

UNIVERSITÉ MOHAMED KHEIDER, BISKRA
FACULTÉ DES SCIENCES EXACTES ET DES SCIENCES DE LA NATURE ET LA
VIE

DÉPARTEMENT DE MATHÉMATIQUES

THÈSE DE DOCTORAT

Discipline: Mathématiques Appliquées

Présentée par

Zoubir KENIOUA

SUR LES MESURES DE RISQUES ET LEURS APPLICATIONS

Sous la direction de Mr. Brahim BRAHIMI

Jury:

Pr. Djabrane YAHIA	MC(A). UMK Biskra	Président
Dr. Halim Zaghoudi	MC(A). UBM Annaba	Examineur
Dr. Abbes Benchaben	MC(A). U Guelma	Examineur
Dr. Fatah BENATIA	MC(A). UMK Biskra	Examineur

Decembre 2017

DÉDICATE

Je dédie cette thèse
à ma mère
qui était elle qui m'encourager
à prendre
cette étape dans ma vie.

REMERCIEMENTS

Mes premiers remerciements vont naturellement à ce qui a encadré cette thèse. C'est donc un grand merci que j'adresse à mon directeur de thèse, Mr. **Brahim Brahimi**, Professeur à l'université Mohamed Khider Biskra, pour sa disponibilité, pour son soutien, pour sa patience, pour m'avoir transmis sa passion pour la recherche. Pour tout cela et pour m'avoir guidé jusqu'au terme de ma thèse, je souhaite lui exprimer ma gratitude.

Je remercie Mr. **Halim Zaghoudi**, Maitre de Conférences-A à l'université Badji Mokhtar Annaba et Mr. **Abbes Benchaben** Maitre de Conférences-A à l'université de Guelma, pour avoir accepté d'être examinateurs de ma thèse. Je vous remercie énormément.

Je remercie également Mr. **Djabrane Yahia**, Maitre de Conférences-A à l'Université Mohamed Khider Biskra, de l'honneur qui'il m'a fait d'avoir accepté de présider mon jury de thèse.

Je remercie Mr. **Fatah Benatia**, Maitre de Conférences-A à l'Université Mohamed Khider Biskra, d'avoir accepté de participer au jury de cette thèse.

Merci à tous les membres du Laboratoire de Mathématiques Appliquées (LMA): doctorants, maîtres de conférences et professeurs pour leurs accueil chaleureux et cela depuis le début de ma thèse.

Et pour finir, un grand merci à ma famille pour son soutien constant et chaleureux pendant toute la réalisation de ce travail.

ABSTRACT

The simplest and most common estimator of the tail index is the Hill estimator (Hill,1975), which works only for Pareto-type tails and has high bias, is also highly non-robust in the presence of extreme values in relation to the assumed model. Recently, there have been some attempts to produce arobust alternatives to Hill estimator. In this thesis, we use the so-called t-Hill estimator of the tail index proposed by Fabián (2001), rather than Hill estimator to derive a robust estimator of the distorted risk premiums for heavy-tailed distributions proposed by Necir and Meraghni (2009). Using the second condition of regular variation, we proved its asymptotic normality. By studying the simulation, we demonstrated that this new estimator is more robust than the Necir and Meraghni (2009) estimator for either small or large samples.

Keywords: Risk measures, Distortion risk premiums, Extreme values, Tail, Robustness.

LE plus simple et la plus courante estimateur de l'indice de queue est destiné à Hill (Hill, 1975), qui ne fonctionne que pour les queues du type de Pareto et a un biais élevé, est également non-robuste en présence de valeurs extrêmes par rapport au modèle supposé. Récemment, il y a eu quelques tentatives pour produire un estimateur alternatif plus robuste que l'estimateur de Hill. Dans cette thèse, nous utilisons le soi-disant estimateur de t-Hill pour l'indice de queue proposé par Fabián (2001), au lieu de Hill pour dériver un estimateur robuste de la prime de risque de distortion pour les distributions à queue lourdes proposés par Necir and Meraghni (2009). En utilisant la deuxième condition de la variation régulière, nous établissons la normalité asymptotique. Grâce à l'étude de simulation, nous avons montré que ce nouvel estimateur est plus robuste que celui de Necir and Meraghni (2009) à la fois pour les petits et les grands échantillons.

Mots clés: Mesures de risques, Prime de risque de distortion, queue de distribution, Valeurs extrêmes, Robustesse.

Table des Matières

Dédicace	ii
Remerciements	iii
Abstract	iv
Résumé	v
Introduction	1
1 Théorie de valeurs extrêmes	5
1.1 Introduction	5
1.1.1 Bref historique	5
1.1.2 Théorie des extrêmes: principe et applications	5
1.2 Les lois de maximums	6
1.3 Problème de la limite des extrêmes	10
1.4 La Distribution Généralisée de Jenkinson-von Mises	13
1.5 Problème du domaine d'attraction du maximum	16
1.5.1 Fonctions à variation régulière et leurs propriétés principales	17
1.5.1.1 Propriétés générales	17
1.5.1.2 Propriétés de base	19
1.5.2 Caractérisation des trois domaines d'attraction	21
1.5.2.1 Domaine d'attraction de Fréchet $\Phi_\gamma (\gamma > 0)$	22
1.5.2.2 Domaine d'attraction de Weibull $(\mathcal{MDA}(\Psi_\gamma), \gamma < 0)$	24
1.5.2.3 Domaine d'attraction de Gumbel $(\mathcal{MDA}(\Lambda), \gamma = 0)$	25
1.5.3 Caractérisation pour différents domaines d'attraction	27
1.6 La loi des excès	28
1.6.1 Principe	28
1.6.2 Théorème de Pickands-Balkema-de Haan	29
2 Estimation de l'Indice de queue	30
2.1 Introduction	30

2.2 Quelques estimateurs de l'indice de queue	31
2.2.1 Estimateur de Pickands ($\gamma \in \mathbb{R}$)	31
2.2.1.1 Construction de l'estimateur de Pickands	32
2.2.2 Estimateur de Hill ($\gamma > 0$)	33
2.2.2.1 Construction de l'estimateur de Hill	33
2.2.3 Estimateurs des moments ($\gamma \in \mathbb{R}$)	34
2.2.3.1 Construction de l'estimateur de moments	35
2.2.4 L'estimateur t-Hill	35
2.2.4.1 Construction de l'estimateur t-Hill	35
2.3 Principales propriétés asymptotiques: Consistance et Normalité	
Asymptotique	37
2.3.1 Estimateur de Pickands ($\gamma \in \mathbb{R}$)	40
2.3.2 Estimateur de Hill	41
2.3.3 Estimateur des moments	41
2.3.4 L'estimateur t-Hill	42
2.3.5 Comment choisir le nombre de statistiques d'ordre	43
2.4 Estimation de quantiles extrêmes	43
3 Mesures de risques	45
3.1 Notations et définitions	45
3.2 Quelques mesures de risque usuelles	51
3.2.1 La variance	51
3.2.2 La Value-at-Risk	51
3.2.3 La Tail Value-at-Risk	53
3.2.3.1 Définitions	53
3.2.3.2 Propriétés	55
3.3 Principes de calcul de primes	55
3.4 Mesure du risque de distorsion	59
3.4.1 Propriétés de mesures de risque de distorsion	62
3.4.2 Exemples de mesures de risques de distorsion	63
3.5 Estimation de mesures de risque de distorsion	65
4 Robust estimator of distortion risk premiums for heavy-tailed losses	73
4.1 Introduction	73
4.2 Defining the estimator	76
4.2.1 Heavy-tailed losses case	77
4.3 Asymptotic distribution	78

4.4 Simulation study	79
4.4.1 Performance and comparative study of $\tilde{\Pi}_n$ and $\hat{\Pi}_n$	79
4.4.2 Comparative robustness study	80
4.5 Concluding notes	82
4.6 Proofs	82
4.6.1 Proof of Theorem 4.3.1	83
Bibliographie	95

Liste des figures

- 1.1 F.d.r $F(x)$ des trois types d'extrêmes, respectivement $(\gamma = 1, \mu = 1, \sigma = 1)$, $(\gamma = -1, \mu = -1, \sigma = 1)$, $(\gamma = 0, \mu = 0, \sigma = 1)$; en comparaison avec la fonction de survie gaussienne de même moyenne et de variance que cel de Gumbel. 14
- 1.2 Fonctions de queues des trois types d'extrêmes, respectivement $(\gamma = 1, \mu = 1, \sigma = 1)$, $(\gamma = -1, \mu = -1, \sigma = 1)$, $(\gamma = 0, \mu = 0, \sigma = 1)$; en comparaison avec la densité gaussienne de même moyenne et de variance que cel de Gumbel. 15
- 1.3 Densités des trois types d'extrêmes, respectivement $(\gamma = 1, \mu = 1, \sigma = 1)$, $(\gamma = -1, \mu = -1, \sigma = 1)$, $(\gamma = 0, \mu = 0, \sigma = 1)$; en comparaison avec la densité gaussienne de même moyenne et de variance que cel de Gumbel. Le panneau de droite est un zoom sur la queue. 17
- 4.1 Plot of the premium estimator $\widetilde{\Pi}_n(\hat{\gamma}, k_1)$ as a function of the sample size n based on 200 Pareto-distributed claims with tail index 0.6. The distortion parameter is 1.12. The horizontal line represents the true value of the premium 2.0487. 81

Liste des tableaux

- 1.1 Lois limites des moyennes et des maxima 13
- 1.2 Exemples des lois qui appartiennent au $\mathcal{MDA}(\Phi_\gamma)$ 24
- 1.3 Exemples des lois qui appartiennent au $\mathcal{MDA}(\Psi_\gamma)$ 25
- 1.4 Exemples des lois qui appartiennent au $\mathcal{MDA}(\Lambda)$ 26

- 3.1 Un Aperçu de Mesures de Risque de Distortion 65

- 4.1 $\hat{\gamma}(k_1^*)$, $\hat{\gamma}^H(k_2^*)$, $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ and $\hat{\Pi}_n(\hat{\gamma}, k_2^*)$ estimators based on 1000 samples of Pareto-distributed claim amounts with tail index 0.6 and distortion parameter $\rho = 1.12$. The exact value of the premium is 2.0487. 80
- 4.2 $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ and $\hat{\Pi}_n(\hat{\gamma}^H, k_2^*)$ are based on 1000 samples of mixture of Pareto distributions with tail index 0.6, $\epsilon = 5\%, 10\%, 15\%, 25\%$ and distortion parameter $\rho = 1.12$. The exact value of the premium is 2.0487. 82

LE risque est omniprésent dans tous les domaines de la vie humaine, mais il est pratique de considérer une définition spécifique en fonction de la zone d'intérêt. Par exemple, dans le contexte de la littérature financière, le terme **risque** désigne la variation liée à la variabilité de la valeur future Artzner et al. [2, 1999].

La quantification du risque ou de l'incertitude dans la valeur future d'un portefeuille est l'une des tâches les plus importantes dans la gestion des risques. Cette quantification est habituellement obtenue en modélisant le rendement incertain en tant que variable aléatoire, à laquelle une certaine fonction est appliquée. Ces fonctions sont habituellement appelées mesures de risque.

Une mesure de risque tente d'attribuer une seule valeur numérique à une perte financière potentielle aléatoire. Les applications des mesures de risque comprennent, par exemple, les exigences de marge dans le commerce financier, les primes de risque d'assurance. D'une manière conventionnelle, les distributions de rendement des portefeuilles sont considérées dans les applications de finance et d'investissement, tandis que les distributions de pertes de portefeuilles sont concernées par les applications d'assurance.

Les mesures de risque sont les outils fondamentaux utilisés dans presque tous les domaines de la gestion des risques. Diverses mesures de risque ont été introduites avec des justifications économiques. Dans cette thèse, nous nous concentrons sur les mesures de risque qui possèdent de bonnes propriétés et capables de saisir le comportement de queue d'un risque. L'une des propriétés largement acceptées pour de bonnes mesures de risque est la soi-disant cohérence introduite par Artzner et al. [2, 1999], qui comprend quatre axiomes: la monotonie, l'homogénéité positive, la subadditivité et l'invariance par

translation.

La mesure de risque de référence, Value-at-Risk (VaR), a été critiquée car il ne pas satisfait la propriété de la subadditivité et ne donne qu'une information ponctuelle au quantile α de la distribution du pertes et aucune information au delà de ce point. C'est pourquoi d'autres mesures de risque ont été proposées. Par exemple, Tail Value-at-Risk (TVaR) ou Conditional Tail Expectation qui est une mesure de risque prometteuse dans le sens qu'elle est cohérente et capable de capturer des risques extrêmes. La VaR et ses alternatives sont des cas particuliers d'une class de mesures de risque dites mesures de risque de distortion.

Dernièrement, la crise financière a exposé les principales lacunes des méthodologies traditionnelles d'évaluation des risques en ce qui concerne la capture du risque d'événements rares mais dommageables, ce qui a permis de rechercher de meilleures approches de la modélisation et de la mesure des risques plus cruciale.

Par sa nature même, le risque d'événements extrêmes (par exemple, des pertes très importantes) est lié à la queue de la distribution. Ainsi, un défi crucial pour obtenir de bonnes estimations de la mesure des risques est de pouvoir estimer la queue de la distribution sous-jacente aussi précisément que possible. Née dans la période de l'entre-deux-guerres des travaux de Fisher et Tippett (1928); Fréchet (1927); Gnedenko (1943); Gumbel (1958), la théorie des valeurs extrêmes propose un cadre statistique pour traiter ces problèmes (voir par exemple Embrechts et al. [33, 1997, Chapitre 3]). Elle a été développée pour estimer la probabilité d'occurrence des événements rares, pour étudier la queue de la distribution de ces phénomènes et estimer un quantile extrême. Pour un large spectre des problématiques que la théorie des valeurs extrêmes a tenté d'explorer dans les dernières décennies nous renvoyons le lecteur aux livres de Leadbetter et al. [58, 1983], Resnick [74, 1987], Embrechts et al. [33, 1997], Beirlant et al. [5, 2004], de Haan and Ferreira [21, 2006], Reiss and Thomas [73, 2007].

