

Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE SAAD DAHLAB DE BLIDA

Faculté des sciences

Département de Mathématiques

THÈSE DE DOCTORAT en SCIENCES

Spécialité : Mathématiques

Option : Statistique

Par

Redouane Frihi

**Estimation de la probabilité de ruine en temps infini
pour des processus non stationnaires**

Devant le jury composé de :

Nadia OUKID	Présidente	Prof.	Univ. Blida 1
Abdelaziz RASSOUL	Directeur de thèse	Prof.	ENSH, Blida
Hamid OULD ROUIS	Co-Directeur	Prof.	Univ. Blida 1
Diffalah LAISSAOUI	Examineur	MCA	Univ. Médéa
Ghania SAIDI	Examinatrice	Prof.	ENSSEA Koléa
Redouane BOUDJEMAA	Examineur	MCA	Univ. Blida 1

Mars 2023

REMERCIEMENTS

Nous remercions Allah le tout-puissant de m'avoir donné la force et le courage pour bien mener et finir ce travail.

Je tiens à remercier Monsieur **AbdelAziz Rassoul**, Professeur à l'Ecole Nationale Supérieure d'Hydraulique de Blida, qui m'a encadré tout au long de cette thèse et qui m'a fait partager ses brillantes intuitions. Qu'il soit aussi remercié pour sa disponibilité permanente et pour les nombreux encouragements qu'il m'a prodigués.

Je remercie Monsieur **Hamid Ouldrouis** professeur à l'université Saad Dahlab de Blida, Co-encadreur pour les nombreux encouragements et précieux conseils.

Mes remerciements s'adressent également aux membres du jury, qui m'ont fait l'honneur de bien vouloir étudier avec attention mon travail.

Cette thèse est le fruit d'une collaboration de plus de cinq années.

J'exprime ma gratitude à ma famille qui m'a toujours soutenue et encouragée dans la voie que je m'étais fixée.

Je remercie particulièrement ma femme qui m'a encouragée pour la réussite de ce projet.



Table des matières

- Introduction Générale** **1**

- 1 Notions de base de la théorie des probabilités** **8**
 - 1.1 Théorèmes de la Convergence des Grands nombres 9
 - 1.1.1 Théorème Central Limite 10
 - 1.2 Distributions à variations régulières 11
 - 1.3 Distributions à queue lourde 14
 - 1.3.1 Distributions sous-exponentielles 15
 - 1.4 Quelques notions sur les processus 15
 - 1.4.1 Processus à accroissements indépendants 16
 - 1.4.2 Processus à accroissements stationnaires 16
 - 1.4.3 Processus de comptage 16
 - 1.4.4 Processus de Poisson 17
 - 1.4.5 Processus de Poisson composé 18
 - 1.4.6 Processus de comptage Non Stationnaire 18
 - 1.4.6.1 Processus de Hawkes 19
 - 1.4.6.2 Processus de Cox avec intensité de bruit de tir 20

- 2 Théorie des valeurs extrêmes** **22**
 - 2.1 Introduction 22
 - 2.1.1 Statistique d'ordre 23
 - 2.1.2 Côté Probabiliste de la théorie des valeurs extrêmes 25
 - 2.2 Domaine d'Attraction 32
 - 2.2.1 Domaine d'attraction de la distribution de Fréchet 34
 - 2.2.2 Domaine d'attraction de la distribution de Weibull 36
 - 2.2.3 Domaine d'attraction de Gumbel 37

2.2.4	Distribution des valeurs extrêmes généralisées et la distribution de Paréto généralisée	40
2.3	Méthodes statistiques pour des événements extrêmes	47
2.3.1	Estimation de l'indice de queue	47
2.3.2	L'estimateur de Pickands	48
2.3.3	Estimateur de Hill	49
2.3.4	Estimateur de Deckers Einmah-de Haan ($\xi \in \mathbb{R}$).	51
2.3.5	Comparaison des différents estimateurs	52
2.3.6	Estimation des constantes de normalisation	52
2.4	Estimation des paramètres de la distribution GPD	53
2.5	Détection du seuil	54
2.5.1	Méthodes graphiques	55
2.5.2	Mean excess-plot(ME-plot))	55
2.5.3	Illustrations graphiques	56
2.5.4	Méthode de bootstrap double	57
2.5.5	Metrique de Kolmogorov-Smirnov	58
2.5.6	Estimateur de Hill (Stabilité de graphe)	58
2.5.7	Hill-plot	61
2.5.8	Illustrations numériques	62
2.5.9	Méthode du maximum de vraisemblance	63
2.5.10	Méthode des moments pondérés	64
3	Quelques résultats sur la Probabilité de ruine pour les processus de risque classiques	65
3.1	Modèle de risque	66
3.1.1	Probabilité de ruine	67
3.1.2	Probabilité de ruine exacte en temps infini	70
3.1.3	Représentation intégro-différentielle du modèle de Cramer-Lundberg	70
3.1.4	Cas particuliers	74
3.2	Processus de risque de queue Sous exponentielle	77
3.2.1	Cas distribution de Fréchet	77
3.2.2	Condition de stationnarité	77
3.3	Calcul de la probabilité de ruine à horizon infini	79
3.3.1	Traditional Estimator of the Ruin Probability	80
3.3.2	POT-based estimator of the ruin probability	80
3.4	Main results	83
3.5	Simulation Study and Application	84
3.5.1	Simulation Study	84
3.5.2	Application	86

3.6	Proofs of the Main results	89
3.6.1	Proof of Theorem 3.1	89
3.6.2	Proof of Theorem 3.2	90
4	Probabilité de ruine pour un processus du risque Non-Stationnaire	95
4.1	Processus de risque avec arrivées Non-Stationnaire	95
4.1.1	Notion du principe de grandes déviations	96
4.1.2	Probabilité de ruine	98
4.2	Réclamation globale	100
4.3	Exemples de processus d'arrivée non renouvelable	100
4.3.1	Processus de Hawkes	101
4.3.2	Processus de Cox avec intensité de bruit de tir	102
4.4	Estimation de la probabilité de ruine par la méthode POT	104
4.4.1	Construction d'un nouvel estimateur de la probabilité de ruine	105
4.5	Illustrations numériques	106
	Conclusions et perspectives	109

TABLE DES FIGURES

2.1	Exemples de densités et distributions associées à la loi des valeurs extrêmes (noir : $\xi = 0$,bleu : $\xi = 1$,rouge : $\xi = -1$)	34
2.2	Echantillon au dessus d'un seuil u	44
2.3	Densité et distribution de Loi Pareto Généralisée (GPD)	46
2.4	Figure Graphiques de Pickands, Hill et DEkker pour 2000 données <i>i.i.d.</i> simulées avec cdf donné par $\bar{F}(x) = x^{-1}, x \geq 1$	52
2.5	<i>Mean residual life plot pour GPD($u = 10, \xi = 0.3, \beta = 1$) et pour échantillon de taille (500,1000,10000)</i>	56
2.6	<i>Mean residual life plot pour GPD($u = 10, \xi = 0.5, \beta = 1$) et pour échantillon de taille (500,1000,10000)</i>	57
2.7	Stabilité des seuils (Threshold stability plots) pour la forme GPD et un paramètre d'échelle modifié avec ($\xi = 0.3$) et de taille ($n = 500, 1000, 10000$)	59
2.8	Stabilité des seuils (Threshold stability plots) pour la forme GPD et un paramètre d'échelle modifié avec($\xi = 0.5$) et de taille $n = (500, 1000, 10000)$	60
2.9	Stabilité des seuils (Threshold stability plots) pour la forme GPD et un paramètre d'échelle modifié avec ($\xi = 0.7$) et de taille $n = (500, 1000, 10000 =$	60
2.10	Hill-plot	61
2.11	Estimateur de Hill en fonction du seuil et du nombre d'excès considérés	62
3.1	Richesses d'une compagnie	66
3.2	Processus des réserves et ruine	67
3.3	Coefficient d'Ajustement	68
3.4	QQ-plot of the Ruin probability against the normal distribution when $\xi = \frac{1}{4}$ (top row) and $\xi = \frac{3}{4}$ (bottom row)	81

TABLE DES FIGURES

3.5	Mean Residual Life Plot	87
3.6	The Histogram for the Danish data	88
3.7	Shape parameter as function of the threshold for the Danish data	88
3.8	The QQ-plot for ruin probability with bootstrap method	89

LISTE DES TABLEAUX

2.1	Domaine d'attraction maximal de la distribution de Fréchet.	36
2.2	Domaine d'attraction maximal de distribution de Weibull.	37
2.3	Domaine d'attraction maximal de la distribution de Gumbel.	39
2.4	Quelques lois et leurs domaines d'attraction.	40
2.5	Fonctions moyennes des excès pour quelques distributions usuelles . .	45
2.6	K.S Metric pour $u=100$	62
2.7	K.S Metric pour $u=10$	62
2.8	Double bootstrap pour $u=100$	63
2.9	Double bootstrap pour $u=10$	63
3.1	Probabilité de ruine pour des remboursements exponentielles avec $\lambda =$ $1/1000000000$ et $\theta = (0.1, 0.3, 0.9)$	74
3.2	Probabilité de ruine pour des remboursements de loi Gamma avec $\alpha =$ 10 $\beta = 1/100000000$ et $\theta = (0.1, 0.3, 0.9)$	75
3.3	Probabilité de ruine pour des remboursements pour un mélange de 2 exponentielles avec $\beta_1 = 1/100000000$ $\beta_2 = 1/1000000000$, $p = 0.7$ et $\theta =$ $(0.1, 0.3, 0.9)$	76
3.4	Regularly varying distribution functions	85
3.5	95% confidence interval for the ruin probability of Pareto, Burr and Log- Gamma Distributions with tail index $\xi = \frac{2}{3}$	85
3.6	95% confidence interval for the ruin probability of Pareto, Burr and Log- Gamma Distributions with tail index $\xi = \frac{3}{4}$ b top 5% of the data	86
3.7	Descriptive Statistics Summary of the dataset of Danish fire	86
4.1	Probabilité de ruine avec les arrivées du processus de Hawkes pour l'in- dice de queue $\xi = \frac{3}{7}$	107
4.2	Probabilité de ruine avec les arrivées du processus de Cox pour l'indice de queue $\xi = \frac{3}{7}$	107

4.3	Probabilité de ruine avec les arrivées du processus de Hawkes pour l'indice de queue $\xi = \frac{2}{3}$	108
4.4	Probabilité de ruine avec les arrivées du processus de Cox pour l'indice de queue $\xi = \frac{2}{3}$	108
4.5	Probabilité de ruine avec les arrivées du processus de Hawkes pour l'indice de queue $\xi = \frac{3}{4}$	108
4.6	Probabilité de ruine avec les arrivées du processus de Cox pour l'indice de queue $\xi = \frac{3}{4}$	108

ملخص : في هذه الأطروحة، نثبت طبيعة الانحراف القياسي لمقدر عملية المخاطر باستخدام توزيعات الذيول الثقيلة لعملية الوصول الثابتة.

تستند طريقتنا إلى نتائج [Balkema and De Haan \(1974\)](#), [Klüppelberg and Stadtmüller \(1998\)](#), [Asmussen et al. \(1999\)](#) et [Johansson \(2003\)](#).

ولهذا الغرض، قمنا بتقديم طريقة (Peak Over threshold) POT فوق عتبة مع احتمالات كبيرة وتباين لامتناه في الوقت اللامحدود. ويتم توضيح أداء مقدرنا الجديد من خلال بعض نتائج المحاكاة لبعض نماذج الخسائر، ونقدم مثالاً على تطبيقه على البيانات الدنماركية لخسائر التأمين الكبيرة للحرائق

في الحالة غير الثابتة، نقترح مقدرًا جديدًا لاحتمال الإفلاس لمبالغ الخسائر ذات الذيول الثقيلة

Résumé :

Dans cette thèse, nous établissons la normalité de la distribution asymptotique de l'estimateur du processus de risque avec des distributions de queues lourdes pour le processus d'arrivée stationnaire. Notre approche est basée sur le résultat de [Balkema and De Haan \(1974\)](#), [Klüppelberg and Stadtmüller \(1998\)](#), [Asmussen et al. \(1999\)](#) et [Johansson \(2003\)](#). À cette fin, nous avons introduit la méthode au dessus d'un seuil POT (Peak Over threshold) avec de grandes réserves initiales et de variance infinie en temps infini. La performance de notre nouvel estimateur est illustrée par quelques résultats de simulation pour certains modèles de pertes et nous fournissons un exemple d'application étendu aux données danoises sur les grandes pertes de l'assurance incendie. Dans le cas non stationnaire, nous proposons un nouvel estimateur de la probabilité de ruine pour montants de sinistres de loi à queue lourde.

Abstract :

In this thesis, we establish the normality of the asymptotic the estimator for the risk process with claims of heavy tails for the stationary arrival process. Our approach is based on the result of [Balkema and De Haan \(1974\)](#), [Klüppelberg and Stadtmüller \(1998\)](#), [Asmussen et al. \(1999\)](#) and [Johansson\(2003\)](#). For this purpose, we have introduced the Peak Over threshold POT method with large initial reserves and infinite variance in infinite time. The performance of our new estimator is illustrated by some simulation results for some loss models and provides an extensive example application to Danish data on large fire insurance losses.

In the non-stationary case, we propose a new estimator of the ruin probability for heavy tail loss amounts.

Abréviations et Notations

Abréviations

TVE	Théorie des valeurs extrêmes.
GEV	Distribution des Valeurs Extrêmes Généralisée.
GPD	Distribution de Pareto Généralisée.
POT	L'approche par dépassements de seuil.
mrlplot	Mean residual life plot .
tcplot	Threshold stability plots
TCL	Théorème central limite.
TVE	Théorie des valeurs extrêmes.
i.i.d	Indépendant et identiquement distribué.
EMV	Estimateur du Maximum de Vraisemblance.
BM	Block Maxima
MSE	L'erreur quadratique moyenne

Notations

F	Fonction de distribution.
F^{\leftarrow}	Inverse généralisée de la fonction de répartition.
F^{-1}	Inverse de la fonction de répartition.
F_0	Distribution de la queue intégrée.
x_r	Point terminal supérieur (ou droit) .
F_n	Fonction de répartition empirique.
F_u	Distribution conditionnelle des excès.
(X_1, \dots, X_n)	Échantillon de taille n de X .
$(X_{(1)}, \dots, X_{(n)})$	Échantillon ordonné.
\mathbb{R}	ensemble de nombre réels
$X_{(i)}$	La $i^{\text{ème}}$ statistique d'ordre.
\mathcal{R}_α	Variation régulière d'indice α
L	Fonction à variation lente
Λ_0	Loi de Gumbel
Φ_ξ	Loi de Fréchet
Ψ_ξ	Loi de weibull
Φ	Probabilité de ruine
\xrightarrow{L}	Convergence en loi
\xrightarrow{P}	Convergence en probabilité
$\xrightarrow{p.s}$	Convergence presque sure

Description du problème

Le sujet de recherche de notre thèse de doctorat se situe dans le domaine de la statistique des valeurs extrêmes. Ici, contrairement à la statistique classique, l'accent est mis sur la modélisation des événements extrêmes, c'est-à-dire des événements de faible fréquence (événements rares), mais dont l'impact est élevé et souvent catastrophique.

Les catastrophes naturelles extrêmes survenues en 2021 et 2022 (un gel hivernal profond, des inondations, des orages violents, des vagues de chaleur et des ouragans majeurs), ont causé des pertes globales de 280 milliards de dollars US en 2021, dont environ 120 milliards de dollars US étaient assurés, et 115 milliards de dollars US en 2022 selon Swiss Re (Le groupe Swiss Re est l'un des principaux fournisseurs mondiaux de réassurance, d'assurance et d'autres formes de transfert de risques fondés sur l'assurance).

L'ouragan Ida (Le premier à atteindre la catégorie 4 de la saison cyclonique 2021 dans l'océan Atlantique nord) a été le principal événement générateur de pertes. L'ouragan Ida a été la catastrophe naturelle la plus coûteuse de l'année, avec des pertes globales de 65 milliards de dollars US (36 milliards de dollars US de pertes assurées), mais plus de la moitié des pertes mondiales sont dues à des périls secondaires.

Ces catastrophes extrêmes ont par ailleurs fait plus de victimes. Avec 32.100 morts en 2022, contre 12.000 à 13.000 en moyenne depuis 2012.

En Europe, les crues soudaines consécutives dues à des précipitations extrêmes ont causé des pertes de 54 milliards de dollars US (46 milliards d'euros), la catastrophe naturelle la plus coûteuse jamais enregistrée en Allemagne.

Les pertes dues aux catastrophes naturelles devraient continuer à augmenter plus

vite que le PIB mondial (Produit Intérieur Brut), compte tenu de l'accroissement de la richesse, de l'urbanisation et du changement climatique.

"En 2021, les pertes assurées liées aux catastrophes naturelles ont à nouveau dépassé la moyenne décennale précédente, poursuivant la tendance à une augmentation annuelle de 5 à 7 % des pertes observée au cours des dernières décennies.

Il semble qu'au moins un événement extrême se produise chaque année, une inondation grave, une tempête d'hiver ou un incendie de forêt, qui entraîne des pertes supérieures à 10 milliards de Dollars US.

Un seul événement de ce type frappant des zones densément peuplées peut avoir un impact considérable sur les pertes annuelles.

Il ne fait aucun doute que les économies et les États modernes ne fonctionneraient pas sans les institutions qui garantissent le remboursement à l'individu, à l'entreprise ou à l'organisation de ses pertes, qui peuvent survenir en raison de catastrophes naturelles ou d'origine humaine, d'incendies, d'inondations, d'accidents, d'émeutes, etc.

Un peu d'histoire de la théorie du risque extrême

Au cours du 20^{ème} siècle, les outils nécessaires pour traiter les questions d'assurance ont été développés. Il s'agit de la théorie des probabilités, de la statistique et des processus stochastiques.

Les mathématiciens suédois [Lundberg \(1903\)](#) et [Cramér\(1930\)](#) ont été des pionniers dans ces domaines. Ils ont réalisé, dans la première moitié du 20^e siècle, que la théorie des processus stochastiques constituait le cadre le plus approprié pour modéliser les sinistres survenant dans une entreprise d'assurance.

Aujourd'hui, le modèle Cramer-Lundberg est l'une des pierres angulaires des mathématiques de l'assurance non-vie.

Il a été modifié et étendu dans des directions très différentes et, plus encore, a motivé des recherches dans divers autres domaines de la théorie appliquée des probabilités, tels que la théorie des files d'attente, les processus de branchement, la théorie du renouvellement, la fiabilité, les modèles de barrage et de stockage, la théorie des valeurs extrêmes et les processus stochastiques.

La théorie des valeurs extrêmes est relativement récente. Ce n'est que dans les années 1980 que les contours de la théorie statistique se sont précisés. La probabilité unidimensionnelle des valeurs extrêmes a été développée par M. [Fréchet \(1927\)](#), [Fisher and Tippett \(1928\)](#), et [Von Mises \(1936\)](#), et résumée dans les travaux de [Gnedenko](#)

(1943). La théorie statistique des valeurs extrêmes est initiée par J. [Pickands \(1975\)](#).

La statistique des valeurs extrêmes nous aide à "tirer des informations des événements désastreux". C'est pourquoi les domaines d'application des statistiques des extrêmes sont très diversifiés. On peut citer les domaines de l'hydrologie (débits des rivières, inondations,...), de la recherche environnementale et de la météorologie (pollution de l'air, pluies torrentielles, vitesses de vent, tempêtes, ...), de la géologie (modélisation des tremblements de terre, éruption volcanique, valeur des diamants,...), la (ré)assurance (calcul des primes, probabilité de ruine d'une société d'assurance,...), la finance (Value-at-Risk, espérance conditionnelle de queue,..), l'informatique (données de trafic du réseau, temps d'attente des serveurs,..), l'ingénierie structurelle, les télécommunications et la biostatistique, entre autres (voir, par exemple, [Reiss et al. \(1997\)](#); [Beirlant et al. \(2004\)](#) et [Castillo et al. \(2005\)](#)).

La statistique des valeurs extrêmes traite de l'estimation des quantités qui sont liées à la queue d'une distribution ou, de manière équivalente, de l'analyse des observations les plus importantes d'un échantillon.

L'analyse des valeurs extrêmes est fermement fondée sur les distributions dites de valeurs extrêmes et leurs domaines d'attraction. Ces distributions apparaissent comme les seules formes limites possibles pour les distributions des maxima dans des échantillons de variables aléatoires indépendantes et identiquement distribuées.

Les sujets importants dans le domaine de la statistique des valeurs extrêmes sont l'estimation de l'indice de queue, les quantiles extrêmes ou les probabilités de petite queue.

Toutefois, la statistique des valeurs extrêmes repose sur le fait que, sous des conditions assez faibles, on peut considérer qu'une classe de fonctions s'adapte à la distribution de la plus grande observation d'un échantillon.

À partir de cette limite, on peut voir que le comportement de la queue d'une fonction de distribution peut être complètement caractérisé par un seul paramètre à valeur réelle ξ , appelé indice de valeur extrême ou paramètre de forme. En fonction du signe de ce paramètre, le domaine d'attraction de l'indice des valeurs extrêmes peut être déterminé.

L'indice de queue est une caractéristique clé des données à queue lourde. Il montre la forme de la queue de la distribution sans faire d'hypothèse sur les paramètres de la distribution entière. Toutes les caractéristiques des variables aléatoires à queue lourde sont basées sur ce paramètre capital.

Plan de la thèse

Dans cette thèse, nous proposons un nouvel estimateur de la probabilité de ruine dans le cas des remboursements extrêmes des sinistres.

Il existe deux façons de modéliser les événements extrêmes à l'aide des modèles des valeurs extrêmes. La première approche consiste à diviser l'échantillon en blocs, puis obtenir le maximum de chaque bloc, ce que l'on appelle la méthode des maxima de blocs. Cette approche est rigoureusement formulée par le théorème de Fisher-Tippett qui remonte à 1928, où les maxima de bloc sont modélisés par une distribution de valeur extrême généralisée (GEV)

$$H_{\xi,u,\beta}(x) = \exp \left\{ - \left(1 + \xi \frac{x-u}{\beta} \right)^{-1/\xi} \right\}, \quad \xi \neq 0, 1 + \xi \frac{x-u}{\beta} > 0 \quad (1)$$

La seconde approche est la méthode des pics au-dessus du seuil (POT), qui consiste à prendre les grandes observations qui dépassent un certain seuil u , où les v.a.r qui dépassent le seuil u sont modélisés par une distribution de Pareto généralisée (GPD). Les modèles POT sont généralement plus utiles pour les applications pratiques. [Reiss et al. \(1997\)](#).

$$G_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \frac{\xi}{\beta} y \right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\beta}\right) & \text{si } \xi = 0 \end{cases} \quad (2)$$

Dans le modèle de ruine, on étudie la stabilité d'un assureur. Partant du capital à un moment donné, on suppose que le capital augmente linéairement dans le temps par des primes annuelles fixes, mais qu'il diminue avec un saut chaque fois qu'un sinistre survient. La ruine survient lorsque le capital est négatif à un moment donné.

La probabilité que cela se produise un jour, en supposant que la prime annuelle ainsi que le processus de génération de sinistres restent inchangés, est une bonne indication pour savoir si l'actif de l'assureur correspond suffisamment bien à son passif.

Les méthodes analytiques de calcul des probabilités de ruine n'existent que pour les distributions de sinistres qui sont des mélanges et des combinaisons de distributions exponentielles. Des algorithmes existent pour les distributions discrètes avec peu de points de masse. Par conséquent, des méthodes asymptotiques ont été développées pour dériver des expansions de la probabilité de ruine lorsque le capital ou la réserve initiale augmente à l'infini.

Dans la littérature de la théorie du risque de l'assurance et la finance, la ruine est l'événement le plus important. Le fondement théorique de la théorie de la ruine, connu sous le nom de modèle de Cramer-Lundberg ou processus de risque classique, a été introduit par [Lundberg \(1903\)](#).

Dans cette thèse, nous considérons un processus de risque classique avec la richesse au temps t donnée par :

$$R_t = u + pt - \sum_{i=1}^{N_t} X_i, \quad t \geq 0 \quad (3)$$

où X_i sont les montants remboursements des sinistres *i.i.d.* de loi de probabilité F à queue lourde dont les deux premiers moments sont finis ou infinis, et indépendants du processus d'arrivée des sinistres $(N_t)_t$ qui est un processus non stationnaire qui satisfait le principe de la grande déviation avec l'intensité $\lambda(t)$, et $p > 0$ est le taux de prime constant que la compagnie d'assurance reçoit, et $u > 0$ est le capitale initiale de la compagnie d'assurance.

Dans le modèle de risque classique de [Lundberg \(1903\)](#), $(N_t)_t$ est supposé suivre un processus de Poisson, qui a des incréments temporels indépendants et stationnaires.

Dans cette thèse, nous supposons que le processus d'arrivée $(N_t)_t$ suit un processus non stationnaire, qui présente des caractéristiques de regroupement et d'auto-excitation et dont les incréments temporels sont dépendants.

Nous présentons à titre d'exemple deux processus d'arrivée des sinistres qui sont :

— Le processus linéaire de Hawkes d'intensité $\lambda(t)$ donné par :

$$\lambda_t = \nu + \int_0^t h(t-s) dN_s, \quad (4)$$

ou bien

$$\lambda_t = \nu + \sum_{\tau < t} h(t-\tau), \quad (5)$$

Où $h : [0, \infty) \rightarrow (0, \infty)$ est intégrable et on suppose que $\|h\|_{L^1} < 1$, qui représente le nombre moyen de descendants directs des ancêtres.

— Processus de Cox avec intensité d'intensité λ_t

$$\lambda_t = \nu(t) + \sum_{\tau^{(1)} < t} g(t-\tau^{(1)}), \quad (6)$$

où $\tau^{(1)}$ le temps d'arrivée d'un processus de Poisson externe homogène avec intensité γ . Ici, $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ est intégrable, c'est-à-dire $\int_0^\infty g(t)dt < \infty$ et $\nu(t)$ est une fonction positive, continue et déterministe telle que $\nu(t) \rightarrow \nu$ avec $t \rightarrow \infty$.

Les processus de Hawkes et de Cox ont été utilisé dans le domaine des assurances (Stabile and Torrisi (2010); Zhu (2013)). En appliquant les techniques des grandes déviations, la probabilité de ruine asymptotique pour les processus de risque en assurance ont été étudiées par (Stabile and Torrisi (2010) pour les sinistres à queue légère et par Zhu (2013) pour les sinistres à queue lourde.

Dans notre Thèse, nous étudions l'utilisation de la méthode des Pics au dessus d'un Seuil (la méthode POT), pour construire un nouvel estimateur de la probabilité de ruine pour un processus de risque avec des montants de sinistres à queues lourdes avec des sinistres arrivent suivant un processus non stationnaire en temps infini. Notre approche est basée sur l'approximation de l'échantillon sur un certain seuil par la distribution de Pareto généralisée.

Enfin, nous donnons les illustrations numériques des résultats.

La thèse comprend un article et trois communications et est divisée en quatre chapitres.

Article : POT-based estimator of the ruin probability in infinite time for loss models : An application to insurance risk.(Chilean Journal of Statistics (<http://soche.cl/chjs/>), Volume 13, Number 2, December 2022)

Communication 1 : "A simulation comparison study of the methods of the threshold selection for the heavy-tailed distributions". 11 international conferences on Extreme Value Analysis(EVA2019, Zagreb, Croatia).
<https://web.math.pmf.unizg.hr/eva2019/>

Communication 2 : "POT Method for Ruin Probability in Infinite Time with Non-Stationary Arrivals and Heavy-Tailed Distribution Claims for Loss Models". 6th International Conference on Advances in Statistics (ICAS2020, Turkey).
<https://icasconference.com/>

Communication 3 : "Value-at-risk prediction using garch model and bayesian extreme value for mixture distributions". International Conference on Mathematics and Applications (ICMA'2021, Algeria). <https://icma21.sciencesconf.org/>

Le reste de la thèse est organisé comme suit. Dans le chapitre 2, nous présentons les définitions de base, les propriétés de l'estimation des lois de probabilité extrêmes. Dans le chapitre 3, nous donnons les définitions et les propriétés de la probabilité de ruine dans le cas stationnaire. Dans le chapitre 4, nous présentons le modèle et fondement

théorique de la probabilité de ruine dans le cas non stationnaire. Nous clôturons ce travail par une conclusion générale.

CHAPITRE 1

NOTIONS DE BASE DE LA THÉORIE DES PROBABILITÉS

Nous présentons dans un premier temps les notions et les propriétés du calcul des probabilités et statistique de bases sur lesquelles repose l'utilisation des lois des valeurs extrêmes.

Définition 1.1 Soit X une variable aléatoire réelle définie sur l'espace de probabilité (Ω, \mathcal{F}, P) à valeur dans l'espace mesurable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

La fonction

$$F(x) = P\{\omega, X(\omega) \leq x\}, x \in \mathbb{R},$$

est appelée la fonction de distribution cumulée (cdf) de la v.a.r X , et la fonction de survie ou la distribution de la queue est donnée par

$$\bar{F}(x) = 1 - F(x).$$

Définition 1.2 Si la fonction réelle non négative F est dérivable de dérivée $f(t), t \in \mathbb{R}$, existe alors pour tout $x \in \mathbb{R}$,

$$F(x) = \int_{-\infty}^x f(t)dt,$$

et la fonction de survie est

$$\bar{F}(x) = \int_x^{\infty} f(t)dt.$$

La fonction $f(t), t \in \mathbb{R}$, est appelée la densité de probabilité (pdf) de la v.a.r X .

Définition 1.3 Soient les v.a.r X_1, X_2, \dots, X_n de cdf commune F sont dites indépendantes et identiquement distribuées (i.i.d.) si, pour toutes parties $B_1, B_2, \dots, B_n \in \mathcal{F}(\mathbb{R})$

$$P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \dots P(X_n \in B_n).$$

ou

$$F(x_1, x_2, \dots, x_n) = F(x_1)F(x_2) \dots F(x_n)$$

Définition 1.4 (*Processus empirique*). Le processus empirique de l'échantillon i.i.d. X_1, X_2, \dots, X_n est définie par

$$e_n(x) = \sqrt{n}(F_n(x) - F(x)), x \in \mathbb{R},$$

où F_n est la fonction de distribution basée sur X_1, X_2, \dots, X_n .

Définition 1.5 Soit X_1, X_2, \dots, X_n une suite de v.a.r i.i.d. de cdf commune F . Pour tout entier $n > 1$, on définit la somme partielle par

$$S_n = \sum_{i=1}^n X_i$$

et sa moyenne correspondante \bar{X}_n donnée par

$$\bar{X}_n = S_n/n.$$

appelée la moyenne empirique.

1.1 Théorèmes de la Convergence des Grands nombres

Théorème 1.1 (*Loi faible des grands nombres*). Soit X_1, X_2, \dots, X_n des variables i.i.d. de moyenne finie $E(X_i) = \mu < \infty$. alors, pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \geq \epsilon) = 0$$

et on note $\bar{X}_n \xrightarrow{P} \mu$ avec $n \rightarrow \infty$.

Théorème 1.2 (*Loi forte des grands nombres*). Soit X_1, X_2, \dots, X_n des variables i.i.d., alors $\bar{X}_n \xrightarrow{p.s.} \mu$ quand $n \rightarrow \infty$.

Remarque 1.1 Formellement, le théorème de loi forte des grands nombres reste valide pour des v.a.r positifs (négatifs) avec une moyenne infinie, c'est-à-dire que dans ce cas, nous avons

$$\bar{X}_n \xrightarrow{p.s.} E(X) = \infty \quad (= -\infty).$$

L'application du théorème de la convergence faible sur $F_n(x)$ donne le resultat suivant.

Corollaire 1.1

$$F_n(x) \xrightarrow{p.s.} F(x) \text{ quand } n \rightarrow \infty, \text{ pour tout } x \in \mathbb{R}.$$

Théorème 1.3 (Glivenko-Cantelli)

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{p.s.} 0, \text{ as } n \rightarrow \infty.$$

Le théorème de la convergence faible d'une suite de v.a.r *i.i.d.* du théorème (1.1) peut être trouvée dans n'importe quel manuel standard de théorie des probabilités.

La convergence faible avec d'autres normalisations et pour le cas non-*i.i.d.* a été traité par exemple dans [Feller \(1991\)](#), [Petrov \(1995\)](#), ou dans le cas des martingales dans [Hall and Heyde \(2014\)](#).

Une meilleure compréhension du comportement limite faible des sommes est donnée par ce qu'on appelle les taux de convergence dans la loi des grand nombres, c'est-à-dire par la détermination de l'ordre de la probabilité de

$$P(|S_n - a_n| > b_n), \quad n \rightarrow \infty,$$

pour des suites de normalisation et de centralisation appropriées $(b_n)_n, (a_n)_n$.

1.1.1 Théorème Central Limite

Le théorème central limite (TCL) est l'un des résultats les plus importants de la théorie des probabilités. Il indique que, sous certaines conditions, la somme d'un grand nombre de variables aléatoires est approximativement normale. Ici, nous énonçons une version du CLT qui s'applique aux variables aléatoires *i.i.d.*

Supposons que X_1, X_2, \dots, X_n sont des variables aléatoires *i.i.d.* avec des valeurs attendues $E(X_i) = \mu < \infty$ et une variance $\text{Var}(X_i) = \sigma^2 < \infty$.

Alors, comme nous l'avons vu ci-dessus, la moyenne de l'échantillon $\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ est de moyenne $E(\bar{X}_n) = \mu$ et de variance $\text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$. Ainsi, la variable aléatoire normalisée

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + \dots + X_n - n\mu}{\sqrt{n}\sigma}$$

est de moyenne $E(Z_n) = 0$ et de variance $\text{Var}(Z_n) = 1$.

Le théorème central limite affirme que le cdf de Z_n converge vers le cdf de la loi normale standard.

Théorème 1.4 Si X_1, X_2, \dots, X_n est une suite de v.a.r *i.i.d.* de moyenne μ et de variance finie

σ^2 , alors

$$\sqrt{n} \left(\frac{\bar{X}_n - \mu}{\sigma} \right) \xrightarrow{d} \mathcal{N}(0, 1) \quad \text{quand } n \rightarrow \infty.$$

Remarque 1.2 Notez qu'une condition nécessaire pour le TLC est que la variance soit finie. Gnedenko et Kolmogorov ont ensuite généralisé le résultat à certaines classes de distributions à variance infinie (distributions dont les queues diminuent comme une loi de puissance $|x|^{-\alpha-1}$, $0 < \alpha < 2$) (queues de Pareto).

Dans ce cas, la somme ne tend pas vers une distribution normale, mais plutôt vers une distribution dite distribution stable. (La distribution normale et la distribution de Cauchy sont des distributions stables particulières).

Une distribution est dite stable si une combinaison linéaire de deux variables aléatoires indépendantes ayant la même distribution.

1.2 Distributions à variations régulières

Les estimations asymptotiques sont partout présentes dans les mathématiques de l'assurance et la finance mathématique.

Dans de nombreux cas, les transformées de Laplace et de Fourier jouent un rôle crucial.

A partir des travaux pionniers de Karamata ([Bingham \(2007\)](#)), la théorie de la variation régulière a maintenant obtenu le statut de connaissance standard pour tout probabiliste ou statisticien.

Nous résumons ci-dessous certains des principaux résultats pertinents pour nos applications. Tout ce que nous discutons, et bien plus encore, se trouve dans le volume encyclopédique de [Bingham \(2007\)](#).

Définition 1.6 (Variation régulière au sens de Karamata)

- (a). On dit qu'une fonction f est à variations régulières à l'infini d'indice $\rho \in \mathbb{R}$, que l'on notera $f \in \mathcal{R}_\rho$, si f est positive à l'infini (i.e. s'il existe A tel que pour tout $x > A, f(x) > 0$) et si pour tout $t > 0$,

$$\lim_{x \rightarrow \infty} \frac{f(tx)}{f(x)} = t^\rho$$

– si $\rho = 0$, c'est-à-dire $f \in \mathcal{R}_0$, alors la fonction f est appelée fonction à variations lentes à l'infini, notée pour la suite L .

- (b). si $\rho = \infty$, on parle de fonction à variations rapides à l'infini. On peut montrer facilement que toute fonction f à variations régulières d'indice $\rho \in \mathbb{R}$ s'écrit,

$$f(x) = x^\rho L(x), L(x) \in \mathcal{R}_0$$

Proposition 1.1 (Resnick (1987))

Soit f est une fonction à variations régulières d'indice ρ , alors pour tout $a < t < b$

$$\lim_{x \rightarrow \infty} \sup_{t \in [a, b]} \left| \frac{f(tx)}{f(x)} - t^\rho \right| = 0.$$

Lemme 1.1 — Si f est à variations régulières d'indice $\rho > 0$, alors $f^\leftarrow(x)$ est à variations régulières d'indice $1/\rho$.

— Si f est à variations régulières d'indice $\rho < 0$, alors $f^\leftarrow(1/x)$ est à variations régulières d'indice $-1/\rho$.

Lemme 1.2 (Bingham (2007)) Si L, L_1, L_2 sont à variations lentes, alors

1. $L^\beta \quad \forall \beta \in \mathbb{R}, \quad L_1 + L_2, \quad L_1 L_2 \quad \text{et} \quad L_1 \circ L_2 \quad (\text{si } L_2(x) \rightarrow \infty \text{ quand } x \rightarrow \infty)$ sont à variations lentes.

2. $\forall \varepsilon > 0$ et $x \rightarrow \infty$

$$\lim_{x \rightarrow \infty} x^\varepsilon L(x) = \infty \quad \text{et} \quad \lim_{x \rightarrow \infty} x^{-\varepsilon} L(x) = 0$$

3. $\frac{\ln L(x)}{\ln x} \rightarrow 0$.

