefore$ $SI = Prt$

$= P \times .05 \times 3 = .15P$

$CI = P[(1 + r)^t - 1]$

$= P[(1.05)^3 - 1]$

$CI - SI = 73.20$

$\Rightarrow P[(1.05)^3 - 1 - 1.5] = 73.20$

$\Rightarrow$ P = Rs. 9600

Example 14: The compound interest on a sum of money for 3 years at 5% is Rs. 1324.05. What is the simple interest?

Solution: Let the principal = Rs. P.

$\therefore$ $SI = P(.05)(3) = .15P$

$CI = P[(1 + .05)^3 - 1]$

$\therefore \quad \frac{SI}{CI} = \frac{.15}{(1.05)^3 - 1} = .951625693$

$\therefore \quad SI = .951625693 \times CI$

$\Rightarrow \quad SI = .951625693 \times 1324.05 = Rs.1260$

Example 15: The difference between 2nd year and 3rd year compound interest on a certain sum at 5% compound interest is Rs. 5.25. Find the sum

Sulution: Let the principal = Rs. P.

Third Year Compound Interest

$= CI \text{ of 3 years} = CI \text{ of 2 years}$

$= P[(1.05)^3 - 1] - P[(1.05)^2 - 1]$

$= P[(1.05)^3 - (1.05)^2]$

$= (1.05)^2 P(.05)$ .....(i)

Second Year Compound Interest

$= P[(1.05)^2 - 1] - P[(1.05) - 1]$

$= (1.05) P(.05)$ .....(ii)

$\therefore \quad (i) - (ii) = 5.25$

$\Rightarrow \quad (1.05)^2 P(.05) - (1.05) P(.05) = 5.25$

$\Rightarrow \quad (1.05)(.05)^2 P - 5.25$

$\Rightarrow$ P = Rs. 2000

Example 16: A sum of money put at compound interest amounts to Rs. 578.40 in 2 years and Rs.614.55 in 3 years. Find the rate of interest.

Solution: In case of compound interest, amount of 2nd years is principal of 2rd year.

$\therefore$ Third year Principal = Rs. 578.40

$$A = \text{Rs. } 614.55$$

$$t = \text{one year}$$

$$614.55 = 578.40(1+r)1$$

$$\Rightarrow \quad r = .0625 = 6.25\%$$

In case compounding is m times in a year, r is replaced by $\frac{r}{m}$ and t is replaced by mt.

$\therefore$ if compounding is 6 monthyly *i.e.* two times in a year, r is replaced by $\frac{r}{2}$ and t is replaced by $2t$.

$$\text{Amount} = P\left(1+\frac{r}{2}\right)^{2t}$$

If compounding is quarterly *i.e.* 4 times in a year, r is replaced by $\frac{r}{4}$ and t is replaced by $4t$.

$\therefore$ $$\text{Amount} = P\left(1+\frac{r}{4}\right)^{4t}$$

If compounding is monthly *i.e.* 12 times in a year then

$\therefore$ $$\text{Amount} = P\left(1+\frac{r}{12}\right)^{12t}$$

If compounding is m times in a year then $\text{Amount} = P\left(1+\frac{r}{m}\right)^{mt}$

But if compounding is continuous, number of times interest compounded in a year will go to infinity *i.e.* limit of m tends towards infinity.

$$\text{Amount} = \lim_{m\to D} P\left(1+\frac{r}{m}\right)^{mt}$$

$$= \lim_{m\to D} P\left[\left(1+\frac{r}{m}\right)^{\frac{m}{r}}\right]^{rt}$$

$$= Pe^{rt} \quad \textit{where} \quad e = \lim_{m\to D}\left(1+\frac{r}{m}\right)^{\frac{m}{r}}$$

Example 17: Accumulate Rs. 5000 for 6 years at 10% per annum compunded quarterly.

Solution: Given P = Rs. 5000

$$r = \frac{.10}{4}, \quad t = 4\times 6 = 24$$

$\therefore$ $$\text{Amount(A)} = 5000\left(1+\frac{.10}{4}\right)^{24} \qquad \text{....(i)}$$

$$= 5000(1.8087) = \text{Rs. } 9043.63$$

It can also be sloved using log and compound interest-tables. Compound interest table gives the compound value of Re. 1 at different rates of interest for different periods. Using log it can writen as

$$\log A = \log 5000 + 24 \log 1.025$$
$$= 3.69897 + 24 \times .01072$$
$$= 3.95634$$
$$\therefore \quad A = \text{Antilog } (3.95634) = \text{Rs. } 9043.63$$

Example 18 : On March 1, 1990, X borrowed Rs. 5000 at 12% compounded quartely. What did he owe on Ist Sept, 1997?

Solution : Given P = Rs. 5000

$$r = \frac{.12}{4} = .03$$

t = 7 years and 2 quarters = 30 quarters

$$\therefore \quad A = 5000\,(1 + .03)^{30} = \text{Rs. } 12136.32$$

Example 19: How long will it take the sum to double if compound continuously at 10% per annum?

Solution : Let Principal = Rs. P.

$$\therefore \quad \text{Amount} = P.e^{rt}$$
$$\Rightarrow 2\,P = Pe^{10t}$$
$$\Rightarrow 2 = e^{10t}$$

Taking log both sides

$$\log 2 = .10t \log e$$
$$\Rightarrow \quad t = \frac{\log 2}{.10 \log e} = \frac{.3010}{0.4343} = 6.93 \text{ years} \qquad [\log e = \log 2.7183 = .4343]$$

Example 20 : Mr. X deposits Rs. 5000 in a book for 3 years. Bank offers interest at the rate of 10% compounded quartely during first year, at 12% per annum compounded monthly during second year and at 15% per annum compounded continuously during 3rd year. Find his balance after 3 years.

Solution : Balance after Ist year

$$= 5000\left(1 + \frac{.10}{4}\right)^{24}$$

= principal for second year

Balance after second year

$$= 5000\left(1 + \frac{.10}{4}\right)^{4}\left(1 + \frac{.12}{12}\right)^{12}$$

= principal for third year

$\therefore$ Balance after 3 years

$$= 5000\left(1 + \frac{.10}{4}\right)^{4}\left(1 + \frac{.12}{12}\right)^{12}(e^{.15(1)})$$

$$= \text{Rs. } 5519.06(1.1268)\,(1.1618) = \text{Rs. } 7225.26$$

Equating the different cash flows involving compound interest

The equivalent value of different cash flows involving compound interest can also be obtained in the same way as in case of simple interest. In this case also, the flows on the left hand side are compounded and the flows on the right hand side are discounted, their present value is required.

Example 21 : In return of a promise to pay Rs. 800 at the end of 10 years, a person agrees to pay Rs. 100 now, Rs. 200 at the end of 6 years and a final payment at the end 12 years. If the rate of interest is 10% per annum what should the final payment be?

Solution : Let the final payment = Rs. x. and the focal date be 10th year.

$$\left.\begin{array}{l}\text{Therefore the future value of Rs.100} + \\ \text{Future value of Rs.200} + \\ \text{Present value of Rs.x}\end{array}\right] = \text{Rs.800}$$

$$\Rightarrow 800 = 100(1+.10)^{10} + 200(1+.10)^{4} + \frac{x}{(1+.10)^{2}}$$

$$\Rightarrow 800 = 259.37 + 292.82 + .8264x$$

$$\Rightarrow \quad x = \text{Rs.}299.85$$

Note-In such cases answer remains the same irrespective of the focal date.

Example 22: The sum of Rs. 2000, Rs. and Rs. 4000 are due at the end of 2, 4 and 8 years respectively. It is proposed to replace the series of payments by a single sum of Rs. 9000 payable at the end of n years. If the rate of interest is 10% per annum effective. Find the value of n.

Solution : One alternative is that the present value of given cash flows should be equal to the present value of Rs. 9000 flowing at the end of n^{th} year.

$$\therefore \qquad \frac{2000}{(1+.10)^{2}} + \frac{3000}{(1+.10)^{4}} + \frac{4000}{(1+.10)^{8}} = \frac{9000}{(1+.10)^{n}}$$

$$\text{or} \qquad 1652.89 + 2049.04 + 1866.03 = \frac{9000}{(1+.10)^{n}}$$

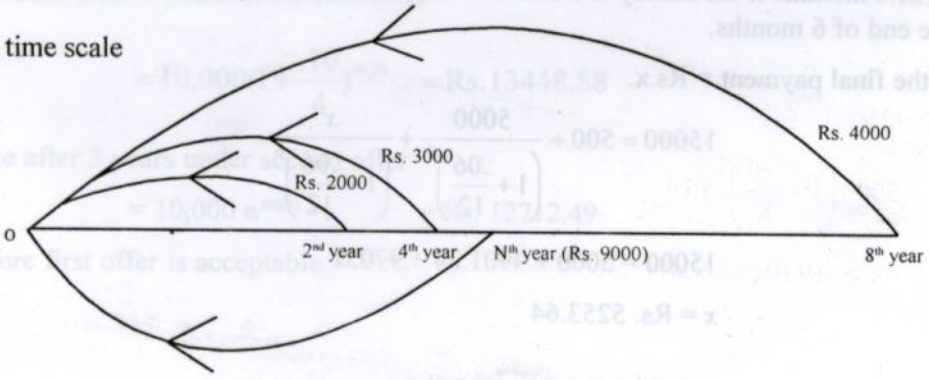
$$\text{or} \qquad (1.10)^{n} = \frac{9000}{5567.95} = 1.61639$$

Using log

$$n \log 1.10 = \log 1.61639$$

$$n = \frac{\log 1.61639}{\log 1.10} = \frac{.2085}{.0414} \qquad n = 5.038 \text{ years}$$

Using time scale



The other alternative is that the focal date may be taken at the end of n^{th} year. So the value of given flows at the end of nth year should be equal to Rs. 9000.

$$\therefore \quad 2000(1.10)^{n-2} + 3000(1.10)^{n-4} + 4000(1.10)^{n-8} = \text{Rs.}9000$$

Same expression is given by multiplying both sides by $(1+.10)^n$ in the first alternative. But it should be solved carefully. Each term is to be evaluated separately using log and not by taking log of both the sides (As log properties are not applicable in case of addition and subtraction).

Example 23 : If the money is worth 5% effective, what equal payments of Rs.x at the end of first year and at the end of third year will equitably replace the obligations: Rs. 30,000 at the end of 2nd year with interest at 6% per annum compounded semi-annually and Rs. 50,000 at the end of 5th year with interest at 4% compounded quarterly.

Solution : If the focal date is taken as zero period, then present value of both set of flows should be equal.

Amount due at the end of 2nd year $= 30{,}000\left(1+\frac{.06}{2}\right)^{2(2)}$

Its present value at 5% effective

$$= 30{,}000\frac{(1+.03)^4}{(1+.05)^2} = \text{Rs. } 30626.09$$

Amount due at the end of 5th year .....(i)

$$= 50{,}000\left(1+\frac{0.4}{4}\right)^{4(5)}$$

Its present value at 5% effective

$$= 50{,}000\frac{(1-.01)^{20}}{(1+.05)^5} = \text{Rs. } 47802.54 \quad .....(ii)$$

The total present value of both these obligations should be equal to the present value of both the payments of Rs. x

$$\therefore \quad \frac{x}{1+.05} + \frac{x}{(1+.05)^3} = 30626.09 + 47802.54$$

or $\quad 1.8162x = 78428.63$

or $\quad x = \text{Rs. } 43182.37$

Example 24 : A trader buys goods of the value of Rs. 15,000. He pays Rs. 5000 down and will pay Rs. 5000 at the end of 4 months. If the money is worth 6% compounded monthly, what kind of payment will be necessary at the end of 6 months.

Solution : Let the final payment = Rs.x.

$$\therefore \quad 15000 = 500 + \frac{5000}{\left(1+\frac{.06}{12}\right)^4} + \frac{x}{\left(1+\frac{.06}{12}\right)^6}$$

or $\quad 15000 = 5000 + 4901.24 + .9705x$

or $\quad x = \text{Rs. } 5253.64$

Nominal and the Effective rate of Interest

Nominal rate of interest is the rate of interest per annum which may be compounded yearly, six monthly, quarterly, monthly or m times in a year or continously. Effective rate of interest is the rate per annum compounded only once in a year. So both these are equal if compounding is only once a year. But if compounding is more than once in a year, effective rate of interest will increase as the frequency of compounding increases (at the decreasing rate). Following are the relationship between nominal (r) and effective (R) rate of interest under two different conditions.

(a) If compounding is m times in a year

then $$P(1+R)^1 = P(1+\frac{r}{m})^{m1}$$

or $$R = (1+\frac{r}{m})^{m-1}$$

(b) If compounding is continuous,

then $$P(1+R)^1 = Pe^{r1}$$

or $$R = e^r - 1$$

Example 25 : A money lender charges interest at the rate of 5 rupees per 100 rupees payable in advance per quarter. What effective rate of interest does he charge per annum?

Solution : Let the money borrowed = Rs. 100

$\therefore$ Interest deducted in advance for the quarter = Rs. 5

So net amount received by borrower = Rs. 95

For borrower interest rate $= \frac{5}{95}$ per quarter

$\therefore$ effective rate per annum will be

$$R = (1+\frac{r}{4})^4 - 1$$

$$= (1+\frac{5}{95})^4 - 1 = 1.2277 - 1 \quad = .2277 \quad = 22.77\%$$

Example 26 : Mr. X deposited Rs. 10,000 in a bank for 3 years. Bank makes two offers-either 10% compounded quarterly or 8% compounded continuously. Which offer is preferable?

Solution : Balance after 3 years under first-offer

$$= 10{,}000(1+\frac{.10}{4})^{4(3)} \quad = \text{Rs.}13448.88$$

Balance after 3 years under second offer

$$= 10{,}000\ e^{.08(3)} \qquad = \text{Rs. } 12712.49$$

Therefore first offer is acceptable.

Alternatively effective rate of interest per annum of both the proposals may be compared, Under first offer effective rate = $(1+\frac{.10}{4})^4 - 1 = .1038 = 10.38\%$

Under second offer, effective rate = $e^{.08} - 1 = .08328 = 8.328\%$

$\therefore$ First offer is acceptable.

Relationship between nominal rates

The same logic as explained above can also be used for establishing relationship between two nominal rates. For example, if interest is compounded quarterly at r_1 percent and the interest is compounded half yearly at r_2 percent, then relationship between the two are given below :

$$P(1+\frac{r_1}{4})^{4t} = P(1+\frac{r_2}{2})^{2t}$$

$$\frac{r_1}{4} = \sqrt{\left(1+\frac{r_2}{2}\right)} - 1$$

$$r_1 = 4\left[\sqrt{\left(1+\frac{r_2}{2}\right)} - 1\right]$$

Example 27: Find the nominal rate r compounded monthly equivalent to 6% compounded semi-annually.

Solution :

$$P(1+\frac{r}{12})^{12t} = P(1+\frac{.06}{2})^{2t}$$

$$\frac{r}{12} = (1.03)^{\frac{1}{6}} - 1$$

$$r = .05926 = 5.926\%$$

LESSON 1

CONSTRUCTION OF FREQUENCY DISTRIBUTION AND GRAPHICAL PRESENTATION

What is frequency distribution

Collected and classified data are presented in a form of frequency distribution. Frequency distribution is simply a table in which the data are grouped into classes on the basis of common characteristics and the number of cases which fall in each class are recorded. It shows the frequency of occurrence of different values of a single variable. A frequency distribution is constructed to satisfy three objectives :

(i) to facilitate the analysis of data,

(ii) to estimate frequencies of the unknown population distribution from the distribution of sample data, and

(iii) to facilitate the computation of various statistical measures.

Frequency distribution can be of two types :

1. Univariate Frequency Distribution.
2. Bivariate Frequency Distribution.

In this lesson, we shall understand the Univariate frequency distribution. Univariate distribution incorporates different values of one variable only whereas the Bivariate frequency distribution incorporates the values of two variables. The Univariate frequency distribution is further classified into three categories :

(i) Series of individual observations,

(ii) Discrete frequency distribution, and

(iii) Continuous frequency distribution.

Series of individual observations, is a simple listing of items of each observation. If marks of 14 students in statistics of a class are given individually, it will form a series of individual observations.

Marks obtained in Statistics :

Roll Nos.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Marks :	60	71	80	41	81	41	85	35	98	52	50	91	30	88

Marks in Ascending Order	Marks in Descending Order
30	98
35	91
41	88
41	85
50	81
52	80
60	71
71	60
80	52
81	50
85	41
88	41
91	35
98	30

Discrete Frequency Distribution: In a discrete series, the data are presented in such a way that exact measurements of units are indicated. In a discrete frequency distribution, we count the number of times each value of the variable in data given to you. This is facilitated through the technique of tally bars.

In the first column, we write all values of the variable. In the second column, a vertical bar called tally bar against the variable, we write a particular value has occurred four times, for the fifth occurrence, we put a cross tally mark (/) on the four tally bars to make a block of 5. The technique of putting cross tally bars at every fifth repetition facilitates the counting of the number of occurrences of the value. After putting tally bars for all the values in the data; we count the number of times each value is repeated and write it against the corresponding value of the variable in the third column entitled frequency. This type of representation of the data is called discrete frequency distribution.

We are given marks of 42 students:

55	51	57	40	26	43	46	41	46	48	33	40	26	40	40	41
43	53	45	53	33	50	40	33	40	26	53	59	33	39	55	48
15	26	43	59	51	39	15	45	26	15						

We can construct a discrete frequency distribution from the above given marks.

Marks of 42 Students

Marks	Tally Bars	Frequency
15	|||	3
26	𝍸	5
33	||||	4

39	||	2
40	𝍸	5
41	||	2
43	|||	3
45	||	2
46	||	2
48	||	2
50	|	1
51	||	2
53	|||	3
55	|||	3
57	|	1
59	||	2
	Total	**42**

The presentation of the data in the form of a discrete frequency distribution is better than arranging but it does not condense the data as needed and is quite difficult to grasp and comprehend. This distribution is quite simple in case the values of the variable are repeated otherwise there will be hardly any condensation.

Continuous Frequency Distribution: If the identity of the units about a particular information collected, is neither relevant nor is the order in which the observations occur, then the first step of condensation is to classify the data into different classes by dividing the entire group of values of the variable into a suitable number of groups and then recording the number of observations in each group. Thus, we divide the total range of values of the variable (marks of 42 students) i.e. 59–15 = 44 into groups of 10 each, then we shall get (42/10) 5 groups and the distribution of marks is displayed by the following frequency distribution:

Marks of 42 Students

Marks (×)	Tally Bars	Number of Students (*f*)
15—25	|||	3
25—35	𝍸 ||||	9
35—45	𝍸 𝍸 ||	12
45—55	𝍸 𝍸 ||	12
55—65	𝍸 |	6
	Total	**42**

The various groups into which the values of a variable are classified are known classes, the length of the class interval (10) is called the width of the class. Two values, specifying the class, are

called the class limits. The presentation of the data into continuous classes with the corresponding frequencies is known as continuous frequency distribution. There are two methods of classifying the data according to class intervals :

(i) exclusive method, and

(ii) inclusive method

In an exclusive method, the class intervals are fixed in such a manner that upper limit of one class becomes the lower limit of the following class. Moreover, an item equal to the upper limit of a class would be excluded from that class and included in the next class. The following data are classified on this basis.

Income (Rs.)	No. of Persons
200—250	50
250—300	100
300—350	70
350—400	130
400—450	50
450—500	100
Total	**500**

It is clear from the example that the exclusive method ensures continuity of the data in as much as the upper limit of one class is the lower limit of the next class. Therefore, 50 persons have their incomes between 200 to 249.99 and a person whose income is 250 shall be included in the next class of 250—300.

According to the inclusive method, an item equal to upper limit of a class is included in that class itself. The following table demonstrates this method.

Income (Rs.)	No.of Persons
200—249	50
250—299	100
300—349	70
350—399	130
400—449	50
450—499	100
Total	**500**

Hence in the class 200—249, we include persons whose income is between Rs. 200 and Rs. 249.

Principles for Constructing Frequency Distributions

Inspite of the great importance of classification in statistical analysis, no hard and fast rules are laid down for it. A statistician uses his discretion for classifying a frequency distribution and sound experience,

wisdom, skill and aptness for an appropriate classification of the data. However, the following guidelines must be considered to construct a frequency distribution:

1. *Type of classes:* The classes should be clearly defined and should not lead to any ambiguity. They should be exhaustive and mutually exclusive so that any value of variable corresponds to only class.

2. *Number of classes:* The choice about the number of classes in which a given frequency distribution should be divided depends upon the following things;

   (i) The total frequency which means the total number of observations in the distribution.
   (ii) The nature of the data which means the size or magnitude of the values of the variable.
   (iii) The desired accuracy.
   (iv) The convenience regarding computation of the various descriptive measures of the frequency distribution such as means, variance etc.

The number of classes should not be too small or too large. If the classes are few, the classification becomes very broad and rough which might obscure some important features and characteristics of the data. The accuracy of the results decreases as the number of classes becomes smaller. On the other hand, too many classes will result in a few frequencies in each class. This will give an irregular pattern of frequencies in different classes thus makes the frequency distribution irregular. Moreover a large number of classes will render the distribution too unwieldy to handle. The computational work for further processing of the data will become quite tedious and time consuming without any proportionate gain in the accuracy of the results. Hence a balance should be maintained between the loss of information in the first case and irregularity of frequency distribution in the second case, to arrive at a suitable number of classes. Normally, the number of classes should not be less than 5 and more than 20. Prof. Sturges has given a formula :

$$k = 1 + 3.322 \log n$$

where k refers to the number of classes and n refers to total frequencies or number of observations. The value of k is rounded to the next higher integer :

If $n = 100$ $\quad k = 1 + 3.322 \log 100 = 1 + 6.644 = 8$

If $n = 10{,}000$ $\quad k = 1 + 3.22 \log 10{,}000 = 1 + 13.288 = 14$

However, this rule should be applied when the number of observations are not very small.

Further, the number or class intervals should be such that they give uniform and unimodal distribution which means that the frequencies in the given classes increase and decrease steadily and there are no sudden jumps. The number of classes should be an integer preferably 5 or multiples of 5, 10, 15, 20, 25 etc. which are convenient for numerical computations.

3. *Size of Class Intervals* : Because the size of the class interval is inversely proportional to the number of classes in a given distribution, the choice about the size of the class interval will depend upon the sound subjective judgment of the statistician. An approximate value of the magnitude of the class interval say i can be calculated with the help of Sturge's Rule :

$$i = \frac{\text{Range}}{1 + 3.22 \log n}$$

where i stands for class magnitude or interval, Range refers to the difference between the largest and smallest value of the distribution, and n refers to total number of observations.

If we are given the following information; $n = 400$, Largest item = 1300 and Smallest item = 340.

then, $$i = \frac{1300 - 340}{1 + 3.22 \log 400} = \frac{960}{1 + 3.222 \times 2.6021} = \frac{960}{9{,}644} = 99.54 \text{ (100 approx.)}$$

Another rule to determine the size of class interval is that the length of the class interval should not be greater than $\frac{1}{4}$th of the estimated population standard deviation. If 6 is the estimate of population standard deviation then the length of class interval is given by: $i \leq 6/4$.

The size of class intervals should be taken as 5 or multiples of 5, 10, 15 or 20 for easy computations of various statistical measures of the frequency distribution, class intervals should be so fixed that each class has a convenient mid-point around which all the observations in that class cluster. It means that the entire frequency of the class is concentrated at the mid value of the class. It is always desirable to take the class intervals of equal or uniform magnitude throughout the frequency distribution.

4. *Class Boundaries:* If in a grouped frequency distribution there are gaps between the upper limit of any class and lower limit of the succeeding class (as in case of inclusive type of classification), there is a need to convert the data into a continuous distribution by applying a correction factor for continuity for determining new classes of exclusive type. The lower and upper class limits of new exclusive type classes are called class boundaries.

If d is the gap between the upper limit of any class and lower limit of succeeding class, the class boundaries for any class are given by:

$$\left.\begin{aligned} &\text{Upper class boundary} = \text{Upper class limit} + \tfrac{1}{2}d \\ &\text{Lower class boundary} = \text{Lower class limit} - \tfrac{1}{2}d \end{aligned}\right\}$$

d/2 is called the correction factor.

Let us consider the following example to understand :

Marks	Class Boundaries
20—24	(20—0.5, 24 + 0.5) i.e., 19.5—24.5
25—29	(25—0.5, 29 + 0.5) i.e., 24.5—29.5
30—34	(30—0.5, 34 + 0.5) i.e., 29.5—34.5
35—39	(35—0.5, 39 + 0.5) i.e., 34.5—39.5
40—44	(40—0.5, 44 + 0.5) i.e., 39.5—44.5

$$\text{Correction factor} = \frac{d}{2} = \frac{35 - 34}{2} = \frac{1}{2} = 0.5$$

5. *Mid-value or Class Mark:* The mid value or class mark is the value of a variable which is exactly at the middle of the class. The mid-value of any class is obtained by dividing the sum of the upper and lower class limits by 2.

Mid value of a class = ½ [Lower class limit + Upper class limit]

The class limits should be selected in such a manner that the observations in any class are evenly distributed throughout the class interval so that the actual average of the observations in any class is very close to the mid-value of the class.

6. *Open End Classes* : The classification is termed as open end classification if the lower limit of the first class or the upper limit of the last class or both are not specified and such classes in which one of the limits is missing are called open end classes. For example, the classes like the marks less than 20 or age above 60 years. As far as possible open end classes should be avoided because in such classes the mid-value cannot be accurately obtained. But if the open end classes are inevitable then it is customary to estimate the class mark or mid-value for the first class with reference to the succeeding class. In other words, we assume that the magnitude of the first class is same as that of the second class.

Example: Construct a frequency distribution from the following data by inclusive method taking 4 as the class interval:

10	17	15	22	11	16	19	24	29	18
25	26	32	14	17	20	23	27	30	12
15	18	24	36	18	15	21	28	33	38
34	13	10	16	20	22	29	19	23	31

Solution: Because the minimum value of the variable is 10 which is a very convenient figure for taking the lower limit of the first class and the magnitude of the class interval is given to be 4, the classes for preparing frequency distribution by the Inclusive method will be 10—13, 14—17, 18—21, 22—25, ... 38—41.

Frequency Distribution

Class Interval	Tally Bars	Frequency (*f*)
10—13	𝍸	5
14—17	𝍸 III	8
18—21	𝍸 III	8
22—25	𝍸 II	7
26—29	𝍸	5
30—33	IIII	4
34—37	II	2
38—41	I	1

Example: Prepare a statistical table from the following :

Weekly wages (Rs.) of 100 workers of Factory A

88	23	27	28	86	96	94	93	86	99
82	24	24	55	88	99	55	86	82	36
96	39	26	54	87	100	56	84	83	46
102	48	27	26	29	100	59	83	84	48
104	46	30	29	40	101	60	89	46	49
106	33	36	30	40	103	70	90	49	50
104	36	37	40	40	106	72	94	50	60
24	39	49	46	66	107	76	96	46	67
26	78	50	44	43	46	79	99	36	68
29	67	56	99	93	48	80	102	32	51

Solution: The lowest value is 23 and the highest 106. The difference between the lowest and highest value is 83. If we take a class interval of 10, nine classes would be made. The first class should be taken as 20—30 instead of 23—33 as per the guidelines of classification.

Frequency Distribution of the Wages of 100 Workers

Wages (Rs.)	Tally Bars	Frequency (f)
20—30	𝍸𝍸 |||	13
30—40	𝍸𝍸 |	11
40—50	𝍸 𝍸 𝍸 |||	18
50—60	𝍸𝍸	10
60—70	𝍸 |	6
70—80	𝍸	5
80—90	𝍸𝍸 |||	14
90—100	𝍸𝍸 ||	12
100—110	𝍸𝍸 |	11
	Total	**100**

Graphs of Frequency Distributions

The guiding principles for the graphic representation of the frequency distributions are same as for the diagrammatic and graphic representation of other types of data. The information contained in a frequency distribution can be shown in graphs which reveals the important characteristics and relationships that are not easily discernible on a simple examination of the frequency tables. The most commonly used graphs for charting a frequency distribution are :

1. Histogram
2. Frequency polygon
3. Smoothed frequency curves
4. Ogives or cumulative frequency curves.

1. Histogram

The term 'histogram' must not be confused with the term 'historigram' which relates to time charts. Histogram is the best way of presenting graphically a simple frequency distribution. The statistical meaning of histogram is that it is a graph that represents the class frequencies in a frequency distribution by vertical adjacent rectangles.

While constructing histogram the variable is always taken on the X-axis and the corresponding frequencies on the Y-axis. Each class is then represented by a distance on the scale that is proportional to its class-interval. The distance for each rectangle on the X-axis shall remain the same in case the class-intervals are uniform throughout; if they are different the width of the rectangles shall also change proportionately. TheY-axis represents the frequencies of each class which constitute the height of its rectangle. We get a series of rectangles each having a class interval distance as its width and the frequency distance as its height. The area of the histogram represents the total frequency.

The histogram should be clearly distinguished from a bar diagram. A bar diagram is one-dimensional where the length of the bar is important and not the width, a histogram is two-dimensional, where both the length and the width are important. However, a histogram can be misleading if the distribution has unequal class intervals and suitable adjustments in frequencies are not made.

The technique of constructing histogram is explained for :

(i) distributions having equal class-intervals, and

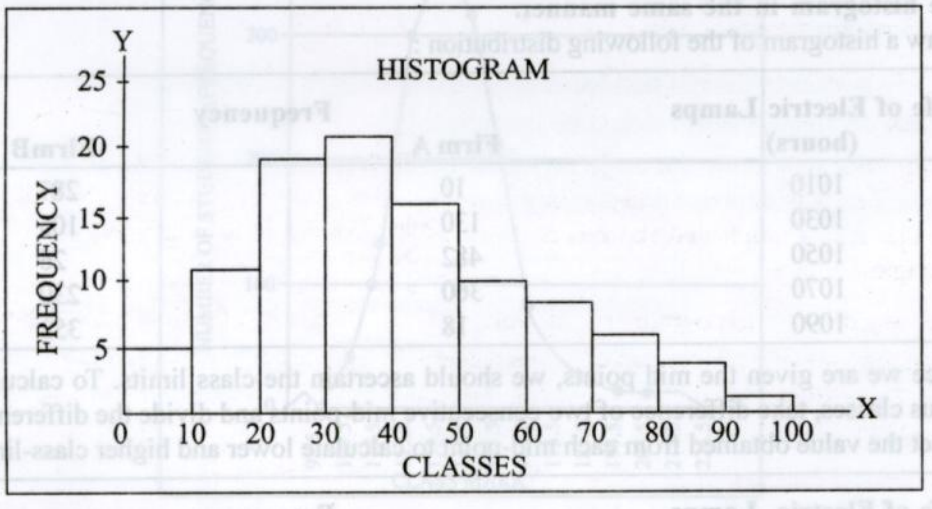
(ii) distributions having unequal class-intervals.

When class-intervals are equal, take frequency on the Y-axis, the variable on the X-axis and construct rectangles. In such a case the heights of the rectangles will be proportional to the frequencies.

Example: Draw a histogram from the following data :

Classes	Frequency
0—10	5
10—20	11
20—30	19
30—40	21
40—50	16
50—60	10
60—70	8
70—80	6
80—90	3
90—100	1

Solution :



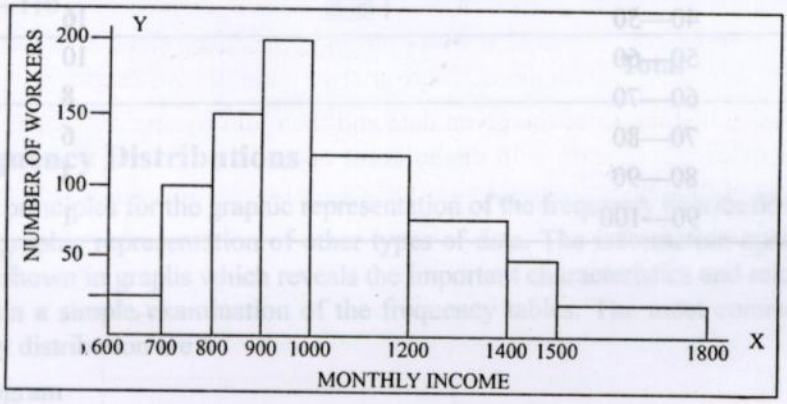
When class-intervals are unequal the frequencies must be adjusted before constructing a histogram. We take that class which has the lowest class-interval and adjust the frequencies of other classes accordingly. If one class interval is twice as wide as the one having the lowest class-interval we divide the height of its rectangle by two, if it is three times more we divide it by three etc., the heights will be proportional to the ratios of the frequencies to the width of the classes.

Example: Represent the following data on a histogram.

Average monthly income of 1035 employees in a construction industry is given below:

Monthly Income (Rs.)	No. of Workers
600—700	25
700—800	100
800—900	150
900—1000	200
1000—1200	240
1200—1400	160
1400—1500	50
1500—1800	90
1800 or more	20

Solution: Histogram showing monthly incomes of workers :



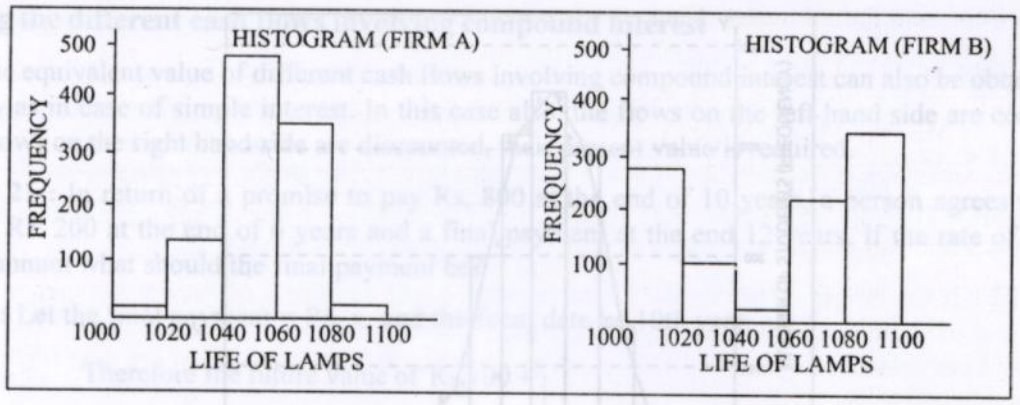
When mid point are given, we ascertain the upper and lower limits of each class and then construct the histogram in the same manner.

