

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



جامعة بجاية  
Tasdawit n Bgayet  
Université de Béjaïa

Université A/Mira de Béjaïa  
Faculté des Sciences Exactes  
Département de Recherche Opérationnelle

## *Mémoire de Master*

En

Mathématiques Appliquées

Spécialité

*Mathématiques Financières*

## Thème

Méthodes Numériques d'Estimation de la  
Probabilité de Ruine

Présenté par :

SMAILI Oussama Abdelkadir

Soutenu le 02 Décembre 2020 devant le jury composé de :

Présidente	Mme LEKADIR Ouiza	M.C.A	Université de Béjaïa.
Promoteurs	Mme HOCINE Safia	M.A.A	Université de Béjaïa.
	Mme ALEM Lala Maghnia	M.A.A	Université de Bouira.
Examineurs	Mme BOUCHEBBAH Kahina	M.A.A	Université de Béjaïa.
	Mme TKHEDMIT Baya	M.C.B	Université de Béjaïa.

Promotion 2019/2020.

## ※ *Remerciements* ※

Je remercie le bon Dieu (sans Lui rien n'aurait pu être possible) tout puissant de m'avoir accordé santé, courage et la volonté pour accomplir ce modeste travail.

Je tiens également à remercier ma Promotrice *M<sup>me</sup>* **S. HOCINE** pour ses précieux conseils, sa disponibilité et ses encouragements qui m'a poussé à donner le meilleur de moi-même tout au long de la préparation de ce mémoire.

Je profite en l'occurrence pour exprimer ma profonde gratitude et mes sincères remerciements à *M<sup>me</sup>* **L. M. ALEM**.

J'adresse mes remerciements à *M<sup>me</sup>* **O. LEKADIR** pour l'honneur qu'elle m'a fait en acceptant de présider le jury de ce mémoire.

Je remercie également *M<sup>me</sup>* **K. BOUCHEBBAH** et *M<sup>me</sup>* **B. TKHEDMIT** de m'avoir fait l'honneur d'accepter de faire partie du jury de ce mémoire.

Je remercier aussi mes parents et proches pour leurs aides et encouragements.

※ *Dédicaces* ※

Je dédie ce modeste travail :

A mon source de courage, mes très chers parents,

A toutes ma famille,

A tous mes professeurs,

A tous mes ami(e)s.

*M. Smaili Oussama Abdelkadir*

---

# TABLE DES MATIÈRES

<b>Introduction générale</b>	<b>1</b>
<b>1 La théorie de la ruine et modèles de risques classiques</b>	<b>3</b>
Introduction . . . . .	3
1.1 Le modèle individuel et le modèle collectif . . . . .	4
1.1.1 Modèle individuel . . . . .	4
1.1.2 Modèle collectif . . . . .	4
1.2 Modèle de risque de la théorie de la ruine . . . . .	5
1.2.1 Processus de réserve et de surplus . . . . .	5
1.2.2 Le processus de Lévy . . . . .	5
1.2.3 Probabilité de ruine . . . . .	6
1.2.4 Processus de renouvellement . . . . .	7
1.3 Modèle de Cramèr-Lundberg . . . . .	7
1.3.1 Cas particulier : Modèle de Lundberg P/P . . . . .	8
1.3.2 Le coefficient de sécurité . . . . .	9
1.3.3 Condition de non ruine . . . . .	9
1.4 Approximation de la probabilité de ruine . . . . .	10
1.4.1 Approximations Cramèr-Lundberg . . . . .	10
1.4.2 Approximations de Cramèr-Lundberg dans le modèle P/P . . . . .	11
1.4.3 Formule de Pollaczek-Khinchine . . . . .	12
1.4.4 Formule de Pollaczek-Khinchine dans le modèle P/P . . . . .	12
1.4.5 Autres approches . . . . .	14
1.5 Conclusion . . . . .	14
<b>2 La méthode de Monte Carlo et Échantillonnage préférentiel (Importance Sampling)</b>	<b>15</b>
2.1 Loi faible et loi forte des grands nombres . . . . .	16
2.1.1 Exemple d'application . . . . .	17

---

2.2	Vitesse de convergence et estimation de l'erreur Monte Carlo . . . . .	18
2.3	Échantillonnage préférentiel (Importance Sampling) . . . . .	23
2.3.1	Convergence de l'estimateur . . . . .	24
2.3.2	Variance de l'estimateur . . . . .	24
2.3.3	Choix de la densité instrumentale . . . . .	24
2.3.4	Diagnostic et taille efficace d'échantillon (effective sample size) .	26
2.4	Conclusion . . . . .	26
<b>3</b>	<b>Estimation de la probabilité de ruine par la méthode d'échantillonnage préférentiel (Importance Sampling)</b>	<b>27</b>
3.1	Méthode Monte Carlo simple . . . . .	30
3.2	Méthode d'échantillonnage préférentiel par changement de mesure . . . . .	32
3.3	Applications numériques . . . . .	33
3.3.1	Sinistres obéissent à une loi exponentielle . . . . .	33
3.3.2	Sinistres obéissent à une loi hyper-exponentielle . . . . .	36
3.3.3	Sinistres obéissent à une loi Gamma . . . . .	38
3.3.4	Sinistres obéissent à une loi Weibull . . . . .	40
3.3.5	Discussion des résultats . . . . .	41
3.4	Conclusion . . . . .	42
	<b>Conclusion générale</b>	<b>43</b>
	<b>Bibliographie</b>	<b>44</b>

---

## TABLE DES FIGURES

1.1	Evolution de la réserve et du surplus au cours du temps . . . . .	6
1.2	Processus de Cramèr-Lundberg (évolution de $X_{(t)}$ dans le temps) . . . . .	8
2.1	Estimation de la valeur $\pi$ , la variance de cette méthode et l'intervalle de confiance. . . . .	20
3.1	Variation de la probabilités de ruine (valeurs exactes et approximations), avec des montants de réclamation exponentielle, en fonction de la reserve initiale $u$ . . . . .	35
3.2	Variation d'erreurs relatives de l'approximation de $\psi(u)$ . . . . .	35
3.3	Variation de la probabilités de ruine, avec des sinistres hyper-exponentielle, en fonction de la reserve initiale $u$ . . . . .	38
3.4	Variation de la probabilités de ruine, avec des montants de réclamation de loi Gamma, en fonction de la reserve initiale $u$ . . . . .	39
3.5	Variation de la probabilités de ruine, avec des montants de réclamation de loi de Weibull, en fonction de la reserve initiale $u$ . . . . .	41

---

## LISTE DES TABLEAUX

2.1	Valeur de $n$ en fonction de $\gamma$ . . . . .	22
3.1	Probabilités de ruine (valeurs exactes et approximations) avec des montants de réclamation exponentielle . . . . .	34
3.2	Erreurs relatives de l'approximation de $\psi(u)$ . . . . .	36
3.3	Approximations de la probabilités de ruine avec des sinistres obéissent à une loi hyper-exponentielle . . . . .	37
3.4	Approximations de la probabilités de ruine avec des sinistres obéissent à une loi Gamma . . . . .	39
3.5	Approximations de la probabilités de ruine avec des sinistres obéissent à une loi de Weibull . . . . .	40

---

# NOTATIONS

$X, Y$  : Variables aléatoires.  
 $\mathcal{N}(0, 1)$  : Distribution gaussienne centrée réduite.  
 $F$  : Fonction de répartition de la densité  $f$ .  
 $f$  : Densité de probabilité.  
 $\mathcal{L}(\cdot)$  : Loi d'une variable aléatoire.  
 $n$  : Nombre des échantillons.  
 $i$  : Indice des itérations d'un algorithme.  
 $i, j$  : Indices généraux.  
 $\pi$  : Densité cible.  
 $D$  : Dilution d'échantillon.  
 $\sim$  : Suivre la loi, de loi...  
 $\mathbb{R}$  : Ensemble des nombres réels.  
 $\mathbb{N}$  : Ensemble des entiers naturels.  
 $\mathbb{R}^d$  : Ensemble des nombres réels de dimension  $d$ .  
 $\mathbb{P}$  : Probabilité.  
 $\log$  : Logarithme népérien.  
v.a : Variable aléatoire.  
i.i.d : Indépendante identiquement distribuée.  
p.s : Presque sûrement.  
I.C : Intervalle de confiance.  
TCL : Théorème Centrale Limite.  
Var : Variance.  
MC : Monte-Carlo.  
CMC : Crude Monte-Carlo.



---

# INTRODUCTION GÉNÉRALE

Les sciences actuarielles sont une discipline relativement récente bien que les bases de l'assurance soient quand à elles très anciennes. Les principales théories en actuariat ont été développées en grande partie dans le dernier siècle. De plus, avec l'amélioration de l'informatique, entre autres, on remarque une progression rapide dans un grand nombre de domaines en actuariat. Un de ces domaines est la théorie de la ruine où la complexité mathématique est souvent un facteur limitant les chercheurs. Cependant, en utilisant des ordinateurs pour des approximations numériques ou encore des simulations, du progrès est encore fait continuellement dans ce domaine.

Le moyen le plus simple d'obtenir une valeur approximative de la probabilité de ruine est d'utiliser la méthode Monte Carlo pour simuler le processus de risque à plusieurs fois rapidement, puis à compter la fréquence relative pendant laquelle le processus passe en dessous de zéro. nous appelons cette méthode Crude Monte Carlo (CMC). Mais ce n'est pas non plus facile. Si la probabilité de ruine est très petit, un grand nombre de simulations est nécessaire, et cela peut prendre beaucoup de temps même avec des ordinateurs modernes.

Une façon d'améliorer la méthode de Monte Carlo consiste à utiliser la méthode d'échantillonnage préférentiel.

Ce travail est composé d'une introduction, trois chapitres, une conclusion et une bibliographie.

Dans le premier chapitre nous présentons quelques notions de base sur les modèles de risque de la théorie de la ruine, nous intéressons en particulier au modèle de Cramér-Lundberg (modèle classique).

Le deuxième chapitre présente une description générale des méthodes de simulation, la méthode de Monte Carlo et la méthode d'échantillonnage préférentiel. Dans un premier temps nous donnons les techniques de Monte-Carlo classiques utilisées pour l'estimation

des quantités données sous forme d'intégration. Par la suite, nous définissons la méthode d'échantillonnage préférentiel et nous donnons quelques notions de base concernant la méthode.

Le troisième chapitre est consacré à la simulation de la probabilité de ruine en utilisant les deux méthodes, la méthode de Monte Carlo et la méthode d'échantillonnage préférentiel. Nous implémentant, pour chaque méthode, un algorithme afin de calculer la probabilité de ruine et on termine par une application numérique.

---

---

# CHAPITRE 1

---

## LA THÉORIE DE LA RUINE ET MODÈLES DE RISQUES CLASSIQUES

### Introduction

La théorie de la ruine appartient aux sciences de la gestion des risques et aux mathématiques appliquées à l'assurance. Il s'agit de l'étude mathématique de modèles stochastiques et dynamiques adaptés aux réserves financières allouées à un portefeuille de contrats d'assurance de type non-vie d'une compagnie d'assurance, l'assurance de type IARD (Incendie, Accidents et Risques Divers).

L'objectif est de définir un cadre permettant la bonne gestion d'un portefeuille de contrat.

On imagine une compagnie d'assurance qui se lance sur un nouveau marché, la business line doit être :

- Viable, la compagnie doit être solvable a tout instant (la réserve ne doit pas tomber en dessous de 0).
- Rentable, la tarification doit permettre à l'assureur d'engranger des bénéfices pour rémunérer les actionnaires et les employés.

La gestion comptable est rendue difficile à cause de l'inversion du cycle de production. La compagnie reçoit les primes avant de verser des prestations, elle doit mettre du cash de côté, c'est ce qu'on appelle le provisionnement.

Cela nécessite une évaluation des risques dans un univers aléatoire, d'où l'intervention de statisticiens et de probabilistes. Cela passe par la définition de deux paramètres :

- Un niveau de capital initial.
- Une tarification.

L'analyse statistique permet de calibrer :

- Une loi pour les montants des sinistres.
- Une loi pour le nombre de sinistres.

L'hypothèse sous jacente est celle d'un portefeuille homogène dans lequel les montants de sinistres sont des variable aléatoire indépendante et identiquement distribuée. L'autre hypothèse est celle d'un modèle collectif quant à la modélisation de la charge totale générée par les sinistres. La modélisation de la charge totale induite par les sinistres impactant un portefeuille de contrats d'assurance sur une période de temps donnée (souvent 1 an) permet de mettre au point la tarification mais aussi de définir le niveau de provision requis pour faire face aux engagements pris par rapport aux assurés. L'étude de sa distribution donne accès à ces quantiles pouvant être vue comme des Value-at-Risk.

Dans ce chapitre, on présente quelques notions sur les modèles du risque classique, par la suite, on introduit la probabilité de ruine et on cite quelques approches utilisées pour approximer la probabilité de ruine dans un modèle de risque classique.

## 1.1 Le modèle individuel et le modèle collectif

### 1.1.1 Modèle individuel

Le modèle individuel modélise la charge total générée par les sinistres individu par individu (voir [5, 16]). La charge totale pour un portefeuille comprenant  $n$  contrats est définie par la formule :

$$S^{Ind} = \sum_{i=1}^N X_i \quad (1.1)$$

où,

- $X_i$  est une variable aléatoire positive qui indique le montant total des sinistres subit par l'assuré  $i$  sur la période d'observation.

Dans ce type de modèle les variables  $X_i$  sont supposées indépendentes mais pas forcément d'une même loi de probabilité.

### 1.1.2 Modèle collectif

Le modèle collectif modélise la charge totale subit par un portefeuille vue, non pas contrat par contrat, mais suivant un nombre total de sinistre tout assuré confondu (voir [5]), le montant global des pertes est :

$$S^{coll} = \sum_{i=1}^N X_i \quad (1.2)$$

où,

- $N$  v.a. pour le nombre de pertes ;
- $X_i$  le montant de la  $i^{\text{ème}}$  perte.

Les variables aléatoires  $X_i$  supposées de même loi indépendentes, de plus la variable  $N$  indépendante des  $X_i$ .

**Remarque 1.1** On parle de modèle collectif car on associe la même loi de probabilité pour les pertes. Par conséquent, le modèle individuel avec les pertes de même loi de probabilité est un cas particulier du modèle collectif quand  $N$  est une constante :

$$P(N = n) = 1.$$

## 1.2 Modèle de risque de la théorie de la ruine

### 1.2.1 Processus de réserve et de surplus

On peut considérer la réserve d'une compagnie d'assurance comme le montant d'argent qu'elle possède pour payer les réclamations de ses clients. Ces derniers paient des primes d'assurance qui sont ajoutées à la réserve de la compagnie, et qui sont sa seule source de revenus.

Le processus de réserve  $R(t)$  ;  $t \geq 0$  d'une compagnie d'assurance prend la forme suivante :

$$R(t) = u + ct - Z_t, \quad (1.3)$$

où :

- $u$  est la réserve initiale de la compagnie ;
- $c$  est le taux de cotisation par unité de temps ;
- $Z_t = \sum_{i=0}^{N(t)} X_i$  est le processus de pertes agrégés.

