PERIYAR INSTITUTE OF DISTANCE EDUCATION

(PRIDE)

PERIYAR UNIVERSITY

SALEM - 636 011.

BACHELOR OF BUSINESS ADMINISTRATION (B.B.A.)

FIRST YEAR

ALLIED - I : BUSINESS MATHEMATICS AND STATISTICS

Prepared by :

Dr. V. NARANEETHA KUMAR
Asst. Professor in Management Studies
Adhiyamaan College of Engineering
Hosur.

BACHELOR OF BUSINESS ADMINISTRATION

FIRST YEAR

ALLIED – I : BUSINESS MATHEMATICS AND STATISTICS

UNIT I	SEQUENCE AND SERIES, MATRICES
UNIT II	MATHEMATICS OF FINANCE
UNIT III	DIAGRAMMATIC PRESENTATION OF DATA (ONE DIMENSION) AND TWO DIMENSION
UNIT IV	MEASURES OF VARIATION
UNIT V	INDEX NUMBERS AND TIME SERIES

INTRODUCTION

Dear Students,

This block consists of five units dealing with the concept of Business Mathematics and Statistics. The first unit deals with the business application of series and sequences and also the matrices construction for the practical situation of the industry.

The second unit describes the mathematical application in finance and calculating methods. The third unit highlights the diagrammatic presentation of data and types of diagram.

The fourth unit is explained the measures of variation. The fifth unit is deals with various types of index number.

PRIDE would be happy if you could make use of this learning material to enrich your knowledge and skills to serve the society.

We assure you of our best service in guiding you throughout the course.

ALLIED – I

BUSINESS MATHEMATICS AND STATISTICS

Unit 1 : Series and Matrices

1.1 Introduction
1.2 Sequence
1.3 Series
1.4 Arithmetic Progression
1.5 Geometric Progression
1.6 Formula for nth term
1.7 Harmonic progressions
1.8 Matrices
1.8.1 Introduction
1.8.2 Definitions
1.8.3 Matrices operations
1.8.4 Determinant of a matrix
1.8.5 Singular matrix
1.8.6 Ad joint of a square matrix
1.8.7 Inverse of the matrix
Summary
Exercise

Unit 2: Mathematics of Finance

2.1 Introduction
2.1.1 Simple and compound interest:
2.1.2 Present Value:
2.1.3 Annuities:
2.1.4 Present value of an annuity
2.2 Sinking Fund Factor
2.3 Discounting
2.3.1 Discount rate
2.3.2 Discount factor
Summary
Exercise

UNIT - I
SEQUENCE AND SERIES, MATRICES

1.1 INTRODUCTION

This chapter will help the readers to understand about the sequence and series. All the definitions are given detaily with examples; also worked examples are given with detailed steps. Formulae also separately given.

Objective

- To learn about business application of series and sequences
- Also the matrices construction for the practical situation of the industry.

1.2 Sequence

An ordered set of numbers is called a sequence

Example

(i) 1, 2, 3, 4…..

(ii) $1^2, 2^2, 3^2, ...$

1.3 Series

A series is the sum of the terms of a sequence. Let we see the following series.

1.4 Arithmetic series or progression

A series, in which the difference of any term and its preceeding term is constant, is called an arithmetic series or progression.

Example

(i) 1, 4, 7, 10………

(ii) 6, 1, - 4,-9,-14…..

In general

a, a+d, a+2d……

Formula for finding nth term

$T_n = a + (n-1) d$

Example 1

Find the 40^{th} term of an AP, whose 9^{th} term is 465 and 20^{th} term is 388

Solution

The formula for finding n^{th} term is $T_n = a + (n-1) d$

The 9^{th} term is 465

$T_9 = a + (9-1) d = 465$

That is

a + 8d = 465 --------------(I)

The 20th term is =

That is ----------(II)

By solving (I) & (II)

$$a+8d=465$$
$$a+19d=388$$
$$-\,-\,-\,-\,-\,-\,-$$
$$-11d=77$$
$$d=-7$$

From (I)

a-56 = 465

a = 521

There fore let we apply "a" and 'b' in the following equation

$$T_{40}=a+39d=521+39(-7)$$
$$T_{40}=248$$

(i.e) the 40th term is = 248

Example 2

The 7th term of an AP is 39 and 17th term is 69. find the series

Solution:

$$T_7=a+6d=39$$
$$a+6d=39\ldots\ldots\ldots\ldots(1)$$
$$T_{17}=a+16d=69$$
$$a+6d=69\ldots\ldots\ldots\ldots(2)$$

Where equation (1) and (2) is the required equation

Example 3:

The rate of monthly salary of a person increases annually in AP. It is known that he was drawing Rs 200 a month. During the 11th year of service and Rs 380 during 29th year . ind his starting salary and the rate of annual increments.

Solution

Let a = starting salary

D = annual increment for each year

a+10d = 200………….(1)

a+28d = 380…..……..(2)

18d = 180

$\therefore$ d = $\frac{180}{18} = 10$

There fore from equation (1)

a = 100

$\therefore$ Starting salary = Rs. 200

Annual increment = Rs. 10

1.5 Geometric Progression

A series in which the ratio of any term to its preceding term is constant, is called a geometric progression

Example

(i) 2, 4, 8,16….

(ii) a,ar,ar^2

1.6 Formula for n^{th} term

$T_n = ar^{n-1}$

Example

Find the three numbers whose sum is 21 and whose product is 216

Solution

Let the number be $\frac{a}{r}, a, ar$

$\therefore$ The sum of three number is

$\frac{a}{r} + a + ar = 21$ ….. (1)

Product of three numbers

$(\frac{a}{r})(a)(ar)$ 216……….. (2)

$a^3 = 216$

a = 6

From (1) $\Rightarrow$

$$\frac{6}{r} + 6 + 6r = 21$$

$$6 + 6r + 6r^2 = 21r$$

$$6r^2 - 15r + 6 = 0$$

If we factorize

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

hen a = 6 and r = 2

The numbers are $\frac{6}{2}, 6, 6(2)$

= 3,6,12

When a = 6 and r = $\frac{1}{2}$

The numbers are

Formula for the sum to n terms

$$s_n = \frac{a(r^n - 1)}{r - 1}$$

Example

Find the sum to n terms of the series

3, 2, 4/3, 8/9…………

Solution

From the above series it is noted

A=3 and r=2/3

$$s_n = \frac{a(r^n - 1)}{r - 1}$$

$$= \frac{3(\left(\frac{2}{3}\right)^n - 1)}{\frac{2}{3} - 1}$$

$$= \frac{9\left(\left(\frac{2}{3}\right)^n - 1\right)}{2-3}$$

$$= s_n = 9[1 - \left(\frac{2}{3}\right)^n]$$

1.7 Harmonic progressions

The harmonic progression is reciprocal of arithmetic progression.

1.8 MATRICES

1.8.1 Introduction

This chapter gives clear idea about the matrices. And all the models of the matrices with definitions along with example. Matrices are extremely useful for writing fast 3D programs. Also they can be used to efficiently keep track of transformations.

The objective of this is:

To explain the usefulness of matrix operation for solving system of equation and also apply matrix operation for solving business problems.

1.8.2 Definitions

Definition of Matrix

A matrix is a rectangular array of row and column.

Example.

Let A denote the matrix

$$A= \begin{bmatrix} 1 & 5 & 6 \\ 3 & 3 & 0 \\ 1 & 0 & 8 \end{bmatrix}$$

This matrix A has three rows and three columns. We say it is a 3x 3 matrix

Definition of Square Matrix

In a given matrix if number of row and columns are same, then we say it's square matrix.

Example

Here number of row is 3 and number of column also 3

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Diagonal Matrix

A diagonal matrix is a square matrix with all the elements are zero except the diagonal elements

Example

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Row Matrix

A matrix with one row is called a row matrix

Example

[7 0 0]

Column Matrix

A matrix with one column is called a column matrix

Example

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Unit or identity Matrix (I)

An identity matrix I is a diagonal matrix with all diagonal element are = 1.

Example

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Scalar Matrix S

A scalar matrix S is a diagonal matrix with all diagonal elementsAre same

Example

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Symmetric Matrix

A square matrix is called symmetric if it is equal to the Opposite of its transpose. Then ai,j '= aji' for all i and j.

Example

$$\begin{bmatrix} 1 & 5 & 7 \\ 5 & 4 & 8 \\ 7 & 8 & 6 \end{bmatrix}$$

A skew-symmetric Matrix

A square matrix is called skew-symmetric if it is equal to the opposite of its transpose. Then ai,j '=-ai,j' for all i and j.

Example

$$\begin{bmatrix} 1 & 5 & -7 \\ -5 & 4 & -8 \\ 7 & 8 & 6 \end{bmatrix}$$

1.8.3 Matrices operations

Addition and Subtraction

Matrices can be added and subtracted only if they are of the same order

Example

Let A = $\begin{bmatrix} 1 & 5 & 6 \\ 3 & 3 & 0 \\ 1 & 0 & 8 \end{bmatrix}$ and B = $\begin{bmatrix} 5 & 6 & 6 \\ 2 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

Then

$$A + B = \begin{bmatrix} 1+5 & 5+6 & 6+6 \\ 3+2 & 3+2 & 0+0 \\ 1+0 & 0+0 & 8+4 \end{bmatrix}$$

Therefore

$$A + B = \begin{bmatrix} 6 & 11 & 12 \\ 5 & 5 & 0 \\ 1 & 0 & 12 \end{bmatrix}$$

In the same way

$$A - B = \begin{bmatrix} 1-5 & 5-6 & 6-6 \\ 3-2 & 3-2 & 0-0 \\ 1-0 & 0-0 & 8-4 \end{bmatrix}$$

There fore

$$A - B = \begin{bmatrix} -4 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 4 \end{bmatrix}$$

Multiplication of the matrix

Two matrix A and B can be multiplied if the number of column in A is equal to the number of rows in B.it can be denoted as AB.

