Course 10/1

Importance sampling

- Importance sampling through change of variable
- Importance sampling through change of random distribution
- Example: Comparison with vs without importance sampling
- Scaling of importance sampling with dimension

Importance sampling through change of variable

Idea for speeding up Monte Carlo calculations

$$\sigma_I^{\,2} = \frac{1}{N}\,\sigma_f^{\,2}$$

We can change variable in the integral without affecting the result. We should aim at finding a new integration variable for which the function to be integrated is smooth.

In this case, the Monte Carlo calculation can be performed with less points.

It is a trade-off: cost of change of variable *vs* gain in Monte Carlo integration.

Importance sampling through change of variable

We rewrite the integral as

$$I = \int_0^1 f(x)\,dx = \int_0^1 \omega(x)\,\frac{f(x)}{\omega(x)}\,dx$$

where the function $\omega(x)$ has the following properties:

1. $\omega(x) > 0$
2. $\int_0^1 \omega(x)\,dx = 1$ normalized
3. $\omega(x)$ varies like $f(x)$

Change of variable

$$\boxed{y(x) = \int_0^x \omega(x')\,dx'}$$

$$\frac{dy}{dx} = \omega(x) \qquad y(x=0) = 0 \qquad y(x=1) = 1$$

Importance sampling through change of variable

Integral to be calculated

$$I = \int_0^1 dx \ \omega(x) \frac{f(x)}{\omega(x)} = \int_0^1 dy \ \frac{f(\ x(y)\ )}{\omega(\ x(y)\ )}$$

Stochastic estimate

$$I \cong \frac{1}{N} \sum_{i=1}^{N} \frac{f(\ x(y_i)\ )}{\omega(\ x(y_i)\ )}$$

Error:

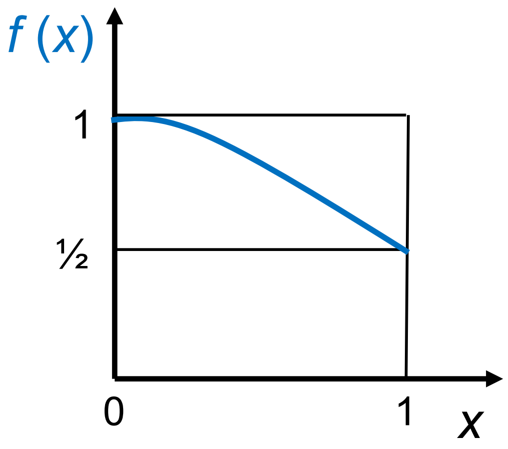
$$\sigma_I^{\ 2} = \frac{1}{N} \left[\ \langle \left(\frac{f}{\omega} \right)^2 \rangle - \langle \frac{f}{\omega} \rangle^2 \right] = \frac{1}{N} \sigma_{f/\omega}^{\ \ 2}$$

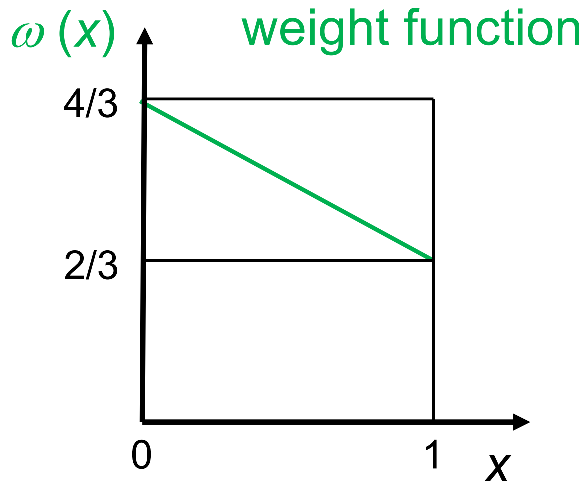
Since $\omega(x)$ varies like $f(x)$, this should result in $\sigma_{f/\omega} \ll \sigma_f$.

Hence, the integration should become more efficient!

