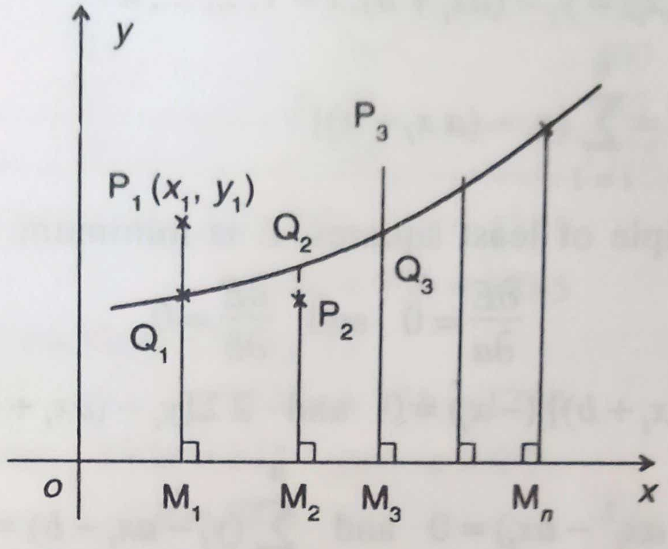
1.6. The principle of least squares

In the previous articles, we have seen two methods of fitting curve, *viz.*, (*i*) the graphical method and (*ii*) the method of group averages. The first one is a rough method and in the second method, the evaluation of constants vary from one grouping to another grouping of data. So, we adopt another method, called *the method of least squares* which gives a unique set of values to the constants in the equation of the fitting curve.



Let (x_i, y_i), $i = 1, 2, ..., n$ be the n sets of observations and let

$$y = f(x) \quad ...(1)$$

be the relation suggested between x and y.

Let (x_i, y_i) be represented by the point P_i. Let the ordinate at P_i meet $y = f(x)$ at Q_i and the x-axis at M_i.

$$M_iQ_i = f(x_i) \quad \text{and} \quad M_iP_i = y_i$$

$$Q_iP_i = M_iP_i - M_iQ_i = y_i - f(x_i), \quad i = 1, 2, ..., n$$

$d_i = y_i - f(x_i)$ is called the residual at $x = x_i$. Some of the d_i's may be positive and some may be negative.

$$E = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} [y_i - f(x_i)]^2 \text{ is the sum of the squares of the residuals.}$$

If $E = 0$, *i.e.*, each $d_i = 0$, then all the n points P_i will lie on $y = f(x)$.

If not, we will choose $f(x)$ such that E is minimum. That is, the best fitting curve to the set of points is that for which E is minimum. This principle is known as the *principle of least squares* or the *least square criterion.*

This principle *does not suggest* to determine the *form of the curve* $y = f(x)$ but it *determines the values of the parameters* or constants of the equation of the curve.

We will consider some of the best fitting curves of the type :

(*i*) a straight line (*ii*) a second degree curve (*ii*) the exponential curve $y = ae^{bx}$ (*iv*) the curve $y = ax^n$.

1.7. Fitting a straight line by the method of least squares

Let (x_i, y_i), $i = 1, 2, ..., n$ be the n sets of observations and let the related relation by $y = ax + b$. Now we have to select a and b so that the straight line is the best fit to the data.

As explained earlier, the residual at $x = x_i$ is

$$d_i = y_i - f(x_i) = y_i - (ax_i + b),\ i = 1, 2, ..., n$$

$$E = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} [y_i - (a\,x_i + b)]^2$$

By the principle of least squares, E is minimum.

$$\therefore \qquad \frac{\partial E}{\partial a} = 0 \text{ and } \frac{\partial E}{\partial b} = 0$$

i.e., $\quad 2\Sigma[y_i - (ax_i + b)]\,(-x_i) = 0$ and $2\,\Sigma[y_i - (ax_i + b)](-1) = 0$

i.e., $\quad \sum_{i=1}^{n} (x_i y_i - ax_i^2 - bx_i) = 0$ and $\sum_{i=1}^{n} (y_i - ax_i - b) = 0$

i.e.,

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \qquad ...(1)$$

and

$$a\sum_{i=1}^{n} x_i + nb = \sum_{i=1}^{n} y_i \qquad ...(2)$$

Since, x_i, y_i are known, equations (1) and (2) give two equations in a and b. Solve for a and b from (1) and (2) and obtain the best fit $y = ax + b$.

Note 1. Equations (1) and (2) are called *normal equations.*

2. Dropping suffix i from (1) and (2), the normal equations are

$$a\Sigma x + nb = \Sigma y \text{ and } a\Sigma x^2 + b\Sigma x = \Sigma xy$$

which are got by taking Σ on both sides of $y = ax + b$ and also taking Σ on both sides after multiplying by x both sides of $y = ax + b$.

3. Transformations like $X = \dfrac{x-a}{h}$, $Y = \dfrac{y-b}{k}$ reduce the linear equation $y = \alpha x + \beta$ to the form $Y = AX + B$. Hence, a linear fit is another linear fit in both systems of coordinates.

Example 14. *By the method of least squares find the best fitting straight line to the data given below :*

x :	5	10	15	20	25
y :	15 16	19	23	26	30

Solution. Let the straight line be $y = ax + b$

The normal equations are $a\Sigma x + 5b = \Sigma y$ ...(1)

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

To calculate Σx, Σx^2, Σy, Σxy we form below the table.

	x	y	x^2	xy
	5	16	25	80
	10	19	100	190
	15	23	225	345
	20	26	400	520
	25	30	625	750
Total	75	114	1375	1885

The normal equations are $75a + 5b = 114$ ...(1)

$$1375a + 75b = 1885 \quad ...(2)$$

Eliminate b; multiply (1) by 15

$$1125a + 75b = 1710 \quad ...(3)$$

(2) – (3) gives, $250a = 175$ or $a = 0{\cdot}7$

Hence $b = 12{\cdot}3$

Hence, the best fitting line is $\mathbf{y = 0{\cdot}7x + 12{\cdot}3}$

Aliter. Let $X = \dfrac{x-15}{5}$, $Y = y - 23$

Let the line in the new variable be $Y = AX + B$ ...(1)

x	y	X	X^2	Y	XY
5	16	– 2	4	– 7	14
10	19	– 1	1	– 4	4
15	23	0	0	0	0
20	26	1	1	3	3
25	30	2	4	7	14
Total ; Σ		0	10	– 1	35

The normal equations are $A\Sigma X + 5B = \Sigma Y$ ...(4)

$$A\Sigma X^2 + B\Sigma X = \Sigma XY \quad ...(5)$$

Therefore, $5B = -1$ $\therefore$ $B = -0{\cdot}2$

$\therefore$ $10A = 35$ $A = 3{\cdot}5$

The equations $Y = 3{\cdot}5X - 0{\cdot}2$

i.e., $$y-23=3\cdot5\left(\frac{x-15}{5}\right)-0\cdot2=0\cdot7x-10\cdot5-0\cdot2$$

i.e., $$\mathbf{Y=0\cdot7x+12\cdot3}$$

which is the same equation as seen before.

