# Numerical Integration and Monte Carlo Integration

Elementary schemes for integration over one variable

Multi-dimensional integration
- dimension-by dimension

Problems with multi-dimensional numerical integrations

Monte Carlo sampling of high-dimensionsl integrals
- includes some aspects of analysis of statistical data

# Numerical integration in one dimension

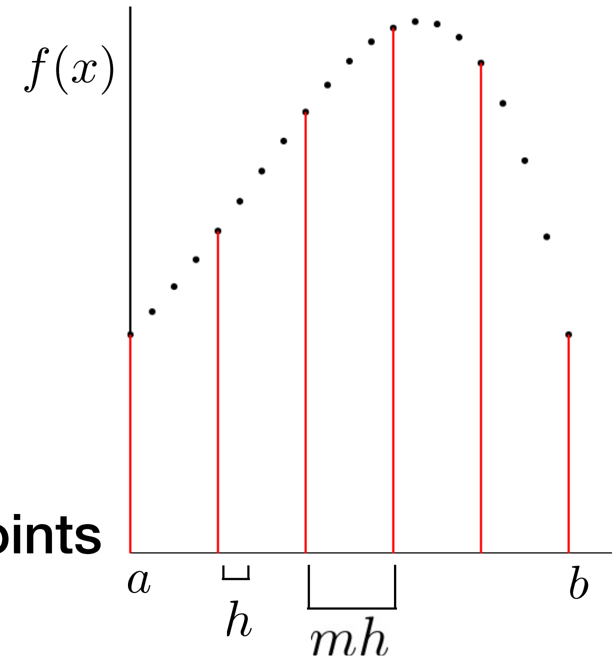
Function of one variable x, assume no singularities

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

Discretize the x-axis

- n+1 equally spaced points including a,b:

$$[a,b] \;\rightarrow\; \{x_0, x_1, \ldots, x_n\} \qquad h \equiv x_i - x_{i-1}$$




Consider groups of m+1 points (m intervals of size h)

Construct the order-m polynomials fitting the m+1 points

$$I = \sum_{i=1}^{n/m} I_i, \qquad I_i = \int_{a+(i-1)mh}^{a+imh} P_i(x)dx$$

Simple formulas exist to construct the polynomials $P_i(x)$

Integrate the polynomials exactly and add up

Leads to simple integration formulas (sums) for small m

Error for one window typically of order $O(h^{m+1})$ or $O(h^{m+2})$

## Simplest case; m=1 (trapezoidal rule)

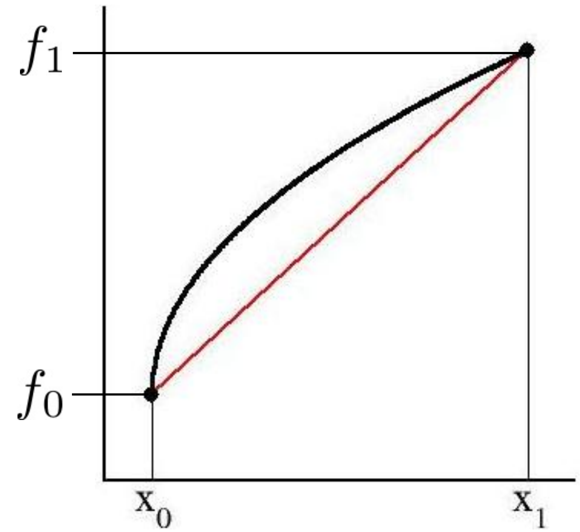
$$f(x_0+\delta) = a + b\delta, \quad 0 \le \delta \le h$$

$$f(x_0) = f_0, \quad f(x_0+h) = f_1$$

$$a = f_0, \quad b = (f_1 - f_0)/h$$

$$I_1 = \int_{x_0}^{x_1} P_1(x)dx = \int_0^h P_1(\delta)d\delta = \left[a\delta + b\delta^2/2\right]_0^h$$

$$= h(a + bh/2) = \frac{h}{2}(f_0 + f_1)$$




Here we also see that the error ("step error") is O($h^3$)

For the total error, we have to sum up step errors from (general m)

$$I = \sum_{i=1}^{n/m} I_i, \quad I_i = \int_{a+(i-1)mh}^{a+imh} P_i(x)dx \qquad n/m = \frac{x_n - x_0}{mh}$$

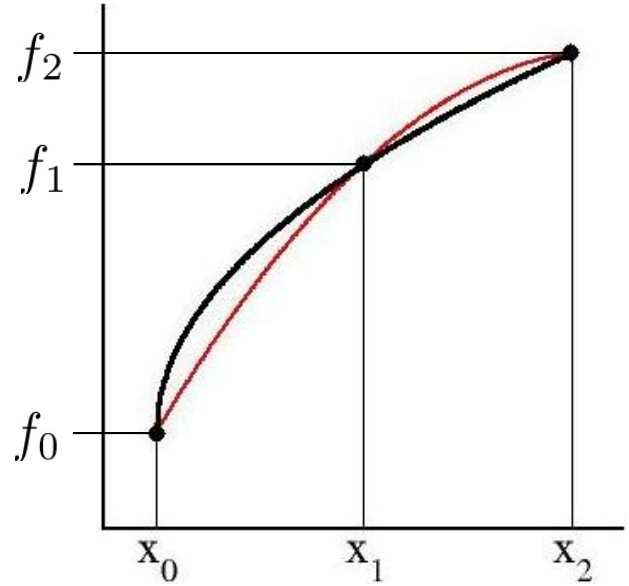
Assuming no "lucky" error cancellations (from sign oscillations)
- error is O($h^{m+2}$)O($h^{-1}$) ~ O($h^{m+1}$)

