Exercises sheet n°1 : Introduction to MC and IS

Exercise 1 - An example in Finance
In financial applications, we have to calculate quantities of the type

$$C = \mathbb{E}\left(\left(e^{\beta G} - K\right)_+\right), \tag{1}$$

G being a standard Gaussian rv and $x_+ = \max(0, x)$. These quantities represent the price of an **option to buy**, commonly called a "call".

Similarly, an **option to sell**, called "put", is defined by

$$P = \mathbb{E}\left(\left(K - e^{\beta G}\right)_+\right). \tag{2}$$

1. Prove the following explicit formula :

$$C = e^{\beta^2/2} N\left(\beta - \frac{\log K}{\beta}\right) - KN\left(-\frac{\log K}{\beta}\right)$$

   where $N(x) = \int_{-\infty}^{x} e^{-u^2/2} \frac{du}{\sqrt{2\pi}}$.

2. Similarly, prove the following explicit formula :

$$P = KN\left(\frac{\log K}{\beta}\right) - e^{\beta^2/2} N\left(\frac{\log K}{\beta} - \beta\right).$$

3. We apply Monte Carlo simulation to estimate C and compare the exact value to results of a simulation based on various sizes of samples in the case $\beta = K = 1$.

	exact value	:	6.72
n=100	estimated 95% CI	:	[0.08,11.39]
	estimated value	:	5.74
n=1000	estimated 95% CI	:	[4.20,10.01]
	estimated value	:	7.1
n=10^4	estimated 95% CI	:	[6.13,8.43]
	estimated value	:	7.28
n=10^5	estimated 95% CI	:	[6.59,7.69]
	estimated value	:	7.14

   We compare these results with those obtained when evaluating an option to sell. We then obtain

	exact value	:	0.23842
n=100	estimated 95% CI	:	[0.166,0.276]
	estimated value	:	0.220
n=1000	estimated 95% CI	:	[0.221,0.258]
	estimated value	:	0.240
n=10000	estimated 95% CI	:	[0.232,0.244]
	estimated value	:	0.238

The approximation is much better than in the case of a call.
Prove theoretically this observation by a calculation of the variance.

4. [**Importance sampling**] We want to apply IS in the case of the calculation of a put (2). Use a Taylor expansion at order 1 near 0 to rewrite P and derive another estimation procedure for P.

   We then obtain

	`exact value`	:	`0.23842`
$n = 100$	`estimated` 95% `CI`	:	`[0.239,0.260]`
	`estimated value`	:	`0.249`
$n = 1000$	`estimated` 95% `CI`	:	`[0.235,0.243]`
	`estimated value`	:	`0.239`
$n = 10^4$	`estimated` 95% `CI`	:	`[0.237,0.239]`
	`estimated value`	:	`0.238`

   We note a significant improvement with respect to the results based on the classical Monte Carlo : for a 10^4-sample, the RE becomes 1% instead of 6%.
   Prove theoretically this observation by a calculation of the variance.

5. [**Control variables**] Find a simple relation between C and P.
   Derive then a new estimation procedure for C from the one of P.
   Conclude.

6. [**Antithetic variables**] Use the fact that the distribution of G is identical to that of $-G$ to calculate the price of a put (2).
   Prove that in that case we reduce the variance of a coefficient almost by 2.

Exercise 2 - Importance sampling on a basic example
Assume that we want to estimate

$$I = \int_0^1 g(x)dx = \int_0^1 \cos\left(\frac{\pi x}{2}\right) dx.$$

1. Express the previous integral in terms of the expectation of a random variable.

2. To apply IS, we approximate g by a second degree polynomial. Since g is even and equals to 0 at $x = 1$ and to 1 at $x = 0$, it is natural to take $\widetilde{f}(x)$ in the form $\lambda(1-x^2)$. Determine the value λ so that the constraint $\int \widetilde{f}(x)dx = 1$ is satisfied.

3. Calculate the variance of $Z = g(Y)f(Y)/\widetilde{f}(Y)$ and show that we have reduced the variance by a factor of 100.

Exercise 3 - A discrete-time Markov chain
To fix ideas and better understand the difficulties to choose a good change of measure, we study the discrete-time Markov chain Y such as the one depicted in Figure 1 with state space $S = \{0, 1, 2, 3\}$ and $0 < a, b, c, d < 1$. The chain starts at 1 and we wish to evaluate the probability that it gets absorbed by state 3, that is to say $I = \mathbb{P}(X(\infty) = 3|X(0) = 1)$. Obviously here, $I = ac/(1 - ad)$. For instance, when a and c are small, the event $\{X(\infty) = 3\}$ becomes rare.

