Numerical Differentiation and Integration

7.1 INTRODUCTION

Consider a function of single variable $y = f(x)$. If the function is known and simple, we can easily obtain its derivative(s) or can evaluate its definite integral. However, if we do not know the function as such or the function is complicated and is given in a tabular form at a set of points $x_0, x_1, \ldots, x_n$, we use only numerical methods for differentiation or integration of the given function. We shall discuss numerical approximation to derivatives of functions of two or more variables in subsequent chapters to follow under partial differential equations. In the next couple of sections, we shall derive and illustrate various formulae for numerical differentiation of a function of a single variable based on finite difference operators and interpolation. Subsequently, we shall develop Newton–Cotes formulae and related trapezoidal rule and Simpson's rule for numerical integration of a function. Finally, we shall present Gaussian Quadrature formulae for evaluating both simple and multiple integrals.*

7.2 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 $p = 0, 1, 2, \ldots$ and so on. To find the derivatives of such a tabular function, we proceed as follows.

Case I: Using forward difference operator Δ and combining Eqs. (6.27) and (6.31) we have

$$hD = \log E = \log (1 + \Delta) \tag{7.1}$$

where D is a differential operator, E a shift operator. In terms of Δ, Eq. (7.1) gives

$$D = \frac{1}{h}\left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \frac{\Delta^5}{5} - \cdots\right) \tag{7.2}$$

Therefore,

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

*[For extrapolation methods, the interested reader may consult Hildebrand (1982).]

In other words,

$$Dy_0 = y_0' = \frac{1}{h}\left(\Delta y_0 - \frac{\Delta^2 y_0}{2} + \frac{\Delta^3 y_0}{3} - \frac{\Delta^4 y_0}{4} + \cdots\right) \tag{7.3}$$

Also, we can easily verify

$$D^2 = \frac{1}{h^2}\left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \cdots\right)^2 = \frac{1}{h^2}\left(\Delta^2 - \Delta^3 + \frac{11}{12}\Delta^4 - \frac{5}{6}\Delta^5 + \cdots\right) \tag{7.4}$$

Thus,

$$D^2 y_0 = \frac{d^2 y_0}{dx^2} = y_0'' = \frac{1}{h^2}\left(\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12}\Delta^4 y_0 - \frac{5}{6}\Delta^5 y_0 + \cdots\right) \tag{7.5}$$

Case II: Using backward difference operator ∇, we have seen in Example 6.5 that $hD = -\log(1 - \nabla)$.

On expansion, we have

$$D = \frac{1}{h}\left(\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \cdots\right) \tag{7.6}$$

we can also verify that

$$D^2 = \frac{1}{h^2}\left(\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \cdots\right)^2 = \frac{1}{h^2}\left(\nabla^2 + \nabla^3 + \frac{11}{12}\nabla^4 + \frac{5}{6}\nabla^5 + \cdots\right) \tag{7.7}$$

Hence,

$$\frac{d}{dx}y_n = Dy_n = y_n' = \frac{1}{h}\left(\nabla y_n + \frac{\nabla^2 y_n}{2} + \frac{\nabla^3 y_n}{3} + \frac{\nabla^4 y_n}{4} + \cdots\right) \tag{7.8}$$

and

$$y_n'' = D^2 y_n = \frac{1}{h^2}\left(\nabla^2 y_n + \nabla^3 y_n + \frac{11}{12}\nabla^4 y_n + \frac{5}{6}\nabla^5 y_n + \cdots\right) \tag{7.9}$$

The formulae (7.3) and (7.5) are useful to calculate the first and second derivatives at the beginning of the table of values in terms of forward differences; while formulae (7.8) and (7.9) are used to compute the first and second derivatives near the end points of the table, in terms of backward differences.

Similar formulae can also be derived for computing higher order derivatives.

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

Recalling Eq. (6.34) in the form

$$y(x_n + ph) = y(x_n) + p\nabla y(x_n) + \frac{p(p+1)}{2!}\nabla^2 y(x_n) + \frac{p(p+1)(p+2)}{3!}\nabla^3 y(x_n) + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y(x_n) + \cdots \tag{7.9a}$$

Let $x = x_n + ph$, then $p = (x - x_n)/h$. Now, differentiating Eq. (7.9a) with respect to x, we get

$$y' = \frac{dy}{dx} = \frac{dy}{dp}\frac{dp}{dx} = \frac{1}{h}\left[\nabla y_n + \frac{2p+1}{2}\nabla^2 y_n + \frac{3p^2+6p+2}{6}\nabla^3 y_n + \frac{4p^3+18p^2+22p+6}{24}\nabla^4 y_n + \cdots\right] \quad (7.9b)$$

Differentiating this result once again with respect to x, we arrive at the second derivative as

$$y'' = \frac{d^2y}{dx^2} = \frac{d}{dp}(y')\frac{dp}{dx} = \frac{1}{h^2}\left[\nabla^2 y_n + (p+1)\nabla^3 y_n + \frac{6p^2+18p+11}{12}\nabla^4 y_n + \cdots\right] \quad (7.9c)$$

Equations (7.9b) and (7.9c) are *Newton's backward interpolation formulae*, 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 formulae can be derived to compute the first- and higher-order derivatives near the beginning of the table of values.