Example: Draw a histogram of the following distribution :

Life of Electric Lamps (hours)	Frequency	
	Firm A	FirmB
1010	10	287
1030	130	105
1050	482	26
1070	360	230
1090	18	352

Solution: Since we are given the mid points, we should ascertain the class limits. To calculate the class limits of various classes, take difference of two consecutive mid-points and divide the difference by 2, then add and subtract the value obtained from each mid-point to calculate lower and higher class-limits.

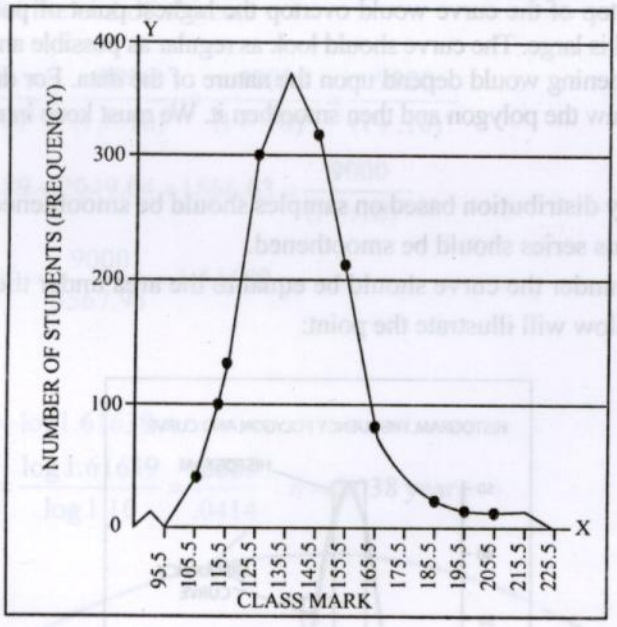
Life of Electric Lamps (hours)	Frequency	
	Firm A	FirmB
1000—1020	10	287
1020—1040	130	105
1040—1060	482	76
1060—1080	360	230
1080—1100	18	352



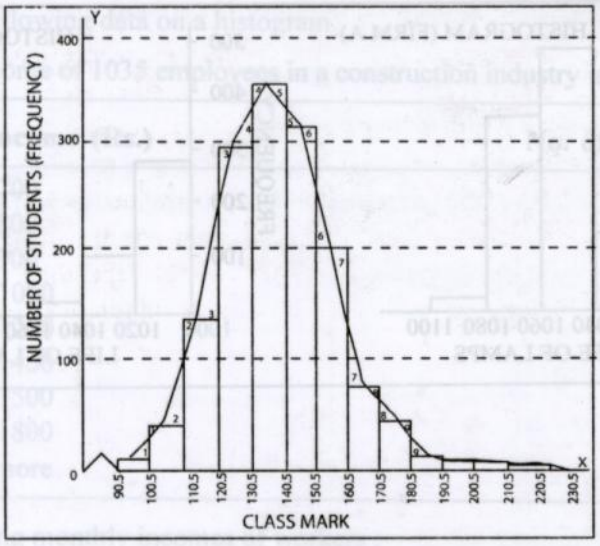
2. Frequency Polygon

This is a graph of frequency distribution which has more than four sides. It is particularly effective in comparing two or more frequency distributions. There are two ways of constructing a frequency polygon.

(i) We may draw a histogram of the given data and then join by straight line the mid-points of the upper horizontal side of each rectangle with the adjacent ones. The figure so formed shall be frequency polygon. Both the ends of the polygon should be extended to the base line in order to make the area under frequency polygons equal to the area under Histogram.



(ii) Another method of constructing frequency polygon is to take the mid-points of the various class-intervals and then plot the frequency corresponding to each point and join all these points by straight lines. The figure obtained by both the methods would be identical.



Frequency polygon has an advantage over the histogram. The frequency polygons of several distributions can be drawn on the same axis, which makes comparisons possible whereas histogram cannot be used in the same way. To compare histograms we need to draw them on separate graphs.

3. Smoothed Frequency Curve

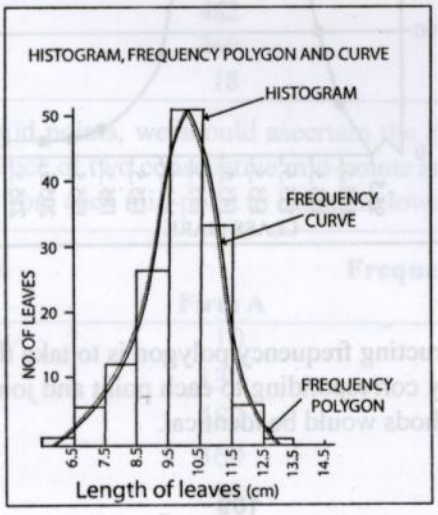
A smoothed frequency curve can be drawn through the various points of the polygon. The curve is drawn by free hand in such a manner that the area included under the curve is approximately the same as that of the polygon. The object of drawing a smoothed curve is to eliminate all accidental variations which exists in the original data, while smoothening, the top of the curve would overtop the highest point of polygon particularly when the magnitude of the class interval is large. The curve should look as regular as possible and all sudden turns should be avoided. The extent of smoothening would depend upon the nature of the data. For drawing smoothed frequency curve it is necessary to first draw the polygon and then smoothen it. We must keep in mind the following points to smoothen a frequency graph:

(i) Only frequency distribution based on samples should be smoothened.

(ii) Only continuous series should be smoothened.

(iii) The total area under the curve should be equal to the area under the histogram or polygon.

The diagram given below will illustrate the point:



4. Cumulative Frequency Curves or Ogives

We have discussed the charting of simple distributions where each frequency refers to the measurement of the class-interval against which it is placed. Sometimes it becomes necessary to know the number of items whose values are greater or less than a certain amount. We may, for example, be interested in knowing the number of students whose weight is less than 65 lbs. or more than say 15.5 lbs. To get this information, it is necessary to change the form of frequency distribution from a simple to a cumulative distribution. In a cumulative frequency distribution, the frequency of each class is made to include the frequencies of all the lower or all the upper classes depending upon the manner in which cumulation is done. The graph of such a distribution is called a cumulative frequency curve or an Ogive. There are two method of constructing ogives, namely:

(i) less than method, and

(ii) more than method.

In less than method, we start with the upper limit of each class and go on adding the frequencies. When these frequencies are plotted we get a rising curve.

In more than method, we start with the lower limit of each class and we subtract the frequency of each class from total frequencies. When these frequencies are plotted, we get a declining curve.

This example would illustrate both types of ogives.

Example: Draw ogives by both the methods from the following data.

Distribution of weights of the students of a college (lbs.)

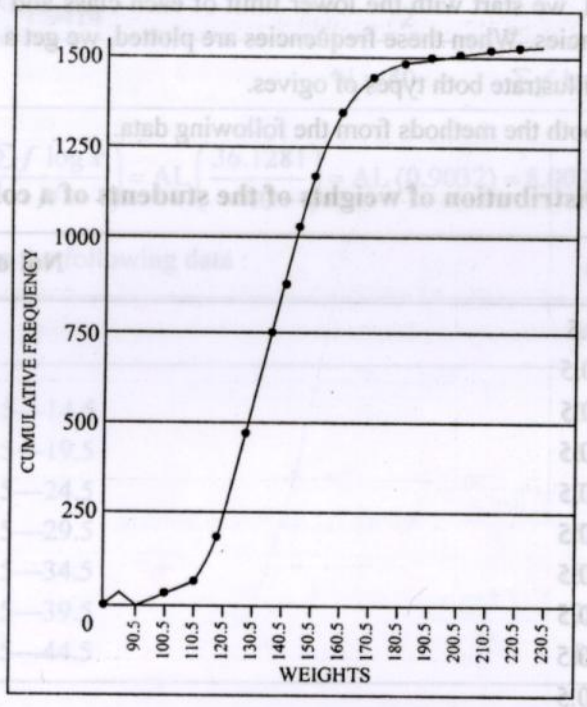
Weights	No. of Students
90.5—100.5	5
100.5—110.5	34
110.5—120.5	139
120.5—130.5	300
130.5—140.5	367
140.5—150.5	319
150.5—160.5	205
160.5—170.5	76
170.5—180.5	43
180.5—190.5	16
190.5—200.5	3
200.5—210.5	4
210.5—220.5	3
220.5—230.5	1

Solution: First of all we shall find out the cumulative frequencies of the given data by less than method.

Less than (Weights)	Cumulative Frequency
100.5	5
110.5	39

120.5	178
130.5	478
140.5	845
150.5	1164
160.5	1369
170.5	1445
180.5	1488
190.5	1504
200.5	1507
210.5	1511
220.5	1514
230.5	1515

Plot these frequencies and weights on a graph paper. The curve formed is called an Ogive.

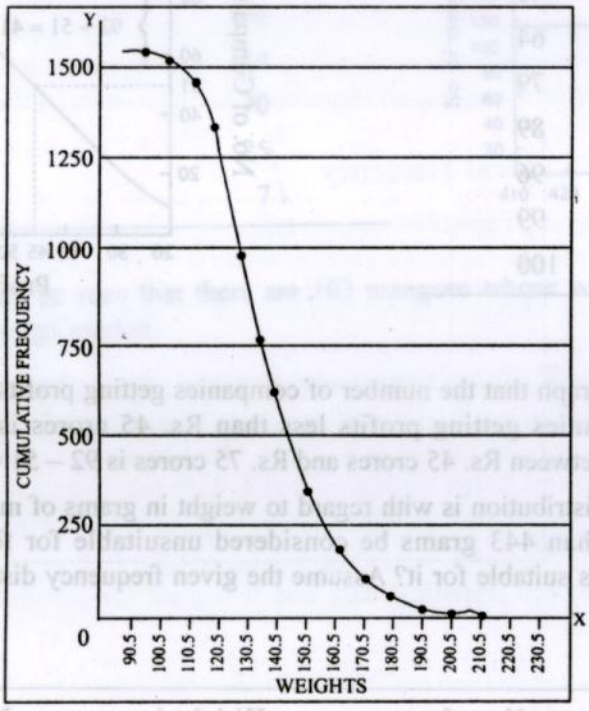


Now we calculate the cumulative frequencies of the given data by more than method.

More than (Weights)	Cumulative Frequencies
90.5	1515
100.5	1510
110.5	1476
120.5	1337

130.5	1037
140.5	670
150.5	351
160.5	146
170.5	70
180.5	27
190.5	11
200.5	8
210.5	4
220.5	1

By plotting these frequencies on a graph paper, we will get a declining curve which will be our cumulative frequency curve or Ogive by more than method.



Although the graphs are a powerful and effective method of presenting statistical data, they are not under all circumstances and for all purposes complete substitutes for tabular and other forms of presentation. The specialist in this field is one who recognizes not only the advantages but also the limitations of these techniques. He knows when to use and when not to use these methods and from his experience and expertise is able to select the most appropriate method for every purpose.

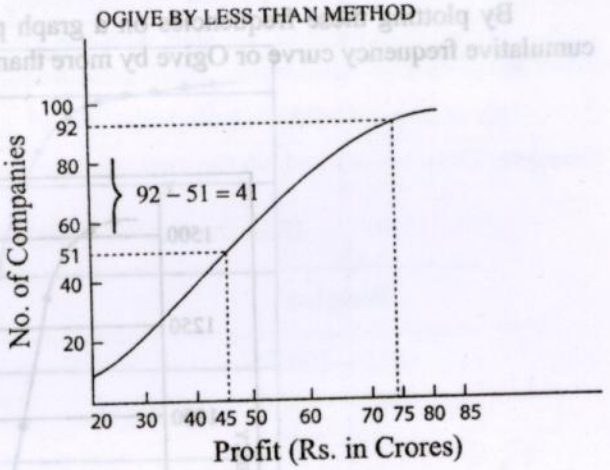
Example: Draw an ogive by less than method and determine the number of companies getting profits between Rs. 45 crores and Rs. 75 crores :

Profits (Rs. crores)	No. of Companies	Profits (Rs. crores)	No. of Companies
10—20	8	60—70	10
20—30	12	70—80	7
30—40	20	80—90	3
40—50	24	90—100	1
50—60	15		

Solution:

OGIVE BY LESS THAN METHOD

Profits (Rs. crores)	No. of Companies
Less than 20	8
Less than 30	20
Less than 40	40
Less than 50	64
Less than 60	79
Less than 70	89
Less than 80	96
Less than 90	99
Less than 100	100



It is clear from the graph that the number of companies getting profits less than Rs.75 crores is 92 and the number of companies getting profits less than Rs. 45 crores is 51. Hence the number of companies getting profits between Rs. 45 crores and Rs. 75 crores is 92 – 51 = 41.

Example: The following distribution is with regard to weight in grams of mangoes of a given variety. If mangoes of weight less than 443 grams be considered unsuitable for foreign market, what is the percentage of total mangoes suitable for it? Assume the given frequency distribution to be typical of the variety:

Weight in gms.	No. of mangoes	Weight in gms.	No. of mangoes
410—419	10	450—459	45
420—429	20	460—469	18
430—439	42	470—479	7
440—449	54		

Draw an ogive of 'more than' type of the above data and deduce how many mangoes will be more than 443 grams.

Solution: Mangoes weighting more than 443 gms. are suitable for foreign market. Number of mangoes weighting more than 443 gms. lies in the last four classes. Number of mangoes weighing between 444 and 449 grams would be

$$\frac{6}{10} \times 54 = \frac{324}{10} = 32.4$$

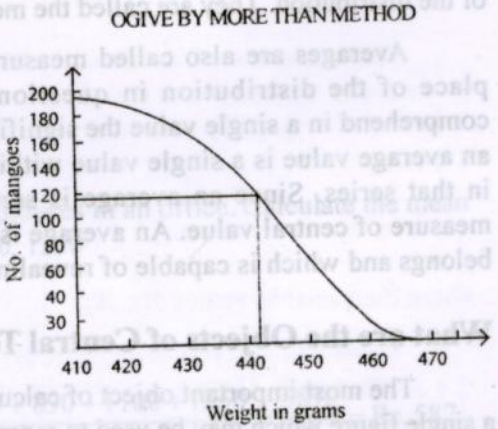
Total number of mangoes weighing more than 443 gms. = 32.4 + 45 + 18 + 7 = 102.4

$$\text{Percentage of mangoes} = \frac{102.4}{196} \times 100 = 52.25$$

Therefore, the percentage of the total mangoes suitable for foreign market is 52.25.

OGIVE BY MORE THAN METHOD

Weight more than (gms.)	No. of Mangoes
410	196
420	186
430	166
440	124
450	70
460	25
470	7



From the graph it can be seen that there are 103 mangoes whose weight will be more than 443 gms. and are suitable for foreign market.

LESSON 2

MEASURES OF CENTRAL TENDENCY

What is Central Tendency

One of the important objectives of statistics is to find out various numerical values which explains the inherent characteristics of a frequency distribution. The first of such measures is averages. The averages are the measures which condense a huge unwieldy set of numerical data into single numerical values which represent the entire distribution. The inherent inability of the human mind to remember a large body of numerical data compels us to few constants that will describe the data. Averages provide us the gist and give a bird's eye view of the huge mass of unwieldy numerical data. Averages are the typical values around which other items of the distribution congregate. This value lie between the two extreme observations of the distribution and give us an idea about the concentration of the values in the central part of the distribution. They are called the measures of central tendency.

Averages are also called measures of location since they enable us to locate the position or place of the distribution in question. Averages are statistical constants which enables us to comprehend in a single value the significance of the whole group. According to Croxton and Cowden, an average value is a single value within the range of the data that is used to represent all the values in that series. Since an average is somewhere within the range of data, it is sometimes called a measure of central value. An average is the most typical representative item of the group to which it belongs and which is capable of revealing all important characteristics of that group or distribution.

What are the Objects of Central Tendency

The most important object of calculating an average or measuring central tendency is to determine a single figure which may be used to represent a whole series involving magnitudes of the same variable.

Second object is that an average represents the entire data, it facilitates comparison within one group or between groups of data. Thus, the performance of the members of a group can be compared with the average performance of different groups.

Third object is that an average helps in computing various other statistical measures such as dispersion, skewness, kurtosis etc.

Essential of a Good Average

An average represents the statistical data and it is used for purposes of comparison, it must possess the following properties.

1. It must be rigidly defined and not left to the mere estimation of the observer. If the definition is rigid, the computed value of the average obtained by different persons shall be similar.
2. The average must be based upon all values given in the distribution. If the item is not based on all value it might not be representative of the entire group of data.
3. It should be easily understood. The average should possess simple and obvious properties. It should be too abstract for the common people.
4. It should be capable of being calculated with reasonable care and rapidity.
5. It should be stable and unaffected by sampling fluctuations.
6. It should be capable of further algebraic manipulation.

Different methods of measuring "Central Tendency" provide us with different kinds of averages. The following are the main types of averages that are commonly used:

1. Mean
   - (i) Arithmetic mean
   - (ii) Weighted mean
   - (iii) Geometric mean
   - (iv) Harmonic mean
2. Median
3. Mode

Arithmetic Mean: The arithmetic mean of a series is the quotient obtained by dividing the sum of the values by the number of items. In algebraic language, if $X_1, X_2, X_3 \ldots\ldots X_n$ are the n values of a variate X, then the Arithmetic Mean $(\overline{X})$ is defined by the following formula:

$$\overline{X} = \frac{1}{n}(X_1 + X_2 + X_3 + \ldots\ldots\ldots\ldots + X_n)$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right) = \frac{\Sigma X}{N}$$

Example : The following are the monthly salaries (Rs.) of ten employees in an office. Calculate the mean salary of the employees: 250, 275, 265, 280, 400, 490, 670, 890, 1100, 1250

Solution: $\overline{X} = \dfrac{\Sigma X}{N}$

$$\overline{X} = \frac{250 + 275 + 265 + 280 + 400 + 490 + 670 + 890 + 1100 + 1250}{10} = \frac{5870}{10} = \text{Rs. } 587$$

Short-cut Method: Direct method is suitable where the number of items is moderate and the figures are small sizes and integers. But if the number of items is large and/or the values of the variate are big, then the process of adding together all the values may be a lengthy process. To overcome this difficulty of computations, a short-cut method may be used. Short cut method of computation is based on an important characteristic of the arithmetic mean, that is, *the algebraic sum of the deviations of a series of individual observations from their mean is always equal to zero.* Thus deviations of the various values of the variate from an assumed mean computed and the sum is divided by the number of items. The quotient obtained is added to the assumed mean to find the arithmetic mean. Symbolically, $\overline{X} = A + \dfrac{\Sigma\, dx}{N}$, where A is assumed mean and dx are deviations = (X – A).

We can solve the previous example by short-cut method.

Computation of Arithmetic Mean

Serial Number	Salary (Rupees) X	Deviations from assumed mean where dx $(X - A)$, $A = 400$
1.	250	–150
2.	275	–125
3.	265	–135
4.	280	–120

5.	400	0
6.	490	+90
7.	670	+270
8.	890	+490
9.	1100	+700
10.	1250	+850
N = 10		$\sum dx = 1870$

$$\overline{X} = A + \frac{\sum dx}{N}$$

By substituting the values in the formula, we get

$$\overline{X} = 400 + \frac{1870}{10} = \text{Rs. } 587$$

Computation of Arithmetic Mean in Discrete series. In discrete series, arithmetic mean may be computed by both direct and short cut methods. The formula according to direct method is:

$$\overline{X} = \frac{1}{n}(f_1 X_1 + f_2 X_2 + \ldots\ldots\ldots\ldots + f_n X_n) = \frac{\sum (f X)}{N}$$

where the variable values X_1, X_2, ... X_n have frequencies f_1, f_2, ... f_n and $N = \sum f$.

Example. The following table gives the distribution of 100 accidents during seven days of the week in a given month. During a particular month there were 5 Fridays and Saturdays and only four each of other days. Calculate the average number of accidents per day.

Days :	Sun.	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.		Total
Number of accidents :	20	22	10	9	11	8	20	=	100

Solution: **Calculation of Number of Accidents per Day**

Day	**No. of Accidents** X	**No. of Days in Month** f	**Total Accidents** fX
Sunday	20	4	80
Monday	22	4	88
Tuesday	10	4	40
Wednesday	9	4	36
Thursday	11	4	44
Friday	8	5	40
Saturday	20	5	100
	100	N = 30	$\sum fX = 428$

$$\bar{X} = \frac{\Sigma f X}{N} = \frac{428}{30} = 14.27 = 14 \text{ accidents per day}$$

The formula for computation of arithmetic mean according to the short cut method is

$$\bar{X} = A + \frac{\Sigma fdx}{N} \text{ where A is assumed mean, } dx = (X - A) \text{ and } N = \Sigma f.$$

We can solve the previous example by short-cut method as given below:

Calculation of Average Accidents per day

Day	X	$dx = X - A$ (where $A = 10$)	f	fdx
Sunday	20	+ 10	4	+ 40
Monday	22	+ 12	4	+ 48
Tuesday	10	+ 0	4	+ 0
Wednesday	9	− 1	4	− 4
Thursday	11	+ 1	4	+ 4
Friday	8	− 2	5	− 10
Saturday	20	+ 10	5	+ 50
			30	+ 128

$$\bar{X} = A + \frac{\Sigma fdx}{N} = 10 + \frac{128}{30} = 14.27 = 14 \text{ accidents per day}$$

Calculation of arithmetic mean for Continuous Series: The arithmetic mean can be computed both by direct and short-cut method. In addition, a coding method or step deviation method is also applied for simplification of calculations. In any case, it is necessary to find out the mid-values of the various classes in the frequency distribution before arithmetic mean of the frequency distribution can be computed. Once the mid-points of various classes are found out, then the process of the calculation of arithmetic mean is same as in the case of discrete series. In case of direct method, the formula to be used:

$$\bar{X} = \frac{\Sigma fm}{N}, \text{ when } m = \text{mid points of various classes and N = total frequency}$$

In the short-cut method, the following formula is applied:

$$\bar{X} = A + \frac{\Sigma fdx}{N} \text{ where } dx = (m - A) \text{ and } N = \Sigma f$$

The short-cut method can further be simplified in practice and is named coding method. The deviations from the assumed mean are divided by a common factor to reduce their size. The sum of the products of the deviations and frequencies is multiplied by this common factor and then it is divided by the total frequency and added to the assumed mean. Symbolically

$$\bar{X} = A + \frac{\Sigma fd'x}{N} \times i, \text{ where } d'x = \frac{m - A}{i} \text{ and } i = \text{common factor}$$

Example. Following is the frequency distribution of marks obtained by 50 students in a test of Statistics:

Marks	Number of Students
0—10	4
10—20	6
20—30	20
30—40	10
40—50	7
50—60	3

Calculate arithmetic mean by;

(i) direct method,

(ii) short-cut method, and

(iii) coding method

Solution: **Calculation of Arithmetic Mean**

X	f	m	fm	$dx = m - A$ (where $A = 25$)	$d'x = \frac{m-A}{i}$ where $i = 10$	fdx	$fd'x$
0—10	4	5	20	− 20	− 2	− 80	− 8
10—20	6	15	90	− 10	− 1	− 60	− 6
20—30	20	25	500	0	0	0	0
30—40	10	35	350	+ 10	+ 1	100	+ 10
40—50	7	45	315	+ 20	+ 2	140	+ 14
50—60	3	55	165	+ 30	+ 3	90	+ 9
	N = 50		$\Sigma fm = 1440$			$\Sigma fdx = 190$	$\Sigma f d'x = +19$

Direct Method:

$$\bar{X} = \frac{\Sigma fm}{N} = \frac{1440}{50} = 28.8 \text{ marks.}$$

Short-cut Method:

$$\bar{X} = A + \frac{\Sigma fdx}{N} = 25 + \frac{190}{50} = 28.8 \text{ marks.}$$

Coding Method:

$$\bar{X} = A + \frac{\Sigma fd'x}{N} \times i = 25 + \frac{19}{50} \times 10 = 25 + 3.8 = 28.8 \text{ marks.}$$

We can observe that answer of average marks i.e. 28.8 is identical by all methods.

Mathematical Properties of the Arithmetic Mean

(i) The sum of the deviation of a given set of individual observations from the arithmetic mean is always zero.

Symbolically, $\Sigma (X - \bar{X}) = 0$. It is due to this property that the arithmetic mean is characterised as the centre of gravity i.e., the sum of positive deviations from the mean is equal to the sum of negative deviations.

(ii) The sum of squares of deviations of a set of observations is the minimum when deviations are taken from the arithmetic average. Symbolically, $\Sigma (X - \bar{X})^2$ = smaller than $\Sigma (X$ – any other value$)^2$.

We can verify the above properties with the help of the following data:

Values	Deviations from $\bar{X}$		Deviations from Assumed Mean	
X	$(X - \bar{X})$	$\Sigma (X - \bar{X})^2$	$(X - A)$	$\Sigma (X - A)^2$
3	– 6	36	– 7	49
5	– 4	16	– 5	25
10	1	1	0	0
12	3	9	2	4
15	6	36	5	25
Total = 45	0	98	– 5	103

$$\bar{X} = \frac{\Sigma X}{n} = \frac{45}{5} = 9, \text{ where A (assumed mean)} = 10$$

(iii) If each value of a variable X is increased or decreased or multiplied by a constant k, the arithmetic mean also increases or decreases or multiplies by the same constant.

(iv) If we are given the arithmetic mean and number of items of two or more groups, we can compute the combined average of these groups by apply the following formula:

$$\bar{X}_{12} = \frac{N_1 \bar{X}_1 + N_2 \bar{X}_2}{N_1 + N_2}$$

where $\bar{X}_{12}$ refers to combined average of two groups,

$\bar{X}_1$ refers to arithmetic mean of first group,

$\bar{X}_2$ refers to arithmetic mean of second group,

N_1 refers to number of items of first group, and

N_2 refers to number of items of second group

We can understand the property with the help of the following examples.

Example. The average marks of 25 male students in a section is 61 and average marks of 35 female students in the same section is 58. Find combined average marks of 60 students.

Solution: We are given the following information,

$$\bar{X}_1 = 61, \quad N_1 = 25, \quad \bar{X}_2 = 58, \quad N_2 = 35$$

Apply $\bar{X}_{12} = \frac{N_1 \bar{X}_1 + N_2 \bar{X}_2}{N_1 + N_2} = \frac{(25 \times 61) + (35 \times 58)}{25 + 35} = 59.25$ marks.

Example: The mean wage of 100 workers in a factory, running two shifts of 60 and 40 workers respectively is Rs.38. The mean wage of 60 workers in morning shift is Rs. 40. Find the mean wage of 40 workers working in the evening shift.

Solution: We are given the following information,

$\bar{X}_1 = 40$, $N_1 = 60$, $\bar{X}_2 = ?$, $N_2 = 40$, $\bar{X}_{12} = 38$, and $N = 100$

Apply $\bar{X}_{12} = \dfrac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$

$$38 = \frac{(60 \times 40) + (40\bar{X}_2)}{60 + 40} \quad \text{or} \quad 3800 = 2400 + 40\bar{X}_2$$

$$\bar{X}_2 = \frac{3800 - 2400}{40} = 35.$$

Example: The mean age of a combined group of men and women is 30 years. If the mean age of the group of men is 32 and that of women group is 27, find out the percentage of men and women in the group.

Solution: Let us take group of men as first group and women as second group. Therefore, $\bar{X}_1 = 32$ years, $\bar{X}_2 = 27$ years, and $\bar{X}_{12} = 30$ years. In the problem, we are not given the number of men and women. We can assume $N_1 + N_2 = 100$ and therefore, $N_1 = 100 - N_2$

Apply $\bar{X}_{12} = \dfrac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$

$$30 = \frac{32N_1 + 27N_2}{100} \qquad (\text{Substitute } N_1 = 100 - N_2)$$

$$30 \times 100 = 32\,(100 - N_2) + 27\,N_2 \qquad \text{or} \qquad 5N_2 = 200$$

$$N_2 = 200/5 = 40\%$$

$$N_1 = (100 - N_2) = (100 - 40) = 60\%$$

Therefore, the percentage of men in the group is 60 and that of women is 40.

(v) Because $\bar{X} = \dfrac{\Sigma X}{N}$

$$\therefore \quad \Sigma X = N.\bar{X}$$

If we replace each item in the series by the mean, the sum of these substitutions will be equal to the sum of the individual items. This property is used to find out the aggregate values and corrected averages. We can understand the property with the help of an example.

Example: Mean of 100 observations is found to be 44. If at the time of computation two items are wrongly taken as 30 and 27 in place of 3 and 72. Find the corrected average.

Solution: $\bar{X} = \dfrac{\Sigma X}{N}$

$$\therefore \quad \Sigma X = N.\bar{X} = 100 \times 44 = 4400$$

Corrected $\Sigma X = \Sigma X$ + correct items − wrong items $= 4400 + 3 + 72 - 30 - 27 = 4418$

$$\text{Corrected average} = \frac{\text{Corrected } \Sigma X}{N} = \frac{4418}{100} = 44.18$$

Calculation of Arithmetic mean in Case of Open-End Classes

Open-end classes are those in which lower limit of the first class and the upper limit of the last class are not defined. In these series, we can not calculate mean unless we make an assumption about the unknown limits. The assumption depends upon the class-interval following the first class and preceding the last class. For example:

Marks	No. of Students
Below 15	4
15—30	6
30—45	12
45—60	8
Above 60	7

In this example, because all defined class-intervals are same, the assumption would be that the first and last class shall have same class-interval of 15 and hence the lower limit of the first class shall be zero and upper limit of last class shall be 75. Hence first class would be 0—15 and the last class 60—75.

What happens in this case?

Marks	No. of Students
Below 10	4
10—30	7
30—60	10
60—100	8
Above 100	4

In this problem because the class interval is 20 in the second class, 30 in the third, 40 in the fourth class and so on. The class interval is increasing by 10. Therefore the appropriate assumption in this case would be that the lower limit of the first class is zero and the upper limit of the last class is 150. In case of other open-end class distributions the first class limit should be fixed on the basis of succeeding class interval and the last class limit should be fixed on the basis of preceding class interval.

If the class intervals are of varying width, an effort should be made to avoid calculating mean and mode. It is advisable to calculate median.

Weighted Mean

In the computation of arithmetic mean, we give equal importance to each item in the series.

Raja Toy Shop sell : Toy Cars at Rs. 3 each; Toy Locomotives at Rs. 5 each; Toy Aeroplane at Rs. 7 each; and Toy Double Decker at Rs. 9 each.

What shall be the average price of the toys sold ? If the shop sells 4 toys one of each kind.

$$\bar{X} \text{ (Mean Price)} = \frac{\Sigma X}{N} = \frac{24}{4} = \text{Rs. } 6.$$

In this case the importance of each toy is equal as one toy of each variety has been sold. While computing the arithmetic mean this fact has been taken care of including the price of each toy once only.

But if the shop sells 100 toys, 50 cars, 25 locomotives, 15 aeroplanes and 10 double deckers, the importance of the four toys to the dealer is not equal as a source of earning revenue. In fact their respective importance is equal to the number of units of each toy sold, i.e. the importance of Toy car is 50; the importance of Locomotive is 25; the importance of Aeroplane is 15; and the importance of Double Decker is 10.

It may be noted that 50, 25, 15,10 are the quantities of the various classes of toys sold. These quantities are called as 'weights' in statistical language. Weight is represented by symbol W and ΣW represents the sum of weights.

While determining the average price of toy sold these weights are of great importance and are taken into account to compute weighted mean.

$$\bar{X}_w = \frac{\Sigma[(W_1X_1)+(W_2X_2)+(W_3X_3)+(W_4X_4)]}{W_1+W_2+W_3+W_4} = \frac{\Sigma WX}{\Sigma W}$$

where, W_1, W_2, W_3, W_4 are weights and X_1, X_2, X_3, X_4 represents the price of 4 varieties of toy.

Hence by substituting the values of W_1, W_2, W_3, W_4 and X_1, X_2, X_3, X_4, we get

$$\bar{X}_w = \frac{(50\times3)+(25\times5)+(15\times7)+(10\times9)}{50+25+15+10}$$

$$\bar{X}_w = \frac{150+125+105+90}{100} = \frac{470}{100} = \text{Rs. } 4.70$$

The table given below demonstrates the procedure of computing the weighted Mean.

Weighted Arithmetic mean of Toys by the Raja Shop.

Toy	Price per toy (Rs.) X	Number Sold W	Price × Weight WX
Car	3	50	150
Locomotive	5	25	125
Aeroplane	7	15	105
Double Decker	9	10	90
		$\Sigma W = 100$	$\Sigma WX = 470$

$$\therefore \quad \bar{X}_w = \frac{\Sigma WX}{\Sigma W} = \frac{470}{100} = \text{Rs. } 4.70$$

Example: The table below shows the number of skilled and unskilled workers in two localities along with their average hourly wages.

	Ram Nagar		Shyam Nagar	
Worker Category	**Number**	**Wages (per hour)**	**Number**	**Wages (per hour)**
Skilled	150	1.80	350	1.75
Unskilled	850	1.30	650	1.25

Determine the average hourly wage in each locality. Also give reasons why the results show that the average hourly wage in Shyam Nagar exceed the average hourly wage in Ram Nagar, even though in Shyam Nagar the average hourly wages of both categories of workers is lower. It is required to compute weighted mean.

Solution :

	Ram Nagar			Shyam Nagar		
	X	W	WX	X	W	WX
Skilled	1.80	150	270	1.75	350	612.50
Unskilled	1.30	850	1105	1.25	650	812.50
Total		1000	1375		1000	1425

$$\overline{X}_w = \frac{1375}{1000} = \text{Rs.}1.375 \qquad \overline{X}_w = \frac{1425}{1000} = \text{Rs.}1.425$$

It may be noted that weights are more evenly assigned to the different categories of workers in Shyam Nagar than in Ram Nagar.