Example

Let $A = \begin{bmatrix} 1 2 3 \\ 3 2 1 \end{bmatrix}$ and $B = \begin{bmatrix} 4 5 \\ 2 3 \\ 1 2 \end{bmatrix}$

$$AB = \begin{bmatrix} 1\times 4+2\times 2+3\times 1 & 1\times 5+2\times 3+3\times 2 \\ 3\times 4+2\times 2+1\times 1 & 3\times 5+2\times 3+1\times 2 \end{bmatrix}$$

$$AB = \begin{bmatrix} 11 & 17 \\ 17 & 13 \end{bmatrix}$$

Example

There are two families A and B. There are two men, three women and one child in family A and one man , one women and two children in family B. The recommended daily allowance for calories is men 2400; women 1900; children1800 and for men proteins 55gms, women 45gms and children 33gms. Represent the above information by matrices using matrix multiplication; calculate the total requirements of calories and proteins for each of the two families

Solution:

Let A = $\begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ and B = $\begin{bmatrix} 2400 & 55 \\ 1900 & 45 \\ 1800 & 33 \end{bmatrix}$

AB = $\begin{bmatrix} 12300 & 278 \\ 7900 & 166 \end{bmatrix}$

The Transpose of a Matrix

For any given matrix A, the matrix whose rows are columns of A and whose columns are rows of A is called the transpose of A and is denoted A^T or A'

Example

If A = $\begin{bmatrix} 4 & 2 & 6 \\ 7 & 3 & 5 \end{bmatrix}$ then A^T = $\begin{bmatrix} 4 & 7 \\ 2 & 3 \\ 6 & 6 \end{bmatrix}$

1.8.4 Determinant of a matrix

Consider the following 2 X 2 matrix

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

There fore the determinant of the matrix A is

$a_{11}\ a_{22} - a_{21}\ a_{12}$

For 3 X 3 matrix

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

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

Example

Find the determinant of the matrix

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

There fore

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

= 2(4-0)-3(16-6)+1(0-3)

= 8-30-3 = -25

1.8.5 Singular matrix

A square matrix 'A' is said to be singular .if its determinant is zero, otherwise it is non-singular

Example

Show that the following is singular matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 4 \\ 2 & 4 & 6 \end{bmatrix}$$

Here $|A|$ =0, therefore it is singular matrix

1.8.6 Adjoint of a square matrix

The transpose of the co-factor matrix is called ad joint matrix

Example

Find adjoint of A

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

Solution

First let we find co factor for all the elements

Co-factor of 1 = $\begin{vmatrix} 2 & -3 \\ -1 & 3 \end{vmatrix}$ = 3

Co-factor of 1 = - $\begin{vmatrix} 1 & -3 \\ 2 & 3 \end{vmatrix}$ = -9

Co-factor of 1= $\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix}$ = -5

Co-factor of 1= - $\begin{vmatrix} 1 & 1 \\ -1 & 3 \end{vmatrix}$ = -4

Co-factor of 2 = $\begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix}$ = 1

Co-factor of -3 = - $\begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix}$ = 3

Co-factor of 2 = $\begin{vmatrix} 1 & 1 \\ 2 & -3 \end{vmatrix}$ = -5

Co-factor of -1 = - $\begin{vmatrix} 1 & 1 \\ 1 & -3 \end{vmatrix}$ = 4

Co-factor of 3 = $\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}$ = 1

Therefore

The co-factor matrix is

Co- factor matrix of A = $\begin{bmatrix} 3 & -9 & -5 \\ -4 & 1 & 3 \\ -5 & 4 & 1 \end{bmatrix}$

Adj A = (co-factor matrix of A)^T

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

1.8.7 Inverse of the matrix

The inverse of the matrix of A^{-1} = $\dfrac{adjA}{|A|}$

Example

Find the inverse of A = $\begin{bmatrix} 2 & 3 & 4 \\ 3 & 2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$

Solution:

We know that $A^{-1} = \dfrac{adjA}{|A|}$

Therefore the co-factor matrix of A = $\begin{bmatrix} -5 & 7 & 1 \\ 10 & -8 & 1 \\ -5 & 10 & -5 \end{bmatrix}$ Also $|A| = 15$

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

Summary

This unit has provided a over all idea about Series and sequence, also about the matrices and calculating methods. Particularly, this unit focused on:

- The calculation of n^{th} term calculation of all AP, GP and HP
- Various types of Matrices along with example
- Inverse matrix calculations

Exercise

1) Write the formula for finding nth term of the AP series
2) What is the relation between AP and HP
3) If A = $\begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$ and B = $\begin{bmatrix} 5 & 6 \\ -2 & 5 \end{bmatrix}$

   Then verify that (AB)' = B' A'
4) What are the properties of transpose matrix?
5) Write the way of finding inverse matrix?
6) Express the matrix A as sum of symmetric and a skew Matrix where

   A = $\begin{bmatrix} 5 & -1 \\ 3 & 4 \end{bmatrix}$
7) Find the sum to n terms of the series 6, 4, 4/3, 8/9…………

NOTES

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

UNIT - II
MATHEMATICS OF FINANCE

2.1 Introduction:

The notation that money has time value is one of the basic concepts of finance. Most businessmen will agree that a rupee in hand today is more valuable to them than a rupee to be received a year from now, because as soon as the funds are received they can be put to work in profitable opportunities. If funds are received late, we have to forgo interest that could be earned in the meantime.

Objective:

- Readers should learn about mathematical application in finance side
- Also should know the calculating methods and relevant meaning

2.1.1 Simple and compound interest:

If an amount of Rs. P is invested today, and a return of I per period is expected on this amount and if, at the end of each period we reinvest the interest earned during the period, then after the end of 1^{st} period we will have an amount , because PXi is the interest earned during the period. Now the amount P(1+i) will earn a return at the rate of i, and so the amount at the end of 2^{nd} period will become $P(1+i)+P(1+i)Xi=P(1+i)(1+i)=p(1+i)^2$. Now the amount $P(1+i)^2$ will earn a return at the rate of i during the third period and so the amount at the end of third period will be P(1+i) 2 +P(1+i) 2Xi= P(1+i) 2(1+i)= $P(1+i)^3$. Extending this logic to n periods, the amount at the end of n periods will be $A= P(1+i)^n$, Where A is known as compound amount and this formula is known as the compound amount formula. This formula states that if an amount of Rs. P is invested today and compound interest of i is considered, then the value of above investment after n periods will be $P(1+i)^n$. In case of simple interest, the investment after n years will be P+PXiXn=P(1+iXn).

2.1.2 Present Value:

We may be interested in the present worth (or present valve) of an amount of A that will be available to us after n periods. Since $A= P (1+i)^n$, it follows that

$$P = \frac{A}{(1+i)^n}$$

Thus, present value of A will always be less than A, because $(1+i)^n$ will always be greater than one, for n>1. The term $1/(1+i)^n$ is known as the discount-factor and tables are available that give the value of the discount factor for various values of i and n.

Example1: (i) Calculate the present valve of Re 1 received 4 years from now at a discount rate of 6 percent.

Solution : Present value is given by

$$P = \frac{A}{(1+i)^n}$$

Here A=1, i=6 and n=4

So $P = \frac{A}{(1.06)^4}$ = Rs. 0.792

(ii) Calculate present value of Rs. 100 received 4 years from now at the discount rate of 6 percent.

Solution: Present value is given by

$$P = \frac{A}{(1+i)^n}$$

= A X discount factor

= 100 X .792 = Rs. 79.20

(iii) Calculate present value of Rs. 100 to be received at the end of the year and Rs. 200 at the end of the 2 years at the discount rate of 6 percent.

Solution: Present value is given by

$$P = \frac{A}{(1+i)^n}$$

$$P = \frac{100}{(1.06)^1} + \frac{200}{(1.06)^2}$$

= 94.34+178.00 = Rs. 272.34

2.1.3 Annuities:

Annuity means equal investment at equal intervals for a designated period of time. The investment is generally assumed to be made at the end of each period. Thus, the basic feature of an annuity is periodic payments. As an example, suppose an investor has planned to invest in the public provident scheme, where by he is required to make a payment of Rs. P at the end each period for n periods and gets an interest of i per annum. The investor is interested in knowing as to how much amount he will get at the end of nth period. Thus, he is interested in the terminal value of the investment he is making over a period of time. The investment plan and the end of nth period is given below:

Period	1^{st}	2^{nd}	3^{rd}	…..	$(n-1)^{th}$	n^{th}
Amount invested	P	P	P	…..	P	P
Number of Years for Which this amount Will earn interest	(n-1)	(n-2)	(n-3)	…..	1	0
The accumulated Amount at the end of n^{th} period	$P(1+i)^{n-1}$	$P(1+i)^{n-2}$	$P(1+i)^{n-3}$	…..	P(1+i)	P

Thus, the total amount at the end of nth period will be:

$A= P (1+i)^{n-1} + P (1+i)^{n-2} + P (1+i)^{n-3} +.....+P (1+i)+P$ (I)

Multiply equation I by (1+i) and we get

$A (1+i) = P (1+i)^{n} + P (1+i)^{n-1} + P (1+i)^{n-2}+...+P (1+i)^{2}+P (1+i)$ (II)

Now subtract (I) from (II) and we get

$$A (1+i)-A = P (1+i)^{n} -P$$

Or $A+ Ai-A=P \{(1+i)^{n}-1\}$

Hence $A = \frac{P\{(1+i)^n - 1\}}{i}$

So, terminal value of annuity = $P\frac{(1+i)^n - 1}{i}$

Tables are available for calculating the annuity factor $\frac{(1+i)^n - 1}{i}$

for various values of i and

Example 2: (i) An investor deposits Rs. 1,000 in a saving institution. Each payment is made at the end of the year. If the payments deposited earn 6 percent interest compounded annually, how much amount he will receive at the end of 10 years.

Solution

Terminal value of annuity = $P\frac{(1+i)^n - 1}{i}$

Terminal value of annuity = $1000\frac{(1+0.06)^{10} - 1}{0.06}$

= $1000\frac{(1.790848 - 1)}{0.06}$

= Rs 13180.80

(ii) An individual plans investing Rs. 100 per year in a savings plan that earns 5 % interest compounded annually; calculate the sum of annuity payments at the end of 10 years?

Solution:

Terminal value of annuity = $100\frac{(1+0.05)^{10} - 1}{0.05}$

= Rs. 1257.7

2.1.4 Present value of an annuity

The present value of an annuity is obtained by the addition of the present value of each of the individual annuity payments.

