மனோன்மணியம் சுந்தரனார் பல்கலைக்கழகம்

Manonmaniam Sundaranar University

Reaccredited with 'A' Grade (CGPA 3.13 Out of 4.0) by NAAC (3rd Cycle)

Tirunelveli - 627 012, Tamilnadu, India.

DIRECTORATE OF DISTANCE

&

CONTINUING EDUCATION

QUANTITATIVE TECHNIQUES

EDITED BY

Dr.M.NITHYA M.COM.,M.Phil.,M.B.A.,TN-SET.,Ph.D

ASSISTANT PROFESSOR (T)

DEPARTMENT OF COMMERCE

MANONMANIAM SUNDARANAR UNIVERISTY

TIRUNELVELI

PART III – MAJOR CORE - 7

QUANTITATIVE TECHNIQUES

Objectives

1. To provide basic knowledge of mathematical techniques as are applicable to business.
2. To provide logical idea to find out practical solutions for the managerial problems.
3. To provide the basic knowledge of statistical techniques as are applicable to business.
4. To enable the students to apply statistical techniques for quantification of data in business.

Unit I:

Analytical geometry- Distance between two points in a plane-slope of a straight line – equation of straight line – point of intersection of two lines – applications (1) demand and supply (2) cost-output (3) break-even analysis

Unit II:

Matrices – meaning – types – algebra of matrices – addition and subtraction – scalar multiplication – Multiplication of matrices-transpose of a matrix –Determinant – minors and co-factors - inverse of a matrix – solving simultaneous linear equations using matrix method.

Unit III:

Measures of Central Tendency – Mean – Median – Mode – Geometric Mean .Measures of Dispersion-Range – Quartile Deviation – Mean Deviation - Standard Deviation – Co-efficient of Variation. Skewness - methods of studying Skewness - Karl Pearson's Co-efficient of Skewness – Bowley's co-efficient of Skewness.

Unit IV:

Correlation – meaning – types-scatter diagram – Karl Pearson's Co-efficient of Correlation- Rank correlation – Concurrent deviation method. Regression analysis – uses-Regression line – Regression equations – least square method - deviations taken from actual mean and assumed mean method.

Unit V:

Index numbers – meaning – types – its problems – Methods of constructing index numbers – unweighted and weighted indices – Index number tests – Consumer price index numbers - Analysis of time series – Meaning – Importance – Components of time series – Secular trend, seasonal, cyclical and irregular variations – Measurement of trend - Graphic method-Semi average method – Moving average method – Method of least square.

Text / Reference Books

1. D.S. Sancheti & V.K. Kapoor, Business Mathematics Sultan Chand and Sons, New Delhi.
2. M. Manoharan & C. Elango, Business Mathematics, Palani Paramount Publications, Palani.

Unit – I - Analytic Geometry

Analytic Geometry is a branch of algebra, a great invention of Descartes and Fermat, which deals with the modelling of some geometrical objects, such as lines, points, curves, and so on. It is a mathematical subject that uses algebraic symbolism and methods to solve the problems. It establishes the correspondence between the algebraic equations and the geometric curves. The alternate term which is used to represent the analytic geometry is "coordinate Geometry".

It covers some important topics such as midpoints and distance, parallel and perpendicular lines on the coordinate plane, dividing line segments, distance between the line and a point, and so on. The study of analytic geometry is important as it gives the knowledge for the next level of mathematics. It is the traditional way of learning the logical thinking and the problem solving skills. In this article, let us discuss the terms used in the analytic geometry, formulas, Cartesian plane, analytic geometry in three dimensions, its applications, and some solved problems.

What is Analytic Geometry?

Analytic geometry is that branch of Algebra in which the position of the point on the plane can be located using an ordered pair of numbers called as Coordinates. This is also called **coordinate geometry** or the **Cartesian geometry**. Analytic geometry is a contradiction to the synthetic geometry, where there is no use of coordinates or formulas. It is considered axiom or assumptions, to solve the problems. But in analytic geometry, it defines the geometrical objects using the local coordinates. It also uses algebra to define this geometry.

Coordinate geometry has its use in both two dimensional and three-dimensional geometry. It is used to represent geometrical shapes. Let us learn the terminologies used in analytic geometry, such as;

- Plane
- Coordinates

Planes

To understand how analytic geometry is important and useful, First, We need to learn what a plane is? If a flat surface goes on infinitely in both the directions, it is called a **Plane**. So, if you find any point on this plane, it is easy to locate it using Analytic Geometry**.** You just need to know the coordinates of the point in X and Y plane.

Coordinates

Coordinates are the two ordered pair, which defines the location of any given point in a plane. Let's understand it with the help of the box below.

	A	B	C
1			
2		x	
3			

In the above grid, The *columns* are labelled as A, B, C, and the *rows* are labelled as 1, 2, 3.

The location of letter x is B2 i.e. Column B and row 2. So, B and 2 are the coordinates of this box, x.

As there are several boxes in every column and rows, but only one box has the point x, and we can find its location by locating the *intersection* of row and column of that box. There are different types of coordinates in analytical geometry. Some of them are as follows:

- Cartesian Coordinates
- Polar Coordinates
- Cylindrical Coordinates
- Spherical Coordinates

Cartesian Coordinates

The most well-known coordinate system is the Cartesian coordinate to use, where every point has an x-coordinate and y-coordinate expressing its horizontal position, and vertical position respectively. They are usually addressed as an ordered pair and denoted as (x, y). We can also use this system for three-dimensional geometry, where every point is represented by an ordered triple of coordinates (x, y, z) in Euclidean space.

Polar Coordinates

In the case of polar coordinates, each point in a plane is denoted by the distance 'r' from the origin and the angle θ from the polar axis.

Cylindrical Coordinates

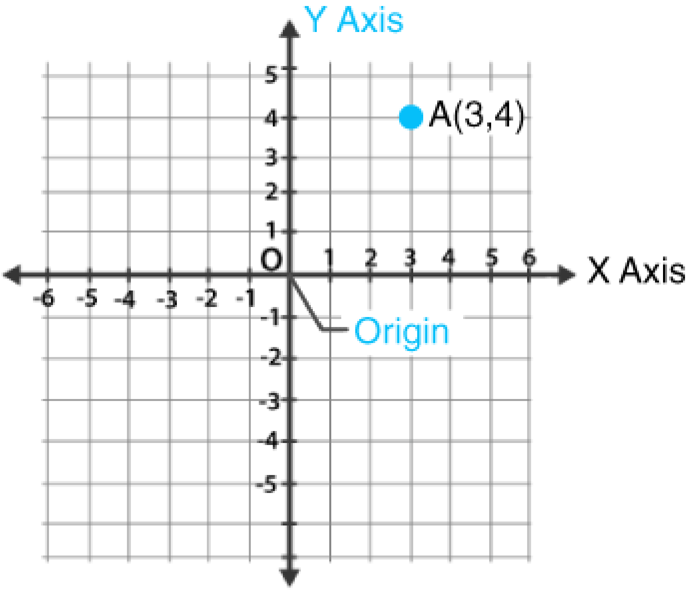
In the case of cylindrical coordinates, all the points are represented by their height, radius from z-axis and the angle projected on the xy-plane with respect to the horizontal axis. The height, radius and the angle are denoted by h, r and θ, respectively.

Spherical Coordinates

In spherical coordinates, the point in space is denoted by its distance from the origin (ρ), the angle projected on the xy-plane with respect to the horizontal axis (θ), and another angle with respect to the z-axis (φ).

Cartesian Plane

In coordinate geometry, every point is said to be located on the coordinate plane or cartesian plane only.



The above graph has *x-axis* and *y-axis* as it's Scale. The x-axis is running across the plane and Y-axis is running at the right angle to the x-axis. It is similar to the box explained above.

Origin: It is the point of intersection of the axis(x-axis and y-axis). Both x and y-axis are zero at this point.

Values of the different sides of the axis:

x-axis – The values at the right-hand side of this axis are positive and those on the left-hand side are negative.

y-axis – The values above the origin are positive and below the origin are negative.

To locate a point: We need two numbers to locate a plane in the order of writing the location of X-axis first and Y-axis next. Both will tell the single and unique position on the plane. You need to compulsorily follow the order of the points on the plane i.e., the x coordinate is always the first one from the pair. (x, y).

If you look at the figure above, point A has a value 3 on the x-axis and value 2 on the Y-axis. These are the rectangular coordinates of Point A represented as (3, 2).

Using the Cartesian coordinates, we can define the equation of a straight lines, equation of planes, squares and most frequently in the three dimensional geometry. The main function of the analytic geometry is that it defines and represents the various geometrical shapes in the numerical way. It also extracts the numerical information from the shapes.

Analytic Geometry Formulas

Graphs and coordinates are used to find measurements of geometric figures. There are many important formulas in analytic Geometry. Since science and engineering involves the study of rate of change in varying quantities, it helps to show the relation between the quantities involved. The branch of Mathematics called “calculus” requires the clear understanding of the analytic geometry. Here, some of the important ones are being used to find the distance, slope or to find the equation of the line.

Distance Formula

Let the two points be A and B, having coordinates to be (x_1, y_1) and (x_2, y_2) respectively.

Thus, the distance between two points is given as-

$$\mathbf{d = \sqrt{[(x_2-x_1)^2+(y_2-y_1)^2]}}$$

Midpoint Theorem Formula

Let A and B are some points in a plane, which is joined to form a line, having coordinates (x_1, y_1) and (x_2, y_2), respectively. Suppose, M(x, y) is the midpoint of the line connecting the point A and B then its formula is given by;

$$\mathbf{M(x, y) = [(x_1+x_2)/2, (y_1+y_2)/2)]}$$

Angle Formula

Let two lines have slope m_1 and m_2 and θ is the angle formed between the two lines A and B, which is represented as;

$$\mathbf{\tan \theta = (m_1-m_2)/ (1+m_1m_2)}$$

Section Formula

Let two lines A and B have coordinates (x_1, y_1) and (x_2, y_2), respectively. A point P the two lines in the ratio of m:n, then the coordinates of P is given by;

- **When the ratio m:n is internal:**

$$\left(\frac{mx_2+nx_1}{m+n}, \frac{my_2+ny_1}{m+n}\right)$$

- **When the ratio m:n is external:**

$$\left(\frac{mx_2-nx_1}{m-n}, \frac{my_2-ny_1}{m-n}\right)$$

Analytic Geometry in Three Dimensions

In this, we consider *triples (*a, b, c) which are real numbers and call this set as three- dimensional number space and denote it by **R'**. All the elements in the triple are called **coordinates.**

Let's see how three-dimensional number space is represented on a geometric space.

In three-dimensional space, we consider three mutually perpendicular lines intersecting in a point O. these lines are designated coordinate axes, starting from 0, and identical number scales are set up on each of them.

Analytic Geometry Applications

Analytic geometry is widely used in the fields such as Engineering and Physics. Also, it is widely used in the fields such as space science, rocket science, aviation, space flights and so on. Analytical geometry has made many things possible like the following:

- We can find whether the given lines are perpendicular or parallel.
- We can determine the mid-point, equation, and slope of the line segment.
- We can find the distance between the points.
- We can also determine the perimeter of the area of the polygon formed by the points on the plane.
- Define the equations of ellipse, curves, and circles.

DISTANCE BETWEEN TWO POINTS

If $A(x1, y1)$ and $B(x2, y2)$ be any two points in the plane. Now find the distance between these two points.

Let P and Q be the foot of the perpendiculars from A and B to the x-axis respectively. AR is drawn perpendicular to BQ. From the diagram,

AR = PQ = OQ - OP = $x2- x1$And BR = BQ - RQ = y2- y1 y B (x2, y2)

From right angle ARB

y2- y1
R

$AB^2 = AR^2+RB^2 = (x_2- x_1)^2 + (y_2- y_1)^2$

$\therefore$ **AB =** $\sqrt{(\boldsymbol{x_2} - \boldsymbol{x_1})^{\boldsymbol{2}} + (\boldsymbol{y_2} - \boldsymbol{y_1})^{\boldsymbol{2}}}$

Problem 1

Find the distance between the points (-4, 0) and (3, 0)

Solution

The points (-4, 0) and (3, 0) lie on the x-axis. Hence

d = $\sqrt{(\boldsymbol{x_2} - \boldsymbol{x_1})^{\boldsymbol{2}} + (\boldsymbol{y_2} - \boldsymbol{y_1})^{\boldsymbol{2}}}$

AB = $\sqrt{(3 - -4)^2 + (0 - 0)^2}$

AB = $\sqrt{(7)^2 + (0)^2}$

AB $\sqrt{49}$

AB =

7

Problem 2

Show that the three points (4, 2), (7, 5) and (9, 7) lie on a straight line

Solution

Let the points be A (4, 2), B (7, 5) and C (9, 7). By the distance formula

$AB^2 = (4-7)^2 + (2-5)^2 = (-3)^2 + (-3)^2 = 9+9 = 18$

$BC^2 = (9-7)^2 + (7-5)^2 = (2)^2 + (2)^2 = 4+4 = 8$

$AC^2 = (9-4)^2 + (7-2)^2 = (5)^2 + (5)^2 = 25+25 = 50$

So, AB = 18 = 9 × 2 = 3 2 ; BC = 8 = 4 × 2 = 2 2 ;

CA= 50 = 25 × 2 = 5 2

This gives AB+ BC = 3 2 + 2 2 = 5 2 Hence the points A, B and C are collinear.

Problem 3

Determine whether the points are vertices of a right triangle A (–3, –4), B (2, 6) and C (–6, 10)

Solution Using the distance formula $d = \square \boldsymbol{x_2 - x_1}^{\mathbf{2}} + \boldsymbol{y_2 - y_1}^{\mathbf{2}}$

$AB^2 = (2- -3)^2 + (6 - -4)^2 = (5)^2 + (10)^2 = 25+100 = 125$

$BC^2 = (-6-2)^2 + (10-6)^2 = (-8)^2 + (4)^2 = 64+16 = 80$

$AC^2 = (-6 - -3)^2 + (10- -4)^2 = (-3)^2 + (14)^2 = 9+196 = 205$

$\therefore AB^2 + BC2 = 125 + 80 = 205 = CA^2$

Hence ABC is a right angled triangle since the square of one side is equal to sum of the squares of the other two sides.

Problem 4

Show that the points (a, a), (- a, - a) and(- a 3 , a 3) form an equilateral triangle.

Solution

Let the points be represented by A= (a, a), B= (- a, - a) and C= (- $a\sqrt{3}$, $\sqrt{3}$) Using the distance formula

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$AB = \sqrt{(a - a)^2 + (-a - a)^2}$$

$$= \sqrt{(-2a)^2 + (-2a)^2}$$

$$= \sqrt{4a^2 + 4a^2}$$

$$= \sqrt{8a^2} = 2\sqrt{2}a$$

$$BC = \sqrt{(-a\sqrt{3} - -a)^2 + (a\sqrt{3} - -a)^2}$$

$$= \sqrt{(-a\sqrt{3} + a)^2 + (\sqrt{3} + a)^2}$$

$$= \sqrt{3a^2 + a^2 - 2a^2\sqrt{3} + 3a^2 + a^2 + 2a^2\sqrt{3}}$$

$$\sqrt{8a^2} = 2\sqrt{2}a$$

$$AC = \sqrt{(a - -a\sqrt{3})^2 + (a - a\sqrt{3})^2}$$

$$= \sqrt{a^2 + 2a^2\sqrt{3} + 3a^2 + a^2 - 2a^2\sqrt{3} + 3a^2}$$

$= \sqrt{8a^2} = 2\sqrt{2}a$

∴ AB = BC= AC

Since all the sides are equal the points form an equilateral triangle

Problem 5

Prove that the points (–7, –3), (5, 10), (15, 8) and (3, –5) taken in order are the corners of a parallelogram.

Solution

Let A, B, C and D represent the points (-7, -3), (5, 10), (15, 8) and (3, -5) respectively.

Using the distance formula $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

$AB^2 = (5- -7)^2 + (10 - -3)^2 = (12)^2 + (13)^2 = 144+169 = 313$

$BC^2 = (15-5)^2 + (8-10)^2 = (10)^2 + (-2)^2 = 100+4 = 104$

$CD^2 = (3-15)^2 + (-5- 15)^2 = (-12)^2 + (-13)^2 = 144+196= 313$

$DA^2 = (3--7)^2 + (-5--3)^2 = (10)^2 + (-2)^2 = 100+4 = 104$

So, AB=CD= 313

*BC=DA= 104

The opposite sides are equal. Hence ABCD is a parallelogram.

Problem 6

Show that the following points (3, –2), (3, 2), (–1, 2) and (–1, –2) taken in order are vertices of a square.

Solution

Let the vertices be taken as A (3, -2), B (3, 2), C (-1, 2) and D (-1, -2).

$AB^2 = (3-3)^2 + (2 - -2)^2 = (0)^2 + (4)^2 = 0+16 = 16$

$BC^2 = (3+1)^2 + (2-2)^2 = (4)^2 + (0)^2 = 16+0 = 16$

$CD^2 = (-1+1)^2 + (2+2)^2 = (0)^2 + (4)^2 = 0+16 = 16$

$DA^2 = (-1-3)^2 + (-2--2)^2 = (-4)^2 + (0)^2 = 16+0 = 16$

AB = BC = CD = DA = 16 = 4. (That is, all the sides are equal.)$AC^2 = (-1-3)^2 + (2 - -2)^2 = (-4)^2 + (4)^2 = 16+16 = 32$

$BD^2 = (-1-3)^2 + (-2-2)^2 = (-4)^2 + (-4)^2 = 16+16 = 32$

AC=BD= 32 =4 2 (that is the diagonals are equal) Hence the points A,B,C and D form a square .

SLOPE OF A STRAIGHT LINE

The measure of steepness and direction of straight line is given by its slope. Slope is usually represented by the letter m.

In the given figure, if the angle of inclination of the given line with the x- axis is ∅, then then the slope of the line is represented by tan ∅.

1. Slope of the line joining two points

$$\boldsymbol{m} = \frac{\boldsymbol{y2 - y1}}{\boldsymbol{x2 - x1}}$$

2. Slope of the line which is parallel to x – axis

$$\boldsymbol{m} = \frac{\boldsymbol{y - y}1}{\boldsymbol{x2 - x1}}$$

$$\boldsymbol{m} = \frac{\mathbf{0}}{\boldsymbol{x2 - x1}}$$

3. Slope of the line which is parallel to y- axis

$$\boldsymbol{m} = \frac{\boldsymbol{y2 - y1}}{\boldsymbol{x - x}}$$

$$\boldsymbol{m} = \frac{\boldsymbol{y2 - y1}}{\mathbf{0}}$$

4. Slope of the line joining the origin and any point

$\boldsymbol{y1}$

$\boldsymbol{m} = \dfrac{}{\boldsymbol{x1}}$

5. Slope of the equation

$$\boldsymbol{m} = \frac{-\boldsymbol{a}}{\boldsymbol{b}}$$

a= co–efficient of x

b=co-efficient of y

Problem 1

Find the slope of the lines joining the points

(i) (-1,3) and (2,5)

(ii) (-2,-1) and (1,3)

Solution

$$m = \frac{\boldsymbol{y2-y1}}{\boldsymbol{x2-x1}}$$

(i) Let A= (-1,3) and B= (2,5)

Slope of the line AB = $m = \dfrac{\mathbf{5-3}}{\mathbf{2--1}}$

= 2/3

(iii) Let A= (-2,-1) and B= (1,3) Slope of the line AB = $m = \dfrac{\mathbf{3-1}}{\mathbf{1-2}}$

=4/3

Problem 2

Find the slope of the line joining the points

(i) (-3,2) and (4,2)

(ii) (2,5) and (2,3)

(iii) (0,0) and (1,2)

Solution

$m = \dfrac{y_2 - y_1}{x_2 - x_1}$

(i) Let A= (-3,2) and B=(4,2)

Slope of the line AB = $m = \dfrac{2-2}{4--3}$

$m = \dfrac{0}{7}$

Note: If m= 0, it shows us that the line is parallel to x- axis

(ii) Let A= (2,5) and B=(2,3)

Slope of the line AB = $m = \dfrac{3-5}{2-2}$

$= \dfrac{-2}{} = 0$

Note: If m= 0, it shows us that the line is parallel to y- axis

(iv) Let A= (0,0) and B= (1,2)

$m = \dfrac{y_1}{x_1}$

Slope of the line AB $= \dfrac{2}{1}$

= 2/1 □ 2

Note : we used $m = \dfrac{y_1}{x_1}$ Formula, because line passes from the origin

Problem 3

Find the slope of the equation of the lines

(i) 2x+5y-4=0

x-

4y=3

y=x+1

Slope of the equation

$$m = \frac{-a}{b}$$

a= co–efficient of

x b=co-efficient of

y

(i) 2x+5y-4=0 a=2 b=5

M = -a/b

M = -2/5

(ii) x-4y=3 a=1 b= -

4M = - 1 / - 4

(iii) y=x+1

-x+y=1 a= -1 b=1

$$m = \frac{-(-1)}{1}$$

$$m = \frac{1}{1} = 1$$

Problem 4

Show that the points (2,-4) (4,-2) and (7,1) are collinear

Solution

Let A= (2,-4) B= (4,-2) and C= (7, 1)

Slope of AB = $\frac{y2 - y1}{x2 - x1}$

$= \frac{-2-(-4)}{4-2}$

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

Slope of BC = $\frac{y2-y1}{x2-x1}$

$= \frac{1-(-2)}{7-4}$

$= \frac{3}{3} = 1$

Slope of AB = slope of BC proved Therefore A,B and C lie on the same line.

Problem 5

Find the value of K if the points (K,3) (-6,4) and (-10,5) are collinear.

Solution

Let A= (K,3), B= (-6,4) and C= (-10,5)

When A,B and C are collinear Slope of AB = slope of BC

$\frac{4-3}{-6-K} = \frac{5-4}{-10-(-6)}$

1 / -6-k = 1 / -4

(cross multiplying)

-6- K= - 4

-K = - 4 + 6

- K=

2 K =

- 2

Problem 6

Show that the line joining the points (-2 , 3) and (4 , 2) is parallel to the line joining the points (3,4) and (- 3 , 5)

Solution

Let A= (-2 , 3) and B= (4 , 2)

C= (3,4) and D= (- 3 , 5)

Slope of AB = $\frac{Y2 - Y1}{X2 - X1} = \frac{2 - 3}{4 - (-2)} = \frac{-1}{6}$

Slope of AB = $\frac{Y2 - Y1}{X2 - X1} = \frac{5 - 6}{3 - 3} = \frac{-1}{-6}$

m1= m2 proved

Hence the line AB is parallel to the line CD

Problem 7

If the line joining the points (3,2) and (2,-3) is parallel to the line joining the points (4,3) and (2,k). Find the K value

Solution

Let A= (3,2) and B= (2,-3)

C= (4,3) and D= (2,K)

Slope of AB = Slope of CD

$$\frac{-3 - 2}{2 - 3} = \frac{K - 3}{2 - 4}$$

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

(cross multiplying)

$10 = -K + 3$

10-3=-K

- K= 7 K=-7

Problem 8

Show that the line joining the points (3,- 4) and (2,1) is perpendicular to the line joining the points (-2 , 2) and (3,3)

Let A= (3,- 4) and B= (2,1)

C= (-2 , 2) and D= (3,3)

Slope of AB = $\frac{Y2 - Y1}{X2 - X1}$ = $\frac{1-(-4)}{2-3}$ = $\frac{5}{-1}$

Slope of CD = $\frac{Y2 - Y1}{X2 - X1}$ = $\frac{3-2}{3-(-2)}$ = $\frac{1}{5}$

m1 ×m2 = −1

$-5 \times \frac{1}{5} = \frac{-5}{5} = -1$ proved

Hence the line AB is perpendicular to the line CD.

Problem 9

If the lines joining the points (-3, 4) and (2,-3) is perpendicular to the line joining the points (3,K) and (2, -3). Find the value of K.

Solution

Let A= (-3, 4) and B= (2,-3)

C= (3,K) and D= (2, -3)

Slope of AB = $\frac{Y2 - Y1}{X2 - X1} = \frac{-3-4}{2-(-3)} = \frac{-7}{-5}$

$\frac{-7}{5}$ X m2 = -1

M2 = $-1 \times \frac{-5}{7}$

M2 = $\frac{5}{7}$

Slope of CD = $\frac{Y2 - Y1}{X2 - X1}$ = M2 = $\frac{5}{7}$

$\underline{5} = \underline{-3- K}$

7 2−3

5 = −3−K

7 −1

(Cross multiplying)

-5= -21-7K

-5+21= -7K

16= -

7K

K=−16

7

EQUATION OF STRAIGHT LINE

If a line is at a distance a and parallel to x-axis, then the equation of the line is **y = ± a**.