Geometric Mean :

In general, if we have n numbers (none of them being zero), then the G.M. is defined as

$$\text{G.M.} = \sqrt{x_1, x_2, \ldots\ldots\ldots\ldots x_n} = (x_1, x_2, \ldots\ldots\ldots\ldots x_n)^{1/n}$$

In case of a discrete series, if $x_1, x_2, \ldots\ldots\ldots\ldots x_n$ occur $f_1, f_2, \ldots\ldots\ldots\ldots f_n$ times respectively and N is the total frequency (i.e. $N = f_1 +, f_2 +, \ldots\ldots\ldots\ldots f_n$), then

$$\text{G.M.} = \sqrt[n]{x_1 f_1, x_2 f_2, \ldots\ldots\ldots\ldots x_n f_n}$$

For convenience, use of logarithms is made extensively to calculate the nth root. In terms of logarithms

$$\text{G.M.} = \text{AL}\left(\frac{\log x_1 + \log x_2 + \ldots\ldots\ldots\ldots \log x_n}{n}\right)$$

$$= \text{AL}\left(\frac{\Sigma \log x}{N}\right), \text{ where AL refers to antilog.}$$

In discrete series, $\text{G.M.} = \text{AL}\dfrac{\Sigma f \log x}{N}$

and in case of continuous series, $\text{G.M.} = \text{AL}\dfrac{\Sigma f \log m}{N}$

Example: Calculate G.M. of the following data :

2, 4, 8

Solution : $\text{G.M.} = \sqrt[3]{2 \times 4 \times 8} = \sqrt[3]{64} = 4$

In terms of logarithms, the question can be solved as follows :

$\log 2 = 0.3010$, $\log 4 = 0.6021$, and $\log 8 = 9.9031$

Apply the formula :

$$\text{G.M.} = \text{AL}\frac{\Sigma \log x}{N} = \text{AL}\frac{1.8062}{3} = \text{AL}\,(0.60206) = 4$$

Example: Calculate geometric mean of the following data :

x	5	6	7	8	9	10	11
f	2	4	7	10	9	6	2

Solution: **Calculation of G.M.**

x	$\log x$	f	$f \log x$
5	0.6990	2	1.3980
6	0.7782	4	3.1128
7	0.8451	7	5.9157
8	0.9031	10	9.0310
9	0.9542	9	8.5878
10	1.0000	6	6.0000
11	1.0414	2	2.0828
		$N = 40$	$\sum f \log x = 36.1281$

$$\text{G.M.} = \text{AL}\left(\frac{\sum f \log x}{N}\right) = \text{AL}\left(\frac{36.1281}{40}\right) = \text{AL}\,(0.9032) = 8.002$$

Example: Calculate G.M. from the following data :

X	f
9.5—14.5	10
14.5—19.5	15
19.5—24.5	17
24.5—29.5	25
29.5—34.5	18
34.5—39.5	12
39.5—44.5	8

Solution: **Calculation of G.M.**

X	m	$\log m$	f	$f \log m$
9.5—14.5	12	1.0792	10	10.7920
14.5—19.5	17	1.2304	15	18.4560
19.5—24.5	22	1.3424	17	22.8208
24.5—29.5	27	1.4314	25	35.7850
29.5—34.5	32	1.5051	18	27.0918
34.5—39.5	37	1.5682	12	18.8184
39.5—44.5	42	1.6232	8	12.9850
			$N = 105$	$\sum f \log m = 146.7410$

$$G.M. = AL\left(\frac{146.7490}{105}\right) = AL\,(1.3976) = 24.98$$

Specific uses of G.M. : The geometric Mean has certain specific uses, some of them are :

(i) It is used in the construction of index numbers,

(ii) It is also helpful in finding out the compound rates of change such as the rate of growth of population in a country.

(iii) It is suitable where the data are expressed in terms of rates, ratios and percentage.

(iv) It is quite useful in computing the average rates of depreciation or appreciation.

(v) It is most suitable when large weights are to be assigned to small items and small weights to large items.

Example: The gross national product of a country was Rs. 1,000 crores 10 years earlier. It is Rs. 2,000 crores now. Calculate the rate of growth in G.N.P.

Solution: In this case compound interest formula will be used for computing the average annual per cent increase of growth.

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

where P_n = principal sum (or any other variate) at the end of the period.

P_o = principal sum in the beginning of the period.

r = rate of increase or decrease.

n = number of years.

It may be noted that the above formula can also be written in the following form :

$$r = \sqrt[n]{\frac{P_n}{P_o}} - 1$$

Substituting the values given in the formula, we have

$$r = \sqrt[10]{\frac{2000}{1000}} - 1 = \sqrt[10]{2} - 1$$

$$= AL\left[\frac{\log 2}{10}\right] - 1 = AL\left[\frac{0.30103}{10}\right] - 1 = 1.0718 - 1 = 0.0718 = 7.18\%$$

Hence, the rate of growth in GNP is 7.18%.

Example: The price of commodity increased by 5 per cent from 2001 to 2002, 8 percent from 2002 to 2003 and 77 per cent from 2003 to 2004. The average increase from 2001 to 2004 is quoted at 26 per cent and not 30 per cent. Explain this statement and verify the arithmetic.

Solution: Taking P_n as the price at the end of the period, P_o as the price in the beginning, we can substitute the values of P_n and P_o in the compound interest formula. Taking $P_o = 100$; $P_n = 200.72$

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

$$200.72 = 100(1+r)^3$$

or $$(1+r)^3 = \frac{200.72}{100} \quad \text{or} \quad 1+r = \sqrt[3]{\frac{200.72}{100}}$$

$$r = \sqrt[3]{\frac{200.72}{100}} - 1 = 1.260 - 1 = 0.260 = 26\%$$

Thus increase is not average of (5 + 8 + 77)/3 = 30 per cent. It is 26% as found out by G.M.

Weighted G.M. : The weighted G.M. is calculated with the help of the following formula :

$$\text{G.M.} = \sqrt{x_1 w_1, x_2 w_2 \ldots\ldots\ldots\ldots x_n w_n}$$

$$= \frac{\log x_1 \times w_1 + \log x_2 \times w_2 + \ldots\ldots\ldots\ldots \log x_n \times w_n}{w_1 + w_2 + \ldots\ldots\ldots w_n}$$

$$= \text{AL}\left[\frac{\Sigma (\log x \times w)}{\Sigma w}\right]$$

Example: Find out weighted G.M. from the following data :

Group	Index Number	Weights
Food	352	48
Fuel	220	10
Cloth	230	8
House Rent	160	12
Misc.	190	15

Solution : **Calculation of Weighted G.M.**

Group	Index Number (x)	Weights (w)	Log x	w log x
Food	352	48	2.5465	122.2320
Fuel	220	10	2.3424	23.4240
Cloth	230	8	2.3617	17.8936
House Rent	160	12	2.2041	26.4492
Misc.	190	15	2.2788	34.1820
		93		225.1808

$$\text{G. M. (weighted)} = \text{AL}\left[\frac{\Sigma w \log x}{\Sigma w}\right] = \text{AL}\,\frac{225.1808}{93} = 263.8$$

Example: A machine depreciates at the rate of 35.5% per annum in the first year, at the rate of 22.5% per annum in the second year, and at the rate of 9.5% per annum in the third year, each percentage being computed on the actual value. What is the average rate of depreciation?

Solution: Average rate of depreciation can be calculated by taking G.M.

Year	X (values taking 100 as base)	log X
I	100 – 35.5 = 64.5	1.8096
II	100 – 22.5 = 77.5	1.8893
III	100 – 9.5 = 90.5	1.9566
		Σ log X = 5.6555

Apply $\text{G.M.} = \text{AL}\left[\frac{\Sigma \log x}{w}\right] = \frac{5.6555}{3} = \text{AL}\,1.8851 = 76.77$

$\therefore$ Average rate of depreciation $= 100 - 76.77 = 23.33\%$.

Example : The arithmetic mean and geometric mean of two values are 10 and 8 respectively. Find the values.

Solution : If two values are taken as a and b, then

$$\frac{a+b}{2} = 10, \quad \text{and} \quad \sqrt{ab} = 8$$

Or $a + b = 20, \quad ab = 64$

then $a - b = \sqrt{(a+b)^2 - 4ab} = \sqrt{(20)^2 - 4 \times 64} = \sqrt{400 - 256} = \sqrt{144} = 12$

Now, we have $a + b = 20$, ...(*i*)

$a - b = 12$ .....(*ii*)

Solving for a and b, we get $a = 4$ and $b = 16$.

Harmonic Mean : The harmonic mean is defined as the reciprocals of the average of reciprocals of all items in a series. Symbolically,

$$\text{H.M.} = \frac{N}{\left(\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \ldots\ldots\ldots\ldots + \frac{1}{x_n}\right)} = \frac{N}{\sum\left(\frac{1}{x}\right)}$$

In case of a discrete series,

$$\text{H.M.} = \frac{N}{\sum\left\{f \times \frac{1}{x}\right\}}$$

and in case of a continuous series,

$$\text{H.M.} = \frac{N}{\sum\left\{f \times \frac{1}{m}\right\}}$$

It may be noted that none of the values of the variable should be zero.

Example: Calculate harmonic mean from the following data: 5, 15, 25, 35 and 45.

Solution :

X	$\frac{1}{X}$
5	0.20
15	0.067
25	0.040
35	0.029
45	0.022
N = 5	$\sum\left(\frac{1}{X}\right) = 0.358$

$$\text{H.M.} = \frac{N}{\sum\left(\frac{1}{x}\right)} = \frac{5}{0.358} = 14 \text{ approx.}$$

Example : From the following data compute the value of the harmonic mean :

x :	5	15	25	35	45
f :	5	15	10	15	5

Solution :

Calculation of Harmonic Mean

x	f	$\frac{1}{x}$	$f \times \frac{1}{x}$
5	5	0.200	1.000
15	15	0.067	1.005
25	10	0.040	0.400
35	15	0.29	0.435
45	5	0.022	0.110
	$\Sigma f = 50$		$\Sigma\left(f \times \frac{1}{x}\right) = 2.950$

$$\text{H.M.} = \frac{N}{\Sigma\left\{f \times \frac{1}{x}\right\}} = \frac{50}{2.95} = 17 \text{ approx.}$$

Example : Calculate harmonic mean from the following distribution :

x	f
0—10	5
10—20	15
20—30	10
30—40	15
40—50	5

Solution : First of all, we shall find out mid points of the various classes. They are 5, 15, 25, 35 and 45. Then we will calculate the H.M. by applying the following formula :

$$\text{H.M.} = \frac{N}{\Sigma\left\{f \times \frac{1}{m}\right\}}$$

Calculation of Harmonic Mean

x (Mid Moints)	f	$\frac{1}{x}$	$f \times \frac{1}{x}$
5	5	0.200	1.000
15	15	0.067	1.005
25	10	0.040	0.400
35	15	0.29	0.435
45	5	0.022	0.110
	$\Sigma f = 50$		$\Sigma\left(f \times \frac{1}{x}\right) = 2.950$

The answer will be 17 (approx).

Application of Harmonic Mean to special cases: Like Geometric means, the harmonic mean is also applicable to certain special types of problems. Some of them are:

(i) If, in averaging time rates, distance is constant, then H.M. is to be calculated.

Example: A man travels 480 km. a day. On the first day he travels for 12 hours @ 40 km. per hour and second day for 10 hours @ 48 km. per hour. On the third day he travels for 1.5 hours @ 32 km. per hour. Find his average speed.

Solution: We shall use the harmonic mean,

$$\text{H.M.} = \frac{N}{\sum\left(\frac{1}{X}\right)} = \frac{3}{\frac{1}{40}+\frac{1}{48}+\frac{1}{32}} = \frac{3}{37/480} = 39 \text{ km. per hour (approx.).}$$

The arithmetic mean would be $\frac{48+40+32}{3} = 40$ km. per hour.

(ii) If, in averaging the price data, the prices are expressed as "quantity per rupee". Then harmonic mean should be applied.

Example: A man purchased one kilo of cabbage from each of four places at the rate of 20 kg., 16 kg., 12 kg., and 10 kg. per rupees respectively. On the average how many kilos of cabbages he has purchased per rupee.

Solution:
$$\text{H.M.} = \frac{N}{\sum\left(\frac{1}{x}\right)} = \frac{4}{\frac{1}{20}+\frac{1}{16}+\frac{1}{12}+\frac{1}{10}} = \frac{4}{71/250} = \frac{4\times 240}{71} = 13.5 \text{ kg. per rupee.}$$

POSITIONAL AVERAGES

Median

The median is that value of the variable which divides the group in two equal parts. One part comprising the values greater than and the other all values less than median. Median of a distribution may be defined as that value of the variable which exceeds and is exceeded by the same number of observation. It is the value such that the number of observations above it is equal to the number of observations below it. Thus we know that the arithmetic mean is based on all items of the distribution, the median is positional average, that is, it depends upon the position occupied by a value in the frequency distribution.

When the items of a series are arranged in ascending or descending order of magnitude the value of the middle item in the series is known as median in the case of individual observation. Symbolically,

$$\text{Median} = \text{size of } \left(\frac{N+1}{2}\right) \text{th item}$$

If the number of items is even, then there is no value exactly in the middle of the series. In such a situation the median is arbitrarily taken to be halfway between the two midddle items. Symbolically,

$$\text{Median} = \frac{\text{size of } \frac{N}{2} \text{ th item} + \text{size of } \left(\frac{N+1}{2}\right) \text{th item}}{2}$$

Example: Find the median of the following series:

(i) 8, 4, 8, 3, 4, 8, 6, 5, 10.

(ii) 15, 12, 5, 7, 9, 5, 11, 28.

Solution :

Computation of Median

(i)		*(ii)*	
Serial No.	***X***	**Serial No.**	***X***
1	3	1	5
2	4	2	5
3	4	3	7
4	5	4	9
5	6	5	11
6	8	6	12
7	8	7	15
8	8	8	28
9	10		
N = 9		N = 8	

For (i) series, $\text{Median} = \text{size of } \left(\frac{N+1}{2}\right) \text{th item} = \text{size of the } \frac{9+1}{2} \text{ th item} = \text{size of 5th item} = 6$

For (ii) series, $\text{Median} = \text{size of } \left(\frac{N+1}{2}\right) \text{th item} = \text{size of the } \frac{8+1}{2} \text{ th item}$

$$= \frac{\text{size of 4th item} + \text{size of 5th item}}{2} = \frac{9+11}{2} = 10$$

Location of Median in Discrete series: In a discrete series, medium is computed in the following manner:

(i) Arrange the given variable data in ascending or descending order.

(ii) Find cumulative frequencies.

(iii) Apply Med. = size of $\left(\frac{N+1}{2}\right)$ th item

(iv) Locate median according to the size i.e., variable corresponding to the size or for next cumulative frequency.

Example: Following are the number of rooms in the houses of a particular locality. Find median of the data:

No. of rooms:	3	4	5	6	7	8
No. of houses:	38	654	311	42	12	2

Solution: **Computation of Median**

No. of Rooms X	No. of Houses f	Cumulative Frequency Cf
3	38	38
4	654	692
5	311	1003
6	42	1045
7	12	1057
8	2	1059

$$\text{Median} = \text{size of } \left(\frac{N+1}{2}\right) \text{ th item} = \text{size of } \frac{1059+1}{2} \text{ th item} = 530\text{th item.}$$

Median lies in the cumulative frequency of 692 and the value corresponding to this is 4

Therefore, Median = 4 rooms.

In a continuous series, median is computed in the following manner:

(i) Arrange the given variable data in ascending or descending order.

(ii) If inclusive series is given, it must be converted into exclusive series to find real class intervals.

(iii) Find cumulative frequencies.

(iv) Apply Median = size of $\frac{N}{2}$ th item to ascertain median class.

(v) Apply formula of interpolation to ascertain the value of median.

$$\text{Median} = l_1 + \frac{\frac{N}{2} - cf_0}{f} \times (l_1 - l_2) \quad \text{or} \quad \text{Median} = l_2 - \frac{\frac{N}{2} - cf_0}{f} \times (l_2 - l_1)$$

where, l_1 refers to lower limit of median class,

l_2 refers to higher limit of median class,

cf_0 refers cumulative frequency of previous to median class,

f refers to frequency of median class,

Example: The following table gives you the distribution of marks secured by some students in an examination:

Marks	No. of Students
0—20	42
21—30	48
31—40	120
41—50	84
51—60	48
61—70	36
71—80	31

Find the median marks.

Solution: **Calculation of Median Marks**

Marks (x)	No. of Students (f)	cf
0—20	42	42
21—30	38	80
31—40	120	200
41—50	84	284
51—60	48	332
61—70	36	368
71—80	31	399

$$\text{Median} = \text{size of } \frac{N}{2}\text{th item} = \text{size of } \frac{399}{2}\text{th item} = 199.5\text{th item.}$$

which lies in (31—40) group, therefore the median class is 30.5—40.5.

Applying the formula of interpolation.

$$\text{Median} = l_1 + \frac{\frac{N}{2} - cf_0}{f} \times (l_1 - l_2)$$

$$= 30.5 + \frac{199.5 - 80}{120} \times (10) = 30.5 + \frac{119.5}{12} = 40.46 \text{ marks.}$$

Related Positional Measures: The median divides the series into two equal parts. Similarly there are certain other measures which divide the series into certain equal parts. There are first quartile, third quartile, deciles, percentiles etc. If the items are arranged in ascending or descending order of magnitude, Q_1 is that value which covers 1/4th of the total number of items. Similarly, if the total number of items are divided into ten equal parts, then, there shall be nine deciles.

Symbolically,

$$\text{First quartile } (Q_1) = \text{size of } \left(\frac{N+1}{4}\right)\text{th item}$$

$$\text{Third quartile } (Q_3) = \text{size of } \frac{3(N+1)}{4} \text{ th item}$$

$$\text{First decile } (D_1) = \text{size of } \left(\frac{N+1}{10}\right) \text{ th item}$$

$$\text{Sixth decile } (D_6) = \text{size of } \frac{6(N+1)}{10} \text{ th item}$$

$$\text{First percentile } (P_1) = \text{size of } \left(\frac{N+1}{100}\right) \text{ th item}$$

Once values of the items are found out, then formulae of interpolation are applied for ascertaining the value of Q_1, Q_3, D_1, D_4, P_{40} etc.

Example: Calculate Q_1, Q_3, D_2 and P_5 from following data:

Marks:	Below 10	10–20	20–40	40–60	60–80	above 80
No. of Students:	8	10	22	25	10	5

Solution: **Calculation of Positional Values**

Marks	No. of Students (*f*)	*C.f.*
Below 10	8	8
10—20	10	18
20—40	22	40
40—60	25	65
60—80	10	75
Above 80	5	80
	N = 80	

$$Q_1 = \text{size of } \frac{N}{4} \text{ th item} = \frac{80}{4} = 20\text{th item}$$

Hence Q_1 lies in the class 20—40, apply

$$Q_1 = l_1 + \frac{\frac{N}{4} - Cf_0}{f} \times i \text{ where } l_1 = 20,\ \frac{N}{4} = 20,\ Cf_0 = 18,\ f = 22 \text{ and } i = (l_2 - l_1) = 20$$

By substituting the values, we get

$$Q_1 = 20 + \frac{(20-18)}{22} \times 20 = 20 + 1.8 = 21.8$$

Similarly, we can calculate

$$Q_3 = \text{size of } \frac{3N}{4} \text{ th item} = \frac{3 \times 80}{4} \text{ th item} = 60\text{th item.}$$

Hence Q_3 lies in the class 40—60, apply

$$Q_3 = l_1 + \frac{\frac{3N}{4} - Cf_0}{f} \times i \quad \text{where } l_1 = 40,\ \frac{3N}{4} = 60,\ Cf_0 = 40,\ f = 25,\ i = 20$$

$$\therefore \quad Q_3 = 40 + \frac{(60-40)}{25} \times 20 = 40 + 16 = 56$$

D_2 = size of $\frac{2N}{10}$ th item = 16th item. Hence D_2 lies in the class 10—20.

$$D_2 = l_1 + \frac{\frac{2N}{10} - Cf_0}{f} \times i \quad \text{where } l_1 = 10,\ \frac{2N}{4} = 16,\ Cf_0 = 8,\ f = 10,\ i = 10.$$

$$D_2 = 10 + \frac{(16-8)}{10} \times 10 = 10 + 8 = 18$$

P_5 = size of $\frac{5N}{100}$ th item = $\frac{5 \times 80}{100}$ th item = 4th item. Hence P_5 lies in the class 0 — 10.

$$P_5 = l_1 + \frac{\frac{5N}{100} - Cf_0}{f} \times i \quad \text{where } l_1 = 0,\ \frac{5N}{100} = 4,\ Cf_0 = 0,\ f = 8,\ i = 10$$

$$P_5 = 0 + \frac{4-0}{8} \times 10 = 0 + 5 = 5.$$

Calculation of Missing Frequencies:

Example: In the frequency distribution of 100 families given below; the number of families corresponding to expenditure groups 20—40 and 60—80 are missing from the table. However the median is known to be 50. Find out the missing frequencies.

Expenditure:	0—20	20—40	40—60	60—80	80—100
No. of families:	14	?	27	?	15

Solution: We shall assume the missing frequencies for the classes 20—40 to be x and 60—80 to y

Expenditure (Rs.)	No. of Families	*C.f.*
0—20	14	14
20—40	x	$14 + x$
40—60	27	$14 + 27 + x$
60—80	y	$41 + x + y$
80—100	15	$41 + 15 + x + y$
	$N = 100 = 56 + x + y$	

From the table, we have $N = \Sigma F = 56 + x + y = 100$.

$\therefore \quad x + y = 100 - 56 + 44,$

Median is given as 50 which lies in the class 40—60, which becomes the median class.

By using the median formula we get:

$$\text{Median} = l_1 + \frac{\frac{N}{2} - Cf_0}{f} \times i$$

$$\therefore \quad 50 = 40 + \frac{50 - (14 + x)}{27} \times (60 - 40) \quad \text{or} \quad 50 = 40 + \frac{50 - (14 + x)}{27} \times 20$$

$$\text{or} \quad 50 - 40 = \frac{36 - x}{27} \times 20 \quad \text{or} \quad 50 - 40 = (36 - x) \times \frac{20}{27}$$

$$\text{or} \quad 10 \times 27 = 720 - 20x \quad \text{or} \quad 270 = 720 - 20x$$

$$\therefore \quad 20x = 720 - 270$$

$$x = \frac{450}{20} = 22.5.$$

By substitution the value of x in the equation,

$$x + y = 44$$

We get, $\quad 22.5 + y = 44$

$$\therefore \quad y = 44 - 22.5 = 21.5.$$

Hence frequency for the class 20—40 is 22.5 and 60—80 is 21.5.

Mode

Mode is that value of the variable which occurs or repeats itself maximum number of times. The mode is the most "fashionable" size in the sense that it is the most common and typical and is defined by Zizek as "the value occurring most frequently in series of items and around which the other items are distributed most densely." In the words of Croxton and Cowden, the mode of a distribution is the value at the point where the items tend to be most heavily concentrated. According to A.M. Tuttle, Mode is the value which has the greater frequency density in its immediate neighbourhood. In the case of individual observations, the mode is that value which is repeated the maximum number of times in the series. The value of mode can be denoted by the alphabet z also.

Example: Calculate mode from the following data:

Sr. Number	:	1	2	3	4	5	6	7	8	9	10
Marks obtained	:	10	27	24	12	27	27	20	18	15	30

Solution:

Marks	No. of Students
10	1
12	1
15	1
18	1

20	1	Mode is 27 marks
24	1	
27	3	
30	1	

Calculation of Mode in Discrete series. In discrete series, it is quite often determined by inspection. We can understand with the help of an example:

X	1	2	3	4	5	6	7
f	4	5	13	6	12	8	6

By inspection, the modal size is 3 as it has the maximum frequency. But this test of greatest frequency is not fool proof as it is not the frequency of a single class, but also the frequencies of the neighbour classes that decide the mode. In such cases, we shall be using the method of Grouping and Analysis table.

Size of shoe	1	2	3	4	5	6	7
Frequency	4	5	13	6	12	8	6

Solution: By inspection, the mode is 3, but the size of mode may be 5. This is so because the neighbouring frequencies of size 5 are greater than the neighbouring frequencies of size 3. This effect of neighbouring frequencies is seen with the help of grouping and analysis table technique.

Grouping table

Size of Shoe	**Frequency**					
	1	2	3	4	5	6
1	4	9	18	22	24	31
2	5					
3	13	19				
4	6	18		26	26	
5	12					
6	8	20	14			
7	6					

When there exist two groups of frequencies in equal magnitude, then we should consider either both or omit both while analysing the sizes of items.

Analysis Table

Column	**Size of Items with Maximum Frequency**
1	3
2	5, 6
3	1, 2, 3, 4, 5
4	4, 5, 6

5	5, 6, 7
6	3, 4, 5

Item 5 occurs maximum number of times, therefore, mode is 5. We can note that by inspection we had determined 3 to be the mode.

Determination of mode in continuous series: In the continuous series, the determination of mode requires one additional step. Once the modal class is determined by inspection or with the help of grouping technique, then the following formula of interpolation is applied:

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2}(l_2 - l_1) \quad \text{or} \quad \text{Mode} = l_2 - \frac{f_1 - f_0}{2f_1 - f_0 - f_2}(l_2 - l_1)$$

l_1 = lower limit of the class, where mode lies.

l_2 = upper limit of the class, where mode lies.

f_0 = frequency of the class proceeding the modal class.

f_1 = frequency of the class, where mode lies.

f_2 = frequency of the class succeeding the modal class.

Example: Calculate mode of the following frequency distribution:

Variable	Frequency
0—10	5
10—20	10
20—30	15
30—40	14
40—50	10
50—60	5
60—70	3

Solution:

Grouping Table

X	1	2	3	4	5	6
0—10	5					
		15				
10—20	10			30		
			25			
20—30	15				39	
		29				39
30—40	14					
			24			

40—50	10		29
		15	
50—60	5		18
		8	
60—70	3		

Analysis Table

Column	Size of Item with Maximum Frequency
1	20—30
2	20—30, 30—40
3	10—20, 20—30
4	0—10, 10—20, 20—30
5	10—20, 20—30, 30—40
6	20—30, 30—40, 40—50

Modal group is 20—30 because it has occurred 6 times. Applying the formula of interpolation.

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2}(l_2 - l_1)$$

$$= 20 + \frac{15-10}{30-10-14}(30-20) = 20 + \frac{5}{6}(10) = 28.3$$

Calculation of mode where it is ill defined. The above formula is not applied where there are many modal values in a series or distribution. For instance there may be two or more than two items having the maximum frequency. In these cases, the series will be known as bimodal or multimodal series. The mode is said to be ill-defined and in such cases the following formula is applied.

Mode = 3 Median – 2 Mean.

Example: Calculate mode of the following frequency data:

Variate Value	Frequency
10—20	5
20—30	9
30—40	13
40—50	21
50—60	20
60—70	15
70—80	8
80—90	3

Solution : First of all, ascertain the modal group with the help of process of grouping.

Grouping Table

X	1	2	3	4	5	6
10—20	5					
		14				
20—30	9			27		
			22			
30—40	13				43	
		34				
40—50	21					54
			41			
50—60	20			56		
		35				
60—70	15				43	
			23			
70—80	8					26
		11				
80—90	3					

Analysis Table

Column	Size of Item with Maximum Frequency
1	40—50
2	50—60, 60—70
3	40—50, 50—60
4	40—50, 50—60, 60—70
5	20—30, 30—40, 40—50, 50—60, 60—70, 70—80
6	30—40, 40—50, 50—60

There are two groups which occur equal number of items. They are 40—50 and 50—60. Therefore, we will apply the following formula:

Mode = 3 median – 2 mean and for this purpose the values of mean and median are required to be computed.

Calculation of Mean and Median

Variate	Frequency	Mid Values	$\left(\frac{m-45}{10}\right)$		
X	f	m	$d'x$	$fd'x$	Cf
10—20	5	15	– 3	– 15	5
20—30	9	25	– 2	– 18	14
30—40	13	35	– 1	– 13	27

40—50	21	45	0	0	48	Median is the
50—60	20	55	+1	+20	68	value of $\frac{N}{2}$th
60—70	15	65	+2	+30	83	item which lies
70—80	8	75	+3	+24	91	in (40—50) group
80—90	3	85	+4	+12	94	
	N = 94			$\Sigma fd' = +40$		

$$\overline{X} = A + \frac{\Sigma\, fd'x}{N} \times i \qquad \text{Med.} = l_1 + \frac{\frac{N}{2} - cf_0}{f} \times i$$

$$= 45 + \frac{40}{94}(10) = 45 + 4.2 = 49.2 \qquad = 40 + \frac{47-27}{21}(10) = 40 + \frac{200}{21} = 49.5$$

Mode = 3 median – 2 mean

= 3 (49.5) – 2 (49.2) = 148.5 – 98.4 = 50.1

Determination of mode by curve fitting: Mode can also be computed by curve fitting. The following steps are to be taken;

(i) Draw a histogram of the data.

(ii) Draw the lines diagonally inside the modal class rectangle, starting from each upper corner of the rectangle to the upper corner of the adjacent rectangle.

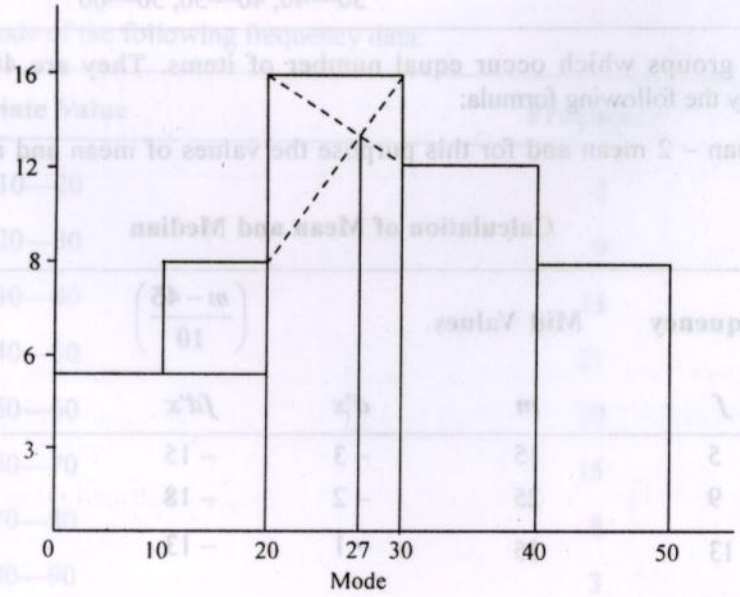
(iii) Draw a perpendicular line from the intersection of the two diagonal lines to the X-axis. The abscissa of the point at which the perpendicular line meets is the value of the mode.

Example: Construct a histogram for the following distribution and, determine the mode graphically:

X	:	0—10	10—20	20—30	30—40	40—50
f	:	5	8	15	12	7

Verify the result with the help of interpolation.

Solution:



$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2}(l_2 - l_1)$$

$$= 20 + \frac{15-8}{30-8-12}(30-20) = 20 + \frac{7}{10}(10) = 27.$$

Example:

Calculate mode from the following data:

Marks	No. of Students
Below 10	4
" 20	6
" 30	24
" 40	46
" 50	67
" 60	86
" 70	96
" 80	99
" 90	100

Solution:

Since we are given the cumulative frequency distribution of marks, first we shall convert it into the normal frequency distribution:

Marks	Frequencies
0—10	4
10—20	6 – 4 = 2
20—30	24 – 6 = 18
30—40	46 – 24 = 22
40—50	67 – 46 = 21
50—60	86 – 67 = 19
60—70	96 – 86 = 10
70—80	99 – 96 = 3
80—90	100 – 99 = 1

It is evident from the table that the distribution is irregular and maximum chances are that the distribution would be having more than one mode. You can verify by applying the grouping and analysing table.

The formula to calculate the value of mode in cases of bio-modal distributions is :

Mode = 3 median – 2 mean.

Computation of Mean and Median:

Marks	Mid-value	Frequency		$\left(\frac{X-45}{10}\right)$	
	(x)	(f)	Cf	(dx)	fdx
0—10	5	4	4	– 4	– 16
10—20	15	2	6	– 3	– 6
20—30	25	18	24	– 2	– 36
30—40	35	22	46	– 1	– 22
40—50	45	21	67	0	0
50—60	55	19	86	1	19
60—70	65	10	96	2	20
70—80	75	3	99	3	9
80—90	85	1	100	4	4
		$\Sigma f = 100$			$\Sigma fdx = -28$

$$\text{Mean} = A + \frac{\Sigma\, fdx}{N} \times i \quad 45 + \frac{-28}{100} \times 10 = 42.2$$

$$\text{Median} = \text{size of } \frac{N}{2} \text{ th item} = \frac{100}{2} = 50\text{th item.}$$

Because 50 is smaller to 67in *C.f.* column. Median class is 40–50

$$\text{Median} = l_1 + \frac{\frac{N}{2} - Cf_0}{f} \times i$$

$$\text{Median} = 40 + \frac{50-46}{21} \times 10 = 40 + \frac{4}{21} \times 10 = 41.9$$

Apply, Mode = 3 median – 2 mean

Mode = 3 × 41.9 – 2× 42.2 = 125.7 – 84.3 = 41.3

Example: Median and mode of the wage distribution are known to be Rs. 33.5 and 34 respectively. Find the missing values.

Wages (Rs.)	No. of Workers
0—10	4
10—20	16
20—30	?
30—40	?
40—50	?
50—60	6
60—70	4
	Total = 230

Solution: We assume the missing frequencies as 20—30 as x, 30—40 as y, and 40—50 as $230 - (4 + 16 + x + y + 6 + 4) = 200 - x - y$.

We now proceed further to compute missing frequencies:

Wages (Rs.) X	No. of workers f	Cumulative frequencies $C.f.$
0—10	4	4
10—20	16	20
20—30	x	$20 + x$
30—40	y	$20 + x + y$
40—50	$200 - x - y$	220
50—60	6	226
60—70	4	230
	N = 230	

Apply, Median $= l_1 + \dfrac{\frac{N}{2} - cf_0}{f} \times (l_2 - l_1)$

$$33.5 = 30 + \frac{115 - (20 + x)}{y} \times (40 - 30)$$

$$y(33.5 - 30) = (115 - 20 - x)10$$

$$3.5y = 1150 - 200 - 10x$$

$$10x + 3.5y = 950 \quad \text{.....(i)}$$

Apply, Mode $= l_1 + \dfrac{f_1 - f_0}{2f_1 - f_0 - f_2} \times (l_2 - l_1)$

$$= 20 + \frac{15 - 8}{30 - 8 - 12}(30 - 20)$$

$$4(3y - 200) = 10(y - x)$$

$$10x + 2y = 800 \quad \text{.....(ii)}$$

Subtract equation (ii) from equation (i),

$$1.5y = 150, \qquad y = \frac{150}{1.5} = 100$$

Substitute the value of $y = 100$ in equation (i), we get

$$10x + 3.5(100) = 950$$

$$10x = 950 - 350$$

$$x = 600/10 = 60.$$

$\therefore$ Third missing frequency $= 200 - x - y = 200 - 60 - 100 = 40.$

LESSON 3

MEASURES OF DISPERSION

Why dispersion?

