# ensIIE. Simulation methods

Abass SAGNA,
abass.sagna@ensiie.fr

Maître de Conférences à l'ensIIE,
Laboratoire de Mathématiques et Modélisation d'Evry
Université d'Evry Val-d'Essonne, UMR CNRS 8071

http://www.math-evry.cnrs.fr/members/asagna/

January 6, 2026

# Plan

# Plan

## Plan

Some situations where their is no analytical formula for $\mathbb{E}(h(X))$, where $X$ is a random vector (or stochastic process) defined on $\mathbb{R}^d$ and $h : \mathbb{R}^d \mapsto \mathbb{R}$.
Examples:

1. $h(x) = \cos(x^2)\mathbb{1}_{[0,2\pi]}(x)$ and $X \sim \mathcal{N}(0,1)$.
2. Let $(X_n)_{n\geq 0}$ be a Markov process, $A \subset \mathbb{R}^d$, $\tau = \inf\{k \geq 0,\ X_k \notin A\}$. Let $m \leq n$ and $\mathcal{F}_m = \sigma(X_0, \ldots, X_m)$. Compute

$$\mathbb{E}\big(h(X_0, \ldots, X_n)\mathbb{1}_{\{\tau > n\}} | \mathcal{F}_m^X\big) \ \text{ or } \ \mathbb{E}\big(h(X_0, \ldots, X_n)\mathbb{1}_{\{\tau < n\}} | \mathcal{F}_m^X\big), \quad (1)$$

where $h$ and $(X_n)_{n\geq 0}$ can be defined as follows:

   1. $h$ may be equal to 1: compute for example $\mathbb{P}(\{\tau > n\}|\mathcal{F}_m^X)$
   2. $h(x_0, \ldots, x_n) = \min_{i=0,\ldots,n} \text{dist}(x_i, A)$,
   3. $h(x_0, \ldots, x_n) = \text{dist}(x_n, A)$,
   4. $(X_{k\wedge n})_{k\geq 0}$ may evolves as $X_0 = x \in \mathbb{R}^d$ and for every $k \geq 0$,

$$X_{k+1} = X_k + \mu(X_k)/n + \sqrt{1/n}\,\sigma(X_k)\, Y_{k+1}$$

where $\mu : \mathbb{R}^d \mapsto \mathbb{R}^d$, $\sigma : \mathbb{R}^d \mapsto \mathbb{R}^{d\times q}$ and $Y_k \overset{iid}{\sim} \mathcal{N}(0; I_q)$.

1. The set $A$ may be equal to the circle of center $x_0$ and radius $\gamma > 0$:
$$A = \mathcal{C}(\gamma) = \{x \in \mathbb{R}^d,\ \text{dist}(x, x_0) \leq \gamma\}.$$
2. In the case $d = 2$, we may choose $X \equiv (X, Y)$:
$$\begin{cases} X_{t_{k+1}} &= X_{t_k}\left((1+\Delta\mu_1) + \sigma_1\sqrt{\Delta}Z_{k+1}\right) \\ Y_{t_{k+1}} &= Y_{t_k}\left((1+\Delta\mu_2) + \rho\,\sigma_2\sqrt{\Delta}Z_{k+1} + \sigma_2\sqrt{1-\rho^2}\sqrt{\Delta}\tilde{Z}_{k+1}\right) \end{cases}$$
where $\mu_1, \mu_2 \in \mathbb{R}$, $\Delta = 1/n$, $\sigma_1, \sigma_2 > 0$, $\rho \in (-1, 1)$; $Z$ and $\tilde{Z}$ are independent with $Z_k \overset{iid}{\sim} \mathcal{N}(0;1)$ and $\tilde{Z}_k \overset{iid}{\sim} \mathcal{N}(0;1)$.
3. Other example: for $x = (x_1, \ldots, x_d)' \in \mathbb{R}^d$, $\mu(x)$ and $\sigma(x)$ may be chosen as
$$\mu(x) = \mu \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \qquad \sigma(x) = \sigma \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & x_d \end{pmatrix}$$
where $\mu \in \mathbb{R}$, $\sigma > 0$.

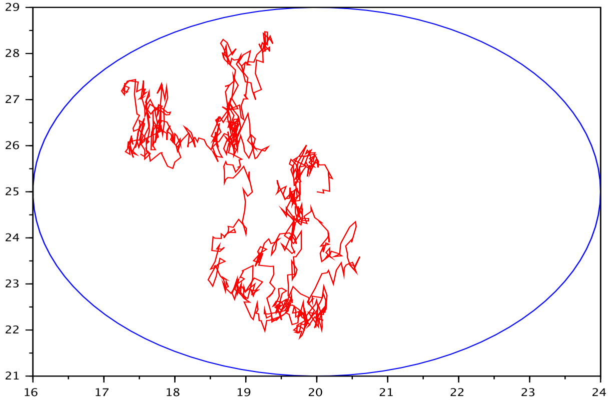



Figure: $\begin{cases} X_{t_{k+1}} = X_{t_k}\left((1+\Delta\mu_1) + \sigma_1\sqrt{\Delta}Z_{k+1}\right) \\ Y_{t_{k+1}} = Y_{t_k}\left((1+\Delta\mu_2) + \rho\,\sigma_2\sqrt{\Delta}Z_{k+1} + \sigma_2\sqrt{1-\rho^2}\sqrt{\Delta}\tilde{Z}_{k+1}\right) \end{cases}$
where $\gamma = 4$, $[x_0, y_0] = [20, 25]$, $[\mu_1, \mu_1] = [0.01, -0.05]$, $[\sigma_1, \sigma_2] = [0.09, 0.2]$, $\rho = 0.1$, $n = 1000$.

