1. Introduction:

During experimental work, we take observations involving two variables as: pressure and temperature, load and deflection, voltage and current etc. and we want to establish relation between these variables.

The problem of finding an equation of an approximating curve, which passes through as many points as possible is called ***curve fitting***.

The method of group averages, the least square method, the method of moments etc. are some methods. Out of all methods, the least square method gives a unique best fit and is highly recommended.

1.1. The Least Square Method:

In curve fitting, for a given data, this method gives the best fit, with prior knowledge of the shape of the curve.

Suppose we have (x_1, y_1), (x_2, y_2) …. (x_n, y_n) be n set of observations in any experiment.

Let us assume that

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

is a relationship between x and y.

When $x = x_1$, the observed value of y is y_1 and the expected or calculated value of y using equation (1) is $f(x_1)$. Which may be slightly differ from y_1. Then the residual is defined as:

$$e_1 = y_1 - f(x_1) \quad (2)$$

Similarly, all other residuals $e_2, e_3 \dots e_n$ defined as:

$$e_2 = y_2 - f(x_2)\,, e_3 = y_3 - f(x_3) \dots\ e_n = y_n - f(x_n) \quad (3)$$

Here some residuals may be positive, some may be negative and some may be zero. In order to give equal importance to both positive and negative residuals, we consider the sum of the squares of residuals (rather than sum of residuals as in the method of group averages). Therefore

$$E = e_1^2 + e_2^2 + \ \dots + e_n^2 = [y_1 - f(x_1)]^2 + [y_2 - f(x_2)]^2 + \cdots + \ [y_n - f(x_n)]^2 \quad (4)$$

Here the quantity E is a measure of how well the curve $y = f(x)$ fits the given data (observations).

Therefore E will be zero if and only if all points of observations lie on the curve given by equation (1). The value of E decreases depending on the closeness of the observed data to the curved assumed. Hence “the best representative curve to the given set of the observed data or observations is one for which E, the sum of the squares of the residuals, is minimum”. This concept is known as ***the principle of least squares.*** In general, we consider straight line, parabola or an exponential curve for fitting the data.

1.2. Fitting a straight line:

Suppose $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ be a set of n observations in any experiment and we wish to fit a straight line to these observations, which is represented as:

$$y = ax + b \quad (5)$$

i.e. we have to determine constants a and b, using the principle of least squares.

For any x_i $(i = 1, 2, \dots n)$, the expected value (computed value) of y is $ax_i + b$, while the observed value of y is y_i.

Hence the residuals are:

$$e_i = observed\ value - expected\ or\ calculated\ value$$

$$e_i = y_i - (ax_i + b)\ for\ i = 1, 2, \dots n \quad (6)$$

The squares: $e_i^2 = [y_i - (ax_i + b)]^2$,

and the sum of the squares of residuals is given as:

$$E = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2 \quad (7)$$

where E is a function of parameters a and b. The necessary condition for E to be minimum gives:

$$\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = 0 \quad (8)$$

The first condition of equation (8) $\frac{\partial E}{\partial a} = 0$, using equation (3) gives

$$\frac{\partial E}{\partial a} = \frac{\partial}{\partial a}\left[\sum_{i=1}^{n} [y_i - (ax_i + b)]^2\right] = 0$$

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

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

$$\sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i = 0$$

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \quad (9)$$

Similarly the second term of equation (8) $\frac{\partial E}{\partial b} = 0$, using equation (3) gives

$$\frac{\partial E}{\partial b} = \frac{\partial}{\partial b}\left[\sum_{i=1}^{n}[y_i - (ax_i + b)]^2\right] = 0$$

$$2\left[\sum_{i=1}^{n}[y_i - (ax_i + b)]\,(-1)\right] = 0$$

$$\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i - b\sum_{i=1}^{n} 1 = 0$$

As

$$\sum_{i=1}^{n} 1 = n$$

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i \qquad (10)$$

Equations (9) and (10) are called normal equations.

To compute E, equation (7) is written as

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

$$E = \sum_{i=1}^{n}[y_i^2 + a^2x_i^2 + b^2 - 2ax_iy_i + 2abx_i - 2by_i]$$

By rearranging the terms

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

$$E = \sum_{i=1}^{n}[y_i^2 - ax_iy_i - by_i - ax_iy_i - by_i + (ax_i + b)^2]$$

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

As $y_i - (ax_i + b) = 0$, the last term of above equation becomes zero

$$\therefore\ E = \sum_{i=1}^{n} y_i^2 - a\sum_{i=1}^{n} x_iy_i - b\sum_{i=1}^{n} y_i \qquad (11)$$

Example: 1

Using the method of least squares, find the straight line $y = ax + b$ that fits the following data.

x	0.5	1.0	1.5	2.0	2.5	3.0
y	15	17	19	14	10	7

Solution: the normal equations of least square fitting that fits a straight line $y = ax + b$ are:

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

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i \ \dots(2)$$

x	y	xy	x^2
0.5	15	7.5	0.25
1.0	17	17.0	1.00
1.5	19	28.5	2.25
2.0	14	28.0	4.00
2.5	10	25.0	6.25
3.0	7	21.0	9.00
$\sum x_i = 10.5$	$\sum y_i = 82$	$\sum x_i y_i = 127$	$\sum x_i^2 = 22.75$

The normal equations (1) and (2) becomes

$$22.75a + 10.5b = 127 \quad \dots(3)$$

$$10.5a + 6b = 82 \dots(4)$$

From equation (4)

$$b = \frac{82 - 10.5a}{6} \ \dots(5)$$

Submit value of equation (5) in equation (3), it becomes …

$$22.75a + 10.5\left(\frac{82 - 10.5a}{6}\right) = 127$$

$$22.75a + 143.5 - 18.375a = 127$$

$$4.375a = -16.5 \ \therefore a = -3.7714 \ \dots(6)$$

Now submit value of equation (6) in equation (5), we have

$$b = \frac{82 - 10.5(-3.7714)}{6} \ \therefore \ b = 20.2667 \ \dots(7)$$

Therefore, equation of line $y = ax + b$ now becomes $\boldsymbol{y = -3.7714x + 20.2667}$

Example: 2 Applying the method of least square find an equation of the form $y = ax + bx^2$ that fits the following data.

x	1	2	3	4	5	6
y	2.6	5.4	8.7	12.1	16.0	20.2

Solution: The required curve that fits the given data is $y = ax + bx^2$, which can be written as $\frac{y}{x} = a + bx$ and by taking $Y = \frac{y}{x}$ it becomes $Y = a + bx$. We can rewrite the data of given table for new variable $Y = \frac{y}{x}$.

x	1	2	3	4	5	6
Y	2.6	2.7	2.9	3.025	3.2	3.367

The corresponding normal equations become:

$$b\sum_{i=1}^{n} x_i^2 + a\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i Y_i \quad \dots (1)$$

$$b\sum_{i=1}^{n} x_i + an = \sum_{i=1}^{n} Y_i \quad \dots (2)$$

From modified data table, we have

x	Y	xY	x^2
1	2.6	2.6	1
2	2.7	5.4	4
3	2.9	8.7	9
4	3.025	12.1	16
5	3.2	16.0	25
6	3.367	20.2	36
$\sum x_i = 21$	$\sum Y_i = 17.792$	$\sum x_i Y_i = 65.0$	$\sum x_i^2 = 91$

Equation (1) and (2) becomes: $91b + 21a = 65 \dots$ (3) $\quad 21b + 6a = 17.792 \dots$ (4)

From equation (4), we have $a = \frac{17.792 - 21b}{6} \dots$ (5)

By substituting value of equation (5) in equation (3) we obtain…

$91b + 21\left(\frac{17.792-21b}{6}\right) = 65 \qquad \therefore 91b + 62.272 - 73.5b = 65 \qquad \therefore 17.5b = 2.728$

$$b = 0.15589 \dots (6)$$

Equation (5) becomes… $a = \frac{17.792-21(0.15589)}{6} = 2.41973 \qquad \therefore a = 2.41973 \dots$ (7)

$Y = a + bx$ becomes $Y = 2.41973 + 0.15589x \dots$ (8)

As $Y = \frac{y}{x}$, equation (8) is finally written as $\boldsymbol{y = 2.41973x + 0.15589x^2}$

Example: 3 Using the method of least squares, find an equation of the form: $y = ax + b$, that fits the following data. Also, calculate the sum of the squares of the residuals E.

x	0	1	2	3	4
y	1	5	10	22	38

Solution: the normal equations of least square fitting that fits a straight line $y = ax + b$ are:

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \quad \dots(1), \qquad a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i \quad \dots(2)$$

x	y	xy	x^2
0	1	0	0
1	5	5	1
2	10	20	4
3	22	66	9
4	38	152	16
$\sum x_i = 10$	$\sum y_i = 76$	$\sum x_i y_i = 243$	$\sum x_i^2 = 30$

The normal equations (1) and (2) becomes

$$30a + 10b = 243 \quad \dots(3)$$

$$10a + 5b = 76 \dots(4)$$

From equation (4)

$$b = \frac{76 - 10a}{5} \quad \dots(5)$$

Submit value of equation (5) in equation (3), it becomes …

$$30a + 10\left(\frac{76 - 10a}{5}\right) = 243, \qquad 30a + 152 - 20a = 243, \qquad \therefore\ 10a = 91$$

$$\therefore a = 9.1 \quad \dots(6)$$

Now submit value of equation (6) in equation (5), we have

$$b = \frac{76 - 10(9.1)}{5}, \qquad \therefore\ b = -3 \quad \dots(7)$$

Therefore, equation of line $y = ax + b$ finally becomes $\boldsymbol{y = 9.1x - 3}$

Example: 4 Using the method of least squares, find an equation of the form: $y = ax + b$, that fits the following data.

x	-2	-1	0	1	2
y	1	2	3	3	4

1.3. Fitting a parabola:

Suppose we have a set of 'n' observations $(x_i, y_i), i = 1, 2, \ldots n$ in any experiment and we wish to fit a parabola to the observed data using least square method.

Let

$$y = ax^2 + bx + c \quad (1)$$

be the equation of parabola that fits the given data. We want to determine constants a, b, c.