The payments of the amount P are being made at the end of each period. Hence, we can find out the present value of each of these payments shown below

Period	1st	2nd	3rd	.....	$(n-1)^{th}$	n^{th}
Amount of payments	P	P	P	.....	P	P
Present value	$\frac{P}{(1+i)}$	$\frac{P}{(1+i)^2}$	$\frac{P}{(1+i)^3}$		$\frac{P}{(1+i)^{n-1}}$	$\frac{P}{(1+i)^n}$

Hence, PV of an annuity is

PV = $\frac{P}{(1+i)} + \frac{P}{(1+i)^2} + \frac{P}{(1+i)^3} + \ldots\ldots\ldots.. + \frac{P}{(1+i)^n}$ --- (I)

Multiply the equation (I) by (1+i)

PV(1+i) = + $\frac{P}{(1+i)^1}$ + $\frac{P}{(1+i)^2}$ +………..+ $\frac{P}{(1+i)^{n-1}}$ ---(II)

Subtract (I) from (II)

There fore PV = $\frac{P(1-(1+i)^{-n}}{i}$

Example:

Calculate the present value of Rupee 1 per year for three year at 6%

Solution

PV = $\frac{P(1-(1+i)^{-n}}{i}$

PV = $\frac{1(1-(1+.06)^{-3}}{.06}$

= 2.673

Example:

A project offers an annual return of rupees 250000, for four years. If the cost of money is 15%, calculate the present value of the project

Solution

PV = $\frac{P(1-(1+i)^{-n}}{i}$

PV = $\frac{250000(1-(1+.15)^{-4}}{.15}$

= Rs. 713750

Example:

Determine the PV of receiving Rs100 for year for ten years with interest rate at 6% per annum

Solution

PV = $\frac{P(1-(1+i)^{-n}}{i}$

PV = $\frac{100(1-(1+.06)^{-10}}{.06}$

$$PV = \frac{100(1-.558335)}{.06}$$

= Rs.736

2.2 SINKING FUND FACTOR

It represents the amount that has to be invested at the end of every year for a period of "n" years at the rate of interest "k", in order to accumulate Re. 1 at the end of the period.

Here is the equation $FVA = A\left[\frac{(1+k)^n - 1}{k}\right]$

We can rewrite as $A = FVA\left[\frac{k}{(1+k)^n - 1}\right]$

The expression $\left[\frac{(1+k)^n - 1}{k}\right]$ is called sinking fund factor

2.3 Discounting

Discounting is the process of finding the present value of an amount of cash at some future date, and along with compounding cash forms the basis of time value of money calculations. The discounted value of a cash flow is determined by reducing its value by the appropriate discount rate for each unit of time between the times when the cash flow is to be valued to the time of the cash flow. Most often the discount rate is expressed as an annual rate.

To calculate the net present value of a single cash flow, it is divided by one plus the interest rate for each period of time that will pass. This is expressed mathematically as raising the divisor to the power of the number of units of time.

As an example, suppose an individual wants to find the net present value of $100 that will be received in five years time. There is a question of how much is it worth presently, and what amount of money, if one lets it grow at the discount rate, would equal $100 in five years.

Let one assume a 12% per year discount rate.

NPV = 100 dollars divided by 1 plus 12% (0.12) divided by 1 plus 12% (0.12), etc.

$$NPV = \frac{100}{(1+0.12)^5}$$

Since 1.12^5 is about 1.762, the net present value is about $56.74.

So, an economy where money loses 12% of its value/purchasing power every year, how will it "feel" like having 100 $ in five years? Well, it will "feel" like having $56.74 now. So the interest rate you be charged at the bank will be very high to compensate for this fact.

2.3.1 Discount rate

The discount rate which is used in financial calculations is usually chosen to be equal to the cost of capital. Some adjustment may be made to the discount rate to take account of risks associated with uncertain cash flows, with other developments.

The discount rates typically applied to different types of companies show significant differences:

- Startups seeking money: 50 - 100 %
- Early Startups: 40 - 60 %
- Late Startups: 30 - 50%
- Mature Companies: 10 - 25%

Reason for high discount rates for startups:

- Reduced marketability of ownerships because stocks are not traded publicly
- Limited number of investors willing to invest
- Startups face high risks
- Over optimistic forecasts by enthusiastic founders.

One method that looks into a correct discount rate is the capital asset pricing model. This model takes in account three variables that make up the discount rate:

1. Risk Free Rate: The percentage of return generated by investing in risk free securities such as government bonds.

2. Beta: The measurement of how a company's stock price reacts to a change in the market. A beta higher than 1 means that a change in share price is more exaggerated then rest of shares in the same market. A beta less than 1 means that the share is stable and not very responsive to changes in the market. Less than 0 means that a share is moving in the opposite of the market change.

3. Equity Market Risk Premium: The return on investment that investors require above the risk free rate.

Discount rate = risk free rate + beta*(equity market risk premium)

2.3.2 Discount factor

The discount factor, P (T), is the number by which a future cash flow to be received at time T must be multiplied in order to obtain the current present value. Thus for a fixed annually compounded discount rate r we have

$$P(T) = \frac{1}{(1+r)^T}$$

For fixed continuously compounded discount rate we have $P(T) = e^{-rT}$

Write a whole number, fraction, or decimal as a percent. Write a percent as a whole number, fraction, or decimal.

Solving Percentage Problems

Identify the rate, base, and portion in percent problems. Use the percentage formula to find the unknown value when two values are known.

Increases and Decreases

Find the amount of increase or decrease in percent problems. Find the new amount directly in percent problems. Find the rate or the base in increase or decrease problems.

Summary

This unit has introduced you to the essentials of present annuity value, simple and compound interest and their important features. Specifically, this unit focused on:

- Finding simple and compound interest
- Finding annuity
- Details about sinking of fund
- Explanation of discounts and percentage

EXERCISE

1) Calculate the present valve of Re 1 received 4 years from now at a discount rate of 8 percent.
2) Calculate the present valve of Re 2 received 6 years from now at a discount rate of 4 percent.
3) Determine the PV of receiving Rs200 for year for ten years with interest rate at 5% per annum
4) Determine the PV of receiving Rs500 for year for ten years with interest rate at 6% per annum

NOTES

UNIT - III

DIAGRAMMATIC PRESENTATION OF DATA
(ONE DIMENSION AND TWO DIMENSION)

3.1 Introduction

Although tabulation is very good technique to present the data, but diagrams are an advanced technique to represent data. As a layman, one cannot understand the tabulated data easily but with only a single glance at the diagram, one gets complete picture often data presented. According to M.J. Moroney. Diagrams register a meaningful Impression almost before we think".

3.1.1 Importance or Utility of Diagrams

1. Diagrams give a very clear picture of data. Even a layman can understand it very easily and in a short time.
2. We can make comparison between different samples very easily. We don't have to use any statistical technique further to compare.
3. This technique can be used universally at any place and at anytime. This technique is used almost in all the subjects and various fields.
4. Diagrams have impressive value also. Tabulated data has not much impression as compared to Diagrams. Good diagrams impress a common man easily.
5. This technique can be used for numerical type of statistical analysis. e.g. to locate Mean, Mode, Median or other values.
6. It does not save only time and energy but also is economical. Not much money is needed to prepare even good diagrams.
7. These give us much more information as compared to tabulation. Technique of tabulation has its own limits.
8. This data is easily remembered. Diagrams, which we see, leave their impression much more lasting than other data techniques.
9. Data can be condensed with diagrams. A simple diagram can present what even cannot be presented by words.

3.1.2 Rules or Directions for Making Diagrams or Essentials of a Good Diagram

While preparing the diagrams we must observe some rules to make these diagrams more impressive and useful.

1. It must be attractive.
2. Its presentation must be proportionate in height and width.
3. It must be Economical in terms of money, energy and time.

4. It must be Intelligible.
5. Scale must be presented along with diagram
6. Size of figure should be such that it may occupy considerable portion of paper
7. It must be self-explanatory. It must indicate nature, place and source of data presented.
8. It must be neat and clean
9. Diagrams are of several types. The diagram drawn must be suitable to data.
10. If some points are to be clarified, footnotes may be given.
11. Different shades, colors can be used to make diagrams more easily understandable
12. Vertical diagram should be preferred to Horizontal diagrams.
13. If possible, suitable title may be given.
14. It must be accurate. Accuracy must not be done away with to make it attractive or impressive.

3.1.3 Limitations of Diagrams

1. Diagrams depict only approximate results. Those are not so accurate.
2. Due to above reason these can't be put for further analysis.
3. If scales are different, two diagrams can't be compared.
4. For false base diagrams, a layman may not make difference.

3.1.4 Types of Diagrams

Diagrams can be classified into following categories:

i. One Dimensional Diagram.
ii. Two Dimensional Diagrams.
iii. Three Dimensional Diagrams.
iv. Pictograms or Picture Diagrams.
v. Cartograms or Maps.

3.1.4.1 One Dimensional Diagrams

In this case only the length dimension is given the importance. These diagrams are either Bar or Line Diagrams.

Merits

1. These and easy to construct.
2. These are easy to understand.
3. Comparison can be made easily by this device.

a. Line Diagrams

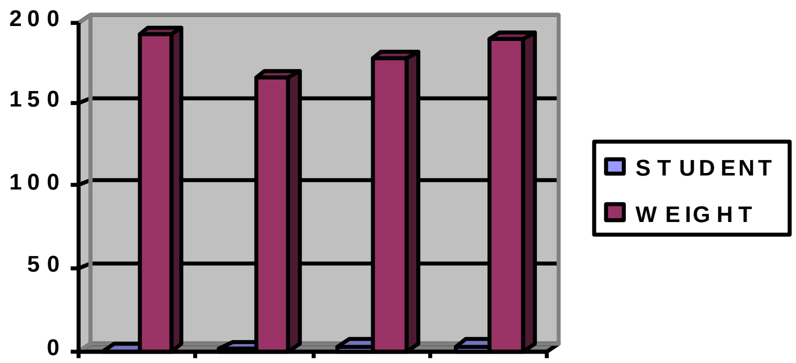
In these diagrams only line is drawn to represent one variable. These lines may be vertical or horizontal. The lines are drawn such that their length is in proportion to value of the terms or items so that comparison may be done easily.

b. Simple Bar Diagram

Like line diagrams these figures are also used where only single dimension i.e., length can present the data. Procedure is almost same, only the thickness of lines is measured. These can also be drawn either vertically or horizontally. Breadth of these lines or bars should be equal. Similarly distance between these bars should be equal the breadth and distance between them should be taken according to space available on the paper.

Example: Average wages of some firms is given below. Represent this by simple bar diagram.

