

Université de Lorraine  
Master 2 IMSD

2023-2024

École des Mines de Nancy  
Ingénierie Mathématique, 3A

# Modélisation stochastique

Madalina Deaconu

# Table des matières

<b>Introduction</b>	<b>4</b>
<b>1 Simulation de variables aléatoires (Rappels)</b>	<b>6</b>
1.1 Introduction . . . . .	6
1.2 Simulation de lois classiques . . . . .	7
1.2.1 Loi de Bernoulli de paramètre $p$ . . . . .	7
1.2.2 Loi binomiale de paramètres $(n, p)$ . . . . .	8
1.2.3 Loi de probabilité discrète . . . . .	9
1.2.4 Loi uniforme sur $\{0, 1, \dots, n - 1\}$ . . . . .	10
1.2.5 Loi uniforme sur $[a, b]$ . . . . .	11
1.2.6 Loi exponentielle de paramètre $\lambda$ . . . . .	11
1.2.7 Loi géométrique de paramètre $p$ . . . . .	12
1.2.8 Loi de Poisson de paramètre $\lambda$ . . . . .	13
1.2.9 Loi gaussienne centrée réduite : Méthode de Box-Muller	15
1.3 Méthodes générales . . . . .	16
1.3.1 Méthode générale pour une variable aléatoire discrète .	16
1.3.2 Méthode de simulation par inversion de la fonction de répartition . . . . .	17
1.3.3 Algorithme par rejet . . . . .	17
1.3.4 Simulation par composition . . . . .	20
1.4 Simulation de vecteurs aléatoires . . . . .	21
1.4.1 Cas indépendant . . . . .	21
1.4.2 Vecteur gaussien . . . . .	21
1.4.3 Cas général . . . . .	23
1.5 Exercices . . . . .	24
<b>2 Méthodes de Monte Carlo</b>	<b>30</b>
2.1 Introduction . . . . .	30
2.1.1 Idée de la méthode pour le calcul d'intégrales . . . . .	30
2.2 Description de la méthode de Monte Carlo . . . . .	31
2.3 Convergence de la méthode et vitesse de convergence . . . . .	31

2.3.1	Convergence . . . . .	31
2.3.2	Vitesse de convergence . . . . .	32
2.4	Une application . . . . .	33
2.5	Loi des grands nombres et estimateur . . . . .	34
2.6	Variance, erreurs et intervalle de confiance . . . . .	34
2.6.1	Variance et erreurs . . . . .	34
2.6.2	Intervalle de confiance par TCL . . . . .	35
2.7	Algorithme de Monte Carlo . . . . .	36
2.8	Choix d'une méthode de Monte Carlo . . . . .	37
2.9	Exemples - Calcul approché de $\frac{\pi}{4}$ . . . . .	38
2.10	Techniques de réduction de variance . . . . .	40
2.10.1	Échantillonnage préférentiel (importance sampling) :	
	choix de la densité $f$ . . . . .	40
2.10.2	Échantillonnage stratifié (stratified sampling) . . . . .	43
2.10.3	Variable de contrôle . . . . .	47
2.10.4	Variables antithétiques . . . . .	49
2.11	Exercices . . . . .	51
<b>3</b>	<b>Processus de renouvellement</b>	<b>53</b>
3.1	Définitions . . . . .	53
3.2	Propriétés trajectorielles . . . . .	55
3.3	Processus de Renouvellement avec Récompense (PRR) . . . . .	56
3.4	Théorème Central Limite . . . . .	57
3.5	Fonction de renouvellement . . . . .	59
3.6	Âge courant et âge résiduel . . . . .	60
3.7	Exercices . . . . .	63

# Introduction

L'objectif de ce cours est d'introduire les principaux outils mathématiques pour illustrer, à travers des problèmes réels, comment la modélisation stochastique et les processus aléatoires permettent de traiter et d'apporter des réponses aux questions complexes issues de nombreux domaines applicatifs.

La modélisation probabiliste est fondamentale dans tous les domaines d'application.

Les outils probabilistes que nous développons, comme les méthodes de Monte Carlo, les chaînes de Markov et les processus de renouvellement, sont des outils généraux utilisés dans de nombreux domaines comme en physique (écoulement d'un fluide, trajectoire d'un avion), en biologie (mutation du génome), en chimie (coagulation des polymères), en climatologie (modélisation du vent, de la pluie, etc.), en informatique, en médecine, en finance (évaluation des produits dérivés), en science du vivant, en assurance, en linguistique, en sociologie, ...

Les modèles que nous étudions sont inspirés du monde de la finance, des files d'attente, etc.

Nous débuterons ce cours par des méthodes numériques basées sur les tirages successifs de nombres aléatoires, méthodes appelées généralement méthodes de Monte Carlo.

En outre, nous développerons l'étude de certains processus stochastiques essentiels dans la modélisation, tels que les chaînes de Markov en temps discret et en temps continu, les processus de renouvellement et les processus de branchement. Nous introduirons aussi les outils probabilistes permettant d'étudier la gestion de stock.

L'ensemble de ces études théoriques sera illustré, lors des travaux pratiques, par des algorithmes et simulations en **Matlab**.



# Chapitre 1

## Simulation de variables aléatoires (Rappels)

### 1.1 Introduction

On présentera dans cette partie un panel de méthodes pour la simulation de variables aléatoires. Nous commençons par rappeler les principes de simulation pour les lois classiques. Ensuite nous introduisons les méthodes générales de simulation pour les variables aléatoires à valeurs discrètes et continues.

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité et  $X$  une variable aléatoire réelle sur cet espace. On notera  $F$  sa fonction de répartition et  $f$  sa densité. Plus précisément, pour tout  $x \in \mathbb{R}$  :

$$F(x) = \mathbb{P}(X \leq x)$$

et

$$f(x) = \frac{d}{dx} F(x).$$

**Principe** *Pour simuler des variables aléatoires de loi quelconque, l'idée est de se ramener à la loi uniforme sur  $[0, 1]$ . La loi uniforme sur  $[0, 1]$  est donc à la base de toute simulation de variable aléatoire. Rappelons ses propriétés.*

Soit  $U$  une variable aléatoire de loi uniforme sur  $[0, 1]$ ,  $U \sim \mathcal{U}[0, 1]$ . Dans ce cas :

$$F_U(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

et

$$f_U(x) = \mathbb{1}_{\{x \in [0, 1]\}}.$$

**Cadre général :** On suppose qu'on dispose d'un générateur de variables aléatoires de loi uniforme sur  $[0, 1]$  indépendantes.

L'hypothèse d'indépendance des valeurs est une des conditions essentielles pour la validité de la plupart des algorithmes présentés dans la suite.

**Exemple 1.** SCILAB possède une fonction `rand()` dont les appels successifs fournissent une suite de variables aléatoires “indépendantes” et identiquement distribuées, de loi uniforme sur  $[0, 1]$ . Nous ne nous intéresserons pas ici à la conception d'une telle fonction.

Les générateurs sont souvent construits à partir d'une relation de congruence sur de nombres de grande dimension et initialisés par exemple à partir de l'horloge de la machine. Il s'agit des générateurs de nombres pseudo-aléatoires, notamment les valeurs obtenues ne sont qu'apparemment indépendantes.

**Remarque 1.** Une variable aléatoire  $U$  de loi uniforme sur  $[0, 1]$  est différente de 0 et 1 presque sûrement. En particulier, si  $\Lambda(dx)$  note la mesure de Lebesgue, on a

$$\mathbb{P}(U \in \{0, 1\}) = \int_{\{0,1\}} \mathbb{1}_{[0,1]} d\Lambda = \Lambda(\{0, 1\}) = 0.$$

**Remarque 2.** En SCILAB la fonction `rand` ne renvoie que des valeurs différentes de 0 et 1.

**Objectif du cours :** Vous apprendre à construire des méthodes pour simuler une variable aléatoire ou un vecteur aléatoire suivant une loi donnée, différents de la loi uniforme sur  $[0, 1]$ .

## 1.2 Simulation de lois classiques

### 1.2.1 Loi de Bernoulli de paramètre $p$

Soit  $p \in [0, 1]$ . On veut simuler une variable aléatoire  $X \sim \mathcal{B}(p)$  de loi de probabilité

$$\mu(x) = p\delta_1(x) + (1 - p)\delta_0(x). \quad (1.1)$$

Pour cela on tire  $U$  une uniforme sur  $[0, 1]$  et on définit

$$X = \begin{cases} 1 & \text{si } U \leq p \\ 0 & \text{sinon,} \end{cases}$$

c'est-à-dire  $X = \mathbb{1}_{\{U \leq p\}}$ .

On propose deux fonctions pour simuler cette variable aléatoire. La première ne rend qu'une réalisation d'une variable aléatoire de la loi de Bernoulli de paramètre  $p$ , la seconde renvoie un  $k$ -échantillon, qui sera utilisé par la suite dans la simulation de la loi binomiale :

```
function x=bernoulli1(p)
// tirage suivant la loi de Bernoulli de parametre p
    x=(rand<p);
end

function x=bernoulli2(k,p)
// tirage d'un k-echantillon suivant la loi de Bernoulli de parametre p
    x=bool2s(rand(k,1) < p);
end;
```

**Exemple 2.** Pile ou face

On veut simuler une variable aléatoire de loi “Pile ou face” *i.e.*  $X \sim \mathcal{B}(\frac{1}{2})$ . On tire  $U$  une uniforme sur  $[0, 1]$  et

$$X = \begin{cases} \text{“Pile”} & \text{si } U \leq \frac{1}{2} \\ \text{“Face”} & \text{sinon,} \end{cases}$$

formellement on peut écrire  $X = \mathbb{1}_{\{U \leq \frac{1}{2}\}}$ .

### 1.2.2 Loi binomiale de paramètres $(n, p)$

Soient  $n \in \mathbb{N}^*$  et  $p \in [0, 1]$ . On veut simuler  $X \sim \mathcal{B}(n, p)$  donc de loi

$$\mu(x) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \delta_k(x). \quad (1.2)$$

On sait qu'une variable aléatoire binomiale peut être représentée comme la somme de  $n$  variables aléatoires indépendantes de Bernoulli de paramètre  $p$ . Plus précisément on a le résultat suivant :

**Lemme 1.** Soient  $(Y_i)_{1 \leq i \leq n}$  des variables aléatoires indépendantes et identiquement distribuées de loi de Bernoulli de paramètre  $p$ , alors  $X = \sum_{i=1}^n Y_i$  suit une loi binomiale de paramètres  $(n, p)$ .



Il suffit donc de simuler  $n$  variables aléatoires indépendantes de loi  $\mathcal{B}(p)$  et d'en faire la somme :

$$X = \sum_{i=1}^n \mathbb{1}_{\{U_i \leq p\}}.$$

```
function x=binomiale(n,p)
// tirage suivant la loi binomiale de parametres (n,p)
x=sum(bernoulli2(n,p));
```

### 1.2.3 Loi de probabilité discrète

Soit  $X$  une variable aléatoire à valeurs dans  $E$  avec  $E = \{x_i | i \in I\}$  et  $I = \mathbb{N}$  ou  $I = \{0, 1, \dots, n\}$ . Considérons le cas  $I$  fini, le second se traitant de la même manière. La loi de probabilité de  $X$  est donnée par

$$\mu(x) = \sum_{i=0}^n p_i \delta_{x_i}(x)$$

où

$$p_i = \mathbb{P}(X = x_i) = \mu(x_i), \forall i \in I \text{ et } \sum_{i=0}^n p_i = 1. \quad (1.3)$$

Notons par  $P_k$  le cumul des  $p_i$ , pour  $0 \leq i \leq k$ , *i.e.*

$$P_k = \sum_{i=0}^k p_i. \quad (1.4)$$

On a alors  $P_0 = p_0$  et  $P_n = 1$ .

Nous souhaitons simuler une variable aléatoire de même loi que  $X$ . Pour ce faire construire une variable aléatoire  $Y$  telle que

$$\mathbb{P}(X = x_i) = \mathbb{P}(Y = x_i) = p_i, \quad \forall i \in I,$$

à l'aide d'une variable aléatoire  $U$  de loi uniforme sur  $[0, 1]$ .

**Algorithme** L'algorithme s'écrit : on tire  $U$  une uniforme sur  $[0, 1]$  et on pose

$$Y = k \text{ si } P_{k-1} < U \leq P_k$$

ce qui s'exprime également sous la forme

$$Y = x_0 \mathbb{1}_{\{U \leq P_0\}} + \sum_{i=1}^n x_i \mathbb{1}_{\{P_{i-1} < U \leq P_i\}}.$$

**Remarque 3.** Nous introduisons cette méthode générale pour l'appliquer dans la suite pour le cas des variables aléatoires discrètes particulières. Nous reviendrons avec des commentaires et justifications dans la partie méthodes générales.

### 1.2.4 Loi uniforme sur $\{0, 1, \dots, n-1\}$

On veut simuler  $X$  de loi uniforme sur  $\{0, 1, \dots, n-1\}$ , donc de loi

$$\mu(x) = \frac{1}{n} \sum_{k=0}^{n-1} \delta_k(x). \quad (1.5)$$

**Méthode 1 :** Nous allons utiliser le résultat suivant :

**Lemme 2.** Si  $U$  est une variable aléatoire de loi uniforme sur  $[0, 1]$  alors la partie entière  $[nU]$  suit la loi uniforme sur  $\{0, 1, \dots, n-1\}$ .

**Remarque 4.** En admettant ce résultat, l'algorithme s'écrit :

```
function x=uniforme1(n)
// tirage suivant la loi uniforme sur {0,...,n-1}
x=floor(n*rand());
end
```

**Démonstration :** On note  $X$  la variable aléatoire rendue par cette fonction. Comme  $\mathbb{P}(0 \leq \text{rand}() < 1) = 1$ , on a  $\mathbb{P}(0 \leq n*\text{rand}() < n) = 1$  et

$$\mathbb{P}(\text{floor}(n*\text{rand}()) \in \{0, 1, \dots, n-1\}) = 1.$$

Donc  $X$  prend ses valeurs dans  $\{0, 1, \dots, n-1\}$ .

Maintenant, soit  $k \in \{0, 1, \dots, n-1\}$  :

$$\begin{aligned} \mathbb{P}(X = k) &= \mathbb{P}(\text{floor}(n*\text{rand}()) = k) = \mathbb{P}(k \leq n*\text{rand}() < k+1) \\ &= \mathbb{P}\left(\frac{k}{n} \leq \text{rand}() < \frac{k+1}{n}\right) = \frac{1}{n}. \end{aligned}$$

Ce qui prouve le résultat souhaité. ■

**Méthode 2 :** En utilisant la méthode générale de simulation d'une variable aléatoire discrète.

### 1.2.5 Loi uniforme sur $[a, b]$

On souhaite simuler  $X$  une variable aléatoire de loi uniforme sur l'intervalle  $[a, b]$  pour  $a < b$ , avec

$$f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x). \quad (1.6)$$

**Lemme 3.** Si  $U$  suit la loi uniforme sur  $[0, 1]$ , alors, si  $a < b$  et  $a, b \in \mathbb{R}$ , la variable aléatoire  $X = a + (b-a)U$  suit la loi uniforme sur  $[a, b]$ .

**Démonstration :** Soit  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  une fonction continue bornée : en faisant le changement de variable  $x = a + (b-a)u$ , on a

$$\mathbb{E}(\varphi(a + (b-a)U)) = \int_0^1 \varphi(a + (b-a)u) du = \int_a^b \varphi(x) \frac{1}{b-a} dx,$$

ce qui signifie que  $X$  est une variable aléatoire de densité  $f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$ , et on reconnaît ainsi la densité de la loi uniforme sur  $[a, b]$ . ■

L'algorithme associé s'écrit alors :

```
function x=uniforme2(a,b)
// tirage suivant la loi uniforme sur [a,b]
x=a+(b-a)*rand();
end
```

### 1.2.6 Loi exponentielle de paramètre $\lambda$

Soit  $\lambda > 0$ . On souhaite simuler une variable aléatoire  $X$  de loi exponentielle de paramètre  $\lambda$ , avec

$$f(x) = \lambda \exp(-\lambda x) \mathbb{1}_{\mathbb{R}_+}(x). \quad (1.7)$$

Sa fonction de répartition est  $F(x) = \mathbb{P}(X \leq x) = 1 - \exp(-\lambda x)$ . Cette fonction est une bijection de  $]0, +\infty[$  dans  $]0, 1[$ , d'inverse

$$G(u) = -\frac{1}{\lambda} \ln(1-u). \quad (1.8)$$

**Lemme 4.** Si  $U$  est de loi uniforme sur  $[0, 1]$ , alors  $G(U)$  suit la loi exponentielle de paramètre  $\lambda$ .

En utilisant ce lemme on déduit l'algorithme suivant pour la simulation de la loi exponentielle :

```

function x=exponentielle(a)
// tirage suivant la loi exponentielle de parametre a
  x=- log(rand())/a;
end

```

Ce procédé annonce la méthode de simulation par inversion de la fonction de répartition, que nous présenterons plus tard.

**Remarque 5.** *Pourquoi a-t-on remplacé  $1 - U$  par  $U$  dans l'algorithme ?*

### 1.2.7 Loi géométrique de paramètre $p$

Soit  $p \in ]0, 1]$ . On veut simuler une variable aléatoire géométrique de paramètre  $p$ , donc de loi

$$\mu(x) = \sum_{k=1}^{+\infty} (1-p)^{k-1} p \delta_k(x). \quad (1.9)$$

Plusieurs méthodes sont possibles.