For a given x_i, let the expected value of y is given by $ax_i^2 + bx_i + c$, while the observed value is y_i. Then the residual e_i in the i[th] value is given by

$$e_i = observed\ value - expected\ or\ calculated\ value$$

$$e_i = y_i - (ax_i^2 + bx_i + c) \quad (2)$$

Summation of all n residual squares gives

$$E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} [y_i - (ax_i^2 + bx_i + c)]^2 \quad (3)$$

where E is a function of parameters a, b, c. The necessary condition for E to be minimum gives:

$$\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = \frac{\partial E}{\partial c} = 0 \quad (4)$$

As $\frac{\partial E}{\partial a} = 0$

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

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

By expansion of summation and rearranging the terms we obtain…

$$a \sum_{i=1}^{n} x_i^4 + b \sum_{i=1}^{n} x_i^3 + c \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i^2 y_i \quad (5)$$

Similarly $\frac{\partial E}{\partial b} = 0$ gives…

$$a \sum_{i=1}^{n} x_i^3 + b \sum_{i=1}^{n} x_i^2 + c \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \quad (6)$$

and $\frac{\partial E}{\partial c} = 0$ gives…

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^2 + cn = \sum_{i=1}^{n} x_i y_i \quad (7)$$

Equations (5), (6) and (7) are called normal equations for parabola and we find out constants a, b, c by solving them. Equation (3) becomes:

$$E = \sum_{i=1}^{n} y_i^2 - a\sum_{i=1}^{n} x_i^2 y_i - b\sum_{i=1}^{n} x_i y_i - c\sum_{i=1}^{n} y_i \quad (8)$$

Example: 5 Fit a parabola to the following data using the method of least squares.

x	1.0	1.2	1.4	1.6	1.8	2.0
y	0.98	1.40	1.86	2.55	2.28	3.20

Solution: Let the equation of the curve of parabolic fit is

$$y = ax^2 + bx + c \ldots \ (1)$$

Corresponding normal equations are

$$a\sum_{i=1}^{n} x_i^4 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i^2 y_i \ldots (2)$$

$$a\sum_{i=1}^{n} x_i^3 + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \ldots (3)$$

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^2 + cn = \sum_{i=1}^{n} x_i y_i \ldots (4)$$

x	y	x^2	x^3	x^4	xy	x^2y
1.0	0.98	1.00	1.000	1.0000	0.980	0.9800
1.2	1.40	1.44	1.728	2.0736	1.680	2.0160
1.4	1.86	1.96	2.744	3.8416	2.604	3.6456
1.6	2.55	2.56	4.096	6.5536	4.080	6.5280
1.8	2.28	3.24	5.832	10.4976	4.104	7.3872
2.0	3.20	4.00	8.000	16.0000	6.400	12.8000
$\sum_{i=1}^{6} x_i = 9$	$\sum_{i=1}^{6} y_i = 12.27$	$\sum_{i=1}^{6} x_i^2 = 14.20$	$\sum_{i=1}^{6} x_i^3 = 23.4$	$\sum_{i=1}^{6} x_i^4 = 39.9664$	$\sum_{i=1}^{6} x_i y_i = 19.848$	$\sum_{i=1}^{6} x_i^2 y_i = 33.3568$

Equations (2), (3), (4) become

$$39.9664a + 23.4b + 14.2c = 33.3568 \ldots \ (5)$$

$$23.4a + 14.2b + 9.0c = 19.848 \ldots \ (6)$$

$$14.2a + 9.0b + 6c = 12.27 \ldots \ (7)$$

Multiply equation (6) with 9 and equation (7) with 14.2, by subtracting the results

$$8.96a - 4.2c = 4.398$$

$$\therefore a = \frac{4.398 + 4.2c}{8.96} \dots (8)$$

Multiply equation (6) with 23.4 and equation (5) with 14.2, by subtracting the results

$$8.96a + 12.6c = 5.2764$$

$$\therefore b = \frac{5.2764 - 12.6c}{8.96} \dots (9)$$

Substituting values of equation (8) and (9) in equation (5), we obtain

$$39.9664\left(\frac{4.398 + 4.2c}{8.96}\right) + 23.4\left(\frac{5.2764 - 12.6c}{8.96}\right) + 14.2c = 33.3568$$

$$c = -1.4464 \; or \; (-1.4471) \dots (10)$$

$$a = -0.1875 \dots (11)$$

$$b = 2.6239 \dots (12)$$

Equation (1) becomes ... $\boldsymbol{y = -0.1875x^2 + 2.6239x - 1.4471}$

Which is required equation.

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

Suppose we have a set of 'n' observations $(x_i, y_i), i = 1, 2, \dots n$ in any experiment and we wish to fit a curve of type $y = ax^b$ to the observed data using least square method. Given that

$$y = ax^b \quad (1)$$

To linearize it take logarithms on both sides of $y = ax^b$

$$log_{10}y = log_{10}a + b \cdot log_{10}x \quad (2)$$

Let $log_{10}y = Y,\; log_{10}a = A, log_{10}x = X \quad (3)$

Therefore equation (2) now becomes ... $Y = A + bX \quad (4)$

Which is required equation, linear in Y and X.

Example: 6 Using the method of least squares, find a relation of the form $y = ax^b$ that fits the given data.

x	2	3	4	5
y	27.8	62.1	110	161

Solution: Let the equation fits is

$$y = ax^b \dots (1)$$

After linearize with $log_{10}y = Y,\ log_{10}a = A, log_{10}x = X$, it becomes

$$Y = A + bX \dots\ (2)$$

Therefore, data in Table 1 is now modified and written in terms of $X = log_{10}x, Y = log_{10}y$

X	0.3010	0.4771	0.6021	0.6990
Y	1.4440	1.7930	2.0414	2.2068

Corresponding normal equations are

$$b\sum_{i=1}^{n} X_i^2 + A\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} X_iY_i \ \dots (3), \qquad b\sum_{i=1}^{n} X_i + An = \sum_{i=1}^{n} Y_i \ \dots (4)$$

X	Y	XY	X^2
0.3010	1.4440	0.4346	0.0906
0.4771	1.7931	0.8555	0.2276
0.6021	2.0414	1.2291	0.3625
0.6990	2.2068	1.5426	0.4886
$\sum X_i = 2.0792$	$\sum Y_i = 7.4853$	$\sum X_i Y_i = 4.0618$	$\sum X_i^2 = 1.1693$

Equations (3) and (4) become …

$$1.1693b + 2.0792A = 4.0618 \dots\ (5), \qquad 2.0792b + 4A = 7.4853 \dots\ (6)$$

Which give $b = 1.9311$ and $A = 0.8678$

as $A = log_{10}a, a = antilog\ (A) = antilog(0.8678) = 7.375$

and the finally equation (1) $y = ax^b$ becomes $\boldsymbol{y = 7.375x^{1.9311}}$

Example: 7 Using the method of least squares, find a relation of the form $y = ax^b$ that fits the given data.

x	1	2	3	4	5
y	0.5	2.0	4.5	8.0	12.5

1.5. Fitting an exponential curve $\boldsymbol{y = ae^{bx}}$:

Suppose we have a set of 'n' observations $(x_i, y_i), i = 1, 2, \dots n$ in any experiment and we wish to fit an exponential curve of the form $y = ae^{bx}$ to the observed data using least square method. Given that

$$y = ae^{bx} \quad (1)$$

To linearize it take logarithms on both sides of $y = ae^{bx}$

$$log_{10}y = log_{10}a + bx \cdot log_{10}e \quad (2)$$

Let $log_{10}y = Y,\ log_{10}a = A,\ b\ log_{10}e = B \quad (3)$

Therefore equation (2) now becomes …

$$Y = A + Bx \text{ or } Y = Bx + A \ \ (4)$$

Which is required equation, linear in Y and X.

Example 6 Using the least square, fit an equation of the form $y = ae^{bx}$ to the given data.

x	1	2	3	4
y	1.65	2.70	4.50	7.35

Let

$$y = ae^{bx} \dots \ (1)$$

$$log_{10}y = log_{10}a + bx \cdot log_{10}e \dots \ (2)$$

$$log_{10}y = Y, \qquad log_{10}a = A, \qquad b\ log_{10}e = B \dots \ (3)$$

Therefore equation (2) now becomes …

$$Y = A + Bx \ \ or \ \ Y = Bx + A \dots \ (4)$$

The normalized equations are:

$$B\sum_{i=1}^{n} x_i^2 + A\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i Y_i \ \ \dots (5),$$

$$B\sum_{i=1}^{n} x_i + An = \sum_{i=1}^{n} Y_i \ \ \dots (6)$$

x	y	$Y = log_{10}y$	x^2	xY
1	1.65	0.2175	1	0.2175
2	2.70	0.4314	4	0.8628
3	4.50	0.6532	9	1.9596
4	7.35	0.8663	16	3.4652
$\sum x_i = 10$		$\sum Y_i = 2.1684$	$\sum x_i^2 = 30$	$\sum x_i Y_i = 6.5051$

Equations (5) and (6) become…

$$30B + 10A = 6.5051, \qquad 10B + 4A = 2.1684$$

$$B = 0.2168, A = 0, \qquad b = 0.4992, a = 1$$

Equation (1) $y = ae^{bx}$ becomes $y = e^{0.4992}$, Which is required equation.

Example: 8 Using the least square method, find a relation of the form $y = ae^{bx}$ that fits the given data.

x	77	100	185	239	285
y	2.4	3.4	7.0	11.1	19.6

Example: 9 Using the least square method, find a relation of the form $y = ae^{bx}$ that fits the given data.

x	1	2	3	4	5
y	0.6	1.9	4.3	7.6	12.6

2. Interpolation:

For a function $y = f(x)$, for a given table of values $(x_k, y_k), k = 1, 2, \dots n$, the process of estimating the value of y, for any intermediate value of x is called ***interpolation***. The method of computing the value of y, for a given value of x, lying outside the table of values of x is known as ***extrapolation***.

To compute trajectory of a rocket flight, we have to solve the Euler's dynamical equations of motion to compute its position and velocity vectors at specified times during the flight. To find out position and velocity vector at some intermediate times, interpolation technique is used.

2.1. Finite Difference Operators:

2.2. Forward Differences:

For a given table of values $(x_k, y_k), k = 1, 2, \dots n$, with equally-spaced intervals on X-axis (abscissas) of a function $y = f(x)$. The forward difference operator Δ is defined as …

$$\Delta y_i = y_{i+1} - y_i \quad i = 0, 1, 2, \dots n-1 \quad (1)$$

$$\Delta y_0 = y_1 - y_0, \quad \Delta y_1 = y_2 - y_1, \dots \quad \Delta y_{n-1} = y_n - y_{n-1} \quad (2)$$

These differences are called first differences of function y, are denoted by symbol Δy_i. Here Δ is called forward difference operator.

The differences of first differences are called 2nd differences (Δ^2), defined as

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0,$$

$$\Delta^2 y_1 = \Delta y_2 - \Delta y_1$$

In general…

$$\Delta^2 y_i = \Delta y_{i+1} - \Delta y_i \quad i = 0, 1, 2, \dots n-1 \quad (3)$$

Similarly

$$\Delta^r y_i = \Delta^{r-1} y_{i+1} - \Delta^{r-1} y_i$$

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
x_0	y_0				
		Δy_0			
x_1	y_1		$\Delta^2 y_0$		
		Δy_1		$\Delta^3 y_0$	
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$
		Δy_2		$\Delta^3 y_1$	
x_3	y_3		$\Delta^2 y_2$		
		Δy_3			
x_4	y_4				

2.2. Backward Differences:

For a given table of values $(x_k, y_k), k = 1, 2, \dots n$, with equally-spaced intervals on X-axis (abscissas) of a function $y = f(x)$.

The backward difference operator ∇ is defined as …

$$\nabla y_i = y_i - y_{i-1} \quad i = n, n-1, \dots 2, 1 \quad (1)$$

$$\nabla y_1 = y_1 - y_0, \quad \nabla y_2 = y_2 - y_1, \dots \quad \nabla y_n = y_n - y_{n-1} \quad (2)$$

These differences are called first differences of function y, are denoted by symbol ∇y_i. Here ∇ is called backward difference operator.

The differences of first differences are called 2nd differences (∇^2), defined as

$$\nabla^2 y_1 = \nabla y_1 - \nabla y_0,$$

$$\nabla^2 y_2 = \nabla y_2 - \nabla y_1$$

In general…

$$\nabla^2 y_i = \nabla y_i - \nabla y_{i-1} \quad i = n, n-1, \dots 2, 1 \quad (3)$$

Similarly

$$\nabla^r y_i = \nabla^{r-1} y_i - \nabla^{r-1} y_{i-1}$$

Or

$$\nabla^k y_i = \nabla^{k-1} y_i - \nabla^{k-1} y_{i-1}$$

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
x_0	y_0				
		∇y_1			
x_1	y_1		$\nabla^2 y_2$		
		∇y_2		$\nabla^3 y_3$	
x_2	y_2		$\nabla^2 y_3$		$\nabla^4 y_4$
		∇y_3		$\nabla^3 y_4$	
x_3	y_3		$\nabla^2 y_4$		
		∇y_4			
x_4	y_4				

2.3. Central Differences:

For a given table of values $(x_k, y_k), k = 1, 2, \ldots n$, with equally-spaced intervals on X-axis (abscissas) of a function $y = f(x)$. The central difference operator δ is defined as the average of the subscripts of the two members of the difference.

$$\delta y_{1/2} = y_1 - y_0 \,, \delta y_{3/2} = y_2 - y_1 \quad (1)$$

$$\delta y_i = y_{i+(1/2)} - y_{i-(1/2)} \quad (2)$$

Higher order differences are defined as..

$$\delta^2 y_i = \delta y_{i+(1/2)} - \delta y_{i-(1/2)} \quad (3)$$

$$\delta^n y_i = \delta^{n-1} y_{i+(1/2)} - \delta^{n-1} y_{i-(1/2)} \quad (4)$$

2.4. Shift Operator: *E*

Let $y = f(x)$ be a function of x, and let x takes consecutive values as $x, x + h, x + 2h,$ etc. the shift operator E is defined as

$$Ef(x) = f(x + h) \quad (1)$$

$$E^2 f(x) = E[Ef(x)] = E[f(x + h)] = f(x + 2h)$$

In general

$$E^n f(x) = f(x + nh)$$

Or in terms of new notations $y_x = f(x)$ then

$$E^n y_x = y_{x+nh}$$

for all real values of n.