If a line is parallel to y-axis at a distance b from y-axis then its equation is **x = ± b**

Point-slope form : The equation of a line having slope m and passing through the point (x0, y0) is given by $\boldsymbol{y - y0 = m(x - x0)}$

Two-point-form : The equation of a line passing through two points x1, y1 and x2, y2 is given

$$\boldsymbol{y - y1 = \frac{y2 - y1}{x2 - x1} \quad x - x1}$$

Slope intercept form : The equation of the line making an intercept c on y-axis and having slope m is given by $\boldsymbol{y = mx + c}$

Note that the value of c will be positive or negative as the intercept is made on the positive or negative side of the y-axis, respectively

Intercept form : The equation of the line making intercepts a and b on x- and y-axis respectively is given by

X + Y =

1A B

General equation of a line

Any equation of the form Ax + By + C = 0, where A and B are simultaneously not zero, is called the general equation of a line.

Problem 1

Find the equation to the line with slope 1/3 and y intercepts 4 Slope – intercept

$\boldsymbol{y} = \boldsymbol{mx} + \boldsymbol{c}$ m = 1/3 c =

4Y = $\underline{1}$ x + 4

3

Y = $\underline{x + 12}$

3

3y= x+12

-x+3y= 12

-x+3y-12=0

Problem 2

Find the equation to the line which passes through (-1,3) and has slope $\underline{1}$ Point slope form

3

$\boldsymbol{y} - \boldsymbol{y0}$ = m($\boldsymbol{x} - \boldsymbol{x0}$)

$\boldsymbol{y} - \boldsymbol{3} = \underline{1\ } \boldsymbol{x}$—

$(-\boldsymbol{1})3$

$\boldsymbol{y} - \boldsymbol{3} = \underline{1}\,\boldsymbol{x} + \boldsymbol{1}$

3

(Cross multiplying)

3y-3=x+1

-x+3y=1+3

-x+3y=4 or

-x+3y-4=0

Problem 3

Find the equation to the line joining the points (0,-3) and (-4,-5) (0,-3) and (-4,-5)

Two point form

$\boldsymbol{y} - \boldsymbol{y1} = \underline{\boldsymbol{y2} - \boldsymbol{y1}} \quad \boldsymbol{x} - \boldsymbol{x1}$

$\boldsymbol{x2} - \boldsymbol{x1}$

Y – (-3) = $\underline{-5 – (-3)}$ X - 0

-4 – 0

Y + 3 = $\underline{-2}$ X

$$Y + 3 = \frac{-4}{\frac{1}{2}} X$$

$2y+6=x$

$-x+2y+6=0$ or $-x+2y=-6$

Problem 4

Find the equation to the line cutting of intercepts -3 and 4 on x and y axis Two intercepts form

$$\frac{x}{a} + \frac{y}{b} = 1 \qquad a= -3 \text{ and } b= 4$$

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

$$\frac{-4x+3y}{12} = 1$$

$-4x+3y= 12$ or $-4x+3y-12=0$

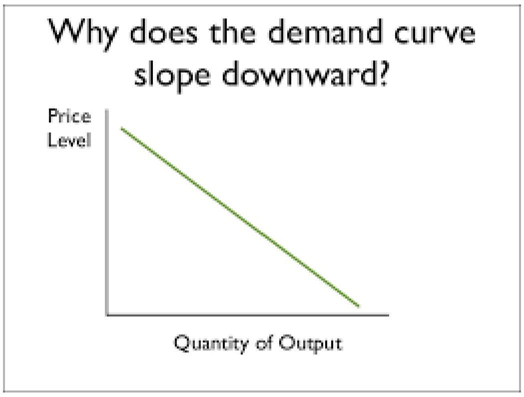
Demand and supply

The law of demand states that there is a negative or inverse relationship between the price and quantity demanded of a commodity over a period of time.

Definition:

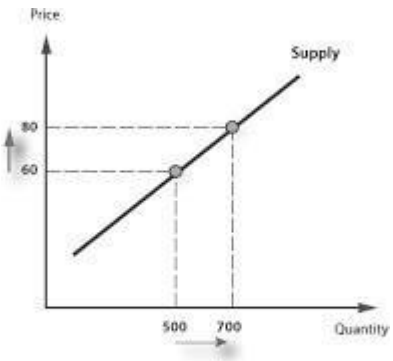
Alfred Marshall stated that "the greater the amount sold, the smaller must be the price at which it is offered, in order that it may find purchasers; or in other words, the amount demanded increases with a fall in price and diminishes with rise in price". According to Ferguson, the law of demand is that the quantity demanded varies inversely with price.

Thus the law of demand states that people will buy more at lower prices and buy less at higher prices, other things remaining the same. By other things remaining the same, we mean the following assumptions.



The demand curve slopes downwards mainly due to the law of diminishing marginal utility. The law of diminishing marginal utility states that an additional unit of a commodity gives a lesser satisfaction. Therefore, the consumer will buy more only at a lower price. **The demand curve slopes downwards because the marginal utility curve also slopes downwards.**

Supply means the goods offered for sale at a price during a specific period of time. It is the capacity and intention of the producers to produce goods and services for sale at a specific price.

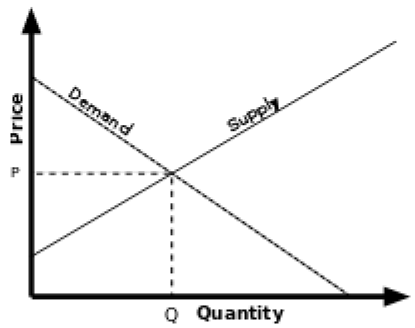


The supply of a commodity at a given price may be defined as the amount of it which is actually offered for sale per unit of time at that price.

The law of supply establishes a direct relationship between price and supply. Firms will supply less at lower prices and more at higher prices. “Other things remaining the same, as the price of commodity rises, its supply expands and as the price falls, its supply contracts”.

Market Equilibrium

When the supply and demand curves intersect, the market is in equilibrium. This is where the quantity demanded and quantity supplied is equal. The corresponding price is the equilibrium price or market-clearing price, the quantity is the equilibrium quantity.



Problem 1

15 tables are sold when the price is Rs 500 and 25 tables are sold when the price is Rs 400.What is equation of the demand curve assuming it to be linear?

Let

X= demandY= price

15	500
25	400

Demand curve points are (15, 500) (25, 400)

X1, Y1 X2, Y2

Equation formula

$\boldsymbol{y - y1} = \underline{\boldsymbol{y2 - y1}} \quad \boldsymbol{x - x1}$

$\boldsymbol{x2 - x1}$

$\boldsymbol{y - 500} = \underline{\boldsymbol{400 - 500}} \quad \boldsymbol{x - 15}$

$\boldsymbol{25 - 15}$

$\boldsymbol{y - 500} = \underline{-100}\ \boldsymbol{x - 15}$

10

$\boldsymbol{y - 500} = \boldsymbol{-10} \quad \boldsymbol{x - 15}$

$\boldsymbol{y - 500} = \boldsymbol{-10x + 150}$

10x+y=150+500

10x+y=650

The demand curve equation is 10x+y=650

Problem 2

When the price is Rs 30 , 100 toys of a particular type are available and when the price is Rs 50, 150 toys of the same type are available in the market

Let

X= supply Y= price

100	30
150	50

Supply curve points are (100, 30) (150, 50)

X1, Y1 X2, Y2

Equation formula

$$\boldsymbol{y - y1 = \frac{y2 - y1}{x2 - x1} \quad x - x1}$$

$$\boldsymbol{y - 30 = \frac{50 - 30}{150 - 100} \quad x - 100}$$

$$\boldsymbol{y - 30} = \frac{20}{50} \quad \boldsymbol{x - 100}$$

$$\boldsymbol{y - 30} = \frac{2}{5} \quad \boldsymbol{-100}$$

Cross multiplying

5(y-30) = 2 (x-100)

5y-150=2x-200

-2x+5y= -200+150

-2x+5y= -50 (or)

2x-5y= 50

The supply curve equation is 2x-5y=50

Analytic Geometry Problems

Example 1:

What is the point of intersection of the axis (X-axis and Y-axis) called?

Solution:

The point of intersection of the axis (X-axis and Y-axis) called **Origin** and X and the Y-axis is 0 at this point.

Example 2:

Find the distance between two points A and B such that the coordinates of A and B are (5, -3) and (2, 1).

Solution:

Given that, the coordinates are:

$A = (5, -3) = (x_1, y_1)$

$B = (2, 1) = (x_2, y_2)$

The formula to find the distance between two points is given as:

Distance, $d = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]}$

$d = \sqrt{[(2 - 5)^2 + (1 - (-3))^2]}$

$d = \sqrt{[(-3)^2 + (4)^2]}$

$d = \sqrt{[9 + 16]}$

$d = \sqrt{(25)}$

$d = 5$

Thus, the distance between two points A and B is 5 units.

Example 3:

Determine the slope of the line, that passes through the point A(5, -3), and it meets y-axis at 7.

Solution:

Given that, the point is A = (5, -3)

We know that, if the line intercepts at y-axis, then $x_2 = 0$

Thus, $(x_2, y_2) = (0, 7)$

The formula to find the slope of a line is:

$m = (y_2 - y_1)/(x_2 - x_1)$

Now, substitute the values

$m = (7 - (-3))/(0 - 5)$

$m = 10/-5$

$m = -2$

Therefore, the slope of the line is -2.

Unit –II - Matrices

Matrices Definition

A matrix is a function that consists of an ordered rectangular array of numbers. The numbers in the array are called the entities or the elements of the matrix. The horizontal array of elements in the matrix is called rows, and the vertical array of elements are called the columns. If a matrix has m rows and n columns, then it is known as the matrix of order m x n.

Introduction:

Sir ARTHUR CAYLEY (1821-1895) of England was the first Mathematician to introduce the term MATRIX in the year 1858. But in the present day applied Mathematics in overwhelmingly large majority of cases it is used, as a notation to represent a large number of simultaneous equations in a compact and convenient manner. Matrix Theory has its applications in Operations Research, Economics and Psychology. Apart

Definition of a Matrix

A rectangular array of numbers or functions represented by the symbol.

$$\begin{matrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ a_{m1} & a_{m2} & a_{mn} \end{matrix}$$ is called a matrix.

The numbers or functions a_{ij} of this array are called elements, may be real or complex numbers, whereas m and n are positive integers, which denotes the number of Rows and number of Columns.

Order of a Matrix

A matrix A with m rows and n columns is said to be of the order m by n (m x n). Symbolically A = (a_{ij}) m x n is a matrix of order m x n. The first subscript i in (a_{ij}) ranging from 1 to m identifies the rows and the second subscript j in (a_{ij}) ranging from 1 to n identifies the columns.

<u>For example</u>

A= 2 3

1

4 5 6 is a Matrix of order 2 × 3

B= 2

1

2 4 is a Matrix of order 2 × 2

C= sin θ cos θ is a Matrix of order 2 ×

2cos θ cos θ

1 2 3

D= 4 5 6 is a Matrix of order 3 × 3

7 8 9

Types of Matrices

SQUARE

MATRIX

When the number of rows is equal to the number of columns, the matrix is called a Square Matrix.

1 3

B= 2 4 is a Matrix of order 2 × 2

1 2 3

D= 4 5 6 is a Matrix of order 3 × 3

7 8 9

ROW MATRIX

A matrix having only one row is called Row Matrix.

For example A= (2 0 1) is a row matrix of order 1 x 3

B=(1 0) is a row matrix or order 1 x 2

COLUMN MATRIX

A matrix having only one column is called Column Matrix. For example

1

$A = \begin{matrix} 2 \\ 3 \end{matrix}$ is a column matrix of order 3 × 1.

$B = \begin{matrix} 2 \\ 1 \end{matrix}$ is a column matrix of order 2 × 1.

ZERO OR NULL MATRIX

A matrix in which all elements are equal to zero is called Zero or Null Matrix and is denoted by O.

$0 = \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}$ is a null matrix of order 2× 2.

$o = \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix}$ is a null matrix of order 3× 2

DIAGONAL MATRIX

A square Matrix in which all the elements other than main diagonal elements are zero is calleda diagonal matrix

For example

$A = \begin{matrix} 1 & 0 \\ 0 & 4 \end{matrix}$ is a diagonal matrix of order 2 and

$B = \begin{matrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{matrix}$ is a diagonal matrix of order 3

SCALAR MATRIX

A Diagonal Matrix with all diagonal elements equal to K (a scalar) is called a Scalar Matrix.

For example

$A = \begin{matrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{matrix}$ is a scalar matrix of order 3 and the value of scalar K=2

UNIT MATRIX OR IDENTITY MATRIX

A scalar Matrix having each diagonal element equal to 1 (unity) is called a Unit Matrix and isdenoted by I.

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ is a unit matrix of order 2

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ is a unit matrix of order 3

Multiplication of a matrix by a scalar

If A = (a_{ij}) is a matrix of any order and if K is a scalar, then the Scalar Multiplication of A by thescalar k is defined as

KA= (Ka_{ij}) for all i, j.

In other words, to multiply a matrix A by a scalar K, multiply every element of A by K.

Negative of a matrix

The negative of a matrix A = $(a_{ij})_{mxn}$ is defined by - A = $(-a_{ij})_{mxn}$ for all i, j and is obtained by changing the sign of every element.

For example

If A= $\begin{pmatrix} 2 & -5 & 7 \\ 0 & 5 & 6 \end{pmatrix}$ then

$$-A = \begin{pmatrix} -2 & 5 & -7 \\ -0 & -5 & -6 \end{pmatrix}$$

Equality of matrices

Two matrices are said to equal when

I. They have the same order and
II. The corresponding elements are equal.

Addition of matrices

Addition of matrices is possible only when they are of same order (i.e., conformal for addition). When two matrices A and B are of same order, then their sum (A+B) is obtained by adding the corresponding elements in both the matrices.

Properties of matrix addition

Let A, B, C be matrices of the same order. The addition of matrices obeys the following

I. Commutative law: $\mathbf{A + B = B + A}$

II. Associative law: $\mathbf{A + (B + C)= (A + B) + C}$

III. Distributive law: $\mathbf{K (A+B) = KA+KB}$, where k is scalar.

Subtraction of matrices

Subtraction of matrices is also possible only when they are of same order. Let A and B be the two matrices of the same order. The matrix A – B is obtained by subtracting the elements of B from the corresponding elements of A.

Multiplication of matrices

Multiplication of two matrices is possible only when the number of columns of the first matrix is equal to the number of rows of the second matrix (i.e. conformable for multiplication)

Let $A = (a_{ij})$ be an m x p matrix, and let

$B = (b_{ij})$ be an p x n matrix.

Then the product AB is a matrix $C = (c_{ij})$ of order m x n,

Properties of matrix multiplication

I. Matrix Multiplication is not commutative i.e. for the two matrices A and B, generally $\mathbf{AB = BA}$.

II. The Multiplication of Matrices is associative i.e., $\mathbf{(AB) C = A(BC)}$

III. Matrix Multiplication is distributive with respect to addition. i.e. if, A, B, C are matrices oforder mxn, n x k, and n x k respectively, then $\mathbf{A(B+C) = AB + AC}$

IV. Let A be a square matrix of order n and I is the unit matrix of same order. then $\mathbf{AI = A = I\,A}$

V. The product $\mathbf{AB = O}$ (Null matrix), does not imply that either A = 0 or B = 0 or both are zero.

Transpose of a matrix

Let $A = (a_{ij})$ be a matrix of order m x n. The transpose of A, denoted by A^T of order n x m is obtained by interchanging rows into columns of A.

For example

If $A = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix}_{2\times3}$

then

$$A^T = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 6 \end{pmatrix}$$

Properties of Matrix Transposition

Let A^T and B^T are the transposed Matrices of A and B and a is a scalar. Then

a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. $(a A)^T = a A^T$
d. $(AB)^T = B^T A^T$ (A and B are conformable for multiplication)

DETERMINANTS

An important attribute in the study of Matrix Algebra is the concept of Determinant, ascribed to a square matrix. Knowledge of Determinant theory is indispensable in the study of Matrix Algebra.

The determinant associated with each square matrix A = (aij) is a scalar and denoted by the symbol det. A or $|A|$ The scalar may be real or complex number, positive, Negative or Zero. A matrix is an array and has no numerical value, but a determinant has numerical value.

Example 1

Evaluate $\begin{vmatrix} 2 & 0 & 4 \\ 5 & -1 & 1 \\ 9 & 7 & 8 \end{vmatrix}$

Solution:

$$\begin{vmatrix} 2 & 0 & 4 \\ 5 & -1 & 1 \\ 9 & 7 & 8 \end{vmatrix} = 2\begin{vmatrix} -1 & 1 \\ 7 & 8 \end{vmatrix} - 0\begin{vmatrix} 5 & 1 \\ 9 & 8 \end{vmatrix} + 4\begin{vmatrix} 5 & -1 \\ 9 & 7 \end{vmatrix}$$

$= 2(-1\times8-1\times7) - 0\ (5\times 8 - 9 \times 1) + 4\ (5\times 7 - (-1) \times 9)$

=2(-8-7) – 0 (40-9) + 4 (35+9)

=-30-0+176 = 146

<u>Properties of Determinants</u>

1. The value of determinant is unaltered, when its rows and columns are interchanged.
2. If any two rows (columns) of a determinant are interchanged, then the value of the determinant changes only in sign.

3. If the determinant has two identical rows (columns), then the value of the determinant is zero.
4. If all the elements in a row or in a (column) of a determinant are multiplied by a constant k(k, $\neq$ 0) then the value of the determinant is multiplied by k.
5. The value of the determinant is unaltered when a constant multiple of the elements of any row (column), is added to the corresponding elements of a different row (column) in a determinant.
6. If each element of a row (column) of a determinant is expressed as the sum of two or more terms, then the determinant is expressed as the sum of two or more determinants of the same order.
7. If any two rows or columns of a determinant are proportional, then the value of the determinant is zero.

Singular Matrix

A square matrix A is said to be singular if det. A = 0, otherwise it is a non-singular matrix.

Example 2

Show that $\begin{matrix} \mathbf{1} & \mathbf{2} \\ \mathbf{2} & \mathbf{4} \end{matrix}$

Solution:

$\begin{matrix} 1 & 2 \\ 2 & 4 \end{matrix} = (1\times 4) – (2\times 2)$

= 4 – 4

= 0

The matrix is singular

Example 3

Show that 2 5 is a non-singular matrix 9

10

Solution:

2 5

9 10 = (2× 10) – (5× 9)

= 20 – 45

= -25 ≠ 0

The given matrix is non singular

INVERSE OF A MATRIX

Minors and Cofactors of the elements of a determinant.

The minor of an element a_{ij} of a determinant A is denoted by M_{ij} and is the determinant obtained from A by deleting the row and the column where a_{ij} occurs.

The cofactor of an element a_{ij} with minor M_{ij} is denoted by C_{ij} and is defined as

$Mi\ j$, $if\ \ i + j\ is\ even$

$Cij\ \ =\ \ Mi\ j\ \ \ , if\ \ i + j\ is\ odd$

Thus, Cofactors are signed minors

In the case of □11 a12 we have

a21 a22

M11 = a22 ; M12 = a21 , M21 = a12 , M22 = a11 Also

C11 = a22, C12 = - a21, C21 = - a12 , C22 = a11

a11 a12 a13

In the case of a21 a22 a23 we have

a31 a32 a33

= □22

a32

M11 a23 C11 = a22 a23

a33 a32 a33

M12 = a23 C12 = a21 a23

□21 a33 a31 a33

a31

$M_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$ $\quad C_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$

$M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$ $\quad C_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$

Adjoint

The transpose of the matrix got by replacing all the elements of a square matrix A by their corresponding cofactors in | A | is called the **Adjoint** of A or **Adjugate** of A and is denoted by Adj A.

Thus, AdjA = $A^T{}_c$

Note

(i) Let A = $\begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{pmatrix}$ then Ac = $\begin{pmatrix} \boldsymbol{d} & -\boldsymbol{c} \\ -\boldsymbol{b} & \boldsymbol{a} \end{pmatrix}$

Adj A = $A^T{}_c$ = $\begin{pmatrix} \boldsymbol{d} & -\boldsymbol{b} \\ -\boldsymbol{c} & \boldsymbol{a} \end{pmatrix}$

Thus the Adjoint of a 2 x 2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ can be written as $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

(ii) Adj I = I, where I is the unit matrix.

(iii) A(AdjA) = (Adj A) A = | A | I

(iv) Adj (AB) = (Adj B) (Adj A)

(v) If A is a square matrix of order 2, then |AdjA| = |A| If A is a square matrix of order 3, then |Adj A| = $|A|^2$

Example 1

Write the Adjoint of the matrix A = $\begin{pmatrix} 4 & -2 \\ 3 & 4 \end{pmatrix}$

Solution

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

Example 2

Find the Adjoint of the matrix A = $\begin{pmatrix} \mathbf{0} & \mathbf{1} & \mathbf{2} \\ \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{3} & \mathbf{1} & \mathbf{1} \end{pmatrix}$

Solution

0 1 2

A = 1 2 3 = Adj A = A^T c

3 1 1

C11 = 2 3 = −1, C12 = − 1 3 = 8, C13 = 1 2 =

−5

1 1 3 1 3 1

— 1 2 0 2 0 1

C21 = = 1, $C22$ = = −6 C23 = − = 3,

1 1 3 1 3 1

C31 = 1 2 = −1, ` $C32$ =0 2 = 2 ` $C33$= 0 1 = −

− 1

2 3 1 3 1 2

∴ Ac = −1 8 −5

1 −63

−1 2 −1

Hence

Adj −1 8 −5 T -1 1 -1

A=

1 − 3 = 8 -6 2

6

−1 2 −1 -5 3 -1

Inverse of a non singular matrix:

The inverse of a non singular matrix A is the matrix B such that AB = BA = I. B is then called the inverse of A and denoted by A^{-1}.

Note

1. A non square matrix has no inverse.
2. The inverse of a square matrix A exists only when $|A| \neq 0$ that is, if A is a singular matrix then A^{-1} does not exist.
3. If B is the inverse of A then A is the inverse of B. That is $B = A^{-1}$ and $A = B^{-1}$.4.

$A A^{-1} = I = A^{-1} A$

5. The inverse of a matrix, if it exists, is unique. That is, no matrix can have more than one inverse.
6. The order of the matrix A^{-1} will be the same as that of A.7. I^{-1}

= I

8. $(AB)^{-1} = B^{-1} A^{-1}$, provided the inverses exist.

9. $A^2 = I$ implies $A^{-1} = A$

10. If AB = C then

11. $A = CB^{-1}$ (b) $B = A^{-1}C$, provided the inverses exist.

12. We have seen that A(AdjA) = (AdjA)A = |A| I

∴ A 1/ (Adj A) = 1/ *A* (Adj A) A= I

This suggests that $A^{-1} = 1 / A$ (Adj A). That is $A^{-1} = 1 / A\ A_c^t$

$(A^{-1})^{-1} = A$, provided the inverse exists.

Problem 1

Find the inverse of A= 5 3 if it exists.

4 2

Solution

$A =$ 5 3 = (5× 2) − (3 × 4) = 10-12 = −2

4 2

∴ A^{-1} exists.

Problem 2

Show that the inverse of the following do not exist:

(i) A = −2 6 3 −9

3 1 −2

(ii) A= 2 7 3

6 2 −4

Solution:

(i) $A =$ −2 6 = 0 ∴ A^{-1} does not exist.

3 −9

3 1 −2

(ii) $A =$ 2 7 3 = 0 ∴ A^{-1} does not exist.

6 2 −4

Problem 3

2 3 4

Find the inverse of A 3 2 1 if it exists

=

1 1 −2

Solution

2 3 4

$A =$ 3 2 1 = 15 ≠ 0 ∴ A-exists.