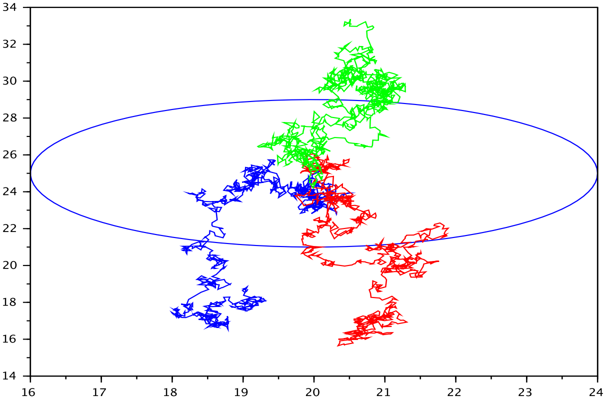



Figure: $\begin{cases} X_{t_{k+1}} = X_{t_k}\left((1+\Delta\mu_1) + \sigma_1\sqrt{\Delta}Z_{k+1}\right) \\ Y_{t_{k+1}} = Y_{t_k}\left((1+\Delta\mu_2) + \rho\,\sigma_2\sqrt{\Delta}Z_{k+1} + \sigma_2\sqrt{1-\rho^2}\sqrt{\Delta}\tilde{Z}_{k+1}\right) \end{cases}$
where $\gamma = 4$, $[x_0, y_0] = [20, 25]$, $[\mu_1, \mu_1] = [0.01, -0.05]$, $[\sigma_1, \sigma_2] = [0.09, 0.2]$, $\rho = 0.1$, $n = 1000$.

In these previous examples, we have to approximate the quantities of interest: $\mathbb{E}(h(X))$, $X \in \mathbb{R}^d$ in general. We may use several methods as

1. deterministic methods like rectangular rule, trapezoidal rule, Simpson's rule (which depend on the dimension of the vector $X$),
2. or use probabilistic approximation method like the Monte Carlo simulation method.
3. The main advantage of the Monte Carlo method is that
    1. it is easy to implement as soon as we may *generate a sample* from $X$: it is approximated by $(h(X_1) + \ldots + h(X_N))/N$ where $(X_1, \ldots, X_N)$ is a sample of $X$ of size $N$
    2. it does not depends on the dimension $d$ of $X$

Our aim is now to say how to generate samples from $X$ in several situations. We will first recall some usual distributions

# Plan

▷ Bernoulli distribution. It models a random experiment with two outcomes: success with probability $p$, failure with probability $q = 1 - p$. If $X \sim \mathrm{Bern}(p)$ then

$$\mathbb{P}(X = 1) = p; \qquad \mathbb{P}(X = 0) = 1 - p,$$

and

$$\mathbb{E}(X) = p, \qquad \mathrm{Var}(X) = p(1 - p).$$

Loi binomiale with parameters $n$ and $p$. It is the distribution of the success numbers from a sample of $n$ independent Bernoulli experiment with parameter $p$. If $X \sim B(n, p)$ then

$$\mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}, \qquad k = 0, \dots, n.$$

We may write $X \sim B(n, p)$ as $X = \sum_{i=1}^n X_i$ with $X_i \sim \mathrm{Bern}(p)$, $i = 1, \dots, n$ so that

$$\mathbb{E}(X) = np, \qquad \mathrm{Var}(X) = np(1 - p).$$

# Poisson distribution

Poisson distribution. $X \sim \mathrm{Poiss}(\lambda)$ if its distribution reads

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

We have

$$\mathbb{E}(X) = \lambda, \qquad \mathrm{Var}(X) = \lambda.$$

Geometric random variable. A random variable $X \sim \mathcal{G}(p)$ if its distribution reads

$$\mathbb{P}(X = k) = (1 - p)^{k-1}p, \qquad k = 1, 2, \ldots$$

and we have

$$\mathbb{E}(X) = \frac{1}{p} \qquad \text{et} \quad \mathrm{Var}(X) = \frac{1-p}{p^2}.$$

Uniform distribution on $]a, b[$, $a < b$. It models a r.v. which behaves uniformly on. If $X \sim U(]a, b[)$ then, its density reads

$$f(x) = \frac{1}{b-a}\mathbf{1}_{]a,b[}(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in ]a, b[ \\ 0 & \text{otherwise.} \end{cases}$$

If $X \sim U([a, b])$ then $\mathbb{E}(X) = (a + b)/2$, $\mathrm{Var}(X) = (b - a)^2/12$.

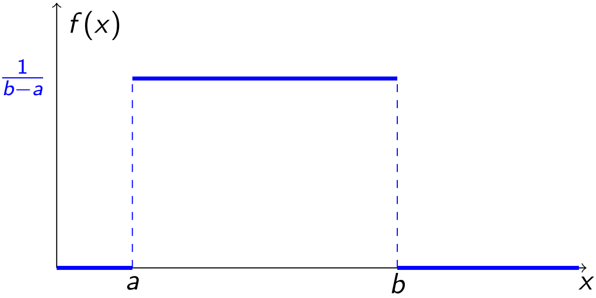
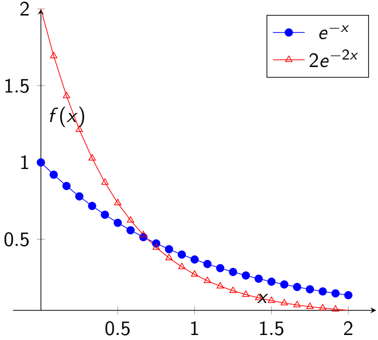



Figure: Density of an uniform distribution on $]a, b[$.

