Chapter

2

INTERPOLATION AND APPROXIMATION

♠♠♠

2.1 INTRODUCTION

The process of construction of y(x) to fit a table of data points is called curve fitting. A table of data may belong to one of the following two categories.

1. Table of values of well-defined functions

Examples of such tables are logarithmic tables, trigonometric tables, interest tables, steam tables etc.

2. Data tabulated from measurements made during an experiment

In such experiments, values of the dependent variable are recorded at various values of the independent variable. There are numerous examples of such experiments-the relationship between stress and strain on a metal strip, relationship between voltage applied and speed on a metal strip, relationship between voltage applied and speed of a fan, relationship between time and temperature raise in heating a given volume of water, relationship between drag force and velocity of a falling body etc can be tabulated by suitable experiments.

In category-1, the table values are accurate because they are obtained from well-behaved functions. This is not the case in category 2 where the relationship between the variable is not well defined. Accordingly, we have two approaches for fitting a curve to a given set of data points.

In the first case, the function is constructed such that it passes through all the data points. This method of constructing a function and estimating values at non-tabular points is called interpolation. The functions are known as interpolation polynomials.

In the second case, the values are not accurate and therefore, it will be meaningless to try to pass the curve through every point. The best strategy would be to construct a single curve that would represent the general trend of the data, without necessarily passing through the individual points. Such functions are called approximating functions. One popular approach for finding an approximate function to fit a given set of experimental data is called least squares regression. The approximate functions are known as least-squares polynomials.

The various methods of interpolation are;

a) Lagrange interpolation
b) Newton's interpolation
c) Newton-Gregory forward interpolation
d) Spline interpolation

2.2 INTERPOLATION WITH UNEQUAL INTERVALS

Interpolation formula for unequally spaced values of x can be obtained from any method given below.

i) Lagrange's interpolation formula
ii) Newton's general interpolation formula with divided differences

2.2.1 Lagrange's Interpolation Formula

Let, $y = f(x)$ takes the value $y_0, y_1, \ldots\ldots, y_n$ corresponding to $x = x_0, x_1, \ldots\ldots, x_n$, then,

$$f(x) = \frac{(x-x_1)(x-x_2)\ldots\ldots(x-x_n)}{(x_0-x_1)(x_0-x_2)\ldots\ldots(x_0-x_n)}y_0 + \frac{(x-x_0)(x-x_2)\ldots\ldots(x-x_n)}{(x_1-x_0)(x_1-x_2)\ldots\ldots(x_1-x_n)}y_1 + \ldots\ldots + \frac{(x-x_0)(x-x_1)\ldots\ldots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\ldots\ldots(x_n-x_{n-1})}y_n \quad \ldots\ldots (1)$$

This is known as Lagrange's interpolation formula for unequal intervals.

Proof:

Let, $y = f(x)$ be a function which takes the values $(x_0, y_0), (x_1, y_1), \ldots\ldots, (x_n, y_n)$. Since there are $n + 1$ pairs of values of x and y, we can represent f(x) by a polynomial in x of degree n. Let this polynomial be of the form,

$$y = f(x) = a_0(x-x_1)(x-x_2)\ldots\ldots(x-x_n) + a_1(x-x_0)(x-x_2)\ldots\ldots(x-x_n) + a_2(x-x_0)(x-x_1)(x-x_3)\ldots\ldots(x-x_n) + \ldots\ldots + a_n(x-x_0)(x-x_1)\ldots\ldots(x-x_{n-1}) \quad \ldots\ldots (2)$$

Replacing $x = x_0, y = y_0$ in (2), we have,

$$y_0 = a_0(x_0-x_1)(x-x_2)\ldots\ldots(x-x_n)$$

$$a_0 = \frac{y_0}{(x - x_1)\,(x - x_2)\,........\,(x - x_n)}$$

Similarly putting $x = x_1, y = y_1$ in (2), we have,

$$a_1 = \frac{y_1}{(x_1 - x_0)\,(x_1 - x_2)\,........\,(x_1 - x_n)}$$

Proceeding the same way, we find $a_2, a_3, a_4 \,........\, a_n$.

Replacing the values of $a_0, a_1, ..., a_n$ in (2), we get (1).

> **NOTE:**
> Lagrange's interpolation formula (1) for n points is a polynomial of degree (n − 1) which is known as the Lagrangian polynomial and is very simple to implement on a computer. This formula can also be used to split the given function into partial fractions.

On dividing both sides of (1) by $(x - x_0)\,(x - x_1)\,.........\,(x - x_n)$, we get,

$$\frac{f(x)}{(x - x_0)(x - x_1)\,.........\,(x_0 - x_n)} = \frac{y_0}{(x_0 - x_1)(x_0 - x_2)\,.........\,(x_0 - x_n)} \cdot \frac{1}{(x - x_0)}$$

$$+ \frac{y_1}{(x_1 - x_0)(x_1 - x_2)\,.........\,(x_1 - x_n)} \cdot \frac{1}{(x - x_1)}$$

$$+ \,.........\, + \frac{y_n}{(x_n - x_0)(x_0 - x_1)\,.........\,(x_n - x_{n-1})} \cdot \frac{1}{(x - x_n)}$$

Example 2.1

Given the values:

x	5	7	11	13	17
f(x)	150	392	1452	2366	5202

Evaluate f(9) using Lagrange's formula.

Solution:

Here;

$x_0 = 5, \quad x_1 = 7, \quad x_2 = 11, \quad x_3 = 13, \quad x_4 = 17$

and, $y_0 = 150, \quad y_1 = 392, \quad y_2 = 1452, \quad y_3 = 2366, \quad y_4 = 5202$

$$f(x) = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} \times y_0$$

$$+ \frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} \times y_1$$

$$+ \frac{(x - x_0)(x - x_1)(x - x_3)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)} \times y_2$$

$$+ \frac{(x - x_0)(x - x_1)(x - x_2)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)} \times y_3$$

$$+ \frac{(x - x_0)(x - x_1)(x - x_2)(x - x_3)}{(x_4 - x_0)(x_4 - x_1)(x_4 - x_1)(x_4 - x_3)} \times y_4$$

Putting x = 9 and replacing the above values in Lagrange's formula, we get,

$$f(9) = \frac{(9 - 7)(9 - 11)(9 - 13)(9 - 17)}{(5 - 7)(5 - 11)(5 - 13)(5 - 17)} \times 150$$

$$+\frac{(9-5)(9-11)(9-13)(9-17)}{(7-5)(7-11)(7-13)(7-17)}\times 392$$

$$+\frac{(9-5)(9-7)(9-13)(9-17)}{(11-5)(11-7)(11-13)(11-17)}\times 1452$$

$$+\frac{(9-5)(9-7)(9-11)(9-17)}{(13-5)(13-7)(13-11)(13-17)}\times 2366$$

$$+\frac{(9-5)(9-7)(9-11)(9-13)}{(17-5)(17-7)(17-11)(17-13)}\times 5202$$

$$=\left(-\frac{50}{3}\right)+\frac{3136}{15}+\frac{3872}{3}+\frac{2366}{3}+\frac{578}{5}$$

$\therefore$ $f(9) = 810$

Example 2.2

Find the polynomial f(x) by using Lagrange's formula and hence find f(3) for

x	0	1	2	5
f(x)	2	3	12	147

Solution:

Here;

$x_0 = 0,\quad x_1 = 1,\quad x_2 = 2,\quad x_3 = 5$

and, $y_0 = 2,\quad y_1 = 3,\quad y_2 = 12,\quad y_3 = 147$

Lagrange's formula is,

$$y=\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}\times y_0$$

$$+\frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}\times y_1$$

$$+\frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}\times y_2$$

$$+\frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}\times y_3$$

$$=\frac{(x-1)(x-2)(x-5)}{(0-1)(0-2)(0-5)}\times 2$$

$$+\frac{(x-0)(x-2)(x-5)}{(1-0)(1-2)(1-5)}\times 3$$

$$+\frac{(x-0)(x-1)(x-5)}{(2-0)(2-1)(2-5)}\times 12$$

$$+\frac{(x-0)(x-1)(x-2)}{(5-0)(5-1)(5-2)}\times 147$$

Hence, $f(x) = x^3 + x^2 - x + 2$

$\therefore$ $f(3) = 3^3 + 3^2 - 3 + 2 = 27 + 9 - 3 + 2 = 35$

Example 2.3

Find the missing term in the following table using interpolation.

x	0	1	2	3	4
f(x)	1	3	9	–	81

Solution:

Since the given data is unevenly spaced, we use Lagrange's interpolation formula.

$$y = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} \times y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} \times y_1$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} \times y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} \times y_3$$

Given that;

$x_0 = 0,\quad x_1 = 1,\quad x_2 = 2,\quad x_3 = 4$

$y_0 = 1,\quad y_1 = 3,\quad y_2 = 9,\quad y_3 = 81$

$$\therefore \quad y = \frac{(x-1)(x-2)(x-4)}{(0-1)(0-2)(0-4)} \times 1 + \frac{(x-0)(x-2)(x-4)}{(1-0)(1-2)(1-4)} \times 3.$$

$$+ \frac{(x-0)(x-1)(x-4)}{(2-0)(2-1)(2-4)} \times 9 + \frac{(x-0)(x-1)(x-2)}{(4-0)(4-1)(4-2)} \times 81$$

When $x = 3$, then,

$$\therefore \quad y = \frac{(3-1)(3-2)(3-4)}{-8} + \frac{3(3-2)(3-4)}{1} + \frac{3(3-1)(3-4)}{-4} \times 9$$

$$+ \frac{3(3-1)(3-2)}{24} \times 81$$

$$= \frac{1}{4} - 3 + \frac{27}{2} + \frac{81}{24} = 31$$

Hence the missing term for $x = 3$ is $y = 31$.

Example 2.4

Using Lagrange's formula, express the function $\frac{3x^2 + x + 1}{(x-1)(x-2)(x-3)}$ is a sum of partial functions.

Solution:

Let us evaluate: $y = 3x^2 + x + 1$ for $x = 1$, $x = 2$ and $x = 3$.

These values are:

x	$x_0 = 1$	$x_1 = 2$	$x_2 = 3$
y	$y_0 = 5$	$y_1 = 15$	$y_2 = 31$

Lagrange's formula is,

$$y = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \times y_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} \times y_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \times y_2$$

Replacing the above values, we get,

$$y = \frac{(x-2)(x-3)}{(1-2)(1-3)} \times 5 + \frac{(x-1)(x-3)}{(2-1)(2-3)} \times 15 + \frac{(x-1)(x-2)}{(3-1)(3-2)} \times 31$$

$$= 2.5(x-2)(x-3) - 15(x-1)(x-3) + 15.5(x-1)$$

Thus, $\frac{3x^2 + x + 1}{(x-1)(x-2)(x-3)}$

$$= \frac{2.5(x-2)(x-3) - 15(x-1)(x-3) + 15.5(x-1)(x-2)}{(x-1)(x-2)(x-3)}$$

$$= \frac{25}{(x-1)} - \frac{15}{(x-2)} + \frac{15.5}{(x-3)}$$

2.2.2 Divided Differences

The Lagrange's formula has the drawback that if another interpolation value were inserted, then the interpolation coefficients are required to be recalculated. This labor of recomposing the interpolation formula which employs what are called 'divided differences". Before deriving this formula, we shall first define these differences.

If (x_0, y_0), (x_1, y_1), (x_2, y_2) .... be given points, then the first divided difference for the arguments x_0, x_1 is defined by the relation

$$[x_0, x_1] \text{ or } \underset{x_1}{\Delta} y_0 = \frac{y_1 - y_0}{x_1 - x_0}$$

Similarly,

$$[x_1, x_2] \text{ or } \underset{x_2}{\Delta} y_0 = \frac{y_2 - y_1}{x_2 - x_1}$$

and, $$[x_2, x_3] \text{ or } \underset{x_3}{\Delta} y_0 = \frac{y_3 - y_2}{x_3 - x_2}$$

The second divided difference for x_0, x_1, x_2 is defined as

$$[x_0, x_1, x_2] \text{ or } \underset{x_1, x_2}{\Delta^2} y_0 = \frac{[x_1, x_2] - [x_0, x_1]}{x_2 - x_0}$$

The third divided difference for x_0, x_1, x_2, x_3 is defined as

$$[x_0, x_1, x_2, x_3] \text{ or } \underset{x_1, x_2, x_3}{\Delta^3} y_0 = \frac{[x_1, x_2, x_3] - [x_0, x_1, x_2]}{x_2 - x_0}$$

Properties of divided differences

1. The divided differences are symmetrical in their arguments *i.e.* independent of the order of the arguments.

$$[x_0, x_1] = \frac{y_0}{x_0 - x_1} + \frac{y_1}{x_1 - x_0} = [x_1, x_0], [x_0, x_1, x_2]$$

$$= \frac{y_0}{(x_0 - x_1)(x_0 - x_2)} + \frac{y_1}{(x_1 - x_0)(x_1 - x_2)} + \frac{y_2}{(x_2 - x_0)(x_2 - x_1)}$$

$$= [x_1, x_2, x_0] \text{ or } [x_2, x_0, x_1] \text{ and so on.}$$

2. The nth divided differences of a polynomial of the nth degree are constant.

Let the arguments be equally spaced so that

$$x_1 - x_0 = x_2 - x_1 = \ldots\ldots\ldots = x_n - x_{n-1} = h$$

Then, $[x_0, x_1] = \dfrac{y_1 - y_0}{x_1 - x_0} + \dfrac{\Delta y_0}{h}$

$$[x_0, x_1, x_2] = \frac{[x_1, x_2] - [x_0, x_1]}{x_2 - x_0}$$

$$= \frac{1}{2h}\left[\frac{\Delta y_1}{h} - \frac{\Delta y_0}{h}\right] = \frac{1}{2!h^2}\Delta^2 y_0$$

and in general,

$$[x_0, x_1, x_2, \ldots\ldots\ldots, x_n] = \frac{1}{n!h^n}\Delta^n y_0$$

If the tabulated function is a nth degree polynomial, then $\Delta^n y_0$ will be constant. Hence, the nth divided differences will also be constant.

