

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
UNIVERSITÉ MOHAMED KHIDER, BISKRA
FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et la
VIE

DÉPARTEMENT DE MATHÉMATIQUES



Thèse présentée en vue de l'obtention du
DIPLÔME : Doctorat en Mathématiques
Option : **Statistique**

Par
Fatima MEDDI

Titre
**Estimation des mesures de risques pour les distributions à queue
lourde**

Membres du jury

Président	Professeur Mezerdi Brahim	Université M.K. Biskra
Rapporteur	Professeur Necir Abdelhakim	Université M.K. Biskra
Examineur	Professeur Fellag Hocine	Université M.M. Tizi-Ouzou
Examineur	Professeur Mohdab Zaher	Université M. Constantine
Examineur	Dr. Meraghni Djamel	Université M.K. Biskra
Examineur	Dr. Remita Mohamed-Riad	Université B.M. Annaba

Janvier 2014

REMERCIEMENTS

Je remercie Dieu le tout-puissant de m'avoir donné la volonté, la force et le courage pour bien mener et finir mon travail de thèse.

Je tiens à remercier infiniment mon encadreur Monsieur le professeur Necir Abdelhakim, d'avoir accepté de m'encadrer, me proposer un sujet de thèse d'actualité ainsi pour son soutien et de m'avoir mis très à l'aise lors de mon travail. Je profite de cette occasion pour lui présenter ma gratitude pour l'effort qu'il a fourni pendant des années à nous transmettre le savoir, l'orientation, les conseils, ainsi être à l'écoute et avoir resté pendant toute la période de notre formation disponible pour nous tous. Tous les mots qu'on va chercher pour lui remercier ne suffiront pas pour lui exprimer le respect, l'estime, l'appréciation et la gratitude qu'on a pour lui.

Je tiens également à présenter un grand merci pour mon enseignant à qui je ressens beaucoup de fierté d'avoir honoré ma soutenance en étant le président du jury, Monsieur le professeur Mezerdi Brahim qui était toujours, et restera le symbole de la réussite dans la recherche scientifique pour moi et mes collègues. Pour sa présence tout d'abord, malgré ses occupations, l'apport, les conseils et les questions efficaces qui ont mettent en valeur mon travail.

Avec un grand honneur j'aimerais présenter mes remerciements et ma gratitude aux membres du jury, Monsieur le Professeur Fellag Hocine, Monsieur le Professeur Mohdab Zaher, Monsieur le dr. Meraghni Djamel et Monsieur le dr. Remita Mohamed-Riad, tout d'abord d'avoir accepté d'examiner ma thèse, qui sans eux cette thèse ne pourra jamais voir le jour. C'est un honneur pour moi d'avoir accepté de lire mon travail attentivement donner de leur temps pour préparer un rapport, pour leurs commentaires, pour bien évaluer ma thèse, pour leurs questions, intérêt et apport qu'ils ont apporté à mon travail.

Je remercie également mes parents et ma petite famille de m'avoir encouragé, soutenu et supporté mon stress durant mon travail.

Table des matières

0.1	Introduction	8
1	Préliminaires mathématiques des valeurs extrêmes	11
1.1	Définitions et théorèmes	11
1.1.1	Distributions empiriques	11
1.1.2	Loi des valeurs extrêmes	14
1.1.3	Domaines d'attraction	16
1.2	Classification des distributions	17
1.2.1	Classe de distributions à queue lourde	17
1.2.2	Classe de distributions à variation régulière	19
1.2.3	Classe de distributions extrémales	22
2	Analyse exploratoire des observations extrêmes	25
2.1	Quantile-quantile plot	29
2.1.1	Introduction	29
2.1.2	QQ-plot d'une distribution exponentielle	30
2.1.3	QQ-plot d'une distribution de Pareto	32
2.2	Mean excess plot	33
2.2.1	Loi des excès	33
2.2.2	Fonction moyenne des excès théorique $e(u)$	35
2.2.3	Fonction moyenne des excès empirique $\hat{e}_n(u)$	35
3	Estimation des statistiques extrêmes de type Pareto	38
3.1	Estimation de l'indice de queue γ	39
3.1.1	Estimateur de Hill $\hat{\gamma}_n^h$	40
3.1.2	Estimateur de Matthys et Beirlant $\hat{\gamma}_n^m$	43
3.1.3	Comparaison entre $\hat{\gamma}_n^h$ et $\hat{\gamma}_n^m$	48
3.2	Estimation des quantiles extrêmes x_p	51
3.2.1	Estimateur des quantiles extrêmes de Weissman \hat{x}_p^w	52
3.2.2	Estimateur des quantiles extrêmes de Matthys \hat{x}_p^m	53
3.2.3	Choix du nombre des valeurs extrêmes k	55

4	Mesures de risques pour les plus grandes revendications	56
4.1	Les outils mathématiques de l'analyse des risques	57
4.1.1	Mesures de risque.	57
4.1.2	Comparaison des risques	62
4.1.3	Principes de calcul de prime	63
4.2	Estimation de la prime de Wang pour des risques extrêmes . .	65
4.2.1	Prime d'hasard proportionnelle	65
4.2.2	Estimation semi-paramétrique	67
4.2.3	Estimateur à biais réduit	69
5	Estimation semi-paramétrique des mesures de risque à dis-	
	torsion pour des pertes à queue lourde	70
5.1	Estimation semi-paramétrique	73
5.1.1	Motivation d'estimation	73
5.1.2	Estimateur de Necir et Meraghni $\Psi[\tilde{Q}_n, g]$	74
5.1.3	Nouveau estimateur à biais réduit $\Psi[\hat{Q}_n, g]$	78
5.2	Etude asymptotique de $\Psi[\hat{Q}_n, g]$	81
5.2.1	Théorème de normalité asymptotique	81
5.2.2	Démonstrations	83
5.2.3	Intervalles de confiance	93
6	Simulations comparatives et exploration des observations ex-	
	trêmes des données réelles sous SOA	94
6.1	Simulations et comparaisons	94
6.1.1	Paramètres de simulation	95
6.1.2	Paramètre de distorsion	101
6.1.3	Paramètres de comparaison	106
6.2	Exploration des observations réelles extrêmes	111
6.2.1	Interprétation des données	112
6.2.2	QQ-plot exponentielle des données	113
6.2.3	QQ-norm de la distribution limite de $\Psi[\hat{Q}_n, g]$	115

Table des figures

3.1	Plot de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d'une loi Pareto de taille $n = 500$ et $\gamma = 2/3$	49
3.2	Plot de l' <i>EQM</i> empirique de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d'une loi Pareto de taille $n = 500$ et $\gamma = 2/3$	49
3.3	Plot de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d'une loi Pareto de taille $n = 1000$ et $\gamma = 2/3$	50
3.4	Plot de l' <i>EQM</i> empirique de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d'une loi Pareto de taille $n = 1000$ et $\gamma = 2/3$	50
5.1	Graphe de $\Psi \left[\hat{Q}_n, g \right]$ vs le nombre des plus grande statistiques d'ordre issue de 100 échantillons d'une distribution de Pareto avec $n = 500$	80
5.2	Graphe de $\Psi \left[\hat{Q}_n, g \right]$ vs le nombre des plus grande statistiques d'ordre issue de 100 échantillons d'une distribution de Pareto avec $n = 1000$	80
6.1	$\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ d'une distribution de Pareto avec $\gamma = 0.6$ et $n = 200$	102
6.2	$\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ d'une distribution de Pareto avec $\gamma = 0.6$ et $n = 500$	103
6.3	$\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ d'une distribution de Pareto avec $\gamma = 0.6$ et $n = 1000$	103
6.4	$\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ d'une distribution de Pareto avec $\gamma = 0.6$ et $n = 1500$	104
6.5	Intervalles de confiance empiriques de $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ pour 100 répétitions de distribution de Pareto, $n = 500, \gamma = 2/3$	105

6.6	Intervalles de confiance empiriques de $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ pour 100 répétitions de distribution de Pareto, $n = 1000, \gamma = 2/3$	106
6.7	QQ-norm de la distribution limite de $\Psi \left(g, \hat{Q}_n \right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 200$	107
6.8	QQ-norm de la distribution limite de $\Psi \left(g, \hat{Q}_n \right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 500$	107
6.9	QQ-norm de la distribution limite de $\Psi \left(g, \hat{Q}_n \right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 1000$	108
6.10	QQ-norm de la distribution limite de $\Psi \left(g, \tilde{Q}_n \right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 200$	108
6.11	QQ-norm de la distribution limite de $\Psi \left(g, \tilde{Q}_n \right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 500$	109
6.12	QQ-norm de la distribution limite de $\Psi \left(g, \tilde{Q}_n \right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 1000$	109
6.13	QQ-plot exponentiel des réclamations de l'année 1997	114
6.14	QQ-plot exponentiel des réclamations de l'année 1998	114
6.15	QQ-plot exponentiel des réclamations de l'année 1999	115
6.16	QQ-norm de la distribution limite de $\Psi \left(g, \hat{Q}_n \right)$ de 200 répétitions des données de l'année 1997	119
6.17	QQ-norm de la distribution limite de $\Psi \left(g, \hat{Q}_n \right)$ de 200 répétitions des données de l'année 1998	119
6.18	QQ-norm de la distribution limite de $\Psi \left(g, \hat{Q}_n \right)$ de 200 répétitions des données de l'année 1999	120

Liste des tableaux

1.1	Quelques distributions pour modéliser les plus grandes revendications des grands sinistres	21
2.1	Table de test de Smirnov-Kolmogorov.	27
4.1	Fonctions de distorsion associées à quelques mesures de risque	62
6.1	200 échantillons générés de distribution de Pareto de taille n=200	110
6.2	200 échantillons générés de distribution de Pareto de taille n=500	111
6.3	200 échantillons générés de distribution de Pareto de taille n=1000	111
6.4	Compte de contrôle et totaux	112

0.1 Introduction

En calcul actuariel les mesures de risque sont d'intérêt considérable. Elles permettent à un assureur de calculer une prime de risque chargée d'une perte aléatoire. Toutefois, l'utilisation du principe de calcul de prime par l'assureur peut être au moins en partie, fondée sur des considérations autres que le risque. Il est alors important de quantifier le degré auquel la prime compense l'assureur pour le risque lié à la perte. Cela peut être fait en choisissant une mesure de risque appropriée et en résolvant le paramètre qui mène à la prime de l'assureur. Lorsque la distribution des pertes est inconnue, cela devient un problème d'estimation statistique. Dans cette thèse, nous étudions l'estimation semi-paramétrique d'une mesure de risque à distorsion de Wang (1996) quand la distribution des pertes est à queue lourde et nous explorons les propriétés asymptotiques de l'estimateur de mesure du risque sous la condition de variation régulière au second ordre.

Le premier chapitre concerne des préliminaires mathématiques où nous commençons par une introduction de concepts de base et de motivation, en présentant par la suite le théorème clé de Gnedenko (1943) des valeurs extrêmes. Nous discutons également la notion des domaines d'attraction d'une distribution des valeurs extrêmes qui sont étroitement liés au concept de variation régulière. Nous spécifions la classe des distributions qui sont utilisés dans cette thèse. A savoir, les classes spécifiques à queue lourde et également la classe extrémal qui apparaissent naturellement dans les applications de la théorie des valeurs extrêmes.

Nous présentons dans *le deuxième chapitre* des outils d'exploration des observations extrêmes. Ces derniers sont caractérisés par la valeur du paramètre de l'indice de queue γ , où plus cet indice élevé en valeur absolue, plus le poids des extrêmes dans la distribution initiale est important. Nous avons commencé par construire un test d'adéquation de Kolmogorov- Smirnov pour ajuster l'échantillon des données en question et donner une estimation numérique pour la valeur de l'indice de queue, γ dans le cas d'une distribution de type de *Pareto*. Ensuite nous présentons deux outils graphiques sous R ; les *Quantile-Quantile plot* et *Mean Excess Plot*, qui sont souvent plus instructifs pour détecter l'existence des valeurs extrêmes. Ces outils peuvent être utilisés comme un moyen d'ajustement dans le cas où on réussit à mieux ajuster l'échantillon des données à une loi de probabilité bien définie, par la détermination de l'indice de queue γ et par la suite on peut utiliser l'échantillon des données dans des résultats théoriques donnés. Dans le cas contraire lorsque la distribution des pertes est inconnue, cela devient un problème d'estimation

statistique où ces outils peuvent servir à savoir si l'échantillon des données suit une distribution à queue lourde, ainsi préciser le domaine d'attraction d'appartenance de la distribution des données en question avant de mettre en œuvre nos estimations à la base supposées, pour des échantillons de variables aléatoires indépendantes et identiquement distribuées de type-*Pareto*.

Lorsque la distribution de l'échantillon est inconnue, l'estimation est la seule alternative. *Le troisième chapitre* traite l'estimation des statistiques extrêmes de type *Pareto* dites aussi queue de la distribution. Ceci est lié directement à l'estimation de l'indice de queue, γ et la détermination du nombre des valeurs extrêmes optimal, k^* influençant sur les qualités asymptotiques de ces estimateurs. Nous avons présenté l'estimateur de l'indice de queue de Hill (1975), $\hat{\gamma}_n^h$ calculé par la méthode du maximum de vraisemblance, associé à l'estimateur des quantiles extrêmes de Weisman (1978), \hat{x}_p^w . Notamment nous avons présenté les estimateurs de Matthys et Beirlant (2003) de l'indice de queue, $\hat{\gamma}_n^m$ et celui des quantiles extrêmes associés, \hat{x}_p^m calculés tout les deux par la méthode de régression. Les résultats asymptotiques sont donnés brièvement et une comparaison graphique entre $\hat{\gamma}_n^h$ et $\hat{\gamma}_n^m$ versus le nombre des valeurs extrêmes k est illustrée. Notre choix s'explique par notre volonté d'introduire une adaptation des estimateurs ainsi proposés aux prochains chapitres.

Nous commençons *le quatrième chapitre*, mesure de risque pour les plus grandes revendications, par présenter les outils mathématiques de l'analyse de risque, en définissant les mesure de risque, la comparaison des risques et les principes de calcul de primes, en donnant les propriétés de bonnes primes. Si la variable de perte a une distribution à queue lourde (i.e., distribution avec une variance infinie). Alors la mesure de risque devrait être plus élevée. L'utilisation des mesures de risque de Wang (1996) s'explique par ses bonnes propriétés pour le calcul des primes concernant les plus grandes revendications. Comme la distribution de risque est dans la plupart du temps inconnue, en se basant sur les estimateurs de Matthys et Beirlant (2003) présentés dans le chapitre 3, nous proposons une estimation semi-paramétrique pour cette mesure de risque par distorsion pour des pertes à queue lourde et uniquement pour la partie des valeurs extrêmes de l'échantillon de risque qu'on note $\hat{\pi}_{u_n}^*$. Ce dans le but de minimiser le biais de l'estimateur de Necir et al. (2007) noté $\hat{\pi}_{u_n}$, ainsi présenté dans ce même chapitre.

Nos principaux résultats sont présentés au *cinquième chapitre*, concernant la motivation de l'estimation semi-paramétrique d'une mesure de risque à distorsion des pertes à queue lourde pour l'échantillon en totalité. Necir et

Meraghni (2009) ont proposé un estimateur asymptotiquement normal en se basant sur des estimateurs des quantiles extrême de Weissman (1978), \hat{x}_p^w qu'on a noté $\Psi[\tilde{Q}_n; g]$. Dans ce chapitre nous proposons un nouvel estimateur à biais réduit pour cette classe de prime de risque en estimant la queue de distribution de l'échantillon par les estimateurs des quantiles extrêmes de Matthys et Beirlant (2003) (voir Brahim et al. 2012). Ce nouvel estimateur noté $\Psi[\hat{Q}_n; g]$, est une généralisation de l'estimateur $\Psi[\tilde{Q}_n; g]$. Le résultat important dans cette thèse est présenté dans le théorème 5.3, démontrant la normalité asymptotique de notre nouvel estimateur de mesure de risque. Ce résultat nous permet d'avoir en même temps des intervalles de confiance théoriques ainsi présentés en fin de ce chapitre.

Nous avons consacré *le sixième chapitre* en première partie pour présenter des simulations comparatives entre notre nouvel estimateur, $\Psi[\hat{Q}_n; g]$ (Brahimi et al. 2012) et l'estimateur $\Psi[\tilde{Q}_n; g]$ (Necir et Meraghni, 2009). Nous avons spécifié des paramètres de simulation qui sont le nombre des valeurs extrêmes k_1 et k_2 où chacun d'eux a été calculé selon les résultats asymptotiques de chacun des estimateurs de l'indice de queue de Hill, $\hat{\gamma}_n^h$ et celui de Matthys et Beirlant, $\hat{\gamma}_k$, à la base lié directement à la construction des estimateurs de mesure de risque $\Psi[\tilde{Q}_n; g]$ et $\Psi[\hat{Q}_n; g]$, respectivement. Une brève étude a été consacré au paramètre de distorsion ρ , jouant un rôle important dans la présentation final des primes et qui peut servir comme un outil de réglage pour obtenir des primes les plus adéquatement possible en traçant des intervalles de confiance en fonction de ce dernier. Notamment, des paramètres de comparaison comme des indicateurs d'écart calculés par la méthode de Bootstrap où nous avons constaté que notre nouvel estimateur $\Psi[\hat{Q}_n; g]$ présente un biais réduit. Dans La deuxième partie du chapitre 6, nous avons présenté une exploration des données réelles regroupées sous la base des données *SOA* des trois années 1997, 1998 et 1999 des sept compagnie d'assurance. Quelques programmes et graphes exécutées sous l'environnement *R* sont présentés pour en illustrer. En traçant le graphique *QQ*-plot des observations réelles nous avons conclu que ces dernières suivent une distribution à queue lourde et en traçant le graphique *QQ*-norm de la distribution limite de notre nouvel estimateur de mesure de risque $\Psi[\hat{Q}_n; g]$ des données précitées calculée par Bootstrap, nous concluons que cette dernière est empiriquement asymptotiquement normal.

Chapitre 1

Préliminaires mathématiques des valeurs extrêmes

Depuis quelques années, la théorie des valeurs extrêmes a reçu beaucoup d'attention aussi bien sur le plan théorique que sur le plan pratique. Les domaines d'applications sont en effet très variés : hydrologie, météorologie, biologie, ingénierie, gestion de l'environnement, finance, assurance, sciences sociales, etc. ; en effet la gestion des risques est devenue aujourd'hui fondamentale dans tous ces domaines. De nouveaux types de risques bouleversent quotidiennement le travail des actuaires et des risk managers. La théorie des valeurs extrêmes peut toutefois leur permettre de mieux modéliser ces risques, comme on va voir dans les prochains chapitres, elle s'intéresse non pas à la modélisation totale d'une distribution mais seulement au queue de cette distribution à partir de lois spécifiques propres que nous allons préciser dans ce qui suit. Pour plus d'approfondissements voir par exemple Embrechts et al. (1997) et Beirlant et al. (2004).

1.1 Définitions et théorèmes

1.1.1 Distributions empiriques

On considère ici une suite de n observations $(X_i)_{1 \leq i \leq n}$ indépendantes et identiquement distribuées. Les statistiques d'ordre représentent ces variables aléatoires "ordonnées" : i.e., pour tout $\omega \in \Omega$, on peut ranger par ordre croissant les réalisations des X_i , i.e. $x_1^{(\omega)} \leq x_2^{(\omega)} \leq \dots \leq x_n^{(\omega)}$.

Définition 1.1. *On appelle statistiques d'ordre les variables aléatoires ordonnées $(X_i)_{1 \leq i \leq n}$ telles que pour tout entiers i , $X_{i,n}(\omega) = x_i^{(\omega)}$. On a donc,*

$X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$, et $X_{i,n}$ est la $i^{\text{ème}}$ statistique d'ordre.

Dans un échantillon de taille n , deux statistiques d'ordre sont particulièrement intéressantes pour l'étude des événements extrêmes; le minimum et le maximum : $X_{1,n} = \min(X_1, \dots, X_n)$ et $X_{n,n} = \max(X_1, \dots, X_n)$.

Loi de $X_{i,n}$

$$F_{i,n} = P\{X_{i,n} \leq x\} = \sum_{r=1}^n \binom{n}{r} (F(x))^r (1 - F(x))^{n-r}$$

Nous en déduisons que la fonction de densité est :

$$f_{i,n}(x) = \frac{n!}{(i-1)!(n-i)!} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x),$$

où $f(x)$ est la densité de probabilité de X_i et F sa fonction de répartition associée. En utilisant la propriété d'indépendance des variables aléatoires X_1, \dots, X_n , on obtient :

Loi de $X_{1,n}$

$$F_{1,n}(x) = P\{X_{1,n} \leq x\} = 1 - (1 - F(x))^n,$$

d'où

$$f_{1,n}(x) = n f(x) (1 - F(x))^{n-1}$$

Loi de $X_{n,n}$

$$F_{n,n}(x) = P\{X_{n,n} \leq x\} = (F(x))^n$$

d'où

$$f_{n,n}(x) = n f(x) (F(x))^{n-1}$$

Remarque 1.1 :

$$P\{X_{n,n} \leq x\} = (F(x))^n \rightarrow 0 \text{ ou } 1 \text{ quand } n \rightarrow \infty.$$

Soit X_1, \dots, X_n un n -échantillon issu d'une fonction de répartition commune F telle que, $F(x) = p(X_i < x)$, $i = 1, \dots, n$ et on note la fonction inverse généralisée de F par Q , telle que,

$$Q(t) = F^{-1}(t) = \inf\{s, F(s) \geq t\}, \quad 0 < t < 1$$

La fonction de répartition empirique : notée F_n et donnée par :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{]-\infty, x[}(X_i), \quad x \in \mathbb{R}$$

Il existe une autre version de la définition de F_n en utilisant les $(s.o)$ comme suit :

$$F_n(x) = \begin{cases} 0 & si, \quad x \leq X_{1,n} \\ \frac{i-1}{n} & si, \quad X_{i-1,n} < x \leq X_{i,n}, \quad 2 \leq i < n \\ 1 & si, \quad x > X_{n,n} \end{cases}$$

La fonction des quantiles empiriques : notée Q_n et donnée par :

$$Q_n(t) = F_n^{-1}(t) = \sup \{s : F_n > t\}, \quad 0 \leq t \leq 1.$$

De même on obtient une autre version de Q_n en utilisant les $(s.o)$ comme suit :

$$Q_n(t) = \begin{cases} X_{i,n} & si \quad \frac{i-1}{n} \leq t \leq \frac{i}{n} \\ X_{n,n} & si \quad \frac{1-1}{n} \leq t \leq \frac{n}{n} \end{cases}$$

La fonction de répartition empirique uniforme : notée G_n et donnée par :

$$G_n(t) = n^{-1} \sum_{i=1}^n \mathbb{I}_{\{\xi_i \leq t\}}, \quad 0 \leq t \leq 1$$

La fonction des quantiles empiriques uniformes : notée V_n et donnée par :

$$V_n = \inf \{s \geq 0 : G_n(t) \geq t\}, \quad 0 < t \leq 1.$$

Le processus empirique uniforme : noté α_n et donné par :

$$\alpha_n(t) = n^{1/2} (G_n(t) - t), \quad 0 \leq t \leq 1$$

Le processus empirique des quantiles uniformes : noté β_n et donné par :

$$\beta_n(t) = n^{1/2} (V_n(t) - t), \quad 0 < t \leq 1$$

Distribution des extrema dans le cas fini : On définit la variable aléatoire M_n , qui traduit le maximum d'un n-échantillon d'une variable aléatoire X , (les variables aléatoires X_i sont indépendantes et suivent la même loi que X) par :

$$M_n = \max (X_i)_{1 \leq i \leq n}$$

On pourrait aussi s'intéresser au minimum et utiliser la relation :

$$\min (X_i)_{1 \leq i \leq n} = - \max (-X_i)_{1 \leq i \leq n}$$

En théorie des valeurs extrêmes, le but visé est de déterminer la loi limite normalisée que suit le maximum (ou le minimum) en fonction de celle de la variable aléatoire X . Notons la fonction de répartition F_X de la variable aléatoire de loi de probabilité P , à savoir $F_X(x) = P(X < x)$. La fonction de répartition de M_n est alors définie par :

$$\begin{aligned} F_{M_n}(x) &= P(M_n \leq x) \\ &= P(X_1 \leq x, \dots, X_n \leq x) \\ &= P(X_1 \leq x) \times \dots \times P(X_n \leq x) \\ &= [F_X(x)]^n \end{aligned}$$

De ces résultats, nous tirons la conclusion que le maximum M_n est une variable aléatoire dont la fonction de répartition est égale à $(F_X)^n$. La fonction de répartition de X étant souvent inconnue et généralement pas possible d'être déterminée. Notons $x_F = \sup \{x \in \mathbb{R} : F_X(x) < 1\}$ le point terminal à droite (right-end point) de la fonction de répartition F_X . Ce point terminal peut être infini ou fini (Embrechts et al. 1997, exemple 3.3.22, p.139). On s'intéresse ici au distribution asymptotique du maximum, en faisant tendre n vers l'infini,

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

On constate que la distribution asymptotique du maximum, donne une loi dégénérée, une masse de Dirac en x_F , puisque pour certaines valeurs de x , la probabilité peut être égale à 1 dans le cas où x_F est fini. Ce fait ne fournit pas assez d'informations, d'où l'idée d'utiliser une transformation afin d'obtenir des résultats plus exploitables pour les loi limites des maxima M_n . La transformation la plus simple est l'opération de standardisation.

1.1.2 Loi des valeurs extrêmes

Comme la fonction de répartition obtenue précédemment conduit à une loi dégénérée lorsque n tends vers l'infini, on recherche une loi non dégénérée

pour le maximum de X . Cette loi limite non dégénérée est fournie par le "théorème des distributions extrêmes" qui donne une condition nécessaire et suffisante pour l'existence d'une loi limite non dégénérée pour le maximum. Ce théorème est proposé par Gnedenko (1943) qui donne la forme des lois limites et Jenkinson (1955) qui en donne l'expression générale.

Théorème 1.1. (Fisher et Tippett, 1928, Gnedenko, 1943) Soit X_1, \dots, X_n une suite de n variables aléatoires réelles *iid* de loi continue P et $M_n = \max(X_i)_{1 \leq i \leq n}$. s'il existe deux suites réelles $(a_n)_{n \geq 1}$ et $(b_n)_{n \geq 1}$ avec $b_n > 0$, et une fonction de répartition non-dégénérée G_γ telle que,

$$\frac{M_n - a_n}{b_n} \xrightarrow{d} G_\gamma \text{ lorsque } n \rightarrow \infty$$

Alors G_γ est nécessairement l'un des trois types suivant :

$$G_\gamma = G_0(x) = \exp(-\exp(-x)), \quad -\infty < x < +\infty$$

$$G_\gamma = G_{1,\gamma}(x) = \begin{cases} 0, & x < 0 \\ \exp(-x^{-1/\gamma}) & x \geq 0, \gamma > 0 \end{cases}$$

$$G_\gamma = G_{2,\gamma}(x) = \begin{cases} \exp(-(-x)^{-1/\gamma}) & x < 0, \gamma < 0 \\ 1 & x \geq 0 \end{cases}$$

On appelle G_0 loi de Gumbel, $G_{1,\gamma}$ loi de Fréchet et $G_{2,\gamma}$ loi de Weibull. En introduisant les paramètres de localisation μ et de dispersion σ dans la paramétrisation des distributions extrêmes, on obtient la forme la plus générale de la distribution des valeurs extrêmes, notée GEV (Generalized Extreme Value Distribution). Elle correspond à :

$$G_{\mu,\sigma,\gamma}(x) = \exp \left\{ - \left[1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\gamma} \right\},$$

où γ est un paramètre de forme (shape parameter) encore appelé indice des valeurs extrêmes ou indice de queue. Plus cet indice est élevé en valeur absolue, plus le poids des extrêmes dans la distribution initiale est important. On parle alors de distribution à "queue épaisses". Trois cas sont possibles

- . $\gamma > 0$, $G_{\mu,\sigma,\gamma}$ suit la loi de Fréchet
- . $\gamma < 0$, $G_{\mu,\sigma,\gamma}$ suit la loi de Weibull
- . $\gamma \rightarrow 0$, $G_{\mu,\sigma,\gamma}$ suit la loi de Gumbel

Remarque 1.2. Si F vérifie le théorème 1.1. On dit alors que F appartient au domaine d'attraction de G_γ et on note $F \in D(G_\gamma)$ selon le signe de γ .

1.1.3 Domaines d'attraction

Nous allons donner des conditions sur la fonction de répartition F pour qu'elle appartienne à l'un des trois domaines d'attraction.

Domaine d'attraction de Fréchet

Théorème 1.2. F appartient au domaine d'attraction de Fréchet avec un indice de valeur extrêmes $\gamma > 0$ si et seulement si $x_F = +\infty$ et $1 - F$ est une fonction à variation régulière d'indice $1/\gamma$. Dans ce cas, un choix possible pour les suites a_n et b_n est :

$$a_n = F^{-1} \left(1 - \frac{1}{n} \right) \quad \text{et} \quad b_n = 0.$$

Remarque 1.3. On déduit du théorème 1.2 que F appartient au domaine d'attraction de Fréchet avec un indice de valeur extrême $\gamma > 0$ si et seulement si $x_F = +\infty$ et $1 - F = x^{-1/\gamma} \ell(x)$, où ℓ est une fonction à variation lente (voir définition 1.6).

Domaine d'attraction de Weibull

Théorème 1.3. F appartient au domaine d'attraction de Weibull avec un indice de valeur extrême $\gamma < 0$ si et seulement si $x_F < +\infty$ et $1 - F^*$ est une fonction à variation régulière d'indice $1/\gamma$ avec

$$F^*(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ F(x_F - x^{-1}) & \text{si } x > 0. \end{cases}$$

Dans ce cas, un choix possible pour les suites a_n et b_n est :

$$a_n = x_F - F^{-1} \left(1 - \frac{1}{n} \right) \quad \text{et} \quad b_n = x_F.$$

Remarque 1.4. On déduit du théorème 1.3 que F appartient au domaine d'attraction de Weibull avec un indice de valeur extrême $\gamma < 0$ si et seulement si $x_F < +\infty$ et $1 - F = (x_F - x)^{-1/\gamma} \ell[(x_F - x)^{-1}]$.

Domaine d'attraction de Gumbel

Rappelons tout d'abord la définition d'une fonction de Von-Mises.

Définition 1.2. Soit F une fonction de répartition de point terminal x_F fini ou infini. S'il existe $z < x$ tel que

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

où $c > 0$ et a une fonction positive absolument continue de densité a' vérifiant $\lim_{x \uparrow x_F} a'(x) = 0$. Alors F est une fonction de Von-Mises et a est sa fonction auxiliaire.

Théorème 1.4. F appartient au domaine d'attraction de Gumbel si et seulement si il existe une fonction de Von-Mises F^* telle que pour $z < x < x_F$ on ait :

$$1 - F(x) = c(x) [1 - F^*(x)] = c(x) \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\},$$

où $c(x) \rightarrow c > 0$ lorsque $x \rightarrow x_F$

1.2 Classification des distributions

1.2.1 Classe de distributions à queue lourde

Définition 1.3. La fonction de répartition F d'une variable aléatoire X est dite à queue lourde (à droite) si pour tout $s > 0$:

$$E [e^{sX} \mathbb{I}_{\{X>0\}}] = \int_{(0,\infty)} e^{sx} dF(x) = \infty.$$

La première classe de distributions à queue lourde que nous considérons est la classe de distributions à queue longue définies comme suit.

Définition 1.4. Une fonction de répartition F de $(0, 1)$ telle que $F(x) < 1$ pour tout $x > 0$ appartient à une classe à queue longue \mathcal{L} , si pour tout $y \in \mathbb{R}$:

$$\lim_{x \rightarrow \infty} \frac{1 - F(x+y)}{1 - F(x)} = 1.$$

Remarque 1.5. La classe de distributions à queue longue peut effectivement être définie pour des distributions sur \mathbb{R} et sont alors dites à *queue longue* sur \mathbb{R} .

Remarque 1.6. Pour toute fonction de répartition F , l'hypothèse donnée $F(x) < 1$ pour tout $x > 0$ signifie que le support de F est sans limite ci-dessus.

La deuxième classe de distributions et très importantes de queue lourde que nous précisons maintenant est appelée la classe de distributions *sous-exponentielles*.

Définition 1.5. Une fonction de répartition F de $(0, 1)$ telle que $F(x) < 1$ pour tout $x > 0$ appartient à une classe sous-exponentielle S si :

$$\lim_{x \rightarrow \infty} \frac{1 - F^{*2}(x)}{1 - F(x)} = 2$$

Remarque 1.7. La notion de sous-exponentialité peut être étendue à des distributions sur \mathbb{R} comme suit : Une fonction de distribution G sur \mathbb{R} est appelée sous-exponentielle sur \mathbb{R} s'il existe $F \in S$ tels que $1 - G(x) \sim 1 - F(x)$ lorsque $x \rightarrow \infty$. La classe S est un sous-ensemble de la classe \mathcal{L} .

Exemples de distribution dans \mathcal{L} et pas dans S peuvent être trouvés dans Embrechts et Goldie (1980) et Pitman (1980). La classe des distributions sous-exponentielle a été introduite par Chistyakov (1964). Certaines propriétés intéressantes de cette classe de distribution sont données ci-dessous.

Lemme 1.1. Soit $F \in S$, alors les propriétés suivantes sont vérifiées :

(i) Pour tout $\epsilon > 0$, $\lim_{x \rightarrow \infty} e^{\epsilon x}(1 - F(x)) = \infty$.

(ii) Pour tout $\epsilon > 0$ donné, il existe une constante $K : K = K(\epsilon) < \infty$ tel que pour tout $x \geq 0$ et $n \in \mathbb{N}_0$:

$$\frac{1 - F^{*n}(x)}{1 - F(x)} \leq K(1 + \epsilon)^n.$$

(iii) Pour tout $n \in \mathbb{N}_0$:

$$\lim_{x \rightarrow \infty} \frac{1 - F^{*n}(x)}{1 - F(x)} = n. \tag{1.1}$$

Lemme 1.1 (i) justifie le nom de sous-exponentielle, comme $1 - F(x)$ décroît vers 0 lorsque $x \rightarrow \infty$ plus lent que n'importe quelle fonction exponentielle $e^{-\epsilon x}$ pour $\epsilon > 0$. Lemme 1.1(ii) est due à Kesten (pour une preuve, voir Athreya et Ney, 1972). Lemme 1.1 (iii) remonte à Chistyakov (1964).

Une distribution qui appartient à S ou à \mathcal{L} est à queue lourde où les valeurs extrêmes sont très dominantes parmi les valeurs de l'échantillon. Les distributions sous-exponentielles sont souvent considérées comme des modèles appropriés de distributions à queue lourde. L'importance de la classe de distributions S dans le cadre des probabilités appliquées, en général, et les mathématiques d'assurance, en particulier, a été réalisée par Teugels (1975). Le rôle potentiel des distributions sous-exponentielles dans la théorie de files d'attente a été reconnu par Pakes (1975) et Borovkov (1976, 1984). Pour plus d'approfondissements sur les distributions sous-exponentielles, nous nous référons à Embrechts et al. (1997), Rolski et al. (1999) et Goldie et Klüppelberg (1998).

1.2.2 Classe de distributions à variation régulière

Définition 1.6. Une fonction mesurable $f : (0, \infty) \rightarrow (0, \infty)$ est dite à variation régulière (à l'infini) avec l'index $\alpha \in \mathbb{R}$ (on écrit $f \in VR_\alpha$) si pour tout $u > 0$:

$$\lim_{x \rightarrow \infty} \frac{f(ux)}{f(x)} = u^\alpha. \quad (1.2)$$

Dans le cas $\alpha = 0$, la fonction f est dite à variation lente (à l'infini) et la notation ℓ est habituellement utilisé pour une telle fonction.

Remarque 1.8 Une fonction g positif et mesurable sur \mathbb{R} est dite à variation régulière sur \mathbb{R} , avec l'indice $\alpha \in \mathbb{R}$ si $g(x) \sim f(x)$ quand $x \rightarrow \infty$ pour une certaine $f \in VR_\alpha$. Nous écrivons $g \in VR_\alpha$ sur \mathbb{R} .

Remarque 1.9. Chaque fonction f à variation régulière avec indice $\alpha \in \mathbb{R}$ peut être représentée pour $x > 0$ comme $f(x) = x^\alpha \ell(x)$ avec $\ell \in VR_0$.

Les distributions avec une queue droite à variation régulière constituent une classe naturelle de distributions à queue lourde qui joue un rôle majeur dans les applications de la théorie des valeurs extrêmes.

Définition 1.7. Une fonction de répartition F sur \mathbb{R} appartient à une classe à variation régulière R s'il existe $\alpha \geq 0$ tel que $1 - F \in VR_{-\alpha}$ sur \mathbb{R} ,

ou d'une manière équivalente :

$$1 - F(x) \sim x^{-\alpha} \ell(x), \text{ quand } x \rightarrow \infty \quad (1.3)$$

pour certaines $\ell \in RV_0$.

La notion de variation régulière apparaît dans divers champs de probabilités appliquées comme dans la théorie des valeurs extrêmes, la théorie de point de processus, la théorie des files d'attente et la théorie du renouvellement de la somme de variables aléatoires. Notons que le travail de de Haan (1970) sur la variation régulière a été d'une grande importance pour le développement de la théorie des valeurs extrêmes moderne.

La famille queue lourde obtenue en limitant la classe de distributions R sur $(0, \infty)$ peut être classée selon que $\alpha = 0$ ou $\alpha > 0$. Quand $\alpha = 0$ dans (1.3), nous avons la définition suivante.

Définition 1.8. Une fonction de répartition F sur $(0, \infty)$ est dite super queue lourde si $1 - F \in VR_0$.

Remarque 1.10. Une distribution F super queue lourde a la représentation $1 - F(x) \sim \ell(x)$ quand $x \rightarrow \infty$ pour une fonction à variation lente qui doit donc satisfaire $\ell(x) = o(1)$, quand $x \rightarrow \infty$.