Corollaire 1.1 (Ferreira et al. (2003), Bingham (2007)) Si f, f_1, f_2 varient régulièrement à l'infini avec indices $\alpha, \alpha_1, \alpha_2$ respectivement, alors :

1. $\lim_{x \rightarrow +\infty} \frac{\ln f(x)}{\ln x} = \alpha$

2. Si $\alpha \neq 0$

$$\lim_{x \rightarrow +\infty} f(x) = \begin{cases} 0 & \text{si } \alpha < 0 \\ +\infty & \text{si } \alpha > 0 \end{cases}$$

3. $f^\beta \in \mathcal{R}_{\beta\alpha}, \beta \in \mathbb{R}$.
4. $f_1 + f_2 \in \mathcal{R}_\alpha$ pour $\alpha = \max(\alpha_1, \alpha_2)$.
5. $f_1 f_2 \in \mathcal{R}_{\alpha_1 + \alpha_2}$.
6. Si, de plus, $f_2 \rightarrow \infty$, alors $(f_1 \circ f_2)(x) = f_1(f_2(x)) \in \mathcal{R}_{\alpha_1 \alpha_2}$.

Les dérivés et les intégrales de fonctions à variations régulières varient également régulièrement, sous certaines conditions de régularité. En général, l'indice de variation régulière augmente de 1 lors de l'intégration, il diminue de 1 lors de la différenciation.

Proposition 1.1 (*Bingham (2007)*) Si L est à variation lente ($L \in \mathcal{R}_0$), alors :

1. Pour $\alpha > -1$:

$$\int_{x_0}^x x^\alpha L(t) dt \sim (\alpha + 1)^{-1} x^{\alpha+1} L(x), x \rightarrow \infty$$

2. Pour $\alpha < -1$:

$$\int_x^{+\infty} x^\alpha L(t) dt \sim -(\alpha + 1)^{-1} x^{\alpha+1} L(x), x \rightarrow \infty.$$

Proposition 1.2 Soit

$$F(x) = \int_{x_0}^x f(t) dt \quad \text{ou} \quad F(x) = \int_x^{+\infty} f(t) dt$$

où f est ultimement monotone (i.e., f à support positif et il existe un x_0 pour lequel elle est monotone sur l'intervalle $[x_0, +\infty[$).

Si $F(x) \sim cx^\alpha L(x)$, si $x \rightarrow +\infty$ avec $c > 0, \alpha \in \mathbb{R}$ et $L \in \mathcal{R}_0$,

alors $f(x) \sim c\alpha x^{\alpha-1} L(x)$, $x \rightarrow +\infty$

Théorème 1.5 ((Représentation de Karamata).

Toutes fonction à variation lentes L s'écrivent sous la forme :

$$L(x) = c(x) \exp \left\{ \int_1^x \frac{\Delta(u)}{u} du \right\}. \quad (1.1)$$

où $c(x) \rightarrow c > 0$ et $\Delta(x) \rightarrow 0$ lorsque $x \rightarrow \infty$. Cette représentation des fonctions à variations lentes est connue sous le nom de représentation de Karamata *Bingham (2007)*.

Si la fonction c est constante, la fonction L est dite normalisée.

Théorème 1.6 (Représentation de Karamata) Si $f \in \mathcal{R}_\rho$ pour un certain $\rho \in \mathbb{R}$, alors

$$h(x) = c(x) \exp \left\{ \int_z^x \frac{\delta(u)}{u} du \right\}, \quad x \geq z,$$

pour un certain $z > 0$ où c et δ sont des fonctions mesurables, $c(x) \rightarrow c_0 \in (0, \infty)$, $\delta(x) \rightarrow \rho$ quand $x \rightarrow \infty$. L'implication inverse est également valable.

Remarque 1.3 1. Sous (a) et (b) ci-dessus, nous avons défini la variation régulière à l'infini, c'est-à-dire pour $x \rightarrow \infty$. De même on peut définir la variation régulière en zéro en remplaçant $x \rightarrow \infty$ par $x \rightarrow 0$, ou pour tout a positif.

En effet, la variation régulière de f en $a > 0$ est définie comme la variation régulière à (l'infini) pour la fonction $f_a(x) = f(a - x^{-1})$. Par exemple, dans le théorème. Dans les cas où la distinction est importante, nous parlerons de variation régulière en a .

2. Les exemples typiques de fonctions à variation lente sont les constantes positives ou les fonctions convergeant vers une constante positive, les logarithmes et les logarithmes itérés. Par exemple, pour tous les réels ρ , les fonctions

$$x^\alpha, x^\alpha \ln(1+x), (x \ln(1+x))^\alpha, x^\alpha \ln(\ln(e+x))$$

sont régulières à ∞ avec l'indice α . Les exemples suivants ne sont pas des fonctions régulières variables

$$2 + \sin x, \quad e^{[\ln(1+x)]},$$

où $[\]$ représente la partie entière.

1.3 Distributions à queue lourde

L'analyse des queues lourdes est une branche de la théorie des valeurs extrêmes consacrée à l'étude des phénomènes régis par des grands mouvements plutôt que par des mouvements réguliers.

Ses applications sont diverses, comme, les réseaux de données, en présence de fichiers de taille importante sur les serveurs du réseau, en finance où les rendements financiers sont à queue lourde.

Une distribution à queue lourde a une queue plus épaisse que la distribution exponentielle. En d'autres termes, une distribution à queue lourde tend vers zéro plus lentement qu'une distribution à queue exponentielle.

On dit que la distribution d'une variable aléatoire X avec une fonction de distribution F a une queue lourde (droite) si la fonction génératrice de moments de X , $M_X(t)$, est infinie pour tous les $t > 0$.

Cela signifie que

$$\int_{-\infty}^{\infty} e^{tx} dF(x) = \infty \quad \text{pour tout } t > 0.$$

Ceci s'écrit également en termes de fonction de distribution de la queue de distribution.

$$\bar{F}(x) \equiv P[X > x]$$

comme

$$\lim_{x \rightarrow \infty} e^{tx} \bar{F}(x) = \infty \quad \text{pour tout } t > 0$$

Dans le cas contraire, on dit que la distribution est à queue légère.

Pour plus de détails les distributions à queues lourdes, voir (Handbook of Heavy Tailed Distributions in Finance) de [Embrechts et al. \(2003\)](#), (Heavy-tail phenomena probabilistic and statistical modeling) de [Resnick \(2007\)](#).

Une distribution F d'une variable aléatoire X est dite à une queue longue à droite si pour tout $t > 0$,

$$\lim_{x \rightarrow \infty} P[X > x + t \mid X > x] = 1$$

ou équivalent

$$\bar{F}(x + t) \sim \bar{F}(x) \quad \text{as } x \rightarrow \infty$$

intuitivement, pour une distribution à queue longue à droite, si la quantité dépasse un certain niveau élevé, la probabilité qu'elle dépasse tout autre niveau plus élevé est presque certain (proche de 1). Toutes les distributions à queue longue sont à queue lourde, mais la réciproque est fausse.

1.3.1 Distributions sous-exponentielles

Définition 1.1 Une distribution de probabilité F , i.e. $F \in \mathcal{S}$ avec support $(0, \infty)$ est sous-exponentielle, si pour tout $n \geq 2$,

$$\lim_{x \rightarrow \infty} \frac{\overline{F^{n*}}(x)}{\bar{F}(x)} = n, \quad n \geq 2.$$

Où $F * G$ désigne la convolution des fonctions de distribution F et G , $F^{*1} = F$ et $F^{*(n+1)} = F^{*n} * F$, (voir [Asmussen and Albrecher \(2010\)](#)).

Lemme 1.1 (Une condition suffisante pour la sous-exponentialité)

Si

$$\limsup_{x \rightarrow \infty} \frac{\overline{F^{2*}}(x)}{\bar{F}(x)} \leq 2, \quad \text{alors } F \in \mathcal{S}.$$

La preuve de ce résultat est donnée dans [Embrechts et al. \(2013\)](#) (Modelling Extremal Events for Insurance and Finance).

1.4 Quelques notions sur les processus

On désigne par (Ω, \mathcal{F}, P) un espace de probabilité, $(\mathcal{E}, \mathcal{B})$ (\mathcal{E} est appelé espace d'état (souvent, on aura $\mathcal{E} = \mathbb{R}, \mathbb{C}, \mathbb{C}^{\mathbb{N}}$, ou un ensemble fini ou dénombrable) un espace mesurable et T est l'ensemble des indices (souvent, on aura $T = \mathbb{R}, \mathbb{R}_+, \mathbb{R}^p, \mathbb{N}, \dots$). T peut

faire référence au temps, à l'espace ou les deux à la fois.

Un processus stochastique (ou aléatoire) est une famille de variables aléatoires (c'est-à-dire, des applications mesurables) définies sur le même espace de probabilité (Ω, \mathcal{F}, P) indexée par T et à valeurs dans $(\mathcal{E}, \mathcal{B})$.

La famille de toutes les distributions finies-dimensionnelles de X s'appelle la loi spatiale du processus. Si $S \subseteq \mathbb{R}$, on parle de loi temporelle.

Pour $\omega \in \Omega$ fixe, l'application $t \mapsto X_t(\omega)$ est appelée trajectoire (ou réalisation) du processus stochastique $(X_t)_{t \in T}$.

Définition 1.2 Un processus stochastique $X = (X_t)_{t \in T}$ est dit stationnaire au sens strict si sa loi de probabilité est invariante par translation i.e., $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ et $(X_{t_1+s}, X_{t_2+s}, \dots, X_{t_n+s})$ ont la même loi, $\forall t_1 < t_2 < \dots < t_n, s \in T$.

Définition 1.3 Un processus stochastique est dit stationnaire au sens large ou faiblement stationnaire si :

1. $E(X_t) = m < \infty$, indépendant de t .
2. $\text{Var}(X_t) = \sigma^2 < \infty$, indépendant de t .
3. $\text{Cov}(X_t, X_s)$ ne dépend que de la différence $|t - s|$.

1.4.1 Processus à accroissements indépendants

Définition 1.4 Un processus X est dit à accroissements indépendants si :

$\forall t_0 < t_1 < \dots < t_n \in T$, les variables aléatoires $(X_{t_1} - X_{t_0}), (X_{t_2} - X_{t_1}), \dots, (X_{t_n} - X_{t_{n-1}})$ sont indépendantes.

1.4.2 Processus à accroissements stationnaires

Définition 1.5 Un processus $(X_t)_{t \in T}$ est dit homogène dans le temps si la loi de $(X_{t+s} - X_t)$ ne dépend que de $s, \forall t$.

1.4.3 Processus de comptage

Définition 1.6 Un processus stochastique $N = (N_t)_{t \in T = \mathbb{R}_+}$ est dit processus de comptage ou processus de dénombrement si :

1. $N_t \in \mathbb{N}, \forall t \in \mathbb{R}_+, N_0 = 0$.
2. Si $s < t$ alors $N_s \leq N_t$ où $N_t - N_s$, représente le nombre d'événements se produisant dans l'intervalle $[s, t]$.

Un des processus de comptage le plus utilisé est le processus de Poisson.

1.4.4 Processus de Poisson

Le processus de Poisson d'intensité λ (réel strictement positif) est un processus de comptage d'occurrences qui vérifie les conditions suivantes :

1. Les nombres d'occurrences dans des intervalles de temps disjoints sont indépendants.
2. La probabilité d'une occurrence dans un petit intervalle de temps est proportionnelle à la longueur de cet intervalle, le coefficient de proportionnalité étant λ .
3. La probabilité qu'il y ait plus d'une occurrence dans un petit intervalle de temps est négligeable.

Ces deux dernières conditions forment la propriété dite des « événements rares ». Mathématiquement, ces propriétés se traduisent, si l'on note $(N_t)_{t \in \mathbb{R}^+}$ le processus de Poisson et \mathbb{P} la probabilité, par :

1. $\forall t_0 = 0 \leq t_1 < \dots < t_k$, les variables aléatoires $(N_{t_k} - N_{t_{k-1}}), \dots, (N_{t_1} - N_{t_0})$ sont indépendantes;
2. $\mathbb{P}(N_{t+h} - N_t = 1) = \lambda h + o(h)$ lorsque $h \rightarrow 0^+$ (étant fixé);
3. $\mathbb{P}(N_{t+h} - N_t > 1) = o(h)$ lorsque $h \rightarrow 0^+$ (t étant fixé).

On déduit des deux dernières égalités que $\mathbb{P}(N_{t+h} - N_t = 0) = 1 - \lambda h + o(h)$.

On démontre, par des résolutions d'équations différentielles d'ordre 1 et par récurrence, que pour un temps t donné (strictement positif), le nombre N_t d'occurrences dans un intervalle de longueur t suit une loi de Poisson d'intensité λt , c'est-à-dire que

$$\mathbb{P}(N_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \forall k \in \mathbb{N}$$

Et on démontre aussi que les temps s'écoulant entre deux incrémentations du processus de comptage (rappelons que la probabilité que le processus de comptage augmente d'un coup de deux unités ou plus est nulle d'après la définition) sont des variables aléatoires indépendantes et de même loi exponentielle de paramètre λ , c'est-à-dire que :

Si $T_n = \inf\{t \geq 0, N_t \geq n\}$, alors les variables aléatoires $S_k = T_k - T_{k-1}$ ($k \in \mathbb{N}^*$) sont indépendantes et $\mathbb{P}(S_k \leq t) = 1 - e^{-\lambda t}$, quel que soit t réel positif ou nul.

Il en résulte que pour tout entier naturel non nul n , la variable aléatoire $T_n = S_1 + \dots + S_n$ suit la loi gamma $\Gamma\left(n, \frac{1}{\lambda}\right)$, dite aussi loi d'Erlang.

1.4.5 Processus de Poisson composé

Supposons que les conditions suivantes sont satisfaites :

- $(N_t)_{t \geq 0}$ est un processus de Poisson d'intensité λ .
- X_1, X_2, \dots sont des variables aléatoires indépendantes et identiquement distribuées, de distribution commune F .
- La suite aléatoire $(X_t)_{t \geq 1}$ et le processus $(N_t)_{t \geq 0}$ sont indépendants.

Le processus aléatoire $(S_t)_{t \geq 0}$, défini par :

$$S_t = \sum_{i=1}^{N_t} X_i,$$

est appelé un processus de Poisson composé d'intensité λ et de distribution F .

On obtient également par conditionnement sur N les premiers moments de S

Proposition 1.2 (*Identité de Wald*).

$$\begin{aligned} E(S_t) &= E(N_t)E(X_t) \\ \text{Var}(S_t) &= \text{Var}(X_t)E(N_t) + E^2(X_t)\text{Var}(N_t) \end{aligned}$$

Preuve.

$$\begin{aligned} E(S_t) &= E(E(S_t | N_t)) = E(NE(X_t)) = E(N_t)E(X_t) \\ \text{Var}(S_t) &= E(\text{Var}(S_t | N_t)) + \text{Var}(E(S_t | N_t)) \\ &= E(N_t \cdot \text{Var}(X_t)) + \text{Var}(N_t \cdot E(X_t)) \\ &= E(N_t)\text{Var}(X_t) + E(X_t)^2 \text{Var}(N_t) \end{aligned}$$

On peut aussi obtenir ce résultat en évaluant les dérivées de la fonction génératrice des moments en 0. ■

1.4.6 Processus de comptage Non Stationnaire

Les processus de comptage non stationnaires sont courants dans de nombreux domaines, tels que la finance, l'économie, l'épidémiologie, la biologie, la météorologie, etc. Dans ces domaines, les événements peuvent être des transactions financières, des infections, des naissances, des décès, des précipitations, etc.

Le comportement des processus de comptage non stationnaires peut être décrit à l'aide de différentes caractéristiques, telles que la tendance, la saisonnalité, la périodicité, les ruptures structurelles, etc. La tendance représente la direction générale dans laquelle les événements se déplacent au fil du temps.

La saisonnalité reflète les variations périodiques des événements en fonction de la période de l'année ou de la semaine. La périodicité peut décrire des cycles plus longs

que la saisonnalité, tels que des cycles économiques ou des phénomènes naturels. Les ruptures structurelles représentent des changements brusques et persistants dans le comportement des événements.

Nous limitons nos études à deux exemples de processus de comptage non stationnaires, à savoir le processus de Hawkes et le processus de Cox.

1.4.6.1 Processus de Hawkes

Le processus ponctuel de Hawkes a été introduit pour la première fois par [Hawkes \(1971\)](#) et est devenu l'un des processus les plus populaires dans l'analyse des données à la fois dans les sciences naturelles et sociales et a été largement appliqué dans la finance ([Embrechts et al. \(2013\)](#)) et les assurances, la sismologie, les neurosciences, la violence des gangs, ou les défaillances bancaires et de nombreux autres domaines, voir, par exemple, [Stabile and Torrisi \(2010\)](#) et [Zhu \(2013\)](#), en raison de sa capacité à étudier l'effet de regroupement (interactions positives).

Le processus de Hawkes peut être utilisé pour modéliser une séquence d'un certain type d'arrivées dans le temps. Chaque arrivée excite le processus dans le sens où la probabilité d'une arrivée ultérieure augmente pendant un certain temps après l'arrivée initiale, c'est-à-dire qu'il a une propriété d'auto-excitation, un effet de regroupement et une longue mémoire.

Formellement, un processus de Hawkes non stationnaire $(\tau_i)_i$ sur $(0, \infty)$ est un processus en grappe de Poisson :

Les centres de grappe de $(\tau_i)_i$ sont donnés par des points particuliers appelés immigrants (également appelés événements de fond ou événements spontanés).

Les immigrants sont distribués selon un processus de Poisson I homogène d'intensité ν . Les autres points du processus sont appelés descendants.

Chaque immigrant τ_i génère un cluster $C_{\tau_i} = C_i$, qui est un ensemble aléatoire formé par les points des générations $j = 0, 1, \dots$ avec la structure de ramification suivante : on dit que l'immigrant τ_i est de la génération 0.

Étant donné les générations $0, 1, \dots, j$ dans C_i , chaque point $\tau \in C_i$ de la génération $j + 1$ génère un processus de Poisson sur (τ, ∞) de la descendance de la génération j avec une fonction d'intensité $h(\cdot, \tau)$, où h est une fonction de Borel non négative appelée taux de fertilité.

Nous supposons que $h(t) \neq 0$ pour éviter le cas trivial, qui est, un processus de Poisson homogène. Ainsi, le processus de Hawkes est une extension non-markovienne du processus de Poisson.

Donnons les immigrants, les clusters centrés

$$C_i - \tau_i = \{\tau - \tau_i, \tau \in C_i\}, \quad \tau_i \in I$$

sont *i.i.d* et indépendant de I . $(\tau_i)_i$ est constitué de l'union de tous les clusters

Le processus de Hawkes à une dimension est un processus ponctuel $(N_t)_t$ qui est caractérisé par son intensité λ_t par rapport à sa filtration naturelle, c'est-à-dire la filtration $\{\mathcal{F}_t^S\}$ générée par le processus agrégé $S_t = \sum_{i=1}^{N_t} X_i$.

$$\lambda_t = \nu + \int_0^t h(t-s) dN_s, \quad (1.2)$$

Un processus ponctuel simple $\{N_t\}$ est un processus linéaire de Hawkes s'il a pour intensité

$$\lambda_t = \nu + \sum_{\tau < t} h(t-\tau), \quad (1.3)$$

Le paramètre de criticité, qui est le nombre moyen de descendants directs par ancêtre, est le suivant

$$\kappa = \|h\|_{L^1} = \int_0^\infty h(x) dx < 1 \quad (1.4)$$

1.4.6.2 Processus de Cox avec intensité de bruit de tir

Le processus de Cox est défini par une suite de variables aléatoires, qui représentent les instants d'occurrence des événements rares. Ces variables aléatoires sont généralement modélisées comme des temps d'arrivée de processus de Poisson, qui sont des processus stochastiques avec des événements aléatoires qui se produisent à un rythme moyen constant.

Cependant, contrairement aux processus de Poisson, le processus de Cox permet de modéliser des événements qui ont une intensité variable dans le temps ou dans l'espace. Cela permet de capturer des phénomènes tels que l'agrégation d'événements dans le temps ou dans l'espace, ainsi que des changements de régime dans l'intensité des événements.

Le processus de Cox est largement utilisé dans des domaines tels que la biologie, l'épidémiologie, la finance, la géostatistique et les réseaux de télécommunications, où il est utilisé pour modéliser des événements rares tels que les mutations génétiques, les épidémies, les fluctuations de prix ou les pannes de réseau.

Un processus Cox $(N_t)_t$ avec intensité λ_t qui suit un processus de bruit de tir

$$\lambda_t = \nu(t) + \sum_{\tau^{(1)} < t} g(t-\tau^{(1)}), \quad (1.5)$$

où $\tau^{(1)}$ le temps d'arrivée d'un processus de Poisson $(N_t^{(1)})_t$ externe homogène avec intensité γ .

Ici, $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ est intégrable, c'est-à-dire $\int_0^\infty g(t) dt < \infty$ et $\nu(t)$ est une fonction po-

sitive, continue et déterministe telle que $\nu(t) \rightarrow \nu$ avec $t \rightarrow \infty$. $\nu(t)$ peut représenter d'autres perturbations qui n'ont pas d'importance asymptotiquement.

La probabilité de ruine pour les demandes d'indemnisation à queue lourde, le processus d'arrivée étant un processus de Cox à bruit de tir, sont connues dans la littérature, voir par exemple le livre [Asmussen and Albrecher \(2010\)](#). Mais les techniques décrites dans la littérature utilisent les caractéristiques très spécifiques du processus de Cox .

2.1 Introduction

La théorie asymptotique des extrêmes d'un échantillon a été développée parallèlement à la théorie de la limite centrale, et en fait les deux théories présentent une certaine ressemblance.

Soit X, X_2, X_3, \dots des variables aléatoires indépendantes et identiquement distribuées. La théorie de la limite centrale s'intéresse au comportement limite des sommes partielles $X_1 + X_2 + \dots + X_n$ quand $n \rightarrow \infty$, tandis que la théorie des extrêmes d'un échantillon s'intéresse au comportement limite des échantillons extrêmes $\max(X_1, X_2, \dots, X_n)$ ou $\min(X_1, X_2, \dots, X_n)$ lorsque $n \rightarrow \infty$.

Les deux théories concernent la défaillance.

Un pneu d'une voiture peut tomber en panne de deux façons. Chaque jour de conduite usera un peu le pneu, et après un long moment, la dégradation accumulée entraînera une défaillance (c'est-à-dire que les sommes partielles dépasseront un certain seuil). Mais en conduisant, on peut aussi heurter accidentellement une glissière de sécurité d'une autoroute en acier ou un trottoir.

Soit ces incidents n'ont aucun effet, soit le pneu est crevé. Dans ce dernier cas, c'est une seule grande observation qui provoque la défaillance, ce qui signifie que les sommes partielles dépassent un certain seuil.

La théorie des valeurs extrêmes (TVE) est capable de prédire l'apparition d'événements rares en dehors de la série de données disponibles.

La théorie des valeurs extrêmes (TVE) concerne les questions probabilistes et statistiques des valeurs très élevées ou très faibles dans des suites de variables aléatoires et dans les processus stochastiques.

L'objectif essentiel de ce chapitre est de présenter les définitions et les résultats principaux sur la théorie des valeurs extrêmes dans le cas unidimensionnel.

2.1.1 Statistique d'ordre

Soit X_1, X_2, X_3, \dots et $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, la statistique d'ordre correspondante. nous désignons $F_{(k)}(x)$ et $f_{(k)}(x)$ comme le cdf et pdf respectivement de $X_{(k)}$, $k = 1, 2, \dots, n$. Nous pouvons écrire

$$f_{(k)}(x) = \frac{n!}{(k-1)!(n-k)!} (F(x))^{k-1} (1-F(x))^{n-k} f(x)$$

La fonction de densité de probabilité conjointe de la statistique d'ordre $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, a la forme

$$f_{1,2,\dots,n}(x_1, x_2, \dots, x_n) = \begin{cases} n! \prod_{k=1}^n f(x_k) & \text{si } -\infty < x_1 < x_2 < \dots < x_n < +\infty \\ 0 & \text{sinon} \end{cases}$$

Il y a une forme simple des distributions du maximum $X_{(n)}$ et du minimum $X_{(1)}$. Les cdfs et les pdfs de la plus petite et la plus grande statistiques d'ordre sont données respectivement comme

$$F_{(1)}(x) = 1 - (1 - F(x))^n \quad (2.1)$$

et

$$F_{(n)}(x) = (F(x))^n, \quad (2.2)$$

Et les densités de $X_{(n)}$ et du minimum $X_{(1)}$ sont simplement déduites comme suivant

$$f_{(1)}(x) = n(1 - F(x))^{n-1} f(x) \quad (2.3)$$

et

$$f_{(n)}(x) = n(F(x))^{n-1} f(x) \quad (2.4)$$

La densité conjointe $f_{(1,n)}(x, y)$ du couple $(X_{(1)}, X_{(n)})$ est donnée par

$$f_{(1,n)}(x, y) = n(n-1)(F(y) - F(x))^{n-2} f(x)f(y), \\ -\infty < x < y < \infty.$$

Exemple 2.1 (*Distribution Exponentielle*) Supposons que X_1, X_2, \dots, X_n sont n i.i.d.. variables aléatoires de cdf F avec

$$F(x) = 1 - e^{-x} \cdot x \geq 0.$$

Les dfs $f_{(1)}(x)$ et $f_{(n)}(x)$ de $X_{(1)}$ et $X_{(n)}$ sont respectivement

$$f_{(1)}(x) = ne^{-nx}, x \geq 0.$$

et

$$f_{(n)}(x) = n(1 - e^{-x})^{n-1} e^{-x}, \quad x \geq 0.$$

On peut voir que $nX_{(1)}$ a la distribution exponentielle.

Remarque 2.1 *L'échantillon ordonné*

$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, n'est pas i.i.d.. Pour $i \neq j$, $X_{(i)}$ et $X_{(j)}$ sont dépendants l'une de l'autre; en particulier si $i < j$ alors $X_{(i)} \leq X_{(j)}$.

Deux statistiques d'ordre sont intéressantes pour l'étude des événements extrêmes, ce sont :

$$X_{(1)} = \min(X_1, X_2, \dots, X_n).$$

qui est la plus petite statistique d'ordre ou statistique du minimum, et

$$X_{(n)} = \max(X_1, X_2, \dots, X_n).$$

qui est la plus grande statistique d'ordre ou statistique du maximum.

On s'intéresse au comportement de ces deux statistiques, donc on concentre aux distributions

$$F_{(1)}(x) = P(X_{(1)} < x) \text{ et } F_{(n)}(x) = P(X_{(n)} < x).$$

Nous nous intéresserons en particulier aux maxima de l'échantillon. Puisque

$$\min(X_1, X_2, \dots, X_n) = -\max(-X_1, -X_2, \dots, -X_n).$$

Définition 2.1 (*Quantile et fonction quantile de la queue*)

La fonction quantile de la cdf F est une généralisation de la fonction inverse de F définie par

$$\mathbb{Q}(s) = F^{\leftarrow}(s) := \inf\{x \in \mathbb{R} : F(x) \geq s\}, 0 < s < 1$$

Avec la convention que le l'infiniment de l'ensemble vide est ∞ . En théorie des valeurs extrêmes, la fonction, dénotée par \mathbb{U} est appelée aussi fonction quantile de la queue. Elle est définie par

$$\mathbb{U}(t) := \mathbb{Q}(1 - 1/t) = (1/\bar{F})^{\leftarrow}(t), 1 < t < \infty.$$

Nous allons voir plus loin, que la fonction \mathbb{U} joue un rôle important dans la théorie des valeurs extrêmes comparable au rôle de la fonction caractéristique dans la théorie des distributions stables.

Définition 2.2 (Quantile empirique et fonctions quantiles de la queue)

La fonction quantile empirique de l'échantillon (X_1, X_2, \dots, X_n) est définie par

$$\mathbb{Q}_n(s) := \inf\{x \in \mathbb{R} : F_n(x) \geq s\}, 0 < s < 1.$$

\mathbb{Q}_n peut être exprimé comme une simple fonction de la statistique d'ordre relative à l'échantillon (X_1, X_2, \dots, X_n) . A savoir,

$$\mathbb{Q}_n(s) := X_{(n-k+1)}, \text{ pour } \frac{n-1}{n} \leq s \leq \frac{n-i+1}{n}, i = 1, 2, \dots, n.$$

Si pour $0 < p < 1$, $X_{([np]+1)}$ est le quantile d'échantillon d'ordre p .

Définition 2.3 (Processus du quantile de la queue).

Soit $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ la statistique d'ordre et $k = k(n)$ vérifiant $k \rightarrow \infty, k/n \rightarrow 0$, quand $n \rightarrow \infty$. On définit le processus de la queue (empirique) par le processus stochastique $\{X_{(n-[ks])}\}_{s \geq 0}$.

Définition 2.4 Le point final supérieur (ou droit) de la cdf F est définie comme suit :

$$x_r = \sup\{x : F(x) < 1\}.$$

Le point final inférieur (ou gauche) de la cdf F est définie comme suit :

$$x_\ell = \inf\{x : F(x) > 0\}.$$

Clairement, x_r et x_ℓ représentent, respectivement, les valeurs minimale et maximale atteintes par la v.a.r X associé à F . Évidemment, $x_\ell = -\infty$ dans le cas d'une v.a.r non borné inférieur, et $x_r = +\infty$ pour une v.a.r non bornée supérieur.

2.1.2 Côté Probabiliste de la théorie des valeurs extrêmes

Une discussion statistique traditionnelle sur la moyenne est basée sur le théorème de la limite centrale et revient donc souvent à la distribution normale comme base de l'inférence statistique.

La formule (2.2) montre que la loi du maximum est relié d'une manière principale avec F , mais cette dernière n'est pas toujours connue, même si elle est connue, la loi du

maximum n'est pas facile à calculer.

De plus, nous avons

$$\lim_{n \rightarrow +\infty} F_{X_{(n)}}(x) = \lim_{n \rightarrow +\infty} \{F(x)\}^n = \begin{cases} 1 & \text{si } F(x) = 1 \\ 0 & \text{si } F(x) < 1 \end{cases}$$

On constate que la distribution asymptotique du maximum donne une loi dégénérée (Elle prend les valeurs 0 et 1 seulement). On cherche alors une loi non dégénérée pour $X_{(n)}$.

Donc on s'intéresse à étudier les comportements asymptotiques du maximum en faisant tendre n vers l'infini.

Supposons qu'il existe deux suites de constantes $a_n > 0$, et b_n réelles ($n = 1, 2, \dots$), telles que

$$\frac{\max(X_1, X_2, \dots, X_n) - b_n}{a_n}$$

a une distribution limite non dégénérée comme $n \rightarrow \infty$, c'est-à-dire,

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) \stackrel{L}{=} H(x), \quad (2.5)$$

Un théorème a été proposé par [Gnedenko \(1943\)](#) connu sous le nom de théorème de Fisher-Tippett-Gnedenko ou théorème de la valeur extrême qui donne la forme des lois limites pour le maximum normalisé.

Ce théorème donne un résultat important pour l'existence d'une loi limite non dégénérée pour le maximum $X_{(n)}$.

Les précédentes versions aient été énoncées par [Fréchet \(1927\)](#) et par [Fisher and Tippett \(1928\)](#). [Jenkinson \(1955\)](#) a donné l'expression générale. Pour plus de détails sur cet sujet, on peut se référer à [Galambos \(1978\)](#).

Nous allons exploiter la condition (2.5). En prenant les logarithmes à gauche et à droite, nous obtenons de (2.5) la relation équivalente tel que pour chaque point de continuité x pour lequel $0 < H(x) < 1$,

$$\lim_{n \rightarrow \infty} n \log F(a_n x + b_n) = \log H(x), \quad (2.6)$$

Il s'ensuit clairement que $F(a_n x + b_n) \rightarrow 1$, pour tout x . d'où

$$\lim_{n \rightarrow \infty} \frac{-\log F(a_n x + b_n)}{1 - F(a_n x + b_n)} = 1,$$

et en fait (2.6) est équivalente à

$$\lim_{n \rightarrow \infty} n(1 - F(a_n x + b_n)) = -\log H(x),$$

où

$$\lim_{n \rightarrow \infty} \frac{1}{n(1 - F(a_n x + b_n))} = \frac{1}{-\log H(x)}.$$

Après nous allons reformuler cette condition en termes de l'inverse des fonctions. Pour toute fonction non décroissante f , on note f^{\leftarrow} son inverse continue à gauche i.e.,

$$f^{\leftarrow}(x) := \inf\{y : f(y) \geq x\}.$$

Lemme 2.1 *Supposons $(f_n)_n$ une suite de fonctions non décroissantes et g est fonction non décroissante. Supposons que pour tout x dans un intervalle ouvert (a, b) de point de continuité de g ,*

$$\lim_{n \rightarrow \infty} f_n(x) = g(x).$$

Soient f_n^{\leftarrow} et g^{\leftarrow} les inverses continues à gauche de f_n et g respectivement. Alors, pour tout x dans l'intervalle $(g(a), g(b))$ de points de continuité de g^{\leftarrow} , nous avons

$$\lim_{n \rightarrow \infty} f_n^{\leftarrow}(x) = g^{\leftarrow}(x).$$

Preuve. Soit x un point de continuité de g^{\leftarrow} . Fixons $\varepsilon > 0$. Nous montrons que pour $n, n_0 \in \mathbb{N}, n \geq n_0$,

$$f_n^{\leftarrow}(x) - \varepsilon \leq g^{\leftarrow}(x) \leq f_n^{\leftarrow}(x) + \varepsilon.$$

Nous allons démontrer l'inégalité à droite; la démonstration du côté gauche est similaire.

Choisissons $0 < \varepsilon_1 < \varepsilon$ tel que $g^{\leftarrow}(x) - \varepsilon_1$ est point de continuité g . ceci est possible puisque les points de continuité de g forment un ensemble dense. Si g^{\leftarrow} est continue en x , $g^{\leftarrow}(x)$ est un point de croissance pour g ;

d'où $g(g^{\leftarrow}(x) - \varepsilon_1) < x$.

Prenons $\delta < x - g(g^{\leftarrow}(x) - \varepsilon_1)$. d'où $g^{\leftarrow}(x) - \varepsilon_1$ est un point de continuité de g , il existe n_0 tel que $f_n(g^{\leftarrow}(x) - \varepsilon_1) < g(g^{\leftarrow}(x) - \varepsilon_1) + \delta < x$ pour $n \geq n_0$.

La définition de la fonction $f_{x^{\leftarrow}} \leftarrow$ puis implique que $g^{\leftarrow}(x) - \varepsilon_1 < f_n^{\leftarrow}(x)$. ■

Nous allons appliquer le lemme (2.1) pour la relation (2.6). Soit la fonction \mathbb{U} la fonction inverse continue à gauche de $1/(1 - F)$.

Notons que $\mathbb{U}(t)$ est définie pour $t > 1$, il suit que (2.6) est équivalente à

$$\lim_{n \rightarrow \infty} \frac{\mathbb{U}(nx) - b_n}{a_n} = G^{\leftarrow}(e^{-1/x}) =: D(x), \tag{2.7}$$

Pour tout x positive.

La relation (2.7) est plus simple que (2.6). Donc, on peut rendre (2.7) plus flexible de la manière suivante :

Théorème 2.1 Soit $a_n > 0$ et b_n deux suites des constantes réelles et H une fonction de distribution non-dégénérée. Les assertions suivantes sont équivalentes :

1.

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) \stackrel{d}{=} H(x)$$

2. Pour tout point de continuité x de H , on a

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + b(t))) = -\log H(x), \quad (2.8)$$

3. Pour tout point de continuité x de H on note $a(t) := a_{[t]}$, et $b(t) := b_{[t]}$ (avec $[t]$ la partie entière de t).

$$\lim_{t \rightarrow \infty} \frac{\mathbb{U}(tx) - b(t)}{a(t)} = D(x), \quad (2.9)$$

Pour tout point de continuité de $D(x) = G^{\leftarrow}(e^{-1/x})$, $a(t) := a_{[t]}$, et $b(t) := b_{[t]}$.

Preuve. L'équivalence de (2) et (3) découle du lemme 2.1. Nous avons déjà vérifié que (1) est équivalent à (2.7). Il suffit donc de prouver que (2.7) implique (3). Soit x un point de continuité de D .

Pour $t \geq 1$,

$$\frac{\mathbb{U}([t]x) - b_{[t]}}{a_{[t]}} \leq \frac{\mathbb{U}(tx) - b_{[t]}}{a_{[t]}} \leq \frac{\mathbb{U}([t]x(1 + 1/[t])) - b_{[t]}}{a_{[t]}}.$$

Le côté droit est éventuellement inférieur à $D(x')$ pour tout point de continuité $x' > x$ avec $D(x') > D(x)$. Puisque D est continue en x , nous obtenons

$$\lim_{t \rightarrow \infty} \frac{\mathbb{U}(tx) - b_{[t]}}{a_{[t]}} = D(x).$$

■

Théorème 2.2 (Fisher and Tippett (1928), Gnedenko (1943)) La classe des distributions des valeurs extrêmes est $H_\xi(ax + b)$ avec $a > 0, b$ réel, où

$$H_\xi(x) = \exp\left(-(1 + \xi x)^{-1/\xi}\right), \quad 1 + \xi x > 0, \quad (2.10)$$

avec ξ est un nombre réel et pour $\xi = 0$, le côté droite est interprété comme $\exp(-e^{-x})$.

Le paramètre ξ dans (2.10) est appelé la valeur indice extrême.

La démonstration de ce théorème est donnée par de Haan and Ferreira (2006).

Preuve. Considérons la classe des fonctions limites D dans (2.9).

Supposons d'abord que 1 est un point de continuité de D .