On définit également le processus de surplus  $\{S(t); t \geq 0\}$

$$S(t) = u - R(t) \quad (1.4)$$

### 1.2.2 Le processus de Lévy

Les processus de Lévy sont des processus stochastiques dont les propriétés sont très intéressantes pour la construction des modèles de risque. pour plus de détails sur les processus de Lévy consulter [3, 26, 12].

**Définition 1.1** *Un processus stochastique  $\{X_t; t \geq 0\}$  défini sur  $\mathbb{R}$  est un processus de Lévy s'il satisfait les conditions suivantes :*

1. *Pour tout choix de  $n \geq 1$  et  $0 \leq t_0 < t_1 < \dots < t_n$  les variables aléatoires  $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$  sont indépendantes.*
2.  $X_0 = 0$ .
3. *La loi de  $X_{s+t} - X_s$  ne dépend pas de  $s$ .*

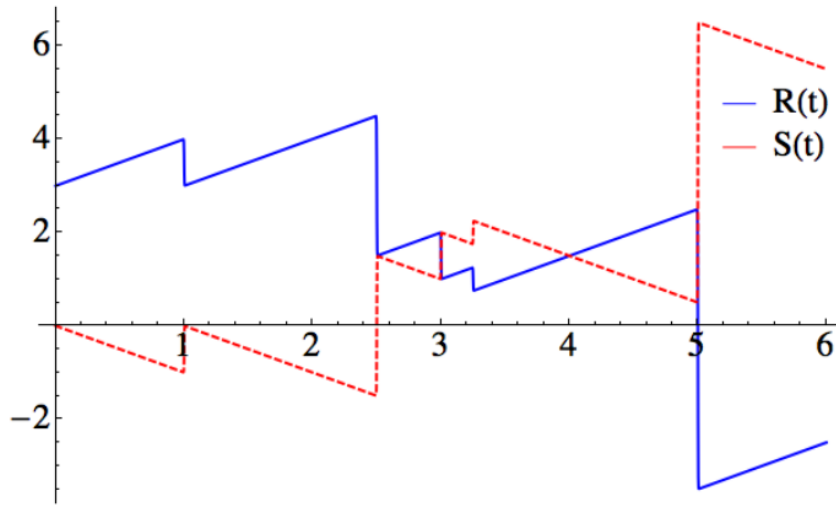


FIGURE 1.1 – Evolution de la réserve et du surplus au cours du temps

4.  $(X_t)$  est continu en probabilité, c'est-à-dire que pour tout  $t \geq 0$  et tout  $\varepsilon > 0$

$$\lim_{s \rightarrow t} P[|X_s - X_t| > \varepsilon] = 0$$

5. Il existe un ensemble  $\Omega_0 \in \mathcal{F}$  tel que  $P\{\Omega_0\} = 1$  et pour tout  $\omega \in \Omega_0$ ,  $X_t(\omega)$  est continue à droite en tout  $t \geq 0$  et a une limite à gauche en tout  $t > 0$ . Un processus aléatoire vérifiant les conditions (1) à (4) est dit processus de Lévy en loi.

### 1.2.3 Probabilité de ruine

Pour tout modèle de risque d'un surplus financier, la première quantité d'intérêt est la probabilité de ruine (voir [19, 13, 15]).

#### Définition 1.2 Instant de ruine

L'instant de la ruine dénotant par  $\tau$  associé à une réserve initiale  $u$ , est défini par :

$$\tau = \inf\{t \geq 0 : R(t) < 0\} = \inf\{t \geq 0; S(t) > u\} \quad (1.5)$$

#### Définition 1.3 Probabilité de ruine à horizon de temps infini

La probabilité de ruine ultime ou probabilité de ruine à horizon de temps infini notée  $\Psi$ , est définie par :

$$\Psi(u) = P(\inf_{t \geq 0} R(t) < 0 : R(0) = u) \quad (1.6)$$

#### Définition 1.4 Probabilité de ruine à horizon de temps fini

La probabilité de ruine avant l'instant  $T$  ou probabilité de ruine à horizon de temps fini est définie par :

$$\Psi(u, T) = P(\inf_{t \in [0, T]} R(t) < 0 : R(0) = u) \quad (1.7)$$

**Définition 1.5** Probabilité de survie ou non ruine

La probabilité complémentaire ou probabilité de non ruine, notée  $\Phi$ , est définie par :

$$\Phi(u) = 1 - \Psi(u) \quad \Phi(u, T) = 1 - \Psi(u, T) \quad (1.8)$$

### 1.2.4 Processus de renouvellement

On considère un modèle général où les interarrivées sont des variables aléatoires indépendantes identiquement distribuées mais ne suivent pas nécessairement la loi exponentielle. Le nombre des sinistres à la date  $t$ , noté  $N(t)$ , est alors appelé processus de renouvellement. C'est un processus de Poisson si de plus les interarrivées suivent la loi exponentielle.[6].

#### Processus de Poisson

Un processus de Poisson est un processus de comptage d'événements entre lesquels s'écoulent des durées indépendantes de distribution exponentielle.

**Définition 1.6** Soit  $\{T_i; i \geq 1\}$  une suite de variables aléatoires indépendantes et identiquement distribuées de loi exponentielle de paramètre  $\mu > 0$ . Notons  $\sigma_n = \sum_{i=1}^n T_i$ . Le processus de comptage  $\{N_t; t \geq 0\}$  défini par :

$$N_{(t)} = \sum_{n \geq 1} 1(\sigma_n \leq t)$$

est un processus de Poisson homogène d'intensité  $\mu$ .

#### Processus de Poisson composé

**Définition 1.7** Soit  $N$  un processus de Poisson d'intensité  $\lambda$ , et  $\{X_i; i \geq 1\}$  des variables aléatoires indépendantes et identiquement distribuées et indépendantes de  $N$ . Le processus  $Z$  défini par  $Z_t = \sum_{i=1}^{N(t)} X_i$  est dit un processus de Poisson composé.

## 1.3 Modèle de Cramèr-Lundberg

Le modèle de risque de Cramèr-Lundberg a été introduit en 1903 par l'actuaire Filip Lundberg, il est la base du fondement de la théorie de la ruine. Il est connue aussi sous le nom du modèle de risque Poisson composé ou encore modèle de risque classique.

Le processus stochastique régissant l'évolution des réserves financières noté  $\{R(t), t \geq 0\}$  avec :

$$R(t) = u + ct - Z_t \quad t \geq 0, \quad (1.9)$$

où  $Z_t = \sum_{i=1}^{N(t)} X_i$  est le montant cumulé des réclamations à l'instant  $t$ . Le processus de risque associé est de la forme

$$S(t) = ct - Z_t.$$

Les hypothèses selon le modèle sont :

- $u > 0$  est la réserve initiale de la compagnie d'assurance ;
- $c > 0$  est le taux de prime ou (cotisation) reçues continuellement dans le temps ;
- $N_t$  est un processus de comptage (Poisson) d'intensité  $\lambda$  du nombre de réclamations ;
- $X_i; i \geq 1$ , sont les montants des réclamations, est représentant une suite de variables aléatoire, strictement positive, i.i.d de fonction de répartition  $F_x$ , de moyenne finie  $\mu$ , et indépendantes de  $N_t$ .

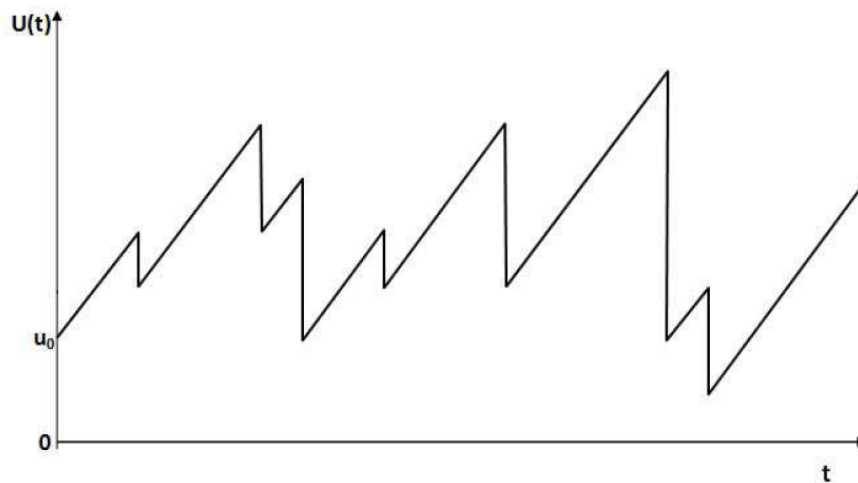


FIGURE 1.2 – Processus de Cramèr-Lundberg (évolution de  $X(t)$  dans le temps)

### 1.3.1 Cas particulier : Modèle de Lundberg P/P

Le modèle de Lundberg est un cas particulier du modèle de risque classique appelé aussi  $P/P$ . Il se caractérise par la distribution exponentielle des montants des réclamations :

$$F_Z(y) = 1 - \exp\left\{-\frac{y}{\mu}\right\}.$$

où  $F_Z$  est la fonction de répartition de la variable aléatoire  $Z$  qui génère les montants des réclamations.



### 1.3.2 Le coefficient de sécurité

Soit un processus  $\{Y_t; t \geq 0\}$  défini par :

$$Y_t = ct - \sum_{i=1}^{N_t} X_i \quad (1.10)$$

Le risque moyen sur un intervalle  $[0, t]$  est égale à :

$$E[Y_t] = ct - \mu E[N_t] = (c - \lambda\mu)t$$

Nous appelons  $\eta = c - \lambda\mu$  le coefficient de sécurité ou chargement de sécurité.

La constante  $\eta$  définit le chargement de sécurité de la compagnie. Il mesure la rentabilité de la compagnie.

Le coefficient  $\lambda\mu$  est interprété comme le montant moyen des sinistres par unité de temps. Il paraît prudent que l'assureur fixe un taux de prime  $c$  supérieur à  $\lambda\mu$  pour que, en moyenne, les primes reçues soient supérieures aux indemnités payées par la compagnie d'assurance. En effet, nous avons la propriété suivante :

**Propriété 1.1** [2] coefficient de sécurité

— Si  $\eta > 0$ , alors, presque sûrement,

$$\lim_{T \rightarrow \infty} R_t = +\infty$$

et

$$\Phi(u) \neq 0$$

— Si  $\eta < 0$ , alors, presque sûrement,

$$\lim_{T \rightarrow \infty} R_t = -\infty$$

et

$$\Psi(u) = 1$$

*L'activité n'est donc rentable.*

### 1.3.3 Condition de non ruine

Considérons le processus de réserve  $R_t$  d'une certaine compagnie d'assurance à l'instant  $t$ , avec le capital initial est  $u$  et les cotisations sont versées par les clientes à un taux instantané  $c$ .

$$R(t) = u + ct - Z_t = u + ct - \sum_{i=1}^{N_t} X_i.$$

Ainsi, le théorème de Wald [24], nous permet d'obtenir :

$$E(R(t)) = u + ct - \lambda\mu t.$$

avec  $\mu$  étant le taux de remboursement.

Une condition qu'il est naturel d'imposer que l'espérance de la fortune d'une compagnie d'assurance soit toujours positive, soit  $c \geq \lambda\mu$ .

Remarquons que la ruine peut se produire nécessairement à l'occasion d'un sinistre, c'est-à-dire à une dates  $T_n$ . On a donc :

$$\Psi(u) = P\{\exists t \geq 0 / R(t) < 0\} = P\{\exists n \geq 0 / R_{(T_n)} < 0\}.$$

Où, on a :

$$R_{(T_n)} = u + cT_n - Z_{(T_n)} = u - \sum_{i=1}^n (X_i - c\tau_i).$$

D'où :

$$\Psi(u) = P\{\exists n \geq 0 \mid \sum_{i=1}^n (X_i - c\tau_i) > \mu\} = 1 - P\{\max_{n \geq 1} \sum_{i=1}^n (X_i - c\tau_i) \leq \mu\}.$$

Par la loi des grands nombres, on sait que  $\frac{\sum_{i=1}^n (X_i - c\tau_i)}{n}$  converge presque sûrement vers :

$$E(X_1 - c\tau_1) = \mu - (c/\lambda)$$

- La ruine est presque sûre, si  $\lambda\mu > c$ , qui converge presque sûrement vers  $+\infty$ .
- La ruine n'est pas presque sûre, si  $\lambda\mu < c$ , qui converge presque sûrement vers  $-\infty$ , c'est-à-dire qu'il existe une probabilité positive de non ruine à horizon infini.
- Si  $\lambda\mu = c$ , la théorie des marches aléatoires montre que  $\limsup_{n \rightarrow \infty} \sum_{i=1}^n (X_i - c\tau_i) = +\infty$  presque sûrement et donc la ruine est presque sûre.

La condition de non-ruine est donc :

$$\lambda\mu < c.$$

On peut alors définir le chargement de sécurité relatif  $\theta$  par :

$$\theta = \frac{c - \lambda\mu}{\lambda\mu} = \frac{c}{\lambda\mu} - 1$$

## 1.4 Approximation de la probabilité de ruine

Dans ce qui suit, nous présentons quelques résultats fondamentaux pour l'évaluation de la probabilité de ruine en temps infini.

### 1.4.1 Approximations Cramér-Lundberg

Soit  $\Psi(u)$ , avec  $u \geq 0$  la probabilité de ruine de modèle 1.9. On suppose que le chargement de sécurité relatif  $\theta = \frac{c - \lambda\mu}{\lambda\mu}$  est strictement positif. On note  $\bar{F}(x) = 1 - F(x)$ , où  $F$  est la fonction de distribution des moments des réclamations.

En utilisant les arguments de renouvellement et en conditionnant par rapport au temps et au montant de la première réclamation, on a la probabilité de ruine qui vérifie l'équation intégrale suivante : [23]

$$\Psi(u) = \frac{\lambda}{c} \int_u^{\infty} \bar{F}(y) dy + \frac{\lambda}{c} \int_u^{\infty} \Psi(u-y) \bar{F}(y) dy.$$

En général, il est très difficile de dériver des expressions explicites de la probabilité de ruine. Cependant, sous certaines conditions convenables, on peut obtenir quelques approximations de cette quantité. Les premiers travaux sur ces approximations ont été réalisés par Cramèr-Lundberg dès 1930. La condition de Cramèr-Lundberg stimule l'existence d'une constante  $k > 0$  satisfaisant l'équation de Lundberg :

$$\int_0^{\infty} e^{kx} \bar{F}(x) dx = \frac{c}{\lambda},$$

qui est équivalente à :

$$\int_0^{\infty} e^{kx} dG(x) = 1 + \theta. \quad (1.11)$$

où  $G(x) = \frac{1}{\mu} \int_0^x s^{kx} \bar{F}(y) dy$  est la distribution équilibre de  $F$ .