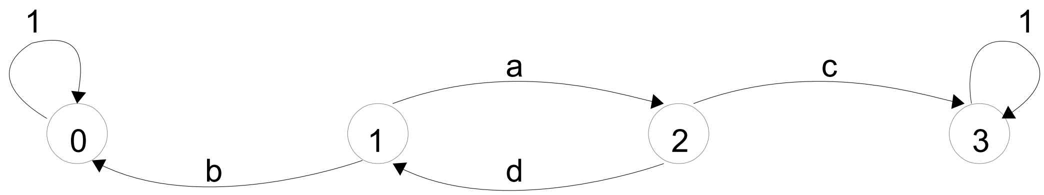

FIGURE 1 – A small discrete-time Markov chain.

For instance, consider the case $a = c = \frac{1}{4}$ and suppose that we decide to make the event of interest $\{X(\infty) = 3\}$ more frequent by changing a to $\widetilde{a} = \frac{1}{2}$ and c to $\widetilde{c} = \frac{3}{4}$. Define $\mathcal{P}$ as the set of all possible paths of X starting at sate 1 :

$$\mathcal{P} = \{\pi = (x_0, x_1, \dots, x_K), K \geqslant 1 \text{ with } x_0 = 1,\ x_K = 0 \text{ or } 3 \text{ and } x_i \notin \{0, 3\} \text{ if } 1 \leqslant i \leqslant K-1\}$$

and $\mathcal{P}_s$ as the set of successful paths (those paths in $\mathcal{P}$ ending with state 3.

1. Observe that
$$\mathcal{P}_s = \{\pi_k,\ k \geqslant 1\}$$
where $\pi_k = (1, (2,1)^k, 2, 3)$ (the notation $(2,1)^k$ meaning that the sequence $(2,1)$ is repeated k times.
2. We have
$$\mathbb{P}(\pi_k) = (ad)^k ac = \left(\frac{1}{4}\,\frac{3}{4}\right)^k \frac{1}{4}\,\frac{1}{4} \quad \text{and} \quad \mathbb{P}(\widetilde{\pi}_k) = \left(\frac{1}{2}\,\frac{1}{4}\right)^k \frac{1}{2}\,\frac{3}{4}.$$
We see that even in such a simple model, finding an appropriate change of measure can be non-trivial.
3. Prove that $\mathbb{P}(\widetilde{\pi}_k) > \mathbb{P}(\pi_k)$ for $k = 0, \dots, 4$ but $\mathbb{P}(\widetilde{\pi}_k) < \mathbb{P}(\pi_k)$ for $k \geqslant 5$.
4. Now consider the following IS scheme : change a to $\widetilde{a} = 1$ and c to $\widetilde{c} = 1 - ad$. Check that
$$L(\pi_k) = \frac{ac}{1 - ad} = I$$
for all k which means that this is the optimal change of measure, the one leading to a zero-variance estimator.

Exercise 4 - Control variables

We want to compute $I = \int_0^1 g(x)dx = \int_0^1 e^x dx$.

1. Express the previous integral in terms of the expectation of a random variable.
2. We approximate g by $1 + x$ and show that
$$\int_0^1 e^x dx = \int_0^1 (e^x - 1 - x)dx + \frac{3}{2}.$$
3. Prove that the variance in this method than reduces significantly.

Exercise 5 - Importance sampling Suppose we want to evaluate the integral $\mu(G) = \int G(x)\mu(dx)$ of a nonnegative and bounded potential function G with respect to some distribution μ on some measurable space $(E, \mathcal{E})$. We associate with a sequence of independent random variables $(X_i)_{i\geqslant 1}$ with common distribution μ the empirical measures $\mu^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$.

1. Check that $\mathbb{E}(\mu^N(G)) = \mu(G)$ and
$$N\mathbb{E}\left(\left(\mu^N(G) - \mu(G)\right)^2\right) = \mu\left((G-\mu(G))^2\right) =: \sigma_\mu(G).$$

2. For any probability measure $\overline{\mu}$ such that $\mu \ll \overline{\mu}$, prove that $\mu(G) = \overline{\mu}(\overline{G})$ with $\overline{G} = G\frac{d\mu}{d\overline{\mu}}$. We let $\overline{\mu}^N := \frac{1}{N}\sum_{i=1}^N \delta_{Y_i}$ be the occupation measure associated with a sequence of N independent random variables $(Y_i)_{i\geqslant 1}$ with common distribution $\overline{\mu}$. Prove that $\mathbb{E}(\overline{\mu}^N(\overline{G})) = \mu(G)$ and
$$\begin{aligned} N\mathbb{E}\left(\left(\overline{\mu}^N(\overline{G}) - \mu(G)\right)^2\right) &= \overline{\mu}\left(\left(\overline{G} - \mu(G)\right)^2\right) =: \sigma_{\overline{\mu}}(\overline{G}) \\ &= \sigma_\mu(G) - \mu\left(G^2\left(1 - \frac{d\mu}{d\overline{\mu}}\right)\right). \end{aligned}$$