**Méthode 1 :** En utilisant la loi exponentielle :

**Lemme 5.** *Si  $U$  suit la loi uniforme sur  $[0, 1]$ , alors  $X = 1 + \left\lceil \frac{\ln U}{\ln(1-p)} \right\rceil$  est une variable aléatoire qui suit la loi géométrique de paramètre  $p$ .*

**Démonstration :** Notons  $X = 1 + \left\lceil \frac{\ln U}{\ln(1-p)} \right\rceil$ . Par définition de la partie entière,  $X$  prend ses valeurs dans  $\mathbb{N}^*$ . Soit maintenant  $k \in \mathbb{N}^*$  :

$$\begin{aligned}
 \mathbb{P}(X = k) &= \mathbb{P}\left(1 + \left\lceil \frac{\ln U}{\ln(1-p)} \right\rceil = k\right) \\
 &= \mathbb{P}\left(\left\lceil \frac{\ln U}{\ln(1-p)} \right\rceil = k - 1\right) \\
 &= \mathbb{P}\left(k - 1 \leq \frac{\ln U}{\ln(1-p)} < k\right) \\
 &= \mathbb{P}(k \ln(1-p) < \ln U \leq (k-1) \ln(1-p)) \\
 &= \mathbb{P}((1-p)^k < U \leq (1-p)^{k-1}) \\
 &= (1-p)^{k-1} - (1-p)^k \\
 &= (1-p)^{k-1} p.
 \end{aligned}$$

■

L'algorithme correspondant s'écrit :

```
function x=geometrique1(p)
// tirage suivant la loi geometrique de parametre p
x=1+floor(log(rand())/log(1-p));
```

**Méthode 2 :** En utilisant la loi de Bernoulli :

**Lemme 6.** *Si  $(X_i)_{i \in \mathbb{N}^*}$  sont des variables aléatoires i.i.d. de loi de Bernoulli de paramètre  $p$ , alors  $N = \min\{i : X_i = 1\}$  est une variable aléatoire qui suit une loi géométrique de paramètre  $p$ .*

L'algorithme correspondant s'écrit.

```
function x=geometrique2(p)
// tirage suivant la loi geometrique de parametre p
x=1;
while rand()>p, x=x+1;
end
```

**Méthode 3 :** En utilisant la méthode générale pour la simulation des variables aléatoires discrètes.

### 1.2.8 Loi de Poisson de paramètre $\lambda$

Soit  $\lambda \in \mathbb{R}_+$ . On veut simuler une variable aléatoire  $X \sim \mathcal{P}(\lambda)$ , donc de loi donnée par

$$\mu(x) = \sum_{k=0}^{+\infty} \exp(-\lambda) \frac{\lambda^k}{k!} \delta_k(x). \quad (1.10)$$

Une variable aléatoire poissonnienne ne prend pas ses valeurs dans un ensemble fini, mais on peut étendre la méthode présentée au paragraphe 1.2.3 au cas où  $X$  prend ses valeurs dans  $\mathbb{N}$ . En fait, la méthode proposée ici annonce la méthode de la fonction inverse.

**Méthode 1 :** Cumul des durées exponentielles.

Une première méthode pour la simulation de variables aléatoires de Poisson est issue d'une des propriétés des processus de Poisson, il s'agit du cumul de durées exponentielles.

On sait que si des événements surviennent à des dates séparées par des durées exponentielles de paramètre  $\lambda$ , le nombre d'événements survenant en une unité de temps suit une loi de Poisson de même paramètre.

**Lemme 7.** Soient  $(X_i)_{i \in \mathbb{N}^*}$  des variables aléatoires i.i.d. exponentielles de paramètre  $\lambda$ . La variable aléatoire définie par

$$N = 0 \text{ si } X_1 \geq 1 \text{ et } N = \max\{i \geq 1 : \sum_{j=1}^i X_j \leq 1\} \text{ sinon,}$$

suit une loi de Poisson de paramètre  $\lambda$ .

**Démonstration :** Soit  $k \in \mathbb{N}$ .

$$\begin{aligned} \mathbb{P}(N = k) &= \mathbb{P}\left(\sum_{j=1}^k X_j \leq 1 < \sum_{j=1}^{k+1} X_j\right) \\ &= \int_{\mathbb{R}_+^{k+1}} \lambda^{k+1} \exp(-\lambda(x_1 + \dots + x_{k+1})) \mathbb{1}_{\{x_1 + \dots + x_k \leq 1\}} \\ &\quad \times \mathbb{1}_{\{x_{k+1} > 1 - (x_1 + \dots + x_k)\}} dx_1 \dots dx_{k+1} \\ &= \int_{\mathbb{R}_+^k} \lambda^k \exp(-\lambda(x_1 + \dots + x_k)) \mathbb{1}_{\{x_1 + \dots + x_k \leq 1\}} \\ &\quad \times \left( \int_{1 - (x_1 + \dots + x_k)}^{+\infty} \lambda \exp(-\lambda x_{k+1}) dx_{k+1} \right) dx_1 \dots dx_k \\ &= \int_{\mathbb{R}_+^k} \lambda^k \exp(-\lambda(x_1 + \dots + x_k)) \mathbb{1}_{\{x_1 + \dots + x_k \leq 1\}} \exp(-\lambda) \\ &\quad \times \exp(\lambda(x_1 + \dots + x_k)) dx_1 \dots dx_k \\ &= \lambda^k \exp(-\lambda) \int_{\mathbb{R}_+^k} \mathbb{1}_{\{x_1 + \dots + x_k \leq 1\}} dx_1 \dots dx_k. \end{aligned}$$

Pour calculer la dernière intégrale, on effectue le changement de variable  $s_1 = x_1$ ,  $s_2 = x_1 + x_2$ ,  $\dots$ ,  $s_k = x_1 + \dots + x_k$  :

$$\begin{aligned} \int_{\mathbb{R}_+^k} \mathbb{1}_{\{x_1 + \dots + x_k \leq 1\}} dx_1 \dots dx_k &= \int_{\mathbb{R}_+^k} \mathbb{1}_{\{s_1 \leq s_2 \leq \dots \leq s_k \leq 1\}} ds_1 \dots ds_k \\ &= \int_0^1 ds_k \int_0^{s_k} ds_{k-1} \dots \int_0^{s_2} ds_1 = \frac{1}{k!}, \end{aligned}$$

ce qui termine la preuve. ■

**Remarque 6.** Il faut donc simuler des variables aléatoires exponentielles de paramètre  $\lambda$  et compter le nombre de simulations nécessaires pour dépasser 1, ou bien simuler des variables aléatoires exponentielles de paramètre 1 et compter le nombre de simulations nécessaires pour dépasser  $\lambda$ .

**Remarque 7.** Soient  $(U_i)_{i \in \mathbb{N}^*}$  des variables aléatoires i.i.d. de loi uniforme sur  $[0, 1]$ , alors si  $X_i = -\frac{1}{\lambda} \ln(U_i)$ , les  $(X_i)_{i \in \mathbb{N}^*}$  sont des variables aléatoires i.i.d. exponentielles de paramètre  $\lambda$ , et

$$\sum_{j=1}^i X_j \leq 1 \Leftrightarrow -\frac{1}{\lambda} \ln \left( \prod_{j=1}^i U_j \right) \leq 1 \Leftrightarrow \prod_{j=1}^i U_j \geq \exp(-\lambda).$$

L'algorithme s'écrit :

```
function x=poisson(a)
// tirage suivant la loi de Poisson de parametre a
test=exp(-a);x=0;prod=rand();
while (prod>=test), x=x+1; prod=prod*rand(); end;
```

**Méthode 2 :** En utilisant la méthode générale pour les variables aléatoires discrètes.

On sait que

$$X \sim \mathcal{P}(\lambda) \Leftrightarrow p_k = \mathbb{P}\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda} \quad (\text{pour } k \in \mathbb{N})$$

ce qui implique que

$$p_{k+1} = \frac{\lambda}{k+1} p_k,$$

et en notant  $P_k$  la somme des  $p_j$  pour  $0 \leq j \leq k$ , ( $P_k = \sum_{j=0}^k p_j$ ), on a

$$P_{k+1} = P_k + \frac{\lambda}{k+1} p_k.$$

On simule donc une variable de Poisson de paramètre  $\lambda$  en prenant

$$X = \sum_{k \geq 0} k \mathbb{1}_{\{P_{k-1} < U \leq P_k\}}$$

avec la convention  $P_{-1} = 0$ .

Cet algorithme est assez simple à programmer.

### 1.2.9 Loi gaussienne centrée réduite : Méthode de Box-Muller

La loi normale n'a pas une densité à support compact et on ne connaît pas d'expression simple de l'inverse de sa fonction de répartition. Nous utiliserons le résultat suivant pour sa simulation :

**Lemme 8.** Soit  $R$  une variable aléatoire de loi exponentielle de paramètre  $1/2$  et  $\theta$  une variable aléatoire de loi uniforme sur  $[0, 2\pi]$ , supposées de plus indépendantes. Alors, si on pose  $X = \sqrt{R} \cos(\theta)$  et  $Y = \sqrt{R} \sin(\theta)$ , les variables aléatoires  $X$  et  $Y$  sont i.i.d. de loi gaussienne centrée réduite.

Ceci conduit à une méthode de simulation simple pour une loi gaussienne, basée sur la simulation de deux variables aléatoires uniformes sur  $[0, 1]$ . C'est la méthode de Box-Muller.

```
function [x,y]=boxmuller
// tirage de deux N(0,1) independantes
r=sqrt(-2*log(rand())); t=2*%pi*rand();
x=r*cos(t); y=r*sin(t);
end
```

## 1.3 Méthodes générales

### 1.3.1 Méthode générale pour une variable aléatoire discrète

Soit  $X$  une variable aléatoire sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ , telle que  $X(\Omega) = \{x_0, x_1, \dots\} = \{x_i\}_{i \in I}$  avec  $I = \mathbb{N}$  ou  $I = \{0, 1, \dots, n\}$ . Pour tout  $i \in I$ ,  $p_i = \mathbb{P}(X = x_i)$ . On rappelle que  $p_i \geq 0$  et  $\sum_{i \in I} p_i = 1$ .

Rappelons les idées générales de la section 1.2.3. L'algorithme suivant simule une variable représentative pour cette loi :

```
function y=simuldiscrete(x,p)
// simulation d'une va de loi discrete,
// x=vecteur des valeurs prises, p=vecteur des probabilités
u=rand(); q=p(1); i=0;
while (u>q); i=i+1; q=q+p(i); end;
y=x(i);
end
```

**Démonstration :** En effet, on a :  $\mathbb{P}(X = x_0) = \mathbb{P}(\text{rand}() \leq p_0) = p_0$  et pour  $k \geq 1$ ,

$$\mathbb{P}(X = x_k) = \mathbb{P}\left(\sum_{i=0}^{k-1} p_i < \text{rand}() \leq \sum_{i=0}^k p_i\right) = p_k.$$

■

**Remarque 8.** Le nombre  $N$  de tests nécessaires satisfait  $N = 1$  ssi  $u \leq p_0$ , et pour  $i > 1$ ,

$$N = i \Leftrightarrow \sum_{j=0}^{i-1} p_j < u \leq \sum_{j=0}^i p_j.$$

On a donc intérêt à réordonner les  $(x_i)_{i \geq 0}$  dans l'ordre des  $(p_i)_{i \geq 0}$  décroissants.



### 1.3.2 Méthode de simulation par inversion de la fonction de répartition

On veut simuler une variable aléatoire  $X$  de fonction de répartition  $F$ .

La méthode utilisée pour simuler une loi exponentielle est en fait générale : dès que l'on sait inverser une fonction de répartition  $F$ , il est très facile de simuler une variable aléatoire de fonction de répartition  $F$ .

**Lemme 9.** *Soit  $U$  une variable aléatoire qui suit la loi uniforme sur  $[0, 1]$ , et  $F$  une fonction de répartition bijective de  $]a, b[$  dans  $]0, 1[$  d'inverse  $F^{-1}$ . Alors  $F^{-1}(U)$  est une variable aléatoire de fonction de répartition  $F$ .*

**Démonstration :** On pose  $X = F^{-1}(U)$ . Cette variable aléatoire prend ses valeurs dans  $]a, b[$ . Remarquons que nécessairement  $F$  est strictement croissante de  $]a, b[$  dans  $]0, 1[$ . Soit  $t \in ]a, b[$  :

$$\mathbb{P}(X \leq t) = \mathbb{P}(F^{-1}(U) \leq t) = \mathbb{P}(U \leq F(t)) = F(t).$$

Donc la fonction de répartition de  $X$  est bien  $F$ . ■

**Remarque 9.** *L'hypothèse de la connaissance de  $F^{-1}$  n'a de sens que si  $F$  est strictement croissante. Cependant, même dans ce cas, il se peut que  $F^{-1}$  n'ait pas d'expression analytique simple, c'est le cas par exemple pour la loi normale.*

Si on note  $\phi(x)$  la fonction de répartition de la loi normale centrée réduite,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt,$$

il n'existe pas de formulation simple de  $\phi(x)$  et encore moins de  $\phi^{-1}(x)$ , la méthode de la fonction inverse ne peut donc pas s'appliquer directement à la loi normale.

Il existe cependant des polynômes donnant de bonnes approximations de  $\phi(x)$  et de  $\phi^{-1}(x)$  qui permettent donc d'appliquer la méthode de la fonction inverse à la loi normale moyennant cette approximation.

### 1.3.3 Algorithme par rejet

On veut simuler une variable aléatoire  $X$  de densité  $f$  et de fonction de répartition  $F$ .

**Exemple 3.** Commençons par un exemple très simple : comment simuler une loi uniforme sur le disque unité  $\{x^2 + y^2 \leq 1\}$  ?

```

function X=disque
// simule un point uniformement sur le disque unite
    X=2*rand(1,2)-[1,1];
while (norm(X)>1),
    X=2*rand(1,2)-[1,1];
end

```

L'idée est la suivante : on tire des points uniformément dans le carré  $[0, 1] \times [0, 1]$ , et on les jette jusqu'à en obtenir un qui tombe dans le disque. La loi du point obtenue est la loi d'un point tiré uniformément dans le carré conditionnellement à être dans le disque, ce qui est encore la loi uniforme sur le disque.

**Remarque 10.** *Quelle est la loi du nombre  $N$  de passages dans la boucle ?*

**Lemme 10.** *Soit  $X$  une variable aléatoire de densité  $f$  (sur  $\mathbb{R}^d$ ) à simuler. On suppose qu'il existe une constante  $k > 0$  et une densité  $g$  (sur  $\mathbb{R}^d$  aussi, facile à simuler) tels que*

$$\forall x, \quad f(x) \leq kg(x).$$

*Soit  $U$  une variable aléatoire de loi uniforme sur  $[0, 1]$  et  $Z$  une variable aléatoire, indépendante de  $U$ , de densité  $g$ . On pose  $V = kUg(Z)$ . Alors, la loi de  $Z$  conditionnellement à l'événement  $\{V < f(Z)\}$  a pour densité  $f$ .*

**Remarque 11.** *Notons que nécessairement  $k \geq 1$  (car  $f, g$  sont des densités).*

**Remarque 12.** *Il est très important de noter qu'on doit choisir la constante  $k$  la plus petite possible pour minimiser le nombre de rejets : plus la majoration est grossière, plus il faut des tirages pour obtenir une valeur acceptable.*

**Démonstration :** On fait la démonstration pour le cas  $d = 1$ . Notons que pour tout  $z \in \mathbb{R}$ ,  $\frac{f(z)}{kg(z)} \leq 1$ . On a tout d'abord

$$\begin{aligned}
 \mathbb{P}(V < f(Z)) &= \mathbb{P}(kUg(Z) < f(Z)) \\
 &= \int_{\mathbb{R}} g(z) \left( \int_0^1 \mathbb{1}_{\{kug(z) < f(z)\}} du \right) dz \\
 &= \int_{\mathbb{R}} g(z) \frac{f(z)}{kg(z)} dz \\
 &= \frac{1}{k}.
 \end{aligned}$$

Évaluons ensuite

$$\begin{aligned}\mathbb{P}(\{Z \leq t\} \cap \{V < f(Z)\}) &= \int_{-\infty}^t g(z) \left( \int_0^1 \mathbb{1}_{\{kug(z) < f(z)\}} du \right) dz \\ &= \int_{-\infty}^t g(z) \frac{f(z)}{kg(z)} dz \\ &= \frac{1}{k} \int_{-\infty}^t f(z) dz.\end{aligned}$$

Ce qui montre que  $\mathbb{P}(Z \leq t | V < f(Z)) = \int_{-\infty}^t f(z) dz$ , donc la loi conditionnelle de  $Z$  sachant que  $\{V < f(Z)\}$  a bien pour densité  $f$ .

Dans  $\mathbb{R}^d$ , on utilise la généralisation de la fonction de répartition. ■

On obtient donc l'algorithme de simulation par rejet (on suppose qu'on possède une fonction `simulg` qui simule une variable aléatoire de densité  $g$ ) :

```
function z=simulf
// simule par rejet une va de densite f
u=rand();
z=simulg;
v=k*u*g(z);
while (v>=f(z));
    u=rand();
    z=simulg;
    v=k*u*g(z);
end;
```

**Démonstration :** Notons  $N$  le nombre de tests fait lors de cette fonction.  $N$  est une variable aléatoire, à valeurs dans  $\mathbb{N}^*$ . Notons  $(U_n)_{n \geq 1}$  la suite des appels à la fonction `rand()`, et  $(Z_n)_{n \geq 1}$  la suite des appels à la fonction `simulg`. Toutes ces variables aléatoires sont indépendantes, les premières de loi uniforme sur  $[0, 1]$ , les secondes de densité  $g$ . On note  $V_n = kU_n g(Z_n)$ , et on note  $X$  la sortie de la fonction.