Exponential distribution. It describes phenomenons like waiting times, lifetime, etc. If $1/\lambda, \lambda > 0$ is the expected lifetime (or the expected waiting time) then, the density of $X$ is defined as

$$f(x) = \lambda e^{-\lambda x}\mathbf{1}_{\{x \geq 0\}} = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

If $X \sim \mathcal{E}(\lambda)$, we have $\mathbb{E}(X) = \frac{1}{\lambda}, \ \mathrm{Var}(X) = \frac{1}{\lambda^2}$.



Gaussian distribution. (Laplace-Gauss or Normal distribution) One of the more used in applied probability. A r.v. $X$ is Gaussian with expected value $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$: $X \sim \mathcal{N}(\mu, \sigma^2)$, if its density reads

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \quad \forall x \in \mathbb{R}.$$

It is said standard if $\mu = 0$ and $\sigma^2 = 1$. If $X \sim \mathcal{N}(\mu, \sigma^2)$,

$$\mathbb{E}(X) = \mu, \qquad \mathrm{Var}(X) = \sigma^2.$$

Proposition. If $X \sim \mathcal{N}(\mu, \sigma^2)$ then

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1), \tag{2}$$

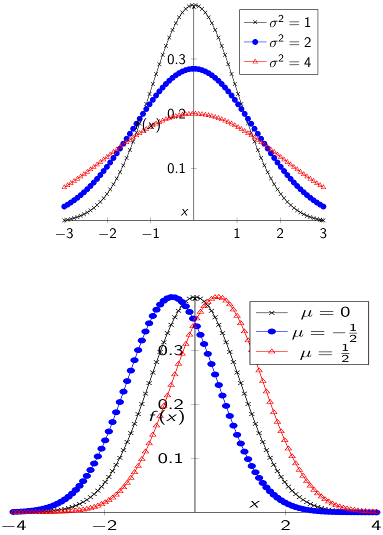



Figure: Density of a $\mathcal{N}(\mu; \sigma^2)$. Figure 1: $\mu = 0$; Figure 2: $\sigma = 1$.

Gamma distribution. A r,v. X has a Gamma distribution with parameters $a > 0$ and $\lambda > 0$: $X \sim \Gamma(a.\lambda)$, if its density reads

$$f(x) = \begin{cases} \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\Gamma(\cdot)$ is the Gamma function definedfor any $a > 0$, by

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx.$$

If $X \sim \Gamma(a, \lambda)$ then

$$\mathbb{E}(X) = a/\lambda \quad \text{and} \quad \mathrm{Var}(X) = a/\lambda^2.$$

The exponential distribution is a particular case of a Gamma distribution:

$$X \sim \mathcal{E}(\lambda) \Longleftrightarrow X \sim \Gamma(1, \lambda).$$

Beta distribution. A r.v. $X$ has a Beta distribution with parameters $a, b > 0$: $X \sim \mathrm{Beta}(a, b)$ if its density reads

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & \text{if } x \in ]0, 1[ \\ 0 & \text{otherwise,} \end{cases}$$

where $B(\cdot, \cdot)$ is the Beta function defined as

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

If $X \sim \mathrm{Beta}(a, b)$ then,

$$\mathbb{E}(X) = \frac{a}{a+b} \quad \text{et} \quad \mathrm{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

Chi-square distribution with $n$ degrees of freedom. Let $X_1, \ldots, X_n$ be $n$ independent $\mathcal{N}(0,1)$-distributed r.v. Then

$$X = \sum_{i=1}^{n} X_i^2$$

has a Chi-square distribution with $n$ degrees of freedom: $X \sim \chi^2(n)$ or $X \sim \chi^2_n$.

The Student with $n$ degrees of freedom. Let $Z \sim \mathcal{N}(0,1)$ and $U \sim \chi^2(n)$ be two independent r.v. Then

$$T = \frac{Z}{\sqrt{U/n}}$$

has the Student with $n$ degrees of freedom: $X \sim T(n)$.

$\rightsquigarrow$ The term

*Generating* a random number $x$ from a r.v. $X$ or

*Simulating* a realization $x$ of $X$ or

*Sampling* a random number $x$ from $X$

consists on mimicking the r.v. in order to generate one possible value (or observation) $X(\omega) = x$ from $X$.

$\rightsquigarrow$ *Example*. Let $X$ be a Bernouilli random variable with success parameter $p$: $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$.

1. When we sample a random number $x$ from $X$, $x = 1$ or $x = 0$.
2. When the sample is of size $N$: $X_1(\omega) = x_1, \ldots, X_N(\omega) := x_N$ are iid with $X_i \stackrel{d}{=} X$, it must be in line with the theoretical results as the Law of Large Numbers: $\bar{X}_N := \frac{X_1 + \ldots + X_N}{N} \stackrel{N \to +\infty}{\longrightarrow} \mathbb{E}(X) = p$, a.s.

$\rightsquigarrow$ There are many simulation techniques: the inversion method, the rejection method, the transformation method, etc ...