Organisation de la thèse

Le présent travail est divisé en quatre chapitres qui sont organisés comme suit:

Chapitre 1

Dans le chapitre 1, les notions fondamentaux de la théorie de valeurs ex-

trêmes sont fournis qui sont nécessaires pour une bonne compréhension des chapitres suivants. Dans la section 1.1 nous donnons un bref historique de la théorie de valeurs extrêmes ainsi leur principe et quelques applications. Nous présentons plusieurs termes et concepts de base, tels que les maxima et les minima de l'échantillon et leurs comportements limités, nécessaires pour un travail ultérieur, ensuite, nous présenterons le théorème célèbre de Fisher-Tippett, qui décrit les trois types possibles de lois limites, les distributions de valeurs extrêmes c'est le problème de loi limite. Dans la section 1.5, par l'intermédiaire de la variation régulière, nous tenterons d'expliquer comment les maxima des échantillons sous la transformation affine appropriée se comportent dans la limite et quelles sont les conditions de convergence sur les distributions sous-jacentes c'est le problème de maximum domaine d'attraction. La fin est consacrée à la distribution généralisée de Pareto, qui fournit une application intéressante à l'analyse de données réelle.

Chapitre 2

Comme nous l'avons déjà mentionné dans le chapitre 1, l'indice de queue γ a une position cruciale dans la théorie de valeurs extrêmes. Lorsque nous travaillons avec des données réelles, il est particulièrement important d'estimer l'indice de queue, car les estimations de la fonction de distribution sous-jacente et des quantiles extrêmes dépendent de γ qu' est essentielle pour les prochains chapitres, c'est l'objective du chapitre 2. Nous présentons quelques estimateurs classiques de l'indice de queue γ ainsi que leurs propriétés asymptotiques. Nous donnons également certaines méthodes de selection du nombre de valeurs extrêmes utilisé lors de l'estimation du paramètre de l'indice de queue. Et enfin nous exposons certaines méthodes d'estimation de quantiles extrêmes.

Chapitre 3

Le chapitre 3 traite des mesures de risque, dans la section 3.1, nous présentons quelques notations et définition pour les sections suivantes. Des mesures de risque bien connues et des relations qui existent entre elles sont présentées dans la section 3.2. Dans la section 3.3 nous discutons de ce qu'est un principe de prime et présentons certains principes de primes importants, des propriétés désirables qu'il peut posséder. La classe des mesures de risque de distorsion est examinée à la section 3.4, nous discutons de ce qu'est un mesure de risque de distorsion et nous donnons des propriétés désirables qu'il peut posséder. La section 3.5 étudie l'estimation empirique et semi-paramétrique par l'intermédiaire de la théorie de valeurs extrêmes.

Chapitre 4

Le Chapitre 4 est constitué d'un article co-écrit avec mon directeur de thèse, s'intitulant Robust estimator of distortion risk premiums for heavy-tailed losses et publié dans la revue: Afrika Statistika, [Brahimi and Kenioua \(2016\)](#). Dans ce chapitre, nous suivons le même approche utilisée par Necir and Meraghni [63], mais au contraire, nous traitons le problème de la robustesse des estimateurs proposés. Pour cela, nous utilisons l'estimateur de l'indice de queue dit t-Hill proposé par Fabián [34], au lieu de l'estimateur de Hill, pour obtenir un estimateur robuste pour la prime de risque de distorsion des pertes. Sous la condition de second ordre de variation régulière, nous établissons sa normalité asymptotique. Par l'étude de simulation, nous montrons que ce nouvel estimateur est plus robuste que de celui proposé par Necir and Meraghni [63] pour les petites et les grandes tailles d'échantillon.

1.1 Introduction

1.1.1 Bref historique

La théorie des valeurs extrêmes (EVT) est essentiellement consacrée à la caractérisation du comportement des observations extrêmes. Les origines de cette théorie remontent aux travaux pionniers de Fréchet [39] en 1927 et Fisher and Tippett [36] en 1928 sur la convergence du maximum d'un échantillon de variables aléatoires i.i.d. À cette époque, les limites possibles pour les distributions de maxima d'échantillons de variables aléatoires indépendantes et identiquement distribuées ont été dérivées. La preuve complète a été proposée par Gnedenko [42] en 1943, dont la preuve fut simplifiée par de Haan [20] en 1976. Ces distributions limites ont été unifiées par un paramétrage donné par von Mises [85] en 1954 et Jenkinson [50] en 1955, menant à la distribution des valeurs extrêmes généralisées. Gumbel [44] en 1958 a été le premier à payer l'attention sur les applications possibles de la théorie des valeurs extrêmes d'un point de vue statistique, il exploite ce résultat dans son célèbre livre « Statistics of Extremes ». Dès lors, des développements majeurs ont été faits en théorie et en applications, son livre étant toujours pertinentes aujourd'hui.

1.1.2 Théorie des extrêmes: principe et applications

Il s'agit dans l'étude des valeurs extrêmes d'analyser l'épaisseur des queues de distributions, ou encore d'étudier les plus grandes observations d'un échantillon pour caractériser sa loi initiale. Ainsi, la théorie des extrêmes vient en

complément de la théorie statistique classique où il est plus commun d'étudier le comportement d'une distribution autour de sa moyenne plutôt que dans le domaine des observations extrêmes souvent appelées événements rares. Nous allons voir que toute la théorie des extrêmes est fondée sur un équivalent au théorème central limite mais pour les queues de distribution. La théorie des valeurs extrêmes joue un rôle important dans de nombreux domaines principalement en relation avec l'analyse du risque comme l'hydrologie, des finances et de l'ingénierie. On peut citer comme exemples :

- l'hydrologie, domaine dans lequel l'estimation du niveau des crues par exemple est particulièrement importante ;
- la finance où l'hypothèse de gaussianité, incompatible avec les observations surtout en hautes fréquences, est de plus en plus fortement remise en cause ;
- l'assurance dont l'une des préoccupations est la prise en compte des grands sinistres ;
- la météorologie où l'étude de la vitesse du vent permet d'évaluer le degré de résistance des matériaux face à la pression exercée par le vent (au cours d'une tempête par exemple) sur les bâtiments ou les structures de génie civil ;
- la macro-économie réelle dans le but de prévoir les crises économiques.

La théorie de valeurs extrêmes est essentiellement née pour répondre à ce genre de questions (voir par exemple Embrechts et al. [33, Chapitre 3]). Elle a été développée pour estimer la probabilité d'occurrence des événements rares et pour étudier la queue de la distribution de ces phénomènes. Pour un large spectre des problématiques que la théorie des valeurs extrêmes a tenté d'explorer dans les dernières décennies nous renvoyons le lecteur aux livres de Leadbetter et al. [58], Resnick [74], Embrechts et al. [33], Beirlant et al. [5], Coles [16], de Haan and Ferreira [21], Reiss and Thomas [73].

1.2 Les lois de maximums

La convergence des extrêmes pour les maximum représente l'équivalent de la convergence des moyens: une autre loi de grands nombres. Si la distribution des moyens autour de leur espérance tend à être gaussienne quand la variance est finie, la distribution des extrêmes converge également vers une limite

particulière. Nous présentons maintenant cette distribution de limite. Soient $\{X_1, X_2, \dots, X_n\}$, n variables aléatoires i.i.d. ayant la fdr $F(x) = \Pr(X < x)$. Considérons un échantillon comprenant n réalisations $\{x_1, x_2, \dots, x_n\}$. Nous les rangeons par ordre croissant, et présentons la convention:

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

Le plus grand de ces réalisations $x_{(n)}$ peut être considéré comme la réalisation d'une nouvelle variable aléatoire $X_{(n)}$. La même idée règne pour les autres observations $x_{(k)}$. Ainsi, nous présentons n nouvelles variables aléatoires avec la convention:

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

Ce qui représente le maximum d'un échantillon de taille n qui est la variable aléatoire donnant la plus grande valeur. Dans la littérature, cette quantité est également notée par $X_{1:n} = M_n$. De même:

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

Représente le minimum d'un échantillon de la taille n qui est la variable aléatoire donnant la plus petite valeur observée. Plus généralement $X_{(k)}$ (ou $X_{k:n}$) est la variable aléatoire attachée au $k^{\text{ème}}$ valeur $x_{(k)}$ obtenu parmi des n réalisations. Ces n nouvelles variables aléatoires peuvent être ordonnées de la façon suivante:

$$X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$$

Deux statistiques d'ordre sont particulièrement intéressantes pour l'étude des événements extrêmes: ce sont les deux statistiques extrêmes, autrement dit le minimum et le maximum de l'échantillon .

Nous voulons déterminer le comportement limite du maximum, i.e, nous voulons caractériser la loi de probabilité de maximum. Tous les résultats pour le minimum d'échantillon peuvent être obtenus à partir de ceux du maximum en utilisant la relation suivante

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

La distribution exacte du M_n peut être obtenue à partir de la fdr F . En effet, considérons la fdr de $X_{(k)}$, que nous noterons par $F_{k:n}$ pour indiquer que l'échantillon est de taille n . Nous avons:

$$F_{k:n}(x) = \Pr(X_{(k)} < x) = \sum_{p=k}^n C_n^p [F(x)]^p [1 - F(x)]^{n-p} \quad (1.1)$$

Utilisons le fait que les variables aléatoires X_1, X_2, \dots, X_n sont indépendantes et ont le même fdr. $F(x)$, nous obtenons la fdr du maximum et du minimum pour les valeurs $k = n$ (loi du maximum) et $k = 1$ (loi du minimum):

$$\begin{cases} F_{n:n}(x) = \Pr(X_{(n)} < x) = \sum_{p=n}^n C_n^p [F(x)]^p [1 - F(x)]^{n-p} \\ F_{1:n}(x) = \Pr(X_{(1)} < x) = \sum_{p=1}^n C_n^1 [F(x)]^p [1 - F(x)]^{n-p} \end{cases}$$

de sorte que

$$\begin{cases} F_{n:n}(x) = [F(x)]^n \\ F_{1:n}(x) = 1 - [1 - F(x)]^n \end{cases} \quad (1.2)$$

et leurs fonctions de densité sont respectivement:

$$\begin{cases} f_{n:n}(x) = n[F(x)]^{n-1} f(x) \\ f_{1:n}(x) = n[1 - F(x)]^{n-1} f(x) \end{cases} \quad (1.3)$$

Nous pouvons également obtenir ces expressions directement. Considérons par exemple le maximum. Alors:

$$\{M_n < x\} \Leftrightarrow \{\max(X_1, X_2, \dots, X_n) < x\} \Leftrightarrow \bigcap_{k=1}^n \{X_k < x\}$$

et, en utilisant l'indépendance des variables aléatoires initiales:

$$\Pr(M_n < x) = \Pr\left(\bigcap_{k=1}^n (X_k < x)\right) \Leftrightarrow \prod_{k=1}^n \Pr(X_k < x)$$

en effet

$$F_{n:n}(x) = [F(x)]^n$$

Par conséquent, nous voyons cela, lorsque la fdr F est connu, la fdr exacte de maximum peut être obtenue aisément.

Cependant, la fdr F est inconnue dans la pratique. Il est donc raisonnable d'étudier le comportement asymptotique de M_n pour $n \rightarrow \infty$, avec l'objectif d'approximer la distribution de M_n par une distribution limite non-dégénérée. Notons par $x_F = \sup\{x \in \mathbb{R}, F(x) < 1\}$, le point terminal de la fdr F avec la convention $\sup\{\emptyset\} = \infty$. Il représente la borne supérieure du support de la loi.

D'après la formule (1.2), Nous pouvons conclure quant à la forme de la loi limite de M_n en faisant tendre n vers l'infini. Nous trouvons

$$\lim_{n \rightarrow \infty} F_{n:n}(x) = \lim_{n \rightarrow \infty} [F(x)]^n = \begin{cases} 1 & \text{si } x \geq x_F \\ 0 & \text{si } x < x_F \end{cases} \quad (1.4)$$

Le résultat (1.4) nous indique que la distribution du maximum M_n est une loi dégénérée¹. Ce résultat fournit un intérêt limité sur le comportement de M_n . On aimerait trouver une distribution de quelque intérêt pour le maximum (c'est-à-dire non-dégénérée).

L'idée est d'appliquer une transformation au maximum M_n afin que le passage à la limite conduise à une distribution non dégénérée. Cette issue est bien connue dans le contexte du théorème de limite centrale. Ce dernier est concerné au comportement asymptotique de la moyenne $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ quand $n \rightarrow \infty$:

Théorème 1.2.1.

Considérons une suite de variables aléatoires i.i.d $\{X_1, X_2, \dots, X_n\}$ avec $E(X_i) = \mu$ et $Var(X_i) = \sigma^2 < \infty$. Donc,

$$\lim_{n \rightarrow \infty} \Pr \left(\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \leq x \right) = \Phi(x) \quad (1.5)$$

où Φ est la f.d.r. de la loi gaussienne standard, $b_n = n\mu$ et $a_n = \sigma\sqrt{n}$.

Le théorème de la limite centrale dit que: lorsque n tend vers l'infini la moyenne tend à être distribuée suivant une loi normale si la variance est finie. Si n est assez grand il semble alors pertinent d'utiliser la loi normale pour modéliser la moyenne. La théorie des valeurs extrêmes se situe dans la même logique, mais étudie la queue de la loi au lieu de sa moyenne. Donc, s'intéresse à une loi non-dégénérée pour le maximum de l'échantillon au lieu de la moyenne. Pour cela, nous avons besoin d'un théorème semblable, c'est-à-dire, nous cherchons des suites de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$, tels que

$$\frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} W \sim G, \quad (1.6)$$

avec W est non-dégénérée, i.e

$$\lim_{n \rightarrow \infty} \Pr(M_n < a_n x + b_n) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x), \quad (1.7)$$

pour tout point x où G est continue.