1 1 −2 1

We have $A^{-1} = \underline{1}$ $A^t{}_c$

A

Now the cofactors are

C11 = 1 = −5, C12 = − 3 1 = 7, C13 = 3 2 = 1

2

1 −2 1 −2 1 1

C21 = -4 = 10, C22 = −2 4 = -8, C23 = -2 3 = 1

3

1 −2 1 −2 1 1

C31 = 3 4 = ` $C$32 = − 2 4 = 10 ` C33 = 2 3 = − 5

−5,

2 1 3 1 2 2

Hence

$$A_C = \begin{matrix} -5 & 7 & 1 \\ 10 & -8 & 1 \\ -5 & 10 & -5 \end{matrix}$$

$$A^t{}_c = \begin{matrix} -5 & 10 & -5 \\ 7 & -8 & 10 \\ 1 & 1 & -5 \end{matrix}$$

$\therefore A^{-1} = \underline{1}$ $A^t A_c$

$$A^{-1} = \frac{1}{15} \begin{matrix} -5 & 10 & -5 \\ 7 & -8 & 10 \\ 1 & 1 & -5 \end{matrix}$$

- **If the value of the determinant of a square matrix is not equal to zero, then it is a non-singular matrix.**
- **If it is a non-singular matrix, then inverse exists.**
- **A^{-1}= $\frac{1}{}$ adjoint of A**
- **$A^{-1}B=X$**

Example

Solve the following equation by matrix inverse method x+2y=6
3x+4y=16

Solution

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 16 \end{bmatrix}$$

A × X= B X= $A^{-1}B$

$A = 1\times \mathbf{4} - \mathbf{3} \times \mathbf{2} = \mathbf{4} - \mathbf{6} = -\mathbf{2}$

$A^{-1} = \frac{1}{} A^t A_c$

Adj A = $\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$

A-1 = $\frac{-1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$

A-1B= $\frac{-1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \text{X} \begin{bmatrix} 6 \\ 16 \end{bmatrix}$

$= \frac{-1}{2}\begin{bmatrix} 4 \times 6 + -2 \times 16 \\ -3 \times 6 + 1 \times 16 \end{bmatrix}$

$= \frac{-1}{} \quad 24 - 32$

2 −18 + 16

= - 1 -8

2 -2

X= 4

1

X = 4

Y 1

x= 4 y=1

Types of Matrices

Depending upon the order and elements, matrices are classified as:

- Column matrix
- Row matrix
- Square matrix
- Diagonal matrix
- Scalar matrix
- Identity matrix
- Zero matrix

Type of matrix	**Definition and Example**
Column matrix	A column matrix is an m × 1 matrix, consisting of a single column of m elements. It is also called a column vector. Example: [41−5]
Row matrix	A row matrix is a 1 × m matrix, consisting of a single row of m elements. It is also called a row vector. Example: [2−10]
Square matrix	A matrix that has an equal number of rows and columns. It is expressed as m × m. Example: Square matrix of order 2 is [18−31]. Square matrix of order 3 is [1−1−4812031].

Diagonal matrix	A square matrix that has non-zero elements in its diagonal part running from the upper left to the lower right or vice versa. Example: [9000−40006]
Scalar matrix	The scalar matrix is a square matrix, which has all its diagonal elements equal and all the off-diagonal elements as zero. Example: [140001400014]
Identity matrix	A square matrix that has all its principal diagonal elements as 1's and all non-diagonal elements as zeros. Example: Identity (Unit) matrix of order 2 is [1001]. Identity matrix of order 3 is [100010001] .
Zero matrix	A matrix whose all entries are zero. It is also called a null matrix. Example: [000000]

Equality of Matrices

Two matrices are said to be equal if-

(i) The order of both the matrices is the same

(ii) Each element of one matrix is equal to the corresponding element of the other matrix

Operations on Matrices

In Chapter 3 of Class 12 Matrices, certain operations on matrices are discussed, namely, the addition of matrices, multiplication of a matrix by a scalar, difference and multiplication of matrices.

Transpose of a Matrix

If $A = [a_{ij}]$ be an $m \times n$ matrix, then the matrix obtained by interchanging the rows and columns of A is called the transpose of A and is denoted by A' or (A^T).

In other words, if $A = [a_{ij}]\ m \times n$, then $A' = [a_{ji}]\ n \times m$.

Symmetric and Skew Symmetric Matrices

A square matrix A = $[a_{ij}]$ is said to be symmetric if the transpose of A is equal to A, that is, $[a_{ij}] = [a_{ji}]$ for all possible values of i and j.

A square matrix A = $[a_{ij}]$ is a skew-symmetric matrix if A′ = – A, that is $a_{ji} = -a_{ij}$ for all possible values of i and j. Also, if we substitute i = j, we have $a_{ii} = -a_{ii}$ and thus, $2a_{ii} = 0$ or $a_{ii} = 0$ for all i's. Therefore, all the diagonal elements of a skew-symmetric matrix are zero.

To understand the symmetric and skew-symmetric matrix in detail, visit here.

Elementary Operation (Transformation) of a Matrix

There are six operations (transformations) on a matrix, three of which are due to rows, and three are due to columns, known as elementary operations or transformations.

1. The interchange of any two rows or two columns.
2. The multiplication of the elements of any row or column by a non zero number.
3. The addition to the elements of any row or column, the corresponding elements of any other row or column are multiplied by any non zero number.

Learn more about the elementary operations of the matrix here.

Invertible Matrices

Suppose a square matrix A of order m, and if there exists another square matrix B of the same order m, such that AB = BA = I, then B is called the inverse matrix of A, and it is denoted by A^{-1}. Also, matrix A is said to be an invertible matrix here.

Applications of Matrices and Determinants

(i) Matrix Inversion Method

Example

Solve the following system of linear equations, using matrix inversion method:

$5x + 2y = 3$, $3x + 2y = 5$.

Solution

The matrix form of the system is AX = B , where $A = \begin{bmatrix} 5 & 2 \\ 3 & 2 \end{bmatrix}, X = \begin{bmatrix} x \\ y \end{bmatrix}, B = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

We find |A| = $\begin{vmatrix} 5 & 2 \\ 3 & 2 \end{vmatrix}$ = 10 - 6= 4 ≠ 0. So, A^{-1} exists and $A^{-1} = \frac{1}{4}\begin{bmatrix} 2 & -2 \\ -3 & 5 \end{bmatrix}$.

Then, applying the formula X = A−1B , we get

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 2 & -2 \\ -3 & 5 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 6-10 \\ -9+25 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} -4 \\ 16 \end{bmatrix} = \begin{bmatrix} \frac{-4}{4} \\ \frac{16}{4} \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}.$$

So the solution is (x = −1, y = 4).

Example

Solve the following system of equations, using matrix inversion method:

2x1 + 3x2 + 3x3 = 5,

x1 – 2x2 + x3 = -4,

3x1 – x2 – 2x3 = 3

Solution

The matrix form of the system is AX = B,where

$$A = \begin{bmatrix} 2 & 3 & 3 \\ 1 & -2 & 1 \\ 3 & -1 & -2 \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, B = \begin{bmatrix} 5 \\ -4 \\ 3 \end{bmatrix}.$$

We find $|A| = \begin{vmatrix} 2 & 3 & 3 \\ 1 & -2 & 1 \\ 3 & -1 & -2 \end{vmatrix} = 2(4+1)-3(-2-3)+3(-1+6) = 10+15+15 = 40 \neq 0$.

So, A^{-1} exists and

$$A^{-1} = \frac{1}{|A|}(\text{adj } A) = \frac{1}{40}\begin{bmatrix} +(4+1) & -(-2-3) & +(-1+6) \\ -(-6+3) & +(-4-9) & -(-2-9) \\ +(3+6) & -(2-3) & +(-4-3) \end{bmatrix}^T = \frac{1}{40}\begin{bmatrix} 5 & 3 & 9 \\ 5 & -13 & 1 \\ 5 & 11 & -7 \end{bmatrix}$$

Then, applying $X = A^{-1}B$, we get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{40}\begin{bmatrix} 5 & 3 & 9 \\ 5 & -13 & 1 \\ 5 & 11 & -7 \end{bmatrix}\begin{bmatrix} 5 \\ -4 \\ 3 \end{bmatrix} = \frac{1}{40}\begin{bmatrix} 25-12+27 \\ 25+52+3 \\ 25-44-21 \end{bmatrix} = \frac{1}{40}\begin{bmatrix} 40 \\ 80 \\ -40 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

So, the solution is (x1 = 1, x2 = 2, x3 = −1) .

Example

If $A = \begin{bmatrix} -4 & 4 & 4 \\ -7 & 1 & 3 \\ 5 & -3 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}$, find the products AB and BA and hence solve the system of equations x − y + z = 4, x – 2y – 2z = 9, 2x + y +3z =1.

Solution

$$\text{We find } AB = \begin{bmatrix} -4 & 4 & 4 \\ -7 & 1 & 3 \\ 5 & -3 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix} = \begin{bmatrix} -4+4+8 & 4-8+4 & -4-8+12 \\ -7+1+6 & 7-2+3 & -7-2+9 \\ 5-3-2 & -5+6-1 & 5+6-3 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix} = 8I_3$$

$$\text{and } BA = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} -4 & 4 & 4 \\ -7 & 1 & 3 \\ 5 & -3 & -1 \end{bmatrix} = \begin{bmatrix} -4+7+5 & 4-1-3 & 4-3-1 \\ -4+14-10 & 4-2+6 & 4-6+2 \\ -8-7+15 & 8+1-9 & 8+3-3 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix} = 8I_3.$$

$$\text{So, we get } AB = BA = 8I_3. \text{That is, } \left(\frac{1}{8}A\right)B = B\left(\frac{1}{8}A\right) = I_3. \text{Hence, } B^{-1} = \frac{1}{8}A.$$

Writing the given system of equations in matrix form, we get

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \\ 1 \end{bmatrix}. \text{ That is, } B\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \\ 1 \end{bmatrix}.$$

$$\text{So, } \begin{bmatrix} x \\ y \\ z \end{bmatrix} = B^{-1}\begin{bmatrix} 4 \\ 9 \\ 1 \end{bmatrix} = \left(\frac{1}{8}A\right)\begin{bmatrix} 4 \\ 9 \\ 1 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} -4 & 4 & 4 \\ -7 & 1 & 3 \\ 5 & -3 & -1 \end{bmatrix}\begin{bmatrix} 4 \\ 9 \\ 1 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} -16+36+4 \\ -28+9+3 \\ 20-27-1 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} 24 \\ -16 \\ -8 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix}$$

Hence, the solution is (x = 3, y = - 2, z = −1).

(ii) Cramer's Rule

Example

Solve, by Cramer's rule, the system of equations

$x1 - x2 = 3$, $2x1 + 3x2 + 4x3 = 17$, $x2 + 2x3 = 7$.

Solution

First we evaluate the determinants

$$\Delta = \begin{vmatrix} 1 & -1 & 0 \\ 2 & 3 & 4 \\ 0 & 1 & 2 \end{vmatrix} = 6 \neq 0, \Delta_1 = \begin{vmatrix} 3 & -1 & 0 \\ 17 & 3 & 4 \\ 7 & 1 & 2 \end{vmatrix} = 12, \Delta_2 = \begin{vmatrix} 1 & 3 & 0 \\ 2 & 17 & 4 \\ 0 & 7 & 2 \end{vmatrix} = -6, \Delta_3 = \begin{vmatrix} 1 & -1 & 3 \\ 2 & 3 & 17 \\ 0 & 1 & 7 \end{vmatrix} = 24.$$

By Cramer's rule, we get $x_1 = \frac{\Delta_1}{\Delta} = \frac{12}{6} = 2,\ x_2 = \frac{\Delta_2}{\Delta} = \frac{-6}{6} = -1,\ x_3 = \frac{24}{6} = 4.$

So, the solution is ($x1 = 2$, $x2 = - 1$, $x3 = 4$).

Example

In a T20 match, Chennai Super Kings needed just 6 runs to win with 1 ball left to go in the last over. The last ball was bowled and the batsman at the crease hit it high up. The ball traversed along a path in a vertical plane and the equation of the path is $y = ax2 + bx + c$ with respect to a xy -coordinate system in the vertical plane and the ball traversed through the points (10,8), (20,16), (30,18) , can you conclude that Chennai Super Kings won the match?

Justify your answer. (All distances are measured in metres and the meeting point of the plane of the path with the farthest boundary line is (70, 0).)

Solution

The path $y = ax2 + bx + c$ passes through the points (10,8), (20,16), (40, 22) . So, we get the system of equations $100a + 10b + c = 8, 400a + 20b + c = 16, 1600a + 40b + c = 22$. To apply Cramer's rule, we find

$$\Delta = \begin{vmatrix} 100 & 10 & 1 \\ 400 & 20 & 1 \\ 1600 & 40 & 1 \end{vmatrix} = 1000\begin{vmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 16 & 4 & 1 \end{vmatrix} = 1000[-2+12-16] = -6000,$$

$$\Delta_1 = \begin{vmatrix} 8 & 10 & 1 \\ 16 & 20 & 1 \\ 22 & 40 & 1 \end{vmatrix} = 20\begin{vmatrix} 4 & 1 & 1 \\ 8 & 2 & 1 \\ 11 & 4 & 1 \end{vmatrix} = 20[-8+3+10] = 100,$$

$$\Delta_2 = \begin{vmatrix} 100 & 8 & 1 \\ 400 & 16 & 1 \\ 1600 & 22 & 1 \end{vmatrix} = 200\begin{vmatrix} 1 & 4 & 1 \\ 4 & 8 & 1 \\ 16 & 11 & 1 \end{vmatrix} = 200[-3+48-84] = -7800,$$

$$\Delta_3 = \begin{vmatrix} 100 & 10 & 8 \\ 400 & 20 & 16 \\ 1600 & 40 & 22 \end{vmatrix} = 2000\begin{vmatrix} 1 & 1 & 4 \\ 4 & 2 & 8 \\ 16 & 4 & 11 \end{vmatrix} = 2000[-10+84-64] = 20000.$$

By Cramer's rule, we get $a = \frac{\Delta_1}{\Delta} = -\frac{1}{60}, b = \frac{\Delta_2}{\Delta} = \frac{7800}{6000} = \frac{78}{60} = \frac{13}{10}, c = \frac{\Delta_3}{\Delta} = -\frac{20000}{6000} = -\frac{20}{6} = -\frac{10}{3}$.

So, the equation of the path is $y = -\frac{1}{60}x^2 + \frac{13}{10}x - \frac{10}{3}$.

When $x = 70$, we get $y = 6$.

So, the ball went by 6 metres high over the boundary line and it is impossible for a fielder standing even just before theboundary line to jump and catch the ball.

Hence the ball went for a super six and the Chennai Super Kings won the match

(iii) Gaussian Elimination Method

Example

Solve the following system of linear equations, by Gaussian elimination method :

$4x + 3y + 6z = 25$, $x + 5y + 7z = 13$, $2x + 9y + z = 1$.

Solution

Transforming the augmented matrix to echelon form, we get

$$\left[\begin{array}{ccc|c} 4 & 3 & 6 & 25 \\ 1 & 5 & 7 & 13 \\ 2 & 9 & 1 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 5 & 7 & 13 \\ 4 & 3 & 6 & 25 \\ 2 & 9 & 1 & 1 \end{array}\right] \xrightarrow[R_3 \to R_3 - 2R_1]{R_2 \to R_2 - 4R_1,} \left[\begin{array}{ccc|c} 1 & 5 & 7 & 13 \\ 0 & -17 & -22 & -27 \\ 0 & -1 & -13 & -25 \end{array}\right]$$

$$\xrightarrow[R_3 \to R_3 \div (-1)]{R_2 \to R_2 \div (-1),} \left[\begin{array}{ccc|c} 1 & 5 & 7 & 13 \\ 0 & 17 & 22 & 27 \\ 0 & 1 & 13 & 25 \end{array}\right] \xrightarrow{R_3 \to 17R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 5 & 7 & 13 \\ 0 & 17 & 22 & 27 \\ 0 & 0 & 199 & 398 \end{array}\right].$$

The equivalent system is written by using the echelon form:

x + 5y + 7z = 13 , … (1)

17y + 22z = 27 , … (2)

199z = 398 . … (3)

From (3), we get $z = \frac{398}{199} = 2$.

Substituting $z = 2$ in (2), we get $y = \frac{27 - 22 \times 2}{17} = \frac{-17}{17} = -1$.

Substituting z = 2, y = -1 in (1), we get x = 13 - 5 $\times (-1) - 7 \times 2 = 4$.

So, the solution is (x =4, y = - 1, z = 2).

Note. The above method of going from the last equation to the first equation is called the **method of back substitution**.

Example

The upward speed *v*(*t*) of a rocket at time t is approximated by v(t) = at2 + bt + c, 0 ≤ t ≤ 100 where a, b, and c are constants. It has been found that the speed at times t = 3, t = 6 , and t = 9 seconds are respectively, 64, 133, and 208 miles per second respectively. Find the speed at time t = 15 seconds. (Use Gaussian elimination method.)

Solution

Since *v*(3) =64, *v*(6) = 133 and *v*(9) = 208 , we get the following system of linear equations

9a +3b + c = 64 ,

36a + 6b + c = 133,

81a + 9b + c = 208 .

We solve the above system of linear equations by Gaussian elimination method.

Reducing the augmented matrix to an equivalent row-echelon form by using elementary row operations, we get

$$[A \mid B] = \left[\begin{array}{ccc|c} 9 & 3 & 1 & 64 \\ 36 & 6 & 1 & 133 \\ 81 & 9 & 1 & 208 \end{array}\right] \xrightarrow{R_2 \to R_2 - 4R_1, R_3 \to R_3 - 9R_1} \left[\begin{array}{ccc|c} 9 & 3 & 1 & 64 \\ 0 & -6 & -3 & -123 \\ 0 & -18 & -8 & -368 \end{array}\right] \xrightarrow{R_2 \to R_2 \div (-3), R_3 \to R_3 \div (-2)} \left[\begin{array}{ccc|c} 9 & 3 & 1 & 64 \\ 0 & 2 & 1 & 41 \\ 0 & 9 & 4 & 184 \end{array}\right]$$

$$\xrightarrow{R_3 \to 2R_3} \left[\begin{array}{ccc|c} 9 & 3 & 1 & 64 \\ 0 & 2 & 1 & 41 \\ 0 & 18 & 8 & 368 \end{array}\right] \xrightarrow{R_3 \to R_3 - 9R_2} \left[\begin{array}{ccc|c} 9 & 3 & 1 & 64 \\ 0 & 2 & 1 & 41 \\ 0 & 0 & -1 & -1 \end{array}\right] \xrightarrow{R_3 \to (-1)R_3} \left[\begin{array}{ccc|c} 9 & 3 & 1 & 64 \\ 0 & 2 & 1 & 41 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Writing the equivalent equations from the row-echelon matrix, we get

9a + 3b + c = 64, 2b + c = 41, c= 1.

By back substitution, we get

$$\text{we get } c = 1,\ b = \frac{(41-c)}{2} = \frac{(41-1)}{2} = 20,\ a = \frac{64-3b-c}{9} = \frac{64-60-1}{9} = \frac{1}{3}.$$

So, we get v (t) = 1/3 t2 + 20t + 1.

Hence, $v(15)$ = 1/3 (225) + 20(15) + 1 = 75 + 300 + 1 = 376

<u>Matrix: Non-homogeneous Linear Equations</u>

Applications of Matrices: Consistency of System of Linear Equations by Rank Method

In second previous section, we have already defined consistency of a system of linear equation. In this section, we investigate it by using rank method. We state the following theorem without proof:

Theorem 1.14 (Rouché - Capelli Theorem)

A system of linear equations, written in the matrix form as $AX = B$, is consistent if and only if the rank of the coefficient matrix is equal to the rank of the augmented matrix; that is, $\rho(A) = \rho([A \mid B])$.

We apply the theorem in the following examples.

Non-homogeneous Linear Equations

Example

Test for consistency of the following system of linear equations and if possible solve:

$x + 2y - z = 3,\ 3x - y + 2z = 1,\ x - 2y + 3z = 3,\ x - y + z + 1 = 0$.

Solution

Here the number of unknowns is 3.

The matrix form of the system is AX = B, where

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & -1 & 2 \\ 1 & -2 & 3 \\ 1 & -1 & 1 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, B = \begin{bmatrix} 3 \\ 1 \\ 3 \\ -1 \end{bmatrix}.$$

The augmented matrix is $[A \mid B] = \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 3 & -1 & 2 & 1 \\ 1 & -2 & 3 & 3 \\ 1 & -1 & 1 & -1 \end{array}\right]$.

Applying Gaussian elimination method on [A | B], we get

$$[A \mid B] \xrightarrow[\substack{R_3 \to R_3 - R_1, \\ R_4 \to R_4 - R_1,}]{R_2 \to R_2 - 3R_1,} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & -7 & 5 & -8 \\ 0 & -4 & 4 & 0 \\ 0 & -3 & 2 & -4 \end{array}\right] \xrightarrow[\substack{R_3 \to (-1)R_3, \\ R_4 \to (-1)R_4}]{R_2 \to (-1)R_2,} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 7 & -5 & 8 \\ 0 & 4 & -4 & 0 \\ 0 & 3 & -2 & 4 \end{array}\right]$$

$$\xrightarrow[R_4 \to 7R_4 - 3R_2]{R_3 \to 7R_3 - 4R_2,} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 7 & -5 & 8 \\ 0 & 0 & -8 & -32 \\ 0 & 0 & 1 & 4 \end{array}\right] \xrightarrow{R_3 \to R_3 \div (-8)} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 7 & -5 & 8 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 4 \end{array}\right] \xrightarrow{R_4 \to R_4 - R_3} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 7 & -5 & 8 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

There are three non-zero rows in the row-echelon form of [A | B].So, ρ ([A | B]) . = 3

So, the row-echelon form of A is $\begin{bmatrix} 1 & 2 & -1 \\ 0 & 7 & -5 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. There are three non-zero rows in it. So ρ(A) = 3.

Hence, ρ(A) = ρ([A | B]) = 3.

From the echelon form, we write the equivalent system of equations

x + 2y –z =3, 7y-5z = 8, z=4, 0=0.

The last equation 0 = 0 is meaningful. By the method of back substitution, we get

z = 4

$7y - 20 = 8 \Rightarrow y = 4$,

$x = 3 - 8 + 4 \Rightarrow x = -1$.

So, the solution is $(x = -1, y = 4, z = 4)$.(Note that A is not a square matrix.)

Here the given system is consistent and the solution is unique.

Example

Test for consistency of the following system of linear equations and if possible solve:

$4x - 2y + 6z = 8, x + y - 3z = -1, 15x - 3y + 9z = 21.$

Solution

Here the number of unknowns is 3.

The matrix form of the system is $AX = B$, where

$$A = \begin{bmatrix} 4 & -2 & 6 \\ 1 & 1 & -3 \\ 15 & -3 & 9 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, B = \begin{bmatrix} 8 \\ -1 \\ 21 \end{bmatrix}.$$

Applying elementary row operations on the augmented matrix$[A \mid B]$, we get

$$[A \mid B] = \left[\begin{array}{ccc|c} 4 & -2 & 6 & 8 \\ 1 & 1 & -3 & -1 \\ 15 & -3 & 9 & 21 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 4 & -2 & 6 & 8 \\ 15 & -3 & 9 & 21 \end{array}\right] \xrightarrow[R_3 \to R_3 - 15R_1]{R_2 \to R_2 - 4R_1,} \left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & -6 & 18 & 12 \\ 0 & -18 & 54 & 36 \end{array}\right]$$

$$\xrightarrow[R_3 \to R_3 \div (-18)]{R_2 \to R_2 \div (-6),} \left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & 1 & -3 & -2 \\ 0 & 1 & -3 & -2 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

So, $\rho(A) = \rho([A \mid B]) = 2 < 3$. From the echelon form, we get the equivalent equations

$x + y - 3z = -1, y - 3z = -2, 0 = 0$.

The equivalent system has two non-trivial equations and three unknowns. So, one of the unknowns should be fixed at our choice in order to get two equations for the other two unknowns. We fix z arbitrarily as a real number t, and we get $y = 3t - 2, x = -1 - (3t - 2) + 3t = 1$. So, the solution is $(x = 1, y = 3t - 2, z = t)$, where t is real . The above solution set is a one-parameter family of solutions.

Here, the given system is consistent and has infinitely many solutions which form a one parameter family of solutions.

Note

In the above example, the square matrix A is singular and so matrix inversion method cannot be applied to solve the system of equations. However,Gaussian elimination method is applicable and we are able to decide whether the system is consistent or not. The next example also confirms the supremacy of Gaussian elimination method over other methods.

Example

Test for consistency of the following system of linear equations and if possible solve:

$x - y + z = -9$, $2x - 2y + 2z = -18$, $3x - 3y + 3z + 27 = 0$.

Solution

Here the number of unknowns is 3.