If $y_o, y_1, \dots y_n$ are the consecutive values of the function y_x, then we can write

$$Ey_0 = y_1, E^2 y_0 = y_2, E^3 y_0 = y_3, \dots \ E^2 y_2 = y_4$$

The inverse shift operator E^{-1} is defined as

$$E^{-1} f(x) = f(x-h), E^{-2} f(x) = f(x-2h), \dots \ E^{-n} f(x) = f(x-nh)$$

2.5. Average Operator μ:

The average operator μ is defined as

$$\mu f(x) = \frac{1}{2}\left[f\left(x+\frac{h}{2}\right) + f\left(x-\frac{h}{2}\right)\right] = \frac{1}{2}\left[y_{x+(h/2)} + y_{x-(h/2)}\right]$$

2.6. Differential Operator D:

The differential operator $Df(x)$ is defined as

$$Df(x) = \frac{d}{dx} f(x) = f'(x)$$

$$D^2 f(x) = \frac{d^2}{dx^2} f(x) = f''(x)$$

2.7. Relations between Operators:

(1) Show that $\Delta = E - 1$ or $E = \Delta + 1$

$$\Delta y_x = y_{x+h} - y_x = E y_x - y_x = (E-1) y_x \quad \therefore \Delta = E - 1 \ or \ E = \Delta + 1$$

(2) Show that $\nabla = \frac{E-1}{E}$

$$\nabla y_x = y_x - y_{x-h} = y_x - E^{-1} y_x = (1 - E^{-1}) y_x = \left(1 - \frac{1}{E}\right) y_x = \left(\frac{E-1}{E}\right) y_x$$

$$\therefore \nabla = \frac{E-1}{E}$$

(3) Show that $\delta = E^{1/2} - E^{-1/2}$

$$\delta y_x = y_{x+(h/2)} - y_{x-(h/2)} = E^{1/2} y_x - E^{-1/2} y_x = \left(E^{1/2} - E^{-1/2}\right) y_x$$

$$\therefore \delta = E^{1/2} - E^{-1/2}$$

(4) Show that $\mu = \frac{1}{2}\left[E^{1/2} + E^{-1/2}\right]$

$$\mu y_x = \frac{1}{2}\left[y_{x+(h/2)} + y_{x-(h/2)}\right] = \frac{1}{2}\left[E^{1/2} y_x + E^{-1/2} y_x\right] = \frac{1}{2}\left[E^{1/2} + E^{-1/2}\right] y_x$$

$$\therefore \mu = \frac{1}{2}\left[E^{1/2} + E^{-1/2}\right]$$

(5) Show that $hD = \log E$

We have $Ey_x = y_{x+h} = f(x+h)$

Using Taylor series expansion we have …

$$Ey_x = Ef(x) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \cdots$$

$$= f(x) + hDf(x) + \frac{h^2}{2!}D^2 f(x) + \cdots$$

$$= \left[1 + \frac{hD}{1!} + \frac{h^2D^2}{2!} + \cdots\right] f(x)$$

$$= e^{hD} f(x)$$

$$\therefore E = e^{hD}$$

By taking log on both sides

$$hD = \log E$$

(6) Show that $hD = \log(1+\Delta) = -\log(1-\nabla)$

As $\Delta = E - 1, \quad E = \Delta + 1$

$$hD = \log E = \log(1+\Delta)$$

We know that $\log x = -\log x^{-1}$ and $\nabla = 1 - E^{-1}$, we have $E^{-1} = 1 - \nabla$

$$\therefore\ hD = \log E = -\log E^{-1} = \log(1-\nabla)$$

2.8. Newton's Forward Difference Interpolation Formula:

Let $y = f(x)$ be a function which takes values $f(x_0), f(x_0+h), f(x_0+2h)$ … corresponding to various equispaced values of x: with spacing h, say x_0, x_0+h, x_0+2h …

Suppose we wish to evaluate $f(x)$ for a value $x_0 + ph$,

where p is any real number $\left[as\ x = x_0 + ph,\ p = \frac{x-x_0}{h}\right]$.

From definition of shift operator E,

$$E^p f(x) = f(x+ph)$$

As $E = 1 + \Delta$, $\qquad E^p = (1+\Delta)^p$

$$\therefore f(x_0+ph) = E^p f(x_0) = (1+\Delta)^p f(x_0)$$

$$= \left[1 + p\Delta + \frac{p(p-1)}{2!}\Delta^2 + \frac{p(p-1)(p-2)}{3!}\Delta^3 + \cdots\right] f(x_0)$$

$$\therefore f(x_0+ph)=f(x_0)+p\Delta f(x_0)+\frac{p(p-1)}{2!}\Delta^2 f(x_0)+\frac{p(p-1)(p-2)}{3!}\Delta^3 f(x_0)$$

$$+\cdots+\frac{p(p-1)\dots(p-n+1)}{n!}\Delta^n f(x_0)+Error$$

In terms of $y=f(x)$ it is written as

$$y_x=y_0+p\Delta y_0+\frac{p(p-1)}{2!}\Delta^2 y_0+\frac{p(p-1)(p-2)}{3!}\Delta^3 y_0+\cdots+\frac{p(p-1)\dots(p-n+1)}{n!}\Delta^n y_0+Error$$

This formula is known as Newton Gregory forward difference formula for interpolation or simply Newton's forward difference formula for interpolation, which gives value of $f(x_0+ph)$ in terms of $f(x_0)$ and its leading differences. Or y_x in terms of y_0. This forward difference interpolation formula is mainly used to interpolating the values of y near the beginning of a set of tabular values and for extrapolating values of y, a short distance backward from y_0.

Example: 10 Evaluate $f(15)$ from the given data table.

x	10	20	30	40	50
$y=f(x)$	46	66	81	93	101

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
10	46				
		20			
20	66		-5		
		15		2	
30	81		-3		-3
		12		-1	
40	93		-4		
		8			
50	101				

$$\therefore y_x=y_0+p\Delta y_0+\frac{p(p-1)}{2!}\Delta^2 y_0+\frac{p(p-1)(p-2)}{3!}\Delta^3 y_0+\frac{p(p-1)(p-2)(p-3)}{4!}\Delta^4 y_0$$

$$x_0=10, y_0=46, \Delta y_0=20, \Delta^2 y_0=-5, \Delta^3 y_0=2, \Delta^4 y_0=-3$$

We want to evaluate $y_x \; for \; x=15. (i.e. y_{15})$. As $h=10, p=\frac{x-x_0}{h}=\frac{15-10}{10}=\frac{5}{10}=0.5$

$$y_{15}=46+0.5(20)+\frac{0.5(-0.5)}{2!}(-5)+\frac{0.5(-0.5)(-1.5)}{3!}(2)+\frac{0.5(-0.5)(-1.5)(-2.5)}{4!}(-3)$$

$$\therefore y_{15}=46+10+0.625+0.125+0.1172$$

$$y_{15}=f(15)=56.8672$$

Example: 11 Find Newton's forward difference interpolating polynomial for the following data. Estimate $f(0.15)$.

x	0.1	0.2	0.3	0.4	0.5
$y = f(x)$	1.40	1.56	1.76	2.00	2.28

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

$$x_0 = 0.1,\ \ y_0 = 1.40,\ \ \Delta y_0 = 0.16,\ \ \Delta^2 y_0 = 0.04,\ \ \mathrm{p} = \frac{x - 0.1}{0.1} = 10x - 1$$

$$\boldsymbol{y = 2x^2 + x + 1.28}$$

For $x = 0.15, \boldsymbol{y = 1.475}$

Theorem:

Differences of a polynomial:

The n^{th} differences of a polynomial of degree n is constant, when the values of the independent variable are given at equal intervals.

For $y_x = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n$, where $a_0 \neq 0, a_1, a_2, \ldots, a_n$ are constants, then

$$\Delta^n y_x = a_0(n!)h^n = constant$$

$$\Delta^{n+1} y_x = 0$$

Example: 12 Estimate the missing figure in the following table:

x	1	2	3	4	5
$y = f(x)$	2	5	7	-	32

Since we are given four entries in the table, the function $y = f(x)$ can be represented by a polynomial of degree three.

Using above theorem,

$$\Delta^3 f(x) = constant, \qquad \Delta^4 f(x) = 0$$

$$\therefore\ \Delta^4 f(x_0) = 0$$

As $\Delta = E - 1, \Delta^4 = (E-1)^4$ & $\Delta^4 f(x_0) = (E-1)^4\ f(x_0) = 0$

$$(E^4 - 4E^3 + 6E^2 - 4E + 1)f(x_0) = 0$$

$$E^4 f(x_0) - 4\,E^3 f(x_0) + 6\,E^2 f(x_0) - 4Ef(x_0) + f(x_0) = 0$$

$$f(x_4) - 4f(x_3) + 6f(x_2) - 4f(x_1) + f(x_0) = 0$$

$$As\ f(x_0) = 2, f(x_1) = 5, f(x_2) = 7, f(x_4) = 32$$

$$32 - 4f(x_3) + (6 \times 7) - (4 \times 5) + 2 = 0$$

$$\therefore 56 - 4f(x_3) = 0$$

$$f(x_3) = 14$$

Example: 13 Evaluate $f(9)$ and $f(18)$ from the given data table.

x	10	20	30	40	50
$y = f(x)$	46	66	81	93	101

$$\boldsymbol{y_9 = 43.5585, y_{18} = 62.5168}$$

Example: 14 The following table gives pressure of a steam at a given temperature. Using Newton's formula, compute the pressure for a temperature of 142°C, 155°C.

Temperature °C	140	150	160	170	180
Pressure kgf/cm^2	3.685	4.854	6.302	8.076	10.225

For $t = 142$°C, $Pressure: P = 3.8987\ kgf/cm^2\ and\ t = 155$°C, $Pressure: P = 5.540\ kgf/cm^2$

2.9. Newton's Backward Difference Interpolation Formula:

To evaluate or interpolate the value of the function $y = f(x)$ near the end of table of values or to extrapolate value of the function a short distance forward from y_n, Newton's backward difference interpolation formula is used.

Let $y = f(x)$ be a function, which takes values $f(x_n), f(x_n - h), f(x_n - 2h) \dots f(x_n - nh) = f(x_0)$ corresponding to equispaced values $x_n, x_n - h, x_n - 2h, \dots\ x_n - nh = x_0$.