3. **An example of potential** G. Roughly speaking, from the equation above, we see that a reduction of variance is obtained as soon as $\overline{\mu}$ is chosen such that $\frac{d\mu}{d\overline{\mu}} < 1$ on regions where G is more likely to take large values. In other words, it is judicious to choose a new reference distribution $\overline{\mu}$ so that the sampled particles $\overline{X}_i$ are more likely to visit regions with high potential. For instance, if $G = \mathbb{1}_A$ is the indicator function of some measurable set $A \in \mathcal{E}$, then prove that
$$\sigma_{\overline{\mu}}(\overline{G}) = \sigma_\mu(G) - \mu\left(1_A\left(1 - \frac{d\mu}{d\overline{\mu}}\right)\right).$$
If we choose $\overline{\mu}$ such that $\frac{d\mu}{d\overline{\mu}} \leqslant 1-\delta$ for any $x \in A$, then check that
$$\overline{\mu}(A) \geqslant \mu(A)/(1-\delta) \qquad \text{and} \qquad \sigma_{\overline{\mu}}(\overline{G}) + \delta\mu(A) \leqslant \sigma_\mu(G).$$

4. **The optimal choice**. Show that the optimal distribution $\overline{\mu}$ is the Boltzmann-Gibbs measure $\overline{\mu}(dx) = \mu(G)^{-1}G(x)\mu(dx)$ in the sense that $\sigma_{\overline{\mu}}(\overline{G}) = 0$. This optimal strategy is clearly hopeless since the normalizing constant $\mu(G)$ is precisely the constant we want to estimate!

5. **A bad choice**. Consider now the distribution $\overline{\mu}$ defined by $\overline{\mu}(dx) = \mu(G^{-2})^{-1}G^{-2}(x)\mu(dx)$, then check that
$$\sigma_{\overline{\mu}}(\overline{G}) \geqslant \mu(G^4)/\mu(G^2) - \mu(G)^2 \geqslant \sigma_\mu(G).$$

Exercise 6 - Simple random walk Let $(\epsilon_n)_{n\geqslant 0}$ be independent and identically distributed random variables with common law $\mathbb{P}(\epsilon_n = 1) = 1 - \mathbb{P}(\epsilon_n = -1) = p \in (0,1)$. We consider the simple random walk X_n on $\mathbb{Z}$ defined by $X_n = \sum_{p=0}^n \epsilon_p$. Suppose we want to evaluate (using a Monte Carlo scheme) the probability that X_n enters a subset $A \subset \mathbb{N}^*$. if we have $p < 1/2$, then the random walk X_n tends to move to the left. One natural way to increase the probability that the random walk visits the set A is to change p by some $\overline{p} \in (p, 1)$. In this case, the random walk Y_n defined as X_n by replacing p by $\overline{p}$ is more likely to move to the right and as a result the event $\{Y_n \in A\}$ is more likely than $\{X_n \in A\}$. The expected value of $f(X_n) = 1\!\!1_A(X_n)$ and the particle approximation mean using the standard Monte Carlo method are given respectively by

$$\mathbb{E}(f(X_n)) = \mathbb{P}(X_n \in A) \qquad \text{and} \qquad \overline{f(X_n)} = \frac{1}{N}\sum_{i=1}^{N} 1\!\!1_A(X_n^i)$$

where $(X_n^i)_{i\leqslant 1}$ is a collection of independent copies of X_n.

1. We let P_n be the distribution of the random sequence $(\epsilon_p)_{0\leqslant p\leqslant n} \in \{-1,+1\}^{n+1}$. Check that
$$P_n(d(u_0,\ldots,u_n)) = (p(1-p))^{(n+1)/2}\,(p/(1-p))^{\sum_{k=0}^n u_k/2}\,.$$
2. We let $\overline{P}_n$ be the distribution of the random sequence $(\overline{\epsilon}_p)_{0\leqslant p\leqslant n} \in \{-1,+1\}^{n+1}$ defined as $(\epsilon_p)_{0\leqslant p\leqslant n} \in \{-1,+1\}^{n+1}$ by replacing p by $\overline{p} \in (0,1)$. Deduce from the first question that $\overline{P}_n \ll P_n$ and
$$\frac{dP_n}{d\overline{P}_n}(u_0,\ldots,u_n) = G_n\left(\sum_{k=0}^{n} u_k\right) \qquad \text{with} \qquad G_n(x) = \left(\frac{p(1-p)}{\overline{p}(1-\overline{p})}\right)^{\frac{n+1}{2}}\left(\frac{p(1-\overline{p})}{\overline{p}(1-p)}\right)^{\frac{x}{2}}.$$
3. Check that $\mathbb{E}(f(X_n)) = \mathbb{E}(f(Y_n)G_n(Y_n))$ for any $f \in \mathcal{B}_b(\mathbb{Z})$.
4. Let $(Y_n^i)_{i\leqslant 1}$ be a collection of independent copies of Y_n. By the Central Limit Theorem, prove that the sequence of random variables
$$\begin{aligned} W_n^N(f) &= \sqrt{N}\left(\overline{f(X_n)} - \mathbb{E}(f(X_n))\right) \\ \overline{W}_n^N(f) &= \sqrt{N}\left(\overline{f(Y_n)G_n(Y_n)} - \mathbb{E}(f(X_n))\right) \end{aligned}$$
converges in law, as $N \to \infty$, to a pair of Gaussian random variables with mean 0 and respective variance $\sigma_n^2(f)$ and $\overline{\sigma}_n^2(f)$ defined by
$$\begin{aligned} \sigma_n^2(f) &= \mathbb{E}\left(f(X_n)^2\right) - \mathbb{E}(f(X_n))^2 \\ \overline{\sigma}_n^2(f) &= \mathbb{E}\left(f(Y_n)^2 G_n(Y_n)^2\right) - \mathbb{E}(f(X_n))^2 \\ &= \sigma_n^2(f) + \mathbb{E}\left(f(X_n)^2\,(G_n(X_n)-1)\right) \end{aligned}$$
5. Prove that for any indicator functions $f = 1\!\!1_A$ with $A \subset \{G_n \leqslant 1/a_n\}$, for some $a_n \geqslant 1$, we have
$$\overline{\sigma}_n^2(f) \leqslant a_n^{-1}\mathbb{P}(X_n \in A) - \mathbb{P}(X_n \in A)^2 \leqslant \sigma_n^2(f).$$