Measures of central tendency, Mean, Median, Mode, etc., indicate the central position of a series. They indicate the general magnitude of the data but fail to reveal all the peculiarities and characteristics of the series. In the other words, they fail to reveal the degree of the spread out or the extent of the variability in individual items of the distribution. This can be explained by certain other measures, known as 'Measures of Dispersion' or Variation.

We can understand variation with the help of the following example :

Series I	*Series II*	*Series III*
10	2	10
10	8	12
10	20	8
$\Sigma X = 30$	30	30
$\bar{X} = \frac{\Sigma X}{N} = \frac{30}{3} = 10$	$\bar{X} = \frac{30}{3} = 10$	$\bar{X} = \frac{30}{3} = 10$

In all three series, the value of arithmetic mean is 10. On the basis of this average, we can say that the series are alike. If we carefully examine the composition of three series. we find the following differences:

(i) In case of Ist series. the value are equal; but in 2nd and 3rd series, the values are unequal and do not follow any specific order.

(ii) The magnitude of deviation, item-wise, is different for the 1st, 2nd and 3rd series. But all these deviations cannot be ascertained if the value of' simple mean is taken into consideration.

(iii) In these three series, it is quite possible that the value of arithmetic mean is 10; but the value of median may differ from each other. This can be understood as follows :

I	*II*	*III*
10	2	8
10 Median	8 Median	10 Median
10	20	12

The value of 'Median' in 1st series is 10, in 2nd series = 8 and in 3rd series = 10. Therefore, the value of the Mean and Median are not identical.

(iv) Even though the average remains the same, the nature and extent of the distribution of the size of the items may vary. In other words, the structure of the frequency distributions may differ even though their means are identical.

What is Dispersion?

Simplest meaning that can be attached to the word 'dispersion' is a lack of uniformity in the sizes or quantities of the items of a group or series. According to Reiglemen, "Dispersion is the extent to which the

magnitudes or quantities of the items differ, the degree of diversity." The word dispersion may also be used to indicate the spread of the data.

In all these definitions, we can find the basic property of dispersion as a value that indicates the extent to which all other values are dispersed about the central value in a particular distribution.

Properties of a good measure of Dispersion

There are certain pre-requisites for a good measure of dispersion:

1. It should be simple to understand.
2. It should be easy to compute.
3. It should be rigidly defined.
4. It should be based on each individual item of the distribution.
5. It should be capable of further algebraic treatment.
6. It should have sampling stability.
7. It should not be unduly affected by the extreme items.

Types of Dispersion

The measures of dispersion can be either 'absolute' or 'relative'. Absolute measures of dispersion are expressed in the same units in which the original data are expressed. For example, if the series is expressed as Marks of the students in a particular subject; the absolute dispersion will provide the value in Marks. The only difficulty is that if two or more series are expressed in different units, the series cannot be compared on the basis of dispersion.

'Relative' or 'Coefficient' of dispersion is the ratio or the percentage of a measure of absolute dispersion to an appropriate average. The basic advantage of this measure is that two or more series can be compared with each other despite the fact they are expressed in different units.

Theoretically, 'Absolute measure' of dispersion is better. But from a practical point of view, relative or coefficient of dispersion is considered better as it is used to make comparison between series.

Methods of Dispersion

Methods of studying dispersion are divided into two types :

(i) *Mathematical Methods:* We can study the 'degree' and 'extent' of variation by these methods. In this category, commonly used measures of dispersion are :
   (a) Range
   (b) Quartile Deviation
   (c) Average Deviation
   (d) Standard deviation and coefficient of variation.

(ii) *Graphic Methods:* Where we want to study only the extent of variation, whether it is higher or lesser a Lorenz-curve is used.

Mathematical Methods

(a) **Range:** It is the simplest method of studying dispersion. Range is the difference between the smallest value and the largest value of a series. While computing range, we do not take into account frequencies of different groups.

Formula : Absolute Range = L-S

$$\text{Coefficient of Range} = \frac{L-S}{L+S}$$

where, L represents largest value in a distribution

S represents smallest value in a distribution

We can understand the computation of range with the help of examples of different series.

(i) **Raw Data:** Marks out of 50 in a subject of 12 students, in a class are given as follows:

12, 18, 20, 12, 16, 14, 30, 32, 28, 12, 12 and 35.

In the example, the maximum or the highest marks obtained by a candidate is '35' and the lowest marks obtained by a candidate is '12'.Therefore, we can calculate range;

$$L = 35 \text{ and } S = 12$$

$$\text{Absolute Range} = L - S = 35 - 12 = 23 \text{ marks}$$

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{35-12}{35+12} = \frac{23}{47} = 0.49 \text{ approx.}$$

(ii) **Discrete Series**

	Marks of the Students in Accounts (out of 50)	No. of students
	(X)	*(f)*
Smallest	10	4
	12	10
	18	16
Largest	20	15
		Total = 45

$$\text{Absolute Range} = 20 - 10 = 10 \text{ marks}$$

$$\text{Coefficient of Range} = \frac{20-10}{20+10} = \frac{10}{30} = 0.34 \text{ approx.}$$

(iii) **Continuous Series**

	X	*Frequencies*
	10—15	4
S = 10	15—20	10
L = 30	20—25	26
	25—30	8

$$\text{Absolute Range} = L - S = 30 - 10 = 20 \text{ marks}$$

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{35-12}{35+12} = \frac{20}{40} = 0.5 \text{ approx.}$$

Range is a simplest method of studying dispersion. It takes lesser time to compute the 'absolute' and 'relative' range. Range does not take into account all the values of a series, *i.e.* it considers only the extreme items and middle items are not given any importance. Therefore, Range cannot tell us anything about the character of the distribution. Range cannot be computed in the case of 'open ends' distribution i.e., a distribution where the lower limit of the first group and upper limit of the higher group is not given.

The concept of range is useful in the field of quality control and to study the variations in the prices of the shares etc.

(b) Quartile Deviation (Q.D.)

The concept of 'Quartile Deviation' does take into account only the values of the 'Upper quartile' (Q_3) and the 'Lower quartile' (Q_1). Quartile Deviation is also called 'inter-quartile range'. It is a better method when we are interested in knowing the range within which certain proportion of the items fall.

'Quartile Deviation' can be obtained as :

(i) Inter-quartile range = $Q_3 - Q_1$

(ii) Semi-quartile range = $\frac{Q_3 - Q_1}{2}$

(iii) Coefficient of Quartile Deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

Calculation of Inter-quartile Range, semi-quartile Range and Coefficient of Quartile Deviation in case of Raw Data

Suppose the values of X are: 20, 12, 18, 25, 32, 10

In case of quartile-deviation, it is necessary to calculate the values of Q_1 and Q_3 by arranging the given data in ascending or descending order.

Therefore, the arranged data are (in ascending order):

X = 10, 12, 18, 20, 25, 32

No. of items = 6

$$Q_1 = \text{the value of } \left(\frac{N+1}{4}\right) \text{th item} = \left(\frac{6+1}{4}\right) = 1.75\text{th item}$$

= the value of 1st item + 0.75 (value of 2nd item – value of 1st item)

$= 10 + 0.75\,(12 - 10) = 10 + 0.75\,(2) = 10 + 1.50 = 11.50$

$$Q_3 = \text{the value of } 3\left(\frac{N+1}{4}\right) \text{th item} = 3\left(\frac{6+1}{4}\right)$$

= the value of 3(7/4)th item = the value of 5.25th item

= the value of 5th item + 0.25 (the value of 6th item *minus* the value of 5th item)

$= 25 + 0.25\,(32 - 25) = 25 + 0.25\,(7) = 26.75$

Therefore,

(i) Inter-quartile range = $Q_3 - Q_1 = 26.75 - 11.50 = 15.25$

(ii) Semi-quartile range = $\frac{Q_3 - Q_1}{2} = \frac{15.25}{2} = 7.625$

(iii) Coefficient of Quartile Deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{26.75 - 11.50}{26.75 + 11.50} = \frac{15.25}{38.25} = 0.39$ approx.

Calculation of Inter-quartile Range, semi-quartile Range and Coefficient of Quartile Deviation in discrete series

Suppose a series consists of the salaries (Rs.) and number of the workers in a factory:

Salaries (Rs.)	*No. of workers*
60	4
100	20
120	21
140	16
160	9

In the problem, we will first compute the values of Q_3 and Q_1

Salaries (Rs.) (x)	*No. of workers* (f)	*Cumulative frequencies* (c.f.)	
60	4	4	
100	20	24	— Q_1 lies in this cumulative
120	21	45	frequency
140	16	61	— Q_3 lies in this cumulative
160	9	70	frequency
	$N = \Sigma f = 70$		

Calculation of Q_1 :

$$Q_1 = \text{size of } \left(\frac{N+1}{4}\right) \text{th item}$$

$$= \text{size of } \left(\frac{70+1}{4}\right) \text{th item} = 17.75\text{th item}$$

17.75 lies in the cumulative frequency 24, which is corresponding to the value Rs. 100

$\therefore\ Q_1 = \text{Rs. } 100$

Calculation of Q_3 :

$$Q_3 = \text{size of } 3\left(\frac{N+1}{4}\right) \text{th item}$$

$$= \text{size of } 3\left(\frac{70+1}{4}\right) \text{th item} = 53.25\text{th item}$$

53.25 lies in the cumulative frequency 61 which is corresponding to Rs. 140

$\therefore\ Q_3 = \text{Rs. } 140$

(i) Inter-quartile range $= Q_3 - Q_1 = \text{Rs. } 140 - \text{Rs. } 100 = \text{Rs. } 40$

(ii) Semi-quartile range $= \dfrac{Q_3 - Q_1}{2} = \left(\dfrac{140-100}{2}\right) = \text{Rs. } 20$

(iii) Coefficient of Quartile Deviation $= \dfrac{Q_3 - Q_1}{Q_3 + Q_1} = \dfrac{140-100}{140+100} = \dfrac{40}{240} = 0.17$ approx.

Calculation of Inter-quartile range, semi-quartile range and Coefficient of Quartile Deviation in the case of continuous series

We are given the following data :

Salaries (Rs.)	*No. of Workers*
10—20	4
20—30	6
30—40	10
40—50	5
	Total = 25

In this example, the values of Q_3 and Q_1 are obtained as follows:

Salaries (Rs.) (x)	*No. of workers* (f)	*Cumulative frequencies* (c.f.)
10—20	4	4
20—30	6	10
30—40	10	20
40—50	5	25
	N = 25	

$$Q_1 = l_1 + \frac{\frac{N}{2} - cf_0}{f} \times i \qquad \left[\frac{N}{4} \text{ is used to find out } Q_1 \text{ group.}\right]$$

Therefore, $\frac{N}{4} = \frac{25}{4} = 6.25$. It lies in the cumulative frequency 10, which is corresponding to class 20—30.

Therefore, Q_1 group is 20—30.

$$Q_1 = 20 + \frac{6.25 - 4}{6} \times 10 = 20 + 3.75 = 23.75$$

where, $l_1 = 20,\ f = 6,\ i = 10,\ \frac{N}{4} = 6.25,$ and $cf_0 = 4$

$$Q_3 = l_1 + \frac{\frac{3N}{4} - cf_0}{f} \times i$$

Therefore, $\frac{3N}{4} = \frac{3 \times 25}{4} = \frac{75}{4} = 18.75$, which lies in the cumulative frequency 20, which is corresponding to class 30 — 40. Therefore Q_3 group is 30—40.

where, $l_1 = 30,\ i = 10,\ \frac{3N}{4} = 18.75,\ cf_0 = 10,$ and $f = 10$

$$Q_3 = 30 + \frac{18.75 - 10}{10} \times 10 = \text{Rs. } 38.75$$

Therefore :

(i) Inter-quartile range = $Q_3 - Q_1$ = Rs. 38.75 – Rs. 23.75 = Rs. 15.00

(ii) Semi-quartile range = $\frac{Q_3 - Q_1}{2} = \frac{38.75 - 23.75}{2} = 7.50$

(iii) Coefficient of Quartile Deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{\text{Rs. } 38.75 - \text{Rs. } 23.75}{\text{Rs. } 38.75 + \text{Rs. } 23.75} = \frac{15}{62.50} = 0.24$

Advantages of Quartile Deviation

Some of the important advantages are :

(i) It is easy to calculate. We are required simply to find the values of Q_1 and Q_3 and then apply the formula of absolute and coefficient of quartile deviation.

(ii) It has better results than range method. While calculating range, we take only the extreme values that make dispersion erratic. In the case of quartile deviation, we take into account middle 50% items.

(iii) The quartile deviation is not affected by the extreme items.

Disadvantages

(i) It is completely dependent on the central items. If these values are irregular and abnormal the result is bound to be affected.

(ii) All the items of the frequency distribution are not given equal importance in finding the values of Q_1 and Q_3.

(iii) Because it does not take into account all the items of the series, considered to be inaccurate of dispersion,

Similarly, sometimes we calculate percentile range, say, 90th and l0th percentile as it gives slightly better measure of dispersion, in certain cases. If we consider the calculations, then

(i) Absolute percentile range = $P_{90} - P_{10}$

(ii) Coefficient of percentile range = $\frac{P_{90} - P_{10}}{P_{90} + P_{10}}$

This method of calculating dispersion can be applied generally in the case of open end series where the importance of extreme values are not considered.

(c) Average Deviation

Average deviation is defined as a value, which is obtained by taking the average of the deviations of various items, from a measure of central tendency, Mean or Median or Mode, after ignoring negative signs.

Generally, the measure of central-tendency, from which the deviations are taken, is specified in the problem. If nothing is mentioned regarding the measure of central tendency specified than deviations are taken from median because the sum of the deviations (after ignoring negative signs) is minimum.

Computation in case of raw data

(i) Absolute Average Deviation about Mean or Median or Mode = $\frac{\Sigma |d|}{N}$

where: N = Number of observations,

|d| = deviations taken from Mean or Median or Mode ignoring signs.

(ii) Coefficient of A.D. = $\frac{\text{Average Deviation about Mean or Median or Mode}}{\text{Mean or Median or Mode}}$

Steps to Compute Average Deviation :

(i) Calculate the value of Mean or Median or Mode

(ii) Take deviations from the given measure of central-tendency and they are shown as d.

(iii) Ignore the negative signs of the deviation that can be shown as $|d|$ and add them to find $\Sigma|d|$.

(iv) Apply the formula to get Average Deviation about Mean or Median or Mode.

Example: Suppose the values are 5, 5, 10, 15, 20. We want to calculate Average Deviation and Coefficient of Average Deviation about Mean or Median or Mode.

Solution : Average Deviation about mean (Absolute and Coefficient).

(x)	*Deviation from mean* d	*Deviations after ignoring signs* $\lvert d \rvert$	
5	− 6	6	$\overline{X} = \frac{\Sigma X}{N} = \frac{55}{5} = 11$
5	− 6	6	where N = 5, ΣX = 55
10	+ 1	1	
15	+ 4	4	
20	+ 9	9	
$\Sigma X = 55$		$\Sigma \lvert d \rvert = 26$	

$$\text{Average Deviation about Mean} = \frac{\Sigma|d|}{N} = \frac{26}{5} = 5.2.$$

$$\text{Coefficient of Average Deviation about Mean} = \frac{\text{Mean Deviation about Mean}}{\text{Mean}} = \frac{5.2}{11} = 0.47$$

Average Deviation (Absolute and Coefficient) about Median

	X	*Deviation from median* d	*Deviations after ignoring negative signs* $\lvert d \rvert$
	5	− 5	5
	5	− 5	5
Median	10	0	0
	15	+ 5	5
	20	+ 10	10
	$N = 5$		$\Sigma \lvert d \rvert = 25$

$$\text{Average Deviation about Median} = \frac{\Sigma|d|}{N} = \frac{25}{5} = 5$$

$$\text{Coefficient of Average Deviation about median} = \frac{\text{A.D. about Mean}}{\text{Mean}} = \frac{5}{10} = 0.5$$

Average Deviation (Absolute and Coefficient) about Mode

	X	*Deviation from mode d*	$\lvert d \rvert$
	5	0	0
Mode	5	0	0
	10	+ 5	5
	15	+ 10	10
	20	+ 15	15
	$N = 5$		$\Sigma \lvert d \rvert = 30$

$$\text{Average deviation about Mode} = \frac{\Sigma |d|}{N} = \frac{30}{5} = 6$$

Coefficient of Average Deviation about Mode

$$= \frac{\text{A.D. about Mode}}{\text{Mode}} = \frac{6}{5} = 1.2$$

Average deviation in case of discrete and continuous series

$$\text{Average Deviation about Mean or Median or Mode} = \frac{\Sigma f |d|}{N}$$

where N = No. of items

|d| = deviations from Mean or Median or Mode, after ignoring negative signs.

$$\text{Coefficent of A.D. about Mean or Median or Mode} = \frac{\text{A.D. about Mean or Median or Mode}}{\text{Value of Mean or Median or Mode}}$$

Example: Suppose we want to calculate coefficient of Average Deviation about Mean from the following descrete series:

X	Frequency
10	5
15	10
20	15
25	10
30	5

Solution: First of all, we shall calculate the value of arithmetic Mean,

Calculation of Arithmetic Mean

X	f	fX	
10	5	50	
15	10	150	$\overline{X} = \frac{\Sigma fX}{N} = \frac{900}{45} = 20$
20	15	300	
25	10	250	
30	5	150	
	N = 45	$\Sigma fX = 900$	

Calculation of Coefficient of Average Deviation about Mean

X	f	*Deviation from mean* d	*Deviations after ignoring negative signs* $\lvert d \rvert$	$\Sigma f \lvert d \rvert$
10	5	− 10	10	50
15	10	− 5	5	50
20	15	0	0	0
25	10	+ 5	5	50
30	5	+ 10	10	50
	$N = 55$			$\Sigma f \lvert d \rvert = 200$

$$\text{Coefficient of Average Deviation about Mean} = \frac{\text{A.D. about mean}}{\text{Mean}} = \frac{4.4}{20} = 0.22$$

$$\text{Average Deviation about Mean} = \frac{\Sigma \lvert d \rvert}{N} = \frac{200}{45} = 4.44 \text{ approx.}$$

In case we want to calculate coefficient of Average Deviation about Median from the following data:

Class Interval	*Frequency*
10—14	5
15—19	10
20—24	15
25—29	10
30—34	5
	N = 45

First of all we shall calculate the value of Median but it is necessary to find the 'real limits' of the given class-intervals. This is possible by subtracting 0.5 from the lower-limits and added to the upper limits of the given classes. Hence, the real limits shall be : 9.5—14.5, 14.5—19.5, 19.5—24.5, 24.5—29.5 and 29.5—34.5

Calculation of Median

Class Interval	f	*Cumulative Frequency*
9.5—14.5	5	5
14.5—19.5	10	15
19.5—24.5	15	30
24.5—29.5	10	40
29.5—34.5	5	45
	$N = 5$	

$$\text{Median} = l_1 + \frac{\frac{N}{2} - cf_0}{f} \times i$$

Where l_1 = lower limit of median group

i = magnitude of median group

f = frequency of median group

Cf_0 = cumulative frequency of the group preceeding median group

$\frac{n}{2}$ = size of median group

$\therefore$ Median size = $\frac{N}{2}$ th item i.e. $\frac{45}{2} = 22.5$

It lies in the cumulative frequency 30, which is corresponding to class 19.5—24.5.

Median group is 19.5—24.5

$$\text{Median} = 19.5 + \frac{22.5 - 15}{15} \times 5 = 19.5 + \frac{7.5}{15} \times 5 = 19.5 + 2.5 = 19.5 + 2.5 = 22$$

Calculation of coefficient of Averagae Deviation about Median

Class Intervals	*Frequency* f	*Mid points* x	*Deviation from median* (22)	*Deviations after ignoring negative signs* $\lvert d \rvert$	$f \lvert d \rvert$
9.5—14.5	5	12	− 10	10	50
14.5—19.5	10	17	− 5	5	50
19.5—24.5	15	22	0	0	0
24.5—29.5	10	27	+ 5	5	50
29.5—34.5	5	32	+ 10	10	50
	$N = 45$				$\Sigma f \lvert d \rvert = 200$

$$\text{Coefficient of Average Deviation about Median} = \frac{\text{A.D. about Median}}{\text{Median}}$$

$$\text{Average Deviation about Mean} = \frac{\Sigma |d|}{N} = \frac{200}{45} = 4.44 \text{ approx.}$$

$$\text{Coefficient of A.D. about Median} = \frac{4.4}{22} = 0.2$$

Advantages of Average Deviations

1. Average deviation takes into account all the items of a series and hence, it provides sufficiently representative results.
2. It simplifies calculations since all signs of the deviations are taken as positive.
3. Average Deviation may be calculated either by taking deviations from Mean or Median or Mode.
4. Average Deviation is not affected by extreme items.

5. It is easy to calculate and understand.
6. Average deviation is used to make healthy comparisons.

Disadvantages of Average Deviations

1. It is illogical and mathematically unsound to assume all negative signs as positive signs.
2. Because the method is not mathematically sound, the results obtained by this method are not reliable.
3. This method is unsuitable for making comparisons either of the series or structure of the series.

This method is more effective during the reports presented to the general public or to groups who are not familiar with statistical methods.

(d) Standard Deviation

The standard deviation, which is shown by greek letter σ (read as sigma) is extremely useful in judging the representativeness of the mean. The concept of standard deviation, which was introduced by Karl Pearson, has a practical significance because it is free from all defects, which exists in a range, quartile deviation or average deviation.

Standard deviation is calculated as the square root of average of squared deviations taken from actual mean. It is also called root mean square deviation. The square of standard deviation i.e. σ^2 is called 'variance'.

Calculation of standard deviation in case of raw data

There are four ways of calculating standard deviation for raw data:

(i) When actual values are considered;
(ii) When deviations are taken from actual mean;
(iii) When deviations are taken from assumed mean; and
(iv) When 'step deviations' are taken from assumed mean.

(i) When the actual values are considered:

$$\sigma = \sqrt{\frac{\Sigma X^2}{N} - (\overline{X})^2}$$

or $$\sigma^2 = \frac{\Sigma X^2}{N} - (\overline{X})^2$$

where, N = Number of the items,

X = Given values in the series,

$\overline{X}$ = Arithmetic mean of the values

We can also write the formula as follows :

$$\sigma = \sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2} \quad \text{where,} \quad \overline{X} = \frac{\Sigma X}{N}$$

Steps to calculate σ

(i) Compute simple mean of the given values.
(ii) Square the given values and aggregate them
(iii) Apply the formula to find the value of standard deviation.

Example: Suppose the values are given 2, 4, 6, 8, 10. We want to apply the formula

$$\sigma = \sqrt{\frac{\Sigma X^2}{N} - (\overline{X})^2}$$

Solution: We are required to calculate the values of $N, \bar{X}, \Sigma X^2$. They are calculated as follows :

X	X^2
2	4
4	16
6	36
8	64
10	100
N = 5	$\Sigma X^2 = 220$

$$\sigma = \sqrt{\frac{220}{5} - (6)^2} = \sqrt{44 - 36} = \sqrt{8} = 2.828$$

Variance $(\sigma^2) = (\sqrt{8})^2 = 8.$

$$\bar{X} = \frac{\Sigma X}{N} = \frac{30}{5} = 6.$$

(ii) When the deviations are taken from actual mean

$$\sigma = \sqrt{\frac{\Sigma x^2}{N}}$$ where, N = no. of items and $x = (X - \bar{X})$

Steps to Calculate σ

(i) Compute the deviations of given values from actual mean i.e., $(X - \bar{X})$ and represent them by x.

(ii) Square these deviations and aggegate them

(iii) Use the formula, $\sigma = \sqrt{\frac{\Sigma x^2}{N}}$

Example. We are given values as 2, 4, 6, 8, 10. We want to find out standard deviation.

X	$(X - \bar{X}) = x$	x^2
2	2 – 6 = – 4	$(-4)^2 = 16$
4	4 – 6 = – 2	$(-2)^2 = 4$
6	6 – 6 = 0	$= 0$
8	8 – 6 = + 2	$(2)^2 = 4$
10	10 – 6 = + 4	$(4)^2 = 16$
$N = 5$		$\Sigma x^2 = 40$

$$\therefore \quad \bar{X} = 6 \quad \left(\frac{\Sigma X}{N} = \frac{30}{5}\right) \text{ and}$$

$$\sigma = \sqrt{\frac{\Sigma x^2}{N}} = \sqrt{\frac{40}{5}} = \sqrt{8} = 2.828$$

(iii) When the deviations are taken from assumed mean

$$\sigma = \sqrt{\frac{\Sigma dx^2}{N} - \left(\frac{\Sigma dx}{N}\right)^2}$$

where, N = no. of items,
dx = deviations from assumed mean i.e., (X – A).
A = assumed mean

Steps to Calculate :

(i) We consider any value as assumed mean. The value may be given in the series or may not be given in the series.

(ii) We take deviations from the assumed value i.e., (X – A), to obtain dx for the series and aggregate them to find Σdx.

(iii) We square these deviations to obtain dx^2 and aggregate them to find Σdx^2.

(iv) Apply the formula given above to find standard deviation.

Example. Suppose the values are given as 2, 4, 6, 8 and 10. We can obtain the standard deviation as:

X	$dx = (X - A)$	x^2
2	$-2 = (2-4)$	4
assumed mean (A) 4	$0 = (4-4)$	0
6	$+2 = (6-4)$	4
8	$+4 = (8-4)$	16
10	$+6 = (10-4)$	36
$N = 5$	$\Sigma dx = 10$	$\Sigma dx^2 = 60$

$$\sigma = \sqrt{\frac{\Sigma dx^2}{N} - \left(\frac{\Sigma dx}{N}\right)^2} = \sqrt{\frac{60}{5} - \left(\frac{10}{5}\right)^2} = \sqrt{12-4} = \sqrt{8} = 2.828$$

(iv) When step deviations are taken from assumed mean

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

where, i = Common factor, N = Number of items, dx (Step-deviations) = $\left(\frac{X-A}{i}\right)$

Steps to Calculate σ :

(i) We consider any value as assumed mean from the given values or from outside.

(ii) We take deviation from the assumed mean *ie.*, (X – A),

(iii) We divide the deviations obtained in step (ii) with a common factor to find step deviations $\left(\frac{X-A}{i}\right)$ and represent them as dx and aggregate them to obtain Σdx.

(iv) We square the step deviations to obtain dx^2 and aggregate them to find Σdx^2.

Example: We continue with the same example to understand the computation of Standard Deviation.

X	$d = (X - A)$	$dx = \left(\frac{d}{i}\right)$ and $i = 2$	dx^2
2	–2	1	1
A = 4	0	0	0
6	+2	1	1
8	+4	2	4
10	+6	3	9
$N = 5$		$\Sigma dx = 5$	$\Sigma dx^2 = 15$

$$\sigma = \sqrt{\frac{\Sigma dx^2}{N} - \left(\frac{\Sigma dx}{N}\right)^2} \times i \qquad \text{where } N = 5,\ i = 2,\ dx = 5,\ \text{and } \Sigma dx^2 = 15$$

$$\sigma = \sqrt{\frac{15}{5} - \left(\frac{5}{5}\right)^2} \times 2 = \sqrt{3-1} \times 2 = \sqrt{2} \times 2 = 1.414 \times 2 = 2.828.$$

Note: We can notice an important point that the standard deviation value is identical by four methods. Therefore any of the four formulae can be applied to find the value of standard deviation. But the suitability of a formula depends on the magnitude of items in a question.

Coefficient of Standard-deviation $= \dfrac{\sigma}{\overline{X}}$

In the above given example, $\sigma = 2.828$ and $\overline{X} = 6$

Therefore, coefficient of standard deviation $= \dfrac{\sigma}{\overline{X}} = \dfrac{2.828}{6} = 0.471$

Coefficient of Variation or C. V.

$$= \frac{\sigma}{\overline{X}} \times 100 = \frac{2.828}{6} \times 100 = 47.1\%$$

Generally, coefficient of variation is used to compare two or more series. If coefficient of variation (C.V.) is more for one series as compared to the other, there will be more variations in that series, lesser stability or consistency in its composition. If coefficient of variation is lesser as compared to other series, it will be more stable, or consistent. Moreover, that series is always better where coefficient of variation or coefficient of standard deviation is lesser.

Example. Suppose we want to compare two firms where the salaries of the employees are given as follows:

	Firm A	**Firm B**
No. of workers	100	100
Mean salary (Rs.)	100	80
Standard-deviation (Rs.)	40	45

Solution: We can compare these firms either with the help of coefficient of standard deviation or coefficient of variation. If we use coefficient of variation, then we shall apply the formula :

$$\text{C.V.} = \left(\frac{\sigma}{\overline{X}} \times 100\right)$$

Firm A	**Firm B**
$\text{C.V.} = \dfrac{40}{100} \times 100 = 40\%$	$\text{C.V.} = \dfrac{45}{80} \times 100 = 56.25\%$
$\overline{X} = 100, \sigma = 40.$	$\overline{X} = 80, \sigma = 45.$

Because the coefficient of variation is lesser for firm A as compared to firm B, therefore, firm A is better.

Calculation of standard-deviation in discrete and continuous series

We use the same formula for calculating standard deviation for a discrete series and a continuous series. The only difference is that in a discrete series, values and frequencies are given whereas in a continuous series, class-intervals and frequencies are given. When the mid-points of these class-intervals are obtained, a continuous series takes shape of a discrete series. X denotes values in a discrete series and mid points in a continuous series.

When the deviations are taken from actual mean

We use the same formula for calculating standard deviation for a continuous series

$$\sigma = \sqrt{\frac{\Sigma f x^2}{N}}$$

where N = Number of items

f = Frequencies corresponding to different values or class-intervals.

x = Deviations from actual mean $(X - \overline{X})$

X = Values in a discrete series and mid-points in a continuous series.

Step to calculate σ

(i) Compute the arithmetic mean by applying the required formula.

(ii) Take deviations from the arithmetic mean and represent these deviations by x.

(iii) Square the deviations to obtain values of x .

(iv) Multiply the frequencies of different class-intervals with x^2 to find fx^2. Aggregate fx^2 column to obtain Σfx^2.

(v) Apply the formula to obtain the value of standard deviation.

If we want to calculate variance then we can compute $\sigma^2 = \frac{\Sigma fx^2}{N}$

Example : We can understand the procedure by taking an example :

Class Intervals	*Frequency* (f)	*Midpoints* (m)	fm
10—14	5	12	60
15—19	10	17	170
20—24	15	22	330
25—29	10	27	270
30—34	5	32	160
	N = 45		$\Sigma fm = 990$

Therefore, $\overline{X} = \frac{\Sigma fm}{N} = \frac{990}{45} = 22$ where, N = 45, $\Sigma fm = 990$

Calculation of Standard Deviation

Class Intervals	f	*Mid points* X	*Deviations from actual median = 22* x $(X-22)$	x^2	fx^2
10—14	5	12	− 10	100	500
15—19	10	17	− 5	25	250
20—24	15	22	0	0	0
25—29	10	27	+ 5	25	250
30—34	5	32	+ 10	100	500
	$N = 45$				$\Sigma fx^2 = 1500$

$$\sigma = \sqrt{\frac{\Sigma f x^2}{N}} \text{ where, } N = 45,\ \Sigma fx^2 = 1500$$

$$\sigma = \sqrt{\frac{1500}{45}} = \sqrt{33.33} = 5.77 \text{ approx.}$$

When the deviations are taken from assumed mean

In some cases, the value of simple mean may be in fractions, then it becomes time consuming to take deviations and square them. Alternatively, we can take deviations from the assumed mean.

$$\sigma = \sqrt{\frac{\Sigma\, fdx^2}{N} - \left(\frac{\Sigma\, fdx}{N}\right)^2}$$

where N = Number of the items,

dx = deviations from assumed mean (X – A),

f = frequencies of the different groups,

A = assumed mean and

X = values or mid points.

Steps to calculate σ

(i) Take the assumed mean from the given values or mid points.

(ii) Take deviations from the assumed mean and represent them by dx.

(iii) Square the deviations to get dx^2.

(iv) Multiply f with dx of different groups to obtain fdx and add them up to get Σfdx.

(v) Multiply f with dx^2 of different groups to obtain fdx^2 and add them up to get Σfdx^2.

(vi) Apply the formula to get the value of standard deviation.

Example : We can understand the procedure with the help of an example.

Class Intervals	*Frequency* f	*Mid points* x	*Deviations from assumed Mean* dx (X–17)	dx^2	fdx	fdx^2
10—14	5	12	– 5	25	– 25	125
15—19	10	17	0	0	0	0
20—24	15	22	+ 5	25	75	375
15—29	10	27	+ 10	100	100	1000
30—34	5	32	+ 15	225	75	1125
	N = 45				Σfdx = 225	Σfdx^2 = 2625

$$\sigma = \sqrt{\frac{\Sigma\, fdx^2}{N} - \left(\frac{\Sigma\, fdx}{N}\right)^2} \text{ where, } N = 45,\ \Sigma\, fdx^2 = 1500,\ \Sigma\, fdx = 225$$

$$\therefore \sigma = \sqrt{\frac{2625}{45} - \left(\frac{225}{45}\right)^2} = \sqrt{58.33 - 25} = \sqrt{33.33} = 5.77 \text{ approx.}$$

When the step deviations are taken from the assumed mean

$$\sigma = \sqrt{\frac{\Sigma\, fdx^2}{N} - \left(\frac{\Sigma\, fdx}{N}\right)^2} \times i$$

where N = Number of the items (Σf),

i = common factor,

f = frequencies corresponding to different groups,

dx = step-deviations $\left(\frac{X - A}{i}\right)$

Steps to calculate σ

(i) Take deviations from the assumed mean of the calculated mid-points and divide all deviations by a common factor (i) and represent these values by dx.

(ii) Square these step deviations dx to obtain dx^2 for different groups.

(iii) Multiply f with dx of different groups to find fdx and add them to obtain Σfdx.