Example 15. *Fit a straight line to the data given below. Also estimate the value of y at x = 2·5.*

x :	*0*	*1*	*2*	*3*	*4*
y :	*1*	*1·8*	*3·3*	*4·5*	*6·3*

Solution. Let the best fit be $y=ax+b$ ...(1)

The normal equations are

$$a\Sigma x+5b=\Sigma y \quad ...(2)$$

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

We prepare the table for easy use.

	x	y	x^2	xy
	0	1·0	0	0
	1	1·8	1	1·8
	2	3·3	4	6·6
	3	4·5	9	13·5
	4	6·3	16	25·2
Total	10	16·9	30	47·1

Substituting in (2) and (3), we get,

$$10a+5b=16\cdot9$$

$$30a+10b=47\cdot1$$

Solving, we get, $a=1\cdot33,\quad b=0\cdot72$

Hence, the equation is $y=1\cdot33x+0\cdot72$

$$y\,(at\; x=2\cdot5)=1\cdot33\times2\cdot5+0\cdot72=4\cdot045$$

Example 16. *Fit a straight line to the following data. Also estimate the value of y at x = 70.*

x :	*71*	*68*	*73*	*69*	*67*	*65*	*66*	*67*
y :	*69*	*72*	*70*	*70*	*68*	*67*	*68*	*64*

Solution. Since the values of x and y are larger, we choose the origins for x and y at 69 and 67 respectively. In other words, we transform x and y.

Let $X=x-69$, and $Y=y-67$

Let $Y=aX+b$ be the best fit. ...(1)

The normal equations are

$$a\Sigma X+8b=\Sigma Y \quad ...(2)$$

$$a\Sigma X^2+b\Sigma X=\Sigma XY \quad ...(3)$$

Calculations :

x	y	X	Y	X^2	XY
71	69	2	2	4	4
68	72	− 1	5	1	− 5
73	70	4	3	16	12
69	70	0	3	0	0
67	68	− 2	1	4	− 2
65	67	− 4	0	16	0
66	68	− 3	1	9	− 3
67	64	− 2	− 3	4	6
Total:		− 6	12	54	12

Substituting in (2) and (3),

$$-6a + 8b = 12 \qquad ...(4)$$

$$54a - 6b = 12$$

Solving, we get, $a = 0{\cdot}424242$, $b = 0{\cdot}181818$

Therefore, $\quad Y = 0{\cdot}4242X + 0{\cdot}1818$

i.e., $\quad y - 67 = 0{\cdot}4242\,(x - 69) + 0{\cdot}1818$

$$y = 0{\cdot}4242x + 37{\cdot}909$$

$$y\,(x = 70) = 0{\cdot}4242 \times 70 + 37{\cdot}909 = \mathbf{67{\cdot}6030}$$

Example 17. *By proper transformation, convert the relation $y = a + bxy$ to a linear form and find the equation to fit the data.*

x :	− 4	1	2	3
y :	4	6	10	8

[*MS. BE 1972*]

Solution. Let $X = xy$ $\therefore$ The equation becomes $y = a + bx$.

The normal equations are

$$a\Sigma X + b\Sigma X^2 = \Sigma Xy \qquad ...(1)$$

and
$$4a + b\Sigma X = \Sigma y \qquad ...(2)$$

	x	y	X	X^2	XY
	− 4	4	− 16	256	− 64
	1	6	6	36	36
	2	10	20	400	200
	3	8	24	576	192
Total:		28	34	1268	364

Using (1) and (2),

$$34a + 1268b = 364$$

$$4a + 34b = 28$$

Solving we get, $b = 0{\cdot}1286$, $a = 5{\cdot}9069$

Therefore, the equation is $y = 5{\cdot}9069 + 0{\cdot}1286\,X$

i.e., $\quad y = 5{\cdot}9069 + 0{\cdot}1286\,xy.$

Using this equation we get $y\,(1 - 0{\cdot}1286x) = 5{\cdot}9069$

i.e., $\quad y = \dfrac{5{\cdot}9069}{1 - 0{\cdot}1286x}$

We tabulate the values to verify:

x :	-4	1	2	3
y :	5·23	6·78	7·95	9·62

Note. If we take $u = \frac{1}{xy}$, $v = \frac{1}{x}$ we get

$v = au + b$. Taking this as linear, we get

$a = 10{\cdot}5,\ b = -0{\cdot}13$

That is $y = 10{\cdot}5 - 0{\cdot}13xy$

i.e., $$y = \frac{10{\cdot}5}{1 + 0{\cdot}13x}$$

Now tabulating, we get

x :	-4	1	2	3
y :	21·87	92·92	0·12	7·55

The values of y are far away from the given values. Perhaps, the *selection of the form* is not correct.

1.8. Fitting a parabola or fitting a second degree curve (by the method of least squares)

Let $(x_i, y_i), i = 1, 2, ..., n$ be n sets of observations of two related variables x and y. Let $y = ax^2 + bx + c$ be the equation which fits them best.

Now, we have to find the constants a, b, c.

For any $x = x_i$, the expected value of y is $ax_i^2 + bx_i + c$ and the corresponding observed value is y_i.

The residual $d_i = y_i - (ax_i^2 + bx_i + c)$

Let E denote the sum of the squares of the residuals.

That is, $$E = \sum_{i=1}^{n} [y_i - (ax_i^2 + bx_i + c)]^2$$

By the principle of least squares, E is minimum for best values a, b, c.

$\therefore$ $$\frac{\partial E}{\partial a} = 0, \quad \frac{\partial E}{\partial b} = 0 \text{ and } \frac{\partial E}{\partial c} = 0$$

$\therefore$ Differentiating E, partially w.r.t. a, b, c and equating to zero, we get,

$$\sum_{i=1}^{n} 2\,[y_i - (ax_i^2 + bx_i + c)]\,(-x_i^2) = 0$$

$$\sum_{i=1}^{n} 2\,[y_i - (ax_i^2 + bx_i + c)]\,(-x_i) = 0$$

$$\sum_{i=1}^{n} 2\,[y_i - (a\,x_i + bx_i + c)]\,(-1) = 0$$

Simplifying, we get

$$a\Sigma\, x_i^4 + b\Sigma x_i^3 + c\Sigma x_i^2 = \Sigma x_i^2 y_i$$

$$a\Sigma x_i^3 + b\Sigma x_i^2 + c\Sigma x_i = \Sigma x_i y_i$$

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

Dropping the suffices, the *normal equations* are

$$a\Sigma x^4 + b\Sigma x^3 + c\Sigma x^2 = \Sigma\, x^2 y \qquad ...(1)$$

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

$$a\Sigma x^2 + b\Sigma x + nc = \Sigma y \qquad ...(3)$$

The three equations (1), (2), (3) give the values of a, b, c. Substituting these values of a, b, c in $y = ax^2 + bx + c$, we get the result.

Note. To obtain the normal equations, we remember the following :

(*i*) In $y = ax^2 + bx + c$, take Σ on both sides.