STUDENT	1	2	3	4
WEIGHT	192	165	177	189



c. Multiple bar Diagrams

This diagram is used. When we have to make comparison between more than two variables. The no. Of variables may be 2. 3 or 4 or more. In case of 2 variables, pair of bars is drawn. Similarly in case of3 variables we draw triple bars. The bars are drawn on the same proportionate basis as in case of simple bars.

d. Sub-divided Bar Diagrams

The data which is presented by multiple bar diagram can be presented by this diagram. In this case we add different variables for a period and draw it on a single bar as shown in following examples. The components must be kept in same order in each bar. This diagram is more efficient if number of components is less i.e. 3 to 5.

e. Percentage Bar Diagram

Like sub-divided bar diagram, in this case also data of one particular period or variable is put on single bar, but in terms of percentages. Components are kept in the same order in each bar for easy comparison.

f. Duo-directional Bar Diagrams

In this case the diagram is on both the sides of base line i.e. to left and right or to above or below sides.

g. Broken Bar Diagram

This diagram is used when value of some variable is very much bigger or smaller than others. In this case the bars with bigger terms or items may be shown broken.

DIAGRAMMATIC PRESENTATION OF DATA (TWO & THREE DIMENSION)

3.1.4.2 Two Dimensional Diagrams

As in single bars it was mentioned that the width of each bar should be equal fm a certain variable or items. But in this case not only the length but the width also is taken proportionately in case of rectangles. But where the items are represented in square terms we use the technique of squares or circ1es. These are also known as Area Diagrams.

a. Rectangles

As mentioned above, not only the length, but also the breadth or width of each item is also taken proportionately.

Example: income of two salaried person is Rs. 10000 and 20000 respectively. Other variables are given below. Present it by Percentage rectangle diagram.

Expenditure	Food	Rent	Traveling	Others
Person1	1300	2700	4000	2000
Person1	3000	7000	6000	4000

Solution: As the income of two Persons are shown below

Expenditure	Person 1			Person2		
	Amount	%	Cum	Amount	%	Cum
Food(F)	1300	13	13	3000	15	15
Rent(R)	2700	27	40	7000	35	50
Traveling(T)	4000	40	80	6000	30	80
Others(O)	2000	20	100	4000	20	100
Total	10000			20000		

b. Squares

As told earlier this technique can be used effectively when given items or terms are squares, preferably having two zeros (00) after every term. Here we take square root of every item and then divide it by a suitable digit or number so as to get the size reduced to be put into the shape of a square on the given space. It is also useful technique when difference between the numbers is large.

c. Simple Circular Diagrams

As in case of squares, the technique is applied if terms are squares and difference between the terms is also large.

d. Sub-divided Circular Diagram

These are also called Pie or Angular Diagrams. We take the total of items and each item is given its proportionate angle taking total as 360°. In this case we may have to compare in terms of totals also, if data belongs to two cases. Note. If there is only single case, we may take any length of radius to suit our space.

3.1.4.3 Three Dimensional Diagrams

In case of two-dimensional diagrams we used only length and width as two dimensions, but in this case we make use of Length, Breadth and Height, all the three. Three-dimensional diagram may be used where cube roots may be taken easily and there is very large difference between the items or terms. These can be given the shape of cube, spheres etc. These are also known as volume diagrams.

3.1.4.4 Pictograms or Picture Diagrams

Dr. Otto Neurath of Vienna adopted this technique of presentation. It is therefore also known as Vienna method. A layman is most attracted by this type of diagrams. A symbol (Picture) is carefully chosen 10 represent the lot.

This symbol must be dear, easily understandable, and of a suitable size. For example we can use following symbols to show.

3.1.4.5 Cartograms or Maps

It is nothing else but map as we study in lower classes. Through it we can present the data, which is to be presented geographically. Location of important cities of India is given Below

3.2 GRAPHICAL PRESENTATION OF DATA

3.2.1 Introduction

Diagrams can present the data in an attractive style but still there is a method more than this. Diagrams are often used for publicity purpose but are not much use in statistical analysis. Hence graphic presentation is more effective and simple process.

3.2.2 Uses or Merits or importance of Graphs

1. It is more effective than diagrams
2. It is economica1 in terms of money, time and energy.
3. It gives us the picture in condensed form.
4. It is free from mathematical calculations.
5. It is most s1litable for comparison.
6. It is helpful in forecasting.
7. It is also helpful in statistical analysis. We can determine Median and Mode by this method.

3.2.3 Rules or Guidelines while Preparing a Graph

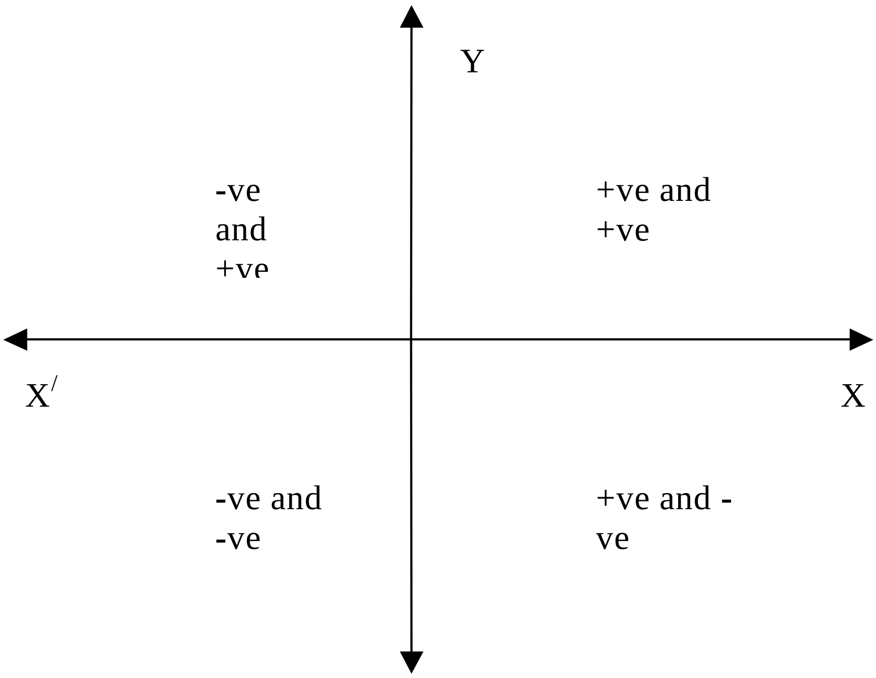
1. It must have proper heading
2. Scale must be provided along with graph.
3. False base line may be used according to need.
4. If required footnotes might be given.
5. While choosing scale, size or the space must be kept in view.
6. If possible Y-axis should be about 50% more than X-axis.
7. Independent variables should be taken on X-axis and dependent variable on Y-axis.
8. In time aeries graph, time should be shown on X-axis and other variable on Y-axis.
9. In frequency distribution, take value of variable on X-axis and frequency on Y-axis

3.2.4 Constitution of Graph Paper

Graphs are drawn on a special type of paper known as graph paper.

Graph papers are divided in small equal squares 1/10 cm. For the construction of graph, two straight lines, are drawn which cut each other at -900o The horizontal line is called 'X' axis and is usually denoted by X'OX. The vertical line is called 'Y' axis and is usually denoted by Y'OY. The point where they cut each other is known as 'Origin'. This origin divides the graph paper in four parts. These parts are known as quadrants. Zero value is taken on the point of origin for both lines. Positive values of X are taken towards right side on horizontal line and of Y towards upper side on vertical line. Negative values of X are taken towards the left side on horizontal line and of Y towards lower side on vertical line.

Positive and Negative Values:



As shown in above diagram, in first quadrant X and Y both have positive values. In second, X is negative and Y positive. In third, X and Y both are negative and in fourth Quadrant X is positive and Y negative. X-axis is also known as absciassa and Y-axis as Ordinate.

Choice of Scale

The scale indicates the unit of a variable that a fixed length of axis would represent. Scale may be different for both the axis. It should be taken in such a way so as to accommodate whole of the data on a given graph paper in a lucid and attractive style.

False Base Line

Sometimes it is difficult to take zero at origin and proceed for the graph as is in the following Example:

If we start with zero in this case, first fifty main divisions will remain empty, without any use. In will make the graph look clumsy. In such cases we use false base lines as shown below. If required, we can also take false base line on x-axis also.

3.2.5 Types or Graphs

There are two types of graphs.

1. Time series Graphs or Historigrams.
2. Frequency Distribution Graphs.

Time Series Graphs or Histograms

In this type of graphs, time is taken along X-axis and the other variables along Y-axis. The number: of variab1es on Y-axis may be one or more than one. These are known as One Variable, Two Variables or Three Variables graphs (Note: This method can only be applied in case the difference is much less as compared to the given data so that there may not 'be any overlapping. Due Co this shortcoming, 1st method as applied in Example 5 is preferred.)

3.3 MEASURES OF CENTRAL TENDENCIES

Measures or Central Tendencies.

Definitions

Different statisticians have defined measures of Central Tendency or Averages differently, which reduce the data to a single representative figure called an average or the measure of central tendency. The commonly used measures of central tendencies are the mean, the median and the mode ,The mode and median will come under Location measure and arithmetic Mean; Geometric Mean and Harmonic Mean are under mathematical measures

3.3.1 Arithmetic Mean (A.M.)

Arithmetic mean is obtained by adding all the terms and by dividing the total by the total number of items.

The AM $\bar{x}$ of n observations $x_1, x_2 \ldots. x_n$ is given by

$$AM\ \bar{x} = \frac{x_{1,}\ x_{2 \ldots\ldots\ldots} x_n}{n}$$

$$\text{i.e AM} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Note

In case of frequency distribution $\frac{x_i}{f_i}$ where i=1,2,3,4….n where f_i is the frequency of the variable x_i, then the AM

$$\bar{x} = \frac{f_1 X_1 + f_2 X_{2 \ldots\ldots\ldots} + f_n X_n}{N}$$

Where N= $\sum_{i=1}^{n} f_i$ = total number of frequencies

Example

Find the arithmetic mean for the following distribution

X	1	2	3	4	5	6
F	5	9	12	17	14	10

Solution

x_i	f_i	$f_i x_i$
1	5	5
2	9	18
3	12	36
4	17	68
5	14	70
6	10	60
	N=67	257

$$\text{Arithmetic Mean } \bar{x} = \sum_{i=1}^{n} \frac{f_i x_i}{N}$$

$$= \frac{257}{67} = 3.836$$

Note

Incase of grouped or continuous frequency distribution the Arithmetic mean

$$\bar{X} = A + \frac{h\sum_{i=1}^{n} f_i d_i}{N}$$

$$\text{Where } d_i = \frac{x_i - A}{h}$$

h-length of the interval

A- assumed mean

Example

Calculate the AM of the marks from the following table

Interval	0-10	10-20	20-30	30-40	40-50	50-60
Marks	12	18	27	20	17	6

Marks	Mid value	No. of Students	$d_i = x_i - 25$	$f_i d_i$
0-10	5	12	-20	-240
10-20	15	18	-10	-180
20-30	25	27	0	0
30-40	35	20	10	200
40-50	45	17	20	340
50-60	55	6	30	180
		N=100		$\sum f_i d_i = 300$

$$AM = \bar{X} = A + \frac{h\sum_{i=1}^{n} f_i d_i}{N}$$

$$= 25 + \frac{300}{100} = 28$$

Median

For a given distribution, median is the value of the variable which divides the distribution in to two equal parts. Incase of frequency distribution $\frac{x_i}{f_i}$ median is obtained by considering the cumulative frequencies.