I. Newton's Divided Difference Formula

Let $y_0, y_1, y_2, \ldots\ldots\ldots, y_n$ be the values of $y = f(x)$ corresponding to the arguments $x_0, x_1, x_2, \ldots\ldots\ldots, x_n$. Then from the definition of divided differences, we have,

$$[x, x_0] = \frac{y - y_0}{x - x_0}$$

So that,

$$y = y_0 + (x - x_0)\,[x, x_0] \qquad \ldots\ldots (1)$$

Again,

$$[x, x_0, x_1] = \frac{[x, x_0] - [x_0, x_1]}{x - x_1}$$

which gives,

$$[x, x_0] = [x_0, x_1] + (x - x_1)\,[x, x_0, x_1]$$

Hence the equation (1) becomes,

$$y = y_0 + (x - x_0)\,[x_0, x_1] + (x - x_0)(x - x_1)\,[x, x_0, x_1] \qquad \ldots\ldots (2)$$

Also, $[x, x_0, x_1, x_2] = \dfrac{[x, x_0, x_1] - [x, x_0, x_2]}{x - x_2}$

which gives $[x, x_0, x_1] = [x_0, x_1, x_2] + (x - x_2)\,[x, x_0, x_1, x_2]$

Replacing this value in equation (2), we get,

$$y = y_0 + (x - x_0)\,[x_0, x_1] + (x - x_0)(x - x_1)\,[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)\,[x, x_0, x_1, x_2]$$

Preceding in this manner, we get,

$$y = y_0 + (x - x_0)\,[x_0, x_1] + (x - x_0)(x - x_1)\,[x_0, x_1, x_2] + (x - x_0)(x - x_1)\ldots\ldots\ldots(x - x_n)\,[x, x_0, x_1, \ldots\ldots\ldots, x_n] + (x - x_0)(x - x_1)(x - x_2)\,[x, x_0, x_1, x_2] + \ldots\ldots\ldots \qquad \ldots\ldots (3)$$

Which is called Newton's general interpolation formula with divided differences.

II. Relation between Divided and Forward Differences

If (x_0, y_0), (x_1, y_1), (x_2, y_2), ... be the given points, then,

$$[x_0, x_1] = \frac{y_1 - y_0}{x_1 - x_0}$$

Also, $\Delta y_0 = y_1 - y_0$

If $x_0, x_1, x_2, ...$ are equispaced,

then, $x_1 - x_0 = h$, so that,

$$[x_0, x_1] = \frac{\Delta y_0}{h}$$

Similarly,

$$[x_1, x_2] = \frac{\Delta y_1}{h}$$

Now,

$$[x_0, x_1, x_2] = \frac{[x_1, x_2] - [x_0, x_1]}{x_2 - x_1} = \frac{\frac{\Delta y_1}{h} - \frac{\Delta y_0}{h}}{2h} \qquad [\because x_2 - x_0 = 2h]$$

$$= \frac{\Delta y_1 - \Delta y_0}{2h^2}$$

Thus, $[x_0, x_1, x_2] = \dfrac{\Delta^2 y_0}{2!h^2}$

Similarly,

$$[x_0, x_1, x_2] = \frac{\Delta^2 y_1}{2!h^2}$$

$$\therefore \quad [x_0, x_1, x_2, x_3] = \frac{\frac{\Delta^2 y_1}{2h^2} - \frac{\Delta^2 y_0}{2h^2}}{x_3 - x_0} = \frac{\Delta^2 y_1 - \Delta^2 y_0}{2h^2(3)} \qquad [\because x_3 - x_0 = 3h]$$

Thus, $[x_0, x_1, x_2, x_3] = \dfrac{\Delta^3 y_0}{3!h^3}$

In general,

$$[x_0, x_1, x_2, ..., x_n] = \frac{\Delta^n y_0}{n!h^n}$$

This is the relation between divided and forward differences.

Example 2.5

Given the values:

x	5	7	11	13	17
f(x)	150	392	1452	2366	5202

Evaluate f(9), using Newton's divided difference formula.

Solution:

Creating difference table from Newton's dividend difference formula as

x	$f(x_n)$	$f(x_n, x_{n+1})$	$f(x_n, x_{n+1}, x_{n+2})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4})$
5	**150**				
		$\frac{392-150}{7-5}$ = **121**			
7	392		$\frac{265-121}{11-5}$ = **24**		
		$\frac{1452-392}{11-7}$ = 265		$\frac{32-24}{13-5}$ = **1**	
11	1452		$\frac{457-265}{13-7}$ = 32		$\frac{1-1}{17-5}$ = **0**
		$\frac{2366-1452}{13-11}$ = 457		$\frac{42-32}{17-7}$ = 1	
13	2366		$\frac{709-457}{17-11}$ = 42		
		$\frac{5202-2366}{17-13}$ = 709			
17	5202				

Here, we have,

$$[x_0, x_1] = 121$$

$$[x_0, x_1, x_2] = 24$$

$$[x_0, x_1, x_2, x_3] = 1$$

$$[x_0, x_1, x_2, x_3, x_4] = 0$$

Then using Newton's Gregory divided difference formula

$$y = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3] + (x - x_0)(x - x_1)(x - x_2)(x - x_3)[x_0, x_1, x_2, x_3, x_4]$$

Then at x = 9

$$y = 150 + (9 - 5)(121) + (9 - 5)(9 - 7)(24) + (9 - 5)(9 - 7)(9 - 11)(1) + 0$$

$$= 150 + 484 + 192 - 16$$

$$\therefore \quad y = 810$$

Hence, the value of f(9) is 810.

Example 2.6

Using Newton's divided differences formula, evaluate f(8) and f(15) given:

Given the values:

x	4	5	7	10	11	13
y = f(x)	48	100	294	900	1210	2028

Solution:

Creating divided difference table

x	$f(x_n)$	$f(x_n, x_{n+1})$	$f(x_n, x_{n+1}, x_{n+2})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4}, x_{n+5})$
4	48					
		$\frac{100-48}{5-4} = 52$				
5	100		$\frac{97-52}{7-4} = 15$			
		$\frac{294-100}{7-5} = 97$		$\frac{21-15}{10-4} = 1$		
7	294		$\frac{220-97}{10-5} = 21$		$\frac{1-1}{11-4} = 0$	
		$\frac{900-294}{10-7} = 202$		$\frac{27-21}{11-5} = 1$		$\frac{0-0}{13-4} = 0$
10	900		$\frac{310-202}{11-7} = 27$		$\frac{1-1}{13-5} = 0$	
		$\frac{1210-900}{11-10} = 310$		$\frac{33-27}{13-7} = 1$		
11	1210		$\frac{409-310}{13-10} = 33$			
		$\frac{2028-1210}{13-11} = 409$				
13	2028					

Here, we have,

$[x_0, x_1] = 52$

$[x_0, x_1, x_2] = 15$

$[x_0, x_1, x_2, x_3] = 1$

$[x_0, x_1, x_2, x_3, x_4] = [x_0, x_1, x_2, x_3, x_4, x_5] = 0$

Using Newton's divided difference formula,

$$f(8) = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3] + 0 + 0$$

$$= 48 + (8 - 4)(52) + (8 - 4)(8 - 5)(15) + (8 - 4)(8 - 5)(8 - 7)(1) + 0 + 0$$

$\therefore \quad y = 448$

Similarly for x = 15,

$$f(15) = 48 + (15 - 4)(52) + (15 - 4)(15 - 5)(15) + (15 - 4)(15 - 5)(15 - 7)(1) + 0 + 0$$

$$= 48 + 572 + 1650 + 880$$

$\therefore \quad f(15) = 3150$

Example 2.7

Using Newton's divided difference formula, find the missing value form the table:

x	1	2	4	5	6
y	14	15	5	...	9

Solution:

Creating dividend difference table

x	$y = f(x_n)$	$f(x_n, x_{n+1})$	$f(x_n, x_{n+1}, x_{n+2})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3})$
1	**14**			
		$\frac{15-14}{2-1} = \mathbf{1}$		
2	15		$\frac{-5-1}{4-1} = -\mathbf{2}$	
		$\frac{5-15}{4-2} = -5$		$\frac{\frac{7}{4}+2}{6-1} = \mathbf{\frac{3}{4}}$
4	5		$\frac{2+5}{6-2} = \frac{7}{4}$	
		$\frac{9-5}{6-4} = 2$		
6	9			

Here, we have,

$[x_0, x_1] = 1$

$[x_0, x_1, x_2] = -2$

$[x_0, x_1, x_2, x_3] = \frac{3}{4}$

Now, using Newton's divided difference formula,

$$y = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3]$$

Then at $x = 5$

$$y = 14 + (5-1)(1) + (5-1)(5-2)(-2) + (5-1)(5-2)(5-4)\left(\frac{3}{4}\right)$$

$$= 14 + 4 - 24 + 9$$

$\therefore \quad y = 3$

Hence, the missing value at $x = 5$ is 3.

2.3 NEWTON'S FORWARD INTERPOLATION FORMULA

Interpolation is the technique of estimating the value of a function for any intermediate value of the independent variable while the process of computing the value of the function outside the given range is called extrapolation.

Let the function $y = f(x)$ takes the values $y_0, y_1, \ldots\ldots, y_n$ correspond to $x_0, x_1, \ldots\ldots, x_n$ of x. Let these values of x be equispaced such that $x_i = x_0 + ih$ ($i = 0, 1, \ldots\ldots$).

Assuming $y(x)$ to be a polynomial of the n^{th} degree in x such that $y(x_0) = y_0$, $y(x_1) = y_1, \ldots\ldots, y(x_n) = y_n$. We can write,

$$y(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2) + \ldots\ldots + a_n(x - x_0)(x - x_1)\ldots\ldots(x - x_{n-1}) \quad \ldots\ldots (1)$$

Putting, $x = x_0, x_1, \ldots\ldots, x_n$ successively in (1), we get,

$$y_0 = a_0,\ y_1 = a_0 + a_1(x_1 - x_0),\ y_2 = a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1)$$

and so on.

From these, we find that $a_0 = y_0$, $\Delta y_0 = y_1 - y_0 = a_1(x_1 - x_0) = a_1 h$

$$\therefore \quad a_1 = \frac{1}{h}\Delta y_0$$

Also, $\Delta y_1 = y_2 - y_1 = a_1(x_2 - x_1) + a_2(x_2 - x_0)\cdot(x_2 - x_1)$

$$= a_1 h + a_2 hh = \Delta y_0 + 2h^2 a_2$$

$$\therefore \quad a_2 = \frac{1}{2h^2}(\Delta y_1 - \Delta y_0) = \frac{1}{2!h^2}\Delta^2 y_0$$

Similarly $a_3 = \frac{1}{3!h^3}\Delta^3 y_0$ and so on.

Replacing values in (1), we get,

$$y(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!h^2}(x - x_0)(x - x_1) + \frac{\Delta^3 y_0}{3!h^3}(x - x_0)(x - x_1)(x - x_2) + \ldots\ldots \quad \ldots\ldots (2)$$

Now, if it is required to evaluate y for $x = x_0 + ph$, then,

$$(x - x_0) = ph,\ x - x_1 = x - x_0 - (x - x_0) = ph - h = (p-1)h$$

$$(x - x_0) = x - x_0 - (x - x_0) = (p-1)h - h = (p-2)h \text{ etc}$$

Hence, $y(x) = y(x_0 + ph) = y_p$, then, (2) becomes,

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0$$

$$+ \text{.........} + \frac{p(p-1)\text{.........}(p-\overline{n-1})}{n!}\Delta^n y_0 \quad \text{..... (3)}$$

It is called Newton's forward interpolation formula as (3) contains y_0 and the forward differences of y_0.

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0$$

$$+ \text{.........} + \frac{p(p-1)\text{.........}(p-\overline{n-1})}{n!}\Delta^n y_0$$

NOTE:

1. This formula is used for interpolating the values of a set of tabulated values and extrapolating values of y a little backward. (*i.e.*, to the left) of y_0.
2. The first two terms of this formula give the linear interpolation which the first three terms give a parabolic interpolation and so on.

2.4 NEWTON'S BACKWARD INTERPOLATION FORMULA

Le the function $y = f(x)$ take the values $y_0, y_1, y_2,$ .... corresponding to the values $x_0, x_0 + h, x_0 + 2h,$ .... of x. Suppose it is required to evaluate f(x) for $x = x^n + ph$, where p is any real number. Then we have,

$$y_p = f(x_n + ph) = Ep\, f(x_n) = (1-\nabla)^{-p} y_n \qquad [\because E^{-1} = 1 - \nabla]$$

$$= \left[1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_0 + \text{.........}\right] y_n$$

[using Binomial theorem]

i.e., $$y_p = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \text{.........} \quad \text{..... (1)}$$

It is called the Newton's backward interpolation formula as (1) contains y_n and backward differences of y_n. This formula is used for interpolating the values of y near the end of a set of tabulted values and also for extrapolating values of y a little ahead (to the right) of y_n.