The matrix form of the system is $AX = B$, where

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ 3 & -3 & 3 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, B = \begin{bmatrix} -9 \\ -18 \\ -27 \end{bmatrix}.$$

Applying elementary row operations on the augmented matrix$[A \mid B]$, we get

$$[A \mid B] = \left[\begin{array}{ccc|c} 1 & -1 & 1 & -9 \\ 2 & -2 & 2 & -18 \\ 3 & -3 & 3 & -27 \end{array}\right] \xrightarrow[R_3 \to R_3 - 3R_1]{R_2 \to R_2 - 2R_1,} \left[\begin{array}{ccc|c} 1 & -1 & 1 & -9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

So, $\rho(A) = \rho([A \mid B]) = 1 < 3$.

From the echelon form, we get the equivalent equations $x - y + z = -9$, $0 = 0$, $0 = 0$.

The equivalent system has one non-trivial equation and three unknowns.

Taking $y = s, z = t$ arbitrarily, we get $x - s + t = -9$; or $x = -9 + s - t$.

So, the solution is $(x = -9 + s - t, y = s, z = t)$, where s and t are parameters.

The above solution set is a two-parameter family of solutions.

Here, the given system of equations is consistent and has infinitely many solutions which form a two parameter family of solutions.

Example

Test the consistency of the following system of linear equations

$x - y + z = -9$, $2x - y + z = 4$, $3x - y + z = 6$, $4x - y + 2z = 7$.

Solution

Here the number of unknowns is 3.

The matrix form of the system of equations is $AX = B$, where

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 1 \\ 3 & -1 & 1 \\ 4 & -1 & 2 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, B = \begin{bmatrix} -9 \\ 4 \\ 6 \\ 7 \end{bmatrix}$$

Applying elementary row operations on the augmented matrix [A|B], we get

$$[A\,|\,B] = \left[\begin{array}{ccc|c} 1 & -1 & 1 & -9 \\ 2 & -1 & 1 & 4 \\ 3 & -1 & 1 & 6 \\ 4 & -1 & 2 & 7 \end{array}\right] \xrightarrow{\substack{R_2 \to R_2 - 2R_1, \\ R_3 \to R_3 - 3R_1, \\ R_4 \to R_4 - 4R_1}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & -9 \\ 0 & 1 & -1 & 22 \\ 0 & 2 & -2 & 33 \\ 0 & 3 & -2 & 43 \end{array}\right]$$

$$\xrightarrow{\substack{R_3 \to R_3 - 2R_2, \\ R_4 \to R_4 - 3R_2}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & -9 \\ 0 & 1 & -1 & 22 \\ 0 & 0 & 0 & -11 \\ 0 & 0 & 1 & -23 \end{array}\right] \xrightarrow{R_3 \leftrightarrow R_4} \left[\begin{array}{ccc|c} 1 & -1 & 1 & -9 \\ 0 & 1 & -1 & 22 \\ 0 & 0 & 1 & -23 \\ 0 & 0 & 0 & -11 \end{array}\right]$$

So, $\rho(A) = 3$ and $\rho([A\,|\,B]) = 4$. Hence $\rho(A) \neq \rho([A\,|\,B])$.

If we write the equivalent system of equations using the echelon form, we get

$x - y + z = -9$, $y - z = 22$, $z = -23$, $0 = -11$.

The last equation is a contradiction.

So the given system of equations is inconsistent and has no solution. By Rouché-Capelli theorem, we have the following rule:

- If there are n unknowns in the system of equations and $\rho(A) = \rho([A\,|\,B]) = n$, then the system $AX = B$, is consistent and has a unique solution.
- If there are n unknowns in the system $AX = B$, and $\rho(A) = \rho([A\,|\,B]) = n - k$, $k \neq 0$ then the system is consistent and has infinitely many solutions and these solutions form a k - parameter family. In particular, if there are 3 unknowns in a system of equations and $\rho(A) = \rho([A\,|\,B]) = 2$, then the system has infinitely many solutions and these solutions form a one parameter family. In the same manner, if there are 3unknowns in a system of equations and $\rho(A) = \rho([A\,|\,B]) = 1$, then the system has infinitely many solutions and these solutions form a two parameter family.
- If $\rho(A) \neq \rho([A\,|\,B])$, then the system $AX = B$ is inconsistent and has no solution.

Example

Find the condition on a, b and c so that the following system of linear equations has one parameter family of solutions: x + y + z = a, x + 2 y + 3z = b, 3x + 5 y + 7z = c.

Solution

Here the number of unknowns is 3.

The matrix form of the system is $AX = B$, where $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 3 & 5 & 7 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

Applying elementary row operations on the augmented matrix [A | B], we get

$$[A\,|\,B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 1 & 2 & 3 & b \\ 3 & 5 & 7 & c \end{array}\right] \xrightarrow[R_3 \to R_3 - 3R_1]{R_2 \to R_2 - R_1,} \left[\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & 1 & 2 & b-a \\ 0 & 2 & 4 & c-3a \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & 1 & 2 & b-a \\ 0 & 0 & 0 & (c-3a)-2(b-a) \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & 1 & 2 & b-a \\ 0 & 0 & 0 & (c-2b-a) \end{array}\right].$$

In order that the system should have one parameter family of solutions, we must have $\rho\,(A) = \rho\,([A, B]) = 2$. So, the third row in the echelon form should be a zero row.

So, $c - 2b - a = 0 \Rightarrow c = a + 2b$.

Example

Investigate for what values of λ and μ the system of linear equations

$x + 2y + z = 7$, $x + y + \lambda z = \mu$, $x + 3y - 5z = 5$ has

(i) no solution (ii) a unique solution (iii) an infinite number of solutions.

Solution

Here the number of unknowns is 3.

The matrix form of the system is $AX = B$, where $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & \lambda \\ 1 & 3 & -5 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, B = \begin{bmatrix} 7 \\ \mu \\ 5 \end{bmatrix}$.

Applying elementary row operations on the augmented matrix [A | B], we get

$$[A\,|\,B] = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 7 \\ 1 & 1 & \lambda & \mu \\ 1 & 3 & -5 & 5 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 7 \\ 1 & 3 & -5 & 5 \\ 1 & 1 & \lambda & \mu \end{array}\right]$$

$$\xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - R_1,} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 7 \\ 0 & 1 & -6 & -2 \\ 0 & -1 & \lambda-1 & \mu-7 \end{array}\right] \xrightarrow{R_3 \to R_3 + R_2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 7 \\ 0 & 1 & -6 & -2 \\ 0 & 0 & \lambda-7 & \mu-9 \end{array}\right].$$

(i) If $\lambda = 7$ and $\mu \neq 9$, then ρ (A) = 2 and ρ ([A | B]) = 3. So ρ (A) $\neq$ ρ ([A | B]). Hence the given system is inconsistent and has no solution.

(ii) If $\lambda \neq 7$ and m is any real number, then ρ (A) = 3 and ρ ([A | B]) = 3.

So ρ (A) = ρ ([A | B]) = 3 = Number of unknowns. Hence the given system is consistent and has a unique solution.

(iii) If $\lambda = 7$ and $\mu = 9$, then ρ(A) = 2 and ρ ([A | B]) = 2.

So, ρ(A) = ρ ([A | B]) = 2 < Number of unknowns. Hence the given system is consistent and has infinite number of solutions.

Matrix: Non-homogeneous Linear Equations

EXERCISE

1. Test for consistency and if possible, solve the following systems of equations by rank method.

(i) $x - y + 2z = 2$, $2x + y + 4z = 7$, $4x - y + z = 4$

(ii) $3x + y + z = 2$, $x - 3y + 2z = 1$, $7x - y + 4z = 5$

(iii) $2x + 2y + z = 5$, $x - y + z = 1$, $3x + y + 2z = 4$

(iv) $2x - y + z = 2$, $6x - 3y + 3z = 6$, $4x - 2y + 2z = 4$

SOLUTION

(i) $x - y + 2z = 2$, $2x + y + 4z = 7$, $4x - y + z = 4$

The system of given equations is

$$x - y + 2z = 2 \quad \text{-----}(1)$$

$$2x + y + 4z = 7 \quad \text{-----}(2)$$

$$4x - y + z = 4 \quad \text{-----}(3)$$

This system has 3 unknowns.

The augmented matrix of the system is $[A \mid B] = \left[\begin{array}{ccc|c} 1 & -1 & 2 & 2 \\ 2 & 1 & 4 & 7 \\ 4 & -1 & 1 & 4 \end{array}\right]$

Applying Gaussian elimination method on $[A \mid B]$ we get

$$[A \mid B] = \left[\begin{array}{ccc|c} 1 & -1 & 2 & 2 \\ 2 & 1 & 4 & 7 \\ 4 & -1 & 1 & 4 \end{array}\right] \xrightarrow[R_3 \to R_3 - 4R_1]{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 2 \\ 0 & 3 & 0 & 3 \\ 0 & 3 & -7 & -4 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 2 \\ 0 & 3 & -7 & -4 \\ 0 & 3 & 0 & 3 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 2 \\ 0 & 3 & -7 & -4 \\ 0 & 0 & 7 & 7 \end{array}\right]$$

From the last echelon form the augmented matrix and the coefficient matrix have three non - zero rows. Hence rank of A and $[A \mid B]$ are same and is equal to the number of unknowns.

That is $\rho(A) = \rho([A \mid B]) = 3 =$ The number of unknowns.

$\therefore$ The system is consistent and has a unique solution. From the last echelon form the equivalent system of equations is

$$x - y + 2z = 2 \quad \text{-----}(4)$$

$$3y - 7z = -4 \quad \text{-----}(5)$$

$$7z = 7 \quad \text{-----}(6)$$

$$z = \frac{7}{7} = 1$$

Substituting $z = 1$ in equation (5)

$$3y - 7 \times 1 = -4 \Rightarrow 3y = -4 + 7$$

$$3y = 3 \Rightarrow y = 1$$

Substituting $y = 1$ and $z = 1$ in equation (4) we get

$$x - 1 + 2 \times 1 = 2 \Rightarrow x = 2 - 1 = 1$$

$\therefore$ The required solutions are $x = 1$, $y = 1$, $z = 1$.

(ii) $3x + y + z = 2$, $x - 3y + 2z = 1$, $7x - y + 4z = 5$

The system of given equations is

$$3x + y + z = 2 \qquad \text{-----}(1)$$

$$x - 3y + 2z = 1 \qquad \text{-----}(2)$$

$$7x - y + 4z = 5 \qquad \text{-----}(3)$$

The number of unknowns is 3.

The augmented matrix of the system is $[A \mid B] = \left[\begin{array}{ccc|c} 3 & 1 & 1 & 2 \\ 1 & -3 & 2 & 1 \\ 7 & -1 & 4 & 5 \end{array}\right]$

Applying Gaussian elimination method on $[A \mid B]$ we get

$$[A \mid B] = \left[\begin{array}{ccc|c} 3 & 1 & 1 & 2 \\ 1 & -3 & 2 & 1 \\ 7 & -1 & 4 & 5 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & -3 & 2 & 1 \\ 3 & 1 & 1 & 2 \\ 7 & -1 & 4 & 5 \end{array}\right] \xrightarrow[R_3 \to R_3 - 7R_1]{R_2 \to R_2 - 3R_1}$$

$$\left[\begin{array}{ccc|c} 1 & -3 & 2 & 1 \\ 0 & 10 & -5 & -1 \\ 0 & 20 & -10 & -2 \end{array}\right] \xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{ccc|c} 1 & -3 & 2 & 1 \\ 0 & 10 & -5 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

In the last echelon form the augmented and the coefficient matrix have two non - zero rows.

Rank of A, $\rho(A) = 2$

Rank of $[A \mid B]$, $\rho([A \mid B]) = 2$

$\rho(A) = \rho([A \mid B]) = 2 < 3$, the number of unknowns.

From the last echelon form the equivalent system of equations is

$$x - 3y + 2z = 1 \qquad \text{-----}(4)$$

$$10y - 5z = -1 \qquad \text{-----}(5)$$

The equivalent system has two non - trivial equations and three unknowns. Let us fix one unknown. Let $z = t$,

$(5) \Rightarrow \quad 10y - 5t = -1 \quad \Rightarrow \quad 10y = -1 + 5t \quad \Rightarrow \quad y = \dfrac{-1 + 5t}{10}$

Put $z = t$ and $y = \dfrac{-1 + 5t}{10}$ in equation (4) we get

$$x - 3\left(\frac{-1 + 5t}{10}\right) + 2t = 1$$

$$x = 1 + 3\left(\frac{-1 + 5t}{10}\right) - 2t$$

$$x = 1 + \frac{15t - 3}{10} - 2t$$

$$x = \frac{10 + 15t - 3 - 20t}{10} = \frac{-5t + 7}{10}$$

$\therefore$ The required solutions are $x = \dfrac{-5t + 7}{10}$, $y = \dfrac{-1 + 5t}{10}$, $z = t$

(iii) $2x + 2y + z = 5$, $x - y + z = 1$, $3x + y + 2z = 4$

The system of given equations is

$$2x + 2y + z = 5 \quad \text{-----}(1)$$

$$x - y + z = 1 \quad \text{-----}(2)$$

$$3x + y + 2z = 4 \quad \text{-----}(3)$$

In this system the number of unknowns is 3.

The augmented matrix of the system is $[A \mid B] = \left[\begin{array}{ccc|c} 2 & 2 & 1 & 5 \\ 1 & -1 & 1 & 1 \\ 3 & 1 & 2 & 4 \end{array}\right]$

Applying Gaussian elimination method on $[A \mid B]$ we get

$$[A \mid B] = \left[\begin{array}{ccc|c} 2 & 2 & 1 & 5 \\ 1 & -1 & 1 & 1 \\ 3 & 1 & 2 & 4 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 2 & 2 & 1 & 5 \\ 3 & 1 & 2 & 4 \end{array}\right] \xrightarrow[R_3 \to R_3 - 3R_1]{R_2 \to R_2 - 2R_1}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 4 & -1 & 3 \\ 0 & 4 & -1 & 1 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 4 & -1 & 3 \\ 0 & 0 & 0 & -2 \end{array}\right]$$

In the last echelon form the augmented matrix has 3 non - zero rows hence its rank is 3.

That is $\rho([A \mid B]) = 3$

Also from the last echelon form the coefficient matrix $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ has two non - zero rows and hence has order 2. That is $\rho(A) = 2$

So $\rho(A) \neq \rho([A \mid B])$

From the last echelon form the equivalent system of equations is

$$x - y + z = 1 \quad \text{-----}(4)$$

$$4y - z = 3$$

$$0 = -2$$

$0 = -2$ is a contradiction

$\therefore$ The given system is inconsistent and has no solution.

(iv) $2x - y + z = 2$, $6x - 3y + 3z = 6$, $4x - 2y + 2z = 4$

The system of given equations is

$$2x - y + z = 2 \quad \text{-----}(1)$$
$$6x - 3y + 3z = 6 \quad \text{-----}(2)$$
$$4x - 2y + 2z = 4 \quad \text{-----}(3)$$

This system has 3 unknowns.

The augmented matrix of the system is $[A\,|\,B] = \left[\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 6 & -3 & 3 & 6 \\ 4 & -2 & 2 & 4 \end{array}\right]$

Applying Gaussian elimination method on $[A\,|\,B]$ we get

$$[A\,|\,B] = \left[\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 6 & -3 & 3 & 6 \\ 4 & -2 & 2 & 4 \end{array}\right] \xrightarrow[R_3 \to \frac{1}{2}R_3]{R_2 \to \frac{1}{3}R_2} \left[\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 2 & -1 & 1 & 2 \\ 2 & -1 & 1 & 2 \end{array}\right] \xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - R_1} \left[\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

From the last echelon form the augmented matrix and the coefficient matrix have only one non - zero row and hence their rank is 1.

That is $\rho([A\,|\,B]) = 1$ and $\rho(A) = 1$

$\rho(A) = \rho([A\,|\,B]) = 1 < 3$ the number of unknowns.

From the echelon form we get the equivalent equations

$$2x - y + z = 2 \quad \text{-----}(4)$$
$$0 = 0 \quad \text{-----}(5)$$
$$0 = 0 \quad \text{-----}(6)$$

The equivalent system has one non trivial equation and three unknowns. Take $y = s$ and $z = t$ we get

$$2x - s + t = 2$$
$$2x = 2 + s - t \Rightarrow x = \frac{s - t + 2}{2}$$

$\therefore$ The solution of the system is $x = \frac{s - t + 2}{2}$, $y = s$, $z = t$

where s and t are parameters. The above solution set is a two parameter family of solutions.

Hence the given system of equations is consistent and has infinitely many solutions which form a two parameter family of solutions.

2. Find the value of k for which the equations $kx - 2y + z = 1$, $x - 2ky + z = -2$, $x - 2y + kz = 1$ have

(i) no solution

(ii) unique solution

(iii) infinitely many solution

SOLUTION

The system of given equations is

$$kx - 2y + z = 1 \quad \text{-----(1)}$$

$$x - 2ky + z = -2 \quad \text{-----(2)}$$

$$x - 2y + kz = 1 \quad \text{-----(3)}$$

The number of unknowns in this system is 3.

The augmented matrix of the system is $[A|B] = \left[\begin{array}{ccc|c} k & -2 & 1 & 1 \\ 1 & -2k & 1 & -2 \\ 1 & -2 & k & 1 \end{array}\right]$

Applying Gaussian elimination method on $[A|B]$ we get

$$[A|B] = \left[\begin{array}{ccc|c} k & -2 & 1 & 1 \\ 1 & -2k & 1 & -2 \\ 1 & -2 & k & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & -2k & 1 & -2 \\ k & -2 & 1 & 1 \\ 1 & -2 & k & 1 \end{array}\right] \xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - kR_1}$$

$$\left[\begin{array}{ccc|c} 1 & -2k & 1 & -2 \\ 0 & -2+2k^2 & 1-k & 1+2k \\ 0 & -2+2k & k-1 & 3 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & -2k & 1 & -2 \\ 0 & 2(k+1)(k-1) & -(k-1) & 2k+1 \\ 0 & 2(k-1) & k-1 & 3 \end{array}\right] \xrightarrow{R_3 \to (k+1)R_3}$$

$$\left[\begin{array}{ccc|c} 1 & -2k & 1 & -2 \\ 0 & 2(k+1)(k-1) & -(k-1) & 2k+1 \\ 0 & 2(k-1)(k+1) & (k-1)(k+1) & 3(k+1) \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2}$$

$$\left[\begin{array}{ccc|c} 1 & -2k & 1 & -2 \\ 0 & 2(k+1)(k-1) & -(k-1) & 2k+1 \\ 0 & 0 & (k-1)(k+1)+(k-1) & 3(k+1)-(2k+1) \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} 1 & -2k & 1 & -2 \\ 0 & 2(k+1)(k-1) & -(k-1) & 2k+1 \\ 0 & 0 & (k-1)(k+2) & (k+2) \end{array}\right]$$

(i) When $k = 1$, the last echelon matrix reduces to the form $\left[\begin{array}{ccc|c} 1 & -2 & 1 & -2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right]$

The augmented matrix $[A|B]$ has 3 non-zero rows and the coefficient matrix A has only one non-zero row.

$$\therefore \quad \rho(A) = 1 \quad \text{and} \quad \rho([A|B]) = 3$$

$$\rho(A) \neq \rho([A|B])$$

In this case the system is inconsistent and has no solution.

Thus for $k = 1$, the system is inconsistent and has no solution.

(ii) When $k \neq 1$ and $k \neq -2$ the last row of the echelon form is a non - zero row. Hence in this case both the augmented matrix $[A|B]$ and the coefficient matrix A have three non zero rows.

$\therefore$ $\rho(A) = \rho([A|B]) = 3 =$ the number of unknowns.

Hence the system is consistent and has unique solution.

Thus for $k \neq 1$ and $k \neq -2$ the given system is consistent and has a unique solution.

(iii) When $k = -2$, the last row of the echelon form reduces to zero. Hence in the echelon form both the augmented matrix $[A|B]$ and the coefficient matrix A have two non - zero rows. Therefore rank of A and rank of $[A|B]$ are same and is equal to 2.

$\rho(A) = \rho([A|B]) = 2 < 3$, the number of unknowns.

Hence in this case the system is consistent and has infinite number of solutions.

Thus $k = -2$, the given system is consistent and has infinite number of solutions.

3. Investigate the values of λ and μ the system of linear equations $2x + 3y + 5z = 9$, $7x + 3y - 5z = 8$, $2x + 3y + \lambda z = \mu$, have

(i) no solution (ii) a unique solution (iii) an infinite number of solutions.

SOLUTION

The system of given equations is

$$2x + 3y + 5z = 9 \quad \text{-----(1)}$$

$$7x + 3y - 5z = 8 \quad \text{-----(2)}$$

$$2x + 3y + \lambda z = \mu \quad \text{-----(3)}$$

In this system the number of unknowns is 3.

The augmented matrix of the system is $[A \mid B] = \left[\begin{array}{ccc|c} 2 & 3 & 5 & 9 \\ 7 & 3 & -5 & 8 \\ 2 & 3 & \lambda & \mu \end{array}\right]$

Applying Gaussian elimination method on $[A \mid B]$ we get

$$[A \mid B] = \left[\begin{array}{ccc|c} 2 & 3 & 5 & 9 \\ 7 & 3 & -5 & 8 \\ 2 & 3 & \lambda & \mu \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 7 & 3 & -5 & 8 \\ 2 & 3 & 5 & 9 \\ 2 & 3 & \lambda & \mu \end{array}\right] \xrightarrow{R_1 \to R_1 - 3R_2} \left[\begin{array}{ccc|c} 1 & -6 & -20 & -19 \\ 2 & 3 & 5 & 9 \\ 2 & 3 & \lambda & \mu \end{array}\right]$$

$$\xrightarrow[R_3 \to R_3 - 2R_1]{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & -6 & -20 & -19 \\ 0 & 15 & 45 & 47 \\ 0 & 15 & \lambda + 40 & \mu + 38 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|c} 1 & -6 & -20 & -19 \\ 0 & 15 & 45 & 47 \\ 0 & 0 & \lambda - 5 & \mu - 9 \end{array}\right]$$

(i) When $\lambda = 5$ and $\mu \neq 9$, the third row of the echelon form is non - zero. Therefore the augmented matrix $[A \mid B]$ has three non zero rows but the coefficient matrix A has two non zero rows.

$$\therefore \quad \rho(A) = 2 \;, \quad \rho([A \mid B]) = 3$$

$$\rho(A) \neq \rho([A \mid B])$$

Hence the system is inconsistent and has no solution.

Thus for $\lambda = 5$ and $\mu \neq 9$, the system is inconsistent and has no solution.

(ii) When $\lambda \neq 5$, the third row of the augmented matrix in the echelon form is non zero. Therefore both the augmented matrix and the coefficient matrix have three non zero rows hence the rank of the augmented matrix and the coefficient matrix are same and is equal to 3.

$$\rho(A) = \rho([A \mid B]) = 3 = \text{the number of unknowns.}$$

Hence the system is consistent and has unique solution.

Therefore $\lambda \neq 5$, the given system is consistent and has a unique solution.

(iii) When $\lambda = 5$ and $\mu = 9$, the last row of the echelon form reduces to zero. Therefore in the last echelon form both the augmented matrix $[A \mid B]$ and the coefficient matrix A have two non zero rows and hence their ranks are equal and is equal to 2.

$$\rho(A) = \rho([A \mid B]) = 2 < 3, \text{ the number of unknowns.}$$

$\therefore$ The system is consistent and has infinite number of solutions.

Thus for $\lambda = 5$ and $\mu = 9$, the given system is consistent and has infinite number of solutions.