Case III: Using central difference operator δ and following the definitions of differential operator D, central difference operator δ and the shift operator E, we have

$$\delta = E^{1/2} - E^{-1/2} = e^{hD/2} - e^{-hD/2} = 2\sinh\frac{hD}{2}$$

Therefore, we find

$$\frac{hD}{2} = \sinh^{-1}\frac{\delta}{2} \quad (7.10)$$

But,

$$\sinh^{-1} x = x - \frac{1}{2}\frac{x^3}{3} + \frac{1\times 3}{2\times 4}\frac{x^5}{5} - \frac{1\times 3\times 5}{2\times 4\times 6}\frac{x^7}{7} + \cdots$$

Using this expansion into Eq. (7.10), we get

$$\frac{hD}{2} = \frac{\delta}{2} - \frac{\delta^3}{6\times 8} + \frac{3}{40\times 32}\delta^5 - \cdots$$

That is,

$$D = \frac{1}{h}\left(\delta - \frac{1}{24}\delta^3 + \frac{3}{640}\delta^5 - \cdots\right)$$

Therefore,

$$\frac{d}{dx}y = y' = Dy = \frac{1}{h}\left(\delta y - \frac{1}{24}\delta^3 y + \frac{3}{640}\delta^5 y - \cdots\right) \quad (7.11)$$

$\sinh^{-1}x = \log(x + \sqrt{x^2+1})$ by taking $\sinh^{-1}x = w$, $x = \sinh w = \frac{e^w - e^{-w}}{2} = \frac{e^{2w}-1}{2e^w}$, $2e^w x = e^{2w}-1$ or $e^{2w} - 2e^w x - 1 = 0$. Put $e^w = t$: $t^2 - 2tx - 1 = 0$

Also

$$D^2 = \frac{1}{h^2}\left(\delta^2 - \frac{1}{12}\delta^4 + \frac{1}{90}\delta^6 - \cdots\right)$$

Hence,

$$y'' = D^2 y = \frac{1}{h^2}\left(\delta^2 y - \frac{1}{12}\delta^4 y + \frac{1}{90}\delta^6 y - \cdots\right) \tag{7.12}$$

For calculating the second derivative at an interior tabular point, we use Eq. (7.12), while for computing the first derivative at an interior tabular point, we in general use another convenient form for D, which is derived as follows.

Multiplying the right-hand side of Eq. (7.11) by

$$\frac{\mu}{\sqrt{1 + (\delta^2/4)}}$$

which is unity and noting the Binomial expansion

$$\left(1 + \frac{1}{4}\delta^2\right)^{-1/2} = 1 - \frac{1}{8}\delta^2 + \frac{3}{128}\delta^4 - \frac{15}{48 \times 64}\delta^6 + \cdots$$

We get

$$D = \frac{\mu}{h}\left(1 - \frac{1}{8}\delta^2 + \frac{3}{128}\delta^4 - \cdots\right)\left(\delta - \frac{1}{24}\delta^3 + \frac{3}{640}\delta^5 - \cdots\right)$$

On simplification, we obtain

$$D = \frac{\mu}{h}\left(\delta - \frac{1}{6}\delta^3 + \frac{4}{120}\delta^5 - \cdots\right)$$

Therefore,

$$y' = Dy = \frac{\mu}{h}\left(\delta y - \frac{1}{6}\delta^3 y + \frac{1}{30}\delta^5 y - \cdots\right) \tag{7.13}$$

Equations (7.12) and (7.13) are known as Stirling's formulae for computing the derivatives of a tabular function. Similar formulae can be derived for computing higher order derivatives of a tabular function. Equation (7.13) can also be written in another useful form as

$$y' = \frac{\mu}{h}\left[\delta y - \frac{1^2}{3!}\delta^3 y + \frac{(1^2)(2^2)}{5!}\delta^5 y - \frac{(1^2)(2^2)(3^2)}{7!}\delta^7 y + \cdots\right] \tag{7.14}$$

In order to illustrate the use of formulae derived so far, for computing the derivatives of a tabulated function, we shall consider the following examples.

Example 7.1 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

Solution Since $x = 0$ and 0.2 appear at and near beginning of the table, it is appropriate to use formulae based on forward differences to find the derivatives. The difference table for the given data is depicted below:

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	2.24	5.76	3.84	0.00
0.2	1.16	2.40	8.00	9.60	3.84	
0.4	3.56	10.40	17.60	13.44		
0.6	13.96	28.00	31.04			
0.8	41.96	59.04				
1.0	101.00					

Using forward difference formula (7.5) for $D^2 f(x)$, i.e.

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

We obtain

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

Also, using the formula (7.3) we have

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

Hence,

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

Example 7.2 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

Solution Since $x = 2.2$ occurs at the end of the table, it is appropriate to use backward difference formulae for derivatives. The backward difference table for the given data is shown as follows:

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				

Using backward difference formulae (7.8) and (7.9) for $y'(x)$ and $y''(x)$, we have

$$y'_n = \frac{1}{h}\left(\nabla y_n + \frac{\nabla^2 y_n}{2} + \frac{\nabla^3 y_n}{3} + \frac{\nabla^4 y_n}{4}\right)$$

Therefore,

$$y'(2.2) = \frac{1}{0.2}\left(1.6359 + \frac{0.2964}{2} + \frac{0.0535}{3} + \frac{0.0094}{4}\right) = 5(1.8043) = 9.0215$$

Also

$$y''_n = \frac{1}{h^2}\left(\nabla^2 y_n + \nabla^3 y_n + \frac{11}{12}\nabla^4 y_n\right)$$

Therefore,

$$y''(2.2) = \frac{1}{(0.2)^2}\left[0.2964 + 0.0535 + \frac{11}{12}(0.0094)\right] = 25\,(0.3585) = 8.9625$$

Example 7.3 From the following table of values, estimate $y'(2)$ and $y''(2)$ using appropriate central difference formula:

x	0	1	2	3	4
y	6.9897	7.4036	7.7815	8.1281	8.4510

Solution The central difference table for the given data is given below:

x	y	δy	$\delta^2 y$	$\delta^3 y$	$\delta^4 y$
0	6.9897				
		0.4139			
1	7.4036		−0.0360		
		0.3779		0.0047	
2	7.7815		−0.0313		0.0029
		0.3466		0.0076	
3	8.1281		−0.0237		
		0.3229			
4	8.4510				