Suppose we wish to calculate $f(x)$ at $x_n + ph$, where p is any real number.

$$\left[as\ x = x_n + ph, \qquad p = \frac{x - x_n}{h}\right]$$

By using shift operator E,

$$f(x_n + ph) = E^p f(x_n) = (E^{-1})^{-p} f(x_n) = (1 - \nabla)^{-p} f(x_n)$$

Using Binomial expansion, we have

$$\therefore f(x_n + ph) = \left[1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 + \frac{p(p+1)(p+2)}{3!}\nabla^3 + \cdots\right] f(x_n)$$

$$\therefore f(x_n + ph) = f(x_n) + p\nabla f(x_n) + \frac{p(p+1)}{2!}\nabla^2 f(x_n) + \frac{p(p+1)(p+2)}{3!}\nabla^3 f(x_n) + \cdots + \frac{p(p+1)\dots(p+n-1)}{n!}\nabla^n f(x_n) + Error$$

In terms of $y = f(x)$ it is written as

$$\therefore y_x = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \cdots + \frac{p(p+1)\dots(p+n-1)}{n!}\nabla^n y_n + Error$$

Example: 15 Evaluate $f(45), f(52)$ from the given data table.

x	10	20	30	40	50
$y = f(x)$	46	66	81	93	101

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
10	46				
		20			
20	66		-5		
		15		2	
30	81		-3		-3
		12		-1	
40	93		-4		
		8			
50	101				

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

$$x_4 = 50, y_4 = 101, \nabla y_4 = 8, \nabla^2 y_4 = -4, \nabla^3 y_4 = -1, \nabla^4 y_4 = -3$$

We want to evaluate $y_x \; for \; x = 45. (i.e. y_{45})$.

As $h = 10, p = \frac{x - x_n}{h} = \frac{45}{10} = \frac{-5}{10} = -0.5$

$$y_{45} = 101 + (-0.5)(8) + \frac{(-0.5)(0.5)}{2!}(-4) + \frac{(-0.5)(0.5)(1.5)}{3!}(-1) + \frac{(-0.5)(0.5)(1.5)(2.5)}{4!}(-3)$$

$$\therefore y_{45} = 101 - 4 + 0.5 + 0.0625 + 0.1172$$

$$y_{45} = f(45) = 97.6797$$

similarly

$$y_{52} = f(52) = 101.8208$$

Example: 16 for the following table of values estimate $f(7.5)$.

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

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
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
		127		6	
7	343		42		
		169			
8	512				

As $\nabla^4 y = 0, \nabla^5 y \ldots = 0$.

$$y_x = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n$$

For $x = 7.5, p = \frac{x - x_n}{h} = \frac{7.5-8}{1} = -0.5,$

$$y_n = 512, \nabla y_n = \nabla y_8 = 169, \nabla^2 y_n = \nabla^2 y_8 = 42, \nabla^3 y_n = \nabla^3 y_8 = 6$$

$$y_{7.5} = 512 + (-0.5)(169) + \frac{(-0.5)(0.5)}{2!}(42) + \frac{(-0.5)(0.5)(1.5)}{3!}(6)$$

$$y_{7.5} = 512 - 84.5 - 5.25 - 0.375 = 421.875$$

$$y_{7.5} = 421.875$$

Example: 17 Estimate $f(0.36), f(0.45)$ using Newton's backward difference interpolating formula for the given data.

x	0.1	0.2	0.3	0.4	0.5
$y = f(x)$	1.40	1.56	1.76	2.00	2.28

For $x = 0.36, y = 1.8992$ For $x = 0.45, y = 2.1350$

Example: 18 The following table gives pressure of a steam at a given temperature. Using Newton's backward formula, compute the pressure for a temperature of 175°C, 183°C.

Temperature °C	140	150	160	170	180
Pressure kgf/cm^2	3.685	4.854	6.302	8.076	10.225

For $t = 175$°C, $Pressure: P = 9.1005\ kgf/cm^2, t = 183$°C, $Pressure: P = 10.9504\ kgf/cm^2$

Example: 19 Evaluate $f(4.8)$ from the following table:

x	1	2	3	4	5
$y = f(x)$	2	5	7	14	32

$$f(4.8) = 27.232$$

Example: 20 The sale in a particular department store for the last five years is given in the table. Estimate the sale for the year 2015, 2017.

Year	2010	2012	2014	2016	2018
Sale (in lakhs)	40	43	48	52	57

Sale in 2015 is 50.1172 lakhs, Sale in 2017 is 54.0547 lakhs

2.10. Lagrange's Interpolation Formula:

Newton's interpolation formula can be used only when the values of the independent variable x are equally spaced and the differences of y must ultimately become small. If the values of the independent variable x are not given at equidistant intervals, then we have to use Lagrange's interpolating formula.

Let $y = f(x)$ be a function which takes the values $y_0, y_1, y_2, \dots, y_n$ corresponding to $x_0, x_1, x_2, \dots, x_n$. Since there are $(n+1)$ values of y corresponding to $(n+1)$ values of x, we can represent the function $f(x)$ by a polynomial of degree n. Suppose the polynomial is of the form …

$$f(x) = A_0 x^n + A_1 x^{n-1} + \cdots + A_n \quad (1)$$

Or

$$\begin{aligned} y = f(x) = {} & a_0 (x - x_1)(x - x_2) \dots (x - x_n) + a_1 (x - x_0)(x - x_2) \dots (x - x_n) \\ & + a_2 (x - x_0)(x - x_1) \dots (x - x_n) + \cdots \\ & + a_n (x - x_0)(x - x_1) \dots (x - x_{n-1}) \qquad (2) \end{aligned}$$

Here, the coefficients a_k are so chosen as to satisfy above equation by the $(n+1)$ pairs (x_i, y_i).

$\therefore x = x_0$, equation (2) becomes

$$y_0 = f(x_0) = a_0 (x_0 - x_1)(x_0 - x_2) \dots (x_0 - x_n)$$

Therefore,

$$a_0 = \frac{y_0}{(x_0 - x_1)\,(x_0 - x_2)\,\dots\,(x_0 - x_n)}$$

Similarly,

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

$$a_i = \frac{y_i}{(x_i - x_1)\,(x_i - x_2)\,\dots(x_i - x_{i-1})\,(x_i - x_{i+1})\,\dots\,(x_i - x_n)}$$

$$a_n = \frac{y_n}{(x_n - x_0)\,(x_n - x_2)\,\dots\,(x_n - x_{n-1})}$$

(3)

Now substitute the values of $a_0, a_1, \dots, a_n$ into equation (2), we get

$$y = f(x) = \frac{(x - x_1)\,(x - x_2)\,\dots\,(x - x_n)}{(x_0 - x_1)\,(x_0 - x_2)\,\dots\,(x_0 - x_n)}\,y_0 + \frac{(x - x_0)\,(x - x_2)\,\dots\,(x - x_n)}{(x_1 - x_0)\,(x_1 - x_2)\,\dots\,(x_1 - x_n)}\,y_1$$

$$+ \cdots + \frac{(x - x_1)\,(x - x_2)\,\dots(x - x_{i-1})\,(x - x_{i+1})\,\dots\,(x - x_n)}{(x_i - x_1)\,(x_i - x_2)\,\dots(x_i - x_{i-1})\,(x_i - x_{i+1})\,\dots\,(x_i - x_n)}\,y_i$$

$$+ \cdots + \frac{(x - x_0)\,(x - x_1)\,\dots\,(x - x_{n-1})}{(x_n - x_0)\,(x_n - x_2)\,\dots\,(x_n - x_{n-1})}\,y_n \qquad (4)$$

Above equation is called Lagrange's equation for interpolation. This formula can be used whether the values $x_0, x_1, x_2, \dots, x_n$ are equally spaced or not.

Equation (4) can be also written as

$$y = f(x) = L_0(x)y_0 + L_1(x)y_1 + \cdots + L_i(x)y_i + \cdots + L_n(x)y_n$$

$$= \sum_{k=0}^{n} L_k(x)y_k$$

$$\therefore\ y = f(x) = \sum_{k=0}^{n} L_k(x)f(x_k) \qquad (5)$$

Where,

$$L_i(x) = \frac{(x - x_1)\,(x - x_2)\,\dots(x - x_{i-1})\,(x - x_{i+1})\,\dots\,(x - x_n)}{(x_i - x_1)\,(x_i - x_2)\,\dots(x_i - x_{i-1})\,(x_i - x_{i+1})\,\dots\,(x_i - x_n)} \qquad (6)$$

From equation (4) we can observe that

$$L_i(x_i) = 1 \text{ and } L_i(x_j) = 0, for\ i \neq j$$

Thus we introduce Kronecker delta notation as

$$L_i(x_j) = \delta_{ij} = \begin{cases} 1, & if\ i = j \\ 0, & if\ i \neq j \end{cases} \quad (7)$$

Further if we introduce a notation

$$\prod(x) = \prod_{i=0}^{n}(x - x_i) = (x - x_0)\,(x - x_1)\dots\,(x - x_n) \quad (8)$$

$\prod(x)$ is a product of $(n+1)$ factors. It's derivative $\prod'(x)$ contains sum of $(n+1)$ terms in each of which one of the factor of $\prod(x)$ will be absent.

$$\prod{}'(x) = (x - x_1)\,(x - x_2)\dots\,(x - x_n) + \,(x - x_0)\,(x - x_2)\dots\,(x - x_n) + \cdots$$
$$+ \,(x - x_0)\,(x - x_1)\dots\,(x - x_{n-1})$$

We define

$$P_k(x) = \prod_{i\neq k}(x - x_i) \quad (9)$$

Which is same as $\prod(x)$ except that the factor $(x - x_k)$ is absent. Then

$$\prod{}'(x) = P_0(x) + \,P_1(x) + \cdots + \,P_n(x) \quad (10)$$

But when $x = x_k$, all terms in the above sum vanishes except $P_k(x_k)$. Hence,

$$\prod{}'(x_k) = \,P_k(x_k) = (x_k - x_0)\,(x_k - x_1)\dots\,(x_k - x_{k-1})\,(x_k - x_{k+1})\dots\,(x_k - x_n) \quad (11)$$

Using equations (9) and equation (11), equation (6) is written as (as i is replace with k)

$$L_k(x) = \frac{P_k(x)}{P_k(x_k)} = \frac{P_k(x)}{\prod'(x_k)} = \frac{\prod(x)}{(x - x_k)\prod'(x)} \quad (12)$$

Therefore Lagrange's interpolation polynomial of degree n can be written as

$$y(x) = f(x) = \sum_{k=0}^{n}\frac{\prod(x)}{(x - x_k)\prod'(x)}\,f(x_k) = \sum_{k=0}^{n}L_k(x)\,f(x_k) = \sum_{k=0}^{n}L_k(x)\,y_k \quad (13)$$

Example: 21 Find Lagrange's interpolation polynomial fitting the points

$y(1) = -3, y(3) = 0, y(4) = 30, y(6) = 132$. Hence find $y(5)$.

x	1	3	4	6
$y(x) = f(x)$	-3	0	30	132

Using Lagrange's interpolation formula, we have

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

$$y(x) = f(x) = \frac{(x-3)\,(x-4)(x-6)}{(1-3)(1-4)\,(1-6)}\,(-3) + \frac{(x-1)\,(x-4)(x-6)}{(3-1)\,(3-4)\,(3-6)}\,(0) + \frac{(x-1)\,(x-3)(x-6)}{(4-1)\,(4-3)\,(4-6)}\,(30) + + \frac{(x-1)\,(x-3)(x-4)}{(6-1)\,(6-3)\,(6-4)}\,(132)$$

$$= \frac{x^3-13x^2+54x-72}{(-30)}\,(-3) + \frac{x^3-11x^2+34x-24}{(6)}\,(0) + \frac{x^3-10x^2+27x-18}{(-6)}\,(30) + \frac{x^3-8x^2+19x-12}{(30)}\,(132)$$

$$= \frac{x^3-13x^2+54x-72}{(10)} + \frac{x^3-10x^2+27x-18}{(10)}\,(-50) + \frac{x^3-8x^2+19x-12}{(10)}\,(44)$$

$$= \frac{1}{10}(-5x^3+135x^2-460x+300) = \frac{1}{2}(-x^3+27x^2-92x+60)$$

Which is the required Lagrange’s interpolation polynomial.

$$y(5) = \frac{1}{2}[-(5^3)+27(5^2)-92(5)+60] = \frac{1}{2}[-125+675-460+60] = \frac{1}{2}[150] = 75$$

$$y(5) = 75$$

Example: 22 Given the following data evaluate $f(3)$ using Lagrange’s interpolating polynomial.

x	1	2	5
$y = f(x)$	1	4	10

Using Lagrange’s interpolation formula, we have

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

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

$$f(3) = 6.4999$$

3.1. Numerical Differentiation:

Consider a function $y = f(x)$ of a single variable x. If the function is known and simple, we can easily obtain its derivative using some mathematical rule or analytical method.

$$y = x^3 \quad \therefore \ \frac{dy}{dx} = 3x^2$$

$$y = e^{2x} \ \therefore \frac{dy}{dx} = 2e^{2x}$$

$$y = \sin 5x \quad \frac{dy}{dx} = 5\cos 5x$$

$$y = x^3 + 4x^2 \quad \frac{dy}{dx} = 3x^2 + 8x$$

If we do not know the function or the function is complicated and it is given in a tabular form at a set of points $x_0, x_1, \dots, x_{n,}$ we use only numerical methods for differentiation or integration of the given function.