Remarquons ensuite que pour les points de continuité $x > 0$.

$$\lim_{t \rightarrow \infty} \frac{\mathbb{U}(tx) - \mathbb{U}(t)}{a(t)} = D(x) - D(1) =: \mathbb{M}(x). \quad (2.11)$$

Prenons $y > 0$ et écrivons

$$\frac{\mathbb{U}(txy) - \mathbb{U}(t)}{a(t)} = \frac{\mathbb{U}(txy) - \mathbb{U}(ty)}{a(ty)} \frac{a(ty)}{a(t)} + \frac{\mathbb{U}(ty) - \mathbb{U}(t)}{a(t)}. \quad (2.12)$$

Posons

$$\mathbb{A}(y) := \lim_{t \rightarrow \infty} \frac{a(ty)}{a(t)}$$

\mathbb{A} existe pour $y > 0$, et pour $x, y > 0$, on a

$$\mathbb{M}(xy) = \mathbb{M}(x)\mathbb{A}(y) + \mathbb{M}(y).$$

Donc si on pose $s := \log x, t := \log y (x, y \neq 1)$, et $K(x) := \mathbb{M}(e^x)$, nous aurons

$$K(t+s) = K(s)\mathbb{A}(e^t) + K(t) \quad (2.13)$$

On peut donc (puisque $K(0) = 0$)

$$\frac{K(t+s) - K(t)}{s} = \frac{K(s) - K(0)}{s} \mathbb{A}(e^t). \quad (2.14)$$

Il y a certainement un t dont H est différentiable, donc par (2.13) K est différentiable partout et

$$K'(t) = K'(0)\mathbb{A}(e^t). \quad (2.15)$$

Posons $Q(t) := K(t)/K'(0)$.

Remarquons que $K'(0)$ ne peut pas être nulle : K ne peut être constante puisque K est non-dégénérée.

Alors $Q(0) = 0, Q'(0) = 1$.

Avec la formule (2.14), on aura

$$Q(t+s) - Q(t) = Q(s)\mathbb{A}(e^t),$$

et par (2.15), on aura

$$Q(t+s) - Q(t) = Q(s)Q'(t)$$

En soustrayant la même expression avec t et s interchangés, on obtient

$$Q(t) \frac{Q'(s) - 1}{s} = \frac{Q(s)}{s} (Q'(t) - 1)$$

avec $s \rightarrow 0$

$$Q(t)Q''(0) = Q'(t) - 1.$$

Il s'ensuit que Q est deux fois différentiable, et par différentiation,

$$Q''(0)Q'(t) = Q''(t).$$

D'où

$$(\log Q')'(t) = Q''(0) =: \xi \in \mathbb{R},$$

pour tout t . Il s'ensuit que (notez que $Q'(0) = 1$)

$$Q'(t) = e^{\xi t}$$

et (puisque $Q(0) = 0$)

$$Q(t) = \int_0^t e^{\xi z} ds$$

Cela signifié que

$$K(t) = K'(0) \frac{e^{\xi t} - 1}{\xi}$$

et

$$D(t) = D(1) + K'(0) \frac{t^\xi - 1}{\xi}.$$

D'où

$$D^+(x) = \left(1 + \xi \frac{x - D(1)}{K'(0)} \right)^{1/\xi}. \quad (2.16)$$

Donc $D(x) = H^{\leftarrow}(e^{-1/x})$, d'où

$$D^+(x) = \frac{1}{-\log H(x)}. \quad (2.17)$$

En Combinant (2.16) et (2.17), nous obtenons la proposition du théorème. Si 1 n'est pas un point de continuité de D , utiliser dans la preuve la fonction $U(tx_0)$ avec x_0 un point de continuité de D . ■

Remarque 2.2 La paramétrisation du théorème (2.2) est due à [Von Mises \(1936\)](#) et [Jenkinson \(1955\)](#).

Ce théorème est un résultat important. Il montre que les fonctions de distribution limite forment une famille simple et explicite à un paramètre, à l'exception des paramètres d'échelle et de localisation.

La figure 2.1 illustre cette famille pour certaines valeurs de ξ .

De plus, elle montre que la classe contient des distributions aux caractéristiques assez différentes.

Considérons les sous-classes $\xi > 0$, $\xi = 0$, et $\xi < 0$ séparément :

1. Pour $\xi > 0$, il est clair que $H_\xi(x) < 1$ pour tout x , c'est-à-dire que l'extrémité droite de la distribution est l'infini.

De plus, quand $x \rightarrow \infty$, $1 - H_\xi(x) \sim \xi^{-1/\xi} x^{-1/\xi}$, c'est-à-dire que la distribution a une queue droite assez lourde, par exemple, les moments d'ordre supérieur ou égal à $1/\xi$ n'existent pas.

2. Pour $\xi = 0$, la queue droite de la distribution est égale à l'infini. La distribution, cependant, est plutôt à queue légère : $1 - H_0(x) \sim e^{-x}$ quand $x \rightarrow \infty$, et tous les moments existent.

3. Pour $\xi < 0$, l'extrémité droite de la distribution est $-1/\xi$, elle a donc une queue courte, vérifiant $1 - H_\xi(-\xi^{-1} - x) \sim (-\xi x)^{-1/\xi}$, quand $x \downarrow 0$.

Remarque 2.3 Une paramétrisation alternative est la suivante :

1. Pour $\xi > 0$ utiliser $H_\xi((x-1)/\xi)$ et obtenir avec $\alpha = 1/\xi > 0$,

$$\Phi_\alpha(x) := \begin{cases} \exp(-x^{-\alpha}), & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Cette classe est souvent appelée la classe des distributions de Fréchet ([Fréchet \(1927\)](#)).

2. La fonction de distribution avec $\xi = 0$,

$$\Lambda_0(x) = \exp(-e^{-x}),$$

pour tous les réels x , est appelée la distribution double-exponentielle ou distribution de Gumbel.

3. Pour $\xi < 0$ use $H_\xi(-(1+x)/\xi)$ et obtenir avec $\alpha = -1/\xi > 0$,

$$\Psi_\alpha(x) := \begin{cases} \exp(-(-x)^\alpha), & x < 0, \\ 0 & x \geq 0. \end{cases}$$

Cette classe est parfois appelée la classe de distribution de Weibull inverse.

Rappelons que si la relation (2.10) est vérifiée avec $H = H_\xi$ pour un certain $\xi \in \mathbb{R}$, on dit

que la fonction de distribution F est dans le domaine d'attraction de H_ξ . On note dans ce cas $F \in \mathcal{D}(H_\xi)$.

Le résultat du Théorème (2.2) conduit à la reformulation suivante du Théorème (2.1)

Exemple 2.2 (Maximum d'une v.a.r exponentielle)

Soit (X_i) une suite de v.a.r exponentielles standards i.i.d.. Alors

$$\begin{aligned} P(X_{(n)} - \ln n \leq x) &= (P(X \leq x + \ln n))^n \\ &= (1 - n^{-1}e^{-x})^n \\ &\rightarrow \exp\{-e^{-x}\} = \Lambda(x), \quad x \in \mathbb{R}. \end{aligned}$$

A titre de comparaison, rappelons que pour les v.a.r. de Gumbel i.i.d. X_i ,

$$P(X_{(n)} - \ln n \leq x) = \Lambda_0(x), \quad x \in \mathbb{R}.$$

Exemple 2.3 (Maxima de la loi de Cauchy)

Soit $(X_n)_n$ une séquence de v.a.r de Cauchy standard i.i.d.

La distribution standard de Cauchy est absolument continue avec la densité

$$f(x) = (\pi(1+x^2))^{-1}, \quad x \in \mathbb{R}.$$

Par la règle de l'Hospital, on obtient

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x)}{(\pi x)^{-1}} = \lim_{x \rightarrow \infty} \frac{f(x)}{\pi^{-1}x^{-2}} = \lim_{x \rightarrow \infty} \frac{\pi x^2}{\pi(1+x^2)} = 1,$$

donnant $\bar{F}(x) \sim (\pi x)^{-1}$. Cela implique

$$\begin{aligned} P\left(X_{(n)} \leq \frac{nx}{\pi}\right) &= \left(1 - \bar{F}\left(\frac{nx}{\pi}\right)\right)^n \\ &= \left(1 - \frac{1}{nx} + o(1)\right)^n \\ &\rightarrow \exp\{-x^{-1}\} = \Phi_1(x), \quad x > 0 \end{aligned}$$

2.2 Domaine d'Attraction

Dans la section précédente, nous avons identifié les distributions de valeurs extrêmes comme les lois limites pour les maxima normalisés des v.a.r i.i.d., voir le théorème (2.1).

Cette section est consacrée à la question suivante.

Étant donné une distribution de valeurs extrêmes H , quelles sont les conditions sur le

df et le cdf qui impliquent que les maxima normalisés $X_{(n)}$ convergent faiblement vers H ?

Comment pouvons-nous choisir les constantes de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$ telles que

$$a_n^{-1} (X_{(n)} - b_n) \xrightarrow{L} H \quad (2.18)$$

Peut-il arriver que des constantes de normalisation différentes impliquent une convergence vers des lois limites différentes?

Définition 2.5 (*Domaine d'attraction maximal*) On dit que la v.a.r X de distribution F appartient au v.a.r domaine d'attraction maximal $\mathcal{D}(H)$ de la distribution des valeurs extrêmes H s'il existe des constantes $a_n > 0$, $b_n \in \mathbb{R}$ telles que (2.16) est vérifiée. Nous écrivons $X \in \mathcal{D}(H)$ ($F \in \mathcal{D}(H)$).

Les cdfs des valeurs extrêmes sont continues sur \mathbb{R} , donc $a_n^{-1} (X_{(n)} - b_n) \xrightarrow{d} H$ est équivalent à

$$\lim_{n \rightarrow \infty} P(X_{(n)} \leq a_n x + b_n) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x), \quad x \in \mathbb{R}.$$

Le résultat suivant est une conséquence immédiate de la Proposition 2.1 et sera utilisé tout au long des sections suivantes.

Proposition 2.1 (*Caractérisation de $\mathcal{D}(H)$*)

Le cdf F appartient au domaine d'attraction maximal de la distribution extrêmes H avec des constantes de normalisation $a_n > 0$, $b_n \in \mathbb{R}$ si et seulement si

$$\lim_{n \rightarrow \infty} n\bar{F}(a_n x + b_n) = -\ln H(x), \quad x \in \mathbb{R}.$$

Si $H(x) = 0$, la limite est interprétée comme ∞ .

Selon le signe de ξ , on définit trois domaines d'attraction :

1. Lorsque $\xi = 0$ la loi des valeurs extrêmes présente une décroissance de type exponentiel dans la queue de la loi, on dit alors que la loi appartient au domaine d'attraction de Gumbel. C'est le cas des lois normale, exponentielle, etc ...
2. Le cas $\xi > 0$ correspond à la loi de Fréchet de paramètre $\alpha = 1/\xi$ dont la fonction de survie décroît comme une fonction puissance, et dans ce cas on dit que la loi appartient au domaine d'attraction de Fréchet. C'est le cas par exemple de la loi de Cauchy ou Pareto.
3. Le cas $\xi < 0$ correspondant à la loi de Weibull de paramètre $\alpha = -1/\xi$ et donc la loi appartient au domaine d'attraction de Weibull. c'est le cas par exemple

de la loi uniforme ou béta. On peut trouver un classement de différents lois par domaine d'attraction dans [Embrechts et al. \(2013\)](#) ou [Coles et al. \(2001\)](#).

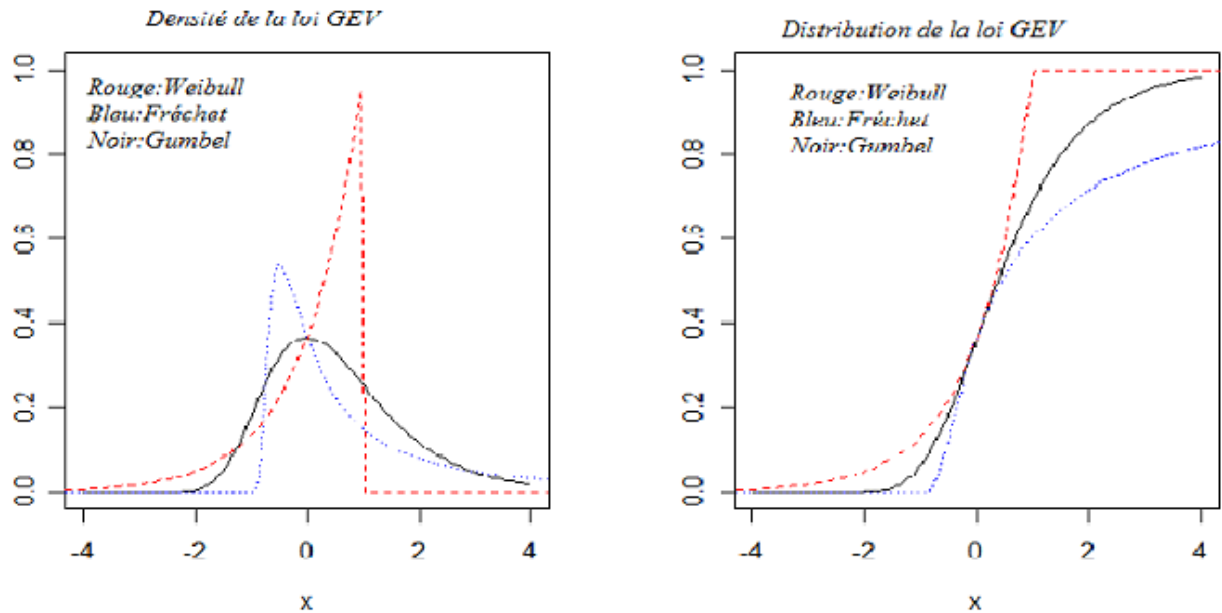


FIGURE 2.1 – Exemples de densités et distributions associées à la loi des valeurs extrêmes (noir : $\xi = 0$,bleu : $\xi = 1$,rouge : $\xi = -1$)

Dans la suite, on va présenter quelques résultats qui permettent de caractériser les trois domaines d'attraction. Pour cela, on va faire un appel à quelques notions principales.

2.2.1 Domaine d'attraction de la distribution de Fréchet

[Gnedenko\(1943\)](#) a énoncé un résultat assure que toute fonction appartenant au domaine d'attraction de Fréchet est une fonction à variations régulières.

Théorème 2.3 Une fonction de répartition F appartient au domaine d'attraction de Fréchet si et seulement si elle s'écrit sous la forme :

$$\bar{F}(x) = x^{-1/\xi}L(x) \tag{2.19}$$

telle que L est une fonction à variations lentes.

Dans ce cas les suites de normalisation peuvent être choisi avec $b_n = 0$ et les a_n au moyen de

la fonction quantile, plus précisément par

$$\begin{aligned} a_n &= F^{\leftarrow}(1 - n^{-1}) = \inf\{x \in \mathbb{R} : F(x) \geq 1 - n^{-1}\} \\ &= \inf\{x \in \mathbb{R} : (1/\bar{F})(x) \geq n\} \\ &= (1/\bar{F})^{\leftarrow}(n) \end{aligned} \tag{2.20}$$

Nous montrons que le domaine d'attraction maximal de Φ_ξ consiste en cdfs F dont la queue droite a une variation régulière d'indice $\alpha = -1/\xi$. Pour $F \in \mathcal{D}(\Phi_\xi)$.

On rappelle que \bar{F} est la fonction de survie définie par $\bar{F}(x) = 1 - F(x)$

[Bingham \(2007\)](#) ont montré que l'équation (2.19) est équivalente à :

$$Q(p) = p^{-\xi} L(p^{-1})$$

où $L \in \mathcal{R}_0$, $p \in [0, 1]$, et $Q(p)$ est la fonction quantile définie par

$$Q(p) = \inf\{x, F(x) \geq \alpha\}$$

Corollaire 2.1 (Von Mises condition)

Soit F une cdf absolument continu de densité f satisfaisant

$$\lim_{x \rightarrow \infty} \frac{xf(x)}{\bar{F}(x)} = \alpha > 0,$$

alors $F \in \mathcal{D}(\Phi_\alpha)$.

Exemple 2.4 (Loggamma distribution)

La distribution loggamma a une queue

$$\bar{F}(x) \sim \frac{\xi^{\beta-1}}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{-\xi}, \quad x \rightarrow \infty, \quad \alpha, \beta > 0.$$

Donc $\bar{F} \in \mathcal{R}_{-\alpha}$ qui est équivalent à $F \in \mathcal{D}(\Phi_\xi)$. Selon la proposition (2.1), nous choisissons a_n l'équation (2.18). En appliquant (2.20) et en prenant les logarithmes, nous trouvons que nous devons résoudre

$$\xi \ln a_n - (\beta - 1) \ln \ln a_n - \ln(\xi^{\beta-1}/\Gamma(\beta)) = \ln n. \tag{2.21}$$

La solution satisfait

$$\ln a_n = \xi^{-1} (\ln n + \ln r_n).$$

où $\ln r_n = o(\ln n)$ as $n \rightarrow \infty$.

On remplace cette expression dans équation (2.21), on obtient

$$\ln r_n = (\beta - 1) \ln(\xi^{-1} \ln n(1 + o(1))) + \ln(\xi^{\beta-1}/\Gamma(\beta)).$$

Cela donne les constantes de normalisation

$$a_n \sim ((\Gamma(\beta))^{-1}(\ln n)^{\beta-1}n)^{1/\xi}.$$

D'où

$$((\Gamma(\beta))^{-1}(\ln n)^{\beta-1}n)^{-1/\xi} X_{(n)} \xrightarrow{\dot{a}} \Phi_\xi.$$

TABLE 2.1 – Domaine d'attraction maximal de la distribution de Fréchet.

Fréchet	$\Phi_\xi(x) = \exp\{-x^{-1/\xi}\}, \quad x > 0, \xi > 0.$
$\mathcal{D}(\Phi_\xi)$	$x_r = \infty, \quad \bar{F}(x) = x^{-1/\xi}L(x), \quad L \in \mathcal{R}_0.$
constantes de normalisation	$a_n = F^{\leftarrow}(1 - n^{-1}) = n^{1/\xi}L_1(n), \quad L_1 \in \mathcal{R}_0, \quad b_n = 0.$
Limite	$X_{(n)} \xrightarrow{L} \Phi_\xi$

2.2.2 Domaine d'attraction de la distribution de Weibull

Gnedenko (1943) et Resnick (1987) ont démontré que par un simple changement de variable dans la fonction de répartition, qu'on peut passer du domaine d'attraction de Fréchet à celui de Weibull.

$$\Psi_\xi(-x^{-1}) = \Phi_\xi(x), \quad x > 0$$

Par conséquent, nous pouvons nous attendre à ce que $\mathcal{D}(\Psi_\xi)$ et $\mathcal{D}(\Phi_\xi)$ soient également étroitement liés. Le théorème suivant confirme cela.

Théorème 2.4 Une fonction de répartition F appartient au domaine d'attraction de Weibull si et seulement si son point terminal x_r est fini et si la fonction de répartition F définie par :

$$F^*(x) = \begin{cases} 0 & \text{si } x < 0 \\ F(x_r - 1/x) & x \geq 0 \end{cases} \quad (2.22)$$

appartient au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes. Ainsi, une fonction de répartition F du domaine d'attraction de Weibull s'écrit :

$$F(x) = 1 - (x_r - x)^{-1/\xi}L((x_r - x)^{-1}), \quad L \in \mathcal{R}_0 \text{ pour } x \leq x_r$$

De manière équivalente, le quantile $Q(p), p \in \mathcal{R}_0$ associé s'écrit :

$$Q(p) = x_r - p^{-\xi} L(1/p) \quad (2.23)$$

Les suites de normalisation (a_n) et (b_n) sont données par $a_n = b_n - \bar{F}^{\leftarrow}(1/n)$ et $b_n = x_r$. Ce domaine d'attraction a été considéré par [Falk \(1995\)](#), [Gardes \(2003\)](#) pour donner un estimateur de point terminal d'une distribution.

Exemple 2.5 (Distribution Uniforme sur $(0, 1)$)

De toute évidence, $x_r = 1$ et $\bar{F}(1 - x^{-1}) = x^{-1} \in \mathcal{R}_{-1}$.

Alors par le Théorème (2.4) nous obtenons $F \in \mathcal{D}(\Psi_1)$.

Puisque $\bar{F}(1 - n^{-1}) = n^{-1}$, nous choisissons $a_n = n^{-1}$.

Cela implique en particulier

$$n(X_{(n)} - 1) \xrightarrow{d} \Psi_1.$$

Exemple 2.6 (Comportement de loi de puissance à l'extrémité droite finie) Soit F un cdf avec une extrémité droite finie x_r et une queue de distribution

$$\bar{F}(x) = K(x_r - x)^\xi, \quad x_r - K^{-1/\xi} \leq x \leq x_r, \quad K, \xi > 0.$$

Alors par le Théorème (2.4) cela garantit que $F \in \mathcal{D}(\Psi_\xi)$.

Les constantes de normalisation a_n peuvent être choisies de telle sorte que $F(x_r - a_n) = n^{-1}$, i.e. $a_n = (nK)^{-1/\xi}$ et, en particulier,

$$(nK)^{1/\xi} (X_{(n)} - x_r) \xrightarrow{d} \Psi_\xi.$$

TABLE 2.2 – Domaine d'attraction maximal de distribution de Weibull.

Weibull	$\Psi_\xi(x) = \exp\{-(-x)^{1/\xi}\}, \quad x < 0, \quad \xi > 0.$
$\mathcal{D}(\Psi_\xi)$	$x_r < \infty, \quad \bar{F}(x_r - x^{-1}) = x^{-1/\xi} L(x), \quad L \in \mathcal{R}_0.$
Constantes de normalisation	$a_n = x_r - F^+(1 - n^{-1}) = n^{-1/\xi} L_1(n), \quad L_1 \in \mathcal{R}_0, \quad b_n = x_r$
Limite	$a_n^{-1} (X_{(n)} - x_r) \xrightarrow{L} \Psi_\xi$

2.2.3 Domaine d'attraction de Gumbel

Le domaine d'attraction de Gumbel contient les lois où la fonction de survie est à décroissance exponentielle, i.e. les lois à queues légères. Le théorème suivant qui a été démontré dans [Resnick \(1987\)](#) donne une caractérisation de ce domaine.

Théorème 2.5 Une fonction de répartition F appartient au domaine d'attraction de Gumbel si et seulement si il existe $t < x_r \leq \infty$ tel que

$$\bar{F}(x) = c(x) \exp \left\{ - \int_t^x \frac{1}{\mathcal{K}(u)} du, t < x \leq x_r \right\}. \quad (2.24)$$

où $c(x) \rightarrow c > 0$ lorsque $x \rightarrow x_r$ et a est une fonction positive et dérivable de dérivée \mathcal{K}' telle que $\mathcal{K}'(x) \rightarrow 0$ lorsque $x \rightarrow x_r$.

Dans ce cas, un choix possible pour les suites (a_n) et (b_n) pour tout $n > 0$ est :

$$a_n = F^{\leftarrow} (1 - n^{-1}) \text{ et } b_n = \frac{1}{\bar{F}(a_n)} \int_{a_n}^{x_r} \bar{F}(s) ds.$$

Nous pouvons maintenant montrer que les fonctions de von Mises appartiennent au domaine d'attraction maximal de la distribution de Gumbel.

De plus, la forme spécifique de F permet de calculer les constantes de normalisation a_n à partir de la fonction auxiliaire \mathcal{K} .

Proposition 2.2 (Fonctions Von Mises et $\mathcal{D}(\Lambda_0)$) Supposons que la df F soit une fonction de von Mises.

Alors $F \in \mathcal{D}(\Lambda_0)$. Un choix possible de constantes de normalisation est

$$b_n = F^{\leftarrow} (1 - n^{-1}) \text{ et } a_n = \mathcal{K}(b_n),$$

où a est la fonction auxiliaire de F .

Proposition 2.3 (Propriété de fermeture de $\mathcal{D}(\Lambda_0)$ sous équivalence de queues)

Soit F et G des cdfs avec le même point droite final, et supposons que $F \in \mathcal{D}(\Lambda_0)$ avec des constantes de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$; c.-à-d.

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \Lambda_0(x), \quad x \in \mathbb{R}.$$

d'où

$$\lim_{n \rightarrow \infty} G^n(a_n x + b_n) = \Lambda_0(x + c), \quad x \in \mathbb{R},$$

si et seulement si F et G ont des queues équivalentes avec

$$\lim_{x \uparrow x_r} \frac{\bar{F}(x)}{\bar{G}(x)} = e^c.$$

Exemple 2.7 (Distribution Normale)

Désignons par Φ la cdf et par φ la densité de la distribution normale standard.

Nous montrons d'abord que Φ vérifie la condition du théorème (2.5).

Une application de la règle de l'Hospital à $\bar{\Phi}(x)/(x^{-1}\varphi(x))$ et en utilisant le fait que

$$\lim_{x \rightarrow \infty} \frac{\bar{\Phi}(x)\varphi'(x)}{\varphi^2(x)} = -1.$$

Nous pouvons déduire les constantes de normalisation en utilisant le rapport suivant :

$$\bar{\Phi}(x) \sim \frac{\varphi(x)}{x} = \frac{1}{\sqrt{2\pi x}} e^{-x^2/2}, \quad x \rightarrow \infty,$$

et interpréter le côté droit comme la queue d'un certain $df G$.

Alors par la Proposition (2.3), Φ et G ont les mêmes constantes de normalisation a_n et b_n .

D'après la proposition (2.2), $b_n = G^{\leftarrow}(1 - n^{-1})$. On cherche donc une solution de $-\ln \bar{G}(b_n) = \ln n$; i.e.

$$\frac{1}{2}b_n^2 + \ln b_n + \frac{1}{2} \ln 2\pi = \ln n.$$

$$b_n = (2 \ln n)^{1/2} - \frac{\ln \ln n + \ln 4\pi}{2(2 \ln n)^{1/2}} + o((\ln n)^{-1/2})$$

comme choix possible pour b_n . Puisque l'on peut prendre $a(x) = \bar{\Phi}(x)/\varphi(x)$ on a que $\mathcal{K}(x) \sim x^{-1}$ et donc

$$a_n = \mathcal{K}(b_n) \sim (2 \ln n)^{-1/2}.$$

Comme les a_n sont uniques et à équivalence près (asymptotique), nous choisissons

$$a_n = (2 \ln n)^{-1/2}$$

Nous en concluons que

$$\sqrt{2 \ln n} \left(X_{(n)} - \sqrt{2 \ln n} + \frac{\ln \ln n + \ln 4\pi}{2(2 \ln n)^{1/2}} \right) \xrightarrow{d} \Lambda_0.$$

Notons que $a_n \rightarrow 0$, c'est-à-dire que la distribution de $X_{(n)}$ devient moins étalée autour de b_n lorsque n augmente.

TABLE 2.3 – Domaine d'attraction maximal de la distribution de Gumbel.

Gumbel	$\Lambda_0(x) = \exp\{-e^{-x}\}, \quad x \in \mathbb{R}.$
$\mathcal{D}(\Lambda_0)$	$x_r \leq \infty, \quad \bar{F}(x) = c(x) \exp\left\{-\int_{x_0}^x \frac{g(t)}{\mathcal{K}(t)} dt\right\}, \quad x_0 < x < x_r.$ où $c(x) \rightarrow c > 0, \quad g(x) \rightarrow 1, \quad \mathcal{K}'(x) \rightarrow 0$ quand $x \uparrow x_r$.
Constantes de normalisation	$a_n = F^{\leftarrow}(1 - n^{-1}), \quad b_n = \mathcal{K}(b_n).$
Limite	$a_n^{-1} (X_{(n)} - b_n) \xrightarrow{L} \Lambda_0$

Nous pouvons résumer le classement de quelques lois par domaine d'attraction

maximal présenté dans le tableau suivant (2.4).

TABLE 2.4 – Quelques lois et leurs domaines d’attraction.

domaines d’attraction	Fréchet ($\xi > 0$)	Weibull ($\xi < 0$)	Gumbel ($\xi = 0$)
Distributions	Cauchy	Uniforme	Gumbel
	Pareto	Beta	Exponentielle
	Student		Lognormale
	Burr		Gamma
			Weibull

La distribution la plus utilisée sur les données financières ou macro-économiques est la distribution de Fréchet. La distribution de Weibull est écartée parce qu’une distribution de Weibull ne peut pas être rare, et que les variations en finance ou en économie sont généralement non bornées en ce sens que les variations extrêmes peuvent toujours être détectées.

La distribution de Gumbel est également extrêmement rare. Cependant, cette dernière distribution est utilisée comme référence et permet ainsi de mesurer, notamment graphiquement, la différence entre les distributions empiriques et la distribution normale. Le résultat du Théorème (2.2) conduit à la réformulation suivante du Théorème (2.1)

2.2.4 Distribution des valeurs extrêmes généralisées et la distribution de Paréto généralisée

Le théorème suivant est l’un des résultats fondamentaux de la théorie des valeurs extrêmes. Il donne, de manière analytique et synthétique, les informations essentielles recueillies dans la section précédente sur les domaines d’attraction maximale. De plus, il constitue la base de nombreuses techniques statistiques.

Théorème 2.6 *Pour $\xi \in \mathbb{R}$ les affirmations suivantes sont équivalentes :*

1. *Il existe des constantes réelles $a_n > 0$ et b_n réelles telles que*

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H_\xi(x) = \exp\left(- (1 + \xi x)^{-1/\xi}\right),$$

Pour tout x avec $1 + \xi x > 0$.

2. *Il existe une fonction positive \mathcal{K} telle que pour $x > 0$,*

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{\mathcal{K}(t)} = D_\xi(x) = \frac{x^\xi - 1}{\xi}, \quad (2.25)$$

où, pour $\xi = 0$, le côté droit est interprété comme $\log x$.

3. Il existe une fonction a positive telle que

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + U(t))) = (1 + \xi x)^{-1/\xi},$$

pour tout x avec $1 + \xi x > 0$.

4. Il existe une fonction positive \mathcal{E} telle que

$$\lim_{u \uparrow x_r} \frac{\bar{F}(u + x\mathcal{E}(u))}{\bar{F}(u)} = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{si } \xi \neq 0. \\ e^{-x} & \text{si } \xi = 0. \end{cases} \quad (2.26)$$

pour tout x pour lequel $1 + \xi x > 0$,

où $x_r = \sup\{x : F(x) < 1\}$. De plus, (2.25) est vrai avec $b_n = U(n)$ et $a_n = \mathcal{K}(n)$. De même, (2.26) est valable avec $\mathcal{E}(t) = 1/(1 - F(t))$.

Preuve. Nous ne donnons ci-dessous que les idées principales afin de montrer que les différentes conditions. Vous pouvez trouver plus de détails dans la littérature, par exemple dans [de Haan and Ferreira \(2006\)](#).

(1) \Leftrightarrow (2).

Pour $\xi = 0$ c'est le théorème 2.5

Pour $\xi > 0$, nous avons $H_\xi(x) = \Phi_\xi(\xi(x + 1/\xi))$ Pour $\xi = 1/\alpha$.

Par le théorème (2.6), (1) est donc équivalente à $\bar{F} \in \mathcal{R}_{-\alpha}$. Par la représentation du théorème pour les fonctions à variations régulières, pour un $z > 0$,

$$\bar{F}(x) = c(x) \exp \left\{ - \int_z^x \frac{1}{\mathcal{K}(t)} dt \right\}, \quad z < x < \infty,$$

où $c(x) \rightarrow c > 0$ and $\mathcal{K}(x)/x \rightarrow \alpha^{-1}$ as $x \rightarrow \infty$ uniforme localement. D'où

$$\lim_{u \rightarrow \infty} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \left(1 + \frac{x}{\alpha}\right)^{-\alpha},$$

qui est (2.26). Si on a (2), on peut choisir $b_n = (1/\bar{F})^{\leftarrow}(n) = U(n)$, d'où $(1/\bar{F})^{\leftarrow}(b_n) \sim n$. et avec $u = b_n$ dans (2.26),

$$\left(1 + \frac{x}{\alpha}\right)^{-\alpha} = \lim_{n \rightarrow \infty} \frac{\bar{F}(b_n + xa(b_n))}{\bar{F}(b_n)} = \lim_{n \rightarrow \infty} n \bar{F}(b_n + xa(b_n)).$$

par conséquent, en vertu de la proposition (2.1), $F \in \mathcal{D}(H_\xi)$ pour $\xi = \alpha^{-1}$.

Le cas $\xi < 0$ peut être traité de la même manière.

(2) \Leftrightarrow (3) Nous nous limitons au cas $\xi \neq 0$, la preuve pour $\xi = 0$ étant analogue. Pour simplifier, nous supposons que F est continue et croissante sur $(-\infty, x_r)$. Si l'on fixe $s = 1/F(u)$, alors (2.26) se traduit par

$$A_s(x) = (s\bar{F}(U(s) + xa(U(s)))^{-1} \rightarrow (1 + \xi x)^{1/\xi}, \quad s \rightarrow \infty.$$

Maintenant pour chaque $s > 0$, $A_s(x)$ est décroissante et pour $s \rightarrow \infty$ converge vers une fonction continue, également $A_s^{\leftarrow}(t)$ converge de manière ponctuelle vers l'inverse de la fonction limite correspondante, i.e.,

$$\lim_{s \rightarrow \infty} \frac{U(st) - U(s)}{a(U(s))} = \frac{t^\xi - 1}{\xi}.$$

Maintenant, l'équation (2.20) est obtenu en utilisant (2.25) pour $t = x$ et $t = y$ et en prenant le quotient. La preuve de la réciproque peut être donnée selon la même méthode.

■

Bien que, pour la modélisation, les types de Λ_0, Φ_α et Ψ_α soient très différents, d'un point de vue mathématique ils sont étroitement liés. En effet, on vérifie immédiatement les propriétés suivantes. Supposons que $X > 0$, alors

$$X \text{ a cdf } \Phi_\alpha \iff \ln X^\alpha \text{ a cdf } \Lambda_0 \iff -X^{-1} \text{ a cdf } \Psi_\alpha.$$

Définition 2.6 (*Distribution de valeurs extrêmes et v.a.r extrême*)

Les cdfs Φ_α, Ψ_α et Λ_0 telles que présentées dans la (2.3) sont appelées distributions de valeurs extrêmes standard, les v.a.r correspondantes sont dites des v.a.r extrêmes standard.

Nous pouvons montrer que la loi limite des extrêmes à toujours la même forme. Les trois formules précédentes peuvent être combinées en une seule paramétrisation :

$$H_\xi(x) = \begin{cases} \exp\left(- (1 + \xi x)^{\frac{-1}{\xi}}\right) & \text{si } \xi \neq 0, 1 + \xi x > 0 \\ \exp(-\exp(-x)) & \xi = 0, -\infty \leq x \leq +\infty \end{cases} \quad (2.27)$$

où H est une fonction de répartition non-dégénérée et ξ est un paramètre qui contrôle la lourdeur de la queue de loi appelé indice des valeurs extrêmes (ou indice de queue).

Cette loi est appelée loi de valeurs extrêmes généralisée (Generalized Extreme Value) que l'on note GEV. La forme la plus générale de la GEV est :

$$H_{\xi,u,\beta}(x) = \exp \left\{ - \left(1 + \xi \frac{x-u}{\beta} \right)^{-1/\xi} \right\}, \quad \xi \neq 0, 1 + \xi \frac{x-u}{\beta} > 0 \quad (2.28)$$

u et β sont respectivement les paramètres de localisation et de dispersion. ξ est le paramètre de forme.

La condition (2.26) a une interprétation probabiliste intéressante. En effet, si X soit une v.a.r avec cdf $F \in \mathcal{D}(H_\xi)$, alors (2.26) se reformule comme suit

$$\lim_{u \uparrow x_r} P \left(\frac{X-u}{\mathcal{E}(u)} > x \mid X > u \right) = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{si } \xi \neq 0 \\ e^{-x} & \text{si } \xi = 0 \end{cases} \quad (2.29)$$

Par conséquent, (2.29) donne une approximation de la distribution pour les excès au-dessus d'un seuil (élevé) u .

Le facteur d'échelle approprié est $\mathcal{E}(u)$.

Cette interprétation s'avérera cruciale dans de nombreuses applications.

Définition 2.7 (Fonction de distribution d'excès, fonction d'excès moyenne)
Soit X un rv avec df F et un point d'extrémité droit x_r . Pour un $u < x_r$ fixe,

$$F_u(x) = P(X - u \leq x \mid X > u), \quad x \geq 0.$$

est l'excès de la v.a.r X (de cdf F) au dessus du seuil u .

La fonction des excès à un seuil jouent un rôle fondamental dans de nombreux domaines.

Par exemple, F_u est connu comme l'excès de vie ou la durée de vie résiduelle de la cdf dans la théorie de la fiabilité et les statistiques médicales.

Dans un contexte d'assurance, F_u est généralement appelé l'excès de perte de la cdf.

Définition 2.8 (Distribution de Pareto généralisée (GPD)) Définir le cdf G_ξ par

$$G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \text{si } \xi \neq 0. \\ 1 - e^{-x} & \text{si } \xi = 0. \end{cases}$$

où

$$\begin{cases} x \geq 0 & \text{si } \xi \geq 0. \\ 0 \leq x \leq -1/\xi & \text{si } \xi < 0. \end{cases}$$

G_ξ est appelée la distribution de Pareto généralisée (GPD) standard.

On peut introduire la famille d'échelles de localisation correspondante $G_{\xi,u,\beta}$ en remplaçant l'argument x ci-dessus par $(x - u)/\beta$ pour $u \in \mathbb{R}, \beta > 0$. Le support doit être ajusté en conséquence. Nous désignons également $G_{\xi,u,\beta}$ par GPD.