Soit  $t \in \mathbb{R}$ . Evaluons

$$\mathbb{P}(X \leq t \text{ et } N = 1) = \mathbb{P}(V_1 < f(Z_1) \text{ et } Z_1 \leq t) = \frac{1}{k} \int_{-\infty}^t f(z) dz$$

par la démonstration précédente. Soit maintenant  $i \geq 2$ . Par indépendance,

et comme précédemment :

$$\begin{aligned}
\mathbb{P}(X \leq t \text{ et } N = i) &= \mathbb{P}(V_1 \geq f(Z_1), V_2 \geq f(Z_2), \dots, V_{i-1} \geq f(Z_{i-1}), V_i < f(Z_i), Z_i \leq t) \\
&= \mathbb{P}(V_1 \geq f(Z_1)) \mathbb{P}(V_2 \geq f(Z_2)) \dots \mathbb{P}(V_{i-1} \geq f(Z_{i-1})) \mathbb{P}(V_i < f(Z_i), Z_i \leq t) \\
&= \left(1 - \frac{1}{k}\right)^{i-1} \frac{1}{k} \int_{-\infty}^t f(z) dz.
\end{aligned}$$

Finalement (rappelons que  $k \geq 1$ ) :

$$\begin{aligned}
\mathbb{P}(X \leq t) &= \sum_{i=1}^{+\infty} \mathbb{P}(X \leq t \text{ et } N = i) \\
&= \sum_{i=1}^{+\infty} \left(1 - \frac{1}{k}\right)^{i-1} \frac{1}{k} \int_{-\infty}^t f(z) dz \\
&= \int_{-\infty}^t f(z) dz,
\end{aligned}$$

donc la densité de  $X$  est bien  $f$ . ■

### 1.3.4 Simulation par composition

**Exemple 4.** Soit  $F$  et  $G$  deux fonctions de répartition sur  $\mathbb{R}$ . On construit une variable aléatoire  $X$  de la façon suivante : on lance une pièce qui tombe sur pile avec probabilité  $1/3$  et sur face avec probabilité  $2/3$ , si pile sort, on tire un nombre au hasard suivant la loi donnée par  $F$ , sinon, on tire un nombre au hasard suivant la loi donnée par  $G$ . Déterminer la fonction de répartition  $H$  de  $X$ .

L'exemple précédent est un exemple de mélange de variables aléatoires. On suppose maintenant qu'on veut simuler une variable aléatoire  $X$  de fonction de répartition  $F = \sum_{i=1}^n \theta_i F_i$ , où les  $\theta_i$  sont des poids :  $\theta_i \geq 0$  et  $\sum_{i=1}^n \theta_i = 1$  et les  $F_i$  sont des fonctions de répartition dont les lois sont faciles à simuler. On suppose qu'on a à notre disposition des fonctions `simulFi` qui simulent des variables aléatoires de fonction de répartition  $F_i$  et une fonction `simulTheta` qui simule une variable aléatoire  $\Theta$  à valeur dans  $\{1, \dots, n\}$  telle que  $\mathbb{P}(\Theta = i) = \theta_i$ .

```
function x = melange
```

```
// simulation par melange
i=simulTheta;
x=simulFi;
end
```

Cette méthode se généralise immédiatement à un nombre infini dénombrable de poids  $(\theta_i)_{i \in \mathbb{N}}$ , et même à un mélange "continu" de lois : on suppose qu'on veut simuler une variable aléatoire  $X$  de densité  $f(z) = \int_{\theta} g(\theta) f_{\theta}(z) d\theta$ , où  $g$  est une densité de probabilité, ainsi que tous les  $f_{\theta}$ . L'algorithme est alors simple : on tire  $\theta$  suivant  $g$ , puis  $x$  suivant  $f_{\theta}$ .

## 1.4 Simulation de vecteurs aléatoires

### 1.4.1 Cas indépendant

Supposons qu'on souhaite simuler  $Z$  un vecteur aléatoire de  $\mathbb{R}^d$ . Si ses composantes sont indépendantes, on est ramené au cas de la simulation de variables aléatoires réelles indépendantes traité dans les sections précédentes (on rappelle que les sorties successives de **rand** donnent des variables aléatoires *indépendantes* de loi uniforme sur  $[0, 1]$ ). Le problème est différent quand les coordonnées ne sont pas indépendantes.

### 1.4.2 Vecteur gaussien

Un vecteur gaussien dans  $\mathbb{R}^d$  est caractérisé par un vecteur moyenne  $m \in \mathbb{R}^d$ , et une matrice de covariance  $K$ , de taille  $d \times d$ , symétrique et positive. On suppose pour l'instant que  $K$  est *définie positive*.

**Théorème 1** (Décomposition de Cholesky). *Si  $K$  est une matrice symétrique définie positive de taille  $d \times d$ , il existe au moins une matrice réelle triangulaire inférieure  $L$  telle que :*

$$K = L^t L.$$

*On peut également imposer que les éléments diagonaux de la matrice  $L$  soient tous strictement positifs, et la factorisation correspondante est alors unique.*

**En pratique** on cherche  $L$  par coefficients indéterminés et identification, colonne par colonne. Voir l'exemple juste après.

**Remarque 13.** *Si  $K$  est seulement semi-définie positive, la décomposition de Cholesky existe encore mais elle n'est plus unique.*

**Lemme 11.** Soit  $T$  un vecteur gaussien de dimension  $d$ , centré et de matrice de covariance  $I_d$  (dont les coordonnées sont des variables aléatoires *i.i.d.* normales centrées réduites). Soit  $m \in \mathbb{R}^d$ , et  $K$  une matrice définie positive de taille  $d \times d$ . Soit  $L$  la matrice triangulaire inférieure donnée par la factorisation de Cholesky de  $K$ .

Alors le vecteur aléatoire  $Z = m + LT$  est un vecteur gaussien de moyenne  $m$  et de matrice de covariance  $K$ .

**Démonstration :** Comme  $L$  est une application linéaire de  $\mathbb{R}^d$  dans  $\mathbb{R}^d$ ,  $Z$  est encore un vecteur gaussien  $d$ -dimensionnel. Pour l'espérance :

$$\mathbb{E}(Z) = \mathbb{E}(m + LT) = m + L\mathbb{E}(T) = m,$$

puisque  $T$  est centré. Pour la matrice de covariance, comme celle de  $T$  est  $I_d$ ,

$$\mathbb{E}((Z - m)^t(Z - m)) = \mathbb{E}(LT^tT^tL) = L\mathbb{E}(T^tT)^tL = LI_d^tL = L^tL = K.$$

■

**Exemple 5.** On veut simuler le vecteur gaussien  $Z$  de  $\mathbb{R}^3$  de moyenne  $m$  et de matrice de covariance  $K$  avec

$$m = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix} \text{ et } K = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 5 & 6 \\ 0 & 6 & 10 \end{pmatrix}.$$

On commence par chercher la matrice de factorisation de Cholesky  $L$  par coefficients indéterminés et identification :

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}.$$

On calcule  $LL^t$  et on identifie, colonne par colonne :

1.  $l_{11}^2 = k_{11} = 1 \Rightarrow l_{11} = 1.$
2.  $l_{11}l_{21} = k_{21} = -1 \Rightarrow l_{21} = -1.$
3.  $l_{11}l_{31} = k_{31} = 0 \Rightarrow l_{31} = 0.$
4.  $l_{21}^2 + l_{22}^2 = k_{22} = 5 \Rightarrow l_{22} = 2.$
5.  $l_{21}l_{31} + l_{22}l_{32} = k_{32} = 6 \Rightarrow l_{32} = 3.$
6.  $l_{31}^2 + l_{32}^2 + l_{33}^2 = k_{33} = 10 \Rightarrow l_{33} = 1.$

Donc

$$Z = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & 3 & 1 \end{pmatrix} T,$$

où  $T$  est un vecteur gaussien de  $\mathbb{R}^3$  dont les trois composantes sont *i.i.d.* centrées réduites.

### 1.4.3 Cas général

1. Le cas d'une *v.a.* discrète à valeurs dans  $\mathbb{R}^d$  se traite comme celui d'une variable aléatoire réelle discrète.
2. La méthode de rejet a été présentée dans la cas d'une variable aléatoire à valeurs dans  $\mathbb{R}^d$ .
3. On peut encore utiliser les lois conditionnelles. Nous allons illustrer cette méthode, dite **méthode récurrente**, par un exemple :

**Exemple 6.** On veut simuler un vecteur  $(X, Y)$  de loi uniforme sur le triangle  $ABC$  avec  $A = (0, 0)$ ,  $B = (1, 1)$  et  $C = (0, 1)$ .

On commence par déterminer la densité de cette loi :

$$f(x, y) = 2\mathbb{1}_{0 \leq x \leq 1} \mathbb{1}_{0 \leq y \leq 1} \mathbb{1}_{x \leq y}.$$

On peut alors calculer la densité de  $X$  :

$$f_X(x) = 2(1 - x)\mathbb{1}_{0 \leq x \leq 1},$$

puis la densité de la loi conditionnelle de  $Y$  sachant  $X$  :

$$f_{Y|X}(y|x) = \frac{1}{1-x} \mathbb{1}_{x \leq y \leq 1}.$$

Pour la simulation, on procède maintenant de la façon suivante : on simule  $X$  suivant sa densité (en utilisant par exemple la méthode de la fonction de répartition), on obtient une valeur  $x$ , puis on simule  $Y$  suivant la densité  $f_{Y|X}(y|x) = \frac{1}{1-x} \mathbb{1}_{x \leq y \leq 1}$  (on reconnaît par exemple une loi usuelle).

**Remarque 14.** *Remarquons que la seconde étape ressemble beaucoup à la simulation par mélange.*

## 1.5 Exercices

**Exercice 1.** Soit  $X$  une v.a. de loi de Bernoulli de paramètre  $p$ .

1. Proposer un algorithme qui construit une réalisation de  $X$ .
2. Proposer un algorithme qui renvoie un  $k$ -échantillon de cette v.a.
3. Tester les codes en **Matlab** ou **Scilab** pour plusieurs valeurs de  $p$  et  $k$ .

**Exercice 2.** Simuler une variable aléatoire  $X$  de loi binomiale de paramètres  $(n, p)$  pour différentes valeurs de  $n$  et  $p$ .

**Exercice 3.** Soit  $X$  une variable aléatoire de loi uniforme sur  $\{1, \dots, n\}$ . Proposer un algorithme en **Matlab** ou **Scilab** pour la simulation de cette loi :

1. En utilisant la simulation d'une loi de probabilité discrète sur un ensemble fini.
2. En utilisant le résultat suivant : Si  $U$  est une variable aléatoire de loi uniforme sur  $[0, 1]$  alors  $[nU] + 1$  est une variable aléatoire de loi uniforme sur  $\{1, \dots, n\}$ .

Simuler les résultats pour différentes valeurs de  $n$ . Comparer les deux approches.

**Exercice 4.** Soit  $X$  une variable aléatoire de loi uniforme sur  $[a, b]$  avec  $a, b \in \mathbb{R}$ ,  $a < b$ . Proposer un algorithme pour la simulation de cette loi et tester le pour différentes valeurs de  $a$  et  $b$ .

**Exercice 5.** Soit  $X$  une variable aléatoire de loi exponentielle de paramètre  $\lambda$ . Proposer un algorithme pour la simulation de cette loi en utilisant la méthode de la fonction inverse.

**Exercice 6.** Soit  $X$  une variable aléatoire géométrique de paramètre  $p$ . Proposer un algorithme pour la simulation de cette loi :

1. En utilisant le résultat : Si  $(Y_i)_{i \in \mathbb{N}^*}$  sont des variables aléatoires *i.i.d.* de loi de Bernoulli de paramètre  $p$ , alors  $N = \min\{i; Y_i = 1\}$  suit une loi géométrique de paramètre  $p$ .
2. En utilisant le résultat : Si  $U$  suit une loi uniforme sur  $[0, 1]$  alors  $G = 1 + \left\lceil \frac{\ln U}{\ln(1-p)} \right\rceil$  suit la loi géométrique de paramètre  $p$ .
3. Comparer les résultats en insistant sur le temps de calcul.

**Exercice 7.** Soit  $X_n$  une variable aléatoire de loi géométrique de paramètre  $p_n = \lambda/n$  et  $Y$  une v.a. de loi exponentielle de paramètre 1.

1. Quelle est la loi de  $[\theta Y] + 1$  où  $[y]$  désigne la partie entière de  $y$  et  $\theta > 0$  ?
2. Simuler un  $N$ -échantillon de  $X_n$  à l'aide d'une v.a.  $Y$  de loi exponentielle de paramètre 1.
3. Établir la convergence (en loi) de la suite  $(X_n/n)_n$  vers la variable  $Y/\lambda$ .



4. Illustrer numériquement cette convergence.

**Exercice 8.** Soit  $Y$  une v.a. de loi exponentielle de paramètre  $\lambda$ , avec  $\lambda > 0$ . Sa densité est donnée par

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{\{x > 0\}}.$$

On définit une v.a.  $Z = 1 + [Y]$ , où, pour tout  $x \in \mathbb{R}$ ,  $[x]$  note la partie entière de  $x$ .

1. Quelle est la loi de  $Z$ ? Justifier votre réponse.
2. Soient  $p \in ]0, 1[$  et  $U$  une v.a. de loi uniforme sur  $[0, 1]$ . Trouver la constante  $\lambda$  telle que

$$1 + \left\lceil -\frac{\ln U}{\lambda} \right\rceil \sim \mathcal{G}(p).$$

3. Utiliser ce résultat pour écrire un algorithme pour la simulation d'une v.a.  $X$  de loi géométrique de paramètre  $p$ .
4. Donner des résultats numériques de votre algorithme pour  $p = 0.1$ ,  $p = 0.5$  et  $p = 0.9$  (4 valeurs pour chaque choix de  $p$ ).

**Exercice 9.** Soit  $X$  une variable aléatoire de loi de Poisson de paramètre  $\lambda$ . Simuler cette variable aléatoire en utilisant le résultat suivant : Soit  $(X_i)_{i \in \mathbb{N}^*}$  une suite de variables aléatoires *i.i.d.* de loi exponentielle de paramètre  $\lambda$ . Alors la variable aléatoire :

$$N = \begin{cases} 0, & \text{si } X_1 \geq 1 \\ \max\{i \geq 1; \sum_{j=1}^i X_j \leq 1\}, & \text{sinon} \end{cases}$$

suit une loi de Poisson de paramètre  $\lambda$ .

**Exercice 10.** Soit  $X$  une variable aléatoire de loi normale de moyenne  $m$  et variance  $\sigma^2$ . Simuler cette variable aléatoire en utilisant la méthode de Box-Muller.

**Exercice 11.** En utilisant la méthode de la fonction de répartition, simuler une loi de Cauchy. Soit  $X$  une variable aléatoire de loi de Cauchy de paramètre  $a > 0$ . Sa densité est donnée par :

$$f(x) = \frac{a}{\pi(x^2 + a^2)}, \quad x \in \mathbb{R}.$$

1. Vérifier que  $f$  est une densité.
2. Calculer sa fonction de répartition  $F$ .
3. En déduire  $G$  l'inverse de la fonction  $F$ .

4. Soit  $U$  une v.a. de loi uniforme sur  $[0, 1]$ . En utilisant la périodicité de la fonction  $\tan$ , montrer que

$$\tan\left(\pi\left(U - \frac{1}{2}\right)\right) = \tan(\pi U), \quad \text{en loi.}$$

5. Écrire un algorithme pour simuler une v.a. de loi de Cauchy de paramètre  $a$ .  
 6. Donner des réalisations de votre code **Matlab** ou **Scilab** pour  $a = 10$  et  $a = 1$  (4 valeurs numériques pour chaque choix de  $a$ ).  
 7. Tracer la vraie densité et un histogramme à l'aide de cet algorithme pour  $a = 10$  et  $a = 1$ . Utiliser un échantillon de taille 10000.

**Exercice 12.** Soit  $X$  une v.a. de densité

$$f(x) = xe^{-\frac{x^2}{2}} \mathbb{1}_{[0, +\infty[}(x).$$

- Montrer que  $f$  est une densité.
- Calculer la fonction de répartition  $F$  de  $X$ .
- Montrer que  $F$  admet une fonction inverse  $G$  et la calculer.
- Écrire un algorithme **Matlab** ou **Scilab** pour simuler  $X$  en utilisant la méthode de simulation par inversion de la fonction de répartition.
- Donner 5 résultats numériques de votre algorithme.
- Représenter graphiquement la densité de  $X$  en utilisant des valeurs entre 0 et 10 avec un pas de 0.1.

**Exercice 13.** Simulation de la gaussienne par rejet par rapport à la double exponentielle. On pose

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \text{ et } g(z) = \frac{1}{2} \exp(-|z|).$$

- Montrer que  $g$  est bien une densité sur  $\mathbb{R}$ .
- Déterminer une constante  $k$  satisfaisant  $\forall z \in \mathbb{R}, f(z) \leq kg(z)$ . On aura intérêt à prendre  $k$  la plus petite possible :  $[k = \sqrt{\frac{2e}{\pi}} = 1,3155]$ .
- Écrire un algorithme de simulation d'une loi gaussienne centrée réduite par rejet.

**Exercice 14.** Soient  $\alpha \geq 1$  et  $\beta > 1$  et notons par  $\mathcal{B}(\alpha, \beta)$  la v.a. de loi Beta. Sa densité est donnée par

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \mathbb{1}_{\{0 < x < 1\}},$$

où  $B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$ .

1. Montrer que pour tout  $x \in ]0, 1[$ ,  $f(x) \leq C_{\alpha, \beta}$  où  $C_{\alpha, \beta} = f(x_{\alpha, \beta})$  avec  $x_{\alpha, \beta} = \frac{\alpha - 1}{\alpha + \beta - 2}$ .
2. En utilisant le point 1, proposer un algorithme de simulation par rejet pour la v.a. de loi Beta.
3. Nous nous plaçons dans le cas  $\alpha = 2$  et  $\beta = 2$ .
  - (i) Dédire la forme de la densité  $f(x)$ .
  - (ii) Calculer  $C_{2,2}$  et écrire le code **Matlab** ou **Scilab** pour simuler une v.a.  $\mathcal{B}(2, 2)$  par rejet.
  - (iii) Produire un échantillon de taille 10 de la v.a. ayant cette loi.
  - (iv) Tracer un histogramme de la loi.