Le cas $\alpha > 0$ dans (1.3) conduit à la célèbre famille de distributions de type Pareto qui joueront un rôle de premier plan dans cette thèse. La définition formelle est la suivante.

Définition 1.9. Une fonction de répartition F sur $(0, \infty)$ est dite de type Pareto avec l'indice $\alpha > 0$ si $1 - F \in VR_{-\alpha}$.

Remarque 1.11. Toute suite de variables aléatoires positives X dont la fonction de distribution F est de type Pareto avec indice $\alpha > 0$ a un moment d'ordre $\delta > 0$ qui satisfait :

$$E[X^\delta] = \delta \int_0^\infty x^{\delta-1} (1 - F(x)) dx \begin{cases} < \infty & \text{si } \delta < \alpha \\ = \infty & \text{si } \delta > \alpha. \end{cases}$$

Il est bien connu que toute distribution super queue lourde ou de type Pareto est sous-exponentielle (par exemple, Corollaire 2(i) de Teugels, 1975). Exemples de distributions dans S mais pas dans R sont la distribution lognormale $LN(\mu, \sigma^2)$ avec $\mu \in \mathbb{R}$ et $\sigma > 0$ et la distribution de Weibull $W(\tau, c)$ avec $\tau \in (0, 1)$ et $c > 0$ (voir le tableau 1.1). La classe R restreinte à des distributions sur $(0, \infty)$ est donc un sous-ensemble de la classe S .

La distribution F	La queue $1 - F$ ou la densité f	Paramètres
Burr ($burr(\alpha, \varkappa, \tau)$)	$1 - F(x) = \left(\frac{\varkappa}{\varkappa + x^\tau}\right)^\alpha$	$\alpha, \varkappa, \tau > 0$
Loggamma ($L\Gamma(\alpha, \beta)$)	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (\log(x))^{\alpha-1} x^{-(\beta+1)}$	$\alpha > 0, \beta > 0$
Pareto ($Par(\alpha, \varkappa)$)	$1 - F(x) = \left(\frac{\varkappa}{\varkappa + x}\right)^\alpha$	$\alpha > 0, \varkappa > 0$
Lognormal ($LN(\mu, \sigma^2)$)	$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} e^{-(\log x - \mu)^2 / (2\sigma^2)}$	$\mu \in \mathbb{R}, \sigma > 0$
Weibull ($W(\tau, c)$)	$1 - F(x) = \exp(-cx^\tau)$	$\tau \in (0, 1), c > 0$

TAB. 1.1 – Quelques distributions pour modéliser les plus grandes revendications des grands sinistres

Pour des preuves et des propriétés supplémentaires sur les fonctions à variation régulière, nous nous référons à Seneta (1976), Bingham et al. (1987) et Geluk et de Haan (1987).

Théorème 1.5. Une fonction $\ell : (0, \infty) \rightarrow (0, \infty)$ est à variation lente si et seulement si elle peut être écrite comme :

$$\ell(x) = c(x) \exp \left\{ \int_{x_0}^x \frac{\varepsilon(u)}{u} du \right\}, \quad x \geq x_0,$$

pour certain $x_0 > 0$, où $\lim_{x \rightarrow \infty} c(x) = c \in (0, \infty)$ et $\lim_{x \rightarrow \infty} \varepsilon(x) = 0$.

Des exemples typiques de fonctions à variation lente sont des constantes positives ou des fonctions convergeant vers une constante positive, logarithmes, puissances des logarithmes et logarithmes itérés.

Remarque 1.12. Il résulte du remarque 1.9 et du théorème 1.5 que la fonction $f : (0, \infty) \rightarrow (0, \infty)$ est à variation régulière avec indice $\alpha \in \mathbb{R}$ si et seulement si elle a la représentation :

$$f(x) = d(x) \exp \left\{ \int_{x_0}^x \frac{\alpha(u)}{u} du \right\}, \quad x > x_0,$$

pour un certain $x_0 > 0$, où $\lim_{x \rightarrow \infty} d(x) = d \in (0, \infty)$ et $\lim_{x \rightarrow \infty} \alpha(x) = \alpha$. Il résulte aussitôt de la Remarque 1.11 que si $f \in VR_\alpha, \alpha \neq 0$, alors :

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

Cela montre en particulier que si $\ell \in VR_0$, alors pour tout $\varepsilon > 0$:

$$\lim x^{-\varepsilon} \ell(x) = 0 \text{ et } \lim x^\varepsilon \ell(x) = \infty.$$

Proposition 1.1.

- (i) Si $f \in VR_\alpha$ ($\alpha \in \mathbb{R}$) alors $f^\rho \in VR_{\rho\alpha}$, pour tout $\rho \in \mathbb{R}$.
- (ii) Si $f_1 \in VR_{\alpha_1}$ ($\alpha_1 \in \mathbb{R}$), $f_2 \in VR_{\alpha_2}$ ($\alpha_2 \geq 0$) et $\lim_{x \rightarrow \infty} f_2(x) = \infty$, alors $f_1 \circ f_2 \in VR_{\alpha_1\alpha_2}$.
- (iii) Si $f_1 \in VR_{\alpha_1}$ ($\alpha_1 \in \mathbb{R}$), $f_2 \in VR_{\alpha_2}$ ($\alpha_2 \in \mathbb{R}$), alors $f_1 f_2 \in VR_{\alpha_1 + \alpha_2}$ et $f_1 + f_2 \in VR_\alpha$ où $\alpha := \max(\alpha_1, \alpha_2)$
- (iv) Si $f_i \in VR_{\alpha_i}$ avec $\alpha_i \in \mathbb{R}$ ($i = 1, \dots, k$) et $r(x_1, \dots, x_k)$ est une fonction rationnelle avec coefficients positifs, alors $r(f_1(x), \dots, f_k(x))$ est à variation régulière.

Le concept d'une fonction variation régulière à 0 est défini formellement comme suit.

Définition 1.10. Une fonction mesurable $f : (0, \infty) \rightarrow (0, \infty)$ est dite à variation régulière (sur la droite) à 0 avec l'indice $\alpha \in \mathbb{R}$ (on écrit $f \in VR_\alpha^0$) si pour tout $u > 0$:

$$\lim_{x \downarrow 0} \frac{f(ux)}{f(x)} = u^\alpha.$$

Dans le cas $\alpha = 0$, la fonction f est dite à variation lente à 0.

Notez que pour chaque $\alpha \in \mathbb{R}$, $f \in VR_\alpha^0$ si et seulement si $g \in VR_{-\alpha}$ où $g(x) := f(1/x)$ pour tout $x > 0$.

1.2.3 Classe de distributions extrémales

Les distributions des classes extrémales sont très utiles dans les applications de la théorie des valeurs extrêmes. En fait, l'ensemble de ces distributions coïncide avec l'ensemble de toutes les distributions qui appartiennent au domaine d'attraction d'une distribution des valeurs extrêmes.

Définition 1.11. Une fonction de répartition F sur \mathbb{R} avec la fonction quantile de queue U appartient à la classe extrême C_γ , $\gamma \in \mathbb{R}$, s'il existe une fonction mesurable $a : (0, \infty) \rightarrow (0, \infty)$ telle que pour tout $u > 0$:

$$\lim_{x \rightarrow \infty} \frac{U(ux) - U(x)}{a(x)} = h_\gamma(u) := \begin{cases} \frac{u^\gamma - 1}{\gamma} & \text{si } \gamma \neq 0 \\ \log u & \text{si } \gamma = 0 \end{cases} \quad (1.4)$$

La fonction a apparaissant dans (1.4) est appelée une fonction auxiliaire pour U . Notez que toute fonction $a_1 : (0, \infty) \rightarrow (0, \infty)$ est une fonction auxiliaire de U si et seulement si $a_1(x) \sim a(x)$ lorsque $x \rightarrow \infty$, ce qui montre que a est unique à une équivalence asymptotique. Nous écrivons $F \in C_\gamma(a)$ si (1.4) est vérifié avec a fonction auxiliaire.

Remarque 1.13. La limite sur la droite de (1.4) n'a pas besoin d'être spécifiée. En effet, si on suppose seulement que $\lim_{x \rightarrow \infty} \{U(ux) - U(x)/a(x)\}$ existe pour tout $u > 0$ et n'est pas constante à u , alors $a \in VR_\gamma$ pour certains $\gamma \in \mathbb{R}$, et la limite est donnée par la fonction h_γ (par exemple, le théorème 1.9 Geluk et de Haan, 1987).

de Haan (1984) a montré que pour chaque $\gamma \in \mathbb{R}$, la condition $F \in C_\gamma$ est nécessaire et suffisante pour que F appartienne au domaine d'attraction de distribution des valeurs extrêmes G_γ , on entend par là qu'il existe des constantes de normalisation $b_n > 0$ et $a_n \in \mathbb{R}$ telle que :

$$b_n^{-1}(X_{n,n} - a_n) \xrightarrow{D} G_\gamma \quad \text{quand } n \rightarrow \infty \quad (1.5)$$

Si $F \in C_\gamma$, $\gamma \in \mathbb{R}$, alors, pour chaque $i \in \mathbb{N}$, la i -ième plus grand statistique d'ordre normalisée $(X_{n-i+1,n} - b_n)/a_n$ converge en distribution quand $n \rightarrow \infty$ à une variable aléatoire non dégénérée. Cela vaut en particulier avec les mêmes constantes de normalisation a_n et b_n comme pour le maximum partiel $X_{n,n}$ dans (1.5).

La condition de classe extrémale (1.4) reformule la condition de valeur extrême (1.5) en termes de la fonction quantile de queue U de F , montrant à ce titre, le rôle éminent joué par U dans la théorie des valeurs extrêmes. Par ailleurs, il s'avère que la relation (1.4) s'inscrit dans le cadre de la variation régulière généralisée (du premier ordre), voir le chapitre 3 de Bingham et al. (1987).

Proposition 1.2. Supposons que $F \in C_\gamma(a)$ pour certains $\gamma \in \mathbb{R}$. Donc $a \in VR_\gamma$ et :

- (i) Si $\gamma > 0$: $a(x)/U(x) \rightarrow \gamma$ quand $x \rightarrow \infty$ et $U \in VR_\gamma$ ou de manière équivalente $1 - F \in VR_{-1/\gamma}$.
- (ii) Si $\gamma < 0$: $U(\infty) < \infty$, $a(x)/(U(\infty) - U(x)) \rightarrow |\gamma|$ quand $x \rightarrow \infty$ et $U(\infty) - U \in VR_\gamma$.
- (iii) Si $\gamma = 0$: $a(x)/U(x) \rightarrow 0$ quand $x \rightarrow \infty$ et dans le cas où $U(\infty)$ est finie, aussi $a(x)/(U(\infty) - U(x)) \rightarrow 0$ quand $x \rightarrow \infty$ ailleurs $|U| \in VR_0$.

Remarque 1.14. On peut montrer à l'aide de la proposition 1.2 que :

$$F \in C_\gamma(a) \implies \begin{cases} a/U \in VR_\gamma & \text{si } \gamma < 0 \text{ et } U(\infty) \neq 0 \\ a/U \in VR_0 & \text{si } \gamma \geq 0 \text{ ou } U(\infty) = 0 \end{cases}$$

Le résultat suivant dû à de Haan (1970) offre des conditions alternatives pour une fonction de répartition F d'appartenir à une classe spécifique extrême.

En particulier, il offre la possibilité de passer de propriétés extrémales de F à celle de sa fonction quantile de queue U et inversement. Voir aussi la section 2.6 de Beirlant et al. (1980).

Théorème 1.6. Soit $\gamma \in \mathbb{R}$ et $a : (0, \infty) \rightarrow (0, \infty)$ une fonction mesurable. Alors les assertions suivantes sont équivalentes :

(i) $F \in C_\gamma(a)$;

(ii) Pour tout $u \in \mathbb{R}$ pour lesquels $1 + \gamma u > 0$:

$$\lim_{x \uparrow x_r} \frac{1 - F(x + h(x)u)}{1 - F(x)} = \frac{1}{h_\gamma^{-1}} =: \eta_\gamma(u)$$

où $h \circ U = a$;

(iii) Pour tout $u \in \mathbb{R}$ pour lesquels $1 + \gamma u > 0$:

$$\lim_{x \rightarrow \infty} x \{1 - F(U(x) + a(x)u)\} = \eta_\gamma(u).$$

Il résulte de la définition de la fonction h_γ que $\eta_\gamma(u) = (1 + \gamma u)^{-1/\gamma}$ pour $1 + \gamma u > 0$ si $\gamma \neq 0$ alors que $\eta_0(u) = e^{-u}$ pour $u \in \mathbb{R}$ et $\gamma = 0$.

S'il existe $\gamma \in \mathbb{R}$ telle que (1.5) est vérifiée pour les constantes de normalisation $a_n = a(n)$ et $b_n = U(n)$, où $a : (0, \infty) \rightarrow (0, \infty)$ est une fonction mesurable, alors $F \in C_\gamma(a)$ puisque le théorème 1.6(ii) se révèle être vrai. L'inverse peut être également montré.

Chapitre 2

Analyse exploratoire des observations extrêmes

Il est important d'examiner les données avant de s'engager dans une analyse statistique détaillée. Dans notre époque de la puissance de calcul quasi illimitée l'exploration graphique des données est de plus en plus importante. Compte tenu des données, un praticien veut utiliser des graphiques qui montrent de manière claire et efficace les caractéristiques des données qui sont pertinentes pour une question de recherche donnée. Avant d'appliquer une étude statistique théorique sur un échantillon de données réelles, il faut d'abord montrer que ces données proviennent d'une famille de loi de probabilité, au préalable considérée comme base dans notre étude de statistique à présenter. Une analyse exploratoire des données est souvent utilisée pour tester la qualité de l'ajustement des observations de l'échantillon aux distributions cibles spécifiques. Quelques outils graphiques tels, ont été largement utilisés pour détecter les comportements sous-exponentielle ou à queue lourde dans les données observées, notamment dans notre cas, des méthodes pour examiner l'existence des valeurs extrêmes. Au début, nous présentons le test d'adéquation de Kolmogorov-Smirnov aux distributions de type Pareto qui pourra servir comme un outil de vérification dans des applications réelles. Aussi, nous nous concentrons sur des techniques statistiques visuelles qui fournissent autant d'informations que possible au sujet de la queue d'une distribution. Ces derniers nous aideront à décider sur un modèle raisonnable de décrire des risques réels extrêmes. Nous nous concentrerons sur les *Quantile-Quantile Plot* et *Mean Excess Plot*, qui sont souvent plus instructif pour nos besoins. En ce qui concerne l'étude asymptotique des ces techniques graphiques voir par exemples les derniers recherches comme Das et Ghosh (2013).

Tests de kolmogorov-smirnov.

Le choix de la loi est très important dans la détermination des caractéristiques statistiques des données utilisées. Le résultat théorique le plus connu est dû à Kolmogorov (1933) et Smirnov (1948), et a conduit à un test statistique éponyme noté KS . Le test de Kolmogorov-Smirnov est un moyen pour mesurer la distance entre la fonction de répartition empirique et la fonction de répartition F . Cantelli et Glivenko ont montré que la statistique de Kolmogorov-Smirnov converge presque sûrement vers 0 (voir par exemple, Serfling, 1980). Il existe plusieurs livres et références sur ce test, on cite à titre d'exemple, Thas (2010), Frank et Massey (1951).

Principe du test

Dans ce test les calculs sur les lois de probabilité se font sur les fonctions de répartition : On mesure l'écart entre la fonction de répartition théorique et la fonction de répartition observée. On considère ainsi une variable aléatoire X de fonction de répartition F_n que l'on veut comparer à une fonction de répartition théorique F continue. On souhaite tester l'hypothèse.

$$H_0 : F_n = F$$

Contre l'hypothèse

$$H_1 : F_n \neq F$$

Si X_1, \dots, X_n est un n -échantillon de X , la fonction de répartition empirique associée à cet échantillon est :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{]-\infty, x[}(X_i), \quad x \in \mathbb{R}$$

C'est une fonction en escalier, croissante, qui saute de $\frac{1}{n}$ en chaque point de l'échantillon. L'écart entre les valeurs observées et les valeurs théoriques du modèle déduites de la fonction de répartition F peut donc être mesuré par la variable aléatoire appelée aussi la statistique de Kolmogorov-Smirnov :

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \tag{2.1}$$

Une valeur élevée de D_n est une indication que la distribution de l'échantillon s'éloigne sensiblement de la distribution de référence $F(x)$, et qu'il est donc peu probable que H_0 soit correcte. Plus précisément, et pour n grand

$$\lim_{n \rightarrow \infty} p(\sqrt{n}D_n \leq c) = 1 - 2 \sum_{i=1}^{+\infty} (-1)^{i-1} \exp(-2i^2 c^2)$$

Ainsi pour tout $0 < \alpha < 1$, on peut déterminer la valeur c telle que, pour n grand

$$P(D_n \leq c/\sqrt{n}) = 1 - \alpha,$$

ce qui revient encore à dire que,

$$P\left(\forall x \in \mathbb{R}, \quad F_n(x) - \frac{c}{\sqrt{n}} \leq F(x) \leq F_n(x) + \frac{c}{\sqrt{n}}\right) = 1 - \alpha$$

L'encadrement

$$F_n(x) - \frac{c}{\sqrt{n}} \leq F(x) \leq F_n(x) + \frac{c}{\sqrt{n}} \quad (2.2)$$

est donc vrai pour tout $x \in \mathbb{R}$, avec la probabilité $1 - \alpha$. la quantité $1 - \alpha$ est appelée *le niveau de confiance de l'encadrement* (2.2) et α est appelée *le risque d'erreur de première espèce*, car c'est la probabilité que (2.2) ne soit pas vraie. Comme nous travaillons sur des échantillons de taille très élevée on note que pour $n > 100$, la valeur critique du test est approximativement de la forme $\frac{c}{\sqrt{n}}$. Les valeurs usuelles de c en fonction de α sont :

α	0.20	0.10	0.05	0.02	0.01
c	1.073	1.223	1.358	1.518	1.629

TAB. 2.1 – Table de test de Smirnov-Kolmogorov.

Calcul de la statistique de Kolmogorov-Smirnov

Nous allons maintenant vérifier que D_n est facile à calculer. Comme F_n est la fonction de répartition empirique observée sur les n observations X_i , à partir de la loi et de ses paramètres, nous disposons de la fonction de répartition théorique F que nous pouvons appliquer sur chaque observations X_i . Le supremum dans la définition de D_n peut être calculé efficacement à partir des statistiques d'ordre $X_{1,n}, \dots, X_{n,n}$ comme suit :

$$D_n = \max_{0 \leq i \leq n} \max \left(\left| F(X_{i,n}) - \frac{i-1}{n} \right|, \left| F(X_{i,n}) - \frac{i}{n} \right| \right) \quad (2.3)$$

Des tests de qualité d'ajustement précis pour les queues extrêmes de distributions empiriques est une question très importante, pertinente dans de nombreux contextes, y compris la géophysique, l'assurance et la finance. Récemment, Chicheportiche et Bouchaud (2012) ont tiré des résultats asymptotiques exacts pour une généralisation du test de Kolmogorov-Smirnov pour des échantillons grand, bien adapté pour tester ces queues extrêmes.

Nous présentons maintenant, le déroulement du test de Kolmogorov-Smirnov, pour une distribution de type Pareto qui pourra être utile dans une application réelle pour tester la queue d'un échantillon donné.

Test de Kolmogorov-Smirnov d'une distribution de type Pareto avec estimation graphique de l'indice γ

Si on dispose d'un échantillon de données réelles X_1, \dots, X_n , et on cherche à vérifier si ce dernier provient d'une loi de distribution de Pareto $F(x) = 1 - x^{-1/\gamma}$, pour tout $0 < \gamma < 3$, et dans le but d'estimer en même temps la valeur du paramètre du queue γ on procède comme suit :

1. **On ordonne l'échantillon dans un ordre croissant.** pour obtenir les statistiques d'ordre correspondantes $X_{1,n}, \dots, X_{n,n}$
2. **On calcul la statistique de Kolmogorov-Smirnov.** pour tout $0 < \gamma < 3$:

$$D = \min_{0 < \gamma < 3} D_n = \min_{0 < \gamma < 3} \left(\max_{0 \leq i \leq n} \max \left(\left| F(X_{i,n}) - \frac{i-1}{n} \right|, \left| F(X_{i,n}) - \frac{i}{n} \right| \right) \right)$$

$$= \min_{0 < \gamma < 3} \left(\max_{0 \leq i \leq n} \max \left(\left| 1 - (X_{i,n})^{-1/\gamma} - \frac{i-1}{n} \right|, \left| 1 - (X_{i,n})^{-1/\gamma} - \frac{i}{n} \right| \right) \right)$$

3. **La précision du niveau de confiance du test.** avec un niveau de confiance, par exemple $\alpha = 0.05$, la valeur critique du test est $\frac{c}{\sqrt{n}} = \frac{1.358}{\sqrt{n}}$, (voir le tableau 2.1).
4. **La prise de décision.** si $D_n \leq \frac{1.358}{\sqrt{n}}$ avec un niveau de confiance $\alpha = 0.05$, on accepte l'hypothèse H_0 que l'échantillon réelle provient de la distribution de pareto de paramètre γ (estimé d'ailleurs par le test *KS*), sinon on rejette.
5. **Pour effectuer le test *KS* sous l'environnement *R*, on peut se servir directement de la fonction *ks.test*.** Comme il est possible de faire le test à la main lorsque il s'agit d'une nécessité particulière, ainsi ça se déroule comme suit.

```
x<-sort(x) # on ordonne l'échantillon des observations  $X_i$ 
u<-(1:n)/n # vecteur des valeurs  $i/n$ 
v<-(0:(n-1))/n # vecteur des valeurs  $(i-1)/n$ 
gama<-seq(1,100)
Dn<-seq(1.100)
for(j in 1:100)
{
```

```

gama[j]<-0.03*j
Fx<-1-(x^(-1/gama[j])) # vecteur des valeurs max entre
#les deux arguments
diff<-pmax(abs(Fx-u), abs(Fx-v))
Dn[j]<-max(diff)
}
D<-min(Dn)
gamma<-gama[which.min(Dn)]
gamma # l'estimation du paramètre  $\gamma$  de la distribution de Pareto
vcr<-1.35/sqrt(n) #valeur critique à ne pas dépasser ( $\frac{c}{\sqrt{n}}$ )
#par D pour accepter que les données proviennent
#d'une distribution de pareto de paramètre gamma
#avec un niveau de confiance  $1 - \alpha = 0.05$ 

```

2.1 Quantile-quantile plot

2.1.1 Introduction

Le *QQ – plot* est une technique de représentation graphique apparue dans les années 1960 (voir par exemple Kimball, 1960). Elle est couramment utilisée de manière informelle pour décider si un échantillon univarié aléatoire de taille n provient d'une distribution spécifiée F . Il s'agit de tracer les quantiles d'échantillon contre les quantiles théoriques de F puis un contrôle visuel est effectué pour voir si oui ou non les points présentent une allure linéaire. Ainsi, la linéarité dans une courbe peut être facilement vérifié par l'œil et peut en outre être quantifiée au moyen d'un coefficient de corrélation. Pour une compréhension complémentaire vous pouvez vous référer aux livres (Beirlant et al. 2004, Embrechts et al. 1997, Thas, 2010).

Concernant des distributions à queue lourde, Il existe beaucoup d'application, introduisant les *QQ – plots* comme un outil graphique d'aide pour un choix convenable d'un modèle (voir par exemple Das et al. 2008).

Dans notre cas on peut aussi citer par exemple les travaux de Cebrian et al. (2004). Ils ont examiné une stratégie de modélisation statistique basée sur la théorie des valeurs extrêmes pour décrire le comportement d'un portefeuille d'assurance, avec un accent particulier sur les gros sinistres. La stratégie est illustrée en utilisant les *QQ – plots* par des réclamations de 1991 – 1992 d'un groupe médical d'une base de données gérée par la Société des Actuaires.

2.1.2 QQ-plot d'une distribution exponentielle

Dans la théorie des valeurs extrêmes, le *QQ-plot* se base sur la distribution exponentielle, qui joue un rôle beaucoup plus important pour notre propos. Le *QQ-plot* sous l'hypothèse d'une distribution exponentielle est la représentation des quantiles de la distribution empiriques sur l'axe des Y contre les quantiles de la fonction de distribution exponentielle sur l'axe des X . Nous commençons par expliquer l'idée du *QQ-plot* pour le modèle exponentiel $Exp(\lambda)$, cette même méthodologie peut alors être explorée et étendue afin de permettre des comparaisons de données empiriques lorsque on veut les ajuster aux modèles tels que le log-normale, pareto ou autres. Nous proposons la distribution exponentielle standard

$$1 - F_1(x) := \exp(-x), x > 0$$

l'exemple-type de la classe de distributions avec la fonction de survie général est

$$1 - F_\lambda(x) := \exp(-\lambda x).$$

nous voulons savoir si la population réelle de distribution F appartient à cette classe, paramétrée par $\lambda > 0$. La réponse est de s'appuyer sur les données X_1, \dots, X_n que nous avons à notre disposition. Il est important de noter que ce paramètre peut être considéré comme un paramètre de nuisance ici car sa valeur n'est pas notre principal point d'intérêt à ce moment. La fonction quantile de la distribution exponentielle a la forme simple

$$Q_\lambda(p) = -\frac{1}{\lambda} \log(1-p), \text{ pour } p \in (0,1).$$

On a de la même façon.

$$Q_1(p) = -\log(1-p), \text{ pour } p \in (0,1).$$

Par conséquent, il existe une relation linéaire simple entre les quantiles d'une distribution exponentielle et les quantiles d'une distribution exponentielle standard.

$$Q_\lambda(p) = \frac{1}{\lambda} Q_1(p) \quad \text{pour } p \in (0,1).$$

A partir de notre échantillon X_1, X_2, \dots, X_n , on estime la fonction quantile Q par le quantile empirique Q_n et dans un système de coordonnées orthogonales, les points de valeurs

$$(-\log(1-p), Q_n(p)).$$

sont tracées pour différentes valeurs de $p \in (0,1)$. Nous attendons alors qu'un motif de ligne droite apparaîtra dans le nuage de points si le modèle exponentiel permet un ajustement statistique plausible pour la population statistique

donnée. Quand un motif de ligne droite est obtenue, la pente de la droite de régression peut être utilisée comme une estimation du paramètre λ . En effet, si le modèle est correct, alors l'équation

$$Q_\lambda(p) = \frac{1}{\lambda} (-\log(1-p)).$$

est valable. Remarquons que le point d'intersection pour le modèle donné doit être égal à 0 quand $Q_\lambda(0) = 0$. En général

$$Q_n(p) = X_{i,n}, \quad \text{pour } \frac{i-1}{n} < p < \frac{i}{n}$$

Un choix très pratique de valeurs de p est donnée par $p_{i,n}$:

$$p_{i,n} \in \left\{ \frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}, 1 \right\}$$

Pour obtenir finalement, les coordonnées

$$(-\log(1-p_{i,n}), X_{i,n}). \quad (2.4)$$

Procédure sous R et explication du graphique.

Pour avoir le graphique *QQ - plot* des données X_1, X_2, \dots, X_n contre les quantiles exponentielles, on exécute le programme suivant sous R.

```
x<-sort(x) # on ordonne l'échantillon des observations  $X_i$ 
p<-(1:n)/(n+1) # vecteur des valeurs  $i/(n+1)$ 
Droite<-lm((-log10(1-p))~x)
coef(Droite)
coef(Droite)[1]
coef(Droite)[2] # estimation du paramètre  $\lambda$ 
qqplot(x,-log10(1-p), xlab="Theoretical Quantiles",
ylab="Sample Quantiles"
, main = "Exponential Q-Q plot")
abline(coef(Droite), col="3", lwd="3")
```

L'intérêt de ce graphique est de nous permettre d'obtenir la forme de la queue de la distribution. Trois cas de figure sont possibles :

1. Les données suivent la loi exponentielle : la distribution présente une queue très légère, les points du graphique présentent une forme linéaire.
2. Les données suivent une distribution à queue lourde "*heavy-tailed distribution*" : le graphique *QQ - plot* est concave
3. Les données suivent une distribution à queue légère "*short-tailed distribution*" : le graphique *QQ - plot* a une forme convexe.

2.1.3 QQ-plot d'une distribution de Pareto

Nous présentons maintenant le *QQ - plot* d'une distribution de Pareto de paramètre $1/\gamma$, qui pourra servir comme un test d'ajustement. Pour le même échantillon de données réelles X_1, \dots, X_n , on veut savoir si ces dernières proviennent d'une distribution de Pareto définie telle que pour tout $\gamma > 0$:

$$F_\gamma(x) = 1 - x^{-1/\gamma}.$$

Ainsi pour $\gamma = 1$

$$F_1(x) = 1 - x^{-1}.$$

La fonction quantile de la distribution de Pareto a la forme simple

$$\log(Q_\gamma(p)) = -\gamma \log(1 - p), \quad \text{pour } p \in (0, 1)$$

et de la même manière on a :

$$\log(Q_1(p)) = -\log(1 - p), \quad \text{pour } p \in (0, 1).$$

Par conséquent, il existe une relation linéaire simple entre les quantiles d'une distribution de Pareto et les quantiles d'une distribution de Pareto standard.

$$\log(Q_\gamma(p)) = -\gamma \log(Q_1(p)), \quad \text{pour } p \in (0, 1).$$

De la même façon qu'on a fait avec la distribution exponentielle on estime Q par Q_n . Dans un système de coordonnées orthogonales, les points de valeurs

$$(-\log(1 - p), \log(Q_n(p)))$$

sont tracées pour différentes valeurs de $p \in (0, 1)$. Nous attendons alors qu'un motif de ligne droite apparaîtra dans le nuage de points si le modèle de Pareto permet un ajustement statistique plausible pour la population statistique donnée. Quand un motif de ligne droite est obtenue, la pente de la droite de régression peut être utilisée comme une estimation du paramètre $1/\gamma$. En effet, si le modèle est correct, alors l'équation

$$\log(Q_\gamma(p)) = \gamma(-\log(1 - p))$$

est valable.

Procédure sous R et estimation graphique de γ si l'ajustement par une distribution pareto est correcte.

En variant les valeur de p sur l'ensemble $\left\{ \frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}, 1 \right\}$, les coordonnées

$$(-\log(1 - p_{i,n}), \log(X_{i,n})) \quad (2.5)$$

nous permettre de tracer le $QQ - plot$ d'une distribution de pareto contre notre échantillon de données X_1, \dots, X_n . On peut facilement obtenir ce graphique sous l'environnement R en exécutant le programme suivant.

```
x<-sort(x) # on ordonne l'échantillon des observations  $X_i$ 
p<-(1:n)/(n+1) # vecteur des valeurs  $i/(n+1)$ 
Droite<-lm((-log10(1-p))~log10(x))
coef(Droite)
coef(Droite)[1]
1/coef(Droite)[2] # estimation numérique de  $\gamma$ 
qqplot(log10(x),-log10(1-p), xlab="Theoretical Quantiles",
ylab="Sample Quantiles"
, main = "Pareto Q-Q plot")
abline(coef(Droite), col="3", lwd="3")
```

2.2 Mean excess plot

Le graphique *mean excess plot* est largement utilisé dans des articles d'applications ou de simulation, il est à la fois un graphique qui aide à déterminer le nombre de valeur extrêmes et examiner si la distribution d'un échantillon est à queue lourde ou non, on explique dans cette section la construction de ce graphique.

2.2.1 Loi des excès

Définition 2.1. *La fonction de répartition des excès au-delà d'un seuil u pour une variable aléatoire X avec fonction de répartition F est définie par :*

$$F_u(x) = P[X - u \leq x | X > u]. \quad (2.6)$$

Cette *distribution des excès* est un outil largement utilisé en hydrologie, en analyse de survie et en sciences de l'environnement et d'une importance majeur dans la pratique actuarielle reconnue (voir par exemple Hogg et Klugman, 1984; Embrechts et al. 2005). Notamment en matière de réassurance, si

on prend par exemple un traité en excédent de perte (excess-of-loss) avec une rétention u sur une demande particulière dans le portefeuille, le réassureur doit payer un montant aléatoire $X - u$, mais seulement si $X > u$. Quand un actuaire veut se prononcer sur un niveau de priorité grâce à la simulation, il a besoin de calculer le montant qui devrait être payé par le client quand un niveau donné u est choisi. Voilà donc une première étape importante pour décider sur la prime. Par exemple, le principe de la prime nette dépend de la taille moyenne des réclamations $E(x)$, où pour le dépassement, l'actuaire a besoin de calculer *La fonction moyenne des excès*. Cette dernière constituera la base de la construction du graphique *mean excess plot*. On donnera plus de détails sur cette fonction dans ce qui suit. Ainsi, on présente tout d'abord, le modèle *POT* et principalement la fonction *GPD* qui nous aidera pour examiner la queue de la distribution de l'échantillon réelle en question et notamment la détermination du seuil u .

Le modèle à dépassement de seuil (Peaks Over Threshold).

Le principe de la méthode *POT* est d'estimer la distribution des excès au-delà d'un seuil u par une classe de distributions de pareto généralisée *GPD* dépendant des deux paramètres γ et $\beta(u)$, étant appropriée pour décrire les propriétés statistiques des excès. *Une variable aléatoire X a une distribution de pareto généralisée (GPD) si elle a une fonction de répartition de la forme*

$$G_{\gamma,\beta}(x) = \begin{cases} 1 - \left(1 + \gamma \frac{x}{\beta}\right)^{-1/\gamma} & \text{si } \gamma \neq 0 \\ 1 - \exp\left(-\frac{x}{\beta}\right) & \text{si } \gamma = 0 \end{cases} \quad (2.7)$$

où $\beta > 0$, et $x \geq 0$ quand $\gamma \geq 0$ et $0 \leq x \leq -\beta/\gamma$ si $\gamma < 0$. Les paramètres γ et β sont désignés respectivement comme les paramètres de la forme et d'échelle.

On a les correspondances suivantes :

$$\begin{aligned} \text{si } \gamma > 0 & \quad G_{\gamma,\beta} \rightarrow \text{loi pareto} \\ \text{si } \gamma < 0 & \quad G_{\gamma,\beta} \rightarrow \text{loi béta} \\ \text{si } \gamma = 0 & \quad G_{\gamma,\beta} \rightarrow \text{loi exponentielle} \end{aligned}$$

Le théorème Pickands-Belkema-de Haan (Embrechts et al. 2005, Theorem 7.20, page 277) fournit la justification théorique de la centralité de la classe des distributions *GPD* pour la modélisation *POT*. Ce résultat montre que, pour une large classe de distributions, la distribution des excès F_u est asymptotiquement équivalente à la loi $G_{\gamma,\beta(u)}$, quand le seuil u s'approche de l'ex-

trémité droite de la distribution F .

$$\lim_{u \rightarrow x_+} \sup_{0 < x < x_+ - u} |F_u(x) - G_{\gamma, \beta(u)}(x)| = 0 \quad (2.8)$$

Le paramètre asymptotique de forme γ est fixe, mais celui de l'échelle $\beta(u)$ peut dépendre de u . La fonction moyenne des excès est un outil couramment utilisé d'aide au choix de u et aussi pour déterminer la pertinence du modèle *GPD* dans la pratique.

2.2.2 Fonction moyenne des excès théorique $e(u)$

Définition 2.2. Pour une variable aléatoire X satisfaisant $E(X) < \infty$, avec fonction de répartition F avec extrémité droite x_F et la queue $\bar{F}(x) = 1 - F(x)$, la fonction moyenne des excès est

$$e(u) := E[X - u \mid X > u] = \frac{\int_u^{x_F} \bar{F}(s) ds}{\bar{F}(u)}, u < x_F. \quad (2.9)$$

Egalement connue comme la fonction moyenne résiduelle de durée de vie, en particulier dans l'analyse de survie. Il a été étudié dès Benktander et Segerdahl (1960). Pour une discussion de ses propriétés voir Hall et Wellner (1981). Notamment le tableau 3.4.7 dans Embrechts et al. (1997) donne la fonction moyenne des excès pour certaine distribution standard s'avéreront utile dans la pratique dans le cas où on connaît la distribution de notre échantillon de données. Dans le cas contraire on aura besoin d'estimer $e(u)$.

2.2.3 Fonction moyenne des excès empirique $\hat{e}_n(u)$

Etant donné un échantillon de données X_1, \dots, X_n . Généralement en pratique on ne connaît pas la distribution de l'échantillon en question. Pour cette raison une estimation naturelle de $e(u)$ est sa fonction empirique $\hat{e}_n(u)$.

Définition 2.3. la fonction moyenne empirique des excès est la somme des excès dépassant un certain seuil élevé noté u , divisé par le nombre de points des données qui dépassent ce seuil, soit :

$$\hat{e}_n(u) = \frac{\sum_{i=1}^n (X_i - u) \mathbb{I}_{\{X_i > u\}}}{\sum_{i=1}^n \mathbb{I}_{\{X_i > u\}}}, u > 0. \quad (2.10)$$

Avec

$$\mathbb{I}_{\{X_i > u\}} = \begin{cases} 1 & \text{si } X_{i,n} > u \\ 0 & \text{sinon} \end{cases}$$

et la convention $0/0=0$.