Unit – III – Measures of Central Tendency

Average – Meaning

Average is a single value that represents group of values

Definition

An Average is a value which a typical or representative of a net of data

Characteristics of a Good Average

- It should be defined clear and unambiguous so that it leads to one and only one interpretation by different persons
- It should be easy to understand and simple to compute and should not involve heavy arithmetical calculations
- It should be based on all the items of the given set of data is compute the average.
- It should be suitable for further algebras mathematical treatment and capable of being used is further statistical computations

Uses of Average

- It is useful to describe the distribution in a concise manner
- It is useful to com pare different distributions
- It is useful to compare various statistical measures such as dispersion, skewness, kurtosis and so on

Functions or An average

- To facilitate Quick understanding of complex data
- To facilitate Comparison
- It establishes mathematical relationship
- Capable of further statistical comparison

Types of Average

- Mathematical Average
- Location Average
- Commercial Average

Objectives of an Average

- To get a single value that describe the features of the entire group
- To provide ground for better comparison
- To provide ground for further statistical computation and analysis

Arithmetic Mean

The arithmetic mean of a series of items is the sum of the values of all items divided by that total number. It is a multinational average and it is the most popular measure of central tendency

Merits of Antiemetic Mean

- Easy to calculate and understand
- It is a perfect average, affect by the value of every item in the series
- It is calculated value and not based on position in the series
- It is determined by a rigid formula. Hence, everyone who computes the average gets the same answer
- I t is used in further calculation
- It gives a good base for comparison

Demerits of Arithmetic Mean

- The mean is unduly affected by the extreme items
- It is unreliable Itmay lead to a false conclusion
- It is not useful for the study of qualities
- It cannot be located by the graphic method

Arithmetic Mean

Individual Series

Find our mean from the following data

Roll No	1	2	3	4	5	6	7	8	9	10
Marks	21	30	28	40	26	34	40	9	15	17

Solution

Roll No	Marks (X)
1	21
2	30
3	28
4	40
5	26
6	34
7	40
8	9
9	15
10	17
N=10	$\sum X$ =300

Formula = X =$\sum$X
/NX =300/10 = 30.
The mean marks =30

Discrete Series

Calculate the arithmetic mean for the wages of workers in a Factory

Wages in Rs.	4	6	8	10	15	16
Workers	5	15	6	7	8	2

Solution

Wages in Rs.	Workers f	fx
4	5	4x5=20
6	15	6x15=96
8	6	8x6=48
10	7	10x7=70
15	8	15x8=120
16	2	16x2=32
	N=∑f=43	∑fx=380

X = ∑fx /N = 380, N=43
= 380 / 43 = 8.837
The average wage of workers = Rs.8.84

Continuous Series Calculate
Arithmetic Mean

Class Intervals	**0-10**	**10-20**	**20-30**	**30-40**	**40-50**
Frequency	**6**	**5**	**8**	**15**	**7**

Class Intervals	Mid-point	Frequency	fm
0-10	5	6	30
10-20	15	5	75
20-30	25	8	200
30-40	35	15	525
40-50	45	7	315
		N=∑f=41	N=∑fm=1145

Arithmetic Mean = X = =∑fm / N
The Arithmetic mean = 27.92

Median

Median is the value of the middle item of a series arranged in ascending or descending order of magnitude. Hence it is the “Middle most” or “Most central” value of a set of number. It divide the series into two equal part, one part containing values greater and the other with values less than the median.

Meaning

The number is that value of the variable which divides the group into two equal parts, one part comprising all values greater and the other, all values less than median.

Merits of Median:

- It is easy to compute and understand
- It eliminates the effect of extreme item
- The value of median can be located graphically
- Demerits of Median
- The calculating media, it is necessary to arrange the data other averages do not need an arrangement
- It is affected more by fluctuation of sampling than the arithmetic mean.
- It is not based on all the items of the series

Individual Series

Arrange the data either ascending or descending order

Median – Size of $\frac{N+1}{2}$ th Item

Find out the median from the following

57	58	61	42	38	65	72	66	80

Solution

Sl.No	Data arranged in ascending order

1	38
2	42
3	57
4	58
5	62
6	65
7	66
8	72
9	80

Median= – Size of $\frac{N + 1}{2}$ th Item

- = Size of 9 + 1
- ------------ the item
-
- 2
- = 10 / 2 = 5th item

Median = 62

Discrete Series

Compute the median for the following distribution of weeks of wagers of 65 employees of the xyz company

Weekly wages in Rs	55	65	785	85	95	105	115
Number of employees	8	10	16	14	10	5	2

Solution

Weekly wages in Rs	No of Employees	Cumulative frequency (cf)
55	8	8
65	10	18
75	16	34
85	14	48
95	10	58
105	5	63
115	2	65

Median = –Size of $\frac{N + 1}{2}$ th Item

= size of $\frac{65 + 1}{2}$ th Item

= 33' which is nearer to 34

Cf of 34 = 75

Median weekly wages = 75

Continuous Series

Calculate the median form the following data

Marks	0-20	20-40	40-60	60-80	80-100
No of Students	5	15	30	8	2

Solution

Marks	No of students	Cumulative frequency
0-20	5	5
20-40	15	20
40-60	30	50
60-80	8	58
80-100	2	60

Median = –ize of $\dfrac{N}{2}$ th Item c size of 60/2th Item

$$=\text{size of } L + \frac{N/2 - cf}{f} \times C$$

$$= 40 + \frac{30 - 20}{2} x\ 20 = 46.47$$

Median marks = 46.676

Mode

Mode is the modal value in the value of the variable which occurs more number of times or most frequently is a distribution. Mode is the value which occurs with the greatest number of frequency in a series

Types of modal

I. **Uni-model**
If there is only one mode in series is called uni-model

II. **Bi-Modal**
If there are two modes in the series, it is called bi-model

III. **Tri-Modal**
If they are three modes in the series, it is Relationship between different Averages Symmetrical is called Tri-model

IV. **Multimodal**
If there are more than three modes in the series it is called multi-mode.
Relationship among mean, median and mode
The three averages are identical, when the distribution is symmetrical. In an asymmetricaldistribution,

the values of mean, median and mode are not equal.
Median = 1/3 (Mean - mode)
Mode = 2 median - 2 mode
Median = Mode * 2/3 (Mean – Mode)

Individual Series

Calculate the mode form the following data of the marks obtain by 10 students

Serial No	1	2	3	4	5	6	7	8	9	10
Marks obtained	60	77	74	62	77	77	70	68	65	80

Solution

Marks obtained by 10 students is here 77 is repeated three times
Therefore the Mode mark is 77

Discrete Series

Calculate the mode form the following data of the wages of workers of are establishment. Find the modal wages

Daily wages in Rs	3	4	6	7	9	10	12	13	15
No of wage earners	2	3	2	6	10	11	12	5	1

Solution Grouping Table

<table>
<tr><th rowspan="2">Daily Wages is Rs.</th><th colspan="6">Frequency of Wages Earners</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th></tr>
<tr><td>3</td><td>2</td><td rowspan="2">5</td><td></td><td rowspan="3">7</td><td></td><td></td></tr>
<tr><td>4</td><td>3</td><td rowspan="2">5</td><td rowspan="3">11</td><td></td></tr>
<tr><td>6</td><td>2</td><td rowspan="2">8</td><td rowspan="3">18</td></tr>
<tr><td>7</td><td>6</td><td rowspan="2">16</td><td rowspan="3">27</td></tr>
<tr><td>9</td><td>10</td><td rowspan="2">21</td><td rowspan="3">33</td></tr>
<tr><td>10</td><td>11</td><td rowspan="2">23</td><td rowspan="2">28</td></tr>
<tr><td>12</td><td>12</td><td>17</td><td>18</td></tr>
</table>

13	5		6			
15	1					

Column	**Size of Item**					
	4	**6**	**7**	**9**	**10**	**12**	**15**
1						**I**	
2				**I**	**I**	**I**	
3					**I**	**I**	
4			**I**	**I**	**I**		
5				**I**	**I**	**I**	
6					**I**	**I**	**I**
			I	**3**	**5**	**4**	**1**

From the analysis table it is known that size10 has been repeated the maximum number of times, thus is, so the modal wages Rs10

Continuous series

Find out the mode from the following series

X	**0-5**	**5-10**	**10-15**	**15-20**	**20-25**	**25-30**	**30-35**
frequency	**1**	**2**	**5**	**14**	**10**	**9**	**2**

Grouping Table

X	Frequency					
	1	2	3	4	5	6
0-5	1	3		8		
5-10	2		7		21	
10-15	5	19				29
15-20	14		24	**33**		
20-25	10	**19**			**21**	
25-30	9		**11**			
30-35	**2**					

Column	**Size of Item**

	0-5	5-10	10-15	15-20	20-25	25-30	30-35
1				**I**			
2			**I**	**I**	**I**		
3				**I**	**I**	**I**	
4				**I**	**I**	**I**	
5		**I**	**I**	**I**	**I**	**I**	**I**
6			**I**				
		1	**3**	**6**	**5**	**3**	**1**

Modal value lies in 15-20 as it occurs most frequently

$$\text{Mode (Z)} = L + \frac{f1 - f0}{2f1 - f0 - f2} \times C$$

$$\text{Mode (z)} = 154 + \frac{14 - 5}{2(14) - 5 - 10} 5$$

$= 15 + 9/13 \text{ X } 5 = 15 + 45/13$

$= 15 + 3.46$

Mode = 18.46

Geometric Mean

Merits of geometric Mean

- Every item in the distribution is included in the calculation
- It can be calculated with mathematical exactness, provided that all the qualities are greater than zero and positive
- Large items have less effect on it than in the arithmetic average.
- It is amenable to further algebraic manipulation

Demerits of Geometirc mean

- It is very difficult to calculate
- It is impossible to use it when any item is zero or negative
- The value of the geometric mean may not correspond with any actual value in the distribution

Uses of Geometric mean

- This average is often used to construct index numbers, where we are chiefly concerned with relative changes over a period of time
- It is the only useful average that can be employed to indicate rate of have

Individual series

$$G.M = \text{Anti ling of } \frac{\sum \log X}{N}$$

Calculate Geometric Mean

50	72	54	82	93

Solution

X	Logt X
50	1.6990
72	1.8573
54	1.7324
82	1.9238
93	1.9685

$$G.M = \text{Anti ling of } \frac{\sum \log X}{N}$$

$$= \frac{9.1710}{5} = 1.8342$$

= Antilog of 1.8342 = 68.26

Discrete Series
Calculate

Geometric mean from the following data

Size of Item	**120**	**125**	**130**	**135**	**136**	**138**	**139**	**140**	**147**
Frequency	**2**	**3**	**3**	**1**	**1**	**7**	**4**	**2**	**8**

Size of Item (X)	**Frequency (f)**	**Log X**	**F log x**
120	2	2.0792	4.1584
125	3	2.0969	6.2907
1360	3	2.1139	6.3417

135	1	2.1303	2.1303
136	2	2.1335	4.2670
138	7	2.1399	14.9793
139	4	2.1430	8.5720
140	2	2.1461	4.2922
147	8	2.1673	17.3384
	N=∑f=32		**N==∑f log x = 68.3700**

$$G.M = \text{Anti ling of } \frac{\sum \log X}{N} = \text{Anti log of } \frac{68.375}{32}$$

= Antilog of 2.1366 = 137   Therefore G.M = 137

Continuous series

Geometric mean from the following data

Yield of wheat	**7.54-10.5**	**10.5-13.5**	**13.5 -16.5**	**16.5-19.5**	**19.5-22.5**	**22.5-25.5**	**25.5-28.5**
No of forms	**5**	**9**	**19**	**23**	**7**	**4**	**1**

Yield of wheat	**Mid Value**	**Log M**	**No. of Forms**	**f log m**
7.54-10.5	9	0.9542	5	4.7710
10.5-13.5	12	1.0792	9	9.7128
13.5 -16.5	15	1.1761	19	22.3459
16.5-19.5	18	1.2553	23	28.8719
19.5-22.5	21	1.3222	7	9.2554
22.5-25.5	24	1.3802	4	5.5208
25.5-28.5	27	1.4314	1	1.4314
			N=68	**∑flog m=81.9092**

G.M = Anti logof ∑f log m

N

= 81.9092/68 = 1.204547
= Antilog of 1.204547 = 16.02 G.M = 16.02

Harmonic Mean
Meaning

Harmonic Mean is the reciprocal of the arithmetic average of the reciprocal of values of various item in the invariable

Merits of Harmonic Mean

- It utilizes all values of a variable
- It is very important to small values
- It is amenable to further algebraic manipulation
- It provides consistent results in problems relating to time and rates than similar averages

Demerits of Harmonic Mean

- It is not very easy to understand
- The method of calculation is difficult
- The presence of both positive and negative items in a series makes it impossible to compute its value. The same difficulty is felt if one or more items are zero
- It is only a summary figure and may not be the actual item in the series.

Individual Series

$$H.M = \frac{N}{\sum 1/x}$$

Find out the Harmonic mean

Family	1	2	3	4	5	6	7	8	9	10
Income	85	70	10	75	500	8	42	250	40	36

Solution

Computation of Harmonic Mean

Family	**Income (X)**	**1 / x**
1	85	0.01176
2	70	0.01429
3	10	0.10000
4	75	0.01333

5	500	0.00200
6	8	0.12500
7	42	0.02381
8	250	0.00400
9	40	0.02500
10	36	0.02778
N =10		**∑1/x = 0.34697**

$$H.M = \frac{N}{\sum 1/x}$$

= 10 /0.34697 = 28.82 H.M = 28.82

Discrete Series

Size of Item	6	7	8	9	10	11
Frequency	4	6	9	5	2	8

Size of Item X	**Frequency f**	**1/x**	**F 1/x**
6	4	0.1667	0.6668
7	6	0.1429	0.8574
8	9	0.1250	1.1250
9	5	0.1111	0.5555
10	2	0.1000	0.20000.7272
11	8	0.0909	
	N=∑f =34		∑f 1/x = 4.1319

$$H.M = \frac{N}{\sum f1/x} = \frac{34}{4.1319} = 8.23$$

Compute Harmonic Mean

Continuous Series

Size	0-10	10-20	20-30	30-40	40-50
Frequency	5	8	12	6	4

Solution

Size	**Frequency f**	**Mid value**	**reciprocal**	**F(1/m)**
0-10	5	5	0.20000	1.00000
10-20	8	15	0.06667	0.53336
20-30	12	25	0.04000	0.48000
-30-40	6	35	0.02857	0.17142
40-50	4	45	0.02222	0.08888

	∑f = 35			**∑f 1/m = 2.27366**

$$H.M = \frac{N}{\sum f 1/m} = \frac{35}{2.27366} = 15.393682$$

Dispersion Meaning

Dispersion is the study of scatterness around an average

Definition

Dispersion is the measures of the variation of the items --- A.L.Bowly

Dispersion is a measure of extent to which the individual items vary --- L.R.Connor

Importance of measuring variation or dispersion

- Testing the Reliability of the Measures of Central Tendency
- Comparing two or more series on the basis of their variability
- Enabling to control the variability
- Facilitating as a Basis for further statistical Analysis

Characteristics of a Measure of Variation

- It is easy to understand and simple to calculate
- It should be rigidly defined
- It should be based on all observations and it should not be affected by extreme observations
- It should be amenable to further algebraic treatment
- It should have sampling stability

Methods of Measuring Dispersion

- Range
- Inter Quartile range
- Quartile Deviation
- Mean Deviation
- Standard Deviation
- Lorenz Curve

Range

Range is the difference between the largest and the smallest value in the distribution. It is the simplest and crudest measure of dispersion

Uses of Range

- It is used in industries for the statistical quality control of the m infected product
- It is used to study the variations such as stock, shares and other commodities

- It facilitates the use of other statistical measures

Advantages of Range

- It is the simplest method of studying variation
- It is easy to understand and the easiest to compute
- It takes minimum time to calculate
- It is accurate

Disadvantages of Range

- Range is completely depended on the two extreme values
- It is subject to fluctuations of considerable magnitude from sample to sample
- It is not suitable for mathematical treatment
- It cannot be applied to open and classes
- Range cannot tell us anything about the character of the distribution

Quartile deviation

Quartile deviation is an absolute measure of dispersion. It is calculated on the basis ofthe difference of upper quartile and the lower Quartile divided by 2.

In the series, four quartiles are there. By eliminating the lowest (25%) items and the highest (25%) items of a series, we can obtain a measure of dispersion and can find out half the distance between the first and the third quartiles.

$$\text{Quartile Deviation (Q.D)} = \frac{Q3 - Q1}{2}$$

$$\text{Co-efficient of Q.D} = \frac{Q3 - Q1}{Q3 + Q1}$$

Merits of Quartile Deviation

- It is simple to calculate and easy to understand
- Risk of extreme item variance is eliminated, as it depend upon the central 50 per cent items
- It can be applied to open and classes

Demerits of quartile Deviation

- Items below Q1 and above Q3 are ignored
- It is not capable of further mathematical treatment
- It is affected much by the fluctuations of sampling
- It is not calculated from a computed average, but from a positional average.

Mean deviation

Mean deviation is the average difference between the items in a distribution computed from the mean, median or mode of that series counting all such deviation as positive. The mean deviation is also known as the average deviation

Mean deviation =∑I D I

-N

Co – efficient of Mean Deviation (M.D) = MD

X or Z or M

Merits of Mean Deviation

- It is clear and easy to understand
- It is based on each and every item of the data It can be calculated from any measure of central tendency and as such as flexible too**.**

Demerits of mean Deviation

- **It** is not suitable for further mathematical processing
- It is rarely used in sociological studies
- It is mathematically unsound and illogical, because the signs are ignored in the calculation of mean deviation

Standard deviation

Standard deviation is the square root of the means of the stranded deviation from the Arithmetic mean. So, it is also known as Root Mean Square Deviation an Average of Second order. Standard deviation is denoted by the small Greek letter ‘σ’ the concept of standard deviation is introduced by Karl Pearson in 1893.

Uses of Standard deviation

- It is used in statistics because it possesses must of the characteristics of an ideal measure of dispersion.
- It is widely used in sampling theory and by biologists.
- It is applied in co-efficient of correlation and in the study of symmetrical frequency distribution

Advantages of standard deviation

- It is rigidly defined determinate
- It is based on all the observations of a series
- It is less affected by fluctuations of sampling and hence stable
- It is amenable to algebraic treatment and is less affected by fluctuations of sampling most other measures of dispersion
- The standard deviation is more appropriate mathematically than the mean deviation, since the negative signs are removed by squaring the deviations rather than by ignoring

Co efficient of Variance

Standard deviation is an absolute measure of dispersion. The corresponding relative measure is known as the co=-efficient of variation. It is used to compare the variability of twoor more series

$$\text{Co –efficient of Standard deviation} = \frac{\sigma}{\bar{X}}$$

$$\text{Co-efficient of Variance (C.V)} = \frac{\sigma}{\bar{x}} \times 100$$

Graphic method of dispersionLorenz Curve

Lorenz Curve is a device used to show the measurement of economic inequalities as in the distribution of income and wealth. It can also be used in business to study the disparities of distribution of profit, wages, turnover, production and the like.

Range

Range = L-S

$$\text{Co-efficient of range} = \frac{L-S}{L+S}$$

Solved Problems

Find the range and co-efficient of range for the heights of 8 students of a class 158,160,165,168,170,173,

Solution

Range = L – S
Given Series the largest value of the series = 173
Smallest value of the series = 158
Range = 178 – 158 = 15

$$\text{Co-efficient of range} = \frac{L-S}{L+S} = 0.045$$

Quartile Deviation

$$\textbf{Quartile Deviation} = \frac{Q3 - Q1}{2}$$

Individual Series

Find out the value of quartile deviation and its co-efficient from the following data

Roll No	1	2	3	4	5	6	7
Marks	20	28	40	30	50	60	52

Solution

Marks arranged in ascending order

20,28,30,40,50,52,60N + 1

Q1 = size of----------------the item

4

= size of 7 + 1

---------- th item = 8/4 th Item
4

= size of 2nd item = 28

3(N + 1)

Q3 = size of ------------------th item

4

3(7+1)
= ....
4

= 3 x 8 / 4 = 24/4 = 6 th item

Size of the 6thitem = 52

Q3 – Q1 52 -28

Q.D ----------- = ---------------
=2 2

=

Q3 - Q1 52-28

Co-efficient of Q.D = -------------------------- = -------- = 24/80
=0.3Q3 + Q1 52+28

Discrete Series

From the following data calculate Quartile deviation and it's co-efficient

x	26	28	32	35	29	24
f	6	7	9	10	7	4

Solution

X	f	cf
24	6	6
26	6	12 Q1
28	7	19
29	7	26
32	9	35 Q3
35	10	45
	N=∑f=45	

$$Q1 = \text{size of } \frac{N+1}{4} \text{ the item}$$

$$= \text{size of } \frac{45+1}{4} \text{th item} = 46/4$$

$$= \text{size of } 11.5^{th} \text{ item} = 26$$

$$Q3 = \text{size of } \frac{3(N+1)}{4} \text{th item}$$

$$= \frac{3(45+1)}{4} = 34.5^{th} \text{ Item} = 32$$

$$\text{Co-efficient of Q.D} = \frac{Q3 - Q1}{Q3 + Q1} = \frac{52-28}{52+28} = 24/80 = 0.3$$

Continuous series

From the following table. Compute the quartile deviation as well as its co-efficient

Size	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	36-40
Frequency	6	10	18	30	15	12	10	6	2

Solution

Weekly wages (x)	**No of workers f**	**cf**
4-8 8-12 12-16 16-20 20-24 24-28 28-32 32-34 34-40	6 10 18 30 15 12 10 6 2	4 16 34 64 79 91 101 107 109
	N==∑f=109	

Q1 = N/4 109/4 = 27.25

Q1 is lies between the class 12-16

N/4 cf

Q1 = L =

27.25 - 16

= 12 +.........................x 4
18

11.25
= 12 + ------------------4
18

= 12 + 45/ 18 = 121 + 2.5 = 14.5

Q1 = 14.5

Q3 = 3N /4 = 3(109)/4 = 81.75

Q3 lies between the class interval 24 – 28

$$Q3 = L + \frac{3N/4 - cf}{f} \times C$$

$$= 24 + \frac{81.75-79}{12} X 4$$

$$= 24 + 0.916$$

Q3 = 24.92

$$Q.D = \frac{Q3 =-Q1}{2} = \frac{24.92 \quad -14.5}{2} = \frac{10.42}{2} = 5.21$$

Mean Deviation
Individual Series

$$M.D = \frac{\sum I\ DI}{N}$$

$$\text{Co efficient of M.D} = \frac{M.D}{\text{Mean or Median or Mode}}$$

Calculate mean deviation a d its coefficient from Arthimetic mean for the following the data

100	150	200	250	300

Solution

X	**IDI = X-X**
100	100
150	50
200	0
250	50
300	100
∑x= 1000	**∑I DI**

Mean = ∑x / N = 1000/5 = 200

$$M.D = \frac{\sum I\ DI}{N} = 300 /5 = 60$$

Co-efficient of M.D = ----------------------- =--------- 0.3

Mean 200

Discrete Series

Find mean deviation from mediation and its co-efficient from the following data

X	10	11	12	13	14
f	3	12	18	12	3

Solution

X	**f**	**cf**
10	3	3
11	12	15
12	18	33
13	12	45
14	3	48

$$\text{Median} = \text{Size of } \frac{N+1}{2} \text{ th item}$$

$$= \text{Size of } \frac{48+1}{2} \text{ th item}$$

$$= \text{Size of } 24.5^{th} \text{ item} = 12$$

X	**f**	**IDI (X –Median)= x-12**	**fIDI**
10	3	2	6
11	12	1	12
12	18	0	0
13	12	1	12
14	3	2	6
			∑fI DI = 36

$$\textbf{Mean Deviation} = \frac{\sum f|D|}{N} = \frac{36}{48} = 0.75$$

M.D 0.75

Co – efficient of M.D = ------------- = -------- = 0.0625

Median 12

Continuous Series

Find the co-efficient of mean deviation from k mean for the following data

Age in years	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No of persons	20	25	32	40	42	35	10	8

Solution

Age in years	**m**	**No of persons f**	**D=m-A**	**fd**	**IDI=m-X**	**fIDI**
0-10	5	20	-30	-600	31.5	630.0
10-20	15	25	-20	-500	21.5	537.5
20-30	25	32	-10	-320	11.5	368.0
30-40	35	40	0	0	1.5	60.0
40-50	45	42	10	420	8.5	357.0
50-60	55	35	20	700	18.5	647.5
60-70	65	10	30	300	28.5	285.0
70-80	75	8	40	320	38.5	308.0
		N = 212		**∑fd=320**		**∑fI DI = 3193.0**

∑fd 320

¯X = A + ---------------- = 35 + ------------- = 35n + 1.5 =

36.5N 212

¯X = 36.5 Years

∑fd 3193

M.D = ------- = --------- = 15.06

N 212

M.D 15.06

Co-efficient of M.D = ----------------- = ------------- = 0.41years

Mean 36.5

Standard Deviation

Individual Series

Compute standard Deviation form the following data of the income of 10 employees of a firm

Monthly income	600	620	640	620	680	670	680	640	700	650

Solution

X	**X=X-¯X**	**x^2**
600	-50	2500
620	-30	900
640	-10	100
620	-30	900
680	30	900
670	20	400
680	30	900
640	-10	100
700	50	2500
650	0	0
∑X = 6500		**∑ X^2 == 9200**

$$\bar{X} = \frac{\sum X}{N} = \frac{6500}{10} = 650$$

σ = √∑ X^2

-------- = √9200 / 10 = 30.3 σ =

30.3N

Discrete Series

Calculate standard deviation from the following data

Marks (X)	10	20	30	40	50	60
No of Students(f)	8	12	20	10	7	3

------------------------- x CF

X	**f**	**fx**	**x = X-X¯= X-30.8**	**x^2**	**Fx^2**
10	8	80	-20.8	432.64	3461.12
20	12	240	-10.8	116.64	1399.68
30	20	600	-0.8	0.64	12.80
40	10	400	9.2	84.64	846.40
50	7	350	19.2	368.64	2580.48
60	3	180	29.2	852.64	2557.92
	N=60	**∑fX =1850**			**∑ Fx^2 =10858.40**

$$\bar{X} = \frac{\sum fX}{N} = \frac{1850}{60} = 30.5$$

σ = √∑ fX^2

-------- = $\sqrt{10858.40}$ / 60 = 13.5 σ = 13.5
N

Continuous series

Calculate standard deviation form the following data

Class	0-10	10-20	20-30	30-40	40-50
Frequency	5	8	15	16	6

Solution

Class	**Mid-point**	**frequency**	**d= $\frac{X-A}{C}$**	**d^2**	**fd**	**Fd^2**
0-10	5	5	-2	4	-10	20
10-20	15	8	-1	1	-8	8
20-30	25	15	0	0	0	0
30-40	35	16	1	1	16	16
40-50	45	6	2	4	12	24
		N=50			**$\sum fd=10$**	**$\sum fd^2=68$**

Assumed mean A= 25Class interval C = 10

Standard Deviation (σ) = $\sqrt{\sum fd^2 - (\sum fd)^2}$

SKEWNESS

Introduction

The term 'Skewness' refers to lack of symmetry, that is, when a distributionis to symmetrical it is called a skewed distribution. It the curve us normal or the data distributed symmetrically or uniformly. Spread will be the same on both sides of the cent repoint and the means median and mode will all have the same value.