Exercises sheet n°2 : Branching Processes

Exercise 1 A culture of blood starts at time 0 with 1 red blood cell. After one minute, a red blood cell dies and is replaced, with the following probabilities, by

- 2 red blood cells with probability 1/4 ;
- 1 red and 1 white with probability 2/3 ;
- 2 white blood cells with probability 1/12.

Every blood cell lives during one minute and gives birth in the same way that its parent. Every white blood cell lives during one minute and dies without reproducing itself.

a) Evaluate the probability that no white blood cell still appeared at time $n + 1/2$ minute.

b) Evaluate the probability that the entire culture disappears.

Exercise 2 A disease is modelized by a branching process with initial size N germs. At every grip of a medicine (1 a day), every germ has the probability $p = \frac{1}{2}$ to disappear. Determine the law of the duration T of the disease (or of the number of used medicine). Same question, when every germ lives an exponential time of average $\frac{1}{\lambda} = 2$ days. Determine also, for $N = 3$, in every case, the mean duration of the disease.

Exercise 3 We consider a population of bacteria of size X_t at time t such that $X_0 = 1$. Between t and $t + \Delta t$, every bacteria is divided in two new bacteria with probability $\lambda\Delta t + o(\Delta t)$, dies with probability $\mu\Delta t + o(\Delta t)$ where $\lambda \neq \mu$ and does not evolve with probability $1 - (\lambda + \mu)\Delta t + o(\Delta t)$.

a) Let $G(s, t) = \mathbb{E}\left(s^{X_t}\right)$ the probability generating function of X_t. Determine a partial differential equation satisfied by satisfied by G and check that the unique solution such that $G(s, 0) = s$ is

$$G(s, t) = \frac{e^{\alpha t}(1 - s) - 1 + s\rho}{\rho e^{\alpha t}(1 - s) - 1 + s\rho}$$

where $\rho = \frac{\lambda}{\mu}$ and $\alpha = \lambda - \mu$. Determine $\mathbb{E}(X_t)$, $p_0(t)$ and the extinction probability of the bacteria.

b) When $\mu = 0$ compare $\mathbb{E}(X_t)$ with the size of the process such that every bacterium divides every λ^{-1} units of time.

c) Determine $\mathbb{E}(X_n)$ and the extinction probability for the discrete process such that at every unit of time, a bacterium divides in two with probability $\frac{\lambda}{\lambda+\mu}$ and dies with probability $\frac{\mu}{\lambda+\mu}$.

Exercise 4 We consider a population such that the number of direct descendents by individual is distributed as a binomial $\mathcal{B}(2, p)$.

a) Assume we start with 1 individual, determine the extinction probability and the probability that there is nobody anymore, for the first time, at the third generation.

b) Assume now that number of individuals at the first generation is Poisson distributed with parameter λ. Prove that, for $p > \frac{1}{2}$, the extinction probability is $\pi = \exp[\lambda(1 - 2p)/p^2]$.

Exercise 5 We consider a population of particles that undergo a shock every minute. Then the particle may divide in 2 (with probability p) or disappear (with probability $1 - p$). We note X_n the population size after n minutes.

a) Determine the extinction probability of the population if $X_0 = 1$ and then if $P([X_0 = k]) = 1/2^k$ for any $k \in \mathbb{N}^*$.

b) We consider now that, independently for every particle, a shock occurs after an exponential time with mean 1mn. Determine the extinction probability.

c) Evaluate in every case the mean size of the population after the n-th minute.