## Second order; m=2 (Simpson's rule)

$$f(x_0+\delta) = a + b\delta + c\delta^2, \quad 0 \le \delta \le 2h$$

$$f(x_0) = f_0, \quad f(x_0+h) = f_1, \quad f(x_0+2h) = f_2$$

$$f_0 = a, \quad f_1 = a + bh + ch^2, \quad f_2 = a + 2bh + 4ch^2$$




Solve for a,b,c, integrate polynomial →

$$I_1 = \int_{x_0}^{x_2} P_1(x)dx = \frac{h}{3}(f_0 + 4f_1 + f_2)$$

What is the order of the error?

- from the polynomial it may seem $O(h^4)$ (from missing integrated $\delta^3$ term)

Write expansion around $x_1$ instead:

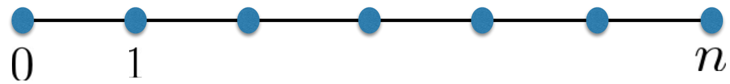
$$f(x_1+\delta) = a + b\delta + c\delta^2 + d\delta^3 + e\delta^4, \quad -h \le \delta \le h$$

When integrated in the symmetric window, all odd powers give 0

Same formula for $I_1$ as above when a,b,c terms included, d term gives 0

Error is $O(h^5)$, becomes $O(h^4)$ for range $[x_0,x_n]$

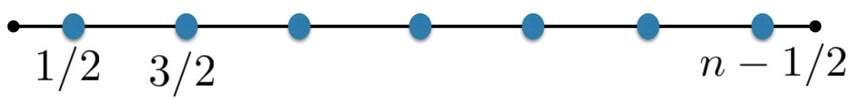
## Extended formulas




$$\int_{x_0}^{x_n} f(x)dx = h(\tfrac{1}{2}f_0 + f_1 + f_2 + f_3 + \ldots f_{n-1} + \tfrac{1}{2}f_n) + O(h^2)$$ trapezoid

$$\int_{x_0}^{x_n} f(x)dx = \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots 4f_{n-1} + f_n) + O(h^4)$$ Simpson

**For integrands with singularities at the end point(s); open interval formulas**

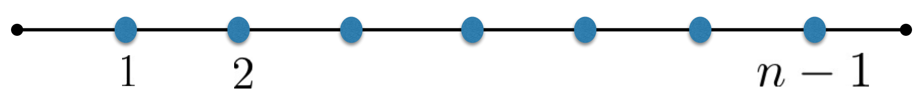



midpoint rule

$$\int_{x_i}^{x_{i+1}} f(x_{i+1/2} + \delta)d\delta = \int_{-h/2}^{h/2} (f_{i+1/2} + b\delta + c\delta^2)d\delta = hf_{i+1/2} + O(h^3)$$

$$\int_{x_0}^{x_n} f(x)dx = h(f_{1/2} + f_{3/2} + \ldots f_{n-3/2} + f_{n-1/2}) + O(h^2)$$

**Alternative:**




$$\int_{x_0}^{x_n} f(x)dx = h(\tfrac{3}{2}f_1 + f_2 + f_3 + \ldots f_{n-2} + \tfrac{3}{2}f_{n-1}) + O(h^2)$$

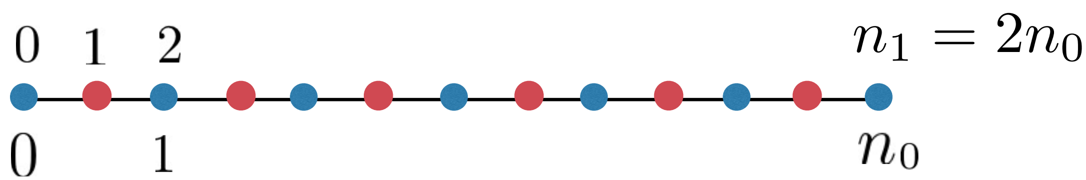
- interior points have $O(h^3)$ errors, sum to $O(h^2)$, end points contribute $O(h^2)$ errors

## Romberg integration

Idea: Use two or more trapezoidal integral estimates, extrapolate

- step sizes (decreasing order) $h_0, h_1, \ldots, h_m$, integral estimates $I_0, I_1, \ldots, I_m$
- use polynomial of order n to fit and extrapolate to h=0
- error for given h scales as $h^2$ (+ higher even powers only)
- use polynomial P(x) with $x=h^2$

Simplest case: 2 points (m=1), using $h_0=(b-a)/n_0$ and $h_1=h_0/2$ ($x_1=x_0/4$)




Function evaluation once only for each point needed

$$I_0 = I_\infty + \epsilon x_0, \quad I_1 = I_\infty + \epsilon x_0/4$$

$$\rightarrow \quad I_\infty = \frac{4}{3} I_1 - \frac{1}{3} I_0 \quad + O(h_0^4) \quad [O(x_0^2)]$$

reducing h by 50%

- error should be 1/4 of previous
- $\varepsilon$ is unknown factor, eliminated

Computation cost doubled, error reduced by two powers of $h_0$!

Generalizes easily to the case of m estimates (Friday)