**Exercice 15.** Soient  $a > 0$  fixé et  $X$  une variable aléatoire sur  $\mathbb{R}_+^*$ , de densité

$$g_a(x) = ax^{a-1} \exp(-x^a), \forall x > 0. \quad (1.11)$$

1. Montrer que  $g_a$  est une densité.
2. Calculer la fonction de répartition  $G_a$  de  $X$ .
3. Montrer que  $G_a$  admet une fonction inverse  $H_a$  et la calculer.
4. Écrire un algorithme **Matlab** pour simuler  $X$  en utilisant la méthode de simulation par inversion de la fonction de répartition.
5. Donner 5 résultats numériques de votre algorithme.
6. Représenter graphiquement la densité de  $X$ .
7. Tracer un histogramme en utilisant 1000 valeurs simulées.

**Exercice 16.** Soit  $0 < a < 1$  fixé. Considérons  $Y$  une variable aléatoire de loi Gamma de paramètre  $a$ , de densité :

$$f_a(x) = \frac{1}{\Gamma(a)} x^{a-1} \exp(-x), \forall x > 0,$$

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

1. Montrer que pour tout  $x > 0$  il existe une constante  $C_a > 0$  telle que

$$f_a(x) \leq C_a g_a(x),$$

où  $g_a(x)$  est la densité donnée dans l'exercice 15. Trouver la valeur minimale de  $C_a$ .

2. En utilisant le point 1, proposer un algorithme de simulation par rejet pour la variable aléatoire de loi Gamma de paramètre  $a$ .
3. Nous nous plaçons dans les cas  $a = 0.1$  et  $a = 0.5$ .
  - (i) Dédire la forme de la densité  $f_a(x)$ .
  - (ii) Écrire le code **Matlab** pour simuler une variable aléatoire de loi Gamma de paramètre  $a$  par rejet. Le code devra rendre en sortie la valeur simulée de la variable aléatoire et le nombre de rejets pour chaque valeur de  $a$ .

- (iii) Produire un échantillon de taille 10 de la variable aléatoire ayant cette loi, pour chaque  $a$ .

**Exercice 17.** Simuler un mélange d'exponentielles  $F(x) = \alpha(1 - \exp(-ax)) + (1 - \alpha)(1 - \exp(-bx))$ , avec  $\alpha \in [0, 1]$ .

**Exercice 18.** Simuler une variable aléatoire de loi  $\frac{1}{2}\delta_0(x) + \frac{1}{2}e^{-x}\mathbb{1}_{\mathbb{R}_+}(x)dx$ .

## Simulation de variables aléatoires

Dans tout ce qui suit  $U$  note une v.a. de loi uniforme sur  $[0, 1]$  et  $(U_n)_{n \geq 1}$  une suite de v.a. i.i.d., avec  $U_1$  de loi uniforme sur  $[0, 1]$ .

Loi	Méthode de simulation
Bernoulli $\mathcal{B}(p)$ , $p \in [0, 1]$	$\mathbb{1}_{\{U \leq p\}}$
Binomiale $\mathcal{B}(n, p)$ , $n \in \mathbb{N}^*$ , $p \in [0, 1]$	$\sum_{i=1}^n \mathbb{1}_{\{U_i \leq p\}}$
Uniforme sur $\{0, \dots, n-1\}$	$[nU]$
Uniforme sur $[a, b]$	$a + (b - a)U$
Exponentielle $\mathcal{E}(\lambda)$	$-\frac{\ln(U)}{\lambda}$
Géométrique $\mathcal{G}(p)$	$1 + \left\lceil \frac{\ln(U)}{\ln(1-p)} \right\rceil$
Poisson $\mathcal{P}(\lambda)$	$\min \left\{ n \in \mathbb{N} : \prod_{i=1}^n U_i \leq e^{-\lambda} \right\}$
Gaussienne $\mathcal{N}(m, \sigma^2)$	$m + \sigma \sqrt{-2 \ln(U_1)} \sin(2\pi U_2)$
Cauchy $\mathcal{C}(a)$	$a \tan(\pi U)$

# Chapitre 2

## Méthodes de Monte Carlo

### 2.1 Introduction

On appelle *Méthode de Monte Carlo* toute méthode numérique basée sur le tirage de nombres aléatoires. Ces méthodes sont utilisées dans de nombreux domaines : en physique nucléaire, en géophysique, en finance, en statistiques, en chimie, etc.

Origine On attribue l'origine de la méthode de Monte Carlo au comte de Buffon qui en 1777 a proposé le problème de l'aiguille de Buffon qui fournit une méthode pour le calcul de  $\pi$  basée sur la réalisation d'expériences répétées.

*Aiguille de Buffon* : Sur un parquet composé de planches parallèles de même largeur  $a$ , on jette des aiguilles de même longueur  $b$  au hasard. La probabilité qu'une aiguille tombe à cheval sur deux planches est, pour  $b < a$ , égale à  $\frac{2b}{\pi a}$ . Ceci donne une approximation de  $\pi$  en répétant le jet des aiguilles.

Le vrai développement de la méthode de Monte Carlo est lié à l'apparition des ordinateurs.

#### 2.1.1 Idée de la méthode pour le calcul d'intégrales

Supposons qu'on souhaite calculer par une méthode numérique l'intégrale

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

pour une fonction intégrable  $f : [0, 1] \mapsto \mathbb{R}$ . On utilise alors des formules du type  $\sum_{i=1}^n \theta_i f(x_i)$  où  $\theta_i \geq 0$  sont tels que  $\sum_{i=1}^n \theta_i = 1$  et  $x_i \in [0, 1]$ .

Suivant le choix de  $\theta_i$  et  $x_i$  on peut construire et proposer différentes méthodes d'approximation.

Une méthode de Monte Carlo est du même type : on choisit  $\theta_i = \frac{1}{n}$  et on tire les  $x_i$  selon la loi uniforme sur  $[0, 1]$ . Cette méthode converge avec une vitesse de l'ordre  $\frac{C}{\sqrt{n}}$ . En dimension 1 cette vitesse peut paraître faible lorsqu'on la compare aux autres méthodes d'intégration déterministe. Mais toutes ces méthodes numériques s'effondrent lorsque la dimension augmente : dans  $\mathbb{R}^d$  il faut  $n^d$  points pour garder une erreur constante. Le gros avantage de la méthode de Monte Carlo est d'être *insensible à la dimension*.

## 2.2 Description de la méthode de Monte Carlo

Pour utiliser une méthode de Monte Carlo il faut tout d'abord mettre sous forme d'une espérance la quantité que l'on cherche à évaluer. En admettant que cela est possible on doit donc calculer une quantité de la forme  $\mathbb{E}(X)$  avec  $X$  une variable aléatoire.

Pour pouvoir évaluer  $\mathbb{E}(X)$  il est souhaitable de savoir simuler une v.a. selon la loi de  $X$ . Il est ainsi possible de simuler une suite des v.a.i.i.d.  $(X_i)_{i \geq 1}$  de même loi que  $X$ . Pour  $n \in \mathbb{N}^*$ , on approche ensuite  $\mathbb{E}(X)$  par la moyenne arithmétique des  $(X_i)_{i \geq 1}$

$$\mathbb{E}(X) \approx \frac{1}{n} \sum_{i=1}^n X_i.$$

## 2.3 Convergence de la méthode et vitesse de convergence

Comme pour toute méthode numérique, nous souhaitons répondre aux deux questions suivantes :

1. Pourquoi la méthode converge ?
2. A quelle vitesse la méthode converge ?

La réponse à ces deux questions est fournie par deux théorèmes fondamentaux de la théorie des probabilités.

### 2.3.1 Convergence

La réponse à la première question est donnée par la loi forte des grands nombres.

**Théorème 1. (*Loi forte des grands nombres*)** Soit  $(X_i)_{i \geq 1}$  des variables aléatoires indépendantes identiquement distribuées et intégrables, de même loi que la variable aléatoire  $X$ . Alors, presque sûrement (et dans  $L^1$ ),

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mathbb{E}(X).$$

**Remarque 1.** Ce résultat précise une limite théorique pour la méthode de Monte Carlo, on ne peut l'utiliser que pour des variables aléatoires intégrables.

### 2.3.2 Vitesse de convergence

Le théorème de la limite centrale précise la convergence. Ce théorème donne le comportement asymptotique de l'erreur, qui finit par ressembler à une loi gaussienne centrée.

**Théorème 2. (*Théorème de la limite centrale*)** Soit  $(X_i)_{i \geq 1}$  une suite de variables aléatoires indépendantes suivant toutes la même loi qu'une variable aléatoire  $X$ . On suppose que  $\mathbb{E}(X^2) < \infty$  et on note  $\sigma^2 = \text{Var}(X)$ . Alors,

$$\frac{\sqrt{n}}{\sigma} \left( \mathbb{E}(X) - \frac{1}{n} (X_1 + \dots + X_n) \right) \text{ converge en loi vers } G,$$

où  $G$  note une variable aléatoire normale centrée et réduite.

**Remarque 2.** De ce théorème on peut déduire que pour  $-\infty \leq a < b \leq +\infty$  on a

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{\sigma}{\sqrt{n}} a \leq \mathbb{E}(X) - \frac{1}{n} (X_1 + \dots + X_n) \leq \frac{\sigma}{\sqrt{n}} b \right\} &= \mathbb{P}(a \leq G \leq b) \\ &= \phi(b) - \phi(a) \\ &= \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \end{aligned}$$

où  $\phi$  note la fonction de répartition de la loi normale centrée et réduite

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy.$$

**Remarque 3.** Le théorème central limite ne permet pas de borner l'erreur puisque le support de la gaussienne est égal à  $\mathbb{R}$  tout entier. On présente alors souvent l'erreur de la méthode de Monte Carlo soit en donnant l'écart type de l'erreur, c'est-à-dire  $\frac{\sigma}{\sqrt{n}}$ , soit en donnant un intervalle de confiance



à 95 % pour le résultat. Ceci signifie que le résultat obtenu se trouve avec une probabilité de 95 % dans l'intervalle donné. Avec les propriétés de la loi normale on a

$$\mathbb{P}\{|G| \leq 1.96\} \approx 0.95$$

ce qui conduit à un intervalle de confiance du type

$$\left[ m - 1.96 \frac{\sigma}{\sqrt{n}}, m + 1.96 \frac{\sigma}{\sqrt{n}} \right],$$

où  $m = \mathbb{E}(X)$ .

**Remarque 4.** On remarque ainsi que, si l'on souhaite diminuer l'erreur, il faut augmenter  $n$  et/ou diminuer  $\sigma$ .

## 2.4 Une application

Nous allons étudier sur un exemple la construction et les propriétés de la méthode de Monte Carlo.

Cadre : Soit  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  une fonction intégrable. On veut calculer une valeur approchée de

$$I = \int_{\mathbb{R}^d} g(x) dx.$$

Cette intégrale peut par exemple provenir d'un problème concret : en fiabilité, calculer la durée moyenne de vie (Mean Time To Failure MTTF) est souvent impossible analytiquement.

Nous admettons que les hypothèses suivantes sont satisfaites.

Hypothèses :

1.  $\int_{\mathbb{R}^d} g^2(x) dx < +\infty$ .
2. Il existe une densité  $f$  sur  $\mathbb{R}^d$  telle que  $\int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx < +\infty$ .
3. On sait simuler une variable aléatoire de densité  $f$  et on a à notre disposition une suite  $(X_i)_{i \geq 1}$  de v.a.i.i.d. de densité  $f$ .

Buts :

1. Donner une valeur approchée  $\hat{I}_n$  de  $I$  en fonction de  $X_1, \dots, X_n$ , pour  $n \in \mathbb{N}^*$ .
2. Écrire l'algorithme.
3. Étudier sa convergence et estimer l'erreur.
4. Améliorer la vitesse et comparer avec d'autres méthodes de calcul d'intégrales.

## 2.5 Loi des grands nombres et estimateur

Pour calculer  $I = \int_{\mathbb{R}^d} g(x)dx$ , l'idée est de l'écrire comme l'espérance d'une fonction de la variable aléatoire  $X$  que l'on sait simuler. On pose pour tout  $i \geq 1$ ,  $Y_i = \frac{g(X_i)}{f(X_i)}$ . Les variables aléatoires  $(Y_i)_{i \geq 1}$  sont indépendantes et de même loi ; leur espérance commune est

$$\mathbb{E}(Y_1) = \int_{\mathbb{R}^d} \frac{g(x)}{f(x)} f(x)dx = I.$$

On introduit naturellement l'*estimateur* suivant :

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n Y_i = \frac{1}{n} \sum_{i=1}^n \frac{g(X_i)}{f(X_i)}.$$

Comme  $\mathbb{E}(\hat{I}_n) = I$ , l'estimateur est *sans biais*. En appliquant la loi forte des grands nombres on déduit :

Soit  $(X_i)_{i \geq 1}$  une suite de v.a.i.i.d. de densité  $f$ . Alors

$$\lim_{n \rightarrow \infty} \hat{I}_n = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \frac{g(X_i)}{f(X_i)} = I = \int_{\mathbb{R}^d} g(x)dx \quad p.s. \text{ et dans } L^1.$$

## 2.6 Variance, erreurs et intervalle de confiance

### 2.6.1 Variance et erreurs

**Exemple 1.** Soit  $(Y_i)_{i \geq 1}$  une suite de variables aléatoires indépendantes identiquement distribuées admettant un moment d'ordre 2. On pose  $\hat{I}_n = \frac{1}{n} \sum_{i=1}^n Y_i$ . Calculer l'espérance et la variance de  $\hat{I}_n$  en fonction de l'espérance et la variance de  $Y_1$ .

Calculons la variance de  $Y_1$  :

$$\mathbb{E}(Y_1^2) = \int_{\mathbb{R}^d} \frac{g^2(x)}{f^2(x)} f(x)dx = \int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx < +\infty \text{ et}$$

$$\text{Var}(Y_1) = \int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx - I^2.$$

On a immédiatement

$$\text{Var}(\hat{I}_n) = \frac{\sigma^2}{n}. \quad (2.1)$$

L'application du théorème central limite implique que, lorsque  $n$  converge vers l'infini,

$$\frac{\sqrt{n}}{\sigma}(\hat{I}_n - I) \text{ converge en loi vers } \mathcal{N}(0, 1),$$

et signifie intuitivement que, pour  $n$  très grand,  $\hat{I}_n - I \approx \frac{\sigma}{\sqrt{n}}G$ , où  $G$  est une gaussienne centrée réduite. Il est donc naturel d'introduire l'erreur standard  $\frac{\sigma}{\sqrt{n}}$  et l'erreur relative  $\frac{1}{I} \frac{\sigma}{\sqrt{n}}$ .

Cependant, la variance  $\sigma^2$  est en général inconnue (en particulier, elle dépend de  $I$ ). On va donc la remplacer par l'estimateur classique de la variance :

**Estimateur de la variance :**

$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \hat{I}_n)^2 = \frac{1}{n-1} \left( \sum_{i=1}^n Y_i^2 - n\hat{I}_n^2 \right).$$

On sait (voir cours de statistique) que  $s_n^2$  converge *p.s.* vers  $\sigma^2$ . Pour estimer l'erreur standard et l'erreur relative, on remplace donc  $\sigma^2$  par  $s_n^2$  :

Nous allons donc utiliser les approximations suivantes :

$$\begin{aligned} \text{Erreur standard} : \frac{\sigma}{\sqrt{n}} &\approx \frac{s_n}{\sqrt{n}} \\ \text{Erreur relative} : \frac{1}{I} \frac{\sigma}{\sqrt{n}} &\approx \frac{1}{\hat{I}_n} \frac{s_n}{\sqrt{n}}. \end{aligned}$$

**En pratique :** Dans l'algorithme de Monte Carlo pour calculer  $\hat{I}_n$ , il sera judicieux d'ajouter le calcul de  $s_n^2$  pour avoir en même temps une estimation de l'erreur.

## 2.6.2 Intervalle de confiance par TCL

On peut maintenant donner des intervalles de confiance pour notre estimation, dans l'utilisation du théorème central limite, on remplace la variance  $\sigma^2$  par son estimation  $s_n^2$  :

$$\mathbb{P}(|\hat{I}_n - I| \leq \varepsilon) = \mathbb{P}\left(\frac{\sqrt{n}}{s_n}|\hat{I}_n - I| \leq \frac{\sqrt{n}}{s_n}\varepsilon\right) \approx 2\Phi\left(\frac{\sqrt{n}}{s_n}\varepsilon\right) - 1.$$

Pour un intervalle de confiance au niveau  $\alpha = 95\%$ , on prend

$$2\Phi\left(\frac{\sqrt{n}}{s_n}\varepsilon\right) - 1 = \alpha = 0.95 \quad \Leftrightarrow \quad \Phi\left(\frac{\sqrt{n}}{s_n}\varepsilon\right) = \frac{\alpha + 1}{2} = 0.975$$

et donc  $\frac{\sqrt{n}}{s_n}\varepsilon = 1.96$ , ou encore  $\varepsilon = 1.96 \frac{s_n}{\sqrt{n}}$ .

L'intervalle de confiance pour  $I$  au niveau 95% est

$$\left[ \hat{I}_n - 1.96 \frac{s_n}{\sqrt{n}}, \hat{I}_n + 1.96 \frac{s_n}{\sqrt{n}} \right].$$

Soit  $\alpha \in [0, 1]$  proche de 1. L'intervalle de confiance au niveau  $\alpha$  est

$$\left[ \hat{I}_n - \Phi^{-1} \left( \frac{\alpha + 1}{2} \right) \frac{s_n}{\sqrt{n}}, \hat{I}_n + \Phi^{-1} \left( \frac{\alpha + 1}{2} \right) \frac{s_n}{\sqrt{n}} \right].$$

Pour le niveau de confiance  $\alpha$  fixé, la largeur de l'intervalle de confiance décroît en  $\frac{\sigma}{\sqrt{n}}$ . On dit que la méthode de Monte Carlo a une vitesse de convergence en  $\frac{1}{\sqrt{n}}$ .