# Plan

Compute the c.d.f $F$ and its generalized inverse $F^{-1}$ defined as

$$F^{-1}(u) := \inf\{x \in \mathbb{R} : \quad F(x) \geq u\}, \qquad \forall u \in ]0,1[,$$

for

1. the bernoulli distribution with success parameter $p$,
2. for a random variable $X$ taking the values in $\{-3, 1, 2\}$ with $\mathbb{P}(X = -3) = 1/6$, $\mathbb{P}(X = 1) = 1/2$, $\mathbb{P}(X = 2) = 1/3$.
3. for a random variable taking values in $\{x_1, \ldots, x_n\}$ with associated probabilities $\{p_1, \ldots, p_n\}$.
4. for a random variable taking values in $\{x_1, \ldots, x_n, \ldots\}$ with associated probabilities $\{p_1, \ldots, p_n, \ldots\}$.
5. for a random variable $X$ with density (the exponential distribution with parameters $\lambda$)

$$f(x) = \lambda\, e^{-\lambda x}, \qquad \lambda > 0.$$

*Proposition. Let $U$ be a r.v., uniformly distributed on $]0,1[$ and let $X$ be a r.v. with cumulative distribution function (cdf) $F$ and (generalized) inverse function $F^{-1}$:*

$$F^{-1}(u) := \inf\{x \in \mathbb{R} : \quad F(x) \geq u\}, \qquad \forall u \in ]0,1[.$$

*Then $X$ and $F^{-1}(U)$ have the same distribution: $X \stackrel{d}{=} F^{-1}(U)$.*

*Proof*. We need to prove that $\forall x \in \mathbb{R}$, $\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(U) \leq x)$. We have

$$\forall u \in ]0,1[,\ \forall x \in \mathbb{R}, \qquad F^{-1}(u) \leq x \iff u \leq F(x).$$

Then

$$\begin{aligned} \mathbb{P}(F^{-1}(U) \leq x) &= \mathbb{P}(U \leq F(x)) \qquad (F \text{ in nondecreasing}) \\ &= F(x) \end{aligned}$$

It follows that the cdf of $F^{-1}(U)$ and $X$ are the same.

# Plan

Let $X$ be a discrete r.v. taking values in $E = \{x_0, \ldots, x_n, \ldots\}$, with cdf $F$. Suppose that the $x_k$ are ordered in a nondecreasing order and denote, $\forall k \geq 0$, $p_k = \mathbb{P}(X = x_k)$ and $c_k = p_0 + \ldots + p_k$. Then, for all $u \in ]0,1[$,

$$F^{-1}(u) = x_0 \mathbb{1}_{\{u \leq c_0\}} + \sum_{k \geq 1} x_k \mathbb{1}_{\{c_{k-1} < u \leq c_k\}}.$$

$\rightsquigarrow$ *When the cardinality $N$ of $E$ is finite* . We stock the values $x_k$ on a table $x$ and those of $c_k$ on a table $c$. To generate a sample $X(\omega)$ of $X$ we use the following algorithm (`rand` generate a r.n. from $U \sim \mathcal{U}(]0,1[)$) :

k $\leftarrow$ 0; u $\leftarrow$ rand
while $(u > c[k])$ and $(k < N)$
$k \leftarrow k + 1$
end
$X(\omega) \leftarrow x[k]$

$\rightsquigarrow$ *Bernoulli distribution*. Let $X$ be Bernoulli r.v. with success probability $p \in [0,1]$: $\mathbb{P}(X=0) = 1-p$ and $\mathbb{P}(X=1) = p$. In this case,

$$F^{-1}(u) = 0 \times \mathbb{1}_{\{u<1-p\}} + 1 \times \mathbb{1}_{\{1-p \le u\}} = \mathbb{1}_{\{1-p\le u\}}.$$

We generate a random number (r.n.) $X(\omega) = x$ from $X$ using the following algorithm:

u $\leftarrow$ rand
if $(u < 1-p)$ $x \leftarrow 0$
else $x \leftarrow 1$

$\rightsquigarrow$ *Poisson distribution*. Let $X$ be a r.v. having a Poisson distribution with parameter $\lambda > 0$, defined as:

$$p_k = \mathbb{P}(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, \ldots$$

We remark that

$$p_k = \frac{\lambda}{k}\, p_{k-1}, \qquad \forall\ k \ge 1.$$

$\rightsquigarrow$ To generate a r.n. from $X$, we first stock the values of the cdf $F(n)$, $n \in \{1, 2, \dots, N\}$, where $N$ is chosen such that $F[N]$ is high (for example $F[N] = 0.999$)

$\rightsquigarrow$ Then, we use the following algorithm ($pN \equiv p_N = \mathbb{P}(X = N)$):

```
u ← rand
if (u ≤ F[N])
   then
      k ← 0
      while (u > F[k]) do
         k ← k + 1
      end
   else
      k ← N, p ← pN, F ← F[N]
      while (u > F) do
         k ← k + 1, p ← λ * p/k, F ← F + p
      end
end
```

*Exercise*. 1. Give a program with the following entries

- a vector $x = [x_1, \ldots, x_n]$ of values taken by a random variable $X$
- a vector $p = [p_1, \ldots, p_n]$ of the associated probabilities

and with output a random number generated from $X$.

We set

$$n = 3, \quad x = [1, 2, 3], \quad p = [1/2, 1/3, 1/6].$$

2. Generate a sample $X_1, \ldots, X_N$ of size $N = 10$ of $X$.
3. Generate a sample $X_1, \ldots, X_N$ of size $N = 10000$ of $X$ and compare, for $i = 1, 2, 3$, the empirical frequency

$$f_i = \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{\{X_k = x_i\}}$$

and compare it with $p_i$.