(*ii*) Multiply by x both sides and then take Σ on both sides.

(*iii*) Multiply both sides by x^2 and then take Σ on both sides.

Example 18. *Fit a parabola, by the method of least squares, to the following data; also estimate y at x = 6.*

x :	1	2	3	4	5
y :	5	12	26	60	97

Solution. Let $y = ax^2 + bx + c$ be the best fit.

Then, the normal equations are

$$a\Sigma x^2 + b\Sigma x + 5c = \Sigma y \qquad ...(1)$$

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

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

We form the table.

	x	y	x^2	x^3	x^4	xy	x^2y
	1	5	1	1	1	5	5
	2	12	4	8	16	24	48
	3	26	9	27	81	78	234
	4	60	16	64	256	240	960
	5	97	25	125	625	485	2425
Total:	15	200	55	225	979	832	3672

Hence the equations (1), (2), (3) become,

$$55a + 15b + 5c = 200 \qquad ...(4)$$

$$225a + 55b + 15c = 832 \qquad ...(5)$$

$$979a + 225b + 55c = 3672 \qquad ...(6)$$

Solving we get, $a = 5{\cdot}7143$, $b = -11{\cdot}0858$ and $c = 10{\cdot}4001$

Hence, the parabola is, $y = 5{\cdot}7143x^2 - 11{\cdot}0858x + 10{\cdot}4001$

$$y\,(x = 6) = 149{\cdot}6001$$

Example 19. *Fit a second degree parabola to the data.*

x :	*1929*	*1930*	*1931*	*1932*	*1933*	*1934*	*1935*
y :	*352*	*356*	*357*	*358*	*360*	*361*	*361*

Let $X = x - 1932$, $Y = y - 357$

Let $Y = aX^2 + bX + C$ be the best fit.

The normals equations are

$$a\Sigma x^2 + b\Sigma X + 7c = \Sigma Y \qquad ...(1)$$

$$a\Sigma X^3 + b\Sigma X^2 + c\Sigma X = \Sigma XY \qquad ...(2)$$

$$a\Sigma X^4 + b\Sigma X^3 + c\Sigma X^2 = \Sigma X^2 Y \qquad ...(3)$$

Calculation Table :

	x	y	X	Y	X^2	X^3	X^4	XY	X^2Y
	1929	352	– 3	– 5	9	– 27	81	15	– 45
	1930	356	– 2	– 1	4	– 8	16	2	– 4
	1931	357	– 1	0	1	– 1	1	0	0
	1932	358	0	1	0	0	0	0	0
	1933	360	1	3	1	1	1	3	3
	1934	361	2	4	4	8	16	8	16
	1935	361	3	4	9	27	81	12	36
Total			0	6	28	0	196	40	6

Hence the normal equations become,

$$28a + 7c = 6 \qquad ...(4)$$

$$28b = 40 \qquad ...(5)$$

$$196a + 28c = 6 \qquad ...(6)$$

$\therefore$ $b = 1{\cdot}4286$, $a = -0{\cdot}21429$, $c = 1{\cdot}7143$

$\therefore$ The equation is $Y = -0{\cdot}21429\,X^2 + 1{\cdot}4286\,X + 1{\cdot}7143$

i.e., $y - 357 = -0{\cdot}21429\,(x - 1932)^2 + 1{\cdot}7143 + 1{\cdot}4286\,(x - 1932)$

i.e., $y = -0{\cdot}21429\,x^2 + 829{\cdot}445\,x - 802265{\cdot}33$

Example 20. *Fit a second degree parabola to the following data, taking y as dependent variable.*

x :	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
y :	*2*	*6*	*7*	*8*	*10*	*11*	*11*	*10*	*9*

Solution. Let $X = x - \bar{x} = x - 5$ and $Y = y - 7$

Let $Y = aX^2 + bX + c$ be the best fit.

x	y	X	Y	X^2	X^3	X^4	XY	X^2Y
1	2	– 4	– 5	16	– 64	256	20	– 80
2	6	– 3	– 1	9	– 27	81	3	– 9
3	7	– 2	0	4	– 8	16	0	0

4	8	−1	1	1	−1	1	−1	1
5	10	0	3	0	0	0	0	0
6	11	1	4	1	1	1	4	4
7	11	2	4	4	8	16	8	16
8	10	3	3	9	27	81	9	27
9	9	4	2	16	64	256	8	32
Total		0	11	60	0	708	51	−9

The normal equations are

$$708a + 60c = -9$$

$$60b = 51$$

$$60a + 9c = 11$$

Solving we get, $b = 0{\cdot}85$, $a = -0{\cdot}2673$, $c = 3{\cdot}0042$

Hence the equation is

$$Y = -0{\cdot}2673\,X^2 + 0{\cdot}85\,X + 3{\cdot}0042$$

$\therefore$ $$y - 7 = -0{\cdot}2673\,(x-5)^2 + 0{\cdot}85\,(x-5) + 3{\cdot}0042$$

i.e., $$\mathbf{y = -0{\cdot}2673\,x^2 + 3{\cdot}523\,x - 0{\cdot}9283}$$

1.9. Fitting an exponential curve

Let (x_i, y_i), $i = 1, 2, \ldots, n$ be the n sets of observations of related data and let $y = ab^x$ be the best fit for the data. Then taking logarithm on both sides,

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

i.e., $$Y = A + Bx \text{ where } Y = \log_{10} y,\; A = \log_{10} a,\; B = \log_{10} b$$

This being linear in x and Y, *we can find A, B* since x and $Y = \log_{10} y$ are known. From A, B, we can get a, b and hence $y = ab^x$ is found out.

1.10. Fitting a curve of the form $y = ax^b$

Since $y = ax^b$, $\log_{10} y = \log_{10} a + b \log_{10} x$

i.e., $$Y = A + bX \text{ where } Y = \log_{10} y,\; X = \log_{10} x \text{ and } A = \log_{10} a$$

Again using this linear fit, we find A, b.

Hence a, b are known. Thus $y = ab^x$ is found out.

Example 21. *From the table given below, find the best values of a and b in the law $y = ae^{bx}$ by the method of least squares.*

x :	0	5	8	12	20
y :	3·0	1·5	1·0	0·55	0·18

Solution. $y = ae^{bx}$

$\therefore$ $$\log_{10} y = \log_{10} a + bx \log_{10} e$$

i.e., $$Y = A + Bx \qquad \ldots(1)$$

The normal equations are

$$B\Sigma x + 5A = \Sigma Y \qquad \ldots(2)$$

$$B\Sigma x^2 + A\Sigma x = \Sigma xY \quad ...(3)$$

	x	y	Y	x^2	xY
	0	3·0	0·4771	0	0
	5	1·5	0·1761	25	0·8805
	8	1·0	0	64	0
	12	0·55	− 0·2596	144	− 3·1152
	20	0·18	− 0·7447	400	− 14·894
Total	45		− 0·3511	633	− 17·1287