Supposons que l'équation 1.11 est vérifiée. La formule asymptotique de la probabilité de ruine est donnée comme suit : Si  $\int_0^{\infty} e^{kx} dG(x) < \infty$ , alors :

$$\Psi(u) \sim \frac{\theta u}{k \int_0^{\infty} y e^{ky} \bar{F}(y) dy} e^{-ku} \quad \text{quand } u \rightarrow \infty,$$

avec,

$$a(x) \sim b(x) \quad \text{quand } x \rightarrow \infty \Leftrightarrow \lim_{x \rightarrow \infty} \frac{a(x)}{b(x)} = 1.$$

Si  $\int_0^{\infty} e^{kx} dG(x) = \infty$ , alors,

$$\Psi(u) = O(e^{-kx}) \quad \text{quand } u \rightarrow \infty,$$

Ainsi, on a l'inégalité de Lundberg :

$$\Psi(u) \leq e^{-ku}, \quad u \geq 0.$$

### 1.4.2 Approximations de Cramèr-Lundberg dans le modèle P/P

La distribution des montants de réclamations dans le modèle de Lundberg est exponentielle, d'où :  $\bar{F}(x) = \exp\{-\frac{x}{\mu}\}$   $x \geq 0$ , on aura la formule de la probabilité de ruine :

$$\Psi(u) = \frac{1}{1+\theta} \exp\{-\frac{\theta u}{\mu(1+\theta)}\} \quad u \geq 0$$

### 1.4.3 Formule de Pollaczak-Khinchine

**Théorème 1.1** [4] Pour tout  $u \geq 0$  :

$$\Psi(u) = (1 - \frac{\lambda\mu}{c}) \sum_{n=1}^{\infty} (\frac{\lambda\mu}{c})^n \overline{(F_X^s)^{*n}}(u). \quad (1.12)$$

où  $\overline{(F_X^s)^{*n}}(u) = 1 - (F_X^s)^{*n}$  et  $(F_X^s)^{*n}$  est la  $n^{\text{ème}}$  convolution de la fonction de distribution complémentaire  $F_X^s$  telle que  $F_X^s$  est la fonction de répartition de  $X$  définie par :

$$F_X^s(x) = \frac{1}{\mu} \int_0^x (1 - F_X(y)) dy, \quad x \geq 0.$$

La formule 1.12 est appelée formule de Pollaczak-Khinchine ou encore formule de Beehan. La représentation en série infinie donnée dans 1.12 est particulièrement utile pour des considérations théoriques. Toutefois, il est également utile d'utiliser des approximations numériques de la probabilité de ruine  $\Psi(u)$ , telle que l'algorithme de Panjer [20].

### 1.4.4 Formule de Pollaczak-Khinchine dans le modèle P/P

En utilisant la formule 1.12, nous allons déduire l'expression exacte de  $\Psi(u)$  pour le modèle de risque de Lundberg (P/P).

$$\Psi(u) = (1 - \rho) \cdot \sum_{i=1}^{\infty} \rho^i \overline{(F_X^s)^{*i}}(u), \rho = \frac{\lambda}{\delta c}$$

Pour des montants de réclamations exponentielles de paramètre  $\delta$  où  $\delta = (\frac{1}{\mu})$ , nous avons :

$$F_X(u) = \begin{cases} 1 - \exp(-\delta u), & \text{si } u \geq 0 \\ 0, & \text{si } u < 0 \end{cases} \quad (1.13)$$

Calculons,

$$F_X^s(u) = \delta \int_0^u (1 - F_X(y)) dy, \quad x \geq 0$$

$$F_X^s(u) = \delta \int_0^u \exp(-\delta y) dy = \delta \left( \frac{-1}{\delta} \right) [\exp(-\delta y)]_0^u = 1 - \exp(-\delta u)$$

Ainsi

$$F_X^s(u) = F_X^s(u), u \in \mathbb{R} \quad (1.14)$$

$(F_X^s)^{*n}$  représente la  $n^{\text{ème}}$  convolution de  $(F_X^s)$  Puisque nous avons l'indépendance des  $n$  variables aléatoires  $Z_i, \{i = 1, 2, \dots, n\}$ .

Nous utiliserons les transformées de Laplace afin de déterminer  $(F_X^s)^{*n}$ . Où  $f_x$  est la densité

de probabilité des montants des réclamations  $Z_i, \{i = 1, 2, \dots, n\}$ .

Ainsi,

$$L_{f_x}(x) = \int_0^{\infty} f_x(t) \exp^{-xt} dt = \int_0^{\infty} \delta \exp^{-(\delta+x)t} dt = \frac{\delta}{\delta+x}$$

D'où,

$$L_{(f_x)^n}(x) = \left[ \frac{\delta}{\delta+x} \right]^n$$

En se référant aux tables des transformées de Laplace [21] nous trouvons que :

$$(f_x)^n(x) = \frac{\delta(\delta x)^{n-1}}{(n-1)!} \exp^{-\delta x}, \quad x \geq 0$$

qui correspond à la densité de probabilité de la loi d'Erlang  $(\delta, n)$  [11] dont la fonction de répartition est donnée par :

$$(f_x)^n(x) = (F_x^s)^n(x) = \frac{\Gamma(n, \delta x)}{(n-1)!} = 1 - \sum_{k=0}^{n-1} \exp^{-\delta x} \frac{(\delta x)^k}{k!}$$

Ce résultat signifie que la somme de  $n$  variables aléatoires indépendantes de distributions exponentielles de même paramètre  $\delta$  est une loi d'Erlang  $(\delta, n)$  [11]. Alors,

$$\overline{(F_X^s)^*n}(x) = -(F_X^s)^*n(x) = \sum_{k=0}^{n-1} \exp^{-\delta x} \frac{(\delta x)^k}{k!}$$

Ainsi

$$\begin{aligned} \Psi(u) &= (1-\rho) \sum_{n=1}^{\infty} \rho^n \sum_{k=0}^{n-1} \exp^{-\delta u} \frac{(\delta u)^k}{k!} \\ &= (1-\rho) \exp^{-\delta u} \sum_{n=1}^{\infty} \rho^n \sum_{k=0}^{n-1} \frac{(\delta u)^k}{k!} \\ &= (1-\rho) \exp^{-\delta u} \sum_{k=0}^{\infty} \frac{(\delta u)^k}{k!} \sum_{n=k+1}^{\infty} \rho^n \\ &= (1-\rho) \exp^{-\delta u} \sum_{k=0}^{\infty} \frac{(\delta u)^k}{k!} \frac{\rho^{k+1}}{1-\rho} \\ &= \rho \exp^{-\delta u} \sum_{k=0}^{\infty} \frac{(\rho \delta u)^k}{k!} \\ &= \rho \exp^{-\delta u} \exp^{\rho \delta u} \\ &= \rho \exp^{-(\delta - \frac{\lambda}{c})u} \end{aligned}$$

Finallement

$$\Psi(u) = \frac{\lambda}{c\delta} \exp^{-(\delta - \frac{\lambda}{c})u}$$

### 1.4.5 Autres approches

En plus de l'approche stochastique pour l'évaluation de la probabilité de ruine, qui possède de large champs dans les modèles de risques, il existe plusieurs autres approches. Ces approches permettent une meilleure considération des faits, car certains faits ignorés dans la modélisation stochastique se retrouvent dans d'autres domaines. C'est le cas par exemple des réactions des assureurs et des assurés dans la théorie des jeux. En général, les solutions proposées pour estimer la probabilité de ruine sont basées sur : les théorèmes limites des marches aléatoires [25, 10], les représentations matricielles avec modèles markoviens [1], la théorie des martingales et inégalités de probabilités (Sato [20]), les méthodes d'optimisation (De Vylder [27]), les transformations analytiques (De Vylder [7]) et la théorie des distributions [29, 28].

## 1.5 Conclusion

Dans ce chapitre, nous avons présenté quelques notions de base sur les modèles de risque de la théorie de ruine. Une attention particulière est portée sur le modèle de Cramér-Lundberg (modèle classique), connu comme le fondement théorique de la théorie du risque dû à plusieurs études, et de nombreux résultats pour la probabilité de ruine existent pour ce modèle. Nous avons présenté également quelques approximations et expressions exactes de la probabilité de ruine pour le modèle de risque classique.

---

---

## CHAPITRE 2

---

# LA MÉTHODE DE MONTE CARLO ET ÉCHANTILLONNAGE PRÉFÉRENTIEL (IMPORTANCE SAMPLING)

### Introduction

Les méthodes de Monte Carlo permettent d'estimer des quantités en utilisant la simulation de variables aléatoires. Les problèmes pouvant être rencontrés comprennent le calcul d'intégrales, les problèmes d'optimisation et la résolution de systèmes linéaires. La simplicité, la flexibilité et l'efficacité pour les problèmes en grande dimension de la méthode en font un outil intéressant, pouvant servir d'alternative ou de référence pour d'autres méthodes numériques.

Supposons que l'on souhaite connaître la valeur d'une certaine quantité  $\delta$ . La première étape de la méthode consiste à écrire le problème sous la forme d'une espérance. Soient une variable aléatoire  $X = (X_1, X_2, \dots, X_d)$  de loi  $v$  sur  $R^d$  (on abrègera cela par  $X \sim v$ ) et une fonction  $h : R^d \rightarrow R$ . Le problème traité par les méthodes de Monte Carlo est l'estimation de

$$\delta = \mathbb{E}_v[h(X)] = \int_{R^d} h(x)v(x)dx. \quad (2.1)$$

La solution standard à ce problème est de simuler une suite  $(X^n)_{n \geq 1} = (X_1, X_2, \dots, X_d, n)_{n \geq 1}$  de variables aléatoires indépendantes identiquement distribuées (i.i.d.) suivant la loi  $v$ , puis d'estimer l'espérance  $\mathbb{E}_v[h(X)]$  par la moyenne empirique, i.e.,

$$\bar{h}_n = \frac{1}{n} \sum_{k=1}^n h(x_k) \quad (2.2)$$

**Exemple 2.1** Soit  $h : [a, b]^d \rightarrow \mathbb{R}$ . On cherche à calculer

$$\delta = \int_{[a, b]^d} h(x_1, x_2, \dots, x_d) dx_1, dx_2, \dots, dx_d.$$

On peut réécrire  $\delta$  sous la forme

$$\delta = (b - a)^d \int_{[a, b]^d} h(x_1, x_2, \dots, x_d) \frac{1}{(b-a)^d} dx_1, dx_2, \dots, dx_d.$$

Si l'on pose  $X = (X_1, X_2, \dots, X_d)$  un  $d$ -uplet de variable i.i.d. suivant la loi uniforme sur  $[a, b]$ , on a alors

$$\delta = (b - a)^d \mathbb{E}[h(X)].$$

De façon générale, si  $\delta = \int_{\mathbb{R}^d} h(x) f(x) dx$ , avec  $f$  le densité de probabilité sur  $\mathbb{R}^d$  et  $h$  une fonction borélienne, alors on peut écrire, sous les hypothèses d'existence de  $\delta$ ,  $\delta = \mathbb{E}[h(X)]$ . [18]

## 2.1 Loi faible et loi forte des grands nombres

Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires réelles i.i.d. de loi  $v$  telle que  $h(x_1)$  soit  $v$  intégrable. Alors, pour tout  $\varepsilon > 0$ ,

$$\mathbb{P}[|\bar{h}_n - \mathbb{E}[h(X_1)]| \leq \varepsilon] \xrightarrow{n \rightarrow +\infty} 1 \text{ (loi faible)}, \quad \mathbb{P}[|\bar{h}_n - \mathbb{E}[h(X_1)]| \xrightarrow{n \rightarrow +\infty} 0] = 1 \text{ (loi forte)}$$

La loi des grands nombres renseigne donc de deux façons sur l'erreur que l'on commet en estimant  $\delta$  par la moyenne empirique  $\bar{h}_n$ . Selon la version faible, il y a une probabilité nulle de commettre une erreur plus grande que  $\varepsilon$ , tandis que la version forte dit qu'une fois que l'erreur d'estimation est inférieure à  $\varepsilon$ , alors cette erreur ne peut que décroître.

La loi des grands nombres n'apporte en revanche pas d'informations pratiques sur l'erreur commise. Ainsi, elle ne permet pas de choisir une taille d'échantillon  $n$  qui assure que la méthode produira une erreur aussi petite que l'on souhaite, ni de dire si pour un échantillon  $X_1, X_2, \dots, X_n$  de taille  $n$  fixée, l'erreur sera arbitrairement petite ou non.

**Théorème 2.1** [18] Soit  $(X_n)_{n \geq 1}$  une suite de v.a.i.i.d à valeur dans  $\mathbb{R}^d, d \in \mathbb{N}, X_1 \sim \mathcal{L}(f)$ . On suppose que  $\varphi(x)$  est une fonction intégrable, ( $\mathbb{E}[\varphi(X)] < +\infty$ ). alors :

$$\frac{1}{n} \sum_{i=1}^n \varphi(x_i) \xrightarrow{n \rightarrow +\infty} \mathbb{E}[\varphi(X)].$$



Sous les hypothèses données précédemment et les lois des grands nombres, l'algorithme qui permet d'estimer une quantité de type  $I = \int_{R_d} \varphi(x)f(x)dx$  est donné comme suit :

---

**algorithme d'estimation.**

---

**Debut**

lire(n);

$s = 0$ ;

**pour**  $i = 1 : n$  **faire**

générer  $X(i) \sim \mathcal{L}(f)$ ; où  $\sim \mathcal{L}$  c'est la loi de density de  $f$

$s = s + \varphi(X(i))$ ;

**fin pour**;

$I_n = \frac{s}{n}$ ;

afficher ( $I_n$ );

**Fin.**

---

### 2.1.1 Exemple d'application

**Exemple 2.2 Estimation de la valeur  $\pi$**

Soit une suite de couple  $(X_n, Y_n)_{n>1}$  de v.a.i.i.d,  $(X_1, Y_1) \sim \mathcal{U}([-1, 1])$ .

Posons  $\varphi(x, y) = 1_{\{x^2+y^2 \leq 1\}}(x, y)$ .

L'aire du cercle d'unité est donnée par :

$$\begin{aligned} I &= \int_{-1}^1 \int_{-1}^1 1_{\{x^2+y^2 \leq 1\}} dx dy \\ &= 4 \int_{-1}^1 \int_{-1}^1 1_{\{x^2+y^2 \leq 1\}} \times \frac{1}{2} 1_{\{-1 \leq x \leq 1\}} \times \frac{1}{2} 1_{\{-1 \leq y \leq 1\}} dx dy \\ &= 4 \int_{-1}^1 \int_{-1}^1 \varphi(x, y) f(x, y) dx dy \end{aligned}$$

D'après la loi des grands nombres on aura,

$$\frac{1}{n} \sum_{i=1}^n \varphi(X_i, Y_i) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}[\varphi(X, Y)].$$

D'où  $I \approx 4 \times \frac{1}{n} \sum_{i=1}^n \varphi(X_i, Y_i)$ .

Mais, nous savons que l'air du cercle d'unité est  $\pi$ , d'où

$$\pi \approx 4 \times \frac{1}{n} \sum_{i=1}^n \varphi(X_i, Y_i).$$

Un des algorithmes d'estimation de la valeur de  $\pi$  est donné comme suit :

---

**Algorithme d'estimation de la valeur  $\pi$ .**

---

**Debut**

lire( $n$ );

$s = 0$ ;

**pour**  $i = 1 : n$  **faire**

générer  $U(i) \sim \mathcal{U}([0, 1])$  ;

générer  $V(i) \sim \mathcal{U}([0, 1])$  ;

$X(i) = 2U(i) - 1$  ;

$Y(i) = 2V(i) - 1$  ;

**si**  $(X(i)^2 + Y(i)^2 \leq 1)$  ;  $s = s + 1$  ;

**fin si** ;

**fin pour** ;

$\pi = 4 * \frac{s}{n}$  ;

$I_n = \frac{s}{n}$  ;

afficher ( $\pi$ ) ;

**Fin.**

---

## 2.2 Vitesse de convergence et estimation de l'erreur Monte Carlo

Il est possible de préciser d'avantage le comportement de la méthode lorsque  $h$  est de carré intégrable par rapport à la mesure  $\nu$ , i.e.,  $\mathbb{E}_\nu[h(X)^2] < \infty$ .  $\bar{h}_n$  étant un estimateur sans biais de  $\delta = \mathbb{E}_\nu[h(X)]$ , i.e  $\mathbb{E}_\nu[\bar{h}_n] = \delta$  son erreur quadratique moyenne est

$$\mathbb{E}[\{\bar{h}_n - \mathbb{E}[h(X)]\}^2] = \frac{\delta^2}{n}, \text{ où } \delta^2 = \text{Var}[h(X)]. \quad (2.3)$$

Autrement dit, la vitesse de convergence de la méthode de Monte Carlo classique est en  $O(n^{-\frac{1}{2}})$ .