Now, using central difference formula (7.13), we shall compute the first derivative

$$y' = \frac{\mu}{h}\left(\delta y - \frac{1}{6}\delta^3 y + \frac{1}{30}\delta^5 y - \cdots\right)$$

In the present example

$$y'(2) = \frac{1}{1}\left(\frac{0.3779 + 0.3466}{2} - \frac{1}{6}\,\frac{0.0047 + 0.0076}{2}\right) = 0.3613$$

To compute the second derivative, we shall use formula (7.12). Thus,

$$y'' = \frac{1}{h^2}\left(\delta^2 y - \frac{1}{12}\delta^4 y + \frac{1}{90}\delta^6 y - \ldots\right)$$

In this example,

$$y''(2) = \frac{1}{1}\left(-0.0313 - \frac{0.0029}{12}\right) = -0.0315$$

Case IV: (Two- and three-point formulae): Retaining only the first term in Eq. (7.3), we can get another useful form for the first derivative as

$$y_i' = \frac{\Delta y_i}{h} = \frac{y_{i+1} - y_i}{h} = \frac{y(x_i + h) - y(x_i)}{h} \tag{7.15}$$

Similarly, by retaining only the first term in Eq. (7.8), we obtain

$$y_i' = \frac{\nabla y_i}{h} = \frac{y_i - y_{i-1}}{h} = \frac{y(x_i) - y(x_i - h)}{h} \tag{7.16}$$

Adding Eqs. (7.15) and (7.16), we have

$$y_i' = \frac{y(x_i + h) - y(x_i - h)}{2h} \tag{7.17}$$

Equations (7.15)–(7.17) constitute two-point formulae for the first derivative. By retaining only the first term in Eq. (7.5), we get

$$y_i'' = \frac{\Delta^2 y_i}{h^2} = \frac{y_{i+2} - 2y_{i+1} + y_i}{h^2} = \frac{y(x_i + 2h) - 2y(x_i + h) + y(x_i)}{h^2} \tag{7.18}$$

Similarly, Eq. (7.9) gives

$$y_i'' = \frac{\Delta^2 y_i}{h^2} = \frac{y(x_i) - 2y(x_i - h) + y(x_i - 2h)}{h^2} \tag{7.19}$$

While retaining only the first term in Eq. (7.12), we obtain

$$y_i'' = \frac{\delta^2 y_i}{h^2} = \frac{\delta y_{i+(1/2)} - \delta y_{i-(1/2)}}{h^2} = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}$$

$$= \frac{y(x_i - h) - 2y(x_i) + y(x_i + h)}{h^2} \tag{7.20}$$

Equations (7.18)–(7.20) constitute three-point formulae for computing the second derivative. We shall see later that these two- and three-point formulae become handy for developing extrapolation methods to numerical differentiation and integration.

7.3 DIFFERENTIATION USING INTERPOLATION

If the given tabular function $y(x)$ is reasonably well approximated by a polynomial $P_n(x)$ of degree n, it is hoped that the result of $P_n'(x)$ will also satisfactorily approximate the corresponding derivative of $y(x)$. However, even if $P_n(x)$ and $y(x)$ coincide at the tabular points, their derivatives or slopes may substantially differ at these points as is illustrated in Fig. 7.1.

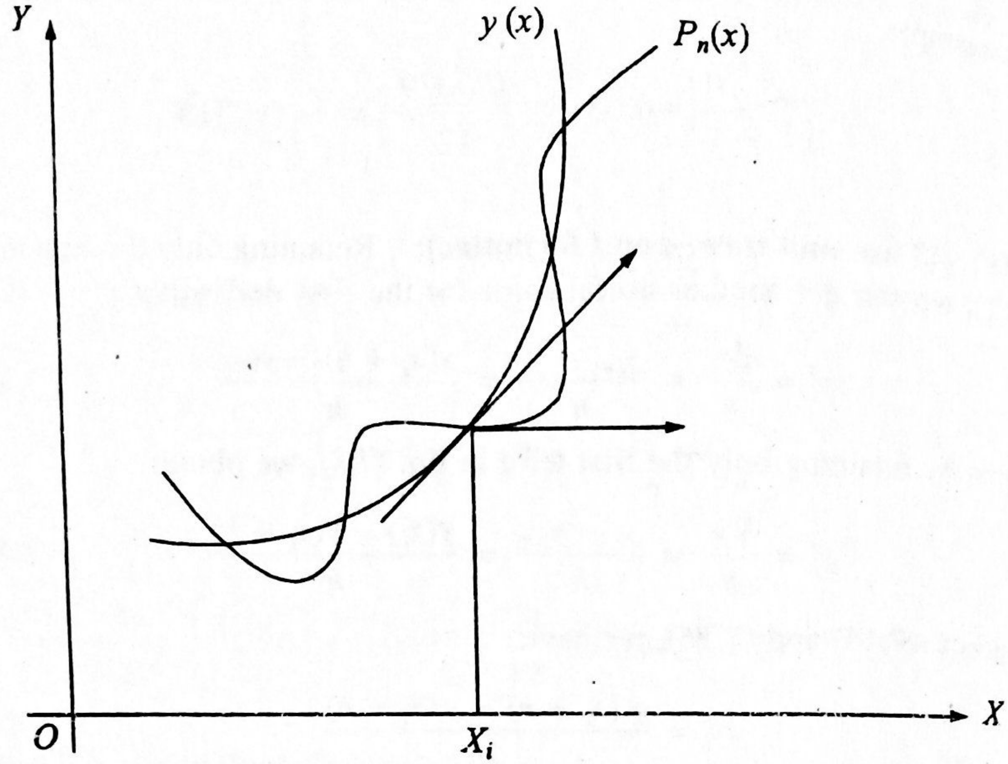

Fig. 7.1 Deviation of derivatives

For higher order derivatives, the deviations may be still worst. However, we can estimate the error involved in such an approximation.