On note donc,

$$G_{\xi,u,\beta}(x) = 1 - \left(1 + \xi \frac{x - u}{\beta}\right)^{-1/\xi},$$

où

$$\begin{cases} x \geq u & \text{si } \xi \geq 0, \\ 0 \leq x \leq -\beta/\xi + u & \text{si } \xi < 0. \end{cases}$$

Définition 2.9 (Fonction de distribution des excès, fonction Moyenne des excès)

Soit X une v.a.r de cdf F et de point final x_r . Pour $u < x_r$ fixe,

$$F_u(x) = P(X - u \leq x | X > u), \quad x \geq 0,$$

est le cdf des excès de la v.a.r X (de cdf F) au dessus du seuil u . Voir figure (2.2)

La fonction

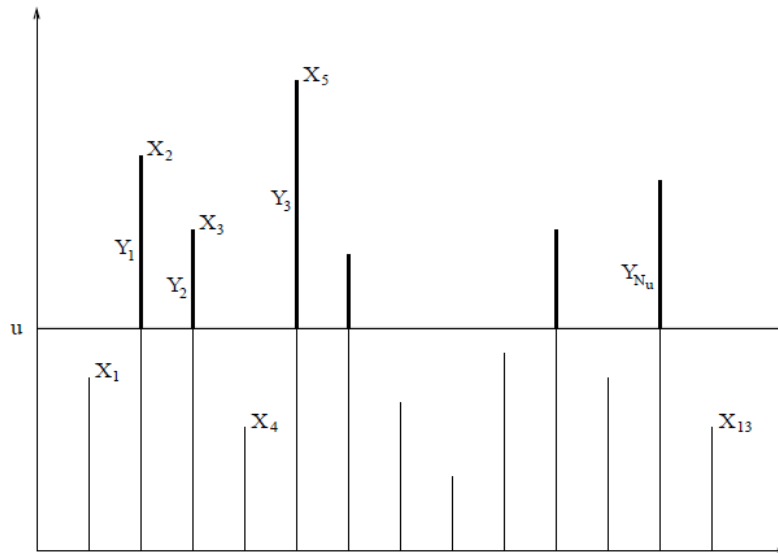


FIGURE 2.2 – Echantillon au dessus d'un seuil u

$$e(u) = E(X - u | X > u)$$

est appelée fonction moyenne des excès de X .

TABLE 2.5 – Fonctions moyennes des excès pour quelques distributions usuelles

Pareto	$\frac{\kappa+u}{\alpha-1}, \alpha > 1$
Burr	$\frac{u}{\alpha\tau-1}(1+o(1)), \alpha\tau > 1$
Loggamma	$\frac{u}{\alpha-1}(1+o(1)), \alpha > 1$
Lognormal	$\frac{\sigma^2 u}{\ln u - \mu}(1+o(1))$
Weibull	$\frac{u^{1-\xi}}{c\xi}(1+o(1))$
Exponential	λ^{-1}
Gamma	$\beta^{-1} \left(1 + \frac{\alpha-1}{\beta u} + o\left(\frac{1}{u}\right)\right)$

Théorème 2.7 (*Propriétés de la distribution GPD*)

1. Supposons X a une loi GPD de paramètres ξ et β . D'où $E(X) < \infty$ si et seulement si $\xi < 1$.

$$E\left(1 + \frac{\xi}{\beta}X\right)^{-m} = \frac{1}{1 + \xi m}, \quad r > -1/\xi,$$

$$E\left(\ln\left(1 + \frac{\xi}{\beta}X\right)\right)^k = \xi^k k!, \quad k \in \mathbb{N},$$

$$EX\left(\bar{G}_{\xi,\beta}(X)\right)^m = \frac{\beta}{(r+1-\xi)(r+1)}, \quad (m+1)/\xi > 0..$$

Si $\xi < 1/m$ avec $m \in \mathbb{N}$, alors

$$E(X^m) = \frac{\beta^m}{\xi^{m+1}} \frac{\Gamma(\xi^{-1} - m)}{\Gamma(1 + \xi^{-1})} m!$$

2. Pour tout $\xi \in \mathbb{R}, F \in \mathcal{D}(H_\xi)$ si et seulement si

$$\lim_{u \uparrow x_r, 0 < x < x_r - u} \sup |F_u(x) - G_{\xi,\beta(u)}(x)| = 0$$

pour une fonction positive β .

3. Supposons $x_i \in D(\xi, \beta), i = 1, 2$, alors

$$\frac{\bar{G}_{\xi,\beta}(x_1 + x_2)}{\bar{G}_{\xi,\beta}(x_1)} = \bar{G}_{\xi,\beta+\xi x_1}(x_2).$$

4. Si N est un processus de Poisson $\mathcal{P}(\lambda)$, indépendant de la suite i.i.d. (X_n) de distribution GPD de paramètres ξ et β . Si on note $M_N = \max(X_1, \dots, X_N)$. Alors

$$P(X_N \leq x) = \exp\left\{-\lambda \left(1 + \xi \frac{x}{\beta}\right)^{-1/\xi}\right\} = H_{\xi;\mu,\psi}(x)$$

où $\mu = \beta\xi^{-1}(\lambda^\xi - 1)$ and $\psi = \beta\lambda^\xi$.

5. Supposons X a une loi GPD de paramètres $\xi < 1$ et β . Alors pour $u < x_r$,

$$e(u) = E(X - u | X > u) = \frac{\beta + \xi u}{1 - \xi}, \quad \beta + u\xi > 0.$$

Pour la démonstration, voir (Balkema and De Haan (1974)) et (Embrechts et al. (2013)). Le théorème résume diverses propriétés qui sont essentielles pour le rôle spécial du GPD dans l'analyse statistique des extrêmes.

La propriété (2) ci-dessus suggère une GPD comme approximation appropriée de l'excès cdf F_u pour de grands u . Ce résultat remonte à Pickands (1975) et est souvent formulé comme suit. Pour une certaine fonction β à estimer à partir des données.

$$\bar{F}_u(x) = P(X - u > x | X > u) \approx \bar{G}_{\xi, \beta(u)}(x), \quad x > 0.$$

On peut aussi considérer pour $x > u$,

$$P(X > x | X > u) \approx \bar{G}_{u, \xi, \beta(u)}(x).$$

La classe des GPD est fermée par rapport aux changements du seuil.

En effet la formule de la propriété (3) montre que la probabilité conditionnelle, étant donné notre v.a.r dépasse x_1 , il dépasse également le seuil $x_1 + x_2$. La formule indique que cette probabilité est aussi du type Pareto généralisé.

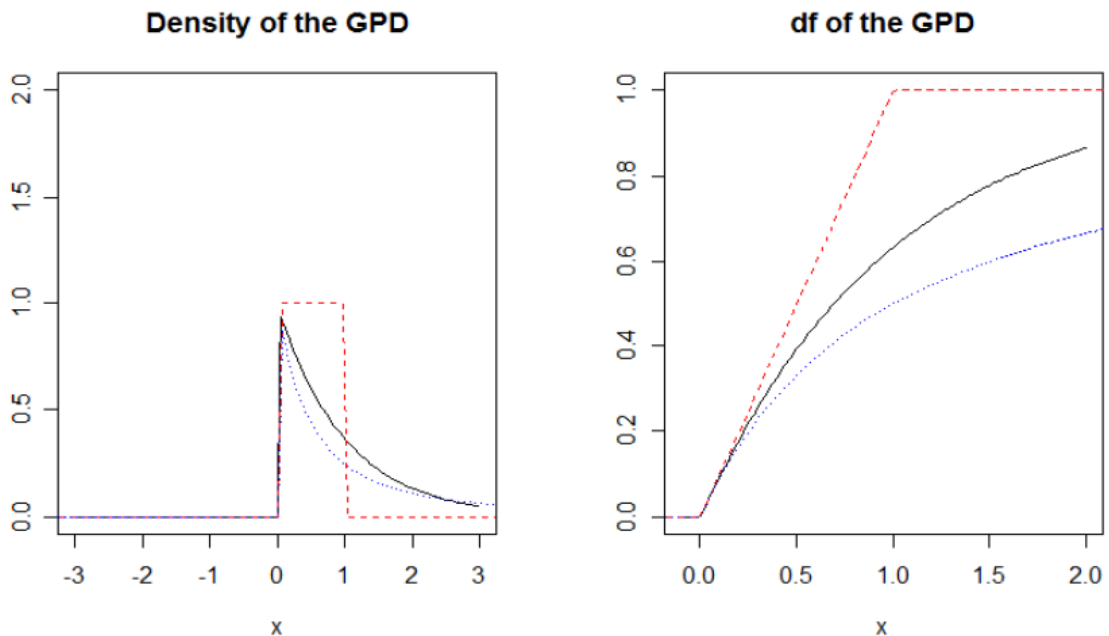


FIGURE 2.3 – Densité et distribution de Loi Pareto Généralisée (GPD)

Dans les sections précédentes, nous avons introduit une multitude de modèles probabilistes afin de décrire d'une manière mathématiquement valable les événements extrêmes dans le cas unidimensionnel. Cependant, le monde réel nous informe souvent sur ces événements par le biais de données statistiques.

Les sinistres d'assurance importants, les crues des rivières les fortes baisses, ou les fortes hausses des valeurs boursières sur une période donnée, les niveaux extrêmes d'indicateurs environnementaux tels que l'ozone ou le monoxyde de carbone, les valeurs de la vitesse du vent sur un site donné, la hauteur des vagues pendant une tempête, ou les valeurs de performance maximale et minimale d'un portefeuille. Tous ces exemples, et bien d'autres encore, ont en commun qu'ils concernent des questions sur les valeurs extrêmes d'un échantillon de données.

2.3 Méthodes statistiques pour des événements extrêmes

2.3.1 Estimation de l'indice de queue

La statistique des valeurs extrêmes repose sur le fait que, dans des conditions assez faibles, on peut considérer qu'une classe de fonctions s'adapte à la distribution de la plus grande observation d'un échantillon.

À partir de ce théorème limite, on peut voir que le comportement de la queue d'une fonction de distribution peut être complètement caractérisé par un seul paramètre à valeur réelle ξ , appelé indice de valeur extrême ou paramètre de forme. En fonction du signe de ce paramètre, le domaine d'attraction de l'extrême de la distribution des valeurs extrêmes peut être divisée en trois sous-classes, à savoir les classes de Fréchet ($\xi > 0$), Gumbel ($\xi = 0$) et Weibull ($\xi < 0$).

Nous nous concentrerons sur le cas où l'indice de valeur extrême est positif, ce qui correspond aux distributions à queue lourde ou de type Pareto.

L'analyse des distributions à queue lourde nécessite des méthodes d'estimation particulières en raison de leurs caractéristiques spécifiques :

Les deux estimateurs sans doute les plus populaires dans la littérature sont les estimateurs de [Hill \(1975\)](#) et l'estimateur de [Pickands \(1975\)](#).

Soit $X_{(1)}, \dots, X_{(n)}$ les statistiques d'ordre associées à l'échantillon X_1, \dots, X_n .

On considère les k valeurs les plus grandes (ou les plus petites), k dépend a priori de n , même si on ne le mentionnera pas dans la notation, l'idée est d'avoir $k \rightarrow \infty$ lorsque $n \rightarrow \infty$, mais sans prendre « trop » de valeurs de l'échantillon, ce qui conduit à imposer $\frac{k}{n} \rightarrow 0$.

Cela implique que se posera la question du choix optimal de k .

En effet, il est indispensable de calculer ces estimateurs sur les queues de distribution. Choisir un k trop élevé engendre le risque de prendre en compte des valeurs qui ne sont pas extrêmes, inversement, un sous-échantillon trop petit ne permet pas aux estimateurs d'atteindre leur niveau de stabilité.

2.3.2 L'estimateur de Pickands

Cet estimateur a été introduit en 1975 par James Pickands (voir [Pickands \(1975\)](#)). L'idée de base de cet estimateur consiste à trouver une condition équivalente à $F \in \mathcal{D}(H_\xi)$ qui permet de déduire le paramètre d'une manière simple. La clé de l'estimateur de Pickands et de ses différentes généralisations est le théorème (2.3) où il a été démontré pour $F \in \mathcal{D}(H_\xi)$.

$U(t) = F^{\leftarrow}(1 - 1/t)$ satisfait à

$$\lim_{t \rightarrow \infty} \frac{U(2t) - U(t)}{U(t) - U(t/2)} = 2^\xi. \quad (2.30)$$

Motivés par la discussion ci-dessus et en utilisant l'équation (2.30), nous définissons maintenant l'estimateur de Pickands

$$\widehat{\xi}_{(k)}^{(P)} = \frac{1}{\ln 2} \ln \frac{X_{(k)} - X_{(2k)}}{X_{(2k)} - X_{(4k)}}. \quad (2.31)$$

Cet estimateur s'avère être faiblement cohérent à condition que $k \rightarrow \infty, k/n \rightarrow 0$ (voir [Dekkers and de Haan \(1993\)](#)) :

$$\widehat{\xi}_{(k)}^{(P)} \xrightarrow{P} \xi, \quad n \rightarrow \infty.$$

Une analyse complète sur $\widehat{\xi}_{(k)}^{(P)}$ se trouve dans [Dekkers and de Haan \(1993\)](#) d'où est tiré le résultat suivant.

Théorème 2.8 *Supposons que $(X_n)_n$ soit une suite i.i.d. avec $df F \in \mathcal{D}(H_\xi), \xi \in \mathbb{R}$. Soit $\widehat{\xi}^{(P)} = \widehat{\xi}_{(k)}^{(P)}$ l'estimateur de Pickands (2.31).*

1. (Convergence faible) Si $k \rightarrow \infty, k/n \rightarrow 0$ for $n \rightarrow \infty$, alors

$$\widehat{\xi}^{(P)} \xrightarrow{P} \xi, \quad n \rightarrow \infty.$$

2. (Convergence forte) If $k/n \rightarrow 0, k/\ln \ln n \rightarrow \infty$ for $n \rightarrow \infty$, then

$$\widehat{\xi}^{(P)} \xrightarrow{p.s.} \xi, \quad n \rightarrow \infty.$$

3. (Normality asymptotique) Sous d'autres conditions sur k et F (voir [Dekkers and de Haan \(1993\)](#), p. 1799),

$$\sqrt{k}(\widehat{\xi} - \xi) \xrightarrow{L} N(0, \Gamma(\xi)), \quad n \rightarrow \infty,$$

où

$$\Gamma(\xi) = \frac{\xi^2 (2^{2\xi+1} + 1)}{(2(2^\xi - 1) \ln 2)^2}.$$

Ce théorème constitue le noyau de tous les résultats obtenus par Dekkers et de Haan sur lesquels sont basés les estimateurs de quantile, de la queue et la construction des intervalles de confiance asymptotiques. Le document cité contient également divers exemples simulés et réels afin de voir la théorie en action.

Par conséquent, on inclut le plus souvent dans l'analyse du graphique dit de Pickands, c'est à dire

$$\left\{ \left(k, \widehat{\xi}_{(k)}^{(P)} \right) : k = 1, \dots, n \right\},$$

afin de permettre un choix en fonction de k . Intuitivement on doit choisir $\widehat{\xi}_{(k)}^{(P)}$ dans la région de k où le graphe est à peu près horizontal.

2.3.3 Estimateur de Hill

Supposons que X_1, \dots, X_n soient *i.i.d.* avec cdf $F \in \mathcal{D}(\Phi_{1/\xi}), \xi > 0$, donc $\bar{F}(x) = x^{-1/\xi} L(x)$, $x > 0$, pour une fonction à variation lente L .

Les distributions avec de telles queues constituent des exemples de premier ordre pour la modélisation des phénomènes à queue lourde. Si par exemple $\xi > 1/2$, alors $E(X_1^2) = \infty$. Ce cas est souvent observé dans la modélisation des données d'assurance.

$$\hat{\xi}_{(k)}^{(H)} = \frac{1}{k} \sum_{i=1}^k \log(X_{(i)} - X_{(k)}) \quad (2.32)$$

On choisit k de sorte $k = k(n) \rightarrow \infty$ et $\frac{k}{n} \rightarrow 0$, comme dans le cas de l'estimateur de Pickands, une suite croissante de statistiques d'ordre supérieur est utilisée.

On peut montrer que $\lim_{k \rightarrow +\infty} \hat{\xi}_{(k)}^{(H)} = \xi$ et que l'estimateur de Hill est de plus asymptotiquement normal :

$$\sqrt{k} \frac{\hat{\xi}_{(k)}^{(H)} - \xi}{\hat{\xi}} \rightarrow \mathcal{N}(0,1)$$

Théorème 2.9 (Propriétés de l'estimation de Hill) Supposons que $(X_n)_n$ est strictement stationnaire avec une distribution marginale F satisfaisante pour certains $\xi > 0$ et $L \in \mathcal{R}_\alpha$,

$$\bar{F}(x) = P(X > x) = x^{-\alpha} L(x), \quad x > 0.$$

Soit $\widehat{\xi}^{(H)} = \widehat{\xi}_{(k)}^{(H)}$ l'estimateur de Hill (2.32).

1. (convergence faible) Supposons que l'une des conditions suivantes soit satisfaite :

- $(X_n)_n$ est i.i.d..
- $(X_n)_n$ est faiblement dépendant (de Haan and Ferreira (2006))
- $(X_n)_n$ est un processus linéaire (Resnick and Stărică (1997)).

Si $k \rightarrow \infty, k/n \rightarrow 0$ pour $n \rightarrow \infty$, alors

$$\widehat{\xi}^{(H)} \xrightarrow{P} \xi.$$

2. (Convergence forte) (Mason(1982)) Si $k/n \rightarrow 0, k/\ln \ln n \rightarrow \infty$ for $n \rightarrow \infty$ et (X_n) est une suite i.i.d., alors

$$\widehat{\xi}^{(H)} \xrightarrow{p.s} \xi.$$

3. (Normalité asymptotique) Si d'autres conditions sur k et F sont satisfaites et que (X_n) est une suite i.i.d., alors

$$\sqrt{k}(\widehat{\xi}^{(H)} - \xi) \xrightarrow{L} N(0, \xi^{-2}).$$

Comme dans le cas de l'estimateur de Pickands, une analyse basée sur l'estimateur de Hill est généralement résumée graphiquement. Le graphe de Hill

$$\left\{ \left(k, \widehat{\xi}_{(k)}^{(H)} \right) : k = 2, \dots, n \right\}.$$

s'avère déterminante pour trouver la valeur optimale de k .

La variance asymptotique de $\widehat{\xi}^{(H)}$ dépend du paramètre inconnu ξ de sorte que pour calculer les intervalles de confiance asymptotiques, un estimateur $\widehat{\xi}^{(H)}$ de ξ , est utilisé.

Exemple 2.8 (Choix de la valeur k)

Considérons le cas particulier

$$\bar{F}(x) = cx^{-\alpha} (1 + x^{-\beta})$$

Pour des constantes positives c, α et β . Nous pouvons choisir

$$a(x) = \beta x^{-\beta},$$

Posons $\rho = -\beta$ in (2.20). avec $U(t) = (ct)^{1/\alpha}(1 + o(1))$, nous obtenons

$$k \sim Cn^{(2\beta)/(2\beta+\alpha)}, \quad k \rightarrow \infty,$$

où C est une constante, dépendant de α, β, c et λ .

De plus, $\lambda = 0$ si et seulement si $C = 0$, donc $k = o(n^{(2\beta)/(2\beta+\alpha)})$.

2.3.4 Estimateur de Deckers Einmah-de Haan ($\xi \in \mathbb{R}$).

C'est un estimateur proposé par Dekkers et al. (1989).

L'un des inconvénients de l'estimateur de Hill est qu'il est essentiellement conçu pour $F \in \mathcal{D}(H_\xi), \xi > 0$.

Cette classe de modèles convient à de nombreuses applications dans le domaine de la finance et de l'assurance, comme nous l'avons déjà souligné précédemment.

Dans Dekkers et al. (1989), l'estimateur de Hill est étendu à la classe entière $H_\xi, \xi \in \mathbb{R}$.

Si $F \in \mathcal{D}(H_\xi), \xi < 0$, l'extrémité droite x_r de F est finie. Pour simplifier, nous supposons que $x_r > 0$. Dekkers et al. (1989) ont proposé la l'estimateur suivant :

$$\widehat{\xi} = 1 + H_n^{(1)} + \frac{1}{2} \left(\frac{(H_n^{(1)})^2}{H_n^{(2)}} - 1 \right)^{-1}, \quad (2.33)$$

où

$$H_n^{(1)} = \frac{1}{k} \sum_{j=1}^k (\ln X_{j,n} - \ln X_{k+1,n})$$

est l'inverse de l'estimateur de Hill et

$$H_n^{(2)} = \frac{1}{k} \sum_{j=1}^k (\ln X_{j,n} - \ln X_{k+1,n})^2.$$

Comme $H_n^{(1)}$ et $H_n^{(2)}$ peuvent être interprétés comme des moments empiriques, $\widehat{\xi}$ est également appelé estimateur de moment de ξ .

Cet estimateur est convergent et asymptotiquement gaussien :

$$\sqrt{k} \frac{\widehat{\xi}^{(k)} - \xi}{\sqrt{1 + \xi^2}} \xrightarrow{L} \mathcal{N}(0, 1)$$

2.3.5 Comparaison des différents estimateurs

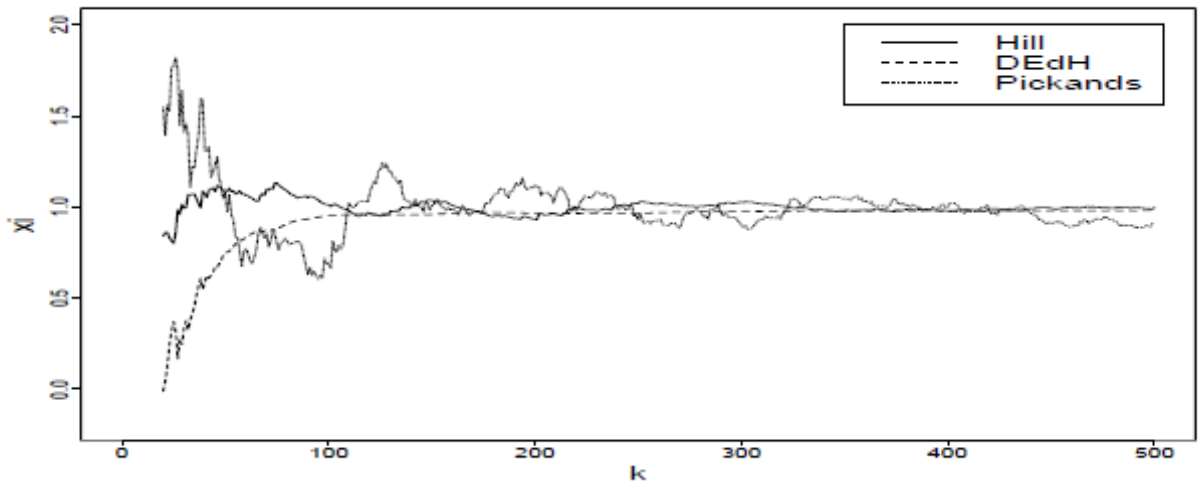


FIGURE 2.4 – Figure Graphiques de Pickands, Hill et DEkker pour 2000 données *i.i.d.* simulées avec cdf donné par $\bar{F}(x) = x^{-1}, x \geq 1$.

On remarque que l'estimateur de Pickands est moins efficace dans l'estimation de l'indice de queue. Par contre, on observe une efficacité de l'estimateur de Hill sur ceux de Pickands et de Deckers Einmahl de Haan.

2.3.6 Estimation des constantes de normalisation

Dans la section précédente, nous avons donné des estimateurs pour le paramètre de forme ξ étant donné des données *i.i.d.* de X_1, \dots, X_n de cdf $F \in \mathcal{D}(H_\xi)$. Rappelons que cette dernière condition est équivalente à

$$a_n^{-1} (X_{(n)} - b_n) \xrightarrow{L} H_\xi$$

pour des constantes de normalisation appropriées $a_n > 0$ et $b_n \in \mathbb{R}$. Nous savons également que cette relation est valable si et seulement si

$$n\bar{F}(a_n x + b_n) \rightarrow -\ln H_\xi(x), \quad n \rightarrow \infty, \quad x \in \mathbb{R}$$

Comme nous l'avons déjà vu dans la section précédente, les constantes de normalisation entrent en jeu dans l'estimation du quantile et de la queue. Nous examinons ci-dessous une méthode d'estimation des constantes de normalisation. Nous avons donné des formules analytiques reliant les suites normatives (a_n) et (b_n) à la queue \bar{F} . Par exemple, dans le cas de Gumbel $\xi = 0$ avec l'extrémité droite $x_r = \infty$.

$$a_n = \mathcal{K}(b_n), \quad b_n = F^{\leftarrow}(1 - n^{-1}),$$

où $\mathcal{K}(\cdot)$ représente la fonction auxiliaire qui peut être prise sous la forme

$$\mathcal{K}(x) = \int_x^\infty \frac{\bar{F}(y)}{\bar{F}(x)} dy.$$

L'estimateur de la queue de la distribution F a déjà été obtenu par Hill pour le modèle exact.

L'estimateur de la queue et l'estimateur du quantile sont donnés dans la proposition suivante.

Proposition 2.4 Soit $\hat{\xi}^{(H)}$ l'estimateur de Hill pour ξ , c'est à dire

$$\hat{\xi}^{(H)} = \frac{1}{k} \sum_{j=1}^k \ln X_{(j)} - \ln X_{(k)}.$$

Soit X_1, \dots, X_n un échantillon de $F \in \mathcal{D}(H_\xi)$, $\xi \geq 0$, et $k = k(n) \rightarrow \infty$ tel que $k/n \rightarrow 0$. Alors pour x suffisamment grand, un estimateur de queue pour $\bar{F}(x)$ devient

$$\hat{\bar{F}}(x) = \frac{k}{n} \left(\frac{x}{X_{(k+1)}} \right)^{-1/\hat{\xi}^{(H)}}.$$

Le quantile x_p défini par $F(x_p) = p \in (0, 1)$ peut être estimé par

$$\hat{x}_p = \left(\frac{n}{k} (1-p) \right)^{-\hat{\xi}^{(H)}} X_{(k+1)}.$$

2.4 Estimation des paramètres de la distribution GPD

La méthodologie présentée jusqu'à présent a été obtenue soit en supposant que les données proviennent d'une GEV; ou appartiennent à son domaine d'attraction maximal. Nous avons basé l'estimation statistique des paramètres pertinents sur le maximum de vraisemblance.

Supposons que X, X_1, \dots, X_n sont *i.i.d.* avec $\text{df } F \in \mathcal{D}(H_\xi)$ pour un certain $\xi \in \mathbb{R}$. Premièrement, on choisit un seuil élevé u et on désigne par

F_Y la distribution des excès au-dessus du seuil u , $F_Y(y) = P(X - u \leq y | X > u)$, pour $0 \leq y < x_r - u$, avec $x_r = \sup\{x \in \mathbb{R}, F(x) < 1\}$ (point final supérieur de F) généralement $x_r = +\infty$.

et si $\beta = \beta(u) = u\xi$, alors $\bar{F}_Y(y)$ est une distribution GPD perturbée, où le cdf de la distribution de Pareto généralisée (GPD) est de la forme

$$G_{\beta,\xi}(y) = \begin{cases} 1 - (1 + \frac{\xi}{\beta}y)^{-\frac{1}{\xi}} & \xi \neq 0 \\ 1 - e^{-\frac{y}{\beta}} & \xi = 0 \end{cases}, y \in \begin{cases} [0, +\infty) & \xi \geq 0 \\ [0, \frac{\beta}{\xi}) & \xi < 0 \end{cases} \quad (2.34)$$

La méthode POT est basée sur le résultat [Pickands \(1975\)](#) qui dit que la distribution des excès au-dessus d'un seuil fixe est approximée par la distribution de Pareto généralisée (GPD). Cela signifie que pour de grandes valeurs de u_n , nous avons

$$\lim_{u_n \rightarrow x_F} \sup_{0 \leq x < x_T - u_n} |F_Y(x) - G_{\xi,\beta}(x)| = 0, \quad (2.35)$$

Voir aussi le théorème (3.4.13) dans [Kluppelberg and Stadtmüller \(1998\)](#).

Soit X_1, \dots, X_n un échantillon aléatoire de taille n . La queue de la distribution F supposée commencer à un certain niveau u supposé suffisamment élevé, et le dépassement $Y_i = X_i - u_n$, pour tous les i tels que $X_i > u_n$ sont approximativement un échantillon aléatoire de loi GPD. Il est clair que F_Y est aussi régulièrement variable à l'infini avec le même indice $-\frac{1}{\xi} < 0$.

2.5 Détection du seuil

L'estimation de F_Y basée sur la valeur du seuil des pics u , nécessite donc le choix d'un seuil à partir duquel commence la queue de la distribution sous-jacente. Si le seuil choisi est trop bas, l'approximation GPD peut ne pas tenir et un biais peut apparaître. Si le seuil est choisi trop haut, la réduction de la taille de l'échantillon augmente la variance des estimations des paramètres.

La tâche consiste à trouver le seuil le plus bas de telle sorte que la GPD s'adapte correctement à l'échantillon de dépassements de ce seuil. Le problème du choix du seuil u est toujours d'un grand intérêt théorique et pratique. Il est souhaitable de disposer d'une procédure de sélection de seuil automatisée et intuitive à utiliser avec l'analyse POT.

La méthode la plus simple est la sélection du seuil a priori, ou fixe, basée sur l'expertise du sujet traité. Diverses règles ont été suggérées, par exemple, la sélection de 10% des valeurs maximales de l'échantillon, voir par exemple [DuMouchel \(1983\)](#), ou de 5% des valeurs maximales de l'échantillon, voir [Kelly and Jiang \(2014\)](#), ou de la racine carrée de la taille de l'échantillon, voir par exemple [Ferreira et al. \(2003\)](#), [Bader et al. \(2018\)](#) et [Silva Lomba and Fraga Alves \(2020\)](#).

De nombreuses méthodes de sélection des seuils sont disponibles dans la littérature,

voir Dekkers et al. (1989), Scarrott and MacDonald (2012), Ivette Gomes et al. (2008), Guillou and Hall (2001), Matthys and Beirlant (2000), et Caeiro and Gomes (2015) pour des revues récentes.

2.5.1 Méthodes graphiques

Le Peak Over Threshold (POT) est une technique qui a été développée par des chercheurs, et elle est fréquemment utilisée dans la structure des valeurs extrêmes. L'approche POT consiste à ajuster un modèle paramétrique pour que ses excès au-dessus d'un seuil u soient assez élevés (de Zea Bermudez and Kotz, 2010) par la loi GPD. En d'autres termes, cette technique permet d'évaluer si le choix du seuil u est adéquat pour être représenté par un modèle asymptotique (e.g. exponentiel ou GPD). Afin de comprendre cette notion, la section suivante fera une brève introduction de quelques méthodes utilisant la technique POT.

2.5.2 Mean excess-plot (ME-plot)

L'une des représentations graphiques de l'analyse POT est le «Mean Excess Plot» (MEP). Il est également connu sous le nom «Mean Residual Life Plot (MRL-plot)» et il représente la variation de la moyenne des excès par rapport à un seuil u . Pour chaque seuil u , on calcule la moyenne et la variance des excès au-dessus de ce seuil.

Le graphe de la fonction moyenne excès est le graphe des points $\{(u, e_n(u)), X_{(1)} < u < X_{(n)}\}$ où $e_n(u)$ est la moyenne des excès au delà du seuil u , définie par :

$$e_n(u) = E(x - u | x > u) = \frac{\beta_u}{1 - \xi} + \frac{\xi}{1 - \xi} u, \quad \xi < 1$$

où β_u est le paramètre d'échelle correspondant au seuil u .

Un estimateur empirique de cette fonction est donné par :

$$\hat{e}_n(u) = \frac{1}{N_u} \sum_{i=1}^{N_u} (x_i - u), \quad x_i > u, \quad 0 < u < +\infty \quad (2.36)$$

où N_u est le nombre de dépassements par rapport à u .

Supposons données les observations (X_1, X_2, \dots, X_n) , on trace graphiquement $\hat{e}(u)$ en fonction de u et on choisit le plus petit u de manière à ce que $\hat{e}(u)$ soit approximativement linéaire pour tout $x > u$.

La fonction moyenne des excès empiriques sous la transformation affine s'écrit pour $x < x_r$, comme suit :

$$\hat{e}(u) = \frac{\hat{\beta}_u + \hat{\xi}u}{1 - \hat{\xi}}, \hat{\beta}_u + \hat{\xi}u > 0 \quad (2.37)$$

Trois cas peuvent alors se présenter :

1. Si à un certain seuil, le graphe de la moyenne des excès (mean excess plot) est marqué par une pente positive. Alors les données suivent la distribution GPD avec un paramètre ξ positif (c'est à dire une distribution de pareto).
2. Si le graphe de la moyenne des excès (mean excess plot) est horizontale. Alors les données suivent une distribution exponentielle..
3. Si le graphe de la moyenne des excès (mean excess plot) est marquée par une pente négative. Alors les données suivent une distribution à queue légère.

2.5.3 Illustrations graphiques

Afin de valider la performance de la méthode MRLP, nous avons une étude de simulation pour deux variables aléatoires de lois GEV et GPD. Nous avons constaté que cette méthode graphique détecte bien le seuil. **mrl-plot pour GPD** :

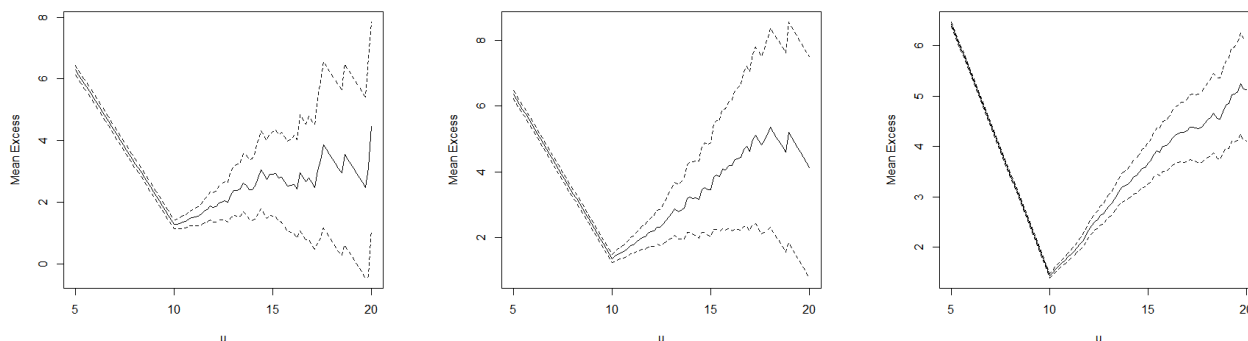


FIGURE 2.5 – Mean residual life plot pour $GPD(u = 10, \xi = 0.3, \beta = 1)$ et pour échantillon de taille (500,1000,10000)

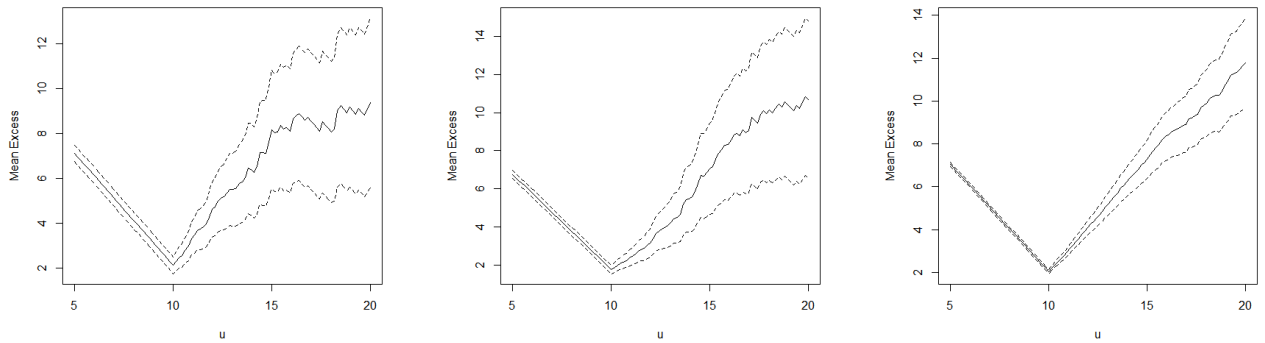


FIGURE 2.6 – Mean residual life plot pour $GPD(u = 10, \xi = 0.5, \beta = 1)$ et pour échantillon de taille (500,1000,10000)

2.5.4 Méthode de bootstrap double

La technique du ré-échantillonnage (Bootstrap) à été introduit par [Efron and Le-Page \(1992\)](#). C'est une méthode de réplcation des échantillons la mieux fondée théoriquement. Elle consiste à créer, à partir d'un échantillon de base, un grand nombre d'échantillon par tirage aléatoire avec remise.

Les méthodes bootstrap peuvent être utilisées pour une analyse entièrement paramétrique, semi-paramétrique et complètement non paramétrique. [Hall \(1990\)](#) et [Daniellson et al. \(2001\)](#) présentent la technique bootstrap et proposent la méthode bootstrap double.