Soit $X_{1,n}, \dots, X_{n,n}$ les statistiques d'ordre correspondantes de l'échantillon X_1, \dots, X_n . Un test graphique pour le comportement de la queue peut s'appuyer sur la fonction moyenne empirique des excès appelé *mean excess plot* donné par le tracé des points :

$$\{(X_{k,n}, \hat{e}_n(X_{k,n})) : 1 < k \leq n\}. \quad (2.11)$$

Les propriétés statistiques des $\hat{e}_n(u)$ peuvent être obtenues en utilisant la théorie des processus empiriques pertinentes, comme expliquées dans Shorack un Wellner (1986). Pour nos besoins, le *mean excess plot* est utilisé uniquement comme une méthode graphique, principalement pour distinguer entre les modèles à queue léger et à queue lourd. Pour une évaluation critique concernant l'utilisation des fonctions moyennes des excès en assurance voir Rytgaard (1996). Pour une application utile du *mean excess plot*, voir la section 6.5.1 dans Embrechts et al. (1997). Si l'approximation *GPD* est valide pour un certain seuil u , alors elle est valide pour tout $x > u$. Dans ce cas le graphique nous permettra de choisir u de manière à ce que $\hat{e}_n(u)$ soit approximativement linéaire. L'estimation graphique de u n'est pas aisée dans la pratique et même lorsque nous travaillons avec des données simulées. On peut aussi conclure que le graphique *mean excess plot* permet de décrire la prédiction du dépassement du seuil lorsqu'un excès se produit. Trois cas peuvent se présenter :

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

Il existe aussi des recherches sur la consistance et la linéarité du graphique *mean excess plot* . Voir par exemple, Yang (1978) et Ghosh et al. (2011) respectivement.

Procédure graphique sous R.

La principale difficulté réside bien sûr dans l'interprétation de l'ordre de la linéarité de $\hat{e}_n(u)$. Seule la pratique peut dire lorsque on observe une modification de la pente de $\hat{e}_n(u)$ pour une certaine valeur de u . Dans les

applications nous préférons souvent des graphiques indiquant la valeur du seuil u , ainsi que le nombre de dépassements utilisées pour l'estimation, mais au final un praticien ne doit jamais s'attendre à un choix unique de u . Si on choisit le seuil u égale à $X_{k,n}$, alors le nombre de dépassements égale à $n - k$, et $\hat{e}_n(X_{k,n})$ peut être écrite comme

$$\hat{e}_n(X_{k,n}) = \frac{1}{n - k} \sum_{i=k}^n (X_{i,n} - X_{k,n}),$$

ou aussi comme

$$\hat{e}_n(X_{k,n}) = \frac{1}{n - k} \sum_{i=1}^{n-k} (X_{k+i,n} - X_{k,n}). \quad (2.12)$$

Dans le but d'examiner la queue de la distribution correspondante, le programme suivant nous permet de tracer un *mean excess plot* sous R pour chaque échantillon de données X_1, \dots, X_n . Par contre on ne précise pas le nombre de dépassements k , car ça rentre pas de nos techniques de détermination de seuil pour notre nouvel estimateur (voir la section 6.1.1).

```
x<-sort(x) # on ordonne l'échantillon des observations Xi
k<-seq(1,n-1)
meplot<-seq(1,n-1)
for (k in 1 :n-1)
{
i<-seq(1,n-k)
meplot[k]<-(sum(x[k+i]-x[k]))/(n-k)
}
U<-x[1 :n-1]
plot(U, meplot, ylim=c(0, 35), xlim=c(1, 16) , col="DarkViolet",
ylab=expression("Empirical Mean Excess"), xlab=expression("u" ) )
```

Chapitre 3

Estimation des statistiques extrêmes de type Pareto

Si on dispose d'un échantillon de données réelles ordonnées $X_{1,n}, \dots, X_{n,n}$. Nous nous intéressons dans ce chapitre à présenter une étude sur l'estimation des valeurs dites extrêmes concernant uniquement une partie de l'échantillon qui commence à partir d'un certain rang donné ($n - k$). Si on arrive à ajuster l'échantillon à une distribution bien définie via la détermination de la valeur de l'indice de queue γ , il suffit donc d'exploiter notre échantillon d'une manière efficace, pour tirer des résultats recherchés comme pour la tarification des primes d'assurance par exemple. On note que plus cet indice est élevé en valeur absolue, plus le poids des extrêmes dans la distribution initiale est important mais lorsque la distribution des pertes est inconnue, cela devient un problème d'estimation statistique. Des outils cités plus haut, comme le *QQ-plot* par exemple nous permettront d'avoir qu'une information partielle sur la distribution de l'échantillon, qui consiste l'appartenance à l'un des trois domaines d'attraction cités dans le théorème 1.1. Une telle appartenance nous donne une idée sur la queue de distribution avant d'entamer l'étape de l'estimation et en d'autre part la raison pour laquelle une estimation *semi-paramétrique* sera imposée. Comme nous travaillons dans notre thème sur l'estimation des mesures de risques à queue lourde, cela nous incite de porter sur l'indice de queue γ et plus précisément pour des valeurs positives, ainsi que l'estimation de la queue elle-même et le nombre des k -plus grandes valeurs qui influencent sur la stabilité des estimateurs de la queue de la distribution et ceux de son indice et notamment sur les résultats asymptotiques de notre nouvel estimateur de mesure de risque. Dans la littérature de la théorie des valeurs extrêmes on trouve plusieurs techniques *semi-paramétriques* pour l'estimation de l'indice de queue, qui engendrent à chaque fois un estimateur des *quantiles extrêmes* correspondant ainsi qu'une procédure du choix

ou d'estimation de k optimal. Dans les sections qui suivent nous exposerons uniquement deux approches d'estimation, l'une est semi-paramétrique basée sur l'estimateur de Hill (1975) et la deuxième à régression exponentielle. Voir Matthys et Beirlant (2003). Notre choix s'explique par notre volonté d'introduire une adaptation des estimateurs ainsi proposés aux prochains chapitres.

3.1 Estimation de l'indice de queue γ

L'indice de queue reflète la forme de la queue de distribution (avec aucune hypothèse sur la forme paramétrique de la queue) et par conséquent joue un rôle clé dans l'analyse des mesures à queue lourde. On l'utilise aussi pour l'estimation des quantiles extrêmes comme il permet d'identifier à peu près si la distribution est à queue lourde ou non, ainsi que pour déterminer le nombre de moments finis. Plusieurs approches existent pour la détermination des estimateurs de l'indice des valeurs extrêmes à savoir entre autre approche paramétrique, non paramétrique, *semi-paramétrique*, par les noyaux et par les excès. On peut citer l'estimateur de Hill (1975) valable pour $\gamma > 0$. Il est considéré comme le plus simple des estimateurs de l'indice de queue. Pour étendre l'estimateur de *Hill* aux deux autres domaines d'attraction, Dekkers et al. (1989) ont proposé un estimateur des moments valable quel que soit $\gamma \in \mathbb{R}$. Les auteurs ont établi la consistance forte, faible et la normalité asymptotique de leur estimateur. Soulignons que la méthode des moments fut initialement utilisée par Hosking et Wallis (1987) qui proposa un estimateur des moments pondérés définit pour tout $\gamma < 1$. Cet estimateur est consistant si $\gamma < 1$ et asymptotiquement gaussien pour $\gamma < 1/2$. Toutefois, Pickands (1975) fût le premier à proposer un estimateur de l'indice de queue plus général que l'estimateur de *Hill*, i.e valable quel que soit le signe de γ . Smith (1985) s'est quant à lui intéressé au comportement des estimateurs du maximum de vraisemblance dans le cas d'une loi *GEV*. Il a été démontré que l'estimateur du maximum de vraisemblance est consistant, asymptotiquement efficace et asymptotiquement normal pour tout $\gamma > -1/2$. Falk (1995) a proposé un complément à l'estimateur du maximum de vraisemblance. Son estimateur, que l'on appelle aussi l'estimateur *néгатif de Hill* est consistant si $\gamma < -1/2$ et asymptotiquement normal si $-1 < \gamma < -1/2$. Comme autres estimateurs, on peut citer l'estimateur du rapport des moments (Danielsson et al. 1996), l'estimateur de Peng (1998), l'estimateur basé sur le *QQ - plot*, l'estimateur basé sur le graphique de la moyenne des excès (Beirlant et al. 1996), l'estimateur construit par des méthodes de régression (Beirlant et al. 2002), d'optimisation avec contrainte (Gsörgö et al. 1985; Bacro et Brito, 1994) ou sans contrainte (Schultze et Steinebach, 1996; Kratz et Resnick,

1996) (voir Embrechts et al. 1997, Beirlant et al. 2004, de Haan et Ferreira, 2006). A côté de cette énumération, on pourrait aussi ajouter les méthodes de correction du biais basées sur des méthodes de régression (Beirlant et al. 1999, 2005; Diebolt et al. 2008) ou sur des méthodes de bootstrap (Gomes et Oliveira, 2001) ou jackknife (Gomes, 1999; Gomes et al., 2005) et les estimateurs à poids ou combinaisons linéaires (Viharos, 1993, 1995; Gomes et al. 2005) qui englobent d'une part les estimateurs construits par des méthodes d'optimisation avec et sans contrainte et d'autres part les estimateurs construits par des méthodes de régression.

D'un point de vue théorique, toutes ces méthodes partagent les mêmes propriétés de consistance et de normalité asymptotique. Cependant, les simulations montrent qu'il y a de grandes différences entre ces différents estimateurs. En général, il n'y a pas une meilleure méthode dans toutes les situations. Les méthodes les plus utilisées sont celles de Hill, Pickands et des moments. Certaines études de comparaison (théorique et par simulation) entre les différentes méthodes peuvent être trouvées dans Rosen et Weissman (1996), Peng (1998), de Haan et Peng (1998), Groeneboom et al. (2003) et Tsourti et Panaretos (2001, 2003).

Tsourti et Panaretos (2001, 2003) pensent que la performance d'une méthode dépend de la distribution de la série étudiée. En d'autres termes, elle dépend de la vraie valeur de l'indice de queue. Ils recommandent l'utilisation de techniques pour déterminer le domaine d'attraction de la loi des valeurs extrêmes, et donc l'intervalle le plus probable pour l'indice des valeurs extrêmes. Pour de plus amples explications sur ces techniques voir El-Adlouni et al. (2007).

3.1.1 Estimateur de Hill $\hat{\gamma}_n^h$

Soient X_1, \dots, X_n des variables aléatoires *iid* de fonction de répartition commune F et soit $X_{1,n}, \dots, X_{n,n}$ les statistiques d'ordre correspondantes. On suppose que F satisfait la condition du domaine d'attraction de Fréchet (voir le théorème 1.1), en terme de variations régulière ceci est équivalent à : pour $\gamma > 0$ à

$$1 - F(x) = x^{-1/\gamma} \ell(x), x > 0$$

Il s'agit d'une caractéristique remarquable que l'estimateur de *Hill* peut être obtenue de plusieurs façons : par exemple, par l'approche de variation régulière (de Haan, 1994), par la méthode de régression (Beirlant et al. 1999), ou en utilisant les quantiles (Beirlant et al. 2004), ou par le maximum de vraisemblance (*Hill*, 1975). Ceci prouve que c'est un estimateur naturel de l'indice des valeurs extrêmes. La méthode du maximum de vraisemblance est la plus populaire et qui sous certaines conditions est la plus efficace. On se

sert des statistiques d'ordre supérieurs à un certain seuil u pour ne garder que les observations les plus grandes de façon à ce quelles suivent approximativement une distribution de *Pareto*. D'autre part on peut écrire :

$$\frac{1 - F(x)}{1 - F(u)} = \left(\frac{x}{u}\right)^{-1/\gamma}, \quad x > u, \quad u \in \mathbb{R}.$$

Si $(n - k)$ désigne le nombre de statistiques d'ordre qui dépassent le seuil u , alors on estime u par $X_{n-k,n}$. En utilisant les plus grande statistiques d'ordre $X_{n-k+1,n}, \dots, X_{n,n}$ la fonction de log-vraisemblance sera alors :

$$\begin{aligned} L(\gamma, X_{n-k+1}, X_{n,n}) &= -k \log(\gamma u) + k \log(1 - F(u)) \\ &\quad - \frac{\gamma}{\gamma + 1} \sum_{i=1}^k (\log X_{n-i+1,n} - \log X_{n-k,n}). \end{aligned}$$

En maximisant la fonction de log-vraisemblance par rapport à γ , on obtient l'estimateur de *Hill* pour $\gamma > 0$,

$$\hat{\gamma}_n^h = \frac{1}{k} \sum_{i=1}^k \log X_{n-i+1,n} - \log X_{n-k,n} \quad (3.1)$$

Mason, (1982) a prouvé la consistance faible de $\hat{\gamma}_n^h$ pour toute suite $k = k(n)$ satisfaisant $k \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$. Deheuvels et al. (1988) on établie la consistance forte pour toute suite $k = k(n)$ satisfaisant $k/\log \log n$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$. De nombreux articles traitent de la normalité asymptotique de l'estimateur de *Hill*; Davis et Resnick (1984) l'ont prouvé en utilisant les domaines d'attraction et sous les conditions de *Von mises*. Sous d'autres hypothèses Gsörgö et Mason (1985) ont démontré cette normalité asymptotique en introduisant l'approximation des processus empiriques par des ponts browniens. Resnick et de Haan (1995) et Geluk et al. (1995) ont montré une équivalence entre la normalité asymptotique et la variation régulière au second ordre moyennant une hypothèse supplémentaire sur $k(n)$. On peut aussi obtenir cette normalité asymptotique comme un cas particulier de la démonstration de Dekkers et al. (1989) à partir de la normalité de leur estimateur des moments (voir par exemple de Wolf, 2001). En pratique, la précision de l'estimation dépend de la sélection de k . L'estimateur de *Hill* a un biais qui tends vers zéro quand k est relativement petit.

La variance asymptotique de l'estimateur de *Hill* vaut $\frac{\gamma^2}{k}$. Cette variance asymptotique tends vers zéro quand k est grand. On remarque que le choix de k est crucial, si on choisi k trop grand ceci engendrera un biais asymptotique important. Par conséquent, on doit rechercher une valeur optimale

de k , qui équilibre entre ces deux vices. Le choix de k optimal représente le principal obstacle dans le processus d'estimation de l'indice de queue (voir la section 6.1.1).

Nous concluons que l'estimateur de *Hill* est inadéquat si la distribution sous-jacente n'a pas de queue à variation régulière, si γ n'est pas positive, la taille de l'échantillon n'est pas assez grand, et la queue n'est pas assez lourde (γ n'est pas grand). Si F est à variation régulière, l'estimateur de *Hill* dépend fortement du type de la fonction à variation lente ℓ , ce qui est généralement inconnue.

Hill-plot.

Pour chaque choix de k , nous obtenons une autre valeur de l'estimateur de *Hill*. Habituellement, on trace les estimations $\hat{\gamma}_n^h$ contre k , ce qui nous donne un graphe appelé *Hill-plot* : $\{(k, \hat{\gamma}_n^h) : 1 \leq k \leq n - 1\}$. Il s'agit de sélectionner graphiquement le nombre d'excès au-delà duquel la valeur de l'indice de queue devient stable. Toutefois, ces graphiques sont généralement loin d'être constant, ce qui rend difficile l'utilisation de l'estimateur en pratique sans autre guide sur la façon de choisir la valeur k . Pour le même échantillon de données X_1, \dots, X_n on pourra obtenir le graphique *Hill-plot* en exécutant le programme suivant sous *R*.

```
x<-sort(x) # on ordonne l'échantillon des observations  $X_i$ 
y=sort(abs(x))
z=log10(y)
k<-seq(1,n-1)
hill<-seq(1,n-1)
for(k in 1 :n-1)
{
l<-seq(1,k)
hill[k]<-((1/k)*sum(z[n-l+1]))-z[n-k] # estimateur de Hill vs  $k$ 
}
plot(k, hill, ylim=c(0.2, 1), type="l", col="6",lwd="1",ylab=expression
(paste(plain(hat(gamma)^h))))
abline(h=gamma, col="red", lwd="2")
```

Matthys et Beirlant (2003) ont présenté une méthode de régression exponentielle non linéaire par maximum de vraisemblance pour donner un estimateur de l'indice de queue à biais réduit. Dans le but d'utiliser cet estimateur dans notre travail, nous expliquons la construction et les propriétés asymptotiques de cet estimateur dans ce qui suit.

3.1.2 Estimateur de Matthys et Beirlant $\hat{\gamma}_n^m$

Soient $U_{j,n}$ ($1 \leq j \leq n$) et $V_{j,k}$ ($1 \leq j \leq k$) les statistiques d'ordre des échantillons des variables aléatoires indépendantes et identiquement distribuées de loi uniforme $(0, 1)$ de taille n et k respectivement, et $E_{j,k}$ ($1 \leq j \leq k$) les statistiques d'ordre d'un échantillon de variables aléatoires indépendantes et identiquement distribuées de loi exponentielle standard de taille k : avec $U(r) := Q(1 - 1/r) = x_{1/r}$,

$$\begin{aligned} X_{n-j+1,n} &\stackrel{d}{=} U(U_{j,n}^{-1}) && \text{pour } j \leq n, \\ \frac{U_{j,n}}{U_{k+1,n}} &\stackrel{d}{=} V_{j,k} && \text{pour } j \leq k < n, \text{ et} \\ -\log(V_{j,k}) &\stackrel{d}{=} E_{k-j+1,k} && \text{pour } j \leq k. \end{aligned}$$

Les $V_{j,k}$ dans la deuxième égalité sont indépendants de $U_{k+1,n}$. Un résultat clé de l'argumentation est la représentation de Rényi des statistiques d'ordre exponentielle standards, qui stipule que

$$E_{k-j+1,k} \stackrel{d}{=} \sum_{i=j}^k \frac{f_{k-i+1}}{i} \quad \text{pour } j \leq k, \quad (3.2)$$

où les f_i sont des variables aléatoires indépendantes et de distribution exponentielles standard identique.

Du théorème 1.1 et le théorème Pickands-Belkema-de Haan, une condition équivalente peut être formulée en termes de la fonction quantile de queue U . de Haan (1970) stipule que F appartient au domaine d'attraction de G_γ si et seulement si il existe une fonction mesurable positive a_U telle que $\forall t > 0$

$$\lim_{r \rightarrow \infty} \frac{U(tr) - U(r)}{a_U(r)} = \begin{cases} \frac{t^\gamma - 1}{\gamma} & \text{si } \gamma = 0 \\ \log t & \text{si } \gamma \neq 0. \end{cases} \quad (3.3)$$

La fonction de normalisation $a_U(r)$ est alors équivalente à $\beta(U(r))$ pour $r \rightarrow \infty$, avec $\beta(u)$ comme dans (2.8). Notez que pour une *GPD* (2.7), nous avons $\beta(u) = \beta + \gamma(u - \mu)$, $U(r) = \mu + \beta(r^\gamma - 1)/\gamma$, et l'égalité dans (3.2) est vraie pour tout $t > 0$ et pour tout $r > 0$ (i.e., non seulement dans la limite pour $r \rightarrow \infty$) avec $a_U(r) = \beta r^\gamma = \beta(U(r))$.

Pour $k < n$ et pour $1 \leq j \leq k$, la condition (3.3) sur la fonction de quantile de queue inspire l'approximation suivante :

$$\begin{aligned} X_{n-j+1,n} - X_{n-k,n} &\stackrel{d}{=} U(U_{j,n}^{-1}) - U(U_{k+1,n}^{-1}) \\ &\stackrel{d}{=} U(V_{j,k}^{-1} U_{k+1,n}^{-1}) - U(U_{k+1,n}^{-1}) \end{aligned}$$

$$\stackrel{d}{\approx} a_{n,k+1} \frac{V_{j,k}^{-\gamma} - 1}{\gamma} \quad (3.4)$$

où $a_{n,k+1}$ désigne $a_U(U_{k+1,n}^{-1})$.

Comme $a_{n,k+1} = a_U(U_{k+1,n}^{-1}) \sim \beta(X_{n-k,n})$, (3.3) est en fait l'équivalent, exprimé en termes de statistiques d'ordre, par l'approximation *GPD* (2.8) avec $u = X_{n-k,n}$.

Pour le log-ratio des espacements des statistiques d'ordre, on obtient alors

$$\log \frac{X_{n-j+1,n} - X_{n-k,n}}{X_{n-j,n} - X_{n-k,n}} \stackrel{d}{\approx} \log \frac{V_{j,k}^{-\gamma} - 1}{V_{j+1,k}^{-\gamma} - 1} \text{ pour } 1 \leq j < k.$$

Le théorème de la valeur moyenne appliquée à la partie droite donne, avec $E_{j,k}^* \in (E_{k-j,k}, E_{k-j+1,k})$ et $V_{j,k}^* = \exp(-E_{j,k}^*)$,

$$\begin{aligned} \log \frac{V_{j,k}^{-\gamma} - 1}{V_{j+1,k}^{-\gamma} - 1} &\stackrel{d}{=} \log(e^{\gamma E_{k-j,k}} - 1) \\ &\stackrel{d}{=} (E_{k-j+1,k}, E_{k-j,k}) \cdot \frac{\gamma e^{\gamma E_{j,k}^*}}{e^{\gamma E_{j,k}^*} - 1} \\ &\stackrel{d}{=} \frac{f_{k-j+1}}{j} \cdot \frac{\gamma}{1 - (V_{j,k}^*)^\gamma} \end{aligned}$$

La dernière égalité découle de (3.2). Estimant $V_{j,k}^*$ par $j/(k+1)$, nous trouvons le modèle de régression exponentielle non linéaire suivant pour les log-ratios des distances :

$$j \log \frac{X_{n-j+1,k} - X_{n-k,n}}{X_{n-j,k} - X_{n-k,n}} \stackrel{d}{\approx} \frac{\gamma}{1 - \left(\frac{j}{k+1}\right)^\gamma} f_{k-j+1} \text{ pour } 1 \leq j < k \quad (3.5)$$

où les f_i sont des variables aléatoires indépendantes et identiquement distribuées de loi exponentielle standard.

Pour une valeur fixe $k < n$, un estimateur du maximum de vraisemblance noté $\hat{\gamma}_n^m$ pour l'indice des valeurs extrêmes peut être facilement mis en œuvre sur la base de (3.6), en calculant les rapports logarithmiques pondérées des espacements $Y_j := j \log [X_{n-j+1,k} - X_{n-k,n} / X_{n-j,k} - X_{n-k,n}]$ pour $1 \leq j < k$ et en maximisant numériquement la log-vraisemblance

$$\mathcal{L}(\gamma; \mathbf{Y}) := \sum_{j=1}^{k-1} \left\{ \log \left(\frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} \right) - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} Y_j \right\}. \quad (3.6)$$

Ce qui donne

$$\hat{\gamma}_n^m = \arg \max_{\gamma > 0} \sum_{j=1}^{k-1} \left\{ \log \left(\frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} \right) - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} Y_j \right\}. \quad (3.7)$$

Sous certaines conditions techniques sur la queue de la distribution, $\hat{\gamma}_n^m$ est convergent et est asymptotiquement normal.

Hypothèse R_ℓ . Il existe un réel constant $\rho \leq 0$ et une fonction de vitesse b satisfaisant $b(r) \rightarrow 0$ quand $r \rightarrow \infty$, telle que pour tout $t \geq 1$, comme $r \rightarrow \infty$,

$$\log \frac{\ell(tr)}{\ell(r)} \sim \begin{cases} b(r) \frac{t^\rho - 1}{\rho} & \text{pour } \rho \neq 0, \\ b(r) \log t, & \text{pour } \rho = 0. \end{cases}$$

Mis à part un changement possible de signe la fonction b est donc à variation régulière d'indice ρ .

Théorème 3.1 (normalité asymptotique de $\hat{\gamma}_n^m$ - Matthys et Beirlant, 2003). Supposons que, dans le cas où $\gamma > 0$,

$$U(r) = r^\gamma \ell(r)$$

avec ℓ satisfaisant (R_ℓ) avec $\rho < 0$. Supposons également que $k, n \rightarrow \infty$ avec $k/n \rightarrow 0$, et dans le cas où $\gamma \neq 0$,

$$\sqrt{k} \frac{\tilde{b}_0(n/k)}{b_0(n/k)} \rightarrow 0$$

Donc

$$\sqrt{k} (\hat{\gamma}_n^m - \gamma) \xrightarrow{L} N(0, \sigma_\gamma^2 / a_\gamma^2),$$

où

$$a_\gamma = \gamma^{-2} \int_0^1 (1 - u^\gamma + u^\gamma \log u^\gamma)^2 (1 - u^\gamma)^{-2} du$$

et σ_γ^2 est égale à la variance des $K_\gamma(U)$ avec U uniformément réparties sur $(0, 1)$ et avec $K_\gamma(U) = (\log u) / \gamma + (1 + \gamma) (di \log u^\gamma) / \gamma^2$, où $di \log u = \int_1^u (\log t) / (1 - t) dt$ ($u \geq 0$) désigne la fonction dilogarithme.

En règle générale, il est toujours nécessaire pour toute estimateur de l'indice de queue de faire un compromis entre la variance, qui diminue lorsque plus de statistiques d'ordre sont utilisés dans l'estimation et le biais, qui augmente lorsque le seuil est estimé à une valeur inférieure. Voir la section suivante, pour plus de détails sur la procédure de choix du seuil d'après Matthys et Beirlant (2003).

Matthys et Beirlant (2003), ont affiné le modèle de régression exponentielle (3.6) avec des concepts de la théorie des valeurs extrêmes sur la variation lente. Nous présentons l'explication pour $\gamma > 0$ qui fournit également une procédure d'estimation adaptative pour sélectionner le seuil optimal dans le but d'adapter l'estimateur $\hat{\gamma}_n^m$ dans notre estimation de mesure de risque.

Correction du biais (Matthys et Beirlant, 2003).

Matthys et Beirlant (2003) ont affiné les modèles de régression exponentielle (3.5) avec des concepts de la théorie des valeurs extrêmes sur la variation lente avec du reste. Une fonction positive est à variation lente à l'infini si

$$\lim_{r \rightarrow \infty} \frac{\ell(tr)}{\ell(r)} = 1, \quad \text{pour tout } t > 0 \quad (3.8)$$

Une fonction f est dite à variation régulière d'indice γ , notée par $f \in R_\gamma$, si $f(x) = x^\gamma \ell(x)$ avec ℓ est à variation lente. R_0 symbolise donc la classe des fonctions à variation lente.

De la théorie des valeurs extrêmes il est bien connu que, si la fonction quantile de queue U vérifie (3.3) d'indice de queue γ , alors la fonction de normalisation a_U varie de façon régulière à l'infini avec le même indice γ . Par ailleurs. Si $\gamma > 0 : U \in R_\gamma$, i.e., $U(r) = r^\gamma \ell(r)$ avec $\ell \in R_0$, et $a_U(r)/U(r) \rightarrow \gamma$ pour $r \rightarrow \infty$, de telle sorte que $a_U/U \in R_0$ (symbolise donc la classe des fonctions à variation lente). Ainsi, pour tous $t, r > 0$,

$$\frac{U(tr) - U(r)}{a_U(r)} = \frac{U(r)}{a_U(r)} \left(\frac{U(tr)}{U(r)} - 1 \right) = \frac{1}{d(r)} \left(t^\gamma \frac{\ell(tr)}{\ell(r)} - 1 \right),$$

avec $d(r) := a_U(r)/U(r) \rightarrow \gamma$ pour $r \rightarrow \infty$, $d \in R_0$;

Ainsi nous obtenons pour tout $\gamma \in \mathbb{R}$ que si U vérifie (3.3), il existe une fonction à variation lente ℓ et une fonction d avec $\pm d \in R_0$ et $d(r) \rightarrow \gamma$ pour $r \rightarrow \infty$, de telle sorte que pour tout $t, r > 0$,

$$\frac{U(tr) - U(r)}{a_U(r)} = \frac{1}{d(r)} \left(t^\gamma \frac{\ell(tr)}{\ell(r)} - 1 \right) \quad (3.9)$$

Inversement, pour $\gamma \neq 0$ le résultat de la limite (3.3) découle immédiatement de (3.9) par la définition (3.8) de variation lente.

Le biais des estimateurs classiques de l'indice de queue et de $\hat{\gamma}_n^m$ introduit découle de prendre les résultats limites (3.3) ou (2.8) comme égalités, alors que ceux-ci ne sont valables réellement que lorsque le seuil $U(r)$ tends vers x_+ . Lorsque $U(r)$ s'éloigne de x_+ , ou lorsque le rapport $\ell(tr)/\ell(r)$ dans (3.9) tend vers 1 avec une vitesse lente, en conséquence l'approximation sera plus pauvre, avec une estimation plus ou moins biaisée pour l'indice de queue. Nous considérons la condition suivante sur ℓ , Qui décrit la variation lente avec reste (voir la section 3.12.1 de Bingham, Goldie et Teugels, 1987).

Les parties à variation lente ℓ dans (3.9) de la plupart des distributions courantes dans les statistiques satisfaire l'hypothèse (R_ℓ).

Pour les espacements des statistiques d'ordre avec $1 \leq j \leq k < n$, (3.9) donne

$$\begin{aligned} X_{n-j+1,n} - X_{n-k,n} &\stackrel{d}{=} U(V_{j,k}^{-1}U_{k+1,n}^{-1}) - U(U_{k+1,n}^{-1}) \\ &\stackrel{d}{=} c_{n,k+1} \left[V_{j,k}^{-\gamma} \frac{l(V_{j,k}^{-1}U_{k+1,n}^{-1})}{l(U_{k+1,n}^{-1})} - 1 \right], \end{aligned}$$

où $c_{n,k+1}$ représente $a_U(U_{k+1,n}^{-1})/d(U_{k+1,n}^{-1})$. Nous obtenons avec l'hypothèse (R_ℓ) l'approximation

$$X_{n-j+1,n} - X_{n-k,n} \stackrel{d}{\approx} c_{n,k+1} \left[V_{j,k}^{-\gamma} \exp\left(b_{n,k+1} \frac{V_{j,k}^{-\rho} - 1}{\rho}\right) - 1 \right], \quad (3.10)$$

où $b_{n,k+1}$ représente $b(U_{k+1,n}^{-1})$.

Nous passons maintenant de la même manière que pour le modèle (3.6) avec le log-ratio des espacements des statistiques d'ordre :

$$\log \frac{X_{n-j+1,n} - X_{n-k,n}}{X_{n-j,n} - X_{n-k,n}} \stackrel{d}{\approx} \log \frac{V_{j,k}^{-\gamma} \exp\left(b_{n,k+1} \frac{V_{j,k}^{-\rho} - 1}{\rho}\right) - 1}{V_{j+1,k}^{-\gamma} \exp\left(b_{n,k+1} \frac{V_{j,k}^{-\rho} - 1}{\rho}\right) - 1} \quad \text{pour } 1 \leq j < k.$$

Pour la partie droite, le théorème de valeur moyenne, avec les mêmes notations pour $E_{j,k}^*$ et $V_{j,k}^*$ citées plus haut, donne

$$\begin{aligned} &\log \frac{V_{j,k}^{-\gamma} \exp\left(b_{n,k+1} \frac{V_{j,k}^{-\rho} - 1}{\rho}\right) - 1}{V_{j+1,k}^{-\gamma} \exp\left(b_{n,k+1} \frac{V_{j,k}^{-\rho} - 1}{\rho}\right) - 1} \\ &\stackrel{d}{=} \log \frac{\exp\left(\gamma E_{k-j+1,k} + b_{n,k+1} \frac{e^{\rho E_{k-j+1,k-1}}}{\rho}\right) - 1}{\exp\left(\gamma E_{k-j,k} + b_{n,k+1} \frac{e^{\rho E_{k-j,k-1}}}{\rho}\right) - 1} \\ &\stackrel{d}{=} (E_{k-j+1,k} - E_{k-j,k}) \cdot \frac{\gamma + b_{n,k+1} e^{\rho E_{j,k}^*}}{1 - \exp\left(-\gamma E_{j,k}^* + b_{n,k+1} \frac{e^{\rho E_{j,k}^* - 1}}{-\rho}\right)} \\ &\stackrel{d}{=} \frac{f_{k-1+1}}{j} \cdot \frac{\gamma + b_{n,k+1} (V_{j,k}^*)^{-\rho}}{1 - (V_{j,k}^*)^{-\gamma} \exp\left(b_{n,k+1} \frac{(V_{j,k}^*)^{-\rho} - 1}{-\rho}\right)} \end{aligned}$$

D'où le modèle de régression exponentielle non linéaire pour les log-ratios des espacements suivant avec $1 \leq j < k$:

$$j \log \frac{X_{n-j+1,n} - X_{n-k,n}}{X_{n-j,n} - X_{n-k,n}} \stackrel{d}{\approx} \frac{\gamma + b_{n,k+1} \left(\frac{j}{k+1}\right)^{-\rho}}{1 - \left(\frac{j}{k+1}\right)^\gamma \exp\left(b_{n,k+1} \frac{\left(\frac{j}{k+1}\right)^{-\rho} - 1}{-\rho}\right)} f_{k-j+1} \quad (3.11)$$

où les f_i sont des variables aléatoires *iid* de loi exponentielles standard. Notez que (3.11) simplifié est le modèle (3.6) lorsque $b_{n,k+1} = 0$.

Le modèle (3.11) permet de calculer les estimations conjointes des paramètres $\gamma, b_{n,k+1}$ et ρ en maximisant numériquement la log-vraisemblance correspondante des log-ratios des espacements pour $1 \leq j < k < n$. Nous noterons ces estimateurs du maximum de vraisemblance par $\hat{\gamma}_{k+1}^B, \hat{b}_{n,k+1}^B$ et $\hat{\rho}_{k+1}^B$, respectivement. Dans la plupart des cas, le nouvel estimateur réduit le biais de $\hat{\gamma}_n^m$ dans une large mesure.

3.1.3 Comparaison entre $\hat{\gamma}_n^h$ et $\hat{\gamma}_n^m$.

Nous présentons une petite analyse comparative entre les deux estimateurs de l'indice de queue $\hat{\gamma}_n^h$ et $\hat{\gamma}_n^m$. Sur la Fig. 3.1 pour $n = 500$ et Fig. 3.3 pour $n = 1000$, nous remarquons que l'estimateur $\hat{\gamma}_n^m$ est très proche de γ par contre $\hat{\gamma}_n^h$ présente des valeurs beaucoup plus inférieures que γ . La taille de l'échantillon n'influence pas sur la qualité des estimateurs $\hat{\gamma}_n^h$ et $\hat{\gamma}_n^m$. Donc le fait d'intégrer l'estimateur $\hat{\gamma}_n^m$ dans notre nouvel estimateur de mesure de risque assure une meilleur estimation de l'indice de queue pour presque toutes les valeurs de k , aussi vu comme avantage dans la détermination de k optimal, qui d'après les graphes, présente une grande marge de choix par rapport à $\hat{\gamma}_n^m$, ce qui nous donne la possibilité de se baser sur d'autre méthodes et techniques de choix de k optimal régissant directement sur la qualité de notre nouvel estimateur de mesure de risque.

La qualité d'un estimateur, pour une taille d'échantillon donnée, se mesure plutôt par l'erreur Quadratique Moyenne (*EQM*). Pour justifier les raisons de notre nouvel estimateur alternatif à biais réduit présenté au chapitre 5, on rappelle que $EQM = Var(\hat{\gamma}_n) + Biais(\hat{\gamma}_n)^2$ (plus l'erreur quadratique moyenne est grande plus le biais est important). Sur la Fig. 3.2 pour $n = 500$ et Fig. 3.4, pour $n = 1000$, nous avons calculer l'*EQM* empirique pour les deux estimateurs $\hat{\gamma}_n^h$ et $\hat{\gamma}_n^m$. Pour tout k on observe que $EQME(\hat{\gamma}_n^m)$ est toujours inférieur à l'*EQME*($\hat{\gamma}_n^h$).

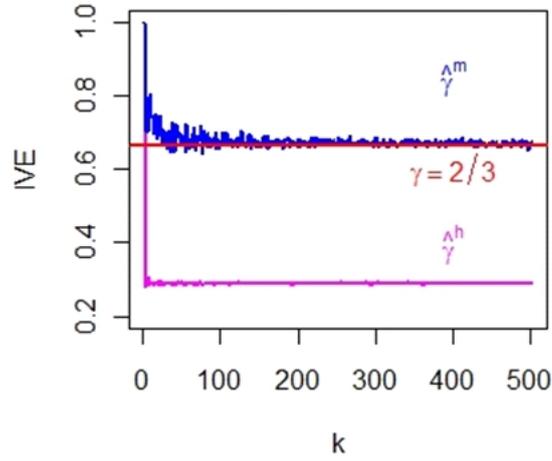


FIG. 3.1 – Plot de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d'une loi Pareto de taille $n = 500$ et $\gamma = 2/3$

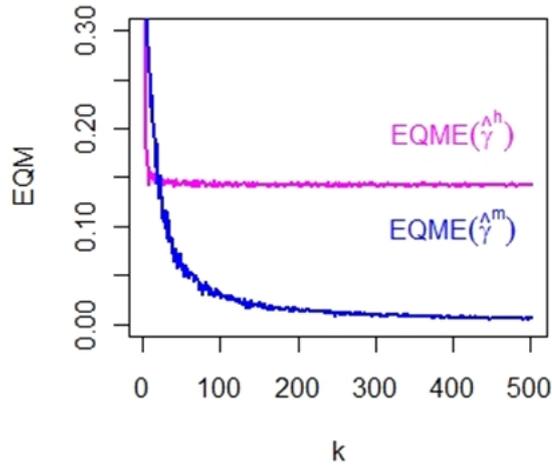


FIG. 3.2 – Plot de l'EQM empirique de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d'une loi Pareto de taille $n = 500$ et $\gamma = 2/3$

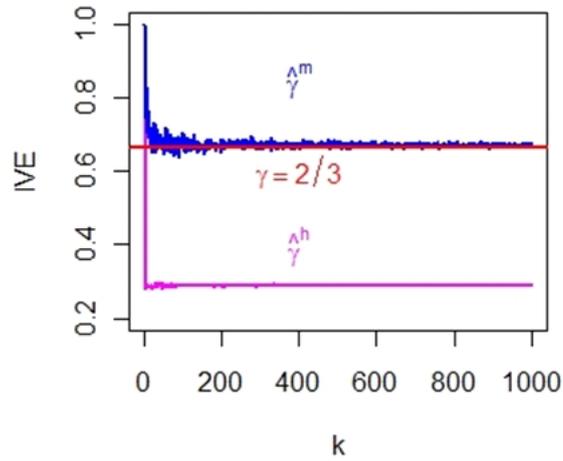


FIG. 3.3 – Plot de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d’une loi Pareto de taille $n = 1000$ et $\gamma = 2/3$

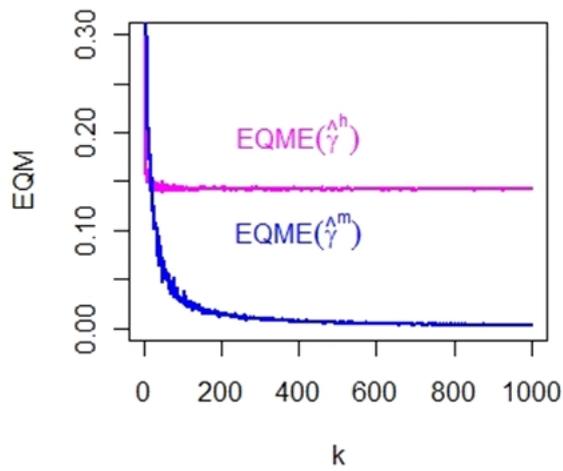


FIG. 3.4 – Plot de l’EQM empirique de $\hat{\gamma}_n^h$ et de $\hat{\gamma}_n^m$ vs k pour 200 échantillons générés d’une loi Pareto de taille $n = 1000$ et $\gamma = 2/3$

Ce qui veut dire que $\hat{\gamma}_n^m$ présente un biais réduit que celui de $\hat{\gamma}_n^w$. L'adaptation de $\hat{\gamma}_n^m$ à notre nouvel estimateur de mesure de risque présente une réduction de biais significative.