On peut d'ores et déjà remarquer que cette vitesse de convergence ne permet pas d'estimer  $\delta$  avec une grande précision, chaque chiffre significatif supplémentaire, nécessitant un coût de simulation 100 fois supérieur.

Lorsque la variance de  $h(X)$  est finie, le théorème Central-Limite permet d'établir que  $\bar{h}_n - \mathbb{E}_\nu[h(X)]$  suit asymptotiquement une loi normale.

**Théorème 2.2** [18] *Théorème central limite*

Soit  $(X^n)_{n \geq 1}$  une suite de variables aléatoires réelles i.i.d. de loi  $v$  telle que  $h[(x_1)]^2$  soit  $v$  intégrable, alors :

$$\sqrt{n}(\bar{h}_n - \mathbb{E}[h(X_1)]) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \delta^2) \Leftrightarrow \forall z \in \mathbb{R}, \mathbb{P}[\sqrt{\frac{n}{\delta^2}}(\bar{h}_n - \mathbb{E}[h(X_1)]) \leq z] \xrightarrow[n \rightarrow +\infty]{} \Phi(z),$$

où  $\Phi$  la fonction de répartition de la loi normale  $\mathcal{N}(0, 1)$ .

L'erreur commise par la méthode de Monte Carlo est aléatoire. Elle ne peut être bornée mais elle peut être quantifiée via un intervalle de confiance. Sous les hypothèse du théorème central limite, pour tout  $q$  réel,

$$\mathbb{P}[\sqrt{\frac{n}{\delta^2}}|\bar{h}_n - \mathbb{E}[h(X)]| \leq q] \rightarrow \int_{-q}^q \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = 2\Phi(q) - 1, \text{ tel que } n \rightarrow +\infty.$$

On en déduit l'intervalle de confiance bilatérale symétrique au niveau de confiance asymptotique  $1 - \alpha$

$$IC_{1-\alpha} = [\bar{h}_n - q_{1-\alpha/2} \sqrt{\frac{\sigma^2}{n}}] \text{ avec } \Phi(q_{1-\alpha/2}) = 1 - \frac{\alpha}{2}.$$

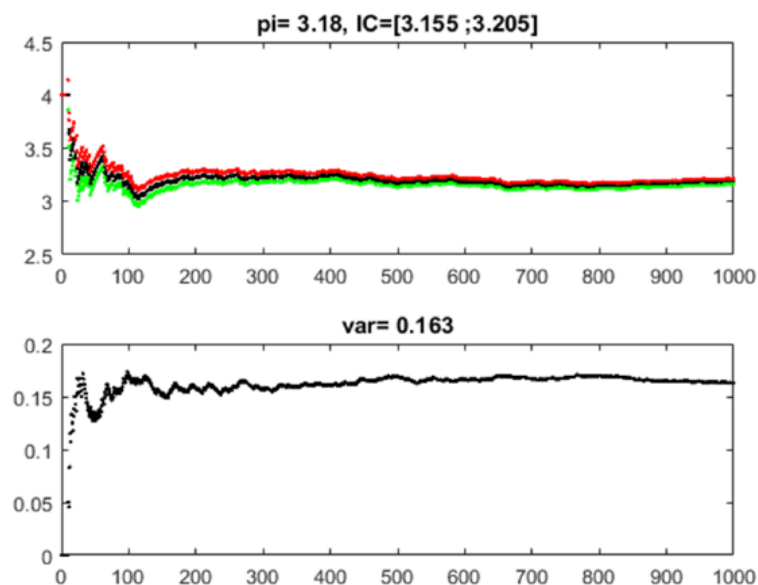
**Remarque 2.1** Le niveau de confiance usuel est  $1 - \alpha = 0.95$ . Alors  $q_{1-\alpha/2} = \Phi^{-1}(0.975) \approx 1.96$ .

L'erreur quadratique moyenne et l'intervalle de confiance asymptotique dépendent tous les deux de la variance  $\sigma^2$ , inconnue en pratique. Lorsque la variance est finie, il est possible d'utiliser l'échantillon  $(X_1, X_2, \dots, X_n)$  pour obtenir un estimateur de  $\sigma^2$ , noté  $\hat{\sigma}_n^2$ .

**Exemple 2.3** Estimation de la variance de  $\pi$  ainsi que le construction de l'intervalle de confiance à  $(1 - \alpha) = 95\%$ .

Dans l'exemple 2.2 de l'estimation de la valeur de  $\pi$ , nous avons posé  $\varphi(x, y) = 1_{\{x^2+y^2 \leq 1\}}(x, y)$  avec  $(X, Y) \sim \mathcal{U}([-1, 1])$ .

L'estimation de la variance et de l'intervalle de confiance de cette méthode sont donnés par l'algorithme suivant et la figure 2.1.

*Algorithme d'estimation de la variance de  $\pi$ .***Debut** $lire(n)$ ; $var = 0$ ;**pour**  $i = 1 : n$  **faire**g nerer  $U(i) \sim \mathcal{U}[0, 1]$  ;g nerer  $T(i) \sim \mathcal{U}[0, 1]$  ;g nerer  $W(i) \sim \mathcal{U}[0, 1]$  ;g nerer  $V(i) \sim \mathcal{U}[0, 1]$  ; $X(i) = 2U(i) - 1$  ; $Y(i) = 2V(i) - 1$  ; $Z(i) = 2T(i) - 1$  ; $R(i) = 2W(i) - 1$  ;**si**  $(X(i)^2 + Y(i)^2 \leq 1)$  **et**  $(Z(i)^2 + R(i)^2 \leq 1)$  **alors** $var = var + [(X(i)^2 + Y(i)^2) - (Z(i)^2 + R(i)^2)]^2$  ;**fin si** ;**fin pour** ; $\hat{\sigma}_n^2 = var / (2 * n)$  ;afficher  $(\hat{\sigma}_n^2)$  ;**Fin.**FIGURE 2.1 – Estimation de la valeur  $\pi$ , la variance de cette m thode et l'intervalle de confiance.

Avec 1000 simulations on obtient un estimateur qui vaut 3.1800 et une variance de 0.1630 pour un intervalle de confiance [3.1550, 3.2050].

Dans cette section nous avons vu que l'estimation d'une quantité de type  $I = \mathbb{E}[\varphi(X)] = \int_{\mathbb{R}^d} \varphi(x)f(x)dx$  génère une erreur  $\varepsilon_n$  d'ordre  $\frac{\sigma^2}{n} = \text{var}[\varphi(X)]/n$ . Donc il est de notre intérêt de réduire la variance afin de minimiser l'erreur de l'estimation d'une part et minimiser le nombre d'itérations nécessaires à l'estimation de la variance d'autre part. En effet, supposons que l'on veuille estimer la quantité suivante :

$$I = \mathbb{E}\varphi(X) = \int_{\mathbb{R}} \varphi(x)f(x)dx.$$

avec :

$$\begin{cases} \varphi = e^{\gamma x}, \gamma \in \mathcal{R}^+ \\ f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, x \in \mathbb{R} \end{cases}$$

Pour l'estimation de cette quantité on doit suivre les étapes suivantes.

1. On tire une suite de v.a.i.i.d  $X_1, X_2, \dots, X_n, X_1 \sim \mathcal{N}(0, 1)$ .
2. On montre que  $\mathcal{E}[\varphi(X)] < +\infty$ .

Nous avons

$$\begin{aligned} \mathbb{E}[\varphi(X)] &= \int_{\mathbb{R}} \varphi(x)f(x)dx \\ &= \int_{-\infty}^{+\infty} e^{\gamma x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{(-\frac{1}{2}(x-\gamma)^2 + \frac{\gamma^2}{2})} dx \\ &= e^{\frac{\gamma^2}{2}} \left( \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\gamma)^2} dx \right) \\ &= e^{\frac{\gamma^2}{2}}. \end{aligned}$$

D'après les étapes (1) et (2) et les lois des grands nombres, la quantité  $I$  peut-être approchée par un estimateur Monte-Carlo de la façon suivante :

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

Nous avons vu précédemment qu'afin d'approcher la quantité  $\mathbb{E}[\varphi(X)]$  à  $\beta$  près avec une confiance de  $(1 - \alpha)$ , nous aurons besoin de générer un nombre des échantillons qui dépend de la variance, i.e.

$$n \simeq \left( \frac{\phi^{-1}(1 - \alpha/2)}{\beta} \right)^2 \sigma^2 \quad (2.4)$$

**Calculons la variance de cette méthode.**

$$\begin{aligned} \sigma^2 &= \text{var}[\varphi(X)] \\ &= \mathbb{E}[\varphi(X)^2] - (\mathbb{E}[\varphi(X)])^2. \end{aligned}$$

avec

$$(\mathbb{E}[\varphi(X)])^2 = e^{\gamma^2}.$$

De la même façon, on trouve

$$\mathbb{E}[\varphi(X)^2] = e^{2\gamma^2}.$$

Donc

$$\begin{aligned} \sigma^2 &= e^{2\gamma^2} - e^{\gamma^2} \\ &= e^{\gamma^2}(e^{\gamma^2} - 1). \end{aligned}$$

Pour  $\beta = 0.01, \alpha = 5\%$  et par substitution dans l'équation 2.4 on trouve

$$n \simeq \left( \frac{1.96}{0.01} \right)^2 e^{\gamma^2} (e^{\gamma^2} - 1).$$

Le tableau (1.1) donne quelque valeurs approximatives de  $n$  en fonction de la valeur  $\gamma$ .

$\gamma$	3	4	5
$n \approx$	$3.10^{12}$	$3.10^{18}$	$2.10^{26}$

TABLE 2.1 – Valeur de  $n$  en fonction de  $\gamma$

A partir du tableau (1.1), nous pouvons voir que la valeur minimale de  $n$  est pour  $\gamma = 3$ . En effet, pour  $\gamma = 3, n \simeq 3.10^{12}$ . Autrement dit,  $n$  est d'ordre 12 ce qui n'est pas réalisable dans la pratique. C'est pourquoi il est important de réduire la variance.

## 2.3 Échantillonnage préférentiel (Importance Sampling)

L'échantillonnage préférentiel est un principe général, particulièrement adapté au cas où la fonction  $h$  est presque nulle (ou nulle) en dehors d'un domaine  $\mathcal{D}$  qui est tel que  $\mathbb{P}[X \in \mathcal{D}]$  soit proche de 0. Cela se produit par exemple lorsque le volume de  $\mathcal{D}$  est petit ou lorsque  $\mathcal{D}$  est défini par les queues de distributions de  $v$ . Il est alors difficile d'obtenir des réalisations de  $X$  dans  $\mathcal{D}$  et donc d'estimer  $\delta$ .

La méthode repose sur l'idée que pour calculer  $\delta$ , au lieu d'échantillonner suivant  $v$ , il est plus intéressant d'échantillonner suivant une loi dont les réalisations seront dans  $\mathcal{D}$  avec forte probabilité, puis de pondérer cet échantillon pour corriger le changement de loi d'échantillonnage.

### Exemple 2.4 Probabilité d'événement rare

On souhaite calculer  $\delta = \mathbb{P}[X > 4.5]$  pour  $X \sim \mathcal{N}(0, 1)$ . Etant donnée  $(X_n)_{n \geq 1}$  une suite de  $n = 10000$  variables aléatoires i.i.d. suivant la loi  $\mathcal{N}(0, 1)$ . on a

$$\bar{h}_n = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i > 4.5\}} = 0$$

La simulation « naïve » s'avère très inefficace car il s'agit d'un événement rare. La solution est de forcer la réalisation de cet événement. Soit  $Z \sim \mathcal{T}\xi(4.5, 1)$  une variable aléatoire suivant la loi exponentielle de paramètre 1 et translatée de 4.5, i.e. de densité donnée pour tout réel  $z$  par

$$g(z) = \frac{e^{z-4.5}}{\int_{4.5}^{+\infty} e^{-x} dx} 1_{\{Z \geq 4.5\}}.$$

On peut écrire

$$\mathbb{P}[X > 4.5] = \int_{4.5}^{+\infty} \phi(x) dx = \int_{4.5}^{+\infty} \frac{\phi(x)}{g(x)} g(x) dx$$

Ainsi, pour une suite  $(Z_n)_{n \geq 1}$  de  $n = 10000$  variables aléatoires i.i.d. suivant la loi de  $Z$ , on a l'estimation

$$\hat{\delta}_n = \frac{1}{n} \sum_{k=1}^n \frac{\phi(Z_k)}{g(Z_k)} 1_{\{Z_k > 4.5\}} = 3.3610^{-6}$$

### Définition 2.1 Estimateur d'échantillonnage préférentiel

Soit  $g$  une densité telle que  $\text{supp}(f) = \{x \in \mathbb{R}^d : f(x) > 0\} \subset \text{supp}(g) = \{x \in \mathbb{R}^d : g(x) > 0\}$ . Alors on peut écrire (pour  $Z \sim g$ )

$$\mathbb{E}_f[h(X)] = \mathbb{E}_g\left[\frac{h(Z)f(Z)}{g(Z)}\right] \quad (2.5)$$

Etant donnée  $(Z_n)_{n \geq 1}$  une suite de variables aléatoires i.i.d. suivant la densité  $g$ , on définit l'estimateur d'échantillonnage préférentiel par

$$\widehat{\delta}_n(g) = \frac{1}{n} \sum_{k=1}^n \frac{f(Z_k)}{g(Z_k)} h(Z_k) = \frac{1}{n} \sum_{k=1}^n w_k h(Z_k)$$

La densité  $g$  est appelée loi instrumentale (ou loi d'importance) et le rapport  $w_k = f(Z_k)/g(Z_k)$  est appelé poids d'importance.

**Remarque 2.2** Le coût de calcul de  $\widehat{\delta}_n(g)$  peut être différent de celui de l'estimateur classique : le coût de simulation suivant  $g$  pouvant être différent du coup de simulation suivant  $f$ . Cela influe donc sur le rapport  $C/C_1$  de la fonction  $R(\bar{h}_n, \widehat{\delta}_n(g))$ .

**Remarque 2.3 Biais de l'estimateur :** Sous l'hypothèse  $\text{supp}(f) \subset \text{supp}(g)$ , on obtient directement que l'estimateur est sans biais en utilisant l'équation 2.5.