# Plan

$\rightsquigarrow$ *The exponential distribution*. If $X$ has an exponential distribution with parameter $\lambda > 0$, with density

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{]0,+\infty[}(x),$$

so that its cdf reads

$$F(x) = (1 - e^{-\lambda x}) \mathbb{1}_{]0,+\infty[}(x),$$

then, for any $u \in ]0,1[$, $F^{-1}(u) = -\frac{\ln(1-u)}{\lambda}$, so that if $U \sim \mathcal{U}(]0,1[)$, then,

$$F^{-1}(U) = -\frac{\ln(1-U)}{\lambda} \stackrel{d}{=} -\frac{\ln(U)}{\lambda} \qquad (\text{since } 1 - U \stackrel{d}{=} U).$$

$\rightsquigarrow$ *The Weibull distribution*. Let $X$ be a Weibull distribution with parameters $(\lambda, a)$, with density

$$f(x) = \lambda a x^{a-1} e^{-\lambda x^a}, \qquad \lambda, a > 0.$$

Its cdf reads

$$F(x) = (1 - e^{-\lambda x^a})\mathbb{1}_{]0,+\infty[}(x).$$

It follows that for any $u \in ]0,1[$, $F^{-1}(u) = \big(-\ln(1-u)/\lambda\big)^{1/a}$, so that if $U \sim \mathcal{U}(]0,1[)$, then,

$$F^{-1}(U) \; = \; \big(-\ln(1-U)/\lambda\big)^{1/a} \; \stackrel{d}{=} \; \big(-\ln(U)/\lambda\big)^{1/a}.$$

$\rightsquigarrow$ As a consequence, if we want to generate a random number from an exponential distribution or a Weibull distribution we just have

- to generate a r.n. $u = U(\omega)$ from a uniform distribution $U \sim \mathcal{U}(]0,1[)$, and
- compute the inverse $F^{-1}(u)$ with respect to the associated distribution.

# Plan

# Plan

⇝ Let $a, b \in \mathbb{R}$ with $a < b$; $S = ]a, b[$ and let $U \sim \mathcal{U}(S)$. Compute $\mathbb{P}(U \in A)$, for any $A \subset S$.

⇝ Let $S$ be Borel set on $\mathbb{R}^d$ and let $U \sim \mathcal{U}(S)$ with density (w.r.t. the Lebesgues measure $\lambda_d$): $f(x) = (1/\lambda_d(S))\,\mathbb{1}_S(x)\lambda_d(dx)$. For any Borel set $A \subset S$,

$$\mathbb{P}(U \in A) = \int_A \frac{1}{\lambda_d(S)}\lambda_d(dx) = \frac{\lambda_d(A)}{\lambda_d(S)} = \frac{|A|}{|S|}.$$

If $d = 2$, we have

$$\mathbb{P}(U \in A) = \frac{\text{area}(A)}{\text{area}(S)}, \qquad A \subset S.$$

*Proposition*. Let $(U_n)_{n\geq 1}$ be sequence of iid r.v. with $U_1 \sim \mathcal{U}(S)$. Let $A \subset S$ and $\tau = \inf\{n \geq 1,\ U_n \in A\}$. Then $U_\tau \sim \mathcal{U}(A)$.

*Proof*. We have for any $B \subset A$,

$$\mathbb{P}(U_\tau \in B) = \sum_{k=1}^{+\infty} \mathbb{P}(U_k \in B; \tau = k)$$

Now, it follows from the independence of the $U_k$'s that

$$\begin{aligned}
\mathbb{P}(U_k \in B; \tau = k) &= \mathbb{P}(\{U_k \in B\} \cap \{U_1 \notin A\} \cap \ldots \cap \{U_{k-1} \notin A\} \cap \{U_k \in A\}) \\
&= \mathbb{P}(\{U_k \in B\} \cap \{U_k \in A\}) \times \mathbb{P}(\{U_1 \notin A\})^{k-1} \\
&= \mathbb{P}(\{U_k \in B\}) \times \mathbb{P}(\{U_1 \notin A\})^{k-1} \quad \text{since} \;\; B \subset A.
\end{aligned}$$

Since

$$\mathbb{P}(\{U_k \in B\}) = \frac{|B|}{|S|} \quad \text{and} \quad \mathbb{P}(\{U_1 \notin A\}) = 1 - \frac{|A|}{|S|}$$

we deduce that for any Borel set $B \subset A$,

$$\mathbb{P}(U_\tau \in B) = \sum_{k=1}^{+\infty} \left(1 - \frac{|A|}{|S|}\right)^{k-1} \frac{|B|}{|S|} = \frac{|B|}{|A|}.$$

As a consequence

$$U_\tau \sim \mathcal{U}(A).$$

*Example*. *Random numbers uniformly distributed on the unit sphere*. Let $X$ be an uniform distribution on the unit sphere $A = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 \leq 1\}$ and let $S$ be the cube $]-1, +1[^3$ on $\mathbb{R}^3$.

- We have $A \subset S$.
- If $U_1, U_2, U_3 \sim \mathcal{U}(]-1, 1[)$, are independent then, $(U_1, U_2, U_3) \sim \mathcal{U}(S)$.
- To sample a r.n. from $\mathcal{U}(A)$ we use the algorithm:

```
do   u1 ← 2*rand -1
     u2 ← 2*rand -1
     u3 ← 2*rand -1
while (u1*u1 + u2*u2 +u3*u3 > 1)
end
U1 ← u1, U2 ← u2 and U3 ← u3
```

# Plan

Let $f$ and $g$ be explicit probability densities in $\mathbb{R}^d$, $c \geq 1$, and suppose that

- we can simulate a r.n. from the density $g$ but not from $f$.
- $f(x) \leq c\, g(x)$, for any $x \in \mathbb{R}^d$.