Soit

$$Q^{(i)}(n, k) = \frac{1}{k} \sum_{j=1}^k (\log(X_{(n-j+1)}) - \log(X_{(n-k)}))^i. \quad (2.38)$$

La méthode optimise k pour l'erreur quadratique moyenne de.

$$Q(n, k) = Q^{(2)}(n, k) - 2(Q^{(1)}(n, k))^2 \quad (2.39)$$

on peut voir que $\hat{\xi} = Q^{(1)}(n, k)$ est l'estimateur de Hill La méthode optimise \hat{k}^* ou $u = X_{(n-k, n)}$ par les étapes suivantes :

1. Sélectionné uniformément $\epsilon \in (0, 1/2)$ et l'ensemble $n_1 = \lfloor n^{1-\epsilon} \rfloor$. Cela assure le Consistance si $n \rightarrow +\infty$.
Évaluation $E[Q(n_1, r)^2 | X_1, \dots, X_n]$ en ré-échantillonnant n_1 taille des échantillons bootstrap de la distribution empirique et le minimiser par rapport à r

$$\hat{k}_1 = \arg \min_r (E[Q(n_1, r)^2 | X_1, \dots, X_n])$$

2. L'ensemble $n_2 = \lfloor \frac{n_1^2}{n} \rfloor$ et minimiser $E[Q(n_2, r)^2 | X_1, \dots, X_n]$ par rapport à r de la même manière qu'à la première étape.

$$\hat{k}_2 = \underset{r}{\operatorname{arg\,min}}(E[Q(n_2, r)^2 | X_1, \dots, X_n])$$

- 3.

$$\hat{k} = \frac{\hat{k}_1^2}{\hat{k}_1} \left[\frac{\log(\hat{k}_1)^2}{(2\log(n_1) - \log(\hat{k}_1))^2} \right]^{\frac{\log(n_1) - \log(\hat{k}_1)}{\log(\hat{k}_1)}}$$

2.5.5 Metrique de Kolmogorov-Smirnov

Cette approche vise à minimiser la distance entre l'indice de la fonction de distribution empirique et la distribution Pareto ajustée avec le paramètre estimé de l'indice de queue. [Danielsson et al. \(2016\)](#) ont proposé la distance pour les quantiles.

Supposons que

$$P(X \geq x) = \bar{F}(x) = \mathbb{A}x^{-\frac{1}{\xi}} + o(x^{-\frac{1}{\xi}})$$

La fonction quantile peut être approximée par

$$q = \left[\frac{\bar{F}(x)}{\mathbb{A}} \right]^{-\xi}$$

La probabilité $\bar{F}(x)$ peut être remplacé par $\frac{j}{n}$, et ξ est estimé par l'estimateur de Hill pour certains k et \mathbb{A} peut être estimée par $\frac{k}{n} (X_{(n-k)})^{\frac{1}{\xi}}$. Le quantile est ainsi estimé par

$$q(j, k) = \left[\frac{k}{j} (X_{(n-k)})^{\frac{1}{\xi}} \right]^{\xi}$$

L'optimal \hat{k}^* pour l'estimateur de Hill qui réduit au minimum la distance entre le quantile empirique et le quantile estimé.

$$\hat{k}^* = \underset{k}{\operatorname{arg\,min}} \left(\sup_{j \in \{1, \dots, S\}} |X_{(n-j)} - q(j, k)| \right)$$

où $S > k$.

Alors le seuil optimal est $u^* = X_{(n-\hat{k}^*)}$.

2.5.6 Estimateur de Hill (Stabilité de graphe)

Encore appelée « stable scale and shape parameters », cette méthode permet de déterminer un seuil requis en ajustant les données à une distribution de GPD en utilisant un seuil différent. La stabilité des paramètres (forme et échelle) peut alors être contrôlée et localisée.

Cette technique est implémentée à l'aide du logiciel *R* avec des packages spécifiques. Ces packages disposent des codes qui permettent de choisir un seuil adéquat en examinant simplement la stabilité des paramètres de forme et d'échelle ξ et β respectivement.

Ce graphe établit un lien direct entre les valeurs des paramètres estimés (ξ et β) et les seuils potentiels u^* . Les paramètres estimés au-dessus des seuils sont ceux pour lesquels le modèle GPD devient valable.

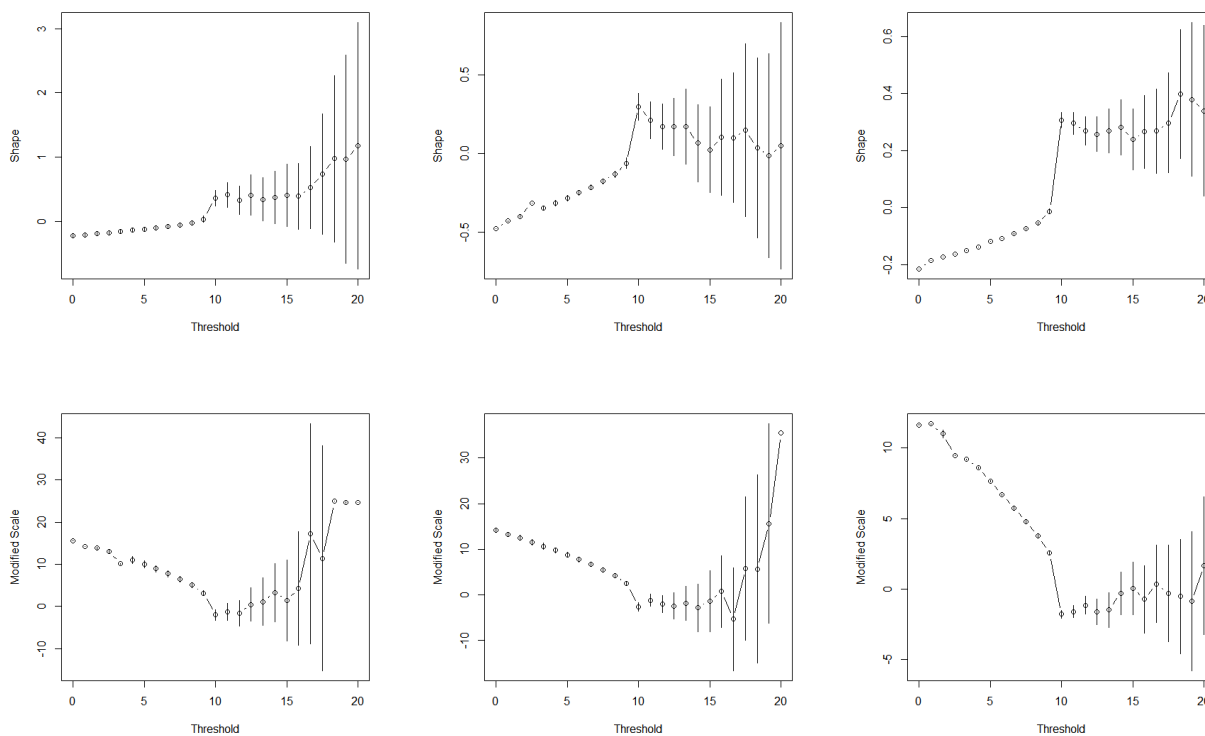


FIGURE 2.7 – Stabilité des seuils (Threshold stability plots) pour la forme GPD et un paramètre d'échelle modifié avec ($\xi = 0.3$) et de taille ($n = 500, 1000, 10000$)

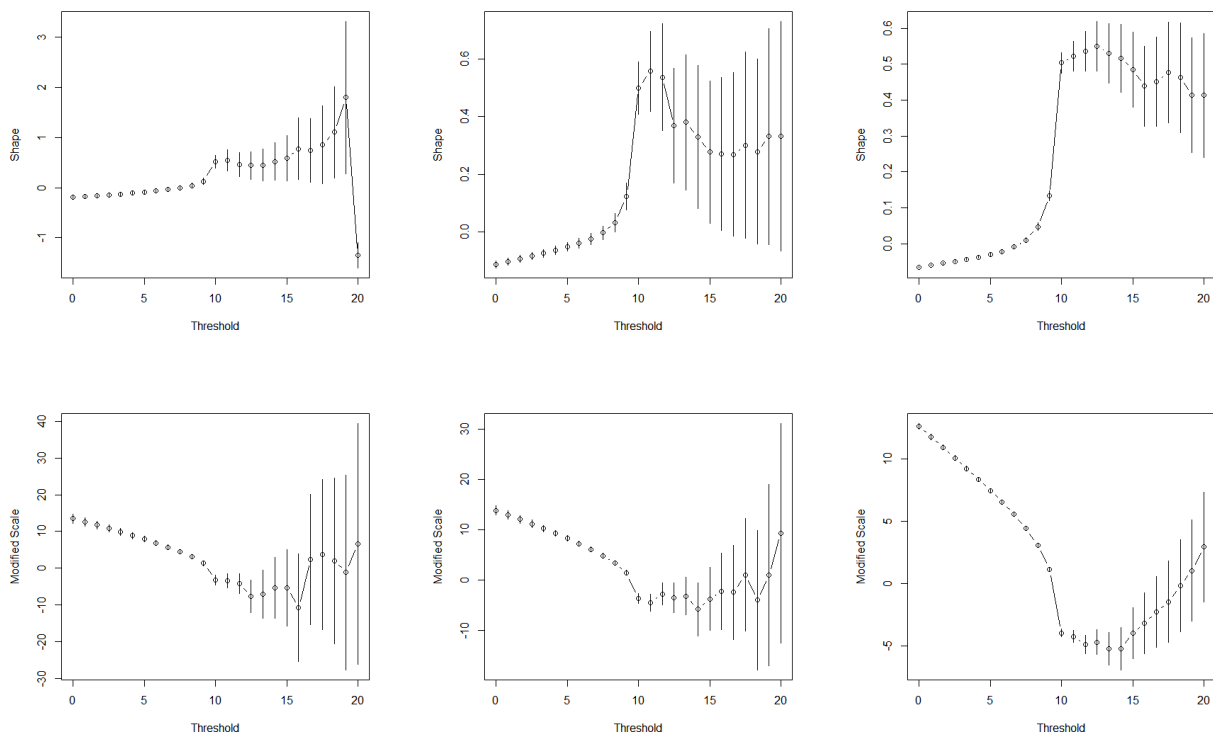


FIGURE 2.8 – Stabilité des seuils (Threshold stability plots) pour la forme GPD et un paramètre d'échelle modifié avec ($\xi = 0.5$) et de taille $n = (500, 1000, 10000)$

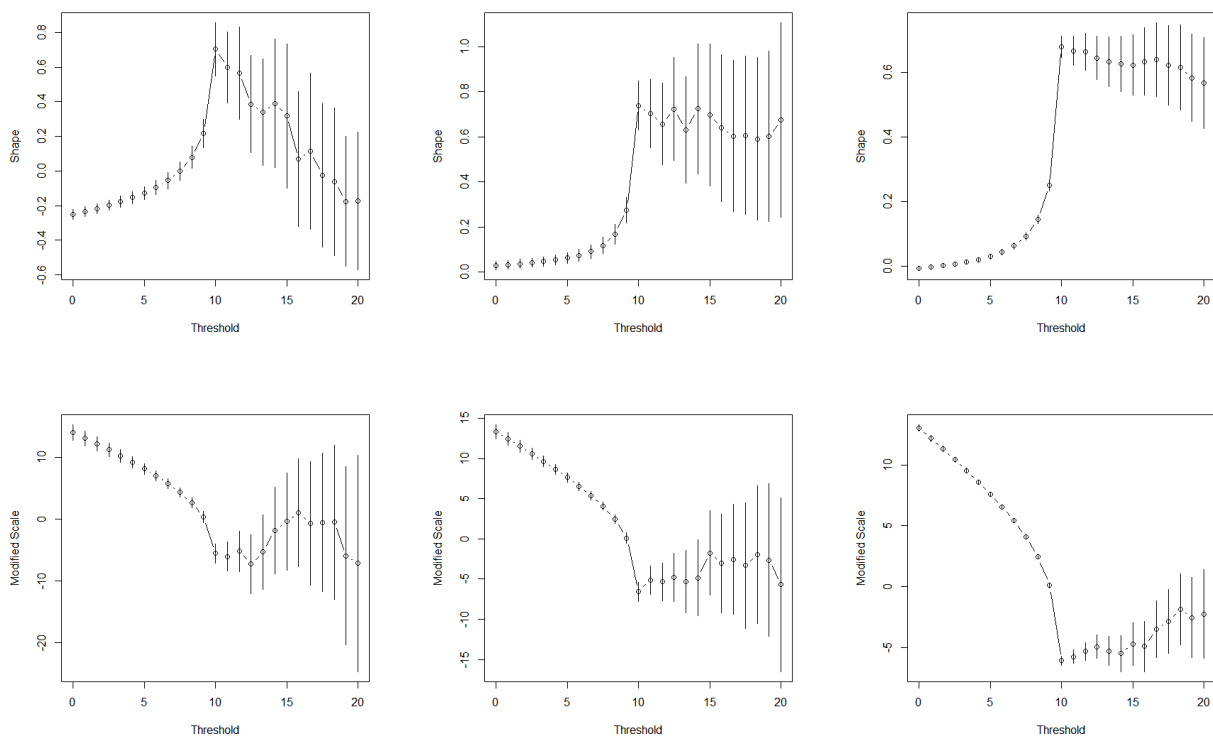


FIGURE 2.9 – Stabilité des seuils (Threshold stability plots) pour la forme GPD et un paramètre d'échelle modifié avec ($\xi = 0.7$) et de taille $n = (500, 1000, 10000)$

Les résultats de la simulation pour la stabilité linéaire en combinant ces deux représentations nous permet de prendre un seuil égale 10 pour les échantillons simulés.

2.5.7 Hill-plot

L'estimateur de Hill est donné par la forme empirique suivante :

$$\hat{\xi} = \frac{1}{k} \sum_{i=1}^k \log(X_{(n+i-1)} - X_{(n-k)}) \quad (2.40)$$

avec k , l'ordre statistique le plus élevé (le nombre des excès) et $\alpha = \frac{1}{\xi}$ est l'indice de la queue de distribution.

Cet estimateur intervient dans la construction du graphe appelé Hill-plot qui Représente ξ en fonction de la statistique d'ordre $(X_{(i)})_{0 \leq i \leq n}$.

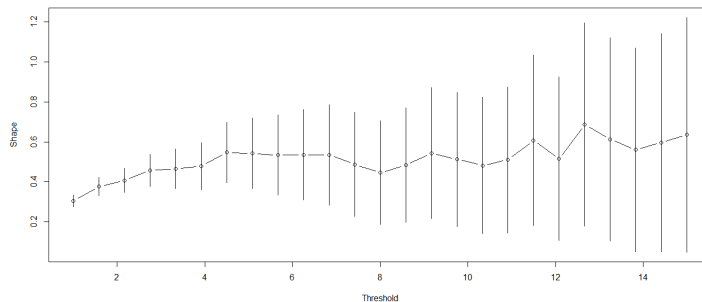


FIGURE 2.10 – Hill-plot

Le graphe Hill-plot, nous permet d'avoir des estimations du paramètre en fonction de l'ordre statistique le plus élevé (nombre des excès), nous choisissons ainsi l'indice le plus stable. Pour cette méthode, il faut calculer un estimateur de $1/\xi$ (valable pour $\xi > 0$). Lorsque nous utilisons tout l'échantillon, sa valeur est biaisée car seuls les dépassements au dessus d'un seuil élevé suivent une loi GPD. L'idée est de calculer cet estimateur avec un grand échantillon et d'enlever au fur et à mesure les observations les plus petites, jusqu'à ce que l'estimateur se stabilise. Le point où il se stabilise est notre seuil (Voir 2.10).

Le Hill-plot nous permet de choisir un seuil élevé pour la construction d'un modèle (GPD).

Le Hill-plot est donc un outil à double utilité :

- L'estimation de l'indice de la queue de la distribution ξ .
- L'estimation du seuil u .

Nous avons tracé le Hill plot en fonction du seuil u et du nombre d'excès k , en mettant en évidence l'intervalle de confiance à 95% autour de cet estimateur. Nous remarquons une zone de stabilité entre 135 et 190 excès. Au delà de 60 excès, l'estimateur n'est plus

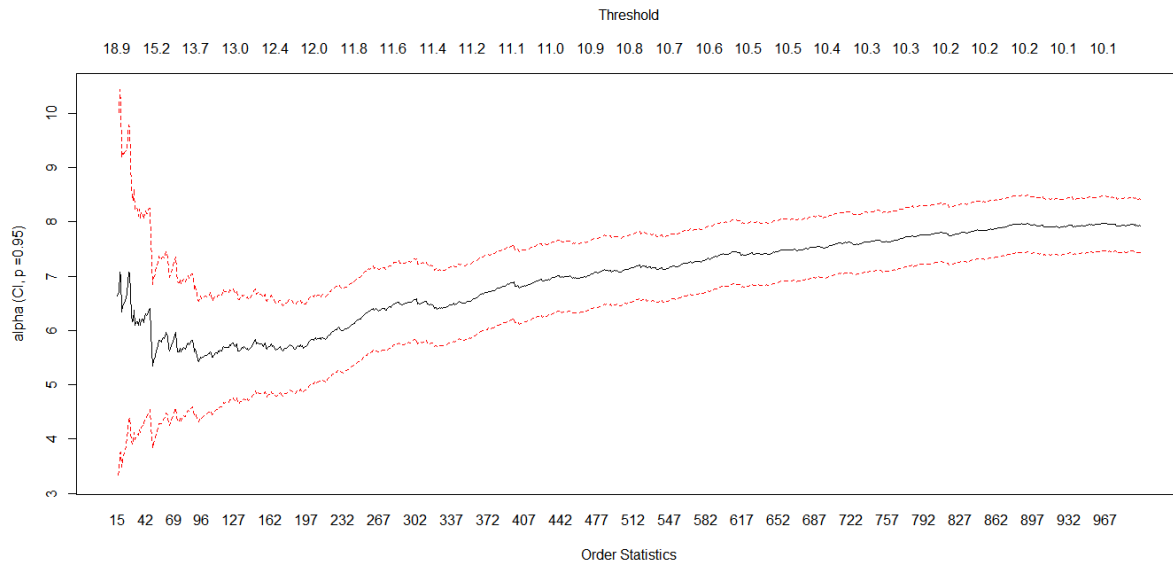


FIGURE 2.11 – Estimateur de Hill en fonction du seuil et du nombre d’excès considérés

du tout stable. Nous considérerons donc que l’adéquation à une GPD débute au niveau du 190ème excès, soit un seuil à 12.

Nous remarquons une zone de stabilité entre 130 et 185 excès. Au delà de 185 excès, l’estimateur n’est plus du tout stable. Nous considérerons donc que l’adéquation à une GEV débute au niveau du 185ème excès, soit un seuil à 11.70.

2.5.8 Illustrations numériques

Les méthodes numériques à savoir la distance de Kolmogorov et double bootstap sont étudiées par simulation pour la loi GPD pour différents tailles et paramètres. Les résultats sont données dans les tableaux suivants :

a) Métrique de Kolmogorov

TABLE 2.6 – K.S Metric pour $u=100$

	n=500	n=1000	n=5000	n=10000
$\xi = 0.3$	100.3794	103.6869	100.8472	102.0041
$\xi = 0.5$	100.9990	100.3415	100.3507	100.1178
$\xi = 0.7$	101.3970	100.7246	100.4318	100.8451

TABLE 2.7 – K.S Metric pour $u=10$

	n=500	n=1000	n=5000
$\xi = 0.3$	12.6704	11.3685	11.9999
$\xi = 0.5$	8.7111	9.6403	9.6305
$\xi = 0.7$	9.5897	9.5329	9.5444

b) Double Bootstrap

TABLE 2.8 – Double bootstrap pour $u=100$

	n=500	n=1000	n=5000
$\xi = 0.3$	100.0272	100.0372	100.0051
$\xi = 0.5$	100.3004	100.0733	100.0001
$\xi = 0.7$	100.0051	100.0001	100.0003

TABLE 2.9 – Double bootstrap pour $u=10$

	n=500	n=1000	n=5000
$\xi = 0.3$	10.0428	10.0118	10.0016
$\xi = 0.5$	10.1204	10.0110	10.0017
$\xi = 0.7$	10.4842	10.0766	10.0054

L'étude de simulation nous a montré que les méthodes estiment bien le seuil, mais la méthode de double bootstrap nécessite plus de temps d'exécution.

2.5.9 Méthode du maximum de vraisemblance

Une fois le seuil optimal est choisi, on construit une nouvelle série d'observations au dessus de ce seuil, et la distribution de ces données suit approximativement une distribution généralisée de Pareto.

La densité de la distribution GPD s'écrit [Embrechts et al. \(2013\)](#)

$$g_{\xi, \beta}(x) = \begin{cases} \beta^{1/\xi} (\beta + \xi x)^{-\frac{1}{\xi}-1} & \text{si } \xi \neq 0 \\ \beta^{-1} \exp(-\frac{x}{\beta}) & \text{si } \xi = 0 \end{cases} \quad (2.41)$$

La fonction de vraisemblance est donnée par :

$$l(\xi, \beta, x_1, \dots, x_n) = \prod_{i=1}^n G_{\xi, \beta}(x_i)$$

Le log de la fonction de vraisemblance, donne

$$\begin{aligned} \log l(\xi, \beta, x_1, \dots, x_n) &= \sum_{i=1}^n G_{\xi, \beta}(x_i) \\ &= -n \log \beta - \left(1 - \frac{1}{\xi}\right) \sum_{i=1}^n \log \left(1 + \frac{\xi}{\beta} x_i\right). \end{aligned} \quad (2.42)$$

Les équations de vraisemblance conditionnelle résultantes peuvent être résolues numériquement. Les équations deviennent simple par une reparamétrisation $(\xi, \beta) \rightarrow$

(ξ, γ) , où $\gamma = -\xi/\beta$. Ceci conduit à la solution

$$\widehat{\xi} = \widehat{\xi}(\gamma) = n^{-1} \sum_{i=1}^n \ln(1 - \gamma X_i),$$

où γ satisfait à

$$h(\tau) = \frac{1}{\gamma} + \frac{1}{n} \left(\frac{1}{\widehat{\xi}(\gamma)} + 1 \right) \sum_{i=1}^n \frac{X_i}{1 - \gamma X_i} = 0$$

Cette dernière équation se résout numériquement de manière itérative pour autant que l'on dispose d'une valeur initiale γ_0 pas trop éloignée de γ .

En pratique cette valeur initiale pourra être obtenue par la méthode des moments ou par la méthode des quantiles.

2.5.10 Méthode des moments pondérés

Définition 2.10 On appelle moment pondéré d'ordre r , le moment défini par :

$$u_r = \mathbb{E} \left[Z(\overline{H}_{\xi, \beta}(Z))^r \right], \quad r \in \mathbb{N}$$

Prenons par exemple Z suit la loi de Pareto généralisée de paramétré (ξ, β) .

Le moment pondéré u_r est égal à :

$$u_r = \frac{\beta}{(r+1)(r+1-\xi)}, \quad r = 0, 1$$

En résolvant l'équation précédente pour $r = 1$ et $r = 2$, on obtient :

$$\hat{\beta} = \frac{2u_0u_1}{u_0 - 2u_1} \text{ et } \hat{\xi} = \frac{u_0}{u_0 - 2u_1}$$

On remplace u_0 et u_1 par leurs estimateurs des moments empiriques pondérés définis par :

$$M_r = \frac{1}{n} \sum_{j=1}^n \left(\prod_{l=1}^r \frac{n-j-l+1}{n-l} \right) X_{(j)} = \frac{1}{n} \sum_{j=1}^n \left(1 - \frac{j}{n+1} \right) X_{(j)}$$

où $X_{j,n}$ est la j ème statistique d'ordre.

On trouve un estimateur de ξ par la méthode des moments pondérés

$$\hat{\xi} = \frac{\frac{1}{n} \sum_{j=1}^n (4 \frac{j}{n+1} - 3) X_{(j)}}{\frac{1}{n} \sum_{j=1}^n (2 \frac{j}{n+1} - 1) X_{(j)}} \quad (2.43)$$

CHAPITRE 3

QUELQUES RÉSULTATS SUR LA PROBABILITÉ DE RUINE POUR LES PROCESSUS DE RISQUE CLASSIQUES

La théorie du risque en général et les probabilités de ruine en particulier sont traditionnellement considérées comme une partie des mathématiques de l'assurance et elles ont constitué un domaine de recherche actif depuis l'époque de Lundberg [Lundberg \(1903\)](#) jusqu'à aujourd'hui. Ce domaine a atteint ces dernières années une maturité mathématique considérable, ce qui a notamment permis d'éliminer l'une des critiques habituelles de ce domaine. Bien que dans la pratique, on utilise généralement des mesures de risque plus simples (et plus grossières) telles que la Valeur-à-Risque (Value-at-risk). Il est largement admis que le raisonnement préconisé par la théorie des ruines reste important pour la gestion moderne du risque. Par exemple, l'évaluation conformes au marché, le rôle de l'effet de diversification temporelle des portefeuilles, qui est l'un des éléments centraux de la théorie de ruine. En outre, la théorie de ruine a des liens méthodologiques fructueux et des applications à d'autres domaines des probabilités appliquées, comme la théorie des files d'attente et la finance mathématique (évaluation des options à barrière, des produits de crédit, etc.). La problématique dans le calcul de la probabilité de ruine d'une compagnie, est que ses réserves financières passent sous la frontière du zéro. Ici, une société ruinée est une compagnie qui ne peut satisfaire à ses engagements vis-à-vis de ses clients.

Dans ce chapitre, nous étudions la probabilité de ruine asymptotique à horizon infini dans le cas où la distribution des remboursements est à queue lourde, et le processus des arrivées des sinistres pour réclamer ses droits de remboursements est un processus de renouvellement et en particulier Poissonien stationnaire.

3.1 Modèle de risque

Pour plusieurs types de risques d'assurance, des distributions à queue lourde sont utilisées pour modéliser les montants des sinistres individuels. Les distributions de Pareto, ou plus généralement les distributions de variation régulière, sont souvent préférées aux distributions log-normales pour s'adapter aux données empiriques. Une façon naturelle d'étudier ces risques pourrait être utilisé pour le modèle classique de risque de renouvellement composé avec une distribution du montant de remboursement des sinistres à queue lourde.

Dans le modèle de risque classique de [Lundberg \(1903\)](#), le processus de risque classique $(R_t)_{t \geq 0}$ est défini comme suit : pour $t \geq 0$

$$R_t = u + pt - S_t, \tag{3.1}$$

où u est le montant de la réserve initiale, $p > 0$ est le taux de revenu des primes. Le montant cumulé des sinistres jusqu'au temps t est décrit par le processus de renouvellement composé

$$S_t = \sum_{i=1}^{N_t} X_i,$$

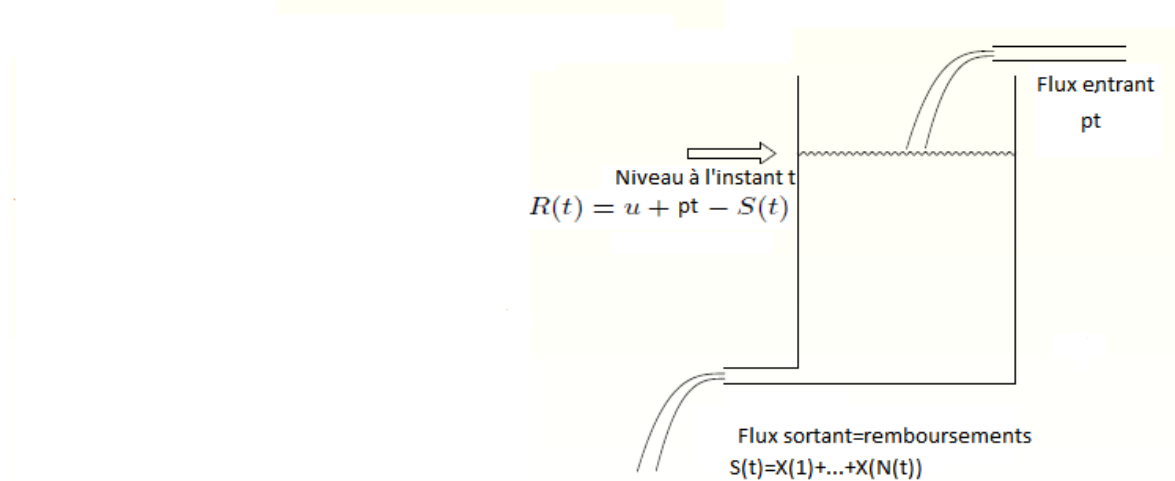


FIGURE 3.1 – Richesses d'une compagnie

Ce modèle est dit de [Lundberg \(1903\)](#) et [Cramér \(1930\)](#), qui est le premier modèle pour modéliser la richesse de l'assurance, ce modèle est dit de Poisson composé ou aussi dit de Cramer-Lundberg, où $X_i, i = 1, 2, \dots$ sont les montants des sinistres, qui sont des variables aléatoires indépendantes non négatives, identiquement distribuées de moyenne μ finie avec $S_t = 0$ si $N_t = 0$.

Le nombre de demandes N_t pour $t \geq 0$ est modélisé par un processus de renouvellement $(N_t)_{t \geq 0}$ défini à partir des temps d'inter-occurrence $(T_k)_{k \geq 1}$ par $N_t = \sum_{k \geq 1} 1_{\{T_k \leq t\}}$.

On suppose que les montants des demandes et les temps d'inter-occurrence sont mutuellement indépendants. On définit le processus de surplus $\{S_t\}_{t \geq 0}$ comme

$$S_t = u - R_t = \sum_{i=1}^{N_t} X_i - pt. \quad (3.2)$$

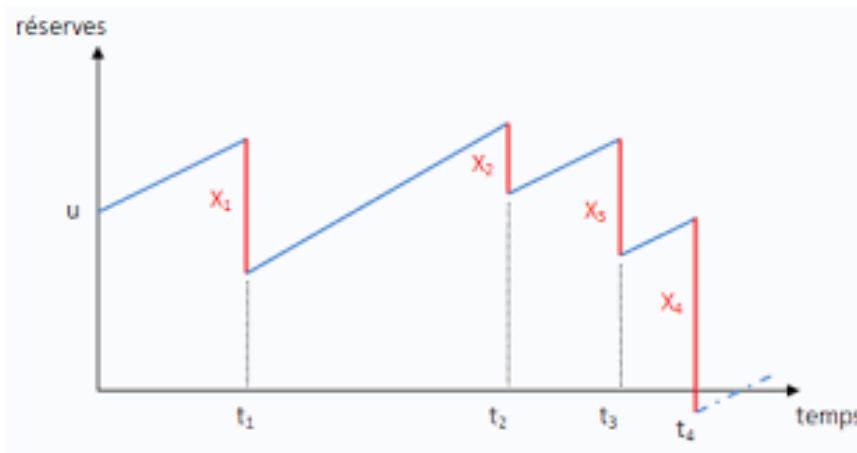


FIGURE 3.2 – Processus des réserves et ruine

3.1.1 Probabilité de ruine

Pour tout modèle de risque d'un surplus financier, la première quantité d'intérêt est la probabilité de ruine.

Définition 3.1 (*Le temps de ruine*)

$$\tau(u) = \inf \{t \geq 0 : R_t < 0\} = \inf \{t \geq 0 : S_t > u\}, \quad (3.3)$$

est le premier instant où le processus de réserve devient négatif ou de manière équivalente le processus de surplus excède le niveau u .

Définition 3.2 La probabilité de ruine à horizon de temps infini notée Φ , est défini par :

$$\Phi(u) = P \left\{ \inf_{t \geq 0} R_t < 0 \mid R_0 = u \right\} \quad (3.4)$$

La probabilité de ruine en temps infini, i.e. la probabilité que le capitale de la compagnie d'assurance descend en dessous de zéro peut alors s'écrire aussi comme

$$\Phi(u) = P \{ \tau(u) < \infty \} = P \{ L > u \}. \quad (3.5)$$

Avec $L = \sup_{0 \leq t < \infty} \{S_t\}$.

La probabilité de ruine à horizon de temps finie est définie par :

$$\Phi(u, T) = P \left\{ \inf_{t \in [0, T]} R_t < 0, R_0 = u \right\} \quad (3.6)$$

La probabilité complémentaire ou la probabilité de non ruine (probabilité de survie), notée ϕ est défini par :

$$\phi(u) = 1 - \Phi(u), \quad (3.7)$$

À partir de maintenant, nous supposons que N_t est un processus de Poisson de paramètre λ , et donc, $E(N_t) = \lambda t$.

Soit

$$\theta = \frac{p - \lambda\mu}{\lambda\mu}$$

θ est dit coefficient de sécurité relatif.

— Si $\theta > 0$, $R(t) \rightarrow +\infty$ p.s. quand $t \rightarrow +\infty$: activité rentable (net profit condition).

La probabilité de non ruine est non nulle.

— Si $\theta < 0$, alors $R(t) \rightarrow -\infty$ p.s. quand $t \rightarrow +\infty$ d'où $\Phi(u) = 1$.

— Si $\theta = 0$, $\lim_{t \rightarrow +\infty} R(t) = -\infty$ et $\overline{\lim}_{t \rightarrow +\infty} R(t) = +\infty$ et $\Phi(u) = 1$.

Soit $\gamma = \sup_z M_X(z) < \infty$ où $M_X(z) = E(e^{zX})$ est la fonction des moments et soit R une solution positive d'équation :

$$1 + (1 + \theta)\mu R = M_X(R), \quad R < \gamma. \quad (3.8)$$

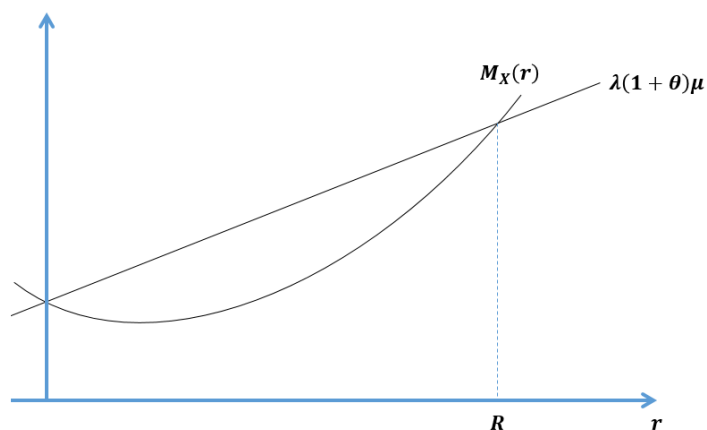


FIGURE 3.3 – Coefficient d'Ajustement

CHAPITRE 3. QUELQUES RÉSULTATS SUR LA PROBABILITÉ DE RUINE POUR LES PROCESSUS DE RISQUE CLASSIQUES

Le coefficient d'ajustement est la plus petite solution strictement positive $R > 0$ de l'équation (3.8). Il est clair que $R = 0$ vérifie l'équation.

Remarque 3.1 Pour certaines lois M_X n'est pas défini (par exemple : Pareto, Log-Normale) et donc le coefficient d'ajustement R n'existe pas.

Si M_X est bien défini, alors le coefficient existe mais n'est pas toujours simple à calculer.

Exemple 3.1 . Si X suit la loi exponentielle de paramètre μ , alors

$$M_X(z) = \begin{cases} \frac{1}{1-\mu z} & \text{si } z < 1/\mu \\ +\infty & \text{sinon.} \end{cases}$$

On doit résoudre l'équation de Lundberg

$$\begin{aligned} \frac{1}{1-\mu z} &= 1 + (1+\theta)\mu z \\ \Leftrightarrow \theta\mu z - (1+\theta)\mu^2 z^2 &= 0 \\ \Leftrightarrow z = 0 \text{ ou } z &= \frac{\theta}{(1+\theta)\mu} \end{aligned}$$

Par conséquent, dans le cas de la loi exponentielle le coefficient de Lundberg est

$$R = \frac{\theta}{(1+\theta)\mu}$$

Exemple 3.2 Si X suit la loi Gamma $G(\alpha = 2, \beta)$ et $\theta = 2$, on a

$$M_X(z) = \begin{cases} \left(\frac{1}{1-\beta z}\right)^\alpha & \text{si } z < 1/\beta \\ +\infty & \text{sinon.} \end{cases}$$

On doit par conséquent résoudre ($\mu = \alpha\beta = 2\beta$ et $1 + \theta = 3$)

$$\begin{aligned} \left(\frac{1}{1-\beta z}\right)^2 &= 1 + 6\beta z \\ \Leftrightarrow 1 &= 1 + 6\beta z - 2\beta z(1 + 6\beta z) + \beta^2 z^2(1 + 6\beta z) \\ \Leftrightarrow 4\beta z - 11\beta^2 z^2 + 6\beta^3 z^3 &= 0 \\ \Leftrightarrow z = 0 \text{ ou } 6\beta^2 z^2 - 11\beta z + 4 &= 0 \end{aligned}$$

Comme $\frac{4}{3\beta} > \frac{1}{2\beta}$ (par ailleurs $\frac{4}{3\beta}$ est en dehors du domaine de M_X), le coefficient de Lundberg est donc $R = \frac{1}{2\beta}$.

3.1.2 Probabilité de ruine exacte en temps infini

On définit $L = \max(S_t - pt \mid t \geq 0)$.

Soit F_L la fonction de répartition de L .

On a

$$\Phi(u) = 1 - F_L(u)$$

$L = L_1 + \dots + L_M$ où L_1 désigne la valeur prise par $S_t - pt$ lorsque le processus passe pour la première fois au-dessus de 0, L_2 la différence entre L_1 et la valeur de $S_t - pt$ où ce record est battu pour la première fois, etc...

— M suit une distribution géométrique de paramètre $q = \theta/(1 + \theta)$.

$$P[K = n] = q^n(1 - q)$$

— La variable aléatoire L peut être exprimé comme

$$L = \sum_{i=1}^M L_i$$

— Les L_i sont i.i.d. de distribution

$$F_{L_1}(x) = F(x)/\mu.$$

Ce qui conduit à la formule de Pollaczek-Khinchine pour la probabilité de ruine.

$$\Phi(u) = 1 - \mathbb{P}(L \leq u) = 1 - \frac{\theta}{1 + \theta} \sum_{n=0}^{\infty} \left(\frac{1}{1 + \theta}\right)^n F_{L_1}^{*n}(u), \quad (3.9)$$

Où $F_{L_1}^{*n}(u)$ est le n^{ime} convolution de la fonction de distribution F_{L_1} .

La formule de Pollaczek-Khinchine (3.9) ne peut être utilisée pour trouver des solutions explicites que pour peu de distributions de montants de sinistres. Pour les autres, afin de calculer la probabilité de ruine, la méthode de Monte Carlo peut être appliquée à (3.4).