Exercise 6 We consider a population of males and females such that every female has one descendent after an exponential time with rate λ : this descendent is a female (resp. a male) with probability p (resp. $1-p$). The lifetimes of the females (resp. males) are exponential with rate μ (resp. ν).
If X_t (resp. Y_t) represents the number of females (resp. males) at time t and if $(X_0, Y_0) = (i, j)$, check that

$$M_X(t) = \mathbb{E}(X_t) = ie^{(\lambda p-\mu)t} \text{ and } M_Y(t) = \mathbb{E}(Y_t) = \frac{i\lambda(1-p)}{\lambda p+\nu-\mu}e^{(\lambda p-\mu)t}+\left(j-\frac{i\lambda(1-p)}{\lambda p+\nu-\mu}\right)e^{-\nu t}.$$

Exercises sheet n°3 : Splitting

An elementary gambler's ruin problem

We consider a simple random walk $X_n = x + \sum_{i=1}^{n} \epsilon_i$ on $E = \mathbb{Z}$, starting at some $x \in \mathbb{Z}$ where $(\epsilon_i)_{i \geqslant 1}$ is a sequence of independent and identically distributed random variables with common law

$$\mathbb{P}(\epsilon_1 = +1) = p \qquad \text{and} \qquad \mathbb{P}(\epsilon_1 = -1) = q$$

with p, $q \in (0,1)$ and $p + q = 1$. If we use the convention $\sum_{\emptyset} = 0$, then we can interpret X_n as the amount of money won or lost by a player starting with $x \in \mathbb{Z}$ euros in a gambling game where he wins and loses 1 euro with respective probabilities p and q. If we let $a < x < b$ be two fixed parameters, one interesting question is to compute the probability that the player will succeed in winning $b - x$ euros, never losing more than $x - a$ euros. More formally this question becomes that of computing the probability that the chain X_n (starting at some $x \in (a, b)$) reaches the set $B = [b, \infty)$ before entering into the set $C = (-\infty, a]$. When $p < q$ (i.e. $p < 1/2$), the random walk X_n tends to move to the left and it becomes less and less likely that X_n will succeed in reaching the desired level B. We further assume that $q > p$. We introduce the stopping time

$$R = \inf\{n \geqslant 0;\ X_n = a\}$$

as well as the first time the chain X_n reaches one of the boundaries

$$T = \inf\{n \geqslant 0;\ X_n \in \{a, b\}\} \leqslant R.$$

Study of $\mathbb{P}_x(R < \infty)$

1. Check that if we have $|x - y| > n$ or $y - x \neq n + 2k$, for some $k \geqslant 1$ then $\mathbb{P}_x(R < \infty) = 0$. The case where $y - x = k - (n-k)$ with $0 \leqslant k \leqslant n$ corresponds to situations where the chain has moved k steps to the right and $n - k$ steps to the left. Prove that $\mathbb{P}_x(X_n = y) = \binom{n}{k} p^k q^{n-k}$.
2. Show that the function α defined by
$$x \in [a, \infty) \mapsto \alpha(x) = \mathbb{P}_x(R < \infty)$$
is the minimal solution of the equation defined for any $x > a$ by $\alpha(x) = p\alpha(x+1) + q\alpha(x-1)$ with the boundary condition $\alpha(a) = 1$.
3. Whenever $p < q$, we recall that the general solution of the equation above has the form $\alpha(x) = A + B(q/p)^x$ with $\alpha(a) = 1 = A + B(q/p)^a$. Deduce from the above that
$$\alpha(x) = 1 + B[(q/p)^x - (q/p)^a] \quad \text{and} \quad \mathbb{P}_x(R < \infty) = 1 \quad \text{for any} \quad x.$$
4. Whenever $p = q$, we recall that the general solution of the equation above has the form $\alpha(x) = Ax + B$ with $\alpha(a) = 1 = A + B(q/p)^a$. Deduce from the above that
$$\alpha(x) = 1 + B[(q/p)^x - (q/p)^a] \quad \text{and} \quad \mathbb{P}_x(R < \infty) = 1 \quad \text{for any} \quad x.$$

Expectation of T

1. Check that for any $n \geqslant 0$ and $\lambda > 0$, we have
$$\mathbb{P}_x(R \geqslant n) = \mathbb{P}_x(X_n \geqslant a) \leqslant e^{-\lambda a}\mathbb{E}_x\left(e^{\lambda X_n}\right) = e^{-\lambda(a-x)}\left(pe^{\lambda} + qe^{-\lambda}\right)^n.$$
2. If we choose $\lambda = \log(q/p)/2 \in (0, \infty)$, then prove that
$$\mathbb{P}_x(R \geqslant n) \leqslant (p/q)^{(x-a)/2}(4pq)^{n/2}.$$
3. Deduce from the above that for $p \neq 1/2$,
$$\mathbb{E}_x(T) \leqslant \mathbb{E}_x(R) = \sum_{n\geqslant 1}\mathbb{P}_x(R \geqslant n) \leqslant \frac{(4pq)^{1/2}}{1-(4pq)^{1/2}}(p/q)^{(x-a)/2}.$$