Example 2.8

Using Newton's backward difference formula, construct an interpolation polynomial of degree 3 for the data:

$f(-0.75) = -0.0718125$, $f(-0.5) = -0.02475$, $f(-0.25) = 0.3349375$,

$f(0) = 1.10100$. Hence find $f\left(-\frac{1}{3}\right)$.

Solution:

Creating the difference table from the given datas;

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$
-0.75	-0.0718125			
		0.0470625		
-0.5	-0.02475		0.312625	
		0.3596875		0.09375
-0.25	0.3349375		0.406375	
		0.7660625		
0	1.10100			

Now, using Newton's backward difference formula

$$y(x) = y_3 + p\nabla y_3 + \frac{p(p+1)}{2!}\nabla^2 y_3 + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_3$$

Taking $x_3 = 0$, $p = \frac{x-0}{h} = \frac{x}{0.25} = 4x$

$$y(x) = 1.10100 + 4x\,(0.7660625) + \frac{4x\,(4x+1)}{2}(0.406375)$$

$$+ \frac{4x\,(4x+1)\,(4x+2)}{6}(0.09375)$$

$$= 1.101 + 3.06425x + 3.251x^2 + 0.81275x + x^3 + 0.75x^2 + 0.125x$$

$\therefore \quad y = x^3 + 4.001x^2 + 4.002x + 1.101$

is the required interpolating polynomial.

At $x = \left(-\frac{1}{3}\right)$,

$$y\left(-\frac{1}{3}\right) = \left(-\frac{1}{3}\right)^3 + 4.001\left(-\frac{1}{3}\right)^2 + 4.002\left(-\frac{1}{3}\right) + 1.101$$

$$= 0.174518$$

Example 2.9

Using Newton's forward formula, find the value of f(1.6) if

x	1	1.4	1.8	2.2
f(x)	3.49	4.82	5.96	6.5

Solution:

Creating difference table

x	y = f(x)	Δy	$\Delta^2 y$	$\Delta^3 y$
1	3.49			
		1.33		
1.4	4.82		-0.19	
		1.14		-0.41
1.8	5.96		-0.6	
		0.54		
2.2	6.5			

We have,

$x = 1.6, \quad x_0 = 1, \quad h = 1.4 - 1 = 0.4$

$x = x_0 + ph$

or, $p = \frac{1.6 - 1}{0.4} = 1.5$

Now, using Newton's forward formula

$$y_{1.6} = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0$$

$$= 3.49 + 1.5(1.33) + \frac{1.5(1.5-1)}{2}(-0.19)$$

$$+ \frac{1.5(1.5-1)(1.5-2)}{6}(-0.41)$$

$$= 3.49 + 1.995 - 0.07125 + 0.025625$$

$\therefore$ $y_{1.6} = 5.439375$

Hence the required value of f(1.6) is 5.439375.

Example 2.10

Given, sin 45° = 0.7071, sin 50° = 0.7660, sin 55° = 0.8192
sin 60° = 0.8660, find sin 52° using Newton's forward formula.

Solution:

Creating difference table from the given data,

$x = \theta$	$y = \sin\theta$	Δy	$\Delta^2 y$	$\Delta^3 y$
45°	**0.7071**			
		0.0589		
50°	0.7660		**-0.0057**	
		0.0532		**-0.0007**
55°	0.8192		-0.0064	
		0.0468		
60°	0.8660			

We have,

$x = 52, \quad x_0 = 45, \quad h = 50 - 45 = 5$

$x = x_0 + ph$

or, $p = \frac{52 - 45}{5} = \frac{7}{5}$

Now, using Newton's forward formula

$$y_{52} = y_{45} + p\Delta y_{45} + \frac{p(p-1)}{2!}\Delta^2 y_{45} + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_{45}$$

$$= 0.7071 + \frac{7}{5}(0.0589) + \frac{\frac{7}{5}\left(\frac{7}{5} - 1\right)}{2}(-0.0057)$$

$$+ \frac{\frac{7}{5}\left(\frac{7}{5} - 1\right)\left(\frac{7}{5} - 2\right)(-0.0007)}{6}$$

$$= 0.7071 + 0.08246 - 0.001596 + 0.0000392$$

$\therefore$ $y_{45} = 0.7880032$

Hence the required value of sin 52° is 0.7880032.

Example 2.11

Apply Newton's backward difference formula to the data below, to obtain a polynomial of degree 4 in x:

x	1	2	3	4	5
y	1	–1	1	–1	1

Solution:

Creating difference table

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
1	1				
		-2			
			4		
2	-1			-8	
		2			16
			-4		
3	1			8	
		-2			
			4		
4	-1				
		2			
5	1				

We have,

$$x_4 = 5, \quad h = 5 - 4 = 1$$

$$x = x_4 + ph$$

or, $$p = \frac{x - x_4}{h} = \frac{x-5}{1} = x - 5$$

Now, using Newton's backward difference formula,

$$y(x) = y_4 + p\nabla y_4 + \frac{p(p+1)}{2!}\nabla^2 y_4 + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_4 + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_4$$

$$= 1 + (x-5)(2) + \frac{4}{2}(x-5)(x-5+1) + \frac{8}{6}(x-5)(x-5+1)(x-5+2) + \frac{16}{24}(x-5)(x-5+1)(x-5+2)(x-5+3)$$

$$= 1 + 2x - 10 + (2x-10)(x-4) + \frac{4}{3}(x-5)(x-4)(x-3) + \frac{2}{3}(x-5)(x-4)(x-3)(x-2)$$

$$= 1 + 2x - 10 + 2x^2 - 18x + 40 + 1.33x^3 - 16x^2 + 62.667 - 80 + 0.667x^4 - 9.333x^3 + 47.33x^2 - 102.667x + 80$$

$\therefore$ $$y(x) = 0.67x^4 - 8.003x^3 + 33.33x^2 - 56x + 31$$

is the required polynomial.

2.5 LINEAR INTERPOLATION

The simplest form of interpolation is to approximate two data points by a straight line. Suppose we are given two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. These two points can be connected linearly as shown in figure 2.1.

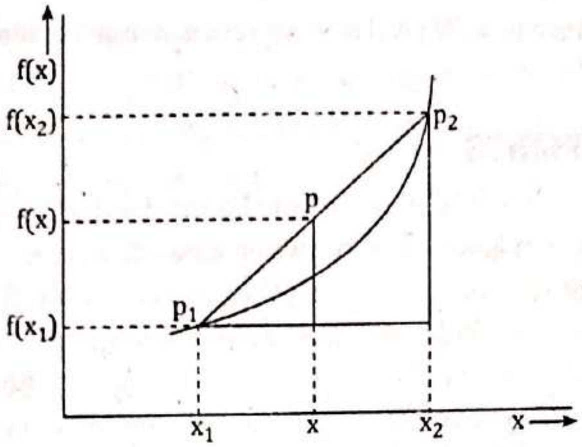

Figure 2.1: Graphical representation of linear interpolation

Using the concept of similar triangles, we can show that,

$$\frac{f(x) - f(x_1)}{x - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

On solving, we get,

$$f(x) = f(x_1) + (x - x_1)\frac{f(x_2) - f(x_1)}{x_2 - x_1} \qquad \text{..... (1)}$$

Equation (1) is called as linear interpolation formula. Note that the term, $\frac{f(x_2) - f(x_1)}{x_2 - x_1}$ represents the slope of the line. Further, note the similarity of equation (1) with the Newton form of polynomial of first order.

$$C_1 = x_1$$

$$a_0 = f(x_1)$$

$$a_1 = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

The coefficient a_1 represents the first derivative of the function.

Example 2.12

The table below gives square roots for integers. Determine the square root of 2.5.

x	1	2	3	4	5
f(x)	1	4.4142	1.7321	2	2.2361

Solution:

The given value of 2.5 lies between the points. Hence,

$x_1 = 2$; $f(x_1) = 1.4142$

$x_2 = 3$; $f(x_2) = 1.7321$

Then, $f(2.5) = 1.4142 + (2.5 - 2.0)\dfrac{1.7321 - 1.4142}{3.0 - 2.0}$

$$\left[\because f(x) = f(x_1) + (x_2 + x_1)\frac{f(x_2) - f(x_1)}{x_2 - x_1}\right]$$

$$= 1.4142 + 0.5 \times 0.3179 = 1.5732$$

The correct answer is 1.5811. The difference is due to the use of a linear model to a non-linear one.

2.6 CUBIC SPLINES

The concept of splines originated from the mechanical drafting tool called "spline" used by designers for drawing smooth curves. It is a slender flexible bar made of wood or some other elastic materials. These curves resembles cubic curves and hence the name "cubic spline" has been given to the piecewise cubic interpolating polynomials. Cubic splines are popular because of their ability to interpolate data with smooth curves. It is believed that a cubic polynomial spline always appears smooth to the eyes.

In the interpolation methods so far explained, a single polynomial has been fitted to the tabulated points. If the given set of points belongs to the polynomial, then this method works well, otherwise the results are rough approximations only. If we draw lines through every two closest points, the resulting graph will not be smooth. Similarly, we may draw a quadratic curve through points A_i, A_{i+1} and another quadratic curve through A_{i+1}, A_{i+2} such that the slopes of the two quadratic curves match at A_{i+1}. The resulting curve looks better but is not quite smooth. We can ensure this by drawing a cubic curve through A_i, A_{i+1} and another cubic through A_{i+1}, A_{i+2} such that the slopes and curvatures of the two curves match at A_{i+1}. Such a curve is called a cubic spline. We may use polynomial of higher order but the resulting graph is not better. As such, cubic splines are commonly used. This technique of spline fitting is of recent origin and has important applications.

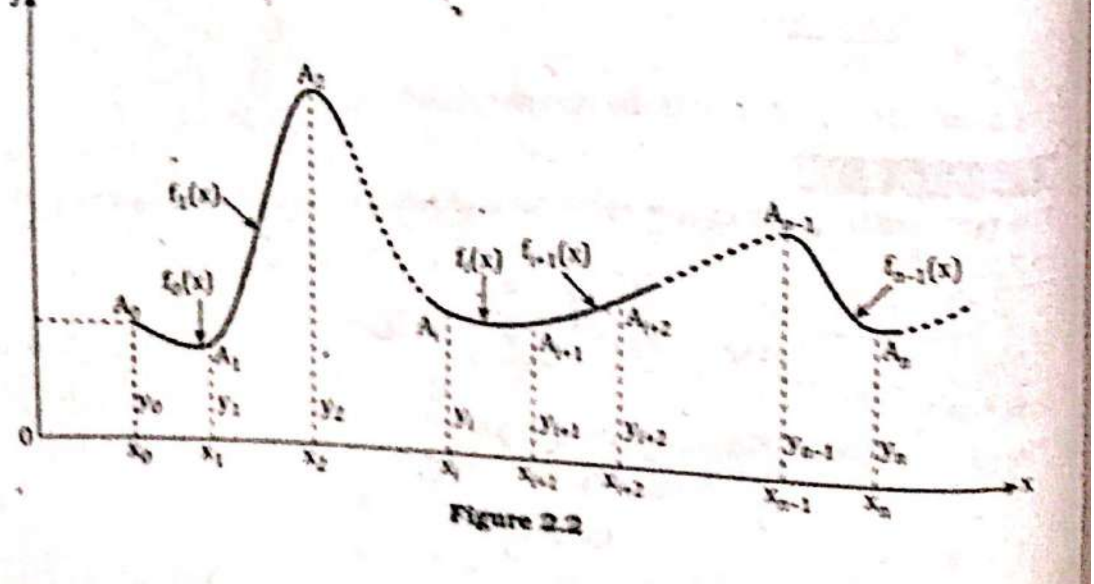

Figure 2.2

Consider the problem of interpolating between the data points (x_0, y_0) (x_1, y_1), ..., (x_n, y_n) by means of spline fitting.

Then the cubic spline $f(x)$ is such that,

i) $f(x)$ is a linear polynomial outside the interval (x_0, x_n).

ii) $f(x)$ is a cubic polynomial in each of the subintervals.

iii) $f'(x)$ and $f''(x)$ are continuous at each point.

Since $f(x)$ is cubic in each of the subintervals $f''(x)$ shall be linear.

$\therefore$ Taking equally spaced values of x so that $x_{i+1} - x_i = h$, we can write,

$$f''(x) = \frac{1}{h}[(x_{i+1} - x)\, f''(x_i) + (x - x_i)\, f''(x_{i+1})]$$

On integrating twice, we get,

$$f(x) = \frac{1}{h}\left[\frac{(x_{i+1} - x)}{3!} f''(x_i) + \frac{(x - x_i)}{3!} f''(x_{i+1})\right] a_i(x_{i+1} - x) + b_i(x - x_i) \quad \text{..... (1)}$$

The constants of integration a_i, b_i are determined by substituting the values of $y = f(x)$ at xi and x_{i+1}. Thus,

$$a_i = \frac{1}{h}\left[y_i - \frac{h^2}{3!} f''(x_i)\right]$$

$$bi = \frac{1}{h}\left[y_{i+1} - \frac{h^2}{3!} f''(x_{i+1})\right]$$

Replacing the values of a_i, b_i and writing $f''(x_i) = M_i$, (1) takes the form,

$$f(x) = \frac{(x_{i+1} - x)^3}{6h} M_i + \frac{(x - x_i)^3}{6h} M_{i+1} + \frac{x_{i+1} - x}{h}\left(y_i - \frac{h^2}{6} M_i\right)$$

$$+ \left(\frac{x - x_i}{h}\right)\left(y_{i+1} - \frac{h^2}{6} M_{i+1}\right) \quad \text{..... (2)}$$

$$\therefore \quad f'(x) = -\frac{(x_{i+1} - x)^2}{2h} M_i + \frac{(x - x_i)^2}{6h} M_{i+1} + \frac{h}{6}(M_{i+1} - M_i) + \frac{1}{h}(y_{i+1} - y_i)$$

To impose the condition of continuity of $f'(x)$, we get,

$$f'(x - \epsilon) = f'(x + \epsilon) \text{ as } \epsilon \to 0$$

$$\therefore \quad \frac{h}{6}(2M_i - M_{i-1}) + \frac{1}{h}(y_i - y_{i-1}) = -\frac{h}{6}(2M_i + M_{i+1}) + \frac{1}{h}(y_{i+1} - y_i)$$

or, $$M_{i-1} + 4M_i + M_{i+1} = \frac{6}{h^2}(y_{i-1} - 2y_i + y_{i+1}),\ i = 1 \text{ to } n - 1 \quad \text{..... (3)}$$

Now, since the graph is linear for $x < x_0$ and $x > x_n$, we have,

$$M_0 = 0,\ M_n = 0 \quad \text{..... (4)}$$

Equation (3) and (4) gives (n + 1) equations in (n + 1) unknowns $M_i(i = 0, 1,$ ...., $n)$ which can be solved.