3.2 Estimation des quantiles extrêmes x_p

La mesure et la gestion des risques exigent qu'on s'intéresse aux queues des distributions et notamment aux pertes que l'on peut subir avec des probabilités faibles mais non négligeables. Généralement on ne connaît jamais dans la plupart des cas la distribution qui modélise notre échantillon de risque, ce qui nous mène automatiquement à l'estimation des quantiles extrêmes. Soient X_1, \dots, X_n la réalisation de n variables aléatoires réelles indépendantes et de fonction de répartition commune F supposée continue. On définit les quantiles extrêmes notés x_p par :

$$x_p = F^{-1}(1 - p) = Q(1 - p) \quad \text{quand } p \rightarrow \infty$$

Il existe plusieurs approches pour estimer les quantiles extrêmes. Celle basée sur la théorie des valeurs extrêmes (Gnedenko, 1943) est donné par :

$$\hat{x}_p = a_n \frac{(np)^{-\hat{\gamma}} - 1}{\hat{\gamma}} + b_n,$$

aussi par la méthode des excès et l'approximation de la loi *GPD* (théorème Pickands-Balkema-de Haan) à savoir :

$$\hat{x}_p = \left(\frac{k}{np} \right)^{\hat{\gamma}} \frac{\hat{\sigma}}{\hat{\gamma}} + u.$$

Plusieurs d'autres méthodes semi-paramétriques existent. Weissman (1978) a examiné l'estimation des quantiles extrêmes pour les trois classes de valeurs extrêmes séparément, en supposant une connaissance préalable sur l'indice de queue γ .

Dans la pratique, toutefois, cette connaissance est souvent pas disponible car il n'est généralement pas évident de savoir à quelle classe appartient la distribution sous-jacente d'un phénomène aléatoire. Pickands (1975) a défini l'estimateur des quantiles associé à son estimateur bien connu de l'indice de queue. Ses propriétés asymptotiques ont été discutés par Dekkers et de Haan (1989). Drees (1996) introduit l'estimateur de Pickands raffiné qui est plus

stable que celui de Pickands. Sur l'estimateur de Pickands raffiné voir aussi Segers (2001).

Un autre estimateur célèbre des quantiles extrêmes basé sur la méthode des moments a été introduit par Dekkers, Einmahl et de Haan (1989). La normalité asymptotique de cet estimateur de quantile sous différentes conditions sur la queue de la distribution est prouvée par Dekkers, Einmahl et de Haan (1989), et par de Haan et Rootzén (1993). On trouve aussi d'autres travaux comme par exemple concernant l'optimisation de l'estimation des quantiles extrêmes (Ferreira et al. 2010), sur la réduction du biais des estimateurs des quantiles extrêmes, (Gomes, 2006 et Li et al. 2010).

Dans ce qui suit, nous présentons deux méthodes classiques, à commencer par l'estimateur de quantile extrême de Weissman (1978) et celui de Matthys et Beirlant (2003) qui est basé sur un modèle de régression exponentielle, ainsi sera la motivation clé pour notre nouvel estimateur de mesure de risque.

3.2.1 Estimateur des quantiles extrêmes de Weissman

$$\hat{x}_p^w$$

On se restreint aux fonctions $F \in D(\text{Fréchet})$ pour lesquelles on a la caractérisation suivante :

$$1 - F(t) = t^{-1/\gamma} L(t), \text{ avec, } \lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1, x > 0$$

En termes de l'inverse généralisé de F on obtient

$$Q(1 - s) = s^{-1/\gamma} \tilde{L}(s) \text{ avec, } \lim_{s \rightarrow \infty} \frac{\tilde{L}(\lambda s)}{\tilde{L}(s)} = 1, \lambda > 0$$

Ce qui permet d'obtenir :

$$\frac{Q(1 - \lambda s)}{Q(1 - s)} = \lambda^{-1/\gamma},$$

ce qui implique

$$Q(1 - \lambda s) = Q(1 - s) \lambda^{-1/\gamma}$$

Si on pose $\lambda s = p$ et $s = k/n$, on obtient :

$$Q(1 - p) = Q(1 - k/n) \left(\frac{np}{k}\right)^{-1/\gamma}$$

Ce qui nous permet d'estimer x_p par :

$$\begin{aligned}\hat{x}_p &= \hat{Q}(1-p) \\ &= Q_n(1-k/n) \left(\frac{np}{k}\right)^{-1/\hat{\gamma}_n^h},\end{aligned}$$

avec $Q_n(1-k/n) = X_{n-k,n}$ est l'estimateur standard de $Q(1-k/n)$ et $\hat{\gamma}_n^h$ représente l'estimateur de Hill (voir section 3.1.1). ce qui donne finalement l'estimateur des quantiles extrêmes de Weissman

$$\hat{x}_p^w = X_{n-k,n} \left(\frac{np}{k}\right)^{-1/\hat{\gamma}_n^h}. \quad (3.12)$$

Cet estimateur est en fonction du nombre de valeurs extrêmes k et de l'estimateur de Hill $\hat{\gamma}_n^h$. Ferreira et Vries (2004) ont proposé un intervalle de confiance pour l'estimateur de Hill aussi bien que l'estimateur de quantile extrêmes de Weissman, au fait que la meilleur vitesse de convergence de l'estimateur de *Hill* est atteint quand sa moyenne est différente de zéro, ceci pour obtenir un choix optimal de k .

3.2.2 Estimateur des quantiles extrêmes de Matthys

$$\hat{x}_p^m$$

Dans cette section nous présentons un deuxième modèle de régression exponentielle présenté dans Matthys et Beirlant (2003), ce qui permet un estimateur simple pour le facteur de l'échelle $a_{n,k+1} = a_U(U_{k+1,n}^{-1})$ qui est nécessaire à la construction ultérieure d'un estimateur pour les quantiles extrêmes. Le calcul est similaire à celle du modèle (3.6), mais au lieu de considérer les log-ratios d'espacements nous employons maintenant (3.4), pour approximer les espacements des statistiques d'ordre eux-mêmes :

$$X_{n-j+1,n} - X_{n-j,n} \stackrel{d}{\approx} a_{n,k+1} \frac{V_{j,k}^{-\gamma} - V_{j+1,k}^{-\gamma}}{\gamma} \quad \text{pour } 1 \leq j < k.$$

Avec la même notation pour $E_{j,k}^*$ et $V_{j,k}^*$ comme ci-dessus, une application de théorème de valeur moyenne pour le deuxième facteur à droite donne

$$\begin{aligned}\frac{V_{j,k}^{-\gamma} - V_{j+1,k}^{-\gamma}}{\gamma} &\stackrel{d}{=} \frac{e^{\gamma E_{k-j+1,k}} - e^{\gamma E_{k-j,k}}}{\gamma} \\ &\stackrel{d}{=} (E_{k-j+1,k} - E_{k-j,k}) e^{\gamma E_{k-j,k}^*} \stackrel{d}{=} \frac{f_{k-j+1}}{j} \cdot (V_{j,k}^*)^{-\gamma}.\end{aligned}$$

En estimant $V_{j,k}^*$ par $j/(k+1)$, nous arrivons à un modèle de régression non linéaire

$$j(X_{n-j+1,n} - X_{n-j,n}) \stackrel{d}{\approx} a_{n,k+1} \left(\frac{j}{k+1} \right)^{-\gamma} f_{k-j+1} \quad \text{pour } 1 \leq j \leq k \quad (3.13)$$

avec f_j ($1 \leq j \leq k$) sont des variables aléatoires *iid* de loi exponentielles standards, comme avant.

Le modèle (3.13) rends possible le calcul d'un estimateur pour $a_{n,k+1}$ par la maximisation de la log-vraisemblance correspondante aux distances $Z_j := j(X_{n-j+1,n} - X_{n-j,n})$ pour $1 \leq j \leq k$:

$$a_{n,k+1} = \frac{1}{k} \sum_{j=1}^k Z_j \left(\frac{j}{k+1} \right)^{\gamma}. \quad (3.14)$$

La condition (3.3) conduit désormais à l'approximation suivante pour les quantiles extrêmes :

$$\begin{aligned} x_p - X_{n-k,n} &\stackrel{d}{=} U(p^{-1}) - U(U_{k+1,n}^{-1}) \\ &\stackrel{d}{=} U((U_{k+1,n}/p) U_{k+1,n}^{-1}) - U(U_{k+1,n}^{-1}) \\ &\stackrel{d}{\approx} \hat{a}_{n,k+1} \frac{(U_{k+1,n}/p)^{\gamma} - 1}{\gamma}. \end{aligned}$$

Nous estimons $U_{k+1,n}$ par sa valeur attendue $(k+1)/(n+1)$, γ par $\hat{\gamma}_n^m$, et $a_{n,k+1}$ en remplaçant $\hat{\gamma}_n^m$ pour γ dans la formule (3.10) :

$$\hat{a}_{n,k+1} := \frac{1}{k} \sum_{j=1}^k j(X_{n-j+1,n} - X_{n-j,n}) \left(\frac{j}{k+1} \right)^{\hat{\gamma}_n^m}$$

Nous arrivons finalement à l'estimateur des quantiles extrêmes

$$\hat{x}_p^m := X_{n-k,n} + \hat{a}_{n,k+1} \frac{\left(\frac{k+1}{p(n+1)} \right)^{\hat{\gamma}_n^m} - 1}{\hat{\gamma}_n^m} \quad \text{pour } k < n \quad (3.15)$$

Matthys et Beirlant (2003) observent que les biais de \hat{x}_p^m et $\hat{\gamma}_n^m$ poussent généralement à un rythme plus lent (avec k) que ceux des moments, en particulier pour les distributions de la classe $\gamma < 0$. Par conséquent, \hat{x}_p^m et $\hat{\gamma}_n^m$ atteignent leur points optimal de l'erreur quadratique moyenne minimale plus profondément dans l'échantillon, c'est à dire à des valeurs de seuil $X_{n-k,n}$ inférieures.

3.2.3 Choix du nombre des valeurs extrêmes k

Tous les estimateurs basés sur des valeurs extrêmes reposent essentiellement sur le nombre k des statistiques d'ordre supérieurs impliqués dans le calcul de l'estimation le fait que ce nombre localise où la queue de la distribution commence.

Les résultats asymptotiques des estimateurs de l'indice de queue sont en général obtenus lorsque $k \rightarrow \infty$ et $k/n \rightarrow 0$, mais en présence d'un échantillon aléatoire de taille finie en pratique, le problème concernant le choix du nombre de valeurs extrêmes n'est pas facile à gérer. Ce nombre k est régi non seulement par la taille de l'échantillon n , mais aussi régi par des paramètres caractérisant F . Lorsque la fonction de distribution sous-jacente est connue, la valeur optimale de k peut être obtenue par la minimisation de l'erreur quadratique moyenne asymptotique (*AMSE*) de l'estimateur considéré.

Dans le cas contraire on peut se baser sur des méthodes de *Bootstrap* pour calculer l'erreur quadratique moyenne empirique. Comme il existe aussi des méthodes graphiques citons par exemple le *Hill-plot* dans le cas d'un estimateur de *Hill*. Pour plus de détails sur le choix de k voir par exemple; Dekkers et de Haan (1993), Drees et Kaufmann (1998), Danielsson et al. (2001), Cheng et Peng (2001), et Neves et Fraga Alves (2004).

On peut conclure que lorsque la distribution sous-jacente dans un échantillon de données observées est inconnue, le nombre de valeurs extrêmes k n'est pas unique le fait qu'il existe plusieurs estimateurs de l'indice de queue γ , où le k optimal correspondant doit assurer une bonne qualité d'estimation concernant les résultats asymptotiques des ces estimateurs.

Pour nos besoins nous nous restreindrons à présenter l'algorithme de *Reiss et Thomas* repris par Neves et Fraga Alves (2004) pour le choix de k optimal. Nous allons l'utiliser dans nos simulations dans le chapitre 6 dans l'estimation des mesures des risques extrêmes de Necir et Meraghni (2009). Nous fournissons ainsi un algorithme qui selon Matthys et Beirlant (2003) consiste à minimiser l'erreur quadratique moyenne de $\hat{\gamma}_n^m$ pour déterminer la valeur optimale de k correspondante dans le but de l'adapter à notre nouvel estimateur de mesure de risque (voir section 5.1.3).

Chapitre 4

Mesures de risques pour les plus grandes revendications

Les principes de primes appliqués par l'assureur peuvent être, au moins en partie, fondée sur des considérations autres que le risque. Il est alors important de quantifier le degré auquel la prime compense l'assureur pour le risque lié à la perte. Cela peut être fait en choisissant une mesure du risque approprié et en résolvant le paramètre qui mène à la prime de l'assureur. Lorsque la distribution des pertes est inconnue, cela devient un problème d'estimation statistique.

L'estimation du prix d'un risque d'assurance est un problème actuariel très important. Une intention particulière est portée sur la prise en compte des valeurs extrêmes sur laquelle repose le calcul des exigences de capitaux propres. Le prix doit refléter la propriété de la distribution de la variable aléatoire décrivant la perte correspondante.

Si la variable de perte a une distribution à queue lourde (i.e., distribution avec une variance infinie), alors, la mesure du risque (comme une mesure de la prime de risque) devrait être plus élevée. Pour fournir des mesures de risque de distributions à queue lourde, les procédures standard de statistiques classiques (quand la variance est finie) ne peut pas être appliquée.

Dans ce chapitre, nous proposons une estimation semi-paramétrique pour une mesure de risque par distorsion pour des pertes à queue lourde et uniquement pour la partie des valeurs extrêmes de l'échantillon de risque, ce dans le but de minimiser le biais pour des estimateurs initialement existant dans la littérature.

4.1 Les outils mathématiques de l'analyse des risques

4.1.1 Mesures de risque.

Définition 4.1. (Mesure de risque). *On appelle mesure de risque toute application ρ associant un risque X à un réel $\rho(X) \in \mathbb{R}_+ (+\infty)$.*

Définition 4.2. (Chargement de sécurité). *Une mesure de risque ρ contient un chargement de sécurité si pour tout risque X on a, $\rho(X) \geq E[X]$.*

Définition 4.3. (Constance). *Une mesure de risque ρ ne contient pas un chargement de risque injustifié si pour toute constante réelle a , $\rho(a) = a$.*

Définition 4.4. (Non chargement excessif). *Une mesure de risque ρ ne dépasse pas la perte maximale si, $\rho(X) \leq \max(X)$.*

Définition 4.5. (Objectivité). *Une mesure de risque ρ est dite objective lorsqu'elle dépend que de F et non de X elle-même.*

Définition 4.6. (L'invariance en loi). *Une mesure de risque ρ est dite invariante en loi si pour tout X et Y telle que $X \stackrel{L}{=} Y$ alors, $\rho(X) = \rho(Y)$.*

Mesure de risque cohérente

La définition d'une mesure de risque est très générale puisque toute fonctionnelle réelle positive d'une variable aléatoire peut être considérée comme étant une mesure de risque. Aussi en pratique on exige de telles mesures qu'elles disposent de propriétés mathématiques dont la transcription conceptuelle permette de les jauger. En pratique, on exige fréquemment qu'une mesure de risque ρ possède une partie des caractéristiques suivantes :

1. *Invariante par translation* : $\rho(X + c) = \rho(X) + c$ pour toute constante c .
2. *Sous-additivité* : $\rho(X + Y) \leq \rho(X) + \rho(Y)$ quels que soient les risques X et Y .
3. *Homogénéité* : $\rho(cX) = c\rho(X)$ pour toute constante positive c .
4. *Monotonie* : $P[X < Y] = 1 \implies \rho(X) \leq \rho(Y)$ quels que soient les risques X et Y .

Ces caractéristiques trouvent une interprétation naturelle dans la situation où la mesure de risque doit permettre de définir un capital de solvabilité d'une société d'assurance.

Ainsi la sous-additivité représente l'effet de la diversification ; une société qui couvre deux risques ne nécessite pas davantage de capitaux que la somme de ceux obtenus pour deux entités distinctes se partageant ces deux risques.

La monotonie traduit quand à elle le fait que si le montant résultant d'un risque est systématiquement (au sens presque sur) inférieur à celui résultant d'un autre risque, le capital nécessaire à couvrir le premier risque ne saurait être supérieur à celui nécessaire pour couvrir le second.

L'association de ces quatre axiomes a donné naissance au concept de cohérence d'une mesure de risque dans Artzner et al. (1999).

Définition 4.7. (Mesure de risque cohérente). *Une mesure de risque invariante par translation, sous-additive, homogène et monotone est dite cohérente.*

Cette notion de cohérence n'est toutefois pas ce que l'on attend à minima d'une mesure de risque. Ainsi certaines mesures de risque parmi les plus exploitées actuellement ne le sont pas. C'est notamment le cas de la Value-at-Risk (VaR) ou encore de la variance.

Mesure de risque comonotone additive

Rappelons qu'un vecteur aléatoire (X_1, X_2) , de fonctions de répartition marginales F_1, F_2 , est un vecteur comonotone s'il existe une variable aléatoire U de loi uniforme sur $[0, 1]$ telle que (X_1, X_2) a la même loi que $(F_1^{-1}(U), F_2^{-1}(U))$

Définition 4.8. (Mesure de risque comonotone). *On appelle mesure de risque comonotone additive toute mesure de risque ρ telle que $\rho(X_1 + X_2) = \rho(X_1) + \rho(X_2)$ pour tout vecteur comonotone $(X_1; X_2)$.*

Une mesure de risque comonotone additive intègre donc le fait que deux risques comonotones ne se mutualisent pas. Pour une mesure utilisée pour déterminer un capital de solvabilité, cette propriété est souhaitable puisque, dès que le risque U se produit, les risques X_1 et X_2 se produisent également avec une ampleur croissante avec celle de U .

Mesure de risque indépendante additive

Rappelons que deux variables aléatoire X_1 et X_2 , de fonctions de répartition marginales F_1, F_2 sont indépendants si la fonction de répartition du couple vérifie $\forall (x_1, x_2)$ alors $F_{X_1 X_2}(x_1, x_2) = F_{X_1}(x_1) F_{X_2}(x_2)$

Définition 4.9. (Mesure de risque indépendant additive). *On appelle mesure de risque indépendant additive toute mesure de risque ρ telle que si X_1 et X_2 sont indépendants alors $\rho(X_1 + X_2) = \rho(X_1) + \rho(X_2)$ pour tout vecteur comonotone $(X_1; X_2)$.*

Cette propriété est une condition raisonnable pour un principe de calcul de prime puisqu'elle s'assure que l'assuré n'a aucun intérêt en divisant le risque demandant l'assurance à plusieurs assureurs.

Mesures de risque usuelles

L'objectif de ce paragraphe est de présenter les mesures de risque les plus usuelles. On s'attardera particulièrement sur la prime de Wang dont l'utilisation sera bien détaillée aux sections et chapitres prochains.

La Value-At-Risk (VaR). La notion de Value-At-Risk ou valeur ajustée au risque s'est originellement développée dans les milieux financiers avant d'être largement reprise dans les problématiques assurantielles.

Définition 4.10. (Value-At-Risk-VaR). *La Value-At-Risk de niveau α associée au risque X est donnée par :*

$$VaR(X, \alpha) = \inf \{x \mid P(X \leq x) \geq \alpha\}.$$

On notera que $VaR(X, \alpha) = F_X^{-1}(\alpha)$ où F_X^{-1} désigne la fonction quantile de la loi de X . Rappelons que, dans le cas général, la fonction quantile est la pseudo-inverse de la fonction de répartition, soit

$$F_X^{-1}(p) = \inf \{x \mid F(x) \geq p\}.$$

Cette mesure de risque a le mérite de reposer sur un concept simple et facilement explicable : $VaR(X, \alpha)$ est le montant qui permettra de couvrir le montant de sinistres engendré par le risque X avec une probabilité α . Ce concept est directement lié à celui de probabilité de ruine puisque si une société, disposant d'un montant de "ressources" égal à $VaR(X, \alpha)$, assure un unique risque X , sa probabilité de ruine est égale à $1 - \alpha$.

La VaR n'est pas cohérente car elle n'est pas sous-additive. Ce résultat peut se démontrer à l'aide d'un contre exemple.

Les Value-At-Risk ont un certain nombre de "bonne" propriétés mathématiques parmi lesquelles le fait que pour toute fonction g croissante et continue à gauche, on a : $VaR_\alpha(g(X)) = g(VaR_\alpha(X))$.

Il découle de cette propriété en prenant $g = F_1^{-1} + F_2^{-1}$ et $X = U$, que les VaR sont comonotones additives puisque pour tout $\alpha \in]0; 1[$, on a $VaR_\alpha((F_1^{-1} + F_2^{-1})(U)) = (F_1^{-1} + F_2^{-1})(VaR_\alpha(U))$.

La Tail-Value-At-Risk (TVaR). Beaucoup d'auteurs et d'articles dans la littérature définissent la Tail Value-at-Risk (*TVaR*). Elle est aussi appelée Conditional Value-at-Risk (*CVaR*) ou Average Value-at-Risk (*AVaR*). Citons par exemple Shapiro et Ruszczyński (2009).

Définition 4.11. (La Tail-Value-At-Risk-VaR). *La Tail-Value-At-Risk de niveau α associée au risque X est donnée par :*

$$TVaR(X, \alpha) = \frac{1}{1 - \alpha} \int_{\alpha}^1 F_X^{-1}(p) dp.$$

On remarque que la *TVaR* peut s'exprimer en fonction de la *VaR* :

$$TVaR(X, \alpha) = VaR(X, \alpha) + \frac{1}{1 - \alpha} E[(X - VaR(X, \alpha))^+].$$

Il vient de cette réécriture que pour tout $\alpha \in]0; 1[$, $TVaR_{\alpha}(X) < +\infty$. Par ailleurs, le deuxième terme du membre de droite représente la perte moyenne au-delà de la *VaR*. La *TVaR* est donc très sensible à la forme de la queue de distribution.

La Tail-Value-At-Risk est cohérente, inclut un chargement de sécurité du fait que pour tout $\alpha \geq 0$, $TVaR(X; \alpha) \geq TVaR(X; 0) = E[X]$, et est comonotone additive, du fait qu'elle est une somme de VaR qui sont elles-mêmes comonotones additives.

Définition 4.12. (Expected Shortfall-ES). *L'expected Shortfall de niveau de probabilité α est la perte moyenne au-delà de la VaR au niveau α , i.e.,*

$$ES_{\alpha}(X) = E[(X - VaR(X, \alpha))^+]$$

On peut remarquer que si X représente la charge brute de sinistres, $ES_{\alpha}(X)$ est le montant de la prime Stop-Loss dont la rétention pour l'assureur est *VaR* au niveau α .

Définition 4.13. (Conditionnal Tail Expectation-CTE). La Conditionnal Tail Expectation de niveau α est le montant de la perte moyenne sachant que celle-ci dépasse la *VaR* au niveau α , i.e.,

$$CTE(X, \alpha) = E[X | X > VaR(X, \alpha)]$$

Cette définition est très proche de celle de la Tail-Value-at-Risk et en particulier, ces deux mesures coïncident lorsque la fonction de répartition F_X du risque X est continue.

Remarquons que si la *TVaR* est comonotone additive, ce n'est en général pas le cas de la *CTE*.

Mesure de risque d'Esscher. La mesure de risque d'Esscher consiste à mesurer le risque comme étant la prime pure, i.e., l'espérance de la transformé d'Esscher du risque initial.

Définition 4.14. (Mesure de risque d'Esscher). On appelle mesure d'Esscher de paramètre $h > 0$ du risque X , la mesure de risque donnée par :

$$Es(X; h) = \frac{E[Xe^{hX}]}{E[e^{hX}]} = \frac{d}{dh} \ln E[e^{hX}].$$

La mesure de risque d'Esscher n'est pas cohérente car elle n'est ni homogène, ni monotone. En revanche, elle contient un chargement de sécurité puisque $Es(X; h)$ est une fonction croissante en h et $Es(X; 0) = E[X]$.

Mesure de risque de Wang. Les mesures de risque de Wang utilisent l'opérateur espérance sur des transformations de la distribution de la variable aléatoire d'intérêt. L'idée est en effet d'alourdir la queue de la distribution de la variable d'intérêt afin d'engendrer un chargement par rapport à la prime pure. Cette transformation de la fonction de répartition sera effectuée à l'aide d'une fonction de distorsion.

Rappelons qu'une fonction de distorsion est une fonction non décroissante $g : [0; 1] \rightarrow [0; 1]$ telle que $g(0) = 0$ et $g(1) = 1$

Définition 4.15. (Mesure de risque de Wang). On appelle mesure de risque de Wang issue de la fonction de distorsion g , la mesure ρ_g définit par :

$$\rho_g(X) = \int_0^\infty g(p[X > x]) dx.$$

Définition 4.16. (Wang-Transform WT). On appelle Wang-Transform la mesure de risque de Wang issue de la fonction de distorsion

$$g_\alpha(x) = \Phi[\Phi^{-1}(x) - \Phi^{-1}(\alpha)].$$

Certaines mesures de risque usuelles telles que la VaR ou la TVaR sont des mesures de risque de Wang, le tableau ci-dessous reprend les fonctions de distorsion correspondantes.

Mesure de risque	Paramètre	fonction de distorsion
Value-at-Risk	VaR_α	$g(x) = 1_{[\alpha, +\infty]}(x)$
Tail-Value-at-Risk	$TVaR_\alpha$	$g(x) = \min(x/(1-\alpha); 1)$
Mesure de risque PH	PH_ξ	$g(x) = x^{1/\xi}$
Wang-transform	WT_α	$g(x) = \Phi[\Phi^{-1}(x) - \Phi^{-1}(\alpha)]$

TAB. 4.1 – Fonctions de distorsion associées à quelques mesures de risque

Les mesures de risque de Wang sont homogènes, invariantes par translation et monotones. Une mesure de Wang n'est en revanche sous-additive que si, et seulement si, la fonction de distorsion dont elle est issue est concave.

Ainsi une mesure de risque de Wang est cohérente si, et seulement si, la fonction de distorsion correspondante est concave, $g \circ \bar{F}_X$ est continue à droite et est donc la fonction de queue d'une variable aléatoire. Ainsi $\rho_g(X)$ est l'espérance mathématique de la variable aléatoire de fonction de queue $g \circ \bar{F}_X$. Une mesure de Wang est additive pour des risques comonotones.

4.1.2 Comparaison des risques

L'objectif de ce paragraphe est de présenter des outils permettant de classer les risques selon leur "dangerosité".

Relation associée à une mesure de risque

Dans le paragraphe précédent, nous avons étudié un certain nombre de mesures de risque. Une idée naturelle pour comparer deux risques X et Y est de choisir une mesure de risque ρ et de comparer $\rho(X)$ et $\rho(Y)$, ce que l'on peut toujours faire puisque \mathbb{R} est ordonnée par la relation d'ordre \leq . Cette démarche nous permet d'introduire la relation définie comme suit.

$$X \prec_\rho Y \text{ si } \rho(X) \leq \rho(Y)$$

La relation issue de la mesure de risque est réflexive et transitive. De plus il est toujours possible de comparer par deux lois de probabilité ou deux variables aléatoires. Cette relation n'est pas une relation d'ordre car elle n'est pas antisymétrique.

Le principal mérite de ce type de relation est qu'il est toujours possible de comparer deux risques. Il faut néanmoins rester prudent car l'on peut avoir simultanément $X \prec_\rho Y$ et $X \succ_{\rho'} Y$ pour deux mesures de risque ρ et ρ' différentes. C'est pour cette raison que l'on préférera se tourner vers des ordres partiels qui permettent de disposer de davantage de propriétés.

Ordre stochastique

Définition 4.17. (Dominance stochastique). *On dit que X domine selon l'ordre stochastique Y ($Y \prec_{st} X$) si pour toute fonction de distorsion g , on a $\rho_g(X) \leq \rho_g(Y)$.*

La relation \prec_{st} est un ordre partiel sur l'ensemble des lois de probabilités.

Ordre convexe

Définition 4.18. (Ordre convexe croissant). *On dit que X est moins dangereux que Y sur la base de l'ordre convexe croissant (\prec_{icx}) et l'on note $X \prec_{icx} Y$ si, pour toute fonction de distorsion g concave, on a $\rho_g(X) \leq \rho_g(Y)$.*

Définition 4.19. (Ordre convexe). *On dira que le risque X est moins dangereux que le risque Y de même prime pure au sens de l'ordre convexe (\prec_{cx}), s'il est moins dangereux au sens de l'ordre convexe croissant, i.e., si $X \prec_{cx} Y \Leftrightarrow X \prec_{icx} Y$ et $E[X] = E[Y]$.*

Pour toute fonction φ convexe croissante et pour autant que les variances existent, on a : $X \prec_{cx} Y \Leftrightarrow E[\varphi(X)] \leq E[\varphi(Y)]$.

L'ordre convexe permet de comparer des variables aléatoires de même espérance, ce que ne permettait pas l'ordre stochastique puisque

$$X \prec_{cx} Y \text{ et } E[X] = E[Y] \Leftrightarrow X =_{loi} Y.$$

4.1.3 Principes de calcul de prime

L'activité d'assurance repose sur le concept de transfert de risque : moyennant une prime, l'assuré, se protège d'un aléa financier. Mesurer le risque assuré s'avère donc inévitable puisque cette information est nécessaire dans le cadre de la tarification pour déterminer les chargements de sécurité à ajouter à la prime pure et dans une approche de solvabilité pour déterminer le niveau des réserves et des fonds propres dont doit disposer l'assureur pour être solvable. Mesurer le risque pour les compagnies d'assurance c'est la détermination de ses besoin en fonds propre, d'où l'importance d'utiliser des mesures de risques efficaces et efficientes qui peuvent être adaptées dans le secteur d'assurance. Pour plus de détails sur le calcul des cotisations et la tarification de l'assurance, voir Goovaerts et Laeven (2007), Bühlman (1980), Kaas, van Heerwaarden et Goovaerts (1994) et Wang (1996). Une considération particulière dans la prochaine section sera donnée à l'estimation de *principe de risque ajusté* dite aussi *mesure de risque de Wang*.

Définition 4.20. (Principe de calcul de prime). *Le principe de calcul de prime est une règle qui assigne un nombre non négatif π (la prime) à n'importe quel risque donné X (une variable aléatoire positive) i.e. déterminer la prime comme un fonctionnel en assignant une valeur $\pi(F_X) \in \mathbb{R} \cup \{\pm\infty\}$ à la distribution de risque F_X .*

Propriétés de bonnes primes. Soit X, Y, Z des risques assurables arbitraires pour les quelles la prime $\pi(X)$ sera mieux définie.

1. *Aucun chargement de sécurité injustifié* : si, pour tout constant $a \geq 0$, $\pi(a) = 0$.
2. *Proportionnalité* : si, pour tout constant $a \geq 0$, $\pi(aX) = a\pi(X)$.
3. *Sous-additivité* : si, $\pi(X + Y) \leq \pi(X) + \pi(Y)$.
4. *Additivité* : si, $\pi(X + Y) = \pi(X) + \pi(Y)$.
5. *Consistance* : si, pour tout $a \geq 0$, $\pi(X + a) = \pi(X) + a$.
6. *Conservation d'ordre stochastique* : si, $X \prec_{st} Y$ implique que $\pi(X) \leq \pi(Y)$.
7. *Compatibilité sous mélange* : pour tout $p \in [0,1]$ et pour tout Z , $\pi(X) = \pi(Y)$ implique $\pi(pF_X + (1-p)F_Z) = \pi(pF_Y + (1-p)F_Z)$.

Quelques principes de base. Parmi les principes de calcul de prime les plus simple on cite :

1. *principe de la valeur espérée* : pour tout $a \geq 0$, $\pi(X) = (1 + a) E(X)$.
2. *principe de la prime mette* : pour tout $a > 0$, $\pi(X) = E(X)$
3. *principe de l'écart-type* : pour tout $a > 0$, $\pi(X) = E(X) + a\sqrt{\text{var}(X)}$
4. *principe de la variance* : pour tout $a > 0$, $\pi(X) = E(X) + a\text{var}(X)$
5. *principe de l'utilité nulle* : soit $u(x)(x \in \mathbb{R})$ une fonction deux fois différentiable avec $u'(x) > 0$. Dans ce cas est $\pi(X)$ est la solution de l'équation $E[u(\pi(X) - X)] = u(0)$.
6. *principe de la valeur moyenne* : soit $u(x)(x \in \mathbb{R}^*)$ une fonction strictement croissante et continue. Donc on définit $\pi(X) = u^{-1}(E[u(X)])$, où u^{-1} note la fonction inverse de u .
7. *principe de la perte maximale* : pour $p \geq 0, q = 1 - p$ on a $\pi(X) = pE(X) + q \max(X)$ où, $\max(X)$ note le point terminal à droite du rang de X .

8. *principe de Bühlmann, Gagliardi, et Strauts (1977)* : Pour tout risque la prime π est définie comme la solution de l'équation $E(u(X - z\pi(X))) = (1 - z)\pi(X)$ où $u(t)$ est une fonction deux fois différentiable avec $u(t) > 0$, $u''(t) \geq 0$, pour $t \in \mathbb{R}^*$. Pour $z = 1$, ce principe coïncide avec le principe d'utilité nulle. Pour $z = 0$ le principe de valeur espéré et un cas spécial.
9. *principe de la variance modifiée* : pour tout $a > 0$, $\pi(X) = E(X) + aVar(X)/E(X)$ si $E(X) > 0$ et $\pi(X) = 0$ si $E(X) = 0$
10. *principe exponentiel* : pour tout $a > 0$, $\pi(X) = a^{-1} \log E(e^{ax})$. Ce principe n'est pas approprié pour les risques Heavy Tailed.
11. *principe de quantile* : $\pi(X) = F_X^{-1}(1 - \varepsilon)$. i.e., la prime la plus petite tel que la probabilité d'une perte est à la plus part ε .
12. *principe de l'écart absolu* : pour tout $a \geq 0$, $\pi(X) = E(X) + ak_X$, où $k_X = E[X - F_X^{-1}(1/2)]$ est la déviation espérée absolue
13. *principe de risque ajusté* : $\pi(X) = \int_0^\infty (1 - F_X(x))^{1/\rho} dx$, où $\rho \geq 1$ appelé coefficient de distorsion.

4.2 Estimation de la prime de Wang pour des risques extrêmes

Définition 4.21. Une fonction de distorsion g est une fonction définie de $[0, 1]$ dans $[0, 1]$ avec $g(0) = 0$ et $g(1) = 1$.

Définition 4.22. Soit X une variable aléatoire positive de fonction de distribution absolument continue F et de fonction de densité f , alors, la fonction définie par $m_X(t) = \frac{f(t)}{1 - F(t)}$ est appelée fonction de taux d'hasard.

4.2.1 Prime d'hasard proportionnelle

Wang (1996) a proposé un principe de prime basé sur une transformation proportionnelle de la fonction d'hasard. Cette approche de tarification des contrats d'assurance considère des pertes de l'assurance comme des variables aléatoires positives et engendre des primes qui sont plus hautes que la valeur attendue de la perte d'assurance. Soit X_1, X_2, \dots, X_n une suite de risque *iid* positifs de fonction distribution commune $F_X(s) = P(X < s)$ avec $S_X(s) = 1 - F_X(s)$ est la fonction de survie. Le principe de prime de Wang (1995, 1996)

est défini comme suit.

$$\pi_g = \int_0^\infty g(S_X(s)) ds,$$

où g est une fonction de distorsion. Si g est une fonction puissance i.e., pour tout $\rho \geq 1$, $g(s) = s^{1/\rho}$, alors la fonction $g(S_X(s))$ est appelé transformation d'hasrad proportionnelle et

$$\pi_\rho = \int_0^\infty (1 - F_X(s))^{1/\rho} ds, \quad (4.1)$$

est appelé prime d'hasrad proportionnelle (*PH premium*) de X , avec $\rho \geq 1$ appelé coefficient de distorsion où coefficient d'aversion. Remarquons que si $F_X(s)$ a une densité, alors la prime donnée dans (4.1) peut être interprétée comme une prime nette d'un autre risque Y dont, sa fonction queue est $\bar{F}_Y(s) = (1 - F_X(s))^{1/\rho}$ et dont sa fonction de taux d'hasrad est proportionnellement inférieur à celle de risque X car $m_Y(t) = d/dt \log \bar{F}_Y(t) = \rho^{-1} m_X(t)$. C'est-à-dire, on peut voir Y comme le risque correspondant à X après le dégonflement de la fonction de taux d'hasrad de X par le facteur constant ρ^{-1} . La prime présentée dans (4.1) appelée aussi prime de risque ajustée et noté *RAP* (Risk-adjusted-Premium) a un chargement de sécurité non négatif, mais aucun chargement de sécurité injustifié, en plus, ce principe de prime est proportionnel, consistant, sous additif et stochastiquement monotone. Par différente approche, plusieurs hauteurs (Necir et Boukhetala, 2004; Necir et al. 2007, ...) se sont intéressés à estimer la prime mentionnée ci-dessous pour tarifier des risques dépassant un seuil élevé u_n sachant que ces risques suivent des distributions à queue lourde. Notons que le seuil u_n doit être aussi grand et dépend de la taille de l'échantillon des revendications produites. La prime de risque ajustée défini dans (4.1) correspondant à des pertes arrivées à partir d'un certain seuil élevé u_n est donnée par :

$$\pi_{u_n} = \int_{u_n}^\infty (1 - F_X(s))^{1/\rho} ds, \quad (4.2)$$

Necir et Boukhetala (2004) ont proposé un estimateur empirique pour (4.2) pour un échantillon de risque *iid* de distribution commune à queue lourde (à variation régulière). Soit $k = k_n$ une suite des entiers satisfaisant

$$1 \leq k \leq n, k \rightarrow \infty \text{ et } k/n \rightarrow \infty$$

telle que : $u_n = Q(1 - k/n)$ où Q est la fonction inverse généralisée de F . Pour tout $n \geq 1$, soit $X_{1,n}, X_{2,n}, \dots, X_{n,n}$ les statistiques d'ordre basées

sur X_1, X_2, \dots, X_n . En substituant $Q(1 - k/n)$ et $F(\cdot)$ par leur estimateurs empiriques $Q_n(1 - k/n)$ et $F_n(\cdot)$ et par une intégration par partie pour $Q_n(0) = X_{1,n}$, $\hat{\pi}_u$ est égal à

$$\hat{\pi}_{u_n} = \sum_{i=1}^k \left(\frac{i}{n}\right)^{1/\rho} \{X_{n-i+1,n} - X_{n-i,n}\} \quad \text{pour } \rho \geq 1 \quad (4.3)$$

Notons que pour $\rho = 1$ les statistiques $n\hat{\pi}_u$ correspondent au plus grande revendications des traités de la réassurance *ECOMOR* (excédent du coût moyen relatif) introduit par Thépaut (1950). Voir Rolski et al. (1999).

