Random Numbers

Random number generator (Uniform)

$$x_{n+1} = (a\,x_n + c) \bmod m$$

$a \rightarrow$ odd, $c \rightarrow$ odd, $m \rightarrow 2^{32}$

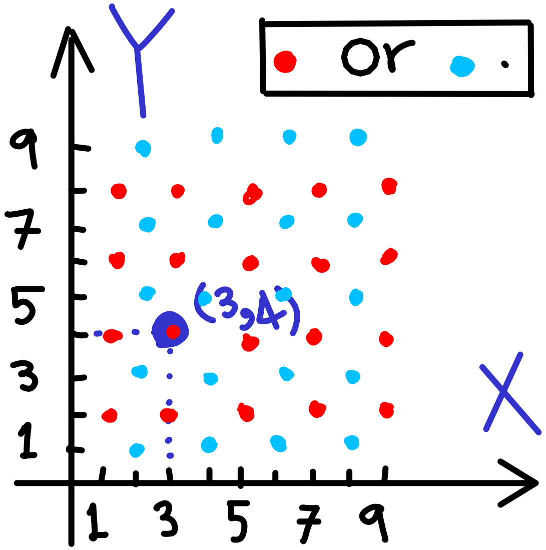
$\longrightarrow$ m should be a big number

Suppose that we want to have a random generator to produce random numbers between 0 and 9. How?

If the random number generated by the above relation is: 581
one can take the right digit of this number. Is this a good way?

answer:

No. Because if X_n is an even number then X_{n+1} will be an odd number, and vice versa. Therefore, if we take the right digit of X_{n+1} then once we have an even number and right after that we will have an odd number. Consequently, there is a correlation between the produced numbers in [0, 9].



If we take two numbers x and y, and make a vector (x,y). Then by plotting these vectors we can not cover all space in 2D between [0,9].
For example (3, 4) has been shown.

One way is to take the digit before the most right digit, i.e. 581.

The most left digit is not good as well. Because its distribution is not uniform and is logarithmic.

Suffling algorithm (suitable for very long simulations)

$$x_{n+1} = (a x_n + c) \bmod m \quad \longrightarrow \quad \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{498} \\ x_{499} \end{bmatrix}$$

1. For example we take $m' = 500$, and we use also the following random number generator. $y_{n+1} = (a' y_n + c') \bmod m'$. Then $0 \leq y_{n+1} < 500$.

2. We produce m' random numbers by using $x_{n+1} = \cdots$ and put them in an array.

3. We produce the random number y_{n+1}, and then take y_{n+1}^{th} element of the array as a desired random number.

4. We make the array empty.

5. Go to 2.

```c
#include <stdio.h>
#include <stdlib.h>
#include <time.h>

int main () {
	int i, n, rr, seed_number;
	time_t t;
	//
	seed_number = time(&t) ;
	//seed_number = 123456789 ;
	n = 5;

	/* Intializes random number generator */
	//srand((unsigned) time(&t));
	srand((unsigned) seed_number);
	printf("%ld\n\n\n\n", time(&t));

	/* Print 5 random numbers from 0 to 49 */
	for( i = 0 ; i < n ; i++ ) {
		rr     = rand();
		printf("%d\n", rr % 100);
		printf("%d\n", rr );
	}

	return(0);
}
```

```
2115143165
65
665142609
9
462857371
71
813125836
36
632321020
20
```

```c
#include <stdio.h>
#include <stdlib.h>
#include <time.h>

int main () {
	int i, n, rr, seed_number;
	time_t t;
	//
	n = 5;

	/* Intializes random number generator */
	//printf("%ld\n\n\n\n", time(&t));

	/* Print 5 random numbers from 0 to 99 */
	for( i = 0 ; i < n ; i++ ) {
		seed_number = time(&t) ;
		srand((unsigned) seed_number);

		rr     = rand();
		printf("%d\n", rr );
		printf("%d\n", rr % 100);
	}

	return(0);
}
```

```
2065990336
36
2065990336
36
2065990336
36
2065990336
36
2065990336
36
```

- Many phenomena in physics reveal Gaussian random distribution. For example velocity distribution of a gass is Gaussian.
- Therefore, we need to Know how non_uniform random numbers can be generated.