Replacing the value of M_i in (2) gives the concerned cubic spline.

Example 2.13

Obtain the cubic spline for the following data:

x	0	1	2	3
y	2	–6	–8	2

Solution:

Since, the points are equispaced with h = 1 and n = 3, the cubic spline can be determined from

$$M_{i-1} + 4M_i + M_{i+1} = 6[y_{i-1} - 2y_i + y_{i+1}], i = 1, 2.$$

$$\therefore \quad M_0 + 4M_1 + M_2 = 6(y_0 - 2y_1 + y_2)$$

$$M_1 + 4M_2 + M_3 = 6(y_1 - 2y_2 + y_3)$$

i.e. $4M_1 + M_2 = 36$ $\quad [\because M_0 = 0, M_3 = 0]$

$$M_1 + 4M_2 = 72$$

On solving, we get,

$$M_1 = 4.8 \text{ and } M_2 = 16.8$$

Now, the cubic spline in $(x_i \le x \le x_{i+1})$ is,

$$f(x) = \frac{1}{6}(x_{i+1} - x)^3 M_i + \frac{1}{6}(x - x_i)^3 M_{i+1} + (x_{i+1})\left(y_i - \frac{1}{6}M_i\right)$$

$$+ (x - x_i)\left(y_{i+1} - \frac{1}{6}M_{i+1}\right)$$

Taking i = 0,

$$f(x) = \frac{1}{6}(1-x)^3 \times 0 + \frac{1}{6}(x-0)^3(4.8) + (1-x)(x-0) + x\left(-6 - \frac{1}{6}\right) \times 4.8$$

$$= 0.8x^3 - 8.8x + 2 \; (0 \le x \le 1)$$

Now taking i = 1, the cubic spline in $(1 \le x \le 2)$ is,

$$f(x) = \frac{1}{6}(2-x)^3 \times (4.8) + \frac{1}{6}(x-1)^3(16.8) + (2-x)$$

$$\left[-6 - \frac{1}{6} \times (4.8)\right] + (x-1)[-8 - 1(16.8)]$$

$$= 2x^3 - 5.84x^2 - 16.8x + 0.8$$

Taking i = 2, the cubic spline in $(2 \le x \le 3)$ is,

$$f(x) = \frac{1}{6}(3-x)^3 \times 4.8 + \frac{1}{6}(x-2)^3(0) + (3-x)[-8 - 1(16.8)]$$

$$+ (x-2)(2 - 1(2))$$

$$= -0.8x^3 + 2.64x^2 + 9.68x - 14.8$$

2.7 CURVE FITTING: REGRESSION

In many applications, it often becomes necessary to establish a mathematical relationship between experimental values. This relationship may be used for either testing existing mathematical models or establishing new ones. The mathematical equation can also be used to predict or forecast values of the dependent variable. The process of establishing such relationships in the form of a mathematical equation is known as regression analysis or curve fitting.

Suppose the values of y for the different values of x are given. If we want to know the effect of x on y, then we may write a functional relationship $y = f(x)$.

The variable y is called the dependent variable and x the independent variable. The relationship may be either linear or non-linear as shown in figure 2.3. The type of relationship to be used should be decided by the experiment based on the nature of scatteredness of data.

It is a standard practice to prepare a scatter diagram as shown in figure 2.4 and try to determine the functional relationship needed to fit the points. The line should best fit the plotted points. This means that the average error introduced by the assumed line should be minimum. The parameters a and b of the various equations shown in figure 2.3 should be evaluated such that the equations best represent the data.

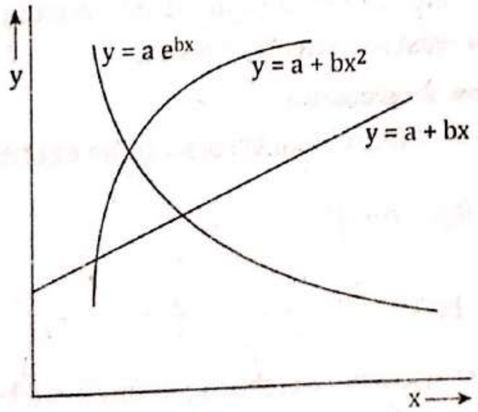

Figure 2.3: Various relationships between x and y

I. Fitting Linear Equations

Fitting a straight line is the simplest approach of regression analysis. Let us consider the mathematical equation for a straight line,

$$y = a + bx = f(x)$$

to describe the data. We know that 'a' is the intercept of the line and 'b' is its slope. Consider a point (x_i, y_i) as shown in figure 2.4. The vertical distance of this points from the line $f(x) = a + bx$ is the error q_i. Then,

$$q_i = y_i - f(x) = y_i - a - bx_i \qquad \text{..... (1)}$$

There are various approaches that could be tried for fitting a 'best' line through the data. They include,

1. Minimize the sum of errors *i.e.*, minimize

$$\Sigma q_i = \Sigma y_i - a - bx_i \qquad \text{..... (2)}$$

2. Minimize the sum of absolute values of errors

$$\Sigma|q_i| = \Sigma|(y_i - a - bx)| \quad \text{..... (3)}$$

3. Minimize the sum of squares of errors

$$\Sigma q_i^2 = \Sigma(y_i - a - bx_i)^2 \quad \text{..... (4)}$$

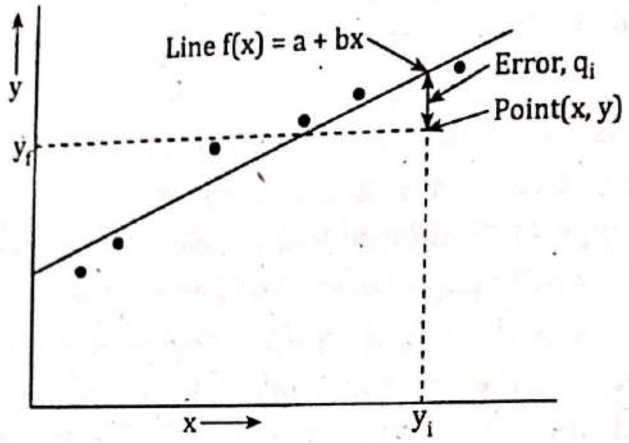

Figure 2.4: Scatter diagram

It can be easily verified that the first two strategies do not yield a unique line for a given set of data. The third strategy overcomes this problems and guarantees a unique line. The technique of minimizing the sum of squares of errors is known as least squares regression.

A. Least Squares Regression

Let the sum of squares of individual errors can be expressed as,

$$Q = \sum_{i=1}^{n} q_i^2 = \sum_{i=1}^{n} [(y_i - f(x_i)]^2$$

$$= \sum_{i=1}^{n} (y_i - a - bx_i)^2 \quad \text{..... (1)}$$

In the method of least squares, we choose a and b such that Q is minimum. Since Q depends on 'a' and 'b', a necessary condition for Q to be minimum is,

$$\frac{\partial Q}{\partial a} = 0 \text{ and } \frac{\partial Q}{\partial b} = 0$$

Then, $\frac{\partial Q}{\partial a} = -2\sum_{i=1}^{n} (y_i - a - bx_i) = 0$

$$\frac{\partial Q}{\partial b} = -2\sum_{i=1}^{n} x_i(y_i - a - bx_i) = 0 \quad \text{..... (2)}$$

Thus, $\Sigma y_i = na + b\Sigma x_i$

$$\Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2$$

These are called normal equations. ..... (3)

Solving for a and b, we get

$$b = \frac{n\Sigma x_i y_i - \Sigma x_i \Sigma y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2}$$

$$a = \frac{\Sigma y_i}{n} - b\frac{\Sigma x_i}{n} = \bar{y} - b\bar{x}$$

when $\bar{x}$ and $\bar{y}$ are the averages of x and y values respectively.

Example 2.14

Fit a straight line to the following set of data:

x	1	2	3	4	5
y	3	4	5	6	8

Solution:

x_i	y_i	x_i^2	x_iy_i
1	3	1	3
2	4	4	8
3	5	9	15
4	6	16	24
5	8	25	40
$\Sigma = 15$	26	55	90

We know,

$$b = \frac{n\Sigma x_iy_i - \Sigma x_i\Sigma y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2}$$

Here, n = 5

so, $$b = \frac{5 \times 90 - 15 \times 26}{5 \times 55 - 15^2} = 1.20$$

Similarly,

$$a = \frac{\Sigma y_i}{n} - \frac{b\Sigma x_i}{n} = \frac{26}{5} - 1.20 \times \frac{15}{5} = 1.60$$

Hence, the linear equation is,

$$y = a + bx = 1.60 + 1.20x$$

Algorithm for Linear Regression

1. Start.
2. Read data values.
3. Compute sum of powers and products.

   $\Sigma x_i, \Sigma y_i, \Sigma x_i^2, \Sigma x_iy_i$
4. Check whether the denominator of the equation for b is zero.
5. Compute b and a.
6. Printout the equation.
7. Interpolate data, if required.
8. Stop.

B. Fitting Transcendental Equations

The relationship between the dependent and independent variables is not always linear (refer figure 2.5)

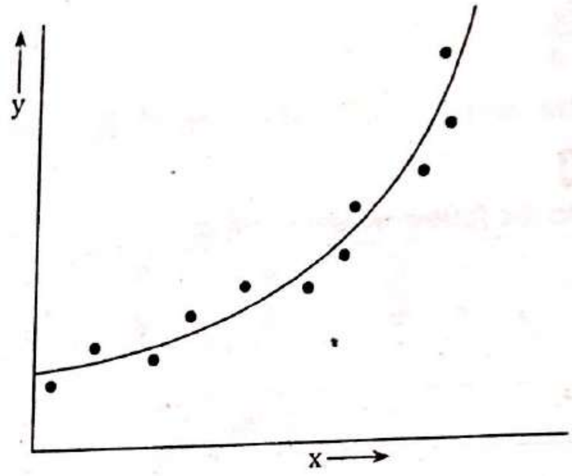

Figure 2.5: Data would fit a non-linear curve better than a linear one.

The non-linear relationship between them may exist in the form of transcendental equations (or higher order polynomials).

For example, the familiar equation for population growth is given by,

$$p = p_0 e^{kt} \qquad \text{..... (1)}$$

where, p_0 is the initial population, k is the growth rate and t is the time.

Another example of non-linear model is the gas law relating to the pressure and volume as given by,

$$P = av^b \qquad \text{..... (2)}$$

Let us consider equation (2) first. If we observe values of p for various values of v, we can then determine the parameters 'a' and 'b'. Using the method of least squares, the sum of squares of all errors can be written as,

$$Q = \sum_{i=1}^{n} [p_i - av_i^b]^2$$

To minimize Q, we have,

$$\frac{\partial Q}{\partial a} = 0 \text{ and } \frac{\partial Q}{\partial b} = 0$$

We can prove that;

$$\Sigma p_i v_i^b = a\Sigma(v_i^b)^2$$

$$\Sigma p_i v_i^b \ln v_i = a\Sigma(v_i^b)^2 \ln v_i$$

These equations can be solved for 'a' and 'b'. But since 'b' appears under the summation sign, an iterative technique must be employed to solve for 'a' and 'b'.

However, this problem can be solved by using the algorithm given in the previous section in the following ways: let us rewrite the equation using the conventional variables x and y as,

$$y = ax^b$$

If we take logarithm on both sides, we get,

$$\ln y = \ln a + b \ln x \qquad \text{..... (3)}$$

This equation is similar in the form to the linear equation and therefore, using the same procedure, we can evaluate the parameters a and b.

$$b = \frac{n\Sigma \ln x_i \ln y_i - \Sigma \ln x_i \Sigma \ln y_i}{n\Sigma(\ln x_i)^2 - (\Sigma \ln x_i)^2} \quad \text{..... (4)}$$

$$\ln a = R = \frac{1}{n}(\Sigma \ln y_i - b\,\Sigma \ln x_i)$$

$$a = e^{R} \quad \text{..... (5)}$$

Similarly, we can linearize the exponential model shown in equation (1) by taking logarithm on both the sides. This would yield,

$$\ln P = \ln p_0 + kt \ln e$$

Since, $\ln e = 1$

We have,

$$\ln P = \ln p_0 + kt \quad \text{..... (6)}$$

This is similar to the linear equation.

$$y = a + bx$$

where, $y = \ln P$

$a = \ln p_0$

$b = k$

$x = t$

We can now easily determine 'a' and 'b' and then p_0 and k.