Using equations (2) and (3),

$$5A + 45B = -0{\cdot}3511$$

$$45A + 633B = -17{\cdot}1287$$

Solving we get $A = 0{\cdot}4815$, $B = -0{\cdot}0613$

$$a = 10^A = 3{\cdot}0304$$

$$b \log_{10} e = B = -0{\cdot}0613$$

$$\therefore \quad b = -0{\cdot}0613 \times \log_e 10 = -0{\cdot}1411$$

Hence, the curve is $\mathbf{y = 3{\cdot}0304\, e^{-0{\cdot}1411x}}$

Example 22. *Fit a curve of the form* $y = ab^x$ *to the data*

x :	1	2	3	4	5	6
y :	151	100	61	50	20	8

Solution. $y = ab^x$

$$\therefore \quad \log_{10} y = \log_{10} a + x \log_{10} b$$

i.e., $Y = A + Bx$

The normal equations are ...(1)

$$B\Sigma x + 6A = \Sigma Y \quad ...(2)$$

$$B\Sigma x^2 + A\Sigma x = \Sigma xY \quad ...(3)$$

	x	y	Y	x^2	xY
	1	151	2·1790	1	2·1790
	2	100	2·0	4	4·0
	3	61	1·7853	9	5·3559
	4	50	1·6990	16	6·7960
	5	20	1·3010	25	6·5050
	6	8	0·9031	36	5·4186
Total	21		9·8674	91	30·2545

Using (2) and (3), we get

$$6A + 21B = 9{\cdot}8674 \quad ...(4)$$

$$21A + 91B = 30{\cdot}2545 \quad ...(5)$$

Solving, $A = 2{\cdot}5010$, $B = -0{\cdot}2447$

Since $\log_{10} a = A$, $a = 10^A = 316{\cdot}9568$

$$b = 10^B = 0{\cdot}5692$$

$\therefore$ the equation is $\mathbf{y = 316{\cdot}9568\,(0{\cdot}5692)^x}$

Example 23. *It is known that the curve $y = ax^b$ fits in the data given below. Find the best values of a and b.*

x :	1	2	3	4	5	6
y :	1200	900	600	200	110	50

Solution. $y = ax^b$

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

i.e., $$Y = A + bX \qquad ...(1)$$

where $Y = \log_{10} y$, $X = \log_{10} x$, $A = \log_{10} a$

The normal equations are

$$b\Sigma X + 6A = \Sigma Y \qquad ...(2)$$

$$b\Sigma X^2 + A\Sigma X = \Sigma XY \qquad ...(3)$$

x	y	X	Y	X^2	XY
1	1200	0·0	3·0792	0·0	0·0
2	900	0·3010	2·9542	0·0906	0·8892
3	600	0·4771	2·7782	0·2276	1·3255
4	200	0·6021	2·3010	0·3625	1·3854
5	110	0·6990	2·0414	0·4886	1·4269
6	50	0·7781	1·6990	0·6054	1·3220
Total		2·8573	14·8530	1·7747	6·3490

Using (2) and (3), we have,

$$6A + 2{\cdot}8573b = 14{\cdot}8530$$

$$2{\cdot}8573A + 1{\cdot}7747b = 6{\cdot}3490$$

Solving, $A = 3{\cdot}3086$, $b = -1{\cdot}7494$

$\therefore$ $a = 10^A = 2035$

Hence, the equation is $\mathbf{y = 2035\, x^{-1{\cdot}7494}}$

f the squares of the residuals in the case

For a constant times another function, we have

$$\Delta\{Cf(x)\} = C\Delta\{f(x)\}$$

$$\Delta[af(x) + bg(x)] = a\Delta[f(x)] + b\Delta[g(x)]$$

INTERPOLATION

Consider the table

x	x_0	x_1	x_2	...	x_n
$f(x)$	$f(x_0)$	$f(x_1)$	$f(x_2)$	...	$f(x_n)$

If the value of $f(y)$ is to be found at some point y in the interval $[x_0, x_n]$ and y is not one of the tabulated points, then the value of $f(y)$ is estimated by using the known values of $f(x)$ at the surrounding points. This process of computing the value of a function inside the given range is called **interpolation**. Simply interpolation means insertion or filling up intermediate terms of a series. If the point y lies outside the domain $[x_0, x_n]$ then the estimation of $f(y)$ is called **extrapolation.** In this chapter we will be mainly concerned with interpolation.

Newton's Forward Interpolation Formula

We know that

$\Delta y_0 = y_1 - y_0$ i.e., $y_1 = y_0 + \Delta y_0 = (1 + \Delta)\, y_0$

$\Delta y_1 = y_2 - y_1$ i.e., $y_2 = y_1 + \Delta y_1 = (1 + \Delta)\, y_1 = (1 + \Delta)^2\, y_0$

$\Delta y_2 = y_3 - y_2$ i.e., $y_3 = y_2 + \Delta y_2 = (1 + \Delta)\, y_2 = (1 + \Delta)^3\, y_0$

In general, $y_n = (1 + \Delta)^n\, y_0$

Expanding $(1 + \Delta)^n$ by using Binomial theorem we have

$$y_n = \left\{1 + n\Delta + \frac{n(n-1)\Delta^2}{2!} + \frac{n(n-1)(n-2)\Delta^3}{3!} + \ldots\right\} y_0$$

$$y_n = f(x_0 + nh) = y_0 + n\,\Delta y_0 + \frac{n(n-1)}{2!}\Delta^2 y_0 + \frac{n(n-1)(n-2)}{3!}\Delta^3 y_0 + \ldots$$

This result is known as **Gregory–Newton forward interpolation (or) Newton's formula for equal intervals.**

Note 1 : This formula is very useful and gives gre accuracy when $x_0 + nh$ is near the beginning of the table

Note 2 : The first two terms of the above formula will the linear interpolation, the third term will give parabolic interpolation and so on.

Example

Find the missing y_x values from the first difference provi

y_x	0	–	–	–	–	–
Δy_x	0	1	2	4	7	11

Solution

Let the unknowns be a, b, c, d, e. Then we have the follow table.

y_x	0	a	b	c	d	e
Δy_x	0	1	2	4	7	11

By definition, we get,

$$a - 0 = 1 \quad \text{i.e.,} \quad a = 1$$

$$b - a = 2 \quad \Rightarrow \quad b = 2 + a = 3$$

$$c - b = 4 \quad \Rightarrow \quad c = 4 + b = 7$$

$$d - c = 7 \quad \Rightarrow \quad d = 7 + c = 14$$

$$e - d = 11 \quad \Rightarrow \quad e = 11 + d = 25$$

Example

A function $f(x)$ is given by the following table. Find f (0. a suitable formula.

x	0	1	2	3	4	5	6
$f(x)$	176	185	194	203	212	220	229

Solution

The difference table is as follows :

$$\therefore y(1946) = 12 + (3.5)(1) + \frac{(3.5)(3.5-1)}{2} \times 6$$

$$+ \frac{(3.5)(3.5-1)(3.5-2)}{6} \times (-6)$$

$$+ \frac{3.5(3.5-1)(3.5-2)(3.5-3)}{24} \times 11$$

$$+ \frac{(3.5)(3.5-1)(3.5-2)(3.5-3)(3.5-4)}{120} \times (-20)$$

$$= 12 + 3.5 + 26.25 - 13.125 + 3.0078 + 0.5469$$

$$= 12 + 3.5 + 26.25 + 3.0078 + 0.5469 - 13.125$$

$$= 32.18$$

∴ The population in the year 1946 is 32.18.