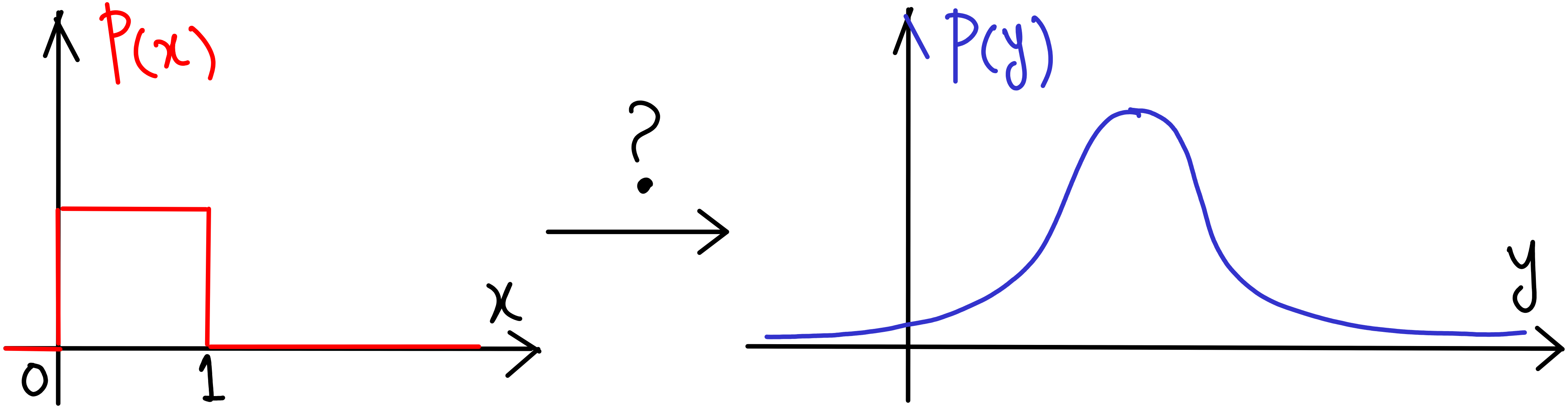
Centeral Limit Theorm (in stat. mech.)

$$X_{n+1} = (a X_n + c) \mod m$$

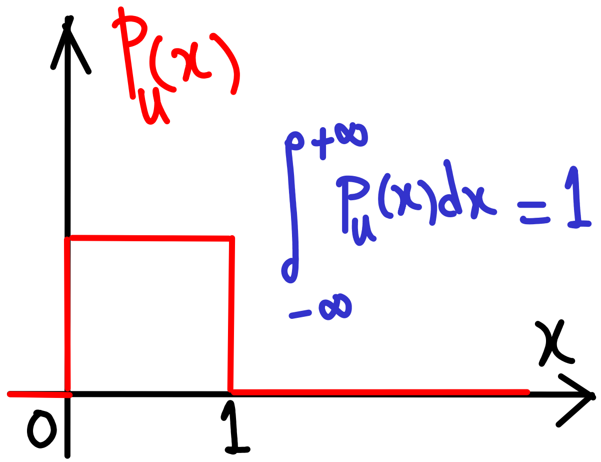
$$\text{If} \begin{cases} X_i\text{s are uncorrelated and random.} \\ y = \frac{1}{N} \sum_{i=1}^{N} X_i \quad \text{(If N is big enough)} \end{cases} \Longrightarrow \boxed{P(y) = \frac{1}{\sqrt{2\pi}\,\sigma_y} e^{-\frac{(y-\langle y \rangle)^2}{2\sigma_y^2}}}$$

$$\langle y \rangle = \langle x \rangle \qquad \sigma_y = \sigma_x / \sqrt{N}$$

How a non-uniform probability distribution can be produced from a uniform probability distribution?



$$\underset{\text{uniform}}{P(x)} = P_u(x) = \begin{cases} 1 & : 0 < x < 1 \\ 0 & : \text{else} \end{cases}$$



We would like to produce y from x by:

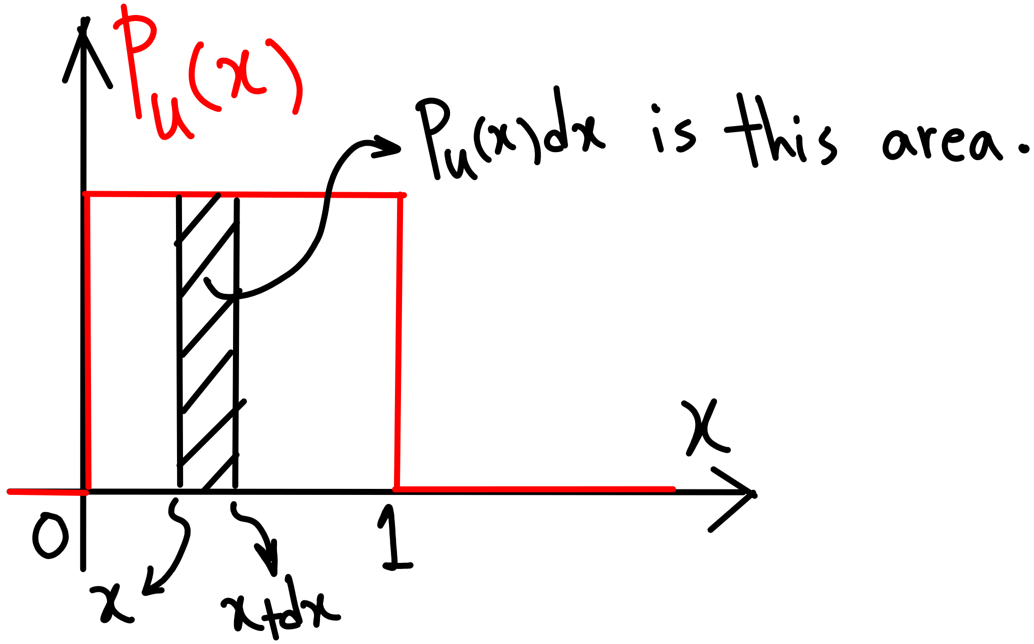
$$y = f(x)$$

in such a way that y possess an arbitrary distribution function:

$$g(y).$$

g can be Gaussian or ... .

How?



Probability of choosing a random number between $(x, x+dx)$: $P_u(x)\,dx$

Corresponds to the:

probability of having a random number y between $(y, y+dy)$: $g(y)\,dy$

$$P_u(x)\,dx = g(y)\,dy$$

$$\int_{-\infty}^{x} P_u(x)\,dx = \int_{-\infty}^{y=f(x)} g(y)\,dy$$

We assume that we know how to perform the integration.

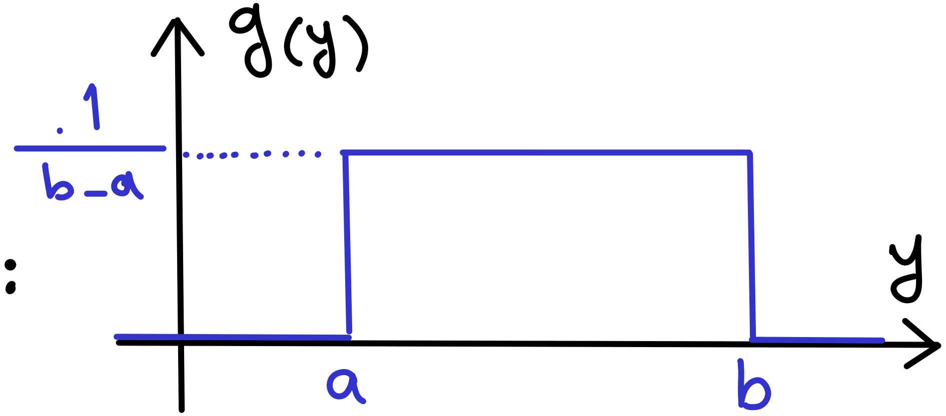
$$x = G(y)$$

$$\Downarrow$$

$$\boxed{y = G^{-1}(x)}$$

Examples:

- Suppose $g(y)$ has this form:



$$g(y) = \begin{cases} \frac{1}{b-a} & a<y<b \\ 0 & \text{else} \end{cases}$$

Indeed, we choose a random number in $(0,1)$ and transform it to a random number in (a,b). How can we do this?

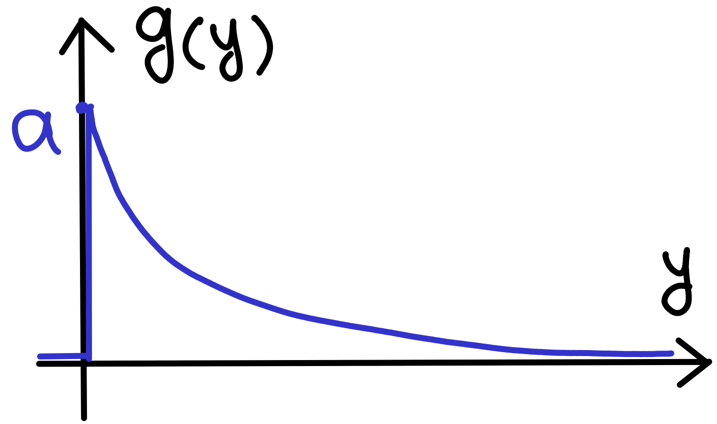
answer:

$$x = \int_{-\infty}^{y} g(y)\,dy \Rightarrow x = \int_{a}^{y} \frac{1}{b-a}\,dy \Rightarrow x = \frac{1}{b-a}\, y \Big|_{a}^{y} \Rightarrow x = \frac{y-a}{b-a}$$

$$y = a + (b-a)x$$

Examples:

- $g(y) = \begin{cases} a e^{-ay} & y > 0 \\ 0 & \text{else} \end{cases}$



answer:

$$x = \int_{-\infty}^{y} g(y)\,dy \Rightarrow x = \int_0^y a e^{-ay} dy \Rightarrow x = a\left(-\frac{1}{a}\right)e^{-ay}\Big|_0^y \Rightarrow x = 1 - e^{-ay}$$

$$e^{-ay} = 1 - x \Rightarrow y = \frac{-1}{a}\ln(1-x) \Rightarrow \qquad y = \frac{1}{a}\ln(1-x)^{-1}$$

$$\int_{-\infty}^{x} P_u(x)\,dx = \int_{-\infty}^{y=f(x)} g(y)\,dy$$

We assume that we know how to perform the integration.

$$x = G(y)$$

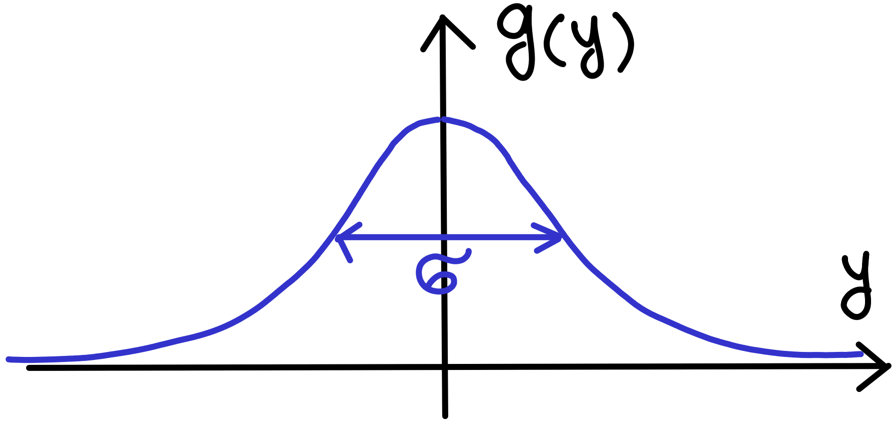
$$\Downarrow$$

$$\boxed{y = G^{-1}(x)}$$

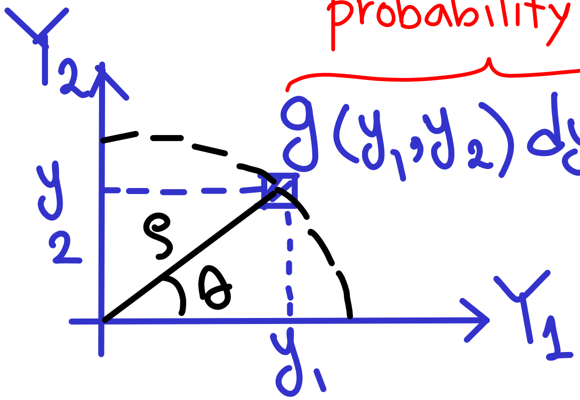
As far as g is integratable and G has inverse this method works.

Examples:

- $$g(y) = \frac{\quad}{\sqrt{2\pi}\,\sigma} e^{-\frac{y^2}{2\sigma^2}}$$



answer: $\int_{-\infty}^{y} g(y)\,dy = ?$



$$\overbrace{g(y_1, y_2)\,dy_1\,dy}^{\text{probability of having } (y_1, y_2)} \underset{\text{if they are independent}}{=} g(y_1)\,g(y_2)\,dy_1\,dy_2 = \frac{1}{2\pi\sigma^2} e^{-\frac{(y_1^2+y_2^2)}{2\sigma^2}} dy_1\,dy_2$$

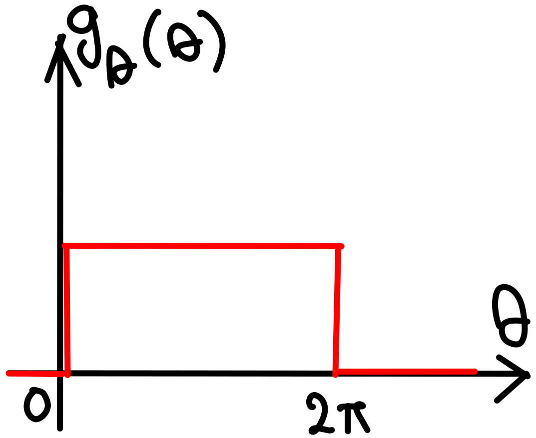
$$\sim \frac{1}{2\pi\sigma^2} e^{-\rho^2/(2\sigma^2)} \rho\,d\rho\,d\theta = g_\theta(\theta)\,g_\rho(\rho)\,d\rho\,d\theta$$

$$\begin{cases} g_\theta(\theta) = \dfrac{1}{2\pi} \\ g_\rho(\rho) = \dfrac{1}{\sigma^2}\,\rho\,e^{-\rho^2/(2\sigma^2)} \end{cases}$$