There is a third form of non-linear model known as saturation growth rate equation as shown below;

$$p = \frac{k_1 t}{k_2 + t} \quad \text{..... (7)}$$

This can be linearized by taking inversion of the terms.

i.e., $$\frac{1}{p} = \left(\frac{k_2}{k_1}\right)\frac{1}{t} = \frac{1}{k_1} \quad \text{..... (8)}$$

This is again similar to the linear equation $y = a + bx$

where, $y = \frac{1}{p}$; $x = \frac{1}{t}$

$a = \frac{1}{k_1}$; $b = \frac{k_2}{k_1}$

Once we obtain 'a' and 'b', they could be transformed back into the original form for the purpose of analysis.

Example 2.15

Given the data table,

x	1	2	3	4	5
y	0.5	2	4.5	8	12.5

Fit a power-function model of the form.

$$y = ax^b$$

Solution:

Given that;

$y = ax^b$

x_i	y_i	$ln\,(x_i)$	$ln\,(y_i)$	$(ln\,x_i)^2$	$(ln\,x_i)\,(ln\,y_i)$
1	0.5	0	-0.6931	0	0
2	2	0.6931	0.6931	0.4805	0.4804
3	4.5	1.0986	1.5041	1.2069	1.6524
4	8	1.3863	2.0794	1.9218	2.8827
5	12.5	1.6094	2.5257	2.5903	4.0649
Sum		4.7874	6.1092	6.1995	9.0804

We get,

$$b = \frac{n\Sigma\, ln\,x_i\, ln\,y_i - \Sigma\, ln\,x_i\, \Sigma\, ln\,y_i}{\Sigma(ln\,x_i)^2 - (\Sigma\, ln\,x_i)^2} = \frac{5 \times 9.0804 - (4.7874) \times (6.1092)}{(5)\,(6.1995) - (4.7874)^2}$$

$$= \frac{45.402 - 29.2472}{30.9975 - 22.9192} = 1.9998$$

$$ln\,a = \frac{1}{n}(\Sigma\, ln\,y_i - b\, \Sigma ln) = \frac{1}{5}(6.1092 - 1.9998 \times 4.7847) = -0.6929$$

$\therefore \quad a = e^{-0.6929} = 0.5001$

Thus, we obtain the power – Function as,

$$y = 0.5001\, x^{1.9998}$$

Note that the data have been derived from the equation,

$$y = \frac{x^2}{2}$$

The discrepancy in the computed coefficients is due to roundoff errors.

C. Fitting Polynomial Function

When a given set of data does not appear to satisfy a linear equations, we can try a suitable polynomial as a regression curve to fit the data. The least squares technique can be readily used to fit the data to a polynomial.

Consider a polynomial of degree m – 1,

$$y = a_1 + a_2x + a_3x^2 + \cdots\cdots + a_mx^{m-1} \quad \text{..... (1)}$$
$$= f(x)$$

If the data contains n sets of x and y values, then the sum of square of the errors is given by,

$$Q = \sum_{i=1}^{n} [(y_i - f(x_i)]^2 \quad \text{..... (2)}$$

Since f(x) is a polynomial and contains coefficients a_1, a_2, a_3, etc. We have to estimate all the m coefficients. As before, we have the following m equations that can be solved for these coefficients

$$\frac{\partial Q}{\partial a_1} = 0$$

$$\frac{\partial Q}{\partial a_2} = 0$$

....

....

....

$$\frac{\partial Q}{\partial a_m} = 0$$

Consider a general term,

$$\frac{\partial Q}{\partial a_j} = -2\sum_{i=1}^{n}[y_i - f(x_i)]\frac{\partial f(x_i)}{\partial a_j} = 0$$

$$\frac{\partial f(x_i)}{\partial a_j} = x_i^{j-1}$$

Thus, we have,

$$\sum_{i=1}^{n}[y_i - f(x_i)]\, x_i^{j-1} = 0, \qquad j = 1, 2, \ldots\ldots, m$$

$$\Sigma[y_i x_i^{j-1} - x_i^{j-1} f(x_i)] = 0$$

Replacing for $f(x_i)$,

$$\sum_{i=1}^{n} x_i^{j-1}(a_1 + a_2x_i + a_3x_i^2 + \ldots\ldots + a_m x_i^{m-1}) = \sum_{i=1}^{n} y_i\, x_i^{j-1}$$

These are m equations ($j = 1, 2, \ldots\ldots, m$) and each summation is for $i = 1$ to n.

$$\begin{aligned}
a_1n + a_2\Sigma x_i + a_3\Sigma x_i^2 + \ldots\ldots + a_m\Sigma x_i^{m-1} &= \Sigma y_i \\
a_1\Sigma x_i + a_2\Sigma x_i^2 + a_3\Sigma x_i^3 + \ldots\ldots + a_m\Sigma x_i^{m} &= \Sigma y_i x_i \qquad \ldots\ldots (3)\\
\vdots \qquad \vdots \qquad \vdots \qquad &\vdots \\
a_1\Sigma x_i^{m-1} + a_2\Sigma x_i^{m} + a_3\Sigma x_i^{m+1} + \ldots\ldots + a_m\Sigma x_i^{2m-2} &= \Sigma y_i x_i^{m-1}
\end{aligned}$$

The set of m equations can be represented in matrix notation as follows:

$$CA = B$$

where,

$$C = \begin{bmatrix} n & \Sigma x_i & \Sigma x_i^2 & \cdots & \Sigma x_i^{m-1} \\ \Sigma x_i & \Sigma x_i^2 & \Sigma x_i^3 & \cdots & \Sigma x_i^{m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \Sigma x_i^{m-1} & \Sigma x_i^{m} & \cdots, & \cdots, & \Sigma x_i^{2m-2} \end{bmatrix}$$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_m \end{bmatrix}, B = \begin{bmatrix} \Sigma y_i \\ \Sigma y_i x_i \\ \Sigma y_i x_2 \\ \vdots \\ \Sigma y_i x_i^{m-1} \end{bmatrix}$$

The element of matrix C is,

$$C(j,k)=\sum_{i=1}^{n}x_i^{j+k-2}\ ;\quad j=1,2,\ldots\ldots,m \text{ and } k=1,2,\ldots\ldots,m.$$

Similarly,

$$B(j)=\sum_{i=1}^{n}y_i x_i^{j-1}\ ;\quad j=1,2,\ldots\ldots,m$$

Example 2.16

Fit a second order polynomial to the data in the table below;

x	1.0	2.0	3.0	4.0
y	6.0	11.0	18.0	27.0

Solution:

The order of polynomial is 2 and therefore we will have 3 simultaneous equations as shown below.

$$a_1 n + a_2\Sigma x_i + a_3\Sigma x_i^2 = \Sigma y_i$$

$$a_1\Sigma x_i + a_2\Sigma x_i^2 + a_3\Sigma x_i^3 = \Sigma y_i x_i$$

$$a_1\Sigma x_i^2 + a_2\Sigma x_i^3 + a_3\Sigma x_i^4 = \Sigma y_i x_i^2$$

The sums of power and products can be evaluated in a tabulator from shown below;

x	y	x^2	x^3	x^4	yx	yx^2
1	6	1	1	1	6	6
2	11	4	8	16	22	44
3	18	9	27	81	54	162
4	27	16	64	256	108	432
Σ = 10	62	30	100	354	190	644

Replacing these values,

$$4a_1 + 10a_2 + 30a_3 = 62$$

op $$10a_1 + 30a_2 + 100a_3 = 190$$

$$30a_1 + 100a_2 + 354a_3 = 644$$

On solving, we get,

$a_1 = 3$

$a_2 = 2$

$a_3 = 1$

Hence the least squares quadratic polynomial is

$$y = 3 + 2x + x^2$$

BOARD EXAMINATION SOLVED QUESTIONS

Use appropriate method of interpolation to get sin θ at 45° from the given table:

θ	10	20	30	40	50
sin θ	0.1736	0.3420	0.5000	0.6428	0.7660

[2013/Fall]

Solution:

Here, the data of $\theta = x$ is equispaced and we have to get $\sin\theta$ at $\theta = x = 45°$ which is near the end of the provided table.

So, we use Newton's backward interpolation formula.

Now, creating difference table from given data

$x = \theta$	$y = \sin\theta$	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
10	0.1736				
		0.1684			
20	0.3420		-0.0104		
		0.1580		-0.0048	
30	0.5000		-0.0152		0.0004
		0.1428		-0.0044	
40	0.6428		-0.0196		
		0.1232			
50	0.7660				

We have,

$$x = 45, \quad h = 50 - 40 = 10, \quad x_n = 50$$

Then,

$$x = x_n + ph$$

or, $\quad 45 = 50 + p10$

∴ $\quad p = -0.5$

Now, using Newton's backward interpolation formula

$$y_p = y_4 + p\nabla y_4 + \frac{p(p+1)}{2!}\nabla^2 y_4 + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_4 + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_4$$

$$= 0.7660 + (-0.5)(0.1232) + \frac{(-0.5)(-0.5+1)(-0.0196)}{2!} + \frac{(-0.5)(-0.5+1)(0.5+2)(-0.0044)}{3!} + \frac{(-0.5)(-0.5+1)(0.5+2)(-0.5+3)(0.0004)}{4!}$$

$= 0.76660 - 0.0616 + 0.00245 + 0.000275 - 0.00001$

$\therefore \quad y_p = 0.7069$

Hence the value of $\sin \theta$ at $\theta = 45°$ is 0.7069.

2. **From the following data:**

x	1	2	3	4	5
y	0.5	2	4.5	8	12.5

Fit a power function model of the from $y = ax^b$. **[2013/Fa**

Solution:

We have the function

$$y = ax^b$$

Taking $\log_{10}$ on both sides

$$\log_{10} y = \log_{10} (ax^b)$$

or, $\log_{10} y = \log_{10} a + b \log_{10} x$

Comparing with the equation,

$$Y = A + bX$$

where, $Y = \log_{10} y$

$A = \log_{10} a$

$X = \log_{10} x$

Forming normal equations as

$$\Sigma Y = nA + b\Sigma X \quad \text{..... (}$$

$$\Sigma XY = A\Sigma X + b\Sigma X^2 \quad \text{..... (}$$

$n = 5$

x	y	$Y = \log_{10} y$	$X = \log_{10} x$	XY	X^2
1	0.5	−0.301	0	0	0
2	2	0.301	0.301	0.0906	0.0906
3	4.5	0.653	0.477	0.3114	0.2276
4	8	0.903	0.602	0.543	0.3624
5	12.5	1.096	0.698	0.765	0.4872
		$\Sigma Y = 2.652$	$\Sigma X = 2.078$	$\Sigma XY = 1.71$	$\Sigma X^2 = 1.168$

Now, equation (1) and (2), we get,

$$2.625 = 5A + 2.078b \quad \text{..... (a)}$$

$$1.71 = 2.078A + 1.168b \quad \text{..... (b)}$$

Solving equation (a) and (b), we get,

$A = -0.299$

or, $a = \text{anti} \log_{10} (-0.299) = 0.5$

and, $b = 1.996$

Hence, $y = 0.5\, x^{1.996}$ is the required function.

3. **If P is pull required to lift a load W by means of a pulley, find a linear law of the form P = mW + C using the following data:**

P	12	15	21	25
W	50	70	100	120

[2013/Spring]

Solution:

We have the function

$P = mW + C$ [$\because Y = a + bX$]

Forming the normal equation

$\Sigma P = nC + m\Sigma W$ ...... (1)

$\Sigma WP = C\Sigma W + m\Sigma W^2$ ...... (2)

$n = 4$

W	P	WP	W^2
50	12	600	2500
70	15	1050	4900
100	21	2100	10000
120	25	3000	14400
$\Sigma Y = 340$	$\Sigma X = 73$	$\Sigma XY = 6750$	$\Sigma X^2 = 31800$

Substituting the obtained value from table to equation (1) and (2), we get,

$73 = 4\,C + 340\,m$ ...... (a)

$6750 = 340\,C + 31800\,m$ ...... (b)

On solving (a) and (b),

$C = 2.275$

$m = 0.187$

Hence a linear law is of the form $P = 0.187\,W + 2.275$.