Considérant le maximum de l'échantillon plutôt que la moyenne, également, un problème double surgit.

D'une part, on doit identifier toutes les fonctions de répartitions non-dégénérées possibles qui peuvent apparaître comme limite dedans (1.7).

¹Une loi est dégénérée si elle place toute la masse de probabilité en un point.

D'autre part, on doit caractériser la fdr F (en termes de conditions nécessaires et suffisantes) pour lesquels il existe des suite $\{a_n; n \geq 1\}$ et $\{b_n; n \geq 1\}$, $a_n > 0$ et $b_n \in \mathbb{R}$, tel que (1.7) est satisfaite.

1.3 Problème de la limite des extrêmes

Le premier problème est de déterminer quelles sont les distributions G peut sembler sur la limite (1.7). Afin de fournir la réponse à cette question, nous devons présenter deux concepts importants:

Définition 1.3.1. *Une fdr F est dite max-stable ou appartient au domaine d'attraction du maximum de G (noté par $\mathcal{MDA}(G)$), s'il existe des suite $\{a_n > 0\}$ et $\{b_n\} \in \mathbb{R}$, tel que*

$$\lim_{n \rightarrow \infty} [F(a_n x + b_n)]^n = G(x), \forall x \quad (1.8)$$

pour tout point x où G est continue et on écrit $F \in \mathcal{MDA}(G)$.

La définition ci-dessus peut être redité comme suit: une fdr non-dégénérée G est max-stable **si et seulement si** G^n est du même type avec G pour tout $n \in \mathbb{N}$.

D'une manière équivalente, une fdr F est dite max-stable si pour une suite de variables aléatoires indépendantes $\{X_1, X_2, \dots, X_n\}$ avec fdr F et pour tout $n \in \mathbb{N}$, il existe $a_n > 0$ et $b_n \in \mathbb{R}$ tel que

$$M_n \stackrel{D}{=} a_n X + b_n \quad (1.9)$$

Définition 1.3.2. *(Fonctions de répartition de même type)*

Deux fdr F_1 et F_2 seraient du même type s'il existe des constantes $a > 0$ et $b \in \mathbb{R}$ tels que

$$F_2(ax + b) = F_1(x). \quad (1.10)$$

Ce la signifie que F_1 et F_2 sont identiques, a une translation et changement d'échelle près.

En plus de ces concepts, nous devons appeler le théorème de convergence des types de Khintchine

Théorème 1.3.3. *(Convergence des types, Khintchine)*

Soit F_n et F_n^* deux suites de fdr sur \mathbb{R} , et G une fdr non dégénérée. Soient $a_n > 0$, b_n une suite de constantes telles que

$$\lim_{n \rightarrow \infty} F(a_n x + b_n) \rightarrow G(x)$$

et soient $\alpha_n > 0$, β_n d'autres suites de constantes de normalisation et G^* une fdr elle aussi non dégénérée.

Alors

$$\lim_{n \rightarrow \infty} F(a_n x + b_n) = G^*(x)$$

pour tout point x de continuité de G et G^* .

si et seulement si il existe des constantes $A > 0, B$, telles que

$$\frac{\alpha_n}{a_n} \longrightarrow A, \quad \frac{\beta_n - b_n}{a_n} \longrightarrow B, \quad (1.11)$$

et l'on a alors

$$\forall x \in \mathbb{R}, G^*(x) = G(Ax + B). \quad (1.12)$$

Remarque 1.3.4.

- Ce théorème établit l'égalité en type entre les limites de fdr obtenues par différentes normalisations affines d'une même fdr initiale.
- Le choix des suites a_n et b_n n'est pas unique: si nous choisissons des suites de normalisation a_n et b_n ou α_n et β_n qui sont asymptotiquement équivalentes, i.e. (1.12) est satisfaite, la fdr F appartient au domaine d'attraction du maximum de deux fdr du même type. Ainsi, une fdr F ne peut pas être dans le domaine d'attraction de deux fdr de types différents.

Pour une preuve, voir Resnick [74, chapitre 0], ou voir Leadbetter et al. [58, chapitre 1].

Le théorème de Fisher and Tippett [36, 1928] fournit la forme de toutes les fonctions de distribution asymptotiques possibles du maximum centré réduit d'un échantillon de taille n . Ce résultat fondamental connu sous le nom théorème des types extrêmes, a été trouvé d'abord par Fisher and Tippett [36] et a été complètement prouvé par Gnedenko [42].

Notons que nous restons dans le contexte d'une suite de n variables i.i.d $\{X_1, X_2, \dots, X_n\}$, et ayant la fdr $F(x)$. Ce théorème utilise la notion de fdr max-stable.

Théorème 1.3.5.

Soient n variables aléatoires i.i.d $\{X_1, X_2, \dots, X_n\}$ ayant la fdr F . S'il existe des suites normalisantes $\{a_n\}$ et $\{b_n > 0\}$ telles que, (1.7) est satisfaite. Alors G ne peut appartenir qu'à l'une des trois familles suivantes:

$$\begin{aligned}
\text{Loi de Fréchet } \Phi_\alpha(x) &= \begin{cases} \exp(-x^{-\alpha}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad \alpha > 0 \\
\text{Loi de Weibull } \Psi_\alpha(x) &= \begin{cases} \exp(-(-x)^\alpha) & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases} \quad \alpha > 0 \\
\text{Loi de Gumbel } \Lambda(x) &= \exp(-\exp(-x)) \quad x \in \mathbb{R}, \alpha = 0
\end{aligned}$$

Pour une preuve de ce théorème, le lecteur pourra se référer aux Resnick [74, Proposition 0.3] ou Embrechts et al. [33, p. 152].

La distribution limite G est unique, en ce sens différentes suites de constantes de normalisation déterminent la même distribution.

Les trois distributions ci-dessus - Fréchet, Weibull et Gumbel - sont appelées distributions standards de valeurs extrêmes. Ce sont des fonctions continues pour tout réel x . Les variables aléatoires distribuées selon l'une des distributions de valeurs extrêmes sont appelées variables aléatoires extrémales.

Ce théorème montre que le maximum convenablement normalisé pour n'importe quelle distribution converge exactement en distribution à une des trois distributions de valeurs extrêmes: Gumbel, Fréchet, et Weibull.

Les trois classes des distributions notées par Φ_α , Ψ_α et Λ s'appellent **les distributions standards de valeurs extrêmes**.

Les constants de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$ s'interprètent comme des paramètres d'échelle et de localisation, tandis que α est le paramètre de forme. Ces trois distributions sont max-stables. Ceci signifie que pour un échantillon des variables aléatoires indépendantes $\{X_1, X_2, \dots, X_n\}$ satisfaisant une de ces distributions, et pour M_n le maximum de ces variables l'expression (1.9) est satisfaite. Plus précisément:

- Si X suit la loi de **Fréchet** de paramètre α alors : $M_n \stackrel{\mathcal{D}}{=} n^{\frac{1}{\alpha}} X$
- Si X suit la loi de **Weibull** de paramètre α alors : $M_n \stackrel{\mathcal{D}}{=} n^{-\frac{1}{\alpha}} X$
- Si X suit la loi de **Gumbel** alors: $M_n \stackrel{\mathcal{D}}{=} X + \ln n$

Il faut bien signaler que chacune des trois lois des valeurs extrêmes peut s'obtenir par une transformation fonctionnelle de l'autre.

Proposition 1.3.6 (Embrechts et al. [33],1997).

Pour tout $\alpha > 0$ et toute variable aléatoire X , il y a équivalence entre:

- X suit une loi de **Fréchet** de paramètre α ;
- $-\frac{1}{X}$ suit une loi de **Weibull** de paramètre α ;

Loi de grands nombres	
Théorème de la limite centrale	Théorème de valeurs extrêmes
Convergence de sommes	Convergence de maxima
$S_n = \sum_{i=1}^n X_i$	$M_n = \max\{X_1, X_2, \dots, X_n\}$
$\lim_{n \rightarrow \infty} \Pr\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq x\right) = \Phi(x)$	$\lim_{n \rightarrow \infty} \Pr(M_n < a_n x + b_n)$
Distribution Gaussienne centrée réduite si μ et σ^2 existes	Distribution de valeurs extrêmes standardes
Densité Gaussienne: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-\frac{t^2}{2}) dt$	Fréchet: $\Phi_\gamma(x) = \exp(-x^{-\frac{1}{\gamma}})$ Weibull: $\Psi_\gamma(x) = \exp(-(-x)^{\frac{1}{\gamma}})$ Gumbel: $\Lambda(x) = \exp(-e^{-x})$
Moivre (1730)-Laplace (1812)	Fisher-Tippett (1928)
Fluctuations de sommes	Fluctuations de maxima

Tableau 1.1: Lois limites des moyennes et des maxima

- $\ln(X^\alpha)$ suit la loi de **Gumbel**.

De ce proposition les lois de **Fréchet** et de **Weibull** sont liées par la relation:

$$\Phi_\alpha\left(-\frac{1}{x}\right) = \Psi_\alpha(x) \tag{1.13}$$

Comparons au problème de la limite centrale, ce problème de la limite extrême n'est pas résolu par la distribution normale standard, mais n'importe quelle fdr non-dégénérée de limite dans (1.7) appartient à une des trois familles possibles de distribution de valeurs extrêmes à savoir: Gumbel, Fréchet ou Weibull.

Il y a une analogie formelle entre le Théorème 1.3.5 pour le maximum et le Théorème 1.2.1 (Théorème de limite centrale pour la moyenne). Cette comparaison entre les deux théorèmes, et les deux formules (1.7) et (1.5), nous permet d'interpréter les paramètres a_n et b_n . Le paramètre b_n correspond à la moyenne μ dans le théorème de limite centrale et est donc un paramètre de position. Le paramètre a_n correspond à σ/\sqrt{n} dans le théorème de la limite centrale et est donc un paramètre d'échelle. Cependant, le paramètre α n'a aucun équivalent dans le théorème de limite centrale. Les éléments de cette analogie sont regroupés dans le tableau 1.1.

1.4 La Distribution Généralisée de Jenkinson-von Mises

Les trois types des distributions de Fréchet, de Weibull, et de Gumbel peuvent être regroupées dans une expression générale en utilisant la représentation

suivante:

$$G_\gamma(x) = \begin{cases} \exp\left(- (1 + \gamma x)^{-\frac{1}{\gamma}}\right), & \text{si } \gamma \neq 0 \\ \exp(-\exp(-x)), & \text{si } \gamma = 0 \end{cases} \quad (1.14)$$

et est définie pour x tel que $1 + \gamma x > 0$.

Le paramètre $\gamma \in \mathbb{R}$ qui apparaît dans l'expression générale indique la forme de la fonction de répartition (le cas $\gamma = 0$ s'obtient en considérant la limite de G_γ quand $\gamma \rightarrow 0$). La figure 1.1 montre la fdr G_γ pour différentes valeurs de γ . Si $\gamma < 0$ la loi G_γ est bornée à droite, si $\gamma > 0$ elle l'est à gauche.

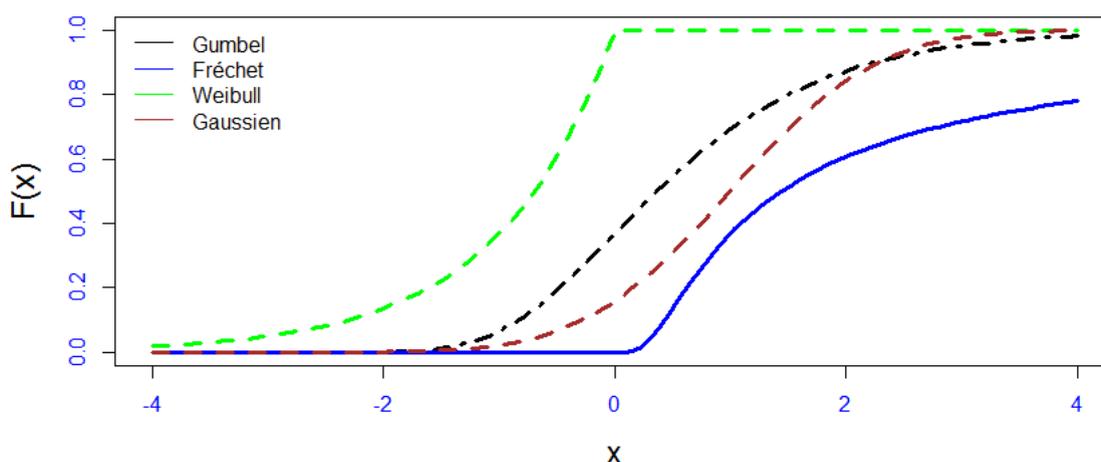


Fig 1.1: F.d.r $F(x)$ des trois types d'extrêmes, respectivement $(\gamma = 1, \mu = 1, \sigma = 1)$, $(\gamma = -1, \mu = -1, \sigma = 1)$, $(\gamma = 0, \mu = 0, \sigma = 1)$; en comparaison avec la fonction de survie gaussienne de même moyenne et de variance que celle de Gumbel.

Le paramètre γ est appelé l'**indice des valeurs extrêmes** ou l'**indice de queue**. Au signe près, c'est l'inverse du paramètre de forme, appelé aussi α , qui apparaît dans les lois de Fréchet et de Weibull. Il contrôle le comportement de la queue de la fdr F , i.e. la fréquence et la gravité des événements rares. Notons que la queue épaisse lorsque γ augmente, donc γ peut être vu comme une jauge de lourdeur de la queue de la fdr sous-jacent. Le signe de γ a une forte influence sur la distribution des extrêmes, et on distingue trois cas:

- $\gamma > 0$, c'est la loi de Fréchet. la fonction de queue est à décroissance polynomiale i.e. La queue de distribution est lourde et $\gamma = \frac{1}{\alpha} > 0$. Plus γ est grand, plus la queue de distribution est lourde (épaisse).