Steps for calculating Median

1. Find $\frac{N}{2}$ where N= $\sum_{i=1}^{n} f_i$ = total number of frequencies
2. see the cumulative frequency just greater than $\frac{N}{2}$
3. corresponding value of x is median

Example

Obtain the median for the following frequency distribution

X	1	2	3	4	5	6	7	8	9
Y	8	10	11	16	20	25	15	9	6

Solution

x_i	f_i	$f_i x_i$
1	8	8
2	10	18
3	11	29
4	16	45
5	20	65
6	25	90
7	15	105
8	9	114
9	6	120
	N=120	

$$\frac{N}{2} = \frac{120}{2} = 60$$

Just greater than 60 is 65

The value of x corresponding to 60 is 5

Hence median is 5

Note

In case of continuous distribution, the value of median is obtained by the following formulae

Median = $l + \frac{h}{f}(\frac{N}{2} - C)$

Where

l = lower limit of the median class

h = magnitude of the median class

f = frequency of the median class

c = is the cumulative of the class preceeding the median

Example

Find the median age of the following number of males in the particular locality

Age	20-30	30-40	40-50	50-60	60-70
No.	3	5	20	10	5

Solution

Age	Mid value x_i	No. of males f_i	C.f
20-30	25	3	3
30-40	35	5	8
40-50	45	20	28
50-60	55	10	38
60-70	65	5	43
		N=43	

$$\text{Here} = \frac{N}{2} = \frac{43}{2} = 21.5$$

The c.f. just greater that 21.5 is 28

l = lower limit of the median class=40

h = magnitude of the median class=10

f = frequency of the median class=20

c = is the cumulative of the class preceeding the median=8

Median = $40+\frac{10}{20}(\frac{43}{2}-8)$ = 46.75

Mode

Mode is the value which occurs most frequently in a set of observations and around which the other items of the set cluster closely. Incase of discrete frequency distribution mode is the value of corresponding to the maximum frequency

Example: Find mode for the following

X	1	2	3	4	5	6	7	8
F	4	9	6	25	22	18	7	3

Solution

Here the maximum frequency 25 is appeared 4 times

Hence mode is 4

Note

Incase of frequency distribution the following formula can use

Mode = $l+\frac{h(f_1-f_0)}{2f_1-f_0-f_2}$

l - lower limit of the modal class

h = magnitude

f_1 = frequency of the model class

f_0 f_2 = are the frequencies of preceeding and succeeding

Example: Find mode for the following frequency value

Class interval	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
frequency	5	8	7	12	28	20	10	10

Solution:

The class interval corresponding to the maximum frequency namely 28 is 30-50. hence the modal class is 40-50

Mode = $l + \frac{h(f_1 - f_0)}{2f_1 - f_0 - f_2}$

l - lower limit of the modal class = 40

h = magnitude = 10

f_1 =frequency of the model class = 28

f_0 f_2 = are the frequencies of preceeding and succeeding = 12,20

Therefore Mode = $40 + \frac{10(28-12)}{2(28)-12-20}$

= 46.67

Empirical relation between Mean, Median and Mode

Mode=3median-2mean

3.3.2 GEOMETRIC MEAN AND HARMONIC MEAN

Geometric mean is the nth root of the product of n items in a given series. Multiplying the values of the items together and then taking it to its root corresponding to the number of items obtain the geometric mean.

The Geometric mean (GM) can be calculated for individual observation in the following manner

Steps for calculating GM

i. Take log values of all the items of a given series.

ii. Multiply each log Value to its respective frequencies.

iii. Add the values and divide by the total number of frequencies.

iv. Take the value of antilog from the antilog table and the result would be the geometric mean.

Example

Calculate the geometric mean of the following data

6.5, 169, 11,112.5, 14.2, 75, 35.5 and 215

Solution

X	Log x
6.5	.8129
169	2.2279
11	1.0414
112.5	2.0512
14.2	1.1523
75	1.18751
35.5	1.5502
215	2.3324
N=8	$\sum \log x = 13.0434$

$$\mathbf{Gm} = \text{antilog} \frac{\sum \log x}{N} = \frac{13.0434}{8} = 42.70$$

Example 2

Calculate GM for the following distribution

weight	100-104	105-109	110-114	115-119	1120-142
Frequency	24	30	45	65	72

Solution

Weight	f	Mid point m	Log(mid point)	f*log(midpoint)
100-104	24	102	2.0086	48.2064
105-109	30	107	2.0294	60.8820
110-114	45	112	2.0492	92.2140
115-119	65	117	2.0682	134.4330
120-124	72	122	2.0684	150.2208
	N=236			$\sum f \log \text{midpoint} = 485.9562$

$$\mathbf{Gm} = \text{antilog} \frac{\sum f\log \text{midpoint}}{N} = 114.6$$

3.3.3 Harmonic Mean

Harmonic mean of a given series is the reciprocal of the arithmetic average of the reciprocals of the values of its various items. Suppose X1,X2,…,Xn are the N items of given series, then the formula for harmonic mean is written as

$$HM = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \text{..........} \frac{1}{x_n}}$$

Example

Calculate HM for the following data

125, 130, 75, 10, 45, 5, .5, .4, 500,150

Solution:

X	$\frac{1}{X}$
125	.00800
130	.00770
75	.01333
10	.10000
45	.02222
5	.20000
0.5	2.0000
0.4	2.5000
500	.00200
150	.00666
N=10	$\sum\frac{1}{X}$ =4.85991

$$HM = \frac{N}{\sum \frac{1}{X}} = 10/4.85991 = 2.06$$

Example2:

Calculate HM from the following data

Income(Rs)	No. of persons
10-20	4
20-30	6
30-40	10
40-50	7
50-60	3

Solution

Income(Rs)	f	m	$\frac{f}{m}$
10-20	4	15	.2667
20-30	6	25	.2400
30-40	10	35	.2857
40-50	7	45	.1556
50-60	3	55	.0545
	N=30		$\sum\frac{f}{m}$ =1.0025

$$HM = \frac{N}{\sum\frac{f}{m}}$$

Summary

This unit has provided a conceptual framework on various types of diagram and Mean with their limitations. Specifically, this unit focused on:

- The need for diagram representation
- One, two and three dimensional diagram
- Arithmetic mean ,GM and HM
- Determining AM,GM and HM

Exercise

1) Prepare line diagram for the following data

STUDENT	1	2	3	4
WEIGHT	192	165	177	189

2) Narrate about the Types of Diagrams

3) Write about one dimensional diagram

4) Write the application of two dimensional diagrams

5) What do you mean by "Measures of Central Tendency?

6) Find the arithmetic mean for the following distribution

X	1	2	3	4	5	6
F	9	5	13	16	14	10

7) Find the median age of the following number of males in the particular locality

Age	25-35	35-45	45-55	55-65	65-75
No.	3	8	17	10	5

8) Calculate HM for the following data

120, 135, 79, 10, 45, 5, .5, .4, 500,150

NOTES

UNIT - IV
MEASURES OF VARIATION

4.1 Introduction

Measure of central tendency namely mean, median and mode give us an idea of the concentration of the observations about the central part of the distribution. Hence if you want to know the complete idea of a distribution we need some other measures. Once such measure is dispersion or variation.

4.2 Measure of dispersion

The following are some important measures dispersion.

1. quartile deviation (Q)
2. Mean deviation (M.D.)
3. Standard deviation (S.D)

Quartile deviation (Q)

The quartile deviation is given by u

Where Q_1 and Q_3 are the first and third quartiles of the distribution respectively

Where Q_1 = size of the items corresponding to c.f. just greater than $\frac{N+1}{4}$

Q_3 = size of the items corresponding to c.f. just greater than $\frac{3(N+1)}{4}$

EXAMPLE:

1. Calculate the quartile deviation and coefficient of quartile deviation from the following data.

Height (in inches)	Frequency
50	10
51	12
52	15
53	10
54	14
55	18
56	6

Solution:

Height in inches	F	c.f.
50	10	10
51	12	22
52	15	37
53	10	47
54	14	61
55	18	79
56	6	85
	85	

Quartile deviation (Q.D) $= \frac{Q_3 - Q_1}{2}$

Now Q_1 size of the item corresponding to the c.f. just greater than $\frac{N+1}{4}$

= size of the item corresponding to the c.f. just greater than $\frac{85}{4}$ ie 21.5

= size of the item corresponding to the c.f. just greater than 21.5 is 22

= 51 inches.

Q_3 = size of the items corresponding to c.f. just greater than $\frac{3(N+1)}{4}$

= size of the item corresponding to c.f. just greater than $\frac{3(85+1)}{4} = 64.5$

= size of the item corresponding to c.f. just greater than 64.5 is 79 corresponding to that value is 55 inches.