*Example*. Let $X \sim \mathcal{N}(0; 1)$.

- $|X|$ has density $f(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \mathbb{1}_{\{x>0\}}$,
- $\forall\, x > 0$, $f(x) \leq c\ g(x)$, with $c = \sqrt{(2e/\pi)}$ and $g(x) = e^{-x}$,
- if $\Theta \in \{-1, +1\}$ with $\mathbb{P}(\Theta = +1) = \mathbb{P}(\Theta = -1) = 1/2$ then $\Theta|X| \sim \mathcal{N}(0; 1)$.
- This is used to sample from a standard normal distribution.

The following algorithm generates a r.v. $X$ with density $f$:

1. generate a r.n $x$ from the density $g$ and a r.n $u$ from $U \sim \mathcal{U}(]0, 1[)$
2. if $c \times u \times g(x) \leq f(x)$, go to 3., otherwise go to 1.
3. return $X(\omega) = x$.

This previous procedure follows from the result below.

*Proposition*. *Let $f$ and $g$ be two densities and let $c \geq 1$ be so that $f \leq cg$. Let $(X_k)_{k\geq 1}$ be an iid sequence of r.v. with density $g$ and let $(U_k)_{k\geq 1}$ be an iid sequence with distribution $\sim \mathcal{U}(]0,1[)$, independent from $X_1$. Then, the r.v. $X_\tau$ where*

$$\tau = \inf\{k \geq 1,\ c\, U_k g(X_k) \leq f(X_k)\}$$

*has density $f$ and $\tau$ has a geometric distribution with success parameter $1/c$.*

*Proof*. *We have for every $x \in \mathbb{R}$,*

$$\mathbb{P}(X_\tau \leq x) = \sum_{k=1}^{+\infty} \mathbb{P}(X_k \leq x; \tau = k)$$

⇝ Now, we have (letting $h(x) = f(x)/(cg(x))$)

$$\mathbb{P}(X_k \le x; \tau = k) = \mathbb{P}(X_k \le x; U_k \le h(X_k)) \times \big(\mathbb{P}(U_1 > h(X_1))\big)^{k-1}$$

and

$$\mathbb{P}(U_1 > h(X_1)) = \int_{-\infty}^{+\infty} g(t)dt \int_{h(t)}^{1} du = 1 - 1/c.$$

⇝ On the other hand,

$$\mathbb{P}(X_k \le x; U_k \le h(X_k)) = \int_{-\infty}^{x} g(t)dt \int_{0}^{h(t)} du = \int_{-\infty}^{x} g(t)h(t)dt = \frac{1}{c}\int_{-\infty}^{x} f(t)dt.$$

It follows that

$$\mathbb{P}(X_\tau \le x) = \frac{1}{c}\int_{-\infty}^{x} f(t)dt \sum_{k=1}^{+\infty}\Big(1 - \frac{1}{c}\Big)^{k-1} = \int_{-\infty}^{x} f(t)dt,$$

so that $Z$ has density $f$.

$\rightsquigarrow$ *The Gamma distribution*. Let $\lambda, a > 0$ and let $X$ be a r.v. with Gamma distribution $\Gamma(\lambda, a)$, with pdf

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \qquad \text{where} \qquad \Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx.$$

We want to generate a r.n. from the distribution of $X \sim \Gamma(\lambda, a)$. First note that if $Z \sim \Gamma(1, a)$, then $X = Z/\lambda \sim \Gamma(\lambda, a)$, so that it is enough to say how to simulate a r.n. from $Z$.

- When $a = n$ is an integer number then $Z \stackrel{d}{=} E_1 + \ldots + E_n$, where the $E_k$'s are iid exponentially distributed r.v. with param. 1: $E_k \sim \mathcal{E}(1)$.
- If $a \in ]0, 1[$ (and $\lambda = 1$), we have $f(x) \leq cg(x)$, where

$$c = \frac{e + a}{ae\Gamma(a)} \quad \text{and} \quad g(x) = \frac{ae}{e+a} \left[ x^{a-1} \mathbb{1}_{]0,1[}(x) + e^{-x} \mathbb{1}_{[1,+\infty[} \right].$$

We can apply the rejection algorithm to generate a r.n. from $Z$.

In fact, if $X$ has pdf $g$ its inverse function reads for every $u \in ]0,1[$,

$$G^{-1}(u) = \left(\frac{e+a}{e}u\right)^{\frac{1}{a}} \mathbb{1}_{]0,\frac{e}{e+a}[}(u) - \ln\left((1-u)\frac{e+a}{ae}\right)\mathbb{1}_{]\frac{e}{e+a},1[}(u).$$

and $h(x) = f(x)/(cg(x))$ reads

$$h(x) = e^{-x}\mathbb{1}_{]0,1[}(x) + x^{a-1}\mathbb{1}_{[1,+\infty[}.$$

Then, to generate a r.n. from $Z \sim \Gamma(1,a)$, $a \in ]0,1[$,

1. we generate a r.n. $V$ from $\mathcal{U}(]0,1[)$, we compute $X = G^{-1}(V)$ and generate another r.n. $U$ from $\mathcal{U}(]0,1[)$, independent from $V$.
2. If $U \leq h(X)$, we set $Z = X$, otherwise, return to step 1.