**En pratique :** On peut aussi parfois majorer la variance, et utiliser l'inégalité de Tchebychev pour obtenir des intervalles de confiance.

## 2.7 Algorithme de Monte Carlo

On suppose qu'on a à notre disposition une procédure `simulf` dont les différents appels simulent une suite de v.a.i.i.d. de densité  $f$ . La fonction `cdfnor("X",0,1,p,1-p)` donne l'inverse de la fonction de répartition de la loi normale centrée réduite en  $p$  :

$$x = \text{cdfnor}("X", 0, 1, p, 1 - p) \Leftrightarrow \mathbb{P}(G \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = p,$$

où  $G$  suit la loi normale centrée réduite.

Dans l'algorithme qui suit, on fixe le niveau de confiance  $\alpha$  (proche de 1), et la largeur de l'intervalle de confiance  $2\delta$  au niveau de confiance  $\alpha$ . Le nombre  $n$  de simulations nécessaires est décidé par un test, et est par conséquent aléatoire :

**Entrées de l'algorithme :**

```
function [n,I,e] = montecarlo(alpha,delta)
// estimation de I par Monte Carlo
// alpha : niveau de confiance alpha
// 2delta : largeur intervalle de confiance
// n : nombre de simulations effectuees
// I : valeur estimee, e:erreur standard
p=(alpha+1)/2; z= cdfnor("X",0,1,p,1-p)};
x=simulf; y=g(x)/f(x);
```

```

S1=y; S2=y^2;
x=simulfr; y=g(x)/f(x);
n=2; S1=S1+y; S2=S2+y^2; V=(S2-S1^2/n)/(n-1);
while (z*sqrt(V/n)>delta);
    x=simulfr; y=g(x)/f(x);
    n=n+1; S1=S1+y; S2=S2+y^2; V=(S2-S1^2/n)/(n-1);
end;
I=S1/n; e=sqrt(V);

```

### Commentaires

On stocke  $\sum_{i=1}^n Y_i$  dans **S1**,  $\sum_{i=1}^n Y_i^2$  dans **S2**, l'estimateur de la variance  $s_n^2$  dans **V**. Le test regarde si la demi-largeur de l'intervalle de confiance après  $n$  simulations,  $z \frac{s_n}{\sqrt{n}}$  est supérieure à celle souhaitée,  $\delta$ , auquel cas on refait une simulation.

### Conclusions

1. L'algorithme est extrêmement facile à mettre en place.
2. Il ne suppose pas de régularité sur  $g$ , autre que l'intégrabilité.
3. L'erreur pour  $n$  simulations est de l'ordre de  $\frac{\sigma}{\sqrt{n}}$ , ce qui justifie l'appellation "méthode de vitesse  $\frac{1}{\sqrt{n}}$ " :
  - Cette vitesse est indépendante de la dimension : c'est donc un avantage quand la dimension est grande, par contre en dimension petite, cette méthode est peu compétitive par rapport aux méthodes d'analyse numérique, en particulier quand la fonction à intégrer est régulière.
  - On voit l'importance de la variance  $\sigma^2$ . Une voie pour améliorer l'erreur sera d'essayer de réduire la variance.
4. C'est une méthode aléatoire : le critère de sortie de l'algorithme est aléatoire.

## 2.8 Choix d'une méthode de Monte Carlo

Supposons qu'on ait le choix entre deux méthodes de Monte Carlo MC1 et MC2.

	Densité	Var	Coût par sim.	Nbr. de sim.	Coût total	Erreur std
MC1	$f_1$	$\sigma_1^2$	$c_1$	$n_1$	$n_1 c_1$	$\frac{\sigma_1}{\sqrt{n_1}}$
MC2	$f_2$	$\sigma_2^2$	$c_2$	$n_2$	$n_2 c_2$	$\frac{\sigma_2}{\sqrt{n_2}}$

Laquelle doit-on choisir ? On rappelle que

$$\sigma^2 = \text{Var}_f \left( \frac{g(X)}{f(X)} \right) = \int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx - \left( \int_{\mathbb{R}^d} g(x) dx \right)^2.$$

**Avec erreurs standards égales :**  $\frac{\sigma_1}{\sqrt{n_1}} = \frac{\sigma_2}{\sqrt{n_2}} \Leftrightarrow \frac{\sigma_1^2}{n_1} = \frac{\sigma_2^2}{n_2}$ .

Pour une même erreur, la meilleure méthode est celle qui coûte le moins cher en calcul : on compare donc les coûts totaux de simulation  $n_i c_i$  :

$$\frac{n_1 c_1}{n_2 c_2} = \frac{\sigma_1^2 c_1}{\sigma_2^2 c_2}.$$

**Avec coûts de simulation égaux :**  $n_1 c_1 = n_2 c_2$ .

Pour un même coût de simulation, la meilleure méthode est celle qui offre la plus petite erreur standard : on compare donc les erreurs standards, ou plutôt les erreurs standards au carré  $\frac{\sigma_i^2}{n_i}$  :

$$\frac{\sigma_1^2 n_2}{\sigma_2^2 n_1} = \frac{\sigma_1^2 c_1}{\sigma_2^2 c_2}.$$

Dans les deux cas, la meilleure méthode est celle qui minimise  $c_i \sigma_i^2$ .

On a vu que dans l'erreur  $\frac{\sigma}{\sqrt{n}}$ , le facteur  $\frac{1}{\sqrt{n}}$  est intrinsèque, et que pour réduire l'erreur, on peut essayer de diminuer la variance  $\sigma^2$ . Il faudra cependant veiller à ce que cette diminution de variance ne soit pas compensée par une augmentation exagérée des coûts de calcul.

## 2.9 Exemples - Calcul approché de $\frac{\pi}{4}$

Nous allons construire plusieurs versions de la méthode de Monte Carlo pour le calcul approché de  $\frac{\pi}{4}$ . Afin de les comparer, nous allons écrire à

chaque fois la variance de l'estimation après  $n$  simulations sous la forme  $\frac{C}{n}$  et comparer les différents  $C$ . L'ensemble des calculs se réalisera avec 4 chiffres après la virgule.

### (1) Calcul approché de $\frac{\pi}{4}$ : Technique du hit or miss

On considère le carré  $[0, 1]^2$ , et on trace le quart de disque  $D$  centré en 0 et de rayon 1.

$$D = \{(x, y) \in \mathbb{R}_+^2 / x^2 + y^2 \leq 1\}.$$

Notons que l'aire du quart de disque  $D$  vaut  $\frac{\pi}{4}$ .

**Méthode de Monte Carlo (1) - Hit or Miss** Soit  $(X_i)_{i \geq 1}$  des v.a.i.i.d. de loi uniforme sur  $[0, 1]^2$ , on pose

$$Y_i = \mathbb{1}_{\{X_i \in D\}}.$$

On a  $\mathbb{E}(Y_1) = \mathbb{P}(X_1 \in D) = \frac{\pi}{4}$ . Pour un nombre  $n$  fixé de simulations l'estimateur

$$\hat{I}_n^{(1)} = \frac{1}{n} \sum_{i=1}^n Y_i$$

est un estimateur de  $\frac{\pi}{4}$ . La variance de l'estimateur est

$$\text{Var}(\hat{I}_n^{(1)}) = \frac{\text{Var}(Y_1)}{n}$$

donc

$$C^{(1)} = \text{Var}(Y_1) = \frac{\pi}{4} \left(1 - \frac{\pi}{4}\right) = 0.1685.$$

### (2) Calcul approché de $\frac{\pi}{4}$ : Moyenne empirique sur $[0, 1]$

L'équation du quart de cercle est donnée par la fonction  $g(x) = \sqrt{1 - x^2}$ ,  $g : [0, 1] \mapsto \mathbb{R}^+$ . Nous avons le choix de la densité. Par exemple, si  $g$  est à support compact le choix naturel est la **la loi uniforme sur ce compact**. On appelle cela la **Méthode de Monte Carlo standard**.

**Méthode de Monte Carlo (2) - Moyenne empirique** Soit  $(X_i)_{i \geq 1}$  des v.a.i.i.d. de loi uniforme sur  $[0, 1]$ , on pose

$$Y_i = g(X_i) = \sqrt{1 - X_i^2}.$$

On a

$$\mathbb{E}(Y_1) = \mathbb{E}(g(X_1)) = \int_0^1 \sqrt{1 - x^2} dx = \frac{\pi}{4}.$$

Pour un nombre  $n$  fixé de simulations l'estimateur

$$\hat{I}_n^{(2)} = \frac{1}{n} \sum_{i=1}^n Y_i$$

est un autre estimateur de  $\frac{\pi}{4}$ . La variance de l'estimateur est

$$\text{Var}(\hat{I}_n^{(2)}) = \frac{\text{Var}(Y_1)}{n}$$

donc

$$\begin{aligned} C^{(2)} &= \text{Var}(Y_1) = \int_0^1 g^2(x) dx - \left(\frac{\pi}{4}\right)^2 \\ &= \int_0^1 (1-x^2) dx - \left(\frac{\pi}{4}\right)^2 = \frac{2}{3} - \left(\frac{\pi}{4}\right)^2 = 0.0498. \end{aligned}$$

### Comparaison

On note que  $C^{(2)} = 0.0498 < C^{(1)} = 0.1685$ . La variance est plus faible dans le second cas, la seconde méthode est donc meilleure.

Ce résultat est vrai dans le cas d'une fonction  $g : [0, 1] \rightarrow [0, 1]$  générale : la technique de la moyenne empirique a une variance plus petite que la technique hit or miss (voir Rubinstein [?], paragraphe 4.2.3).

## 2.10 Techniques de réduction de variance

**But :** On a vu que l'erreur dans une méthode de Monte Carlo pour le calcul de  $I = \int_0^1 g(x) dx$  et une densité  $f$ , est de la forme

$$\frac{\sigma}{\sqrt{n}} \quad \text{avec} \quad \sigma^2 = \text{Var}_f \frac{g(X)}{f(X)} = \int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx - I^2.$$

Le facteur en  $\frac{1}{\sqrt{n}}$  étant intrinsèque à la méthode de Monte Carlo, on cherche à réduire la variance  $\sigma^2$ , pour diminuer l'erreur commise, à nombre de simulation fixé.

### 2.10.1 Échantillonnage préférentiel (importance sampling) : choix de la densité $f$

Dans l'estimation précédente, on a la choix de la densité  $f$ . Si le support de  $g$  est un ensemble  $E$  de mesure de Lebesgue finie de  $\mathbb{R}^d$ , la première idée



pour choisir  $f$  est de prendre simplement la loi uniforme sur  $E$  : c'est ce qu'on appelle la **Méthode de Monte Carlo standard**. Ce n'est cependant pas nécessairement le meilleur choix possible.

Pour une suite de variables aléatoires i.i.d.  $(X_i)_{i \geq 1}$  de densité  $f$  et  $Y_i = g(X_i)/f(X_i)$  pour tout  $i \geq 1$ , on a vu que l'erreur était contrôlée par la variance

$$\text{Var}(Y_1) = \sigma^2 = \int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx - I^2.$$

**Théorème 3.** *La variance minimale est égale à  $(\int_{\mathbb{R}^d} |g(x)| dx)^2 - I^2$ , et est atteinte pour*

$$f_0(y) = \frac{|g(y)|}{\int_{\mathbb{R}^d} |g(x)| dx}.$$

**Démonstration :** Pour montrer que  $f_0$  réalise bien le minimum de la variance, il suffit de vérifier que pour toute densité  $f$  sur  $\mathbb{R}^d$  :

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx &\geq \int_{\mathbb{R}^d} \frac{g^2(x)}{|g(x)|} dx \int_{\mathbb{R}^d} |g(x)| dx \\ &= \left( \int_{\mathbb{R}^d} |g(x)| dx \right)^2. \end{aligned}$$

Mais, par l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} \left( \int_{\mathbb{R}^d} |g(x)| dx \right)^2 &= \left( \int_{\mathbb{R}^d} \frac{|g(x)|}{\sqrt{f(x)}} \sqrt{f(x)} dx \right)^2 \\ &\leq \int_{\mathbb{R}^d} \frac{|g(x)|^2}{f(x)} dx \int_{\mathbb{R}^d} f(x) dx \\ &= \int_{\mathbb{R}^d} \frac{g^2(x)}{f(x)} dx. \end{aligned}$$

La valeur du minimum se vérifie facilement. ■

Remarquons que, dans le cas particulier où  $g > 0$ , la densité optimale est  $f = g/I$ , auquel cas la variance serait nulle. Ce choix est bien sûr d'intérêt purement théorique, puisque  $I$  est inconnu. Cependant...

**Échantillonnage préférentiel :** Afin de réduire la variance, on essaie d'adapter la densité  $f$  à la fonction à intégrer  $g$  : on a intérêt à prendre  $f$  grande dans les régions où  $|g|$  est grande et petite dans les régions où  $|g|$  est petite.

### (3) Calcul approché de $\frac{\pi}{4}$ : Échantillonnage préférentiel

On reprend la méthode de la moyenne empirique, en essayant de choisir une densité  $f$  meilleure que la densité de la loi uniforme sur  $[0, 1]$  : on peut prendre par exemple une fonction affine décroissante sur  $[0, 1]$ , comme  $f(x) = \frac{3}{2} - x$ . Alors

$$\begin{aligned}\int_0^1 \frac{g^2(x)}{f(x)} dx &= \int_0^1 \frac{1-x^2}{\frac{3}{2}-x} dx \\ &= \int_0^1 \left( x + \frac{3}{2} - \frac{5}{4} \frac{1}{\frac{3}{2}-x} \right) dx \\ &= \frac{1}{2} + \frac{3}{2} - \frac{5}{4} \ln(3) \\ &= 0,6267.\end{aligned}$$

et donc  $\text{Var}_f \frac{g(X)}{f(X)} = \int_0^1 \frac{g^2(x)}{f(x)} dx - I^2 = 0.0099$ .

### Méthode de Monte Carlo (3) - Échantillonnage préférentiel

Soit  $(X_i)_{i \geq 1}$  des v.a.i.i.d. de loi  $f(x) = \frac{3}{2} - x$  sur  $[0, 1]$ . On définit, pour  $i \geq 1$

$$Y_i = \frac{g(X_i)}{f(X_i)}.$$

On a

$$\mathbb{E}(Y_1) = \mathbb{E} \left[ \frac{g(X_1)}{f(X_1)} \right] = \int_0^1 g(x) dx = \frac{\pi}{4}.$$

Pour un nombre  $n$  fixé de simulations l'estimateur

$$\hat{I}_n^{(3)} = \frac{1}{n} \sum_{i=1}^n Y_i$$

est un autre estimateur sans biais de  $I$ . La variance de cet estimateur est

$$\text{Var}(\hat{I}_n^{(3)}) = \frac{\text{Var}(Y_1)}{n}$$

donc

$$C^{(3)} = \text{Var}(Y_1) = \int_0^1 \frac{g^2(x)}{f(x)} dx - I^2 = 0.0099.$$

### Comparaison

On note que  $C^{(3)} = 0.0099 < C^{(2)} = 0.0498 < C^{(1)} = 0.1685$ . La variance est plus faible dans la méthode par échantillonnage préférentiel, elle est donc la meilleure parmi ces trois méthodes.

## 2.10.2 Échantillonnage stratifié (stratified sampling)

L'idée ici est de découper le domaine d'intégration en un nombre fini  $m$  de parties  $D_1, D_2, \dots, D_m$  deux-à-deux disjointes :

$$I = \int_{\mathbb{R}^d} g(x) dx = \sum_{i=1}^m I_i, \text{ avec } I_i = \int_{D_i} g(x) dx,$$

et d'estimer séparément chacun des  $I_{D_i}$ .

Soit  $f$  une densité sur  $\mathbb{R}^d$  fixée. On va comparer les deux estimateurs suivants

- **Estimateur de la moyenne empirique.** Soient  $(X_j)_{j \geq 1}$  des v.a.i.i.d. de densité  $f$ , l'estimateur de la moyenne empirique avec  $n$  estimations est comme précédemment

$$\hat{I}_n = \frac{1}{n} \sum_{j=1}^n \frac{g(X_j)}{f(X_j)}.$$

On note la variance de cet estimateur  $V_1$ .

- **Estimateur avec stratification.** On note  $p_i = \int_{D_i} f(x) dx$ , et, sur  $D_i$ , on renormalise  $f$  pour obtenir une densité  $f_i = \frac{f}{p_i} \mathbb{1}_{D_i}$ . Pour chaque morceau  $D_i$ , on estime  $I_i$  avec  $n_i$  simulations : soient  $(X_j^i)_{j \geq 1}$  des v.a.i.i.d. de densité  $f_i$ , l'estimateur correspondant est  $\hat{I}_{n_i}^i = \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{g(X_j^i)}{f_i(X_j^i)}$ . L'estimateur de  $I$  par échantillonnage stratifié avec  $\sum_{i=1}^m n_i$  simulations s'obtient en sommant ces  $m$  estimations :

$$\hat{I}_n^{strat} = \sum_{i=1}^m \hat{I}_{n_i}^i = \sum_{i=1}^m \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{g(X_j^i)}{f_i(X_j^i)}.$$

On note la variance de cet estimateur  $V_2$ .

Afin de comparer ces deux estimations, on fait le même nombre de simulations pour les deux :  $n = \sum_{i=1}^m n_i$ , et on compare les variances.