Quartile deviation (Q.D) $= \frac{Q_3 - Q_1}{2} = \frac{55-51}{2} = 2$ inches

Coefficient of quartile deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{55-51}{55+51} = .0377$

inches

Note: if it is for the continuous frequency distribution the following steps are required

1. Find $\frac{N}{4}$ where N = $\sum f_i$ is the total frequency
2. See the c.f. just greater than $\frac{N}{4}$
3. The class corresponding to this c.f. contains Q1
4. $Q_1 = l + \frac{h}{f}\left[\frac{N}{4} - c\right]$ where

   l = lower limit of the class containing Q_1

   f = frequency of the class containing Q_1

   h = magnitude of the class containing Q_1

   c = c.f. of the class preceeding the class containing Q_1

for computing Q_3

1. see the c.f. just greater than $\frac{3N}{4}$
2. the class corresponding to the c.f. contains

$$Q_3 = l + \frac{h}{f}\left[\frac{3N}{4} - c\right]$$

l = lower limit of the class containing Q_3

f = frequency of the class containing Q_3

h = magnitude of the class containing Q_3

c = c.f. of the class preceeding the class containing Q3`

EXAMPLE:

Calculate the quartile deviation of the marks of 39 students given below

Marks	0-5	5-10	10-15	15-20	20-25	25-30
No. of Students	4	6	8	12	7	2

Solution

Marks	No. of students	c.f.
0-5	4	4
5-10	6	10
10-15	8	18
15-20	12	30
20-25	7	37
25-30	2	39
	$\sum f$ = 39	

$$\frac{N}{4} = \frac{39}{4} = 9.75$$

To find Q_1 :

The c.f. just greater than 9.75 is 10 corresponding to that class is 5-10

$$\therefore \quad Q_1 = l + \frac{h}{f}\left[\frac{N}{4} - c\right]$$

l = 5

f = 6

h = 5

c = 4

$$Q_1 = 5 + \frac{5}{6}\left[\frac{39}{4} - 4\right] = 9.79$$

to find Q_3

$\left[\frac{3N}{4}\right]$ = 29.25

the c.f. just greater than 29.5 is 30 and corresponding class is 15-20

$$\therefore \quad Q_3 = l + \frac{h}{f}\left[\frac{3N}{4} - c\right]$$

l =15

f = 12

h = 5

c = 18

$$Q_3 = 15 + \frac{5}{12}\left[\frac{3\times 39}{4} - 18\right] = 19.69$$

Quartile deviation (Q.D) $= \frac{Q_3 - Q_1}{2} = \frac{19.69 \; - 9.79}{2} = 4.95$

4.3 MEAN DEVIATION AND STANDARD DEVIATION

If $\frac{x_i}{f_i}$ i= 1,2,3,… n is a frequency distribution then the mean deviation from the average A is given by

$$A = \frac{1}{N}\sum f_i \left|x_i - A\right|$$

$$where \sum f_i = N$$

Standard deviation is given by $\sigma = \sqrt{\frac{1}{N}\left(\sum_i f_i \left(x_i - \bar{x}\right)^2\right)}$

The square of the deviation is called variance $\sigma^2 = \frac{1}{N}\left(\sum_i f_i \left(x_i - \bar{x}\right)^2\right)$

EXAMPLE

1. find the mean deviation from the mean for the following distribution

Marks	0-10	10-20	20-30	30-40	40-50
No of the students	5	8	15	16	6

SOLUTION:

Marks	Mid value(x_i)	f_i	$d_i=x_i-25$	$f_i\,d_i$	$\lvert x_i-\bar{x}\rvert$	$f_i\lvert x_i-\bar{x}\rvert$
0-10	5	5	-20	-100	22	110
10-20	15	8	-10	-80	12	96
20-30	25	15	0	0	2	30
30-40	35	16	10	160	8	128
40-50	45	6	20	120	18	108
		$\sum f_i=50$		$\sum f_i d_i=100$		472

Mean $\bar{x}=A+\frac{\sum f_i d_i}{N}=25+\frac{100}{50}=27$

Mean deviation from the mean = $\frac{1}{N}\sum_i f_i\lvert x_i-\bar{x}\rvert$ = $\frac{1}{50}\times 472=9.44$

EXAMPLE 2

Calculate the mean deviation from the median of the following distribution

Marks	0-10	10-20	20-30	30-40	40-50
No of the students	5	8	15	16	6

SOLUTION:

Marks	Mid value(x_i)	f_i	c.f.	$\lvert x_i-M_d\rvert$	$f_i\lvert x_i-M_d\rvert$
0-10	5	5	5	23	115
10-20	15	8	13	13	104
20-30	25	15	28	3	45
30-40	35	16	44	7	112
40-50	45	6	50	17	102
		$\sum f_i=50$			478

Here $\frac{N}{2}=\frac{50}{2}=25$

The c.f. just greater than 25 is 28 and the corresponding class is 20-30 . hence 20 -30 is median class.

$$median = l + \frac{h}{f}\left(\frac{N}{2} - c\right)$$

$$l = 20, h = 10, \ f = 15 and$$

$$c = 13$$

$$20 + \frac{10}{15}\left(\frac{50}{2} - 13\right) = 28$$

mean deviation from the median = $\frac{1}{N}\sum_i f_i |x_i - M_d|$ =9.56

EXAMPLE:

Calculate the mean and standard deviation from the following table giving the age distribution of 542 members

Age	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No of members	3	61	132	153	140	51	2

SOLUTION:

Age	x_i	f_i	$d_i = \frac{x_i - 55}{10}$	$f_i d_i$	$f_i d_i^2$
20-30	25	3	-3	-9	27
30-40	35	61	-2	-122	244
40-50	45	132	-1	-132	132
50-60	55	153	0	0	0
60-70	65	140	1	140	140
70-80	75	51	2	102	204
80-90	85	2	3	6	18
		N=542		-15	765

Now, $\bar{x} = A + \frac{h}{N}\sum_i f_i d_i$

$$\bar{x} = 55 + \frac{10}{542}(-15) = 54.723$$

Variance:

$$\sigma^2 = h^2\left[\frac{1}{N}\sum_i f_i d_i^{\,2} - \left(\frac{1}{N}\sum_i f_i d_i\right)^2\right] = 141.067$$

$\therefore$ Standard deviation $\sigma = \sqrt{141.067} = 11.87$

4.4 CORRELATION AND REGRESSION

4.4.1 Introduction

This unit shows you how statistics can summarize the relationship between two factors based on a bivariate data set with two columns of numbers. The correlation will tell you how strong the relationship is, and regression will help you predict one factor from the other.

Objectives

After reading this unit, you will be able to:

- Define correlation coefficient with its properties.
- Calculate correlation coefficient and interpret.
- Appreciate the role of regression.
- Formulate the regression equation and use it for estimation and prediction.

4.4.2 Correlation Analysis

Definition: Correlation is the relation that exists between two or more variable. If two variable are related to each other in such a way that change in one creates a corresponding change in the other, then the variable are said to be correlated.

Examples

1. Relationship between the heights and weights
2. Relationship between the quantum of rainfall and the yield of wheat

Correlation Analysis is a statistical technique used to measure the degree and direction of relationship between the variables.

4.4.3 Karl Pearson's coefficient of correlation

Definition- Given a set of pairs of observation relating to two variables X and Y, Coefficient of Correlation between X and Y, denoted by the symbol 'r' is defined as-

$$r = \frac{Cov.(X,Y)}{\sigma_x, \sigma_y}$$

Where, Cov.(X,Y) =Covariance of X and Y

σ_x = Standard Deviation of X variable

σ_y = Standard Deviation of variable Y

Properties of coefficient of correlation

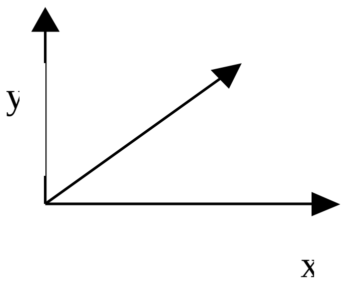
1. **Independent of choice of origin -** The coefficient of correlation (r) is independent of the choice of origin. In other words, the value of 'r' is not affected even if each of the individual values of X and Y is increased or decreased by some constant.
2. **Independent of choice of scale -** The coefficient of correlation (r) is independent of the choice of scale of observations. In other words, the value of 'r' is not affected even if each of the individual values of X or Y is multiplied or divided by some constant.
3. **Independent of Units of Measurement -** The correlation coefficient r is a pure number and is independent of the units of measurement. This means that if, for example, X represents height in inches and Y weight in lbs., then the correlation coefficient between X and Y will neither be in inches nor in lbs. or any other unit, but only a number.
4. The correlation coefficient r lies between -1 and +1.
5. The coefficient of correlation is the geometric mean of two regression coefficients $r = \sqrt{b_{xy} \times b_{yx}}$

Value of correlation

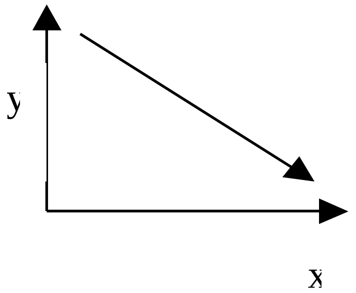
Correlation values always it will lie between

(a) If r = +1

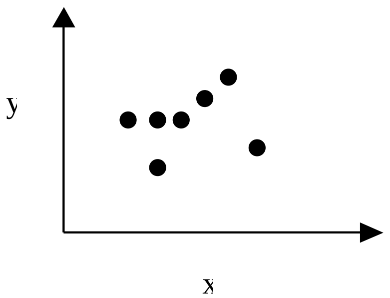
Then there is a perfect positive correlation between the variables, this we can show diagrammatically as follows



b) If r = -1 then There exists perfect negative correlation between the variable , diagrammatically as follows



(c) If r=0, then there is no relation between the variables. only the cluster of points we can find without any relation



Example

Find the correlation between the following variables

Add expenses	7	10	9	4	11	5	3
Sales(in thousands)	12	14	13	5	15	7	4

Solution

Add Expenses(X)	Sales(Y)	$(X-\bar{X})$	$(Y-\bar{Y})$	$(X-\bar{X})(Y-\bar{Y})$	$(X-\bar{X})^2$	$(Y-\bar{Y})^2$
7	12	0	2	0	0	4
10	14	3	4	12	9	16
9	13	2	3	6	4	9
4	5	-3	-5	15	9	25
11	15	4	5	20	16	25
5	7	-2	-3	6	4	9
3	4	-4	-6	24	16	36
7	10			83	58	124

From the above table we can get the following

$$\bar{X} = \frac{\sum X}{n} = 7 \text{ and } \bar{Y} = \frac{\sum Y}{n} = 10$$

$$\sum(X-\bar{X})(Y-\bar{Y}) = 83,\ \sum(X-\bar{X})^2 = 58 \text{ and } \sum(Y-\bar{Y})^2 = 124$$

$$\text{Cov(X Y)} = 83,\ \sigma_X = \sqrt{58} \text{ and } \sigma_Y = \sqrt{58}$$

$$r = \frac{\text{cov}(XY)}{\sigma_X \sigma_Y} = \frac{83}{\sqrt{58 \times 124}} = 0.97$$

4.4.3 Rank correlation

Let ,i=1, 2,…n be the ranks of 'i' individuals in two characteristic A and B respectively. The coefficient of correlation between the ranks x_i's and y_i's is called the rank correlation coefficient between the characteristics A and B for that group of individuals and its given by

$$r = 1 - \frac{\sum_{i=1}^{n} d_i^2}{n\left(n^2 - 1\right)},$$

Where $d_i = x_i - y_i$, this formula is called spearman's formula.