Study of $\mathbb{P}_x(T < R)$

1. Show that for any $a < x < b0$, the stochastic process $M_n = (q/p)^{X_n}$ is a $\mathbb{P}_x$-martingale with respect to the filtration $F_n = \sigma(X_0, \ldots, X_n)$ and if $p < q$, then $\mathbb{P}_x$-a.s. on the event $\{T \geqslant n\}$, we have that
$$\mathbb{E}_x\left(|M_{n+1} - M_n||F_n\right) \leqslant 2(q/p)^b(q-p).$$
2. Since we have $\mathbb{E}_x(T) < \infty$ and $\mathbb{E}_x\left(|M_{n+1} - M_n||F_n\right)\mathbb{1}_{\{T\geqslant n\}} < c$ for some finite constant, prove by a well-known martingale theorem of Doob that $\mathbb{E}_x(M_T) = \mathbb{E}_x(M_0) = (q/p)^x$ and deduce that for any $x \in [a, b]$
$$(q/p)^x = (q/p)^b\mathbb{P}_x(T < R) + (q/p)^a(1 - \mathbb{P}_x(T < R)).$$
Finally conclude that for any $p \neq q$, we have
$$\mathbb{P}_x(T < R) = \frac{(q/p)^x - (q/p)^a}{(q/p)^b - (q/p)^a}. \tag{3}$$
3. Using the strong Markov property, check that for any p and q, the function $\beta(x) = \mathbb{P}_x(T < R) = \mathbb{E}_x(\mathbb{1}_b(X_T))$ satisfies the equation
$$\beta(x) = p\beta(x+1) + q\beta(x-1)$$
for any $x \in (a, b)$ with the boundary conditions $(\beta(a), \beta(b)) = (0, 1)$.
For $p \neq q$, check that the function (3) is the unique solution and for $p = q = 1/2$, prove that the solution is given for any $x \in [a, b]$ by
$$\mathbb{P}_x(T < R) = (x-a)/(b-a).$$

Splitting algorithm Assume that we want to fix the intermediate thresholds B_n in such a way that the transition probability between two successive thresholds equals θ i.e.
$$\mathbb{P}_{b_n}\left(X_{T_{n+1}} = b_{n+1}\right) = \theta,$$
where $T_n = \inf\{k \geqslant 0;\ X_k \in \{a, b_n\}\}$.

1. Show that the optimal solution is given recursively by
$$b_{n+1} = b_n + \frac{\log\left(1 + (\theta - 1)\left(\frac{p}{q}\right)^{a-b_n}\right) - \log(\theta)}{\log(p/q)}.$$
2. Deduce that as b_n goes to infinity, if $q > p$,
$$b_{n+1} \sim b_n - \frac{\log(\theta)}{\log(p/q)}.$$

Practical on Scilab

1 Illustrative examples

1.1 Crude Monte Carlo

Assume that we want to calculate $E := \mathbb{E}(e^{\beta G})$ where G is a standard Gaussian rv.

1. Compute the exact value of E.
2. Propose an algorithm using the Monte Carlo scheme to evaluate

$$\mathbb{E}\left(e^{\beta G}\right)$$

   with $\beta = 5$ and G a standard Gaussian rv.

3. Determine also a 95%-confidence interval.

1.2 Methods to reduce the variance

We want to evaluate $I = \int_0^1 e^x dx$.
Propose algorithms using

1. crude Monte Carlo method
2. control variables method (see Exercise 4 of the sheet of exercises n˚1)
3. antithetic variables method

to evaluate by several ways I.
For each method, determine also a 95%-confidence interval.

2 An example in finance

Let us study Exercise 1 of the sheet of exercises n˚1. Here we take $\beta = K = 1$.
Determine

1. crude Monte Carlo estimations of C and P ;
2. an estimation of C based on control variables and the first estimation of P ;
3. an estimation of P with IS method.
4. an estimation of P with antithetic variables method.

For each method, determine also a 95%-confidence interval.
Conclude.

3 An example in queuing theory

Let us study a M/M/1 queue. See next Section for some reminders on queuing theory.
Take for example, $\lambda = 0.1$ and $\mu = 0.12$. Determine

1. a crude Monte Carlo estimation of $\mathbb{P}(Q \geqslant L)$, $L = [\![1 : 5 : 150]\!]$;
2. a splitting estimation of $\mathbb{P}(Q \geqslant L)$, $L = [\![1 : 5 : 150]\!]$.

For each method, determine also a 95%-confidence interval.
Conclude.