Dans le premier cas, la variance pour  $n$  simulations est :

$$\begin{aligned}
V_1 &:= \text{Var}(\hat{I}_n) \\
&= \frac{1}{n} \text{Var}_f \left( \frac{g(X)}{f(X)} \right) \\
&= \frac{1}{n} \int_D \left( \frac{g(x)}{f(x)} - I \right)^2 f(x) dx \\
&= \frac{1}{n} \sum_{i=1}^m p_i \int_{D_i} \left( \frac{g(x)}{f(x)} - I \right)^2 f_i(x) dx \\
&= \frac{1}{n} \sum_{i=1}^m p_i \int_{D_i} \left( \frac{g(x)}{f(x)} - \frac{I_i}{p_i} + \frac{I_i}{p_i} - I \right)^2 f_i(x) dx \\
&= \frac{1}{n} \sum_{i=1}^m \frac{1}{p_i} \int_{D_i} \left( \frac{g(x)}{f_i(x)} - I_i \right)^2 f_i(x) dx \\
&\quad + \frac{1}{n} \sum_{i=1}^m p_i \int_{D_i} \left( \frac{I_i}{p_i} - I \right)^2 f_i(x) dx \\
&= \frac{1}{n} \sum_{i=1}^m \frac{1}{p_i} \text{Var}_{f_i} \left( \frac{g(X_1^i)}{f_i(X_1^i)} \right) + \frac{1}{n} \sum_{i=1}^m \frac{I_i^2}{p_i} - \frac{1}{n} I^2.
\end{aligned}$$

Pour l'estimateur de la moyenne empirique de  $I_i$ , la variance pour  $n_i$  simulations est  $\frac{1}{n_i} \text{Var}_{f_i} \left( \frac{g(X_1^i)}{f_i(X_1^i)} \right) = \frac{1}{n_i} \sigma_i^2$ . Pour l'estimation stratifiée, avec  $n_i$  simulations pour chaque morceau  $I_i$ , la variance est, vu que les estimations des différents morceaux sont indépendantes,

$$\text{Var}(\hat{I}_n^{strat}) := V_2 = \sum_{i=1}^m \frac{1}{n_i} \text{Var}_{f_i} \left( \frac{g(X_1^i)}{f_i(X_1^i)} \right) = \sum_{i=1}^m \frac{1}{n_i} \sigma_i^2,$$

à comparer à  $V_1$  quand  $n = \sum_{i=1}^m n_i$ . On minimise donc cette variance en les  $n_i$ , sous la contrainte  $n = \sum_{i=1}^m n_i$ . On trouve (extrema liés)

$$n_i = n \frac{\sigma_i^2}{\sum_{j=1}^m \sigma_j^2},$$

ce qui donne une variance minimale pour l'estimation stratifiée avec  $n$  simulations de

$$V_2 = \frac{m}{n} \sum_{j=1}^m \sigma_j^2.$$

Si on choisit maintenant les  $D_i$  de sorte que les  $p_i = \int_{D_i} f(x)dx$  soient tous égaux à  $1/m$ , il vient

$$\begin{aligned} V_1 &= \frac{1}{n} \text{Var}_f \left( \frac{g(X)}{f(X)} \right) \\ &= \frac{m}{n} \sum_{i=1}^m \sigma_i^2 + \frac{1}{n} \left( m \sum_{i=1}^m I_i^2 - I^2 \right) \\ &= V_2 + \frac{1}{n} \left( m \sum_{i=1}^m I_i^2 - I^2 \right). \end{aligned}$$

Mais, sur  $\mathbb{R}^m$ , on a  $\|x\|_1 \leq \sqrt{m}\|x\|_2$ , et donc  $m \sum_{i=1}^m I_i^2 - I^2 \geq 0$ . La variance de l'estimateur avec échantillonnage stratifié est plus petite que celle de l'estimateur de départ.

**Échantillonnage stratifié :** Afin de diminuer la variance par échantillonnage stratifié, pour une densité fixée  $f$  : on partitionne le support de  $g$  en domaines  $(D_i)_{1 \leq i \leq m}$  de même masse pour  $f$ , puis on estime séparément l'intégrale de  $g$  sur chacun des  $D_i$ . Dans l'idéal, le nombre  $n_i$  de simulations pour le morceau  $D_i$  doit être proportionnel à  $\sigma_i^2 = \text{Var}_{f_i} \left( \frac{g(X_1^i)}{f_i(X_1^i)} \right)$  : on pourra estimer grossièrement ces variances afin d'avoir une idée du choix des  $n_i$ .

L'idée est la même que pour l'échantillonnage préférentiel, mais pour une densité fixée : pour diminuer la variance, il faut faire davantage de simulations dans les domaines  $D_i$  où la variance  $\sigma_i^2 = \text{Var}_{f_i} \left( \frac{g(X_1^i)}{f_i(X_1^i)} \right)$  est grande.

#### (4) Calcul approché de $\frac{\pi}{4}$ : échantillonnage stratifié

On reprend pour  $f$  la densité de la loi uniforme sur  $[0, 1]$ , et  $g(x) = \sqrt{1-x^2}$ . On fait une stratification en  $m = 2$  morceaux :

$$D_1 = [0, 1/2]$$

$$D_2 = ]1/2, 1]$$

$$f_1 = 2\mathbb{1}_{D_1}$$

$$f_2 = 2\mathbb{1}_{D_2}$$

$$I_1 = \int_{D_1} g = \frac{\pi}{12} + \frac{\sqrt{3}}{8} = 0.4783 \quad I_2 = \int_{D_2} g = \frac{\pi}{6} - \frac{\sqrt{3}}{8} = 0.3071$$

$$= \int_{D_1} \frac{g^2}{f_1}(x)dx = \frac{11}{48} = 0.2292 \quad \int_{D_2} \frac{g^2}{f_2}(x)dx = \frac{5}{48} = 0.1042$$

$$\sigma_1^2 = \text{Var}_{f_1}(g/f_1)$$

$$\sigma_2^2 = \text{Var}_{f_2}(g/f_2)$$

$$= \int_{D_1} \frac{g^2}{f_1}(x)dx - I_1^2 = 0.0004 \quad = \int_{D_2} \frac{g^2}{f_2}(x)dx - I_2^2 = 0.0099.$$

#### Méthode de Monte Carlo (4) - Échantillonnage stratifié

On va construire un estimateur pour  $I_1$  avec  $n_1$  simulations et  $f_1$  comme densité et un estimateur pour  $I_2$  avec  $n_2$  simulations et densité  $f_2$ .

L'estimateur pour  $I$  avec  $n = n_1 + n_2$  simulations est

$$\hat{I}_n^{(4)} = \hat{I}_{n_1}^{1, strat} + \hat{I}_{n_2}^{2, strat}$$

où l'estimateur de  $I_1$  est

$$\hat{I}_{n_1}^{1, strat} = \frac{1}{n_1} \sum_{j=1}^{n_1} \frac{g(X_j^1)}{f_1(X_j^1)},$$

et  $(X_j^1)_{j \geq 1}$  est une suite de v.a.i.i.d. de densité  $f_1$ , et, pour  $I_2$ ,

$$\hat{I}_{n_2}^{2, strat} = \frac{1}{n_2} \sum_{j=1}^{n_2} \frac{g(X_j^2)}{f_2(X_j^2)},$$

où  $(X_j^2)_{j \geq 1}$  est une suite de v.a.i.i.d. de densité  $f_2$ .

Evaluons le nombre de simulations  $n_1$  et  $n_2$  pour  $n$  fixé.

Comme, pour  $i = 1, 2$ , la variance de l'estimateur de  $I_i$ ,  $\hat{I}^{i, strat}$  est notée par  $\sigma_i^2$  et donnée par

$$\begin{aligned}\sigma_1^2 &= \text{Var}_{f_1} \left( \frac{g(X_1^1)}{f_1(X_1^1)} \right) = \int_{D_1} \frac{g^2(x)}{f_1(x)} dx - I_1^2 = \int_0^{\frac{1}{2}} \frac{1-x^2}{2} dx - I_1^2 = 0,0004, \\ \sigma_2^2 &= \text{Var}_{f_2} \left( \frac{g(X_1^2)}{f_2(X_1^2)} \right) = \int_{D_2} \frac{g^2(x)}{f_2(x)} dx - I_2^2 = \int_{\frac{1}{2}}^1 \frac{1-x^2}{2} dx - I_2^2 = 0,0099.\end{aligned}$$

Pour optimiser la stratification, on doit donc répartir les  $n$  simulations en

$$n_1 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} n = 0,0381n, \quad n_2 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} n = 0,9691n.$$

**Remarque** Pour  $n = 10^4$  on obtient  $n_1 = 381$  et  $n_2 = 9691$ . Ce qui montre qu'il faut faire beaucoup plus de simulations sur  $D_2$  pour réduire la variance.

On obtient une réduction de variance (par rapport à (2)) de l'ordre de

$$V_1 - V_2 = \frac{1}{n} (2(I_1^2 + I_2^2) - I^2), \text{ d'où } C^{(4)} = 0.0207.$$

### Comparaison

On a donc les valeurs ordonnées :

$$C^{(1)} = 0,1685 > C^{(2)} = 0,0498 > C^{(4)} = 0,0207 > C^{(3)} = 0,0099.$$

On voit que **l'échantillonnage préférentiel (la méthode (3))** est jusqu'à présent la meilleure méthode.

### 2.10.3 Variable de contrôle

L'idée de la méthode par variable de contrôle est la suivante : on veut estimer une intégrale  $I = \mathbb{E}(Y)$ , et on connaît une v.a.  $C$  de moyenne  $\mu$  corrélée à  $Y$ , appelée *variable de contrôle*. Pour  $\beta > 0$ , on pose

$$Y(\beta) = Y - \beta(C - \mu),$$

de sorte que  $\mathbb{E}(Y(\beta)) = \mathbb{E}(Y) = I$ .

On estime alors  $I$  par la moyenne empirique de v.a.i.i.d. de même loi que  $Y(\beta)$ , et la variance pour  $n$  simulations est alors  $\frac{1}{n} \text{Var}(Y(\beta))$ , avec

$$\text{Var}(Y(\beta)) = \text{Var}(Y - \beta(C - \mu)) = \text{Var}(Y) + \beta^2 \text{Var}(C) - 2\beta \text{Cov}(Y, C), \quad (2.2)$$

qui peut être inférieure à  $\text{Var}(Y)$  si  $\beta$  est bien choisi.

**Lemme 1.** *Le minimum en la variable  $\beta$  de la fonction  $\text{Var}(Y(\beta))$  donnée en (2.2) est atteint en*

$$\beta^* = \frac{\text{Cov}(Y, C)}{\text{Var}(C)}$$

*et la variance minimale est*

$$\text{Var}(Y(\beta^*)) = (1 - \rho_{Y,C}^2) \text{Var}(Y).$$

**Remarque 5.** *Donc plus  $C$  est corrélée à  $Y$ , plus la réduction de variance est importante.*

**Variable de contrôle :** On construit une v.a.  $C$ , de moyenne  $\mu$  connue, et corrélée à  $Y$

$$Y(\beta) = Y - \beta(C - \mu), \quad \beta^* = \frac{\text{Cov}(Y, C)}{\text{Var}(C)}, \quad \text{Var}(Y(\beta^*)) = (1 - \rho_{Y,C}^2) \text{Var}(Y).$$

**En pratique :** La méthode présente deux difficultés. Premièrement, on ne connaît pas forcément une variable de contrôle  $C$ , c'est-à-dire une v.a. de moyenne connue et corrélée à  $Y$ . Deuxièmement, même si on connaît une variable de contrôle  $C$ , la valeur optimale  $\beta^*$  dépend de la covariance entre  $C$  et  $Y$  qui est souvent inconnue et par conséquent doit être estimée.

**Remarque 6.** *Cette méthode se généralise au cas d'un vecteur de contrôle  $C = (C_1, C_2, \dots, C_m)^t$  de moyenne  $\mu \in \mathbb{R}^m$ . Pour  $\beta \in \mathbb{R}^m$ , on pose  $Y(\beta) = Y - \beta^t(C - \mu)$ , et*

$$\text{Var}(Y(\beta)) = \text{Var}(Y) + \beta^t \Gamma_c \beta - 2 \text{Cov}(Y, C)^t \beta,$$

*où  $\Gamma_c$  est la matrice de covariance de  $C$ , et  $\text{Cov}(Y, C)$  est le vecteur dont les composantes sont les  $\text{Cov}(Y, C_i)$ . Le choix optimal pour  $\beta$  est alors :*

$$\beta^* = \Gamma_c^{-1} \text{Cov}(Y, C) \text{ et } \text{Var}(Y(\beta^*)) = (1 - R_{Y,C}^2) \text{Var}(Y),$$

$$\text{où } R_{Y,C} = \frac{\text{Cov}(Y, C)^t \Gamma_c^{-1} \text{Cov}(Y, C)}{\text{Var}(Y)}.$$

#### (5) Calcul approché de $\frac{\pi}{4}$ : Variable de contrôle

Considérons  $U$  de densité uniforme,  $Y = \sqrt{1 - U^2}$  et comme variable de contrôle  $C = (1 - U)^2$ . En faisant les calculs nous obtenons :

$$\mathbb{E}(Y) = \frac{\pi}{4}, \quad \text{Var}(Y) = \frac{2}{3} - \left(\frac{\pi}{4}\right)^2 = 0.0498, \quad \mathbb{E}(C) = \frac{1}{3}, \quad \text{Var}(C) = \frac{4}{45}, \quad \rho_{Y,C} = 0.8006.$$

Si on prend  $\beta^* = \frac{\text{Cov}(C, Y)}{\text{Var}(C)} = 0.5996$ , on trouve  $\text{Var}(Y(\beta^*)) = 0.0179$ , d'où

$$C^{(5)} = 0.0179.$$



### 2.10.4 Variables antithétiques

Idée de la méthode de variables antithétiques : nous supposons qu'il existe deux variables aléatoires  $Y$  et  $Z$  telles que  $I = \mathbb{E}(Y) = \mathbb{E}(Z)$ , avec  $I$  l'intégrale à estimer. On remarque que :

$$I = \mathbb{E} \left( \frac{1}{2}(Y + Z) \right)$$

et

$$\text{Var} \left[ \frac{1}{2}(Y + Z) \right] = \frac{1}{4} \text{Var}(Y) + \frac{1}{4} \text{Var}(Z) + \frac{1}{2} \text{Cov}(Y, Z).$$

Si  $Y$  et  $Z$  sont négativement corrélées, alors on peut espérer faire diminuer la variance.

Supposons par exemple qu'on doit estimer  $I = \int_0^1 g(x)dx$ . Soit  $(U_i)_{i \geq 1}$  une suite de v.a.i.i.d. de loi uniforme sur  $[0, 1]$ . L'estimateur de la moyenne empirique  $\frac{1}{n} \sum_{i=1}^n g(U_i)$  pour  $n$  simulations a pour variance  $\frac{1}{n} V_1$  avec

$$V_1 = \text{Var}(g(U)).$$

Comme  $x \rightarrow 1 - x$  laisse la mesure  $dx$  invariante sur  $[0, 1]$ , on peut prendre comme variables antithétiques  $Y = g(U)$  et  $Z = g(1 - U)$  avec  $U$  de loi uniforme sur  $[0, 1]$ . L'estimateur pour  $n$  simulations s'écrit :

$$\hat{I}_n^a = \frac{1}{2n} \sum_{i=1}^n (g(U_i) + g(1 - U_i)),$$

et a une variance  $\frac{1}{n} V_2$ , avec

$$V_2 = \frac{1}{2} (\text{Var}(g(U)) + \text{Cov}(g(U), g(1 - U))).$$

Mais attention, si le nombre de simulations est le même, le temps de calcul est deux fois plus important dans le second cas, donc pour que la méthode des variables antithétiques soit intéressante, il faut que  $V_2 < \frac{1}{2} V_1$ .

**Proposition 1.** *Si la fonction  $g : [0, 1] \mapsto \mathbb{R}$  est de classe  $C^1$  et monotone avec  $g(0) \neq g(1)$ , alors  $V_2 < \frac{1}{2} V_1$ .*

**Démonstration :** On suppose  $g$  croissante, donc  $g(1) > g(0)$ . On doit montrer que  $\text{Cov}(g(U), g(1 - U)) < 0$ , autrement dit que

$$\int_0^1 g(u)g(1 - u)du < I^2.$$

On pose  $\phi(x) = \int_0^x g(1-t)dt - xI$ . Alors  $\phi(0) = \phi(1) = 0$  et

$$\phi'(x) = g(1-x) - I,$$

est une fonction décroissante. Le théorème de la moyenne assure que  $\phi'(0) > 0$  et  $\phi'(1) < 0$ , et donc que  $\phi(x) > 0$  pour tout  $x \in ]0, 1[$ . Ce qui conduit à

$$0 < \int_0^1 \phi(x)g'(x)dx = - \int_0^1 \phi'(x)g(x)dx = - \int_0^1 g(u)g(1-u)du + I^2.$$

Le cas décroissant se traite de la même manière. ■

**Variables antithétiques :** Pour une fonction  $g$  monotone sur  $[0, 1]$ , si  $(U_i)_{i \geq 1}$  est une suite de v.a.i.i.d. de loi uniforme sur  $[0, 1]$ , l'estimateur avec variables antithétiques

$$\hat{I}_n^a = \frac{1}{2n} \sum_{i=1}^n (g(U_i) + g(1 - U_i))$$

est plus efficace (en terme de réduction de variance) que l'estimateur  $\frac{1}{n} \sum_{i=1}^n g(U_i)$ .

## (6) Calcul de $\frac{\pi}{4}$ : Variables antithétiques

Appliquer cette méthode au calcul de  $\pi/4$ .

## 2.11 Exercices

**Exercice 1.** Implémenter les méthodes de la technique de hit or miss et celle de la moyenne empirique. Donner des intervalles de confiance à l'aide du TCL.