3.2. Differentiation using Difference operators:

We assume that the function $y = f(x)$ is given for the values of the independent variable $x = x_0 + ph$, for $p = 0, 1, 2, \dots$ and so on.

3.3. Differentiation using Forward Difference Operators:

We know that $hD = \log E$ and $E = 1 + \Delta$ (1)

Where D is a differential operator and E is a shift operator.

From equation (1), we have,

$$hD = \log(1 + \Delta) \quad (2)$$

we know that $\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots$

$$hD = \log(1 + \Delta) = \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \frac{\Delta^5}{5} - \cdots$$

$$\therefore D = \frac{1}{h}\left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \frac{\Delta^5}{5} - \cdots\right) \quad (3)$$

Therefore,

$$Df(x_0) = f'(x_0) = \frac{1}{h}\left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \frac{\Delta^5}{5} - \cdots\right) f(x_0)$$

$$Df(x_0) = f'(x_0) = \frac{1}{h}\left[\Delta f(x_0) - \frac{\Delta^2 f(x_0)}{2} + \frac{\Delta^3 f(x_0)}{3} - \frac{\Delta^4 f(x_0)}{4} + \frac{\Delta^5 f(x_0)}{5} - \cdots\right]$$

$$= \frac{d}{dx} f(x_0)$$

Or in terms of $y = f(x)$

$$Dy_0 = \frac{dy_0}{dx} = y_0' = \frac{1}{h}\left[\Delta y_0 - \frac{\Delta^2 y_0}{2} + \frac{\Delta^3 y_0}{3} - \frac{\Delta^4 y_0}{4} + \frac{\Delta^5 y_0}{5} - \cdots\right] \quad (4)$$

Similarly

$$D^2 = \frac{1}{h^2}\left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4}\right)^2 = \frac{1}{h^2}\left[\Delta^2 - \Delta^3 + \frac{11}{12}\Delta^4 - \frac{5}{6}\Delta^5 + \cdots\right] \quad (5)$$

$$D^2 y_0 = \frac{d^2 y_0}{dx^2} = y_0'' = \frac{1}{h^2}\left[\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12}\Delta^4 y_0 - \frac{5}{6}\Delta^5 y_0 + \cdots\right] \quad (6)$$

3.4. Differentiation using Backward Difference Operators:

We know that $hD = -\log(1-\nabla)$ (7)

On expansion, we have

$$D = \frac{1}{h}\left(\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \cdots\right) \quad (8)$$

$$D^2 = \frac{1}{h^2}\left(\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \cdots\right)^2 = \frac{1}{h^2}\left(\nabla^2 + \nabla^3 + \frac{11}{12}\nabla^4 + \frac{5}{6}\nabla^5 + \cdots\right) \quad (9)$$

Hence,

$$Dy_n = \frac{dy_n}{dx} = y_n' = \frac{1}{h}\left(\nabla y_n + \frac{\nabla^2 y_n}{2} + \frac{\nabla^3 y_n}{3} + \frac{\nabla^4 y_n}{4} + \cdots\right) \quad (10)$$

$$D^2 y_n = \frac{d^2 y_n}{dx^2} = y_n'' = \frac{1}{h^2}\left[\nabla^2 y_n + \nabla^3 y_n + \frac{11}{12}\nabla^4 y_n + \frac{5}{6}\nabla^5 y_n + \cdots\right] \quad (11)$$

Formulae (4), (6) are useful to calculate the first and second derivatives at the beginning of the table of values in terms of forward differences; while formulae (10) & (11) are used to compute the first and second derivative near the end points of the table, in terms of backward differences.

To compute the derivatives of tabular function at points not found in the table, we can proceed as follows:

$$y(x_n + ph) = y(x_n) + p\nabla y(x_n) + \frac{p(p+1)}{2!}\nabla^2 y(x_n) + \frac{p(p+1)(p+2)}{3!}\nabla^3 y(x_n) + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y(x_n) + \cdots \quad (12)$$

or

$$y(x_n + ph) = y(x_n) + p\nabla y(x_n) + \frac{(p^2+p)}{2}\nabla^2 y(x_n) + \frac{(p^3+3p^2+2p)}{6}\nabla^3 y(x_n) + \frac{(p^4+6p^3+11p^2+6p)}{24}\nabla^4 y(x_n) + \cdots \quad (12)$$

Let $x = x_n + ph$, then $p = (x - x_n)/h$. Now differentiating equation (12) with respect to x, we get

$$y' = \frac{dy}{dp}\frac{dp}{dx} = \frac{1}{h}\left[\nabla y_n + \frac{2p+1}{2}\nabla^2 y_n + \frac{3p^2+6p+2}{6}\nabla^3 y_n + \frac{4p^3+18p^2+22p+6}{24}\nabla^4 y_n + \cdots\right] \quad (13)$$

Where $\frac{dp}{dx} = \frac{1}{h}$

$$y'' = \frac{d^2y}{dx^2} = \frac{d}{dp}(y')\frac{dp}{dx} = \frac{1}{h^2}\left[\nabla^2 y_n + (p+1)\nabla^3 y_n + \frac{6p^2+18p+11}{12}\nabla^4 y_n + \cdots\right] \quad (14)$$

Equations (13) and (14) are Newton's backward interpolation formula, which can be used to compute the first and second derivatives of a tabular function near the end of the table.

Similar expressions of Newton's forward interpolation formula can be derived to compute derivatives near the beginning of the table of values.

Example: 23 Compute $f''(0)$ and $f'(0.2)$ from the following tabular data.

x	0.0	0.2	0.4	0.6	0.8	1.0
$f(x)$	1.00	1.16	3.56	13.96	41.96	101.00

Since $x = 0$ and $x = 0.2$ appear at and near beginning of the table, it is appropriate to use forward difference formula to find the derivatives.

The forward difference table

x	$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$	$\Delta^4 f(x)$	$\Delta^5 f(x)$
0.0	1.00					
		0.16				
0.2	1.16		2.24			
		2.40		5.76		
0.4	3.56		8.00		3.84	
		10.40		9.60		0.00
0.6	13.96		17.60		3.84	
		28.00		13.44		
0.8	41.96		31.04			
		59.04				
1.0	101.00					

We have differentiation formula using forward differences

$$D^2 y_0 = \frac{d^2 y_0}{dx^2} = y_0'' = \frac{1}{h^2}\left[\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12}\Delta^4 y_0 - \frac{5}{6}\Delta^5 y_0 + \cdots\right]$$

Or for any x:

$$D^2f(x) = \frac{d^2f(x)}{dx^2} = f''(x) = \frac{1}{h^2}\left[\Delta^2 f(x) - \Delta^3 f(x) + \frac{11}{12}\,\Delta^4 f(x) - \frac{5}{6}\,\Delta^5 f(x)\right]$$

Here we have $h = 0.2$

for $x = x_0 = 0,\ \Delta^2 f(0) = 2.24,\ \Delta^3 f(0) = 5.76,\ \Delta^4 f(0) = 3.84,\ \Delta^5 f(0) = 0.00$

$$f''(0) = \frac{1}{h^2}\left[\Delta^2 f(0) - \Delta^3 f(0) + \frac{11}{12}\,\Delta^4 f(0) - \frac{5}{6}\,\Delta^5 f(0)\right]$$

$$f''(0) = \frac{1}{(0.2)^2}\left[2.24 - 5.76 + \frac{11}{12}\,(3.84) - \frac{5}{6}\,(0)\right] = 0.0$$

for $x = x_1 = 0.2,\ \Delta f(0.2) = 2.40,\ \ \Delta^2 f(0.2) = 8.00,\ \Delta^3 f(0.2) = 9.60,\ \Delta^4 f(0.2) = 3.84$

$$Df(x_1) = f'(x_1) = \frac{1}{h}\left[\Delta f(x_1) - \frac{\Delta^2 f(x_1)}{2} + \frac{\Delta^3 f(x_1)}{3} - \frac{\Delta^4 f(x_1)}{4}\right]$$

$$f'(0.2) = \frac{1}{0.2}\left[2.40 - \frac{8.00}{2} + \frac{9.60}{3} - \frac{3.84}{4}\right] = 3.2$$

Example: 24 find $y'(2.2)$ and $y''(2.2)$ from the table

x	1.4	1.6	1.8	2.0	2.2
$y(x)$	4.0552	4.9530	6.0496	7.3891	9.0250

Here 2.2 occurs at the end of the table, it is appropriate to use backward difference formula for derivatives.

The backward difference table:

x	$y(x)$	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
1.4	4.0552				
		0.8978			
1.6	4.9530		0.1988		
		1.0966		0.0441	
1.8	6.0496		0.2429		0.0094
		1.3395		0.0535	
2.0	7.3891		0.2964		
		1.6359			
2.2	9.0250				

Here $h = 0.2$, for $x = x_n = x_4 = 2.2, \nabla y_4 = 1.6359, \nabla^2 y_4 = 0.2964, \nabla^3 y_4 = 0.0535, \nabla^4 y_4 = 0.0094$

$$Dy_n = \frac{dy_n}{dx} = y_n' = \frac{1}{h}\left(\nabla y_n + \frac{\nabla^2 y_n}{2} + \frac{\nabla^3 y_n}{3} + \frac{\nabla^4 y_n}{4} + \cdots\right)$$

$$Dy_4 = y_4' = \frac{1}{h}\left(\nabla y_4 + \frac{\nabla^2 y_4}{2} + \frac{\nabla^3 y_4}{3} + \frac{\nabla^4 y_4}{4}\right)$$

$$y_4' = \frac{1}{0.2}\left(1.6359 + \frac{0.2964}{2} + \frac{0.0535}{3} + \frac{0.0094}{4}\right) = 5(1.8043) = 9.0215$$

$$D^2 y_4 = y_4'' = \frac{1}{h^2}\left[\nabla^2 y_4 + \nabla^3 y_4 + \frac{11}{12}\nabla^4 y_4\right]$$

$$y_4'' = \frac{1}{(0.2)^2}\left[0.2964 + 0.0535 + \frac{11}{12}(0.0094)\right] = 25(0.3585) = 8.9629$$

3.5. Numerical Integration:

Consider the definite integral

$$I = \int_{x=a}^{b} f(x)\,dx \quad (1)$$

Where $f(x)$ is known either explicitly or is given as a table of values corresponding to some values of x. Here we assume that the function is smooth and integrable in the given interval.

3.6. Newton-Cotes integration formula:

Newton-Cotes integration formula based on interpolation is used to form basis for trapezoidal rule and Simpson's rule of numerical integration. Here we shall approximate the given tabulated function, by a polynomial $P_n(x)$ and then integrate this polynomial.