Définition 4.23. *On dit que $1 - F$ est à variation régulières au second ordre avec $1/\gamma$ est le premier paramètre et $\rho \leq 0$ est le deuxième paramètre, telle que : il existe une fonction $A(t) \rightarrow 0$ quand $t \rightarrow \infty$ de signe constant à l'infini si pour tout $x > 0$,*

$$\lim_{t \rightarrow 0} (A(t))^{-1} \left\{ \frac{Q(1 - tx)}{Q(1 - t)} - x^{-\gamma} \right\} = x^{-\gamma} \frac{x^{\rho\gamma} - 1}{\rho\gamma}. \quad (4.4)$$

Si $\rho = 0$, on interprète $\frac{x^{\rho\gamma} - 1}{\rho\gamma}$ comme $\log x$ (voir par exemple, de Haan et Stadtmüller, 1996)

Sous la condition (4.4) et d'autres conditions considérées, la normalité asymptotique de $\hat{\pi}_u$ est prouvée dans Necir et Boukhetala (2004).

4.2.2 Estimation semi-paramétrique

Nous nous concentrons dans nos travaux sur l'approche des quantiles extrêmes pour l'estimation de la prime d'hasard proportionnelle. Nous présenterons celui de Necir et al. (2007), qui est basé sur l'estimateur de quantiles extrêmes de Weissman dans le but de calculer un autre estimateur basé sur les estimateurs des quantiles extrêmes de Matthys et Beirlant (2003). Ce dernier estimateur à biais réduit sera la base de construction de notre nouvel estimateur de prime d'hasard proportionnelle dans une version plus complète qu'on va présenter au le chapitre 5.

Des études empiriques ont montré que les données financières et actuarielles présentent des distributions à queues lourdes ou de type pareto comme les distributions de Burr, Student, Lévy-satble et log-gamma. Ces dernières sont connues pour être des modèles appropriés pour modéliser les grandes réclamations d'assurance, les fluctuations des prix, etc. Voir par exemple, Beirlant

et al. (2001); Reiss et Thomas (2007) et Rolski et al. (1999). La queue de distribution est donc donnée par :

$$1 - F(x) = x^{-1/\gamma} \ell_F(x), \text{ pour tout } x > 0, \quad (4.5)$$

où $\gamma > 0$ et ℓ_F est une fonction à variation lente.

Pour tout $n \geq 1$, soit $X_{1,n}, X_{2,n}, \dots, X_{n,n}$ les statistiques d'ordre basées sur X_1, X_2, \dots, X_n . Selon Necir et al (2007) π_{u_n} s'écrira comme suit

$$\pi_{u_n} = - \int_0^{k/n} (s)^{1/\rho} dQ(1-s),$$

où $Q(1-s)$ est noté x_s , ce qui donne :

$$\pi_{u_n} = - \int_0^{k/n} (s)^{1/\rho} dx_s, \quad (4.6)$$

En effectuant une intégration par partie, ensuite remplaçant x_s par l'estimateur de quantiles extrêmes de Weissman donné dans (3.12) par :

$$\hat{x}_s^w = X_{n-k,n} \left(\frac{ns}{k} \right)^{-1/\hat{\gamma}_n^h},$$

où $\hat{\gamma}_n^h$ est l'estimateur de Hill donné dans (3.1). Nous obtenons après calcul,

$$\hat{\pi}_{u_n} = \left(\frac{k}{n} \right)^{1/\rho} \frac{\hat{\gamma}_n^h}{1/\rho - \hat{\gamma}_n^h} X_{n-k,n}, \quad \text{pour } \hat{\gamma}_n^h < 1/\rho \quad (4.7)$$

Le théorème suivant décrit la normalité asymptotique de l'estimateur $\hat{\pi}_{u_n}$.

Théorème 4.1 (normalité asymptotique de $\hat{\pi}_{u_n}$ - Necir et al. 2007).

Supposons que (4.4) est vérifiée avec $s^{1/\rho} Q(1-p) \rightarrow 0$ quand $p \rightarrow 0$ pour tout $\hat{\gamma}_n^h < 1/\rho$. Soit $k = k_n$ tel que $k \rightarrow \infty$, $k/n \rightarrow 0$ avec $(k/n)^{1/2} A(k/n) \rightarrow 0$ quand $n \rightarrow \infty$, donc,

$$\frac{(k/n)^{-1/\rho} k^{1/2}}{X_{n-k,n}} \{ \hat{\pi}_{u_n} - \pi_{u_n} \} \xrightarrow{D} N(0, \sigma_1^2(\rho, \gamma)),$$

$$\text{où } \sigma^2(\rho, \gamma) = \frac{\rho\gamma(\gamma^2 + 2\gamma^3\rho + \gamma^4\rho^2)}{(1 - \rho\gamma)^2}.$$

Preuve du Théorème 4.1. Voir Necir et al. (2007).

4.2.3 Estimateur à biais réduit

Dans le but de calculer un estimateur de π_{u_n} à biais réduit, nous utilisons l'estimateur des quantiles extrêmes de Mattyhs et Beirlant (2003) donné dans (3.15) par :

$$\hat{x}_s^m = X_{n-k,n} + \hat{a}_{n,k} \frac{\left(\frac{k}{sn}\right)^{\hat{\gamma}_n^m} - 1}{\hat{\gamma}_n^m} \quad \text{pour } k < n$$

Avec $\hat{a}_{n,k}$ donnée par

$$\hat{a}_{n,k} = \frac{1}{k} \sum_{j=1}^k j (X_{n-j+1,n} - X_{n-j,n}) \left(\frac{j}{k+1}\right)^{\hat{\gamma}_n^m},$$

et $\hat{\gamma}_n^m$ donnée dans (3.7) par :

$$\hat{\gamma}_n^m = \arg \max_{\gamma > 0} \sum_{j=1}^{k-1} \left\{ \log \left(\frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} \right) - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} Y_j \right\}.$$

On substituant la valeur de \hat{x}_s^m dans (4.6), on obtient l'estimateur attendu de π_{u_n} :

$$\hat{\pi}_{u_n}^*(X) = - \int_0^{k/n} (s)^{1/\rho} d\hat{x}_s^m,$$

Nous effectuons une intégration par partie et après calcul et simplification nous obtenons :

$$\hat{\pi}_{u_n}^* = \left(\frac{k}{n}\right)^{1/\rho} \frac{\rho \hat{a}_{n,k}}{1 - \rho \hat{\gamma}_n^m} \quad (4.8)$$

Cet estimateur est un cas particulier des estimateurs donnés dans Vandewalle et Beirlant (2006) et sera ainsi la base de construction de notre nouvel estimateur de la prime de risque ajustée concernant la partie des risques extrêmes. La normalité asymptotique de cet estimateur est vérifiée. Pour éviter les répétitions la preuve peut être déduite directement de notre démonstration donnée au chapitre 5.

Chapitre 5

Estimation semi-paramétrique des mesures de risque à distorsion pour des pertes à queue lourde

Nous nous basons dans ce chapitre sur des estimateurs des quantiles extrêmes de Mathys et Beirlant (2003), pour proposer un nouvel estimateur de la prime de risque de distorsion lorsque les pertes suivent des distributions à queue lourde ainsi, sous certaines conditions nous avons établi sa normalité asymptotique. Cet estimateur est meilleur en termes de biais que celui de Necir et Meraghmi (2009). Le calcul de cet estimateur dépend du choix du nombre optimal des valeurs extrêmes k utilisés dans l'estimation de l'indice de queue $\hat{\gamma}_n^m$ (Mathys et Beirlant, 2003) qui sera noté par la suite $\hat{\gamma}_k$. Un tel problème peut être résolu en utilisant la méthode heuristique de Reiss et Thomas détaillée dans Neves et Fraga Alves (2004).

Primes de risque de distorsion

De nombreux principes de calcul des primes sont introduites dans la littérature des sciences actuarielles (par exemple, Denuit et al. (1985); Furman et Zitikis, 2008a). Les plus populaires sont la prime nette, la variance et des primes à base d'écart type, la valeur exposée au risque, l'espérance conditionnelle, et la transformation des risques proportionnels. La plupart des primes sont des cas particuliers de la prime de distorsion de Wang (1996), qui est définie pour une variable de perte $X \geq 0$, avec fonction de distribution F par :

$$\Psi[F, g] = \int_0^\infty g(1 - F(x))dx,$$

où $g : [0, 1] \rightarrow [0, 1]$ est une fonction de distorsion, non décroissante et telle que $g(0) = 0$ et $g(1) = 1$. En terme de fonction quantile $Q(s) = \inf \{x : F(x) \geq s\}$, $0 < s < 1$. Cette classe des primes de risque prend la forme suivante

$$\Psi[Q; g] = \int_0^1 Q(1 - s) dg(s). \quad (5.1)$$

La plupart des fonctions de distorsion g sont concaves, ce qui rend les primes de distorsion correspondantes $\Psi[F, g]$ cohérentes (Artzner et al. 1999 et Wirth et Hardy, 1999). Dans cette thèse, nous sommes motivés par les pertes liées aux risques pour des distributions à queue lourde. Ensuite, il est assez naturel de supposer que les fonctions de distorsion g sont telles que $t \rightarrow g(1/t)$ est à variation régulière à l'infini avec l'indice de variation régulière $\rho \geq 1$, c'est-à-dire,

$$g(1/t) = t^{-1/\rho} \ell_g(t), \quad (5.2)$$

où $t \rightarrow \ell_g(t)$ est à variation lente à l'infini, c'est-à-dire que $\ell_g(tx)/\ell_g(t) \rightarrow 1$ quand $t \rightarrow \infty$, pour tout $x > 0$. Quelques exemples de fonctions de distorsion habituelles satisfaisant la condition de variation régulière (5.2) sont énumérés ci-dessous :

– Prime nette

$$g(x) = x \text{ avec } \rho = 1, \ell_g(x) = 1;$$

– Tail value-at-risk

$$g(x) = \min\left(\frac{x}{p}, 1\right) \text{ avec } \rho = 1, \ell_g(x) = 1;$$

– Transformée de risques proportionnels (PHT).

$$g(x) = x^{1/\beta} \text{ avec } \rho = \beta, \ell_g(x) = 1;$$

– Dual power function principle ($0 < \alpha < 1$)

$$g(x) = 1 - (1 - x)^\alpha \text{ avec } \rho = 1/\alpha, \ell_g(x) = x^{-\alpha} - (x^{-\alpha} - 1)^\alpha;$$

– Gini principle ($0 < \alpha \leq 1$)

$$g(x) = (1 + \alpha)x - \alpha x^2 \text{ avec } \rho = 1, \ell_g(x) = (1 + \alpha) - \alpha x;$$

– Beta-distortion risk premium (Wirch et Hardy, 1999)

$$g(x) = \int_0^x \frac{t^a (1-t)^b}{\beta(a,b)} dt \text{ avec } \rho = 1/(a+1), \ell_g(x) = \frac{1}{(a+1)\beta(a,b)};$$

où

$$\beta(a,b) := \int_0^1 \frac{t^a (1-t)^b}{\beta(a,b)} dt$$

– MINMAXVAR2 risk premium (Madan and Schoutens, 2010)

$$g(x) = 1 - \left(1 - x^{\frac{1}{1+\lambda}}\right)^{1+\tau}, \quad \lambda > 0, \tau > 0$$

$$\text{avec } \rho = \frac{1+\lambda}{1+\tau}, \quad \ell_g(x) = x^{-\frac{1+\tau}{1+\lambda}} - \left(x^{-\frac{1}{1+\lambda}} - 1\right)^{1+\tau}.$$

Jones et Zitikis (2003) ont remarqué que la contrepartie empirique de $\Psi[F; g]$ est une combinaison linéaire des statistiques d'ordre, communément appelé L -statistique. Cela ouvre un domaine fécond pour le développement de résultats d'inférence statistique, qui ont été activement étudiés par un certain nombre de chercheurs. Plus précisément, soient X_1, \dots, X_n une suite de variables aléatoires indépendantes et soient $X_{1:n} \leq \dots \leq X_{n:n}$ les statistiques d'ordre correspondantes. L'estimateur empirique de la prime de risque $\Psi[Q, g]$ est obtenue par la substitution de Q sur le côté droit de l'équation (5.1) par son homologue empirique

$$Q_n(s) := \inf \{x \in \mathbb{R} : F_n(x) \geq s\}, \quad 0 < s \leq 1,$$

associée à la fonction de répartition empirique définie sur la droite réelle, définie par

$$F_n(x) := n^{-1} \sum_{i=1}^n \mathbb{I}(X_i \leq x),$$

avec $\mathbb{I}(\cdot)$ étant la fonction indicatrice. Après un calcul simple, on obtient la formule :

$$\Psi[Q_n; g] = \int_0^1 Q_n(1-s) dg(s) = \sum_{i=1}^n a_{i,n}(g) X_{n-i+1:n},$$

dont le côté droit est un L -statistique avec les coefficients,

$$a_{i,n}(g) := g\left(\frac{i}{n}\right) - g\left(\frac{i-1}{n}\right). \quad (5.3)$$

Pour la littérature récente sur l'inférence statistique pour les primes de distortion, nous nous référons à Jones et Zitikis (2003), Jones et Zitikis (2007),

Centeno and Andrade (2005), Furman et Zitikis (2008a, 2008b), Brazauskas et al. (2008), Greselin et al. (2009), Necir et Meraghi (2009), Necir et Meragni (2010), Brahim et al. (2011), Peng et al. (2012).

Jones et Zitikis (2003) utilisent la théorie asymptotique pour les L -statistiques pour prouver que, pour les distributions sous-jacentes avec un nombre suffisant de moments finis et sous certaines conditions de régularité sur la fonction de distorsion g , l'estimateur empirique d'une prime de risque $\Psi[F_n; g]$ de $\Psi[F; g]$ est fortement consistant et asymptotiquement normal. Plus précisément

$$n^{1/2} (\Psi[F_n; g] - \Psi[F; g]) \xrightarrow{d} \mathcal{N}(0, \sigma_F^2), \text{ quand } n \rightarrow \infty, \quad (5.4)$$

à condition que la variance asymptotique

$$\sigma_F^2 := \int_0^1 \int_0^1 (\min(s, t) - st) g'(s) g'(t) dQ(1-s) dQ(1-t),$$

est finie. Les mêmes auteurs ont également discuté des estimateurs des primes d'hasard proportionnelle qui correspond à la fonction de distorsion de puissance $g(x) = x^{1/\rho}$, $\rho \geq 1$. Dans ce cas, la normalité asymptotique (5.4) est vérifiée pour tout $1 < \rho < 2$, à condition que

$$\mathbf{E}(X^\nu) < \infty \text{ pour certains } \nu > 2\rho/(2 - \rho).$$

5.1 Estimation semi-paramétrique

5.1.1 Motivation d'estimation

Des études empiriques ont montré que les données financières et actuarielles suivent des distributions à queue lourde ou de type Pareto. La classe des distributions à variation régulière est une sous-classe importante de distributions à queue lourde, il comprend des distributions comme Pareto, Burr, student, Lévy-stable, et log-gamma, qui sont connues pour être des modèles appropriés pour modéliser les grandes réclamations d'assurance, les grandes fluctuations des prix, log-retours, etc. (voir, par exemple, Beirlant et al. 2001; Reiss et Thomas, 2007 et Rolski et al. 1999). La queue de distribution de la plupart de ces modèles peut être écrite comme suit

$$1 - F(x) = x^{-1/\gamma} \ell_F(x), \text{ pour chaque } x > 0, \quad (5.5)$$

où $\gamma > 0$ et ℓ_F est une fonction à variation lente.

Supposons que $1 - F(x) = x^{-1/\gamma}$, $x \geq 1$. Ensuite, la prime de risque PHT , qui est supérieur ou égal au risque moyen, doit être finie pour toute $\rho \geq 1$. Cela

signifie que : $1 \leq \rho < 1/\gamma$. D'autre part, pour $\gamma > 1/2$, nous avons $1 \leq \rho < 2$. Toutefois, le moment d'ordre $\mathbf{E}(X^2)$ est infini. Dans ce cas, $2\rho/(2-\rho) > 2$, ce qui implique que $\mathbf{E}(X^\nu)$ est infini pour toute $\nu > 2\rho/(2-\rho)$. Ainsi, l'intervalle $\gamma \in (1/2, 1)$ n'est pas couvert par le théorème central limite (5.4) et donc nous devons chercher une autre approche pour gérer cette situation.

En faisant usage de la théorie des valeurs extrêmes (voir par exemple de Haan et Ferreira, 2006), Necir et Meraghni (2009) ont proposé un estimateur alternatif pour $\Psi[F, g]$ et ont établi sa normalité asymptotique pour tout $\gamma \in (1/2, 1)$. Bien que cet estimateur jouit de la propriété de normalité asymptotique, il a encore un problème dû au fait que, il est basé sur l'estimateur des quantiles extrêmes de Weissman connus pour être largement biaisée. Plusieurs estimateurs de quantile extrêmes à biais réduits sont proposés dans la littérature comme alternatifs à ceux de Weissman, voir, par exemple, Feureverger et Hall (1999), Beirlant et al. (2002), Gomes et Martins (2002), Gomes et martins (2004), Caeiro et al. (2004), Caeiro et al. (2009), Peng et Qi (2004), Matthyhs et Beirlant (2003), Gomes et Figueiredo (2006) et Gomes et Pestana (2004). En fait, nous nous intéressons aux estimateurs de quantiles extrêmes à biais réduit avec de belles propriétés asymptotiques qui nous fournissent des approximations gaussiennes menant à des intervalles de confiance pour $\Psi[F, g]$. Notre choix est celui proposé par Matthyhs et Beirlant (2003).

Dans la prochaine section, nous présentons une brève introduction sur l'estimateur de la prime de risque dans le cas des pertes à queue lourde qui est proposées par Necir et Meraghni (2009) et puis on va définir un nouvel estimateur à biais réduit de cette classe des primes de risque. Nous énonçons nos principaux résultats dans la section 5.2.

5.1.2 Estimateur de Necir et Meraghni $\Psi[\tilde{Q}_n, g]$

Comme il est mentionné dans la section 5.1.1, l'utilisation des quantiles empiriques pour estimer la prime de risque $\Psi[F, g]$ ne garantie pas la normalité asymptotique. Par conséquent, Necir et Meraghni (2009) ont adopté une autre méthode d'estimation basée sur les quantiles extrêmes. L'idée principale est la suivante : Nous définissons l'estimateur de Q comme suit:

$$\tilde{Q}_n(1-s) := \begin{cases} Q_n^W(1-s) & \text{si } 0 < s < k/n \\ Q_n(1-s) & \text{si } k/n \leq s < 1, \end{cases}$$

où $Q_n^W(1-s) := (k/n)^{\widehat{\gamma}_n^h} X_{n-k:n} s^{-\widehat{\gamma}_n^h}$, $s \downarrow 0$, est l'estimateur de Weissman des quantiles extrêmes (Weissman, 1978), avec

$$\widehat{\gamma}_n^h := k^{-1} \sum_{i=1}^k \log \left(\frac{X_{n-i+1:n}}{X_{n-k:n}} \right),$$

étant l'estimateur connu de Hill de l'indice de queue γ (Hill, 1975). La séquence des entier $k = k_n$, représente le nombre de statistiques d'ordre supérieurs utilisées dans le calcul de $\widehat{\gamma}_n^h$, satisfaisant

$$1 < k < n, \quad k \rightarrow \infty \text{ et } k/n \rightarrow 0 \text{ quand } n \rightarrow \infty. \quad (5.6)$$

En remplaçant $Q(1-s)$ par $\tilde{Q}_n(1-t)$ dans la formule (4.1), Necir et Meraghni (2009) ont proposé un estimateur alternatif pour $\Psi[Q; g]$ comme suit :

$$\Psi[\tilde{Q}_n; g] := \int_0^1 \tilde{Q}_n(1-s) dg(s) = \int_0^{k/n} Q_n^W(1-s) dg(s) + \int_{k/n}^1 Q_n(1-s) dg(s)$$

qui, par un calcul simple égale à

$$\Psi[\tilde{Q}_n; g] = g(k/n) \frac{X_{n-k:n}}{1 - \rho \widehat{\gamma}_n^h} + \sum_{i=k+1}^n a_{i,n}(g) X_{n-i+1:n}, \quad (5.7)$$

où $a_{i,n}(g)$ est celle définie dans (5.3) à condition que $\rho \widehat{\gamma}_n^h \in (0, 1)$. Par ailleurs, les mêmes auteurs ont montré que, sous des hypothèses de régularité appropriées, pour toute $\gamma \in (1/2, 1)$, cet estimateur est asymptotiquement normal

Loi asymptotique

La normalité asymptotique de $\Psi[\tilde{Q}_n; g]$ est liée à celle de Hill, qui est équivalente à la condition de variation régulière de second ordre liées à (5.5) (voir par exemple, Geluk al. 1997). Une fonction de distribution F remplit la condition de variation régulière de second ordre avec le paramètre du second ordre $\omega \leq 0$, s'il existe une fonction positive ou négative $a(\cdot)$ avec $\lim_{t \rightarrow \infty} a(t) = 0$, telle que

$$\lim_{t \rightarrow \infty} (a(t))^{-1} \left(\frac{1 - F(tx)}{1 - F(t)} - x^{-1/\gamma} \right) = x^{-1/\gamma} \frac{x^{\omega/\gamma} - 1}{\omega/\gamma}, \quad \text{pour tout } x > 0 \quad (5.8)$$

Si $\omega = 0$, interpréter $\frac{x^{\omega/\gamma} - 1}{\omega/\gamma}$ comme $\log x$. En terme de la fonction inverse généralisée Q .

$$\lim_{s \downarrow 0} (A(s))^{-1} \left(\frac{Q(1-sx)}{Q(1-s)} - x^{-\gamma} \right) = x^{-\gamma} \frac{x^\omega - 1}{\omega}$$

où $A(s) := \gamma^2 a(Q(1-s))$ (voir par exemple, de Haan et Standtmüller, 1996 ou Théorème 3.2.9, page 48 de Haan et Ferreira, 2006). Par exemple, si l'on considère le cas particulier où $1 - F(x) = cx^{-1/\gamma}(1+x^{-\eta})$, pour des constantes positives c, γ et η , on peut choisir $a(t) = \eta t^{-\eta}$ et ω serait égal à $-\eta$ dans (5.8). Comme $Q(1-s) = (cs)^{-\gamma}(1+o(1))$, pour s assez petit, alors nous obtenons

$$A(s) = \eta\gamma^2 (cs)^{\eta\gamma} (1+o(1)), \quad \text{quand } s \downarrow 0.$$

Le résultat principal de Necir et Meraghni (2009) est donné par le théorème suivant où ils établissent la normalité asymptotique de $\Psi[\tilde{Q}_n; g]$.

Théorème 5.1. (Normalité asymptotique de $\Psi[\tilde{Q}_n, g]$ - Necir et Meraghni, 2009). Soit F une fonction de répartition vérifiant (5.7) avec $\gamma > 1/2$ et supposant que $Q(\cdot)$ est continument différentiable sur $[0, 1]$. Soit $k = k_n$, telle que $k \rightarrow \infty$, $k/n \rightarrow 0$ et $k^{1/2}(A(k/n)) \rightarrow 0$ quand $n \rightarrow \infty$. pour tout $1 \leq \rho < 1/\gamma$ nous avons

$$\frac{\sqrt{n} \left(\Psi[\tilde{Q}_n; g] - \Psi[Q; g] \right)}{(k/n)^{-1/2} g(k/n) Q(1-k/n)} \xrightarrow{d} \mathcal{N}(0, \sigma^2(\gamma; \rho)), \quad \text{quand } n \rightarrow \infty,$$

où

$$\sigma^2(\rho, \gamma) = \frac{\gamma^2(\gamma^2\rho^2 - 2\gamma^2\rho^3 + 4\gamma\rho^2 - 2\gamma\rho + \rho^2 - 2\rho + 1)}{(\gamma\rho - 1)^4} + \frac{2\gamma^2(\rho + \gamma\rho - 1)}{(\gamma\rho - 1)^2(\rho + 2\gamma\rho - 2)}$$

Preuve du théorème 5.1. Nous nous restreindrons à présenter les trois lemmes clés pour la preuve du théorème 5.1, ainsi, nous vous recommandons de voir Necir et Meraghni (2009) pour le déroulement de la démonstration. D'après les hauteurs, pour plus de commodité, ils ont mis pour $\alpha \in \mathbb{R}$,

$$J_\alpha(x) := - \int_x^1 t^\alpha dQ(1-t), \quad \text{pour } 0 < x < 1,$$

et

$$\sigma_\alpha^2(x) := \int_x^1 \int_x^1 (\min(s, t) - st) s^\alpha t^\alpha dQ(1-s) dQ(1-t).$$

Lemme 5.1. Soit F une fonction de distribution satisfaisant (5.4). Pour tout $\alpha < \gamma$ nous avons

$$\lim_{x \downarrow 0} \frac{J_\alpha(x)}{x^\alpha Q(1-x)} = \frac{\gamma}{\gamma - \alpha}.$$

Lemme 5.2. Soit F une fonction de distribution satisfaisant (5.4) avec $\gamma > 1/2$. Pour tout $\alpha \in \mathbb{R}$ et pour tout $0 < x < 1$, nous avons

$$\sigma_\alpha^2(x) = \int_x^1 J_\alpha^2(t) dt + xJ_\alpha^2(x) - J_{\alpha+1}^2(x),$$

où $J_\alpha^2(\cdot)$ représente $(J_\alpha^2(\cdot))^2$. De plus, pour tout $-1/2 < \alpha < 0$ nous avons

$$\lim_{x \downarrow 0} \frac{\sigma_\alpha^2}{x^{2\alpha+1} Q^2(1-x)} = \frac{2\gamma^2}{(\alpha - \gamma)(2\alpha + 1 - 2\gamma)},$$

à condition que $J_{\alpha+1}(0) < \infty$.

Pour le troisième lemme, ils introduisent une suite ξ_1, ξ_2, \dots de variables aléatoires indépendantes uniformes sur $[0, 1]$. Pour tout entier n , la fonction quantiles empirique uniforme est définie par $V_n(t) = \xi_{i,n}$, pour $(i-1)/n < t \leq i/n, i \in \mathbb{N}$, avec $V_n(0) = \xi_{1,n}$ où $\xi_{1,n} \leq \dots \leq \xi_{n,n}$ désignent les statistiques d'ordre basées sur ξ_1, \dots, ξ_n . Soit $\{v_n(s), 0 < s < 1\}_{n \in \mathbb{N}}$ une suite de variables aléatoires à valeurs dans l'intervalle ouvert d'extrémités $s \in (0, 1)$ et $1 - V_n(1-s)$.

Lemme 5.3. Soit F une fonction de distribution satisfaisant (5.4) et supposons que $Q(\cdot)$ est continûment différentiable sur $[0, 1]$. Soit $k = k_n$ tel que $k \rightarrow \infty$ et $k/n \rightarrow 0$. Alors, pour tout $\epsilon > 0$ fixe assez petit et pour tout n assez grand, nous avons

$$\sup_{k/n \leq s \leq \epsilon} \left| \frac{Q'(1-v_n(s))}{Q'(1-s)} - 1 \right| = o_p(1),$$

et

$$\sup_{\epsilon \leq s \leq 1} \left| Q'(1-v_n(s)) - Q'(1-s) \right| = o(1), \quad \text{p.s}$$

Intervalles de confiance

Le corollaire suivant est plus pratique que le théorème 5.1 dans le sens où il fournit des intervalles de confiance de $\Psi[Q; g]$ d'une manière directe.

Corollaire 5.1. (Necir et Meraghni, 2009) Sous les hypothèses du théorème 5.2, nous avons

$$\frac{n^{1/2} \left(\Psi[\tilde{Q}_n; g] - \Psi[Q; g] \right)}{(k/n)^{1/\rho-1/2} X_{n-k,n} \sigma(\rho; \hat{\gamma}_n^h)} \xrightarrow{D} N(0,1), \quad \text{quand } n \rightarrow \infty$$

Calcul des bornes. Tout d'abord, fixer l'indice d'aversion au risque $\rho \leq 1$ et soit pour $0 < \zeta < 1$, z_ζ est le $(1 - \zeta)$ -quantile de la distribution normale $\mathcal{N}(0,1)$. Supposons que, pour n assez grand, nous avons une réalisation (x_1, \dots, x_n) de l'échantillon (X_1, \dots, X_n) des montants de réclamation avec fonction de répartition F remplissant toutes les hypothèses du théorème 4.2. Les $(1 - \zeta)$ -intervalles de confiance de $\Psi[Q; g]$ seront obtenus via les deux prochaines étapes :

1. Choisissez le nombre de valeurs extrêmes optimal k^* des plus grandes statistiques d'ordre utilisées pour calculer l'estimateur de Hill (Voir section 6.1.1) et déterminer la valeur associée de $X_{n-k^*,n}$.
2. Calculer les valeurs correspondantes de $\Psi[\tilde{Q}_n; g]$ et $\sigma_\rho(\rho; \hat{\gamma}_n^h)$ désignées respectivement par Ψ^* et σ_ρ^* . Ensuite, les $(1 - \zeta)$ -intervalles de confiance pour $\Psi[Q; g]$ seront

$$\left(\Psi^* - z_{\zeta/2} \frac{\sigma_\rho^*(k^*/n)^{1/\rho-1/2} X_{n-k^*,n}}{\sqrt{n}}, \Psi^* + z_{\zeta/2} \frac{\sigma_\rho^*(k^*/n)^{1/\rho-1/2} X_{n-k^*,n}}{\sqrt{n}} \right)$$

5.1.3 Nouveau estimateur à biais réduit $\Psi[\hat{Q}_n, g]$

La construction de notre nouvel estimateur est basée sur le fait que la queue est modélisée par l'estimateur des quantiles extrêmes de Matthys et Beirlant (2003). Notre idée consiste à utiliser la classe des estimateurs de quantiles extrêmes précités pour estimer la queue de distribution et une estimation empirique pour calculer les risques ordinaires en dessous du seuil $X_{n-k,n}$. L'estimateur des quantiles extrêmes de Matthys-Beirlant est défini comme suit

$$Q_n^{\text{MB}}(1-s) = X_{n-k:n} + \hat{a}_{n,k} \frac{(k/n)^{\hat{\gamma}_k} s^{-\hat{\gamma}_k} - 1}{\hat{\gamma}_k}, \quad 0 < s < k/n,$$

Où $\hat{\gamma}_k$ est un estimateur du maximum de vraisemblance de l'indice de queue $\gamma > 0$, donné dans et (3.7) défini par

$$\hat{\gamma}_k = \arg \max_{\gamma > 0} \sum_{i=1}^{k-1} \left\{ \log \frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} Y_j \right\},$$

et

$$\hat{a}_{n,k} = \frac{1}{k} \sum_{i=1}^k j (X_{n-j+1:n} - X_{n-j:n}) \left(\frac{j}{k+1} \right)^{\hat{\gamma}_k},$$

avec

$$Y_j := j \log \frac{X_{n-j+1:n} - X_{n-k:n}}{X_{n-j:n} - X_{n-k:n}}, \quad 1 \leq j < k.$$

Au lieu de \hat{Q}_n , nous définissons l'estimateur suivant pour Q :

$$\hat{Q}_n(1-s) := \begin{cases} Q_n^{\text{MB}}(1-s) & \text{for } 0 < s < \frac{k}{n} \\ Q_n(1-s) & \text{for } \frac{k}{n} \leq s < 1, \end{cases}$$

Ensuite, l'estimateur à biais réduit de $\Psi[Q; g]$ est

$$\Psi[\hat{Q}_n; g] = \int_0^1 \hat{Q}_n(1-s) dg(s),$$

qui par un simple calcul nous donne :

$$\Psi[\hat{Q}_n; g] = g(k/n) X_{n-k:n} + g(k/n) \frac{\rho \hat{a}_{n,k}}{1 - \rho \hat{\gamma}_k} + \sum_{i=k+1}^n a_{i,n}(g) X_{n-i+1:n}, \quad (5.9)$$

où les $a_{i,n}$ sont comme dans (5.3). Observer que $a_{n,k} \sim \gamma Q(1 - k/n)$. Par conséquent, il est facile de montrer que $a_{n,k} = \hat{\gamma}_k X_{n-k:n} + o_p(1)$. Il est clair que le choix de $a_{n,k} = \hat{\gamma}_k X_{n-k:n}$, dans (5.9), coïncide avec celui de Necir et Meraghni (2009).

Ce nouveau estimateur est définie en terme des paramètres ρ et k et sa construction est basée sur les statistiques d'ordre supérieurs ($X_{i,n}, k \leq i < n$) et sa normalité asymptotique est possible du fait qu'il est construit en fonction d'un estimateur de queue asymptotiquement normal $\hat{\gamma}_k$.

Petite simulation de $\Psi[\hat{Q}_n; g]$

Afin de mieux comprendre le comportement de l'estimateur $\Psi[\hat{Q}_n, g]$, nous avons simulé des échantillons de variables aléatoires *iid* suivant une loi de Pareto. En particulier, nous étudions l'évolution de l'estimateur en fonction de k , où, les graphiques montrent que, lorsque le nombre de valeurs extrême est relativement plus petit que n , l'estimateur $\Psi[\hat{Q}_n; g]$ présente des valeurs aberrantes et à partir de $k/n \approx 0, 1$, $\Psi[\hat{Q}_n; g]$ commence à s'approcher de plus en plus de la valeur théorique de $\Psi[F; g]$. Nous concluons, qu'il est possible de jouer sur le choix du nombre des valeurs extrêmes k vu la grande marge ou les valeurs de notre estimateur paru proche des valeurs réelles de la prime théorique. Pour des échantillons inférieurs, l'estimateur $\Psi[\hat{Q}_n; g]$ est aussi une bonne statistique à utiliser, en étant prudent dans notre choix de k (voir section 6.1.1).

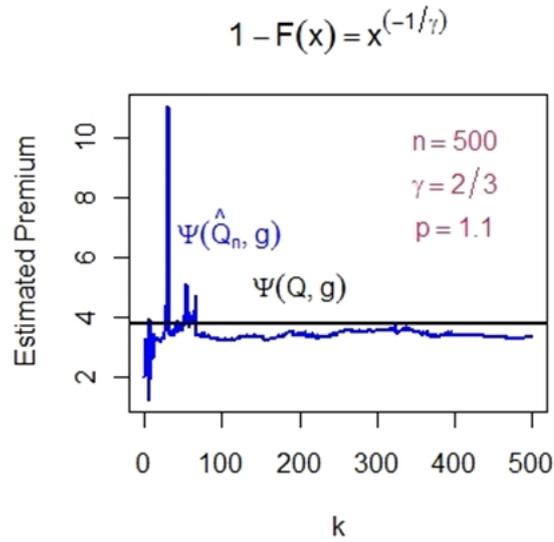


FIG. 5.1 – Graphe de $\Psi \left[\hat{Q}_n, g \right]$ vs le nombre des plus grande statistiques d'ordre issue de 100 échantillons d'une distribution de Pareto avec $n = 500$.

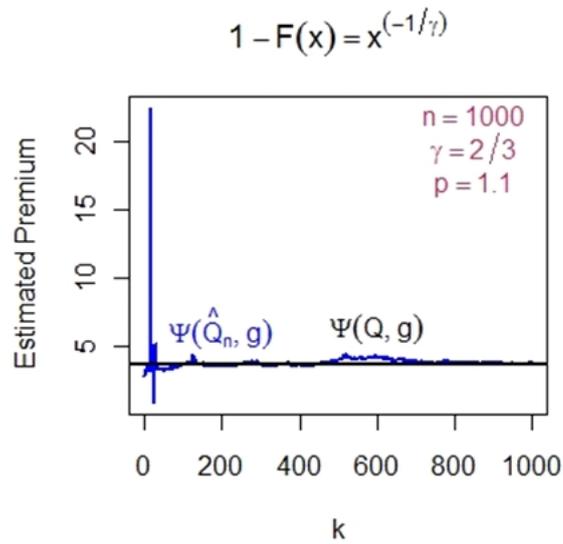


FIG. 5.2 – Graphe de $\Psi \left[\hat{Q}_n, g \right]$ vs le nombre des plus grande statistiques d'ordre issue de 100 échantillons d'une distribution de Pareto avec $n = 1000$.