4. **Estimate the value of sin θ at θ = 25 using Newton-Gregory divided difference formula with the help of the following table:**

θ	10	20	30	40	50
sin θ	0.1736	0.3420	0.5	0.6428	0.7660

[2013/Spring]

Solution:

From Newton-Gregory divided difference formula for 5 data points, we have, Table

$x=\theta$	$y = f(x_n)$	$f(x_n, x_{n+1})$	$f(x_n, x_{n+1}, x_{n+2})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3})$	$f(x_n, x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4})$
10	0.1736				
		$\frac{0.3420-0.1736}{20-10} = \mathbf{0.0168}$			
			$\frac{0.0158-0.0168}{30-10} = \mathbf{-0.05 \times 10^{-3}}$		
20	0.3420				
				$\frac{(-8\times10^{-5})-(-0.05\times10^{-3})}{40-10} = \mathbf{-1 \times 10^{-6}}$	
		$\frac{0.5-0.3420}{30-20} = 0.0158$			
					$\frac{(-0.5\times10^{-6})-(-1\times10^{-6})}{50-10} = \mathbf{0.0125 \times 10^{-6}}$
			$\frac{0.0142-0.0158}{40-20} = -8 \times 10^{-5}$		
30	0.5				
				$\frac{(-9.5\times10^{-5})-(-8\times10^{-5})}{50-20} = -0.5 \times 10^{-6}$	
		$\frac{0.6428-0.5}{40-30} = 0.0142$			
			$\frac{0.0123-0.0142}{50-30} = -9.5 \times 10^{-5}$		
40	0.6428				
		$\frac{0.7660-0.6428}{50-40} = 0.0123$			
50	0.7660				

We have,

$$[x_0, x_1] = 0.0168$$
$$[x_0, x_1, x_2] = -0.05 \times 10^{-3}$$
$$[x_0, x_1, x_2, x_3] = -1 \times 10^{-6}$$
$$[x_0, x_1, x_2, x_3, x_4] = 0.0125 \times 10^{-6}$$

Then, using Newton's Gregory divided difference formula

$$\begin{aligned} y &= y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] \\ &\quad + (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3] \\ &\quad + (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3, x_4] \\ &= 0.1736 + (25 - 10)(0.0168) + (25 - 10)(25 - 20)(-0.05 \times 10^{-3}) \\ &\quad + (25 - 10)(25 - 20)(25 - 30)(-1 \times 10^{-6}) \\ &\quad + (25 - 10)(25 - 20)(25 - 30)(25 - 40)(0.0125 \times 10^{-6}) \\ &= 0.1736 + 0.252 - (3.75 \times 10^{-3}) + (3.75 \times 10^{-4}) + (7.031 \times 10^{-5}) \end{aligned}$$

$\therefore \quad y = 0.4222$

Hence, the value of $\sin\theta$ at $\theta = 25$ is 0.4222.

5. **Find the missing term in the following table using suitable interpolation**

X	0	1	2	3	4
Y	1	3	9	?	81

[2014/F]

Solution:

To find the missing term from the given table, we use linear interpolation method at $x = 3$.

Here,

$$x_1 = 2, \qquad f(x_1) = 9$$
$$x_2 = 4, \qquad f(x_2) = 81$$

Now, from linear interpolation

$$y(x) = f(x_1) + (x - x_1)\frac{[f(x_2) - f(x_1)]}{x_2 - x_1}$$

$$= 9 + (3 - 2)\frac{(81 - 9)}{4 - 2} = 9 + 36 = 45$$

$\therefore \qquad y = 45$

Hence, the required missing term is 45.

Next Method

Here, the provided data is unevenly spaced so, using Lagrange's interpolation formula:

$$y = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)}y_0 + \frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)}y_1$$
$$+ \frac{(x - x_0)(x - x_1)(x - x_3)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)}y_2 + \frac{(x - x_0)(x - x_1)(x - x_2)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)}y_3$$

We have,

$$x_0 = 0 \qquad y_0 = 1$$
$$x_1 = 1 \qquad y_1 = 3$$
$$x_2 = 2 \qquad y_2 = 9$$
$$x_3 = 4 \qquad y_3 = 81$$

Substituting the values

$$y = \frac{(x - 1)(x - 2)(x - 4)}{(0 - 1)(0 - 2)(0 - 4)}(1) + \frac{(x - 0)(x - 2)(x - 4)}{(1 - 0)(1 - 2)(1 - 4)}(3)$$
$$+ \frac{(x - 0)(x - 1)(x - 4)}{(2 - 0)(2 - 1)(2 - 4)}(9) + \frac{(x - 0)(x - 1)(x - 2)}{(4 - 0)(4 - 1)(4 - 2)}(81)$$

When $x = 3$, then

$$y = \frac{(3 - 1)(3 - 2)(3 - 4)}{-8} + 3(3 - 2)(3 - 4) + \frac{3(3 - 1)(3 - 4)}{-4}(9)$$
$$+ \frac{3(3 - 1)(3 - 2)}{24}(81)$$

$$= \frac{1}{4} - 3 + \frac{27}{2} + \frac{81}{4}$$

$$= 31$$

Hence the missing term for $x = 3$ is $y = 31$.

6. **The following table gives the heights, x(cm) and weights y(kg) of five persons.**

x	175	165	160	155	145
y	68	58	55	52	48

Assuming the linear relationship between x and y, obtain the regression line (x or y). Also obtain x value for y = 40. **[2014/Fall]**

Solution:

For linear relationship between x and y, we have,

$y = a + bx$

Forming the normal equation ..... (1)

$\Sigma Y = na + b\Sigma X$ ..... (2)

$\Sigma xy = a\Sigma x + b\Sigma x^2$

$n = 5$

x	y	xy	x^2
175	68	11900	30625
165	58	9570	27225
160	55	8800	25600
155	52	8060	24025
145	48	6960	21025
$\Sigma x = 800$	$\Sigma Y = 281$	$\Sigma xy = 45290$	$\Sigma x^2 = 128500$

Substituting the values obtained in equation (1) and (2), we get,

$281 = 5a + 800b$ ..... (a)

$45290 = 800a + 128500b$ ..... (b)

Solving (a) and (b), we get,

$a = -49.4$

$b = 0.66$

We get,

$y = -49.4 + 0.66x$

which is the required linear regression equation.

Now, for y = 40

$40 = -49.4 + 0.66x$

$\therefore \quad x = 135.45.$

7. **The following table gives the population of a town during the last six censuses. Estimate the increase in the population during the period from 1976 to 1978.**

Year	1941	1951	1961	1971	1981	1991
Population	12	15	20	27	39	52

[2014/Spring, 2015/Fall, 2015/Spring, 2016/Fall]

Solution:

Here, the data of given year is equispaced and we have to estimate data at 1976 and 1978 which is near the end of the table.

So, we use Newton's backward interpolation formula.

Now, creating difference table from given data.

x = year	y = Population	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$	$\nabla^5 y$
1941	12					
		3				
1951	15		2			
		5		0		
1961	20		2		3	
		7		3		−10
1971	27		5		−7	
		12		−4		
1981	39		1			
		13				
1991	52					

At $x = 1976 \quad x_n = 1991, \quad h = 1991 - 1981 = 10$

Then,

$x = x_n + ph$

or, $1976 = 1991 + 10p$

$\therefore p = -1.5$

Now, using Newton's backward interpolation formula,

$$y_p = y_5 + p\nabla y_5 + \frac{p(p+1)}{2!}\nabla^2 y_5 + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_5 + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_5 + \frac{p(p+1)(p+2)(p+3)(p+4)}{5!}\nabla^5 y_5$$

$$= 52 + (-1.5)(13) + \frac{(-1.5)(-1.5+1)}{2!}(1) + \frac{(-1.5)(-1.5+1)(-1.5+2)}{3!}(-4) + \frac{(-1.5)(-1.5+1)(-1.5+2)(-1.5+3)}{4!}(-7) + \frac{(-1.5)(-1.5+1)(-1.5+2)(-1.5+3)(-1.5+4)}{5!}(-10)$$

$$= 52 - 19.5 + 0.375 - 0.25 - 0.1640 - 0.1171$$

$\therefore y_p = 32.3439$

Again,

At $x = 1978, \quad x_n = 1991, \quad h = 1991 - 1981 = 10$

Then,

$x = x_n + ph$

or, $1978 = 1991 + 10p$

$\therefore p = -1.3$

Then, using Newton's backward interpolation formula and substituting the values, we get,

$$y_p = 52 + (-1.3)(13) + \frac{(-1.3)(-1.3+1)}{2!}(1)$$
$$+ \frac{(-1.3)(-1.3+1)(-1.3+2)}{3!}(-4)$$
$$+ \frac{(-1.3)(-1.3+1)(-1.3+2)(-1.3+3)}{4!}(-7)$$
$$+ \frac{(-1.3)(-1.3+1)(-1.3+2)(-1.3+3)(-1.3+4)}{5!}(-10)$$
$$= 52 - 16.5 + 0.195 - 0.182 - 0.1353 - 0.1044$$
$$\therefore \quad y_p = 34.8733$$

Now, Increase in population during the period of 1976–1978 is given by

$$\Delta\text{Population} = y_{p\text{ at }1978} - y_{p\text{ at }1976}$$
$$= 34.8733 - 32.3439 = 2.5294$$

8. The pressure and volume of a gas are related by the equation $PV^{\gamma} = C$, where γ and C being constants. Fit this equation to the following set of observations.

P (kg/cm^2)	0.5	1.0	1.5	2.0	2.5	3.0
V (literes)	1.62	1.00	0.75	0.62	0.52	0.46

[2016/Spring, 2015/Fall, 2014/Spring]

Solution:

Given equation $PV^{\gamma} = C$

Taking log on both sides,

$$\log(PV^{\gamma}) = \log C$$

or, $\log P + \log V^{\gamma} = \log C$

or, lop $P = \log C - \gamma \log V$

or, $Y = A - \gamma X$

where, $Y = \log P$

$A = \log C$

$X = \log V$

Forming normal equations

$$\Sigma Y = nA - \gamma \Sigma X \quad \text{..... (1)}$$
$$\Sigma XY = A\Sigma X - \gamma \Sigma X^2 \quad \text{..... (2)}$$

$n = 6$

P	V	$Y = \log P$	$X = \log V$	XY	X^2
0.5	1.62	−0.301	0.209	−0.0629	0.0436
1.0	1.00	0	0	0	0
1.5	0.75	0.176	−0.124	−0.0218	0.0153
2.0	0.62	0.301	−0.207	−0.0623	0.0428
2.5	0.52	0.397	−0.283	−0.1123	0.0800
3.0	0.46	0.477	−0.337	−0.1607	0.1135
		$\Sigma Y = 1.050$	$\Sigma X = -0.742$	$\Sigma XY = -0.420$	$\Sigma X^2 = 0.2952$

Substituting obtained values in equation (1) and (2), we get,

$1.050 = 6A + 0.742\gamma$ ..... (a)

$-0.420 = -0.742A - 0.2952\gamma$ ..... (b)

Solving equation (a) and (b), we get,

$A = -0.00137$

$\gamma = 1.426$

and, $A = \log C$

or, $C = \text{antilog}_{10}(A) = 10^{-0.00137}$

$\therefore$ $C = 0.9968$

Hence, required equation is $PV^{1.426} = 0.9968$

9. For the following set of data, fit a parabolic curve using least square method and find f(2).

x	0.5	1	1.5	4.5	6.5	7.5
f(x)	2.5	2.7	3.5	6.5	8.4	9.5

[2015/Spring]

Solution:

We have, $y = a + bx + cx^2$ as a parabolic curve.

Forming normal equations as,

$\Sigma y = na + b\Sigma x + c\Sigma x^2$ ..... (1)

$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$ ..... (2)

$\Sigma x^2y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$ ..... (3)

$n = 6$

x	f(x) = y	x^2	x^3	x^4	xy	x^2y
0.5	2.5	0.25	0.125	0.0625	1.25	0.625
1	2.7	1	1	1	2.7	2.7
1.5	3.5	2.25	3.375	5.0625	5.25	7.875
4.5	6.5	20.25	91.125	410.06	29.25	131.625
6.5	8.4	42.25	274.62	1785	54.6	354.9
7.5	9.5	56.25	421.87	3164	71.25	534.375
$\Sigma x =$ 21.5	$\Sigma y =$ 33.1	$\Sigma x^2 =$ 122.25	$\Sigma x^2 =$ 792.11	$\Sigma x^4 =$ 5365.18	$\Sigma xy =$ 164.3	$\Sigma x^2y =$ 1032.1

Substituting the values obtained in equations (1) and (2) and (3), we get,

$33.1 = 6a + 21.5b + 122.25c$ ..... (a)

$164.3 = 21.5a + 122.25b + 792.11c$ ..... (b)

$1032.1 = 122.25a + 792.11b + 5365.18c$ ..... (c)

Solving equation (a), (b) and (c), we get,

$a = 1.8840$

$b = 1.0218$

$c = -0.0014$

$\therefore$ Required parabolic curve is $y = 1.8840 + 1.0218x - 0.0014x^2$

Now, to find f(2),

$$f(2) = 1.8840 + 1.0218 \times 2 - 0.0014 \times (2)^2$$

$\therefore$ $f(2) = 3.922$.

10. Use Newton's divided difference formula to find f(3) form the following data:

x	0	1	2	4	5	6
f(x)	1	14	15	5	6	19

[2016/Spring]

Solution:

From Newton's divided difference formula for 6 data points.

Let, $y = f(x_n)$

$y_a = f(x_n, x_{n+1})$

$y_b = f(x_n, x_{n+1}, x_{n+2})$

$y_c = f(x_n, x_{n+1}, x_{n+2}, x_{n+3})$

$y_d = f(x_n, x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4})$

$y_e = f(x_n, x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4}, x_{n+5})$

Table:

x	y	y_a	y_b	y_c	y_d	y_e
0	1					
		$\frac{14-1}{1-0} = 13$				
1	14		$\frac{1-13}{2-0} = -6$			
		$\frac{15-14}{2-1} = 1$		$\frac{-2+6}{4-0} = 1$		
2	15		$\frac{-5-1}{4-1} = -2$		$\frac{1-1}{5-0} = 0$	
		$\frac{5-15}{4-2} = -5$		$\frac{2+2}{5-1} = 1$		$\frac{0-0}{6-0} = 0$
4	5		$\frac{1+5}{5-2} = 2$		$\frac{1-1}{6-1} = 0$	
		$\frac{6-5}{5-4} = 1$		$\frac{6-2}{6-2} = 1$		
5	6		$\frac{13-1}{6-4} = 6$			
		$\frac{19-6}{6-5} = 13$				
6	19					

We have,

$[x_0, x_1] = 13$

$[x_0, x_1, x_2] = -6$

$[x_0, x_1, x_2, x_3] = 1$

$[x_0, x_1, x_2, x_3, x_4] = 0$

$[x_0, x_1, x_2, x_3, x_4, x_5] = 0$

Now, Newton's divided difference formula

$$y = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2]$$
$$+ (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3]$$
$$+ (x - x_0)(x - x_1)(x - x_2)(x - x_3)[x_0, x_1, x_2, x_3, x_4]$$
$$+ (x - x_0)(x - x_1)(x - x_2)(x - x_3)(x - x_4)[x_0, x_1, x_2, x_3, x_4, x_5]$$

At $x = 3$, we have,

$$y = 1 + (3 - 0)\,13 + (3 - 0)(3 - 1)(-6) + (3 - 0)(3 - 1)(3 - 2)(1) + 0 + 0$$
$$= 1 + 39 - 36 + 6$$

$\therefore \quad y = 10$

Hence, the value of $y = f(x)$ at $x = 3$ is 10.