(iv) Multiply f with dx^2 of different groups to find fdx^2 for different groups and add them to obtain Σfdx^2.

(v) Apply the formula to get standard deviation.

Example : Suppose we are given the series and we want to calculate standard deviation with the help of step deviation method. According to the given formula, we are required to calculate the value of $i, N, \Sigma fdx$ and Σfdx^2.

Class Intervals	*Frequency*	*Mid point*	*Deviations from assumed mean (22)*	$i = 5$ $\left(\frac{X - A}{i}\right)$			
	f	X	x	dx	dx^2	fdx	fdx^2
10—14	5	12	− 10	− 2	4	− 10	20
15—19	10	17	− 5	− 1	1	− 10	10
20—24	15	22	+ 0	0	0	0	0
25—29	10	27	+ 5	+ 1	1	10	10
30—34	5	32	+ 10	+ 2	4	10	20
	N = 45					$\Sigma fdx = 0$	$\Sigma fdx^2 = 60$

$$\sigma = \sqrt{\frac{\Sigma fd^2x}{N} - \left(\frac{\Sigma\, fdx}{N}\right)^2} \times i \text{ where, } N = 45,\ i = 5,\ fdx = 0,\ fdx^2 = 60$$

$$\therefore \quad \sigma = \sqrt{\frac{60}{45} - \left(\frac{0}{45}\right)^2} \times 5 = \sqrt{\frac{4}{3}} \times 5 = \sqrt{1.33} \times 5 = 1.154 \times 5 = 5.77 \text{ approx.}$$

Advantages of Standard Deviation

(i) Standard deviation is the best measure of dispersion because it takes into account all the items and is capable of future algebric treatment and statistical analysis.

(ii) It is possible to calculate standard deviation for two or more series.

(iii) This measure is most suitable for making comarisons among two or more series about varibility.

Disadvantages

(i) It is difficult to compute.

(ii) It assigns more weights to extreme items and less weights to items that are nearer to mean. It is because of this fact that the squares of the deviations which are large in size would be proportionately greater than the squares of those deviations which are comparatively small.

Mathematical properties of standard deviation (σ)

(i) If deviations of given items are taken from arithmetic mean and squared then the sum of squared deviation should be minimum, i.e., $\left(X-\overline{X}\right)^2 = \text{Minimum}$.

(ii) If different values are increased or decreased by a constant, the standard deviation will remain the same. Whereas if different values are multiplied or divided by a constant than the standard deviation will be multiplied or divided by that constant.

(iii) Combined standard deviation can be obtained for two or more series with below given formula:

$$\sigma_{12} = \sqrt{\frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_1d_1^2 + N_2d_2^2}{N_1 + N_2}}$$

where: N_1 represents number of items in first series,

N_2 represents number of items in second series,

σ_1^2 represents variance of first series,

σ_2^2 represents variance of second series,

d_1 represents the difference between $\overline{X}_{12} - \overline{X}_1$

d_2 represents the difference between $\overline{X}_{12} - \overline{X}_2$

$\overline{X}_1$ represents arithmetic mean of first series,

$\overline{X}_2$ represents arithmetic mean of second series,

$\overline{X}_{12}$ represents combined arithmetic mean of both the series.

Example : Find the combined stnadard deviation of two series, from the below given information :

	First Series	**Second Series**
No. of items	10	15
Arithmetic means	15	20
Standard deviation	4	5

Solution : Since we are considering two series, therefore combined standard deviation is computed by the following formula :

$$\sigma_{12} = \sqrt{\frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_1d_1^2 + N_2d_2^2}{N_1 + N_2}}$$

where : $N_1 = 10,\quad N_2 = 15,\quad \overline{X}_1 = 15,\quad \overline{X}_2 = 20,\quad \sigma_1 = 4,\quad \sigma_2 = 5$

$$\overline{X}_{12} = \frac{\overline{X}_1 N_1 + \overline{X}_2 N_2}{N_1 + N_2}$$

or $$\overline{X}_{12} = \frac{(15\times 10)+(20\times 15)}{10+15} = \frac{150+300}{25} = \frac{450}{25} = 18$$

$$d_1 = \left(\overline{X}_{12} - \overline{X}_1\right) = 18 - 15 = 3 \text{ and } d_2 = \left(\overline{X}_{12} - \overline{X}_2\right) = 18 - 20 = -2$$

By applying the formula of combined standard deviation, we get :

$$\sigma_{12} = \sqrt{\frac{10(4)^2 + 15(5)^2 + 10(18-15)^2 + 15(18-20)^2}{10+15}}$$

$$= \sqrt{\frac{(10\times 16)+(15\times 25)+(10\times 9)+(15\times 4)}{25}}$$

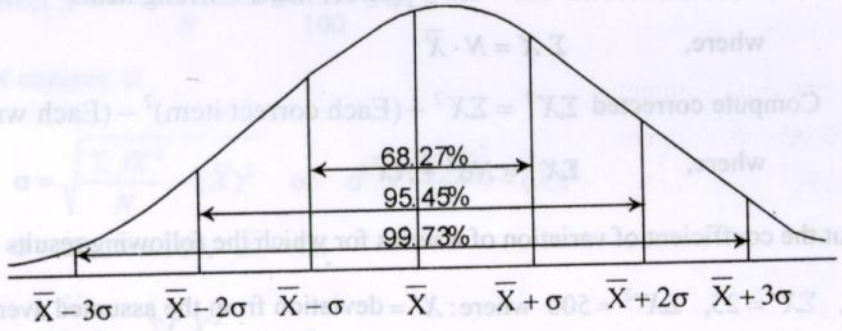
$$= \sqrt{\frac{160+375+90+60}{25}} = \sqrt{\frac{685}{25}} = \sqrt{27.4} = 5.2 \text{ approx.}$$

(iv) Standard deviation of n natural numbers can be computed as :

$$\sigma = \sqrt{\frac{1}{12}\left(N^2 - 1\right)} \qquad \text{where,} \qquad \text{N represents numbers number of items.}$$

(v) For a symmetrical distribution

$\overline{X} \pm \sigma$ covers 68.27% of items,
$\overline{X} \pm 2\sigma$ covers 95.45% of items,
$\overline{X} \pm 3\sigma$ covers 99.73% of items,



Example : You are heading a rationing department in a State affected by food shortage. Local investigators submit the following report :

Daily calorie value of food available per adult during current period :

Area	Mean	Standard deviation
A	2,500	400
B	2,000	200

The estimated requirement of an adult is taken at 2,800 calories daily and the absolute minimum is 1,350. Comment on the reported figures, and determine which area, in your opinion, need more urgent attention.

Solution : We know that $\bar{X} \pm \sigma$ covers 68.27% of items, $\bar{X} \pm 2\sigma$ covers 95.45% of items, $\bar{X} \pm 3\sigma$ covers 99.73%. In the given problem if we take into consideration 99.73%. i.e., almost the whole population, the limits would be $\bar{X} \pm 3\sigma$.

For Area A these limits are :

$$\bar{X} + 3\sigma = 2{,}500 + (3 \times 400) = 3{,}700$$
$$\bar{X} - 3\sigma = 2{,}500 - (3 \times 400) = 1{,}300$$

For Area B these limits are :

$$\bar{X} + 3\sigma = 2{,}000 + (3 \times 200) = 2{,}600$$
$$\bar{X} - 3\sigma = 2{,}000 - (3 \times 200) = 1{,}400$$

It is clear from above limits that in Area A there are some persons who are getting 1300 calories, i.e. below the minimum which is 1,350. But in case of area B there is no one who is getting less than the minimum. Hence area A needs more urgent attention.

(vi) Relationship between quartile deviation, average deviation and standard deviation is given as:

Quartile deviation = 2/3 Standard deviation

Average deviation = 4/5 Standard deviation

(vii) We can also compute corrected standard deviation by using the following formula :

$$\text{Correct}\sigma = \sqrt{\frac{\text{Correct } \Sigma X^2}{N} - (\text{correct } \bar{X})^2}$$

(a) Compute corrected $\bar{X} = \dfrac{\text{Corrected } \Sigma X}{N}$

where, corrected $\Sigma X = \Sigma X$ + correct items – wrong items

where, $\Sigma X = N \cdot \bar{X}$

(b) Compute corrected $\Sigma X^2 = \Sigma X^2 + (\text{Each correct item})^2 - (\text{Each wrong item})^2$

where, $\Sigma X^2 = N\sigma^2 + N\bar{X}^2$

Example : Find out the coefficient of variation of a series for which the following results are given :

$N = 50$, $\Sigma X' = 25$, $\Sigma X'^2 = 500$ where : X' = deviation from the assumed average 5.

(b) For a frequency distribution of marks in statistics of 100 candidates, (grouped in class inervals of 0—10, 10—20) the mean and standard deviation, were found to be 45 and 20. Later it was discovered that the score 54 was misread as 64 in obtaining frequency distribution. Find out the correct mean and correct standard deviation of the frequency destribution.

(c) Can coefficient of variation be greater than 100%? If so, when?

Solution : (a) We want to calculate, coefficient of variation, which is $= \dfrac{\sigma}{\bar{X}} \times 100$

Therefore, we are required to calculate mean and standard deviation.

Calculation of simple mean

$$\bar{X} = A + \frac{\Sigma X'}{N} \quad \text{where,} \quad A = 5, \quad N = 50, \quad \Sigma X' = 25$$

$$\therefore \qquad \bar{X} = 5 + \frac{25}{50} = 5.5$$

Calculation of standard deviation

$$\sigma = \sqrt{\frac{\Sigma X'^2}{N} - \left(\sqrt{\frac{\Sigma X'}{N}}\right)^2} = \sqrt{\frac{500}{50} - \left(\frac{25}{50}\right)^2} = \sqrt{5 - 0.25} = \sqrt{4.75} = 2.179$$

Calculation of Coefficient of variation

$$\text{C.V.} = \frac{\sigma}{\bar{X}} \times 100 = \frac{2.179}{5.5} \times 100 = \frac{217.9}{5.5} = 39.6\%$$

(b) Given $\bar{X} = 45$, $\sigma = 20$, $N = 100$, wrong value = 64, correct value = 54

Since this is a case of continuous series, therefore, we will apply the formula for mean and standard devation that are applicable in a continuous series.

Calculation of correct Mean

$$\bar{X} = \frac{\Sigma fx}{N} \text{ or } N\bar{X} = \Sigma fX$$

By substituting the values, we get $100 \times 45 = 4500$

$$\text{Correct } \Sigma fX = 4500 - 64 + 54 = 4490$$

$$\therefore \quad \text{Correct } \bar{X} = \frac{\text{Correct } \Sigma fx}{N} = \frac{4490}{100} = 44.9$$

Calculation of correct σ

$$\sigma = \sqrt{\frac{\Sigma fX^2}{N} - (\bar{X})^2} \quad \text{or} \quad \sigma^2 = \frac{\Sigma fX^2}{N} - (\bar{X})^2$$

where, $\sigma = 20$, $N = 100$, $\bar{X} = 45$.

$$(20)^2 = \frac{\Sigma f X^2}{100} - (45)^2$$

$$\text{or} \qquad 400 = \frac{\Sigma f X^2}{100} - 2025$$

$$\text{or} \quad 400 + 2025 = \frac{\Sigma f X^2}{100}$$

$$\text{or} \quad 2425 \times 100 = \Sigma fX^2 = 242500$$

$$\therefore \text{ Correct } \Sigma fX^2 = 242500 - (64)^2 + (54)^2 = 242500 - 4096 + 2916 = 242500 - 1180 = 241320$$

$$\text{Correct } \sigma = \sqrt{\frac{\text{Correct } \Sigma fX^2}{N} - (\text{correct } \bar{X})^2}$$

$$= \sqrt{\frac{241320}{100} - (49.9)^2} = \sqrt{2413.20 - 2016.01} = \sqrt{397.19} = 39.9 \text{ approx.}$$

(c) The formulae for the computation of coefficient of variation is $= \left\{\frac{\sigma}{\bar{X}} \times 100\right\}$

Hence, coefficient of variation can be greater than 100% only when the value of standard deviation is greater than the value of mean.

This will happen when data contains a large number of small items and few items are quite large. In such a case the value of simple mean will be pulled down and the value of standard deviation will go up.

Similarly, if there are negative items in a series, the value of mean will come down and the value of standard deviation shall not be affected because of squaring the deviations.

Example : In a distribution of 10 observations, the value of mean and standard deviation are given as 20 and 8. By mistake, two values are taken as 2 and 6 instead of 4 and 8. Find out the value of correct mean and variance.

Solution : We are given; N = 10, $\bar{X} = 20, \sigma = 3$

Wrong values = 2 and 6 and Correct values = 4 and 8

Calculation of correct Mean

$$\bar{X} = \frac{\Sigma X}{N} \text{ or } \bar{X} = \Sigma X$$

$$\Sigma X = 10 \times 20 = 200$$

But ΣX is incorrect. Therefore we shall find correct ΣX.

Correct $\Sigma X = 200 - 2 - 6 + 4 + 8 = 204$

$$\text{Correct Mean} = \frac{\text{Correct } \Sigma X}{N} = \frac{204}{10} = 20.4$$

Calculation of correct variance

$$\sigma = \sqrt{\frac{\Sigma X^2}{N} - (\bar{X})^2}$$

or $$\sigma^2 = \frac{\Sigma X^2}{N} - (\bar{X})^2$$

or $$(8)^2 = \frac{\Sigma X^2}{10} - (20)^2$$

or $$64 + 400 = \frac{\Sigma X^2}{10}$$

or $$\Sigma X^2 = 4640$$

But this is wrong and hence we shall compute correct ΣX^2

$$\text{Correct } \Sigma X^2 = 4640 - 2^2 - 6^2 + 4^2 + 8^2$$

$$= 4640 - 4 - 36 + 16 + 64 = 4680$$

$$\text{Correct } \sigma^2 = \frac{\text{Correct } \Sigma X^2}{N} - \text{Correct } (\bar{X})^2$$

$$= \frac{4680}{10} - (20.4)^2 = 468 - 416.16 = 51.84$$

Revisionary Problems

Example : Compute (a) Inter-quartile range, (b) Semi-quartile range, and (c) Coefficient of quartile deviation from the following data :

Farm Size (acres)	*No. of firms*	*Farm Size (acres)*	*No. of firms*
0—40	394	161—200	169
41—80	461	201—240	113
81—120	391	241 and over	148
121—160	334		

Soultion :

In this case, the real limits of the class intervals are obtained by subtracting 0.5 from the lower limits of each class and adding 0.5 to the upper limits of each class. This adjustment is necessary to calculate median and quartiles of the series.

Farm Size (acres)	*No. of firms*	*Cumulative frequency (c.f.)*
– 0.5—40.5	394	394
40.5—80.5	461	855
80.5—120.5	391	1246
120.5—160.5	334	1580
160.5—200.5	169	1749
200.5—240.5	113	1862
240.5 and over	148	2010
	N = 2010	

$$Q_1 = l_1 + \frac{N/4 - c.f_0}{f} \times i$$

$\because$ $\frac{n}{4} = \frac{2010}{4} = 502$th item

Q_1 lies in the cumulative frequency of the group 40.5–80.5, and

l_1=40.5, f =461, i = 40, $c.f_0$ = 394, $\frac{n}{4}$ = 502.5

$\therefore$ $Q_1 = 40.5 + \frac{502.5 - 394}{461} \times 40 = 40.5 + 9.4 = 49.4$ acres

Similarly,

$$Q_3 = l_1 + \frac{\frac{3n}{4} - c.f_0}{4} \times i$$

$$\because \quad \frac{3n}{4} = \frac{3 \times 2010}{4} \times 1507.5 \text{ th item}$$

Q_3 lies in the cumulative frequency of the group 121-160, where the real limits of the class interval are 120.5-160.5 and $l_1 = 120.5$, $i = 40$, $f = 334$, $\frac{3n}{4} = 1507.5$, $c.f. = 1246$

$$\therefore \quad Q_3 = 120.5 + \frac{1507.5 - 1246}{334} \times 40 = 120.5 + 31.3 = 151.8 \text{ acres}$$

Inter-quartile range $= Q_3 - Q_1 = 151.8 - 49.9 = 101.9$ acres

$$\text{Semi-quartile range} = \frac{Q_3 - Q_1}{2} = \frac{151.8 - 49.9}{2} = 50.95 \text{ approx.}$$

$$\text{Coefficient of quartile deviation} = \frac{Q_3 - Q_1}{Q_3 - Q_1} = \frac{151.8 - 49.9}{151.8 - 49.9} = \frac{101.9}{201.7} = 0.5 \text{ approx}$$

Example : Calculate mean and coefficient of mean deviation about mean from the following data :

Marks less than	No. of students
10	4
20	10
30	20
40	40
50	50
60	56
70	60

Solution :

In this question, we are given less than type series alongwith the cumulative frequencies. Therefore, we are required first of all to find out class intervals and frequencies for calculating mean and coefficient of mean deviation about mean.

Marks	*No. of students*	*Mid points*	*Deviations from assumed Mean (A = 35)*	*Step Deviation i = 10*		*Deviation from mean (35) (ignoring signs)*	
	f	X	X'	$dx = \left(\frac{X - A}{i}\right)$	$\lvert dx \rvert$	fdx	$f\lvert dx \rvert$
0—10	4	5	–30	–3	3	–12	12
10—20	6	15	–20	–2	2	–12	12

20—30	10	25	−10	−1	1	−10	10
30—40	20	35	0	0	0	0	0
40—50	10	45	+10	+1	1	+10	10
50—60	6	55	+20	+2	2	+12	12
60—70	4	65	+30	+3	3	+12	12
	N = 60						$\Sigma fdx = 0$ $\Sigma f \lvert dx \rvert = 68$

$$\bar{X} = A + \frac{\Sigma f\, dx}{N} \times i$$

where, $N = 60$, $A = 35$, $i = 10$, $\Sigma fdx = 0$

$$\therefore \quad \bar{X} = 35 + \left(\frac{0}{60} \times 10\right) = 35$$

$$\text{M.D. about mean} = \frac{\Sigma f \lvert dx \rvert}{N} \times i = \frac{68}{60} \times 10 = 11.33$$

$$\text{Coefficient of M.D. about mean} = \frac{\text{M.D. about mean}}{\text{mean}} = \frac{11.33}{35} = 0.324 \text{ approx.}$$

Example : Calculate standard deviation from the following data :

Class Interval	*frequency*
−30 to −20	5
−20 to −10	10
−10 to 0	15
0 to 10	10
10 to 20	5
	N = 45

Solution:

Calculation of Standard Deviation

Class Intervals	*Frequency* f	*Mid points* X	*Deviations from assumed Mean (A = −5)* X'	*Step Derivations when i = 10* $dx = \left(\frac{X-A}{i}\right)$	dx^2	fdx	fdx^2
−30 to −20	5	−25	−20	−2	4	−10	20
−20 to −10	10	−15	−10	−1	1	−10	10
−10 to 0	15	−5	+0	0	0	0	0
0 to 10	10	5	+10	1	1	10	10
10 to 20	5	15	+20	2	4	10	20
	N = 45					$\Sigma fdx = 0$	$\Sigma fdx^2 = 60$

$$\sigma = \sqrt{\frac{\Sigma\, fdx^2}{N} - \left(\frac{\Sigma\, fdx}{N}\right)^2} \times i$$

where, $N = 45$, $i = 10$, $\Sigma fdx = 0$, $\Sigma fdx^2 = 60$

$$\therefore \quad \sigma = \sqrt{\frac{60}{45} - \left(\frac{0}{45}\right)^2} \times 10 = \sqrt{\frac{60}{45}} \times 10 = \sqrt{1.33} \times 10 = 1.153$$

Example : For two firms A and B belonging to same industry, the following details are available :

	Firm A	**Firm B**
Number of Employees :	100	200
Average wage per month :	Rs. 240	Rs. 170
Standard deviation of the wage per month :	Rs. 6	Rs. 8

Find (i) Which firm pays out larger amount as monthly wages?

(ii) Which firm shows greater variability in the distribution of wages?

(iii) Find average monthly wages and the standard deviation of wages of all employees for both the firms.

Solution : (i) For finding out which firm pays larger amount, we have to find out ΣX.

$$X = \frac{\Sigma X}{N} \quad \text{or} \quad \Sigma X = NX$$

Firm A : $N = 100$, $X = 240$ $\quad \therefore\ \Sigma X = ,00 \times 240 = 24000$

Firm B : $N = 200$, $X = 170$ $\quad \therefore\ \Sigma X = 200 \times 170 = 34000$

Hence firm B pays larger amount as monthly wages.

(ii) For finding out which firm shows greater variability in the distribution of wages, we have to calculate coefficient of variation.

Firm A : $\text{C.V.} = \frac{\sigma}{X} \times 100 = \frac{6}{240} \times 100 = 2.50$

Firm B : $\text{C.V.} = \frac{\sigma}{X} \times 100 = \frac{8}{170} \times 100 = 4.71.$

Since coefficient of variation is greater for firm B, hence it shows greater variability in the distribution of wages.

(iii) Combined wages : $\bar{X}_{12} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$

where, $N_1 = 100$, $\bar{X}_1 = 240$, $N_2 = 200$, $\bar{X}_2 = 170$

Hence $\bar{X}_{12} = \frac{(100 \times 240) + (200 \times 170)}{100 + 200} = \frac{24000 + 34000}{300} = 193.33.$

Combined Standard Deviation :

$$\sigma_{12} = \sqrt{\frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_1d_1^2 + N_2d_2^2}{N_1 + N_2}}$$

where $N_1 = 100,\ N_2 = 200,\ \sigma_1 = 6,\ \sigma_2 = 8,\ d_1 = (\overline{X}_1 - \overline{X}_{12}) = 240 - 193.3 = 46.7$

and $d_2 = (\overline{X}_2 - \overline{X}_{12}) = 170 - 193.3 = -23.3$

$$\sigma_{12} = \sqrt{\frac{(100)(36) + (200)(64) + (100)(46.7)^2 + (200)(-23.3)^2}{100 + 200}}$$

$$= \sqrt{\frac{3600 + 12800 + 218089 + 108578}{300}} = \sqrt{\frac{451643}{300}} = 38.8.$$

Example : From the following frequency distribution of heights of 360 boys in the age-group 10-20 years, calculate the :

(i) arithmetic mean;

(ii) coefficient of variation; and

(iii) quartile deviation

Height (cms)	*No. of boys*	*Height (cms)*	*No. of boys*
126—130	31	146—150	60
131—135	44	151—155	55
136—140	48	156—160	43
141—145	51	161—165	28

Solution : **Calculation of $\overline{X}$, Q.D., and C.V.**

Heights	*m.p.* X	f	$(X-143)/5$ dx	fdx	fdx^2	*c.f.*
126—130	128	31	-3	-93	279	31
131—135	133	44	-2	-88	176	75
136—140	138	48	-1	-48	48	123
141—145	143	51	0	0	0	174
146—150	148	60	+1	+60	60	234
151—155	153	55	+2	+10	220	289
156—160	158	43	+3	+129	387	332
161—165	163	28	+4	+112	448	360
	$N = 45$			$\Sigma fdx = 182$	$\Sigma fdx^2 = 1618$	

(i) $\bar{X} = A + \frac{\Sigma fdx}{N} \times i$ where, $N = 360$, $A = 143$, $i = 5$, $\Sigma fdx = 182$

$\therefore\ \bar{X} = 143 + \frac{182}{360} \times 5 = 143 + 2.53 = 145.53.$

(ii) $\text{C.V.} = \frac{\sigma}{X} \times 100$

$$\sigma = \sqrt{\frac{\Sigma\, fdx^2}{N} - \left(\frac{\Sigma\, fdx}{N}\right)^2} \times i = \sqrt{\frac{1618}{360} - \left(\frac{182}{360}\right)^2} \times 5$$

$$= \sqrt{4.494 - 0.506} \times 5 = 2.00 \times 5 = 10$$

$$\text{C.V.} = \frac{10}{145.53} \times 100 = 6.87 \text{ percent}$$

(iii) $\text{Q.D.} = \frac{Q_3 - Q_1}{2}$

$Q_1 = \text{Size of } \frac{N}{4} \text{th observation} = \frac{360}{4} = 90\text{th observation}$

Q_1 lies in the class $136 - 140$. But the real limits of this class is $135.5 - 140.5$

$$Q_1 = l_1 + \frac{N/4 - cf_0}{f} \times i = 135.5 + \frac{90 - 75}{48} \times 5 = 135.5 + 1.56 = 137.06$$

$Q_3 = \text{Size of } \frac{3N}{4} \text{th observation} = 3 \times \frac{360}{4} = 270\text{th observation.}$

Q_3 lies in the class $151 - 155$. But the real limit of this class is $150.5 - 155.5$

$$Q_3 = l_1 + \frac{3N/4 - cf_0}{f} \times i = 150.5 + \frac{270 - 234}{55} \times 5 = 150.5 + 3.27 = 153.77$$

$$\text{Q.D.} = \frac{Q_3 - Q_1}{2} = \frac{153.77 - 137.06}{2} = 8.355$$

UNIT - II

LESSON 1

CORRELATION

In the earlier chapters we have discussed univariate distributions to highlight the important characteristics by different statistical techniques. Univariate distribution means the study related to one variable only. We may however come across certain series where each item of the series may assume the values of two or more variables. The distributions in which each unit of series assumes two values is called bivariate distribution. In a bivariate distribution, we are interested to find out whether there is any relationship between two variables. The correlation is a statistical technique which studies the relationship between two or more variables and correlation analysis involves various methods and techniques used for studying and measuring the extent of relationship between the two variables. When two variables are related in such a way that a change in the value of one is accompanied either by a direct change or by an inverse change in the values of the other, the two variables are said to be correlated. In the correlated variables an increase in one variable is accompanied by an increase or decrease in the other variable. For instance, relationship exists between the price and demand of a commodity because keeping other things equal, an increase in the price of a commodity shall cause a decrease in the demand for that commodity. Relationship might exist between the heights and weights of the students and between amount of rainfall in a city and the sales of raincoats in that city.

These are some of the important definitions about correlation.

Croxton and Cowden says, "When the relationship is of a quantitative nature, the appropriate statistical tool for discovering and measuring the relationship and expressing it in a brief formula is known as correlation".

A.M. Tuttle says, "Correlation is an analysis of the covariation between two or more variables."

W.A. Neiswanger says, "Correlation analysis contributes to the understanding of economic behaviour, aids in locating the critically important variables on which others depend, may reveal to the economist the connections by which disturbances spread and suggest to him the paths through which stabilizing forces may become effective.

L.R. Conner says, "If two or more quantities vary in sympathy so that the movements in one tends to be accompanied by corresponding movements in others than they are said be correlated.

Utility of Correlation

The study of correlation is very useful in practical life as revealed by these points.

1. With the help of correlation analysis, we can measure in one figure, the degree of relationship existing between variables like price, demand, supply, income, expenditure etc. Once we know that two variables are correlated then we can easily estimate the value of one variable, given the value of other.

2. Correlation analysis is of great use to economists and businessmen, it reveals to the economists the disturbing factors and suggest to him the stabilizing forces. In business, it enables the executive to estimate costs, sales etc. and plan accordingly.

3. Correlation analysis is helpful to scientists. Nature has been found to be a multiplicity of inter-related forces.

Difference between Correlation and Causation

The term correlation should not be misunderstood as causation. If correlation exists between two variables, it must not be assumed that a change in one variable is the cause of a change in other variable. In simple words, a change in one variable may be associated with a change in another variable but this change need not necessarily be the cause of a change in the other variable. When there is no cause and effect relationship between two variables but a correlation is found between the two variables such correlation is known as "spurious correlation" or "nonsense correlation". Correlation may exist due to the following:

1. *Pure change correlation:* This happens in a small sample. Correlation may exist between incomes and weights of four persons although there may be no cause and effect relationship between

incomes and weights of people. This type of correlation may arise due to pure random sampling variation or because of the bias of investigator in selecting the sample.

2. *When the correlated variables are influenced by one or more variables.* A high degree of correlation between the variables may exist, where the same cause is affecting each variable or different cause affecting each with the same effect. For instance, a degree of correlation may be found between yield per acre of rice and tea due to the fact that both are related to the amount of rainfall but none of the two variables is the cause of other.

3. *When the variable mutually influence each other so that neither can be called the cause of other.* At times it may be difficult to say that which of the two variables is the cause and which is the effect because both may be reacting on each other.

Types of Correlation

Correlation can be one of the following :

(i) Positive and Negative.

(ii) Simple and Multiple.

(iii) Partial and Total.

(iv) Linear and Non-Linear (Curvilinear)

Positive and Negative Correlation

Positive or direct Correlation refers to the movement of variables in the same direction. The correlation is said to be positive when the increase (decrease) in the value of one variable is accompanied by an increase (decrease) in the value of other variable also. Negative or inverse correlation refers to the movement of the variables in opposite direction. Correlation is said to be negative, if an increase (decrease) in the value of one variable is accompanied by a decrease (increase) in the value of other.

Simple and Multiple Correlation

Under simple correlation, we study the relationship between two variables only i.e., between the yield of wheat and the amount of rainfall or between demand and supply of a commodity. In case of multiple correlation, the relationship is studied among three or more variables. For example, the relationship of yield of wheat may be studied with both chemical fertilizers and the pesticides.

Partial and Total Correlation

There are two categories of multiple correlation analysis. Under partial correlation, the relationship of two or more variables, is studied in such a way that only one dependent variable and one independent variable is considered and all others are kept constant. For example, coefficient of correlation between yield of wheat and chemical fertilizers excluding the effects of pesticides and manures is called partial correlation. Total correlation is based upon all the variables simultaneously.

Linear and Non-Linear Correlation

When the amount of change in one variable tends to keep a constant ratio to the amount of change in the other variable then the correlation is said to be linear. But if the amount of change in one variable does not bear a constant ratio to the amount of change in the other variable then the correlation is said to be non-linear. The distinction between linear and non-linear correlation is dependent upon the consistency of the ratio of change between the variables.

Methods of Studying Correlation

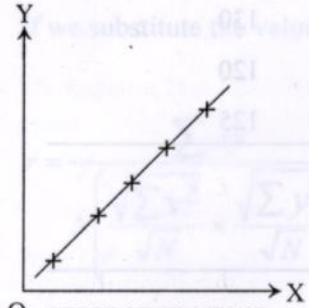
There are different methods which helps us to find out whether the variables are related or not.

1. Scatter Diagram Method.
2. Karl Pearson's Coefficient of correlation.

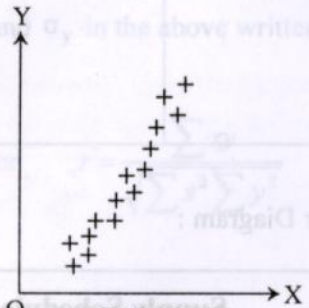
We shall discuss these methods one by one.

(1) **Scatter Diagram:** Scatter diagram is drawn to visualise the relationship between two variables. The values of more important variable is plotted on the X-axis while the values of the other variable are plotted on the Y-axis. On the graph, dots are plotted to represent different pairs of data. When dots are plotted to represent all the pairs, we get a scatter diagram. The way the dots scatter gives an indication of the kind of relationship exists between the two variables. While drawing scatter diagram, it is not necessary to take at the point of sign the zero values of X and Y variables, but the minimum values of the variables considered may be taken.

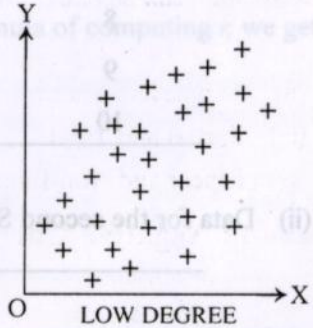
When there is a positive correlation between the variables, the dots on the scatter diagram run from left hand bottom to the right hand upper corner. In case of perfect positive correlation all the dots will lie on a straight line.



PERFECT POSITIVE CORRELATION

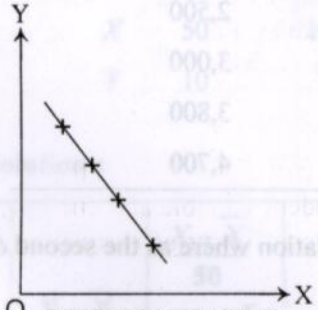


HIGH DEGREE POSITIVE CORRELATION

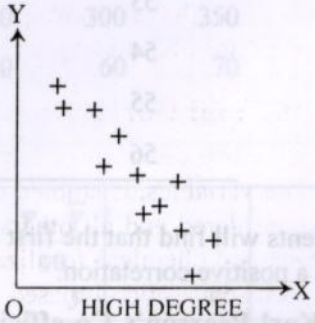


LOW DEGREE POSITIVE CORRELATION

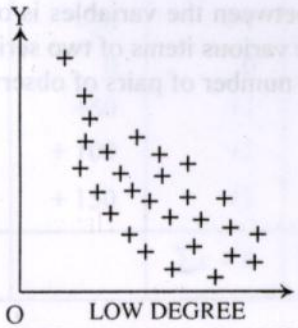
When a negative correlation exists between the variables, dots on the scatter diagram run from the upper left hand corner to the bottom right hand corner. In case of perfect negative correlation, all the dots lie on a straight line.



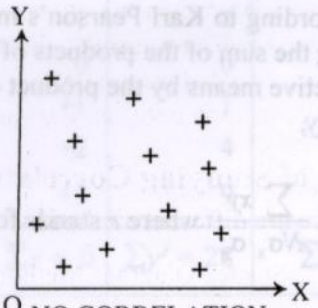
PERFECT NEGATIVE CORRELATION



HIGH DEGREE NEGATIVE CORRELATION



LOW DEGREE NEGATIVE CORRELATION



NO CORRELATION

If a scatter diagram is drawn and no path is formed, there is no correlation. Students are advised to prepare two scatter diagrams on the basis of the following data :

(i) Data for the first Scatter Diagram :

Demand Schedule

Price (Rs.)	Commodity Demand (units)
6	180
7	150
8	130
9	120
10	125

(ii) Data for the second Scatter Diagram :

Supply Schedule

Price (Rs.)	Commodity Supply
50	2,000
51	2,100
52	2,200
53	2,500
54	3,000
55	3,800
56	4,700

Students will find that the first diagram indicate a negative correlation where as the second diagram shall reveal a positive correlation.

(2) **Karl Pearson's Co-efficient of Correlation.** Karl Pearson's method, popularly known as Pearsonian co-efficient of correlation, is most widely applied in practice to measure correlation. The Pearsonian co-efficient of correlation is represented by the symbol r.