5.2 Etude asymptotique de $\Psi[\hat{Q}_n, g]$

5.2.1 Théorème de normalité asymptotique

Nous présentons maintenant la fonction quantile de queue définie par

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

Notez que la condition de variation régulière (5.5) implique que,

$$\lim_{t \rightarrow \infty} \mathbb{U}(tx) / \mathbb{U}(t) = x^\gamma, \quad \text{pour tout } x > 0. \quad (5.10)$$

La condition de variation régulière elle-même n'est pas suffisante pour établir les distributions asymptotiques. À cette fin, nous renforçons la condition (5.10), en le suivant : la fonction de répartition F est dite pour remplir la condition de variation régulière du second ordre avec le paramètre du second ordre $\eta \leq 0$ si il existe une fonction $t \rightarrow A(t)$, qui converge vers zéro lorsque t tend vers l'infini, est de signe constant pour les grandes valeurs de t et de telle sorte que

$$\lim_{t \rightarrow \infty} \frac{\frac{\mathbb{U}(tx) - \frac{x^\gamma - 1}{\gamma}}{\mathbb{U}(t)}}{A(t)} = \int_1^x s^{\gamma-1} \int_1^s v^{\eta-1} dv ds, \quad (5.11)$$

Pour plus de détails, voir la section 2.3 Page 43 (de Haan and Ferreira, 2006).

Théorème 5.2. *Supposons que la condition du second ordre (5.11) est vérifiée avec la fonction $A(\cdot)$ et soit $k = k_n$ une suite d'entiers satisfaisant (5.6) avec $\sqrt{k}A(n/k) \rightarrow 0$ quand $n \rightarrow \infty$. Alors il existe une suite de ponts browniens $\{B_n(t), 0 \leq t \leq 1\}$ telle que pour tout n grand*

$$a_\gamma \sqrt{k} (\hat{\gamma}_k - \gamma) \approx \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \int_0^1 t dK_\gamma^{(1)}(t) - \sqrt{\frac{n}{k}} \int_0^1 B_n \left(\frac{k}{n} t \right) dK_\gamma^{(1)}(t),$$

et pour une séquence

$$a_{n,k} = \gamma U(n/k),$$

nous l'avons

$$\begin{aligned} \sqrt{k} \left(\frac{\hat{a}_{n,k}}{a_{n,k}} - 1 \right) &\approx \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \int_0^1 t dK_\gamma^{(2)}(t) - \sqrt{\frac{n}{k}} \int_0^1 B_n \left(\frac{k}{n} t \right) dK_\gamma^{(2)}(t) \\ &\quad + \gamma \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right), \end{aligned}$$

où

$$K_\gamma^{(1)}(t) := \frac{1}{\gamma} \log t + \frac{1+\gamma}{\gamma^2} \int_1^{t^\gamma} \frac{\log u}{1-u} du, (t \geq 0),$$

$$K_\gamma^{(2)}(t) := (\gamma + 1) \log t + a_\gamma^{-1} K_\gamma^{(1)}(t),$$

et

$$a_\gamma := \gamma^{-2} \int_0^1 (1 - u^\gamma + u^\gamma \log u^\gamma)^2 (1 - u^\gamma)^{-2} du. \quad (5.12)$$

Corollaire 5.2. *Sous les hypothèses du théorème 5.2, nous avons*

$$\left(\sqrt{k}(\hat{\gamma}_k - \gamma), \sqrt{k} \left(\frac{\hat{a}_{n,k}}{a_{n,k}} - 1 \right) \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma_\gamma^2 \right),$$

où

$$\Sigma_\gamma^2 := (\sigma_{ij}^2)_{i,j=1,2}$$

est la matrice de variance-covariance asymptotique défini par

$$\sigma_{ij}^2 = \sigma_{ji}^2 = \mathbf{Cov} [K_\gamma^{(i)}(\xi), K_\gamma^{(j)}(\xi)], i \neq j, \sigma_1^2 = \text{Var} [K_\gamma^{(1)}(\xi)]$$

et

$$\sigma_2^2 = \text{Var} [K_\gamma^{(2)}(\xi)] + \gamma^2,$$

avec ξ est une variable aléatoire uniforme sur $(0, 1)$

Théorème 5.3. *Supposons que la condition de variation régulière du second ordre (5.11) détient avec certains $\gamma \in (1/2, 1)$, et soit la fonction quantile Q continue sur l'intervalle $(0, 1)$. En outre, soit $k = k_n$ des entiers satisfaisant la condition (5.6) tel que*

$$\sqrt{k}A(n/k) \rightarrow 0 \text{ quand } n \rightarrow \infty.$$

Alors, pour tout $1 < \rho < 1/\gamma$, avec,

$$K_\gamma^{(3)}(t) = \frac{\rho^2}{(1 - \rho\gamma)^2} K_\gamma^{(1)}(t) + \frac{\rho}{1 - \rho\gamma} K_\gamma^{(2)}(t).$$

Nous avons :

$$\begin{aligned} \frac{\sqrt{n} \left(\Psi[\hat{Q}_n; g] - \Psi[Q; g] \right)}{(k/n)^{-1/2} g(k/n) a_{n,k}} &= \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \int_0^1 t dK_\gamma^{(3)}(t) \\ &\quad - \sqrt{\frac{n}{k}} \int_0^1 B_n \left(\frac{k}{n} t \right) dK_\gamma^{(3)}(t) \\ &\quad + \frac{1}{1 - \rho\gamma} \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \\ &\quad - \frac{\int_0^{1-k/n} g'(1-t) B_n(t) dQ(t)}{(k/n)^{-1/2} g(k/n) a_{n,k}}, \end{aligned}$$

Corollaire 5.3. *Sous les hypothèses du théorème 5.3, nous avons*

$$\frac{\sqrt{n} \left(\Psi[\hat{Q}_n; g] - \Psi[Q; g] \right)}{(k/n)^{-1/2} g(k/n) a_{n,k}} \xrightarrow{d} \mathcal{N} \left(0, D^2(\gamma, \rho) \right), \text{ quand } n \rightarrow \infty,$$

où

$$\begin{aligned} D^2(\gamma, \rho) &= \mathbf{Var} \left[K_\gamma^{(3)}(\xi) \right] + \frac{1}{(1 - \rho\gamma)^2} + \frac{2\rho^2}{1 - \rho - \gamma\rho} \times \\ &\quad \left(\frac{\rho + 3\gamma\rho - 3}{(1 - \gamma\rho)(\rho + 2\gamma\rho - 2)} + \int_0^1 t dK_\gamma^{(3)}(t) \right). \end{aligned}$$

5.2.2 Démonstrations

Preuve du théorème 5.2

Distribution asymptotique de $\hat{\gamma}_k$. Tout d'abord, pour plus de commodité, nous avons mis

$$\mathcal{L}(\gamma; \mathbf{Y}) := \sum_{i=1}^{k-1} \left\{ \log \frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma}{\gamma} Y_j \right\},$$

avec $\mathbf{Y} := (Y_1, \dots, Y_{k-1})$. Par construction, l'estimateur de l'indice de queue $\hat{\gamma}_k$ de γ , proposé par Matthys et Beirlant (2003) et est la valeur de γ qui

maximise la quantité $\mathcal{L}(\gamma; \mathbf{Y})$. En d'autres termes, $\hat{\gamma}_k$ est la solution de l'équation suivante :

$$\frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} = 0. \quad (5.13)$$

Faire usage d'un développement de Taylor de la fonction $\gamma \rightarrow \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma}$ autour de $\hat{\gamma}_k$, nous obtenons

$$\left[\frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} \right]_{\gamma=\hat{\gamma}_k} - \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} = \frac{1}{k} \left[\frac{\partial^2 \mathcal{L}(\gamma; \mathbf{Y})}{\partial^2 \gamma} \right]_{\gamma=\hat{\gamma}_k^*} (\hat{\gamma}_k - \gamma).$$

Comme $\hat{\gamma}_k$ est la solution de l'équation (5.13), donc

$$(\hat{\gamma}_{k+1} - \gamma) \frac{1}{k} \left[\frac{\partial^2 \mathcal{L}(\gamma; \mathbf{Y})}{\partial^2 \gamma} \right]_{\gamma=\hat{\gamma}_k^*} = - \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma}.$$

D'autre part, Matthys et Beirlant (2003) ont montré que, sous la condition du second ordre (5.11),

$$\frac{1}{\sqrt{k}} \frac{\partial^2 \mathcal{L}(\gamma; \mathbf{Y})}{\partial^2 \gamma} = -a_\gamma + o_p(1), \text{ quand } n \rightarrow \infty,$$

où a_γ est celui défini en (5.12). Par conséquent, nous avons

$$a_\gamma \sqrt{k} (\hat{\gamma}_k - \gamma) \approx \frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma}.$$

Notre tâche consiste maintenant à rapprocher, asymptotiquement, la quantité $\frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma}$ par une suite de variables aléatoire gaussiennes. Plus précisément, nous allons représenter cette quantité par les processus quantile uniformes définis ci-après, et alors faire usage de la faible approximation donnée dans le théorème 2.1 de Csörgő et al. (1986). À cette fin, soit ξ_1, ξ_2, \dots une suite de variables aléatoires indépendantes uniformes sur $(0, 1)$. Pour tout $n \geq 1$ définissons

$$\mathbb{G}_n(t) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}(\xi_i \leq t), \quad 0 \leq t \leq 1,$$

la fonction de distribution empirique uniforme, défini comme étant le droit continue, basée sur ξ_1, \dots, ξ_n via la fonction indicatrice $\mathbb{I}(\cdot)$. La fonction quantile empirique uniforme est donnée par

$$\mathbb{V}_n(t) := \inf \{s : \mathbb{G}_n(s) \geq t\}, \quad 0 \leq s \leq 1, \quad \mathbb{V}_n(0) := \mathbb{V}_n(0+),$$

à savoir \mathbb{V}_n est définie comme étant l'inverse continue à gauche de la fonction de répartition empirique \mathbb{G}_n . En terme de la statistique d'ordre $\xi_{1:n} \leq \dots \leq \xi_{n:n}$, se rapportant à l'échantillon ξ_1, \dots, ξ_n , nous avons $\mathbb{V}_n(t) = \xi_{i:n}$, $(i-1)/n < t \leq i/n$. Le processus de quantile correspondant uniforme est définie par :

$$\beta_n(t) := \sqrt{n}(t - \mathbb{V}_n(t)), \quad 0 \leq t \leq 1. \quad (5.14)$$

Les deux suites de statistiques d'ordre $X_{1:n} \leq \dots \leq X_{n:n}$ et $\xi_{1:n} \leq \dots \leq \xi_{n:n}$ sont liées par une égalité en distribution comme, $\{X_{n-j+1:n}\}_{j=1}^n \stackrel{d}{=} \{\mathbb{U}(\xi_{j:n}^{-1})\}_{j=1}^n$. Matthys et Beirlant (2003) ont également montré que, pour tous n grands

$$\frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} = \frac{1}{\sqrt{k}} \sum_{j=2}^{k-1} c_{k,j} H\left(\frac{\xi_{j:n}}{\xi_{k+1:n}}\right) - \sqrt{k} \mu_k + o_p(1),$$

où $H(t) = -\log(t^{-\gamma} - 1)$,

$$c_{k,j} = -(k+1) \int_{\frac{j-1}{k+1}}^{\frac{j}{k+1}} J(s) ds \quad \text{et} \quad \mu_k = \sum_{j=2}^{k-1} c_{k,j} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} H(t) dt.$$

Il est facile de montrer que $\mu_k = \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) H(t) dt + O(k^{-1})$. Il s'ensuit donc que

$$\frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} = \frac{1}{\sqrt{k}} \sum_{j=2}^{k-1} c_{k,j} H\left(\frac{\xi_{j:n}}{\xi_{k+1:n}}\right) - \sqrt{k} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) H(t) dt + o_p(1).$$

Dans la suite, par souci de confusion, afin d'éviter les grandes formules, nous avons mis $\omega_{n,k}(t) = \frac{\mathbb{V}_n(\frac{k+1}{n}t)}{\xi_{k+1:n}}$.

Maintenant, nous faisons usage de la fonction quantile empirique $\mathbb{V}_n(\cdot)$ et réécrire le premier terme du côté droit de l'équation précédente en

$$\frac{k+1}{\sqrt{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) H(\omega_{n,k}(t)) dt.$$

En utilisant le théorème de la valeur moyenne de la fonction $t \rightarrow H(t)$, nous obtenons $H(\omega_{n,k}(t)) - H(t) = H'(\vartheta_{n,k}(t))(\omega_{n,k}(t) - t)$, où $\vartheta_{n,k}(t)$ est une suite de variables aléatoires appartenant à un intervalle ouvert avec les points terminaux $\omega_{n,k}(t)$ et t . Ainsi nous arrivons à

$$\frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} \approx \frac{k+1}{\sqrt{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) H'(t) (\omega_{n,k}(t) - t) dt + R_n^{(1)}, \quad (5.15)$$

où

$$R_n^{(1)} := \frac{k+1}{\sqrt{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) \left(\frac{H'(\vartheta_{n,k}(t))}{H'(t)} - 1 \right) H'(t) (\omega_{n,k}(t) - t) dt + o_p(1).$$

Nous pouvons traiter ce terme reste $R_n^{(1)}$ comme l'a fait Necir et Meraghni (2009) dans leur théorème 1 et montrer qu'il tend vers zéro en probabilité quand $n \rightarrow \infty$ en omettant les détails. Du Lemme de Smirnov (voir le lemme 2.2.3 dans de Haan et Ferreira, 2006, à la page 41), nous avons $\xi_{k+1:n} \approx k/n$, il s'ensuit que le premier terme de la partie droite de l'équation (5.15) est

$$\approx \frac{n}{\sqrt{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) \left(\mathbb{V}_n \left(\frac{k+1}{n} t \right) - \xi_{k+1:n} t \right) dH(t).$$

Observez maintenant que l'expression précédente est égale à

$$\begin{aligned} \frac{n}{\sqrt{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) \left(\mathbb{V}_n \left(\frac{k+1}{n} t \right) - \frac{k+1}{n} t \right) dH(t) \\ - \frac{n}{\sqrt{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) \left(\xi_{k+1:n} t - \frac{k+1}{n} t \right) dH(t). \end{aligned}$$

Faisant usage de la procédure quantile empirique uniforme $\beta_n(\cdot)$, défini dans (5.14), nous écrivons

$$\begin{aligned} \frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} \approx -\sqrt{\frac{n}{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) \beta_n \left(\frac{k+1}{n} t \right) dH(t) \\ + \sqrt{\frac{n}{k}} \beta_n \left(\frac{k+1}{n} \right) \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} t J(t) dH(t). \end{aligned}$$

D'après le théorème 2.1 de Csörgő et al. (1986), sur un espace de probabilité convenable, il existe une suite de ponts browniens $\{B_n(t), 0 \leq t \leq 1\}$ de telle sorte que, pour chaque $0 \leq \delta < 1/2$, nous avons, comme $n \rightarrow \infty$

$$\sup_{1/n \leq s \leq 1-1/n} \frac{|\beta_n(s) - B_n(s)|}{(s(1-s))^{1/2-\delta}} = O_p(n^{-\tau}). \quad (5.16)$$

En faisant usage de ce rapprochement, nous obtenons facilement

$$\begin{aligned} \frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} = -\sqrt{\frac{n}{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) B_n \left(\frac{k+1}{n} t \right) dH(t) \\ + \sqrt{\frac{n}{k}} B_n \left(\frac{k+1}{n} \right) \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} t J(t) dH(t) + R_n^{(2)}. \end{aligned}$$

Là encore, montrant que le terme reste $R_n^{(2)}$ tend vers zéro en probabilité. Il est facile de montrer que

$$\sqrt{\frac{n}{k}} \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} J(t) B_n\left(\frac{k+1}{n}t\right) dH(t) \approx \sqrt{\frac{n}{k}} \int_0^1 J(t) B_n\left(\frac{k}{n}t\right) dH(t)$$

et

$$\sqrt{\frac{n}{k}} B_n\left(\frac{k+1}{n}\right) \int_{\frac{1}{k+1}}^{\frac{k}{k+1}} tJ(t) dH(t) \approx \sqrt{\frac{n}{k}} B_n\left(\frac{k}{n}\right) \int_0^1 tJ(t) dH(t).$$

Finalement, nous obtenons

$$\frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} \approx \sqrt{\frac{n}{k}} B_n\left(\frac{k}{n}\right) \int_0^1 tJ(t) dH(t) - \sqrt{\frac{n}{k}} \int_0^1 J(t) B_n\left(\frac{k}{n}t\right) dH(t).$$

en mettant $dK_\gamma^{(1)}(t) = -J(t) dH(t)$, nous obtenons

$$a_\gamma \frac{1}{\sqrt{k}} \frac{\partial \mathcal{L}(\gamma; \mathbf{Y})}{\partial \gamma} \approx \sqrt{\frac{n}{k}} \int_0^1 B_n\left(\frac{k}{n}t\right) dK_\gamma^{(1)}(t) - \sqrt{\frac{n}{k}} B_n\left(\frac{k}{n}\right) \int_0^1 t dK_\gamma^{(1)}(t).$$

Cela achève la preuve de la première partie du théorème 5.2

Distribution asymptotique de $\hat{a}_{n,k}$. Rappelons que $a_{n,k} = \gamma \mathbb{U}(n/k)$. en mettant $\tilde{a}_{n,k} = \gamma \mathbb{U}(\xi_{k+1:n}^{-1})$ nous écrivons

$$\frac{\hat{a}_{n,k}}{a_{n,k}} - 1 = \left(\frac{\hat{a}_{n,k}}{\tilde{a}_{n,k}} - 1 \right) + \frac{\hat{a}_{n,k}}{\tilde{a}_{n,k}} \left(\frac{\tilde{a}_{n,k}}{a_{n,k}} - 1 \right). \quad (5.17)$$

Ensuite, nous montrons que chacun des deux termes de (5.17) sont des variables aléatoires asymptotiquement gaussiennes. Faisant usage de la condition du second ordre 5.11), nous obtenons

$$\frac{X_{n-j+1:n} - X_{n-j:n}}{a_{n,k}} \approx \frac{\left(\frac{\xi_{j:n}}{\xi_{k+1:n}} \right)^{-\gamma} - \left(\frac{\xi_{j+1:n}}{\xi_{k+1:n}} \right)^{-\gamma}}{\gamma}, \text{ pour } 1 \leq j \leq k.$$

Nous avons également

$$\begin{aligned} \frac{\hat{a}_{n,k}}{a_{n,k}} &= \frac{1}{k} \sum_{j=1}^k j \frac{X_{n-j+1:n} - X_{n-j:n}}{a_{n,k}} \left(\frac{j}{k+1} \right)^{\hat{\gamma}_k} \\ &\approx \frac{1}{k} \sum_{j=1}^k j \frac{\left(\frac{\xi_{j:n}}{\xi_{k+1:n}} \right)^{-\gamma} - \left(\frac{\xi_{j+1:n}}{\xi_{k+1:n}} \right)^{-\gamma}}{\gamma} \left(\frac{j}{k+1} \right)^{\hat{\gamma}_k} =: \Delta_{n,k}. \end{aligned}$$

Cette dernière expression peut être réécrite en

$$\gamma^{-1} \frac{k+1}{k} \xi_{k+1:n}^\gamma \sum_{j=1}^k (\xi_{j:n}^{-\gamma} - \xi_{j+1:n}^{-\gamma}) \left(\frac{j}{k+1} \right)^{\widehat{\gamma}_{k+1}}.$$

Il est facile de vérifier que

$$\begin{aligned} & \sum_{j=1}^k (\xi_{j:n}^{-\gamma} - \xi_{j+1:n}^{-\gamma}) \left(\frac{j}{k+1} \right)^{\widehat{\gamma}_{k+1}} \\ &= \sum_{j=1}^k \left[\left(\frac{j}{k+1} \right)^{\widehat{\gamma}_{k+1}} - \left(\frac{j-1}{k+1} \right)^{\widehat{\gamma}_{k+1}} \right] \xi_{j:n}^{-\gamma} - \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_{k+1}} \xi_{k+1:n}^{-\gamma}. \end{aligned}$$

donc

$$\Delta_{n,k} = \gamma^{-1} \frac{k+1}{k} \xi_{k+1:n}^\gamma \sum_{j=1}^k \left[\left(\frac{j}{k+1} \right)^{\widehat{\gamma}_{k+1}} - \left(\frac{j-1}{k+1} \right)^{\widehat{\gamma}_{k+1}} \right] \xi_{j:n}^{-\gamma - \gamma^{-1}} \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_k},$$

qui égale à

$$\gamma^{-1} (\widehat{\gamma}_k + 1) \frac{k+1}{k} \left(\frac{n}{k+1} \right)^{\widehat{\gamma}_{k+1}} \xi_{k+1:n}^\gamma \sum_{j=1}^k \xi_{j:n}^{-\gamma} \int_{\frac{j-1}{n}}^{\frac{j}{n}} t^{\widehat{\gamma}} dt - \gamma^{-1} \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_k}.$$

En utilisant la fonction des quantiles empirique uniforme $\mathbb{V}_n(\cdot)$, nous réécrivons $\Delta_{n,k}$ comme

$$\gamma^{-1} (\widehat{\gamma}_k + 1) \frac{k+1}{k} \left(\frac{n}{k+1} \right)^{\widehat{\gamma}_{k+1}} \xi_{k+1:n}^\gamma \int_0^{k/n} s^{\widehat{\gamma}_k} \mathbb{V}_n^{-\gamma}(s) ds - \gamma^{-1} \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_k},$$

qui, par intégration par parties, est égale à

$$\gamma^{-1} (\widehat{\gamma}_k + 1) \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_k} \xi_{k+1:n}^\gamma \int_0^1 s^{\widehat{\gamma}_k} \mathbb{V}_n^{-\gamma}(ks/n) ds - \gamma^{-1} \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_k}.$$

donc

$$\begin{aligned} \Delta_{n,k} &= \gamma^{-1} \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_k} \left[(\widehat{\gamma}_k + 1) \xi_{k+1:n}^\gamma \int_0^1 t^{\widehat{\gamma}_k} \mathbb{V}_n^{-\gamma}(kt/n) dt - 1 \right]. \\ &=: \gamma^{-1} \left(\frac{k}{k+1} \right)^{\widehat{\gamma}_k} (\Delta_{n,k}^* - 1). \end{aligned}$$

Par le fait que $\left(\frac{k}{k+1}\right)^{\widehat{\gamma}_k} = 1 + O_p(k^{-1})$, Nous en déduisons que

$$\left(\widehat{a}_{n,k}/a_{n,k} - 1\right) = \left(1 + O_p(k^{-1})\right) \gamma^{-1} \left(\Delta_{n,k}^* - 1\right) - 1.$$

Par conséquent

$$\sqrt{k} \left(\frac{\widehat{a}_{n,k}}{a_{n,k}} - 1\right) = \sqrt{k} \gamma^{-1} \left(\Delta_{n,k}^* - 1 - \gamma\right) + \sqrt{k} O_p(k^{-1}).$$

Comme $\sqrt{k} O_p(k^{-1}) \rightarrow 0$ en probabilité, il suffit de montrer que $\sqrt{k} \left(\Delta_{n,k}^* - 1 - \gamma\right)$ sont des variables aléatoires asymptotiquement gaussiennes centrées.

$$\Delta_{n,k}^* = (\widehat{\gamma}_k + 1) \left(n\xi_{k+1:n}/k\right)^\gamma \left(k/n\right)^\gamma \int_0^1 t^{\widehat{\gamma}_k} \mathbb{V}_n^{-\gamma}(kt/n) dt.$$

Il est facile de vérifier que

$$\Delta_{n,k}^* - 1 - \gamma =: S_n T_{n1} + T_{n2} + T_{n3} + T_{n4},$$

où

$$S_n = (\widehat{\gamma}_k + 1) \left(\frac{k}{n}\right)^\gamma \int_0^1 t^{\widehat{\gamma}_k} \mathbb{V}_n^{-\gamma}\left(\frac{k}{n}t\right) dt,$$

$$T_{n1} = \left(n\xi_{k+1:n}/k\right)^\gamma - 1,$$

$$T_{n2} = (\widehat{\gamma}_k + 1) \left(\frac{k}{n}\right)^\gamma \int_0^1 \left(t^{\widehat{\gamma}_k} - t^\gamma\right) \mathbb{V}_n^{-\gamma}\left(\frac{k}{n}t\right) dt,$$

$$T_{n3} = (\widehat{\gamma}_k + 1) \left(\frac{k}{n}\right)^\gamma \int_0^1 t^\gamma \left[\mathbb{V}_n^{-\gamma}\left(\frac{k}{n}t\right) - \left(\frac{k}{n}t\right)^{-\gamma}\right] dt,$$

et

$$T_{n4} = \widehat{\gamma}_k - \gamma.$$

Ensuite, nous montrerons que T_{ni} , $i = 1, 2, 3$ sont des variables aléatoires asymptotiquement gaussiennes quand $S_n \xrightarrow{p} \gamma + 1$. D'abord, nous considérons que le second terme T_{n2} . Une fois de plus, compte tenu du théorème de la valeur moyenne,

$$\mathbb{V}_n^{-\gamma}\left(\frac{k}{n}t\right) - \left(\frac{k}{n}t\right)^{-\gamma} = -\gamma \left[\zeta_{n,k}(t)\right]^{-\gamma-1} \left(\mathbb{V}_n\left(\frac{k}{n}t\right) - \frac{k}{n}t\right),$$

où $\zeta_{n,k}(t)$ est une suite de variables aléatoires appartenant à l'intervalle ouvert avec les points terminaux $\mathbb{V}_n(kt/n)$ et kt/n . Alors

$$T_{n2} = -(\widehat{\gamma}_k + 1) \left(\frac{k}{n}\right)^\gamma \int_0^1 t^{\widehat{\gamma}_k} \left(\frac{k}{n}t\right)^{-\gamma-1} \left[\mathbb{V}_n\left(\frac{k}{n}t\right) - \left(\frac{k}{n}t\right)\right] dt + R_n^{(2)},$$

où

$$R_n^{(2)} : = -(\widehat{\gamma}_k + 1) \left(\frac{k}{n}\right)^\gamma \times \int_0^1 t^{\widehat{\gamma}_k} \left[\left(\frac{n\zeta_{n,k}(t)}{kt}\right)^{-\gamma-1} - 1 \right] \left(\frac{k}{n}t\right)^{-\gamma-1} \left[\mathbb{V}_n \left(\frac{k}{n}t\right) - \left(\frac{k}{n}t\right) \right] dt.$$

Par des arguments similaires à ceux utilisés ci-dessous, on peut facilement montrer que $\sqrt{k}R_n^{(2)} \xrightarrow{P} 0$, donc les détails sont omis. En faisant usage de la procédure quantile empirique uniforme $\beta_n(\cdot)$, et en utilisant le fait que $\widehat{\gamma}_k \xrightarrow{P} \gamma$, il devient clair que

$$\sqrt{k}T_{n2} = (1 + o_p(1))(\gamma + 1) \sqrt{\frac{k}{n}} \int_0^1 t^{-1} \beta_n \left(\frac{k}{n}t\right) dt.$$

En utilisant des arguments similaires, nous montrons aussi que

$$\sqrt{k}T_{n1} = -(1 + o_p(1))\gamma \sqrt{n/k} \beta_n \left(\frac{k+1}{n}\right).$$

Un calcul élémentaire donne

$$\sqrt{k}T_{n3} = -\sqrt{k} \frac{\widehat{\gamma}_k - \gamma}{\widehat{\gamma}_k - \gamma + 1}.$$

Comme $\widehat{\gamma}_k \xrightarrow{P} \gamma$, puis $\sqrt{k}T_{n3} = -(1 + o_p(1))\sqrt{k}(\widehat{\gamma}_k - \gamma)$. Compte tenu de l'approximation faible (4.16), nous obtenons

$$\begin{aligned} \sqrt{k}T_{n1} &= -\gamma \sqrt{n/k} B_n(k/n) + o_p(1), \\ \sqrt{k}T_{n2} &= (\gamma + 1) \sqrt{n/k} \int_0^1 t^{-1} B_n(kt/n) dt + o_p(1), \end{aligned}$$

et

$$\sqrt{k}T_{n3} = -\sqrt{n/k} \int_0^1 B_n(kt/n) dK_\gamma^{(1)}(t) + \sqrt{n/k} B_n(k/n) \int_0^1 t dK_\gamma^{(1)}(t) + o_p(1).$$

Il est facile de vérifier que

$$S_n = T_{n2} + T_{n3} + \gamma^{-1} + 1.$$

Comme $\sqrt{k}T_{ni}$, $i = 1, 2, 3$ sont des variables aléatoires asymptotiquement gaussiennes, puis

$$T_{ni} \xrightarrow{P} 0.$$

Il s'ensuit que $S_n \xrightarrow{p} \gamma^{-1} + 1$. Finalement

$$\begin{aligned} \sqrt{k} \left(\frac{\widehat{a}_{n,k}}{a_{n,k}} - 1 \right) &= (\gamma + 1) \sqrt{n/k} \int_0^1 t^{-1} B_n(kt/n) dt \\ &\quad - (\gamma + 1) \sqrt{n/k} B_n(k/n) \\ &\quad + \sqrt{n/k} \int_0^1 B_n(kt/n) dK_\gamma^{(1)}(t) \\ &\quad - \sqrt{n/k} B_n(k/n) \int_0^1 t dK_\gamma^{(1)}(t) + o_p(1). \end{aligned}$$

Cela peut être réécrit sous la forme simple

$$\sqrt{k} \left(\frac{\widehat{a}_{n,k}}{a_{n,k}} - 1 \right) = \sqrt{\frac{n}{k}} \int_0^1 B_n\left(\frac{k}{n}t\right) dK_\gamma^{(2)}(t) - \sqrt{\frac{n}{k}} B_n\left(\frac{k}{n}\right) \int_0^1 t dK_\gamma^{(2)}(t) + o_p(1),$$

où $K_\gamma^{(2)}(t) := (\gamma + 1) \log t + a_\gamma^{-1} K_\gamma^{(1)}(t)$. Pour le second terme de l'équation précédente (5.17), nous avons d'abord à partir de l'équation précédente $\widehat{a}_{n,k}/a_{n,k} = 1 + o_p(1)$, et

$$\sqrt{k} \left(\frac{\widetilde{a}_{n,k}}{a_{n,k}} - 1 \right) = \gamma \sqrt{n/k} B_n\left(\frac{k}{n}\right) + o_p(1).$$

Ce dernier complète la preuve du théorème 5.2.

Preuve Corollaire 5.2. La preuve est obtenue par des calculs simples et élémentaires.

Preuve Théorème 5.3. Ecrivons

$$\Psi[\widehat{Q}_n; g] - \Psi[Q; g] = \Delta_{n1} + \Delta_{n2},$$

où

$$\Delta_{n1} := g(k/n) X_{n-k,n} + g(k/n) \frac{\rho \widehat{a}_{n,k}}{1 - \rho \widehat{\gamma}_k} - \int_0^{k/n} Q(1-s) dg(s),$$

et

$$\Delta_{n2} := \sum_{i=k+1}^n a_{i,n}(g) X_{n-i+1,n} - \int_{k/n}^1 Q(1-s) dg(s).$$

Il est facile de vérifier que

$$\begin{aligned}
\frac{\sqrt{n}\Delta_{n1}}{(k/n)^{-1/2} g(k/n) a_{n,k}} &= \gamma^{-1}\sqrt{k} \left(\frac{X_{n-k,n}}{Q(1-k/n)} - 1 \right) + \frac{\rho}{1-\rho\widehat{\gamma}_k} \sqrt{k} \left(\frac{\widehat{a}_{n,k}}{a_{n,k}} - 1 \right) \\
&+ \frac{\rho^2}{(1-\rho\widehat{\gamma}_k)(1-\rho\gamma)} \sqrt{k} (\widehat{\gamma}_k - \gamma) \\
&+ \sqrt{k} \left(\frac{1}{\gamma(1-\rho\gamma)} - \int_0^{k/n} \frac{Q(1-s)}{g(k/n) a_{n,k}} dg(s) \right) \\
&= \sum_{i=1}^4 \Delta_{n1}^{(i)}.
\end{aligned}$$

Compte tenu du théorème 5.2, nous écrivons

$$\begin{aligned}
\Delta_{n1}^{(1)} &\approx \sqrt{n/k} B_n \left(\frac{k}{n} \right), \\
\Delta_{n1}^{(2)} &\approx \frac{\rho}{1-\rho\gamma} \\
&\times \left\{ \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \int_0^1 tdK_\gamma^{(2)}(t) - \sqrt{\frac{n}{k}} \int_0^1 B_n \left(\frac{k}{n}t \right) dK_\gamma^{(2)}(t) + \gamma \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \right\}, \\
\Delta_{n1}^{(3)} &\approx \frac{\rho^2}{(1-\rho\gamma)^2} \\
&\times \left\{ \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \int_0^1 tdK_\gamma^{(1)}(t) - \sqrt{\frac{n}{k}} \int_0^1 B_n \left(\frac{k}{n}t \right) dK_\gamma^{(1)}(t) \right\}.
\end{aligned}$$

donc

$$\begin{aligned}
\frac{\sqrt{n}\Delta_{n1}}{(k/n)^{-1/2} g(k/n) a_{n,k}} &\approx \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right) \int_0^1 tdK_\gamma^{(3)}(t) \\
&- \sqrt{\frac{n}{k}} \int_0^1 B_n \left(\frac{k}{n}t \right) dK_\gamma^{(3)}(t) \\
&+ \frac{1}{1-\rho\gamma} \sqrt{\frac{n}{k}} B_n \left(\frac{k}{n} \right),
\end{aligned}$$

où

$$K_\gamma^{(3)}(t) = \frac{\rho^2}{(1-\rho\gamma)^2} K_\gamma^{(1)}(t) + \frac{\rho}{1-\rho\gamma} K_\gamma^{(2)}(t).$$

Il a été montré dans Necir et Meraghni (2009), en utilisant la condition du second ordre (5.10), que

$$\frac{\sqrt{n}\Delta_{n2}}{(k/n)^{-1/2} g(k/n) a_{n,k}} \approx - \frac{\int_0^{1-k/n} g'(1-t) B_n(t) dQ(t)}{(k/n)^{-1/2} g(k/n) a_{n,k}},$$

et $\Delta_{n1}^{(4)}$ tend vers zéro quand $n \rightarrow \infty$. Cela permet d'obtenir la preuve du théorème 5.3.

Preuve Corollaire 5.3. La preuve se fait également par des calculs simples et élémentaires.

5.2.3 Intervalles de confiance

Sous les hypothèses du corollaire 5.3, nous avons

$$\frac{n^{1/2} \left(\Psi[\hat{Q}_n; g] - \Psi[Q; g] \right)}{(k/n)^{-1/2} g(k/n) a_{n,k} D(\gamma, \rho)} \xrightarrow{D} N(0,1), \quad \text{quand } n \rightarrow \infty$$

Tout d'abord, fixer l'indice d'aversion au risque $\rho \leq 1$ et soit pour $0 < \zeta < 1$, z_ζ est le $(1 - \zeta)$ -quantile de la distribution normale $\mathcal{N}(0,1)$. Supposons que, pour n assez grand, nous avons une réalisation (x_1, \dots, x_n) de l'échantillon (X_1, \dots, X_n) des montants de réclamation avec fonction de répartition F remplissant toutes les hypothèses du corollaire 4.3. Les $(1 - \zeta)$ -intervalles de confiance de $\Psi[Q, g]$ seront obtenus via les deux prochaines étapes :

1. Choisissez le nombre de valeurs extrêmes optimal k^* des plus grandes statistiques d'ordre utilisées pour calculer l'estimateur de Matthys et Beirlant (Voir section 6.1.1).
2. Calculer les valeurs correspondantes de $\Psi[\hat{Q}_n; g]$ et $D(\gamma, \rho)$ désignées respectivement par Ψ^* et D^* . Ensuite, les $(1 - \zeta)$ -intervalles de confiance pour $\Psi[Q; g]$ seront

$$\left(\Psi^* - z_{\zeta/2} \frac{D^* (k^*/n)^{-1/2} g(k^*/n) a_{n,k^*}}{\sqrt{n}}, \Psi^* + z_{\zeta/2} \frac{D^* (k^*/n)^{-1/2} g(k^*/n) a_{n,k^*}}{\sqrt{n}} \right)$$

Chapitre 6

Simulations comparatives et exploration des observations extrêmes des données réelles sous SOA

Dans ce chapitre nous allons présenter dans une première section une simulation comparative entre les deux estimateurs de mesure de risque $\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ déjà présentés dans les chapitres précédents, en étudiant tout d'abord des paramètres nécessaires pour la simulation ainsi que calculer des paramètres de comparaison effectué par la méthode de bootstrap.

Dans une deuxième section nous allons juste préciser quelques étapes à entamer avant d'effectuer une application réelle. Nous avons analysé des données prise sous la base des données de *SOA* des sept compagnies d'assurance présentant des réclamations des années 1997, 1998 et 1999 afin de montrer que ces derniers représentent un échantillon de risque à queue lourde comme une condition indispensable pour la mise en œuvre de notre estimation. Ainsi nous avons calculé la distribution limite de notre nouveau estimateur de mesure de risque par bootstrap pour vérifier la normalité asymptotique empiriquement.