**Exercice 2.** Dans cet exercice nous cherchons à approcher l'intégrale suivante :

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

Nous allons proposer une méthode de Monte Carlo ainsi que diverses techniques de réduction de variance pour évaluer cette intégrale. Pour chaque méthode il est demandé d'écrire le code **Scilab** correspondant et de donner des résultats numériques de ce code.

1. Exprimer cette intégrale sous la forme  $\mathbb{E}[f(X)]$  avec  $f$  une fonction à préciser et  $X$  une variable aléatoire suivant une loi uniforme sur  $[0, 1]$ . Donner la valeur exacte de  $I$ . Proposer une méthode de Monte Carlo pour ce calcul et justifier son application.
2. Calculer  $\text{Var}[f(X)]$  et préciser la variance de l'estimateur Monte Carlo associé.
3. Nous cherchons à approcher  $f$  par un polynôme de degré 2. Construire une fonction paire, positive sur  $[0, 1]$ , qui vaut 0 en 1 et 1 en 0. Trouver une densité  $g : [0, 1] \rightarrow \mathbb{R}_+$  à partir de cette fonction.
4. Si  $Y$  est une variable aléatoire de densité  $g$  notons par  $Z$  la variable aléatoire définie par

$$Z = \frac{2 \cos\left(\frac{\pi Y}{2}\right)}{3 - Y^2}.$$

Montrer que  $\mathbb{E}(Z) = \mathbb{E}(f(X))$ . Calculer  $\text{Var}(Z)$ .

5. Proposer une nouvelle méthode de Monte Carlo pour le calcul de  $I$  en utilisant le point précédent. En admettant que  $\text{Var}(Z) = 0.001$ , commenter la réduction de variance en comparant les deux estimateurs de Monte Carlo. Préciser l'amélioration engendrée sur la taille de l'intervalle de confiance.
6. Calculer la fonction de répartition  $F_Y$  de  $Y$ .
7. Soit  $H : [0, 1] \rightarrow [0, 1]$  définie par  $H(x) = 2 \cos\left(\frac{1}{3} \arccos(-x) + \frac{4\pi}{3}\right)$ . Vérifier que  $F_Y \circ H(x) = H \circ F_Y(x) = x$  pour tout  $x \in [0, 1]$ . Écrire la fonction inverse  $F_Y^{-1}$ . (Indication :  $\cos 3x = 4 \cos^3 x - 3 \cos x, \forall x \in \mathbb{R}$ .)
8. Quelle est la loi de la variable aléatoire  $F_Y^{-1}(X)$ ? En utilisant cette remarque proposer une méthode de simulation pour  $Y$ .

**Exercice 3.** Dans cet exercice nous cherchons à approcher numériquement par des méthodes probabilistes l'intégrale suivante :

$$I = \int_0^4 (x^3 + x)e^{-x^2} dx.$$

Nous allons proposer une méthode de Monte Carlo ainsi que diverses techniques de réduction de variance pour évaluer cette intégrale. Pour chaque méthode il est demandé d'écrire le code **Scilab** correspondant et de donner des résultats numériques de ce code.

1. Exprimer  $I$  sous la forme d'une espérance,  $I = \mathbb{E}(h(X))$  avec  $h$  une fonction et  $X$  une variable aléatoire de loi  $f$ , à préciser.
2. Proposer une méthode de Monte Carlo pour ce calcul et justifier son application. Calculer la valeur exacte de  $I$  et  $\text{Var}[h(X)]$ .
3. Proposer une nouvelle méthode de Monte Carlo pour le calcul de  $I$  en utilisant l'échantillonnage préférentiel. Calculer la variance de ce nouvel estimateur et commenter la réduction de variance par rapport à la méthode standard, proposée en 2.
4. Proposer une méthode de réduction de variance par échantillonnage stratifié. Calculer la variance associée. (Vous pouvez par exemple prendre  $D_1 = [0, 2]$  et  $D_2 = ]2, 4]$  si vous avez fait le choix d'une variable aléatoire à densité sur  $[0, 4]$ .)
5. Soit  $Y$  une variable aléatoire de densité  $g(x) = \frac{1}{4}(1 + x)$  sur  $[0, 2]$ . Notons par  $Z$  la variable aléatoire

$$Z = 2e^{-Y}.$$

Montrer que  $\mathbb{E}(Z) = I$ . Proposer une méthode de Monte Carlo en utilisant la variable  $Z$ . Quelle méthode de réduction de variance utilise-t-on ? Est-ce qu'on réduit réellement la variance avec cette méthode ? Justifier.

6. Classer ces méthodes dans l'ordre décroissant. Commenter votre choix de méthode par rapport à la réduction de variance.

# Chapitre 3

## Processus de renouvellement

Nous allons introduire dans ce chapitre les notions de processus de renouvellement et de processus de renouvellement avec récompense ainsi que leurs propriétés. Commençons par introduire deux exemples.

**Exemple 1.** Dans une machine, la durée de vie d'un certain composant est modélisée par une variable aléatoire  $X$  à valeurs dans  $\mathbb{R}_+$ . On remplace ce composant dès qu'il est en panne, et on peut se demander combien de fois on va devoir le remplacer pendant les 10 prochaines années.

**Exemple 2.** Des clients se présentent à un bureau de poste, à des instants d'arrivée qu'on suppose aléatoires. Peut-on étudier la longueur de la file d'attente ? Combien de guichets doit-on ouvrir pour optimiser le service ?

### 3.1 Définitions

Pour modéliser l'évolution au cours du temps d'une certaine quantité aléatoire nous introduisons la notion de processus stochastique.

**Définition 1.** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité. Un *processus stochastique*  $(Y_t)_{t \geq 0}$  sur cet espace est une famille de variables aléatoires à valeurs dans  $\mathbb{R}$ , indexée par le temps  $t \in \mathbb{R}_+$ .

Pour un  $\omega \in \Omega$  donné,  $t \mapsto Y_t(\omega)$  représente l'évolution d'une quantité au cours du temps. Autrement dit,  $\omega \mapsto (t \mapsto Y_t(\omega))$  est une variable aléatoire, à valeurs dans l'ensemble des fonction de  $\mathbb{R}_+$  dans  $\mathbb{R}$ .

À titre d'exemple,  $Y_t$  peut décrire la température au temps  $t$ , ou le prix d'une action au temps  $t$ , ou le nombre de gens qui se sont présentés au guichet avant  $t$ .

Soit  $(X_i)_{i \in \mathbb{N}^*}$  une suite de v.a.i.i.d. à valeurs dans  $\mathbb{R}_+^*$ . On définit

$$S_0 = 0 \text{ et pour } n \geq 1, S_n = \sum_{i=1}^n X_i.$$

**Définition 2.** On définit le *processus de renouvellement*  $(N_t)_{t \geq 0}$  associé aux temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$  en posant :  $N_0 = 0$  et pour  $t > 0$ ,

$$N_t = \sum_{i=1}^{+\infty} \mathbb{1}_{\{S_i \leq t\}}.$$

Nous donnons dans la remarque suivante les premières propriétés du processus de renouvellement  $(N_t)_{t \geq 0}$ .

**Remarque 1.**

1. Pour tout  $\omega \in \Omega$ ,  $t \mapsto N_t(\omega)$  est une fonction croissante de  $\mathbb{R}_+$  dans  $\mathbb{N}$ . On dit que  $(N_t)_{t \geq 0}$  est un processus croissant.
2. Comme  $X_1$  est à valeurs dans  $\mathbb{R}_+^*$ ,  $\mathbb{P}(X_1 = 0) = 0$ , et donc il n'y a pas d'arrivées simultanées. En particulier, la suite  $(S_n)_{n \geq 1}$  des instants de renouvellement est strictement croissante et le processus de renouvellement  $(N_t)_{t \geq 0}$  ne fait que des sauts de hauteur 1.
3. Comme  $X_1$  est à valeurs dans  $\mathbb{R}_+$ , soit elle est intégrable et  $\mathbb{E}(X_1) \in [0, +\infty[$ , soit elle n'est pas intégrable, et alors on pose  $\mathbb{E}(X_1) = +\infty$ .

Revenons aux exemples 1 et 2 et explicitons les notions pour ces cas précis.

Pour l'exemple 1, si  $X_i$  représente la durée de vie du composant  $i$  (celui en fonctionnement après  $i - 1$  remplacements), alors  $S_n$  représente l'instant de la  $n$ -ième panne ou l'instant du  $n$ -ième renouvellement, et  $N_t$  le nombre de composants utilisés durant l'intervalle  $[0, t]$ .

Dans l'exemple 2, si  $X_i$  représente le laps de temps qui s'écoule entre l'arrivée du client  $i - 1$  et celle du client  $i$ ,  $S_n$  représente l'instant d'arrivée du client  $n$ , et  $N_t$  le nombre de clients qui se sont présentés durant l'intervalle  $[0, t]$ .

**Remarque 2.** Il est équivalent de connaître le processus de renouvellement  $(N_t)_{t \geq 0}$ , la suite  $(X_i)_{i \in \mathbb{N}^*}$  des temps inter-arrivées, ou la suite  $(S_n)_{n \geq 1}$  des instants de renouvellement. En effet,

1. Si on connaît les temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$ , on définit les instants de renouvellement  $(S_n)_{n \geq 1}$ , puis le processus de renouvellement  $(N_t)_{t \geq 0}$ .
2. Si on connaît les instants de renouvellement  $(S_n)_{n \geq 1}$ , on peut construire le processus de renouvellement  $(N_t)_{t \geq 0}$ , et on retrouve les temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$  en remarquant que

$$X_n = S_n - S_{n-1}.$$

3. Si on connaît le processus de renouvellement  $(N_t)_{t \geq 0}$ , on retrouve les instants de renouvellement  $(S_n)_{n \geq 1}$  en remarquant que

$$S_n = \inf\{t \geq 0 : N_t \geq n\}.$$

## 3.2 Propriétés trajectorielles

Nous démontrons dans cette section un théorème qui décrit le comportement en temps long d'un processus de renouvellement.

**Théorème 1.** Soit  $(N_t)_{t \geq 0}$  un processus de renouvellement associé aux temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$  et aux instants de renouvellement  $(S_n)_{n \geq 1}$ . Supposons que  $\mathbb{E}(X_1) > 0$ , alors

- (i) Presque-sûrement,  $\lim_{n \rightarrow +\infty} \frac{S_n}{n} = \mathbb{E}(X_1)$ .
- (ii) Presque-sûrement, pour tout  $t \geq 0$ ,  $N_t < +\infty$ .
- (iii) Presque-sûrement,  $\lim_{t \rightarrow +\infty} N_t = +\infty$ .
- (iv) Presque-sûrement,  $\lim_{t \rightarrow +\infty} \frac{N_t}{t} = \frac{1}{\mathbb{E}(X_1)}$ .

**Démonstration :**

- (i) On applique simplement une version de la loi forte des grands nombres :

**Théorème 2.** Soient  $(X_i)_{i \in \mathbb{N}^*}$  des v.a.i.i.d. à valeurs dans  $\mathbb{R}_+$ . Alors, avec probabilité 1,

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n X_i = \mathbb{E}(X_1).$$

Ce théorème est vrai même si  $\mathbb{E}(X_1) = +\infty$ .

(ii) D'après le point (i), comme  $\mathbb{E}(X_1) > 0$ , il existe une partie  $\tilde{\Omega} \subset \Omega$ , de probabilité 1 telle que pour tout  $\omega \in \tilde{\Omega}$ ,

$$\lim_{n \rightarrow +\infty} S_n(\omega) = +\infty.$$

Soit  $\omega \in \tilde{\Omega}$  fixé,  $t > 0$  fixé. Il existe  $n \geq 1$  tel que  $S_n(\omega) \geq t \iff N_t(\omega) \leq n$ . D'où le résultat annoncé.

(iii) Soit  $M \in \mathbb{N}^*$  fixé. On remarque que  $N_{S_M} = M$  (voir Exercices), et comme  $t \rightarrow N_t$  est croissant, pour tout  $t \geq S_M$ ,  $N_t \geq M$ . Ce qui nous conduit au résultat.

(iv) Soit  $t > 0$  fixé. On a toujours  $S_{N_t} \leq t < S_{N_t+1}$ , d'où, si  $N_t > 0$ ,

$$\frac{S_{N_t}}{N_t} \leq \frac{t}{N_t} \leq \frac{S_{N_t+1}}{N_t+1} \frac{N_t+1}{N_t}.$$

Comme  $N_t$  tend vers  $+\infty$ , le point (i) assure que

$$\lim_{t \rightarrow +\infty} \frac{S_{N_t}}{N_t} = \lim_{t \rightarrow +\infty} \frac{S_{N_t+1}}{N_t+1} = \mathbb{E}(X_1),$$

et le point (iii) assure que  $\lim_{t \rightarrow +\infty} \frac{N_t+1}{N_t} = 1$ . En mettant ensemble ces résultats nous obtenons la limite souhaitée. ■

### 3.3 Processus de Renouvellement avec Récompense (PRR)

Soit maintenant une suite  $(X_i, Z_i)_{i \in \mathbb{N}^*}$  i.i.d. à valeurs dans  $\mathbb{R}_+^* \times \mathbb{R}$ .

**Définition 3.** On appelle *processus de renouvellement avec récompense* le processus

$$R_t = \sum_{i=1}^{N_t} Z_i,$$

où  $N_t$  est le processus de renouvellement associé aux temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$ .

Reprenons les exemples 1 et 2. Dans ces cas :

1. Dans l'exemple 2,  $Z_n$  peut représenter le temps de service du client numéro  $n$  dans la file d'attente, et alors  $R_t$  représente le temps total de service avant l'instant  $t$ .



2. Dans l'exemple 1,  $Z_1$  peut représenter le coût de remplacement d'un composant, et alors  $R_t$  représente le coût d'entretien du système jusqu'au temps  $t$ .

**Remarque 3.** *Attention, on suppose que les  $(X_i, Z_i)_{i \in \mathbb{N}^*}$  sont indépendants, mais par contre on ne suppose pas que  $X_1$  et  $Z_1$  sont indépendants.*

**Remarque 4.** *Si  $Z_i = 1$ , le processus de récompense coïncide avec le processus de renouvellement.*

Nous pouvons décrire le comportement du processus de renouvellement avec récompense en temps grand, par la proposition suivante.

**Proposition 1.** *Soit  $(X_i, Z_i)_{i \in \mathbb{N}^*}$  i.i.d. à valeurs dans  $\mathbb{R}_+^* \times \mathbb{R}$ . On suppose que  $\mathbb{E}(X_1) < +\infty$  et que  $\mathbb{E}(|Z_1|) < +\infty$ . Alors, presque sûrement,*

$$\lim_{t \rightarrow +\infty} \frac{R_t}{t} = \frac{\mathbb{E}(Z_1)}{\mathbb{E}(X_1)}.$$

**Démonstration :** On écrit simplement  $\frac{R_t}{t} = \frac{R_t}{N_t} \frac{N_t}{t}$ . On a déjà vu que, presque sûrement,

$$\lim_{t \rightarrow +\infty} \frac{N_t}{t} = \frac{1}{\mathbb{E}(X_1)}.$$

Maintenant,  $\frac{R_t}{N_t} = \frac{1}{N_t} \sum_{i=1}^{N_t} Z_i$  et comme  $N_t \rightarrow +\infty$ , on peut appliquer encore une fois la loi des grands nombres pour montrer que, p.s.

$$\lim_{t \rightarrow +\infty} \frac{R_t}{N_t} = \mathbb{E}(Z_1),$$

ce qui termine la preuve. ■

### 3.4 Théorème Central Limite

Nous introduisons dans cette partie un résultat de type théorème central limite pour le processus de renouvellement.

**Théorème 3.** *Soit  $(X_i)_{i \in \mathbb{N}^*}$  une suite de v.a.i.i.d. à valeurs dans  $\mathbb{R}_+^*$ . On suppose que  $\mathbb{E}(X_1^2) < +\infty$  et on note  $\mu = \mathbb{E}(X_1)$  et  $\sigma^2 = \text{Var}(X_1)$ . On note  $(N_t)_{t \geq 0}$  le processus de renouvellement associé aux temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$ . Alors on a la convergence en loi suivante :*

$$\lim_{t \rightarrow +\infty} \frac{\mu^{3/2} \sqrt{t}}{\sigma} \left( \frac{N_t}{t} - \frac{1}{\mu} \right) = G,$$

où  $G$  note une v.a. normale  $\mathcal{N}(0, 1)$ .

