Simpson's Rule and Friends

Newton–Cotes Formulas

These are a generalisation of Simpson's Rule. We take $n+1$ equally-spaced values of $f(x)$ for $x = a$, $x = a+h$, $x = a+2h$, $\ldots$, $x = a+nh = b$ [where $h = (b-a)/n$], interpolate f by a polynomial of degree n, and integrate that between a and b to estimate $\int_a^b f(x)\,\mathrm{d}x$.

In each case, the result is $(b-a)$ times a weighted average of f_0, f_1, f_2, $\ldots$, f_n, where f_k is short-hand for $f(x_k)$, where $x_k = a+kh$ [so that $x_0 = a$ and $x_n = b$.

n	Name	Formula	Error	I
1	Trapezoidal	$\frac{1}{2}(b-a)(f_0+f_1)$	$(b-a)^3 f''(\xi)/12$	0.750000
2	Simpson	$\frac{1}{6}(b-a)(f_0+4f_1+f_2)$	$(b-a)^5 f^{(4)}(\xi)/2880$	0.694444
3	Three-eighths	$\frac{1}{8}(b-a)(f_0+3f_1+3f_2+f_3)$	$(b-a)^5 f^{(4)}(\xi)/6480$	0.693750
4	Bode	$\frac{1}{90}(b-a)(7f_0+32f_1+12f_2+32f_3+7f_4)$	$(b-a)^7 f^{(6)}(\xi)/195360$	0.693175
6	Weddle	$\frac{1}{20}(b-a)(f_0+5f_1+f_2+6f_3+f_4+5f_5+f_6)$	messy	0.693149

Weddle's formula is not strictly a Newton-Cotes formula, but the coefficients have been slightly tweaked from that to be much simpler. The column labelled 'Error' gives the error in the formula; note that [apart from Weddle] it is always of the form $(b-a)^k f^{(m)}(\xi)/N$, where $f^{(m)}$ indicates the m-th derivative of f, and ξ is some value between a and b. [Not proved here!]

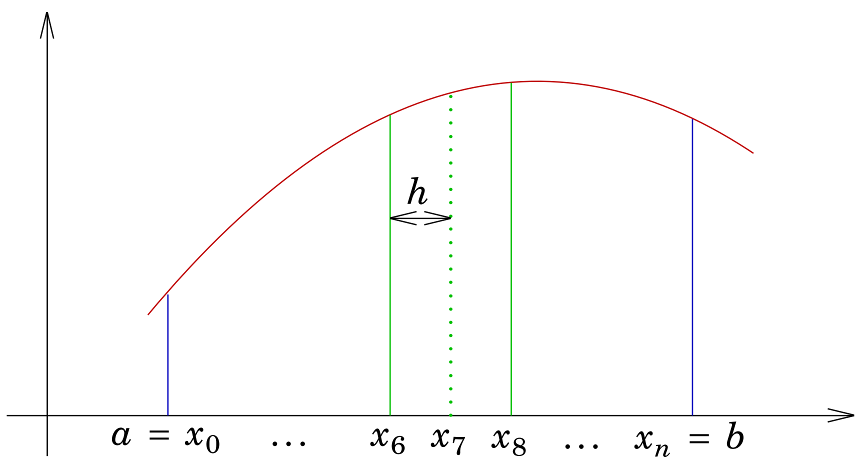
The column labelled I gives the estimate, according to that formula, of the integral $I = \int_1^2 1/x\,\mathrm{d}x = \log 2 \approx 0.693147$.

What should we do to get better estimates?

OK, we've tried, say, Simpson's Rule, and either we don't know how good the estimate is [because $f^{(4)}$ is too messy to contemplate] or we know it's not good enough for our purposes. What then?

What we must not do is go to formulas with larger and larger n. High-degree polynomials are not our friends. Note that Bode's Rule is sensible only if f has a 'nice' sixth derivative. This is OK when $f(x) = 1/x$ and $x \geqslant 1$; it very soon looks silly for 'real' functions.