Importance sampling: Example I

$$I = \int_0^1 \frac{1}{1+x^2}\,dx = \frac{\pi}{4} \cong 0.78540$$





Let us take as weight function the linear function: $\omega(x) = \frac{1}{3}(4 - 2x)$

Properties:

1. $\omega(x) > 0$ for $0 < x < 1$
2. $\int_0^1 \frac{1}{3}(4-2x)\,dx = 1$ normalized
3. $\omega(x)$ varies like $f(x)$: $\frac{f(0)}{\omega(0)} = \frac{3}{4}$ $\frac{f(1)}{\omega(1)} = \frac{3}{4}$

Importance sampling: Example I

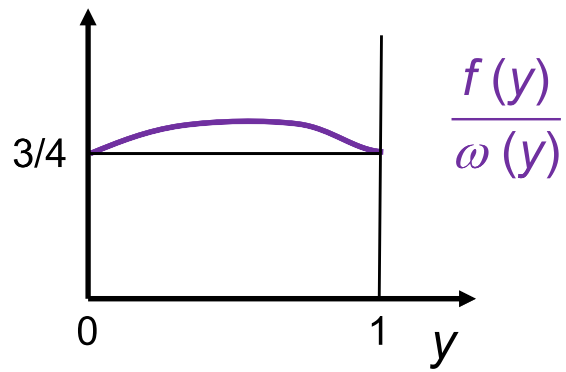
Change of variable

$$y(x) \;=\; \int_0^x \omega\,(x'\,)\,dx' \;=\; \frac{4}{3}\,x - \frac{1}{3}\,x^2 \;=\; \frac{1}{3}\,x\,(\,4 - x\,)$$

which needs to be inverted to give $\quad x\,(\,y\,) \;=\; 2 - \sqrt{4 - 3y}$

New integral

$$I \;=\; \int_0^1 dy\;\frac{f\,(\,x(y)\,)}{\omega\,(\,x(y)\,)}$$



Stochastic estimate

$$I \;\cong\; \frac{1}{N}\sum_{i\,=\,1}^{N}\frac{f\,(\,x(y_i)\,)}{\omega\,(\,x(y_i)\,)}$$

Importance sampling: Example I

$$I = \int_0^1 \frac{1}{1+x^2}\,dx = \frac{\pi}{4} \cong 0.78540$$

	$w(x)=1$		$w(x)=\frac{1}{3}(4-2x)$	
N	I	σ_I	I	σ_I
10	0.81491	0.04638	0.79982	0.00418
20	0.73535	0.03392	0.79071	0.00392
50	0.79606	0.02259	0.78472	0.00258
100	0.79513	0.01632	0.78838	0.00194
200	0.78677	0.01108	0.78529	0.00140
500	0.78242	0.00719	0.78428	0.00091
1000	0.78809	0.00508	0.78524	0.00064
2000	0.78790	0.00363	0.78648	0.00045
5000	0.78963	0.00227	0.78530	0.00028
$\pi/4$	0.78540		0.78540	

Change of random distribution

Without importance sampling

$$I = \int_0^1 f(x)\, dx$$

With importance sampling

$$I = \int_0^1 dx \;\; \omega(x) \;\; \frac{f(x)}{\omega(x)}$$

$$= \int_0^1 dy \;\; \frac{f(\, x(y)\,)}{\omega(\, x(y)\,)}$$

Stochastic estimate

$$I \cong \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

x_i uniformly distributed

Stochastic estimate

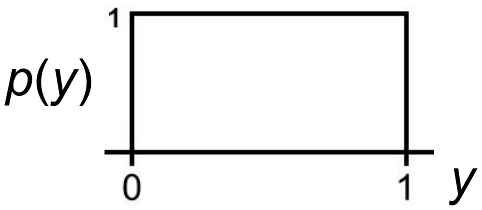
$$I \cong \frac{1}{N} \sum_{i=1}^{N} \frac{f(\, x(y_i)\,)}{\omega(\, x(y_i)\,)}$$

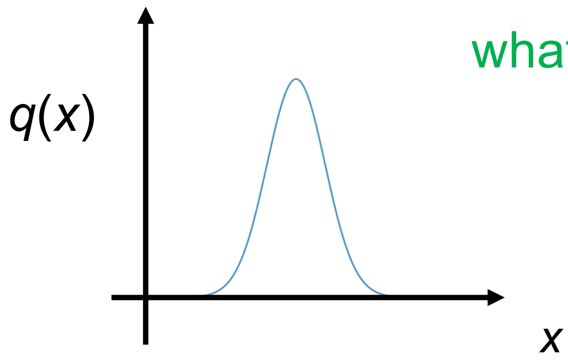
y_i uniformly distributed and
the $x(y_i)$ result from the inverse of

$$y(x) = \int_0^x \omega(x')\, dx'$$

What is the distribution of the $x(y_i)$ values?