### 2.3.1 Convergence de l'estimateur

Les variables aléatoires  $(Y_n)_{n \geq 1} = (h(Z_n)f(Z_n)/g(Z_n))_{n \geq 1}$  sont i.i.d. et d'espérance finie sous  $g$ . La loi forte des grands nombre donne

$$\widehat{\delta}_n(g) \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}[Y_1] = \mathbb{E}_g\left[\frac{h(Z)f(Z)}{g(Z)}\right] = \mathbb{E}_f[h(X)] \quad (2.6)$$

### 2.3.2 Variance de l'estimateur

Les variables aléatoires  $(Z_n)_{n \geq 1}$  étant i.i.d. suivant la densité  $g$ , la variance de l'estimateur s'écrit

$$\mathbb{V}ar_g[\widehat{\delta}_n(g)] = \frac{1}{n} \left\{ \mathbb{E}_g\left[\frac{h^2(Z)f^2(Z)}{g^2(Z)}\right] - \widehat{\delta}_n^2 \right\} = \frac{1}{n} \int_{\mathcal{D}} \frac{\{h(x)f(x) - \widehat{\delta}_n g(x)\}^2}{g(x)} dx \quad (2.7)$$

En pratique, un estimateur naturel de  $\sigma_1^2 = \mathbb{V}ar[Y_1]$  est

$$\widehat{\delta}_1^2(g) = \frac{1}{n-1} \sum_{k=1}^n \{w_k h(Z_k) - \widehat{\delta}_n(g)\}^2$$

### 2.3.3 Choix de la densité instrumentale

La méthode d'échantillonnage préférentiel repose sur le fait qu'il n'y a pas de l'unicité de l'écriture intégrale et de la loi par rapport à laquelle on intègre pour calculer  $\delta$ . Cela laisse beaucoup de liberté quant au choix de  $g$  parmi les lois faciles à simuler. L'égalité 2.7 nous renseigne sur le choix de bonnes lois instrumentales. La première égalité indique que  $g$  est d'autant meilleure que  $\mathbb{E}_g[h^2(X)f^2(X)/g^2(X)]$  est petite (borne supérieure pour la variance). Mais elle illustre également un inconvénient majeur de l'estimateur d'échantillonnage préférentiel. Elle n'a pas nécessairement de variance finie ! On obtient un estimateur de variance finie que lorsque  $\mathbb{E}_g[h^2(X)f^2(X)/g^2(X)]$  est finie. Ainsi



les densités instrumentales  $g$  telles que  $f/g$  n'est pas borné (i.e.,  $g$  a des queues de distribution plus légères que celles de  $f$ ) sont à proscrire. Dans ce cas, les poids d'importance  $f(Z_k)/g(Z_k)$ ,  $k = 1, 2, \dots, n$ , sont très variables et seuls certains termes dans la définition de  $\widehat{\delta}_n(g)$  ont un poids significatif. Autrement dit, la méthode d'estimation est instable :  $\widehat{\delta}_n(g)$  varie brusquement d'une itération à l'autre. Lorsque  $\text{Var}_f[h(X)] < +\infty$ , une condition suffisante pour garantir que  $\widehat{\delta}_n(g)$  soit de variance finie est de choisir une densité  $g$  telle que  $f/g$  soit bornée.

La seconde égalité dans 2.7 indique qu'en prenant  $g$  proportionnel à  $hf$  on obtient un estimateur de variance presque nulle, voire nulle si la fonction  $h$  est de signe constant. Plus précisément, la loi instrumentale optimale en terme de variance est donnée par le résultat suivant.

**Proposition 2.3** *L'estimateur d'échantillonnage préférentiel de variance minimale,  $\widehat{\delta}_n(g)$ , est obtenu pour la densité instrumentale*

$$g^*(z) = \frac{|h(z)|f(z)}{\int_{\text{supp}(f)} |h(x)|f(x)dx}, z \in \mathbb{R}^d$$

Ce résultat n'a qu'un intérêt limité en pratique. Lorsque  $h > 0$ , la densité instrumentale optimale est  $g^* = hf/\mathbb{E}_f[h(X)]$  et  $\text{Var}_{g^*}[\widehat{\delta}_n(g^*)] = 0$ . Néanmoins, cela requiert de connaître  $\delta = \mathbb{E}_f[h(X)]$  qui est justement la quantité d'intérêt. La proposition 2.3 fournit en revanche une stratégie pour choisir  $g$  : un candidat pertinent est tel que  $|h|f/g$  soit quasi constant et de variance finie.

**Estimateur auto-normalisé** : Il est possible que les densités  $f$  et/ou  $g$  ne soient connues qu'à une constante de normalisation près, i.e.,  $f = c_f \widetilde{f}$  et  $g = c_g \widetilde{g}$ . Dans ce cas, on travaille avec une version de l'estimateur dite autonormalisée.

**Définition 2.2** *Sous les mêmes hypothèses que la définition 2.1 et étant donnée  $(Z_n)_{n \geq 1}$  une suite de variables aléatoires i.i.d. suivant  $g$ , l'estimateur auto-normalisée est défini par*

$$\widehat{\delta}_n(\widetilde{g}) = \frac{\sum_{k=1}^n \widetilde{w}_k h(Z_k)}{\sum_{k=1}^n \widetilde{w}_k} \quad (2.8)$$

$$= \frac{\sum_{k=1}^n w_k h(Z_k)}{\sum_{k=1}^n w_k} \quad (2.9)$$

**Proposition 2.4** *L'estimateur 2.8 converge presque sûrement vers  $\delta$ .*

À la différence de l'estimateur d'échantillonnage préférentiel (2.6), l'estimateur auto-normalisé est biaisé. Il présente néanmoins l'avantage d'être exact lorsque la fonction  $h$  est constante (ce qui n'est pas le cas pour (2.6)). Certes, il n'y a pas d'intérêt à calculer des intégrales pour des fonctions  $h$  constantes, en revanche cette propriété permet d'éviter des comportements atypiques.

**Exemple 2.5** *Supposons que l'on souhaite estimer  $\delta = \mathbb{E}_f[h(X)+C]$  avec  $C$  une constante. Soit  $(Z_n)_{n \geq 1}$  une suite de variables aléatoires i.i.d. suivant  $g$ . Alors*

$$\frac{1}{n} \sum_{k=1}^n w(Z_k)\{h(Z_k) + C\} = \frac{1}{n} \sum_{k=1}^n w_k h(Z_k) + \frac{C}{n} \sum_{k=1}^n w_k \neq C + \frac{1}{n} \sum_{k=1}^n w_k h(Z_k)$$

*Autrement dit, l'estimateur d'échantillonnage préférentiel de  $\delta$  n'est pas égal à l'estimateur d'échantillonneur préférentiel de  $\mathbb{E}_f[h(X)]$  auquel on ajoute  $C$ . L'estimation ne préserve pas les propriétés de l'espérance. En revanche, si on considère la version auto-normalisée la propriété attendue est bien vérifiée :*

$$\frac{\sum_{k=1}^n w_k \{h(Z_k) + C\}}{\sum_{k=1}^n w_k} C + \frac{1}{n} \frac{\sum_{k=1}^n w_k h(Z_k)}{\sum_{k=1}^n w_k} h(Z_k)$$

### 2.3.4 Diagnostique et taille efficace d'échantillon (effective sample size)

S'il est possible de comparer les performances en terme de variance de l'échantillonnage préférentiel avec la méthode classique, cette comparaison est fortement sensible à la qualité d'estimation de la variance et donc au choix de  $g$ . Il peut être préférable de mener d'autres diagnostics en amont, notamment pour s'assurer que le choix de  $g$  est judicieux. Par exemple, lorsque l'on sait calculer explicitement les poids d'importances ( $f$  et  $g$  sont connues explicitement et pas à une constante près), on peut utiliser que  $\mathbb{E}_g[f(Z)/g(Z)] = 1$  et donc si la moyenne empirique des poids  $\bar{w}_n$  diffère beaucoup de 1, le choix de  $g$  n'est pas bon.

On utilise généralement la taille efficace d'échantillon pour juger le choix de  $g$  :

$$ESS = \frac{(\sum_{k=1}^n w_k)^2}{\sum_{k=1}^n w_k^2} = n \frac{\bar{w}_n^2}{w_n^2}$$

ESS représente le nombre d'observations apportant effectivement de l'information. Ainsi si la taille efficace est très petite devant  $n$ , on considèrera que l'estimation n'est pas fiable. Il existe de nombreuses définitions de la taille efficace d'un échantillon dans la littérature. Il n'est pas possible dans faire une description exhaustive dans ce cours. On retiendra néanmoins, qu'il ne s'agit pas d'un outil de diagnostic idéal, il permet aumieux d'identifier un problème concernant les poids d'importance mais jamais de valider la méthode.

## 2.4 Conclusion

Dans ce chapitre, nous avons définie les principes de la méthode Monte Carlo et la méthode d'échantillonnage préférentiel, ainsi que un bref aperçu sur l'estimateur. Nous avons exposé aussi quelques exemples célèbres pour chaque méthode ainsi que leur efficacité.

---

---

## CHAPITRE 3

---

# ESTIMATION DE LA PROBABILITÉ DE RUINE PAR LA MÉTHODE D'ÉCHANTILLONNAGE PRÉFÉRENTIEL (IMPORTANCE SAMPLING)

### Introduction

Le premier but de la théorie de la ruine a été de modéliser l'évolution de la richesse de la compagnie d'assurance par un processus stochastique  $\{U(t), t \geq 0\}$ , d'évaluer sa probabilité de ruine, c'est-à-dire la probabilité que le scénario introduisant un échec se réalise.

Considérons le processus de réserves associé au modèle de risque classique  $\{U(t), t \geq 0\}$ , défini par :

$$U(t) = u + ct - Z(t), \quad t \geq 0, \quad (3.1)$$

où  $U(0) = u$  est le surplus initial,  $c$  est le taux de primes constant par unité de temps, il est défini de telle sorte qu'il inclut un chargement de sécurité positif afin que la ruine ne soit pas presque certaine.

Le processus stochastique  $\{Z(t), t \geq 0\}$  avec  $Z(t) = \sum_{i=1}^{N(t)} Z_i$  représente le montant cumulé des sinistres, avec la convention  $Z(t) = 0$  si  $N(t) = 0$ .

Le nombre de réclamations ( ou de sinistres )  $N(t)$  survenus jusqu'au temps  $t \geq 0$  est représenté par un processus de Poisson de paramètre  $\lambda > 0$ .  $\{Z_i\}_{i \in \mathbb{N}^*}$  est une suite de variables aléatoires indépendantes et identiquement distribuées et indépendante du processus de Poisson  $\{N(t), t \geq 0\}$ , de fonction de répartition  $F_Z$  et de moyenne  $\mu$  supposée finie.

L'objectif d'une telle modélisation est de calculer les différentes mesures de ruine (par exemple la probabilité de ruine ou l'espérance de la valeur actualisée du déficit à la ruine) afin d'évaluer quantitativement le risque d'un portefeuille d'assurances.

Dans ce modèle, le processus du montant total des sinistres agrégés,  $\{Z(t), t \geq 0\}$ , où  $Z(t) = \sum_{i=1}^{N(t)} Z_i$  ( $b < a \Rightarrow \sum_a^b = 0$ ), est un processus de Poisson composé.

Le processus du nombre de sinistres,  $\{N(t), t \geq 0\}$ , est un processus de Poisson, donc les temps inter-sinistres,  $\{W_j, j \in \mathbb{N}^+\}$ , forment une suite de variables aléatoires indépendantes et identiquement distribuées de loi exponentielle, où  $W_j$  représente le temps entre le  $(j-1)^e$  et le  $j^e$  sinistre ( $W_1$  est le moment du premier sinistre).

Les variables aléatoires  $\{W_j, j \in \mathbb{N}^+\}$  sont distribuées comme la variable aléatoire canonique  $W$  avec fonction de densité de probabilité  $f_W(t) = \lambda e^{-\lambda t}$  et fonction de répartition  $F_W(t) = 1 - e^{-\lambda t}$ . On peut interpréter  $\frac{1}{\lambda}$  comme étant le temps moyen entre deux sinistres. Le moment où survient le  $j^e$  sinistre est noté

$$T_j = W_1 + W_2 + \dots + W_j$$

Les variables aléatoires des montants des sinistres  $\{Z_j, j \in \mathbb{N}^+\}$ , où  $Z_j$  est le montant du  $j^e$  sinistre, sont supposées indépendantes, positives et distribuées identiquement à la variable aléatoire canonique  $Z$  avec  $f_Z$  comme fonction de densité de probabilité et  $F_Z$  comme fonction de répartition.

Le temps de la ruine en sachant que le surplus initial est  $u$ , noté  $\tau_u$ , est défini comme le premier temps où  $U(t) < 0$  et  $\tau_u = \infty$  si  $U(t) \geq 0, \forall t \geq 0$ . Bien que ce ne soit pas la seule définition possible de la ruine, cette dernière est de loin la plus commune et elle est simple à interpréter. Tomber en ruine signifie, dans ce cas, que le niveau de surplus passe sous zéro. La probabilité de ruine sur un horizon de temps infini  $\psi(u)$  est donc

$$\psi(u) = \mathbb{P}(\{U(T_1) < 0\} \cup \{U(T_2) < 0\} \cup \dots \cup \{U(T_n) < 0\}),$$

qui peut être exprimée en fonction du temps de la ruine

$$\psi(u) = \mathbb{P}(\tau_u < \infty). \quad (3.2)$$

Une autre approche fréquemment utilisée pour définir  $\psi(u)$  est celle utilisant la perte cumulée. Le processus de perte,  $\{L_j, j \in \mathbb{N}^+\}$ , où  $L_j = Z_j - cW_j$  est la perte nette lors du  $j^e$  sinistre.

La marche aléatoire de la perte nette cumulée est  $\{V_j, j \in \mathbb{N}^+\}$ , où  $V_0 = 0$  et

$$V_j = \sum_{i=1}^j L_i, \quad j \in \mathbb{N}^+. \quad (3.3)$$

Comme mentionné précédemment, le taux de prime inclut un chargement positif, c'est-à-dire respecte l'inégalité  $E[cW - Z] > 0$ . Le taux de prime  $c$  peut donc être exprimé comme

$$c = (1 + \theta) \frac{E[Z]}{E[W]}$$

où  $\theta > 0$  est le chargement de sécurité relatif.

Compte tenu du chargement positif, la dérive de  $\{V_j, j \in \mathbb{N}^+\}$  est négative.

Le processus de la perte nette cumulée maximum associé à  $\{V_j, j \in \mathbb{N}^+\}$  est défini par  $\{S_j, j \in \mathbb{N}\}$  où

$$S_j = \max_{i=0,1,\dots,j} \{V_i\} \text{ et } S_\infty = \max_{i \in \mathbb{N}} \{V_i\}. \quad (3.4)$$

La probabilité de ruine est donc équivalente à la probabilité que la perte maximum cumulée dépasse le surplus initial, soit

$$\begin{aligned} \psi &= \mathbb{P}(\max\{V_1, V_2, \dots\} > u) \\ &= \mathbb{P}(S_\infty > u) \end{aligned}$$

Il est aussi possible d'étudier la probabilité de ruine sur un horizon fini

$$\psi_n(u) = \mathbb{P}(\tau_u < n)$$

où  $n$  est l'horizon de temps considéré.