Definition

'Skewness or symmetry is the attribute of a frequency distribution that extends furtheron one side of the class with the highest frequency on the other**--- Simpson and Kafka**

When a series is not symmetrical it is said to be asymmetrical or skewed **-Croxton and cowden**

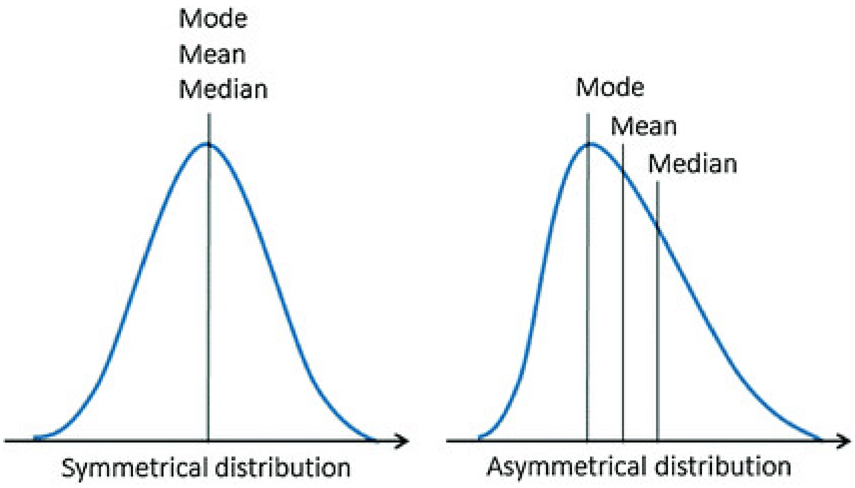
Skewness of a Distribution

When a distribution is not symmetrical it is called a skewed Distribution.
The analysis of presence of skewness in a distribution implies two main tasks. They are

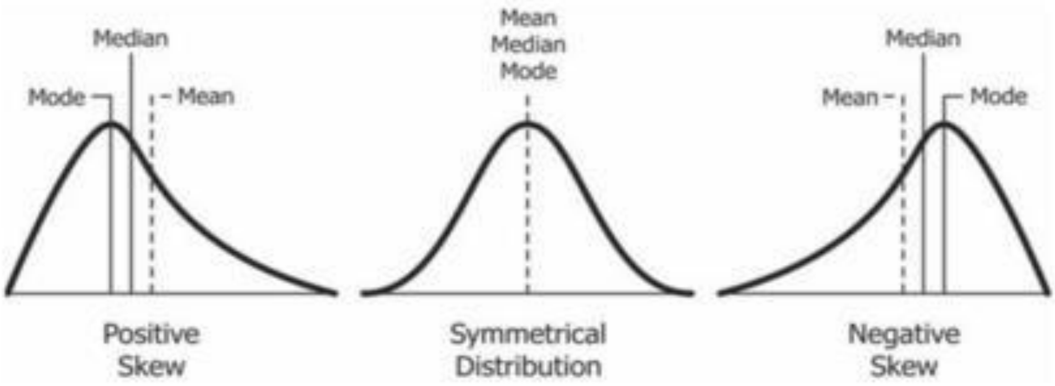
I. Determination of the sign of skewness and testing of skewness and
II. Determination of the extent of skewness

Symmetrical Distribution

In a symmetrical distribution, the values of mean, median and mode are coinciding. The spreadof the frequencies is the same on both sides of the centre point of the curve.



Skewed Distribution



A distribution which is not symmetrical is called a skewed distribution it is called skewed distribution. It may be either positively Skewed or negatively skewed Distribution

I) Positively Skewed Distribution

In a frequency distribution positively skewed distribution the curve has longer tail to the rights and its value of the mean is highest and the made is least. The median lies in between the two. That is $\bar{X} > M > Z$

II) Negatively Skewed Distribution

In a frequency distribution if the curve has long tail to the left then it is negatively skewed distribution in which value of mode is higher and mean is the least. The median lies in between the two. That is $\bar{X} < M < Z$

Various measures of Skewness

Skewness can be measured absolutely or relatively. Absolutely measures are called measures of skewness and relative measures are called the co-efficient of skewness

Absolute measures of skewness

i) The Karl Peason's Coefficient of Skewness
ii) The Bowley's Co efficient of Skewness

iii) The Kelly’s Coefficient of Skewness
iv) Measure of Skewness based on moments

Karl Pearson’s Co-efficient of Skewness

This method is based upon the difference between mean and mode and the difference is divided by standard deviation to give a relative measures.

Bowley’s Coefficient of Skewness

Bowelys measure is based on quartiles, in a symmetrical distribution first and third quartiles are equidistant from the median

Objectives of Skewness

I) To find out the direction and extent of asymmetry in a series.
II) To compare two or more series with regards to skewness.
III) To study the nature of variation of the items about the central value.

Karl Pearsons Coefficient of Skewness

Calculate Karl Pearson’s coefficient of skewness for the following data

25	15	23	40	27	25	23	25	20

Solution

Size of the Item	**Deviation d=X-A**	**D^2**
25 15 23 40	-2 --12 -4 13	4 144 16 169
27	0	0
25	-2	4
23	-4	16
25	-2	4
20	-7	49
	∑d = -20	∑d2 = 406

$$\bar{X} = A + \frac{\sum d}{N} = 27 + \frac{-29}{9} = 27 - 2.22 = 24.78$$

$$(\sigma) = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} = \sqrt{406 / 9 - (-20/9)^2}$$

$=\sqrt{45.11 - (2.2)^2} = 6.3$

In the given series, 25 is repeated three times
Mode is 25

$$SKp = \frac{\bar{X} - \text{Mode}}{\sigma} = \frac{24.78 - 25}{6.3} = 0.03$$

Bowley's Co-efficient of Skewness

Calculate the coefficient of Skewness

Age	0-10	10-20	20-30	30-40	40-50
No of Persons	8	11	26	9	6

Solution

Age	**No of Persons (f)**	**cf**
0-10	8	8
10-20	11	19
20-30	26	45
30-40	9	54
40-50	6	60
	N = 60	

N/4 = 60/4 = 15 lies between the class interval 10 – 20

$$Q1 = L + \frac{N/4 - cf}{F} \times C$$

$$= 10 + \frac{15 - 8}{11} \times 10 = 10 + \frac{7}{11} \times 10 = 10 + 6.36 = 16.36$$

Q3 = 3N /4 = 3 (60)/4 = 45 lies between the CL 20 -30

$$Q3 = L + \frac{3N/4 - cf}{F} \times C = 20 + \frac{45 - 19}{26} \times 10 = 20 + \frac{28}{26} \times 10$$

$=20 + 10 = 30$

N/2 = 60 /2 = 30 lies between the CL 20 - 30

$$\text{Median} = L + \frac{N/2 - cf}{F} \times C$$

$$= 20 + \frac{30 - 19}{26} \times 10 = 20 + \frac{11}{26} \times 10 = 20 + \frac{110}{26}$$

$= 20 + 4.23 = 24.23$

Bowley's Co – efficient of Skewness

$$SK_p = \frac{Q_3 + Q_1 - 2\text{Median}}{Q_3 - Q_1} = \frac{30 + 16.36 - 2(24.23)}{30 - 16.36} = \frac{48.38 - 48.46}{13.64} = \frac{-2.1}{13.64}$$

$= -0.15$

Kelly's Co efficient of Skewness

From the data given below calculate Kelly's co-efficient of skewnessMedian =130

P_{20} = 27

P_{90} =242

Solution

$$SK_p = \frac{P_{90} + P_{10} - 2\ \text{Median}}{P_{90} - P_{10}}$$

$$= \frac{242 + 27 - (2 \times 130)}{242 - 27}$$

$$= \frac{269 - 260}{215}$$

$$= \frac{9}{215} = 0.042$$

N N

$$= \sqrt{68/50 - (10/50)^2} \times 10$$

$$= \sqrt{1.36 - (0.2)^2} \times 10$$

= 11.49

Unit – IV – Correlation

Meaning

Correlation is the study of the natural relationship between two or more variables. Hence, it should be noted that the detection and analysis of correlation between two statistical variables requires relationship of some sort which associates the observation in pairs each of which is a value of the two variables

Definition

The relationship that exists between two variables ---Smith Correlation analysis deals with the association between two or more variables. ---Tuite

Uses of Correlation

I) Correlation is very useful in physical and social sciences. Business and economics
II) Correlation analysis is very useful in economics to study the relationship between price and demand
III) It is also useful in business to estimates costs, value, price and other related variables
IV) Correlation is the basis of the concept of regression
V) Correlation analysis help in calculation the sampling once.

Types of Correlation

- Positive correlation
- Negative Correlation
- Simple Correlation
- Multiple Correlations
- Partial Correlation
- Linear Correlation
- Non –Linear Correlation

Positive Correlation

Correlation is said to be positive when the values of two variables move in the same direction, so that an increase in the value of one variable is accompanied by an increase in the value of the other variable or a decrease in the value of one variable is followed by a decreasein the value of the other variable.

Negative Correlation

Correlation is said to be negative when the values of two variables move in opposite direction, so that an increase in the values of one variable is followed by a decrease in the value of the other and vice-versa

Simple Correlation

When only two variables are stated, it is said to be simple correlation

Multiple Correlations

When more than two variables are stated simultaneously, the correlation is said to be multiple

Partial Correlation

Partial correlation coefficient provides a measure of relationship between a dependent variable and a particular independent variable when all other variables involved are keptconstant analysis to yield and rainfall; it becomes a problem relating to simple correlation Linear Correlation. The correlation is said to be linear, if the amount of change is one variable tends to beara constant ratio to the amount of change in the other

Non Linear Correlation

The correlation is nonlinear, if the amount of change in one variable does not bear aconstant ratio to the amount of change in the other related variable.

Methods of studying correlation
Graphical method

- Scatter diagram
- Simple graph method

Mathematical Methods

- Karl Pearson's Co-efficient of correlation
- Spearman's Rank Correlation coefficient
- Concurrent deviation method
- Method of least square

Scatter diagram method

It is a method of studying correlation between two related variables. The two variablesX and Y will be taken upon the X and Y axes of a graph paper. For each part of X and Y values, we mark a dot and we go as many points as the numbers of observation.

Graphical method

In this method curves are drawn for separate series on a graph paper. By examining the direction and closeness of the two curves we can offer whether prompt variances are related. If both the curves are moving in the same direction correlation is said to be positive. On the contrary, if the curves are moving in the opposite directions is said to be negative.

Karl Pearson's Co-efficient of correlation

Karl Pearson, a great statistician introduced a mathematical method for measuring the magnitude of relationship between two variables. This method. Known as Pearson Coefficient of correlation is widely used. It is denoted by the symbol "r'

Spearman's Rank Correlation Co-efficient

In 1904, a famous British psychologist Charles Edward Spearman found out the method of Co-efficient of correlation of rank. Rank correlation is applicable to individual observation. This measure is useful in dealing with qualitative characteristics. The result, by using ranking method, is only approximate.

Concurrent deviation method

Under this method, the direction of change in X variable and y variable is taken into account to find out the deviation for each term the change in the value of the variable form its preceding value which may be + or –

Co-efficient of correlation

Find the Karl Pearson's Coefficient of Correlation

X	6	2	10	4	8
Y	9	11	5	8	7

Solution

X	**Y**	**X^2**	**Y^2**	**XY**
6	9	36	81	54
2	11	4	121	22
10	5	100	25	50
4	8	16	64	32
8	7	64	49	56
30	40	220	340	214

Coefficient of Correlation $= \dfrac{N\sum XY - (\sum X)(\sum y)}{\sqrt{N\sum X^2 - (\sum x)^2} \cdot \sqrt{N\sum Y^2 - (\sum Y)^2}}$

$$r = \frac{(5 \times 214) - (30 \times 40)}{\sqrt{5 \times 220 - (30)^2}\ \sqrt{5 \times 340 - (40)^2}}$$

$$= \frac{1070 - 1200}{\sqrt{1100-900}\ \sqrt{1700-1600}} = \frac{-130}{\sqrt{200}\ \sqrt{100}} = \frac{-130}{14.14(10)} = -0.9194$$

Karl Pearson's Coefficient of Correlation:

Karl Pearson's method of calculating coefficient of correlation is based on the covariance of the two variables in a series. This method is widely used in practice and the coefficient of correlation is denoted by the symbol "***r***". If the two variables under study are X and Y, the following formula suggested by Karl Pearson can be used for measuring the degree of relationship of correlation.

$$r = \frac{Covariance\ (x,y)}{S.D.(x)S.D.(y)}$$

$$r = \frac{Cov(X,Y)}{\sigma_x\ \sigma_y}$$

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}} \quad \text{where } X = x - \bar{x},\ Y = y - \bar{y}$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

$$r = \frac{\Sigma(X-\overline{X})(Y-\overline{Y})}{\sqrt{\Sigma(X-\overline{X})^2}\sqrt{(Y-\overline{Y})^2}}$$

Where, $\overline{X}$ = mean of X variable

$\overline{Y}$ = mean of Y variable

$$r = \frac{\Sigma f(dx)(dy) - \frac{\Sigma fdx)(\Sigma fdy)}{N}}{\sqrt{\Sigma(fdx)\wedge 2 - \frac{(\Sigma fdx)\wedge 2}{N}}\sqrt{\Sigma(dy)\wedge 2 - \frac{(\Sigma fdy)\wedge 2}{N}}}$$

$d_x = X - A$

$d_y = Y - A$

Above different formula's can be used in different situation depending upon the information given in the problem.

From following information find the correlation coefficient between advertisement expenses and sales volume using Karl Pearson's coefficient of correlation method.

Firm	1	2	3	4	5	6	7	8	9	10
Advertisement Exp. (Rs. In Lakhs)	11	13	14	16	16	15	15	14	13	13
Sales Volume (Rs. In Lakhs)	50	50	55	60	65	65	65	60	60	50

Solution:

Let us assume that advertisement expenses are variable X and sales volume arevariable Y.

Calculation of Karl Pearson's coefficient of correlation

Firm	X	Y	$x = X - \dot{X}$	x^2	$y = Y - \dot{Y}$	y^2	xy
1	11	50	-3	9	-8	64	24
2	13	50	-1	1	-8	64	8
3	14	55	0	0	-3	9	0
4	16	60	2	4	2	4	4
5	16	65	2	4	7	49	14
6	15	65	1	1	7	49	7
7	15	65	1	1	7	49	7
8	14	60	0	0	2	4	0
9	13	60	-1	1	2	4	-2
10	13	50	-1	1	-8	64	8
	140	**580**		**22**		**360**	**70**
	$\sum X$	**$\sum Y$**		**$\sum x2$**		**$\sum y2$**	**$\sum xy$**

$\dot{X} = \frac{\sum X}{n} = \frac{140}{10} = 14$ $\quad \dot{Y} = \frac{\sum Y}{n} = \frac{580}{10} = 58$

$$\boldsymbol{r} = \frac{\sum xy}{\sqrt{\sum x2 \sum y2}} = \frac{70}{\sqrt{22*360}} = \frac{70}{88.9944} = \mathbf{0.7866}$$

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}} \quad \text{where } X = x - \bar{x},\ Y = y - \bar{y}$$

Interpretation: From the above calculation it is very clear that there is high degree of **positive correlation** i.e. ***r* = 0.7866,** between the two variables. i.e. Increase in advertisement expenses leads to increased sales volume.

Illustration 01:

Find the correlation coefficient between age and playing habits of the following students using Karl Pearson's coefficient of correlation method.

Age	15	16	17	18	19	20
Number of students	250	200	150	120	100	80
Regular Players	200	150	90	48	30	12

Solution:

To find the correlation between age and playing habits of the students, we need tocompute the percentages of students who are having the playing habit.

Percentage of playing habits = No. of Regular Players / Total No. of Students * 100

Now, let us assume that ages of the students are variable X and percentages of playing habits are variable Y.

Calculation of Karl Pearson’s coefficient of correlation

Age (X)	No of Students	Regular Players	Percentage of Playing Habits (Y)	X - Ẋ	(X - Ẋ)^2	Y - Ẏ	(Y - Ẏ)^2	(X - Ẋ)(Y - Ẏ)
15	250	200	80	-2.5	6.25	30	900	-75
16	200	150	75	-1.5	2.25	25	625	-37.5
17	150	90	60	-0.5	0.25	10	100	-5
18	120	48	40	0.5	0.25	-10	100	-5
19	100	30	30	1.5	2.25	-20	400	-30
20	80	12	15	2.5	6.25	-35	1225	-87.5
105			**300**		**17.5**		**3350**	**-240**
∑X			**∑Y**		**∑x2**		**∑y2**	**∑xy**

$$\dot{X} = \frac{\sum X}{n} = \frac{105}{6} = 17.5 \qquad \dot{Y} = \frac{\sum Y}{n} = \frac{300}{6} = 50$$

$$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2}\sqrt{(Y-\bar{Y})^2}}$$

$$r = \frac{\sum(X-\dot{X})(Y-\dot{Y})}{\sqrt{\sum(X-X)2 \sum(Y-Y)2}} = \frac{-240}{\sqrt{17.5*3350}} = \frac{-240}{242.126} = \mathbf{-0.9912}$$

Interpretation: From the above calculation it is very clear that there is high degree of **negative correlation** i.e. ***r* = -0.9912,** between the two variables of age and playing habits. i.e. Playing habits among students decreases when their age increases.

Illustration 02:

Find Karl Pearson’s coefficient of correlation between capital employed and profit obtained from the following data.

Capital Employed (Rs. In Crore)	10	20	30	40	50	60	70	80	90	100
Profit (Rs. In Crore)	2	4	8	5	10	15	14	20	22	50

Solution:

Let us assume that capital employed is variable X and profit is variable Y.

Calculation of Karl Pearson’s coefficient of correlation

X	Y	X^2	Y^2	XY
10	2	100	4	20
20	4	400	16	80
30	8	900	64	240
40	5	1600	25	200
50	10	2500	100	500
60	15	3600	225	900
70	14	4900	196	980
80	20	6400	400	1600
90	22	8100	484	1980
100	50	10000	2500	5000
550	**150**	**38500**	**4014**	**11500**
∑X	**∑Y**	**$∑X^2$**	**$∑Y^2$**	**∑XY**

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n(\sum X2) - (\sum X)2][n(\sum Y2) - (\sum Y)2]}}$$

$$r = \frac{(10 * 11500) - (550 * 150)}{\sqrt{[(10*38500)-(550^2)]\ [(10*4014)-(150^2)]}}$$

$$r = \frac{(1,15,000) - (82,500)}{\sqrt{[(3,85,000)-(3,02,500)]\ [(40,140)-(22,500)]}}$$

$$r = \frac{32,500}{\sqrt{(82,500)\ (17,640)}} = \frac{32,500}{\sqrt{1455300000}}$$

$$r = \frac{32,500}{38148} = \mathbf{0.8519}$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2]\ [n(\sum y^2) - (\sum y)^2]}}$$

Illustration 03:

A computer while calculating the correlation coefficient between the variable X and Y obtained the following results:

N = 30; $\sum X = 120$ $\sum X^2 = 600$ $\sum Y = 90$ $\sum Y^2 = 250$ $\sum XY = 335$

It was, however, later discovered at the time of checking that it had copied down two pairs of observations as: (X, Y) : (8, 10) (12, 7)

While the correct values were: (X, Y) : (8, 12) (10, 8)

Obtain the correct value of the correlation coefficient between X and Y.

Solution:

Correct $\sum X$ = 120 – 8 – 12 + 8 + 10 = 118

Correct $\sum X^2$ = $600 - 8^2 - 12^2 + 8^2 + 10^2$

= 600 – 64 – 144 + 64 + 100 = 556

Correct $\sum Y$ = 90 – 10 – 7 + 12 + 8 = 93

Correct $\sum Y^2$ = $250 - 10^2 - 7^2 + 12^2 + 8^2$

= $250 - 100 - 49 + 144 + 64$ = 309

Correct $\sum XY$ = $335 - (8*10) - (12*7) + (8*12) + (10*8)$

= $335 - 80 - 84 + 96 + 80$ = 347

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n(\sum X2) - (\sum X)2][n(\sum Y2) - (\sum Y)2]}}$$

$$r = \frac{(30 * 347) - (118 * 93)}{\sqrt{[(30*556)-(118^2)][(30*309)-(93^2)]}}$$

$$r = \frac{(10,410) - (10,974)}{\sqrt{[(16,680)-(13,924)][(9270)-(8649)]}}$$

$$r = \frac{-564}{\sqrt{(2,756)(621)}} = \frac{-564}{\sqrt{1711476}}$$

$$r = \frac{-564}{1308.2339} = \mathbf{-0.4311}$$

Therefore, the correct value of correlation coefficient between X and Y is moderately negative correlation of -0.4311.

Illustration 04:

Coefficient of correlation between X and Y is 0.3. Their covariance is 9. The variance ofX is 16. Find the standard devotion of Y series.

Solution:

Given information:

r = 0.3 Cov (X, Y) = 9 Var (X) = 16

$$r = \frac{Co(X,Y)}{\sqrt{Var(X) * Var(Y)}} \qquad 0.3 = \frac{9}{\sqrt{16 * Var(Y)}} \qquad 0.3 = \frac{9}{4 * \sqrt{Var}(Y)}$$

$$0.3 * 4 = \frac{9}{SD(Y)} \qquad 1.2 = \frac{9}{SD(Y)} \qquad SD(Y) = \frac{9}{1.2} = 7.5$$

Therefore the standard deviation of Y series = σ(Y) = 7.5

Illustration 05:

Calculate correlation coefficient from the following two-way table, with X representing the average salary of families selected at random in a given area and Y representing the average expenditure on entertainment.

Expenditure on Entertainment (Rs. '000)	**Average Salary (Rs. '000)** 100-150	150-200	200-250	250-300	300-350
0 – 10	5	4	5	2	4
10 – 20	2	7	3	7	1
20 – 30	-	6	-	4	5
30 – 40	8	-	4	-	8
40 – 50	-	7	3	5	10

Solution:

Let us assume that Average Salary is variable X and Expenditure on Entertainment is variable Y.

In case of grouped data, we need to follow the assumed mean method to calculate Karl Pearson's Coefficient of Correlation. Following steps are followed to compute correlation.