Suppose we have given data $(x_i, y_i), i = 0, 1, \dots, n$ at equispaced points with spacing $h = x_{i+1} - x_i$.

Suppose we use Lagrangian approximation to represent the polynomial then we have

$$f(x) \approx \sum L_k(x)\, y(x_k) \quad (2)$$

With associated error given by

$$E(x) = \frac{\prod(x)}{(n+1)!}\, y^{(n+1)}(\xi) \quad (3)$$

Where

$$L_k(x) = \frac{\prod(x)}{(x - x_k)\prod'(x_k)} \quad (4)$$

With $\prod(x) = (x - x_0)\,(x - x_1)\dots(x - x_n) \quad (5)$

and $\prod'(x_k) = (x_k - x_0)\,(x_k - x_1)\dots(x_k - x_{k-1})\,(x_k - x_{k+1})\dots(x_k - x_n) \quad (6)$

Then, we obtain an equivalent integration formula to define integral of equation (1) in the form

$$\int_{x=a}^{b} f(x)\,dx = \int_{x=a}^{b} \left[\sum L_k(x)\,y(x_k)\right] dx = \sum \left[\int_{x=a}^{b} L_k(x)\,dx\right] y(x_k) = \sum c_k\,y(x_k)$$

$$\therefore \int_{x=a}^{b} f(x)\,dx = \sum c_k\,y(x_k) \quad (7)$$

Where c_k are the weighting coefficients given by

$$c_k = \int_{x=a}^{b} L_k(x)\,dx \quad (8)$$

These c_k are called Cotes numbers. The equispaced nodes are defined by the limits:

$$x_0 = a, \qquad x_n = b, \qquad h = \frac{x_n - x_0}{n} = \frac{b-a}{n}, \qquad x_k = x_0 + kh \quad (9)$$

From equation (9),

$$x_k - x_0 = kh, \qquad x_k - x_1 = (k-1)\,h, \qquad x_k - x_n = (k-n)\,h \quad (10)$$

Now we change variable from x to p such that

$$x = x_0 + ph \quad (11)$$

$$x - x_0 = ph, \qquad x - x_1 = (p-1)\,h, \qquad x - x_2 = (p-2)\,h, \qquad x - x_n = (p-n)\,h \quad (12)$$

Therefore equation (5) becomes

$$\prod(x) = (x-x_0)\,(x-x_1)\ldots(x-x_n) = ph\,(p-1)\,h\ldots(p-n)\,h$$

$$\therefore \prod(x) = h^{n+1}\,p\,(p-1)\ldots\,(p-n) \quad (13)$$

and

$$L_k(x) = \frac{\prod(x)}{(x-x_k)\prod'(x_k)}$$

$$\therefore\ L_k(x) = \frac{(x-x_0)\,(x-x_1)\ldots(x-x_{k-1})\,(x-x_k)\,(x-x_{k+1})\ldots(x-x_n)}{(x-x_k)\,[(x_k-x_0)\,(x_k-x_1)\ldots(x_k-x_{k-1})\,(x_k-x_{k+1})\ldots(x_k-x_n)]}$$

$$= \frac{(x-x_0)\,(x-x_1)\ldots(x-x_{k-1})\,(x-x_{k+1})\ldots(x-x_n)}{(x_k-x_0)\,(x_k-x_1)\ldots(x_k-x_{k-1})\,(x_k-x_{k+1})\ldots(x_k-x_n)}$$

$$= \frac{ph\,(p-1)h\ldots\big(p-(k-1)\big)h\,\big(p-(k+1)\big)h\ldots(p-n)h}{kh\,(k-1)h\ldots\big(k-(k-1)\big)h\,\big(k-(k+1)\big)h\ldots(k-n)h}$$

$$= \frac{p\,(p-1)\,...\,(p-k+1)\,(p-k-1)\,...\,(p-n)}{k\,(k-1)\,...\,(k-k+1)\,(k-k-1)\,...\,(k-n)}$$

$$= \frac{p\,(p-1)\,...\,(p-k+1)\,(p-k-1)\,...\,(p-n)}{k\,(k-1)\,...\,(1)\,(-1)\,...\,(k-n)}$$

(Here in denominator: $(-1)\,(-2)\,...\,(k-n) = (-1)1\,(-1)2\,...\,(-1)(n-k) = (-1)^{n-k}(n-k)!$)

$$\therefore\ L_k(x) = (-1)^{(n-k)}\ \frac{p\,(p-1)\,...\,(p-k+1)\,(p-k-1)\,...\,(p-n)}{k!\,(n-k)!} \quad (14)$$

As $x = x_0 + ph$, by taking derivative $dx = h\,dp$

Now substitute value of equation (14) in to equation (8) and by changing limits from 0 to n, it becomes

$$c_k = \int_{x=a}^{b} L_k(x)\,dx = \int_{o}^{n} (-1)^{(n-k)}\ \frac{p\,(p-1)\,...\,(p-k+1)\,(p-k-1)\,...\,(p-n)}{k!\,(n-k)!}\ h\,dp$$

$$c_k = \frac{(-1)^{(n-k)}\,h}{k!\,(n-k)!} \int_{o}^{n} p\,(p-1)\,...\,(p-k+1)\,(p-k-1)\,...\,(p-n)\,dp \quad (15)$$

Note: to calculate c_k skip $(p-k)$ term in integration of equation (15)

The error in approximating integral given by equation (7) can be obtained by substituting equation (13) in equation (3), we have

$$E_n = \frac{h^{n+2}}{(n+1)!} \int_{o}^{n} p\,(p-1)\,...\,(p-n)\,y^{(n+1)}\,(\xi)\,dp \quad (16)$$

Where $x_0 < \xi < x_n$.

Case: 1 For $n = 1$, (possible values of $k = 0, 1$)

(i) For $n = 1, k = 0$

Skip $p - k$ in equation (15) i.e. skip $p - k = p - 0 = p$ in calculations.

Integrand: $p\,(p-1)\,...\,(p-k+1)\,(p-k-1)\,...\,(p-n)$, becomes $(p-1)$ for $n = 1$.

$$c_k = \frac{(-1)^{(n-k)}\,h}{k!\,(n-k)!} \int_{o}^{n} p\,(p-1)\,...\,(p-k+1)\,(p-k-1)\,...\,(p-n)\,dp$$

Becomes

$$c_0 = \frac{(-1)^{(1-0)}\,h}{0!\,(1-0)!} \int_{o}^{1} (p-1)\,dp$$

$$c_0 = \frac{(-1)h}{1(1)} \int_o^1 (p-1)\, dp = -h \left[\frac{p^2}{2} - p\right]_0^1 = -h \left[\frac{1}{2} - 1\right] = -h \left[-\frac{1}{2}\right] = \frac{h}{2}$$

$$c_0 = \frac{h}{2} \quad (17)$$

(ii) For $n = 1, k = 1$

Skip $p - k$ in equation (15) i.e. skip $p - k = p - 1$ in calculations.

Integrand: $p\,(p-1) \dots (p-k+1)\,(p-k-1) \dots (p-n)$, becomes p for $n = 1$.

$$c_k = \frac{(-1)^{(n-k)}\, h}{k!\,(n-k)!} \int_o^n p\,(p-1) \dots (p-k+1)\,(p-k-1) \dots (p-n)\, dp$$

Becomes

$$c_1 = \frac{(-1)^{(1-1)}\, h}{1!\,(1-1)!} \int_o^1 p\, dp$$

$$c_1 = \frac{(1)\, h}{1(1)} \int_o^1 p\, dp = h \left[\frac{p^2}{2}\right]_0^1 = h \left[\frac{1}{2}\right] = \frac{h}{2}$$

$$c_1 = \frac{h}{2} \quad (18)$$

Equation (16)

$$E_n = \frac{h^{n+2}}{(n+1)!} \int_o^n p\,(p-1) \dots (p-n)\, y^{(n+1)}(\xi)\, dp \quad (16)$$

Becomes

$$E_1 = \frac{h^{1+2}}{(1+1)!} \int_o^1 p\,(p-1)\, y^{(1+1)}(\xi)\, dp$$

$$E_1 = \frac{h^3}{2}\, y''(\xi) \int_o^1 p\,(p-1)\, dp$$

$$= \frac{h^3}{2}\, y''(\xi) \int_o^1 (p^2 - p)\, dp = \frac{h^3}{2}\, y''(\xi) \left[\frac{p^3}{3} - \frac{p^2}{2}\right]_0^1 = \frac{h^3}{2}\, y''(\xi) \left[\frac{1}{3} - \frac{1}{2}\right]$$

$$= \frac{h^3}{2}\, y''(\xi) \left[\frac{2-3}{6}\right] = \frac{h^3}{2}\, y''(\xi) \left[-\frac{1}{6}\right]$$

$$\therefore\ E_1 = -\frac{h^3}{12}\, y''(\xi) \quad (19)$$

Therefore equation (7) becomes

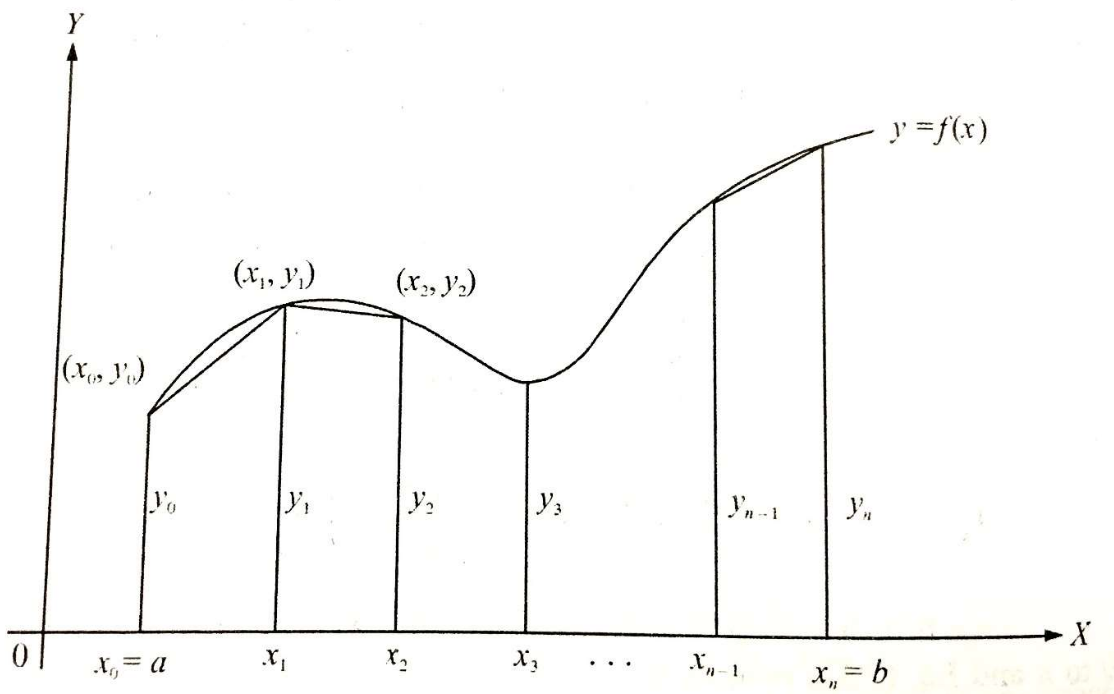
$$\int_{x=a}^{b} f(x)\,dx = \sum c_k\, y(x_k) \;\; or \; \int_{x_0}^{x_1} f(x)\,dx = \sum_{k=0}^{1} c_k\, y(x_k)$$

$$\int_{x_0}^{x_1} f(x)\,dx = c_0\, y_0 + c_1\, y_1 + Error$$

As $c_0 = c_1 = \frac{h}{2}$

$$\int_{x_0}^{x_1} f(x)\,dx = \frac{h}{2}\,(y_0 + y_1) - \frac{h^3}{12}\, y''\,(\xi) \quad (20)$$

This equation represents the Trapezoidal rule in the interval $[x_0,\ x_1]$ with the error term. Geometrically it represents an area between the curve $y = f(x)$ and the $X - axis$ between points $x = x_0\ (= a)$ & $x = x_1\ (= b)$. This area is approximated by the trapezium formed by the curve replacing the curve with its secant line drawn between end points (x_0, y_0) and (x_1, y_1).