Dans le modèle de risque classique, en utilisant des techniques de résolution d'équations intégrées-différentielles, il est possible d'obtenir une équation pour la probabilité de non ruine  $\phi(u) = 1 - \psi(u)$

$$\phi(u) = \phi(0) + \frac{\lambda}{c} \int_0^u \phi(u-x)(1 - F_X(x))dx, \quad (3.5)$$

où  $\phi(0) = \frac{\theta}{1+\theta}$ .

Dans le cas particulier du modèle de risque classique où la distribution de  $Z$  est exponentielle de paramètre  $\frac{1}{\mu}$ , alors il est possible de simplifier (3.5) en la forme suivante

$$\phi(u) = 1 - \frac{1}{1+\theta} e^{-\left(\frac{\theta}{1+\theta}\right)\mu u}, \quad (3.6)$$

d'où

$$\psi(u) = \frac{1}{1+\theta} e^{-\left(\frac{\theta}{1+\theta}\right)\mu u}. \quad (3.7)$$

Dans le modèle de risque classique, l'équation de Lundberg est donnée par

$$E[e^{x(Z-cW)}] = 1 \quad (3.8)$$

qui est une équation fondamentale à la théorie de la ruine. La plus petite solution en  $x$  strictement positive de l'équation (3.8) est appelée coefficient d'ajustement ou coefficient de Lundberg et est notée  $\rho$ . Le coefficient de Lundberg est notamment utilisé dans le calcul de la borne exponentielle de Lundberg

$$\mathbb{P}(\tau < \infty) \leq e^{-\rho u}, \quad u \geq 0 \quad (3.9)$$

et dans l'expression exacte de Gerber pour la probabilité de ruine

$$\mathbb{P}(\tau < \infty) = \frac{e^{-\rho u}}{E[e^{-\rho U_{\tau u}} | \tau_u < \infty]}. \quad (3.10)$$

On note que le coefficient de Lundberg peut donner une idée du niveau de dangerosité d'un portefeuille. En effet, plus  $\rho$  est grand, moins le portefeuille est dangereux. Voir [2, 26, 8] pour les détails.

Pour obtenir l'équation de Lundberg, nous considérons le processus en temps discret implicite au processus en temps continu  $\{U(t), t \geq 0\}$ . Définissons  $\{\tilde{U}_k, k = 0, 1, 2, \dots\}$ , où  $\tilde{U}_0 = u$  et  $\tilde{U}_k = U(T_k)$  dénote le surplus directement après le  $k^{\text{ème}}$  sinistre, c'est-à-dire

$$\tilde{U}(t) = u + \sum_{j=1}^k (cW_j - Z_j), \quad k \in \mathbb{N}^+, \quad (3.11)$$

Soit  $\{\tilde{Y}_k, k = 0, 1, 2, \dots\}$ , où  $\tilde{Y}_k = e^{-\rho \tilde{U}_k}$ . De l'expression de  $\{\tilde{Y}_k, k = 0, 1, 2, \dots\}$ , on constate que le processus est adapté à la filtration engendrée par la suite de vecteurs aléatoires  $\{(Z_k, W_k), k = 0, 1, 2, \dots\}$ . En effet, la connaissance des réalisations de la suite de vecteurs aléatoires  $\{(Z_k, W_k), k = 0, 1, 2, \dots\}$ , permet d'obtenir les réalisations de  $e^{-\rho \tilde{Y}}$ .

### 3.1 Méthode Monte Carlo simple

La méthode la plus simple et intuitive pour simuler la probabilité de ruine,  $\psi(u)$ , est probablement la méthode Monte Carlo simple (voir [2]). L'idée derrière cette méthode est de simuler un nombre fixé de parcours du processus de surplus et de calculer le ratio du nombre de parcours menant à la ruine sur le nombre de parcours simulés. Ce ratio correspond évidemment à l'approximation de la probabilité de ruine dans ce cas. Cependant, pour certains parcours le niveau de surplus,  $U(t)$ , tendra vers l'infini lorsque  $t$  tend vers l'infini sans que le parcours ne soit tombé en ruine. Un critère doit donc être fixé pour éviter que la simulation entre dans une boucle infinie.

Deux critères sont proposés. Premièrement, on peut simuler chaque parcours jusqu'à un nombre fixé de sinistres  $n$ . Les avantages principaux de ce critère sont qu'il est facile à implémenter dans la programmation et qu'il permet un bon contrôle du temps de simulation. Par contre, il est possible que le niveau du surplus soit très près de zéro après le nombre fixé de sinistres  $n$ . Ainsi, la ruine pourrait survenir dans les prochains sinistres alors arrêter le parcours n'est pas souhaitable.

Pour régler cet inconvénient, un autre critère peut être utiliser, simuler jusqu'à ce que le surplus passe au-dessus d'une barrière fixée. L'inconvénient principal de ce critère est que le temps de simulation devient impossible à contrôler (il est possible qu'un parcours reste entre zéro et la barrière très longtemps avant d'être absorbé par une des barrières).

Soit  $\psi_n(u)$ , la probabilité qu'il y ait ruine avant ou lors du  $n^{\text{ième}}$  sinistre. Comme

$$\lim_{n \rightarrow \infty} \psi_n(u) = \psi(u), \quad (3.12)$$

on peut facilement justifier l'utilisation du premier critère si le nombre de sinistres simulés dans un parcours est suffisamment grand. La méthode d'estimation consiste donc à estimer  $\psi(u)$  par  $\psi_n(u)$ , lui même estimé par des simulations de Monte Carlo en utilisant l'algorithme suivant :

**Algorithme 1.**

1. Générer  $\{(Z_k^{(j)}, W_k^{(j)}), k = 1, 2, \dots, n\}$ ,  
 $\{S_k^{(j)}, k = 1, 2, \dots, n\}$  et  $\{U^{(j)}(T_k^{(j)}), k = 1, 2, \dots, n\}$ .
2. Calculer les parcours  $\{T_k^{(j)}, k = 1, 2, \dots, n\}$ ,  $\{V_k^{(j)}, k = 1, 2, \dots, n\}$ .
3. Si pour un parcours la ruine est observée, alors assigner  
 $\sigma_u^{(j)} = \inf_{k \in \{1, 2, \dots, n\}} \{k, S_k^{(j)} > u\}$  et  $\tau_u^{(j)} = T_{\sigma_k^{(j)}}$ .  
 sinon, assigner  $\sigma_u^{(j)} = \infty$  et  $\tau_u^{(j)} = \infty$ .
4. Répéter les étapes précédentes  $m$  fois, c'est-à-dire pour  $j = 1, 2, \dots, m$ .

L'approximation de  $\psi_n(u)$  est donnée par

$$\psi(u) = \frac{1}{m} \sum_{j=1}^m \mathbf{1}_{\{\sigma_u^{(j)} \leq n\}} = \frac{1}{m} \sum_{j=1}^m \mathbf{1}_{\{Z_u^{(j)} \leq u\}} = \frac{1}{m} \sum_{j=1}^m \mathbf{1}_{\{\tau_u^{(j)} \leq T_n^{(j)}\}},$$

où  $\mathbf{1}_{\{\tau_u^{(j)} \leq T_n^{(j)}\}} = 1$  si la ruine a été observée après un des  $n$  premiers sinistres lors du  $j^{\text{ème}}$  parcours du processus de surplus. Par la loi des grands nombres et (3.12),  $\psi_n^{CMC, m}(u)$  converge vers  $\psi(u)$ , lorsque  $n$  et  $m$  deviennent grands.

La méthode Monte Carlo simple a une variance de  $\sigma_Z^2 = \psi(u)(1 - \psi(u))$ , qui tend vers zéro lorsque  $u \rightarrow \infty$ .

Par contre, comme  $\psi(u)$  est très petit pour des valeurs élevées de surplus initial, l'erreur  $\sigma_Z$  est considérablement plus grande que  $\psi(u)$ , ce qui se traduit par

$$\frac{\sigma_Z}{\psi(u)} = \frac{\sqrt{\psi(u)(1 - \psi(u))}}{\psi(u)} \rightarrow \infty \quad \text{pour } u \rightarrow \infty.$$

Ceci est dû au fait que le nombre de parcours menant à la ruine diminue lorsque  $u$  augmente. Donc, la méthode de Monte Carlo donne une estimation de la probabilité de ruine dont l'erreur relative n'est pas bornée. De plus, cette méthode introduit un biais puisqu'elle approxime  $\psi_n(u)$  qui est plus petit que  $\psi(u)$ , puisque le fait qu'un parcours n'est pas fait ruine dans les  $n$  premiers sinistres n'implique pas que le parcours ne fera jamais ruine. De plus, lorsque  $u$  augmente,  $\psi(u)$  diminue de telle sorte que  $n$  doit être fixé suffisamment grand pour que le nombre de ruine observée puisse être significatif.

## 3.2 Méthode d'échantillonnage préférentiel par changement de mesure

Une autre technique mathématique importante qui sera utilisée dans ce qui suit est le changement de mesure. Comme son nom l'indique, cette technique permet de remplacer la distribution d'une variable aléatoire par une autre qui accorde plus de poids aux réalisations d'intérêt. L'avantage de la méthode est qu'il est possible d'ajuster les poids dans le calcul de l'espérance d'une fonction de la variable aléatoire et d'obtenir la même valeur. Cette méthode est particulièrement utile pour évaluer différentes quantités reliées à des événements très peu probables. Comme les probabilités de ruine diminuent lorsque le surplus initial augmente, le changement de mesure est utile, notamment, pour évaluer les probabilités de ruine lorsque  $u$  est très grand. De plus, les changements de mesure permettent d'obtenir des expressions exactes pour  $\psi(u)$  qui peut être évaluée relativement simplement en utilisant la simulation. Lorsqu'on utilise la simulation en combinaison avec un changement de mesure, on parle d'échantillonnage préférentiel. Pour plus de détails sur la technique de changement mesure et leurs applications en simulation, voir, [26, 9, 17, 22].

Cette technique offerent une alternative à l'expression exacte de Gerber (voir [8])pour développer une méthode d'approximation de divers mesures de risque. Une méthode basée sur la simulation sous une nouvelle mesure est dite une méthode d'échantillonnage préférentiel. L'avantage d'une telle méthode provient du fait que comme la dérive du processus de surplus est négative sous la nouvelle mesure de probabilité  $\mathbb{P}^{(\rho\delta)}$ , la ruine est certaine (c'est-à-dire  $\mathbb{P}(\tau < \infty) = 1$ ) sous la nouvelle mesure.

La méthode d'échantillonnage préférentiel par changement de mesure tire avantage de ce résultat et simule chaque parcours jusqu'à la ruine (et non pas pour un nombre fixé de sinistres comme dans la méthode Monte Carlo simple.



L'algorithme général utilisé pour simuler de la distribution conjointe de  $(Z, W)$  sous la nouvelle mesure  $\mathbb{P}^{(\rho\delta)}$  est le suivant :

**Algorithme 2.**

1. *Simuler*  $\{(Z_k^{(j)}, W_k^{(j)}), k \in \mathbb{N}^+\}$ .
2. *Calculer*  $\{U^{(j)}(T_k^{(j)}), k \in \mathbb{N}^+\}$ , où  $T_k^{(j)} = W_1^{(j)} + W_2^{(j)} + \dots + W_k^{(j)}$ .
3. *Arrêter les étapes 1 et 2 lorsque la ruine survient. Noter le numéro du sinistre causant la ruine,  $\sigma_u^{(j)}$ .*
4. *Noter le temps de la ruine pour le  $j^e$  parcours,  $\tau_u^{(j)} = T_{\sigma_u^{(j)}}^{(j)}$ .*
5. *Noter le déficit à la ruine pour le  $j^e$  parcours,  $U^{(j)}(\tau_u^{(j)}) = u - V_{\sigma_u^{(j)}}^{(j)}$ .*
6. *Répéter les étapes 1 – 5  $m$  fois, c'est-à-dire pour  $j = 1, 2, \dots, m$ .*

L'algorithme permet d'identifier les parcours pour lesquels la ruine a été observée et d'approximer  $E[e^{-\rho_0 U(\tau_u)} \mid \tau_u < \infty]$  dans (3.10) par la moyenne empirique de  $e^{-\rho_0 U(\tau_u)}$  calculée seulement sur ces parcours, ce qui permet de calculer la probabilité de ruine  $\psi(u)$ .

### 3.3 Applications numériques

Cette section est consacrée à la présentation des différents résultats numériques et graphiques obtenus lors de l'étude de l'approximation de la probabilité de ruine dans le modèle de risque classique. En appliquant les algorithmes définis précédemment, on détermine la probabilité de ruine par simulation pour différentes distributions des montants de réclamation.

#### 3.3.1 Sinistres obéissent à une loi exponentielle

1. La densité de probabilité de la loi exponentielle est donnée par :

$$f(z; \frac{1}{\mu}) = \frac{1}{\mu} e^{-\frac{z}{\mu}}, \quad \mu > 0, \quad z \geq 0.$$

2. Le chargement de sécurité relatif  $\theta$  est donné par :

$$\theta = \frac{c}{\lambda\mu} - 1.$$

3. Le coefficient de Lundberg  $\rho$  est donné par :

$$\rho = \frac{\theta}{(\theta + 1) * \mu} = \frac{c - \lambda\mu}{c\mu}.$$

4. La formule explicite de la probabilité de ruine pour des sinistres exponentielle est donnée par :

$$\psi(u) = \frac{1}{1 + \theta} e^{-\left(\frac{\theta}{(1+\theta)\mu}\right)u}. \quad (3.13)$$

Dans le cas des sinistres obéissent à une loi exponentielle, la probabilité de ruine admet une formule explicite donné par (3.13), Cela nous donne la possibilité d'effectuer une comparaison numérique entre la valeur exacte de la probabilité de ruine et son approximation par simulation. Après le calcul de la probabilité de ruine en se basant sur la formule exacte et l'exécution des deux algorithmes définis précédemment, en prenant comme paramètres d'entrées ( $\lambda = 1, c = 1$ , et  $\mu = 0.5$ ), les résultats obtenus sont présentés dans le Tableau (3.1) et illustrés dans la Figure(3.1).

$u$	$\psi_{CM}(u)$	$\psi_{IS}(u)$	$\psi_{Exacte}(u)$
0.1	0.4733300	0.4728949	0.4756147
0.2	0.4543700	0.4533243	0.4524187
0.3	0.4269400	0.4302472	0.4303539
0.4	0.4100500	0.4077440	0.4093653
0.5	0.3886200	0.3894317	0.3894003
1	0.3055500	0.3034354	0.3032653
2	0.1845800	0.1849059	0.1839397
3	0.1110500	0.1118471	0.1115650
4	0.0692500	0.0682083	0.0676676
5	0.0412900	0.0411083	0.0410425
10	0.0033100	0.0034107	0.0033689
20	0.0000300	$2.2667936 \cdot 10^{-5}$	$2.2699964 \cdot 10^{-5}$
30	0	$1.4994302 \cdot 10^{-7}$	$1.5295116 \cdot 10^{-7}$
40	0	$1.0207837 \cdot 10^{-9}$	$1.0305768 \cdot 10^{-9}$
50	0	$6.8211580 \cdot 10^{-12}$	$6.9439719 \cdot 10^{-12}$

TABLE 3.1 – Probabilités de ruine (valeurs exactes et approximations) avec des montants de réclamation exponentielle

On définit l'erreur relative d'une méthode comme suit

$$ER_{\text{méthode}} = |\psi_{Exacte}(u) - \psi_{\text{méthode}}(u)| / \psi_{Exacte}(u)$$

L'erreur relative associée à la probabilité de ruine pour les deux méthodes de simulation sont données par le Tableau (3.2) et illustrés dans la Figure(3.2).