$$P_u(x_1)\, P_u(x_2)\, dx_1\, dx_2 = g(y_1)\, g(y_2)\, dy_1\, dy_2$$

$$\int_{-\infty}^{x_1} P_u(x_1)\,dx_1 \int_{-\infty}^{x_2} P_u(x_2)\,dx_2 = \int_{-\infty}^{y_1} g(y_1)\,dy_1 \int_{-\infty}^{y_2} g(y_2)\,dy_2$$

$$x_1\, x_2 = \boxed{\int_{\theta=0}^{2\pi} g_\theta(\theta)\,d\theta}\ \boxed{\int_{\rho=\rho}^{\infty} \frac{1}{\sigma^2}\,\rho\, e^{-\rho^2/(2\sigma^2)}\, d\rho}$$



$$\begin{array}{l} a = 0 \\ b = 2\pi \end{array} \Rightarrow \underbrace{y}_{\theta} = a + (b-a)\,x_1 \Rightarrow \boxed{\theta = 2\pi\, x_1}$$

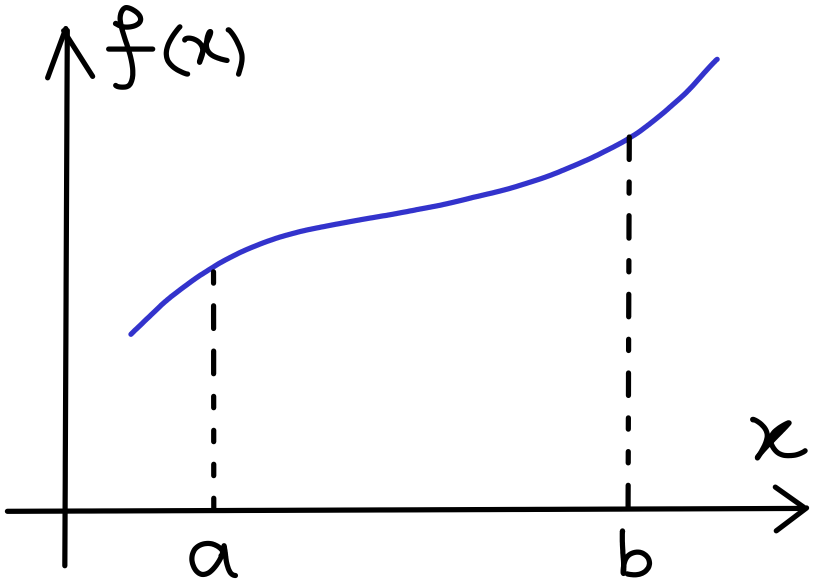
$$x_2 = \int_{\rho}^{\infty} \frac{1}{\sigma^2}\, \rho\, e^{-\underbrace{\rho^2/(2\sigma^2)}_{y}}\, d\rho \Rightarrow x_2 = \int_{z=\rho^2/(2\sigma^2)}^{\infty} dz\, e^{-z} = -e^{-z}\Big|_{z=\frac{\rho^2}{2\sigma^2}}^{z=\infty}$$

$$x_2 = e^{-\rho^2/(2\sigma^2)} \Longrightarrow \boxed{\rho = +2\sigma^2 \ln x_2^{-1}}$$

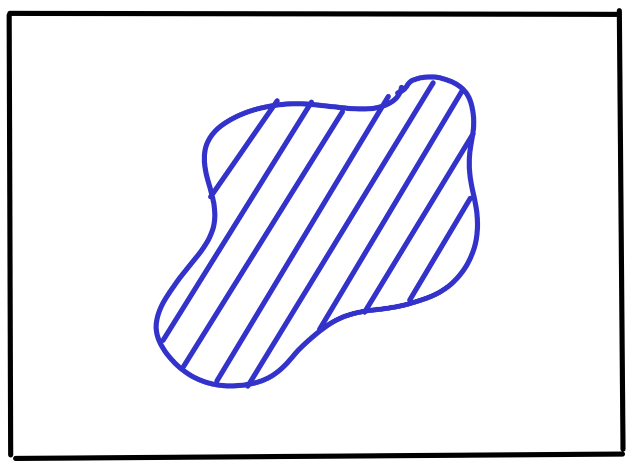
$$\begin{cases} y_1 = \rho \cos\theta \\ y_2 = \rho \sin\theta \end{cases}, \begin{cases} \theta = 2\pi x_1 \\ \rho = +2\sigma^2 \ln x_2^{-1} \end{cases} \Longrightarrow \boxed{\begin{aligned} y_1 &= +2\sigma^2 \cos(2\pi x_1) \ln x_2^{-1} \\ y_2 &= +2\sigma^2 \sin(2\pi x_1) \ln x_2^{-1} \end{aligned}}$$