Change of random distribution





what is $q(x)$?

Relationship between y and x :

$$y(x) = \int_0^x \omega(x')\,dx' \;\Rightarrow\; \frac{dy}{dx} = \omega(x) \;\Rightarrow\; dy = \omega(x)\,dx$$

but $p(y) = 1$ $\Rightarrow$ $p(y)\,dy = \omega(x)\,dx$

Conservation of probability!

Indeed, $\omega(x)$ can be seen as a probability density :

1. $\omega(x) > 0$
2. $\omega(x)$ is normalized

$$\boxed{q(x) = \omega(x)}$$

The x variables are distributed according to $\omega(x)$

Change of random distribution

With importance sampling

$$\int_0^1 \frac{f(\,x(y)\,)}{\omega(\,x(y)\,)}\,dy \quad = \quad \int_0^1 \frac{f(x)}{\omega(x)}\,\omega(x)\,dx$$

Stochastic estimate

$$I \;\cong\; \frac{1}{N}\sum_{i=1}^{N} \frac{f(\,x(y_i)\,)}{\omega(\,x(y_i)\,)}$$

y_i uniformly distributed

Stochastic estimate

$$I \;\cong\; \frac{1}{N}\sum_{i=1}^{N} \frac{f(x_i)}{\omega(x_i)}$$

x_i distributed according to $\omega(x)$

Change of random distribution

Without importance sampling

$$I = \int_0^1 f(x)\, dx$$

Stochastic estimate

$$I \cong \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

x_i uniformly distributed

With importance sampling

$$I = \int_0^1 \frac{f(x)}{\omega(x)}\, \omega(x)\, dx$$

Stochastic estimate

$$I \cong \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{\omega(x_i)}$$

x_i distributed according to $\omega(x)$

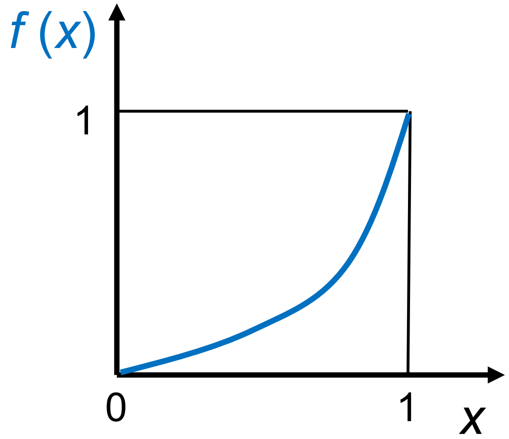
$\omega(x)$ has to be chosen in such a way that it behaves like $f(x)$.

In other terms, the regions where $f(x)$ is large should be sampled more frequently.

In this way, the integral is evaluated in a more efficient way.

Importance sampling

Importance sampling: Example II



$$f(x) = \frac{1}{e-1}\,(e^{x} - 1)$$

Exact result

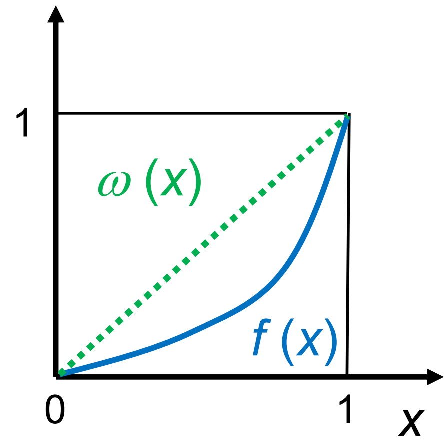
$$I = \int_0^1 f(x)\,dx = \frac{1}{e-1}\,[\,e^{x} - x\,]_0^1 = \frac{e-2}{e-1} \cong 0.418$$

Error with $\omega(x) = 1$

$$\sigma_{f,\omega=1}^{\;2} = \int_0^1 [\,f(x) - \underbrace{\langle f \rangle}_{=I}\,]^2 dx = (\,0.286\,)^2$$

Hence, $\sigma_{I,\omega=1} = \frac{1}{\sqrt{N}}\,\sigma_{f,\omega=1} = \frac{1}{\sqrt{N}}\,0.286$

Importance sampling: Example II



With importance sampling

$\omega(x) = 2x$ (where the 2 comes from the normalization)

The integral can be rewritten but gives the same result…

$$I = \int_0^1 \frac{f(x)}{2x}\, 2x\, dx = \int_0^1 \frac{f(\, x(y)\,)}{2x(y)}\, dy$$