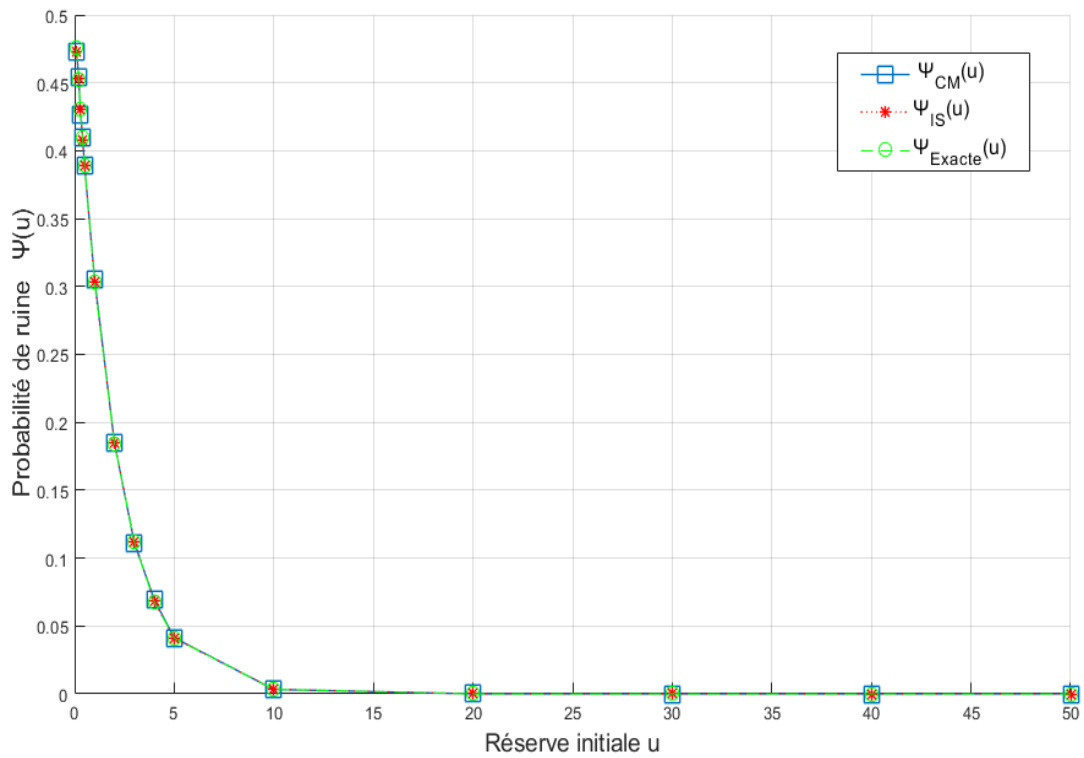


FIGURE 3.1 – Variation de la probabilités de ruine (valeurs exactes et approximations), avec des montants de réclamation exponentielle, en fonction de la réserve initiale  $u$

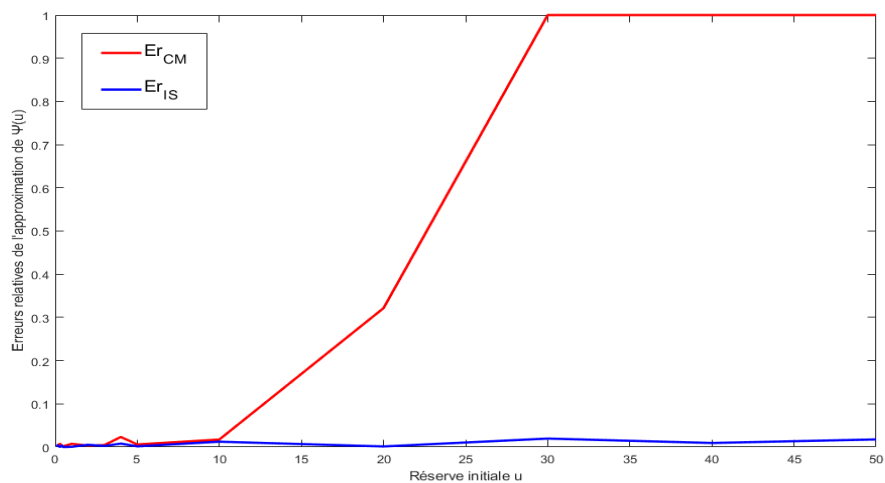


FIGURE 3.2 – Variation d'erreurs relatives de l'approximation de  $\psi(u)$

$u$	$ER_{CM}$	$ER_{IS}$
0.1	$4.8037039 \cdot 10^{-3}$	$5.7185086 \cdot 10^{-3}$
0.2	$4.3130200 \cdot 10^{-3}$	$2.0017107 \cdot 10^{-3}$
0.3	$7.9329768 \cdot 10^{-3}$	$2.481165133 \cdot 10^{-4}$
0.4	$1.6724019 \cdot 10^{-3}$	$3.9605108 \cdot 10^{-3}$
0.5	$2.0040851 \cdot 10^{-3}$	$8.0609506 \cdot 10^{-5}$
1	$7.5335685 \cdot 10^{-3}$	$5.6094979 \cdot 10^{-4}$
2	$3.4809197 \cdot 10^{-3}$	$5.2531172 \cdot 10^{-3}$
3	$4.6168574 \cdot 10^{-3}$	$2.5283484 \cdot 10^{-3}$
4	$2.3384269 \cdot 10^{-2}$	$7.9903612 \cdot 10^{-3}$
5	$6.0303512 \cdot 10^{-3}$	$1.6038350 \cdot 10^{-3}$
10	$1.7504886 \cdot 10^{-2}$	$1.2398958 \cdot 10^{-2}$
20	$3.2158794 \cdot 10^{-1}$	$1.4109441 \cdot 10^{-3}$
30	1	$1.9667325 \cdot 10^{-2}$
40	1	$9.4155368 \cdot 10^{-3}$
50	1	$1.7686404 \cdot 10^{-2}$

TABLE 3.2 – Erreurs relatives de l'approximation de  $\psi(u)$ 

### 3.3.2 Sinistres obéissent à une loi hyper-exponentielle

La loi hyper-exponentielle est une loi de probabilité continue mélangeant plusieurs lois exponentielles. Elle dépend de trois paramètres :  $n$  le nombre de lois exponentielles indépendantes,  $(\mu_i, 1 \leq i \leq n)$  les paramètres de ces lois exponentielles et  $(p_i, 1 \leq i \leq n)$  une pondération de ces lois, vérifiant  $\sum_{i=1}^n p_i = 1$ .

La densité de probabilité de la loi hyper-exponentielle est donnée par :

$$f(z; n, \mu_i, p_i) = \sum_{i=1}^n p_i \mu_i e^{-\mu_i z}, \quad \mu > 0, \quad z \geq 0.$$

La fonction de répartition est donnée par :

$$F(z; n, \mu_i, p_i) = 1 - \sum_{i=1}^n p_i e^{-\mu_i z}, \quad z \geq 0.$$

L'espérance est la somme pondérée des espérances des lois exponentielles :

$$\mathbb{E}(Z) = \sum_{i=1}^n \frac{p_i}{\mu_i}$$

Le chargement de sécurité relatif  $\theta$  est donné par :

$$\theta = \frac{c\mathbb{E}(W)}{\mathbb{E}[Z]} - 1 = \frac{c}{\lambda\mathbb{E}[Z]} - 1.$$

Après l'exécution des deux algorithmes définis précédemment, en prenant comme paramètres d'entrées ( $\lambda = 1, c = 1, \mathbf{H}(n = 2, \mu_1 = 3, \mu_2 = 4, p_1 = 4, p_2 = 1 - 4)$ ), les résultats obtenus sont présentés dans le Tableau (3.3) et illustrés dans la Figure(3.3).

$u$	$\psi_{CM}(u)$	$\psi_{IS}(u)$
0.1	0.5952212	0.5999477
0.2	0.5873451	0.5824673
0.3	0.5742477	0.5708522
0.4	0.5550442	0.5554640
0.5	0.5449557	0.5402105
1	0.4908849	0.4899449
2	0.3918584	0.3812405
3	0.3040707	0.2907337
4	0.2370796	0.2320610
5	0.1952212	0.1820108
10	0.0615044	0.0588381
20	0.0061061	0.0059475
30	0.0002654	0.0002330
40	0	$1.2882210 \cdot 10^{-9}$
50	0	$5.8485143 \cdot 10^{-14}$

TABLE 3.3 – Approximations de la probabilités de ruine avec des sinistres obéissent à une loi hyper-exponentielle

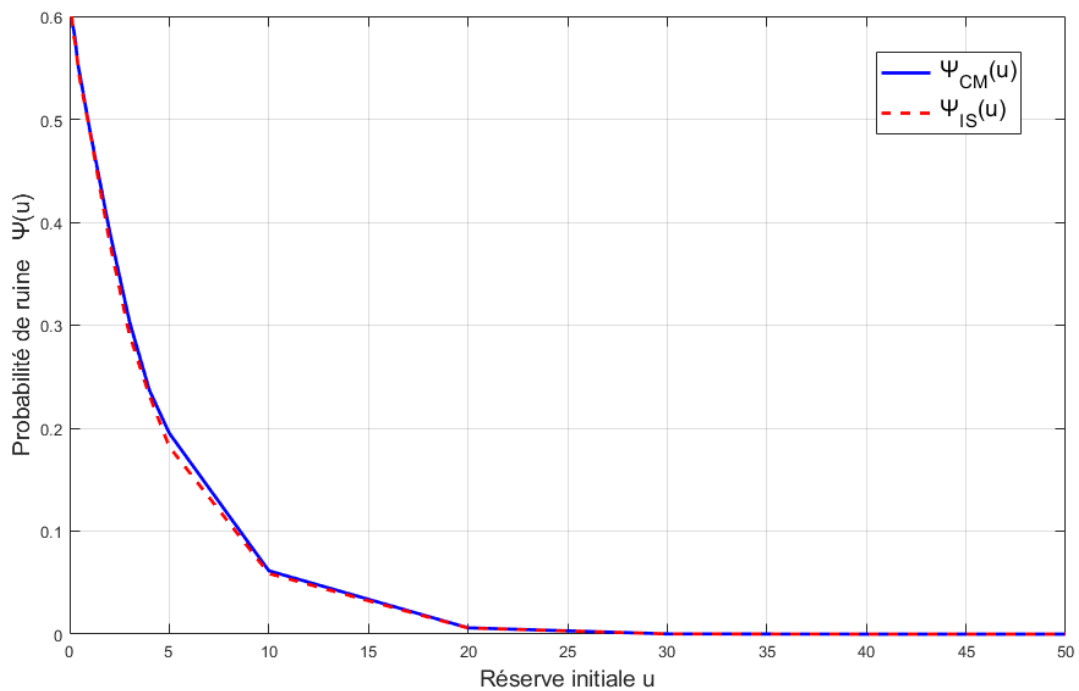


FIGURE 3.3 – Variation de la probabilités de ruine, avec des sinistres hyper-exponentielle, en fonction de la reserve initiale  $u$

### 3.3.3 Sinistres obéissent à une loi Gamma

La densité de probabilité de la loi Gamma est donnée par :

$$f(z; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\beta z}, \quad z \geq 0 \quad \alpha, \beta > 0.$$

L'espérance est donné par :

$$\mathbb{E}(Z) = \frac{\alpha}{\beta}$$

Le chargement de sécurité relatif  $\theta$  est donné par :

$$\theta = \frac{c\mathbb{E}(W)}{\mathbb{E}[Z]} - 1 = \frac{c}{\lambda\mathbb{E}[Z]} - 1 = \frac{c\beta}{\lambda\alpha} - 1.$$

L'exécution des deux algorithmes nous donne, pour des sinistres obéissent à une loi Gamma( $\alpha = 2, \beta = 1.5$ ), les résultats du Tableau (3.4) et la Figure(3.4).

$u$	$\psi_{CM}(u)$	$\psi_{IS}(u)$
0.1	0.1560	0.1514093
0.2	0.1450	0.1442300
0.3	0.1200	0.1218902
0.4	0.1150	0.1097527
0.5	0.0860	0.0952798
1	0.0720	0.0761579
2	0.0280	0.0216714
3	0.0120	0.0126672
4	0.0060	0.0078720
5	0.0010	0.0010161
10	0	$2.1005051 \cdot 10^{-6}$
20	0	$6.0175844 \cdot 10^{-10}$
30	0	$5.6251865 \cdot 10^{-15}$

TABLE 3.4 – Approximations de la probabilités de ruine avec des sinistres obéissent à une loi Gamma

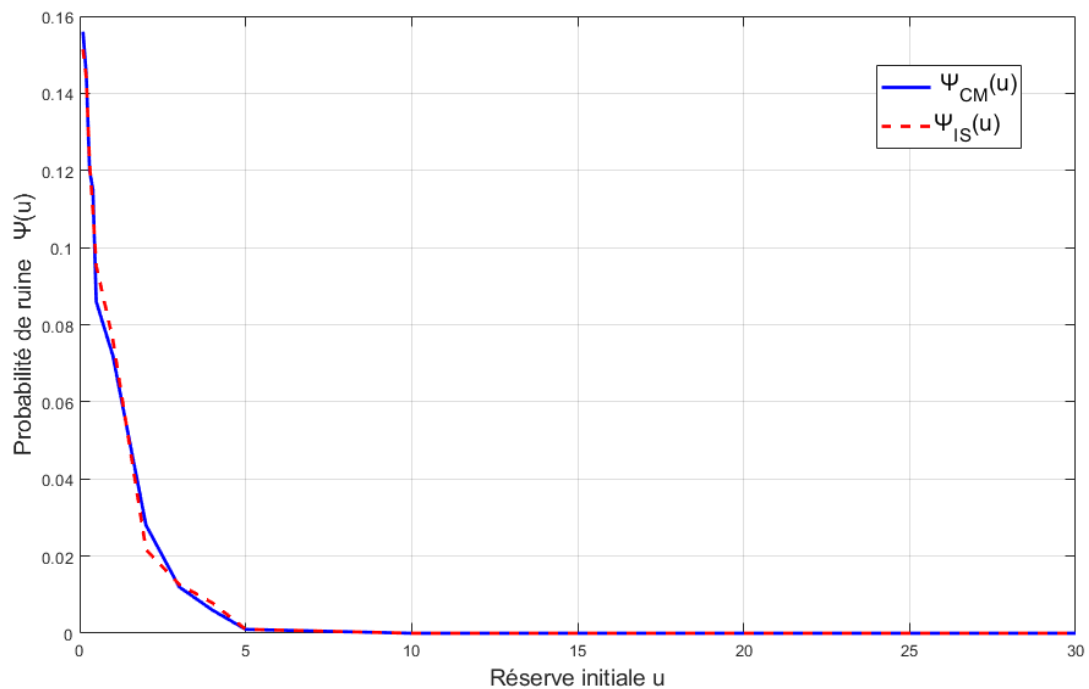


FIGURE 3.4 – Variation de la probabilités de ruine, avec des montants de réclamation de loi Gamma, en fonction de la reserve initiale  $u$

### 3.3.4 Sinistres obéissent à une loi Weibull

La densité de probabilité de la loi de Weibull est donnée par :

$$f(z; \alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{z}{\alpha}\right)^{\beta-1} e^{-\left(\frac{z}{\alpha}\right)^\beta}, \quad z \geq 0, \alpha, \beta > 0.$$

La fonction de répartition est donnée par :

$$F(z; \alpha, \beta) = 1 - e^{-\left(\frac{z}{\alpha}\right)^\beta}, \quad z \geq 0.$$

L'espérance est donné par :

$$\mathbb{E}(Z) = \alpha \Gamma\left(1 + \frac{1}{\beta}\right).$$

Le chargement de sécurité relatif  $\theta$  est donné par :

$$\theta = \frac{c\mathbb{E}(W)}{\mathbb{E}[Z]} - 1 = \frac{c}{\lambda\mathbb{E}[Z]} - 1.$$

Pour des sinistres de loi de Weibull ( $\alpha = 2, \beta = 0.5$ ), l'exécution des deux algorithmes nous donne les résultats du Tableau (3.5) et la Figure(3.5).