- $\gamma < 0$, c'est la loi de Weibull. La queue de distribution est bornée et $\gamma = -\alpha^{-1} < 0$.
- $\gamma = 0$, c'est la loi de Gumbel. La fonction de queue est à décroît exponentielle i.e. la queue est légère. C'est un cas de limite, entre les deux situations précédentes, où $\gamma \rightarrow 0$. La figure 1.2 illustre ces trois types.

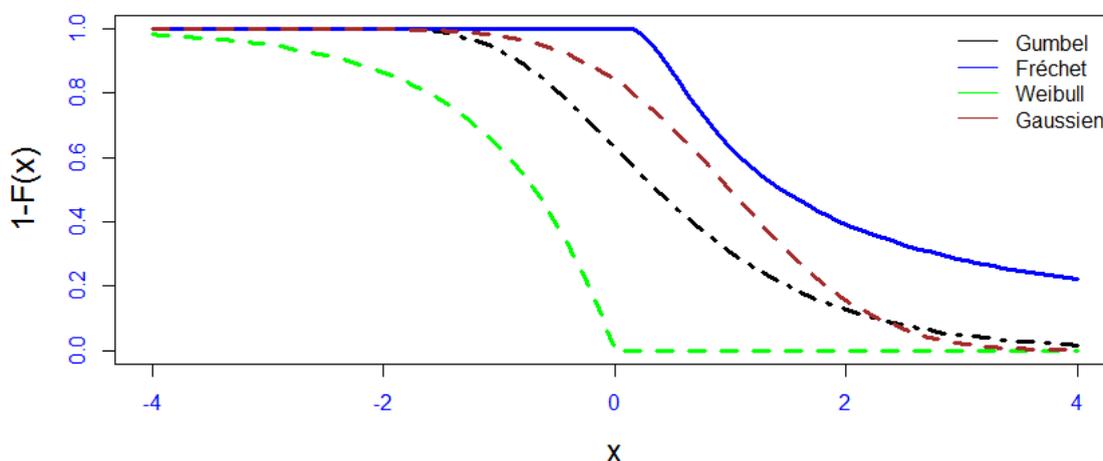


Fig 1.2: Fonctions de queues des trois types d'extrêmes, respectivement $(\gamma = 1, \mu = 1, \sigma = 1)$, $(\gamma = -1, \mu = -1, \sigma = 1)$, $(\gamma = 0, \mu = 0, \sigma = 1)$; en comparaison avec la densité gaussienne de même moyenne et de variance que cel de Gumbel.

La paramétrisation (1.14) est due à von Mises [85] (1936) et à Jenkinson [50] (1955) et est s'appelle **distribution de valeurs extrêmes généralisée standard** (GEVD), elle englobe les trois types de distributions extrêmes Fréchet, Weibull et Gumbel. En effet, on a la proposition suivante:

Proposition 1.4.1. (de Haan and Ferreira [21])

Soient $G_\gamma (\gamma \in \mathbb{R})$ la distribution de valeurs extrêmes généralisée standard et $\Phi_\alpha, \Psi_\alpha, \Lambda$ les distribution des valeurs extrêmes standards avec $\alpha > 0$. On a

$$\begin{cases} \gamma > 0 : G_\gamma \left(\frac{1}{\gamma} (x - 1) \right) = \Phi_{\frac{1}{\gamma}}(x) \mathbf{1}_{(x \geq 0)} \\ \gamma < 0 : G_\gamma \left(-\frac{1}{\gamma} (x + 1) \right) = \Psi_{-\frac{1}{\gamma}}(x) \mathbf{1}_{(x < 0)} + \mathbf{1}_{(x \geq 0)} \\ \gamma = 0 : G_0(x) = \Lambda(x) \mathbf{1}_{\mathbb{R}}. \end{cases}$$

où $\mathbf{1}_A$ est la fonction indicatrice de l'ensemble A .

Ainsi, pour $H \in \{\Phi_\alpha, \Psi_\alpha, \Lambda\}$, il existe γ tel que H et G_γ soient des mêmes type. Par conséquent, les deux représentations des distributions de valeurs

extrêmes ainsi que leurs types caractérisent le même ensemble de distributions, ainsi toutes les deux sont équivalentes. On peut obtenir une version plus générale de G_γ , en introduisant un paramètre de localisation $\mu \in \mathbb{R}$ et un paramètre d'échelle $\sigma > 0$ (Weissman [90, 1978]):

$$G_{\gamma,\mu,\sigma}(x) = G_\gamma\left(\frac{x-\mu}{\sigma}\right) = \begin{cases} \exp\left(-\left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\gamma}}\right]\right), & 1 + \gamma\frac{x-\mu}{\sigma} > 0 \quad \text{si } \gamma \neq 0 \\ \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right), & x \in \mathbb{R} \quad \text{si } \gamma = 0 \end{cases} \quad (1.15)$$

Nous pouvons interpréter les paramètres μ et σ d'une manière intéressante, comparant les équations (1.7) et (1.15):

$$\lim_{n \rightarrow \infty} \Pr\left(\frac{M_n - b_n}{a_n} \leq x\right) = \exp\left(-\left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\gamma}}\right]\right)$$

Cette égalité prouve que les paramètres μ et σ sont les limites des constants de normalisations $b_n \in \mathbb{R}$ et $a_n > 0$ quand n est grand. Les fonctions de densités (dérivé de la f.d.r. $G_\gamma(x)$), g_γ , correspondantes s'écrivent:

$$g_\gamma(x) = \begin{cases} \frac{1}{\sigma} \left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1+\gamma}{\gamma}}\right] \exp\left(-\left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\gamma}}\right]\right) & 1 + \gamma\frac{x-\mu}{\sigma} > 0 \quad \text{si } \gamma = 0 \\ \frac{1}{\sigma} \exp\left(\frac{x-\mu}{\sigma}\right) \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) & x \in \mathbb{R} \quad \text{si } \gamma \neq 0 \end{cases}$$

La figure 1.3 représente quelques densités pour γ fixé en comparaison avec la densité gaussienne.

Les figures 1.2 et 1.3 expliquent pourquoi les fonctions de répartition dans le domaine d'attraction de Fréchet sont habituellement dites à queues lourdes, alors que ceux du domaine d'attraction de Gumbel sont dites à queues légères et de domaine d'attraction de Weibull sont dites à queues bornées.

1.5 Problème du domaine d'attraction du maximum

Après que nous avons identifié la forme générale de toutes les lois limites possibles, nous devons résoudre la deuxième problème, qui est appelé le problème du domaine d'attraction. Ceci peut être décrit de la manière suivante:

On suppose que G est une distribution limite possible pour la suite $\left(\frac{M_n - b_n}{a_n}\right)$, Quelles sont les conditions nécessaires et/ou suffisantes sur la fdr F de X pour obtenir précisément cette fdr limite G ? L'ensemble de ces distributions constituent le domaine d'attraction de G . Ces conditions étant basées sur la notion de fonctions à variations régulières et lentes.

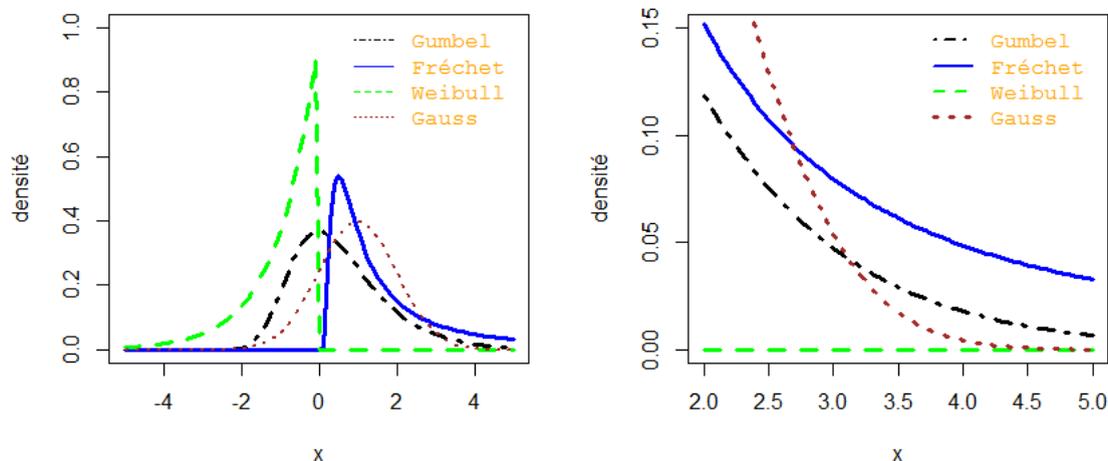


Fig 1.3: Densités des trois types d'extrêmes, respectivement $(\gamma = 1, \mu = 1, \sigma = 1)$, $(\gamma = -1, \mu = -1, \sigma = 1)$, $(\gamma = 0, \mu = 0, \sigma = 1)$; en comparaison avec la densité gaussienne de même moyenne et de variance que celle de Gumbel. Le panneau de droite est un zoom sur la queue.

1.5.1 Fonctions à variation régulière et leurs propriétés principales

La théorie de fonctions à variation régulière est un outil analytique essentiel pour traiter les domaines d'attraction (jouent un rôle important dans l'exposition suivante). Dans cette section, nous présenterons les propriétés de base des fonctions susmentionnées qui seront utilisées dans ce qui suit. En général, les fonctions à variation régulière sont des fonctions qui se comportent asymptotiquement comme des fonctions de puissance.

1.5.1.1 Propriétés générales

Définition 1.5.1. (*Fonctions à variation régulière*)

Une fonction mesurable $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est à variation régulière à l'infini (écrit $h \in RV_\rho$) si

$$\forall t > 0, \exists \lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} \in \mathbb{R}_+$$

Dans ce cas $\exists \rho \in \mathbb{R}$ tel que

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = t^\rho (t > 0)$$

Le nombre ρ est appelé l'indice de variation régulière (de h) et h est dite à variation régulière à l'infini d'indice ρ .

Evidemment $x \mapsto h(x)$ est variation régulière à 0 si et seulement si $x \mapsto h(\frac{1}{x})$ variation régulière à ∞ .

Remarque 1.5.2. *Le concept peut être étendu à la variation régulière à tout point x_0 autres que l'infini (ou l'origine) on remplace $h(x)$ par $h(x_0 - \frac{1}{x})$ dans la définition ci-dessus. Il est donc évident qu'il suffit d'étudier la théorie de la variation régulière à l'infini.*

Définition 1.5.3. *(Fonctions à variation lente)*

Une fonction mesurable $\ell : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est à variation lente à ∞ (s'écrit $\ell \in RV_0$) si elle est à variation régulière avec indice $\rho = 0$.

La connexion entre les fonctions à variation lente et régulière est fournie par le théorème de caractérisation suivant:

Proposition 1.5.4. *(Théorème de caractérisation)*

Si

$$h(tx) \sim g(t)h(x) \text{ quand } x \rightarrow \infty \quad (1.16)$$

1. (1.16) est satisfaite pour tout $t > 0$,
2. Il existe $\rho \in \mathbb{R}$ tel que $g(t) = t^\rho \forall t > 0$, et
3. $h(x) = x^{-\rho}\ell(x)$ où $\ell(x)$ est une fonction à variation lente à l'infini.

Du ce résultat, il est clair que pour étudier les variations régulières, il suffit d'étudier les propriétés des fonctions à variation lente. Des exemples de telles fonctions sont:

- Les constantes positives et toute fonction qui a une limite strictement positive et finie, donc la classe des fonctions à variation lentes (et donc régulièrement) est clairement non vide à cause de cet exemple trivial.
- La fonction logarithme ($\ell : x \mapsto \log x$).
- Les fonctions ($\ell : x \mapsto \prod_{k=1}^n (\log_k x)^{\xi_k}$, $\xi_k \in \mathbb{R}$), $\xi_k \in \mathbb{R}$ et \log_k est le k .ième itération de logarithme.
- La class de fonctions de Hall définie par:

$$\exists M > 0, \forall x \geq M, \ell(x) = a + by^{-c}(1 + o(1)),$$

avec $(a, c) \in \mathbb{R}_+^2$ et $b \in \mathbb{R}$.

1.5.1.2 Propriétés de base

Théorème 1.5.5. (*Représentation de Karamata*)

Toute fonction à variations lentes ℓ s'écrit sous la forme

$$\ell(x) = c(x) \exp \left\{ \int_1^x \frac{\varepsilon(t)}{t} dt \right\}$$

où $c > 0$ et ε sont deux fonctions mesurables telles que

$$\begin{aligned} \lim_{x \rightarrow \infty} c(x) &= c \neq 0 \\ \lim_{x \rightarrow \infty} \varepsilon(x) &= 0 \end{aligned}$$

Dans le cas où la fonction $c(\cdot)$ est constante, la fonction $\ell(\cdot)$ correspondante est dite normalisée. Si la fonction ℓ est normalisée alors elle est dérivable de dérivée ℓ' définie pour tout x par

$$\ell'(x) = \frac{\varepsilon(x)\ell(x)}{x}$$

Proposition 1.5.6. (*Bingham et al. [8, prop. 1.3.6]*)

Si ℓ, ℓ_1, ℓ_2 varient lentement, il en est de même pour

1. $\ell^\beta(x) \forall \beta \in \mathbb{R}$, $\ell_1(x) + \ell_2(x)$, $\ell_1(x)\ell_2(x)$ et $\ell_1 \circ \ell_2(x)$ (si $\ell_2(x) \rightarrow \infty$ quand $x \rightarrow \infty$)
2. $\forall \varepsilon > 0$ et $x \rightarrow \infty$

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

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

Remarque 1.5.7. C'est-à-dire que ℓ est "petit" par rapport à toute fonction de puissance, x^ε .