According to Karl Pearson's method, co-efficient of correlation between the variables is obtained by dividing the sum of the products of the corresponding deviations of the various items of two series from their respective means by the product of their standard deviations and the number of pairs of observations. Symbolically,

$$r = \frac{\sum xy}{N\sigma_x \sigma_y}$$ where r stands for coefficient of correlation ....(*i*)

where $x_1, x_2, x_3, x_4 \ldots\ldots x_n$ are the deviations of all items of the first variable from the mean,

$y_1, y_2, y_3 \ldots\ldots y_n$ are the deviations of all items of the second variable from the mean,

Σxy is the sum of products of these corresponding deviations. N refers to the number of pairs, σ_x stands for the standard deviation of X variable and σ_y stands for the standard deviation of Y variable.

$$\sigma_x = \sqrt{\frac{\sum x^2}{N}} \quad \text{and} \quad \sigma_y = \sqrt{\frac{\sum y^2}{N}}$$

If we substitute the value of σ_x and σ_y in the above written formula of computing r, we get

$$r = \frac{\sum xy}{N\left(\frac{\sqrt{\sum x^2}}{\sqrt{N}} \times \frac{\sqrt{\sum y^2}}{\sqrt{N}}\right)} \quad \text{or} \quad r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}} \qquad \ldots.(ii)$$

Degree of correlation varies between + 1 and –1; the result will be + 1 in case of perfect positive correlation and – 1 in the case of perfect negative correlation.

Computation of correlation coefficient can be simplified by dividing the given data by a common factor. In such a case, the final result is not multiplied by the common factor because coefficient of correlation is independent of change of scale and origin.

Illustration: Calculate Co-efficient of Correlation from the following data :

X	50	100	150	200	250	300	350
Y	10	20	30	40	50	60	70

Solution :

X	$X - \overline{X}$	$\frac{X - \overline{X}}{50}$ x	x^2	Y	$Y - \overline{Y}$	$\frac{Y - \overline{Y}}{10}$ y	y^2	xy
50	–150	–3	9	10	–30	–3	9	9
100	–100	–2	4	20	–20	–2	4	4
150	–50	–1	1	30	–10	–1	1	1
200	0	0	0	40	0	0	0	0
250	+50	+1	1	50	+ 10	+1	1	1
300	+ 100	+2	4	60	+20	+2	4	4
350	+ 150	+3	9	70	+30	+3	9	9
		$\Sigma x = 0$	$\Sigma x^2 = 28$			$\Sigma x = 0$	$\Sigma y^2 = 28$	$\Sigma xy = 28$

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

By substituting the values, we get $r = \frac{28}{\sqrt{28 \times 28}} = \frac{28}{28} = 1$

Hence there is perfect positive correlation.

Illustration: A sample of five items is taken from the production of a firm, length and weight of the five items are given below:

Length (inches)	3	4	6	7	10
Weight (ounces)	9	11	14	15	16

Calculate Karl Pearson's correlation co-efficient between length and weight and interpret the value of correlation coefficient.

Solution : $\overline{X} = \frac{\sum X}{N} = \frac{30}{5} = 6$ and $\overline{Y} = \frac{\sum Y}{N} = \frac{65}{5} = 13$

X	$X - \overline{X}$ x	x^2	Y	$Y - \overline{Y}$ y	y^2	xy
3	−3	9	9	−4	16	12
4	−2	4	11	−2	4	4
6	0	0	14	+1	1	0
7	+1	1	15	+2	4	2
10	+4	16	16	+3	9	12
$\Sigma x = 30$	0	30	$\Sigma y = 65$	0	34	30

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}} \quad \text{where,} \quad \Sigma xy = 30,\ \Sigma x^2 = 30,\ \text{and } \Sigma y^2 = 34$$

$$\therefore \quad r = \frac{30}{\sqrt{30 \times 34}} = \frac{30}{\sqrt{1020}} = +\mathbf{0.939\ Ans.}$$

The value of r indicates that there exists a high degree positive correlation between lengths and weights.

Illustration: From the following data, compute the co-efficient of correlation between X and Y:

	X Series	Y Series
Number of items	15	15
Arithmetic Mean	25	18
Square of deviation from Mean	136	138

Summation of product deviations of X and Y from their Arithmetic Means = 122

Solution: Denoting deviations of X and Y from their arithmetic means by x and y respectively, the given data are : $\Sigma x^2 = 136$, $\Sigma xy = 122$, and $\Sigma y^2 = 138$

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}} = \frac{122}{\sqrt{136 \times 138}} = \frac{122}{137} = 0.89 \text{ Ans.}$$

Short-cut Method: To avoid difficult calculations due to mean being in fraction, deviations are taken from assumed means while calculating coefficient of correlation. The formula is also modified for standard deviation because deviations are taken from assumed means. Karl Perason's formula for short-cut method is given below:

$$r = \frac{\Sigma dxdy - \frac{\Sigma dx . \Sigma dy}{N}}{\sqrt{\left\{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}\right\}\left\{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}\right\}}} \quad \text{or} \quad r = \frac{N\Sigma dxdy - \Sigma dx \times \Sigma dy}{\sqrt{\left\{N\Sigma dx^2 - (\Sigma dx)^2\right\}\left\{N\Sigma dy^2 - (\Sigma dy)^2\right\}}}$$

Illustration: Compute the coefficient of correlation from the following data :

Marks in Statistics	20	30	28	17	19	23	35	13	16	38
Marks in Mathematics	18	35	20	18	25	28	33	18	20	40

Solution:

Marks in Statistics X	**$(X - 30)$ dx**	**dx^2**	**Marks in Maths Y**	**$Y - 30$ dy**	**dy^2**	**$dxdy$**
20	–10	100	18	–12	144	+120
30	0	0	35	+5	25	0
28	–2	4	20	–10	100	+20
17	–13	169	18	–12	144	+156
19	–11	121	25	–5	25	+55
23	–7	49	28	–2	4	+14
35	+5	25	33	+3	9	+15
13	–17	289	18	–12	144	+204
16	–14	196	20	–10	100	+140
38	+8	64	40	+10	100	+80
N = 10	–61	1017		–45	795	804

$$r = \frac{N\,\Sigma\, dxdy - \Sigma\, dx . \Sigma\, dy}{\sqrt{\left\{N\,\Sigma\, dx^2 - (\Sigma\, dx)^2\right\}\left\{N\,\Sigma\, dy^2 - (\Sigma\, dy)^2\right\}}}$$

where dx — deviations of X series from an assumed mean 30.

dy — deviations of Y series from an assumed mean 30.

$\Sigma\, dx^2$ — sum of the squares of the deviations of X series from an assumed mean.

$\Sigma\, dy^2$ — sum of the squares of the deviations of Y series from an assumed mean.

$\Sigma\, dxdy$ — sum of the products of the deviations of X and Y series from an assumed mean.

$$r = \frac{10 \times 804 - (-61)(-45)}{\sqrt{10 \times 1017 - (-61)^2}\sqrt{10 \times 795 - (-45)^2}}$$

or $$r = \frac{8040 - 2745}{\sqrt{(10170 - 3721)(7950 - 2025)}} = \frac{5295}{\sqrt{6449 \times 5925}} = 0.856$$

Direct Method of Computing Correlation Coefficient

Correlation coefficient can also be computed from given X and Y values by using the below given formula:

$$r = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

The above given formula gives us the same answer as we are getting by taking deviations from actual mean or arbitrary mean.

Illustration: Compute the coefficient of correlations from the following data :

Marks in Statistics	20	30	28	17	19	23	35	13	16	38
Marks in Mathematics	18	35	20	18	25	28	33	18	20	40

Solution :

Marks in Statistics X	Marks in Mathematics Y	X^2	Y^2	XY
20	18	400	324	360
30	35	900	1225	1050
28	20	784	400	560
17	18	289	324	306
19	25	361	625	475
23	28	529	784	644
35	33	1225	1089	1155
13	18	169	324	234
16	20	256	400	320
38	40	1444	1600	1520
$\Sigma X = 239$	$\Sigma Y = 255$	$\Sigma X^2 = 6357$	$\Sigma Y^2 = 7095$	$\Sigma XY = 6624$

Substitute the computed values in the below given formula,

$$r = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}} = \frac{(10 \times 6624) - (239)(255)}{\sqrt{10 \times 6357 - (239)^2}\sqrt{10 \times 7095 - (255)^2}}$$

$$= \frac{66240 - 60945}{\sqrt{63570 - 57121}\sqrt{70950 - 65095}} = \frac{5295}{\sqrt{6449}\sqrt{5925}} = \frac{5295}{6181.45} = 0.856$$

Coefficient of Correlation in a Continuous Series

In case of a continuous series, we assume that every item which falls within a given class interval falls exactly at the middle of that class. The formula, because of the presence of frequencies is modified as follows :

$$r = \frac{\Sigma\, fdxdy - \frac{\Sigma\, fdx.\Sigma\, fdy}{\Sigma f}}{\sqrt{\left\{\Sigma\, fdx^2 - \frac{(\Sigma\, fdx)^2}{\Sigma f}\right\}\left\{\Sigma\, dy^2 - \frac{(\Sigma\, fdy)^2}{\Sigma f}\right\}}}$$

Various values shall be calculated as follows :

(i) Take the step deviations of variable X and denote it by *dx*.

(ii) Take the step deviations of variable Y and denote it by *dy*.

(iii) Multiply dx dy and the respective frequency of each cell and write the figure obtained in the right-hand upper corner of each cell.

(iv) Add all the cornered values calculated in step (iii) to get $\Sigma dxdy$

(v) Multiply the frequencies of the variable *X* by the deviations of *X* to get Σfdx.

(vi) Take the squares of the deviations of the variable X and multiply them by the respective frequencies to get Σfdx^2.

(vii) Multiply the frequencies of the variable *Y* by the deviations of *Y* to get Σfdy.

(viii) Take the squares of the deviations of the variable Y and multiply them by the respective frequencies to get Σfdy^2.

(ix) Now substitute the values of $\Sigma fdxdy$, Σfdx, Σfdx^2, Σfdy, Σfdy^2 in the formula to get the value of *r*.

Illustration: The following table gives the ages of husbands and wives at the time of their marriages. Calculate the correlation coefficient between the ages of husbands and wives.

Ages of Husbands

Age of Wives	20—30	30—40	40—50	50—60	60—70	Total
15—25	5	9	3	—	—	17
25—35	—	10	25	2	—	37
35—45	—	1	12	2	—	15
45—55	—	—	4	16	5	25
55—65	—	—	—	4	2	6
Total	5	20	44	24	7	100

Solution :

Age of wives (Y)	Age group	Mid points	Deviations from assumed mean 40	Step deviations c.f. = 10	20-30	30-40	40-50	50-60	60-70	f	fdy	fdy²	∑fdxdy
					Age of Husbands (X)								
					Age group	20-30	30-40	40-50	50-60	60-70				
					Mid points	25	35	45	55	65				
					Deviations from assumed mean 40	−20	−10	0	+10	+20				
					Step deviations c.f. = 10	−2	−1	0	+1	+2				
	15-25	20	−20	−2	5 [20]	9 [18]	3 [0]			17	−34	68	38
	25-35	30	−10	−1	–	10 [10]	25 [0]	2 [−2]		37	−37	37	8
	35-45	40	0	0	–	1 [0]	12 [0]	2 [0]		15	0	0	0
	45-55	50	+10	1	–	–	4 [0]	16 [16]	5 [10]	25	25	25	26
	55-56	60	+20	2	–	–	–	4 [8]	2 [8]	6	12	24	16
	Age group	Mid points	Deviations from assumed mean 40	Step deviations c.f. = 10	5	20	44	24	7	∑f = 100	∑fdy = −34	∑fdy² = 154	∑fdxdy = 88
					−10	−20	0	24	14	∑fdx = 8			
					20	20	0	24	28	∑fdx² = 92			
					20	28	0	22	18	∑fdxdy = 88			

$$r = \frac{\Sigma\, fdxdy - \dfrac{\Sigma\, fdx.\,\Sigma\, fdy}{\Sigma\, f}}{\sqrt{\left\{\Sigma\, fdx^2 - \dfrac{(\Sigma\, fdx)^2}{\Sigma\, f}\right\}\left\{\Sigma\, dy^2 - \dfrac{(\Sigma\, fdy)^2}{\Sigma\, f}\right\}}}$$

$$= \frac{88 - \dfrac{(+8)\,(-34)}{100}}{\sqrt{\left\{92 - \dfrac{(8)^2}{100}\right\}\left\{154 - \dfrac{(-34)^2}{100}\right\}}} = \frac{\dfrac{9072}{100}}{\sqrt{\dfrac{91.36}{100} + \dfrac{142.44}{100}}} = \frac{90.72}{\sqrt{91.36 + 14.244}} = +0.79$$

Properties of Coefficient of Correlation

Following are some of the important properties of r :

(1) The coefficient of correlation lies between −1 and + 1 $(-1 \le r \le +1)$

(2) The coefficient of correlation is independent of change of scale and origin of the variable X and Y.

(3) The coefficient of correlation is the geometric mean of two regression coefficients.

$$r = \sqrt{bxy \times dyx}$$

Merits of Pearson's coefficient of correlation : The correlation of coefficient summarizes in one figure the degree and direction of correlation. Value varies between +1 and –1.

Demerits of Pearson's coefficient of correlation : It always assumes linear relationship between the variables; in fact the assumption may be wrong. Secondly, it is not easy to interpret the significance of correlation coefficient. The method is time consuming and affected by the extreme items.

LESSON 2

REGRESSION ANALYSIS

The statistical technique correlation establshes the degree and direction of relationship between two or more variables. But we may be interested in estimating the value of an unknown variable on the basis of a known variable. If we know the index of money supply and price-level, we can find out the degree and direction of relationship between these indices with the help of correlation technique. But the regression technique helps us in determining the general price-level for a given supply of money. Similarly if we know that the price and demand of a commodity are correlated we can find out the demand for that commodity for a fixed price. Hence, the statistical tool with the help of which we can estimate or predict the unknown variable from known variable is called regression. The meaning of the term "Regression" is the act of returning or going back. This term was first used by Sir Francis Galton in 1877 when he studied the relationship between the height of fathers and sons. His study revealed a very interesting relationship. All tall fathers tend to have tall sons and all short fathers short sons but the average height of the sons of a group of tall fathers was less than that of the fathers and the average height of the sons of a group of short fathers was greater than that of the fathers. The line describing this tendency of going back is called "Regression Line". Modern writers have started to use the term estimating line instead of regression line because the expression estimating line is more clear in character. According to Morris Myers Blair, regression is the measure of the average relationship between two or more variables in terms of the original units of the data.

Regression analysis is a branch of statistical theory which is widely used in all the scientific disciplines. It is a basic technique for measuring or estimating the relationship among economic variables that constitute the essence of economic theory and economic life. The uses of regression analysis are not confined to economics and business activities. Its applications are extended to almost all the natural, physical and social sciences. The regression technique can be extended to three or more variables but we shall limit ourselves to problems having two variables in this lesson.

Some of the uses of the regression analysis are given below:

(i) Regression Analysis helps in establishing a functional relationship between two or more variables. Once this is established it can be used for various analytic purposes.

(ii) With the use of electronic machines and computers, the medium of calculation of regression equation particularly expressing multiple and non-linear relations has been reduced considerably.

(iii) Since most of the problems of economic analysis are based on cause and effect relationship, the regression analysis is a highly valuable tool in economic and business research.

(iv) The regression analysis is very useful for prediction purposes. Once a functional relationship is established the value of the dependent variable can be estimated from the given value of the independent variables.

Difference between Correlation and Regression

Both the techniques are directed towards a common purpose of establishing the degree and direction of relationship between two or more variables but the methods of doing so are different. The choice of one or the other will depend on the purpose. If the purpose is to know the degree and direction of relationship, correlation is an appropriate tool but if the purpose is to estimate a dependent variable with the substitution of one or more independent variables, the regression analysis shall be more helpful. The points of difference are discussed below:

(i) **Degree and Nature of Relationship :** The correlation coefficient is a measure of degree of covariability between two variables whereas regression analysis is used to study the nature of relationship between the variables so that we can predict the value of one on the basis of another. The reliance on the estimates or predictions depend upon the closeness of relationship between the variables.

(ii) **Cause and Effect Relationship:** The cause and effect relationship is explained by regression analysis. Correlation is only a tool to ascertain the degree of relationship between two variables and we can not say that one variable is the cause and other the effect. A high degree of correlation between price and demand for a commodity or at a particular point of time may not suggest which is the cause and which is the effect. However, in regression analysis cause and effect relationship is clearly expressed, one variable is taken as dependent and the other an independent.

The variable which is the basis of prediction is called independent variable and the variable that is to be predicted is called dependent variable. The independent variable is represented by X and the dependent variable by Y.

Principle of Least Squares

Regression refers to an average of relationship between a dependent variable with one or more independent variables. Such relationship is generally expressed by a line of regression drawn by the method of the "Least Squares". This line of regression can be drawn graphically or derived algebraically with the help of regression equations. The condition usually stipulated in regression analysis is that the sum of the squares of the deviations of the observed y values from the fitted line shall be minimum. This is known as the least squares or minimum squared error criterion.

A line fitted by the method of least squares is the *line of best fit.* The line satisfies the following conditions:

(i) The algebraic sum of deviations above the line and below the line are equal to zero.

$$\Sigma (x - x_c) = 0 \text{ and } \Sigma (y - y_c) = 0$$

Where x_c and y_c are the values derived with the help of regression technique.

(ii) The sum of the squares of all these deviations is less than the sum of the squares of deviations from any other line, we can say

$$\Sigma (x - x_c)^2 \text{ is smaller than } \Sigma (x - A)^2 \text{ and}$$
$$\Sigma (y - y_c)^2 \text{ is smaller than } \Sigma (y - A)^2$$

Where A is some other value or any other straight line.

(iii) The line of regression (best fit) intersect at the mean value of the variables i.e., $\bar{x}$ and $\bar{y}$

(iv) When the data represent a sample from a larger population, the least square line is the best estimate of the population line.

Methods of Regression Analysis

We can study regression by the following methods

1. Graphic method (regression lines)
2. Algebraic method (regression equations)

We shall discuss these methods in detail.

1. **Graphic Method:** When we apply this method different points are plotted on a graph paper representing different pairs of variables. These points give a picture of a seatter diagram with several points spread over. A regression line may be drawn between these points either by free hand or by a scale in such a way that the squares of the vertical or horizontal distances between the points and the line of regression is minimum. It should be drawn in such a manner that the line leaves equal number of points on both sides. However, to ensure this is rather difficult and the method only renders a rough estimate which can not be completely free from subjectivity of person drawing it. Such a line can be a straight line or a curved line depending upon the scatter of points and relationship to be established. A non-linear free hand curve will have more element of subjectivity and a straight line is generally drawn. Let us understand it with the help of an example:

Example

Height of fathers (Inches)	Height of sons (Inches)
65	68
63	66
67	68
64	65
68	69
62	66
70	68
66	65
68	71
67	67
69	68
71	70

Solution: The diagram given below shows the height of fathers on x–axis and the height of sons on y–axis. The line of regression called the regression of y on x is drawn between the scatter dots.

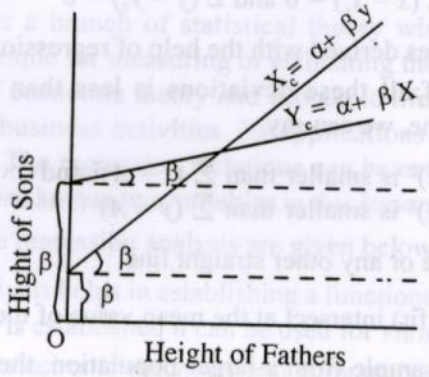


Fig. 1

Another line of regression called the regression line of x on y is drawn amongst the same set of seatter dots in such a way that the squares of the horizontal distances between dots are minimised.

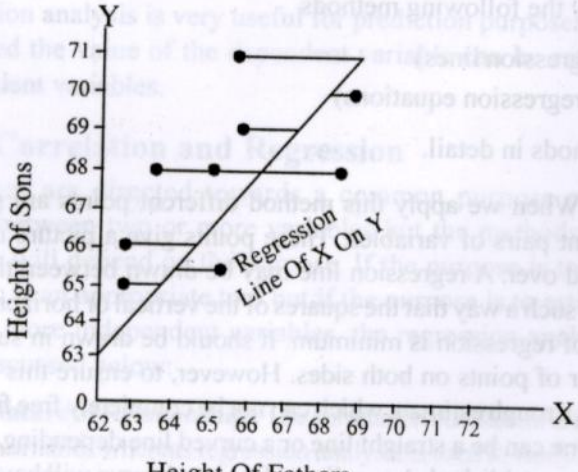


Fig. 2

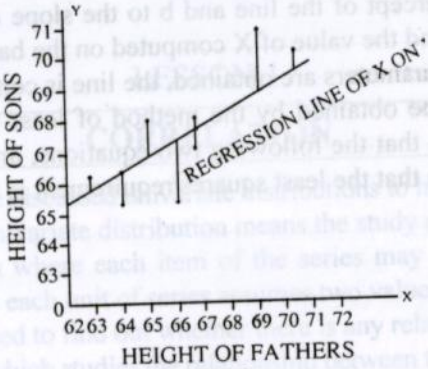


Fig. 3

It is clear that the position of the regression line of x on y is not exactly like that of the regression line of y on x. In the following figure both the regression of y on x and x on y are exhibited.

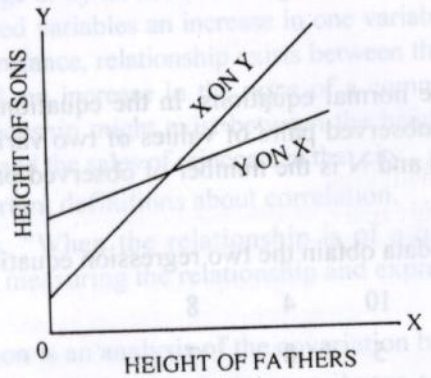


Fig. 4

When there is either perfect positive or perfect negative correlation between the two variables, the two regression lines will coincide and we will have only one line. The farther the two regression lines from each other, the lesser is the degree of correlation and vice-versa. If the variables are independent, correlation is zero and the lines of regression will be at right angles. It should be noted that the regression lines cut each other at the point of average of x and y, i.e., if from the point where both the regression lines cut each other a perpendicular is drawn on the x–axis, we will get the mean value of x series and if from that point a horizontal line is drawn on the y–axis we will get the mean of y series.

2. **Algebraic Method:** The algebraic method for simple linear regression can be understood by two methods :

(i) Regression Equations.

(ii) Regression Coefficients.

Regression Equations: These equations are known as estimating equations. Regression equations are algebraic expressions of the regression lines. As there are two regression lines, there are two regression equations :

(i) x on y is used to describe the variations in the values of x for given changes in y.

(ii) y on x is used to describe the variations in the values of y for given changes in x.

The regression equations of y on x is expressed as

$$Y_c = a + bx$$

The regression equations of x on y is expressed as

$$X_c = a + bx$$

In these equations a and b are constants which determine the position of the line completely. These constants are called the parameters of the line. If the value of any of these parameters is changed, another line is determined.

Parameter a refers to the intercept of the line and b to the slope of the line. The symbol Y_c and X_c refers to the values of Y computed and the value of X computed on the basis of independent variable in both the cases. If the values of both the parameters are obtained, the line is completely determined. The values of these two parameters a and b can be obtained by the method of least squares. With a little algebra and differential calculus it can be shown that the following two equations, are solved simultaneously, will give values of the parameters a and b such that the least squares requirement is fulfilled ;

For regression equation $y_c = a + bx$

$$\Sigma y = Na + b\Sigma x$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$

For regression equation $x_c = a + by$

$$\Sigma x = Na + b\Sigma y$$

$$\Sigma xy = a\Sigma y + b\Sigma y^2$$

These equations are called the normal equations. In the equations Σx, Σy, Σxy, Σx^2, Σy^2 indicate totals which are computed from the observed pairs of values of two variables x and y to which the least squares estimating line is to be fitted and N is the number of observed pairs of values. Let us understand with the help of an example.

Example: From the following data obtain the two regression equations :

x	:	6	2	10	4	8
y	:	9	11	5	8	7

Solution : ***Computation of 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
$\Sigma x = 30$	$\Sigma y = 40$	$\Sigma xy = 214$	$\Sigma x^2 = 220$	$\Sigma x^2 = 340$

Regression line of Y on X is expressed by the equation of the form

$$Y_c = a + bx$$

To determine the values of a and b, the following two normal equations are solved

$$\Sigma y = Na + b\Sigma x$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$

Substituting the valu , we get

$$40 = 5a + 30b \quad \text{.....(i)}$$

$$214 = 30a + 220b \quad \text{.....(ii)}$$

Multiplying equation (i) by 6, we get

$$240 = 30a + 180b \quad \text{.....(iii)}$$

$$214 = 30a + 220b \quad \text{.....(iv)}$$

Deduct equation (iv) from (iii)

$$-40b = +26$$

$$\therefore \quad b = -0.65$$

Substitute the value of b in equation (i)

$$40 = 5a + 30(-0.65)$$

$$5a = 40 + 19.5 \text{ or } a = 11.9$$

Substitute the values of a and b in the equation

Regression line of Y on X is

$$Y_c = 11.9 - 0.65x$$

Regression line of X on Y is

$$X_c = a + by$$

The corresponding normal equations are

$$\Sigma x = Na + b\Sigma y$$

$$\Sigma xy = a\Sigma y + b\Sigma y^2$$

Substituting the values

$$30 = 5a + 40b \quad \text{.....(i)}$$

$$214 = 40a + 340b \quad \text{.....(ii)}$$

Multiply equation (i) by 8

$$240 = 40a + 320b \quad \text{.....(iii)}$$

$$214 = 40a + 340b \quad \text{.....(iv)}$$

Deduct equation (iv) from (iii)

$$-20b = 26 \text{ or } b = -1.3$$

Substitute the value of b in equation (i)

$$30 = 5a + 40(-1.3)$$

$$5a = 30 + 52 \text{ or } a = 16.4$$

Substitute the values of a and b in the equation. Regression line of X on Y is

$$X_c = 16.4 - 1.3y$$

Regression Coefficient: In the regression equation b is the regression coefficient which indicates the degree and direction of change in the dependent variable with respect to a change in the independent variable. In the two regression equations :

$$X_c = a + bxy$$

$$Y_c = a + byx$$

Where bxy and byx are known as the regression coefficients of the two equations. These coefficients can be obtained independently without using simultaneous normal equations with these formulae :

Regression coefficients of x on y is

$$bxy = r\frac{\sigma_x}{\sigma_y}$$

$$bxy = \frac{\Sigma xy}{N\sigma_x\sigma_y} \times \frac{\sigma_x}{\sigma_y} = \frac{\Sigma xy}{N\sigma_y^{\,2}}$$

$$bxy = \frac{\Sigma\, xy}{\Sigma\, y^2} \qquad \text{where} \quad x = X - \overline{X} \quad \text{and} \quad y = Y - \overline{Y}$$

Regression Coefficient of Y on X is

$$byx = r\frac{\sigma_y}{\sigma_x}$$

$$byx = \frac{\Sigma xy}{N\sigma_x\sigma_y} \times \frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{N\sigma_x^{\,2}}$$

$$byx = \frac{\Sigma xy}{\Sigma x^2} \qquad \text{where } x = X - \overline{X} \quad \text{and } y = Y - \overline{Y}$$

Example : Calculate the regression coefficients from the data given below:

	Series x	*Series y*	
Average	25	22	
Standard deviation	4	5	$r = 0.8$

Solution : The coefficient of regression of y on x is

$$bxy = r\frac{\sigma_x}{\sigma_y} = 0.8 \times \frac{4}{5} = +0.64$$

The coefficient of regression of y on x is

$$byx = r\frac{\sigma_y}{\sigma_x} = 0.8 \times \frac{5}{4} = 1.00$$

Example : Calculate the following from the below given data:

(a) the two regression equations,

(b) the coefficient of correlation and

(c) the most likely marks in Statistics when the marks in Economics are 30

Marks in Economics :	25	28	35	32	31	36	29	38	34	32
Marks in Statistics :	43	46	49	41	36	32	31	30	33	39

Solution : Calculation of Regression Equations and Correlation Coefficient

Marks in Eco (X)	$(X-\bar{X})$ x	x^2	Marks in Stats (Y)	$(Y-\bar{Y})$ y	y^2	xy
25	– 7	49	43	+ 5	25	– 35
28	– 4	16	46	+ 8	64	– 32
35	+ 3	9	49	+ 11	121	+ 33
32	0	0	41	+ 3	9	0
31	– 1	1	36	– 2	4	+ 2
36	+ 4	16	32	– 6	36	– 24
29	– 3	9	31	– 7	49	+ 21
38	+ 6	36	30	– 8	64	– 48
34	+ 2	4	33	– 5	25	– 10
32	0	0	39	+ 1	1	0
$\Sigma X = 320$	$\Sigma = 0$	$\Sigma x^2 = 140$	$\Sigma Y = 380$	$\Sigma y = 0$	$\Sigma y^2 = 398$	$\Sigma xy = -92$

Regression equation X on Y

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$

$$byx = \frac{\Sigma xy}{\Sigma y^2} = \frac{-93}{398} = -0.234$$

$$\bar{X} = \frac{\Sigma X}{N} = \frac{320}{10} = 32 \quad \text{and} \quad \bar{Y} = \frac{\Sigma Y}{N} = \frac{380}{10} = 38$$

Substituting the values

$$X - 32 = -0.234\,(Y - 38)$$

$$X - 32 = -0.234Y + 8.892$$

or $$X = 40.892 - 0.234Y$$

Regression equation Y on X

$$(Y - \bar{Y}) = byx\,(X - \bar{X})$$

$$bxy = \frac{\Sigma xy}{\Sigma x^2} = \frac{-93}{140} = -0.664$$

$\bar{X} = 32, \quad \bar{Y} = 38, \quad b = -0.664$

$\therefore$ $$Y - 38 = -0.664\,(X - 32)$$

$$= -0.664Y + 21.248$$

or $$Y = 59.248 - 0.664X$$

(b) Correlation Coeficient $(r) = \sqrt{bxy \times byx} = \sqrt{-0.234 \times -0.664} = -0.394$

Since both the regression coefficients are negative, value of r must also be negative.

(c) Likely marks in statistics when marks in Econimics are 30.

$$Y = -0.664\,X + 59.248 \text{ where } X = 30$$

$$Y = (-0.664 \times 30) + 59.248 = 39.328 \text{ or } 39$$

Example: The following scores were worked out from a test in Mathematics and English in an annual examnation.

	Scores in Mathematics (x)	English (y)	
Mean	39.5	47.5	
Standard deviation	10.8	16.8	r = + 0.42

Find both the regression equations. Using these regression estimate find the value of Y for X =50 and the value of X for Y = 30.

Solution : Regression equation of X on Y

$$X - \overline{X} = r\frac{\sigma_x}{\sigma_y}(Y - \overline{Y})$$

where $\overline{Y} = 47.5$, $\overline{X} = 39.5$, $r = 0.42$, $\sigma_x = 10.8$ and $\sigma_y = 16.8$

By substituting values, we get

$$X - 39.5 = 0.42\frac{10.8}{16.8}(Y - 47.5)$$

$$= 0.27\,(Y - 47.5) = 0.27\,Y - 12.82$$

or $$X = 0.27Y - 12.82 + 39.5 = 0.27Y + 26.68$$

when $Y = 30$

Value of $X = (0.27 \times 30 + 26.68) = 34.78$

Regression equation of Y on X

$$Y - \overline{Y} = r\frac{\sigma_y}{\sigma_x}(X - \overline{X})$$

where $X = 39.5$, $\overline{Y} = 47.52$, $r = 0.42$, $\sigma_x = 10.8$ and $\sigma_y = 16.8$

$$Y - 47.5 = 0.42\frac{16.8}{10.8}(X - 39.5)$$

$$Y - 47.5 = 0.653\,(X - 39.5) = 0.653\;X - 25.79$$

or $$Y = 0.653\,X - 25.79 + 47.5 = 0.653X + 21.71$$

When $X = 50$

Value of $Y = (0.653 \times 50 + 21.71) = 32.65 + 21.71 = 54.36$

Thus the regression equations are

$$X_c = 0.27y + 26.68$$

$$Y_c = 0.653x + 21.71$$

Value of X when Y = 30 is 34.78

Value of Y when X = 50 is 54.36

When actual mean of both the variables X and Y come out to be in fractions, the deviation from actual means create a problem and it is advisable to take deviations from the assumed mean. Thus when devitations are taken from assumed means, the value of bxy and byx is given by

$$bxy = \frac{\Sigma dxdy - \frac{(\Sigma dx) \times (\Sigma dy)}{N}}{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}} \quad \text{where} \quad dx = (X - A) \text{ and } dy = (Y - A)$$

The regression equation is :

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$

Similarly the regression equation of Y on X is

$$Y - \bar{Y} = byx\,(X - \bar{X})$$

$$byx = \frac{\Sigma dxdy - \frac{(\Sigma dx) \times (\Sigma dy)}{N}}{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}}$$

Example. You are given the data relating to purchases and sales. Compute the two regression equations by method of least squares and estimate the likely sales when the purchases are 100.