Example:

Find rank correlation coefficient for proficiency in mathematics and physics

Ranks is Maths	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Rank in Physics	1	10	3	4	5	7	2	6	8	11	15	9	14	12	16	13

Solution:

Ranks in Maths(x_i)	Ranks in Physics (y_i)	$d_i = x_i - y_i$	d_i^2
1	1	0	0
2	10	-8	64
3	3	0	0
4	4	0	0
5	7	-1	1
6	7	-1	1
7	2	5	25
8	6	2	4
9	8	1	1
10	11	-1	1
11	15	-4	16
12	9	3	9
13	14	-1	1
14	12	2	1
15	16	-1	1
16	13	3	9
		0	136

Rank correlation coefficient

$$R = 1-\frac{6\sum d_i^2}{n\left(n^2-1\right)} = 1-\frac{6\times 136}{16\left(256-1\right)}$$

$$= 1-\frac{816}{4080} = 1-0.2=0.8$$

4.4.5 Regression Analysis

Regression is the measure of average relationship between two or more variables in terms of the original units of the data. For example, after having established that two variables(say sales and advertising expenditure)are correlated, one may find out the average relationship between the two to estimate the unknown values of dependent variable(say sales) from the known values of independent variable(say advertising expenditure.

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 variables from the known values of independent variables.

4.4.6 Distinction between correlation and regression

Correlation differs from Regression in the following respect:

Correlation	Regression
Correlation measures degree and direction of relationship between the variable	Regression measures the nature and extent of average relationship between two or more variables in terms of the original units of the data.
It is a relative measure showing association between variables	It is a absolute measure of relationship
Correlation coefficient is independent of choice of both origin and scale.	Regression coefficient is Independent of choice of origin and not scale.
Expression of relationship Between the variable ranges from -1 to +1.	Expression of relationship Between the variable may be in any of the forms like Y=a+bX Y= $a+bX+cX^2$
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.

4.4.7 Regression Lines

For simple linear regression model, there are two regression lines as follows:

1) Regression line of X on Y

2) Regression line of Y on X

Regression line of X on Y: this line can be expressed as follows:

$$(X-\overline{X})=b_{xy}(Y-\overline{Y})$$

or $(X-\bar{X})=r\frac{\sigma_x}{\sigma_y}(Y-\bar{Y})$

$[Since\ b_{xy}=r\frac{\sigma_x}{\sigma_y}]$

in the same way **Regression line of Yon X**

$$(Y-\bar{Y})=b_{yx}(X-\bar{X})$$

or $(Y-\bar{Y})=r\frac{\sigma_Y}{\sigma_X}(X-\bar{X})$

$[Since\ b_{yx}=r\frac{\sigma_y}{\sigma_x}]$

Where, $\bar{X}$ =Arithmetic Mean of X series

$\bar{Y}$ = Arithmetic Mean of Y series

σ_x =Standard deviation of X series

σ_y = Standard deviation of Series

r = Coefficient of correlation between two variables X and Y.

Note: we have another way to find b_{yx} and b_{xy}

$b_{yx}=\frac{\sum(Y-\bar{X})(X-\bar{X})}{\sum(X-\bar{X})^2}$ And $b_{xy}=\frac{\sum(Y-\bar{X})(X-\bar{X})}{\sum\left(Y-\bar{Y}\right)^2}$

Example:

From the following data, find

(i) The two lines of regression equations

(ii) The co-efficient of correlation between the marks in physics and chemistry

(iii) the most likely marks in chemistry when marks in economics are 30

Marks in physics	25	28	35	32	31	36	29	38	34	32
Marks in chemistry	43	46	49	41	36	32	31	30	33	39

Solution:

X	Y	$X-\bar{X}=X-32$	$Y-\bar{Y}=Y-38$	$(X-\bar{X})^2$	$(Y-\bar{Y})^2$	$(X-\bar{X})(Y-\bar{Y})$
25	43	-7	5	49	25	-35
28	46	-4	8	16	64	-32
35	49	3	11	9	121	33
32	41	0	3	0	9	0
31	36	-1	-2	1	4	2
36	32	4	-6	16	36	-24
29	31	-3	-7	9	49	21
38	30	6	-8	36	64	-48
34	33	2	-5	4	25	-10
32	39	0	1	0	1	0
320	380	0	0	140	398	-93

$$\bar{X}=\frac{\sum X}{n}=\frac{320}{10}=32$$

and = $\bar{Y}=\frac{\sum Y}{n}=\frac{380}{10}=38$

$$b_{yx}=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}=\frac{-93}{140}$$

=-.6643

There fore

X onY equation is

$$=(x-\bar{x})=b_{xy}(y-\bar{y})$$

$$=(x-32)=-.2337(y-38)$$

x = -.2337y+40.8806

Y on X equation is

$$(y-\bar{y})=b_{yx}(x-\bar{x})$$

$$(y-38) = -.6643(x-32)$$

$$= \quad y = .6643x + 59.2576$$

The co-efficient of correlation

$$r^2 = b_{yx} \times b_{xy} = .1552$$

$$r = \pm\sqrt{b_{yx} \times b_{xy}} = \pm\sqrt{.1552} = \pm.394$$

Now we have to find the most likely marks in chemistry(y) when marks in physics(x) are 30.

We use the line of regression of **y on x**

$$y = -.6643x + 59.2576$$

$$= -.6643(30) + 59.2576$$

$$= 39.32$$

$$\cong 39$$

Summary

This unit has introduced you to the essentials of Measures of variation , correlation and regression with their important features. Specifically, this unit focused on:

- Finding quartile deviation
- Mean deviation
- Standard deviation
- Meaning and role of correlation.
- Understanding correlation through scatter diagram.
- Definition and properties of Pearson's correlation coefficient.
- How to compute correlation coefficient.
- Basics of regression analysis.
- Linear regression model and statistical validation.

EXERCISE:

1) Calculate the quartile deviation and coefficient of quartile deviation from the following data.

Weight (in KG)	frequency
59	10
53	12
52	18
58	10
54	19
55	18
60	6

2) Calculate the quartile deviation of the marks of 29 students given below

Marks	0-5	5-10	10-15	15-20	20-25	25-30
No. of Students	4	6	8	2	7	2

3) Find the mean deviation from the mean for the following distribution

Marks	0-10	10-20	20-30	30-40	40-50
No of the students	8	5	25	6	6

4) Calculate the mean deviation from the median of the following distribution

Marks	0-10	10-20	20-30	30-40	40-50
No of the students	25	8	5	6	6

5) Following figures give the rainfall in inches and production in tons, for Rabi and Paddy crops for number of years. Find the coefficient of correlation between rainfall and total production:

Rainfall	20	22	24	26	28	30	32
Rabi production	15	18	20	32	40	39	40
Paddy production	15	17	20	18	20	21	15

6) Calculate coefficient of correlation from the following data:

X	100	200	300	400	500	600	700
Y	0.3	0.5	0.6	0.8	1.0	1.1	1.3

7) Percentage of marks of 10 students in ME and BE Examination are follows:

ME	65	58	40	67	72	48	54	76	54	66
BE	70	75	62	45	78	60	40	64	45	61

Under similar conditions, how much a student securing 76 marks in BE may expect in ME?

NOTES

....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....

UNIT - V
INDEX NUMBERS AND TIME SERIES

5.1 Index Numbers

Definition:

An index number is a statistical measure designed to show changes in a variable or a group of related variables with respect to time geographical location or other characteristics.

5.1.1 Types of Index Number

There are various types of index numbers but in brief we shall take three kinds and they are

1. Price index

For measuring the value of money the general price index is used. It is an index number which compares the prices for a group of commodities at a certain time or at a place with prices of a base period. There are wholesale price index numbers and retail price index numbers. The wholesale price index reveals the changes in the general price level of a country. Retail price index reveals the changes in the retail prices of commodities, such as, consumption goods, bank deposits, bonds, etc.

2. Quantitative Index

Quantity index numbers study the changes in the volume of goods produced or consumed; for instance, industrial productions, agricultural production, import, export, etc. They are useful and helpful to study the output in an economy.

3. Value Index

These index numbers compare the total value of a certain period with the total value of the base period. Here the total value is equal to the price of each, multiplied by the quantity; for instance, indices of profits, sales, inventories, etc.

Formula

The problem is perhaps of greater theoretical then practical importance. In general, choice of the formula to be used depends upon the data available and the nature, purpose and scope of the enquiry.

Notations:

Base Year: The year selected for which comparison i.e. the year with reference to which comparisons are made. It is denoted by 0,

Current Year: The year for which comparisons are sought or required.

P_0 = Price of a commodity in the Base Year.

P_1 = Price of a commodity in the Current Year.

q_0 = Quantity of a commodity consumed or purchased during: the Base Year.

q_1 = Quantity of a commodity consumed or purchased in the Current Year.

w = Weight assigned to a commodity according to its relative importance in the group.