Lorsque $u = 0$ il est facile d'obtenir la formule exacte

$$\Phi(0) = \frac{1}{1 + \theta} = \frac{\lambda\mu}{p}, \quad (3.10)$$

3.1.3 Représentation intégrô-différentielle du modèle de Cramer-Lundberg

Théorème 3.1 (Équations intégrô-différentielles pour la probabilité de non ruine)

Soit le processus de surplus $(R_t)_{t \geq 0}$ suivant (3.1). Alors, $\phi(u)$ est continu sur \mathbb{R}_+ avec des

dérivées droite et gauche $\phi'_+(u)$ et $\phi'_-(u)$, respectivement. De plus, nous avons

$$\phi'_+(u) = \frac{\lambda}{p} \left(\phi(u) - \int_0^u \phi(u-y) dF(y) \right), \quad u \geq 0, \quad (3.11)$$

et

$$\phi'_-(u) = \frac{\lambda}{p} \left(\phi(u) - \int_0^{u-} \phi(u-y) dF(y) \right), \quad u > 0. \quad (3.12)$$

Preuve. On voit facilement que la ruine ne se produit pas jusqu'au temps t_1 et qu'elle ne se produira pas après t_1 si et seulement si $X_1 \leq u + pt_1$ et la ruine ne se produit pas sur l'intervalle de temps $(t_1, +\infty)$ avec le surplus initial $u + pt_1 - X_1$. De plus, notons que t_1 est distribuée suivant la loi exponentielle de moyenne $1/\lambda$. Ainsi, nous avons $\mathbb{P}[t_1 > t] = e^{-\lambda t}$ pour tout $t \geq 0$. Ainsi, par la loi des probabilités totales, on obtient

$$\phi(u) = e^{-\lambda t} \varphi(u + pt) + \int_0^t \lambda e^{-\lambda s} \int_0^{u+ps} \varphi(u + ps - y) dF(y) ds \quad (3.13)$$

pour tout $u \geq 0$ et $t \geq 0$. En laissant $t \downarrow 0$ dans (3.13), nous concluons que $\phi(u)$ est continu à droite sur \mathbb{R}_+ . De plus, nous pouvons réécrire (3.13) comme suit

$$\begin{aligned} p \frac{\phi(u + pt) - \phi(u)}{pt} &= \frac{1 - e^{-\lambda t}}{t} \varphi(u + pt) \\ &\quad - \frac{1}{t} \int_0^t \lambda e^{-\lambda s} \int_0^{u+ps} \phi(x + cs - y) dF(y) ds. \end{aligned} \quad (3.14)$$

Ceci montre que $\phi(u)$ est différentiable à droite sur \mathbb{R}_+ . En laissant $t \downarrow 0$ dans (3.14), on obtient (3.11). Ensuite, pour $u > 0$ et $t \leq u/p$, on peut réécrire (3.13) sous la forme

$$\begin{aligned} \phi(u - pt) &= e^{-\lambda t} \phi(u) \\ &\quad + \int_0^t \lambda e^{-\lambda s} \int_0^{u-p(t-s)} \phi(u - p(t-s) - y) dF(y) ds. \end{aligned} \quad (3.15)$$

En laissant $t \downarrow 0$ dans (3.15), nous déduisons que $\varphi(x)$ est continu à gauche sur $(0, +\infty)$. En réécrivant (3.15) sous la forme

$$\begin{aligned} p \frac{\phi(u) - \phi(x - pt)}{pt} &= \frac{1 - e^{-\lambda t}}{t} \phi(u) \\ &\quad - \frac{1}{t} \int_0^t \lambda e^{-\lambda s} \int_0^{x-c(t-s)} \varphi(x - p(t-s) - y) dF(y) ds \end{aligned} \quad (3.16)$$

montre que $\phi(u)$ est différentiable à gauche sur $(0, +\infty)$. Si on laisse $t \downarrow 0$ dans (3.16), on obtient (3.12). ■

Théorème 3.2 Pour tout $u \geq 0$,

$$\phi(u) = \phi(0) + \frac{\lambda}{p} \int_0^u \phi(u-y)(1-F(y)) dy.$$

Preuve. L'équation (3.12), pour $u > 0$, peut être écrite comme suit :

$$\phi'(u) = \frac{\lambda}{p} (\phi(u) - \phi(u) * f(u)). \quad (3.17)$$

Soit $\mathcal{L}\{\phi\}(s) = \int_0^{+\infty} e^{-sx} \phi(x) dx$ la transformée de Laplace de ϕ .

En appliquant la transformée de Laplace à l'équation (3.17), on aura

$$s\mathcal{L}\{\phi\}(s) - \phi(0) = \frac{\lambda}{p} \mathcal{L}\{\phi\}(s) - \frac{\lambda}{p} \mathcal{L}\{\phi\}(s) \cdot s\mathcal{L}\{F\}(s) \quad (3.18)$$

puisque

$$(\mathcal{L}\{\phi\})'(s) = -\phi(0) + s\mathcal{L}\{\phi\}(s)$$

Et

$$\mathcal{L}\{f\}(s) = \mathcal{L}\{F'\}(s) = -F(0) + s\mathcal{L}\{F\}(s) = s\mathcal{L}\{F\}(s).$$

car $F(0) = 0$.

En divisant par s l'équation (3.18), on obtient

$$\mathcal{L}\{\phi\}(s) = \frac{\phi(0)}{s} + \frac{\lambda}{p} \left(\mathcal{L}\{\phi\}(s) \left(\frac{1}{s} - \mathcal{L}\{F\}(s) \right) \right).$$

Or,

$$\mathcal{L}\{\phi(0)\}(s) = \frac{\phi(0)}{s},$$

et

$$\frac{1}{s} - \mathcal{L}\{F\}(s) = \mathcal{L}\{1-F\}(s)$$

On obtient pour tout $s > 0$

$$\mathcal{L}\{\phi\}(s) = \mathcal{L}\{\phi(0)\}(s) + \frac{\lambda}{p} (\mathcal{L}\{\phi\} \mathcal{L}\{1-F\})(s),$$

Ce qui donne le resultat final en utilisant la transformée de la Laplace inverse.

$$\phi(u) = \phi(0) + \frac{\lambda}{p} \int_0^u \phi(u-y)(1-F(y)) dy.$$

■

Théorème 3.3 (Équation intégrale pour la probabilité de ruine à l'horizon infini)

CHAPITRE 3. QUELQUES RÉSULTATS SUR LA PROBABILITÉ DE RUINE POUR
LES PROCESSUS DE RISQUE CLASSIQUES

Soit le processus de surplus $(R_t(u))_{t \geq 0}$ suivant (3.1) et $p > \lambda\mu$. Alors, $\Phi(u)$ satisfait l'équation intégrale

$$\Phi(u) = \frac{\lambda}{p} \left(\int_u^{+\infty} (1 - F(y)) dy + \int_0^u \Phi(u - y)(1 - F(y)) dy \right), \quad u \geq 0. \quad (3.19)$$

Preuve.

En intégrant (3.11) sur $[0, u]$ on obtient

$$\begin{aligned} p(\phi(u) - \phi(0)) &= \lambda \int_0^x \phi(z) dz - \lambda \int_0^u \int_0^z \phi(z - y) dF(y) dz \\ &= \lambda \int_0^u \phi(z) dz - \lambda \int_0^u \int_y^u \phi(z - y) dz dF(y) \\ &= \lambda \int_0^u \phi(z) dz - \lambda \int_0^u \int_0^{u-y} \phi(z) dz dF(y) \\ &= \lambda \int_0^u \phi(z) dz - \lambda \int_0^u \int_0^{u-z} dF(y) \phi(z) dz \\ &= \lambda \int_0^u \phi(z)(1 - F(u - z)) dz \\ &= \lambda \int_0^u \phi(u - y)(1 - F(y)) dy. \end{aligned} \quad (3.20)$$

Ainsi, nous obtenons

$$p(\phi(u) - \phi(0)) = \lambda \int_0^u \phi(u - y)(1 - F(y)) dy, \quad u \geq 0. \quad (3.21)$$

Par convergence monotone, en fait $u \rightarrow +\infty$ dans (3.21), on obtient

$$p \left(\lim_{u \rightarrow +\infty} \phi(u) - \phi(0) \right) = \lambda\mu \lim_{u \rightarrow +\infty} \phi(u). \quad (3.22)$$

Par conséquent, nous avons

$$\phi(0) = \left(1 - \frac{\lambda\mu}{p}\right) \lim_{u \rightarrow +\infty} \phi(u). \quad (3.23)$$

Puisque $p > \lambda\mu$, $\lim_{u \rightarrow +\infty} \phi(u) = 1$. Par conséquent, d'après (3.22), on obtient

$$\phi(0) = 1 - \frac{\lambda\mu}{p} \quad (3.24)$$

Enfin, en substituant $\phi(u) = 1 - \Phi(u)$ dans (3.21) et en utilisant (3.24), on obtient

$$p\Phi(u) = \lambda\mu - \lambda \int_0^u (1 - F(y)) dy + \lambda \int_0^u \Phi(u - y)(1 - F(y)) dy, \quad u \geq 0. \quad (3.25)$$

et en utilisant le fait que $\mu = \int_0^{+\infty} (1 - F(y))dy$, nous aurons donc le résultat (3.19). ■

3.1.4 Cas particuliers

1. Loi exponentielle :

Un cas particulier du modèle de risque classique est le modèle de [Lundberg \(1903\)](#). Ce modèle se caractérise par la distribution exponentielle des montants des réclamations, c'est-à-dire :

$$F_X(y) = 1 - e^{-\lambda y} \quad (3.26)$$

où F_X est la fonction de répartition de la variable aléatoire X qui génère le montant des réclamations.

En remplaçant dans l'équation (3.19) F_X par la loi exponentielle, nous montrons que la probabilité de ruine a la forme suivante

$$\Phi(u) = \frac{\lambda\mu}{p} \exp(-\lambda\mu) \quad (3.27)$$

ou bien

$$\Phi(u) = \frac{1}{1+\theta} \exp\left(\frac{-\theta\lambda u}{1+\theta}\right) \quad (3.28)$$

Exemple 3.3 *La probabilité de ruine dans le cas d'une loi exponentielle ou un mélange de lois exponentielles est relativement simple à calculer.*

TABLE 3.1 – Probabilité de ruine pour des remboursements exponentielles avec $\lambda = 1/1000000000$ et $\theta = (0.1, 0.3, 0.9)$.

u	0	10^3	10^6	10^8	10^9	10^{10}
				$\theta = 0.1$		
$\Phi(u)$	0.99009	0.99009	0.99000	0.98034	0.89676	0.36786
				$\theta = 0.3$		
$\Phi(u)$	0.97087	0.97087	0.9705910	0.94300	0.72555	0.05275
				$\theta = 0.9$		
$\Phi(u)$	0.92592592	0.92592524	0.92524030	0.85981750	0.44144502	0.00056179

2. Loi Gamma :

Exemple 3.4 *Grandell and Segerdahl (1971)* ont montré que pour la distribution Gamma du montant du sinistre avec une moyenne de 1 et $\alpha < 1$, la valeur exacte de la probabilité de ruine peut être calculée par la formule, c'est à dire :

$$\alpha > 0, \beta > 0 \quad f_X(x) = \frac{\beta^{(\alpha)}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x \geq 0$$

$$\Phi(u) = \frac{\theta(1 - R/\alpha) \exp(-Ru)}{1 + (1 + \theta)R - (1 + \theta)(1 - R/\alpha)} + \frac{\alpha\theta \sin(\alpha\pi)}{\pi} \cdot \mathbb{J},$$

Où

$$\mathbb{J} = \int_0^\infty \frac{x^\alpha \exp\{-(x+1)\alpha u\}}{[x^\alpha \{1 + \alpha(1 + \theta)(x+1)\} - \cos(\alpha\pi)]^2 + \sin^2(\alpha\pi)} dx.$$

TABLE 3.2 – Probabilité de ruine pour des remboursements de loi Gamma avec $\alpha = 10$
 $\beta = 1/10000000$ et $\theta = (0.1, 0.3, 0.9)$.

u	0	10^3	10^6	10^8	10^9	10^{10}
				$\theta = 0.1$		
$\Phi(u)$	0.99009	0.99009	0.99000	0.97507	0.82883	0.16327
				$\theta = 0.3$		
$\Phi(u)$	0.97087	0.97087	0.97058	0.92804	0.57392	0.00470
				$\theta = 0.9$		
$\Phi(u)$	0.9174311	0.91743	0.91667	0.80632	0.202120	0.00000098715

3. Mélange de deux lois exponentielles :

Exemple 3.5 La distribution du montant soit un mélange de deux exponentielles avec les paramètres β_1, β_2 , et de poids $p, (1 - p)$, est donnée par :

$$f_X(y) = p\beta_1 e^{-\beta_1 y} + (1 - p)\beta_2 e^{-\beta_2 y} \quad (3.29)$$

En utilisant l'inversion de la transformée de Laplace, on peut obtenir une formule explicite, voir *Panjer and Willmot (1992)* :

$$\Phi(u) = \frac{1}{(1 + \theta)(r_2 - r_1)} \{(\rho - r_1) \exp(-r_1 u) + (r_2 - \rho) \exp(-r_2 u)\},$$

Où

$$r_1 = \frac{\rho + \theta(\beta_1 + \beta_2) - \left[\rho + \theta(\beta_1 + \beta_2) \right]^2 - 4\beta_1\beta_2\theta(1 + \theta)}{2(1 + \theta)}$$

$$r_2 = \frac{\rho + \theta(\beta_1 + \beta_2) + \left[\rho + \theta(\beta_1 + \beta_2) \right]^2 - 4\beta_1\beta_2\theta(1 + \theta)}{2(1 + \theta)}$$

et

$$q = \frac{p\beta_1^{-1}}{p\beta_1^{-1} + (1-p)\beta_2^{-1}}, \quad \rho = \beta_1(1-q) + \beta_2q.$$

TABLE 3.3 – Probabilité de ruine pour des remboursements pour un mélange de 2 exponentielles avec $\beta_1 = 1/10000000$, $\beta_2 = 1/100000000$, $p = 0.7$ et $\theta = (0.1, 0.3, 0.9)$.

u	0	10^3	10^6	10^8	10^9	10^{10}
$\theta = 0.1$						
$\Phi(u)$	0.76237	0.76237	0.75914	0.57987	0.059025	0
$\theta = 0.3$						
$\Phi(u)$	0.7475728	0.74756	.7442107	0.55915	0.049511	0
$\theta = 0.9$						
$\Phi(u)$	0.71296	0.71295	0.70931	0.1281	0.032825	0

3.2 Processus de risque de queue Sous exponentielle

Considérons le modèle de risque classique (3.2), on s'intéresse au cas où X_i ont des queues lourdes, une fonction de distribution F est une sous exponentielle i.e $F \in \mathcal{S}$ si :

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_1 + X_2 > x)}{\mathbb{P}(X_1 > x)} = 2 \quad (3.30)$$

Où X_1, X_2 sont des v.a *i.i.d* avec une fonction de distribution F (Voir le lemme (1.1)).

Notons $F(x) = \mathbb{P}(X_1 \geq x)$ et supposons que $E(X_1) < \infty$, et on défini :

$$F_0(x) = \frac{1}{E(X_1)} \int_0^x \bar{F}(y) dy \quad (3.31)$$

par Goldie and Resnick (1988) ont montré que si $F \in \mathcal{S}$ est satisfait à certaines condition de lissage, alors F appartient au domaine d'attraction maximum de la distribution Fréchet ou la distribution de Gumbel.

3.2.1 Cas distribution de Fréchet

Dans le premier cas (distribution de Fréchet), \bar{F} varie régulièrement i.e :

$$\bar{F}(x) = \frac{L(x)}{x^{\alpha+1}}$$

pour certains $\alpha > 0$ et nous écrivons comme $\bar{F} \in \mathcal{R}\{-\alpha - 1\}$, $\alpha > 0$.

Où $L : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ est fonction à variation lente i.e :

$$\lim_{x \rightarrow \infty} \frac{L(\gamma x)}{L(x)} = 1, \quad \text{pour tout } \gamma > 0.$$

3.2.2 Condition de stationnarité

$T_i = \tau_i - \tau_{i-1}$ est la durée de l'intervalle de temps entre deux arrivées consécutives du processus ponctuel τ_{i-1} et τ_i , représentent le processus ponctuel.

Si T_i sont *i.i.d* avec moyenne $E[T_1]$ alors N_t est une processus de renouvellement et assume la condition habituelle de :

$$\theta = \frac{\mu}{pE[T_1]} < 1 \quad (3.32)$$

Si N_t est une processus de Poisson d'intensité λ , on a

$$\theta = \frac{\lambda\mu}{p} < 1 \quad (3.33)$$

alors, il est bien connu que (Voir [McNeil \(1997\)](#), [Asmussen and Albrecher \(2010\)](#)) :

$$\lim_{u \rightarrow \infty} \frac{\Phi(u)}{\bar{F}_0(u)} = \frac{\theta}{1 - \theta} \quad (3.34)$$

Si on pose $\rho = \lambda/p$ et $\mu_X = \mu$, l'équation (3.34) devient

$$\lim_{u \rightarrow \infty} \frac{\Phi(u)}{\bar{F}_0(u)} = \frac{\mu_X}{\rho - \mu_X} \quad (3.35)$$

avec $\Phi(u) = \mathbb{P}(\tau_u < \infty)$ est la probabilité de ruine à horizon infinie.

Dans la section suivante, nous présentons notre principale contribution qui était l'objet d'une publication dans un journal international (The Chilean Journal of Statistics (ChJS) <http://soche.cl/chjs/issues.html>) Frihi et al. (2022).

Nous avons construit un nouvel estimateur pour la probabilité de ruine. Nous avons montré que cet estimateur est asymptotiquement consistant et il est de loi normale.

Ces résultats ont été prouvés dans le cas où les sinistres arrivent suivant un processus de Poisson et la loi des montants de remboursement est à queue lourde à l'horizon infini, pour une large réserve initiale.

Notre approche est basée sur le résultat de Pickands (1975 et Balkema and De Haan (1974) qui donne une approximation de la distribution des excès par une distribution de Paréto généralisée.

3.3 Calcul de la probabilité de ruine à horizon infini

We are interested in the case when X_1, X_2, \dots, X_n have heavy tails.

Let X be a positive random variable with distribution function F , where

$$\bar{F}(x) = Cx^{-\frac{1}{\xi}}(1 + x^{-\delta}\mathbb{L}(x)), \text{ as } x \uparrow +\infty, \quad (3.36)$$

and $\bar{F}(x) = 1 - F(x)$ represent the tail function of the distribution F .

For $\xi \in (0, 1)$, $\delta > 0$, and some real constant C , with \mathbb{L} a slowly varying function at infinity, that $\frac{\mathbb{L}(tx)}{\mathbb{L}(x)} \rightarrow 1$, as $x \uparrow +\infty$ for $t > 0$. For more details on these function, see, e.g., chapter 0 in Resnick (1987) or Seneta (2006).

F is regularly varying tail at $+\infty$ with index ξ , that $\frac{\bar{F}(tx)}{\bar{F}(x)} \rightarrow t^{-\frac{1}{\xi}}$, as $x \uparrow +\infty$ uniformly for $t > 0$.

F is called heavy tail with tail index $\xi > 0$ (e.g., Pareto, Burr, log-Normal, log-Gamma, Student, etc.), see, e.g., (Dekkers and de Haan (1993), Dekkers et al.1989).

Notice that when $\xi \in (0, 1/2)$, we have $\mu_X = E(X) < \infty$ and $E(X^2) < \infty$. But if $\xi \in (1/2, 1)$, we have $E(X^2) = \infty$.

Any F with a regularly varying tail is subexponential, i.e. $\frac{\bar{F}^{*2}(x)}{\bar{F}(x)} \rightarrow 2$, as $x \uparrow +\infty$. Here $F * G$ denotes the convolution of the distribution functions F and G , $F^{*1} = F$ and $F^{*(n+1)} = F^{*n} * F$, (see Asmussen and Albrecher (2010))

If F is subexponential, it has been shown that for large initial reserve u , the ruin probability $\Phi(u)$, (see, e.g., Klüppelberg and Stadtmüller (1998), Resnick (1987), Asmussen and Albrecher (2010), Kalashnikov and Konstantinides (2000) and Zhu (2013)), given by

$$\lim_{u \rightarrow +\infty} \frac{\Phi(u)}{\bar{F}_0(u)} = \frac{\mu_X}{\rho - \mu_X}, \quad (3.37)$$

where $\rho = p/\lambda > \mu_X$ F_0 denotes the stationary excess distribution or the integrated tail distribution,

$$F_0(x) = \frac{1}{\mu_X} \int_0^x \bar{F}(x) dx, \quad x \geq 0, \quad (3.38)$$

We can then write the approximation of the ruin probability $\Phi(u)$ as follows.

$$\Phi(u) \simeq \frac{1}{(\rho - \mu_X)} \int_u^{+\infty} \bar{F}(x) dx \text{ for a large initial reserve } u \quad (3.39)$$

3.3.1 Traditional Estimator of the Ruin Probability

Let X_1, \dots, X_n be a random sample of size n of cdf F . The Tail of the distribution F is assumed to start at some level u_n supposed sufficiently high, and we set $Y_i = \max(X_i - u_n, 0)$ with cdf F_Y . By replacing F and μ_X by its empirical estimators \hat{F}_n and $\hat{\mu}_X$ respectively in equation 3.39, the estimator is written as follows :

$$\hat{\Phi}_n(u_n) = \frac{1}{(\rho - \hat{\mu}_X)} \int_{u_n}^{+\infty} \hat{F}_n(x) dx \text{ for a large initial reserve } u_n \quad (3.40)$$

In section 3.6, we prove that the estimator given in (9) is asymptotically normal provided that it has a finite second moment.

To illustrate the performance and normality of $\hat{\Phi}_n(u_n)$, we draw samples from the Pareto distribution $\bar{F}(x) = x^{-1/\xi}$, $x > 1$, $\xi > 0$ using R packages (actuar, evir, extremefit, tea, boot) to generate samples, and to estimate parameters of the GPD distribution. The qq-plot and bootstrap functions were used to adjust graphically the normality of the ruin probability. We choose two values $\xi = 1/4$ in which case we have the two moments are finite and $\xi = 3/4$ in which case we have the second moment is infinite as shown in figure 1. We can notice that the ruin probability is well fitted by the normal distribution for $\xi = 1/4$ (finite variance). On the other hand, if $\xi = 3/4$ (infinite variance), the ruin probability is far from the normal distribution (See Figure 3.4).

3.3.2 POT-based estimator of the ruin probability

If X_1, X_2, \dots, X_n be positive *i.i.d.* rv's with cdf F under assumption (3.36) and $u_n = O_+(n^{\alpha\xi})$ for some $\alpha \in (1/(1 + 2\delta\xi), 1)$, where $\delta > 0$, and if $Y_i = X_i - u_n$, for all i such that $X_i > u_n$ be positive *i.i.d.* rv's with cdf F_Y .

By definition

$$\bar{F}(y + u_n) = \bar{F}_Y(y)\bar{F}(u_n), \quad (3.41)$$

for $p_n = P(X_1 > u_n) = \bar{F}(u_n)$, the estimation of p_n may be done using

$$\hat{p}_n = \widehat{\bar{F}}(u_n) = \frac{1}{n} \sum_{i=0}^n 1_{\{X_i > u_n\}} = \frac{N_{u_n}}{n}, \quad (3.42)$$

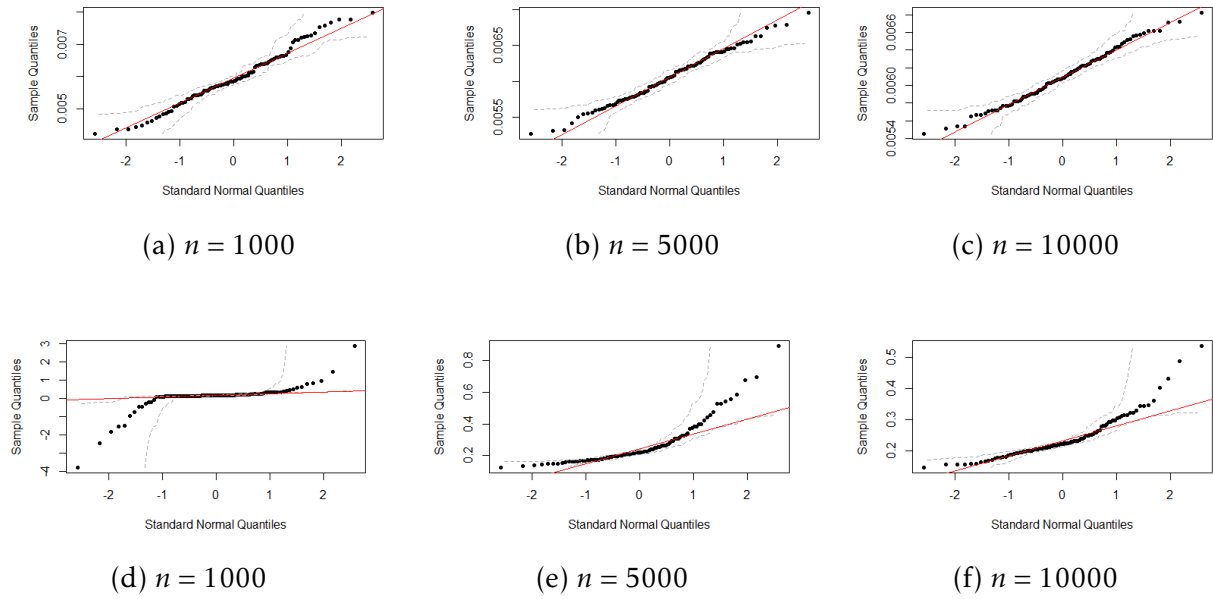


FIGURE 3.4 – QQ-plot of the Ruin probability against the normal distribution when $\xi = \frac{1}{4}$ (top row) and $\xi = \frac{3}{4}$ (bottom row)

$N := N_{u_n} = \sum_{i=1}^n \mathbf{1}_{\{X_i > u_n\}}$, where N is the number of X_i which exceed u_n , we have a binomial distribution, i.e, $N \rightsquigarrow \mathcal{B}(n, p_n)$.

When $np_n = n^{1-\alpha} O_+(1)$, it follows that from the Central Limit Theorem that,

$$\frac{\sqrt{n}}{\sqrt{p_n(1-p_n)}}(\hat{p}_n - p_n) \xrightarrow{L} \mathcal{N}(0, 1) \text{ as } n \uparrow +\infty. \quad (3.43)$$

Further, let $\mu_n = E(X_1 \mathbf{1}_{\{X_1 \leq u_n\}})$ with its empirical estimator $\hat{\mu}_n = n^{-1} \sum_{i=1}^n X_i \mathbf{1}_{\{X_i \leq u_n\}}$ and $\sigma_n^2 = \text{Var}(X_1 \mathbf{1}_{\{X_1 \leq u_n\}})$. Then, from the Central Limit Theorem, we have,

$$\frac{\sqrt{n}}{\sigma_n}(\hat{\mu}_n - \mu_n) \xrightarrow{L} \mathcal{N}(0, 1) \text{ as } n \uparrow +\infty \quad (3.44)$$

(see e.g., [Smith \(1987\)](#)). For large values of u_n , we use $\widehat{F}_Y(y) \approx \bar{G}_{\hat{\xi}_n, \hat{\beta}_n(u_n)}(y)$, for appropriate estimates $\hat{\xi}_n$, and $\hat{\beta}_n(u_n)$. Note that β_n will be estimated separately, i.e. $\beta_n = \xi u_n$ will be not used, see, e.g., [Smith \(1987\)](#).

Notice that the maximum likelihood estimates $\hat{\xi}_n$ and $\hat{\beta}_n$ are consistent and converges in probability, see e.g., [Beirlant and Teugels \(1987\)](#) and Section 5.2, [Smith \(1987\)](#).

For $\xi > -0.5$ it can be shown that maximum likelihood regularity conditions are fulfilled and that maximum likelihood estimates $(\hat{\xi}_n, \hat{\beta}_n)$ based on a sample of N excesses of a threshold u_n are asymptotically normally distributed, see e.g., [Smith \(1985\)](#), [Smith](#)

(1987), and [Fard and Holmquist \(2013\)](#) as follows

$$\sqrt{np_n}Q^{\frac{1}{2}}\begin{pmatrix} \hat{\beta}_n - \beta \\ \hat{\xi}_n - \xi \end{pmatrix} \xrightarrow{L} \mathcal{N}(0, I), \text{ as } n \uparrow +\infty, \quad (3.45)$$

where

$$Q^{-1} = (1 + \xi) \begin{pmatrix} 2\beta^2 & -\beta \\ -\beta & 1 + \xi \end{pmatrix}.$$

Under the assumption that $\sqrt{np_n}u_n\mathbb{L}(u_n) \rightarrow 0$ and that $x^{-\delta}\mathbb{L}(x)$ is non-increasing, this condition is met if $u_n = O_+(n^{\alpha\xi})$ with $\alpha > \frac{1}{(1+2\delta\xi)}$, where \mathbb{L} is slowly varying at $+\infty$, (see [Johansson \(2003\)](#)). If we use the equation (3.41), we can write

$$\int_{u_n}^{+\infty} \bar{F}(x)dx = \int_0^{+\infty} \bar{F}(y + u_n)dy = p_n \int_0^{+\infty} \bar{F}_Y(y)dy = p_n\mu_Y$$

Then, the simplified form is

$$\Phi(u) \simeq p_n \frac{\mu_Y}{(\rho - \mu_X)}$$

For large values of u_n , and under the assumption $\rho = \frac{p}{\lambda} > \mu_X$.

By the replacement of the μ_X and μ_Y by its empirical estimators, $\hat{\mu}_X$ and $\hat{\mu}_Y$ respectively, we can define an empirical estimator of $\Phi(u_n)$ as follows

$$\hat{\Phi}_n(u_n) = \hat{p}_n \frac{\hat{\mu}_Y}{(\rho - \hat{\mu}_X)} \quad (3.46)$$

For $\xi \in (0, 1/2)$, X_1 has finite variance ($\sigma_n^2 < \infty$), in this case $\mu_X = E(X_1)$ and $\mu_Y = E(Y_1)$ is estimated by the sample mean $\hat{\mu}_X = \bar{X} = n^{-1} \sum_{i=1}^n X_i$, and $\hat{\mu}_Y = \bar{Y} = n^{-1} \sum_{i=1}^n Y_i$ respectively. The estimator of $\Phi_n(u_n)$ become as follows :

$$\hat{\Phi}_n^{(1)}(u_n) = \hat{p}_n \frac{\bar{Y}}{(\rho - \bar{X})} \quad (3.47)$$

Using the central limit theorem, the asymptotic normality of $\hat{\Phi}(u_n)$ is established in the [Theorem 3.1](#).

An alternative way to estimate μ_X may be made by GPD's approximation (see [Johansson \(2003\)](#)). Indeed, for each $n \geq 1$, we have

$$\mu_X = \int_0^{\infty} x dF(x) = \mu_n^* + \kappa_n, \quad (3.48)$$

where

$$\mu_n^* = \int_0^{u_n} x dF(x) \text{ and } \kappa_n = \int_{u_n}^{\infty} x dF(x).$$

For $\xi \in (1/2, 1)$, X_1 has an infinite variance ($\sigma_n^2 = \infty$) Johansson's estimator (see, [Johansson \(2003\)](#)) of μ_X is given by

$$\hat{\mu}_{X,n}^J = \hat{\mu}_n^* + \hat{\kappa}_n, \quad (3.49)$$

where

$$\hat{\mu}_n^* = \frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{\{X_i \leq u_n\}} \text{ and } \hat{\kappa}_n = \hat{p}_n \left(u_n + \frac{\hat{\beta}_n}{1 - \hat{\xi}_n} \right),$$

and μ_Y is estimated by the approximate of the tail distribution function \bar{F} by the GPD distribution \bar{G} , and after the integrating, we obtain the formula

$$\hat{\mu}_{Y,n} = \frac{\hat{\beta}_n}{1 - \hat{\xi}_n}. \quad (3.50)$$

Then, the estimator of $\Phi_n(u_n)$ become as follows :

$$\hat{\Phi}_n^{(2)}(u_n) = \hat{p}_n \frac{\hat{\mu}_{Y,n}}{(\rho - \hat{\mu}_{X,n}^J)} \quad (3.51)$$

In the end, an asymptotic estimator of $\Phi_n(u_n)$ for any $0 < \xi < 1$ has the following form :

$$\hat{\Phi}_n(u_n) := \begin{cases} \hat{\Phi}_n^{(1)}(u_n) & \text{for } 0 < \xi \leq 1/2 \\ \hat{\Phi}_n^{(2)}(u_n) & \text{for } 1/2 < \xi < 1 \end{cases} \quad (3.52)$$

In the following section we present our main results.

3.4 Main results

In this section, we present the main results of this paper. The three theorems provides the consistency and the asymptotic normality of our estimator.

Our main first result is the asymptotic normality of $\hat{\Phi}_n(u_n)$, when $E(X) < \infty$ and $E(X^2) < \infty$ (i.e., $0 < \xi \leq 1/2$). This result is a straight application of the central limit theorem.

Théorème 3.1 *Let F be a cdf fulfilling (3.36) with $\xi \in (0, 1/2)$, we have*

$$\sqrt{n}(\hat{\Phi}_n^{(1)}(u_n) - \Phi(u_n)) \xrightarrow{D} \mathcal{N}(0, \sigma^2), \quad \text{as } n \uparrow +\infty. \quad (3.53)$$

where the asymptotic variance σ^2 is given by the formula

$$\sigma^2 = \frac{1}{(\rho - \mu_X)^4} (\sigma_1^2 (\rho - \mu_2 - u_n)^2 + \sigma_2^2 (\mu_1 - u_n)^2),$$

and

$$\begin{aligned}\mu_1 &= E(\bar{X}_1(X_i \geq u_n)), \\ \mu_2 &= E(\bar{X}_1(X_i \leq u_n)), \\ \sigma_1^2 &= \text{Var}(\bar{X}_1(X_i \geq u_n)), \\ \sigma_2^2 &= \text{var}(\bar{X}_1(X_i \leq u_n)).\end{aligned}$$

The second main result is the almost pointwise sure convergence and asymptotic normality of $\hat{\Phi}_n(u_n)$, when $E(X^2) = \infty$ (i.e., $1/2 < \xi < 1$). In this case, the central limit theorem is not applicable.

Théorème 3.2 *Let F be a cdf fulfilling (3.36) with $\xi \in (1/2, 1)$. Suppose that \mathbb{L} is locally bounded in $[x_F, +\infty)$ for $x_F \geq 0$ and $x \mapsto x^{-\delta}\mathbb{L}(x)$ is non-increasing near infinity, for some $\delta > 0$. For any $u_n = O_+(n^{\alpha\xi})$ with $\alpha \in (1/(1 + 2\delta\xi), 1)$, we have*

$$\Phi(u_n) = \hat{\Phi}_n^{(2)}(u_n) + o_P(1), \quad \text{as } n \uparrow +\infty.$$

Théorème 3.3 *Let F be as in Theorem 3.2, then for any $u_n = O(n^{\alpha\xi})$ with $\alpha \in (1/(1 + 2\delta\xi), 1)$, we have*

$$\frac{\sqrt{n}}{\sigma} \left(\hat{\Phi}_n^{(2)}(u_n) - \Phi(u_n) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad \text{as } n \uparrow +\infty,$$

where

$$\sigma^2 = \theta_1 \left(p_n^2 \sigma_n^2 + p_n(1 - p_n)\theta_2 + \frac{p_n(2\xi^2 - \xi + 1)(1 + \xi)}{(1 - \xi)^2} \theta_3 \right) = O_+(1)$$

with

$$\theta_1 = \frac{\beta^2}{(1 - \xi)^2 (\rho - \mu_X)^4}, \quad \theta_2 = (\rho - u_n)^2, \quad \theta_3 = (\rho - \mu_n - p_n u_n)^2.$$

The proofs of the above theorems are given in the Appendix.

3.5 Simulation Study and Application

3.5.1 Simulation Study

In this section, we have applied the result of the 3.2 and 3.3 we first begin to fix a significance level $\eta \in (0, 1)$ and calculate $z_{\frac{\eta}{2}}$ for the $(1 - \frac{\eta}{2})$ quantile of the standard normal distribution $\mathcal{N}(0, 1)$. If we use a realization of the random variables X_1, \dots, X_n , which follow a cdf F satisfying the conditions of the theorem 3.3, we construct a level $1 - \eta$ confidence interval for $\Phi(u)$ as follows.

$$\Phi(u_n) \in \left[\bar{\Phi}_n^{(2)}(u_n) - \frac{1}{\sqrt{n}} z_{\frac{\eta}{2}} \hat{\sigma}_{\bar{\Phi}_n^{(2)}(u_n)}, \bar{\Phi}_n^{(2)}(u_n) + \frac{1}{\sqrt{n}} z_{\frac{\eta}{2}} \hat{\sigma}_{\bar{\Phi}_n^{(2)}(u_n)} \right],$$

where $\hat{\Phi}_n^{(2)}(u_n)$ and $\hat{\sigma}_{\hat{\Phi}_n^{(2)}(u_n)}^2$ are the empirical mean and variance respectively.

We realized a simulation study to validate the performance of the ruin probability estimation and the asymptotic normality of our proposed estimator. All numerical evaluations and graphics presented here were done in the R software using the packages POT, evir, extremefit, and actuar for Extreme values. To this end, we simulated samples of the Pareto, Burr and Log-Gamma distributions, whose tail index parameters are summarized in Table 3.4.