1. Identify the mid-point of the class intervals for variable X and Y.
2. Chose an assumed mean from the mid-point identified above for both X and Y.
3. To simplify further, deviation from assumed mean is computed by dividing deviation by a common factor.
4. Add the values in cell, row-wise and column-wise, to compute frequencies (f). Sum of either row-wise or column-wise represent the value of N.
5. Obtain the product of d_X and d_Y and the corresponding frequencies (f) in each cell. Write the figure thus obtained in the right corner of each cell which represent the value of fd_Xd_Y.

Y	**X**	**100 - 150**	**150 – 200**	**200 – 250**	**250 – 300**	**300 – 350**	**f**	**dy**	**fdy**	**fdy^2**	**fdxdy**
	Mid Point	**125**	**175**	**225**	**275**	**325**					
0 – 10	5	**20** **5**	**8** **4**	**0** **5**	**-4** **2**	**-16** **4**	20	-2	-40	80	8

10 – 20	**15**	**4** **2**	**7** **7**	**0** **3**	**-7** **7**	**-2** **1**	20	-1	-20	20	2
20 – 30	**25**	**-** **-**	**0** **6**	**-** **-**	**0** **4**	**0** **5**	15	0	0	0	0
30 – 40	**35**	**-16** **8**	**-** **-**	**0** **4**	**-** **-**	**16** **8**	20	1	20	20	0
40 – 50	**45**	**-** **-**	**-14** **7**	**0** **3**	**10** **5**	**40** **10**	25	2	50	100	36
f		15	24	15	18	28	**100 = N**		**10**	**220**	**46**
d_X		-2	-1	0	1	2			**∑fdy**	**∑fdy 2**	**∑fdx dy**
fd_X		-30	-24	0	18	56	20	**∑fdx**			
fd_X^2		60	24	0	18	112	214	**$∑fdx^2$**			
fd_Xdy		8	1	0	-1	38	46	**∑fdxdy**			

d_X = Mid Point of Series X – Assumed Mean of Series X = MP(X) - 225dy

= Mid Point of Series Y – Assumed Mean of Series Y = MP(Y) - 25

$$r = \frac{n\sum fdxdy - \sum fdx \sum fdy}{\sqrt{[n(\sum fdx^2) - (\sum fdx)^2][n(\sum fdy^2) - (\sum fdy)^2]}} = \frac{(100*46) - (20*10)}{\sqrt{[(100*214) - (20)^2][(100*220) - (10)^2]}}$$

$$r = \frac{(4{,}600) - (200)}{\sqrt{[21{,}400 - 400][22{,}000 - 100]}} = \frac{4{,}400}{\sqrt{[21{,}000]*[21{,}900]}} = \frac{4{,}400}{21{,}445.2792} = \mathbf{0.2052}$$

Interpretation: From the above calculation it is very clear that there is low degree of **positive correlation** i.e. ***r* = 0.2052,** between the two variables of salary and expenditure. It means average salary of income have slightly or low influence over entertainment expenditure.

Rank Correlation

Spearman's Rank Coefficient of Correlation:

When quantification of variables becomes difficult such beauty of female, leadership ability, knowledge of person etc, then this method of rank correlation is useful which was developed by British psychologist Charles Edward Spearman in 1904. In this method ranks are allotted to each element either in ascending or descending order. The correlation coefficient between these allotted two series of ranks is popularly called as "Spearman's Rank Correlation" and denoted by "***R***".

To find out correlation under this method, the following formula is used.

$$\mathbf{R = 1 -} \frac{6\sum D^2}{N^3 - N}$$

where, D =Difference of the ranks between paired items in two series.

N = Number of pairs of ranks

In case of tie in ranks or equal ranks:

In some cases it may be possible that it becomes necessary to assign same rank to two or more elements or individual or entries. In such situation, it is customary to give each individual or entry an average rank. For example, if two individuals are ranked equal to 5^{th} place, then both of them are allotted with common rank (5+6)/2 =

5.5 and if three are ranked in 5^{th} place, then they are given the rank of (5+6+7)/3 = 6. It means where two or more individuals are to be ranked equal, the rank assigned for the purpose of calculating coefficient of correlation is the average of the ranks with these individual or items or entries would have got had they differed slightly with eachother.

Where equal ranks are assigned to some entries, an adjustment factor is to be added to the value of $6\sum D^2$ in the above formula for calculating the rank coefficient correlation. This adjustment factor is to be added for every repetition of rank.

Adjustment factor = $\frac{1}{12}(m_1^3 - m_1)$ where, m = number of items whose rank are common

For example, if a particular rank repeated two times then m=2 and if it repeats three times then m= 3 and so on.

Hence the above formula can be re-written as follows:

$$R = 1 - \frac{6 * [\sum D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + \ldots\ldots]}{N^3 - N}$$

Illustration 01:

Find out spearman's coefficient of correlation between the two kinds of assessment of graduate students' performance in a college.

Name of students	A	B	C	D	E	F	G	H	I
Internal Exam	51	68	73	46	50	65	47	38	60
External Exam	49	72	74	44	58	66	50	30	35

Solution:

Calculation of Spearman's Rank Coefficient of Correlation

Name	**Internal Exam**	**Ranks (R_1)**	**External Exam**	**Ranks (R_2)**	**D = $R_1 - R_2$**	**D^2**
A	51	5	49	6	-1	1
B	68	2	72	2	0	0

C	73	1	74	1	0	0
D	46	8	44	7	1	1
E	50	6	58	4	2	4
F	65	3	66	3	0	0
G	47	7	50	5	2	4
H	36	9	30	9	0	0
I	60	4	35	8	-4	16
$\sum D^2$ =						**26**

$$R = 1 - \frac{6\sum D^2}{N^3 - N} = 1 - \frac{6*26}{9^3 - 9} = 1 - \frac{156}{729 - 9} = 1 - \frac{156}{720} = 1 - 0.2167 = 0.7833$$

Interpretation: From the above calculation it is very clear that there is high degree of **positive correlation** i.e. **R = 0.7833,** between two exams. It means there is a high degree of positive correlation between the internal exam and external exam of the students.

Illustration 02:

The coefficient of rank correlation of the marks obtained by 10 students in statistics and accountancy was found to be 0.8. It was later discovered that the difference in ranks in the two subjects obtained by one of the students was wrongly taken as 7 instead of 9. Find the correct coefficient of rank correlation.

Solution:

$$R = 1 - \frac{6\sum D^2}{N^3 - N} \Rightarrow 0.8 = 1 - \frac{6\sum D^2}{10^3 - 10} \Rightarrow 0.8 = 1 - \frac{6\sum D^2}{990} \Rightarrow \frac{6\sum D^2}{990} = 1 - 0.8 \Rightarrow$$

$$\frac{6\sum D^2}{990} = 0.2 \Rightarrow 6\sum D^2 = 0.2 * 990 \Rightarrow \sum D^2 = 198/6 \Rightarrow \sum D^2 = 33$$

But this is not correct $\sum D^2$ therefore we need to compute correct value Correct

$\sum D^2 = 33 - 7^2 + 9^2 = 65$

Hence, correct value of rank coefficient of correlation is:

$$R = 1 - \frac{6\sum D^2}{N^3 - N} = 1 - \frac{6 * 65}{990} = 1 - \frac{390}{990} = 1 - 0.394 = \mathbf{0.606}$$

Illustration 03:

Ten competitors in a beauty contest are ranked by three judges in the following order:

1st Judge	1	6	5	10	3	2	4	9	7	8
2nd Judge	3	5	8	4	7	10	2	1	6	9
3rd Judge	6	4	9	8	1	2	3	10	5	7

Use the rank correlation coefficient to determine which pairs of judges has the nearest approach to common tastes in beauty.

Solution:

In order to find out which pair of judges has the nearest approach to common tastes in beauty, we compare rank correlation between the judgements of

1. 1st Judge and 2nd Judge
2. 2nd Judge and 3rd Judge
3. 1st Judge and 3rd Judge

Calculation of Spearman's Rank Coefficient of Correlation

Rank by 1st Judge (R_1)	**Rank by 2nd Judge (R_2)**	**Rank by 3rd Judge (R_3)**	**$D^2 = (R_1 - R_2)^2$**	**$D^2 = (R_2 - R_3)^2$**	**$D^2 = (R_1 - R_3)^2$**
1	3	6	4	9	25
6	5	4	1	1	4
5	8	9	9	1	16
10	4	8	36	16	4
3	7	1	16	36	4
2	10	2	64	64	0
4	2	3	4	1	1
9	1	10	64	81	1
7	6	5	1	1	4
8	9	7	1	4	1
N = 10	**N = 10**	**N = 10**	**$\sum D^2 = 200$**	**$\sum D^2 = 214$**	**$\sum D^2 = 60$**

1. 1st Judge and 2nd Judge: $R = 1 - \frac{6\sum D^2}{N^3 - N} = 1 - \frac{6*200}{10^3 - 10} = 1 - \frac{1200}{990} = 1 - 1.2121 = \mathbf{-0.2121}$

2. 2nd Judge and 3rd Judge: $R = 1 - \frac{6\sum D^2}{N^3 - N} = 1 - \frac{6*214}{10^3 - 10} = 1 - \frac{1284}{990} = 1 - 1.297 = \mathbf{-0.297}$

3. 1st Judge and 3rd Judge: $R = 1 - \frac{6\sum D^2}{N^3 - N} = 1 - \frac{6*60}{10^3 - 10}$

$$= 1 - \frac{360}{990}$$

$$= 1 - 0.3636$$

$$= \mathbf{0.6364}$$

Interpretation: From the above calculation it can be observed that coefficient of correlation is positive in the judgement of the first and third judges. Therefore, it can be concluded that first and third judges have the nearest approach to common tastes in beauty.

Illustration 04:

From the following data, compute the rank correlation.

X	82	68	75	61	68	73	85	68
Y	81	71	71	68	62	69	80	70

Solution:

In the problem we find there are repetitions of ranks. Value of X = 68 repeated 3 times and Value of Y = 71 repeated 2 times. Therefore we need to compute adjustment factor to be added to the value of $\sum D^2$.

Calculation of Spearman’s Rank Coefficient of Correlation

X	**Y**	**R1**	**R2**	**D= R1 – R2**	**D^2**
82	81	2	1	1	1
68	71	6	3.5	2.5	6.25
75	71	3	3.5	-0.5	0.25
61	68	8	7	1	1
68	62	6	8	-2	4
73	69	4	6	-2	4
85	80	1	2	-1	1
68	70	6	5	1	1
$\sum D^2$					**18.5**

$$R = 1 - \frac{6 * [\sum D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m)]}{N^3 - N}$$

When value X repeated three times, m=3,

Adjustment factor (1) = $\frac{1}{12}(3^3\text{-}3) = \frac{1}{12} * (27\text{-}3) = \frac{1}{12} * 24 = 2$

When value Y repeated two times, m=2,

Adjustment factor (2) = $\frac{1}{12}(2^3\text{-}2) = \frac{1}{12} * (8\text{-}2) = \frac{1}{12} * 6 = 0.5$

$$R = 1 - \frac{6 * [18.5 + 2 + 0.5]}{8^3 - 8} = 1 - \frac{6 * 21}{512 - 8} = 1 - \frac{126}{504} = 1 - 0.25 = \mathbf{0.75}$$

Spearman's Rank Coefficient of Correlation = 0.75, which indicates there is highdegree of positive correlation.

Properties of Coefficient of Correlation:

1. The coefficient of correlation always lies between – 1 to +1, symbolically it can written as $-1 \leq r \leq 1$.
2. The coefficient of correlation is independent of change of origin and scale.
3. The coefficient of correlation is a pure number and is independent of the units of measurement. It means if X represent say height in inches and Y represent say weights in kgs, then the correlation coefficient will be neither in inches nor in kgs but only a pure number.
4. The coefficient of correlation is the geometric mean of two regression coefficient, symbolically $r = \sqrt{bxy * byx}$
5. If X and Y are independent variables then coefficient of correlation is zero.

Two judges ina beauty contest rank the 12 entries as follows

X	1	6	5	10	3	2	4	9	7	8
Y	6	4	9	8	1	2	3	10	5	7

Rank X	Rank Y	D=R(X)-(Y)	D^2
1	6	-5	25
6	4	2	4
5	9	-4	16
10	8	2	4
3	1	2	4
2	2	0	0
4	3	1	1
9	10	-1	1

7 8	5 7	2 1	4 1
N=10			$\sum D^2 = 60$

Regression

A study of measuring the relationship between associated variables, wherein one variable is dependent on another independent variable, called as Regression. It is developed by Sir Francis Galton in 1877 to measure the relationship of height betweenparents and their children.

Regression analysis is a statistical tool to study the nature and extent of functional relationship between two or more variables and to estimate (or predict) the unknown values of dependent variable from the known values of independent variable.

The variable that forms the basis for predicting another variable is known as the Independent Variable and the variable that is predicted is known as dependent variable. For example, if we know that two variables price (X) and demand (Y) are closely related we can find out the most probable value of X for a given value of Y or the most probable value of Y for a given value of X. Similarly, if we know that the amount of tax and the rise in the price of a commodity are closely related, we can find out the expected price for a certain amount of tax levy.

Uses of Regression Analysis:

1. It provides estimates of values of the dependent variables from values of independent variables.
2. It is used to obtain a measure of the error involved in using the regression line asa basis for estimation.
3. With the help of regression analysis, we can obtain a measure of degree of association or correlation that exists between the two variables.
4. It is highly valuable tool in economies and business research, since most of theproblems of the economic analysis are based on cause and effect relationship.

Distinction between Correlation and Regression

Sl No	Correlation	Regression
1	It measures the degree and direction of relationship between the variables.	It measures the nature and extent of average relationship between two or more variables in terms of the originalunits of the data
2	It is a relative measure showing association between the variables.	It is an absolute measure of relationship.

3	Correlation Coefficient is independent of change of both origin and scale.	Regression Coefficient is independent of change of origin but not scale.
4	Correlation Coefficient is independent of units of measurement.	Regression Coefficient is not independent of units of measurement.
5	Expression of the relationship between the variables ranges from –1 to +1.	Expression of the relationship between the variables may be in any of the forms like: Y = a + bX Y = a + bX + cX^2
6	It is not a forecasting device.	It is a forecasting device which can be used to predict the value of dependent variable from the given value of independent variable.
7	There may be zero correlation such as weight of wife and income of husband.	There is nothing like zero regression.

Regression Lines and Regression Equation:

Regression lines and regression equations are used synonymously. Regression equations are algebraic expression of the regression lines. Let us consider two variables: X & Y. If y depends on x, then the result comes in the form of simple regression. If we take the case of two variable X and Y, we shall have two regression lines as the regression line of X on Y and regression line of Y on X. The regression line of Y on X gives the most probable value of Y for given value of X and the regression line of X on Y given the most probable value of X for given value of Y. Thus, we have two regression lines. However, when there is either perfect positive or perfect negative correlation between the two variables, the two regression line will coincide, i.e. wewill have one line. If the variables are independent, r is zero and the lines of regression are at right angles i.e. parallel to X axis and Y axis.

Therefore, with the help of simple linear regression model we have the following two regression lines

1. Regression line of Y on X: This line gives the probable value of Y (Dependent variable) for any given value of X (Independent variable).

   Regression line of Y on X : $Y - \dot{Y} = b_{yx} (X - \dot{X})$

   OR : Y = a + bX

2. Regression line of X on Y: This line gives the probable value of X (Dependent variable) for

any given value of Y (Independent variable).

Regression line of X on Y : $X - \dot{X} = b_{xy} (Y - \dot{Y})$

OR : $X = a + bY$

In the above two regression lines or regression equations, there are two regression parameters, which are “a” and “b”. Here “a” is unknown constant and “b” which is also denoted as “b_{yx}” or “b_{xy}”, is also another unknown constant popularly called as regression coefficient. Hence, these “a” and “b” are two unknown constants (fixed numerical values) which determine the position of the line completely. If the value of either or both of them is changed, another line is determined. The parameter “a” determines the level of the fitted line (i.e. the distance of the line directly above or below the origin). The parameter “b” determines the slope of the line (i.e. the changein Y for unit change in X).

If the values of constants “a” and “b” are obtained, the line is completely determined. But the question is how to obtain these values. The answer is provided by the method of least squares. With the little algebra and differential calculus, it can be shown that the following two ***normal equations***, if solved simultaneously, will yield the values of the parameters “a” and “b”.

Two normal equations:

X on Y	**Y on X**
$\sum X = Na + b\sum Y$	$\sum Y = Na + b\sum X$
$\sum XY = a\sum Y + b\sum Y^2$	$\sum XY = a\sum X + b\sum X^2$

This above method is popularly known as direct method, which becomes quite cumbersome when the values of X and Y are large. This work can be simplified if instead of dealing with actual values of X and Y, we take the deviations of X and Yseries from their respective means. In that case:

Regression equation Y on X:

$Y = a + bX$ will change to $(Y - \dot{Y}) = b_{yx} (X - \dot{X})$

Regression equation X on Y:

$X = a + bY$ will change to $(X - \dot{X}) = b_{xy} (Y - \dot{Y})$

In this new form of regression equation, we need to compute only one parameter i.e. “b”. This “b” which is also denoted either “b_{yx}” or “b_{xy}” which is called as regression coefficient.

Regression Coefficient:

The quantity “b” in the regression equation is called as the regression coefficient or slope coefficient. Since there are two regression equations, therefore, wehave two regression coefficients.

1. Regression Coefficient X on Y, symbolically written as “b_{xy}”
2. Regression Coefficient Y on X, symbolically written as “b_{yx}”

Different formula’s used to compute regression coefficients:

Method	**Regression Coefficient X on Y**	**Regression Coefficient Y on X**
Using the correlation coefficient (r) and standard deviation (σ)	$b_{xy} = r \frac{\sigma x}{\sigma y}$	$b_{yx} = r \frac{\sigma y}{\sigma x}$
Direct Method: Using sum of X and Y	$b_{xy} = \frac{N\sum XY - \sum X \sum Y}{N\sum Y^2 - (\sum Y)^2}$	$b_{yx} = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2}$
When deviations are taken from arithmetic mean	$b_{xy} = \frac{\sum xy}{\sum y^2}$ where $x = X - \dot{X}$ and $y = Y - \dot{Y}$	$b_{yx} = \frac{\sum xy}{\sum x^2}$ where $x = X - \dot{X}$ and $y = Y - \dot{Y}$

Properties of Regression Coefficients:

1. The coefficient of correlation is the geometric mean of the two regression coefficients. Symbolically $r = \sqrt{b_{xy} * b_{yx}}$
2. If one of the regression coefficients is greater than unity, the other must be less than unity, since the value of the coefficient of correlation cannot exceed unity. For example if bxy = 1.2 and byx = 1.4 “r” would be = $\sqrt{1.2 * 1.4} = 1.29$, which is not possible.
3. Both the regression coefficient will have the same sign. i.e. they will be either positive or negative. In other words, it is not possible that one of the regression coefficients is having minus sign and the other plus sign.
4. The coefficient of correlation will have the same sign as that of regression coefficient, i.e. if regression coefficient have a negative sign, “r” will also have negative sign and if the regression coefficient have a positive sign, “r” would also be positive. For example, if bxy = -0.2 and byx = -0.8 then $r = -\sqrt{0.2 * 0.8} =$ **– 0.4**
5. The average value of the two regression coefficient would be greater than the value of coefficient of correlation. In symbol (bxy + byx) / 2 > r. For example, if bxy = 0.8 and byx = 0.4 then average of the two values = (0.8 + 0.4) / 2 = 0.6 and the value of r = r = $\sqrt{0.8 * 0.4} = 0.566$ which less than 0.6
6. Regression coefficients are independent of change of origin but not scale.

Illustration 01:

Find the two regression equation of X on Y and Y on X from the following data:

X	:	10	12	16	11	15	14	20	22
Y	:	15	18	23	14	20	17	25	28

Solution:

Calculation of Regression Equation

X	**Y**	X^2	Y^2	**XY**
10	15	100	225	150
12	18	144	324	216
16	23	256	529	368
11	14	121	196	154
15	20	225	400	300
14	17	196	289	238
20	25	400	625	500
22	28	484	784	616
120	**160**	**1,926**	**3,372**	**2,542**
$\sum X$	$\sum Y$	$\sum X^2$	$\sum Y^2$	$\sum XY$

Here N = Number of elements in either series X or series Y = 8

Now we will proceed to compute regression equations using normal equations.

Regression equation of X on Y: **X = a + bY**

The two normal equations are:

$$\sum X = Na + b\sum Y$$

$$\sum XY = a\sum Y + b\sum Y^2$$

Substituting the values in above normal equations, we get

$$120 = 8a + 160b \quad \text{..... (i)}$$

$$2542 = 160a + 3372b \quad \text{..... (ii)}$$

Let us solve these equations (i) and (ii) by simultaneous equation method Multiply equation (i) by 20 we get 2400 = 160a + 3200b

Now rewriting these equations:

$$2400 = 160a + 3200b$$

$$2542 = 160a + 3372b$$

$$(-) \qquad (-) \qquad (-)$$

-142	=	-172b

Therefore now we have -142 = -172b, this can rewritten as 172b = 142

Now, $b = \frac{142}{172} = 0.8256$ (rounded off)

Substituting the value of b in equation (i), we get

120	=	8a	+	(160 * 0.8256)
120	=	8a	+	132 (rounded off)
8a	=	120	-	132
8a	=	-12		
a	=	-12/8		
a	=	-1.5		

Thus we got the values of a = -1.5 and b = 0.8256 Hence the required regression equation of X on Y:

X = a + bY => **X = -1.5 + 0.8256Y**

Regression equation of Y on X: Y = a + bX

The two normal equations are:

$$\sum Y = Na + b\sum X$$

$$\sum XY = a\sum X + b\sum X^2$$

Substituting the values in above normal equations, we get

160	=	8a	+	120b	..... (iii)
2542	=	120a	+	1926b	..... (iv)

Let us solve these equations (iii) and (iv) by simultaneous equation method Multiply equation (iii) by 15 we get 2400 = 120a + 1800b

Now rewriting these equations:

2400	=	120a	+	1800b
2542	=	120a	+	1926b
(-)		(-)		(-)
-142	=			-126b

Therefore now we have -142 = -126b, this can rewritten as 126b = 142

Now, $b = \frac{142}{126} = 1.127$ (rounded off)

Substituting the value of b in equation (iii), we get

160	=	8a	+	(120 * 1.127)
160	=	8a	+	135.24
8a	=	160	-	135.24
8a	=	24.76		
a	=	24.76/8		
a	=	3.095		

Thus we got the values of a = 3.095 and b = 1.127 Hence the required regression equation of Y on X:

Y = a + bX => **Y = 3.095 + 1.127X**

Illustration 02:

After investigation it has been found the demand for automobiles in a city depends mainly, if not entirely, upon the number of families residing in that city. Below are the given figures for the sales of automobiles in the five cities for the year 2019 and the number of families residing in those cities.

City	No. of Families (in lakhs): X	Sale of automobiles (in '000): Y
Belagavi	70	25.2
Bangalore	75	28.6
Hubli	80	30.2
Kalaburagi	60	22.3
Mangalore	90	35.4

Fit a linear regression equation of Y on X by the least square method and estimate the sales for the year 2020 for the city Belagavi which is estimated to have 100 lakh families assuming that the same relationship holds true.