To find the population in the year 1948 :

To find $y(1948)$.

i.e., $x_0 + nh = 1948$

$$1911 + n \cdot 10 = 1948$$

$$n = 3.7$$

$$y(1948) = 12 + 3.7 + \frac{(3.7)(3.7-1)}{2} \times 6$$

$$+ \frac{3.7(3.7-1)(3.7-2)}{6} \times (-6)$$

$$+ \frac{3.7(3.7-1)(3.7-2)(3.7-3)}{24} \times 11$$

$$+ \frac{3.7(3.7-1)(3.7-2)(3.7-3)(3.7-4)}{120} \times (-20)$$

$$= 12 + 3.7 + 29.97 - 16.983 + 5.4487 + 0.5944$$

$$= 34.73$$

The population in the year 1948 is 34.73.

Increase in the population during the period 1946 to 1948.

= Population in 1948 – Population in 1946

= 34.73 – 32.18

= 2.55 thousands

Example 3

The following table gives the population of a town during the last six census. Estimate, using Newton's interpolation formula, the increase in the population during the period 1946 to 1948.

Year	1911	1921	1931	1941	1951	1961
Population (in thousands)	12	13	20	27	39	52

Solution

The difference table is as follows :

x	$y = f(x)$	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
1911(x_0)	12 (y_0)	Δy_0				
		1	($\Delta^2 y_0$)			
1921	13		6	($\Delta^3 y_0$)		
		7		−6	($\Delta^4 y_0$)	
1931	20		0		11	($\Delta^5 y_0$)
		7		5		−20
1941	27		5		−9	
		12		−4		
1951	39		1			
		13				
1961	52					

Here $x_0 = 1911$, $h = 10$, $y_0 = 12$

By Newton's formula we have

$$y(x_0 + nh) = y_0 + n\Delta y_0 + \frac{n(n-1)}{2!}\Delta^2 y_0 + \frac{n(n-1)(n-2)}{3!}\Delta^3 y_0 +$$

$$y(1946) = ?$$

i.e., $x_0 + nh = 1946$

i.e., $1911 + n \cdot 10 = 1946$ *i.e.*, $n = 3.5$

$$\therefore y(1946) = 12 + (3.5)(1) + \frac{(3.5)(3.5-1)}{2} \times 6 + \frac{(3.5)(3.5-1)(3.5-2)}{6} \times (-6) + \frac{3.5(3.5-1)(3.5-2)(3.5-3)}{24} \times 11 + \frac{(3.5)(3.5-1)(3.5-2)(3.5-3)(3.5-4)}{120} \times (-20)$$

$$= 12 + 3.5 + 26.25 - 13.125 + 3.0078 + 0.5469$$

$$= 12 + 3.5 + 26.25 + 3.0078 + 0.5469 - 13.125$$

$$= 32.18$$

∴ The population in the year 1946 is 32.18.

To find the population in the year 1948 :

To find $y(1948)$.

i.e., $x_0 + nh = 1948$

$$1911 + n \cdot 10 = 1948$$

$$n = 3.7$$

$$\therefore y(1948) = 12 + 3.7 + \frac{(3.7)(3.7-1)}{2} \times 6 + \frac{3.7(3.7-1)(3.7-2)}{6} \times (-6) + \frac{3.7(3.7-1)(3.7-2)(3.7-3)}{24} \times 11 + \frac{3.7(3.7-1)(3.7-2)(3.7-3)(3.7-4)}{120} \times (-20)$$

$$= 12 + 3.7 + 29.97 - 16.983 + 5.4487 + 0.5944$$

$$= 34.73$$

The population in the year 1948 is 34.73.

Increase in the population during the period 1946 to 1948.

= Population in 1948 – Population in 1946

= 34.73 – 32.18

= 2.55 thousands

BACKWARD DIFFERENCES

We use another operator called the backward difference operator ∇ and is defined by

$$\nabla y_n = y_n - y_{n-1}$$

For $n = 0, 1, 2, \ldots$ we get

$$\nabla y_0 = y_0 - y_{-1}$$

$$\nabla y_1 = y_1 - y_0$$

$$\nabla y_2 = y_2 - y_1, \text{ and so on.}$$

The second backward difference is

$$\nabla^2 y_n = \nabla (\nabla y_n)$$

$$= \nabla (y_n - y_{n-1})$$

$$= \nabla y_n - \nabla y_{n-1}$$

$$= (y_n - y_{n-1}) - (y_{n-1} - y_{n-2})$$

$$= y_n - 2y_{n-1} + y_{n-2}$$

Similarly the third backward difference is

$$\nabla^3 y_n = \nabla^2 y_n - \nabla^2 y_{n-1}$$

$$= (y_n - 2y_{n-1} + y_{n-2}) - (y_{n-1} - 2y_{n-2} + y_{n-3})$$

$$= y_n - 3y_{n-1} + 3y_{n-2} - y_{n-3} \text{ and so on.}$$

Note : The backward difference can also be defined in following way, where h being the length of the interval.

i.e., $\nabla f(x) = f(x) - f(x-h)$

$$\nabla^2 f(x) = \nabla [\nabla f(x)]$$

$$= \nabla [f(x) - f(x-h)]$$

$$= \{f(x) - f(x-h)\} - \{f(x-h) - f(x-2h)\}$$

$$= f(x) - 2f(x-h) + f(x-2h) \text{ and so on.}$$

The differences are shown below

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
$x_{-4} = x_0 - 4h$	y_{-4}				
		∇y_{-3}			
$x_{-3} = x_0 - 3h$	y_{-3}		$\nabla^2 y_{-2}$		
		∇y_{-2}		$\nabla^3 y_{-1}$	
$x_{-2} = x_0 - 2h$	y_{-2}		$\nabla^2 y_{-1}$		$\nabla^4 y_0$
		∇y_{-1}		$\nabla^3 y_0$	
$x_{-1} = x_0 - h$	y_{-1}		$\nabla^2 y_0$		
		∇y_0			
$x_0 = y_0$	y_0				

Newton's Backward Interpolation Formula

We know that $\nabla y_1 = y_1 - y_0$ (or) $(1 - \nabla)\, y_1 = y_0$

(or)

$$y_1 = (1 - \nabla)^{-1}\, y_0 \quad \dots (1)$$

Also we know that $y_1 = (1 + \Delta)\, y_0 \quad \dots (2)$

[By definition of forward difference operator]

From (1) and (2) we get,

$$(1 - \nabla)^{-1} \equiv (1 + \Delta)$$

Hence

$$y_n = (1 + \Delta)^n\, y_0 = (1 - \nabla)^{-n}\, y_0$$

$$= \left[1 + n\nabla + \frac{n\,(n+1)}{2!}\nabla^2 + \frac{n\,(n+1)\,(n+2)}{3!}\nabla^3 + \dots\right] y_0$$

i.e., $y\,(x_0 + nh)$

$$= y_0 + n\nabla y_0 + \frac{n\,(n+1)}{2!}\nabla^2 y_0 + \frac{n\,(n+1)\,(n+2)}{3!}\nabla^3 y_0 + \dots$$

This is **Gregory-Newton backward** interpolation formula.