11. By the method of least square methods, find the straight line that best fits the following data:

x	1	2	3	4	5
y	14	27	40	55	68

[2016/Spring]

Solution:

Using the function $y = a + bx$ to find straight line

Forming the normal equation,

$\Sigma y = na + b\Sigma x$ ..... (1)

$\Sigma xy = a\Sigma x + b\Sigma x^2$ ..... (2)

$n = 5$

x	y	xy	x^2
1	14	14	1
2	27	54	4
3	40	120	9
4	55	220	16
5	68	340	25
$\Sigma x = 15$	$\Sigma Y = 204$	$\Sigma xy = 748$	$\Sigma x^2 = 55$

Substituting the values obtained in equation (1) and (2), we get,

$204 = 5a + 15b$ ..... (a)

$748 = 15a + 55b$ ..... (b)

Solving (a) and (b), we get,

$a = 0$

$b = 13.6$

Hence, $y = 0 + 13.6x$ is the equation of best fit.

12. The growth of bacteria (N) in a culture after t hours is given by the following table:

Time t(hr)	0	1	2	3	4
Bacteria (N)	32	47	65	92	132

If the relationship between bacteria N and time t is of the form $N = ab^t$. Using least square approximation estimate the N at t = 5 hr.
[2017/Spring]

Solution:

Given that;

$$N = ab^t$$

Taking N = log on both sides

$$\log_{10} N = \log_{10} (ab^t)$$

or, $\log_{10} N = \log_{10} a + \log_{10} b^t$

or, $\log_{10} N = \log_{10} a + t \log_{10} b$

Comparing with the equation,

$$Y = A + BX$$

where, $Y = \log_{10} N$

$A = \log_{10} a$

$X = t$

$B = \log_{10} b$

Now, forming normal equations

$\Sigma Y = nA + B\Sigma X$

$\Sigma XY = A\Sigma X + B\Sigma X^2$ ..... (1)

$n = 5$ ..... (2)

t	N	$Y = \log_{10} N$	$X = t$	XY	X^2
0	32	1.505	0	0	0
1	47	1.672	1	1.672	1
2	65	1.812	2	3.624	4
3	92	1.963	3	5.889	9
4	132	2.120	4	8.48	16
		$\Sigma Y = 9.072$	$\Sigma X = 10$	$\Sigma XY = 19.665$	$\Sigma X^2 = 30$

Substituting the obtained values in equation (1) and (2), we get,

$9.072 = 5A + 10B$ ..... (a)

$19.665 + 10A + 30B$ ..... (b)

On solving (a) and (b), we get,

$A = 1.5102$

$B = 0.1521$

Then,

$A = \log_{10} a$

or, $a = \text{anitlog}_{10}(A) = 10^{1.5102} = 32.374$

and, $B = \log_{10} b$

or, $b = 10^{0.1521} = 1.419$

So the relation between bacteria N and time t is

$$N = 32.374 \times 1.419^t$$

Now, at t = 5 hr,

$$N = 32.374 \times 1.419^5$$

$\therefore$ $N = 186.25$

13. **The following table given the percentage of criminals for different age groups. Using interpolation formula, find the percentage of criminals under the age of 35.**

Under age	25	30	40	50
% of criminals	52	67.3	84.1	94.4

[2017/Spring]

Solution:

The provided data in the table is unevenly spaced, so using Lagrange's interpolation formula.

$$y = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} y_0 + \frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} y_1$$

$$+ \frac{(x - x_0)(x - x_1)(x - x_3)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{(x - x_0)(x - x_1)(x - x_2)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)} y_3$$

We have,

$x_0 = 25$ $\quad$ $y_0 = 52$

$x_1 = 30$ $\quad$ $y_0 = 67.3$

$x_2 = 40$ $\quad$ $y_0 = 84.1$

$x_3 = 50$ $\quad$ $y_0 = 94.4$

When, x = 35, then,

$$y = \frac{(35 - 30)(35 - 40)(35 - 50)}{(25 - 30)(25 - 40)(25 - 50)}(52)$$

$$+ \frac{(35 - 25)(35 - 40)(35 - 50)}{(30 - 25)(30 - 40)(30 - 50)}(67.3)$$

$$+ \frac{(35 - 25)(35 - 30)(35 - 50)}{(40 - 25)(40 - 30)(40 - 50)}(84.1)$$

$$+ \frac{(35 - 25)(35 - 30)(35 - 40)}{(50 - 25)(50 - 30)(50 - 40)}(94.4)$$

$\therefore$ $= -10.4 + 50.475 + 42.05 - 4.72$

$y = 77.405$

Hence the percentage of criminal under age of 35 is 77.405%.

14. **Find the number of students securing marks between 50-55 using appropriate interpolation technique.**

Marks obtained	20-30	30-40	40-50	50-60
No. of students	10	20	30	40

[2017/Fall]

Solution:

Total number of students = 10 + 20 + 30 + 40 = 100

Here, the data of marks obtained, x is equispaced at interval of 10 and 50-50 lies at near end of the provided table.

So, we use Newton's backward interpolation formula.

Now, creating difference table

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
20	0				
		10			
30	10		10		
		20		0	
40	30		10		0
		30		0	
50	60		10		
		40			
60	100				

We have,

$x = 50 \quad x_n = 55, \quad h = 60 - 50 = 10$

Then,

$x = x_n + ph$

or, $50 = x_n + 10p$

$\therefore \quad p = -0.5$

Now, using Newton's backward interpolation formula

$$y_p = y_4 + p\nabla y_4 + \frac{p(p+1)}{2!}\nabla^2 y_4 + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_4 + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_4$$

$$= 100 - 0.5 \times 40 + \frac{(-0.5)(-0.5+1)}{2!}(10) + 0 + 0$$

$$= 100 - 20 - 1.25$$

$\therefore \quad y_p = 78.75 \approx 79$

Hence, number of students securing marks between 50-55 are;

$= 79 - 60$

$= 19$ students

15. **The voltage V across a capacitor at time t seconds is given by following table**

Time t(sec)	0	2	4	6	8	10
Voltage (V)	150	63	28	12	5.6	124

If the relationship between voltage V and time t is of the form $V = e^{kt}$. Using least-square approximation. Estimate the voltage at t = 2.6 sec.

[2017/Fall]

Solution:

Given that;

$$V = e^{kt}$$

Taking $\log_{10}$ on both sides,

$$\log_{10} V = \log_{10} (e^{kt})$$

or, $\log_{10} V = (kt) \log_{10} (e)$

or, $\log_{10} V = [k \log_{10} (e)]t$

Comparing with

$$Y = A + BX$$

where, $Y = \log_{10} V$

$B = k \log_{10} (e)$

$X = t$

$A = 0$

Now, forming normal equations,

$$\Sigma Y = nA + B\Sigma X \quad \text{..... (1)}$$

$$\Sigma XY = A\Sigma X + B\Sigma X^2 \quad \text{..... (2)}$$

Since, A = 0, equations become

$$\Sigma Y = B\Sigma X \text{ and } \Sigma XY = B\Sigma X^2$$

n = 6

x = t	V	$Y = \log_{10} V$	XY	X^2
0	150	2.176	0	0
2	63	1.799	3.598	4
4	28	1.447	5.788	16
6	12	1.079	6.474	36
8	5.6	0.748	5.984	64
10	124	2.093	20.930	100
$\Sigma X = 30$		$\Sigma Y = 9.342$	$\Sigma XY = 42.774$	$\Sigma X^2 = 220$

Substituting the obtained values in equation (1) and (2), we get,

$$9.342 = B30$$

or, $B = \frac{9.342}{30} = 0.3114$

$$B = k \log_{10} (e)$$

or, $k = \frac{B}{\log_{10}(e)} = \frac{0.3114}{\log_{10}(e)} = 0.717$

Hence, $V = e^{0.717t}$ is the required relation.

And, at t = 2.63 seconds

$$V = e^{0.717 \times 2.6} = 6.45 \text{ volts}$$

16. Determine the constants a and b by the method of least squares such that $y = ae^{bx}$.

X	2	4	6	8	10
Y	4.077	11.084	30.128	81.897	222.62

[2018/Spring]

Solution:

Given that;

$$y = ae^{bx}$$

Taking log on both sides,

$$\log_{10} y = \log_{10}(ae^{bx})$$

or, $\log_{10} y = \log_{10}(a) + \log_{10}(e^{bx})$

or, $\log_{10} y = \log_{10}(a) + bx \log_{10}(e)$

or, $\log_{10} y = \log_{10} a + (b \log_{10} e)\, x$

Comparing with

$$Y = A + BX$$

where, $Y = \log_{10} y$

$A = \log_{10} a$

$B = b \log_{10} e$

Now, forming normal equations,

$$\Sigma Y = nA + B\Sigma X \quad \text{..... (1)}$$

$$\Sigma XY = A\Sigma X + B\Sigma X^2 \quad \text{..... (2)}$$

$n = 5$

X	y	$Y = \log_{10}(y)$	XY	X^2
2	4.077	0.610	1.22	4
4	11.084	1.044	4.176	16
6	30.128	1.478	8.868	36
8	81.897	1.913	15.304	64
10	222.62	2.347	23.47	100
$\Sigma X = 30$		$\Sigma Y = 7.392$	$\Sigma XY = 53.038$	$\Sigma X^2 = 220$

Substituting the obtained values in equation,

$$7.392 = 5A + 30B \quad \text{..... (a)}$$

$$53.038 = 30A + 220B \quad \text{..... (b)}$$

On solving (a) and (b), we get,

$A = 0.175$

$B = 0.217$

Now,

$A = \log_{10}(a)$

$\therefore \quad a = \text{anitlog}_{10}(A) = 10^{0.175} = 1.496$

and, $\quad B = b\log_{10}(e)$

$\therefore \quad b = \dfrac{B}{\log_{10}(e)} = \dfrac{0.217}{\log_{10}(e)} = 0.499$

Hence, the required relation is $y = 1.496\, e^{0.499x}$.

17. **From the following, find the number of students who obtained less than 45 marks.**

Marks	30-40	40-50	50-60	60-70	70-80
No. of students	31	42	51	35	31

[2018/Spring, 2019/Spring]

Solution:

First we prepare the cumulative frequency table

Marks (x)	40	50	60	70	80
Students (y)	31	73	124	159	190

Now, creating the difference table

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
40	31				
		42			
50	73		9		
		51		-25	
60	124		-16		37
		35		12	
70	159		-4		
		31			
80	190				

To find the number of students with marks less than 45.

Taking, $x = 45, \quad x_0 = 40, \quad h = 50 - 40 = 10$

$x = x_0 + ph$

$\therefore \quad p = 0.5$

Now using Newton's forward interpolation formula,

$$y_{45} = y_3 + p\nabla y_0 + \frac{p(p-1)}{2!}\nabla^2 y_{40} + \frac{p(p-1)(p-2)}{3!}\nabla^3 y_{40} + \frac{p(p-1)(p-2)(p-3)}{4!}\nabla^4 y_{40}$$

$$= 31 + 0.5 \times 42 + \frac{0.5(-0.5)}{2} \times 9 + \frac{0.5(-0.5)(-1.5)}{6}(-25) + \frac{0.5(-0.5)(-1.5)(-2.5)}{24}(37)$$

$$= 31 + 21 - 1.125 - 1.5625 - 1.4453$$

$\therefore \quad y_{45} = 47.87$

Here, the number of students with marks less than 40 is 31 and the number of students with marks less than 45 is 47.87 = 48.

Also, students securing marks in between 40 and 45 = 48 − 31 = 17.

18. Generate a Lagrange's interpolating polynomial for the function $y = \cos \pi x$, taking the pivotal points $0, \frac{1}{4}$ and $\frac{1}{2}$. **[2018/Fall]**

Solution:

Given that;

$x_0 = 0$; $y_0 = \cos \pi x_0 = 1$

$x_1 = \frac{1}{4} = 0.25$; $y_1 = \cos \pi x_1 = 0.707$

$x_2 = \frac{1}{2} = 0.5$; $y_2 = \cos \pi x_2 = 0$

Now, using Lagrange's interpolation formula

$$y = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2$$

Substituting the values

$$y = \frac{(x - 0.25)(x - 0.5)}{(0 - 0.25)(0 - 0.5)}(1) + \frac{(x - 0)(x - 0.5)}{(0.25 - 0)(0.25 - 0.5)}(0.707)$$

$$+ \frac{(x - 0)(x - 0.25)}{(0.5 - 0)(0.5 - 0.25)}(0)$$

$$= \frac{(x - 0.25)(x - 0.5)}{\left(\frac{1}{8}\right)} + \frac{(x - 0)(x - 0.5)}{-0.0625}(0.707) + (0)$$

$$= 8(x - 0.25)(x - 0.5) - 11.312x(x - 0.5)$$

$$= (x - 0.5)(8x - 2 - 11.312x)$$

$$= x(-3.312x - 2) - 0.5(-3.312x - 2)$$

$$= -3.312x^2 - 0.344x + 1$$

is the required Lagrange's interpolating polynomial for the given function.