Trapezoidal rule.

Case: 2 For $n = 2$, (possible values of $k = 0, 1, 2$)

(i) For $n = 2, k = 0$

Skip $p - k$ in equation (15) i.e. skip $p - k = p - 0 = p$ in calculations.

For $n = 2$, integrand: $p\,(p-1)\ldots(p-k+1)\,(p-k-1)\ldots(p-n)$, becomes $(p-1)\,(p-2)$.

$$c_k = \frac{(-1)^{(n-k)}\, h}{k!\,(n-k)!} \int_o^n p\,(p-1)\ldots(p-k+1)\,(p-k-1)\ldots(p-n)\, dp$$

Becomes

$$c_0 = \frac{(-1)^{(2-0)}\, h}{0!\,(2-0)!} \int_o^2 (p-1)\,(p-2)\,dp$$

$$c_0 = \frac{(-1)^2 h}{1(2)} \int_o^2 (p-1)\,(p-2)\,dp = \frac{h}{2}\int_o^2 (p^2 - 3p + 2)\ dp = \frac{h}{2}\left[\frac{p^3}{3} - \frac{3p^2}{2} + 2p\right]_0^2$$

$$= \frac{h}{2}\left[\frac{2^3}{3} - \frac{3(\,2^2)}{2} + 2(2)\right] = \frac{h}{2}\left[\frac{8}{3} - \frac{12}{2} + 4\right] = \frac{h}{2}\left[\frac{8}{3} - 2\right] = \frac{h}{2}\left[\frac{2}{3}\right] = \frac{h}{3}$$

$$c_0 = \frac{h}{3} \quad (21)$$

(ii) For $n = 2, k = 1$

Skip $p - k$ in equation (15) i.e. skip $p - k = p - 1$ in calculations.

For $n = 2$ integrand: $p\,(p-1) \dots (p-k+1)\,(p-k-1) \dots (p-n)$, becomes $p\,(p-2)$.

$$c_k = \frac{(-1)^{(n-k)}\, h}{k!\,(n-k)!} \int_o^n p\,(p-1) \dots (p-k+1)\,(p-k-1) \dots (p-n)\,dp$$

Becomes

$$c_1 = \frac{(-1)^{(2-1)}\, h}{1!\,(2-1)!} \int_o^2 p(p-2)\,dp$$

$$c_1 = \frac{(-1)\,h}{1(1)} \int_o^2 [p^2 - 2p]\,dp = -h\left[\frac{p^3}{3} - 2\frac{p^2}{2}\right]_0^2 = -h\left[\frac{2^3}{3} - 2^2\right] = -h\left[\frac{8}{3} - 4\right] = -h\left[-\frac{4}{3}\right]$$

$$= \frac{4h}{3}$$

$$c_1 = \frac{4h}{3} \quad (22)$$

(iii) For $n = 2, k = 2$

Skip $p - k$ in equation (15) i.e. skip $p - k = p - 2$ in calculations.

For $n = 2$ $integrand$: $p\,(p-1) \dots (p-k+1)\,(p-k-1) \dots (p-n)$, becomes $p\,(p-1)$.

$$c_k = \frac{(-1)^{(n-k)}\, h}{k!\,(n-k)!} \int_o^n p\,(p-1) \dots (p-k+1)\,(p-k-1) \dots (p-n)\,dp$$

Becomes

$$c_2 = \frac{(-1)^{(2-2)}\, h}{2!\,(2-2)!} \int_o^2 p\,(p-1)\,dp$$

$$c_2 = \frac{(1)\,h}{2(1)} \int_o^2 \; [p^2 - p]\, dp = \frac{h}{2}\left[\frac{p^3}{3} - \frac{p^2}{2}\right]_0^2 = \frac{h}{2}\left[\frac{2^3}{3} - \frac{2^2}{2}\right] = \frac{h}{2}\left[\frac{8}{3} - \frac{4}{2}\right] = \frac{h}{2}\left[\frac{8}{3} - 2\right] = \frac{h}{2}\left[\frac{2}{3}\right]$$

$$c_2 = \frac{h}{3} \quad (23)$$

Therefore equation (7) becomes for $n = 2$

$$\int_{x=a}^{b} f(x)\, dx = \sum c_k\, y(x_k) \;\; or \int_{x_0}^{x_2} f(x)\, dx = \sum_{k=0}^{2} c_k\, y(x_k)$$

$$\int_{x_0}^{x_2} f(x)\, dx = \; c_0\, y_0 + \; c_1\, y_1 + c_2\, y_2 + Error$$

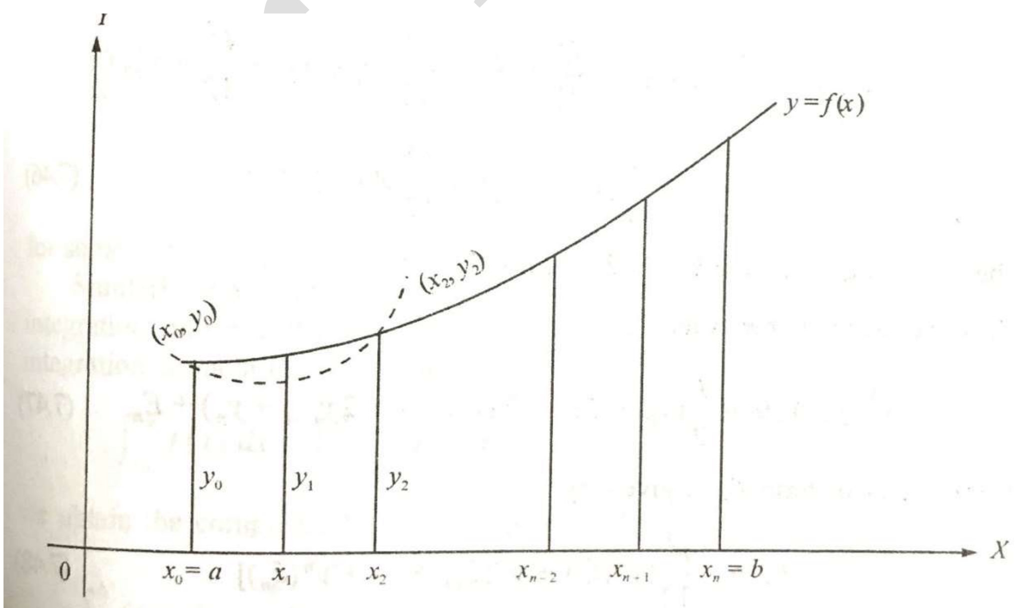
As $c_0 = \frac{h}{3},\; c_1 = \frac{4h}{3},\; c_2 = \frac{h}{3}$

$$\int_{x_0}^{x_2} f(x)\, dx = \frac{h}{3}\,(y_0 + \; 4y_1 + y_2)$$

With error,

$$\int_{x_0}^{x_2} f(x)\, dx = \frac{h}{3}\,(y_0 + \; 4y_1 + y_2) - \frac{h^5}{90}\, y^{(iv)}\,(\xi) \quad (24)$$

This is known as Simpson's 1/3rd rule. Geometrically, this equation represent the area between the curve $y = f(x)$, the $X - axis$ and the ordinates at $x = x_0$ and x_2 after replacing the arc of the curve between (x_0, y_0) and (x_2, y_2) by an arc of a quadratic polynomial as shown in the figure. Thus Simpson's1/3 rule is based on fitting three points with a quadratic.



Simpson's rule.

Similarly for $n = 3$, equation (7) becomes

$$\int_{x_0}^{x_3} f(x)\,dx = \frac{3}{8}h\,(y_0 + 3y_1 + 3y_2 + y_3) - \frac{3h^5}{80}\,y^{(iv)}\,(\xi) \quad (25)$$

This is known as Simpson's 3/8 rule, which is based on fitting four points by a cubic.

3.7. The Trapezoidal Rule (composite Form):

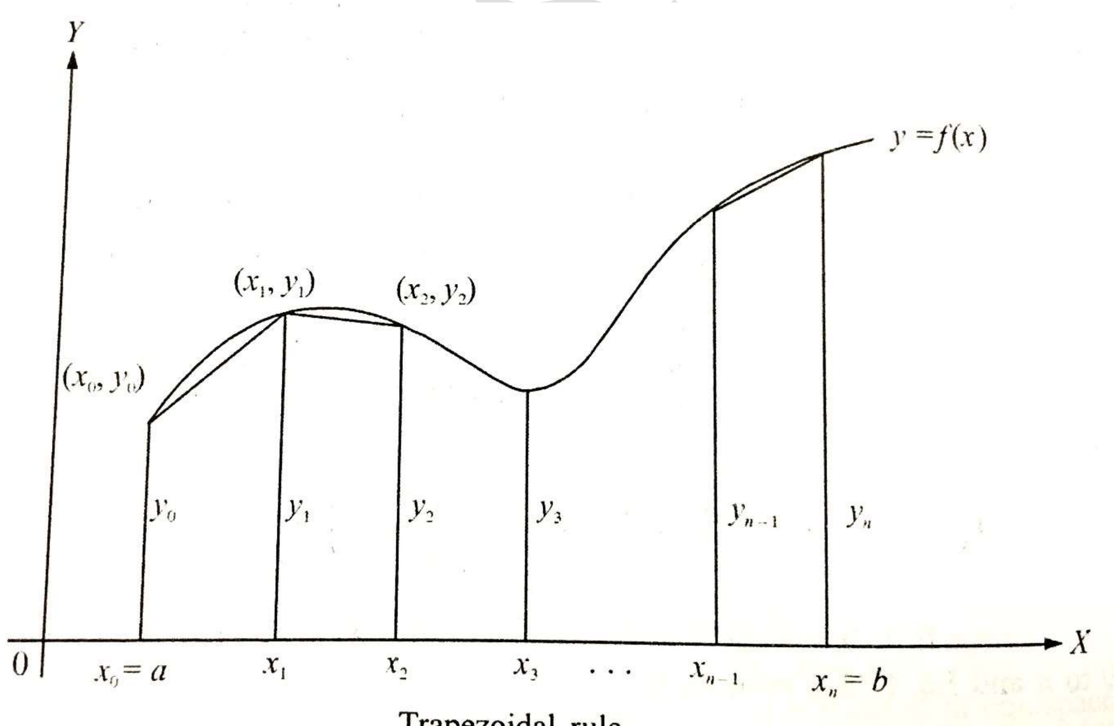
The Newton-Cotes formula is based on approximating $y = f(x)$ between (x_0, y_0) and (x_1, y_1) by a straight line, thus forming a trapezium, is called Trapezoidal Rule.

To evaluate the definite integral

$$I = \int_{x=a}^{b} f(x)\,dx \quad (1)$$

We divide the interval $[a, b]$ into n sub-intervals, each of size $h = (b-a)/n$ and denote the sub-intervals by $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$, such that $x_0 = a$, $x_n = b$ and $x_k = x_0 + kh, k = 1, 2, \dots, n-1$. Thus we can write above integral as a sum (of sub-integrals) as

$$I = \int_{x=a}^{b} f(x)\,dx = \int_{x_0}^{x_n} f(x)\,dx = \int_{x_0}^{x_1} f(x)\,dx + \int_{x_1}^{x_2} f(x)\,dx + \cdots + \int_{x_{n-1}}^{x_n} f(x)\,dx \quad (2)$$



Trapezoidal rule.