Remarque 1.5.8. Appliquant le théorème de représentation, la représentation d'une fonction h variant régulièrement avec index $\rho \in \mathbb{R}$ est obtenue:

$$\begin{aligned} h(x) &= x^\rho \ell(x) = c(x) x^\rho \exp \left\{ \int_1^x \frac{\varepsilon(t)}{t} dt \right\} = c(x) \exp \left\{ \rho \ln x + \int_1^x \frac{\varepsilon(t)}{t} dt \right\} \\ &= c(x) \exp \left\{ \int_1^x \frac{1}{t} (\rho + \varepsilon(t)) dt \right\} = c(x) \exp \left\{ \int_1^x \frac{\rho(t)}{t} dt \right\}, \forall x > 0, \end{aligned}$$

où $x \mapsto c(x)$ et $x \mapsto \rho(x)$ satisfont

$$\begin{aligned} \lim_{x \rightarrow \infty} c(x) &= c \neq 0 \\ \lim_{x \rightarrow \infty} \rho(x) &= \rho \end{aligned}$$

Proposition 1.5.9. (de Haan and Ferreira [21], Bingham et al. [8, prop. 1.5.7])

Si h, h_1, h_2 varient régulièrement à l'infini avec indices ρ, ρ_1, ρ_2 respectivement, alors:

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

2. Si $\rho \neq 0$

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

3. $h^\beta \in RV_{\beta\rho}, \beta \in \mathbb{R}$.

4. $h_1 + h_2 \in RV_\rho$ pour $\rho = \max(\rho_1, \rho_2)$.

5. $h_1 h_2 \in RV_{\rho_1 + \rho_2}$.

6. Si, de plus, $h_2 \rightarrow \infty$, alors $(h_1 \circ h_2)(x) = h_1(h_2(x)) \in RV_{\rho_1 \rho_2}$.

Les dérivés et les intégrales de fonctions à variations régulières varient également régulièrement, sous certaines conditions de régularité. Grosso modo, l'indice de variation régulière augmente de 1 lors de l'intégration, il diminue de 1 lors de la différenciation. Les résultats précis sont désignés sous le nom de théorème de Karamata et le théorème de densité monotone, respectivement.

Proposition 1.5.10. (Théorème de Karamata) (Bingham et al. [8, prop. 1.5.8 et 1.5.10])

Si ℓ est à variation lente ($\ell \in RV_0$), alors:

1. Pour $\rho > -1$:

$$\int_{x_0}^x x^\rho \ell(t) dt \sim (\rho + 1)^{-1} x^{\rho+1} \ell(x), \quad x \rightarrow \infty,$$

2. Pour $\rho < -1$:

$$\int_x^{+\infty} x^\rho \ell(t) dt \sim -(\rho + 1)^{-1} x^{\rho+1} \ell(x), \quad x \rightarrow \infty.$$

Proposition 1.5.11. (Théorème de densité monotone) (Bingham et al. [8, Th.1.7.2])

Soit

$$H(x) = \int_{x_0}^x h(t) dt \text{ ou } H(x) = \int_x^{+\infty} h(t) dt$$

où h est ultimement monotone². Si

$$H(x) \sim cx^\rho \ell(x), \quad x \rightarrow \infty$$

avec $c > 0, \rho \in \mathbb{R}$ et $\ell \in RV_0$ alors

$$h(x) \sim c\rho x^{\rho-1} \ell(x), \quad x \rightarrow \infty \quad (1.17)$$

Remarque 1.5.12. Si $c > 0$: nos hypothèses dire d'une part que $H \in RV_\rho$. Cependant, (1.17) n'implique pas que $h \in RV_{\rho-1}$, sauf dans le cas où $c\rho > 0$.

Pour plus d'approfondissements de la théorie des fonctions à variations régulières voir Seneta [77], Bingham et al. [8], Resnick [74, Section 0.4] et de Haan and Ferreira [21, Appendix B].

1.5.2 Caractérisation des trois domaines d'attraction

Pour chaque loi d'attraction, on rappelle la fonction de répartition de la loi limite G et on donne les conditions nécessaires et/ou suffisantes d'appartenance à son domaine d'attraction ainsi que des valeurs des paramètres a_n et b_n . Nous présenterons aussi quelques exemples de lois appartenant à chaque domaine d'attraction.

Définissons d'abord quelques concepts, la fonction de quantile et la fonction quantile de queue. Étant donné une fdr F , la fonction $\bar{F}(x) = 1 - F(x)$ est la queue de la fdr $F(x)$ et est appelée fonction de survie ou fonction de queue.

Définition 1.5.13. La fonction quantile de F , s'écrit $Q(1-s) = F^{\leftarrow}(1-s)$, est définie comme suit:

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

La quantité $x_{1-s} = Q(1-s) = F^{\leftarrow}(1-s)$ alors défini le $(1-s)$ -quantile de F .

Définition 1.5.14. La fonction quantile de queue U est définie comme suit:

$$U(t) = Q\left(1 - \frac{1}{t}\right), \quad t \in [1, \infty[$$

où, on dit que $x_{1-s} = Q(1-s) = F^{\leftarrow}(1-s) = U(1/s)$ est un quantile extrême quand s tend vers 0 ($s \downarrow 0$).

² une fonction h à support positif est ultimement monotone s'il existe un x_0 pour lequel elle est monotone sur l'intervalle $[x_0; +\infty[$

1.5.2.1 Domaine d'attraction de Fréchet Φ_γ ($\gamma > 0$)

Si une variable aléatoire est de variation régulière, alors elle appartient au \mathcal{MDA} de Fréchet; si cette variable aléatoire est de variation rapide, alors elle appartient au \mathcal{MDA} de Gumbel. Nous examinons maintenant ces deux cas. Rappelons que la distribution de valeurs extrêmes standard de Fréchet est définie comme suit:

$$\Phi_\gamma(x) = \begin{cases} \exp(-x^{-\alpha}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad \alpha > 0$$

Les lois appartenant à ce domaine d'attraction sont caractérisées par une queue à décroissance lente (polynômiale) à l'infini, et un point terminal $x_F = +\infty$. elles sont dites aussi distributions à queue lourde. Une caractérisation de ce domaine d'attraction noté $\mathcal{MDA}(\Phi_\gamma)$ est donnée par la condition nécessaire et suffisante suivante:

Théorème 1.5.15. (*Resnick [74]*)

La fdr F appartient au domaine d'attraction de Fréchet (Φ_α) avec $\gamma > 0$ **si et seulement si** $x_F = +\infty$ et la fonction de queue $\bar{F}(x)$ est à variation régulière d'indice $-\frac{1}{\gamma}$ qui s'écrit $\bar{F} \in RV_{-1/\gamma}$ i.e.

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\frac{1}{\gamma}}$$

Autrement dit, une fdr F appartenant au domaine d'attraction de Fréchet s'écrit sous la forme:

$$F(x) = x^{-\frac{1}{\gamma}} l(x), \quad \text{où } l \in RV_0$$

Proposition 1.5.16 (de Haan and Ferreira [21]).

Si $F \in \mathcal{MDA}(\Phi_\gamma)$, alors avec les constantes de normalisation $a_n = U(n) = F^{-1}\left(1 - \frac{1}{n}\right)$ et $b_n = 0$ la suite $(a_n^{-1}M_n, n \geq 1)$ converge en loi vers une variable aléatoire de f.d.r Φ_α i.e.

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

La condition dans le théorème 1.5.15 peut s'écrire en termes de fonction quantile et quantile de queue.

Proposition 1.5.17. (*de Haan and Ferreira [21]*)

les assertions suivantes sont équivalentes:

1. F est à queue lourde

$$F \in \mathcal{MDA}(\Phi_\gamma)$$

2. F est une fonction à variations régulières à l'infini d'indice $-\frac{1}{\gamma}$

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\frac{1}{\gamma}}$$

3. $Q(1 - s)$ est une fonction à variations régulières à zéro d'indice $-\gamma$

$$\lim_{t \rightarrow 0} \frac{Q(1 - tx)}{Q(1 - t)} = x^{-\gamma}$$

4. U est une fonction à variations régulières à l'infini d'indice γ

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\gamma$$

Les conditions mentionnées dans la proposition précédente sont appelées conditions de variation régulière de premier ordre. De plus, dans l'analyse du comportement des estimateurs dans le domaine de la théorie de valeurs extrêmes, on utilise fréquemment des propriétés de fonctions à variation régulière plus fortes. Nous définissons l'hypothèse de second ordre qui renforce la condition de variation régulière de premier ordre sur F , dans le Chapitre 2.

Remarque 1.5.18. Afin de vérifier le théorème de Fisher-Tippet on calcule F^n . On a dans ce cas $a_n = F^{-1}\left(1 - \frac{1}{n}\right) = n^{\frac{1}{\alpha}}$. Ceci conduit à

$$\lim_{n \rightarrow \infty} [F(a_n x + b_n)]^n = \lim_{n \rightarrow \infty} [F(n^{\frac{1}{\alpha}} x)]^n = \lim_{n \rightarrow \infty} \left(1 - \frac{x^{-\alpha}}{n}\right)^n$$

alors

$$\lim_{n \rightarrow \infty} [F(a_n x + b_n)]^n = \exp(-x^{-\alpha}) = \Phi_\alpha(x)$$

Une condition suffisante bien connu peut être donné en termes de fonction de hasard

$$h(x) = \frac{F'(x)}{\overline{F}(x)}$$

Proposition 1.5.19. (Condition de von Mises)

Supposons $x_F = \infty$ et $f(x)$ exist. si

$$\lim_{x \rightarrow \infty} xh(x) = \lim_{x \rightarrow \infty} \frac{xf(x)}{\overline{F}(x)} = \frac{1}{\gamma}$$

alors $F \in \mathcal{MDA}(\Phi_\gamma)$

Le tableau 1.2 présente des exemples de lois qui appartiennent au ce domaine d'attraction. Embrechts et al. [33] donnent les suites $a_n > 0$ et b_n qui assurent la convergence en distribution dans le Théorème 1.3.5 pour la loi correspondante.

Distributions	$1 - F(x)$	Indice de queue γ	$\ell(x)$
Pareto(α) $\alpha > 0$	$x^{-\alpha}$	$\frac{1}{\alpha}$	1
Burr(η, τ, λ) $\eta, \tau, \lambda > 0$	$\left(\frac{\eta}{\eta+x^\tau}\right)^\lambda x > 0$	$\frac{1}{\lambda\tau}$	$\left(\frac{\eta}{1+\frac{\eta}{x^\tau}}\right)^\lambda$
Pareto – Généralisé(σ, ξ) $\sigma, \xi > 0$	$\left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}} x > 0$	ξ	$\left(\frac{\sigma}{\xi}\right)^{\frac{1}{\xi}} \left(1 + \frac{\sigma}{\xi x}\right)^{-\frac{1}{\xi}}$
Fréchet(α) $\alpha > 0$	$1 - \exp(-x^{-\alpha}) x > 0$	$\frac{1}{\alpha}$	$1 - \frac{x^{-\alpha}}{2} + o(x^{-\alpha})$

Tableau 1.2: Exemples des lois qui appartiennent au $\mathcal{MDA}(\Phi_\gamma)$

1.5.2.2 Domaine d’attraction de Weibull ($\mathcal{MDA}(\Psi_\gamma)$, $\gamma < 0$)

Rappelons que la distribution de valeurs extrêmes standard de Weibull est définie comme suit:

$$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha) & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases} \quad \alpha > 0$$

Théorème 1.5.20 (Resnick [74]).

Une fdr F appartient au domaine d’attraction de la loi de Weibull Ψ avec $\gamma < 0$ **si et seulement si** $x_F < \infty$ et $F_*(x) = F(x_F - x^{-1}) \in \mathcal{MDA}(\Phi_{(-\gamma)})$ c’est-à-dire que F_* est une fonction à variation régulière d’indice $\frac{1}{\gamma}$ à l’infini et on note ($F_* \in RV_{\frac{1}{\gamma}}$).

Par conséquence, une fdr F du domaine d’attraction de Weibull **si et seulement si** $x_F < \infty$ et $F(x) = 1 - (x_F - x)^{-\frac{1}{\gamma}} l((x_F - x)^{-\frac{1}{\gamma}})$, où la fonction l est à variation lente.

Proposition 1.5.21 (de Haan and Ferreira [21]).

Si $F \in \mathcal{MDA}(\Psi_\gamma)$, alors avec les suites de normalisation $a_n = x_F - U(n) = x_F - F^{-1}\left(1 - \frac{1}{n}\right)$ et $b_n = x_F$, la suite $(a_n^{-1}(M_n - x_F), n \geq 1)$ converge en loi vers une variable aléatoire de fonction de répartition Ψ_α i.e.

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \Psi_\gamma(x), \forall x \in \mathbb{R}$$

Proposition 1.5.22. (Conditions de von Mises) Supposons $x_F < \infty$ et $f(x)$ existe pour $x < x_F$. Si

$$\lim_{x \rightarrow x_F} \frac{(x_F - x) f(x)}{1 - F(x)} = -\frac{1}{\gamma}$$

alors $F \in \mathcal{MDA}(\Psi_\gamma)$

Le tableau 1.3 présente des exemples de lois qui appartiennent au ce domaine d’attraction. Embrechts et al. [33] donnent les suites $a_n > 0$ et b_n qui assurent la convergence en distribution dans le théorème 1.3.5 pour la loi correspondante.