**Démonstration :** Soit  $x \in \mathbb{R}$  :

$$\begin{aligned}\mathbb{P}\left(\frac{\mu^{3/2}\sqrt{t}}{\sigma}\left(\frac{N_t}{t} - \frac{1}{\mu}\right) < x\right) &= \mathbb{P}\left(N_t < \frac{t}{\mu} + \frac{x\sigma\sqrt{t}}{\mu^{3/2}}\right) \\ &= \mathbb{P}(S_{k_t} > t),\end{aligned}$$

où  $k_t = \left\lceil \frac{t}{\mu} + \frac{x\sigma\sqrt{t}}{\mu^{3/2}} \right\rceil$ . En utilisant le théorème central limite pour la suite  $(X_i)_{i \in \mathbb{N}^*}$ , on a la convergence en loi suivante :

$$\lim_{n \rightarrow +\infty} \frac{\sqrt{n}}{\sigma} \left( \frac{S_n}{n} - \mu \right) = G, \text{ où } G \sim \mathcal{N}(0, 1).$$

Comme  $k_t$  tend vers  $+\infty$  quand  $t$  tend vers  $+\infty$ , on a

$$\begin{aligned}\mathbb{P}(S_{k_t} > t) &= \mathbb{P}\left(\frac{\sqrt{k_t}}{\sigma} \left( \frac{S_{k_t}}{k_t} - \mu \right) > \frac{\sqrt{k_t}}{\sigma} \left( \frac{t}{k_t} - \mu \right)\right) \\ &\sim 1 - \Phi(\alpha_t) = \Phi(-\alpha_t),\end{aligned}$$

où  $\alpha_t = \frac{\sqrt{k_t}}{\sigma} \left( \frac{t}{k_t} - \mu \right)$ . Il suffit donc de voir que lorsque  $t$  tend vers  $+\infty$ ,  $\alpha_t$  tend vers  $-x$ . Pour  $t$  assez grand on déduit que

$$\begin{aligned}k_t &= \frac{t}{\mu} + \frac{x\sigma\sqrt{t}}{\mu^{3/2}} + o(\sqrt{t}) \\ \frac{k_t}{t} &= \frac{1}{\mu} + \frac{x\sigma}{\mu^{3/2}\sqrt{t}} + o\left(\frac{1}{\sqrt{t}}\right) \\ \frac{t}{k_t} &= \mu \left( \frac{1}{1 + \frac{x\sigma}{\sqrt{t}\mu} + o\left(\frac{1}{\sqrt{t}}\right)} \right) = \mu \left( 1 - \frac{x\sigma}{\sqrt{t}\mu} + o\left(\frac{1}{\sqrt{t}}\right) \right).\end{aligned}$$

Donc, nous obtenons d'une part

$$\frac{t}{k_t} - \mu \sim -x\sigma\sqrt{\frac{\mu}{t}}$$

et d'autre part

$$\frac{\sqrt{k_t}}{\sigma} \sim \frac{1}{\sigma}\sqrt{\frac{t}{\mu}}.$$

En reprenant la définition de  $\alpha_t$  on trouve le résultat. ■

### 3.5 Fonction de renouvellement

La variable  $N_t$ , qui représente le nombre de renouvellements dans l'intervalle de temps  $[0, t]$ , est une variable aléatoire dont l'espérance mathématique est un indicateur important. Cette espérance est appelée fonction de renouvellement et est notée :

$$M(t) = \mathbb{E}(N_t).$$

**Définition 4.** Soit  $(N_t)_{t \geq 0}$  un processus de renouvellement associé aux temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$ . On appelle *fonction de renouvellement* et on la note  $M(t)$ , l'espérance mathématique de la variable aléatoire  $N_t$  :

$$t \mapsto M(t) = \mathbb{E}(N_t).$$

Dans la proposition suivante nous donnons une représentation de  $M(t)$  en fonction de  $(S_n)_{n \geq 1}$ .

**Proposition 2.** Soit  $F_n$  la fonction de répartition de  $S_n$ , date du  $n^{\text{ième}}$  événement (instant de renouvellement). Alors, pour tout  $t \geq 0$  on a

$$M(t) = \sum_{n \geq 1} F_n(t) = \sum_{n \geq 1} \mathbb{P}(S_n \leq t) < \infty. \quad (3.1)$$

**Démonstration :** Par la définition de la fonction de renouvellement on a :

$$M(t) = \mathbb{E}(N_t) = \sum_{n \geq 1} n \mathbb{P}(N_t = n).$$

Calculons la loi de  $N_t$  en utilisant les  $F_n$ . Nous avons

$$\mathbb{P}(N_t = n) = \mathbb{P}(N_t \geq n) - \mathbb{P}(N_t \geq n + 1).$$

Nous remarquons que l'événement  $\{N_t \geq n\}$  est équivalent avec l'événement  
 $\{ \text{à la date } t, \text{ il y a eu au moins } n \text{ renouvellements} \}$

qui est lui même équivalent à l'événement

$\{ \text{la date du } n^{\text{ième}} \text{ renouvellement est inférieure à } t \}$ .

Ou encore, en utilisant les exercices, on déduit :

$$\mathbb{P}(N_t \geq n) = \mathbb{P}(S_n \leq t) = F_n(t), \quad (3.2)$$

d'où  $\mathbb{P}(N_t = n) = F_n(t) - F_{n+1}(t)$ . Ce qui nous conduit à :

$$M(t) = \sum_{n \geq 1} n(F_n(t) - F_{n+1}(t)) = \sum_{n \geq 1} F_n(t). \quad (3.3)$$

■

Nous pouvons aussi obtenir ce résultat avec la remarque suivante :

**Remarque 5.** En utilisant le résultat général pour une variable aléatoire discrète, si  $Y : \Omega \mapsto \mathbb{N}$  alors son espérance s'écrit :

$$\mathbb{E}(Y) = \sum_{n \geq 1} \mathbb{P}(Y \geq n),$$

nous obtenons en l'appliquant à  $N_t$  :

$$\mathbb{E}(N_t) = \sum_{n \geq 1} \mathbb{P}(N_t \geq n) = \sum_{n \geq 1} \mathbb{P}(S_n \leq t) = \sum_{n=1}^{+\infty} F_n(t).$$

### 3.6 Âge courant et âge résiduel

Nous abordons ici quelques notions définissant diverses notions concernant le temps.

**Définition 5.** L'âge courant à l'instant  $t$  est le temps qui s'écoule depuis le dernier renouvellement, il est égal à :

$$t - S_{N_t}.$$

L'âge résiduel à l'instant  $t$  est le temps qui reste avant le prochain renouvellement, il est égal à :

$$S_{N_t+1} - t.$$

La *vie totale* à l'instant  $t$  est le temps qui s'écoule entre le  $N_t^{\text{ième}}$  et le  $(N_t + 1)^{\text{ième}}$  renouvellements. C'est aussi la somme entre l'âge courant et l'âge résiduel, elle est égale à :

$$X_{N_t+1} = S_{N_t+1} - S_{N_t}.$$

Nous avons le résultat suivant :

**Proposition 3.** Soit  $(N_t)_{t \geq 0}$  un processus de renouvellement associé aux temps inter-arrivées  $(X_i)_{i \in \mathbb{N}^*}$ , et notons  $\mathbb{E}(X_1) = \mu$ . Alors :

$$\mathbb{E}(S_{N_t+1}) = \mu(M(t) + 1). \quad (3.4)$$

**Démonstration :** Calculons, en utilisant les définitions

$$\begin{aligned} \mathbb{E}(S_{N_t+1}) &= \mathbb{E}(X_1 + \dots + X_{N_t+1}) \\ &= \mathbb{E}(X_1) + \mathbb{E} \left[ \sum_{i=2}^{N_t} X_i \right] \\ &= \mu + \mathbb{E} \left[ \sum_{i=2}^{+\infty} X_i \mathbb{1}_{\{i \leq N_t+1\}} \right] \\ &= \mu + \sum_{i=2}^{+\infty} \mathbb{E} [X_i \mathbb{1}_{\{i \leq N_t+1\}}]. \end{aligned} \quad (3.5)$$

En utilisant l'équivalence des événements  $\{N_t \geq i-1\}$  et  $\{X_1 + \dots + X_{i-1} \leq t\}$  on peut écrire

$$\mathbb{1}_{\{N_t \geq i-1\}} = \mathbb{1}_{\{S_{i-1} \leq t\}} = \mathbb{1}_{\{X_1 + \dots + X_{i-1} \leq t\}} \quad (3.6)$$

et en revenant avec ce résultat dans les termes présents dans le calcul (3.5) on obtient, en utilisant l'indépendance et l'égalité en loi des  $(X_i)_{i \in \mathbb{N}^*}$  :

$$\begin{aligned} \mathbb{E} [X_i \mathbb{1}_{\{i \leq N_t + 1\}}] &= \mathbb{E} [X_i \mathbb{1}_{\{X_1 + \dots + X_{i-1} \leq t\}}] \\ &= \mathbb{E}(X_i) \mathbb{E} [\mathbb{1}_{\{X_1 + \dots + X_{i-1} \leq t\}}] \\ &= \mu \mathbb{P}(S_{i-1} \leq t) \\ &= \mu F_{i-1}(t). \end{aligned} \quad (3.7)$$

En utilisant la proposition 2 et ce résultat on déduit que (3.5) devient :

$$\begin{aligned} \mathbb{E}(S_{N_t+1}) &= \mu + \sum_{i=2}^{+\infty} \mu F_{i-1}(t) \\ &= \mu \left( 1 + \sum_{i=2}^{+\infty} F_{i-1}(t) \right) \\ &= \mu \left( 1 + \sum_{i=1}^{+\infty} F_i(t) \right) \\ &= \mu(1 + M(t)). \end{aligned} \quad (3.8)$$

Ce qui finit la preuve. ■

Nous présentons dans la proposition suivante une formule pour la fonction de renouvellement.

**Proposition 4.** *Pour tout  $t \geq 0$  nous avons  $M(t) < \infty$  et :*

$$\lim_{t \rightarrow +\infty} \frac{M(t)}{t} = \frac{1}{\mu}. \quad (3.9)$$

*En outre, si les temps inter-arrivées  $(X_i)_{i \geq 1}$  sont des v.a. à densité et si on note  $f$  leur densité, alors la fonction  $M$  est continue et vérifie l'équation de renouvellement :*

$$M(t) = \int_0^t (1 + M(t-s))f(s)ds, \forall t \geq 0. \quad (3.10)$$

**Démonstration :** On remarque que :

$$\mathbb{P}(S_i \leq t) = \mathbb{E}(\mathbb{1}_{\{S_i \leq t\}}) \leq \mathbb{E}[e^{t-S_i}] = e^t [\mathbb{E}(e^{-X_1})]^i. \quad (3.11)$$

Comme  $X_1$  est une v.a. strictement positive on a  $\mathbb{E}(e^{-X_1}) < 1$  ce qui conduit à

$$M(t) \leq e^t \sum_{i=1}^{+\infty} [\mathbb{E}(e^{-X_1})]^i < \infty.$$

Considérons maintenant la situation  $(X_i)_{i \geq 1}$  des v.a. à densité  $f$ . Pour  $i \geq 2$  on a :

$$\begin{aligned} \mathbb{P}(S_i \leq t) &= \mathbb{P}(X_1 + \dots + X_i \leq t) \\ &= \int_{t_1 + \dots + t_i \leq t} f(t_1) \dots f(t_i) dt_1 \dots dt_i \\ &= \int_0^t f(t_i) \left( \int_{t_1 + \dots + t_{i-1} \leq t - t_i} f(t_1) \dots f(t_{i-1}) dt_1 \dots dt_{i-1} \right) dt_i \\ &= \int_0^t f(s) \mathbb{P}(S_{i-1} \leq t - s) ds. \end{aligned}$$

Ce qui nous permet de réécrire les choses sous la forme :

$$\begin{aligned} M(t) &= \sum_{i=1}^{+\infty} \mathbb{P}(S_i \leq t) \\ &= \mathbb{P}(X_1 \leq t) + \sum_{i \geq 2} \int_0^t f(s) \mathbb{P}(S_{i-1} \leq t - s) ds \\ &= \int_0^t f(s) ds + \int_0^t f(s) \sum_{i \geq 1} \mathbb{P}(S_i \leq t - s) ds. \end{aligned}$$

Par ailleurs, on sait avec la proposition 2 que  $M(t) = \sum_{i \geq 1} \mathbb{P}(S_i \leq t)$ .

Ceci nous conduit au résultat recherché :

$$M(t) = \int_0^t (1 + M(t - s)) f(s) ds, \forall t \geq 0. \quad (3.12)$$

■

### 3.7 Exercices

**Exercice 1.** Soit  $(N_t)_{t \geq 0}$  un processus de renouvellement associé à  $(X_i)_{i \geq 1}$ , et  $S_n = X_1 + \dots + X_n$ . Montrer que  $N_{S_n} = n$ , et que  $S_{N_t} \leq t$ .

**Exercice 2.** Soit  $(N_t)_{t \geq 0}$  un processus de renouvellement associé à  $(X_i)_{i \geq 1}$ , et  $S_n = X_1 + \dots + X_n$ .

1) A-t-on

$$(i) \quad N_t = n \text{ ssi } S_n \leq t < S_{n+1} ?$$

$$(ii) \quad N_t < n \text{ ssi } S_n > t ?$$

$$(iii) \quad N_t \leq n \text{ ssi } S_n \geq t ?$$

$$(iv) \quad N_t > n \text{ ssi } S_n < t ?$$

2) Le temps vaut  $t$ . Exprimer le délai écoulé depuis la dernière arrivée. Exprimer le temps d'attente entre la dernière arrivée et la prochaine arrivée.

**Exercice 3.** Soit  $(N_t)_{t \geq 0}$  un processus de renouvellement associé à  $(X_n)_{n \geq 1}$ . Montrer que

$$\mathbb{E}(X_{N_t+1}) \leq \mathbb{E}(X_1)\mathbb{E}(N_t + 1).$$

Pourquoi n'est-il pas clair que  $\mathbb{E}(X_{N_t+1}) = \mathbb{E}(X_1)$  ?

**Exercice 4.** Donner un exemple de v.a. à valeurs dans  $\mathbb{N}$  d'espérance infinie, et une v.a. à densité à valeurs dans  $\mathbb{R}_+$  d'espérance infinie.

**Exercice 5.** Dans un avion, un composant est remplacé à coût  $c_2$  à chaque fois qu'il arrive à l'âge  $T$ , sauf s'il tombe en panne avant, auquel cas il est remplacé à coût  $c_1 > c_2$ . Les durées des vies des composants successifs sont des v.a.i.i.d.  $(Y_n)_{n \geq 1}$ , de densité  $f$  et de fonction de répartition  $F$ .

1. A l'aide d'un processus de renouvellement avec récompense associé à une suite de vecteurs i.i.d.  $(X_n, Z_n)$ , exprimer le coût moyen  $C_t(T)$  par unité de temps sur la période d'utilisation  $[0, t]$ .
2. Donner une valeur approchée de  $C_t(T)$  lorsque  $t$  est grand.
3. Dans le cas où  $Y_1 \sim \mathcal{E}(\lambda)$ , comment choisir  $T$  pour un coût minimal ? Commenter le résultat. [ $T = +\infty$ ]
4. Dans le cas où  $Y_1 \sim \mathcal{U}([0, 1])$ , comment choisir  $T$  pour un coût minimal ? [exprimer  $T$  en fonction de  $c_1$  et  $c_2$ ]

5. Supposons  $T$ ,  $c_1$ ,  $c_2$  donnés, et  $f(x) = 2x1_{[0,1]}(x)$ . Simuler  $C_{100}(T)$ . Comment faire un calcul approché de  $E[C_{100}(T)]$  ?

**Exercice 6.** La production d'une centrale électrique (en Mwh) sur l'intervalle  $[0, t]$  est donnée par

$$P_t = \int_0^t \phi(Y_s) ds$$

où  $Y_s$  est l'âge d'un composant primordial, et où  $\phi$  est une fonction décroissante de  $\mathbb{R}_+$  dans  $\mathbb{R}_+$ . Les durées de vie des composants successifs sont données par des v.a. positives  $(X_n)_n$  i.i.d. de densité  $f$ .

Le remplacement d'un composant coûte  $c$  \$, et chaque Mwh produit rapporte  $p$  \$.

Afin d'optimiser le profit, on décide de remplacer le composant à chaque fois qu'il dépasse l'âge  $A$ .

1. Soit  $Z_n$  la durée d'utilisation du composant numéro  $n$ . Exprimer  $P_n$ , la production de Mwh lors de l'utilisation du composant numéro  $n$ .
2. Exprimer le profit moyen par unité de temps  $\pi_t(A)$  réalisé sur la période  $[0, t]$ .
3. Donner une valeur approchée de  $\pi_t(A)$  lorsque  $t$  est grand.
4. Dans le cas d'un composant de très mauvaise qualité, on a  $\phi(\alpha) = e^{-\alpha}$  et  $X_1 \sim \mathcal{E}(1)$ . Comment choisir  $A$  ?
5. Dans un cas plus favorable on a  $\phi(\alpha) = 1$  et  $X_1 - 10 \sim \mathcal{E}(1)$ . Comment choisir  $A$  ?
6. On suppose que  $\phi(\alpha) = e^{-\alpha}$  et  $X_1 \sim \mathcal{E}(1)$ . Simuler  $\pi_{1000}(10)$ .

**Exercice 7.** Soit  $(X_n)_{n \geq 1}$  une suite de v.a.i.i.d. à valeurs dans  $\mathbb{R}_+^*$ . Notons par  $(N_t)_{t \geq 0}$  le processus de renouvellement associé aux temps inter-arrivées  $(X_n)_{n \geq 1}$ . Pour tout  $t \geq 0$  on note  $M(t) := \mathbb{E}(N_t)$  la fonction de renouvellement.

- (2) Considérons le cas où les variables  $X_n$  suivent la loi uniforme sur  $[0, 1]$ . En utilisant l'équation de renouvellement trouver une forme explicite de  $M(t)$  pour  $t \leq 1$ .
- (3) Supposons maintenant que les variables  $X_n$  suivent la loi gamma de paramètres  $(1, 2)$ , de densité  $t \exp(-t)$  sur  $\mathbb{R}_+$ .  
(3.1) En remarquant que la loi gamma de paramètre  $(1, 2)$  est aussi la loi de la somme de deux variables exponentielles indépendantes de paramètre 1, vérifier que

$$\mathbb{P}(N_t = n) = \mathbb{P}(R_t = 2n) + \mathbb{P}(R_t = 2n + 1),$$

où  $R_t$  note un processus de Poisson de paramètre 1.



**(3.2)** D  duire que

$$M(t) = \frac{\mathbb{E}(R_t)}{2} - \frac{1}{2} \sum_{n=0}^{+\infty} \mathbb{P}(R_t = 2n + 1).$$

**(3.3)** Montrer que  $M$  est de la forme  $M(t) = \frac{t}{2} + \frac{e^{-2t}-1}{4}$ . V  rifier ensuite que  $M$  est bien solution de l  quation de renouvellement.