6.1 Simulations et comparaisons

Pour prouver que notre nouvel estimateur $\Psi \left[\hat{Q}_n, g \right]$ est à biais réduit, nous effectuerons une étude de simulation à l'aide du logiciel statistique *R* (voir Ihka et Gentleman, 1996). Nous allons préciser des paramètres de simulation ainsi que d'autres de comparaison

6.1.1 Paramètres de simulation

Dans cette section, nous présenterons les paramètres influençant directement sur les valeurs des estimateurs des mesures de risque $\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ qui sont le nombre de valeurs extrêmes k et l'estimateur de l'indice de queue γ , ainsi que le paramètre de distorsion ρ . Etant donné que le nombre de valeur extrêmes est lié directement à la propriété de la normalité asymptotique de l'estimateur de l'indice de queue à la base utilisé dans la construction de l'estimateur de mesure de risque. Nous avons généré des échantillons de taille 500 et 1000 issues d'une distribution de Pareto et nous avons calculé le nombre des valeurs extrêmes dans les deux cas des deux estimateurs d'indice de queue correspondants ainsi que les valeurs des deux estimateurs de mesure de risque $\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$.

Sélection du nombre des valeurs extrêmes k_1 correspondant à $\hat{\gamma}_n^h$ concernant l'estimateur $\Psi \left[\tilde{Q}_n, g \right]$.

Reiss et Thomas (1997) ont proposé une méthode heuristique pour choisir le nombre des plus grandes valeurs à utiliser dans l'estimation de l'indice de queue. Cette méthodologie, incorporée dans le package "Xtremes", sélectionne k de manière automatique adoptant comme optimale la valeur k^* qui minimise

$$\frac{1}{k} \sum_{i \leq k} i^\beta |\hat{\gamma}_{i,n} - \text{med}(\hat{\gamma}_{1,n}, \dots, \hat{\gamma}_{k,n})|, 0 \leq \beta \leq \frac{1}{2} \quad (6.1)$$

Pour l'estimateur de Hill. Si la seule information existante sur l'indice de queue, c'est que $\gamma > 0$ ou s'il est raisonnable de s'assurer que $\gamma > 1$. Donc on prend $\beta = 0$ dans (6.1). Alors que pour $0 < \gamma \leq 1$, on choisi $\beta = 0.3$ dans (6.1).

Calcul de k_1 et $\hat{\gamma}_n^h$ sous R . Soit un échantillon d'observations *iid*, X_1, \dots, X_n . Dans un premier temps nous allons calculer toutes les valeurs de l'estimateur de Hill en fonction des valeurs $k < n$ prises entre $[n/3]$ et $[n - (n/3)]$, ensuite on cherche à minimiser la quantité (6.1) par rapport à k dans le but de trouver k_1 . Le programme suivant sous R calcul directement la valeur de k_1 et donne aussi la valeur de l'estimateur $\hat{\gamma}_n^h$ correspondante.

```
y=sort(x) # on ordonne l'échantillon des observations  $X_i$ 
z=log10(y)
n=? # taille de l'échantillon  $n$ 
```

```

b1=floor(n/3) # valeur inférieure prise par k
b2=floor(n-(n/3)) # valeur supérieure prise par k
hills<-seq(1,b2) # valeurs de  $\hat{\gamma}_n^h$  pour différentes valeur de k
for(k in 1 :b2)
{
l<-seq(1,k)
hills[k]<-((1/k)*sum(z[n-l+1]))-z[n-k]
}
x2<-seq(1,b2)
f0 <- fonction (x) (1/x)*sum((x2^0.3)*abs(hills[x2]-median(hills[1 :x])))
xmin0 <-optimize(f0, c(3, b1), tol = 0.0001)
xmin0 # names(xmin) pour savoir les nom du tableau
K*=floor(xmin0$minimum) # nombre de valeurs extrêmes  $k_1$ 
hill<-hills[K*] # la valeur de l'estimateur de Hill  $\hat{\gamma}_n^h$  en  $k_1$ 

```

Sélection de nombre des valeurs extrêmes k_2 correspondant à $\hat{\gamma}_k$ concernant $\Psi [\hat{Q}_n, g]$

D'après Matthys et Beirlant (2003) (voir la page 876, 878). Pour $\gamma > 0$, le seuil optimal pour $\hat{\gamma}_k$ peut être trouver en minimisant l'erreur quadratique moyenne asymptotique par rapport à k

$$EQMA = \frac{1}{k} \frac{\sigma_\gamma^2}{a_\gamma^2} + b^2 (n/k) \frac{I_{\gamma,\rho}^2}{a_\gamma^2}.$$

Cette quantité est proportionnelle à

$$Q = k^{-1} \sigma_\gamma^2 + b^2 (n/k) I_{\gamma,\rho}^2,$$

où σ_γ^2 est la variance de $K_\gamma(U)$ avec U uniformément distribuée sur (0,1), avec

$$K_\gamma(U) = (\log u) / \gamma + (1 + \gamma) (di \log u^\gamma) / \gamma^2$$

où

$$di \log u = \int_1^u \log t / (1 - t) dt \quad (u \geq 0)$$

et

$$I_{\gamma,\rho} = \rho^{-1} \int_0^1 (J(u) / (1 - u^\gamma)) (1 - u^{-\rho}) du$$

où

$$J(u) = \gamma^{-2} (u^\gamma - 1 - \gamma(1 + \gamma) u^\gamma \log u), \quad 0 < u < 1$$

où $b_{n,k+1}$ est synonyme de $b(U_{k+1,n}^{-1})$ qui est une fonction de vitesse satisfaisant l'hypothèse (R_ℓ) . Pour $0 < u < 1$, on a :

$$Q = k^{-1} \text{var} \left[(\log u) / \gamma + (1 + \gamma) \left(\int_1^{u^\gamma} \log t / (1 - t) dt \right) / \gamma^2 \right] \\ + b^2(n/k) \rho^{-1} \int_0^1 \left(\frac{\gamma^{-2} (u^\gamma - 1 - \gamma(1 + \gamma) u^\gamma \log u)}{1 - u^\gamma} \right) (1 - u^{-\rho}) du$$

On cherche à minimiser l'estimateur de la quantité Q par rapport à k . Ceci en remplaçant $b(n/k)$, γ , et ρ par leur estimateurs $\hat{b}_{n,k+1}^B$, $\hat{\gamma}_{k+1}^B$, $\hat{\rho}_{k+1}^B$. Ces derniers sont calculés numériquement par la méthode du maximum du vraisemblance à partir du modèle (3.11). Où la fonction log-vraisemblance $\mathcal{L}(b_{n,k+1}, \gamma, \rho; \mathbf{Y})$ est donnée par :

$$\sum_{j=1}^{k-1} \left\{ \log \left(\frac{1 - \left(\frac{j}{k+1}\right)^\gamma \exp \left(b_{n,k+1} \frac{\left(\frac{j}{k+1}\right)^{-\rho} - 1}{-\rho} \right)}{\gamma + b_{n,k+1} \left(\frac{j}{k+1}\right)^{-\rho}} \right) - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma \exp \left(b_{n,k+1} \frac{\left(\frac{j}{k+1}\right)^{-\rho} - 1}{-\rho} \right)}{\gamma + b_{n,k+1} \left(\frac{j}{k+1}\right)^{-\rho}} Y_j \right\}$$

Pour des raisons de commodités on pose $\rho = -1$, $\mathcal{L}(b_{n,k+1}, \gamma, \rho; \mathbf{Y})$: devient

$$\sum_{j=1}^{k-1} \left\{ \log \left(\frac{1 - \left(\frac{j}{k+1}\right)^\gamma \exp b_{n,k+1} \left(\frac{j-k-1}{k+1}\right)}{\gamma + b_{n,k+1} \left(\frac{j}{k+1}\right)} \right) - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma \exp \left(b_{n,k+1} \frac{j-k-1}{k+1} \right)}{\gamma + b_{n,k+1} \left(\frac{j}{k+1}\right)} Y_j \right\} \quad (6.2)$$

D'après Matthys et Beirlant (2003), pour des raisons de commodités on prend $\hat{\rho}_{k+1}^B = -1$, l'estimateur de Q est le suivant :

$$\hat{Q} = \frac{1}{k} \text{var} \left(\frac{(\log u)}{\hat{\gamma}_{k+1}^B} + \frac{(1 + \hat{\gamma}_{k+1}^B)}{(\hat{\gamma}_{k+1}^B)^2} \left(\int_1^{u^{\hat{\gamma}_{k+1}^B}} \frac{\log t}{1 - t} dt \right) \right) + \hat{b}_{n,k+1}^2 \times \\ \left(\int_0^1 \left(\frac{(\hat{\gamma}_{k+1}^B)^{-2} \left(u^{\hat{\gamma}_{k+1}^B} - 1 - \hat{\gamma}_{k+1}^B (1 + \hat{\gamma}_{k+1}^B) u^{\hat{\gamma}_{k+1}^B} \log u \right)}{1 - u^{\hat{\gamma}_{k+1}^B}} \right) \times (1 - u) du \right)^2$$

Pour faciliter le calcul, on met

$$I_2 = \int_0^1 \left(\frac{(\hat{\gamma}_{k+1}^B)^{-2} \left(u^{\hat{\gamma}_{k+1}^B} - 1 - \hat{\gamma}_{k+1}^B (1 + \hat{\gamma}_{k+1}^B) u^{\hat{\gamma}_{k+1}^B} \log u \right)}{1 - u^{\hat{\gamma}_{k+1}^B}} \right) \times (1 - u) du$$

et

$$I_1 = \int_1^{u^{\hat{\gamma}_{k+1}^B}} \frac{\log t}{1-t} dt.$$

Pour calculer I_1 faisant un changement de variable $t = 1 - x$, alors I_1 devient

$$I_1 = \int_1^{1-u^{\hat{\gamma}_{k+1}^B}} \frac{\log(1-x)}{x} dx$$

Cette fonction est appelée fonction di-logarithmique d'Euler. Nous avons :

$$\left| 1 - u^{\hat{\gamma}_{k+1}^B} \right| = \sqrt{\left(1 - u^{\hat{\gamma}_{k+1}^B} \right)^2} = 1 - u^{\hat{\gamma}_{k+1}^B},$$

et d'autre part on a

$$u \geq 0 \implies 1 - u^{\hat{\gamma}_{k+1}^B} < 1 \implies \left| 1 - u^{\hat{\gamma}_{k+1}^B} \right| < 1.$$

Comme $\left| 1 - u^{\hat{\gamma}_{k+1}^B} \right| < 1$, nous avons :

$$I_1 = \sum_{n=1}^{\infty} \frac{\left(1 - u^{\hat{\gamma}_{k+1}^B} \right)^n}{n^2}.$$

Nous remarquons que I_1 est une série numérique restant à démontrer sa convergence. Si on met

$$v_n = \frac{\left(1 - u^{\hat{\gamma}_{k+1}^B} \right)^n}{n^2},$$

alors,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{v_{n+1}}{v_n} &= \lim_{n \rightarrow \infty} \frac{\left(1 - u^{\hat{\gamma}_{k+1}^B} \right)^{n+1} n^2}{\left(1 - u^{\hat{\gamma}_{k+1}^B} \right)^n (n+1)^2} \\ &= \lim_{n \rightarrow \infty} \frac{\left(1 - u^{\hat{\gamma}_{k+1}^B} \right)^{n+1} n^2}{n^2 + n + 1} \\ &= 1 - u^{\hat{\gamma}_{k+1}^B} < 1. \end{aligned}$$

Donc la série I_1 est convergente, sa somme partielle $S_n = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{\left(1 - u^{\hat{\gamma}_{k+1}^B} \right)^i}{i^2}$ sera calculée numériquement sous R . \hat{Q} devient :

$$\hat{Q} = \frac{1}{k} \text{var} \left(\frac{(\log u)}{\hat{\gamma}_{k+1}^B} + \frac{(1+\hat{\gamma}_{k+1}^B)}{(\hat{\gamma}_{k+1}^B)^2} \times \sum_{n=1}^{\infty} \frac{(1-u^{\hat{\gamma}_{k+1}^B})^n}{n^2} \right) + \hat{b}_{n,k+1}^2 \times I_2^2 \quad (6.3)$$

En minimisant numériquement \hat{Q} par rapport à k sous l'environnement R , nous obtenons le nombre de valeurs extrêmes recherché que l'on note par la suite k_2 .

Calcul de k_2 et $\hat{\gamma}_k$ sous R . Soit un échantillon d'observations *iid*, X_1, \dots, X_n . Nous expliquerons dans ce qui suit l'algorithme à adopter pour déterminer le nombre des valeurs extrêmes k_2 . Dans le cas de

$$b(r) = Cr^\rho (1 + o(1)) (r \rightarrow \infty)$$

pour certains $c > 0, np_n \rightarrow c$ (fini, ≥ 0) et $\log p_n = o(n^\epsilon)$ pour $\epsilon > 0 (n \rightarrow \infty)$, on prends $b(r) = 1/r$. Nous avons,

$$b_{n,k+1} = b(U_{k+1,n}^{-1}) = 1/U_{k+1,n}^{-1} = 1 - F(X_{n-k,n})$$

1. Pour une distribution de Pareto d'indice $\gamma > 0$, $1 - F(x) = x^{-1/\gamma}$. Nous avons

$$b_{n,k+1} = (X_{n-k,n})^{1/\gamma} \implies \hat{b}_{n,k+1} = (X_{n-k,n})^{1/\hat{\gamma}_{k+1}^B}$$

2. Pour chaque $k < n$, on doit calculer les estimateurs $\hat{b}_{n,k+1}^B, \hat{\gamma}_{k+1}^B, \hat{\rho}_{k+1}^B$, en maximisant la fonction log-vraisemblance (6.2). En remplaçant $b_{n,k+1}$ par $(X_{n-k,n})^{1/\gamma}$. La fonction log-vraisemblance (6.2) devient,

$$\sum_{j=1}^{k-1} \log \left\{ \left(\frac{1 - \left(\frac{j}{k+1}\right)^\gamma \exp(X_{n-k,n})^{1/\gamma} \left(\frac{j-k-1}{k+1}\right)}{\gamma + (X_{n-k,n})^{1/\gamma} \left(\frac{j}{k+1}\right)} \right) - \frac{1 - \left(\frac{j}{k+1}\right)^\gamma \exp\left((X_{n-k,n})^{1/\gamma} \frac{j-k-1}{k+1}\right)}{\gamma + (X_{n-k,n})^{1/\gamma} \left(\frac{j}{k+1}\right)} Y_j \right\}. \quad (6.4)$$

Dans ce cas la maximisation se fait uniquement par rapport à γ pour trouver $\hat{\gamma}_{k+1}^B$.

3. En substituant la valeur de $\hat{b}_{n,k+1}$ par $(X_{n-k,n})^{1/\hat{\gamma}_{k+1}^B}$ dans (6.3), \hat{Q} devient

$$\hat{Q} = \frac{1}{k} \text{var} \left(\frac{(\log u)}{\hat{\gamma}_{k+1}^B} + \frac{(1+\hat{\gamma}_{k+1}^B)}{(\hat{\gamma}_{k+1}^B)^2} \times \sum_{n=1}^{\infty} \frac{(1-u^{\hat{\gamma}_{k+1}^B})^n}{n^2} \right) + \left((X_{n-k,n})^{1/\hat{\gamma}_{k+1}^B} \right)^2 \times I_2^2 \quad (6.5)$$

4. Pour chaque $k < n$, nous avons obtenue une valeur $\hat{\gamma}_{k+1}^B$. En minimisant la quantité \hat{Q} dans (6.5) par rapport à k , nous pourrions sélectionner le nombre des valeurs extrêmes recherchée que l'on note par la suite k_2 .
5. En substituant la valeur de k_2 dans (3.6), la log-vraisemblance devient

$$\mathcal{L}(\gamma; \mathbf{Y}) := \sum_{j=1}^{k_2-1} \left\{ \log \left(\frac{1 - \left(\frac{j}{k_2+1}\right)^\gamma}{\gamma} \right) - \frac{1 - \left(\frac{j}{k_2+1}\right)^\gamma}{\gamma} Y_j \right\} \quad (6.6)$$

L'estimateur $\hat{\gamma}_k$ peut être calculé en maximisant la log-rvraisemblance (6.6). Cette log-rvraisemblance est en fonction de γ . Pour se faire, on choisit une valeur initiale de γ par une utilisation au préalable des outils graphiques d'exploration des données de notre échantillon X_1, \dots, X_n . Ensuite on va chercher la valeur de l'estimateur $\hat{\gamma}_k$ en maximisant (6.6) en balayant les valeurs g telle que $0 < g < 2\gamma$.

Le programme sous R est donnée par :

```

y<-sort(x) # on ordonne l'échantillon des observations Xi
gamma=0.6 # à titre d'exemple faut avoir une valeur initial de γ
b1=floor(n/3)
b2=floor(n-(n/3))
integrand <- fonction(u) {(((u^gamma)-(u^(gamma+1)))/(1-(u^gamma)))
*log10(u)}
b=n-1
Int=1 :b2
Q<-seq(1,b2)
VAR<-seq(1,b2)
gh<-seq(1,b2)
for(k in 2 :b2)
{
g<-seq(1,100)
LV<-seq(1,100)
x3<-seq(1,k-1)
x5<-seq(1,100)
g[x5]<-x5*((2*gamma)/100)
f1 <- fonction (x) -sum(log10 # la log-vraisemblance (6.4)
(((1-((x3/(k+1))^g[x]))*(exp(((y[n-k])^(-1/g[x]))*
((x3-k-1)/(k+1)))))/(g[x]+(((y[n-k])^(-1/g[x]))
*(x3/(k+1))))))-(((1-((x3/(k+1))^g[x]))*(exp(((y[n-k])^(-1/g[x]))
*((x3-k-1)/(k+1)))))/(g[x]+(((y[n-k])^(-1/g[x]))*(x3/(k+1))))))*

```

```

(x3*log10((y[n-x3+1]-y[n-k])/(y[n-x3]-y[n-k])))
xmin1 <-optimize(f1, c(1, 100), tol = 0.0001)
F=floor(xmin1$minimum)
gh[k]<-g[F]
gama=gh[k]
Int0=integrate(integrand, lower=0, upper=1)
Int[k]=Int0$val
Valeurs<-seq(1,n)
x4<-seq(1,1000)
x1<-seq(1,n)
SN<-seq(1,n)
for(m in 1 :n)
{
SN[m]<-sum(((1-u[m]^gh[k])^x4)/(x4^2))
}
Valeurs[x1]<-(log10(u[x1])/gh[k])+(((1+gh[k])/(gh[k]^2))*SN[x1])
Valeurs
VAR[k]<-var(Valeurs)
}
x9<-seq(2,b2)
Q[x9]<-((1/x9)*VAR[x9])+((((y[n-x9])^(-1/gh[x9]))^2)*((1/(2*gh[x9]^2))
+(((1+gh[x9])/gh[x9])*Int[x9]))^2) # les valeurs de  $\hat{Q}$  dans (6.5)
Km<-which.min(Q[b1 :b2])+(b1-1) # la valeur du nombre  $k_2$ 
g<-seq(1,100)
Lv<-seq(1,100)
x6<-seq(1,Km-1)
x5<-seq(1,100)
g[x5]<-x5*((2*gamma)/100)
f2 <- fonction (x) (-(sum((log10((1- # la log-vraisemblance (6.6)
((x6/(Km+1))^g[x])/g[x]))-(((1-((x6/(Km+1))^g[x])/g[x])
*((x6)*log10((y[n-x6+1]-y[n-Km])/(y[n-x6]-y[n-Km]))))))))
xmin2 <-optimize(f2, c(1, 100), tol = 0.0001)
H=floor(xmin2$minimum)
mattys<-g[H] # la valeur de l'estimateur  $\hat{\gamma}_k$ 

```

6.1.2 Paramètre de distorsion

Notez que la prime développée est particulièrement sensible au choix de ρ ($\rho \geq 1$), ainsi on devrait être prudent dans sa sélection. Par conséquent, il est important de mettre en place un outil graphique pour le réglage de

ce paramètre. Brahim et al. (2011) ont mis l'accent sur le paramètre de distorsion ρ en présentant son estimation ainsi qu'une étude asymptotique de sa distribution limite.

Nous générons dans cette section des échantillons de différentes tailles de distribution de Pareto avec $\gamma = 0.6$ (figures 6.1, 6.2, 6.3 et 6.4). Les estimateurs sont bons pour les petites valeurs de ρ mais quand $\rho > 1.5$, ils s'écartent de la valeur réelle.

Globalement, pour $\rho \in [1, 1.5]$, $\Psi[\hat{Q}_n, g]$ est meilleure que $\Psi[\tilde{Q}_n, g]$. On a aussi remarqué que la taille de l'échantillon a peu d'influence sur notre estimation à ce niveau d'études.

Cet outil graphique peut servir, comme un outil d'aide à la décision dans une application réelle et nous permet d'avoir une première idée sur l'estimateur par rapport au paramètre de distorsion ρ , qui sert également à déterminer la vraie valeur appropriée de la prime avant l'estimation, mais par premièrement la fixation d'une part, de l'indice de queue selon la distribution de l'échantillon des données.

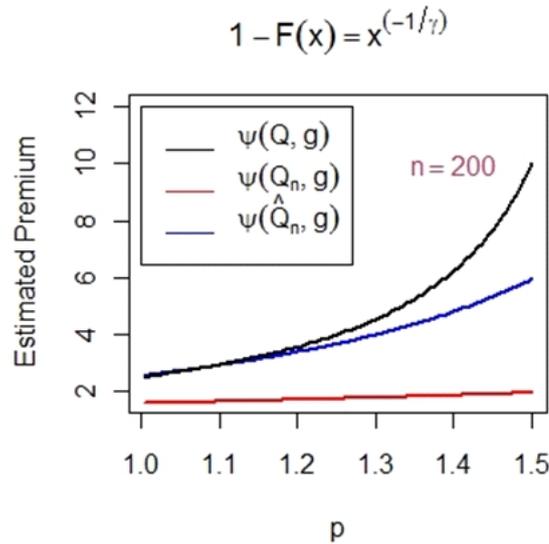


FIG. 6.1 – $\Psi[\tilde{Q}_n, g]$ et $\Psi[\hat{Q}_n, g]$ vs le paramètre de distorsion ρ d'une distribution de Pareto avec $\gamma = 0.6$ et $n = 200$

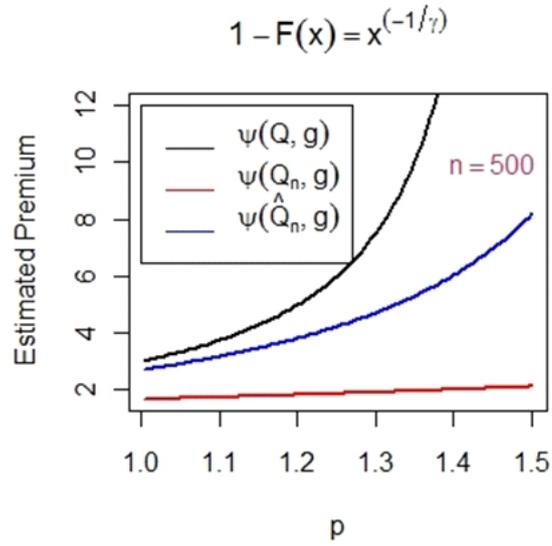


FIG. 6.2 – $\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ d’une distribution de Pareto avec $\gamma = 0.6$ et $n = 500$

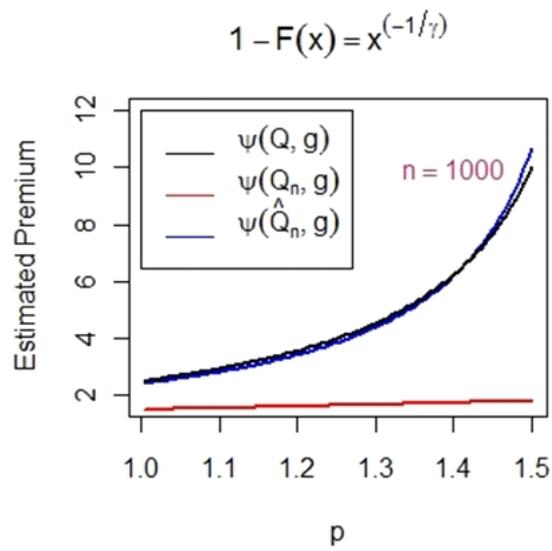


FIG. 6.3 – $\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ d’une distribution de Pareto avec $\gamma = 0.6$ et $n = 1000$

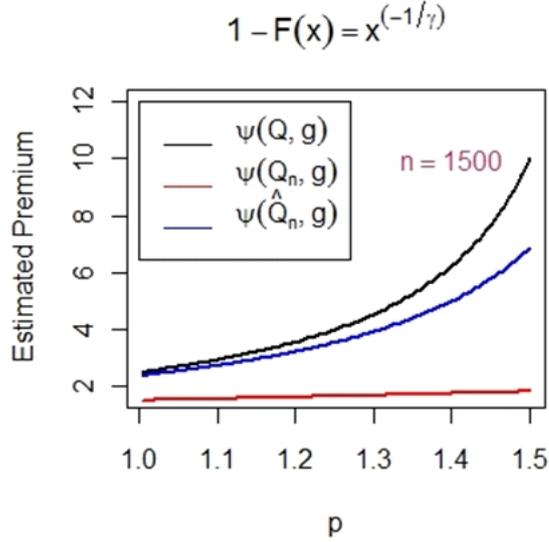


FIG. 6.4 – $\Psi \left[\tilde{Q}_n, g \right]$ et $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ d’une distribution de Pareto avec $\gamma = 0.6$ et $n = 1500$

Ré-échantillonnage. La méthode du bootstrap introduite par Efron (1979) est une procédure de ré-échantillonnage très générale pour estimer les distributions des statistiques basées sur des observations indépendantes. Elle a été largement appliquée par les financiers et les actuaires. Pour un traitement complet, voir le manuel standard d’Efron et Tibshirani (1993) et de Davison et Hinkley (1997). L’idée de base du bootstrap est de créer des échantillons de ré-échantillonnage (avec remplacement) à partir de l’échantillon initial. La procédure de base du bootstrap peut être expliquée comme suit : Supposons que nous avons un échantillon aléatoire *iid*, $\mathbf{X} = (X_1, X_2, \dots, X_n)$ issue d’une fonction de distribution F inconnue, et nous nous intéressons à une statistique $\Psi \left[\hat{Q}_n, g \right]$ que nous noterons $\Psi_g \left(\hat{Q}_n \right)$ pour faciliter l’écriture. Les pseudo-échantillons sont générés par échantillonnage, avec remplacement. L’échantillon généré, notée \mathbf{X}^* , qu’on appelle un échantillon bootstrap est indiqué par l’exposant *. La statistique d’intérêt, en utilisant cet échantillon généré est ensuite notée $\Psi_g^* = \Psi \left(\hat{Q}_n \left(\mathbf{X}^* \right) \right)$. Nous répétons l’exercice R fois pour différents R échantillons de bootstrap $\mathbf{X}_1^*, \dots, \mathbf{X}_R^*$, chacun de taille n . A partir de chaque échantillon, nous générons la statistique d’intérêt, c’est- $\Psi_{g,j}^*$

du j -emme échantillon du bootstrap, ce qui donne

$$\Psi_{g,j}^* = \Psi \left(\hat{Q}_n(\mathbf{X}_j^*) \right), j = 1, \dots, R$$

Enfin, l'estimation bootstrap de la statistique $\Psi_g \left(\hat{Q}_n \right)$ est donnée par

$$(1/R) \sum_{j=1}^R \Psi_{g,j}^*.$$

La distribution empirique des valeurs résultantes $\Psi_{g,1}^*, \Psi_{g,2}^*, \dots, \Psi_{g,R}^*$ est une approximation de la fonction de distribution de $\Psi_g \left(\hat{Q}_n \right)$. Cette approximation est bonne si R est très grand. Lorsque nous ordonnons l'échantillon des estimateurs $\Psi_{g,1}^*, \Psi_{g,2}^*, \dots, \Psi_{g,R}^*$, les $(1 - 2\alpha)$ - intervalles de confiance de $\Psi \left[\hat{Q}_n, g \right]$ seront

$$\left(\Psi_{g,R(\alpha/2)}^*, \Psi_{g,R(1-\alpha/2)}^* \right)$$

(prenant $\alpha = 0.05$).

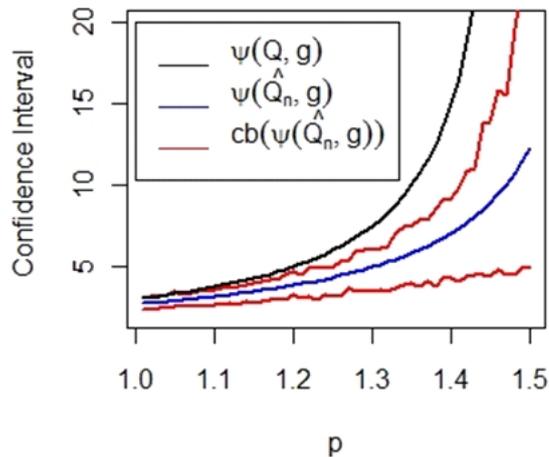


FIG. 6.5 – Intervalles de confiance empiriques de $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ pour 100 répétitions de distribution de Pareto, $n = 500, \gamma = 2/3$

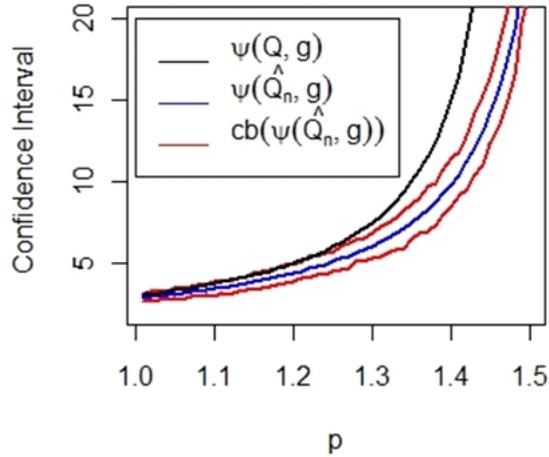


FIG. 6.6 – Intervalles de confiance empiriques de $\Psi \left[\hat{Q}_n, g \right]$ vs le paramètre de distorsion ρ pour 100 répétitions de distribution de Pareto, $n = 1000$, $\gamma = 2/3$

Sur les figures 4.3 et 4.4, nous avons tracé versus le paramètre de distorsion ρ les intervalles de confiance empiriques de notre nouvel estimateur $\Psi \left[\hat{Q}_n, g \right]$ calculées par bootstarp à partir de 100 répétitions d'un échantillon issu d'une distribution de Pareto d'indice de queue $\gamma = 2/3$ (comme étant un échantillon issue d'une distribution inconnue dans les applications réelles) où nous avons remarqué que l'estimateur est bien encadré par ses bornes de confiance pour tout valeur de ρ .

Plus ρ s'approche de 1 on remarque que les bornes de confiance ainsi que l'estimateur $\Psi \left[\hat{Q}_n, g \right]$ et la valeur théorique $\Psi [Q, g]$ se rapproche l'un à l'autre

6.1.3 Paramètres de comparaison

Nous avons simulé 200 échantillons issues d'une distribution de Pareto d'indice de queue $\gamma = 2/3$ et de paramètre de distorsion $\rho = 1.1$, pour différentes valeurs de n (200, 500, 1000) (comme étant un échantillon issue d'une distribution inconnue dans les applications réelles).

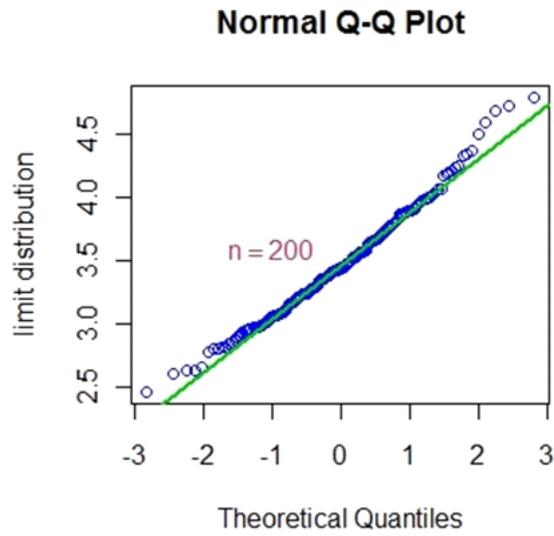


FIG. 6.7 – QQ-norm de la distribution limite de $\Psi\left(g, \hat{Q}_n\right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 200$

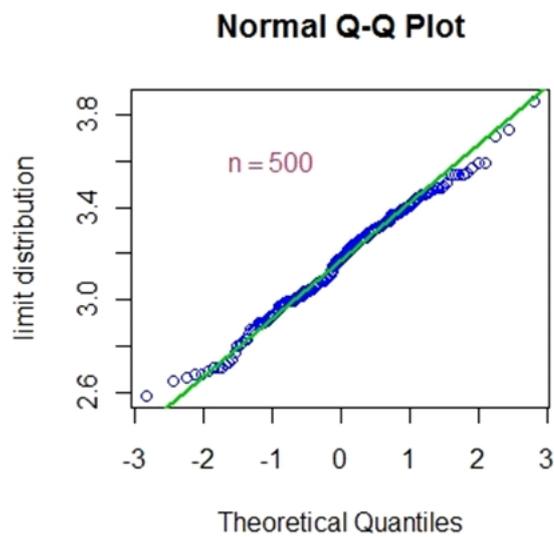


FIG. 6.8 – QQ-norm de la distribution limite de $\Psi\left(g, \hat{Q}_n\right)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 500$

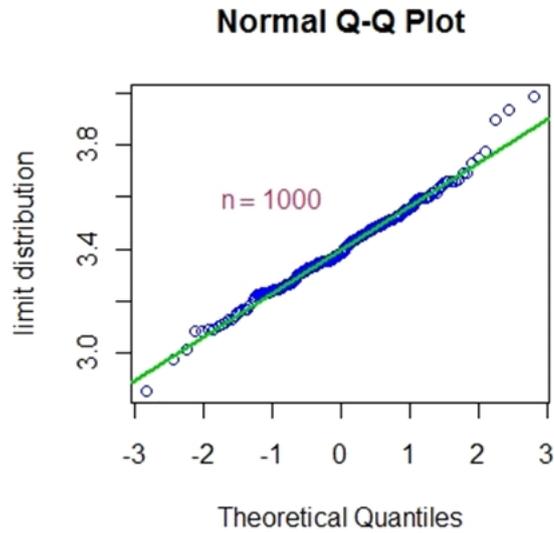


FIG. 6.9 – QQ-norm de la distribution limite de $\Psi(g, \hat{Q}_n)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 1000$

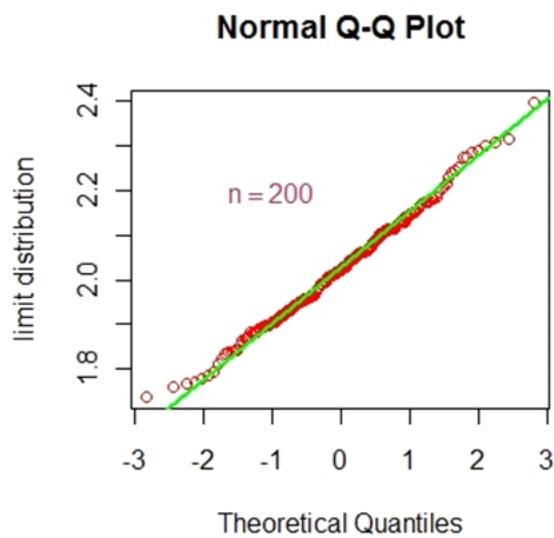


FIG. 6.10 – QQ-norm de la distribution limite de $\Psi(g, \tilde{Q}_n)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 200$

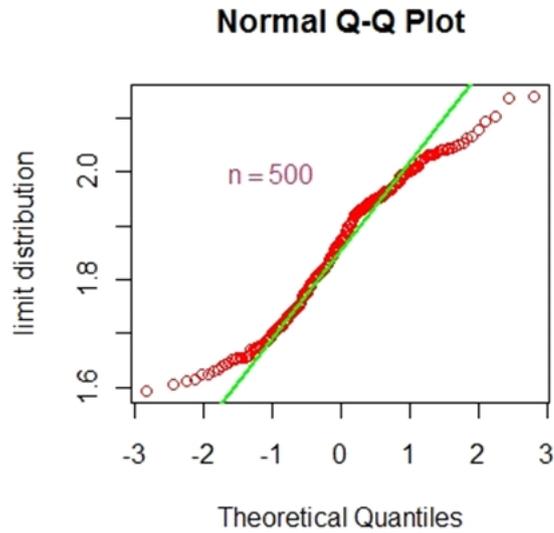


FIG. 6.11 – QQ-norm de la distribution limite de $\Psi(g, \tilde{Q}_n)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 500$

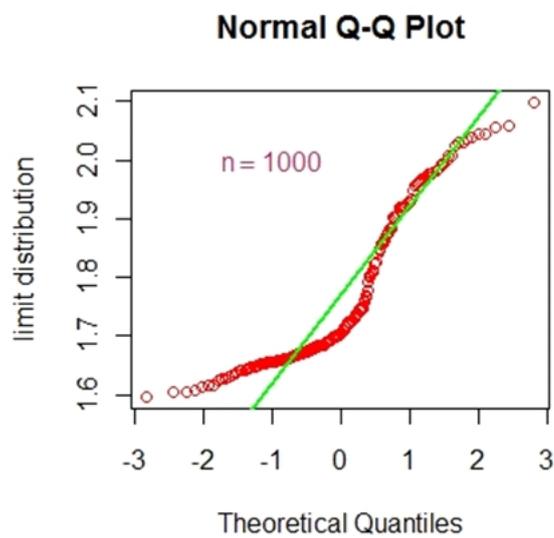


FIG. 6.12 – QQ-norm de la distribution limite de $\Psi(g, \tilde{Q}_n)$ de 200 échantillon de Pareto, $\gamma = 2/3, n = 1000$

Sur les figures 6.5, 6.6 et 6.7 les points du QQ -norm sont parfaitement alignés, ce qui signifie que cette distribution limite empirique de notre nouvel estimateur de mesure de risque $\Psi(g, \hat{Q}_n)$ est issue d'une loi normale.

Les figures 6.8, 6.9 et 6.10 représentent la distribution limite empirique de l'estimateur $\Psi(g, \tilde{Q}_n)$ de Necir et Meraghni (2009) pour $n = 200, 500, 1000$ respectivement où nous remarquons que plus n est grand, les points sont de plus en plus non alignés, ce qui signifie empiriquement que la normalité asymptotique n'est pas vraiment tenue sans passer par la vérification des conditions théoriques mentionnées dans le théorème 4.1.