Note: 1. This formula is very useful and gives greater accuracy when $x_0 + nh$ is near the end of the table.

2. The Gregory-Newton formula may also be used to extrapolate beyond the interval in which y is unknown.

Example 18

The following table gives the values of a function at equal intervals.

x	0.0	0.5	1.0	1.5	2.0
$f(x)$	0.3989	0.3521	0.2420	0.1295	0.0540

Evaluate $f(1.8)$.

Solution

The difference table is as given below.

x	$y = f(x)$	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
0.0	0.3989				
		–0.0468			
0.5	0.3521		–0.0633		
		–0.1101		0.0609	
1.0	0.2420		–0.0024		–0.0215
		–0.1125		0.0394	$(\nabla^4 y_0)$
1.5	0.1295		0.0370	$(\nabla^3 y_0)$	
		–0.0755	$(\nabla^2 y_0)$		
2.0	0.0540	(∇y_0)			
(x_0)	(y_0)				

Here $x_0 = 2.0$, $y_0 = 0.0540$, $h = 0.5$.

As we have to find the value of $f(1.8)$ near the lower end, we apply Newton's Backward interpolation formula,

$$y(x_0 + nh) = y_0 + n\nabla y_0 + \frac{n(n+1)}{2!}\nabla^2 y_0 + \frac{n(n+1)(n+2)}{3!}\nabla^3 y_0 + \ldots$$

$$y(1.8) = ?$$

i.e., $x_0 + nh = 1.8 \Rightarrow 2.0 + n(0.5) = 1.8$

$$n = \frac{1.8 - 2.0}{0.5} = -0.4$$

$$\therefore y(1.8) = 0.0540 + (-0.4)(-0.0755) + \frac{(-0.4)(-0.4+1)}{2!}(0.0370)$$

$$+ \frac{(-0.4)(-0.4+1)(-0.4+2)}{3!}(0.0394) + \text{negligible terms}$$

$$= 0.0540 + 0.0302 - 0.0044 - 0.00252 = 0.07728$$

Hence $\boxed{f(1.8) = \mathbf{0.07728}}$

Example 2

Given

x	1	2	3	4	5	6	7	8
$f(x)$	1	8	27	64	125	216	343	512

Estimate $f(7.5)$. Use Newton's formula.

Solution

The difference table is as given below.

x	$y = f(x)$	$\nabla f(x)$	$\nabla^2 f(x)$	$\nabla^3 f(x)$	$\nabla^4 f(x)$
1	1				
		7			
2	8		12		
		19		6	
3	27		18		0
		37		6	
4	64		24		0
		61		6	
5	125		30		0
		91		6	
6	216		36		0 ($\nabla^4 x_0$)
		127		6 ($\nabla^3 x_0$)	
7	343		42 ($\nabla^2 x_0$)		
		167 (∇x_0)			
8 (x_0)	512 (y_0)				

Here $x_0 = 8$, $y_0 = 512$, $h = 1$ and $x = 7.5$.

Newton's Backward interpolation formula is

$$\boxed{y(x_0 + nh) = y_0 + n\nabla y_0 + \frac{n(n+1)}{2!}\nabla^2 y_0 + \ldots}$$

$$y(7.5) = ?$$

i.e., $x_0 + nh = 7.5 \Rightarrow 8 + n(1) = 7.5$

$$n = 7.5 - 8 = -0.5$$

$$\boxed{n = -0.5}$$

$$\therefore y(7.5) = 512 + (-0.5)(169) + \frac{(-0.5)(-0.5+1)}{2!}(42)$$

$$+ \frac{(-0.5)(-0.5+1)(-0.5+\ldots}{3!}$$

$$= 512 - 84.5 + (-0.5)(0.5)(21) + \frac{(-0.5)(0.5)(1.5\ldots}{6}$$

$$= 512 - 84.5 - 5.25 - 0.375$$

$$f(7.5) = 421.875$$

$$\therefore \boxed{f(7.5) = 421.875}$$

Example 3

Result 4 : The n^{th} divided differences of a polynomial of n^{th} degree are constant.

❑ NEWTON'S DIVIDED DIFFERENCE FORMULA

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

$$f(x, x_0) = \frac{f(x) - f(x_0)}{x - x_0}$$

i.e., $$f(x) = f(x_0) + (x - x_0) f(x, x_0)$$

Now $$f(x, x_0, x_1) = \frac{f(x, x_0) - f(x_0, x_1)}{x - x_1}$$

$$f(x, x_0) = f(x_0, x_1) + (x - x_1) f(x, x_0, x_1)$$

Substituting (2) in (1), we get,

$$f(x) = f(x_0) + (x - x_0) \{f(x_0, x_1) + (x - x_1) f(x, x_0, x_1)\}$$
$$= f(x_0) + (x - x_0) f(x_0, x_1) + (x - x_0)(x - x_1) f(x, x_0, x_1)$$

Now, $$f(x, x_0, x_1, x_2) = \frac{f(x, x_0, x_1) - f(x_0, x_1, x_2)}{x - x_2}$$

i.e., $$f(x, x_0, x_1) = f(x_0, x_1, x_2) + (x - x_2) f(x, x_0, x_1, x_2)$$

Substituting (4) in (3)

$$f(x) = f(x_0) + (x - x_0) f(x_0, x_1) + (x - x_0)(x - x_1) f(x_0, \ldots$$
$$+ (x - x_0)(x - x_1)(x - x_2) f(x, x_0, \ldots$$

Proceeding in this manner, we get,

$$f(x) = y_0 + (x - x_0)\, f(x_0, x_1) + (x - x_0)(x - x_1)\, f(x_0, x_1, x_2) + (x - x_0)(x - x_1)(x - x_2)\, f(x, x_0, x_1, x_2) + \ldots + (x - x_0)(x - x_1) \ldots (x - x_n)\, f(x, x_0, x_1, \ldots, x_n)$$

which is called **Newton's divided difference interpolation formula for unequal intervals.**

The above formula can also be written as

$$y = f(x) = y_0 + (x - x_0) \Delta f(x_0) + (x - x_0)(x - x_1) \Delta^2 f(x_0) + (x - x_0)(x - x_1)(x - x_2) \Delta^3 f(x_0) + \ldots \quad \ldots (A)$$

Note : Let $x_0, x_1, x_2, \ldots x_n$ are equally spaced such that $x_1 - x_0$, $x_2 - x_1, \ldots, x_{n-1} - x_n = h$.