Distributions	$1 - F(x_F - \frac{1}{x})$	Indice de queue γ	$\ell(x)$
Weibull	$1 - \exp(-x^{-\alpha})$	$-\frac{1}{\alpha}$	$1 - \frac{x^{-\alpha}}{2} + o(x^{-\alpha})$
Uniforme	$\frac{1}{x}, x > 1$	-1	1
ReverseBurr (λ, β, τ)	$\left(\frac{\beta}{\beta+x^\tau}\right)^\lambda, x > 0$	$-\frac{1}{\lambda\tau}$	$\beta^\lambda (1 - \lambda\beta x^{-\tau} + o(x^{-\alpha}))$

Tableau 1.3: Exemples des lois qui appartiennent au $\mathcal{MDA}(\Psi_\gamma)$

1.5.2.3 Domaine d'attraction de Gumbel ($\mathcal{MDA}(\Lambda), \gamma = 0$)

Rappelons que la distribution de valeurs extrêmes standard de Gumbel est définie comme suit:

$$\Lambda(x) = \exp(-\exp(-x)) \quad x \in \mathbb{R}, \alpha = 0$$

Le cas de ce domaine est plus délicat à traiter car il n'y a pas de lien entre la queue de distribution et les fonctions à variations lentes comme c'est le cas pour les deux autres domaines d'attraction. Dans la littérature il n'y a pas de conditions nécessaires et suffisantes faciles à mettre en oeuvre. Cependant Von Mises (1936) a donné une caractérisation plus simple, illustrée dans le théorème suivant Resnick [74, Proposition 1.4].

Définition 1.5.23. (*Fonction de Von Mises*) Une fdr $F_\#$ avec point terminal x_F est appelée une fonction de Von Mises s'il a la représentation suivante: Supposons il existe des $z_o < x_F$ tel que, pour $z_o < x < x_F$ et $c > 0$

$$1 - F_\#(x) = c \exp \left\{ - \int_{z_o}^x \frac{1}{\varphi(t)} dt \right\}$$

où $\varphi(\cdot)$ est une fonction positive et absolument continue sur (z_o, x_F) avec densité $\varphi'(x)$ et $\lim_{x \rightarrow x_F} \varphi'(x) = 0$. La fonction φ est appelée fonction auxiliaire de $F_\#$.

Théorème 1.5.24. Une fdr F appartient au domaine d'attraction de Gumbel **si et seulement** s'il existe une fonction de Von Mises $F_\#$ telle que pour $x \in (z_o, x_F)$

$$1 - F(x) = c(x) (1 - F_\#(x)) = c(x) \exp \left\{ - \int_{z_o}^x \frac{1}{\varphi(t)} dt \right\}$$

et $\lim_{x \rightarrow x_F} c(x) = c > 0$

Proposition 1.5.25 (de Haan and Ferreira [21]).

Si $F \in \mathcal{MDA}(\Lambda)$, alors avec les suites de normalisation $a_n = \varphi(b_n)$ et

$b_n = U(n)$, la suite $(a_n^{-1}(M_n - U(n)), n \geq 1)$ converge en loi vers une variable aléatoire de $\text{fdr}\Lambda$ i.e

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

Proposition 1.5.26. (Condition de Von Mises)

Soit F une fonction de distribution et x_F son point terminal. Supposons que $F''(x)$ existe et $F'(x)$ est positive au voisinage de x_F . Si

$$\lim_{x \rightarrow x_F} \left(\frac{1 - F(x)}{F'(x)} \right)' = 0$$

ou

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

alors, $F \in \mathcal{MDA}(\Lambda)$

Le tableau 1.4 présente des exemples de lois qui appartiennent au ce domaine d'attraction. Embrechts et al. [33] donnent les suites $\{a_n > 0\}$ et $\{b_n\}$ qui assurent la convergence en distribution dans le théorème 1.3.5 (page) pour la loi correspondante.

Distributions	Fonction de queue
Gumbel	$\exp(-x), x \in \mathbb{R}$
Logistic	$\frac{1}{1 + \exp(x)}$
Exp $(\lambda) \lambda > 0$	$\exp(-\lambda x), x \geq 0, \lambda > 0$
$N(0, 1)$	$\frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt, x \in \mathbb{R}$

Tableau 1.4: Exemples des lois qui appartiennent au $\mathcal{MDA}(\Lambda)$.

Concluons sur le comportement des queues de distribution:

- Le demaine d'attraction de la loi de Fréchet $\mathcal{MDA}(\Phi_\gamma)$ contient toutes les distributions caractérisées par une queue à décroissance polynomiale à l'infini, et un point terminal x_F infini. Elle sont dites aussi à queues lourdes.
- Le demaine d'attraction de la loi de Gumbel $\mathcal{MDA}(\Lambda)$ contient toutes les distributions caractérisées par une queue à décroissance exponentielle à l'infini. Les exemples de telles distributions sont les distributions exponentielles avec les queues peu lourdes (légères) et plus épaisses que gaussiens.
- Le demaine d'attraction de la loi de Weibull $\mathcal{MDA}(\Psi_\gamma)$ contient toutes les distributions avec des supports bornés à droite i.e. $x_F < +\infty$. Elles sont dites aussi à queues bornées.

1.5.3 Caractérisation pour différents domaines d'attraction

Une caractérisation de $\mathcal{MDA}(G_\gamma)$ pour le cas général $\gamma \in \mathbb{R}$ est fournie par le théorème suivant qui est l'un des résultats de base dans la théorie des valeurs extrêmes. De manière concise et analytique, il donne les informations essentielles recueillies dans la section précédente sur les domaines d'attraction maximum.

Théorème 1.5.27. (de Haan and Ferreira [21], Théorème 1.1.6).

Soit $\gamma \in \mathbb{R}$ et considérons la fdr G_γ dans (1.14). Les assertions suivantes sont équivalentes.

- Il existe deux constantes réelles $a_n > 0$ et b_n , telles que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_\gamma(x), \quad (1.18)$$

Pour chaque point de continuité x de G_γ . Rappelons que, si cette relation se vérifie, on dit que la fdr F appartient au max-domaine d'attraction de G_γ .

- Il existe une fonction positive a telle que, pour $x > 0$

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = D_\gamma(x) = \frac{x^\gamma - 1}{\gamma} \quad (1.19)$$

où, Pour $\gamma = 0$, le membre de droite s'interprète comme sa limite lorsque $\gamma \rightarrow 0$. qu'est égale à $\log x$.

- Il existe une fonction positive a telle que

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + U(t))) = -\log G(x) = (1 + \gamma x)^{-\frac{1}{\gamma}} \quad (1.20)$$

pour chaque point de continuité x de G , où $1 + \gamma x > 0$.

- Il existe une fonction positive g telle que

$$\lim_{t \rightarrow x^F} \frac{\bar{F}(t + xg(t))}{\bar{F}(t)} = (1 + \gamma x)^{-1/\gamma}, \quad (1.21)$$

Pour tout x , où $1 + \gamma x > 0$. De plus, (1.18) tient avec $b_n = U(n)$ et $a_n = a(n)$. En outre, (1.21) tient avec $g(t) = a\left(\frac{1}{F(t)}\right)$.

Proposition 1.5.28.

Soit $\gamma \in \mathbb{R}$. Il y a équivalence entre $F \in \mathcal{MDA}(G_\gamma)$ et pour tous $x > 0, y > 0, y \neq 1$,

$$\lim_{s \rightarrow \infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^\gamma - 1}{y^\gamma - 1} & \text{si } \gamma \neq 0 \\ \frac{\log x}{\log y} & \text{si } \gamma = 0 \end{cases} \quad (1.22)$$

Nous utiliserons ce résultat pour construire l'estimateur de Pickand de γ .

1.6 La loi des excès

L'approche basée sur la GEV a été critiquée dans la mesure où l'utilisation d'un seul maxima conduit à une perte d'information contenue dans les autres grandes valeurs de l'échantillon. Pour pallier ce problème, la méthode POT (Peaks-over-Threshold) ou méthode des excès au-delà d'un seuil élevé a été introduite dans Pickands [71].

1.6.1 Principe

La méthode des excès au-delà d'un seuil (ou Peak Over Threshold, POT) repose sur le comportement des valeurs observées au-delà d'un seuil donné. Plus précisément, elle consiste à observer non pas le maximum ou les plus grandes valeurs mais toutes les valeurs des réalisations qui excèdent un certain seuil élevé. L'idée de base de cette approche consiste à choisir un seuil suffisamment élevé et à étudier les excès au-delà de ce seuil.

Soit X une variable aléatoire de fdr F et u un réel suffisamment grand appelé seuil. Notons par

$$N_u = \text{card} \{i : i = 1, \dots, n, X_i > u\}$$

le nombre de dépassements du seuil par les $X_{i \leq n}$ et les $Y_{j \leq N_u}$ tels que

$$Y_j = X_i - u$$

sont les excès correspondants. On cherche à partir de la loi F de X à définir une loi conditionnelle F_u par rapport au seuil u pour les variables aléatoires dépassant ce seuil. On définit alors la loi conditionnelle des excès F_u par:

$$F_u(x) = \mathbb{P}(X - u \leq y | X > u) = \frac{F(u + y) - F(u)}{1 - F(u)}, y \geq 0. \quad (1.23)$$

La méthode POT s'appuie sur le Théorème de Pickands-Balkema-de Haan (Balkema and de Haan [3], 1974 ; Pickands [71], 1975), (voir Embrechts et al. [33, Théorème 3.4.5], 1997). Ce Théorème démontre que lorsque le seuil u est proche du point terminal x_F , la distribution conditionnelle des excès est asymptotiquement une Distribution de Pareto Généralisé (GPD: Generalised Pareto Distribution) de fonction de répartition:

$$H_\gamma(x) = \begin{cases} 1 - (1 + \gamma x)^{-\frac{1}{\gamma}}, & \text{si } \gamma \neq 0 \\ 1 - \exp(-x), & \text{si } \gamma = 0 \end{cases} \quad (1.24)$$

$$\text{où } \begin{cases} x \geq 0, & \text{si } \gamma \geq 0 \\ 0 \leq x \leq -\frac{1}{\gamma}, & \text{si } \gamma < 0 \end{cases}$$

H_γ est appelée Distribution de Pareto Généralisée standard. On peut obtenir une forme plus générale de $H_\gamma(x)$ en introduisant des paramètres d'échelle $\sigma > 0$ et de position $\mu \in \mathbb{R}$, respectivement:

$$H_{\gamma,\mu,\sigma}(x) = H_\gamma\left(\frac{x - \mu}{\sigma}\right)$$

Le cas $\gamma = 0$ peut être vu comme le cas limite lorsque que $\gamma \rightarrow 0$, on retrouve une loi exponentielle. Le paramètre de forme γ apparaissant dans $H_{\gamma,\sigma}$ est le même que celui découlant de la distribution généralisée des valeurs extrêmes. La fdr $H_{\gamma,0,\sigma}$ joue un rôle important dans la partie suivante. On la note par:

$$H_{\gamma,\sigma}(x) = 1 - \left(1 + \gamma \left(\frac{x}{\sigma}\right)\right)^{-\frac{1}{\gamma}}, x \in D(\gamma, \sigma) \quad (1.25)$$

$$\text{où } x \in D(\gamma, \sigma) = \begin{cases} [0, \infty), & \text{si } \gamma \neq 0 \\ \left[0, -\frac{\sigma}{\gamma}\right], & \text{si } \gamma < 0 \end{cases}$$

1.6.2 Théorème de Pickands-Balkema-de Haan

Le Théorème suivant représente le lien entre $H_{\gamma,\sigma}$ et G_γ , Embrechts et al. [33, Théorème 3.4.13(b)].

Théorème 1.6.1 (Pickands, Balkema, de Haan).

Pour chaque $\gamma \in \mathbb{R}$ on a:

$$F \in \mathcal{MDA}(G_\gamma) \iff \exists \sigma(u) > 0 : \lim_{u \rightarrow x_F} \sup_{0 \leq y \leq x_F - u} |F_u(y) - H_{\gamma,\sigma(u)}| = 0 \quad (1.26)$$

Ce théorème signifie que si F vérifie le théorème de Fisher et Tippet 1.3.5, i.e si F appartient au domaine d'attraction $\mathcal{MDA}(G_\gamma)$ alors il existe une fonction $\sigma(\cdot)$ positive et un réel γ tels que la loi des excès F_u peut être uniformément approchée par une distribution de Pareto généralisée (GPD) $H_{\gamma,\sigma}$. En résumant:

GEV	GPD
$G_\gamma, \gamma \in \mathbb{R}$	$H_{\gamma,\sigma}, \gamma \in \mathbb{R}, \sigma > 0$
Décrit la limite normalisé	Apparaît comme la distribution limite
des distributions de maxima.	des excès au-delà des seuils élevés.

2.1 Introduction

SOIT X_1, X_2, \dots, X_n une suite de variables aléatoires i.i.d. ayant pour fonction de répartition F et notons $X_{1,n}, X_{2,n}, \dots, X_{n,n}$ les statistiques d'ordre associés à cet échantillon. Dans le [Chapitre 1](#), nous avons connu que le comportement asymptotique des maximum $X_{n,n}$ (après renormalisation) est bien connu et peut être modélisé par la distribution des valeurs extrêmes ayant pour paramètre γ . Ce paramètre clé est appelé indice de valeurs extrêmes ou indice de queue. Il mesure le poids de la queue droite de la fdr sous-jacent F , ce qui nous permet de comprendre et de décrire le comportement des valeurs extrêmes d'une population. Selon son signe, trois domaines d'attraction sont possible pour F : Fréchet ($\gamma > 0$), Gumbel ($\gamma = 0$) et Weibull ($\gamma < 0$). Donc, l'estimation de ce paramètre constitue une tâche importante dans la théorie de valeurs extrêmes (L'estimation de l'indice de queue est importante pour de nombreux aspects), puisqu'elle est un point de départ pour l'inférence statistique sur les valeurs extrêmes d'une population. Avec son estimation, il est possible d'estimer d'autres paramètres d'événements extrêmes comme le point terminal à droite x^F de la fdr sous-jacent F , les quantiles extrêmes, la période de retour et la probabilité de dépassement d'un niveau élevé ainsi que l'existence des moments.