More sensible is to go for a 'repeated' ['composite'] Simpson's Rule:



Divide the interval $a \leqslant x \leqslant b$ into n strips. Note that n **must** be even! Then treat the integral from a to b as the sum of $\frac{1}{2}n$ integrals, successively from $a = x_0$ to x_2, from x_2 to x_4, x_4 to x_6, ..., x_{n-2} to $x_n = b$. For each of these sub-integrals, $b-a$ is replaced by $2h$, where $h = (b-a)/n$ is the strip width. Thus

$$I = \int_a^b f(x)\,\mathrm{d}x \approx \tfrac{1}{3}h(f_0+4f_1+f_2)+\tfrac{1}{3}h(f_2+4f_3+f_4)+\tfrac{1}{3}h(f_4+4f_5+f_6)+\ldots,$$

or

$$I \approx \tfrac{1}{3}h(f_0+4f_1+2f_2+4f_3+2f_4+\ldots+4f_{n-1}+f_n).$$

The term in parentheses is often summarised as 'ends plus four times the odds plus twice the evens'.

What happens to the errors? Well, we now have separate contributions from each sub-integral. These are

$$(2h)^5 f^{(4)}(\xi_1)/2880 + (2h)^5 f^{(4)}(\xi_2)/2880 + (2h)^5 f^{(4)}(\xi_3)/2880 + \ldots$$

where ξ_1 lies between x_0 and x_2, where ξ_2 lies between x_2 and x_4, and so on. If we replace each $f^{(4)}(\xi_i)$ by some average value, and remember that there are $\frac{1}{2}n$ contributions, we see that the total error is

$$\tfrac{1}{2}n \times 2^5 \times h^5 \times f^{(4)}(\xi)/2880 \;=\; (b-a)h^4 f^{(4)}(\xi)/180.$$

[remembering that $nh = b-a$].

So, *provided that* $f^{(4)}$ is not too badly-behaved, we can make the error as small as we like by making h small.

Note the warning! You must **not** use Simpson's Rule if f has bad behaviour anywhere between a and b, inclusive. You are simply wasting your time. You *must* do something clever to f first—we'll see some ideas later.

'Standard operating procedure' is to start with $n = 2$, producing some I_2 as our first approximation to I. Then try with $n = 4$, producing I_4 as a better approximation, then with $n = 8$, producing I_8, and so on, doubling the number of strips each time. We can track how the approximations are converging, if at all. The big advantage of doubling is that we can re-use the old values; they all become 'evens' in the new estimate. Don't be tempted to short-cut this process; if you go straight to, say, $n = 8$, you have no idea how good or bad the approximation is until you do $n = 16$, and meanwhile you could have worked out I_4 virtually for free in the process. *Very* occasionally, it makes sense to start with, say, $n = 6$. **Never** start with n odd—you will have a function value left over, and it will never work.

Example

We work out our standard integral, $I = \int_1^2 1/x \, dx$. Firstly, the 'ends': these are $f(a) = 1/1 = 1$ and $f(b) = 1/2 = 0.500000$, so 'ends' $= 1.500000$. This will stay the same throughout the calculation. For I_2, there are no 'evens', and one 'odd' $= f(1.5) \approx 0.666667$. Further, the strip width is 0.5, so

$$I_2 \approx \tfrac{1}{3} \times 0.5 \times (1.500000 + 4 \times 0.666667 + 2 \times 0) \approx 0.694445$$

[the discrepancy in the sixth dp compared with our previous result is, of course, a rounding error caused by using 6dp throughout.]

On to $n = 4$. 'Ends' are the same, 1.500000. The new 'evens' are the old ('odds' plus 'evens') ≈ 0.666667. The new 'odds' are $f(1.25) + f(1.75) \approx 0.800000 + 0.571429 = 1.371429$. The new strip width is 0.25, so

$$I_4 \approx \tfrac{1}{3} \times 0.25 \times (1.500000 + 4 \times 1.371429 + 2 \times 0.666667) \approx 0.693254.$$

On to $n = 8$. 'Ends' are the same, 1.500000. The new 'evens' are the old ('odds' plus 'evens') $\approx 1.371429 + 0.666667 = 2.038096$. The new 'odds' are $f(1.125) + f(1.375) + f(1.625) + f(1.875) \approx 0.888889 + 0.727273 + 0.615385 + 0.533333 = 2.764880$. The new strip width is 0.125, so

$$I_8 \approx \tfrac{1}{3} \times 0.125 \times (1.500000 + 4 \times 2.764880 + 2 \times 2.038096) \approx 0.693155.$$

This is converging nicely to $\log 2 \approx 0.693147$. Note that the total work done—nine evaluations of f and a bit of arithmetic [and in serious work, it's working out f that takes all the time]—has given us better results than Bode's Rule [five evaluations] and very nearly as good as Weddle's Rule [seven]; but we *know* how well the composite Simpson's Rule is working. *And* there is a simple way to do better, for almost no work. …