As shown in the above figure, the area under the curve in each sub-interval is approximated by a trapezium. The integral I, which represents an area between the curve $y = f(x)$, the $X - axis$

and the ordinates at $x = x_0$ and $x = x_n$ is obtained by adding all the trapezoidal areas in each sub-interval.

We have trapezoidal rule as:

$$\int_{x_0}^{x_1} f(x)\, dx = \frac{h}{2}(y_0 + y_1) - \frac{h^3}{12} y''(\xi) \quad (3)$$

Therefore,

$$I = \int_{x_0}^{x_n} f(x)\, dx = \frac{h}{2}(y_0 + y_1) - \frac{h^3}{12} y''(\xi_1) + \frac{h}{2}(y_1 + y_2) - \frac{h^3}{12} y''(\xi_2)$$

$$+ \cdots + \frac{h}{2}(y_{n-1} + y_n) - \frac{h^3}{12} y''(\xi_n) \quad (4)$$

Where $x_{k-1} < \xi_k < x_k$, for $k = 1, 2, \dots, n-1$.

Equation (4) is finally written as:

$$I = \int_{x_0}^{x_n} f(x)\, dx = \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n) + E_n \quad (5)$$

Where the error term E_n is given by

$$E_n = -\frac{h^3}{12}\left[y''(\xi_1) + y''(\xi_2) + \cdots + y''(\xi_n)\right] \quad (6)$$

Equation (7) represents the trapezoidal rule over $[x_0, x_n]$, which is also called composite form of the trapezoidal rule. The error given by equation (6) is called the global error. If we assume that $y''(x)$ is continuous over $[x_0, x_n]$ then there exists some ξ in $[x_0, x_n]$ such that $x_n = x_0 + nh$ and

$$E_n = -\frac{h^3}{12}\left[n\, y''(\xi)\right] = -\frac{x_n - x_0}{12} h^2\, y''(\xi) = O(h^2) \quad (7)$$

3.8. Simpson's Rules (Composite forms):

To derive composite from of Simpson's rule, we shall divide the interval of integration $[a,\ b]$ into an even number of sub-intervals say $2N$.Each of width $(b-a)/2N$, therefore we have $x_0 = a$, $x_1,\ \dots,\ x_{2N} = b$ and $x_k = x_0 + kh, k = 1, 2, \dots, (2N-1)$. thus the definite integral I can be written as

$$I = \int_{x=a}^{b} f(x)\, dx = \int_{x_0}^{x_{2N}} f(x)\, dx = \int_{x_0}^{x_2} f(x)\, dx + \int_{x_2}^{x_4} f(x)\, dx + \cdots + \int_{x_{2N-2}}^{x_{2N}} f(x)\, dx \quad (8)$$

We have Simpson's 1/3 rule as

$$I = \int_{x_0}^{x_{2N}} f(x)\, dx = \frac{h}{3}(y_0 + 4y_1 + y_2) + \frac{h}{3}(y_2 + 4y_3 + y_4)$$

$$+\cdots+\ \frac{h}{3}\ (y_{2N-2}+\ 4y_{2N-1}+y_{2N})\ -\ \frac{N}{90}\ h^5\ y^{(iv)}\ (\xi)$$

By rearranging the terms,

$$I=\int_{x_0}^{x_{2N}} f(x)\ dx$$

$$=\ \frac{h}{3}\ [y_0+\ 4(y_1+y_3+\cdots+\ y_{2N-1})+2(y_2+\ y_4+\cdots+\ y_{2N-2})+\ y_{2N}]+Error \qquad (9)$$

This formula is called composite form of Simpson's 1/3 rule. The error term E (called global error) is

$$-\ \frac{N}{90}\ h^5\ y^{(iv)}\ (\xi)=-\frac{x_{2N}-x_0}{180}\ h^4\ y^{(iv)}\ (\xi) \quad (10)$$

for some ξ in $[x_0, x_{2N}]$. Thus in Simpson's 1/3 rule the global error is of $O(h^4)$.

Simpson's 3/8 rule:

To derive Simpson's 3/8 rule we divide the interval of integration into n sub-intervals, where n is divisible by 3, and applying the integration formula to each of the integral given below

$$I=\int_{x=a}^{b} f(x)\ dx=\int_{x_0}^{x_n} f(x)\ dx=\int_{x_0}^{x_3} f(x)\ dx+\int_{x_3}^{x_6} f(x)\ dx+\cdots+\int_{x_{n-3}}^{x_n} f(x)\ dx \quad (11)$$

We obtain the composite form of Simpson's 3/8 rule as

$$I=\int_{x_0}^{x_n} f(x)\ dx=\ \frac{3}{8}\ h\begin{bmatrix} y(a)+\ 3y_1+3y_2+2y_3+\ 3y_4+\ 3y_5+2y_6 \\ +\cdots+\ 2y_{n-3}+3y_{n-2}+3y_{n-1}+y(b)\end{bmatrix} \quad (12)$$

With the global error E given by

$$E=-\frac{x_n-x_0}{80}\ h^4\ y^{(iv)}\ (\xi) \qquad (13)$$

From equation (10) and equation (13) it is seen that the global error in Simpson's 1/3 and 3/8 rules are of the same order. By considering magnitudes of the error terms we can say that Simpson's 1/3 rule is superior to Simpson's 3/8 rule.

Example: 25 Find the approximate value of

$$y=\int_0^{\pi} \sin x\, dx$$

using (i) trapezoidal rule, (ii) Simpson's 1/3 rule by dividing the range of integration into six equal parts. Calculate the percentage error from its true value in both the cases.

We shall divide the range of integration $[0, \pi]$ into six equal parts. As $h = \frac{b-a}{n} = \frac{\pi-0}{6} = \frac{\pi}{6}$, each part of width $\frac{\pi}{6}$. We prepare a table

x	0	$\pi/6$	$2\pi/6 = \pi/3$	$3\pi/6 = \pi/2$	$4\pi/6 = 2\pi/3$	$5\pi/6$	$6\pi/6 = \pi$
$y = \sin x$	0.0	0.5	0.8660	1.0	0.8660	0.5	0.0

(i) Applying trapezoidal rule, we have $h = \frac{b-a}{n} = \frac{\pi-0}{6} = \frac{\pi}{6}$

$$y = \int_0^{\pi} \sin x \, dx = \frac{h}{2} \left[y_0 + y_6 + 2(y_1 + y_2 + y_3 + y_4 + y_5)\right]$$

$$= \frac{\pi/6}{2} [0 + 0 + 2\,(0.5 + 0.8660 + 1.0 + 0.8660 + 0.5)] = \frac{\pi}{12} [2(3.732)]$$

$$= \frac{\pi}{6} \times 3.732 = 1.9540$$

(ii) Applying Simpson's 1/3 rule,

$$y = \int_0^{\pi} \sin x \, dx = \frac{h}{3} \left[y_0 + y_6 + 4(y_1 + y_3 + y_5) + 2(y_2 + y_4)\right]$$

$$= \frac{\pi/6}{3} [0 + 0 + 4(0.5 + 1.0 + 0.5) + 2(0.8660 + 0.8660)] = \frac{\pi}{18} [4(2) + 2(1.732)]$$

$$= \frac{\pi}{18} \times 11.464 = 2.008$$

Actual value of the integration is

$$y = \int_0^{\pi} \sin x \, dx = [-\cos x]_0^{\pi} = -[\cos \pi - \cos 0] = -[-1 - 1] = 2$$

Hence in the case of trapezoidal rule

$$The\ percentage\ of\ error = \frac{2 - 1.9540}{2} \times 100 = 2.3\ \%$$

While in the case of Simpson's 1/3 rule

$$The\ percentage\ of\ error = \frac{2 - 2.008}{2} \times 100 = -0.04\ \%\ or\ 0.04\ \%$$

Example: 26 From the following data, estimate the value of $y = \int_1^5 \log x \, dx$

using Simpson's 1/3 rule.

x	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
$y = \log x$	0.0000	0.4055	0.6931	0.9163	1.0986	1.2528	1.3863	1.5041	1.6094

Here, we have $n = 0, 1, 2, \ldots 8, h = \frac{b-a}{n} = \frac{5-1}{8} = \frac{4}{8} = 0.5$

Using Simpson's 1/3 rule,

$$y = \int_1^5 \log x \, dx = \frac{h}{3}\left[y_0 + y_8 + 4(y_1 + y_3 + y_5 + y_7) + 2(y_2 + y_4 + y_6)\right]$$

$$= \frac{0.5}{3}\left[0 + 1.6094 + 4\,(0.4055 + 0.9163 + 1.2528 + 1.5041) + 2\,(0.6931 + 1.0986 + 1.3863)\right]$$

$$= \frac{0.5}{3}\left[1.6094 + 4\,(4.0787) + 2\,(3.178)\right] = 4.0467$$

Example: 27 evaluate the integral

$$I = \int_0^1 \frac{dx}{1+x^2}$$

using (i) trapezoidal rule, (ii) Simpson's 1/3 rule by taking $h = \frac{1}{4}$, Compute approximate value of π.

As $h = \frac{1}{4} = 0.25$

x	0	0.25	0.50	0.75	1.0
$y = \frac{1}{1+x^2}$	1	0.9412	0.8000	0.6400	0.5000

(i) Using trapezoidal rule

$$I = \int_0^1 \frac{dx}{1+x^2} = \frac{h}{2}\left[y_0 + y_4 + 2(y_1 + y_2 + y_3)\right]$$

$$= \frac{0.25}{2}\left[1 + 0.5 + 2(0.9412 + 0.8 + 0.64)\right] = 0.125[1.5 + 2(2.312)] = 0.7828$$

(ii) Using Simpson's 1/3 rule

$$I = \int_0^1 \frac{dx}{1+x^2} = \frac{h}{3}\left[y_0 + y_4 + 4(y_1 + y_3) + 2(y_2)\right]$$

$$= \frac{0.25}{3}\left[1 + 0.5 + 4(0.9412 + 0.64) + 2(0.8)\right] = 0.7854$$

We know that the analytical solution of

$$\int_0^1 \frac{dx}{1+x^2} = [\tan^{-1} x]_0^1 = \frac{\pi}{4}$$

Therefore by equating the result of Simpson's 1/3 rule with above result, we have

$$\frac{\pi}{4} = 0.7854$$

$$\pi = 3.1416$$

Example: 28 A missile is launched from a ground station. The acceleration during its first 80 seconds of flight, as recorded, is given in the following table:

$t\ (s)$	0	10	20	30	40	50	60	70	80
$a\ (m/s^2)$	30	31.63	33.34	35.47	37.75	40.33	43.25	46.69	50.67

Compute the velocity of the missile when $t = 80\ s$, using Simpson's 1/3 rule.

We know that acceleration is the rate of change of velocity,

$$a = \frac{dv}{dt} \quad or \quad v = \int_0^{80} a\ dt$$

Using Simpson's 1/3 rule

$$v = \frac{h}{3}\left[y_0 + y_8 + 4(y_1 + y_3 + y_5 + y_7) + 2(y_2 + y_4 + y_6)\right]$$

Here $h = 10\ s$

$$v = \frac{10}{3}\left[30 + 50.67 + 4(31.63 + 35.47 + 40.33 + 46.69) + 2(33.34 + 37.75 + 43.25)\right]$$

$$v = 3086.1\ m/s$$

Or

$$v = 3.0861\ km/s$$

Which is the required velocity.

x	0.0000	0.2618	0.5236	0.7854	1.0472	1.3090	1.5708
y	4.0000	3.3307	2.2857	1.6000	1.2308	1.0529	1.0000

$$h = \frac{b-a}{n} = \frac{\pi/2}{6} = \frac{\pi}{12} = 0.2618$$