Numerical Integration:
a Stochastic Method

Performing integrals

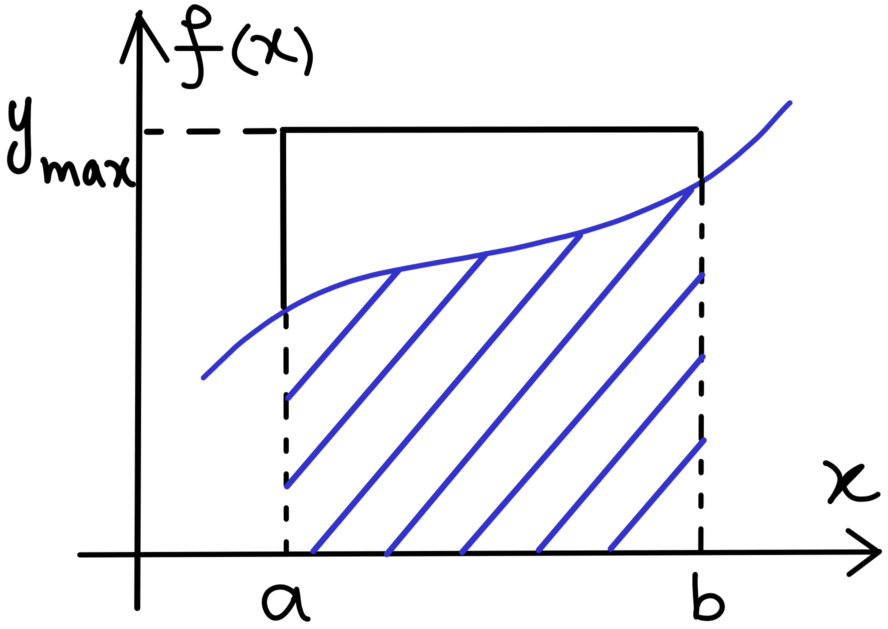


$$I = \int_{x=a}^{x=b} f(x)\,dx = ?$$



What is the area of the pool inside a garden?

$$A_P = \frac{\text{shelep}}{\text{shelep} + \text{telep}} A_G$$



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

$$a < x < b \qquad 0 < y < y_{max}$$

loop N

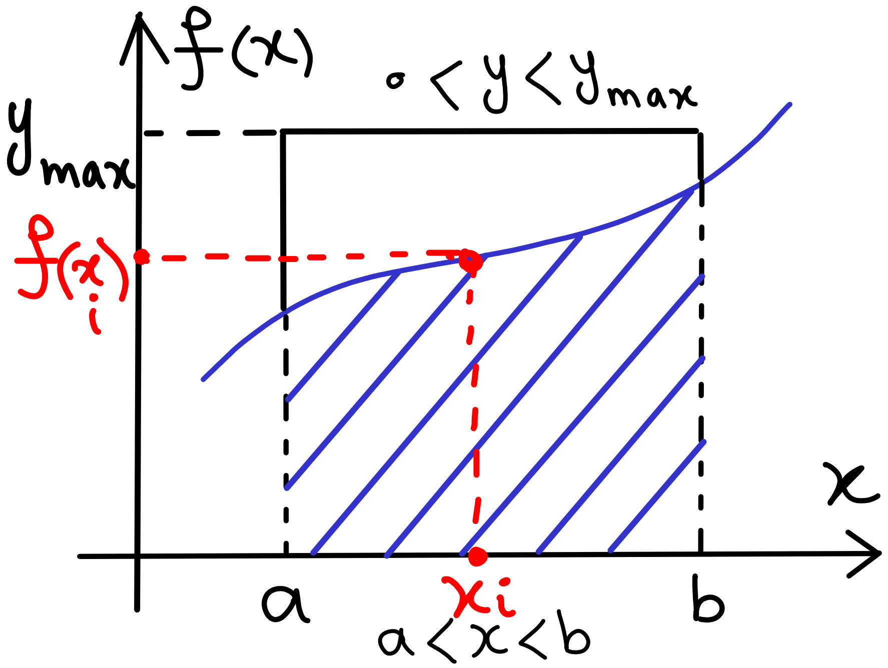
$a < \underset{\text{random}}{x} < b$

$0 < \underset{\text{random}}{y} < y_{max}$

If $f(x) > y$ then shelep += 1;

$$\Longrightarrow \quad I = \underbrace{(b-a)\,y_{max}}_{A_G} \frac{\text{Shelep}}{N}$$

This is not a fast algorithm.