TABLE 3.4 – Regularly varying distribution functions

Distribution	Tail $\bar{F}(x)$ or density $f(x)$	Parameters	Index Parameter ($\frac{1}{\xi}$)
Pareto	$\bar{F}(x) = x^{-c}, x > 1$	$c > 0$	c
Burr	$\bar{F}(x) = (1 + x^c)^{-k}, x > 0$	$c, k > 0$	ck
Log-Gamma	$f(x) = \frac{c^k}{\Gamma(k)} (\ln(x))^{k-1} x^{-c-1}, x > 0$	$c, k > 0$	c

We generated 200 samples of sizes 500, 1000 and 5000 from the previous distributions with the index values $\xi = 2/3$ and $\xi = 3/4$.

For each simulated sample, we obtain a value of the estimators $\hat{\Phi}_n^{(2)}(u_n)$. The overall estimated $\hat{\Phi}_n^{(2)}(u_n)$ is then taken as the empirical mean of the values in the 200 repetitions with its confidence interval. We also obtain the bias and the root mean squared error (RMSE) of the estimator of $\hat{\Phi}_n^{(2)}(u_n)$. We summarize the results in Table 3.5 and Table 3.6.

TABLE 3.5 – 95% confidence interval for the ruin probability of Pareto, Burr and Log-Gamma Distributions with tail index $\xi = \frac{2}{3}$

Distribution	n	$\hat{\Phi}(u)$	Bias	RMSE	Confidence Interval
$\Phi(u) = 0.08186737$					
Pareto	500	0.079678	-0.002188	0.027843	(0.073641, 0.085715)
	1000	0.081419	-0.000447	0.022333	(0.077098, 0.08574157)
	5000	0.081760	0.000106	0.020226	(0.080826, 0.082695)
$\Phi(u) = 0.065366$					
Burr	500	0.061051	-0.004314	0.008490	(0.051344, 0.070759)
	1000	0.061616	-0.003749	0.007504	(0.056035, 0.067198)
	5000	0.064166	-0.001199	0.006906	(0.061623, 0.066710)
$\Phi(u) = 0.07368063$					
Log-Gamma	500	0.071909	-0.001771	0.025905	(0.059811, 0.084006)
	1000	0.075159	0.001479	0.023896	(0.072138, 0.078181)
	5000	0.074933	0.001253	0.019920	(0.072138, 0.077729)

TABLE 3.6 – 95% confidence interval for the ruin probability of Pareto, Burr and Log-Gamma Distributions with tail index $\xi = \frac{3}{4}$ b top 5% of the data

Distribution	n	$\hat{\Phi}(u)$	Bias	RMSE	Confidence Interval
$\Phi(u) = 0.177327$					
Pareto	500	0.161809	-0.015517	0.054281	(0.141244, 0.182373)
	1000	0.168686	-0.008639	0.044726	(0.154511, 0.182861)
	5000	0.179787	0.002460	0.036559	(0.176026, 0.183547)
$\Phi(u) = 0.1052539$					
Burr	500	0.094451	-0.010802	0.007272	(0.057324, 0.122166)
	1000	0.099181	-0.006072	0.004516	(0.082715, 0.115647)
	5000	0.100817	-0.004436	0.004436	(0.094665, 0.106969)
$\Phi(u) = 0.1576236$					
Log-Gamma	500	0.171419	0.013795	0.053793	(0.150337, 0.192500)
	1000	0.159431	0.001807	0.038889	(0.150499, 0.168363)
	5000	0.157925	0.000302	0.035215	(0.155589, 0.160261)

3.5.2 Application

We illustrate an application to the Danish fire insurance data (in millions DKK). The Danish data on large fire insurance claims are widely used and provides an exceptional example of the use of the extreme value theory in a significant application context, see [McNeil \(1997\)](#), [Mikosch and Samorodnitsky \(2000\)](#).

The full Danish data of Reinsurance comprise 2167 fire losses from Thursday 3rd January 1980 until Monday 31st December 1990.

[Mikosch and Samorodnitsky \(2000\)](#) confirm that a homogeneous Poisson process is an appropriate model for the arrivals of the Danish fire insurance data for shorter periods of time such as one year with the parameter $\lambda = 1/1.85$.

[McNeil \(1997\)](#) was adjusted the danish fire insurance data, he concluded that the Lognormal distribution to be good, but the Pareto distribution did not fit the data well, while the GPD distribution was acceptable.

We restrict our attention to the 2156 losses exceeding one million.

The descriptive statistical study of the Danish data is summarized in [Table 3.7](#).

TABLE 3.7 – Descriptive Statistics Summary of the dataset of Danish fire

min	Mean	Std	First Quartile	Median	Third Quartile	max
1	1.85	8.5274	1.321	1.778	2.967	263.250

All numerical evaluations and graphics were presented in the R software using the packages POT, evir, extremefit, and actuar for Extreme values. We also used the functions qqnorm for the adjustment with the normal distribution. For the resampling, we used the package boot.

The generalized Pareto distribution can be fitted to data on excesses of high thresholds by a variety of methods including the maximum likelihood method (ML) to estimate the parameters and the Mean Residual Life Plot (MRLP) to select the threshold. The MRLP function (see again 3.5) is the plot of $\{(u_n, e_n(u_n)), X_{(1)} < u_n < X_{(n)}\}$ where $X_{(1)}$, and $X_{(n)}$ are the first and nth order statistics and $e_n(u)$ is the sample mean excess function gives by the sum of the excesses over the threshold u_n divided by the number of data points which exceed the threshold u_n .

$$e_n(u_n) = \frac{1}{N} \sum_{i=1}^N (X_i - u_n), X_i > u_n$$

The interpretation of the mean excess plot as explained in [Beirlant et al. 1996](#) and [Embrechts \(2000\)](#). There is evidence of a straightening out of the plot above a threshold of 1, 10 and perhaps 20 (See Figure ??).

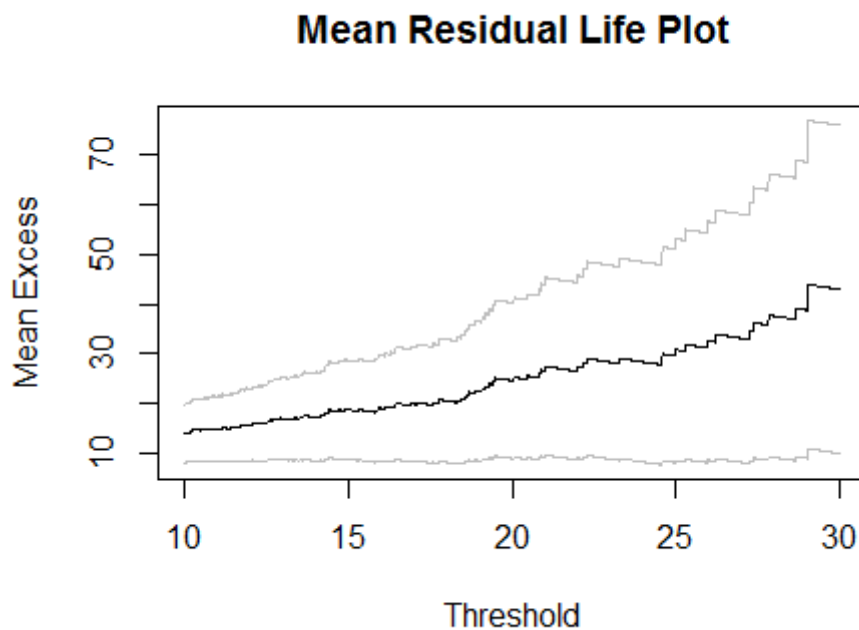


FIGURE 3.5 – Mean Residual Life Plot

The points show an upward trend, then this is a sign of heavy-tailed behavior. The Histogram (Figure 3.6) shows that the data may perhaps is a heavy-tailed distribution. It is Noted that we must take the threshold $u \geq 20$ to have the shape parameter $\xi > 0.5$, which fits our case (see again Figure 3.7). We look at standard choices of curve fitted to the whole dataset. We use the GPD with the threshold $u = 20$, the maximum likelihood estimators of the parameters are given by the shape $\hat{\xi} = 0.6840479$ and the scale $\hat{\beta} = 9.6316941$.

The parameter estimates are the same as those given by [McNeil \(1997\)](#) for the GPD

Histogram of danishClaims\$DANISH

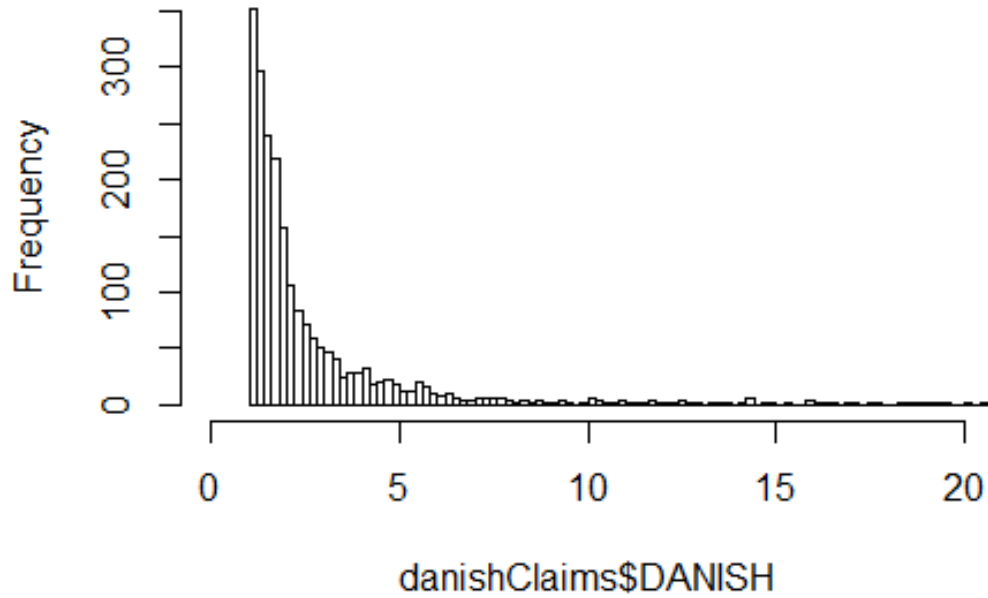


FIGURE 3.6 – The Histogram for the Danish data

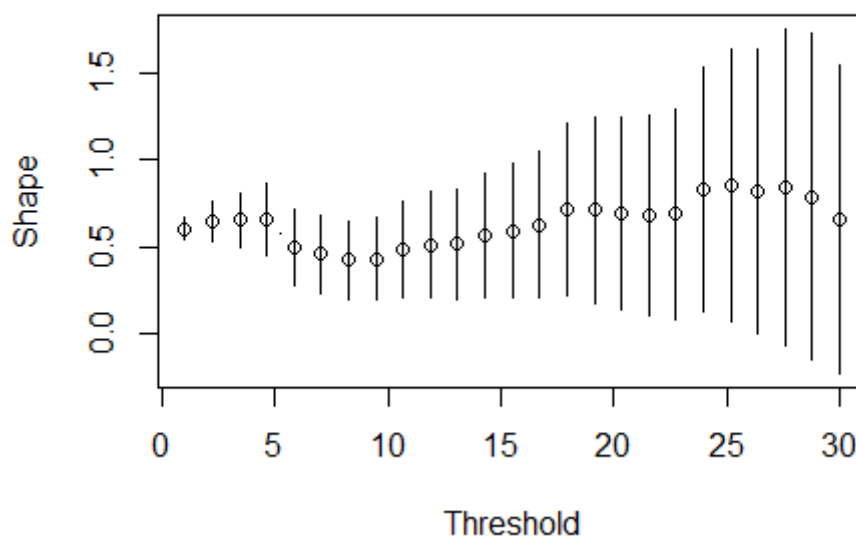


FIGURE 3.7 – Shape parameter as function of the threshold for the Danish data

distribution.

The QQ-plot in Figure 3.8 indicate that the ruin probability of the Danish data via bootstrap sampling (random sampling with replacement), see e.g., [Burkhalter and Lio \(2021\)](#) is well fitted by the normal distribution when the losses exceeding a threshold $u = 20$.

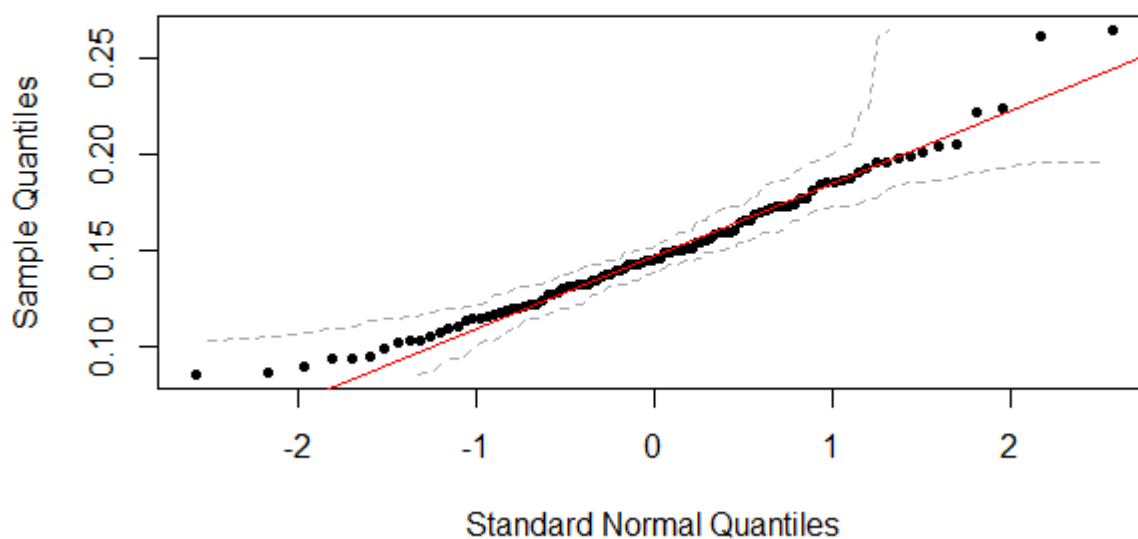


FIGURE 3.8 – The QQ-plot for ruin probability with bootstrap method

3.6 Proofs of the Main results

3.6.1 Proof of Theorem 3.1

Preuve. The proof of this theorem is based on the multivariate Delta method theorem which generalizes the central limit theorem.

If we take

$$Z_n = \begin{pmatrix} Z_n^1 \\ Z_n^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n X_i 1_{\{X_i \geq u_n\}} \\ \frac{1}{n} \sum_{i=1}^n X_i 1_{\{X_i \leq u_n\}} \end{pmatrix} = \begin{pmatrix} \bar{X} 1_{\{X \geq u_n\}} \\ \bar{X} 1_{\{X \leq u_n\}} \end{pmatrix}.$$

Therefore the asymptotic normality of Z_n follows directly from the classical Central Limit Theorem

$$\sqrt{n}(Z_n - \mu) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma), \text{ as } n \uparrow +\infty,$$

where $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$,

And if we take $g(s_1, s_2) = \frac{s_1 - u_n}{\rho - (s_1 + s_2)}$, we obtain

$$\sqrt{n}(g(Z_n) - g(\mu)) = \sqrt{n} \left(\hat{p}_n \frac{\bar{Y}}{(\rho - \bar{X})} - p_n \frac{\mu_Y}{\rho - \mu_X} \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \nabla g(\mu)^T \Sigma \nabla g(\mu) \right), \text{ as } n \uparrow +\infty.$$

where

$$\nabla g(\mu) = \begin{pmatrix} \frac{\partial g}{\partial y_1}(\mu) \\ \frac{\partial g}{\partial y_2}(\mu) \end{pmatrix}$$

and

$$\sigma^2 := \nabla g(\mu)^T \Sigma \nabla g(\mu) = \frac{1}{(\rho - \mu_X)^4} \left(\sigma_1^2 (\rho - \mu_2 - u_n)^2 + \sigma_2^2 (\mu_1 - u_n)^2 \right).$$

■

3.6.2 Proof of Theorem 3.2

Preuve. Let us write

$$\begin{aligned} \hat{\Phi}_n^{(2)}(u_n) - \Phi(u_n) &= \left(\hat{p}_n \frac{\hat{\mu}_{Y,n}}{(\rho - \hat{\mu}_{X,n}^J)} - p_n \frac{\mu_Y}{(\rho - \mu_X)} \right) \\ &= \frac{(\hat{p}_n \hat{\mu}_{Y,n} - p_n \mu_Y)(\rho - \mu_X) + (\hat{\mu}_{X,n}^J - \mu_X) p_n \mu_Y}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} \end{aligned}$$

If we note $A_n = (\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)(\hat{\Phi}(u_n) - \Phi(u_n))$.

For GPD's approximation, Johansson has proposed (see, [Johansson, 2003](#)) the estimators of $\hat{\mu}_{X,n}^J$ and $\hat{\mu}_Y$ as follows

$$\hat{\mu}_{X,n}^J = \hat{\mu}_n^* + \hat{\kappa} = \hat{\mu}_n^* + \hat{p}_n \left(u_n + \frac{\hat{\beta}}{1 - \hat{\xi}} \right)$$

and $\hat{\mu}_Y = \frac{\hat{\beta}}{1 - \hat{\xi}}$ (Y has a GPD distribution). This expression may be rewritten as follows

$$\begin{aligned} A_n &= \left(\frac{\hat{p}_n \hat{\beta}_n}{(1 - \hat{\xi}_n)} - \frac{p_n \beta}{1 - \xi} \right) \left(\rho - \mu - p_n \left(u_n + \frac{\beta}{1 - \xi} \right) \right) \\ &\quad + \left(\hat{\mu} + \hat{p}_n \left(u_n + \frac{\hat{\beta}_n}{(1 - \hat{\xi}_n)} \right) - \mu - p_n \left(u_n + \frac{\beta}{1 - \xi} \right) \right) \frac{p_n \beta}{1 - \xi} \end{aligned}$$

After rearrangement, we can write

$$A_n = \left(\frac{\hat{p}_n \hat{\beta}_n}{(1 - \hat{\xi}_n)} - \frac{p_n \hat{\beta}_n}{(1 - \hat{\xi}_n)} + \frac{p_n \hat{\beta}_n}{(1 - \hat{\xi}_n)} - \frac{p_n \beta}{1 - \xi} \right) \left(\rho - \mu_n - p_n \left(u_n + \frac{\beta}{1 - \xi} \right) \right) \\ + \left((\hat{\mu}_n - \mu_n) + (\hat{p}_n - p_n) \left(u_n + \frac{\hat{\beta}_n}{(1 - \hat{\xi}_n)} \right) + p_n \left(\frac{\hat{\beta}_n}{(1 - \hat{\xi}_n)} - \frac{\beta}{1 - \xi} \right) \right) \frac{p_n \beta}{1 - \xi}$$

Using the approximations of [Johansson \(2003\)](#), we can deduce

$$A_n = \frac{p_n \beta}{1 - \xi} (\hat{\mu}_n - \mu_n) + \frac{\beta}{1 - \xi} (\rho - \mu_n) (\hat{p}_n - p_n) \\ + \frac{p_n (\rho - \mu_n - p_n u_n)}{(1 - \xi)} (\hat{\beta}_n - \beta) \\ + \frac{p_n \beta (\rho - \mu_n - u_n p_n)}{(1 - \xi)^2} (\hat{\xi}_n - \xi) + o_{\mathbb{P}}(1) \quad (3.54)$$

then, we can rewrite

$$\hat{\Phi}_n^{(2)}(u_n) - \Phi(u_n) = \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (A_{n,1} + A_{n,2} + A_{n,3} + A_{n,4}) + o_{\mathbb{P}}(1)$$

where

$$A_{n,1} = \frac{p_n \beta}{1 - \xi} \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{\mu}_n - \mu_n), \\ A_{n,2} = \frac{\beta}{1 - \xi} (\rho - \mu_n) \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{p}_n - p_n) \\ A_{n,3} = \frac{p_n (\rho - \mu_n - p_n u_n)}{(1 - \xi)} \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{\beta}_n - \beta) \\ A_{n,4} = \frac{p_n \beta (\rho - \mu_n - u_n p_n)}{(1 - \xi)^2} \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{\xi}_n - \xi)$$

We need the following proposition to complete the demonstration of the Theorem [3.2](#).

Proposition 3.1 *Let F_X be a df fulfilling (3.36) with $\xi \in (0, 1)$, $\delta > 0$ and real c . suppose that \mathbb{L} is locally bounded in $[x_F, \infty)$ for somme $x_F \geq 0$. Then fr large n enough and for any $u_n = O(n^{\alpha\xi})$, $\alpha \in (0, 1)$, we have*

$$\begin{aligned}\sigma_n^2 &= O_{\mathbb{P}}(n^{\alpha(2\xi-1)}), \\ (\hat{\mu}_n - \mu_n) &= O_{\mathbb{P}}\left(\frac{\sigma_n}{\sqrt{n}}\right), \\ (\hat{p}_n - p_n) &= O_{\mathbb{P}}\left(\frac{\sqrt{p_n(1-p_n)}}{\sqrt{n}}\right), \\ (\hat{\beta}_n - \beta) &= O_{\mathbb{P}}\left(\frac{1}{\sqrt{p_n n}}\right), \\ (\hat{\xi}_n - \xi) &= O_{\mathbb{P}}\left(\frac{1}{\sqrt{p_n n}}\right)\end{aligned}$$

Preuve. For the proof of the proposition (see, e.g., [Johansson \(2003\)](#)). We can deduce that

$$\begin{aligned}(\hat{\mu}_n - \mu_n) &= o_{\mathbb{P}}(1), \text{ as } n \uparrow +\infty, \\ (\hat{p}_n - p_n) &= o_{\mathbb{P}}(1), \text{ as } n \uparrow +\infty, \\ (\hat{\beta}_n - \beta) &= o_{\mathbb{P}}(1), \text{ as } n \uparrow +\infty, \\ (\hat{\xi}_n - \xi) &= o_{\mathbb{P}}(1), \text{ as } n \uparrow +\infty.\end{aligned}$$

■ Under the assumption that $\rho > \mu_X$, It follows that

$$\begin{aligned}\frac{p_n \beta}{1 - \xi} \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{\mu}_n - \mu_n) &= o_{\mathbb{P}}(1) \text{ as } n \uparrow +\infty, \\ \frac{\beta}{1 - \xi} (\rho - \mu_n) \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{p}_n - p_n) &= o_{\mathbb{P}}(1) \text{ as } n \uparrow +\infty, \\ \frac{p_n(\rho - \mu_n - p_n u_n)}{(1 - \xi)} \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{\beta}_n - \beta) &= o_{\mathbb{P}}(1) \text{ as } n \uparrow +\infty,\end{aligned}$$

and

$$\frac{p_n \beta (\rho - \mu_n - u_n p_n)}{(1 - \xi)^2} \frac{1}{(\rho - \hat{\mu}_{X,n}^J)(\rho - \mu_X)} (\hat{\xi}_n - \xi) = o_{\mathbb{P}}(1) \text{ as } n \uparrow +\infty$$

Hence, we have

$$\hat{\Phi}_n^{(2)}(u_n) - \Phi(u_n) = o_{\mathbb{P}}(1)$$

This completes the proof of Theorem 3.2. ■

Proof of Theorem (3.3)

The Lemma A.2 in Johansson (2003) gives that, the estimators conditional on N , $\hat{\mu}_n$ is independent of $(\hat{\beta}_n, \hat{\xi}_n)$. And if we use that

$$\sqrt{np_n}Q^{\frac{1}{2}} \begin{pmatrix} \hat{\beta}_n - \beta \\ \hat{\xi}_n - \xi \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, I), \text{ as } n \uparrow +\infty,$$

where

$$Q^{-1} = (1 + \xi) \begin{pmatrix} 2\beta^2 & -\beta \\ -\beta & 1 + \xi \end{pmatrix},$$

$$\frac{\sqrt{n}}{\sqrt{p_n(1-p_n)}}(\hat{p}_n - p_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \text{ as } n \uparrow +\infty,$$

and

$$\frac{\sqrt{n}}{\sigma_n}(\hat{\mu}_n - \mu_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \text{ as } n \uparrow +\infty.$$

Using the result (3.54) in the proof of the Theorem 2 and if we note $M = \frac{1}{(\rho - \hat{\mu}_{X,n}^I)(\rho - \mu_X)}$,

we can write

$$\begin{aligned} \sqrt{n} \left(\hat{\Phi}_n^{(2)}(u_n) - \Phi(u_n) \right) &= M \left[\frac{\sigma_n \sqrt{n} p_n \beta}{\sigma_n (1 - \xi)} (\hat{\mu}_n - \mu_n) + \frac{\sqrt{n} \sqrt{p_n(1-p_n)} \beta}{(1 - \xi) \sqrt{p_n(1-p_n)}} (\rho - \mu_n) (\hat{p}_n - p_n) \right. \\ &+ \frac{\sqrt{n p_n} \sqrt{p_n}}{(1 - \xi)} (\rho - \mu_n - p_n u_n) (\hat{\beta}_n - \beta) \\ &\left. + \frac{\sqrt{n p_n} \sqrt{p_n} \beta}{(1 - \xi)^2} (\rho - \mu_n u_n p_n) (\hat{\xi}_n - \xi) \right] + o_{\mathbb{P}}(1) \end{aligned} \quad (3.55)$$

By the law of large numbers the random variable $\frac{1}{(\rho - \hat{\mu}_{X,n}^I)(\rho - \mu_X)}$ converge in probability to $\frac{1}{(\rho - \mu_X)^2}$. Also, by Slutsky's Theorem, we can justify the asymptotic normality of $\sqrt{n} \left(\hat{\Phi}_n^{(2)}(u_n) - \Phi(u_n) \right)$. We can deduce the asymptotic variance as follows

$$\begin{aligned} \sigma^2 &= \frac{\beta^2}{(1 - \xi)^2 (\rho - \mu_X)^4} \left[p_n^2 \sigma_n^2 + p_n (1 - p_n) (\rho - u_n)^2 + 2 p_n (1 + \xi) (\rho - \mu_n - p_n u_n)^2 \right. \\ &\left. + \frac{p_n (1 + \xi)^2}{(1 - \xi)^2} (\rho - \mu_n - p_n u_n)^2 - \frac{2 p_n (1 + \xi)}{(1 - \xi)} (\rho - \mu_n - p_n u_n)^2 \right] \end{aligned} \quad (3.56)$$

If we note $\theta_1 = \frac{\beta^2}{(1-\xi)^2(\rho-\mu_X)^4}$, $\theta_2 = (\rho - u_n)^2$, $\theta_3 = (\rho - \mu_n - p_n u_n)^2$,
we can rewrite σ^2 as follows

$$\sigma^2 = \theta_1 \left(p_n^2 \sigma_n^2 + p_n(1-p_n)\theta_2 + \frac{p_n(2\xi^2 - \xi + 1)(1 + \xi)}{(1-\xi)^2} \theta_3 \right) = O_+(1)$$

CHAPITRE 4

PROBABILITÉ DE RUINE POUR UN PROCESSUS DU RISQUE NON-STATIONNAIRE

Dans ce chapitre, nous étudions la probabilité de ruine asymptotique à horizon fini et infini pour les processus de risques à queue lourdes (donc sous-exponentielles) où les sinistres arrivent suivant un processus non stationnaire en utilisant le principe de la grande déviation.

Nous donnons deux exemples de processus des arrivées non renouvelables, à savoir, le processus de Hawkes et le processus de Cox avec intensité de bruit de tir. Nous proposons un nouvel estimateur de la probabilité de ruine pour des remboursements à queues lourdes et qui arrivent suivant un processus de Hawkes ou un processus de Cox en appliquant la méthode POT.

4.1 Processus de risque avec arrivées Non-Stationnaire

Soit le modèle de risque classique de [Lundberg \(1903\)](#), le processus de risque classique $(R_t)_{t \geq 0}$ pour $t \geq 0$

$$R_t = u + pt - S_t,$$

où u est le montant de la réserve initiale, $p > 0$ est le taux de revenu des primes. Le montant cumulé des sinistres jusqu'au temps t est décrit par le processus de renouvellement composé

$$S_t = \sum_{i=1}^{N_t} X_i,$$

On suppose que X_1, X_2, \dots sont *i.i.d* de cdf F à queue lourde et $(N_t)_t$ est le nombre des sinistres qui est un processus non stationnaire.

4.1.1 Notion du principe de grandes déviations

Les premiers résultats rigoureux concernant le principe des grandes déviations sont dus au mathématicien [Cramér \(1930\)](#), qui les a appliqués pour modéliser les problèmes d'assurance. Du point de vue des compagnies d'assurances, les revenus sont à taux constant par mois (les mensualités), mais les dépenses sont aléatoires.

Le principe de grandes déviations, en théorie des probabilités, concerne le comportement asymptotique de queues de suite de loi de probabilités. Quelques premières idées de la théorie ont été données par Laplace et Cramér.

La théorie des grandes déviations concerne la décroissance exponentielle des distributions de probabilité de certains types d'évènements extrêmes ou de queue, lorsque le nombre d'observations est arbitrairement grand.

Pour que la compagnie soit bénéficiaire après une durée de plusieurs mois, la somme totale de revenus doit être supérieure aux dépenses totales. Ainsi pour estimer les mensualités, il faut se demander : « Quelles mensualités q doit-on choisir pour que, après N mois, les dépenses totales $S_t = \sum_{i=1}^{N_t} X_i$, soient inférieures à Nq ? » Ce qui est clairement la même question posée par la théorie des grandes déviations. Cramer a donné une solution à cette question pour des variables aléatoires *i.i.d* avec une fonction de taux sous forme de série entière.

Définition 4.1 Une suite $(P_n)_{n \in \mathbb{N}}$ des mesures de probabilité sur E satisfait le principe de grande déviation (ou Large Deviation Principle (LDP)). avec une fonction de taux $J : E \rightarrow \mathbb{R}$ où J est non négative, et continue à gauche et pour tout ensemble mesurable A , nous avons :

$$-\inf_{x \in A^\circ} J(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log P_n(A) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_n(A) \leq -\inf_{x \in \bar{A}} J(x) \quad (4.1)$$

avec A° est l'intérieur de A et \bar{A} est sa fermeture.

A titre d'exemple, considérons le cas où \bar{X}_n est la moyenne de n variables indépendantes identiquement distribuées de loi gaussienne centrée réduite. La variable \bar{X}_n suit une loi normale d'espérance 0 et de variance $1/n$, si bien que pour tout intervalle A .

$$\mathbb{P}(\sqrt{n}\bar{X}_n \in A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-x^2/2} dx.$$

La valeur typique de \bar{X}_n est de l'ordre de $1/\sqrt{n}$. Soit $\delta > 0$. La probabilité de l'événement

$\{|\bar{X}_n| \geq \delta\}$ tend vers 0, et plus précisément

$$\mathbb{P}\left(|\bar{X}_n| \geq \delta\right) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} e^{-x^2/2} dx,$$

si bien que

$$\frac{1}{n} \log \mathbb{P}\left(|\bar{X}_n| \geq \delta\right) \xrightarrow{n \rightarrow +\infty} -\frac{\delta^2}{2}$$

Ainsi, avec une petite probabilité (de l'ordre de $e^{-n\delta^2/2}$), $|\bar{X}_n|$ dévie de son comportement typique en prenant des grandes valeurs. Cela explique la terminologie "grandes déviations". L'hypothèse suivante est l'hypothèse principale de ce chapitre.

Hypothèse 4.1 (i) $(N_t)_t$ satisfait le principe de grande déviation avec fonction de taux J de telle sorte que $J(x) = 0$ si et seulement si $x = \lambda$.

(ii) J est croissante sur $[\lambda, \infty)$ et décroissante sur $[0, \lambda]$.

(iii) Si $T_i = \tau_i - \tau_{i-1}$ est la longueur de l'intervalle de temps entre deux temps d'arrivée consécutifs du processus ponctuel τ_{i-1} et τ_i où τ_i représente le i ème temps d'arrivée du processus ponctuel. Si T_i sont i.i.d., avec une moyenne $E[T_1]$, alors $(N_t)_t$ est un processus de renouvellement et on suppose la condition habituelle de bénéfice net

$$\rho = \frac{\mu_X}{pE(T_1)} < 1 \quad (4.2)$$

(iv) Il existe $\theta > 0$ de telle sorte que $E[\exp(\theta \sum_{i=1}^n T_i)] < \infty$ pour tout $n \in \mathbb{N}$.

Sous l'hypothèse précédent, les deux lemmes suivantes sont applicables.

Lemme 4.1 *Zhu (2013)*

Sous l'hypothèse (4.2) pour tout $\epsilon, \epsilon' > 0$, il existe une constante $M > 0$ telle que :

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \left\{ p \sum_{i=1}^n T_i \leq n \left(\frac{p}{\lambda} + \epsilon \right) + M \right\}\right) > 1 - \epsilon' \quad (4.3)$$

Preuve. En remplaçant ϵ par p et M par pM dans l'équation ci-dessus, on peut prouver que

$$\limsup_{M \rightarrow \infty} \mathbb{P}\left(\bigcup_{n=1}^{\infty} \left\{ \sum_{i=1}^n T_i > n \left(\frac{1}{\lambda} + \epsilon \right) + M \right\}\right) = 0..$$

On peut voir que $\{N_t \leq n\} = \{\sum_{i=1}^n T_i > t\}$ pour tout $n \in \mathbb{N}$ et $t \in \mathbb{R}^+$ et aussi pour tout $\lambda' < \lambda$ fixe, il existe un certain $\delta' > 0$ tel que $J(\lambda') - \delta' > 0$ et pour t assez grand.

$$\begin{aligned} & \limsup_{M \rightarrow \infty} \left(\bigcup_{n=1}^{\infty} \sum_{i=1}^n T_i > n \left(\frac{1}{\lambda} + \epsilon \right) + M \right) \\ & \leq \limsup_{M \rightarrow \infty} \sum_{n=1}^{\infty} \mathbb{P} \left(\sum_{i=1}^n T_i > n \left(\frac{1}{\lambda} + \epsilon \right) + M \right) \\ & \leq \limsup_{M \rightarrow \infty} \sum_{n \geq N} \exp \left(- \left(n \left(\lambda^{-1} + \epsilon \right) + M \right) \left[1 - \left(\frac{\lambda}{\lambda + \epsilon} \right) - \delta' \right] \right) = 0 \end{aligned}$$

■

Lemme 4.2 *Sous l'hypothèse (4.2) et supposons en outre que $F_0 \in \mathcal{S}$ (i.e., F_0 sous exponentielle),*

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}(\sup_{n \geq 1} \{n(p\lambda^{-1} - \epsilon) - p \sum_{i=1}^n T_i\} \geq u)}{\bar{F}_0} = 0. \quad (4.4)$$

pour tout $\epsilon > 0$ assez petit.

Où \bar{F}_0 est la distribution de la queue intégrée .

Preuve. Voir [Zhu \(2013\)](#) ■

4.1.2 Probabilité de ruine

[Asmussen et al. \(1999\)](#) ont prouvé que (3.37) est vrai si les deux lemmes précédents sont vérifiés. Notre tâche principale ici est donc de prouver le lemme (4.1) et le lemme (4.2) sous les hypothèses précédentes.

Notez que le lemme (4.1) est vrai si $(T_i)_{i \geq 1}$ est stationnaire et ergodique.

Et c'est le seul endroit où [Asmussen et al. \(1999\)](#) ont utilisé l'hypothèse de stationnarité et d'ergodicité. C'est pourquoi tant que nous pouvons prouver le lemme (4.1), nous pouvons abandonner l'hypothèse de stationnarité et d'ergodicité.

Dans la suite, nous donnons les résultats asymptotiques des probabilités de ruine dans le cas où les deux lemmes (4.1 et 4.2) sont vérifiés.

Théorème 4.1 [Zhu\(2013\)](#)

Sous l'hypothèse (4.2) et supposons que $F_0 \in \mathcal{S}$, nous aurons

$$\lim_{u \rightarrow \infty} \frac{\Phi(u)}{\bar{F}_0(u)} = \frac{\rho}{1 - \rho} \quad (4.5)$$

Remarque 4.1 *Dans le théorème (4.1), nous pouvons remplacer l'hypothèse de grande déviation de $(N_t)_t$ par une hypothèse de grande déviation de $(\frac{1}{n} \sum_{i=1}^n T_i)$.*

Mais généralement si $(N_t)_t$ satisfait le principe de grande déviation avec la fonction de taux $J(x)$ si et seulement si $(\frac{1}{n} \sum_{i=1}^n T_i)$ satisfait le principe de grande déviation avec la fonction de taux $xJ(\frac{1}{x})$.

C'est pourquoi il faut supposer la grande déviation de $(N_t)_t$ dans l'hypothèse (4.2) est que lorsque $(N_t)_t$ n'est pas un processus de renouvellement (non stationnaire), les temps inter-arrivées ne sont pas *i.i.d* et il est généralement plus facile et plus naturel d'établir la grande déviation pour $(N_t)_t$, ce qui vérifiée par deux exemples, processus Hawkes, et processus Cox avec intensité du bruit de tir.

Ensuite, considérons les probabilité de ruine à horizon fini.

Si $e(u) = E[X_1 - u | X_1 > u]$ est la fonction de moyenne en excès et $\Phi(u, T) = P(\tau(u) \leq T)$, $T > 0$, la probabilité de ruine à l'horizon fini.

Exemple 4.1 (i) (Distribution à variation régulière) si $\bar{F}(u) = \frac{L(u)}{u^{\alpha+1}}$, $\alpha \in (0, \infty)$, i.e., $\bar{F} \in \mathcal{R}_{(-\alpha-1)}$, alors $e(u) \sim \frac{u}{\alpha}$.

(ii) (Distribution de Lognormal) si $F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{(\log u - \mu)}{\sigma}} \exp(-x^2/2) dx$, alors $F \in \Lambda_0$ et $e(u) \sim \frac{\sigma^2 u}{\log u - \mu}$.