For non-equidistant tabular pairs (x_i, y_i), $i = 0, \ldots, n$ we can fit the data by using either Lagrange's interpolating polynomial or by using Newton's divided difference interpolating polynomial. In view of economy of computation, we prefer the use of the latter polynomial. Thus, recalling the Newton's divided difference interpolating polynomial for fitting this data as

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

$$+ \prod_{i=0}^{n-1} (x - x_i)y[x_0, x_1, \ldots, x_n] \tag{7.21}$$

Assuming that $P_n(x)$ is a good approximation to $y(x)$, the polynomial approximation to $y'(x)$ can be obtained by differentiating $P_n(x)$. Using product rule of differentiation, the derivative of the products in $P_n(x)$ can be seen as follows:

$$\frac{d}{dx}\prod_{i=0}^{n-1} (x - x_i) = \sum_{i=0}^{n-1} \frac{(x - x_0)(x - x_1)\cdots(x - x_{n-1})}{x - x_i}$$

Thus, $y'(x)$ is approximated by $P_n'(x)$ which is given by

$$P_n'(x) = y[x_0, x_1] + [(x - x_1) + (x - x_0)]\, y[x_0, x_1, x_2] + \cdots$$
$$+ \sum_{i=0}^{n-1} \frac{(x - x_0)(x - x_1)\cdots(x - x_{n-1})}{x - x_i}\, y[x_0, x_1, \cdots, x_n] \tag{7.22}$$

The error estimate in this approximation can be seen from the following.

In Section 6.6.3, we have seen that if $y(x)$ is approximated by $P_n(x)$, the error estimate is shown to be

$$E_n(x) = y(x) - P_n(x) = \frac{\Pi(x)}{(n+1)!}\, y^{(n+1)}(\xi) \tag{7.23}$$

Its derivative with respect to x can be written as

$$E_n'(x) = y'(x) - P_n'(x) = \frac{\Pi'(x)}{(n+1)!} y^{(n+1)}(\xi) + \frac{\Pi(x)}{(n+1)!} \frac{d}{dx} y^{(n+1)}(\xi) \tag{7.24}$$

Since $\xi(x)$ depends on x in an unknown way the derivative

$$\frac{d}{dx} y^{(n+1)}(\xi)$$

cannot be evaluated. However, for any of the tabular points $x = x_i$, $\Pi(x)$ vanishes and the difficult term drops out. Thus, the error term in Eq. (7.24) at the tabular point $x = x_i$ simplifies to

$$E_n'(x_i) = \text{Error} = \Pi'(x_i)\frac{y^{(n+1)}(\xi)}{(n+1)!} \tag{7.25}$$

for some ξ in the interval I defined by the smallest and largest of $x, x_0, x_1, \ldots, x_n$ and

$$\Pi'(x_i) = (x_i - x_0)\cdots(x_i - x_n) = \prod_{\substack{j=0 \\ j\neq i}}^{n} (x_i - x_j) \tag{7.26}$$

The error in the rth derivative at the tabular points can indeed be expressed analogously.

To understand this method better, we consider the following example.

Example 7.4 Find $y'(0.25)$ and $f''(0.25)$ from the following data using the method based on divided differences:

x	0.15	0.21	0.23	0.27	0.32	0.35
$y = f(x)$	0.1761	0.3222	0.3617	0.4314	0.5051	0.5441

Solution We first construct divided difference table for the given data as follows:

x	y	1st divided difference	2nd divided difference	3rd divided difference	4th divided difference	5th divided difference
$x_0 = 0.15$	0.1761					
		2.4350				
$x_1 = 0.21$	0.3222		−5.7500			
		1.9750		15.6250		
$x_2 = 0.23$	0.3617		−3.8750		−44.23	
		1.7425		8.1064		172.2
$x_3 = 0.27$	0.4314		−2.9833		−9.79	
		1.4740		6.7358		
$x_4 = 0.32$	0.5051		−2.1750			
		1.3000				
$x_5 = 0.35$	0.5441					

Using Newton's divided difference formula (7.21), we have

$$y(x) = p_5(x) = y[x_0] + (x - x_0)\, y[x_0, x_1] + (x - x_0)(x - x_1)\, y[x_0, x_1, x_2]$$
$$+ (x - x_0)(x - x_1)(x - x_2)\, y[x_0, x_1, x_2, x_3]$$
$$+ (x - x_0)(x - x_1)(x - x_2)(x - x_3) y[x_0, x_1, x_2, x_3, x_4]$$

Now, using values from the above table of divided differences, we obtain

$$y(x) = 0.1761 + (x - 0.15)\, 2.4350 + (x - 0.15)\,(x - 0.21)\,(-5.75)$$
$$+ (x - 0.15)\,(x - 0.21)\,(x - 0.23) 15.625$$
$$+ (x - 0.15)\,(x - 0.21)\,(x - 0.23)\,(x - 0.27)\,(-44.23)$$
$$+ (x - 0.15)\,(x - 0.21)\,(x - 0.23)\,(x - 0.27)\,(x - 0.32)\, 172.2 \quad (1)$$

Differentiating Eq. (1) with respect to x, we get

$$y'(x) = 2.4350 - (2x - 0.36)5.75 + 15.625(3x^2 - 1.18x + 0.1143)$$
$$- 44.23(4x^3 - 2.58x^2 + 0.5472x - 38.105 \times 10^{-3})$$
$$+ 172.2(5x^4 - 4.72x^3 + 1.6464x^2 - 0.2515x + 14.15 \times 10^{-3}) \quad (2)$$