4 Comparison between IS and Splitting on the simple random walk on $\mathbb{Z}$

In this section, we want to compare numerically on Scilab the IS and Splitting methods in the setting of the simple random walk on $\mathbb{Z}$. The goal is to estimate the probability that the line reaches length b before returning at 0.

4.1 General framework

4.2 Importance Sampling

Following exercise 6 of the sheet of exercises n°1, we define a new random variable to simulate and the corresponding likelihood ratio.

4.3 Splitting

Following the exercise of the sheet of exercises n°3, we define the optimal thresholds and run N simple random walks starting at 0. As soon as a queue reaches the next threshold before returning to 0, it is duplicated in R sub queues that evolve from this threshold and so on. The unbiased estimator of the probability under concern is then given by

$$\widehat{P}_{Splitt} = \frac{1}{N}\sum_{i_0=1}^{N}\frac{1}{R^M}\sum_{i_1=1}^{R}\cdots\sum_{i_M=1}^{R}\mathbb{1}_{i_0}\mathbb{1}_{i_0 i_1}\ldots\mathbb{1}_{i_0 i_1\ldots i_M}$$

where $\mathbb{1}_{i_0 i_1 \ldots i_j}$ represents the result of j-th trial (i.e. it is equals to 1 if the queue reaches B_j, 0 esle).

4.4 Practical on Scilab

Write a program for both algorithms (IS and Splitting) to compare their performances (accuracy estimation, cost...) for the simple random walk.
Check also that crude simulation fails to propose an estimator in that case.

5 Comparison between IS and Splitting on the M/M/1 queue

In this section, we want to compare numerically on Scilab the IS and Splitting methods in queuing theory. The goal is to estimate the probability that the line reaches length L_0 before returning at 0.

5.1 General framework

See [1] or [2] for more details.

A queue is constituted by

a) an **arrival flow** that represents the instants of arrival of "customers". We consider in general that the times between two successive arrivals are iid rvs. Then arrival flow is a stationary renewal process. A simple and commonly used case is the one with exponential inter arrivals ; the process is then a Poisson process.

b) a **service** characterized by
 * a service duration : a customer that starts his service will be immobilized a random duration with known distribution,
 * a number of counter.

c) **service rules** that indicate how the service is proceeding :
 * system with or without line (in a system without line, there is no queue ; a customer that can not be served at his arrival is lost),
 * service order : Fist In First Out (FIFO) (ex : line in the Post office), Last In First Out (LIFO) (ex : print line an the photocopier)
 * several classes of customers clients (definition of priority customers)
 * capacity of the queue
 * at his arrival, if the line is too long, a customer may quit the line with a probability depending on the length of the queue and other parameters. . .
. . .

A queue is characterized by its Kendall notation

$$A/B/C/\ldots$$

A represents the arrival flow, B the service time, C the number of counters. Then we add complementary information like policies. . . We use the following convention :
 * M (like Markov) corresponds to a Poisson flow for the arrivals and to an exponential time service.
 * D (like deterministic) corresponds to constant inter arrival times and to a fix time service for every customer.
 * G (like general) corresponds to general distributions.

5.2 M/M/1 queue

This is the simplest and most studied queue.
 * The arrivals correspond to a Poisson process with rate λ (the inter arrival times

are iid rvs with parameter λ.

* The service time of the customers is exponentially distributed with parameter μ.

* There is a unique counter and the customers are served according to their order of arrival. There is no capacity limitation.

Let N_t be the number of customers in the queue at time t. N_t is an homogenous integer Markov process.

Proposition 5.1 *We have*
(i) $\mathbb{P}_x(N_t = x) = 1 - (\lambda + (x \wedge 1)\mu) + o(t)$ *;*
(ii) the intensity of the process is given by

$$i(x) = \lambda + (x \wedge 1)\mu \qquad \textit{for} \;\; x \geqslant 0;$$

(iii) the transition matrix of the embedded chain is given by

$$\begin{cases} P(x, x+1) = \frac{\lambda}{\lambda + (x \wedge 1)\mu} \\ P(x, x-1) = \frac{(x \wedge 1)\mu}{\lambda + (x \wedge 1)\mu}. \end{cases}$$

The study of the transience of the process N_t amounts to that of the embedded chain. Let $\rho = \frac{\lambda}{\mu}$ the process intensity.