Let $x = x_0 + nh$

$$\text{Now, } f(x_0, x_1) = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{y_1 - y_0}{h} = \frac{1}{h} \Delta y_0 \quad \ldots (1)$$

$$f(x_0, x_1, x_2) = \frac{f(x_1, x_2) - f(x_0, x_1)}{x_2 - x_0}$$

$$= \frac{\dfrac{f(x_2) - f(x_1)}{x_2 - x_1} - \dfrac{1}{h} \Delta y_0}{2h}$$

$$= \frac{y_2 - y_1 - \Delta y_0}{2h^2} = \frac{\Delta y_1 - \Delta y_0}{2h^2}$$

$$= \frac{\Delta^2 y_0}{2!\, h^2} \quad \ldots (2)$$

Substituting (1), (2) in (A), we get,

$$y(x_0 + nh) = y_0 + nh \cdot \frac{1}{h} \Delta y_0 + nh \frac{(n-1)h}{2!\, h^2} \Delta^2 y_0 + \ldots$$

$$= y_0 + n \Delta y_0 + \frac{n(n-1)}{2!} \Delta^2 y_0 + \ldots$$

which is **Newton's forward interpolation formula for equal intervals.**

Given $\log_{10} 654 = 2.8156$, $\log_{10} 658 = 2.8182$, $\log_{10} 659 = 2.8189$ and $\log_{10} 661 = 2.8202$. Find the value of $\log_{10} 656$ using Newton's divided difference formula. [OR] the following table find f (656).

x	654	658	659	661
y	2.8156	2.8182	2.8189	2.8202

Solution

x	$y = f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$
654	2.8156			
		$\frac{2.8181 - 2.8156}{658 - 654} = 0.00065$		
658	2.8182		$\frac{0.0007 - 0.00065}{659 - 654} = 0.00001$	
		$\frac{2.8181 - 2.8182}{659 - 658} = 0.00070$		$\frac{-0.000016 - 0.0\ldots}{661 - 6\ldots} = -0\ldots$
659	2.8189		$\frac{0.00065 - 0.0007}{661 - 658} = -0.000016$	
		$\frac{2.8202 - 2.8189}{661 - 659} = 0.00065$		
661	2.8202			

By Newton's divided difference formula we have

$$f(x) = y_0 + (x - x_0)\Delta f(x_0) + (x - x_0)(x - x_1)\Delta^2 \ldots$$
$$+ (x - x_0)(x - x_1)(x - x_2)\Delta^3 f(x_0)$$

$f(656) = ?$

Here $x_0 = 654$, $y_0 = 2.8156$, $x_1 = 658$, $x_2 = 659$,

$\Delta f(x_0) = 0.00065$, $\Delta^2 f(x_0) = 0.00001$, $\Delta^3 f(x_0) = -0.0\ldots$

Substituting these values in (1) we get,

$$f(656) = 2.8156 + (656-654)(0.00065) + (656-654)(656-658)(0.00001) + (656-654)(656-658)(656-659)(-0.0000037)$$

$$f(656) = 2.8168156$$

Example 2

Find f(1), f(5) and f(9) using Newton's divided difference formula from the following data.

x	0	2	3	4	7	8
$y = f(x)$	4	26	58	112	466	668

Solution

$y = f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$	$\Delta^4 f(x)$
4 y_0				
	$\frac{26-4}{2-0} = 11$ $\Delta f(x_0)$			
26		$\frac{32-11}{3-0} = 7$ $\Delta^2 f(x_0)$		
	$\frac{58-26}{3-2} = 32$		$\frac{11-7}{4-0} = 1$ $\Delta^3 f(x_0)$	
58		$\frac{54-32}{4-2} = 11$		$\frac{1-1}{7-0} = 0$ $\Delta^3 f(x_0)$
	$\frac{112-58}{4-3} = 54$		$\frac{16-11}{7-2} = 1$	
112		$\frac{118-54}{7-3} = 16$		$\frac{1-1}{8-2} = 0$
	$\frac{466-112}{7-4} = 118$		$\frac{21-16}{8-3} = 1$	
466		$\frac{202-118}{8-4} = 21$		
	$\frac{668-466}{8-7} = 202$			
668				

Newton's divided difference formula is

$$y = f(x) = y_0 + (x-x_0)\Delta f(x_0) + (x-x_0)(x-x_1)\Delta^2 f(x_0) + (x-x_0)(x-x_1)(x-x_2)\Delta^3 f(x_0) \quad \dots (1)$$

Here $x_0 = 0, y_0 = 4, x_1 = 2, x_2 = 3.$

$\Delta f(x_0) = 11,\ \Delta^2 f(x_0) = 7,\ \Delta^3 f(x_0) = 1$

Substituting these values in (1) we get,

$$f(x) = 4 + (x-0)\,11 + x\,(x-0)\,(x-2)(7)$$

$$+ (x-0)\,(x-$$

$$= 4 + 11x + (x^2 - 2x)7 + x\,(x^2 - 5x + 6)(1)$$

$$= x^3 + 7x^2 - 5x^2 + 11x + 6x - 14x + 4$$

$$f(x) = x^3 + 2x^2 + 3x + 4$$

Substituting $x = 1$, in (2) we get

$$f(1) = (1)^3 + 2(1)^2 + 3(1) + 4 = 1 + 2 + 3 + 4 =$$

Substituting $x = 5$, in (2) we get

$$f(5) = (5)^3 + 2(5)^2 + 3(5) + 4 = 125 + 50 + 15 +$$

Substituting $x = 9$, in (2) we get

$$f(9) = (9)^3 + 2(9)^2 + 3(9) + 4 = 729 + 162 + 2$$

x	**1**	**5**	**9**
f(x)	**10**	**194**	**922**

INTERPOLATION

(FINITE DIFFERENCES)

❑ INTERPOLATION WITH UNEQUAL INTERVALS ❑

Lagrange's Interpolation Formula for unequal intervals

Let $f(x_0), f(x_1), \ldots, f(x_n)$ be the values of the function $= f(x)$ corresponding to the arguments $x_0, x_1, \ldots, x_n$, not necessarily equally spaced.

Let $f(x)$ be a polynomial in x of degree n. Then we can represent $f(x)$ as

$$f(x) = a_0(x-x_1)(x-x_2)\ldots(x-x_n)$$
$$+ a_1(x-x_0)(x-x_2)\ldots(x-x_n) + \ldots$$
$$+ a_n(x-x_0)(x-x_1)\ldots(x-x_{n-1}) \qquad \ldots (1)$$

where $a_0, a_1, \ldots, a_n$ are constants.

Now we have to determine the $(n+1)$ constants $a_0, a_1, \ldots a_n$.