$\rightsquigarrow$ *The Beta distribution*. The Beta distribution with parameters $a, b > 0$ has pdf

$$f(x) = B(a,b)x^{a-1}(1-x)^{b-1}\mathbb{1}_{]0,1[}(x) \quad \text{with} \quad B(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}.$$

For $a, b > 1$, we have

$$f(x) \le cg(x) \quad \text{where} \quad c = \left(\frac{a-1}{a+b-2}\right)^{a-1}\left(\frac{b-1}{a+b-2}\right)^{b-1}$$

$$\text{and} \qquad g(x) = \mathbb{1}_{]0,1[}(x).$$

Then, to generate a r.n. from $Z \sim B(a,b)$,

1. we generate a r.n. $X$ from $\mathcal{U}(]0,1[)$ and generate another r.n. $U$ from $\mathcal{U}(]0,1[)$, independent from $X$.
2. If $U \le h(X)$, we set $Z = X$, otherwise, return to step 1.

# Plan

$\rightsquigarrow$ Some times, the random variable to generate reads as a function of easy simulable random variables. This method is specific to some random variables and we are going to give examples of the Gamma distribution and the Gaussian distribution.

$\rightsquigarrow$ Let $T = (T_1, \ldots, T_d) : \mathbb{R}^d \mapsto \mathbb{R}^d$ be a diffeomorphism whose inverse has Jacobian matrix

$$J(z) = \Big(\frac{d}{dz_j} T_i^{-1}(z)\Big)_{1 \le i,j \le d}.$$

It follows that if $Z = T(X)$, where $X$ is an $\mathbb{R}^d$-valued random vector with pdf $f_X$, then, the pdf of $Z$ reads

$$f_Z(z) = f_X(T^{-1}(z)) \times |\det(J(z))|, \qquad z \in \mathbb{R}^d.$$

# Plan

$\rightsquigarrow$ *Example of the Gamma distribution*. Let $\lambda > 0$ and $a_i > 0$, $i = 1, \ldots, n$. Let $X_i \overset{\text{iid}}{\sim} \Gamma(\lambda, a_i)$, $i = 1, \ldots, n$. Then, we know that $Z = X_1 + \ldots + X_n \sim \Gamma(\lambda, a_1 + \ldots + a_n)$. Suppose that $a_i = 1$ for any $i$.

- Since $\Gamma(\lambda, 1) \sim \text{Exp}(\lambda)$, we can represent $Z \sim \Gamma(\lambda, n)$ as: $Z = X_1 + \ldots + X_n$, with $X_i \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$.
- We know from the inversion method that if $X_i \sim \text{Exp}(\lambda)$ then $X_i \overset{d}{=} -\ln(U_i)/\lambda$, where $U_i \sim \mathcal{U}(]0, 1[)$.
- Then, $Z = X_1 + \ldots + X_n \sim \Gamma(\lambda, n)$ can be written as

$$Z = T(U_1, \ldots, U_n) = -\frac{1}{\lambda} \sum_{i=1}^{n} \ln(U_i) = -\frac{1}{\lambda} \ln\Big(\prod_{i=1}^{n} U_i\Big).$$

# Plan

Let $Z = (Z_1, Z_2)$ be a two dimensional Gaussian vector. The following result, known as the *Box-Muller* method, say how to simulate $Z$ from independent uniform random variables.

*Proposition*. Let $U_i \overset{\text{iid}}{\sim} \mathcal{U}(]0,1[)$, $i = 1, 2$. Then,

$$(Z_1, Z_2) = T(U_1, U_2) = \left(\sqrt{-2\ln(U_1)}\cos(2\pi U_2), \sqrt{-2\ln(U_1)}\sin(2\pi U_2)\right)$$

is a pair of indep. standard Normal distribution: $Z_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$, $i = 1, 2$.

*Proof*. Note that the transformation $T : ]0,1[^2 \mapsto \mathbb{R}^2$ is one to one and if $z_1 = \sqrt{-2\ln(u_1)}\cos(2\pi u_2)$ and $z_2 = \sqrt{-2\ln(u_1)}\sin(2\pi u_2)$, we have $z_1^2 + z_2^2 = -2\ln(u_1)$ and $z_2/z_1 = \tan(2\pi u_2)$. Then $(z = (z_1, z_2))$

$$(u_1, u_2) = \left(T_1^{-1}(z), T_2^{-1}(z)\right) = \left(e^{-(z_1^2+z_2^2)/2}, (2\pi)^{-1}\arctan(z_2/z_1)\right).$$

The Jacobian matrix

$$J(z) = \begin{pmatrix} \frac{\partial T_1^{-1}(z)}{\partial z_1} & \frac{\partial T_1^{-1}(z)}{\partial z_2} \\ \frac{\partial T_2^{-1}(z)}{\partial z_1} & \frac{\partial T_2^{-1}(z)}{\partial z_2} \end{pmatrix} = \begin{pmatrix} -z_1 e^{-(z_1^2+z_2^2)/2} & -z_2 e^{-(z_1^2+z_2^2)/2} \\ -\frac{1}{2\pi}\frac{z_2}{z_1^2+z_2^2} & \frac{1}{2\pi}\frac{z_1}{z_1^2+z_2^2} \end{pmatrix}$$

It follows that $\det(J(z)) = \frac{1}{2\pi} e^{-(z_1^2+z_2^2)/2}$. Then

$$f_Z(z) = f_{(U_1,U_2)}\left(T^{-1}(z)\right) |\det(J(z))| = \frac{1}{2\pi} e^{-(z_1^2+z_2^2)/2} = \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \frac{1}{\sqrt{2\pi}} e^{-z_2^2/2}$$

so that $(Z_1, Z_2) \sim \mathcal{N}(0_2, I_2)$: $0_2 = (0,0)$, $I_2$ is the $2 \times 2$ identity matrix.