(iii) Distribution de Weibull si $F(u) = \exp(-u^\alpha)$, où $\alpha \in (0, 1)$, alors $F \in \Lambda_0$ et $F_0 \in \mathcal{S}$ et $e(u) \sim \frac{u^{1-\alpha}}{\alpha}$.

Remarque 4.2 Il est bien connu que si $F \in \Lambda_0$, (i.e le domaine maximal d'attraction de la distribution de Gumbel), alors :

$$\lim_{u \rightarrow \infty} \frac{\bar{F}(u + xe(u))}{\bar{F}(u)} = \exp(-x), \quad x \in \mathbb{R} \quad (4.6)$$

Lemme 4.3 Pour tout $y_0 < \infty$, $\lim_{x \rightarrow \infty} \frac{\bar{G}(x+y)}{\bar{G}(x)} = 1$ uniformément pour $y \in [0, y_0]$ pour tout $G \in \mathcal{S}$.

Théorème 4.2 Sous l'hypothèse (4.2) supposons que $F_0 \in \mathcal{S}$, nous avons pour tout $T > 0$

(i) Si $\bar{F} \in \mathcal{R}_{(-\alpha-1)}$

$$\lim_{u \rightarrow \infty} \frac{\Phi(u, e(u)T)}{\bar{F}_0(u)} = \frac{\rho}{1-\rho} \left[1 - \left(1 + (1-\rho) \frac{T}{\alpha} \right)^{-\alpha} \right] \quad (4.7)$$

(ii) Si $F \in \Lambda_0$

$$\lim_{u \rightarrow \infty} \frac{\Phi(u, e(u)T)}{\bar{F}_0(u)} = \frac{\rho}{1-\rho} [1 - \exp(-(1-\rho)T)] \quad (4.8)$$

4.2 Réclamation globale

Soit $S_t = \sum_{i=1}^{N_t} X_i$ la réclamation globale jusqu'au temps t , où comme précédemment nous supposons que X_i sont v.a positive *i.i.d* considérons les hypothèses suivantes.

Hypothèse 4.2 (i) $E[N_t] < \infty$ pour toute t et $E[N_t] \rightarrow \infty$ quand $t \rightarrow \infty$

(ii) $\frac{N_t}{E[N_t]} \rightarrow 1$, quand $t \rightarrow \infty$

(iii) Il existe $\epsilon, \sigma > 0$ tel que

$$\sum_{k > (1+\sigma)E[N_t]} P(N_t > k)(1 + \epsilon)^k \rightarrow 0, \text{ quand } t \rightarrow \infty \quad (4.9)$$

[Klüppelberg and Stadtmüller \(1998\)](#) ont montré dans l'hypothèse (4.2) pour un temps fixe t , on a

$$\mathbb{P}(S_t - E[S_t] > x) \sim E[N_t]P(X_1 \geq x) \quad (4.10)$$

de manière uniforme pour $x \geq \gamma E[N_t]$ et pour tout $\gamma > 0$.

Remarque 4.3 [Klüppelberg and Stadtmüller \(1998\)](#) ont montré un résultat légèrement plus fort que (4.10) et elle est valable en supposant que les montants de réclamations de sinistre X_i , sont *i.i.d* de fonction de distribution $\bar{F} \in \mathcal{R}_{(-\alpha, -\beta)}$ avec $1 < \alpha \leq \beta \leq \infty$.

Il est généralement facile de vérifier (i) et aussi sous les hypothèse du théorème (4.1), $\frac{N_t}{t} \rightarrow \lambda$ où $(N_t)_t$ satisfait le principe de la grande déviation avec la fonction de taux $J(x)$ qui est non nulle pour $x \neq \lambda$.

Par conséquent, nous pouvons prouver que $\frac{E[N_t]}{t} \rightarrow \lambda$ quand $t \rightarrow \infty$, alors (ii) est satisfait.

De plus (iii) peut être remplacé par (iii'), donné par :

pour tout $\lambda' > 0$, $c_{\mu'} = \inf_{x \geq \mu'} \frac{J(x)}{x} > 0$.

4.3 Exemples de processus d'arrivée non renouvelable

Nous supposons également que N_t commence avec une histoire passée vide, c'est-à-dire $N(-\infty, 0] = 0$. Selon la définition, le processus Hawkes est non stationnaire et est en général même non markovien (sauf si $h(\cdot)$ est une fonction exponentielle). En outre, il n'a pas de structure générative. Ainsi, les conditions énoncées dans [Asmussen and Albrecher \(2010\)](#) ne s'appliquent pas ici.

4.3.1 Processus de Hawkes

Le processus de Hawkes est un processus ponctuel simple qui a une propriété auto-excitante, en effet de regroupement et une longue mémoire (voir le paragraphe (1.4.6.1) du chapitre (1)). Il a été introduit pour la première fois par [Hawkes \(1971\)](#) et a été largement appliqué dans la finance, la sismologie, les neurosciences, la modélisation de l'ADN et bien d'autres domaines. Un processus ponctuel simple N_t est un processus linéaire de Hawkes s'il a une intensité [Hawkes \(1971\)](#).

$$\lambda_t = \nu + \sum_{\tau < t} h(t - \tau) \quad (4.11)$$

$h(\cdot) : [0, \infty) \rightarrow (0, \infty)$ est intégrable et $\|h\|_{L^1} < 1$. Nous supposons également que N_t commence avec une histoire passée vide, c'est-à-dire $N(-\infty, 0] = 0$. Selon la définition, le processus Hawkes est non stationnaire et est en général même non markovien (sauf si $h(\cdot)$ est une fonction exponentielle). En outre, il n'a pas de structure générative. Ainsi, les conditions énoncées dans [Asmussen and Albrecher \(2010\)](#) ne s'appliquent pas ici.

$$\lim_{t \rightarrow \infty} \frac{N_t}{t} = \lambda = \frac{\nu}{1 - \|h\|_{L^1}}, \quad (4.12)$$

[Bordenave and Torrisi \(2007\)](#) ont prouvé le principe de la grande déviation pour (N_t) , c'est-à-dire, le Lemme (4.1) est vérifié.

Par conséquent il est donc naturel que nous puissions étudier les probabilités de ruine avec des lois sous-exponentielles lorsque le processus d'arrivée est un processus de Hawkes linéaire non stationnaire.

Lemme 4.4 [Bordenave and Torrisi \(2007\)](#)

(N_t) satisfait au principe de grande déviation avec fonction de taux,

$$J(x) = \begin{cases} x \log \left(\frac{x}{\nu + x \|h\|_{L^1}} \right) - x + x \|h\|_{L^1} + \nu & \text{si } x \in [0, \infty) \\ +\infty & \text{sinon.} \end{cases} \quad (4.13)$$

Remarque 4.4 En effet, dans [Bordenave and Torrisi \(2007\)](#), ils ont exprimé la fonction de taux I d'une autre manière, moins explicite. L'expression de la fonction de taux dans la lemme (4.1) a été donné pour la première fois dans [Zhu \(2013\)](#).

Lemme 4.5 .

$$\frac{E(N_t)}{t} \rightarrow \frac{\nu}{1 - \|h\|_{L^1}} \text{ avec } t \rightarrow \infty.$$

Supposons la condition de bénéfice net $p > \frac{\nu \mu_X}{1 - \|h\|_{L^1}}$.

Si X_i , ont des queues légères, alors [Stabile and Torrisi\(2010\)](#) ont obtenu les asymptotiques

pour la probabilité de ruine à horizon infini $\Phi(u)$ et la probabilité de ruine à horizon fini $\Phi(u, z)$ pour tout $z > 0$. Comme indiqué dans [Stabile and Torrisi2010](#) le cas où X_i sont à queue lourde est ouvert.

Proposition 4.1 Si la condition de bénéfice net $p > \mu_X \frac{\nu}{1 - \|h\|_{L^1}}$.

— (Horizon Infini)

$$\lim_{u \rightarrow \infty} \frac{\Phi(u)}{\bar{F}_0} = \frac{\nu \mu_X}{p(1 - \|h\|_{L^1}) - \nu \mu_X} \quad (4.14)$$

— (Horizon-Fini) Pour tout $T > 0$,

$$\lim_{u \rightarrow \infty} \frac{\Phi(u, uz)}{\bar{F}_0(u)} = \begin{cases} \frac{\nu \mu_X}{p(1 - \|h\|_{L^1} - \nu \mu_X)} \left[1 - \left(1 + \left(\frac{p(1 - \|h\|_{L^1}) - \nu \mu_X}{c(1 - \|h\|_{L^1})} \right) \frac{T}{\alpha} \right)^{-\alpha} \right] & \text{si } \bar{F} \in \mathcal{R}_{(-\alpha-1)} \\ \frac{\nu \mu_X}{p(1 - \|h\|_{L^1} - \nu \mu_X)} \left[1 - e^{-\frac{c(1 - \|h\|_{L^1}) - \nu \mu_X}{p(1 - \|h\|_{L^1})} T} \right] & \text{si } F \in \Lambda_0 \end{cases} \quad (4.15)$$

— (Réclamations globales). Pour le temps fixe t ,

$$\mathbb{P}(S_t - E(S_t) > 0) \sim E(N_1) \mathbb{P}(X_1 \geq x), \quad (4.16)$$

de manière uniforme pour $x \geq \gamma E(N_t)$ pour tout $\gamma > 0$.

Preuve. Voir [Zhu \(2013\)](#) ■

4.3.2 Processus de Cox avec intensité de bruit de tir

Nous considérons un processus Cox $(N_t)_t$ avec intensité λ_t qui suit un processus de bruit de tir

$$\lambda_t = \nu(t) + \sum_{\tau^{(1)} < t} g(t - \tau^{(1)}), \quad (4.17)$$

où $\tau^{(1)}$ le temps d'arrivée d'un processus de Poisson $(N_t^{(1)})_t$ externe homogène avec intensité γ .

Ici, $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ est intégrable, c'est-à-dire $\int_0^\infty g(t) dt < \infty$ et $\nu(t)$ est une fonction positive, continue et déterministe telle que $\nu(t) \rightarrow \nu$ avec $t \rightarrow \infty$. $\nu(t)$ peut représenter d'autres perturbations qui n'ont pas d'importance symptomiquement (voir ((1.4.6.2) du chapitre (1)).

La probabilité de ruine pour les demandes d'indemnisation à queue lourde, le processus d'arrivée étant un processus de Cox à bruit de tir, sont connues dans la littérature, voir par

exemple le livre [Asmussen and Albrecher \(2010\)](#). Mais les techniques décrites dans la littérature utilisent les caractéristiques très spécifiques du processus de Cox .

Comme $(N_t^{(1)})_t$ est un processus de Poisson avec intensité γ , selon la définition de λ_t , il est facile de voir que

$$\frac{N_t}{t} \rightarrow \nu + \gamma \|g\|_{L^1}, \text{ avec } t \rightarrow \infty. \quad (4.18)$$

Lemme 4.6 $(N_t)_t$ satisfait à un principe de grande déviation avec fonction de taux,

$$J(x) = \begin{cases} \sup_{\theta \in \mathbb{R}} \{ \theta x (e^\theta - 1) \nu - \gamma (e^{(\theta-1)} \|g\|_{L^1} - 1) \} & \text{si } x \in [0, \infty) \\ +\infty & \text{sinon.} \end{cases} \quad (4.19)$$

Lemme 4.7 $\frac{E(N_t)}{t} \rightarrow \nu + \gamma \|g\|_{L^1}$, quand $t \rightarrow \infty$.

Proposition 4.2 supposons la condition de bénéfice net $p > E[X](\nu + \gamma \|g\|_{L^1})$.

— (Horizon infini)

$$\lim_{u \rightarrow \infty} \frac{\Phi(u)}{\bar{F}_0(u)} = \frac{(\nu + \gamma \|g\|_{L^1}) \mu_X}{p - (\nu + \gamma \|g\|_{L^1}) \mu_X}. \quad (4.20)$$

— (Horizon Fini) Pour tout $T > 0$,

$$\lim_{u \rightarrow \infty} \frac{\Phi(u, uz)}{\bar{F}_0(u)} = \begin{cases} \frac{(\nu + \gamma \|g\|_{L^1}) \mu_X}{p - (\nu + \gamma \|g\|_{L^1}) \mu_X} \left[1 - \left(1 + \left(1 - \frac{(\nu + \gamma \|g\|_{L^1}) \mu_X}{p} \right)^{\frac{T}{\alpha}} \right)^{-\alpha} \right] & \text{si } \bar{F} \in \mathcal{R}_{(-\alpha-1)} \\ \frac{(\nu + \gamma \|g\|_{L^1}) \mu_X}{p - (\nu + \gamma \|g\|_{L^1}) \mu_X} \left[1 - e^{-(p - (\nu + \gamma \|g\|_{L^1}) \mu_X) \frac{T}{p}} \right] & \text{si } B \in \\ \text{Lambda.} & \end{cases} \quad (4.21)$$

— (Réclamations globales) Pour un temps t fixe,

$$P(S_t - E(S_t) > x) \sim E[N_t] \mathbb{P}(X_1 \geq x), \quad (4.22)$$

uniformément pour $x \geq \gamma E(N_t)$ pour tout $\gamma > 0$.

Preuve. Voir [Zhu \(2013\)](#) ■

4.4 Estimation de la probabilité de ruine par la méthode POT

Zhu (2013) a montré que la probabilité de ruine pour des sinistres qui arrivent suivant un processus de Hawkes ou un processus de Cox est donnée comme suit.

Proposition 4.1 1. Si $(N_t)_t$ est un processus de Hawkes sur $(0, \infty)$ (respectivement un processus de Cox) et suppose la condition de bénéfice net $\rho_1 = \frac{p(1-\kappa)}{v} > \mu_X$ (respectivement $\rho_2 = \frac{p}{(v+\gamma\sigma)} > \mu_X$).

2.

$$\lim_{u \rightarrow +\infty} \frac{\Phi(u)}{\bar{F}_0(u)} = \frac{\mu_X}{\rho_i - \mu_X} \quad 1 \leq i \leq 2 \quad (4.23)$$

Preuve. Voir *Zhu (2013)* ■

La proposition suivante donne l'approximation de la probabilité de ruine $\Phi(u)$ pour une grande réserve initiale u .

Corollaire 4.1 Si $(N_t)_t$ est un processus de Hawkes sur $(0, \infty)$ (respectivement un processus de Cox) et suppose la condition de bénéfice net $\rho_1 = \frac{p(1-\kappa)}{v} > \mu_X$ (respectivement $\rho_2 = \frac{p}{(v+\gamma\sigma)} > \lambda$), la probabilité de ruine est donnée par.

$$\Phi(u) = \frac{1}{\rho_i - \mu_X} \int_u^{+\infty} \bar{F}(x) dx \quad 1 \leq i \leq 2 \quad (4.24)$$

Preuve.

$$\lim_{u \rightarrow +\infty} \frac{\Phi(u)}{\bar{F}_0(u)} = \frac{\mu_X}{\rho_i - \mu_X} \quad 1 \leq i \leq 2 \quad (4.25)$$

Alors, pour une grande réserve initiale u , c'est-à-dire qu'il existe u_0 , que pour tout $u > u_0$, on peut donc écrire

$$\Phi(u) = \frac{\mu_X}{\rho_i - \mu_X} \bar{F}_0(u)$$

Or \bar{F}_0 est la distribution de la queue intégrée donnée par

$$\bar{F}_0(u) = \frac{1}{\mu_X} \int_u^{+\infty} \bar{F}(x) dx$$

Alors

$$\Phi(u) = \frac{1}{\rho_i - \mu_X} \int_u^{+\infty} \bar{F}(x) dx \quad 1 \leq i \leq 2$$

■

Dans la section suivante, nous allons présenter le nouvel estimateur de la probabilité de ruine dans le cas où les sinistres arrivent suivant le processus de Hawkes ou processus de Cox avec des réclamations à queue lourde en temps infini pour une large réserve initiale u .

Donc tous les résultats utilisés dans le calcul de la probabilité de ruine dans le cas stationnaire sont repris dans le cas non stationnaires vu l'expression de la probabilité déduite dans la formule (4.24). Nous utilisons en particuliers la convergence normale de la moyenne de l'échantillon des réclamations vers la loi normale en utilisant l'approximation de la partie au dessus du seuil u par la loi GPD basée sur les résultats de [Johansson \(2003\)](#). La construction de l'estimateur de la probabilité de ruine est basée sur les résultats proposés par [Zhu \(2013\)](#).

4.4.1 Construction d'un nouvel estimateur de la probabilité de ruine

Si X_1, X_2, \dots, X_n sont des rv positifs i.i.d.. avec cdf F satisfaisant l'hypothèse (3.36).

Si $Y_i = X_i - u_n$, pour tout i tel que $X_i > u_n$ être des rv positifs i.i.d.. avec cdf F_Y .

Par définition, c'est le nombre de ces X_i qui dépassent u_n . Nous avons vu que

$$N := N_{u_n} = \sum_{i=1}^n \mathbf{1}_{\{X_i > u_n\}},$$

suit la distribution binomiale $\text{Bin}(p_n, n)$ avec le paramètre $p_n = \mathbf{P}[X_1 > u_n]$, qui est égal à $\bar{F}(u_n)$. Nous avons vu aussi que l'estimateur suivant de $\bar{F}(u_n + y)$:

$$\widehat{\bar{F}}(u_n + y) = \widehat{\bar{F}}(u_n) \widehat{\bar{F}}_{u_n}(y) = \widehat{p}_n \bar{G}_{\widehat{\xi}_n, \widehat{\beta}_n}(y). \quad (4.26)$$

En outre, si $\mu_n^* = E(X_1 \mathbf{1}_{\{X_1 \leq u_n\}})$ avec son estimateur empirique $\widehat{\mu}_n^* = n^{-1} \sum_{i=1}^n X_i \mathbf{1}_{\{X_i \leq u_n\}}$ et $\gamma_n^2 = \text{Var}(X_1 \mathbf{1}_{\{X_1 \leq u_n\}})$.

Pour une grande valeur de u_n , nous utilisons l'approximation suivante $\widehat{\bar{F}}_Y(y) \approx \bar{G}$.

Si on pose $\tau_n = \int_{u_n}^{+\infty} \bar{F}(x) dx$, et avec l'utilisation de l'équation (4.26), nous pouvons écrire

$$\tau_n = \int_{u_n}^{+\infty} \bar{F}(x) dx = \int_0^{+\infty} \bar{F}(y + u_n) dy = p_n \int_0^{+\infty} \bar{F}_Y(y) dy.$$

Alors, la forme simplifiée de la probabilité de ruine est donnée par

$$\Phi(u) \simeq \frac{\tau_n}{(\rho_i - \mu)}, \quad i = 1, 2. \quad (4.27)$$

Pour de grandes valeurs de u_n , et sous l'hypothèse $\rho_i > \mu$, pour $1 \leq i \leq 2$.

L'approximation de la moyenne μ_X en utilisant la loi de GPD (voir [Johansson \(2003\)](#)), $n \geq 1$, nous a permis d'écrire

$$\mu_X = \int_0^{u_n} \bar{F}(x) dx + \int_{u_n}^{\infty} \bar{F}(x) dx = \mu_n^* + \kappa_n.$$

Par conséquent, l'estimateur de Johansson de μ est donné par

$$\widehat{\mu}_n^{(J)} = \widehat{\mu}_n^* + \widehat{\kappa}_n, \quad (4.28)$$

Où

$$\widehat{\mu}_n^* = \frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{\{X_i \leq u_n\}} \quad \text{et} \quad \widehat{\kappa}_n = \widehat{p}_n \left(u_n + \frac{\widehat{\beta}_n}{1 - \widehat{\xi}_n} \right),$$

avec \widehat{p}_n est donné par l'équation (??) comme estimateur de p_n et $\widehat{\xi}_n, \widehat{\beta}_n$ sont des estimateurs de ξ, β respectivement donnés dans l'équation (3.45). Il est naturel de définir un estimateur empirique de $\Phi_n(u_n)$ par l'expression suivante

$$\widehat{\Phi}_n(u_n) = \frac{\widehat{\tau}_n}{\left(\rho_i - \widehat{\mu}_n^{(J)} \right)}, \quad i = 1, 2. \quad (4.29)$$

Cet estimateur ainsi construit est bien défini puisque $\rho_i > \mu_X$ pour $1 \leq i \leq 2$.

Comme dans le cas stationnaire, si $\xi \in (0, 1/2)$, X_1 a une variance finie ($\sigma^2 < +\infty$) et on a

$$\widehat{\Phi}_n(u_n) = \widehat{p}_n \frac{\bar{Y}_n}{\left(\rho_i - \widehat{\mu}_n^{(J)} \right)}, \quad i = 1, 2. \quad (4.30)$$

Par contre si $\xi \in (1/2, 1)$, X_1 a une variance infinie ($\sigma^2 = +\infty$) et le théorème de la limite centrale n'est pas applicable.

4.5 Illustrations numériques

Nos résultats font référence à un processus de risque avec des arrivées de sinistres non stationnaires, qui est un processus de Hawkes ou un processus de Cox où les sinistres ont une distribution à queue lourde.

Dans cette section de simulation, nous supposons que la fonction de distribution des montants de remboursements est une distribution de Pareto $\bar{F}(x) = x^{-1/\xi}$, $x > 1$.

Notre étude de simulation se fait lorsque la variance des créances est finie, c'est-à-dire que le paramètre de forme ξ est dans l'intervalle $(0, 1/2)$, pour illustrer ce cas nous avons pris $\xi = 3/7$ et lorsque la variance est infinie, c'est-à-dire que le paramètre de forme ξ est dans l'intervalle $(1/2, 1)$, nous avons pris pour illustration $\xi = 2/3$ et $\xi = 3/4$.

De plus, nous supposons que le taux de fertilité du processus de Hawkes et Cox est $h(t) = g(t) = \delta e^{-\gamma t}$ avec $\delta = 0,4$ et $\gamma = 2$. Pour satisfaire les hypothèses du théorème précédent, on peut facilement réaliser que l'intensité des immigrants $\nu = 2$, le nombre moyen de points dans tout processus de descendance $\kappa = \sigma = \int_0^{+\infty} h(t) dt$, et le taux de prime $p = 3$ doivent

CHAPITRE 4. PROBABILITÉ DE RUINE POUR UN PROCESSUS DU RISQUE
NON-STATIONNAIRE

être pris de telle sorte que les contraintes du théorème précédent soient satisfaites.

Dans le cas du processus de Cox, nous prenons $\sigma = 2$.

Nous avons réalisé une étude de simulation à l'aide du logiciel statistique R pour valider les performances de notre méthode d'estimation de la probabilité de ruine.

Nous avons généré 100 échantillons de tailles 5000 et 10000 à partir des distributions précédentes avec les valeurs d'indice $\xi = 3/7$, $\xi = 2/3$ et $\xi = 3/4$.

Les packages du logiciel de programmation R utilisés pour simuler des échantillons des distributions de Pareto et Pareto généralisé, sont le package *actuar*, *evir*, *tea* et *eva*. Pour sélectionner le seuil de la loi GPD, faire l'ajustement non paramétrique et de l'inférence statistique sur le paramètre de loi GPD, nous avons utilisé *evir*. Pour simuler des processus de Hawkes et processus de Cox, nous avons utilisé le package *hawkes*.

Pour chaque échantillon simulé, nous obtenons une valeur des estimateurs $\hat{\Phi}(u)$. L'estimateur global $\hat{\Phi}(u)$ est alors considéré comme la moyenne empirique des valeurs dans les 100 répétitions. Nous obtenons également le biais et l'erreur quadratique moyenne (RMSE) de l'estimateur de $\hat{\Phi}(u)$. Nous résumons ces résultats dans les tableaux 4.1 à 4.6. Dans les tableaux ci-dessous, nous rapportons quelques illustrations numériques.

TABLE 4.1 – Probabilité de ruine avec les arrivées du processus de Hawkes pour l'indice de queue $\xi = \frac{3}{7}$

$\Phi(u) = 0.09161031$			
n	$\hat{\Phi}(u)$	$ \Phi(u) - \hat{\Phi}(u) $	RMSE
5000	0.08782737	0.0003782934	0.00665037
10000	0.09433485	0.0002724541	0.009415543

TABLE 4.2 – Probabilité de ruine avec les arrivées du processus de Cox pour l'indice de queue $\xi = \frac{3}{7}$

$\Phi(u) = 0.00801111$			
n	$\hat{\Phi}(u)$	$ \Phi(u) - \hat{\Phi}(u) $	RMSE
5000	0.007917437	0.00009367351	0.003718936
10000	0.007973199	0.00003791125	0.00269465

CHAPITRE 4. PROBABILITÉ DE RUINE POUR UN PROCESSUS DU RISQUE
NON-STATIONNAIRE

TABLE 4.3 – Probabilité de ruine avec les arrivées du processus de Hawkes pour l'indice de queue $\xi = \frac{2}{3}$

$\Phi(u) = 0.1664357$			
n	$\hat{\Phi}(u)$	$ \Phi(u) - \hat{\Phi}(u) $	RMSE
5000	0.164896	0.0015396	0.1533261
10000	0.1672634	0.000827	0.0469

TABLE 4.4 – Probabilité de ruine avec les arrivées du processus de Cox pour l'indice de queue $\xi = \frac{2}{3}$

$\Phi(u) = 0.1455442$			
n	$\hat{\Phi}(u)$	$ \Phi(u) - \hat{\Phi}(u) $	RMSE
5000	0.141575	0.00396921	0.02744071
10000	0.143824	0.001162201	0.01762983

TABLE 4.5 – Probabilité de ruine avec les arrivées du processus de Hawkes pour l'indice de queue $\xi = \frac{3}{4}$

$\Phi(u) = 0.3664412$			
n	$\hat{\Phi}(u)$	$ \Phi(u) - \hat{\Phi}(u) $	RMSE
5000	0.3787894	0.01234817	0.2723891
10000	0.3772105	0.01076923	0.1132433

TABLE 4.6 – Probabilité de ruine avec les arrivées du processus de Cox pour l'indice de queue $\xi = \frac{3}{4}$

$\Phi(u) = 0.3204444$			
n	$\hat{\Phi}(u)$	$ \Phi(u) - \hat{\Phi}(u) $	RMSE
5000	0.3162182	0.00422617	0.0422617
10000	0.3171442	0.0033	0.0379676

CONCLUSIONS ET PERSPECTIVES

Dans cette thèse, l'utilisation de la théorie des valeurs extrêmes est proposée. Nous avons proposé un nouvel estimateur pour la probabilité de ruine dans le cas où les montants de sinistres sont à queue lourde et arrivent suivant un processus stationnaire et non stationnaire via une estimation semi-paramétrique.

Notre approche est basée sur une variante de la théorie des valeurs extrêmes, appelée la méthode POT. Nous avons démontré la consistance et la normalité de l'estimateur de cette probabilité de ruine dans le cas stationnaire.

L'estimateur proposé dans le cas stationnaire est validé via une étude de simulation. Pour tester la performance de notre approche dans le cas stationnaire, nous avons calculé la probabilité de ruine pour des sinistres d'assurance incendie danois et utilisé la méthode bootstrap pour prouver que cette dernière est normale.

L'estimateur de la probabilité de ruine proposé dans le cas non stationnaire est validé par une étude de simulation. Les résultats obtenus sont très satisfaisantes et elles méritent d'être approfondis en étudiant la consistance et la normalité asymptotique de l'estimateur proposé.

Parmi les perspectives, de ce travail est l'étude du cas de dépendance entre les montants des sinistres.

La modélisation conjointe des réserves financières allouer à plusieurs branches d'activité conduit à calculer des probabilités de ruine multivariées. Cela modélise les interactions entre les différentes branches. Dans ce cas nous faisons introduire la théorie des processus multivariés et les distributions de probabilités conjointes et les copules.

Autre perspective est d'inclure une dépendance entre le temps inter-arrivée des sinistres et les montants de sinistres, car dans l'assurance contre les inondations, l'occurrence de plusieurs inondations dans des intervalles de temps courts peut générer de gros dégâts et donc de gros montants de prestation à cause de l'accumulation de l'eau.

Par contre dans l'assurance contre les séismes, c'est l'inverse. Dans une zone à risque, plus le temps entre deux séismes est grand, plus le second séisme est important du fait de l'accumulation de l'énergie.

Notre proposition fournit des résultats intéressants et constitue un outil qui peut être utile pour les chercheurs en actuariat et les compagnies d'assurance lorsqu'ils modélisent par exemple les dommages causés par les incendies et les tempêtes.

Notre proposition a quelques limitations telles que le choix du seuil qui reste un problème de recherche théorique et pratique, et nous devons avoir un grand échantillon de sinistres et un grand capital initial. Ces limitations ouvrent des portes pour des recherches plus approfondies, qui seront prises en compte par des auteurs dans des travaux futurs.

Bibliographie

Soren Asmussen and Hansjorg Albrecher. *Ruin probabilities*, volume 14. World scientific, 2010.

Søren Asmussen, Claudia Klüppelberg, and Karl Sigman. Sampling at subexponential times, with queueing applications. *Stochastic processes and their applications*, 79 (2) :265–286, 1999.

Brian Bader, Jun Yan, and Xuebin Zhang. Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate. 2018.

August A Balkema and Laurens De Haan. Residual life time at great age. *The Annals of probability*, 2(5) :792–804, 1974.

Jan Beirlant and Jozef L Teugels. Asymptotic normality of hill's estimator. *Extreme Value Theory (Oberwolfach, 1987)*, 51 :148–155, 1987.

Jan Beirlant, Petra Vynckier, and Jozef L Teugels. Tail index estimation, pareto quantile plots regression diagnostics. *Journal of the American statistical Association*, 91(436) : 1659–1667, 1996.

Jan Beirlant, Yuri Goegebeur, Johan Segers, and Jozef L Teugels. *Statistics of extremes : theory and applications*, volume 558. John Wiley & Sons, 2004.

NH Bingham. Regular variation and probability : The early years. *Journal of computational and applied mathematics*, 200(1) :357–363, 2007.

Charles Bordenave and Giovanni Luca Torrisi. Large deviations of poisson cluster processes. *Stochastic Models*, 23(4) :593–625, 2007.

Ruth Burkhalter and Yuhlong Lio. Bootstrap control charts for the generalized pareto distribution percentiles. *Chilean Journal of Statistics (ChJS)*, 12(1), 2021.

Frederico Caeiro and M Ivette Gomes. Threshold selection in extreme value analysis. *Extreme value modeling and risk analysis : Methods and applications*, pages 69–87, 2015.

Enrique Castillo, Ali S Hadi, Narayanaswamy Balakrishnan, and José-Mariá Sarabia. *Extreme value and related models with applications in engineering and science*. 2005.

Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.

CH Cramér. On the mathematical theory of risk, forsakringsaktiebolaget skandias festskrift. *Centraltryckeriet, Stockholm*, 1930.

Jon Danielsson, Laurens de Haan, Liang Peng, and Casper G de Vries. Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate analysis*, 76(2) :226–248, 2001.

Jon Danielsson, Lerby Murat Ergun, Laurens de Haan, and Casper G de Vries. Tail index estimation : Quantile driven threshold selection. *Available at SSRN 2717478*, 2016.

Laurens de Haan and Ana Ferreira. *Extreme value theory : an introduction*, volume 3. Springer, 2006.

P de Zea Bermudez and Samuel Kotz. Parameter estimation of the generalized pareto distribution—part i. *Journal of Statistical Planning and Inference*, 140(6) :1353–1373, 2010.

Arnold LM Dekkers and Laurens de Haan. Optimal choice of sample fraction in extreme-value estimation. *Journal of Multivariate Analysis*, 47(2) :173–195, 1993.

- Arnold LM Dekkers, John HJ Einmahl, and Laurens De Haan. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, pages 1833–1855, 1989.
- William H DuMouchel. Estimating the stable index α in order to measure tail thickness : A critique. *the Annals of Statistics*, 11(4) :1019–1031, 1983.
- Bradley Efron and Raoul LePage. *Introduction to bootstrap*. Wiley & Sons, New York, 1992.
- P Embrechts, F Lindskog, A McNeil, and S Rachev. Handbook of heavy tailed distributions in finance. *Chapter, 8* :329–384, 2003.
- Paul Embrechts. *Extremes and integrated risk management*. Risk Books, 2000.
- Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events : for insurance and finance*, volume 33. Springer Science & Business Media, 2013.
- Michael Falk. On testing the extreme value index via the pot-method. *The Annals of Statistics*, 23(6) :2013–2035, 1995.
- Mir Nabi Pirouzi Fard and Björn Holmquist. Powerful goodness-of-fit tests for the extreme value distribution. *Chilean Journal of Statistics*, 4(1) :55–67, 2013.

William Feller. *An introduction to probability theory and its applications, Volume 2*, volume 81. John Wiley & Sons, 1991.

Ana Ferreira, Laurens de Haan, and Liang Peng. On optimising the estimation of high quantiles of a probability distribution. *Statistics*, 37(5) :401–434, 2003.

Ronald Aylmer Fisher and Leonard Henry Caleb Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical proceedings of the Cambridge philosophical society*, volume 24, pages 180–190. Cambridge University Press, 1928.

Maurice Fréchet. Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Math. Polon.*, 6 :93–116, 1927.

Redouane Frihi, Abdelaziz Rassoul, and Hamid Ould Rouis. Pot-based estimator of the ruin probability in infinite time for loss models. *Chilean Journal of Statistics*, 13(2) :1–19, 2022.

Janos Galambos. The asymptotic theory of extreme order statistics. Technical report, 1978.

Laurent Gardes. *Estimation d'une fonction quantile extrême*. PhD thesis, Université Montpellier II-Sciences et Techniques du Languedoc, 2003.

- Boris Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, pages 423–453, 1943.
- Charles M Goldie and Sidney Resnick. Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution. *Advances in Applied Probability*, 20(4) :706–718, 1988.
- Jan Grandell and C-O Segerdahl. A comparison of some approximations of ruin probabilities. *Scandinavian Actuarial Journal*, 1971(3-4) :143–158, 1971.
- Armelle Guillou and Peter Hall. A diagnostic for selecting the threshold in extreme value analysis. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 63(2) :293–305, 2001.
- Peter Hall. Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of multivariate analysis*, 32(2) :177–203, 1990.
- Peter Hall and Christopher C Heyde. *Martingale limit theory and its application*. Academic press, 2014.
- Alan G Hawkes. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society : Series B (Methodological)*, 33(3) :438–443, 1971.

- Bruce M Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, pages 1163–1174, 1975.
- M Ivette Gomes, Lígia Rodrigues, Björn Vandewalle, and Clara Viseu. A heuristic adaptive choice of the threshold for bias-corrected hill estimators. *Journal of Statistical Computation and Simulation*, 78(2) :133–150, 2008.
- Arthur F Jenkinson. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348) :158–171, 1955.
- Joachim Johansson. Estimating the mean of heavy-tailed distributions. *Extremes*, 6 :91–109, 2003.
- Vladimir Kalashnikov and Dimitrios Konstantinides. Ruin under interest force and subexponential claims : a simple treatment. *Insurance : Mathematics and Economics*, 27(1) : 145–149, 2000.
- Bryan Kelly and Hao Jiang. Tail risk and asset prices. *The Review of Financial Studies*, 27(10) :2841–2871, 2014.
- Claudia Klüppelberg and Ulrich Stadtmüller. Ruin probabilities in the presence of heavy-tails and interest rates. *Scandinavian Actuarial Journal*, 1998(1) :49–58, 1998.

- F Lundberg. Approximerad framställning av sannolikehetsfunktionen, återförsäkering av kollektivrisker, almqvist & wiksell, uppsala, 1903.
- David M Mason. Laws of large numbers for sums of extreme values. *The Annals of Probability*, pages 754–764, 1982.
- Gunther Matthys and Jan Beirlant. Adaptive threshold selection in tail index estimation. *Extremes and Integrated Risk Management*, pages 37–49, 2000.
- Alexander J McNeil. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin : The Journal of the IAA*, 27(1) :117–137, 1997.
- Thomas Mikosch and Gennady Samorodnitsky. Ruin probability with claims modeled by a stationary ergodic stable process. *Annals of Probability*, pages 1814–1851, 2000.
- Harry H Panjer and Gordon E Willmot. *Insurance risk models*. Number 517/P19i. 1992.
- Valentin V Petrov. Limit theorems of probability theory ; sequences of independent random variables. Technical report, 1995.
- James Pickands. Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131, 1975.

- Rolf-Dieter Reiss, Michael Thomas, and RD Reiss.
Statistical analysis of extreme values, volume 2. Springer, 1997.
- SI Resnick. Extreme values, regular variation, and point processes, *appl. Probab. Ser. Appl. Probab. Trust*, 4 :33, 1987.
- Sidney Resnick and Cătălin Stărică. Smoothing the hill estimator. *Advances in Applied Probability*, 29(1) :271–293, 1997.
- Sidney I Resnick. *Heavy-tail phenomena : probabilistic and statistical modeling*. Springer Science & Business Media, 2007.
- Carl Scarrott and Anna MacDonald. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT-Statistical journal*, 10(1) :33–60, 2012.
- Eugene Seneta. *Regularly varying functions*, volume 508. Springer, 2006.
- Jessica Silva Lomba and Maria Isabel Fraga Alves.
L-moments for automatic threshold selection in extreme value analysis. *Stochastic Environmental Research and Risk Assessment*, 34(3-4) :465–491, 2020.
- Richard L Smith. Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1) :67–90, 1985.

- Richard L Smith. Estimating tails of probability distributions. *The annals of Statistics*, pages 1174–1207, 1987.
- Gabriele Stabile and Giovanni Luca Torrisi. Risk processes with non-stationary hawkes claims arrivals. *Methodology and Computing in Applied Probability*, 12 :415–429, 2010.
- Richard Von Mises. La distribution de la plus grande de n valeurs. *Rev. math. Union interbalcanique*, 1 :141–160, 1936.
- Lingjiong Zhu. Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance : Mathematics and Economics*, 53(3) :544–550, 2013.