Which immediately gives

$$y'(0.25) = 2.4350 - 0.805 + 0.10625 + 2.432 \times 10^{-3} - 7.5338 \times 10^{-3} = 1.7312$$

Now, differentiating Eq. (2) once again with respect to x, we obtain

$$y''(x) = 3444x^3 - 2969.112x^2 + 888.99696x - 91.700456$$

which gives at once

$$y''(0.25) = 53.8125 - 185.5695 + 222.24924 - 91.700456 = -1.208216$$

7.4 RICHARDSON'S EXTRAPOLATION METHOD

To improve the accuracy of the derivative of a function, which is computed by starting with an arbitrarily selected value of h, Richardson's extrapolation method is often employed in practice, as explained below:

Suppose we use two-point formula (7.17) to compute the derivative of a function, then we have

$$y'(x) = \frac{y(x+h) - y(x-h)}{2h} + E_T = F(h) + E_T$$

where E_T is the *truncation error*. Using Taylor's series expansion, we can see that

$$E_T = c_1 h^2 + c_2 h^4 + c_3 h^6 + \cdots$$

The idea of *Richardson's extrapolation* is to combine two computed values of $y'(x)$ using the same method but with two different step sizes usually h and $h/2$ to yield a higher order method. Thus, we have

$$y'(x) = F(h) + c_1 h^2 + c_2 h^4 + \cdots$$

and

$$y'(x) = F\left(\frac{h}{2}\right) + c_1 \frac{h^2}{4} + c_2 \frac{h^4}{16} + \cdots$$

Here, c_i are constants, independent of h, and $F(h)$ and $F(h/2)$ represent approximate values of derivatives. Eliminating c_1 from the above pair of equations, we get

$$y'(x) = \frac{4F\left(\frac{h}{2}\right) - F(h)}{3} + d_1 h^4 + O(h^6) \tag{7.27}$$

Now, assuming

$$F_1\left(\frac{h}{2}\right) = \frac{4F\left(\frac{h}{2}\right) - F(h)}{3} \tag{7.28}$$

Equation (7.27) reduces to

$$y'(x) = F_1\left(\frac{h}{2}\right) + d_1 h^4 + O(h^6)$$

Thus, we have obtained a fourth-order accurate differentiation formula by combining two results which are of second-order accurate. Now, repeating the above argument, we have

$$y'(x) = F_1\left(\frac{h}{2}\right) + d_1 h^4 + O(h^6)$$

$$y'(x) = F_1\left(\frac{h}{4}\right) + \frac{d_1 h^4}{16} + O(h^6)$$

Eliminating d_1 from the above pair of equations, we get a better approximation as

$$y'(x) = F_2\left(\frac{h}{4}\right) + O(h^6)$$

Again, using Eq. (7.30), we obtain

$$F_2\left(\frac{h}{2^2}\right) = \frac{4^2 F_1\left(\frac{h}{2^2}\right) - F_1\left(\frac{h}{2}\right)}{4^2 - 1} = 400.00195 \qquad (5)$$

The above computation can be summarized in the following table:

h	F	F_1	F_2
0.0128	428.0529		
		399.5327	
0.0064	406.6627		400.00195
		399.9726	
0.0032	401.6452		

Thus, after two steps, it is found that $y'(0.05) = 400.00195$ while the exact value is

$$y'(0.05) = \left(\frac{1}{x^2}\right)_{x=0.05} = \frac{1}{0.0025} = 400$$

7.5 NUMERICAL INTEGRATION

Consider the definite integral

$$I = \int_{x=a}^{b} f(x)\, dx \qquad (7.31)$$

where $f(x)$ is known either explicitly or is given as a table of values corresponding to some values of x, whether equispaced or not. Integration of such functions can be carried out using numerical techniques. Of course, we assume that the function to be integrated is smooth and Riemann integrable in the interval of integration. In the following section, we shall develop Newton–Cotes formulae based on interpolation which form the basis for trapezoidal rule and Simpson's rule of numerical integration.

7.6 NEWTON–COTES INTEGRATION FORMULAE

In this method, as in the case of numerical differentiation, we shall approximate the given tabulated function, by a polynomial $P_n(x)$ and then integrate this polynomial. Suppose, we are given the data (x_i, y_i), $i = 0\,(1)\,n$, at equispaced points with spacing $h = x_{i+1} - x_i$, we can represent the polynomial by any standard interpolation polynomial. Suppose, we use Lagrangian approximation given by Eq. (6.45), then we have

$$f(x) \approx \sum L_k(x) y(x_k) \qquad (7.32)$$

with associated error given by

$$E(x) = \frac{\Pi(x)}{(n+1)!} y^{(n+1)}(\xi) \tag{7.33}$$

where

$$l_k(x) = \frac{\Pi(x)}{(x - x_k)\Pi'(x_k)} \tag{7.34}$$

and

$$\Pi(x) = (x - x_0)(x - x_1) \dots (x - x_n) \tag{7.35}$$

Then, we obtain an equivalent integration formula to the definite integral (7.31) in the form

$$\int_a^b f(x)\, dx = \sum_{k=1}^{n} c_k y(x_k) \tag{7.36}$$

where c_k are the weighting coefficients given by

$$c_k = \int_a^b L_k(x)\, dx \tag{7.37}$$

which are also called *Cotes numbers*. Let the equispaced nodes are defined by

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

so that $x_k - x_1 = (k-1)h$ etc. Now, we shall change the variable x to p such that, $x = x_0 + ph$, then we can rewrite Eqs. (7.35) and (7.34) respectively as

$$\Pi(x) = h^{n+1} p(p-1) \dots (p-n) \tag{7.38}$$

and

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

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

or

$$L_k(x) = (-1)^{(n-k)} \frac{p(p-1)\cdots(p-k+1)(p-k-1)\cdots(p-n)}{k!(n-k)!} \tag{7.39}$$