Purchases :	62	72	98	76	81	56	76	92	88	49
Sales :	112	124	131	117	132	96	120	136	97	85

Solution : ***Calculations of Regression Equations***

Purchases	(X – 76)		Sales	(Y – 120)		
X	dx	dx^2	Y	dy	dy^2	dxy
62	–14	196	112	– 8	64	112
72	–4	16	124	+4	16	–16
98	+22	484	131	+11	121	+242
76	0	0	117	–3	9	0
81	+5	25	132	+12	144	+60
56	–20	400	96	–24	576	+480
76	0	0	120	0	0	0
92	+16	256	136	+16	256	+256
88	+12	144	97	–23	529	–276
49	–27	729	85	–35	1225	+945
	$\Sigma dx = -10$	$\Sigma dx^2 = 2250$		$\Sigma dy = -50$	$\Sigma dy^2 = 2940$	$\Sigma dxdy = 1803$

$$\bar{X} = A + \frac{\Sigma dx}{N} = 76 - \frac{10}{10} = 75 \quad \text{and} \quad \bar{Y} = A + \frac{\Sigma dy}{N} = 120 - \frac{50}{10} = 115$$

Regression Cefficients : X on Y

$$bxy = \frac{\Sigma dxdy - \frac{(\Sigma dx) \times (\Sigma dy)}{N}}{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}} = \frac{1803 - \frac{(-10) \times (-50)}{10}}{2940 - \frac{(-50)^2}{10}} = \frac{1753}{2690} = 0.652$$

Y on X

$$byx = \frac{\Sigma dxdy - \frac{(\Sigma dx) \times (\Sigma dy)}{N}}{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}} = \frac{1803 - \frac{(-10) \times (-50)}{10}}{2250 - \frac{(-10)^2}{10}} = \frac{1753}{2240} = 0.78$$

Regression equation : X on Y

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$

Substituting the values

$$X - 75 = 0.652\,(Y - 115) = 0.652Y - 74.98$$

or $X = 0.652Y + 0.02$

when $X = 100$

$$Y = 0.78 \times 100 + 56.5 = 134.5$$

Regression equation : Y on X

$$Y - \bar{Y} = \quad (X - \bar{X})$$

$$Y - 115 = 0.78\,(X - 75) = 0.78\,X - 58.5$$

$$Y = 0.78X + 56.5$$

UNIT - III

LESSON

INDEX NUMBERS

Economic activities have constant tendency to change. Prices of commodities which are the total result of number of economic activities also have a tendency to fluctuate. The problem of change in prices is very important. But it is not very simple to study this problem and derive conclusions because price of different commodities change by different degrees. Hence, there is a great need for a device which can smoothen the irregularities in the prices to obtain a conclusion. This need is satisfied by Index Numbers which makes use of percentages and average for achieving the desired objective. Index Number is a device for comparing the general level of the magnitude of a group of distinct but related variables in two or more situations. Index Numbers are used to feel the pulse of the economy and they reveal the inflationary or deflationary tendencies. In reality, Index Numbers are described as barometers of economic activity because if one wants to have an idea as to what is happening in an economy, he should check the important indicates like the index numbers of industrial production, agricultural production, business activity etc.

The various definitions of Index Numbers are discussed under three heads:

(i) Measure of change
(ii) Device to measure change
(iii) A series representing the process of change.

According to *Maslow,* it is a numerical value characterising the change in complex economic phenomenon over a period of time.

Spiegal explains an index number is a statistical measure designed to show changes in variable or a group of related variables with respect to time, geographical location or other characteristics.

Gregory and Ward describes it as a measure over time designed to show average change in the price, quantity or value of a group of items.

Croxton and Cowden says Index numbers are devices for measuring differences in the magnitude of a group of related variables.

B.L. Bowley describes Index Numbers as a series which reflects in its trend and fluctuations the movements of some quantity to which it is related.

Blair puts Index Numbers are specialised kinds of an average.

Index Numbers have the following features :

(i) Index numbers are specialised averages which are capable of being expressed in percentage.
(ii) Index numbers measure the changes in the level of a given phenomenon.
(iii) Index numbers measure the effect of changes over a period of time.

Index Numbers are indispensable tools of economic and business analysis. Their significance can be appreciated by following points :

1. Index number helps in measuring relative changes in a set of items.
2. Index numbers provide a good basis of comparison because they are expressed in abstract unit distinct from the unit of element.
3. Index numbers help in framing suitable policies for business and economic activities.
4. Index numbers help in measuring the general trend of the phenomenon.
5. Index numbers are used in deflating. They are used to adjust the original data for price changes or to adjust wages for cost of living changes.

6. The utility of index numbers has increased a great deal because of the method of splicing whereby the index prepared on any base can be adjusted with reference to any other base.
7. As a measure of average change in a group of elements the index numbers can be used for forecasting future events. Whereas a trend line gives an average rate of change in a single phenomenon, it indicates the trend for a group of commodities.
8. It is helpful in a study of comparative purchasing power of money in different countries of the world.
9. Index numbers of business activities throw light on the economic progress made by various countries.

Problems in the Construction of Index Numbers

While constructing Index Number, the following problems arise :

1. *The purpose of Index:* Before constructing an Index Number, it is necessary to define precisely the purpose for which they are to be constructed. A single Index can not fulfil all the purposes. Index Numbers are specialised tools which are more efficient and useful when properly used. If the purpose is not clear, the data used may be unsuitable and the indices obtained may be misleading. If it is desired to construct a Cost of Living Index of labour class, then only those item will be included which are required by the labour class.

2. *Selection of the items:* The list of commodities included in the Index numbers is called the 'Regimen'. Because it may not be possible to include all the items, it becomes necessary to decide what items are to be included. Only those items should be selected which are representative of the data, e.g. in a consumer Price Index for working class, items like scooters, cars, refrigerators, cosmetics, etc. find no place. There is no hard and fast rule regarding the inclusion of number of commodities while constructing Index Numbers. The number of commodities should be such as to permit the influence of the inertia of large numbers. At the same time the numbers should not be so large as to make the work of computation uneconomical and even difficult. The number of commodities should therefore be reasonable. The following points should be considered while selecting the items to be included in the Index :

(i) The items should be representative.
(ii) The items should be of a standard quality.
(iii) Non-tangible items should be excluded.
(iv) The items should be reasonable in number.

3. *Price Quotations:* It is neither possible nor necessary to collect prices of the commodities from all markets in the country where it is dealt with, we should take a sample of the markets. Selection must be made of the representative places and persons. These places should be well known for trading these commodities.

It is necessary to select a reliable agency from where price quotations are obtained.

4. *Selection of the Base period:* In the construction of Index Numbers, the selection of the base period is very important step since the base period serves as a reference period and the prices for a given period are expressed as percentages of those for the base year, it is therefore necessary that

(i) the base period should be normal and
(ii) it should not be too far in the past.

There are two methods by which base period can be selected (i) Fixed base method and (ii) Chain base method.

Fixed base Method: According to this any year is taken as a base. Prices during the year are taken equal to 100 and the prices of other years are shown as percentages of those prices of the base year. Thus indices for 1998, 99,2000, and 2001 are calculated with 1997 as base year, such indices will be called as fixed base indices.

Chain base Method: According to this method, relatives of each year are calculated on the basis of the prices of the preceding year. The Chain base Index Numbers are called as Link Relatives e.g., if index numbers are constructed for 1997, 98, 99, 2000 and 2001 then for 1998, 1997 will be the base and for 1999, 1998, will be the base and so on.

5. *The choice of an average*: An Index number is a technique of 'averaging' all the changes in the group of series over a period of time, the main problem is to select an average which may be able to summaries the change in the component series adequately. Median, Mode and Harmonic Mean are never used in the construction of index numbers. A choice has to be made between the Arithmetic Mean and the Geometric Mean. Merits and demerits of the two are then to be compared. Theoretically A.M. is superior to the A.M. in many respects but due to difficulty in its computation, it is not widely used for this purpose.

6. *Selection of appropriate weights*: The term weight refers to the relative importance of the different items in the construction of index numbers. All items are not of equal importance and hence it is necessary to find out some suitable methods by which the varying importance of the different items is taken into account. The system of weighting depends upon the purpose of index numbers, but they ought to reflect the relative importance of the commodities in the regimen. The system may be either arbitrary or rational. The weightage may be according to:

(1) the value of quantity produced, or
(2) the value of quantity consumed, or
(3) the value or quantity sold or put to sale.

There are two methods of assigning weights.

(i) Implicit and
(ii) Explicit.

Implicit: Under this method, the commodity to which greater importance has to be given is repeated a number of times i.e., a number of varieties of such commodities are included in the index numbers as separate items.

Explicit: In this case, the weights are explicitly assigned to commodities. Only one kind of a commodity is included in the construction of Index Numbers but its price relative is multiplied by the figure of weights assigned to it. There has to be some logic in assigning such type of weights.

Methods of Constructing Index Numbers

The index number for this purpose is divided into two heads :

(1) Unweighted Indices; and
(2) Weighted Indices.

Each one of these is further sub-divided under two categories :

(i) Simple aggregative ; and
(ii) Average of price relatives.

Unweighted Index Numbers

(i) *Simple aggregative method:* Under this method the total of the current year prices for various commodities is divided by the total of the base year prices and the quotient is multiplied by 100.

Symbolically,

$$P_{01} = \frac{\Sigma P_1}{\Sigma P_0} \times 100$$

where P_{01} represents the Price Index, P_1 represents prices of current year and Po prices of base year.

Illustration: From the following data construct the index for 2006 taking 2000 at base year.

Commodity	*Prices in* 2000 (*Rs.*)	*Prices in* 2006 (*Rs.*)
A	30	30
B	35	50
C	45	75
D	45	70
E	25	40

Solution: Construction of Price Index.

Commodity	*Prices in* 2000 (*Rs.*)	*Prices in* 2006 (*Rs.*)
A	30	30
B	35	50
C	45	75
D	45	70
E	25	40
	$\Sigma p_0 = 180$	$\Sigma p_1 = 265$

$$\text{Price Index for 2006 with 200 as base} = \frac{\text{sum of prices in } 2006}{\text{sum of prices in } 2000} \times 100$$

Symbolically

$$P_{01} = \frac{\Sigma P_1}{\Sigma P_0} \times 100 = \frac{265}{180} \times 100 = 147.2$$

Hence there is an increase of 47.2% in prices of commodities during the year 2006 as compared to 2000.

(*ii*) *Average of Price Relative Method:* Under this method, first calculate the price relatives for the various items included in the index and then average the price relatives by using any of the measures of the central value, i.e. A.M.; the median; the mode; the Geometric mean or the Harmonic mean.

(*i*) When arithmetic mean is used

$$P_{01} = \frac{\Sigma\left\{\frac{P_1}{P_0} \times 100\right\}}{N}$$

(*ii*) When geometric mean is used

$$P_{01} = AL\left\{\frac{\left\{\Sigma \log\left(\frac{P_1}{P_0} \times 100\right)\right\}}{N}\right\}$$

where N refers to the number of items whose price relatives are averaged.

Illustration: Calculate Index Numbers for 2005, 2006 and 2007 taking 2004 as base from the following data by average of relatives method.

Commodity	*2004*	*2005*	*2006*	*2007*
A	2	5	4	3
B	8	11	13	6
C	4	5	6	8
D	6	4	5	7
E	5	4	6	3

Solution : **Construction of Index Numbers based on Mean of Relatives.**

Commodity	2004		2005		2006		2007	
	p_0		p_1	$\frac{P_2}{P_0} \times 100$	p_2	$\frac{P_3}{P_0} \times 100$	p_3	
A	2	100	5	250.0	4	200.0	3	150.0
B	8	100	11	137.5	13	162.5	6	75.0
C	4	100	5	125.0	6	150.0	8	200.0
D	6	100	4	66.7	5	83.3	7	116.7
E	5	100	4	80.0	6	120.0	3	60.0
		500		659.2		715.8		601.7

P_{01} = Index with 2004 as base and 2005 as current year

$$P_{01} = \frac{\Sigma\left\{\frac{P_1}{P_0} \times 100\right\}}{N} = \frac{659.2}{5} = 131.83$$

P_{02} = Index with 2004 as base and 2006 as current year

$$P_{02} = \frac{\Sigma\left\{\frac{P_2}{P_0} \times 100\right\}}{N} = \frac{715.8}{5} = 143.16$$

P_{03} = Price Index with 2004 as base arid 2007 as current year

$$P_{03} = \frac{\Sigma\left\{\frac{P_3}{P_0} \times 100\right\}}{N} = \frac{601.7}{5} = 120.33$$

2.Weighted Index Numbers

(i) *Aggregative Method:* These indices are of the simple aggregative type with the only difference that the weights are assigned to the various items included in the index. This method in fact can be described as an extension of the simple aggregative method in the sense that the weights are assigned to the different commodities included in the index. There are various methods by which weights can be assigned and hence a large number of formulae for constructing Index Numbers have been devised. Some commonly used methods suggested by different authorities are as follows :

(i) Laspeyre's method.
(ii) Paasche's method.
(iii) Fisher's ideal method.
(iv) Marshall Edgeworth method.
(v) Kelly's method.
(vi) Dorbish and Bowley's method.

(i) *Laspeyre's Method.*

Laspeyre suggested that for the purposes of calculating Price Indices, the quantities in the base year should be used as weights. Hence the formula for computing price Index number would be :

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

where P_{01} refers to Price Index,

p refers to price of each commodity,

q refers to quantity of each commodity,

o base year,

1 current year, and

Σ refers to the summation of items.

The steps for calculating Index Numbers are :

(a) Multiply the price of each commodity for current year with its respective Quantity for the base year ($P_1 \times q_0$) and then find out the total of this product $\Sigma (P_1 q_0)$.

(b) Multiply the price of each commodity for the base year with the respective quantity for the base year ($P_0 \times q_0$) and then find out the total of these products for different commodities $\Sigma (P_0 q_0)$.

(c) Divide $\Sigma (P_1 q_0)$ with $\Sigma (P_0 q_0)$ and multiply the quotient by 100. On the other hand, if Quantity Index is to be calculated, the prices of base year will be used as weights. Symbolically,

$$Q_{01} = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times 100$$

Illustration. Compute Price Index and Quantity Index from data given below by Lespeyre's method.

Items	Base year		Current year	
	Quantity	Price	Quantity	Price
A	6 units	40 paise	7 units	30 paise
B	4 units	45 paise	5 units	50 paise
C	5 units	90 paise	1.5 units	40 paise

Solution: Computation of Price and Quantity Indices.

Items	Base year q_0	Base year p_0	Current year q_1	Current year p_1	p_0q_0	p_1q_0	p_0q_1	p_1q_1
A	6	40	7	30	240	180	280	210
B	4	45	5	50	180	200	225	250
C	5	90	1.5	40	45	20	135	60
					$\Sigma p_0q_0 = 465$	$\Sigma p_1q_0 = 400$	$\Sigma p_0q_1 = 640$	$\Sigma p_1q_1 = 520$

$$\text{Price Index } (P_{01}) = \frac{\Sigma p_1q_0}{\Sigma p_0q_0} \times 100 = \frac{400}{465} \times 100 = 86.00$$

$$\text{Quantity Index } (Q_{01}) = \frac{\Sigma q_1p_0}{\Sigma q_0p_0} \times 100 = \frac{640}{665} \times 100 = 137.63$$

(*ii*) *Paasches Method:* Under this method of calculating Price Index the quantities of the current year are used as weights.

$$\text{Symbolically Price Index or } P_{01} = \frac{\Sigma p_1q_1}{\Sigma p_0q_1} \times 100$$

Steps of construction Index according to Paasche's method are :

(i) Calculate the product of the current year prices of different commodities and their respective quantities for the current year ($p_1 \times q_1$) and find out the total of the product of different commodities $\Sigma(p_1 \times q_1)$.

(ii) Calculate the product of p_0 and q_1 of different commodities and aggregate them $\Sigma(p_0q_1)$.

(iii) Divide $\Sigma(p_1 q_1)$ with $\Sigma(p_0q_1)$ and multiply the quotient by 100 to obtain Price Index. Similarly, quantity index is calculated using the current year price as weights. Symbolically,

$$Q_{01} = \frac{\Sigma q_1p_1}{\Sigma q_0p_1} \times 100$$

Illustration: From the data of previous illustration, calculate (i) Price Index (ii) Quantity Index by Paasche's method.

Items	Base year q_0	Base year p_0	Current year q_1	Current year p_1	p_0q_0	p_1q_0	p_0q_1	p_1q_1
A	6	40	7	30	240	180	280	210
B	4	45	5	50	180	200	225	250
C	5	90	1.5	40	45	20	135	60
Total					465	400	640	520

$$\text{Price Index } P_{01} = \frac{\Sigma p_1q_1}{\Sigma p_0q_1} \times 100 = \frac{520}{640} \times 100 = 81.5$$

$$\text{Quantity Index } Q_{01} = \frac{\Sigma q_1p_1}{\Sigma q_0p_1} \times 100 = \frac{520}{400} \times 100 = 130$$

(*iii*) *Fisher's Ideal Index:* Laspeyre has used base year quantities as weights whereas Paasche's has used Current year quantities as weights for the computation of Index Number of prices. Fisher suggested that both the current year quantities and the base year quantities should be used but geometric mean of the two be calculated and that figure should be the Index Number. Symbolically,

$$\text{FIsher's Price Index } P_{01} = \sqrt{\left(\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100\right)\left(\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100\right)} = \sqrt{\left(\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}\right) \times \left(\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}\right)} \times 100$$

$$\text{Fisher's Index} = \sqrt{\text{Laspeyre's Index} \times \text{Paasches's Index}}$$

On the other hand if quantity Indices by this method are to be calculated the geometric mean of the Index Number of quantities with base year prices as weights and Index Number of Quantities with current year as weights be found out. Symbolically,

$$\text{Fisher's Quantity Index } Q_{01} = \sqrt{\left(\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}\right) \times \left(\frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}\right)} \times 100$$

Illustration. Construct Index Number of Prices and Quantities from the following data using Fisher's method (2000 = 100).

Commodity	2000 Price	2000 Qty.	2007 Price	2007 Qty.
A	2	8	4	6
B	5	10	6	5
C	4	14	5	10
D	2	19	2	13

Solution: **Calculation of Price and Production Indices.**

Items	2000 Price (p_o)	2000 Qty.(q_o)	2007 Price (p_1)	2007 Qty.(q_1)	p_oq_o	p_1q_1	p_1q_o	p_oq_1
A	2	8	4	6	16	24	32	12
B	5	10	6	5	50	30	60	25
C	4	14	5	10	56	50	70	40
D	2	19	2	13	38	26	38	26
Total					160	130	200	103

$$P_{01} = \sqrt{\left(\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}\right) \times \left(\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}\right)} \times 100 = \sqrt{\frac{200}{160} \times \frac{130}{103}} \times 100 = 125.6$$

$$Q_{01} = \sqrt{\left(\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}\right) \times \left(\frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}\right)} \times 100 = \sqrt{\frac{103}{160} \times \frac{130}{200}} \times 100 = 64.7$$

(iv) Marshall & Edgeworth's Method: In this method also both current year as well as base year prices and quantities are considered. The formula is as follows :

$$P_{01} = \frac{\Sigma(q_0 + p_1)p_1}{\Sigma(q_0 + q_1)p_0} \times \frac{\Sigma q_0 p_1 + \Sigma q_1 p_1}{\Sigma q_0 p_0 + \Sigma q_1 p_0} \times 100$$

and Quantity Index is calculated by the formula

$$Q_{01} = \frac{\Sigma(p_0 + p_1)q_1}{\Sigma(p_0 + p_1)q_0} \times 100 = \frac{\Sigma p_0 q_1 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_1 q_0} \times 100$$

(v) Kelly's Method: Truman Kelly has suggested the following formula for constructing Index Number.

$$P_{01} = \frac{\Sigma p_1 q}{\Sigma p_0 q} \times 100 \qquad \text{where} \quad q = \frac{q_0 + q_1}{2}$$

where q refers to the average quantity of two periods. This is also known as fixed aggregative method.

(vi) Dorbish & Bowley's Method: Dorbish & Bowley have suggested the simple arithmetic mean of Lespeyre's and Paasche's formula. Symbolically.

$$P_{01} = \frac{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} + \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}{2} \times 100$$

(II) *Weighted Average of Price Relatives* :

This method is also known as the Family Budget Method. Weights are values ($p_0 q_0$) of the base year in this method. The Index Number for the current year is calculated by dividing the sum of the products of the current year's price relatives and base year values by the total of the weights, i.e., the weighted arithmetic average of the price relatives gives the required index numbers. Symbolically,

$$\text{Weighted Index number of the current year} = \frac{\Sigma IV}{\Sigma V}$$

where I stands for Price Relatives of the current year and V stands for the values of the base year.

Illustration: From the data given below, calculate the Weighted Index Number by using weighted average of Relatives.

Commodities	Units	Base Yr. Qty.	Base Year's Price	Current Yr. Price
A	Quintal	7	16	19.6
B	Kg.	6	2	3.2
C	Dozen	16	5.6	7.0
D	Metre	21	1.5	1.4

Solution :

$$\text{The Price relative of the current year} = \frac{\text{Current year's Price}}{\text{Base Year's Price}} \times 100$$

The value of the base year = Quantity of base year × Price of the base year

Commodities	Price Relatives $I = (\frac{p_1}{p_0} \times 100)$	Value of Weights i.e. $V = p_0q_0$	Weights x Price Relatives $V \times I$
A	122.5	112.0	13,720
B	160.0	12.0	1,920
C	125	89.6	11,200
D	93.3	31.5	2,939
		$\Sigma V = 245.1$	$\Sigma IV = 29,779$

Weighted Index Number of the Current year $= \frac{\Sigma IV}{\Sigma V} = \frac{29779}{245} = 121.5$ Ans.

In weighted average of relatives, the Geometric mean may be used instead of arithmetic mean. The weighted geometric mean of relatives is calculated by applying logarithms to the relatives. When this mean is used, then formula is:

$$P_{01} = \text{Antilog}\left\{\frac{\Sigma\ V.\log I}{\Sigma\ V}\right\} \text{ where } I = \frac{p_1}{p_0} \times 100 \text{ and } V = p_0q_0$$

Illustration: Find out price index by weighted average of price relatives from the following commodities using geometric mean :

Commodities	P_0	q_0	P_1
X	3.0	20	4.0
Y	1.5	40	1.6
Z	1.0	10	1.5

Solution : **Calculation of Index Number**

Commodities	P_0	q_0	P_1	$(p_0 q_0)$ V	$\left(\frac{p_1}{p_0} \times 100\right)$ I	Log I	V. log I
X	3.0	20	4.0	60	$133.33 \left(\frac{4}{3} \times 100\right)$	2.1249	127.494
Y	1.5	40	1.6	60	$106.7 \left(\frac{1.6}{1.5} \times 100\right)$	2.0282	121.692
Z	1.0	10	1.5	10 =	$150.0 \left(\frac{1.5}{1.0} \times 100\right)$	2.1761	21.761
				$\Sigma V = 130$			$\Sigma V \log I = 270.947$

By applying the formula :

$$P_{01} = AL\left\{\frac{\Sigma V.\log I}{\Sigma V}\right\} = AL\left(\frac{270.947}{130}\right) = AL\,2.084 = 121.3$$

Tests of Adequacy of Index Numbers

Since several formulas have been suggested for the construction of index numbers, then the question arises which method of index number is the most suitable in a given situation. These are some tests to choose an appropriate index :

(*i*) *Unit Test:* It requires that the method of constructing index should be independent of the units of the problem. All the methods except simple aggregative method satisfy this test.

(*ii*) *Circular Test:* This test was suggested by Westerguard and C. M. Walsch. It is based on the shift ability of the base. Accordingly, the index should work in a circular fashion i.e., if an index number is computed for the period 2 on the base period 1, and still another index number is computed for period 3 on the base period 2. Then the product should be equal to one.

$$P_{01} \times P_{12} \times P_{23} \quad \times P\ no = 1$$

Only simple aggregative and fixed weight aggregative method satisfy the test.

If the test is applied to simple aggregative method, we will get

$$\frac{\Sigma p_1}{\Sigma p_0} \times \frac{\Sigma p_2}{\Sigma p_1} \times \frac{\Sigma p_3}{\Sigma p_2} = 1$$

The test is met by simple geometric mean of price relatives and the weighted aggregative of fixed weights.

(*iii*) *Time Reversal Test:* According to Prof. Fisher the formula for calculating an index number should be such that it gives the same ratio between one point of time and the other, no matter which of the two time is taken as the base. In other words, when the data for any two years are treated by the same method, but with the base reversed, the two index numbers should be reciprocals of each other.

$P_{01} \times P_{10} = 1$ (omitting the factor 100 from each index).

Where P_{01} denotes the index for current year 1 based on the base year 0 and P_{10} is for current year 0 on the base year 1.

It can be easily verified that simple geometric mean of price relatives index, weighted aggregative formula, weighted geometric mean of relatives and Marshall Edgeworth and Fisher's ideal method satisfies the test.

Let us see how Fishers ideal method satisfies the test.

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}$$

By changing time from 0 to 1 and 1 to 0

$$P_{01} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}}$$

Now $P_{01} \times P_{10} = 1$

Substitute the value of P_{01} and P_{10}

$$P_{01} \times P_{10} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}} = 1$$

(iv) Factor Reversal Test: It says that the product of a price index and the quantity index should be equal to value index. In the words of Fisher, just as each formula should permit the interchange of the two times without giving inconsistent results similarly it should permit interchanging the prices and quantities without giving inconsistent results which means two results multiplied together should give the true value ratio. The test says that the change in price multiplied by change in quantity should be equal to total change in value. If P_{01} is a price index for the current year with reference to base year and Q_{01} is the quantity index for the current year.

$$\text{Then } P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

This test is satisfied only by Fisher's ideal index method.

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}$$

Changing p to q and q to p.

$$Q_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}$$

$$\therefore \quad P_{01} \times Q_{10} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} = \sqrt{\frac{(\Sigma q_1 p_1)^2}{(\Sigma p_0 q_0)^2}} = \frac{(\Sigma p_1 q_1)}{(\Sigma p_0 q_0)}$$

In other words, factor reversal test is based on the following analogy. If the price per unit of a commodity increases from Rs.10 in 2000 to Rs. 15 in 2007, and the quantity of consumption changes from 100 units to 140 units during the same period, the price and quantity in 2007 are 15 and 140 respectively. The values of consumption (p x q) were Rs. 1000 in 2000 and Rs. 2100 in 2007 giving a value ratio.

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0} = \frac{2100}{1000} = 2.1$$

'Thus we find that the product of price ratio and quantity ratio equals the value ratio :

$$1.5 \times 1.4 = 2.1$$

Chain Base Index

The index of a given year on a given fixed base is not affected by changes in the prices or the quantities of any other year. On the other hand, in the chain base method, the value of each period is related with the immediately proceeding period and not with any fixed period. To construct index numbers by chain base method, a series of index numbers are computed for each year with preceding year as the base. These numbers are known as Link relatives. The link relatives when multiplied successively known as the chaining process give link to a common base. The products obtained are expressed as % and give the required index number. The steps of chain base index are :

(i) Express the figures of each period as a % of the preceding period to obtain Link Relatives (LR)

(ii) These link relatives are chained together by successive multiplication to get chain indices by the formula:

$$\text{Chain Base Index (CBI)} = \frac{\text{Current year LR} \times \text{Preceding year Chain Index}}{100}$$

(iii) The chain index can be converted into a fixed base index by this formula :

$$\text{Fixed Base Index (FBI)} = \frac{\text{Current year CBI} \times \text{Previous year FBI}}{100}$$

Chain relatives are computed from link relatives whereas fixed base relatives are computed directly from the original data. The results obtained by fixed base and chain base index invariably are the same.

We shall understand the process by taking some examples.

Illustration: Construct Index Numbers by chain base method from the following data of wholesale prices.

Year:	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Prices:	75	50	65	60	72	70	69	75	84	80

Solution :

Computation of Chain Index

Year	Price	Link Relatives	Chain Base Index	Fixed Base Index
1998	75	100	100	100
1999	50	$\frac{50}{75}\times 100 = 66.67$	$\frac{66.67\times 100}{100} = 66.67$	$\frac{50}{75}\times 100 = 66.67$
2000	65	$\frac{65}{50}\times 100 = 130$	$\frac{130\times 66.67}{100} = 86.67$	$\frac{65}{75}\times 100 = 86.67$
2001	60	$\frac{60}{65}\times 100 = 92.31$	$\frac{92.31\times 86.67}{100} = 80.00$	$\frac{60}{75}\times 100 = 80$
2002	72	$\frac{72}{60}\times 100 = 120$	$\frac{120\times 80}{100} = 96.00$	$\frac{72}{75}\times 100 = 96$

2003	70	$\frac{70}{72}\times 100 = 97.22$	$\frac{97.22\times 96}{100} = 93.33$	$\frac{70}{75}\times 100 = 93.33$
2004	69	$\frac{69}{70}\times 100 = 98.57$	$\frac{98.57\times 93.33}{100} = 92.00$	$\frac{69}{75}\times 100 = 92$
2005	75	$\frac{75}{69}\times 100 = 108.69$	$\frac{108.69\times 92}{100} = 100.00$	$\frac{75}{75}\times 100 = 100$
2006	84	$\frac{84}{75}\times 100 = 112$	$\frac{112\times 100}{100} = 112.00$	$\frac{84}{75}\times 100 = 112$
2007	80	$\frac{80}{84}\times 100 = 95.24$	$\frac{95.24\times 112}{100} = 106.67$	$\frac{80}{75}\times 100 = 106.67$

It may be seen that index by chain base and fixed base method are same.

Illustration: Construct chain index numbers from the link relatives given below:

Year	:	2003	2004	2005	2006	2007
LinkRelatives	:	100	105	95	115	102

Solution : **Calculations for Chain Base Index**

Year	Link Relatives	Chain Index Number
2003	100	100
2004	105	$\frac{105}{100}\times 100 = 105.00$
2005	95	$\frac{95}{100}\times 105 = 99.75$
2006	115	$\frac{115}{100}\times 99.75 = 114.7$
2007	102	$\frac{102}{100}\times 114.75 = 137.64$

Consumer Price Index Numbers

The consumer price index, known as cost of living index, is calculated to know the average change over time in the prices of commodities consumed by the consumers. The need to construct consumer price index arises because the general index numbers fail to give an exact idea of the effect of the change in the general price level on the cost of living of different classes of people, because a given change in the level of prices affect different classes of people in different manners. Different people consume different commodities and if same commodities then in different proportions. The consumer price index helps us in determining the effect of rise and fall in prices on different classes of consumers living in different area. The consumer price index is significant because the demand for higher wages is based on the cost of living index and the wages and salaries in most nations are adjusted according to this index. We should understand that the cost of living index does not measure the actual cost of living nor the fluctuations in the cost of living due to causes other than the change in price level but its object is to find out how much more

the consumers of a particular class have to pay for a certain basket of goods and services. That is why the term cost of living index has been replaced by the term price of living index, cost of living price index or consumer price index.

The significance of studying the consumer price index is that it helps in wage negotiations and wage contracts. It also helps in preparing wage policy, price policy, rent control, taxation and general economic policies. This index is also used to find out the changing purchasing power of different currencies.

Consumer Price Index can be prepared by two methods :

(i) Aggregative Method ;

(ii) Weighted Relatives Method.

When, aggregative method is used to prepare consumer price index, the aggregative expenditure for current year and base year are calculated and we use the formula given below:

$$\text{Consumer Price Index} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

When weighted relatives method is used then the family budgets of a large number of people for whom the index is meant are carefully studied and the aggregative expenditure of an average family on various items is estimated. In other words, the weights are calculated by multiplying the base year quantities and prices ($p_0 q_0$). The price relatives for all the commodities are prepared and multiplied by the weights. By applying the formula, we can calculate Consumer price Index.

$$\text{Consumer Price Index} = \frac{\sum IV}{\sum V} \text{ where } I = \frac{p_1}{p_0} \times 100 \text{ and } V = p_0 q_0$$

Example: Prepare the Consumer price Index for 2007 on the basis of 2000 from the following data by both methods.

Commodities	Quantities Consumed 2000	Prices 2000	Prices 2007
A	6	5.75	6.00
B	6	5.00	8.00
C	1	6.00	9.00
D	6	8.00	10.00
E	4	2.00	1.50
F	1	20.00	15.00

Solution : **Consumer Price Index by Aggregative Method**

Commodities	q_0	p_0	p_1	p_1q_0	p_0q_0
A	6	5.75	6.00	36.00	34.50
B	6	5.00	8.00	48.00	30.00
C	1	6.00	9.00	9.00	6.00
D	6	8.00	10.00	60.00	48.00
E	4	2.00	1.50	6.00	8.00
F	1	20.00	15.00	15.00	20.00
				$\Sigma p_1q_0 = 174$	$\Sigma p_0q_0 = 146.5$

$$\text{Conusumer Price Index} = \frac{\Sigma p_1q_0}{\Sigma p_0q_0} \times 100 = \frac{173}{146.5} 100 = 118.77$$

Consumer Price Index by Weighted Relatives

Commodities	q_0	p_0	p_1	I	V	IV
A	6	5.75	6.00	104.34	34.50	3600
B	6	5.00	8.00	160.00	30.00	4800
C	1	6.00	9.00	150.00	6.00	900
D	6	8.00	10.00	125.00	48.00	6000
E	4	2.00	1.50	75.00	8.00	600
F	1	20.00	15.00	75.00	20.00	1500
					$\Sigma V = 146.5$	$\Sigma IV = 17400$

$$\text{Conusumer Price Index} = \frac{\Sigma IV}{\Sigma V} = \frac{17400}{146.5} = 118.7$$

Limitations of Index Numbers

1. They are only approximate indicators of the relatives level of a phenomenon.
2. Index number are good for achieving one objective, may be unsuitable for the other.
3. Index numbers can be manipulated in a manner as to draw the desired conclusion.

LESSON

ANALYSIS OF TIME SERIES

When quantitative data are arranged in the order of their occurrence, the resulting statistical series is called a time series. The quantitative values are usually recorded over equal time interval daily, weekly, monthly, quarterly, half yearly, yearly, or any other time measure. Monthly statistics of Industrial Production in India, Annual birth-rate figures for the entire world, yield on ordinary shares, weekly wholesale price of rice, daily records of tea sales or census data are some of the examples of time series. Each has a common characteristic of recording magnitudes that vary with passage of time.

Time series are influenced by a variety of forces. Some are continuously effective other make themselves felt at recurring time intervals, and still others are non-recurring or random in nature. Therefore, the first task is to break down the data and study each of these influences in isolation. This is known as decomposition of the time series. It enables us to understand fully the nature of the forces at work. We can then analyse their combined interactions. Such a study is known as time-series analysis.

Components of time series

A time series consists of the following four components:

1. Basic or Secular or Long-time trend;
2. Seasonal variations;
3. Business cycles or cyclical movement; and
4. Erratic or Irregular fluctuations.

These components provide a basis for the explanation of the past behaviour. They help us to predict the future behaviour. The major tendency of each component or constituent is largely due to casual factors. Therefore a brief description of the components and the causal factors associated with each component should be given before proceeding further.