$$p(x) = \begin{cases} 1 & a < x < b \\ 0 & \text{else} \end{cases}$$

$\downarrow$ distribution function

$$\int_{-\infty}^{+\infty} p(x)\,dx = b - a$$

$$I = \int_{x=a}^{x=b} f(x)\,dx = (b-a)\frac{\int_{-\infty}^{+\infty} f(x)\,p(x)\,dx}{\int_{-\infty}^{+\infty} p(x)\,dx} = (b-a)\langle f \rangle$$

$$\langle f \rangle = \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

$$\text{Loop}_{N}\ \Big|\ I = f(\text{Random}(a,b)) + I;$$

$$I = \frac{I}{N} * (b-a)$$

Simple sampling Monte Carlo

fluctuations of f

$$\underset{\text{error}}{\Delta} = \frac{\sigma_f}{\underbrace{\sqrt{N}}_{\text{number of samples}}}, \quad \sigma_f = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}$$

During the run of the code we can find Δ and control the error.

$$I = \int_{x=a}^{x=b} f(x)\,dx = (b-a)\frac{\int_{-\infty}^{+\infty} f(x)\,p(x)\,dx}{\int_{-\infty}^{+\infty} p(x)\,dx} = (b-a)\langle f \rangle$$

$$\Delta = \frac{\sigma_f}{\sqrt{N}}$$

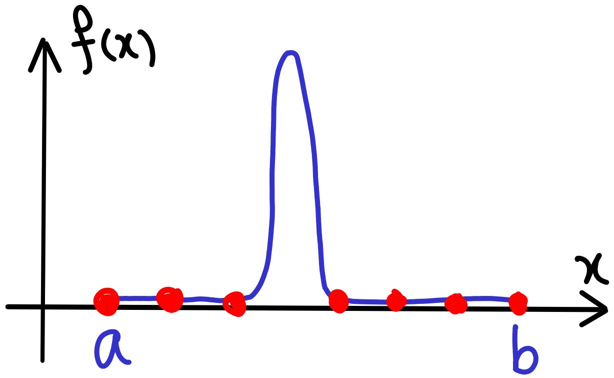
How error (Δ) can be reduced?

- by increasing N.
- by reducing σ_f. $\longrightarrow$ How?

$$\sigma_f = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}$$

$\downarrow$

We must reduce $\langle f^2 \rangle - \langle f \rangle^2$.



If the number of sampling is not enough we may mis the peak, that contains the main contribution to the integration.