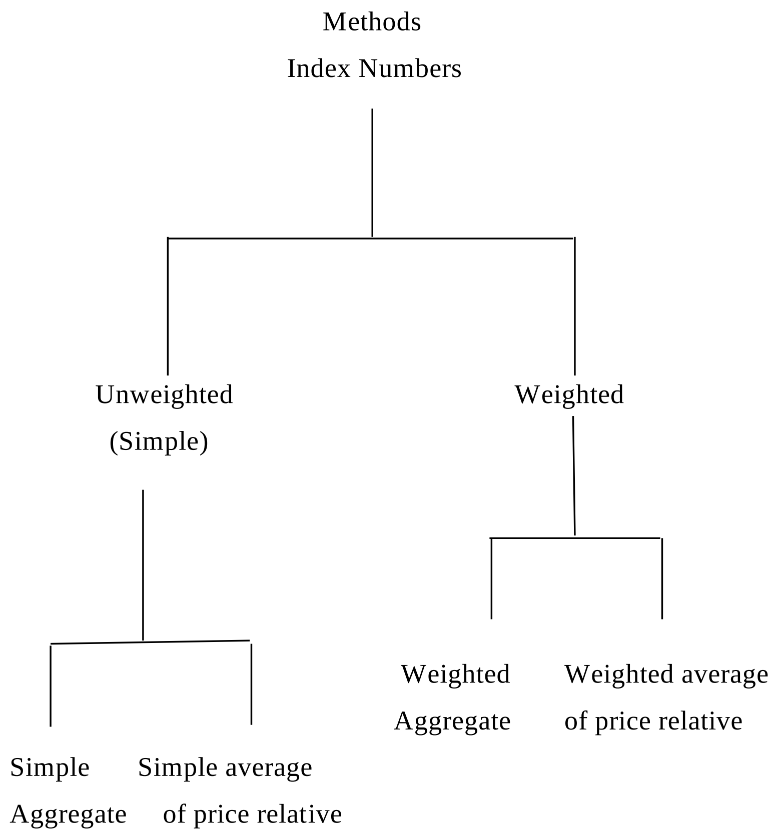
P_{01}= Price Index Number for the base year with reference to the base year.

P_{10}= Price Index Number for the current year with reference to the current year.

q_{01} =Quantity Index Number for the current year with reference to the base year.

q_{10} =Quantity Index Number for the base year with reference to the current year.

The various methods of construction of index numbers are given below:



5.1.2 Unweighted:

This is the simplest method of constructing the index numbers, the prices of the different commodities of the current year are added and the totals is divided by the sum of the prices of the base year commodity and multiply by 100. Symbolically

$$P_{01} = \frac{\sum P_1 \times 100}{\sum P_0}$$

Where

= price index no. for the current year with reference to the base year

SP$_1$= aggregate of prices for the current year

SP$_0$= aggregate of prices for the base year

5.1.3 Simple Average method:

In this method the price relative of each item is calculated separately and then averaged. A price relative is the price of the current year expressed as a percentage of the price of the base year.

$$P_{01} = \frac{\frac{P_1 \times 100}{P_0}}{N} = \frac{\sum P}{N}$$

Where N is Number of items.

When geometric mean is employed, instead of the arithmetic mean the formula is

$$P_{01} = anti\log \frac{\sum \log\left(\frac{P_1 \times 100}{P_0}\right)}{N} = anti\log \frac{\sum \log P}{N}$$

Where P = $\frac{P_1 \times 100}{P_0}$

5.1.4 Weighted index numbers:

Weighted index numbers are two types.

The Index numbers we have studied above are Unweighted and they assign equal importance to all the items included in the index. The purpose of weighting is to make the index numbers more representative and to give more importance to them.

Weighted index aggregate numbers: According to this method, prices themselves are weighted by quantities; i.e., pxq. Thus Physical quantities are used as weights. There are various methods of assigning weights, and thus various formulas have been formed for the construction of index numbers.

Some of the important names of the formula are given below

1. Laspeyres method
2. Paasche's method
3. Dorbish and Bowley's method
4. Fisher's ideal method
5. Marshall-Edgeworth method
6. Kelly's method and
7. Walsch's method

5.1.5 Test of Consistency of index numbers

Several formulae have been studied for the construction of index numbers. The question arises as to which formula is appropriate to a given problem. A number of tests have been developed and the important among them are:

1. Time Reversal Test

Reversibility is an important property that an index number should posses. A good index number should satisfy the time reversal test. In the words of Irving Fisher, "The formula for calculating an index number should be such that it gives the same ratio between one point of comparison and the other, no matter which of the two is taken as the base: or putting it in another way, the index number reckoned forward should be reciprocal of the one reckoned backward" One of the advantages claimed in favor of Fisher's formula is that it makes the index number reversible.

2. Factor Reversal Test

Another basic test is that the formula for index number ought to permit interchanging the prices and quantities without giving inconsistence results.

Unit Test

This test requires that the index number formula should be independent of the units, in which the prices or quantities of various commodities are quoted. This test is satisfied by all the formula except the Simple Aggregative index.

Circular Test

This is another test for the adequacy of an index number. It is an extension of time reversal test. According to this test the index should work in a circular fashion. This test is satisfied only by the index number formula based on

1. Simple Geometric mean of the price relatives
2. Kelly's fixed base method

5.1.6 Chain Base Method

The base may be fixed or changing so for we have used the fixed base method in various formula. In this fixed base method, the base remains constant throughout. On the other hand, the chain base method, the relative for each year found out from the prices of the immediately preceding year. Thus the base changes from year to year. The following formula is used for finding out the chain index.

$$\text{chain index} = \frac{\text{Link relative of the current Year} \times \text{Prevoius Year chain index}}{100}$$

5.1.7 Fixed base Method

Here base Year does not change, any change in the commodities will involve the entire index number to be recast.

Consumer Price index or cost of living index numbers

There are two methods of constructing consumer price index. They are

1. Aggregate expenditure method or aggregate method
2. Family budget method or method of weighted relatives

5.1.8 Aggregate expenditure method

This method is based upon the Laspeyre's method. It is widely used. The quantities of commodities consumed by a particular group in the base year are the weight. The Formula is

Consumer price index number = $\frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$

5.1.9 Family Budget Method

Here an aggregate expenditure of an average family on various items are estimated and it is value weighted the formula is consumer price index = $\frac{\sum pv}{\sum V}$

Where P= $\frac{P_1}{P_0} \times 100$

and V=Value of weight

5.2 TIME SERIES ANALYSIS

Forecasting, or predicting, is an essential tool in any decision- making process. Its uses vary from organization to organization. The quality of the forecasts management can make is strongly related to the information that can be extracted and used from past data. Time-series is recording the data periodically.

Objectives

After reading this unit, you will be able to:

- Forecast using time series models.
- Measure the forecast error to assess the accuracy of the models.

5.2.1 Components of Time Series

There are four components of time series analysis:

1. Secular trend
2. Cyclical fluctuation
3. Seasonal variation
4. Irregular variation

Secular Trend the value of the variable tends to increase or decrease over a long period of time. The steady increase in the cost of living recorded by the Consumer Price Index is an example of secular trend. From year to individual year, the cost of living varies a great deal, but if we examine a long-term period, we see that the trend is toward a steady increase.

Cyclical fluctuation. The most common example of cyclical fluctuation is the business cycle. Over time, there are years when the business cycle hits a peak above the trend line. At other times, business activity is likely to slump, hitting a low point below the trend line. At other times, business activity is likely to slump, hitting a low point below the trend line. The time between hitting peaks or falling to low points is at least 1 year, and it can be many as 15 or 20 years.

Seasonal variation. As we might expect from the name, seasonal variation involves patterns of change within a year that tend to be repeated from year to year. For example, a physician can expect a substantial increase in the number of flu cases every winter and of poison ivy every summer. Because these are regular patterns, they are useful in forecasting the future.

Irregular variation is the fourth type of change in time-series analysis. In many situations, that value of a variable may be completely unpredictable, changing in a random manner.

Thus far, we have referred to a time series as exhibiting one or another of these four types of variation. In most instances, however, a time series will contain several of these components. Thus, we can describe the overall variation in a single time series in terms of these four different kinds of variation .In the following sections, we will examine the four components and the ways in which we measure each.

5.2.2 Trend analysis

Of the four components of a time series, secular trend represents the long-term direction of the series. One way to describe the trend component is to fit a line visually to a set of points on a graph. Any given graph, however, is subject to slight different interpretations by different individuals. The following are methods of measuring trend.

(1) Graphic or free hand method

(2) Semi average method

(3) Moving average method

(4) Method of least square method

(1) Graphical method

Given data are plotted on a graph paper and trend line is fitted to the data just by inspection

(2) Semi-average method

The given data are divided into two equal parts. If there is odd number of data the middle year is left and two equal parts are formed. An average of each part is secured and the two points thus obtained are centered corresponding to the middle period and shown on the graph. A straight line is drawn through these two points which describes the trend. By projecting the line it is possible to forecast the future values.

(3) Moving average method

In this method, first the period of moving average is selected which may be 3,4,5,6...etc.the moving totals of the given data are obtained. Each of these figures is divided by the period of moving average which gives the trend value.

(4) Method of least square

With help of this model we can fit a straight line or a second degree parabolic equation

the equation of the straight line or linear equation is

there fore the normal equation is

$$\sum y = na + \sum x$$
$$\sum xy = a\sum x + b\sum x^2$$

If the deviation is taken from the middle period so that $\sum x = 0$

then the values 'a' and 'b' can be obtained from the following

Example:

A small company that manufactures portable TV in Carolina. Since they started the Company, the number of table they have sold is represented by this time series:

year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
sales	42	50	61	75	92	111	120	127	140	138

(a) Find the linear equation that describes the trend in the number of TV sold by company.

(b) Estimate their sales of TV in 1998.

Solution:

Year	x	y	xY	x^2
1987	-9	42	-378	81
1988	-7	50	-350	49
1989	-5	61	-305	25
1990	-3	75	-225	9
1991	-1	92	-92	1
1992	1	111	111	1
1993	3	120	360	9
1994	5	127	635	25
1995	7	140	980	49
1996	9	138	1242	81
Total	0	956	1,978	330

$$a = \overline{Y} = \frac{956}{10} = 95.6$$

$$b = \frac{\sum xY}{\sum x^2} = \frac{1,978}{330} = 5.9939$$

$$y = 95.6 + 5.9939x \ (where 1991.5 = 0 \, and \ x \ units = 0.5 year)$$

Summary

This unit has provided a conceptual framework on various forecasting techniques with their strengths and limitations. Specifically, this unit focused on:

- The need for index number and its calculations.
- Components of time series
- Trend projection using least square line.

Exercises:

1) The number of faculty-owned personal computers at the University of Ohio increased dramatically between 1990 and 1995:

Year	1990	1991	1992	1993	1994	1995
Number of PCs	50	110	350	1,020	1,950	3,710

(a). Develop a linear estimating equation that best describes these data.

(b) Estimate the number of PCs that will be in use at the university in 1999,

2) Below are the figures of production (thousands) of a sugar factory:

Year	1999	2001	2202	2003	2005
Production (thousands)	700	600	400	900	900

Required:

(a) Fit Straight Line Trend by the method of the least squares and tabulate the values.

(b) Estimate the likely production for the year 2007.

NOTES

NOTES

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....

....