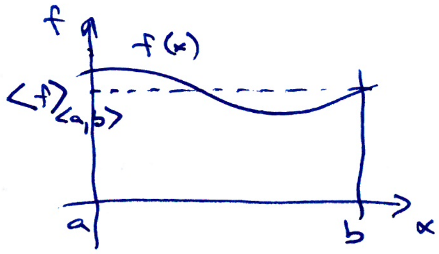
Monte Carlo integration

- although the power of the MC method shows up for multi-dimensional integrals we will illustrate it on the 1D integral for which, of course, it is much better to use Newton-Cotes ~~m~~ethod or Gauss quadratures

- we can consider integration as a kind of averaging proces and write



$$I = \int_0^1 f(x)\,dx \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

where x_i are uniformly distributed on $\langle 0,1 \rangle$

or generally

$$I = \int_a^b f(x)\,dx \approx \frac{(b-a)}{N} \sum_{i=1}^{N} f(x_i) \quad \text{with } x_i \in \langle a,b \rangle$$

(note that $(b-a)/N = h$ and if x_i would be equidistant we would get almost trapezoidal rule)

- because x_i is a random variable, we consider $f(x_i)$ also as a random variable and thus according to the CLT we get for the probable error

$$r \approx 0{,}67\,\sigma_I = 0{,}67\,\frac{(b-a)\,\sigma_f}{\sqrt{n}}$$

where σ_f^2 is variance of the $f(x)$ and it is usually estimated as

$$\sigma_f^2 \approx \langle f^2 \rangle - \langle f \rangle^2 = \frac{1}{N} \sum_{i=1}^{N} f(x_i)^2 - \left(\frac{1}{N} \sum_{i=1}^{N} f(x_i) \right)^2$$

- for d-dimensional integral we use

$$\int_\Omega f(x_1, \dots x_d)\,dx_1 \dots dx_d \approx \frac{|\Omega|}{N} \sum_{i=1}^{N} f(\underbrace{x_1^{(i)}, \dots x_d^{(i)}}_{i\text{-th random point in } dD})$$

- comparison with other methods:
  - in general we can use in 1D any method and the error usually behaves as $\mathcal{O}(h^k)$ where h is a typical distance between x_i and x_{i+1} and k is the order of the method (e.g. $k=2$ for the trapezoidal rule or $k=4$ for the Simpson rule)
  - if we use such a method to evaluate the d-dimensional integral then we need $n^d = N$ points where $n \sim \frac{1}{h}$ and thus the error will be $\mathcal{O}(h^k) \approx \mathcal{O}\left(\frac{1}{n^k}\right) \approx \mathcal{O}\left(\frac{1}{N^{k/d}}\right)$
  - we see that the MC method is comparable for $\frac{1}{N^{k/d}} \approx \frac{1}{\sqrt{N}}$ or $\frac{k}{d} \approx \frac{1}{2}$ or $d \approx 2k$

    for the trapezoidal rule we get $d \approx 4$

    for the Simpson rule $d \approx 8$ etc.

- error depends also on σ_f
  - if $f(x)$ is close to a constant function σ_f is very small (for a constant $\sigma_f = 0$) and the MC method works very well
  - on the other hand if e.g. $f(x)$ is localised in a relatively small part of $\langle a,b\rangle$ (usual case in multi-dimen. integrals) σ_f can be very large

    but we can increase accuracy by choosing a suitable weight function $w(x)$ satisfying $w(x) > 0$ on $\langle a,b\rangle$ (Ω) and $\int_a^b w(x)\,dx = 1$ which will be „similar" to $f(x)$

- rewriting

$$I = \int_a^b f(x)\,dx = \int_a^b \frac{f(x)}{w(x)}\, w(x)\,dx = \int_0^1 g(y)\,dy$$

where $y(x) = \int_a^x w(t)\,dt \qquad dy = w(x)\,dx$

$$g(y) = \frac{f(x(y))}{w(x(y))} \qquad y(a) = 0\ ,\ y(b) = 1$$

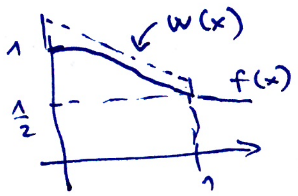
we can approximate $I \approx \frac{1}{N}\sum_{i=1}^{N} g(y_i)$

with uniformly distributed y_i

- the error can be much smaller if $\sigma_g \ll \sigma_f$

- the problem is the inversion of $y(x)$ which is actually the CDF of $w(x)$ distribution

Example:



$$f(x) = \frac{1}{1+x^2}$$

if we use
$$w(x) = \frac{4}{3} - \frac{2}{3}x$$
satisfying $\int_0^1 w(x)\,dx = 1$

then $y(x) = \frac{1}{3}x(4-x)$

and $x(y) = 2 - \sqrt{4-3y}$

this „simple" modification gives the result of 1 order better!

- if we cannot invert $y(x)$ we can still use the trick but we use directly

$$I = \int_a^b \frac{f(x)}{w(x)}\, w(x)\,dx \approx \frac{1}{N}\sum_{i=1}^{N} \frac{f(x_i)}{w(x_i)}$$

where x_i must be generated in such a way to have probability distribution given by $w(x)$ satisfying $\int_a^b w(x)\,dx = 1$

- if $w(x)$ is normalized to A then we have to use $I = \frac{A}{N}\sum_{i=1}^{N}\frac{f(x_i)}{w(x_i)}$