Solution:

Calculation of Regression Equation

City	X	Y	X^2	XY
Belagavi	70	25.2	4900	1764
Bangalore	75	28.6	5625	2145
Hubli	80	30.2	6400	2416
Kalaburagi	60	22.3	3600	1338
Mangalore	90	35.4	8100	3186
	375	**141.7**	**28,625**	**10,849**
	$\sum$**X**	$\sum$**Y**	$\sum X^2$	$\sum$**XY**

Regression equation of Y on X: **Y = a + bX**

The two normal equations are:

$$\sum Y = Na + b\sum X$$

$$\sum XY = a\sum X + b\sum X^2$$

Substituting the values in above normal equations, we get

141.7 = 5a + 375b .... (i)

10849= 375a + 28625b .... (ii)

Let us solve these equations (i) and (ii) by simultaneous equation method Multiply equation (i) by 75 we get 10627.5 = 375a + 28125b

Now rewriting these equations:

10627.5	=	375a	+	28125b
10849	=	375a	+	28625b
(-)		(-)		(-)
-221.5	=			-500b

Therefore now we have -221.5 = -500b, this can rewritten as 500b = 221.5

Now, $b = \frac{221.5}{500} = 0.443$

Substituting the value of b in equation (i), we get

141.7 = 5a + (375 * 0.443)

141.7 = 5a + 166.125

5a = 141.7 - 166.125

5a = -24.425

a = -24.425/5

a = -4.885

Thus we got the values of a = -4.885 and b = 0.443 Hence, the required regression equation of Y on X:

Y = a + bX => **Y = -4.885 + 0.443X**

Estimated sales of automobiles (Y) in city Belagavi for the year 2020, where number of families (X) are 100(in lakhs):

Y = -4.885 + 0.443X

Y = -4.885 + (0.443 * 100)

Y = -4.885 + 44.3

Y = 39.415 (‘000)

Means sales of automobiles would be 39,415 when number of families are 100,00,000

Illustration 03:

From the following data obtain the two regression lines:

Capital Employed (Rs. in lakh):	7	8	5	9	12	9	10	15
Sales Volume (Rs. in lakh):	4	5	2	6	9	5	7	12

Solution:

Calculation of Regression Equation

X	**Y**	X^2	Y^2	**XY**
7	4	49	16	28
8	5	64	25	40
5	2	25	4	10
9	6	81	36	54
12	9	144	81	108
9	5	81	25	45
10	7	100	49	70
15	12	225	144	180
75	**50**	**769**	**380**	**535**
$\sum X$	$\sum Y$	$\sum X^2$	$\sum Y^2$	$\sum XY$

Regression line/equation of X on Y: $(X - \dot{X}) = b_{xy}(Y - \dot{Y})$

$\dot{X} = \frac{\sum X}{n} = \frac{75}{8} = 9.375$

$\dot{Y} = \frac{\sum Y}{n} = \frac{50}{8} = 6.25$

Regression coefficient of X on Y:

$$b_{xy} = \frac{n\sum XY - \sum X \sum Y}{n\sum Y^2 - (\sum Y)^2}$$

$$b_{xy} = \frac{(8*535) - (75*50)}{(8*380) - (50)^2}$$

$$= \frac{4280 - 3750}{3040 - 2500}$$

$$= \frac{530}{540} = 0.9815$$

$(X - \dot{X}) = b_{xy}(Y - \dot{Y})$

⇨ $X - 9.375 = 0.9815(Y - 6.25)$

⇨ $X - 9.375 = 0.9815Y - 6.1344$

⇨ $X = 9.375 - 6.1344 + 0.9815Y$

⇨ $\mathbf{X = 3.2406 + 0.9815Y}$

Regression line/equation of Y on X: $(Y - \dot{Y}) = b_{yx}(X - \dot{X})$

$\dot{X} = \frac{\sum X}{n} = \frac{75}{8} = 9.375$

$\dot{Y} = \frac{\sum Y}{n} = \frac{50}{8} = 6.25$

Regression coefficient of Y on X:

$$b_{yx} = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$b_{xy} = \frac{(8*535) - (75*50)}{(8*769) - (75)^2}$$

$$= \frac{4280 - 3750}{6152 - 5625}$$

$$= \frac{530}{527} = 1.0057$$

$(Y - \dot{Y}) = b_{yx}(X - \dot{X})$

⇨ $Y - 6.25 = 1.0057(X - 9.375)$

⇨ $Y - 6.25 = 1.0057X - 9.4284)$

⇨ $Y = 6.25 - 9.4284 + 1.0057X$

⇨ $\mathbf{Y = -3.1784 + 1.0057X}$

Illustration 04:

From the following information find regression equations and estimate the production when the capacity utilisation is 70%.

	Average (Mean)	Standard Deviation
Production (in lakh units)	42	12.5
Capacity Utilisation (%)	88	8.5
Correlation Coefficient (*r*)	0.72	

Solution:

Let production be variable X and capacity utilisation be variable Y. Regression equation of production based on based on capacity utilisation shall be given by X on Y and regression equation of capacity utilisation of production shall be given by Y on X, which can be computed as given below:

Given Information: $\dot{X} = 42$ $\dot{Y} = 88$ $\sigma_X = 12.5$ $\sigma_y = 8.5$ $r = 0.72$

Regression coefficient of X on Y:

$$b_{xy} = r\frac{\sigma x}{\sigma y} = 0.72 * \frac{12.5}{8.5} = 1.0588$$

Regression coefficient of Y on X:

$$b_{yx} = r\frac{\sigma y}{\sigma x} = 0.72 * \frac{8.5}{12.5} = 0.4896$$

Regression Equation of X on Y:

$(X - \dot{X}) = b_{xy}\,(Y - \dot{Y})$

⇨ X – 42 = 1.0588 (Y – 88)

⇨ X = 42 – 93.1744 + 1.0588Y

⇨ **X = -51.1744 + 1.0588Y**

Regression Equation of Y on X:

$(Y - \dot{Y}) = b_{yx}\,(X - \dot{X})$

⇨ Y – 88 = 0.4896 (X – 42)

⇨ Y = 88 – 20.5632 + 0.4896X

⇨ **Y = 67.4368 + 0.4896X**

Estimation of the production when the capacity utilization is 70% is regression equation X on Y, where Y = 70

Regression Equation of X on Y: $(X - \dot{X}) = b_{xy} (Y - \dot{Y})$

$X = -51.1744 + 1.0588Y$

$= -51.1744 + (1.0588 * 70)$

$= -51.1744 + 74.116$

= **22.9416**

Therefore, the estimated production would be **22,94,160** units when there is a capacity utilisation of 70%.

Illustration 05:

The following data gives the age and blood pressure (BP) of 10 sports persons.

Name	:	A	B	C	D	E	F	G	H	I	J
Age (X)	:	42	36	55	58	35	65	60	50	48	51
BP (Y)	:	98	93	110	85	105	108	82	102	118	99

i. Find regression equation of Y on X and X on Y (Use the method of deviation from arithmetic mean)
ii. Find the correlation coefficient (r) using the regression coefficients.
iii. Estimate the blood pressure of a sports person whose age is 45.

Solution:

Calculation of Regression Equation

Name	Age (X)	BP (Y)	$x=X-\dot{X}$ $x=X-50$	$y=Y-\dot{Y}$ $y=Y-100$	x^2	y^2	xy
A	42	98	-8	-2	64	4	16
B	36	93	-14	-7	196	49	98
C	55	110	5	10	25	100	50
D	58	85	8	-15	64	225	-120
E	35	105	-15	5	225	25	-75
F	65	108	15	8	225	64	120
G	60	82	10	-18	100	324	-180
H	50	102	0	2	0	4	0
I	48	118	-2	18	4	324	-36
J	51	99	1	-1	1	1	-1

	500	**1,000**	**0**	**0**	**904**	**1,120**	**-128**
	$\sum \mathbf{X}$	$\sum \mathbf{Y}$	$\sum \mathbf{x}$	$\sum \mathbf{y}$	$\sum \mathbf{x}^2$	$\sum \mathbf{y}^2$	$\sum \mathbf{xy}$

$\dot{X} = \frac{\sum X}{n} = \frac{500}{10} = 50$ $\quad$ $\dot{Y} = \frac{\sum Y}{n} = \frac{1000}{10} = 100$

$\dot{X} = \frac{\sum X}{n} = \frac{500}{10} = 50$ $\quad$ $\dot{Y} = \frac{\sum Y}{n} = \frac{1000}{10} = 100$

Regression coefficients can be computed using the following formula:

$b_{xy} = \frac{\sum xy}{\sum y^2}$ $\quad$ $b_{yx} = \frac{\sum xy}{\sum x^2}$ $\quad$ where $x = X - \dot{X}$ and $y = Y - \dot{Y}$

Regression equation of X on Y:

$(X - \dot{X}) = b_{xy} (Y - \dot{Y})$

⇨ $X - 50 = -0.1143 (Y - 100)$

⇨ $X - 50 = -0.1143Y + 11.43$

⇨ $X = 50 + 11.43 - 0.1143Y$

⇨ $\mathbf{X = 61.43 - 0.1143Y}$

Regression equation of Y on X:

$(Y - \dot{Y}) = b_{yx} (X - \dot{X})$

⇨ $Y - 100 = -0.1416 (X - 50)$

⇨ $Y - 100 = -0.1416X + 7.08$

⇨ $Y = 100 + 7.08 - 0.1416X$

⇨ $\mathbf{Y = 107.08 - 0.1416X}$

Computation of coefficient of correlation using regression coefficient:

$r = \sqrt{bxy * byx} = -\sqrt{0.1143 * 0.1416} = -\sqrt{0.01618488} = -0.1272$

Therefore, we have low degree of negative correlation between age and blood pressure of sports person.

Estimation of the blood pressure (Y) of a sports person whose age is X=45 can be calculated using regression equation Y on X:

Regression equation of Y on X:

$(Y - \dot{Y}) = b_{yx} (X - \dot{X})$

⇨ $Y = 107.08 - 0.1416X = 107.08 - (0.1416 * 45) = 107.08 - 6.372 =$ **<u>100.708</u>**

It means estimated blood pressure of a sports person is 101 (rounded off) whose age is 45.

Illustration 06:

There are two series of index numbers, ***P*** for price index and ***S*** for stock of commodity. The mean and standard deviation of *P* are 100 and 8 and *S* are 103 and 4 respectively. The correlation coefficient between the two series is 0.4. With these data, work out a linear equation to read off values of *P* for various values of *S*. Can the same equation be used to read off values of *S* for various values of *P*?

Solution:

Let us assume that *P*=Price Index be variable X an *S*=Stock of Commodity be variable Y. Linear equation to read off values of P for various values of S would be regression equation of X on Y. Regression coefficient is to be computed using mean and standard deviation.

From the problem we can list out the given information:

$\dot{X} = 100$ $\dot{Y} = 103$ $\sigma_X = 8$ $\sigma_y = 4$ $r = 0.4$

Regression equation of X on Y:

$$(X - \dot{X}) = b_{xy}\,(Y - \dot{Y})$$

$$\Rightarrow (X - \dot{X}) = r\frac{\sigma x}{\sigma y}(Y - \dot{Y})$$

$$\Rightarrow (X - 100) = \left(0.4 * \frac{8}{4}\right)(Y - 103)$$

$$\Rightarrow (X - 100) = 0.8\,(Y - 103)$$

$$\Rightarrow (X - 100) = 0.8Y - 82.4$$

$$\Rightarrow X = 100 - 82.4 + 0.8Y$$

$$\Rightarrow \mathbf{X = 17.6 + 0.8Y}$$

Linear equation to read off values of *P* for various values of *S* is X = 17.6 + 0.8Y

To read off values of S for various values of P we need regression equation of Y on X and therefore above linear equation cannot be used. Hence, the following regression equation of Y on X be computed:

$$(Y - \dot{Y}) = b_{yx}\,(X - \dot{X})$$

$$\Rightarrow (Y - \dot{Y}) = r\frac{\sigma y}{\sigma x}(X - \dot{X})$$

$$\Rightarrow (Y - 103) = 0.4 * \frac{4}{8}(X - 100)$$

$$\Rightarrow (Y - 103) = 0.2\,(X - 100)$$

$$\Rightarrow Y - 103 = 0.2X - 20$$

$$\Rightarrow Y = 103 - 20 + 0.2X$$

$$\Rightarrow \mathbf{Y = 83 + 0.2X}$$

Hence, the linear equation to read off values of S for various values of *P* is Y = 83 + 0.2X

Review of Correlation and Regression Analysis:

In correlation analysis, when we are keen to know whether two variables under study are associated or correlated and if correlated what is the strength of correlation. The best measure of correlation is proved by Karl Pearson’s Coefficient of Correlation. However, one severe limitation of this method is that it is applicable only in case of a linear relationship between two variables. If two variables say X and Y are independent or not correlated then the result of correlation coefficient is zero.

Correlation coefficient measuring a linear relationship between the two variables indicates the amount of variation one variable accounted for by the other variable. A better measure for this purpose is provided by the square of the correlation coefficient, known as “coefficient of determination”. This can be interpreted as the ratio between the explained variance to total variance:

$$r^2 = \frac{\text{Explained variance}}{\text{Total Variance}}$$

Similarly, Coefficient of non-determination = $(1 - r^2)$.

Regression analysis is concerned with establishing a functional relationship between two variables and using this relationship for making future projection. This can be applied, unlike correlation for any type of relationship linear as well as curvilinear. The two lines of regression coincide i.e. become identical when r= -1 or +1 in other words, there is a perfect negative or positive correlation between the two variables under discussion if r = 0, then regression lines are perpendicular to each other.

Unit – V – Index Numbers

As index number is a specialized average designed to measure the change in a group ofrelated variable over a period of time. It was first constructed in the year

Concept

In its simplest form on Index number is a Ratio of two numbers expressed as percent.

Definition

Index number devices for measuring difference in the magnitude of a group of related variables - Croxton and Cowden

Characteristics of Index number

- They are specialized average
- They measure the net change in a group of related variables
- They measure the effect of changes over a period of time
- They help comparison of groups of variables directly

Uses of Index Number

- Index number is most widely used statistical devises
- Index numbers are used to measure the relative changes
- They are widely used in the evaluation of business and economic conditions
- It is useful for better comparison
- It is a good guide for the progress of every country
- It is useful for better comparison
- It is useful to know trends and techniques
- For forecasting future activities

Types of Index Numbers

- Price Index
- Quantity Index
- Value Index

Methods of Index Number

- Unweighted Index Number
- Simple Aggregative method
- Simple average of price Relative method

Weighted Index Number

$\sum p1q0$

1. Laapeyre's index number $P01 = \text{---------------}100$

$\sum P0q0$

2. PAASCHEY'S Method

$\sum p1q1$

. $P01 = \text{---------------------}100$

$\sum P0q0$

3. Bowley and Dorfish method

$\sum p1q0 \sum p1q0$

$P01 = \text{------------} + \text{.....................}$

$\sum P0q0 \sum P0q0$

4. Fisher's Ideal method or Fishers Price Index Number

$P01 = \sqrt{L \times P}$

Consumer price Index Number (or) Cost of living index

Consumer price Index is designed to measure the change in the cost of living of workers because of change in the retail price. A change in the price level affects the cost of living of the people. People con some different types of commodities. So there is need to construct consumer's price index. Consumer price index can be used in different places for many purposes.

Uses of cost of living index

- It is useful in fixing the wages
- It is useful to know the purchasing power of money
- By using the cost of living index the Government determines the price and other variables
- It is useful the analysis of price situations

Limitations of Index numbers

- If the chosen base year is not a normal one, the purpose is list
- Every index number has its own purpose. No index number can serve all purpose
- These are only appropriate indications of the relative level.

Solved Problems

Construct an index number for 2014 taking2013 as base from the following data

Commodity	price in 2013(Rs)	Price in 2014 (Rs)
A	50	60
B	40	80
C	70	110
D	90	70
E	50	40

Solution

Commodity	price in 2013(Rs)	Price in 2014 (Rs)
A	50	60
B	40	80
C	70	110
D	90	70
E	50	40
	$\sum p_0$ = 300	**$\sum p_1$ = 360**

$\sum$p1360

Price Index P01 = -------------- X 100 = -------------- X 100 = 120

$\sum$p0 300

This means that we compared to 2013, in 2014 there is a net increase in the prices of commodities to the extent of 20 %

Compute Fisher's Ideal Index from the following data

Item	2000	2003

	Price	Quantity	Price	Quantity
A	6	50	10	56
B	2	200	2	120
C	4	60	6	60
D	10	30	12	24
E	8	40	12	36

Solution

COMMODITY	P0	Q0	P1	Q1	p1q0	P0q0	p1q1	P0q1
A	6	50	10	56	500	300	560	336
B	2	100	2	120	200	200	240	240
C	4	60	6	60	360	240	360	240
D	10	30	12	24	360	300	288	240
E	8	40	12	36	480	320	432	288
					∑ p1q0=1900	∑ P0q0 =1360	∑ p1q1 =1880	∑ P0q1 =1344

P01 = √ L X P

∑p1q0 ∑p1q0

P01 = √............x ....

∑P0q0 ∑P0q0

= √ 1,3971 X 1.3988 X 100

= 1.39796 X 100 = 139.80

Analysis of time Series

An arrangement of statistical data in accordance with time of occurrence or inchronological order is called a time series.

Definition

A time series is a set of observation arranged in chronological order. MorrisHamberg requirement of a time seriesData must be available for a long period of time. Data must consist of a homogeneous set of values belonging to different time periods. The time gap between the variables or composite of variables must be as For Passible equal.

Uses of Time series

i. It helps in understanding the past behaviors and in establishing the future behavior
ii. It helps in planning and forecasting the future operation
iii. It facilitates comparison between data of one period with those of another period
iv. It helps in evaluating current accomplishment
v. It is useful in forecasting the trade cycles

Time Series Models

Mathematical Models and Multiplicative method

In classical analysis, it is assumed that some types of relationship exist among the four components of time series

1. Additive Model

According to this model, the time series is expressed as

Y = T + S + C + I

Y = the value of original time series

T = Time Value

S = Seasonal variation

C = Cyclical Variation Irregular fluctuation

Multiplicative Model

According this model, the time series is expressed as

Y = Y X S X C X I

Time series Analysis

- Time series analysis is the analysis of identifying different components such as trend, seasonal, cyclical and irregular in a given time series data.
- Components of time series
- Time series data contain variations of the following types
  1. Secular Trend
  2. Seasonal Variation
  3. Cyclical Variations
  4. Irregular variation

1. Secular Trend

A secular trend or long-term trend refers to the movements of the series reflecting continuous growth or decline over a long period of time. There are many types of trend. Some trends rise upward and some fall downward

2. Seasonal Variation

Is that periodic investment in business activities within the year recurring periodically year after year**?**

Generally, seasonal variation appear at weekly, monthly or quarterly intervals

3. Cyclical Variation

Up and down movements are different from seasonal fluctuations, in that they extend over longer period of time – usually two or more years. Business time series is influenced by the wave-like changes of prosperity and depression.

Causes

Changes of property and depressionUses

I) Useful to study the character of business fluctuation
II) Useful to take timely decision in maintaining the business during different stages
III) Helps in facing recession and utilizing the booms

Un-Secular Variation

Irregular variation refers to such variation in business activity which do not repeat in a definite pattern. They are also called 'erratic 'accidental or random variations which are generally non-recurring and unpredictable

Causes

War, food, revolution, strike, lockouts and the like

Measurement of Secular of secular trend

a) Free hand Graphic Method
b) In this method we must plot the original data on the graph. Draw a smooth curve carefully which will show the direction of the trend. The time is taken on the horizontal axis I(X) and the value of the variable on the vertical axis (Y)

Merits

a. It is the simplest and easiest method
b. It can be applied to all types of trends
c. It is useful to understand the character of time series

Demerits

a. It is subject to personal bias
b. Its results depend upon the judgments of the person who draw the time
c. It does not help to measures trend

Semi – Average Method

In this method the original data are divided into two equal parts and average are calculated for both the parts. These averages are called semi average. Trend line is drawn withthe help of the semi averages

Merits

I) It is simple and easier to understand
II) Everyone will get the same trend like
III) We can predict the future values based on the intermediate values

Demerits

I) It is affected by the limitations of arithmetic mean
II) It is not enough for forecasting the future trend

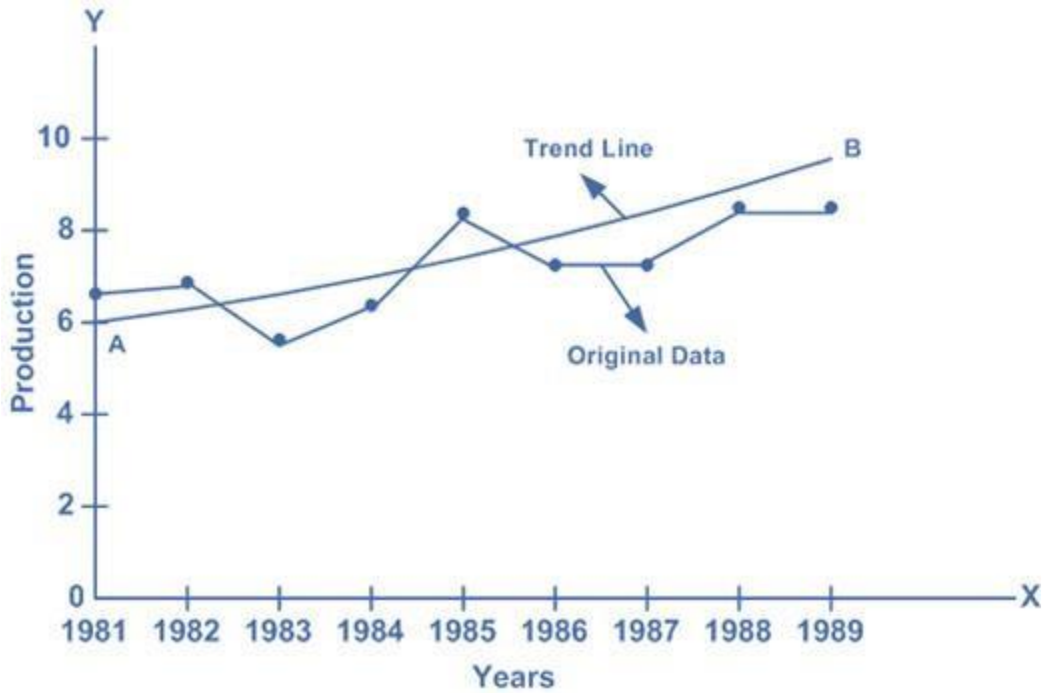
Moving average method

In this method, the average value of a number of years or months or weeks is taken into account and placed it at the Centre of the time span and it is is the normal or trend value for the middle period.

Solved Problems

Free-Hand Trend line1985

Year	1981	1982	1983	1984	1985	1986	1987	1988	1989
Production tons	0	2	40	6	8	10	6	8	8



Moving Average

1.Find the 3 yearly moving average from the following time series data

Year	1998	1999	2000	2001	2002	2003	2004	2005
Sales(In tones)	30.1	45.4	39.3	41.4	42.2	46.4	46.6	49.2

Solution

Year	**Sales (in tons)**	**3yearly Moving Total**	**3 Yearly moving value**
1998	30.1	-------	-------
1999	45.4	114.8	38.27
2000	39.3	126.1	42.03
2001	41.4	122.9	40.97
2002	42.2	130.0	43.33
2003	46.4	135.2	45.07
2004	46.6	142.2	47.40
2005	49.2	-------	--------

1.Calculate the 5 yearly moving average from the following data

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Number of students	705	685	703	687	705	689	715	685	725	730

Solution

Year	No of students	5 yearly moving total	Moving value
1998	**705**	**--**	**--**
1999	**685**	**--**	**--**
2000	**703**	**3485**	**697.0**
2001	**687**	**3469**	**693.8**
2002	**705**	**3499**	**699.8**
2003	**689**	**3481**	**696.2**
2004	**715**	**3519**	**703.8**
2005	**685**	**3544**	**708.8**
2006	**725**	**---**	**--**
2007	**730**	**--**	**--**

2. Calculate the Four – Yearly moving average for the following data

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Production (in ‘000Tons)	464	515	518	467	502	540	557	570	586	612

Solution

Year	**Production in ‘000 Tons**	**4yearly Moving**	**Combined Total**	**Moving Average**
1998	464			
1999	515			
		1964		
2000	518		3966	495.75
		2002		
2001	467		4029	503.63

		2027		
2002	502		4093	511.63
		2066		
2003	540		4236	529.50
		2170		
2004	557		4424	553.00
		2254		
2005	570		4580	572.50
		2326		
2006	586			
2007	612			

1.Compute the trend from the following by the method of least square method

Year	2000	2001	2002	2003	2004
Production in Lakhs	830	920	710	900	1690

Solution

Computation of trend Values

Year	**Production in Lakhs (Y)**	**Deviation from 2002(X)**	**XY**	**X^2**
2000	830	-2	-1600	4
2001	920	-1	-920	1
2002	710	0	0	0
2003	900	1	900	1
2004	1690	2	3380	4
	$\sum y = 5050$	**$\sum x = 0$**	**$\sum xy$=1700**	**$\sum x^2$ =10**

Since $\sum x = 0$

$\sum y$

a = ---- = 5050 /

5 = 1010N

b = $\sum xy$

$\sum x^2$

$Y = a + bx$

$= = 1010 + 170x$

$Y_c = 1010 +$

$170\ x$

When x= -2

$Y_{2000} = 1010 + 170(-2) = 1010 - 340 = 670$

When X = -1

$Y_{2001} = 1010 + 170(-1) = 1010 - 170 = 840$

When X = 0

$Y_{2002} = 1010 + 170(0) = 1010 - 0 = 1010$

When X =1

$Y_{2003} = 1010 + 170(-1) = 1010 + 170 = 1180$

When X=2

$Y_{2003} = 1010 + 170(2) = 1010 + 340 = 1350$