Proposition 5.2 *A positive measure invariant for P is given by*

$$m(x) = m(0)\frac{P(0,1)\ldots P(x-1,x)}{P(1,0)\ldots P(x,x-1)} = m(0)\rho^x \frac{\lambda + (x \wedge 1)\mu}{\lambda}.$$

Here we are interested by the case $\lambda < \mu$.The previous measure is then bounded and we get the existence of an invariant probability π given by

$$\pi(n) = \rho^n (1 - \rho), \quad n \geqslant 0.$$

Proposition 5.3 *The performance parameters are given by*
– the flow (arrival or departure) d is λ ;
– the counter use rate is ρ ;
– the average number L of customers in the system is

$$L = \mathbb{E}_\pi(N_t) = \frac{\rho}{1 - \rho};$$

– the average number L_q of customers in the queue is

$$L_q = \frac{\rho^2}{1 - \rho};$$

– the sojourn time in the system is

$$W = \frac{1}{\mu(1 - \rho)} = \frac{1}{\mu} + \frac{\rho}{\mu(1 - \rho)};$$

– the sojourn time in the queue is

$$W_q = \frac{\rho}{\mu(1 - \rho)}.$$

Proof First, the arrival flow is clearly λ and $d = \lambda$. Second, if a customer enters the system with a queue of length n, its sojourn time T_q in the queue will be null if $n = 0$ and the sum of n iid exponential distributed rvs with parameter μ if $n > 0$. As a consequence,

$$\begin{aligned}
\mathbb{P}(T_q \leqslant t) &= \sum_{n\geqslant 0} \mathbb{P}(T_q \leqslant t \ \text{ and } n \text{ customers in the queue}) \\
&= \sum_{n\geqslant 0} \mathbb{P}(T_q \leqslant t \ \mid \ n \text{ customers in the queue})\, \mathbb{P}(n \ \text{ customers in the queue}) \\
&= \pi(0) + \sum_{n\geqslant 1} \int_0^t \frac{\mu^n x^{n-1}}{n!} e^{-\mu x} dx \rho^n (1-\rho) \\
&= 1 - \rho + \rho(1 - e^{-\mu(1-\rho)t}) = 1 - \rho e^{-\mu(1-\rho)t}
\end{aligned}$$

Then

$$W_q = \mathbb{E}(T_q) = \int_0^{+\infty} \mathbb{P}(T_q \geqslant t) dt = \frac{\rho}{\mu(1-\rho)}.$$

We then use the following relations

$$L = L_q + L_s \quad \text{and} \quad W = W_q + \frac{1}{\mu}$$

and the law of Little applied to the system, to the queue or to the counter

$$L = d\, W, \quad L_q = d\, W_q \quad \text{and} \quad L_s = \frac{1}{\mu}\, d.$$

$\square$

5.3 Importance Sampling

From [3], the optimal change is given by

$$\begin{cases} \lambda^* = \mu \\ \mu^* = \lambda \end{cases}$$

We study N queues starting 1 according to the arrival and service rates λ^* and μ^*. The unbiased estimator of the probability under concern is then given by

$$\widehat{P}_{IS} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{Y_i \geqslant L_0} L(Y_i)$$

Let

$$p_\lambda = \frac{\lambda^*}{\lambda^* + \mu^*} \frac{\lambda + \mu}{\lambda} \quad \text{and} \quad p_\mu = \frac{\mu^*}{\lambda^* + \mu^*} \frac{\lambda + \mu}{\mu}.$$

The likelihood ratio should be updated at each new event by

$$L = \begin{cases} L \times p_\lambda = L \times \frac{\lambda^*}{\lambda^* + \mu^*} \frac{\lambda+\mu}{\lambda} \\ L \times p_\mu = L \times \frac{\mu^*}{\lambda^* + \mu^*} \frac{\lambda+\mu}{\mu}. \end{cases}$$

5.4 Splitting

We define the optimal thresholds and run N queues starting at 1. As soon as a queue reaches the next threshold before returning to 0, it is duplicated in R sub queues that evolve from this threshold and so on. The unbiased estimator of the probability under concern is then given by

$$\widehat{P}_{Splitt} = \frac{1}{N} \sum_{i_0=1}^{N} \frac{1}{R^M} \sum_{i_1=1}^{R} \dots \sum_{i_M=1}^{R} \mathbb{1}_{i_0} \mathbb{1}_{i_0 i_1} \dots \mathbb{1}_{i_0 i_1 \dots i_M}$$

where $\mathbb{1}_{i_0 i_1 \dots i_j}$ represents the result of j-th trial (i.e. it is equals to 1 if the queue reaches B_j, 0 esle).

5.5 Practical on Scilab

Write a program for both algorithms (IS and Splitting) to compare their performances (accuracy estimation, cost...) on the M/M/1 queue.
For example, take $\lambda = 0.4$ and $\mu = 1$.
Check also that crude simulation fails to propose an estimator in that case.

Références

[1] Nicolas Bouleau. *Processus stochastiques et applications*, volume 1420 of *Actualités Scientifiques et Industrielles [Current Scientific and Industrial Topics]*. Hermann, Paris, 1988.

[2] Peter Tjerk de Boer. Analysis and efficient simulation of queueing models of telecommunication systems. 2000.

[3] P.E. Heegaard. Speed-up techniques for simulation. *Telektronikk ISSN 0085-7130*, 91(2-3) :195–207, 1995.