En fonction de l'importance, il ya une vaste littérature sur l'estimation de l'indice de queue. La connaissance de γ est donc nécessaire pour résoudre un certain nombre de problèmes dans l'analyse de valeur extrême, comme l'estimation des quantiles extrêmes de X , ce qui a fait de son estimation un sujet central dans la littérature. Les deux premiers estimateurs de ce paramètre

ont été proposés en 1975 par Hill [47] et Pickands [71]. Puis d'autres estimateurs ont été suggérés, comme l'estimateur du maximum de vraisemblance ou l'estimateur des moments (de Haan and Ferreira [21]). Drees and Kaufmann [32] a montré que les estimateurs de γ sont, en général, des fonctions régulières des grandes statistiques d'ordre. Ces estimateurs sont qualifiés d'estimateurs semi-paramétriques. Le terme semi-paramétrique reflète le fait que nous faisons seulement des hypothèses partielles sur la fdr F inconnue. La plupart des méthodes classiques précédentes consistent à utiliser les k statistiques d'ordre supérieures $X_{n-k,n}, \dots, X_{n,n}$. Cela semble naturel car l'information de la queue est seulement contenue dans la partie extrême supérieure de l'échantillon. Le choix de k dépend de n , donc $k = k_n$ et doit vérifier les conditions suivantes

$$k = k_n \longrightarrow \infty \quad \text{lorsque} \quad n \longrightarrow \infty \quad (2.1)$$

$$\frac{k_n}{n} \longrightarrow 0 \quad \text{lorsque} \quad n \longrightarrow \infty \quad (2.2)$$

La condition (2.1) est exigée afin d'utiliser un nombre assez grand de données pour l'estimation. La condition (2.2) vise à assurer que l'estimation est effectuée à partir de données qui appartiennent à la queue de la fdr. La statistique $X_{n-k,n}$ est alors dite statistique d'ordre intermédiaire.

Une suite d'entiers k_n est dite intermédiaire si elle satisfait (2.1) et (2.2). Par conséquent, les statistiques d'ordre $X_{n-k:n}$, avec k satisfaisant (2.1) et (2.2) sont appelées **statistiques d'ordre intermédiaire**.

2.2 Quelques estimateurs de l'indice de queue

Dans le paragraphe suivant, nous présentons quelques estimateurs classiques pour l'indice de valeurs extrêmes γ .

2.2.1 Estimateur de Pickands ($\gamma \in \mathbb{R}$)

L'estimateur de Pickands (Pickands, 1975) est le premier estimateur suggéré pour le paramètre γ . L'intérêt de cet estimateur est d'être défini pour $\gamma \in \mathbb{R}$

Définition 2.2.1. *L'estimateur de Pickands est défini par:*

$$\hat{\gamma}_{n,k}^{(P)} = \frac{1}{\ln 2} \ln \left(\frac{X_{n-k+1,n} - X_{n-2k+1,n}}{X_{n-2k+1,n} - X_{n-4k+1,n}} \right)$$

2.2.1.1 Construction de l'estimateur de Pickands

Une justification plus formelle de l'estimateur de Pickands est fournie par Embrechts et al. (1997). Selon eux, l'idée de base derrière l'estimateur est de trouver une condition équivalente à $F \in \mathcal{MDA}(G_\gamma)$ qui peut être facilement appliquée à l'estimation du paramètre $\gamma \in \mathbb{R}$. Le point clé est la propriété de caractérisation de $\mathcal{MDA}(G_\gamma)$ comme mentionné au chapitre 1 (Théorème 1.5.27, l'assertion 1.19). Donc, d'après l'assertion 1.19 du Théorème 1.5.27, on a

$$F \in \mathcal{MDA}(G_\gamma) \iff \lim_{t \rightarrow \infty} \{U(tx) - U(t)\}/a(t) = D_\gamma(x) = \frac{x^\gamma - 1}{\gamma}$$

existe pour tout $x > 0$.

D'après la Proposition 1.5.28, nous connaissons la relation entre $D_\gamma(x)$ et $D_\gamma(y)$. Par conséquent, nous pouvons écrire

$$\frac{D_\gamma(x)}{D_\gamma(y)} = \lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{U(ty) - U(t)} = \frac{x^\gamma - 1}{y^\gamma - 1},$$

pour $x, y > 0, y \neq 1$.

Lorsque nous choisissons les paramètres $x = 2$ et $y = 1/2$, les valeurs suivantes sont retenues

$$\lim_{t \rightarrow \infty} \frac{U(2t) - U(t)}{U(t) - U(t/2)} = \frac{2^\gamma - 1}{1 - (1/2)^\gamma} = \frac{2^\gamma - 1}{2^\gamma - 1} = 2^\gamma \quad (2.3)$$

La fonction de quantile de queue $U(t) = Q(1 - 1/t)$, $t \geq 0$ définit le t -ième quantile supérieur de la distribution F , qui correspond au quantile empirique $X_{n-k:n}$, où $k/n = 1/t$. Il s'ensuit que

$$\begin{aligned} U(t) &= U(n/k) \sim X_{n-k:n} \\ U(t/2) &= U(n/2k) \sim X_{n-2k:n} \\ U(2t) &= U(2n/k) \sim X_{n-k/2:n} \end{aligned}$$

Par conséquent, nous pouvons réécrire (2.3) sous la condition que $k/n = 1/t \rightarrow 0$ quand

$$\lim_{n \rightarrow \infty} \frac{X_{n-k/2:n} - X_{n-k:n}}{X_{n-k:n} - X_{n-2k:n}} \sim 2^\gamma.$$

Motivés par la discussion, nous définissons maintenant l'estimateur de Pickands

$$\gamma_{k,n}^{(P)} = \frac{1}{\log 2} \log \left(\frac{X_{n-k:n} - X_{n-2k:n}}{X_{n-2k:n} - X_{n-4k:n}} \right).$$

Une amélioration de l'estimateur de Pickands est proposée par Drees [31]. Il s'agit d'une combinaison convexe des estimateurs de Pickands obtenus pour différentes valeurs de k . Cet estimateur, appelé estimateur de Drees-Pickands, qui a été généralisé plus tard par Johan [51].

2.2.2 Estimateur de Hill ($\gamma > 0$)

Un des estimateurs les plus populaires de l'indice de queue est l'estimateur de Hill [47], qui est un estimateur non paramétrique simple basé sur des statistiques d'ordre. Hill a proposé une approche de l'estimation du maximum de vraisemblance conditionnelle (MLE) en maximisant la fonction de vraisemblance conditionnelle. Introduit en 1975 par Hill [47], l'estimateur de Hill, défini pour $\gamma > 0$, est probablement l'estimateur le plus étudié dans la littérature.

Définition 2.2.2. *L'estimateur de Hill, noté $\hat{\gamma}_{n,k}^{(H)}(k)$, construit à partir des k plus grandes statistiques d'ordre, est défini par*

$$\begin{aligned}\hat{\gamma}_{n,k}^{(H)}(k) &= \frac{1}{k} \sum_{j=1}^k \log X_{n-j+1:n} - \log X_{n-k:n} \\ &= \frac{1}{k} \sum_{j=1}^k j \log X_{n-j+1:n} - \log X_{n-j:n}\end{aligned}\tag{2.4}$$

2.2.2.1 Construction de l'estimateur de Hill

L'estimateur de Hill est le plus connu de tous les estimateurs connus de l'indice de valeur extrême γ et il a été conçu sous l'hypothèse que $\gamma > 0$. Il existe plusieurs approches pour obtenir la formule, tout aussi intéressante (voir, par exemple, Beirlant et al. [5], de Haan and Ferreira [21] ou Embrechts et al. [33]). Son plus grand inconvénient, est que nous ne pouvons évaluer l'indice de queue que lorsque la distribution sous-jacente a une queue lourde. Nous avons décidé de démontrer ici l'approche MLE, pour laquelle nous avons déjà les instruments mathématiques nécessaires.

L'approche MLE. Considérons la notation comme dans (1.23) et supposons la variable aléatoire X avec fdr F afin que

$$P_u(X > x) = F_u(x) \approx Cx^{-1/\gamma}, x \geq u > 0, \gamma > 0,$$

Quand on choisit la substitution $Y = \log(X/u)$, pour $x/u \geq 1$, il en résulte

que

$$P(Y > y) \approx \exp(-y)/\gamma, y \geq 0.$$

Ainsi, Y a approximativement une distribution exponentielle $\mathcal{Exp}(1/\gamma)$ avec $\mathbb{E}[Y] = \gamma$. L'estimateur MLE pour l'indice exponentiel est la moyenne de l'échantillon, à partir de laquelle nous obtenons immédiatement

$$\hat{\gamma}_{k,n} = \frac{1}{k} \sum_{j=1}^k Y_j = \frac{1}{k} \sum_{j=1}^k \log \left(\frac{X_{j,n}}{u} \right) = \frac{1}{k} \sum_{j=1}^k \log X_{j,n} - \log u$$

où k est le nombre de $X_{i,n} \geq u$. Puisque nous considérons généralement que le seuil u est une statistique d'ordre appropriée ($X_{n-k:n}$), on arrive à l'estimateur de Hill

$$\hat{\gamma}_{k,n}^{(H)}(k) = \frac{1}{k} \sum_{j=1}^k \log X_{n-j+1,n} - \log X_{n-k,n}$$

où k est une suite intermédiaire.

Cet estimateur peut aussi être exprimé sous la forme d'une moyenne de log-espacements

$$\hat{\gamma}_{n,k}^{(H)}(k) = \frac{1}{k} \sum_{j=1}^k j \log X_{n-j+1:n} - \log X_{n-j:n}$$

2.2.3 Estimateurs des moments ($\gamma \in \mathbb{R}$)

Alors que l'estimateur de Hill est adapté pour les lois dans le domaine d'attraction de la loi de Fréchet i.e $\gamma > 0$, l'estimateur suivant, appelé estimateur des moments, a été proposé par Dekkers et al. [25] est valable pour tous les domaines d'attraction, i.e $\gamma \in \mathbb{R}$. Le problème avec l'estimateur de Hill lorsque $\gamma \leq 0$ est que la limite de U en l'infini peut être négative. Dans ce cas, les X_i peuvent prendre des valeurs négatives et leur logarithme n'existe pas. Nous nous plaçons, dans ce paragraphe, dans le cas où $U(\infty) = \lim_{t \rightarrow \infty} U(t) > 0$. Il faut toutefois savoir que cette hypothèse peut influencer le comportement de l'estimateur.

Définition 2.2.3. (*Estimateurs des moments*). On définit l'estimateur des moments par

$$\gamma_{n,k}^{(M)}(k) = \gamma_{1,n,k}^{(M)}(k) + \gamma_{2,n,k}^{(M)}(k)$$

où

$$\gamma_{1,n,k}^{(M)}(k) = M_n^{(1)}$$

et

$$\gamma_{2,n,k}^{(M)}(k) = 1 + \frac{1}{2} \left(\frac{(M_n^{(1)})^2}{M_n^{(2)}} - 1 \right)^{-1}$$

avec

$$M_n^{(\ell)} = \sum_{i=1}^k \left(\log X_{n-i+1:n} - \log X_{n-k:n} \right)^\ell, \quad (\ell = 1, 2)$$

où $M_n^{(\ell)}$ est le $\ell^{\text{ième}}$ moment empirique.

2.2.3.1 Construction de l'estimateur de moments

Il a une forme un peu plus compliquée que l'estimateur de Hill et une motivation encore plus élaborée, que nous ne présentons donc pas ici (pour les lecteurs intéressés, voir p.142 dans [Beirlant et al., 2004](#)).

2.2.4 L'estimateur t-Hill

L'estimateur le plus simple et le plus courant de l'indice de queue est l'estimateur Hill. Cet estimateur ne fonctionne que pour les queues de type Pareto et a un biais élevé. Des recherches récentes ont montré que l'estimateur Hill est non robuste en présence de valeurs aberrantes par rapport au modèle supposé. Comme le montre Brazauskas and Serfling [13], de petites erreurs dans l'estimation de l'indice de queue peuvent entraîner de grandes erreurs dans l'estimation des quantiles. Certaines tentatives ont été faites pour produire des estimateurs robustes de l'indice de queue. Ceux-ci inclus: Peng and Welsh [70], Juárez and Schucany [55], Vandewalle et al. [84], Kim and Lee [57]. Stehlík et al. [80] a utilisé une nouvelle procédure d'estimation robuste, introduite par Fabián [34], connue sous le nom méthode de t-estimation ou t-score, pour proposer un nouvel estimateur pour l'indice de queue, appelé estimateur de t-Hill donné dans la définition suivante.

Définition 2.2.4. (L'estimateur t-Hill)

L'estimateur t-Hill donné en fonction de la moyenne harmonique de k grandes valeurs extrêmes de l'échantillon

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

2.2.4.1 Construction de l'estimateur t-Hill

La construction de cet estimateur est basé sur la la méthode t-estimation (en Anglais: score moment estimation, t-score ou t-estimation) introduite par

Fabián [34]. Ce dernier est plus robuste que l'estimateur de Hill $\hat{\gamma}^H$ (voir Stehlík et al. [80]).

Une fonction de score standard est généralement considérée comme une base pour l'inférence statistique. Si un paramètre va être estimé, nous avons généralement un groupe de traduction étroitement lié à sa fonction de score. Pour un paramètre à estimer θ , le score standard est le gradient par rapport à θ de logarithme de la fonction de vraisemblance, i.e

$$S(\theta, X) = \frac{\partial}{\partial \theta} \log L(\theta, X).$$

Soit \mathfrak{X} le support de la distribution F avec densité f , continûment différentiable par rapport à $x \in \mathfrak{X}$ et soit $v : \mathfrak{X} \rightarrow \mathbb{R}$ donné par

$$\text{t-score: } v(x) = \begin{cases} x & \text{pour } \mathfrak{X} = \mathbb{R} \\ \log(x - a) & \text{pour } \mathfrak{X} = (0, \infty) \end{cases}$$

Alors, le score transformé ou le t-score Fabián [34] est défini par

$$T(x) = \frac{-1}{f(x)} \frac{d}{dx} \left(\frac{1}{v'(x)} f(x) \right) \quad (2.5)$$

Qui exprime un changement relatif de composantes de base de la densité, la densité divisé par le jacobien de la fonction v .