19. The voltage V across a capacitor at a time T seconds is given by the following table. Use the principle of least squares to fit the curve of the form. $V = \alpha e^{\beta T}$ to the data.

T	0	2	4	6	8
V	150	63	28	12	5.6

[2013/Fall, 2019/Spring]

Solution:

Given that;

$V = \alpha e^{\beta T}$

Taking log on both sides

$\log_{10} V = \log_{10}(\alpha e^{\beta T})$

or, $\log_{10} V = \log_{10} \alpha + \beta T \log_{10} e$

Comparing with

$Y = A + BX$

where, $Y = \log_{10} V$

$A = \log_{10} \alpha$

$X = T$

$B = \beta \log_{10} e$

Forming normal equations

$\Sigma Y = nA + B\Sigma X$ ..... (1)

and, $\Sigma XY = A\Sigma X + B\Sigma X^2$ ..... (2)

$n = 5$

X = T	V	Y = $\log_{10}$ V	XY	X^2
0	150	2.176	0	0
2	63	1.799	3.598	4
4	28	1.447	5.788	16
6	12	1.079	6.474	36
8	5.6	0.748	5.984	64
ΣX = 20		ΣY = 7.249	ΣXY = 21.844	$\Sigma X^2 = 120$

Substituting the obtained values,

$7.249 = 5A + 20B$ ..... (a)

$21.844 = 20A + 120B$ ..... (b)

On solving (a) and (b), we get,

$A = 2.165$

$B = -0.178$

Then,

$A = \log_{10} \alpha$

or, $\alpha = \text{anitlog}_{10} (A) = 10^{2.165} = 146.217$

and, $B = \beta \log_{10} e$

or, $\beta = \frac{B}{\log_{10} e} = \frac{-0.178}{\log_{10} (e)} = -0.409$

Hence, the required curve is $V = 146.217\ e^{-0.409T}$.

20. Fit a curve of the form: $y = \frac{1}{a + bx}$ by using the method of least square with the following data points.

x	1	2	3	4	5
f(x)	3.33	2.20	1.52	1.00	0.91

[2018/Fall]

Solution:

Given that;

$y = \frac{1}{a + bx}$

or, $\frac{1}{y} = a + bx$

Comparing with $Y = A + BX$

$$Y = \frac{1}{y}, A = a, B = b, X = x$$

Forming normal equations

$$\Sigma X = nA + B\Sigma X \quad \text{---(1)}$$

$$\Sigma XY = A\Sigma X + B\Sigma X^2 \quad \text{---(2)}$$

$n = 5$

x	f(x) = y	$Y = \frac{1}{y}$	XY	X^2
1	3.33	0.3	0.3	1
2	2.20	0.454	0.908	4
3	1.52	0.657	1.971	9
4	1.00	1	4	16
5	0.91	1.098	5.49	25
$\Sigma X = 15$		$\Sigma Y = 3.509$	$\Sigma XY = 12.669$	$\Sigma X^2 = 55$

Substituting the obtained values,

$$3.509 = 5A + 15B \quad \text{..... (a)}$$

$$12.669 = 15A + 55B \quad \text{..... (b)}$$

On solving (a) and (b), we get,

$$A = 0.0592$$

$$B = 0.2142$$

Then, $Y = 0.0592 + 0.2142x$

or, $\frac{1}{y} = 0.0592 + 0.2142x$

$$\therefore \quad y = \frac{1}{0.0592 + 0.2142x}$$

is the required curve of best fit.

21. The function y = f(x) is given at the points (7, 3), (8, 1) (9, 1) and (10, 9). Find the value of y = for x = 9.5 using Lagrange Interpolation formula. **[2019/Fall]**

Solution:

Given that;

$x_0 = 7$, $y_0 = 3$

$x_1 = 8$, $y_1 = 1$

$x_2 = 9$, $y_2 = 1$

$x_3 = 10$, $y_3 = 9$

Now, using Lagrange's interpolation formula

$$y = f(x) = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} y_0 + \frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} y_1 + \frac{(x - x_0)(x - x_1)(x - x_3)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{(x - x_0)(x - x_1)(x - x_2)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)} y_3$$

At x = 9.5,

$$y = \frac{(9.5-8)(9.5-9)(9.5-10)(3)}{(7-8)(7-9)(7-10)} + \frac{(9.5-7)(9.5-9)(9.5-10)(1)}{(8-7)(8-9)(8-10)}$$
$$+ \frac{(9.5-7)(9.5-8)(9.5-10)(1)}{(9-7)(9-8)(9-10)} + \frac{(9.5-7)(9.5-8)(9.5-9)(9)}{(10-7)(10-8)(10-9)}$$
$$= 0.1875 - 0.3125 + 0.9375 + 2.8125$$

∴ $y = 3.625$ is the required answer.

22. The following table shows pressure and specific volume of dry saturated steam.

V	38.4	20	8.51	4.44	3.03
P	10	20	50	100	150

Fit a curve of the form: $PV^{\alpha} = \beta$ by using least square method.

[2019/Fall]

Solution:

Given that;

$$PV^{\alpha} = \beta$$

Taking log on both sides

$$\log_{10}(PV^{\alpha}) = \log_{10}\beta$$

or, $\log_{10} P + \alpha \log_{10} V = \log_{10}\beta$

or, $\log_{10} P = \log_{10}\beta - \alpha \log_{10} V$

Comparing with

$$Y = A + BX$$

where, $Y = \log_{10} P$

$A = \log_{10}\beta$

$B = -\alpha$

$X = +\log_{10} V$

Now, forming normal equations

$$\Sigma Y = nA + B\Sigma X \quad \text{..... (1)}$$

and, $$\Sigma XY = A\Sigma X + B\Sigma X^2 \quad \text{..... (2)}$$

$n = 5$

V	P	$Y = \log_{10} P$	$X = \log_{10} V$	XY	X^2
38.4	10	1	1.548	1.584	2.509
20	20	1.301	1.301	1.692	1.692
8.51	50	1.698	0.929	1.577	0.863
4.44	100	2	0.647	1.294	0.418
3.03	150	2.176	0.481	1.046	0.231
		$\Sigma Y = 8.175$	$\Sigma X = 4.942$	$\Sigma XY = 7.193$	$\Sigma X^2 = 5.713$

Substituting the obtained values,

$$8.175 = 5A + 4.942B \quad \text{..... (a)}$$
$$7.193 = 4.942A + 5.713B \quad \text{..... (b)}$$

On solving (a) and (b), we get,

$A = 2.693$

$B = -1.071$

Then,

$A = \log_{10}(\beta)$

$\beta = \text{anitlog}_{10}(A) = 10^A = 10^{2.693} = 493.173$

and, $B = -\alpha$

or, $\alpha = 1.071$

Hence, $PV^{1.071} = 493.173$ is the required curve of best fit.

23. From following experimental data, it is known that the relation connects V and t as $V = at^b$. Find the possible values of a and b.

V	350	400	500	600
P	61	26	7	2.6

[2020/Fall]

Solution:

Given that;

$V = at^b$

Taking $\log_{10}$ on both sides,

$\log_{10} V = \log_{10}(at^b)$

or, $\log_{10} V = \log_{10} a + \log_{10} t^b$

or, $\log_{10} V = \log_{10} a + b \log_{10} t$

Comparing with

$Y = A + BX$

where, $Y = \log_{10} V$

$A = \log_{10} a$

$B = b$

$X = \log_{10} t$

Forming normal equations

$\Sigma Y = nA + B\Sigma X$ ..... (1)

and, $\Sigma XY = A\Sigma X + B\Sigma X^2$ ..... (2)

$n = 4$

V	T	$Y = \log_{10} V$	$X = \log_{10} t$	XY	X^2
350	61	2.544	1.785	4.541	3.186
400	26	2.602	1.414	3.679	1.999
500	7	2.698	0.845	2.279	0.714
600	2.6	2.778	0.414	1.150	0.171
		$\Sigma Y = 10.622$	$\Sigma X = 4.458$	$\Sigma XY = 11.649$	$\Sigma X^2 = 6.07$

Substituting the obtained values

$$10.622 = 4A + 4.458B \quad \text{..... (a)}$$

$$11.649 = 4.458A + 6.07B \quad \text{..... (b)}$$

On solving (a) and (b), we get,

$$A = 2.846$$

$$B = -0.171$$

Then,

$$A = \log_{10} a$$

or, $a = \text{antilog}_{10}(A) = 10^{2.846} = 701.455$

and, $B = b = -0.171$

Hence, $V = 701.455t^{-0.171}$ is the required solution.

24. **The following table gives the viscosity of oil as the function of temperature. Use Lagrange's interpolation formula to find the viscosity of oil at a temperature of 140°C.**

T(°C)	110	130	160	190
Viscosity	10.8	8.1	5.5	4.8

[2020/Fall]

Solution:

Given that;

$x_0 = 110 \qquad y_0 = 10.8$

$x_1 = 130 \qquad y_1 = 8.1$

$x_2 = 160 \qquad y_2 = 5.5$

$x_3 = 190 \qquad y_3 = 4.8$

Now, using Lagrange's interpolation formula

$$y = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}y_1 + \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}y_3$$

At $x = 140$,

$$y = \frac{(140-130)(140-160)(140-190)}{(110-130)(110-160)(110-190)}(10.8) + \frac{(140-110)(140-160)(140-190)}{(130-110)(130-160)(130-190)}(8.1) + \frac{(140-110)(140-130)(140-190)}{(160-110)(160-130)(160-190)}(5.5) + \frac{(140-110)(140-130)(140-160)}{(190-110)(190-130)(190-160)}(4.8)$$

$$= -1.35 + 6.75 + 1.833 - 0.2$$

$\therefore$ $y = 7.033$ is the required viscosity of oil.

25. Write short notes on cubic spline.
[2013/Spring, 2017/Fall, 2018/Spring, 2019/Spring]

Solution: See the topic 2.6.

26. Write short notes on: An algorithm for Lagrange's interpolation polynomial. **[2014/Fall, 2018/Fall]**

Solution:

Algorithm for Lagrange's interpolation polynomial.

1.	Read x, n
2.	For i = 1 to (n + 1) in steps of 1 do read x_i, f_i end for
3.	Sum ← 0
4.	For i = 1 to (n + 1) in steps of 1 do
5.	Prodfunc ← 1
6.	For j = 1 to (n + 1) in steps of 1 do

7.	If (j + 1) then
	Prodfunc ← Prodfunc × $(x - x_j)/(x_i - x_j)$
	end for

8.	Sum ← sum + f_i × prodfunc Remarks: sum is the value of f at x end for
9.	Write x, sum s
10.	Stop.

27. Write short notes on: Linear interpolation. **[2015/Fall]**

Solution: See the topic 2.5.

28. Write short notes on: Numerical differentiation. **[2016/Fall]**

Solution:

It is the process of calculating the value of the derivative of a function at some assigned value of x from the given set of values (x_i, y_i). To compute $\frac{dy}{dx}$, we first replace the exact relation $y = f(x)$ by the best interpolating polynomial $y = \phi(x)$ and then differentiate the latter as many times as we desire. The choice of the interpolation formula to be used, will depend on the assigned value of x at which $\frac{dy}{dx}$ is desired.

If the value of x are equi-spaced and is required near the beginning of the table, we employ Newton's forward formula. If it is required near the end of the table, $\frac{dy}{dx}$ is calculated by means of Stirlings or Bessel's formula. If the values of x are not equi-spaced, we use Lagrange's formula or Newton's divided difference formula to represent the function.

Hence, corresponding to each of the interpolation formula, we can derive a formula for finding the derivative. While using this formula it must be observed that the table of values defines the function at these points only and does not completely define the function and the function may not be differentiable at all. As such, the process of numerical differentiation should be used only if the tabulated values are such that the differences of some order are constants. Otherwise, errors are bound to creep in which go on increasing as derivatives of higher order are found. This is due to the fact that the difference between f(x) and the approximating polynomial $\phi(x)$, may be small at the data points but $f'(x) - \phi'(x)$ may be large.

1. **Forward difference formulae**

$$\left(\frac{dy}{dx}\right)_{x_0} = \frac{1}{h}\left[\Delta y_0 - \frac{1}{2}\Delta^2 y_0 + \frac{1}{3}\Delta^3 y_0 - \frac{1}{4}\Delta^4 y_0 + \cdots\cdots\right]$$

$$\left(\frac{d^2y}{dx^2}\right)_{x_0} = \frac{1}{h^2}\left[\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12}\Delta^4 y_0 - \cdots\cdots\right]$$

and so on.

2. Backward difference formula

$$\left(\frac{dy}{dx}\right)_{x_n} = \frac{1}{h}\left[\nabla y_n - \frac{1}{2}\nabla^2 y_n + \frac{1}{3}\nabla^3 y_n - \frac{1}{4}\nabla^4 y_n + \cdots\cdots\right]$$

$$\left(\frac{d^2y}{dx^2}\right)_{x_n} = \frac{1}{h^2}\left[\nabla^2 y_n - \nabla^3 y_n + \frac{11}{12}\nabla^4 y_n - \cdots\cdots\right]$$

and so on.