Indicateurs d'écart

Dans ce qui suit nous nous intéressons à faire une simulation comparative entre les deux estimateurs $\Psi[\tilde{Q}_n, g]$ et $\Psi[\hat{Q}_n, g]$ selon certains indicateurs d'écart comme l'erreur absolue (AE), l'erreur standard (SE), l'erreur quadratique moyenne (MSE) et le biais.

L'écart-type où erreur standard est donné par

$$\sigma(\Psi_g(\hat{Q}_n)) = \sqrt{\frac{1}{R-1} \sum_{b=1}^R \left(\Psi_{g,b}^* - \frac{1}{R} \sum_{b=1}^R \Psi_{g,b}^* \right)^2}.$$

Le calcul de biais empirique est donnée par :

$$Bias(\Psi_g(\hat{Q}_n)) = \frac{1}{R} \sum_{b=1}^R \Psi_{g,b}^* - \Psi_g(\hat{Q}_n)$$

k_1	61	k_2	66
$\hat{\gamma}_n^h$	0.2621680	$\hat{\gamma}_k$	0.6403333
$\Psi[\tilde{Q}_n, g]$	2.0260260	$\Psi[\hat{Q}_n, g]$	3.4775150
$AE(\Psi[\tilde{Q}_n, g])$	1.7239740	$AE(\Psi[\hat{Q}_n, g])$	0.2724850
$SE(\Psi[\tilde{Q}_n, g])$	0.1219114	$SE(\Psi[\hat{Q}_n, g])$	0.4278235
$RMSE(\Psi[\tilde{Q}_n, g])$	0.3825634	$RMSE(\Psi[\hat{Q}_n, g])$	0.4498007
$Bias(\Psi[\tilde{Q}_n, g])$	0.3626189	$Bias(\Psi[\hat{Q}_n, g])$	0.1388040

TAB. 6.1 – 200 échantillons générés de distribution de Pareto de taille $n=200$

k_1	238	k_2	168
$\hat{\gamma}_n^h$	0.2609875	$\hat{\gamma}_k$	0.5956667
$\Psi[\tilde{Q}_n, g]$	1.8585050	$\Psi[\hat{Q}_n, g]$	3.1658920
$AE(\Psi[\tilde{Q}_n, g])$	1.8914950	$AE(\Psi[\hat{Q}_n, g])$	0.5841080
$SE(\Psi[\tilde{Q}_n, g])$	0.1335753	$SE(\Psi[\hat{Q}_n, g])$	0.2382059
$RMSE(\Psi[\tilde{Q}_n, g])$	0.1725476	$RMSE(\Psi[\hat{Q}_n, g])$	0.2382532
$Bias(\Psi[\tilde{Q}_n, g])$	0.1092261	$Bias(\Psi[\hat{Q}_n, g])$	0.0047487

TAB. 6.2 – 200 échantillons générés de distribution de Pareto de taille $n=500$

k_1	557	k_2	333
$\hat{\gamma}_n^h$	0.2773872	$\hat{\gamma}_k$	0.6266667
$\Psi[\tilde{Q}_n, g]$	1.7657810	$\Psi[\hat{Q}_n, g]$	3.3985000
$AE(\Psi[\tilde{Q}_n, g])$	1.9842200	$AE(\Psi[\hat{Q}_n, g])$	0.3515000
$SE(\Psi[\tilde{Q}_n, g])$	0.12935680	$SE(\Psi[\hat{Q}_n, g])$	0.1712464
$RMSE(\Psi[\tilde{Q}_n, g])$	0.1622788	$RMSE(\Psi[\hat{Q}_n, g])$	0.1713426
$Bias(\Psi[\tilde{Q}_n, g])$	0.0979859	$Bias(\Psi[\hat{Q}_n, g])$	0.0057426

TAB. 6.3 – 200 échantillons générés de distribution de Pareto de taille $n=1000$

D'après nos simulations ainsi présentées sur les tableaux 6.1, 6.2 et 6.3 où le calcul par la méthode de ré-échantillonnage où 200 échantillons sont générés d'une distribution de Pareto de taille $n = (200, 500, 1000)$ avec $\gamma = 2/3$ et $\rho = 1.1$, nous avons constaté que dans la plus part des cas en variant la taille de l'échantillon, l'estimateur de l'indice de queue de Mattyhs et Beirlant (2003) $\hat{\gamma}_k$ présente des valeurs plus proches de la valeur théorique γ que l'estimateur de Hill $\hat{\gamma}_n^h$, notamment notre nouvel estimateur de mesure de risque $\Psi[\hat{Q}_n, g]$ est meilleur que celui de Necir et Meraghni (2009) $\Psi[\tilde{Q}_n, g]$ en terme d'erreur absolue. Concernant l'erreur standard et l'erreur quadratique moyenne, on a remarqué une légère différence sans signification, par contre notre nouvel estimateur $\Psi[\hat{Q}_n, g]$ présente un biais réduit par rapport à l'estimateur $\Psi[\tilde{Q}_n, g]$.

6.2 Exploration des observations réelles extrêmes

Au cours des dernières années, les actuaires sont devenus de plus en plus conscients du risque potentiel inhérent à de très grandes tailles de sinistres dus à des événements catastrophiques. Un assureur doit compenser les pertes dues

aux paiements pour les revendications des détenteurs de polices par le biais d'une prime appropriée. L'importance du travail est donnée sur la génération d'un modèle qui prend en compte le fait qu'une partie des revendications des échantillons sont considérés comme des valeurs extrêmes avec une probabilité plus ou moins petite. Nous effectuons dans cet section une mise en œuvre réelle des données sous la base des données de *Society of actuaries (SOA)*, où la motivation de cette étude était d'explorer et d'analyser les facteurs qui influent sur la queue de distribution des sinistres pour pouvoir proposer un modèle approprié qui mène dans la suite à engendrer des primes commerciaux les plus adéquatement possible qui répondent au attentes des gestionnaires de risque.

6.2.1 Interprétation des données

La base de données contient des revendications de trois ans : 1997, 1998 et 1999 et contient tous les demandeurs. Les données de sept assureurs sont représentées, la base de données contient 4.294.030 demandeurs et les frais totaux payés d'environ 7.1 milliards de dollars. Les données de chaque année sont présentées séparément pour fournir des tailles de fichiers plus gérables. Bien sûr, l'utilisateur peut ajouter des fichiers à travailler avec les données combinées. Chaque enregistrement contient un champ pour l'année d'indemnisation pour permettre l'identification de l'année si les fichiers sont combinés. Compte de contrôle et les totaux pour chaque fichier sont présentés sur ce tableaux (les données sont disponible à <http://www.soa.org>).

	demandeurs	total des charges payées
réclamations 1997	1.241.438	2.003.162.217.76\$
réclamations 1998	1.460.854	2.466.093.740.87\$
réclamations 1999	1.591.738	2.599.356.657.86\$

TAB. 6.4 – Compte de contrôle et totaux

Dans notre cas, nous considérons que les réclamations représentent l'échantillon de risque à étudier loin de passer par des détails comme le sexe ou l'âge des demandeurs ou même d'autre détails fournis par la base de données en questions. Pour cette raison précisément, à partir des tableaux prise en compte pour notre application et dans l'absence du montant y étant réclamé par chaque demandeur nous avons extraire des tableaux bien déterminés qui

répondent à nos besoins d'étude à effectuer où nous avons considéré la taille de sinistre et la fréquence d'avoir été réclamé par les demandeurs en essayant d'uniformiser ces données présentés à la base sous forme de classe par rapport au taille de sinistre qui regroupe un nombre de demandeur, ainsi la totalité de prix réclamé par ces derniers selon la classe de prix correspondante.

6.2.2 QQ-plot exponentielle des données

Procédure sous R. Ce programme permet de tracer le QQ-plot.

```

dataIn=read.table("E :\\DOCTORAT\\
these-meddi\\1997.txt",header=T,dec=".")
freq=dataIn[,1]
valeur=dataIn[,2]
nc=length(freq)
valeurs<-seq(1,nc)
x0<-seq(1,nc)
valeurs[x0]<-valeur[x0]/freq[x0]
freq
valeurs
repFreq=function(freq,valeurs)
{output=numeric(0)
i=1
while (i <=length(freq)){temp=numeric(0)
temp=rep(valeurs[i],freq[i])
output=c(output,temp)
i = i + 1}
return(output)}
monvect=repFreq(freq,valeurs)
n=length(monvect)
yr<-sort(monvect)
p<-seq(1,n)
for(i in 1 :n){
p[i]<-i/(n+1)}
Droite<-lm(-log10(1-p)~yr)
coef(Droite)
coef(Droite)[1]
coef(Droite)[2]
qqplot(-log10(1-p),yr, xlab="Theoretical Qunatiles",
ylab="Sample Quantiles", main="Exponential Q-Q plot")
abline(coef(Droite), col="red", lwd="1")

```

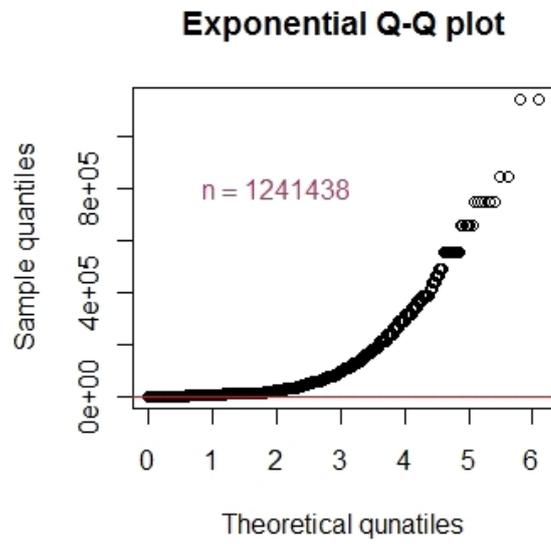


FIG. 6.13 – QQ-plot exponentiel des réclamations de l'année 1997

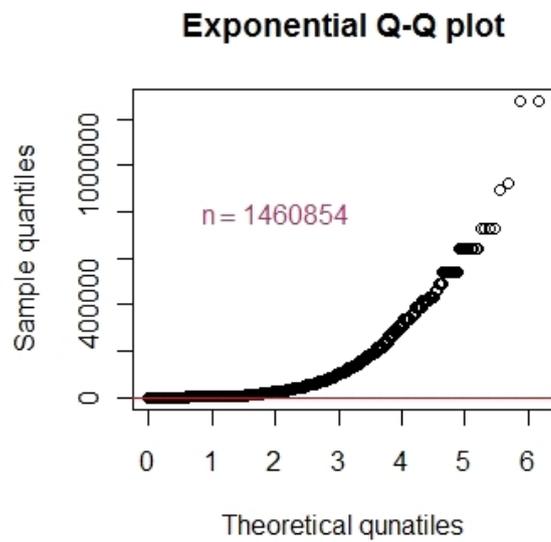


FIG. 6.14 – QQ-plot exponentiel des réclamations de l'année 1998

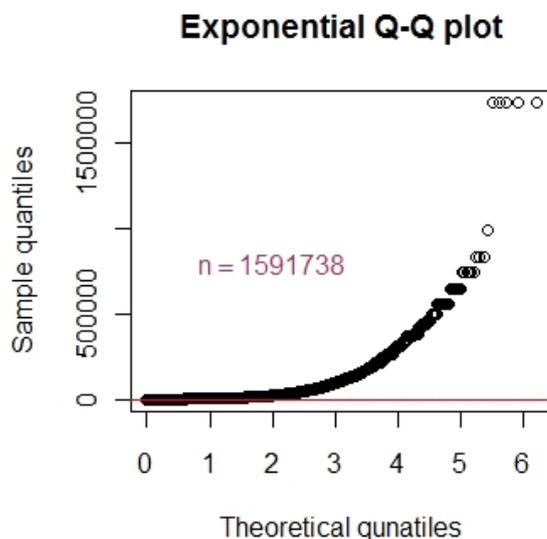


FIG. 6.15 – QQ-plot exponentiel des réclamations de l'année 1999

Sur les figures 6.11, 6.12 et 6.13, nous avons présenté les QQ – $plot$ exponentielle des réclamations des années 1997, 1998 et 1999 respectivement. Nous remarquons que les graphiques QQ – $plot$ sont concaves, ce qui signifie que l'échantillon de risque est issu d'une distribution à queue lourde.

6.2.3 QQ-norm de la distribution limite de $\Psi[\hat{Q}_n, g]$

Procédure sous R. Nous présentons ici un programme sous R afin de tracer le QQ – $norm$ de la distribution limite de l'estimateur $\Psi[\hat{Q}_n, g]$.

```

dataIn=read.table("E :\\DOCTORAT\\
these-meddi\\1997.txt",header=T,dec=".")
freq=dataIn[,1]
valeur=dataIn[,2]
nc=length(freq)
valeurs<-seq(1,nc)
x0<-seq(1,nc)
valeurs[x0]<-valeur[x0]/freq[x0]
freq
valeurs
repFreq=function(freq,valeurs)

```

```

{output=numeric(0)
i=1
while (i <=length(freq)){temp=numeric(0)
temp=rep(valeurs[i],freq[i])
output=c(output,temp)
i = i + 1}
return(output)}
monvect=repFreq(freq,valeurs)
n=length(monvect)
yr<-sort(monvect)
n
u<-runif(n)
p=1.1
a=n-2
b=n-1
b1=floor(n/3)
b2=floor(n-(n/3))
z=log10(y)
R=200
EST0=rep(0,R)
KA1=rep(0,R)
HILL=rep(0,R)
EST1=rep(0,R)
KA2=rep(0,R)
MATTYS=rep(0,R)
EST2=rep(0,R)
for( s in 1 :R){
yboot<-sample(y,length(y),replace=T)
Y<-sort(yboot)
Z<-log10(Y)
x1<-seq(1,n)
estb0<-sum((((x1/n)^(1/p))-(((x1-1)/n)^(1/p))))*Y[n-x1+1])
hills<-seq(1,b2)
for(k in 1 :b2){
l<-seq(1,k)
hills[k]<-((1/k)*sum(Z[n-l+1]))-Z[n-k]}
x2<-seq(1,b2)
f0 <- function (x) (1/x)*sum((x2^0.3)*abs(hills[x2]
-median(hills[1 :x])))
xmin0 <-optimize(f0, c(3, b1), tol = 0.0001)
K1=floor(xmin0$minimum)

```

```

hill<-hills[K1]
c1=K1+1
x0<-seq(c1,b)
estb1<-sum((((x0/n)^(1/p))-((x0-1)/n)^(1/p)))
*Y[n-x0+1])+(((K1/n)^(1/p))*(Y[n-K1]/(1-(p*hill))))
integrand <- function(u) {(((u^gama)-(u^(gama+1)))/
(1-(u^gama)))*log10(u)}
b=n-1
Int=1 :b2
Q<-seq(1,b2)
VAR<-seq(1,b2)
gh<-seq(1,b2)
for(k in 2 :b2){
g<-seq(1,100)
LV<-seq(1,100)
x3<-seq(1,k-1)
x5<-seq(1,100)
g[x5]<-x5*((2*gamma)/100)
f1 <-function (x) -sum(log10(((1-((x3/(k+1))^g[x]))*
(exp(((Y[n-k])^(-1/g[x]))*(x3-k-1)/(k+1)))))/(g[x]+
(((Y[n-k])^(-1/g[x]))*(x3/(k+1)))))-((((1-((x3/(k+1))^g[x]))
*(exp(((Y[n-k])^(-1/g[x]))*(x3-k-1)/(k+1)))))/(g[x]
+(((Y[n-k])^(-1/g[x]))*(x3/(k+1)))))*(x3*log10((y[n-x3+1]
-y[n-k])/(y[n-x3]-y[n-k])))))
xmin1 <-optimize(f1, c(1, 100), tol = 0.0001)
F=floor(xmin1$minimum)
gh[k]<-g[F]
gama=gh[k]
Int0=integrate(integrand, lower=0, upper=1)
Int[k]=Int0$val
Valeurs<-seq(1,n)
x4<-seq(1,1000)
x1<-seq(1,n)
SN<-seq(1,n)
for(m in 1 :n){
SN[m]<-sum(((1-u[m]^gh[k])^x4)/(x4^2))}
Valeurs[x1]<-(log10(u[x1])/gh[k])+
(((1+gh[k])/(gh[k]^2))*SN[x1])
VAR[k]<-var(Valeurs)}
x9<-seq(2,b2)
Q[x9]<-((1/k)*VAR[x9])+((((Y[n-x9])^(-1/gh[x9]))^2)

```

```

*((1/(2*gh[x9]^2))+(((1+gh[x9])/gh[x9])*Int[x9]))^2)
K2<-which.min(Q[b1 :b2])+(b1-1)
g<-seq(1,100)
Lv<-seq(1,100)
x6<-seq(1,K2-1)
x5<-seq(1,100)
g[x5]<-x5*((2*gamma)/100)
f2 <- function (x) (-(sum((log10((1-((x6/(K2+1))
^g[x]))/g[x]))-(((1-((x6/(K2+1))^g[x]))/g[x])
*((x6)*log10((y[n-x6+1]-y[n-K2])/(y[n-x6]-y[n-K2]))))))))
xmin2 <-optimize(f2, c(1, 100), tol = 0.0001)
H=floor(xmin2$minimum)
mattys<-g[H]
x7<-seq(1,K2)
x8<-seq(K2+1,b)
estb2<-(((K2/n)^(1/p))*Y[n-K2])+sum((((x8/n)
^(1/p))-(((x8-1)/n)^(1/p)))*Y[n-x8+1])+((((K2/n)^(
1/p))*(((n*K2)+n)/((K2*n)+K2))^(mattys))*
(sum((1/K2)*(x7)*(Y[n-x7+1]-Y[n-x7])
*((x7/(K2+1))^(mattys))/((1/p)-mattys)))
EST0[s]=estb0
KA1[s]=K1
HILL[s]=hill
EST1[s]=estb1
KA2[s]=K2
MATTYS[s]=mattys
EST2[s]=estb2}
qqnorm(EST2, col= "blue", ylab="limit distribution" );
qqline(EST2,col="3", lwd="2")
text(-1, 3.6, expression(n == 500), col="Maroon")
qqnorm(EST1, col= "red", ylab="limit distribution" );
qqline(EST1,col="green", lwd="2")
text(-1, 2, expression(n == 1000), col="Maroon")

```

Les figures 6.14, 6.15 et 6.16 montrent que les points des graphiques des QQ -norm des valeurs des distributions limites de l'estimateur $\Psi(g, \hat{Q}_n)$, pour les années 1997, 1998 et 1999 respectivement sont parfaitement alignés, ce qui signifie que cette distribution limite empirique de notre nouvel estimateur de mesure de risque $\Psi(g, \hat{Q}_n)$ est issue d'une loi normale.

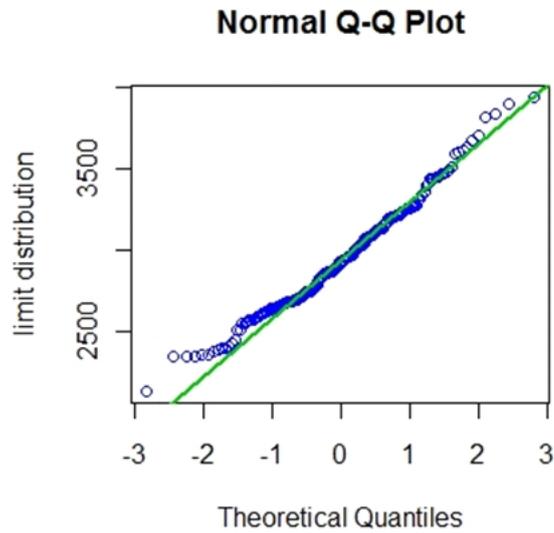


FIG. 6.16 – QQ-norm de la distribution limite de $\Psi(g, \hat{Q}_n)$ de 200 répétitions des données de l'année 1997

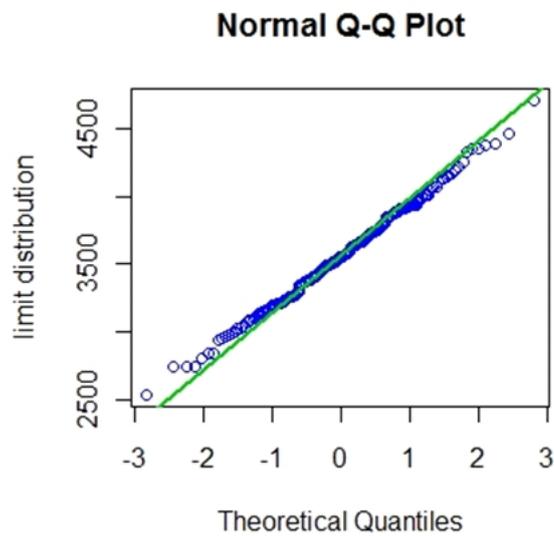


FIG. 6.17 – QQ-norm de la distribution limite de $\Psi(g, \hat{Q}_n)$ de 200 répétitions des données de l'année 1998

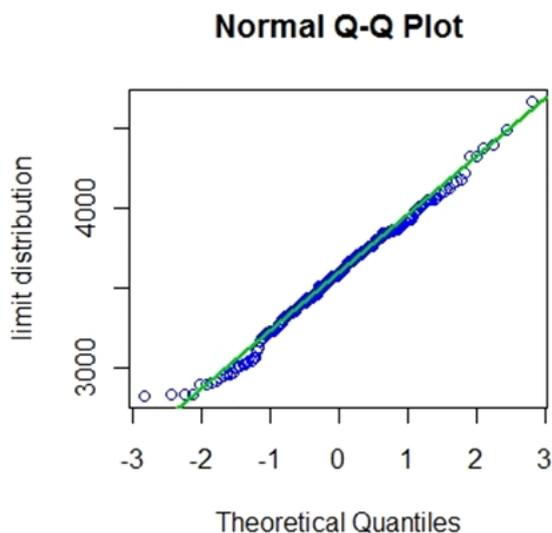


FIG. 6.18 – QQ-norm de la distribution limite de $\Psi(g, \hat{Q}_n)$ de 200 répétitions des données de l'année 1999

Commentaire. Nous avons utilisé le bootstrap pour avoir une idée sur le comportement asymptotique de l'estimateur de façon empirique sans conditions précédemment établies sur les données pour son application, cela par le calcul de la distribution limite empirique, l'écart type et les intervalles de confiance. D'un point de vue pratique, cette méthode numérique semble être plus facile que l'étude théorique de la normalité asymptotique pour deux raisons. La première est qu'elle n'est pas limitée par les conditions théoriques qui peuvent rendre à chaque fois, en appliquant le théorème de normalité asymptotique presque impossible si l'échantillon de risque, ne vérifie pas certains d'entre eux, parce que les actuaires ont besoin de procédures qui pourraient convenir directement dans leur étude de cas et ne pas chercher à s'adapter à un modèle théorique donnée précédemment sous certaines conditions, le plus souvent difficiles à vérifier. La technique de bootstrap donne plus de généralité et de flexibilité dans la pratique pour déterminer les propriétés de notre nouvel estimateur. Pour un point de vue asymptotique, la taille de l'échantillon est fixé indépendamment de la condition de faire tendre n vers l'infini, ce qui rend notre étude plus pratique. Nous concluons que notre étude empirique n'est pas seulement une alternative importante à l'étude asymptotique, mais il peut aussi remédier aux insuffisances de l'étude théorique asymptotique.

Bibliographie

- [1] Artzner, P., Delbaen, F., Eber, J. M. and Heath, D., 1999. Coherent measures of risk. *Math. Finance*. 9(3), 203-228.
- [2] Athreya, K.B. and Ney, P.E., 1972. *Branching Processes*. Springer-Verlag, New York.
- [3] Beirlant, J., Diercks, G., Guillou, A. and Stărică, C., 2002. On exponential representations of log-spacings of extreme order statistics. *Extremes*. 5(2), 157-180.
- [4] Beirlant, J., Figueiredo, F., Gomes, M.I. and Vandewalle, B., 2008. Improved reduced-bias tail index and quantile estimators. *J. Statist. Plann. Inference*. 138(6), 1851-1870.
- [5] Beirlant, J., Goegebeur, Y., Segers, J. & Teugels, J., 2004, *Statistics of Extremes - Theory and Applications*, Wiley, England.
- [6] Beirlant, J., Matthys, G., Diercks, G., 2001. Heavy-tailed distributions and rating. *Astin Bull.* 31(1), 37-58.
- [7] Benkatander, G., Segerdahl, C., 1960. On the analytical representation of claim distributions with special reference to excess of loss reinsurance. In : *XVIth International Congress of Actuaries*, Brussels.
- [8] Bingham, N.H., Goldie, C.M., and Teugels, J.L., 1987. Regular Variation. *Encyclopedia of Mathematics and its Applications* 27. Cambridge University Press, Cambridge.
- [9] Borovkov, A.A., 1984. *Asymptotic Methods in Queueing Theory*. John Wiley & Sons, Chichester.
- [10] Borovkov, A.A., *Stochastic Processes in Queueing Theory*, 1976. *Applications of Mathematics* 4. Springer-Verlag, New York.
- [11] Brahim, B., Meddi, F. and Necir, A., 2012. Bias-corrected estimation in distortion risk premiums for heavy-tailed losses. *Journal Afrika Statistika*. Vol. 7, pages 474-490.
- [12] Brahim, B., Meraghni, D., Necir, A. and Zitikis, R., 2011. Estimating the distortion parameter of the proportional hazard premium for heavy-tailed losses. *Insurance Math. Econom.* 49, 325-334.

- [13] Brazauskas, V., Jones, B. L.; Puri, M.L. and Zitikis, R., 2008. Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference*. 138, 3590-3604.
- [14] Caeiro, F., Figueiredo, F. and Gomes, M.I., 2004. Bias reduction of a tail index estimator through an external estimation of the second order parameter. *Statistics*. 38(6), 497-510.
- [15] Caeiro, F., Gomes, M.I. and Rodrigues, L.H., 2009. Reduced-bias tail index estimators under a third-order framework. *Comm. Statist. Theory Methods*. 38(6-7), 1019-1040.
- [16] Cebrian, A. C. , Denuit, M., Lambert, P. 2004. Generalized Pareto fit to the Society of Actuaries large claims database. *North American Actuarial Journal*. Vol. 8. PP. 18-36.
- [17] Centeno, M.L. and Andrade e Silva, J., 2005. Applying the proportional hazard premium calculation principle. *Astin Bulletin*. 35, 409-425.
- [18] Chicheportiche, R. and Bouchaud, J. P. 2012. Weighted Kolmogorov-Smirnov test : Accounting for the tails. *Phys. Rev. E*. Vol 86. Issue 4.
- [19] Chistyakov, V.P., 1964. A theorem on sums of independent positive random variables and its application to branching random processes. *Theory Proba. Appl.*,p(4), pp. 640-648.
- [20] Csörgő, M., Csörgő, S., Horváth, L. and Mason, D.M., 1986. Weighted empirical and quantile processes. *Ann. Probab.* 14(1), 31-85.
- [21] Das B., and Resnick, S. I. 2008. QQ plots, Random sets and data from a heavy tailed distribution. *Stochastic Models*. Vol. 24. PP. 103-132.
- [22] Das, B., and Ghosh, S, 2013. Weak limits for exploratory plots in the analysis of extremes. *Bernoulli Society for Mathematical Statistics and Probability*. Vol 19, Number 1. PP. 308-343
- [23] de Haan, L. and Ferreira, A., 2006. *Extreme Values Theory : An introduction*. Springer.
- [24] de Haan, L. de and U. Stadtmüller 1996. Generalized regular variation of second order. *J. Austral. Math. Soc. (A)* 61, 381-395.
- [25] de Haan, L., 1970. On regular Variation and its Application to the Weak Convergence of Sample Extremes. MC Tract 32. Mathematical Centre, Amsterdam.
- [26] de Haan, L., 1984. Slow variation and characterization of domains of attraction. In J. Tiago de Oliveira (Ed.), *Statistical Extremes and Applications*, pp. 31-48.

- [27] Demuit, M., Dhaene, J., Goovaerts, M.J. and Kaas, R., 2005. Actuarial Theory for Dependent Risk : Measures, Orders and Models. Wiley, New York.
- [28] Embrechts and C.M. Goldie, 1980. On closure and factorization theorems for subexponential and related distributions. J. Austral. Math. Soc., Ser. A, 29(2), pp. 243-256.
- [29] Embrechts, P., Kluppelberg, C. & Mikosch, T., 1997, Modelling Extremal Events for Insurance and Finance, Springer, Berlin.
- [30] Embrechts, P., McNeil, A. J., Frey, R., 2005. Quantitative Risk Management : Concept, Techniques, and Tools. Princeton University Press.
- [31] Feureverger, A. and Hall, P., 1999. Estimating a tail exponent by modeling departure from a Pareto distribution. Ann. Statist. 27(2), 760-781.
- [32] Fisher R.A., Tippett L.H.C. 1928, Limiting forms of the frequency distribution of the largest or smallest of a sample Porc. Cambridge Philos. Soc., 24, pp. 180-190.
- [33] Frank, J. Massey. Jr. 1951. The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association. Vol. 46, No. 253
- [34] Furman, E. and Zitikis, R., 2008a. Weighted premium calculation principles. Insurance Math. Econom. 42, 459-465.
- [35] Furman, E. and Zitikis, R., 2008b. Weighted risk capital allocations. Insurance Math. Econom. 43, 263-269.
- [36] Geluk, J. and de Haan,L., 1987. Regular Variation, Extensions and Tauberian Theorems. CWI Tract 40. Center for Mathematics and Computer Science, Amsterdam.
- [37] Goldie, C.M. and Kluppelberg, C., 1998. Subexponential distributions. In R.J. Adler, R.E. Feldman, and M.S. Taqqu (Eds.), A practical Guide to Heavy Tails : Statistical Techniques and Applications, pp. 435-459.[14]
- [38] Gomes, M.I. and Figueiredo, F., 2006. Bias reduction in risk modelling : semi-parametric quantile estimation. Test. 15(2), 375-396.
- [39] Gomes, M.I. and Martins, M.J., 2002. "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. Extremes. 5(1), 5-31.
- [40] Gomes, M.I. and Martins, M.J., 2004. Bias reduction and explicit semi-parametric estimation of the tail index. J. Statist. Plann. Inference. 124(2), 361-378.
- [41] Gomes, M.I. and Pestana, D., 2007. A sturdy reduced-bias extreme quantile (VaR) estimator. J. Amer. Statist. Assoc. 102(477), 280-292.

- [42] Goovaerts, Marc J. & Roger J. A. Laeven (2008). Actuarial risk measures for financial derivative pricing, *Insurance : Mathematics and Economics* 42, 540-547.
- [43] Greselin, F., Puri, M.L. and Zitikis, R., 2009. L-functions, processes, and statistics in measuring economic inequality and actuarial risks. *Statistics and Its Interface.* 2, 227-245.
- [44] Hall, W., Wellner, J., 1981. Mean residual life. *Statistics and Related Topics*, 169-184.
- [45] Hans Bühlmann ,B. Gagliardi, H.U and Gerber, E. Straub. 1977. Some inequalities for stop-loss premiums, *ASTIN Bulletin*
- [46] Hans Bühlmann, 1980. An economic premium principle, *ASTIN Bulletin*
- [47] Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann. Statist.* 3(5), 1163-1174.
- [48] Hogg, R., Klugman, S., 1984. *Loss distributions.* Wiley, New York.
- [49] Jones, B.L. and Zitikis, R., 2003. Empirical estimation of risk measures and related quantities. *North American Actuarial Journal.* 17, 44-54.
- [50] Jones, B.L. and Zitikis, R., 2007. Risk measures, distortion parameters, and their empirical estimation. *Insurance Math. Econom.* 41, 279-297.
- [51] Kaas, R., van Heerwaarden, A.E. and Goovaerts, M.J., 1994, Ordering of actuarial risks
- [52] Kimball, B., 1960. On the choice of plotting positions on probability paper, *Journal of the American Statistical Association* 55. PP. 546-560.
- [53] Kolmogorov, A. N. 1933. *Giornale dell'Istituto degli Attuari* 4, 83.
- [54] Madan, D. B. and Schoutens, W., 2010. Conic financial and the corporate finance. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1547022
- [55] Mason, D., 1982. Laws of large numbers for sums of extreme values. *Ann. Probab.* 10(3), 754-764.
- [56] Matthys, G. and Beirlant, J., 2003. Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica.* 13, 853-880.
- [57] Matthys, G., Delafosse, E., Guillou, A. and Beirlant, J., 2004. Estimating catastrophic quantile levels for heavy-tailed distributions. *Insurance Math. Econom.* 34(3), 517-537.
- [58] Necir, A. and Boukhetala. K. 2004. Estimating the Risk-Adjusted Premium for The Largest Claims Reinsurance Covers. *COMPSTAT'2004. Proceeding in Computational Statistics.* ISBN 3-7908-1554-3, Physica-Verlag, Heidelberg, New York.

- [59] Necir, A. and Meraghni, D., 2009. Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance Math. Econom.* 45, no. 1, 49-58.
- [60] Necir, A. and Meraghni, D., 2010. Estimating L-functionals for heavy-tailed distributions and applications. *Journal of Probability and Statistics* 2010, ID 707146.
- [61] Necir, A., Meraghni, D. and Fatima Meddi, 2007. Statistical estimate of the proportional hazard premium of loss. *Scandinavian Actuarial Journal*. Issue 3 pages 147-161.
- [62] Neves, C. and Fraga Alves, M.I., 2004. Reiss and Thomas' automatic selection of the number of extremes. *Comput. Statist. Data Anal.* 47(4), 689-704.
- [63] Pack, A.G., 1975. On the tails of waiting-time distributions. *J. Appl. Probab.*,12(3), pp. 555-564.
- [64] Peng, L. and Qi, Y., 2004. Estimating the first- and second-order parameters of a heavy-tailed distribution. *Aust. N. Z. J. Stat.* 46(2), 305–312.
- [65] Peng, L., Qi, Y., Wang, R. and Yang, J. 2012. Jackknife empirical likelihood method for some risk measures and related quantities. *Insurance Math. Econom.* 51(1), 142-150.
- [66] Pitman, E.J.G., 1980. Subexponential distribution functions. *J. Austral. Math. Soc., Ser. A*, 29(3), pp. 337-347
- [67] Reiss, R.-D. and Thomas, M., 2007. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd ed. Birkhäuser Verlag, Basel, Boston, Berlin.
- [68] Rolski, T., Schimidli, H., Schimd, V. and Teugels, J.L., 1999. *Stochastic Processes for Insurance and Finance*. John Wiley and Sons, New York.
- [69] Rytgaard, M. 1996. Simulation experiments on the mean residual life function $m(x)$. In : *Proceedings of the XXVII ASTIN Colloquium*, Copenhagen, Denmark, vol. 1, pp. 59-81. [296, 298].
- [70] Seneta, E., 1976. *Functions of Regular Variation*. Lecture Notes in Mathematics 508. Springer-Verlag, Berlin.
- [71] Serfling, R. J. 1981. *Approximation Theorems of Mathematical Statistics*. John Wiley Sons Inc., New York, ISBN 0-471-02403-1.
- [72] Shapiro, A., Dentcheva, D. and A. Ruszczyński. 2009. *Lectures on Stochastic Programming : Modeling and Theory*. MPS-SIAM, Philadelphia-USA.
- [73] Shorack, G. R. and Wellner, J. A. 1986. *Empirical Processes with Applications to Statistics*. Wiley, New York. [83, 195, 292, 596, 534].

- [74] Smirnov, N. 1948. The Annals of Mathematical Statistics 19, 279.
- [75] Smith, R., 1985, Maximum likelihood estimation in a class of non regular cases, *Biometrika*, 72, 67–90.
- [76] Souvik, G., Sidney, I. Resnick. S. I. 2011. When Does the Mean Ecess Plot Look Linear ?. *Stochastic Models*, 27 :4, 705-722.
- [77] Teugels, J.L., 1975. The class of subexponential distributions. *Ann. Probab.*, 3(6), pp. 1000-1011.
- [78] Thas, O., 2010. Comparing distributions. Springer Series in Statistics
- [79] Vandewalle, B. and Beirlant, J., 2006. On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance : Mathematics and Economics*, 38, 441-459.
- [80] Wang S., 2002. A risk measure that goes beyond coherence, Actes du 12e colloque AFIR, Cancun.
- [81] Wang, S.S., 1996. Premium calculation by transforming the layer premium density. *Astin Bulletin*. 26, 71-92.
- [82] Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* 73(364), 812-815.
- [83] Wirch, J.L. and Hardy, M.R., 1999. A synthesis of risk measures for capital adequacy. *Insurance Math. Econom.* 25, 337-347.
- [84] Yang, G., 1978. Estimation of a biometric function. *The annals of Statistics* 6 (1), 112-116

Abstract. Recently Necir and Meraghni (2009) proposed an asymptotically normal estimator for distortion risk premiums when losses follow heavy-tailed distributions. In this thesis, we propose a bias-corrected estimator of this class of risk premiums and establish its asymptotic normality. Our considerations are based on the high quantile estimator given by Matthys and Beirlant (2003).

Key words : Bias reduction, High quantiles, Hill estimator, L-statistics, Order statistics, Risk Measure, Second order regular variation, Tail index.

Résumé. Récemment Necir et Meraghni (2009) ont proposé un estimateur asymptotiquement normal pour les primes de risque de distorsion lorsque les pertes suivent des distributions à queues lourdes. Dans cette thèse, nous proposons un estimateur à biais réduit, pour cette classe de primes de risque et nous établissons sa normalité asymptotique. Nos considérations sont basées sur l'estimateur des quantiles extrêmes introduits par Matthys and Beirlant (2003).

Mots clés : Réduction du biais, Quantiles extrêmes, Estimateur de Hill, L-statistiques, Les statistiques d'ordre, Mesure du risque, Variation régulière du deuxième ordre, Indice de queue.