Putting $x = x_0$ in (1), we get

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

i.e.,
$$a_0 = \frac{f(x_0)}{(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n)} \qquad \ldots (2)$$

Putting $x = x_1$ in (1), we get

$$f(x_1) = a_1(x_1-x_0)(x_1-x_2)\ldots(x_1-x_n)$$

$$\text{i.e.,} \quad a_1 = \frac{f(x_1)}{(x_1 - x_0)(x_1 - x_2) \dots (x_1 - x_n)}$$

Similarly

$$a_2 = \frac{f(x_2)}{(x_2 - x_0)(x_2 - x_1) \dots (x_2 - x_n)}$$

$$\vdots \qquad \vdots$$

$$a_n = \frac{f(x_n)}{(x_n - x_0)(x_n - x_1) \dots (x_n - x_{n-1})}$$

Substituting (2), (3), (4), (5) in (1), we get

$$f(x) = \frac{(x - x_1)(x - x_2) \dots (x - x_n)}{(x_0 - x_1)(x_0 - x_2) \dots (x_0 - x_n)} f(x_0) + \frac{(x - x_0)(x - x_2) \dots (x - x_n)}{(x_1 - x_0)(x_1 - x_2) \dots (x_1 - x_n)} f(x_1) + \dots + \frac{(x - x_0)(x - x_1) \dots (x - x_{n-1})}{(x_n - x_0)(x_n - x_1) \dots (x_n - x_{n-1})} f(x_n)$$

If we denote $f(x_0), f(x_1), \dots f(x_n)$ by $y_0, y_1, \dots, y_n$ we get

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

which is **Lagrange's interpolation** formula.

Using Lagrange's formula to calculate $f(3)$ from the following table. [A.U. Nov., Dec. 2007]

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

Solution

Here $x_0 = 0, \quad x_1 = 1, \quad x_2 = 2, \quad x_3 = 4, \quad x_4 = 5, \quad x_5 = 6$

$y_0 = 1, \quad y_1 = 14, \quad y_2 = 15, \quad y_3 = 5, \quad y_4 = 6, \quad y_5 = 19$

We know that Lagrange's formula is

$$y(x) = \frac{(x-x_1)(x-x_2)(x-x_3)(x-x_4)(x-x_5)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)(x_0-x_4)(x_0-x_5)} y_0$$

$$+ \frac{(x-x_0)(x-x_2)(x-x_3)(x-x_4)(x-x_5)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)(x_1-x_4)(x_1-x_5)} y_1$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_3)(x-x_4)(x-x_5)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)(x_2-x_4)(x_2-x_5)} y_2$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_2)(x-x_4)(x-x_5)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)(x_3-x_4)(x_3-x_5)} y_3$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_2)(x-x_3)(x-x_5)}{(x_4-x_0)(x_4-x_1)(x_4-x_2)(x_4-x_3)(x_4-x_5)} y_4$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_2)(x-x_3)(x-x_4)}{(x_5-x_0)(x_5-x_1)(x_5-x_2)(x_5-x_3)(x_5-x_4)} y_5$$

$$
\begin{aligned}
y(x) = {} & \frac{(x-1)(x-2)(x-4)(x-5)(x-6)}{(0-1)(0-2)(0-4)(0-5)(0-6)}\,(1) \\
& + \frac{(x-0)(x-2)(x-4)(x-5)(x-6)}{(1-0)(1-2)(1-4)(1-5)(1-6)} \\
& + \frac{(x-0)(x-1)(x-4)(x-5)(x-6)}{(2-0)(2-1)(2-4)(2-5)(2-6)} \\
& + \frac{(x-0)(x-1)(x-2)(x-5)(x-6)}{(4-0)(4-1)(4-2)(4-5)(4-6)} \\
& + \frac{(x-0)(x-1)(x-2)(x-4)(x-6)}{(5-0)(5-1)(5-2)(5-4)(5-6)} \\
& + \frac{(x-0)(x-1)(x-2)(x-4)(x-5)}{(6-0)(6-1)(6-2)(6-4)(6-5)}
\end{aligned}
$$

$$
\begin{aligned}
y(x) = {} & \frac{(x-1)(x-2)(x-4)(x-5)(x-6)}{(-1)(-2)(-4)(-5)(-6)} \\
& + (14)\,\frac{(x)(x-2)(x-4)(x-5)(x-6)}{(1)(-1)(-3)(-4)(-5)} \\
& + \frac{(x)(x-1)(x-4)(x-5)(x-6)}{(2)(1)(-2)(-3)(-4)} \\
& + \frac{(x)(x-1)(x-2)(x-5)(x-6)}{(4)(3)(2)(-1)(-2)} \\
& + \frac{(x)(x-1)(x-2)(x-4)(x-6)}{(5)(4)(3)(1)(-1)} \\
& + \frac{(x)(x-1)(x-2)(x-4)(x-5)}{6(5)(4)(2)(1)}
\end{aligned}
$$

Substitute $x = 3$ we get

$$y(3) = -\frac{(3-1)(3-2)(3-4)(3-5)(3-6)}{240}$$

$$+\frac{(3)(3-2)(3-4)(3-5)(3-6)}{60}(14)$$

$$-\frac{(3)(3-1)(3-4)(3-5)(3-6)}{48}(15)$$

$$+\frac{(3)(3-1)(3-2)(3-5)(3-6)}{48}(5)$$

$$-\frac{(3)(3-1)(3-2)(3-4)(3-6)}{60}(6)$$

$$+\frac{(3)(3-1)(3-2)(3-4)(3-5)}{240}(19)$$

$$= -\frac{(2)(1)(-1)(-2)(-3)}{240} + \frac{3(1)(-1)(-2)(-3)(14)}{60}$$

$$-\frac{3(2)(-1)(-2)(-3)(15)}{48} + \frac{3(2)(1)(-2)(-3)(5)}{48}$$

$$-\frac{3(2)(1)(-1)(-3)(6)}{60} + \frac{3(2)(1)(-1)(-2)(19)}{240}$$

$$= \frac{12}{240} - \frac{252}{60} + \frac{540}{48} + \frac{180}{48} - \frac{108}{60} + \frac{228}{240}$$

$$= \frac{12 - 1008 + 2700 + 900 - 432 + 228}{240} = \frac{2400}{240}$$

$y(3) = 10$

Example 2

Using Lagrange's interpolation formula, find the value corresponding to $x = 10$ from the following table :

x	5	6	9	11
y	12	13	14	16

Solution

The Lagrange's interpolation formula is

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

Here $x = 10,\quad x_0 = 5,\quad x_1 = 6,\quad x_2 = 9,\quad x_3$

$y_0 = 12,\quad y_1 = 13,\quad y_2 = 14\quad y_3 = 16$

$$f(10) = \frac{(10-6)(10-9)(10-11)}{(5-6)(5-9)(5-11)}(12) + \frac{(10-5)(10-9)(10-11)}{(6-5)(6-9)(6-11)}(13) + \frac{(10-5)(10-6)(10-11)}{(9-5)(9-6)(9-11)}(14) + \frac{(10-5)(10-6)(10-9)}{(11-5)(11-6)(11-9)}(16)$$

$$= \frac{4\cdot 1\cdot 1}{1\cdot 4\cdot 6}(12) - \frac{5\cdot 1\cdot 1}{1\cdot 3\cdot 5}(13) + \frac{5\cdot 4\cdot 1}{4\cdot 3\cdot 2}(14) + \frac{5\cdot 4\cdot}{6\cdot 5\cdot}$$

$$\boxed{f(10) = 14.63}$$

Example

data.