Error with $\omega(x) = 2x$

$$\sigma_{f,\omega=2x}^{\;2} = \int_0^1 \left[\frac{f(\, x(y)\,)}{2x(y)} - \left\langle \frac{f(\, x(y)\,)}{2x(y)} \right\rangle \right]^2 dy$$

$$= \int_0^1 \left[\frac{f(x)}{2x} - I \right]^2 2x\, dx = (\, 0.0523\,)^2$$

Hence, $\sigma_{I,\omega=2x} = \frac{1}{\sqrt{N}}\, \sigma_{f,\omega=2x} = \frac{1}{\sqrt{N}}\, 0.0523$

Importance sampling: Example II

Summary

Without importance sampling: $\omega(x) = 1$

$$I \cong \frac{1}{N}\sum_{i=1}^{N} f(x_i) = 0.418 \pm \frac{0.286}{\sqrt{N}}$$

x_i uniformly distributed

for 1% accuracy,
$N = 4700$

With importance sampling: $\omega(x) = 2x$

$$I \cong \frac{1}{N}\sum_{i=1}^{N} \frac{f(x_i)}{2x_i} = 0.418 \pm \frac{0.0523}{\sqrt{N}}$$

x_i distributed according to $\omega(x)$

for 1% accuracy,
$N = 155$

Without importance sampling, we need 30× more evaluations to achieve the same accuracy as with importance sampling.

Scaling of importance sampling with dimension

In general this depends on the function to be integrated. To illustrate the scaling, we here extend the previous example to higher dimensions.

We assume that the function in *d* dimensions corresponds to a product of *d* times the same one-dimensional function.

$$F\,(\,x_1,\,x_2,\,\ldots,\,x_d\,) = f\,(x_1)\cdot f\,(x_2)\cdots f\,(x_d)$$

Then, we naturally take for the weight function:

$$W\,(\,x_1,\,x_2,\,\ldots,\,x_d\,) = \omega\,(x_1)\cdot \omega\,(x_2)\cdots \omega\,(x_d)$$

Variance

$$\sigma_{d,\omega}^{\;2} = \int_0^1 \left(\frac{F}{W}\right)^2 W\,(\,x_1,\,\ldots,\,x_d\,)\,dx_1\cdots dx_d \;-\; \left[\int_0^1 \left(\frac{F}{\cancel{W}}\right)\cancel{W}\,(\,x_1,\,\ldots,\,x_d\,)\,dx_1\cdots dx_d\right]^2$$

$$= \left[\int_0^1 \left(\frac{f\,(x)}{\omega\,(x)}\right)^2 \omega\,(x)\,dx\right]^d \;-\; (\,I^{\,d}\,)^2 \;=\; [\;\sigma_{f,\omega}^{\;2} + I^{\,2}\,]^{\,d} - I^{\,2d}$$

Scaling of importance sampling with dimension

Ratio between variances with and without importance sampling

$$\frac{\sigma_{d,\omega}^{2}}{\sigma_{d,\omega=1}^{2}} = \frac{[\,\sigma_{f,\omega}^{2} / I^{2} + 1\,]^{d} - 1}{[\,\sigma_{f,\omega=1}^{2} / I^{2} + 1\,]^{d} - 1} \overset{d \to \infty}{=} 10^{\,d \log_{10}\left[\frac{1 + \sigma_{f,\omega}^{2} / I^{2}}{1 + \sigma_{f,\omega=1}^{2} / I^{2}}\right]}$$

For the $f(x)$ and $\omega(x)$ seen in Example II, we obtain:

$$\frac{1 + \sigma_{f,\omega}^{2} / I^{2}}{1 + \sigma_{f,\omega=1}^{2} / I^{2}} \cong \frac{1 + (0.0523 / 0.418)^{2}}{1 + (0.286 / 0.418)^{2}} \cong 0.69179$$

$$\log_{10} 0.69179 = -0.16$$

Scaling

$$\frac{\sigma_{d,\omega}^{2}}{\sigma_{d,\omega=1}^{2}} \cong 10^{-0.16\,d}$$

For instance, for $d = 50$, this gives 10^{-8} ,
i.e. 8 orders of magnitude!

Course 10/1

Importance sampling

- Importance sampling through change of variable
- Importance sampling through change of random distribution
- Example: Comparison with vs without importance sampling
- Scaling of importance sampling with dimension