$u$	$\psi_{CM}(u)$	$\psi_{IS}(u)$
0.1	0.7880	0.7879587
0.2	0.7870	0.7700770
0.3	0.7680	0.7677631
0.4	0.7590	0.7504663
0.5	0.7340	0.7384563
1	0.6860	0.6845396
2	0.5050	0.5073290
3	0.4060	0.4152037
4	0.3160	0.3349702
5	0.2690	0.2466911
10	0.0310	0.0375859
20	0	4.0433939 $10^{-4}$
30	0	3.2703101 $10^{-7}$
40	0	4.8377708 $10^{-10}$
50	0	5.5295099 $10^{-13}$

TABLE 3.5 – Approximations de la probabilités de ruine avec des sinistres obéissent à une loi de Weibull



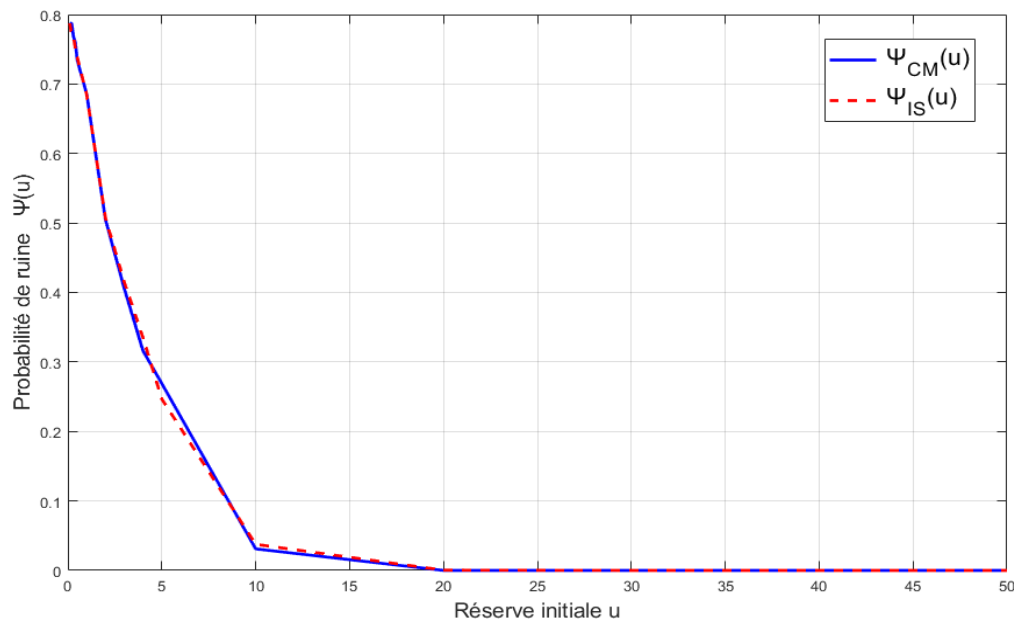


FIGURE 3.5 – Variation de la probabilités de ruine, avec des montants de réclamation de loi de Weibull, en fonction de la réserve initiale  $u$

### 3.3.5 Discussion des résultats

Le Tableau(3.1) montre que toutes les valeurs de probabilité de ruine obtenus pour une réserve initiale inférieure à 10 sont presque identique, ce qui signifie que les valeurs simulées sont très proches de la valeur exacte de la probabilité de ruine. Par contre, pour  $u > 20$ , la probabilité de ruine obtenue par la méthode de Monte Carlo est nulle, et celle obtenue par la méthode d'échantillonnage préférentiel est très petite mais différente de 0 et elle est proche de la valeur exacte de la probabilité de ruine. Dans la Figure(3.1) on remarque que les deux courbes sont similaire pour  $u \leq 10$ , ce qui signifie que les résultats obtenus par les deux méthodes de simulation sont très proches de la valeur exacte de la probabilité de ruine. Mais dès que la réserve initiale dépasse 10 l'erreur relative à la méthode de Monte Carlo augmente avec l'augmentation de la réserve initiale, ce qui signifie que, pour  $u > 10$ , les résultats obtenus par la simulation de Monte Carlo s'éloignent, de la valeur exacte de la probabilité de ruine, à fur et à mesure qu'on augmente la réserve initiale.

Les figures (Figure(3.1),Figure(3.3),Figure(3.4),Figure(3.5)) montrent que, pour toutes les distributions des montants de réclamation considérés, la variation de la probabilité de ruine est inversement proportionnelle à la variation de la réserve initiale. En augmentant la réserve initiale, la probabilité de ruine se rapproche à fur et à mesure de 0.

En augmentant la réserve initiale  $u$ , la méthode de Monte Carlo devient de plus en plus imprécis.

La simulation avec méthode de Monte Carlo se fait jusqu'à un nombre fixé de sinistres  $n$ . Mais après le nombre fixé de sinistres  $n$  il est possible que le niveau du surplus soit très près de zéro, le fait qu'un parcours n'est pas fait ruine dans les  $n$  premiers sinistres n'implique pas que le parcours ne fera jamais ruine.

### 3.4 Conclusion

Dans ce chapitre, on a présenté le modèle de risque classique pour lequel on a effectué les deux méthodes de simulation, dont la méthode de Monte Carlo et la méthode d'échantillonnage préférentiel. Par la suite, on a implémenté un algorithme de simulation pour chaque méthode. Enfin une application numérique qui consiste à l'exécution des deux algorithmes de simulation, la présentation et la discussion des résultats obtenus.

---

# CONCLUSION GÉNÉRALE

Dans le domaine de la gestion du risque en assurance, la probabilité de ruine dans les modèles du risque représente un axe de recherche très développé où plusieurs approches ont été élaborées, et la recherche d'approximations pour la probabilité de ruine dans les modèles de risque a été l'un des points principaux en mathématiques de l'assurance.

Dans ce travail, nous sommes intéressés aux méthodes de simulation de la probabilité de ruine dans un modèle de risque classique, il s'agit de la méthode de Monte Carlo et la méthode d'échantillonnage préférentiel.

D'abord, nous avons présenté quelques notions principales de la théorie de la ruine, nous nous sommes intéressés au modèle de risque classique, en particulier à une mesure très importante dans la théorie de la ruine, cette mesure est la probabilité de ruine, pour laquelle nous avons présenté quelques approximations et expressions exactes. Par la suite nous avons donné un aperçu général sur la simulation Monte Carlo et la méthode d'échantillonnage préférentiel. Nous avons implémenté deux algorithmes de simulation, le premier en utilisant la méthode Monte Carlo et le deuxième en utilisant la méthode d'échantillonnage préférentiel. Nous avons utilisé ces deux algorithmes pour différentes distributions des montants des sinistres. Nous avons présenté également des résultats numériques et graphiques.

Les résultats présentés et discutés montrent que la méthode de Monte Carlo est très utile quand la probabilité de ruine n'est pas très petite, mais quand la probabilité de ruine se rapproche de 0, cette méthode ne donne plus de bon résultat. Contrairement à la méthode d'échantillonnage préférentiel, cette dernière donne de bonnes approximations de la probabilité de ruine (précision jusqu'à  $10^{-15}$ ).

C'est très difficile de trouver une bonne transformation pour la méthode d'échantillonnage préférentiel afin de passer sous la nouvelle mesure. En temps d'exécution, cette méthode est très lente parce qu'elle simule chaque parcours jusqu'à la ruine, et non pas pour un

nombre fixé de sinistres comme dans la méthode Monte Carlo.

La méthode de Monte Carlo peut également servir de moyen de comparaison à la méthode d'échantillonnage préférentiel, il suffit de choisir un nombre petit de  $u$ , qui représente la réserve initiale, pour que la probabilité de ruine soit grand, et vérifiez que les deux méthodes donnent presque les mêmes estimations, confirmant ainsi que la méthode d'échantillonnage préférentiel est utile et peut être utilisé pour estimer la probabilité de ruine des grandes valeurs de la réserve initiale.

---

# BIBLIOGRAPHIE

- [1] S. Asmussen. Conditioned limit theorems relating a random walk to its associate, with applications to risk reserve processes and the gi/g/1 queue. *Advances in Applied Probability, JSTOR*, 14 :143–170, 1982.
- [2] S. Asmussen and H. Albrecher. *Ruin probabilities*. World Scientific, Second Ed., New Jersey, 2010, 2010.
- [3] J. Bertoin and R. Doney. Cramér’s estimate for lévy processes. *Stat. Probab., Lett* 21 :363–365, 1994.
- [4] J. Cai. Ruin probabilities under dependent rates of interest. *Journal of Applied Probability* 39 312-323., 39 :312–323, 2002.
- [5] A. Charpentier and M. Denuit. *Principe Fondamentaux de Théorie de Risque*. Mathématiques de l’assurance non vie, Economica, 2004.
- [6] C. Coccozza. *Processus stochastique et fiabilité des systemes*. pringer-Verlag, Berlin Heidelberg, 1997.
- [7] F. DeVylder and E. Marceau. Explicite analytic ruin probabilities for bounded claims. *Insurance : Mathematics and Economics*, 16 :79–105, 1995.
- [8] H.U. Gerber. An introduction to mathematical risk theory. *Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania Philadelphia*, 8, 1979.
- [9] P. Glasserman. *Monte Carlo methods in Financial engineering*, volume 53. Springer Science and Business Media, 2003.
- [10] J. Grandell. *Aspects of Risk Theory*. Springer Series in Statistics. Springer-Verlag, 1991.
- [11] S. Li and J. Garrido. On ruin for the erlang(n) risk process. *Insurance : Mathematics and Economics*,, 18 :391–408, 2004.
- [12] X. S. lin. Tail of compond distribution and excess time. *Journal of applied probability*, 4 :184–195, 1996.

- [13] S. Loisel. Cours de gestion des risques d'assurances et de théorie de la ruine. *Cours en ISFA 3 ème année, Université de Lyon 1*.
- [14] F. Lundberg. Some supplementary researches on the collective risk theory. *Skandinavisk Aktuarietidskrift*, 15 :137–158, 1932.
- [15] Marceau. *Modélisation et évaluation quantitative des risques en actuariat, Modèles sur une période*. Springer, 2013.
- [16] Marceau and Devylder. Explicite analytic ruin probabilities for bounded claims. *Insurance, Mathematics and Economics*, 1996.
- [17] H. Pham. Some applications and methods of large deviations in finance and insurance. In *Paris-Princeton Lectures on Mathematical Finance*, Springer, pages 191–244, 2007.
- [18] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer-Verlag New York, 2 edition, 2004.
- [19] Janssen. S and Manca. R. *insurance and reliability*. Springer, 2007.
- [20] K. I. Sato. Lévy processes and infinitely divisible distributions. *Cambridge University Press*, 1999.
- [21] J. L. Schiff. *The Laplace Transform, Theory and Applications*. Springer-Verlag New York Inc., 1999.
- [22] K. Sigman. Rare event simulation and importance sampling. *Lecture Notes on Monte Carlo Simulation*, 2007.
- [23] B. Silvermann. Density estimation for statistic and data analysis. *Chapman and Hall, london*, 1986.
- [24] A. K. Stuart, H. P. Harry, and E. W. Gordon. *Loss Models : From Data to Decisions*. Wiley–Blackwell, 2012.
- [25] J.L. Teugels. Estimation of ruin probabilities. *Insurance : Mathematics and Economics* 1, 1 :163–175, 1982.
- [26] T. J. Teugels V. Schmidt, V. Rolski and H. Schmidli. Stochastic processes for insurance and finance. *Wiley*, 1999.
- [27] F. De Vylder, M. Goovaerts, and E. Marceau. The bi-atomic minimal solution of schmitter's problem. *Insurance : Mathematics and Economics*, pages 59–78, 1997.
- [28] G. E. Willmot. Refinements and distributional generalizations of lundberg's inequality. *Insurance : Mathematics and Economics*, pages 49–63, 1994.
- [29] G. E. Willmot and X. S. Lin. Simplified bounds on the tails compound distribution. *Journal of Applied Probability*, pages 127–133, 1997.
- [30] B. Ycart. *Modèles et Algorithmes Markoviens*, volume 39 of *Mathématiques et Applications*. Springer-Verlag Berlin Heidelberg, 1 edition, 2002.

## ***RÉSUMÉ***

---

La théorie de la ruine est un des domaines des sciences actuarielles où la complexité mathématique est un facteur limitant les chercheurs. Dans ce mémoire, on s'intéresse donc à des méthodes numériques permettant d'approximer la probabilité de ruine pour les différentes quantités d'intérêt. Cependant, avant d'aborder le cœur du sujet, on fournit une revue de la littérature concernant la théorie de la ruine et on étudie certaines mesures de ruine en temps fini et infini pour des modèles de risque. On présente aussi les bases mathématiques nécessaires à la compréhension de ce mémoire pour toute personne ayant des connaissances de bases en science actuarielle et en statistiques. Puis, le cœur de ce travail, l'évaluation numérique de mesures de ruine à l'aide de deux méthodes numériques basées sur la simulation, respectivement, (1) la méthode de Monte Carlo simple, et (2) la méthode basée sur l'échantillonnage préférentiel. Nous discuterons également de la qualité respective de chaque méthode. En particulier, nous montrerons que la méthode basée sur l'échantillonnage préférentiel fournit des résultats sans biais et avec une erreur relative bornée. On présentera aussi plusieurs illustrations numériques.

---

**Mots clés :** Théorie de la ruine, Simulation, Échantillonnage Préférentiel, Monte Carlo.

## ***ABSTRACT***

---

Ruin theory is a field in actuarial science where researchers are often impeded by mathematical complexity. In this thesis, we look at some numerical methods that can be used to alleviate this problem. Before getting to the core of this work, we provide a review of the literature concerning ruin theory and we study some finite and infinite time ruin measures within risk models. We also present the mathematical background necessary to understand this memoir for anyone with a basic understanding of actuarial science and statistics. The main focus of this work is the computation of ruin measures via two different methods based on simulations, namely (1) the crude Monte Carlo method, and (2) the importance sampling method based on change of measure techniques. Another topic that is discussed is the quality of the approximation of each method. In particular, we show that the importance sampling method provides unbiased approximations and bounded relative errors. We also present numerous numerical illustrations.

---

**Key words :** Ruin Theory, Simulation, Importance Sampling, Monte Carlo.