1. Basic or secular or long-time trend: Basic trend underlines the tendency to grow or decline over a period of years. It is the movement that the series would have taken, had there been no seasonal, cyclical or erratic factors. It is the effect of such factors which are more or less constant for a long time or which change very gradually and slowly. Such factors are gradual growth in population, tastes and habits or the effect on industrial output due to improved methods. Increase in production of automobiles and a gradual decrease in production of food grains are examples of increasing and decreasing secular trend.

All basic trends are not of the same nature. Sometimes the predominating tendency will be a constant amount of growth. This type of trend movement takes the form of a straight line when the trend values are plotted on a graph paper. Sometimes the trend will be constant percentage increase or decrease. This type takes the form of a straight line when the trend values are plotted on a semi-logarithmic chart. Other types of trend encountered are "logistic", "S-curyes", etc.

Properly recognising and accurately measuring basic trends is one of the most important problems in time series analysis. Trend values are used as the base from which other three movements are measured. Therefore, any inaccuracy in its measurement may vitiate the entire work. Fortunately, the causal elements controlling trend growth are relatively stable. Trends do not commonly change their nature quickly and without warning. It is therefore reasonable to assume that a representative trend, which has characterized the data for a past period, is prevailing at present, and that it may be projected into the future for a year.

2. Seasonal Variations: The two principal factors liable for seasonal changes are the climate or weather and customs. Since, the growth of all vegetation depends upon temperature and moisture, agricultural activity is

confined largely to warm weather in the temperate zones and to the rainy or post-rainy season in the torried zone (tropical countries or sub-tropical countries like India). Winter and dry season make farming a highly seasonal business. This high irregularity of month to month agricultural production determines largely all harvesting, marketing, canning, preserving, storing, financing, and pricing of farm products, Manufacturers, bankers and merchants who deal with farmers find their business taking on the same seasonal pattern which characterise the agriculture of their area.

The second cause of seasonal variation is custom, education or tradition. Such traditional days as Dewali, Christmas, Id etc., product marked variations in business activity, travel, sales, gifts, finance, accident, and vacationing.

The successful operation of any business requires that its seasonal variations be known, measured and exploited fully. Frequently, the purchase of seasonal item is made from six months to a year in advance. Departments with opposite seasonal changes are frequently combined in the same firm to avoid dull seasons and to keep sales or production up during the entire year.

Seasonal variations are measured as a percentage of the trend rather than in absolute quantities. The seasonal index for any month (week, quarter etc.) may be defined as the ratio of the normally expected value (excluding the business cycle and erratic movements) to the corresponding trend value. When cyclical movement and erratic fluctuations are absent in a time series, such a series is called normal. Normal values thus are consisting of trend and seasonal components. Thus when normal values are divided by the corresponding trend values, we obtain seasonal component of time series.

3. Business Cycle: Because of the persistent tendency for business to prosper, decline, stagnate, recover; and prosper again, the third characteristic movement in economic time series is called the business cycle. The business cycle does not recur regularly like seasonal movement, but moves in response to causes which develop intermittently out of complex combinations of economic and other considerations.

When the business of a country or a community is above or below normal, the excess deficiency is usually attributed to the business cycle. Its measurement becomes a process of contrast occurrences with a normal estimate arrived at by combining the calculated trend and seasonal movements. The measurement of the variations from normal may be made in terms of actual quantities or it may be made in such terms as percentage deviations, which is generally more satisfactory method as it places the measure of cyclical tendencies on comparable base throughout the entire period under analysis.

4. Erratic or Irregular Component: These movements are exceedingly difficult to dissociate quantitatively from the business cycle. Their causes are such irregular and unpredictable happenings such as wars, droughts, floods, fires, pestilence, fads and fashions which operate as spurs or deterrents upon the progress of the cycle. Examples of such movements are: high activity in middle forties due to erratic effects of 2nd world war, depression of thirties throughtout the world, export boom associated with Korean War in 1950. The common denominator of every random factor is that it does not come about as a result of the ordinary operation of the business system and does not recur in any meaningful manner.

Mathematical Statement of the Composition of Time Series

A time series may not be affected by all type of variations. Some of these type of variations may affect a few time series, while the other series may be effected by all of them. Hence, in analysing time series, these effects are isolated. In classical time series analysis it is assumed that any given observation is made up of trend, seasonal, cyclical and irregular movements and these four components have multiplicative relationship.

Symbolically:

$$O = T \times S \times C \times I$$

where O refers to original data,

T refers to trend,

S refers to seasonal variations,

C refers to cyclical variations and

I refers to irregular variations.

This is the most commonly used model in the decomposition of time series.

There is another model called Additive model in which a particular observation in a time series is the sum of these four components.

$$O = T + S + C + I$$

To prevent confusion between the two models, it should be made clear that in Multiplicative model S, C and I are indices expressed as decimal percents whereas in Additive model S, C and I are quantitative deviations about trend that can be expressed as seasonal, cyclical and irregular in nature.

If in a multiplicative model T = 500, S = 1.4, C = 1.20 and I= 0.7 then

$$O = T \times S \times C \times I$$

By substituting the values we get

$$O = 500 \times 1.4 \times 1.20 \times 0.7 = 608$$

In additive model, T = 500, S = 100, C = 25, I = –50

$$O = 500 + 100 + 25 - 50 = 575$$

The assumption underlying the two schemes of analysis is that whereas there is no interaction among the different constituents or components under the additive scheme, such interaction is very much present in the multiplicative scheme. Time series analysis, generally, proceed on the assumption of multiplicative formulation.

Methods of Measuring Trend

Trend can be determined by: (i) moving averages method and (ii) least-squares method. They are explained below.

(i) Method of Moving Averages : The moving average is a simple and flexible process of trend measurement which is quite accurate under certain conditions. This method establishes a trend by means of a series of averages covering overlapping periods of the data.

The process of successively averaging, say, three years data. and establishing each average as the moving average value of the central year in the group, should be carried throughout the entire series. For a five item, seven item or other moving averages, the same procedure is followed: the average obtained each time being considered as representive of the middle period of the group.

The choice of a 5-year, 7-year, 9-year, or other moving average is determined by the length of period necessary to eliminate the effects of the business cycle and erratic fluctuations. A good trend must be free from such movements, and if there is any definite periodicity to the cycle, it is well to have the moving average to cover one cycle period. Ordinarily, the necessary periods will range between three and ten years for general business series but even longer periods are required for certain industries.

In the preceding discussion, the moving averages of odd number of years were representatives of the middle years. If the moving average covers an even number of years, each average will still be representative of the midpoint of the period covered, but this mid-point will fall half way between the two middle years. In the case of a four year moving average, for instance each average represents a point half way between the second and third years. In such a case, a second moving average may be used to 'recentre' the averages. If the first moving averages gives averages centering half-way between the years, a further two-point moving average will recentre the data exactly on the years.

This method, however, is valuable in approximating trends in a period of transition when the mathematical lines or curves may be inadequate. This method provides a basis for testing other types of trends, even though the data are not such as to justify its use otherwise.

Example: Calculate 5-yearly moving average trend for the time series given below.

Year	:	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Quantity	:	239	242	238	252	257	250	273	270	268	288	284
Year	:	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	
Quantity	:	282	300	303	298	313	317	309	329	333	327	

Solution :

Year	Quantity	5-yearly moving total	5-yearly moving average
1980	239		
1981	242		
1982	238	1228	245.6
1983	252	1239	247.8
1984	257	1270	254.0
1985	250	1302	260.4
1986	273	1318	263.6
1987	270	1349	269.8
1988	268	1383	276.6
1989	288	1392	278.4
1990	284	1422	284.4
1991	282	1457	291.4
1992	300	1467	293.4
1993	303	1496	299.2
1994	298	1531	306.2
1995	313	1540	308.0
1996	317	1566	313.2
1997	309	1601	320.2
1998	329	1615	323.0
1999	333		
2000	327		

To simplify calculation work: Obtain the total of first five years data. Find out the difference between the first and sixth term and add to the total to obtain the total of second to sixth term. In this way the difference between the term to be omitted and the term to be included is added to the preceding total in order to obtain the next successive total.

Example : Fit a trend line by the method of four-yearly moving average to the following time series data.

Year	:	1991	1992	1993	1994	1995	1996	1997	1998
Sugar production (lakh tons)	:	5	6	7	7	6	8	9	10
Year	:	1999	2000	2001	2002				
Sugar production (lakh tons)	:	9	10	11	11				

Solution :

Year	Sugar Production (lakh tons)	4 yearly moving total	4 yearly moving average	To recenter trend values	
				2 yearly centred total	2 yearly moving average
1.	2.	3.	4.	5.	6.
1991	5				
1992	6				
1993	7	25	6.25	12.75	6.375
1994	7	26	6.50	13.50	6.75
1995	6	28	7.00	14.50	7.25
1996	8	30	7.50	15.75	7.875
1997	9	33	8.25	17.25	8.625
1998	10	36	9.00	18.50	9.25
1999	9	38	9.50	19.50	9.75
2000	10	40	10.00	20.25	10.125
2001	11	41	10.25		
2002	11				

Merits

1. This is a very simple method.
2. The element of flexibility is always present in this method as all the calculations have not to be altered if same data is added. It only provides additional trend values.
3. If there is a coincidence of the period of moving averages and the period of cyclical fluctuations, the fluctuations automatically disappear.
4. The pattern of moving average is determined in the trend of data and remains unaffected by the choice of method to be employed.
5. It can be put to utmost use in case of series having strikingly irregular trend.

Limitations

1. It is not possible to have a trend value for each and every year. As the period of moving average increases, there is always an increase in the number of years for which trend values cannot be calculated and known. For example, in a five yearly moving average, trend value cannot be obtained for the first two years and last two years, in a seven yearly moving average for the first three years and last three years and so on. But usually values of the extreme years are of great interest.
2. There is no hard and fast rule for the selection of a period of moving average.
3. Forecasting is one of the leading objectives of trend analysis. But this objective remains unfulfilled because moving average is not represented by a mathematical function.
4. Theoretically it is claimed that cyclical fluctuations are ironed out if period of moving average coincide with period of cycle, but in practice cycles are not perfectly periodic.

(ii) Method of Least Squares: If a straight line is fitted to the data it will serve as a satisfactory

trend, perhaps the most accurate method of fitting is that of least squares. This method is designed to accomplish two results.

(*i*) The sum of the vertical deviations from the straight line must equal zero.

(*ii*) The sum of the squares of all deviations must be less than the sum of the squares for any other conceivable straight line.

There will be many straight lines which can meet the first condition. Among all different lines, only one line will satisfy the second condition. It is because of this second condition that this method is known as the method of least squares. It may be mentioned that a line fitted to satisfy the second condition, will automatically satisfy the first condition.

The formula for a straight-line trend can most simply be expressed as

$$Y_c = a + bX$$

where X represents time variable, Y_c is the dependent variable for which trend values are to be calculated, a and b are the constants of the straight line to be found by the method of least squares.

Constant a is the Y-intercept. This is the difference between the point of the origin (O) and the point when the trend line and Y-axis intersect. It shows the value of Y when X= 0, constant b indicates the slope which is the change in Y for each unit change in X.

Let us assume that we are given observations of Y for n number of years. If we wish to find the values of constants a and b in such a manner that the two conditions laid-down above are satisfied by the fitted equation.

Mathematical reasoning suggests that, to obtain the values of constants a and b according to the Principle of Least Squares, we need to solve the following two equations.

$$\Sigma Y = na + b\Sigma Y \quad \ldots(i)$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 \quad \ldots(ii)$$

Solution of the two normal equations yield the following values for the constants *a* and *b* :

$$b = \frac{n\Sigma XY - \Sigma X \Sigma Y}{n\Sigma X^2 - (\Sigma X)^2}$$

and $$a = \frac{\Sigma Y - b\Sigma X}{n}$$

Least Squares Long Method: It makes use of the above mentioned two normal equations without attempting to shift the time variable to convenient mid-year. This method is illustrated by the following example.

Example : Fit a linear trend curve by the least-squares method to the following data :

Year	Production (Kg.)
1995	3
1996	5
1997	6
1998	6
1999	8

2000	10
2001	11
2002	12
2003	13
2004	15

Solution: The first year 1995 is assumed to be 0, 1996 would become 1, 1997 would be 2 and so on. The various steps are outlined in the following table.

Year	Production Y	X	XY	X^2
1	2	3	4	5
1995	3	0	0	0
1996	5	1	5	1
1997	6	2	12	4
1998	6	3	18	9
1999	8	4	32	16
2000	10	5	50	25
2001	11	6	66	36
2002	12	7	84	49
2003	13	8	104	64
2004	15	9	135	81
Total	89	45	506	285

The above table yields the following values for various terms mentioned below:

$n = 10$, $\Sigma X = 45$, $\Sigma X^2 = 285$, $\Sigma Y = 89$, and $\Sigma XY = 506$

Substituting these values in the two normal equations, we obtain

$$89 = 10a + 45b \quad ...(i)$$

$$506 = 45a + 285b \quad ...(ii)$$

Multiplying equation (*i*) by 9 and equation (*ii*) by 2, we obtain

$$801 = 90a + 405b \quad ...(iii)$$

$$1012 = 90a + 570b \quad ...(iv)$$

Subtracting equation *(iii)* from equation *(iv)*, we obtain

$$211 = 165\ b \text{ or } b = 211/165 = 1.28$$

Substituting the value of b in equation *(i)*, we obtain

$$89 = 10a + 45 \times 1.28$$

$$89 = 10a + 57.60$$

$$10a = 89 - 57.6$$

$$10a = 31.4$$

$$a = 31.4/10 = 3.14$$

Substituting these values of a and b in the linear equation, we obtain the following trend line

$Y_c = 3.14 + 1.28\ X$

Inserting various values of X in this equation, we obtain the trend values as below :

Year	Observed Y	a	b × X	y_c (Col. 3 + Col. 4)
1	2	3	4	5
1995	3	3.14	1.28 × 0	3.14
1996	5	3.14	1.28 × 1	4.42
1997	6	3.14	1.28 × 2	5.70
1998	6	3.14	1.28 × 3	6.98
1999	8	3.14	1.28 × 4	8.26
2000	10	3.14	1.28 × 5	9.54
2001	11	3.14	1.28 × 6	10.82
2002	12	3.14	1.28 × 7	12.10
2003	13	3.14	1.28 × 8	13.38
2004	15	3.14	1.28 × 9	14.66

Least Squares Method: We can take any other year as the origin, and for that year X would be 0. Considerable saving of both time and effort is possible if the origin is taken in the middle of the whole time span covered by the entire series. The origin would than be located at the mean of the X values. Sum of the X values would then equal 0. The two normal equations would then be simplified to

$$\Sigma Y = Na \qquad \ldots(i)$$

or $$a = \frac{\Sigma Y}{N}$$

and $$\Sigma XY = b\Sigma X^2 \qquad or \qquad b = \frac{\Sigma XY}{\Sigma X^2} \qquad \ldots(ii)$$

Two cases of short cut method are given below. In the first case there are odd number of years while in the second case the number of observations are even.

Illustration: Fit a straight line trend on the following data :

Year	1996	1997	1998	1999	2000	2001	2002	2003	2004
Y	4	7	7	8	9	11	13	14	17

Solution: Since we have 9 observations, therefore, the origin is taken at 2000 for which X is assumed to be 0.

Year	y	X	XY	X^2
1996	4	–4	–16	16
1997	7	–3	–21	5
1998	7	–2	–14	4
1999	8	–1	–8	1

2000	9	0	0	0
2001	11	1	11	1
2002	13	2	26	4
2003	14	3	42	9
2004	17	4	68	16
Total	90	0	88	60

Thus $n = 9$, $\Sigma Y = 90$, $\Sigma Y = 0$, $\Sigma XY = 88$, and $\Sigma X^2 = 60$

Substituting these values in the two normal equations, we get

$90 = 9a$ or $a = 90/9$ or $a = 10$

$88 = 60b$ or $b = 88/60$ or $b = 1.47$

$\therefore$ Trend equation is: $Y_c = 10 + 1.47\,X$

Inserting the various values of X, we obtain the trend values as below.

Years	Observed Y	X	a	b × X	Y_c (*Col.* 4 + *Col.* 5)
1996	4	–4	10	1.47 × –4 =5.88	4.12
1997	7	–3	10	1.47 × –3 =4.41	5.59
1998	7	–2	10	1.47 × –2 =2.84	7.06
1999	8	–1	10	1.47 × –1 =1.47	8.53
2000	9	0	10	1.47 × 0 = 0	10.00
2001	11	1	10	1.47 × 1 = 1.47	11.47
2002	13	2	10	1.47 × 2 = 2.94	12.94
2003	14	3	10	1.47 × 3 =4.41	14.41
2004	17	4	10	1.47 × 4 =5.88	15.88

Example: Fit a straight line trend to the data which gives number of passenger cars sold (millions)

Year	1995	1996	1997	1998	1999	2000	2001	2002
No. of cars (millions)	6.7	5.3	4.3	6.1	5.6	7.9	5.8	6.1

Solution :

Here are two mid-years viz; 1998 and 1999. The mid-point of the two years is assumed to be 0 and the time of six months is treated to be the unit. On this basis the calculations are as shown below:

Year	*Observed Y*	*X*	*XY*	X^2
1995	6.7	–7	–46.9	49
1996	5.3	–5	–26.5	25
1997	4.3	–3	–12.9	9
1998	6.1	–1	–6.1	1
1999	5.6	1	5.6	1

2000	7.9	3	23.7	9
2001	5.8	5	29.0	25
2002	6.1	7	42.7	49
Total	47.8	0	8.6	168

From the above computations, we get the following values.

$n = 8$, $\Sigma X = 47.8$, $\Sigma X = 0$. $\Sigma XY = 8.6$, $\Sigma X^2 = 168$

Substituting these values in the two normal equations, we obtain

$47.8 = 8a$ or $a = 47.8/8$ or $a = 5.98$

and $8.6 = 168\,b$ or $= 8.6/168$ or $b = 0.051$

The equation for the trend line is: $Y_c = 5.98 + 0.051$

Trend values generated by this equation are below.

Years	*Observed* Y	X	a	b	Y_c (*Col.* 4 + *Col.* 5)
1995	6.7	–7	5.98	0.051 × –7 = –.357	5.623
1996	5.3	–5	5.98	0.051 × –5 = –.255	5.725
1997	4.3	–3	5.98	0.051 × –3 = –.153	5.827
1998	6.1	–1	5.98	0.051 × –1 = –.051	5.939
1999	5.6	1	5.98	0.051 × 1 = 0.051	6.031
2000	7.9	3	5.98	0.051 × 3 = 0.153	6.133
2001	5.8	5	5.98	0.051 × 5 = 0.255	6.235
2002	5.1	7	5.98	0.051 × 7 = 0.357	6.337

This question paper contains 8 printed pages]

Your Roll No.....

आपका अनुक्रमांक.....

7928

B.Com/II JS

Paper VI - BUSINESS MATHEMATICS AND STATISTICS

Part A-Business Mathematics

(Admission of 2006 onwards)

Time : 1 Hour Maximum Marks : 25

समय : 1 घण्टा पूर्णांक : 25

(Write your Roll No. on the top immediately on receipt of this question paper.)

(इस प्रश्न–पत्र के मिलते ही ऊपर दिए गए निर्धारित स्थान पर अपना अनुक्रमांक लिखिए।)

Note :- (*i*) Answer may be written either in English or in Hindi; but the same medium should be used throughout the paper.

(*ii*) The maximum marks printed on the question paper are applicable for the students of the Cat. 'B'. These marks will, however, be scaled down proportionately in respect of the students of Cat. 'A' at the time of posting of posting of awards for compilation of result.

टिप्पणी : (*i*) इस प्रश्न–पत्र का उत्तर अंग्रेजी **या** हिंदी किसी एक भाषा में दीजिए; लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए।

(*ii*) प्रश्नपत्र पर अंकित पूर्णांक श्रेणी 'B' के विद्यार्थियों के लिए अनुप्रयोज्य हैं। तथापि ये अंक श्रेणी 'A' के विद्यार्थियों के संबंध में उनके परिणाम के संकलन के लिए नियुक्त अधिनिर्णय के समय पर, उनके आनुपातिक रूप में कम होंगे।

Attempt All questions.

Use of simple calculator and Log Table is allowed.

सभी प्रश्नों के उत्तर दीजिए।

सादा कैल्कुलेटर और लॉग टेबल के प्रयोग की अनुमति है।

1. An automobile company uses three grades of steel G-I, G-II, and G III for producing three models of car M-I, M-II and M-III. The steel requirement (in tons) for each model of car is given below :

Car Type

		M-I	M-II	M-III
Steel	G-I	2	3	4
Grade	G-II	1	1	2
	G-III	3	2	1

Determine the number of cars of each type which can be produced using 29, 13 and 16 tons of steel of three types respectively. 8

एक कार विनिर्माता कम्पनी कार के तीन मॉडल – M-I, M-II और M-III बनाने के लिए स्टील की तीन कोटि – G-I, G-II और G-III प्रयुक्त करती है। कार के प्रत्येक मॉडल के लिए स्टील की आवश्यकता (टनों में) इस प्रकार है :

कार–मॉडल

		M-I	M-II	M-III
स्टील	G-I	2	3	4
कोटि	G-II	1	1	2
	G-III	3	2	1

क्रमशः तीन प्रकार के 29, 13 और 16 टनों का उपयोग करते हुए बनाई जा सकने वाली प्रत्येक प्रकार की कारों की संख्या निर्धारित कीजिए।

Or

(अथवा)

A shop sells 20 table fans, 30 ceiling fans and 10 pedestal fans in a month. The price of a table fan is Rs. 300, of a ceiling fan is Rs. 400 and of pedestal fan is Rs. 500. The cost to the shop is Rs. 220 for one table fan, Rs. 325 for a ceiling fan and Rs. 400 for a pedestal fan. Find the monthly profit of the shop. 8

एक दुकान एक महीने में 20 टेबल फैन, 30 सीलिंग फैन और 10 पैडस्टल फैन बेचती है। एक टेबल फैन की 300 रू., सीलिंग फैन की 400 रू., और पैडस्टल फैन की 500 रू., कीमत है। दुकान को एक टेबल फैन की 220 रू., सीलिंग फैन की 325 रू., और पैडस्टल फैन की 400 रू. लागत आती है। दुकान का मासिक लाभ ज्ञात कीजिए।

2. Attempt any two :

(i) There are two concepts of interest-Simple interest and compound interest. Explain the two. It will take how many years for a sum to quadruple itself at 10% per year simple interest and compound interest ?

(ii) A machine depreciates at the rate of 10% per annum for the first two years and then 7% per annum for the next 3 years. Depreciation being calculated on the diminishing value. If the value of the machine be Rs. 10,000 initially, find the average rate of depreciation and the depreciated value of the machine at the end of fifth year.

(iii) The sum of Rs. 2,000, Rs. 3,000, Rs. 4,000 are due at the end of 2, 4 and 8 years respectively. It is proposed to replace this series of payments by a single sum of Rs. 9,000 payable at the end of n years. If the rate of interest is 10% per annum effective, find n.

(iv) Mr. X deposited Rs. 10,000 in a bank for 3 years offering interest rate of 6% compounded half yearly during first year, at the rate of 12% compounded quarterly during second year and at 10% compounded continuously during 3rd year.

कोई **दो** कीजिए :

(*i*) ब्याज की दो संकल्पनाएँ हैं–साधारण ब्याज और चक्रवृद्धि ब्याज। दोनों को स्पष्ट कीजिए।

किसी रकम को 10% वार्षिक साधारण ब्याज और चक्रवृद्धि ब्याज पर चौगुना होने में कितने वर्ष लगेंगे ?

(*ii*) एक मशीन को पहले दो वर्षों में 10% वार्षिक की पद पर मूल्यह्रास होता है और फिर अगले तीन वर्षों में 7% वार्षिक पर। मूल्यह्रास ह्रासमान मूल्य पर परिकलित किया जाता है। यदि प्रारंभ में मशीन का मूल्य 10,000 रूपये था, तब मूल्यह्रास की औसत दर तथा पाँचवें वर्ष के अंत में मशीन का मूल्यह्रासित मूल्य ज्ञात कीजिए।

(*iii*) 2,000 रू. 3,000 रू. और 4,000 रू. की राशि क्रमशः 2, 4 और 8 वर्षों के अंत में देय होगी। भुगतान की इस श्रृंखला को 9,000 रू. की एक एकल राशि से प्रतिस्थापित करने का प्रस्ताव है, जो n वर्षे के अंत में देय होगी। यदि प्रभावी ब्याज दर 10% वार्षिक है, तब n ज्ञात कीजिए।

(iv) श्री 'एक्स' ने बैंक में 3 वर्ष के लिए 10,000 रू. जमा कराए, जहाँ पहले वर्ष अर्ध–वार्षिक चक्रवृद्धि परिकलित 6% ब्याज मिलेगा, दूसरे वर्ष 12% दर पर तिमाही चक्रवृद्धि परिकलित ब्याज मिलेगा और तीसरे वर्ष 10% लगातार चक्रवृद्धि परिकलित ब्याज मिलेगा। 3 वर्ष के पश्चात् शेष ज्ञात कीजिए।

3. (i) For a company the total cost function C and the total revenue function R are given by

$$C = x^3 - 12x^2 + 48x + 11$$

and $R = 83x - 4x^2 - 21$

where x is output. Obtain the output for which profit is maximum and the maximum profit. 5

एक कम्पनी का कुल लागत फलन C और कुल आगम फलन R इस प्रकार दिए गए हैं :

$$C = x^3 - 12x^2 + 48x + 11$$

$$R = 83x - 4x^2 - 21$$

जहाँ x निर्गत है। वह निर्गत प्राप्त कीजिए जहाँ लाभ अधिकतम है और अधिकतम लाभ भी ज्ञात कीजिए।

Or

(अथवा)

The demand curve for a commodity is given by:

$$x = 20 - 2p - p^2$$

where P and x are the price and the quantity in demand respectively.

Find the elasticity of demand for P = 25. For which demand the elasticity will be unity ? 5

एक वस्तु का माँग वक्र इस प्रकार है :

$$x = 20 - 2p - p^2$$

जहाँ P और x क्रमशः कीमत और माँगी गई मात्रा है।

P = 25 के लिए माँग की लोच ज्ञात कीजिए। किस माँग के लिए लोच एक होगी ?

(ii) It is known that marginal revenue function of a commodity is given by

$$MR = 10 - 5x^2$$

Find the revenue function and demand function. 4

एक वस्तु का सीमांत आगम फलन इस प्रकार दिया गया है :

$$MR = 10 - 5x^2$$

आगम फलन और माँग फलन ज्ञात कीजिए।

Or

(अथवा)

Determine consumer surplus under perfect competition if the demand and supply law is given by $P = 25 - x^2$ and $P = 2x + 1$ respectively. 4

माँग और पूर्ति नियम क्रमशः इस प्रकार दिया गया है :

$P = 25 - x^2$ और $P = 2x + 1$

पूर्ण प्रतियोगिता के अंतर्गत उपभोक्ता अधिशेष निर्धारित कीजिए।

This question paper contains 8 printed pages]

Your Roll No.....

आपका अनुक्रमांक.....

5367

B.Com./II **J-I**

Paper VI- BUSINESS MATHEMATICS AND STATISTICS

(Part A-Business Mathematics)

(NC-Admission of 2006 onwards)

Time : 1 Hour Maximum Marks : 25

समय : 1 घण्टा पूर्णांक : 25

(Write your Roll No. on the top immediately on receipt of this question paper.)

(इस प्रश्न–पत्र के मिलते ही ऊपर दिए गए निर्धारित स्थान पर अपना अनुक्रमांक लिखिए।)

Note :- Answer may be written either in English or in Hindi; but the same medium should be used throughout the paper.

इस प्रश्न–पत्र का उत्तर अंग्रेजी या हिन्दी किसी एक भाषा में दीजिए; लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए।

Attempt All questions. Log tables and other tables shall be provided on demand. Use of simple calculator is allowed.

सभी प्रश्न कीजिए। लॉग टेबल तथा अन्य टेबल माँगने पर दी जायेंगी। साधारण कैल्कुलेटर के प्रयोग की अनुमति है।

1. What is a Matrix ? Use this tool for solving the following problem :

Three firms A, B and C supplied 40, 35 and 25 truck loads of stones and 10, 5 and 8 truck loads of sand respectively to a contractor. If the cost of stone and sand are Rs. 1,200 and Rs. 500 per truck load respectively, find the total amount paid by the contractor to each of these firms. 8

मैट्रिक्स क्या होती है ? निम्नलिखित समस्या को हल करने के लिए इस उपकरण को उपयोग कीजिए:

तीन फर्म A, B और C ने एक ठेकेदार को क्रमशः 40, 35, और 25 ट्रक भार पत्थर और 10, 5, 8 ट्रक भार रेत की पूर्ति की। यदि पत्थर और रेत की लागत क्रमशः 1,200 रु. और 500 रु. प्रति ट्रक भार है, तब ठेकेदार द्वारा इन फर्मों में से प्रत्येक को दी गई कुल राशि ज्ञात कीजिए।

Or

(अथवा)

Food I has 3 units of Vitamin A, 9 units of Vitamin B and 12 units of Vitamin C. Food II has 6, 9 and 15 units respectively and Food III has 9, 0, 9 units respectively. 33 units of Vitamin A, 27 of B and 60 of C are required. Find the amount of three foods that will provide exactly these amounts. Solve by matrix inverse method. 8

खाद्य I में 3 यूनिट विटामिन A, 9 यूनिट विटामिन B और 12 यूनिट विटामिन C हैं। खाद्य II में क्रमशः 6, 9 और 15 हैं और खाद्य III में क्रमशः 9, 0 और 9 यूनिट हैं। विटामिन A के 33 यूनिट, B के 27 यूनिट और C के 60 यूनिट की आवश्यकता है। इन तीन खाद्यों की वह मात्रा ज्ञात कीजिए जो ठीक इन मात्राओं की पूर्ति करेगी। मैट्रिक्स प्रतिलोम विधि द्वारा हल कीजिए।

2. Attempt any two : 4x2=8

(*i*) The compound interest on a certain sum of money for two years is Rs. 126 and simple interest for two years at the same rate is Rs. 120. Find the rate of interest.

(*ii*) A machine depreciates at the rate of 8% of its value at the beginning of a year. The machine was purchased for Rs. 1,00,000. The scrap value realized when sold was Rs. 43,440. Find the number of years the machine was sold ?

(*iii*) A certain sum of money at compound interest becomes Rs. 6,655 in 3 years and Rs. 7,320.50 in four years. Find the rate of interest.

(*iv*) What is effective rate of interest ?

Which is better investment-9% converted monthly or 9.1% converted semi-annually ?

कोई **दो** कीजिए:

(*i*) एक निश्चित रकम पर दो वर्ष के लिए चक्रवृद्धि ब्याज 126 रु. है और उसी दर पर दो वर्ष के लिए साधारण ब्याज 120 रु. है। ब्याज दर ज्ञात कीजिए।

(*ii*) एक मशीन का वर्ष के प्रारम्भ में अपने मूल्य के 8% की दर पर मूल्यह्रास होता है। मशीन 1,00,000 रु. में खरीदी गई थी। बेचने पर 43,440 रु. स्क्रेप मूल्य प्राप्त हुआ। ज्ञात कीजिए कि मशीन का कितने वर्ष उपयोग हुआ ?

(*iii*) एक रकम चक्रवद्धि ब्याज पर 3 वर्ष में 6,655 रु. हो जाती है और चार वर्ष में 7,320.50 रु.। ब्याज दर ज्ञात कीजिए।

(*iv*) प्रभावी ब्याज दर से क्या तात्पर्य है ?

बेहतर निवेश कौनसा है – 9% पर मासिक परिवर्तित अथवा 9.1% पर अर्धवार्षिक परिवर्तित ?

3. (A) (*i*) The cost function of a manufacturing company is given by:

$$C = 0.1x^3 - 3x^2 + 60x.$$

Find the value of x at which slope of average cost is zero. Also verify that AC = MC at this value of x 3

एक विनिर्माण कम्पनी का लागत फलन है :

$$C = 0.1x^3 - 3x^2 + 60x.$$

x का वह मान ज्ञात कीजिए जिस पर औसत लागत की ढाल शून्य है। x के इस मान पर AC=MC का भी सत्यापन कीजिए।

(*ii*) Find the marginal revenue function for a monopolist if the demand function is :

$$x = 20 - 10P$$ 2

मांग फलन $x = 20 - 10P$ होने पर एकाधिकारी के लिए सीमांत आगम फलन ज्ञात कीजिए।

Or

(अथवा)

Out of the following two functions, one is demand function and the other is the supply function. Determine, which one is demand function and which one is the supply function :

$$P = \sqrt{100 - x^2}$$

$$x = 2P - 10$$

Where P is price per unit and x is quantity. Also determine the elasticities of demand and supply at equilibrium price. 5

निम्नलिखित दो फलनों में से एक माँग फलन है और दूसरा पूर्ति फलन। निर्धारित कीजिए कि माँग फलन कौनसा और पूर्ति फलन कौनसा है :

$$P = \sqrt{100 - x^2}$$

$$x = 2P - 10$$

जहाँ P प्रति इकाई कीमत है और x मात्रा है। समतोल कीमत पर माँग और पूर्ति की लोच भी निर्धारित कीजिए।

(B) A firm has the following marginal revenue and marginal cost function :

$$MR = 50 - 4x \text{ and}$$

$$MC = -20 + 2x.$$

The firm suffers a loss of Rs. 200 when no units are sold.

(*i*) Determine total profit function.

(*ii*) Find break-even points.

(*iii*) Determine the profit enclosed between these two break-even points.

एक फर्म का सीमांत आगत और सीमांत लागत फलन निम्नलिखित है :

$$MR = 50 - 4x$$

$$MC = -20 + 2x.$$

फर्म द्वारा कोई इकाई न बेचने पर 200 रु. की हानि होती है :

(*i*) कुल लाभ फलन निर्धारित कीजिए।

(*ii*) लाभ–अलाभ प्रतिच्छेद बिन्दु ज्ञात कीजिए।

(*ii*) इन दोनों लाभ–अलाभ प्रतिच्छेद बिन्दुओं में आवेष्टित लाभ निर्धारित कीजिए।

Or

(अथवा)

Find consumer surplus for the demand function $x = 525 - 20p - p^2$ if the quantity demanded is 264 units. 4

माँग फलन $x = 525 - 20p - p^2$ के लिए उपभोक्ता अधिशेष ज्ञात कीजिए यदि 264 इकाइयों की माँग है।