Also, noting that $dx = h\,dp$. The limits of the integral in Eq. (7.37) change from 0 to n and Eq. (7.37) reduces to

$$c_k = \frac{(-1)^{n-k} h}{k!(n-k)!} \int_0^n p(p-1) \cdots (p-k+1)(p-k-1) \cdots (p-n)dp \tag{7.40}$$

The error in approximating the integral (7.36) can be obtained by substituting (7.38) into Eq. (7.33) in the form

$$E_n = \frac{h^{n+2}}{(n+1)!}\int_0^n p(p-1)\cdots(p-n)y^{(n+1)}(\xi)\,dp \tag{7.41}$$

where $x_0 < \xi < x_n$. For illustration, let us consider the cases for $n = 1, 2$. From Eq. (7.40), we get

$$c_0 = -h\int_0^1 (p-1)\,dp = \frac{h}{2}, \qquad c_1 = h\int_0^1 p\,dp = \frac{h}{2}$$

and Eq. (7.41) gives

$$E_1 = \frac{h^3}{2}y''(\xi)\int_0^1 p(p-1)\,dp = -\frac{h^3}{12}y''(\xi)$$

Thus, the integration formula corresponding to integral (7.36) is found to be

$$\int_{x_0}^{x_1} f(x)\,dx = c_0 y_0 + c_1 y_1 + \text{Error} = \frac{h}{2}(y_0 + y_1) - \frac{h^3}{12}y''(\xi) \tag{7.42}$$

This equation represents the Trapezoidal rule in the interval $[x_0, x_1]$ with error term. Geometrically, it represents an area between the curve $y = f(x)$, the x-axis and the ordinates erected at $x = x_0$ $(=a)$ and $x = x_1$ as shown in Fig. 7.2. This area is approximated by the trapezium formed by replacing the curve with its secant line drawn between the end points (x_0, y_0) and (x_1, y_1).

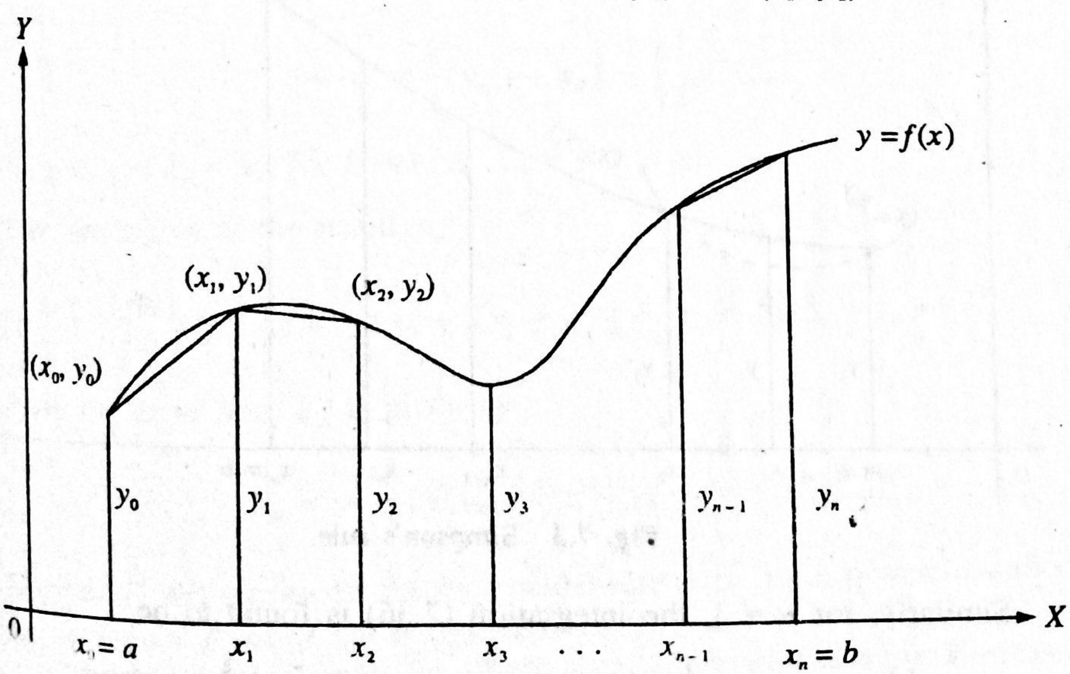

Fig. 7.2 Trapezoidal rule.

For $n = 2$, Eq. (7.40) gives

$$c_0 = \frac{h}{2}\int_0^2 (p-1)(p-2)\,dp = \frac{h}{3}$$

$$c_1 = -h\int_0^2 p(p-2)\,dp = \frac{4}{3}h$$

$$c_2 = \frac{h}{2}\int_0^2 p(p-1)\,dp = \frac{h}{3}$$

and the error term is given by

$$E_2 = -\frac{h^5}{90}y^{(iv)}(\xi)$$

Thus, for $n = 2$, the integration (7.36) takes the form

$$\int_{x_0}^{x_2} f(x)\,dx = c_0y_0 + c_1y_1 + c_2y_2 + \text{Error}$$

$$= \frac{h}{3}(y_0 + 4y_1 + y_2) - \frac{h^5}{90}y^{(iv)}(\xi) \tag{7.43}$$

This is known as *Simpson's 1/3 rule*. Geometrically, this equation represents 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 Fig. 7.3. Thus Simpson's 1/3 rule is based on fitting three points with a quadratic.