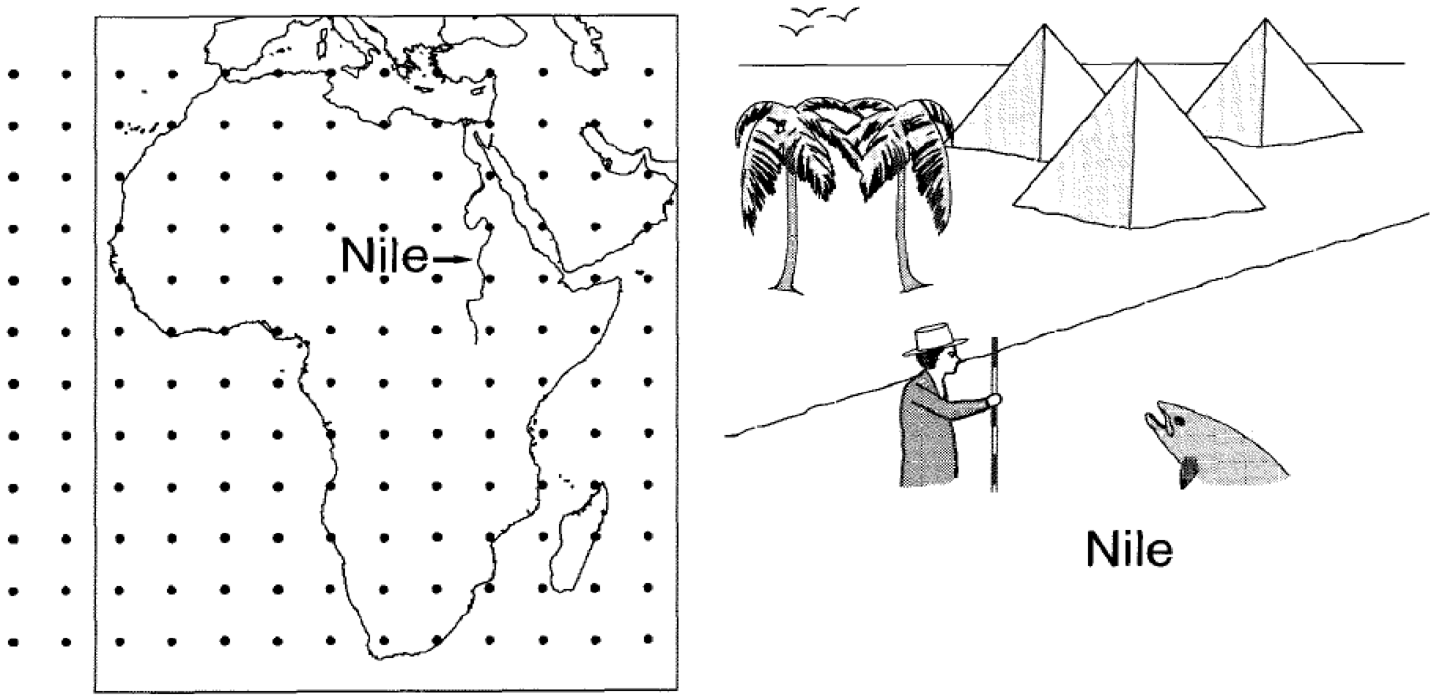


Figure 3.1: Measuring the depth of the Nile: a comparison of conventional quadrature (left), with the Metropolis scheme (right).

Frenkel and Smit, page 28.

$$I = \int_{x=a}^{x=b} f(x)\,dx = (b-a)\frac{\int_{-\infty}^{+\infty} f(x)\,p(x)\,dx}{\int_{-\infty}^{+\infty} p(x)\,dx} = (b-a)\langle f \rangle$$

$$\Delta = \frac{\sigma_f}{\sqrt{N}} \qquad \sigma_f = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}$$

To reduce σ_f, i.e. reducing $\langle f^2 \rangle - \langle f \rangle^2$, we must pick a non-uniform $p(x)$.

$$I = \int_{x=a}^{x=b} f(x)\,dx = \int g(x)\,dx \; \frac{\int \frac{f(x)}{g(x)}\, g(x)\,dx}{\int g(x)\,dx}$$

$$I = \int_{x=a}^{x=b} f(x)\,dx = \int g(x)\,dx \; \frac{\int \frac{f(x)}{g(x)}\, g(x)\, dx}{\int g(x)\, dx}$$

$$I = \int_{x=a}^{x=b} f(x)\,dx = \int g(x)\,dx \; \left\langle \frac{f}{g} \right\rangle_g$$

We know how to perfom

averaging over distribution function g

g is not uniform. Therefore, when $g(x)$ is big the probability of choosing x must be big and vice versa.

$$I = \int_{x=a}^{x=b} f(x)\,dx = \int g(x)\,dx \left\langle \frac{f}{g} \right\rangle_g$$

$$\Delta = \int g(x)\,dx \, \frac{\sigma_{f/g}}{\sqrt{N}} \qquad \sigma_{f/g} = \sqrt{\left\langle \left(\frac{f}{g}\right)^2 \right\rangle - \left\langle \frac{f}{g} \right\rangle^2}$$

Now the error Δ depends on the fluctuations of f/g.

This tells us that g should be chosen in such a way that f/g has small fluctuations. This happens if g behaves similar to f.

What is the meaning of $\left\langle \frac{f}{g} \right\rangle_g$ in algorithmic language?

Loop N

$x = \text{Random}$ ← uniform

$y = \text{Random}(g(x))$

$I = f(y)/g(y) + I$

$$I = \frac{I}{N} * \boxed{\int g(x)\,dx}$$

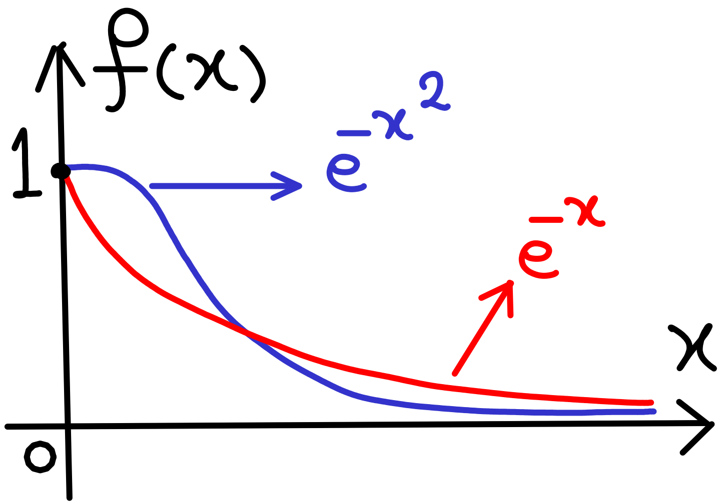
← We knew how to perform this.

Importance Sampling Monte Carlo

Exercise :

Calculate the following integral with both simple sampling and importance sampling Monte Carlo. For the importance sampling one can use g as $g(x) = e^{-x}$.

$$I = \int_0^2 e^{-x^2}\,dx$$



	SS				IS
N	I	ΔI	ΔI_m	t	
100					
200					
500					
1000					