$\rightsquigarrow$ *Generating a Gaussian vector $X \sim \mathcal{N}(0_d, I_d)$*. One way of generating a r.n. from $X \sim \mathcal{N}(0_d, I_d)$ is

- to call $d$ times the function generating a gaussian pair $(Z_1, Z_2) \sim \mathcal{N}(0_2, I_2)$ (using for example the *Box-Muller* method)
- and to set $X = (Z_1^1, \ldots, Z_1^d)$, where $Z_1^i$ is the value of $Z_1$ at the $i$-th call of the function generating $(Z_1, Z_2)$.

$\rightsquigarrow$ *Drawing a Gaussian vector* $Z \sim \mathcal{N}(\mu, \Sigma)$, *where* $\mu \in \mathbb{R}^\ell$ *and* $\Sigma$ *is a* $\ell \times d$ *matrix*. We can write $Z = \mu + \Sigma^{1/2} X$, where $X \sim \mathcal{N}(0_d, I_d)$. We have seen how to draw a sample from $X$. It remains to say how to compute $\Sigma^{1/2}$. Several methods exist but we recall here the two main methods.

1. In the non degenerated case where $\Sigma$ is positive-definite we may use the *Cholesky* decomposition: find a lower triangular matrix $L$ so that $LL^T = \Sigma$ and $\Sigma^{1/2} = L$.
2. In general (including the degenerated case) we may use $L = UA^{1/2}$ where $A$ and $U$ are obtained from the spectral (or eigenvalue) decomposition $\Sigma = UAU^{-1}$ of $\Sigma$. Then, $\Sigma^{1/2} = L$.

# Plan

$\rightsquigarrow$ Let $X$ be an $\mathbb{R}^d$-valued random variable with pdf $f = \sum_{n\geq 0} p_n f_n$, where $\mathbb{Q} := (p_n, n \geq 0)$ is a probability on $\mathbb{N}$ and $f_n$ is pdf for every $n \geq 0$.

$\rightsquigarrow$ Let $(X_n)_{n\geq 0}$ be an iid sequence of r.v. such that for every $n \geq 0$, $X_n$ has pdf $f_n$.

$\rightsquigarrow$ Let $\nu : \Omega \mapsto \mathbb{N}$ be a r.v. with distribution $\mathbb{Q}$, independent from $(X_n)_{n\geq 0}$.

*Proposition*. The random variable $X_\nu$ has pdf $f$.

*Proof*. For any Borel set $A \subset \mathbb{R}^d$, we have

$$\begin{aligned}
\mathbb{P}(X_\nu \in A) &= \sum_{n\geq 0} \mathbb{P}(\nu = n)\mathbb{P}(X_\nu \in A|\nu = n) \\
&= \sum_{n\geq 0} p_n \int_A f_n(x)dx \\
&= \int_A \sum_{n\geq 0} p_n f_n(x)dx = \int_A f(x)dx.
\end{aligned}$$

*Example*. Let $(p_1, p_2, p_3) = (1/6, 1/3, 1/2)$ and $X$ be a random variable with pdf

$$f(x) = p_1 f_1(x) + p_2 f_2(x) + p_3 f_3(x)$$

where

$$f_1(x) = \mathbb{1}_{]0,1]}(x), \quad f_2(x) = \frac{1}{2}(2x-1)\mathbb{1}_{]1,2]}(x), \quad f_3(x) = \frac{2}{3}(-3x+9)\mathbb{1}_{]2,3]}(x).$$

Propose an algorithm to generate a sample from $X$.

*Example*. Let $X_1 \sim \mathcal{N}(-3, 1)$ and $X_2 \sim \mathcal{N}(3, 1)$ be two independent r.v. with resp. pdf $f_1$ and $f_2$. Let $X$ be a r.v. with pdf

$$f(x) = p_1 f_1(x) + p_2 f_2(x), \qquad p_1, p_2 \in [0, 1],\ p_1 + p_2 = 1.$$

- Plot the graphs of $f$ for $(p_1, p_2) = (\frac{1}{2}, \frac{1}{2})$, $(p_1, p_2) = (\frac{1}{4}, \frac{3}{4})$, $(p_1, p_2) = (\frac{3}{4}, \frac{1}{4})$.
- Plot in the same graph (w.r.t $(p_1, p_2)$) the densities estimates from a sample of $f$.

## Exercise to do in the practical session

Consider the file `ToBeGropuedRandomly.xlx` and implement a code that compose randomly 5 groups of 3 individuals where, at each step $k \leq 5$ and for each group,

- 1 individual of the group is chosen randomly from the uniform distribution among the $5 - k + 1$ first individuals of the file
- the 2 other individual of the group are chosen randomly from the uniform distribution among the remaining individuals

Display the groups with their set of individuals.

# Plan

1. Annie Millet. *Méthodes de Monte-Carlo*: https://samos.univ-paris1.fr/archives/ftp/cours/millet/montecarlo.pdf

2. Otten, R.H.J.M., and van Ginneken, L.P.P.P. 1989, The Annealing Algorithm(Boston: Kluwer).

3. Gilles Pagès. Numerical Probability: an Introduction with Applications to Finance, 587p., Springer, août 2018

4. Matthias Winkel. *Simulation*: http://www.stats.ox.ac.uk/ winkel/ASim11.pdf