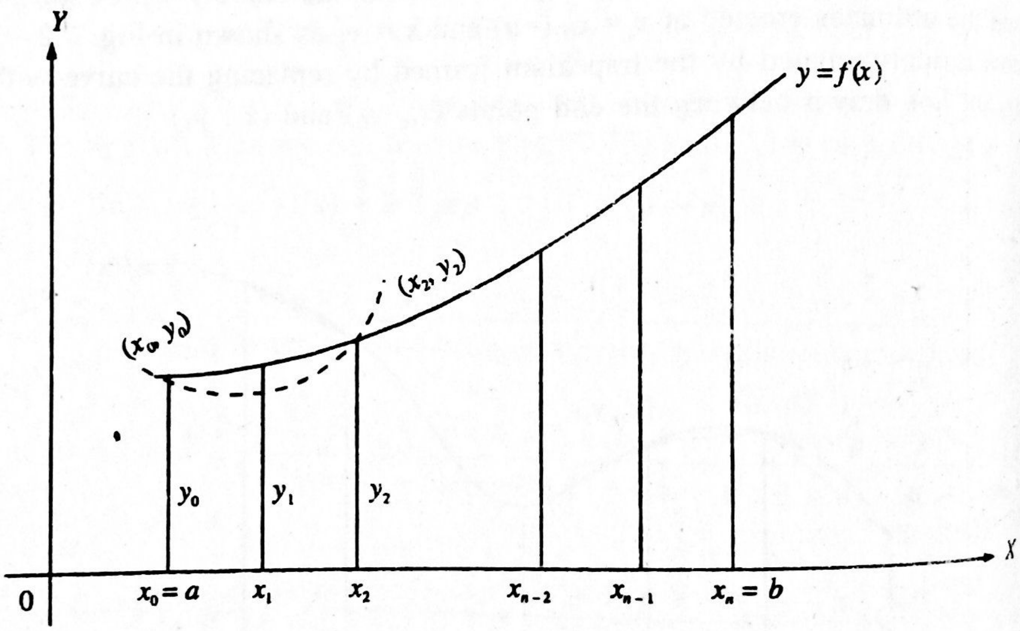

Fig. 7.3 Simpson's rule.

Similarly, for $n = 3$, the integration (7.36) is found to be

$$\int_{x_0}^{x_3} f(x)\,dx = \frac{3}{8}h(y_0 + 3y_1 + 3y_2 + y_3) - \frac{3}{80}h^5y^{(iv)}(\xi) \tag{7.44}$$

This is known as *Simpson's 3/8 rule*, which is based on fitting four points by a cubic. Still higher order Newton–Cotes integration formulae can be derived for large values of n. But for all practical purposes, Simpson's 1/3 rule is found to be sufficiently accurate.

we also derive Rectangular and Mid-Point Rule.

7.6.1 The Trapezoidal Rule (Composite Form)

The Newton–Cotes formula (7.42) 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*. In order to evaluate the definite integral

$$I = \int_a^b f(x)\,dx$$

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

$$I = \int_{x_0}^{x_n} f(x)\,dx = \int_{x_0}^{x_1} f(x)\,dx + \int_{x_1}^{x_2} f(x)\,dx + \cdots + \int_{x_{n-1}}^{x_n} f(x)\,dx \tag{7.45}$$

As shown in Fig. 7.2, 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.

Now, using the trapezoidal rule as expressed in Eq. (7.42) into Eq. (7.45), we get

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

$$+ \cdots + \frac{h}{2}(y_{n-1} + y_n) - \frac{h^3}{12}y''(\xi_n) \tag{7.46}$$

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

Thus, we arrive at the result

$$\int_{x_0}^{x_n} f(x)\,dx = \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n) + E_n \tag{7.47}$$

where the error term E_n is given by

$$E_n = -\frac{h^3}{12}[y''(\xi_1) + y''(\xi_2) + \cdots + y''(\xi_n)] \tag{7.48}$$

Equation (7.47) represents the trapezoidal rule over $[x_0, x_n]$, which is also called the *composite form of the trapezoidal rule.*

The error term given by Eq. (7.48) is called the *global error*. However, 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

$$E_n = -\frac{h^3}{12}[ny''(\xi)] = -\frac{x_n - x_0}{12}h^2 y''(\xi) \tag{7.49}$$

Then the global error can be conveniently written as $O(h^2)$.

7.6.2 Simpson's Rules (Composite Forms)

In deriving Eq. (7.43), the Simpson's 1/3 rule, we have used two sub-intervals of equal width. In order to get a composite formula, we shall divide the interval of integration $[a, b]$ into an even number of sub-intervals say $2N$, each of width $(b-a)/2N$, thereby we have $x_0 = a, x_1, \ldots, x_{2N} = b$ and $x_k = x_0 + kh$, $k = 1, 2, \ldots, (2N-1)$. Thus, the definite integral I can be written as

$$I = \int_a^b f(x)\,dx = \int_{x_0}^{x_2} f(x)\,dx + \int_{x_2}^{x_4} f(x)\,dx + \cdots + \int_{x_{2N-2}}^{x_{2N}} f(x)\,dx \quad (7.50)$$

Applying Simpson's 1/3 rule as in Eq. (7.43) to each of the integrals on the right-hand side of Eq. (7.50), we obtain

$$I = \frac{h}{3}[(y_0 + 4y_1 + y_2) + (y_2 + 4y_3 + y_4) + \cdots + (y_{2N-2} + 4y_{2N-1} + y_{2N})] - \frac{N}{90}h^5 y^{(iv)}(\xi)$$

That is,

$$\int_{x_0}^{x_{2N}} f(x)\,dx = \frac{h}{3}[y_0 + 4(y_1 + y_3 + \cdots + y_{2N-1}) + 2(y_2 + y_4 + \cdots + y_{2N-2}) + y_{2N}] + \text{Error term} \quad (7.51)$$

This formula is called *composite Simpson's 1/3 rule*. The error term E, which is also called *global error*, is given by

$$E = -\frac{N}{90}h^5 y^{(iv)}(\xi) = -\frac{x_{2N} - x_0}{180} h^4 y^{(iv)}(\xi) \quad (7.52)$$

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

Similarly in deriving composite 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 (7.44) to each of the integral given below

$$\int_{x_0}^{x_n} f(x)\,dx = \int_{x_0}^{x_3} f(x)\,dx + \int_{x_3}^{x_6} f(x)\,dx + \cdots + \int_{x_{n-3}}^{x_n} f(x)\,dx$$

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

$$\int_a^b f(x)\,dx = \frac{3}{8}h[y(a) + 3y_1 + 3y_2 + 2y_3 + 3y_4 + 3y_5 + 2y_6 + \cdots + 2y_{n-3} + 3y_{n-2} + 3y_{n-1} + y(b)] \quad (7.53)$$

with the global error E given by

$$E = -\frac{x_n - x_0}{80}h^4 y^{(iv)}(\xi) \quad (7.54)$$

It may be noted from Eqs. (7.52) and (7.54), the global error in Simpson's 1/3 and 3/8 rules are of the same order. However, if we consider the magnitudes of the error terms, we notice that Simpson's 1/3 rule is superior to Simpson's 3/8 rule. For illustration, we consider few examples.

Example 7.6 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.

Solution We shall at first divide the range of integration $(0, \pi)$ into six equal parts so that each part is of width $\pi/6$ and write down the table of values:

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

Applying trapezoidal rule, we have

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

Here, h, the width of the interval is $\pi/6$. Therefore,

$$y = \int_0^{\pi} \sin x \, dx = \frac{\pi}{12} [0 + 0 + 2(3.732)] = \frac{3.1415}{6} \times 3.732 = 1.9540$$

Applying Simpson's 1/3 rule (7.41), we have

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

$$= \frac{\pi}{18} [0 + 0 + (4 \times 2) + (2)(1.732)] = \frac{3.1415}{18} \times 11.464 = 2.0008$$

But the actual value of the integral is

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

Hence, in the case of trapezoidal rule

$$\text{The percentage of error} = \frac{2 - 1.954}{2} \times 100 = 2.3$$

While in the case of Simpson's rule the percentage error is

$$\frac{2 - 2.0008}{2} \times 100 = 0.04 \qquad \text{(sign ignored)}$$

Example 7.7 From the following data, estimate the value of

$$\int_1^5 \log x \, dx$$

using Simpson's 1/3 rule. Also, obtain the value of h, so that the value of the integral will be accurate up to five decimal places.

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

Solution We have from the data, $n = 0, 1, \ldots, 8$, and $h = 0.5$. Now using Simpson's 1/3 rule,

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

$$= \frac{0.5}{3}[(0 + 1.6094) + 4(4.0787) + 2(3.178)]$$

$$= \frac{0.5}{3}(1.6094 + 16.3148 + 6.356)$$

$$= 4.0467$$

The error in Simpson's rule is given by

$$E = \frac{x_{2N} - x_0}{180} h^4 y^{(iv)}(\xi) \qquad \text{(ignoring the sign)}$$

Since

$$y = \log x, \qquad y' = \frac{1}{x}, \qquad y'' = -\frac{1}{x^2}, \qquad y''' = \frac{2}{x^3}, \qquad y^{(iv)} = -\frac{6}{x^4}$$

$$\max_{1 \le x \le 5} y^{(iv)}(x) = 6, \qquad \min_{1 \le x \le 5} y^{(iv)}(x) = 0.0096$$

Therefore, the error bounds are given by

$$\frac{(0.0096)(4)h^4}{180} < E < \frac{(6)(4)h^4}{180}$$

If the result is to be accurate up to five decimal places, then

$$\frac{24h^4}{180} < 10^{-5}$$

That is, $h^4 < 0.000075$ or $h < 0.09$. It may be noted that the actual value of integral is

$$\int_1^5 \log x \, dx = [x \log x - x]_1^5 = 5 \log 5 - 4$$

Example 7.8 Evaluate the integral

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

using (i) trapezoidal rule, (ii) Simpson's 1/3 rule by taking $h = 1/4$. Hence, compute the approximate value of π.

Solution At first, we shall tabulate the function as

x	0	1/4	1/2	3/4	1
$y = \frac{1}{1+x^2}$	1	0.9412	0.8000	0.6400	0.5000

using trapezoidal rule, and taking $h = 1/4$

$$I = \frac{h}{2}[y_0 + y_4 + 2(y_1 + y_2 + y_3)] = \frac{1}{8}[1.5 + 2(2.312)] = 0.7828 \qquad (1)$$

using Simpson's 1/3 rule, and taking $h = 1/4$, we have

$$I = \frac{h}{3}[y_0 + y_4 + 4(y_1 + y_3) + 2y_2] = \frac{1}{12}[1.5 + 4(1.512) + 1.6] = 0.7854 \qquad (2)$$

But the closed form solution to the given integral is

$$\int_0^1 \frac{dx}{1 + x^2} + [\tan^{-1} x]_0^1 = \frac{\pi}{4} \qquad (3)$$

Equating (2) and (3), we get $\pi = 3.1416$.

Example 7.9 Compute the integral

$$I = \sqrt{\frac{2}{\pi}} \int_0^1 e^{-x^2/2} \, dx$$

using Simpson's 1/3 rule, taking $h = 0.125$.

Solution At the outset, we shall construct the table of the function as required.

x	0	0.125	0.250	0.375	0.5	0.625	0.750	0.875	1.0
$y = \sqrt{\frac{2}{\pi}} e^{-x^2/2}$	0.7979	0.7917	0.7733	0.7437	0.7041	0.6563	0.6023	0.5441	0.4839