Pour la distribution de Pareto, l'estimation du score standard avec $v = id$ et $S(X, \gamma) = \frac{1}{\gamma} - \log x$ est liée au MLE. Ce dernier est non-robuste en présence des valeurs aberrantes. Pour l'approche t-estimation, en utilisant la fonction $v(x) = \log(x - 1)$ nous obtenons d'après (2.5)

$$S_F(x) = \gamma \left(1 - \frac{\gamma + 1}{\gamma x} - \log x \right)$$

Cela nous permet de définir l'estimateur t-Hill par

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

Pour plus de détail voir Stehlík et al. [80], Stehlík et al. [81] ou Fabián and Stehlík [35].

Remarque 2.2.5. *Notons que les estimateurs Hill et t-Hill sont des membres d'une class d'estimateurs introduite par Henry [46] est dite Harmonic*

Moment Tail Index Estimators et est défini par

$$\hat{\gamma}_{\beta,n}^{(HM)}(k) = \frac{1}{1-\beta} \left\{ \left[\frac{1}{k} \sum_{j=1}^k \left(\frac{X_{n-k:n}}{X_{n-j+1:n}} \right)^{1-\beta} \right]^{-1} - 1 \right\}$$

Donc, si $\beta = 2$ alors $\hat{\gamma}_{2,n}^{(HM)}(k) = \hat{\gamma}_n^{(tH)}(k)$, et si $\beta \rightarrow 1$ on a $\lim_{\beta \rightarrow 1} \hat{\gamma}_{2,n}^{(HM)}(k) = \hat{\gamma}_n^{(H)}(k)$.

2.3 Principales propriétés asymptotiques: Consistance et Normalité Asymptotique

Alors que la consistance (en probabilité ou presque sûrement) de l'un des estimateurs présentés ci-dessus ne dépend que du comportement de la suite intermédiaire $k = k_n$ i.e. satisfait (2.1) et (2.2). quand $n \rightarrow \infty$ et que $F \in \mathcal{D}(G_\gamma)$, d'autres propriétés, à savoir la normalité asymptotique des estimateurs, nécessite des conditions plus délicates sur la fonction de distribution F et donc sur la fonction de quantile de queue $U(\cdot)$. Cette condition supplémentaire est connue sous le nom de propriété de variation régulière étendue du second ordre et, comme pour la condition de premier ordre, nous nous référerons à cette nouvelle propriété comme condition de second ordre.

Définition 2.3.1. (*Condition de second ordre*). On dit que la fonction U (ou la fdr F associé) satisfait la condition de second ordre si, pour une fonction positive a et pour une fonction positive ou négative A , avec $\lim_{t \rightarrow \infty} A(t) = 0$,

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx) - U(t)}{a(t)} - D_\gamma(x)}{A(t)} = H(x), \quad x > 0 \quad (2.6)$$

Comme la fonction a de la condition du premier ordre, on appelle A la fonction auxiliaire du second ordre.

Nous devons maintenant déterminer quelles fonctions $H(x)$ sont admissibles à la limite dans (2.6). Suivant de Haan and Ferreira [21, 2006], nous pouvons alors énoncer le résultat suivant pour la fonction $H(x)$:

Théorème 2.3.2. (*de Haan and Ferreira(2006)*, Théorème 2.3.3 et Corollaire 2.3.4) Supposons que la relation (2.6) soit vérifiée et que la fonction H ne soit pas un multiple de D_γ et qu'elle ne soit pas identiquement nulle. Il existe alors des fonctions a , positive et A , positive ou négative, et un paramètre

$\tau \leq 0$ tel que

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx) - U(t)}{a(t)} - D_\gamma(x)}{A(t)} = H_{\gamma, \tau}(x) = \frac{1}{\tau} \left(\frac{x^{\gamma+\tau} - 1}{\gamma + \tau} - \frac{x^\gamma - 1}{\gamma} \right) \quad (2.7)$$

pour $x > 0$. τ est un paramètre de second ordre contrôlant la vitesse de convergence de la condition de premier ordre. Pour les cas $\gamma = 0$ et/ou $\tau = 0$, $H_{\gamma, \tau}(x)$ est compris comme égal à la limite respective dans (2.7), par des arguments de continuité. De plus, $A(t)$ est tel que

$$\lim_{t \rightarrow \infty} \frac{A(tx)}{A(t)} = x^\tau,$$

i.e, $|A| \in \mathcal{RV}_\tau$.

La fonction A , décrivant la vitesse de convergence de la condition du premier ordre, varie régulièrement avec l'indice τ . Donc, si $\tau < 0$, on a une vitesse algébrique de convergence et si $\tau = 0$, la vitesse de convergence est plus lente (logarithmique, par exemple). Par conséquent, la vitesse auquel la suite intermédiaire k_n tend vers l'infini doit être conforme à la vitesse de convergence dans la condition de premier ordre, telle que quantifiée par la fonction de vitesse A .

Notons que la condition de second ordre implique la condition de domaine d'attraction.

Pour les distributions à queue lourde $\gamma > 0$, la condition de second ordre est habituellement écrite comme

Théorème 2.3.3. *Soit $F \in \mathcal{MDA}(G_\gamma)$, $\gamma > 0$, on dit que F est à variation régulière du second ordre à l'infini si elle satisfait la condition suivante:*

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\tau - 1}{\tau} \quad (2.8)$$

où $\tau \leq 0$ et $A(t)$ est une fonction de signe constant telle que $A(t) \rightarrow 0$ quand $t \rightarrow \infty$.

Remarque 2.3.4. *La condition de second ordre formulée en fonction de U est aussi équivalente aux deux conditions suivantes:*

- En fonction de F : Il existe une constante réelle $\tau \leq 0$ et une fonction de signe constant $A_F(t) \rightarrow 0$ quand $t \rightarrow \infty$, telles que pour tout $x > 0$

$$\lim_{t \rightarrow \infty} \frac{\frac{\overline{F}(tx) / \overline{F}(t)}{A_F(t)} - x^{-1/\gamma}}{A_F(t)} = x^{-1/\gamma} \frac{x^{\tau/\gamma} - 1}{\gamma\tau}, \quad (2.9)$$

- En fonction de Q : Il existe une constante réelle $\tau \leq 0$ et une fonction $A_Q(s) \rightarrow 0$ quand $s \rightarrow 0$, telles que, pour tout $x > 0$

$$\lim_{s \rightarrow 0} \frac{\frac{Q(1-sx)}{Q(1-s)} - x^{-\gamma}}{A_Q(s)} = x^{-\gamma} \frac{x^\tau - 1}{\tau}, \quad (2.10)$$

où $A_Q(s) := \gamma^2 a(Q(s))$, (voir de Haan and Stadtmüller, 1996 ou Théorème 3.2.9 dans de Haan and Ferreira, 2006)

- Le paramètre τ est appelé paramètre du second ordre. Ce paramètre contrôle la vitesse de convergence de $U(tx)/U(t)$ vers x^γ quand $t \rightarrow \infty$ ou de manière équivalente celle de $\bar{F}(tx)/\bar{F}(t)$ vers $x^{-1/\gamma}$ quand $t \rightarrow \infty$.
- Pour $\tau = 0$, $(x^\tau - 1)/\tau$ s'interprète comme sa limite lorsque $\tau \rightarrow 0$, i.e. $\lim_{\tau \rightarrow 0} (x^\tau - 1)/\tau$ et est égale à $\log x$.

Notons par E l'un des estimateurs présentés dans la section précédent. Sous (2.6), ou (2.7) dans le cas $\gamma > 0$, on peut alors donner la normalité asymptotique de $\hat{\gamma}_{n,k}^{(E)}$. Tel qu'il est possible de garantir l'existence des constantes $b_E \in \mathbb{R}$ et $v_E > 0$, telles que

$$\hat{\gamma}_{n,k}^{(E)} \stackrel{d}{=} \gamma + v_E N_k^E / \sqrt{k} + b_E A(n/k) + o_p A(n/k) \quad (2.11)$$

avec $N_k^{(E)}$ une variable aléatoire de loi asymptotique normale standard. De de Haan and Ferreira(2006), on obtient le théorème suivant:

Théorème 2.3.5. (Normalité asymptotique). Si $F \in \mathcal{MDA}(G_\gamma)$, $\gamma \in \mathbb{R}$ satisfait une condition de variation régulière du second ordre d'indices γ et $\rho < 0$ avec pour fonction auxiliaire A (Théorème 2.3.2), et si (k_n) est une suite intermédiaire telle que

$$\lim_{n \rightarrow +\infty} \sqrt{k} A\left(\frac{n}{k}\right) = \lambda < \infty$$

alors

$$\sqrt{k}(\hat{\gamma}_{n,k}^{(E)} - \gamma) \xrightarrow{d} Z \sim \mathcal{N}(\lambda b_E, v_E)$$

où b_E est un terme de biais dépendant à la fois de γ et du paramètre du second ordre ρ et v_E est un terme de variance qui ne dépend que de γ et $\mathcal{N}(\mu, \sigma^2)$ désignant une v.a. de loi normal avec moyen μ et variance σ^2 .

Les valeurs b_E et v_E sont communément appelées le biais asymptotique et la variance asymptotique de $\hat{\gamma}_{n,k}^{(E)}$ respectivement. On peut trouver des expressions pour b_E et v_E dans de Haan and Ferreira [21, 2006, chapitre 3].

2.3.1 Estimateur de Pickands ($\gamma \in \mathbb{R}$)

Les propriétés de l'estimateur de Pickands ($\gamma_{n,k}^{(P)}$) ont été étudiées par Pickands [71] et étendues par Dekkers and De Haan [26]. Pickands [71] démontre la consistance faible de son estimateur et Dekkers and De Haan [26] ont prouvé que cet estimateur est fortement consistant (et donc faiblement consistant aussi) et asymptotiquement Normal-distribué.

Théorème 2.3.6. (*Consistance faible, Pickands 1975*). *Supposons que $F \in \mathcal{MDA}(G_\gamma)$ et (k_n) est une suite intermédiaire, alors*

$$\gamma_{k,n}^P \xrightarrow{P} \gamma \quad n \rightarrow \infty$$

Remarque 2.3.7. *Au cours de la preuve de ce résultat, il est montré que*

$$\frac{X_{n-k+1:n} - X_{n-2k+1:n}}{X_{n-2k+1:n} - X_{n-4k+1:n}} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \frac{4^\gamma - 1}{2^\gamma - 1} = 2^\gamma$$

ce qui explique la normalisation par $\ln 2$ dans l'expression de $\hat{\gamma}_P$. Par ailleurs, $(2^\gamma - 1)/\gamma$ et $(4^\gamma - 1)/\gamma$ sont respectivement la médiane et le troisième quartile de la loi de Pareto généralisée. Ainsi, l'estimateur de Pickands estime γ via les quantiles de la loi limite.

Théorème 2.3.8. (*Consistance forte, Dekkers et de Haan, 1989*).

Supposons que $F \in \mathcal{MDA}(G_\gamma)$ et (k_n) est une suite intermédiaire telle que $k/\log \log n \rightarrow \infty$ quand $n \rightarrow \infty$, alors

$$\gamma_{k,n}^P \xrightarrow[p.s.]{} \gamma \quad n \rightarrow \infty$$

Théorème 2.3.9. (*Normalité asymptotique*)

Si $F \in \mathcal{MDA}(G_\gamma)$, $\gamma \in \mathbb{R}$ satisfait la condition de variation régulière du second ordre (2.7) (Théorème 2.3.2), et si (k_n) est une suite intermédiaire telle que

$$\lim_{n \rightarrow +\infty} \sqrt{k} A\left(\frac{n}{k}\right) = \lambda < \infty$$

alors

$$\sqrt{k}(\gamma_{k,n}^P - \gamma) \xrightarrow{d} \mathcal{N}(0, \sigma_P^2) \quad n \rightarrow \infty$$

$$\text{où } \sigma_P^2 = \frac{\gamma^2(2^{2\gamma+1}+1)}{(2(2^\gamma-1)\log 2)^2}.$$

2.3.2 Estimateur de Hill

Mason [60] a démontré la consistance faible et Deheuvels et al. [24] ont établi la consistance forte.

Théorème 2.3.10. (*Consistance faible, de Haan et Ferreira, 2006, Théorème 3.2.5*). Supposons $F \in \mathcal{MDA}(\Phi_\gamma)$ avec $\gamma > 0$, et k une suite intermédiaire, alors

$$\gamma_{k,n}^H \xrightarrow{\mathbb{P}} \gamma \quad n \rightarrow \infty$$

Théorème 2.3.11. (*Consistance forte*) Supposons $F \in \mathcal{MDA}(\Phi_\gamma)$ avec $\gamma > 0$, et k une suite intermédiaire telle que $k/\log(\log n) \rightarrow \infty$ quand $n \rightarrow \infty$, alors

$$\gamma_{k,n}^H \xrightarrow{p.s.} \gamma \quad n \rightarrow \infty$$

La normalité asymptotique est due entre autre à Davis and Resnick [19], Csörgő et al. [18], Haeusler and Teugels [45] et Smith [79] Smith (1987). Le Théorème suivant correspond au Theorem 3.2.5. dans de Haan and Ferreira [21].

Théorème 2.3.12. (*Normalité asymptotique*). Soit $F \in \mathcal{MDA}(G_\gamma)$, $\gamma > 0$ et k une suite intermédiaire telle que $\sqrt{k}A(n/k) \rightarrow \lambda$ quand $n \rightarrow \infty$ alors,

$$\sqrt{k}(\gamma_{k,n}^H - \gamma) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \sigma^2\right) \quad n \rightarrow \infty.$$

2.3.3 Estimateur des moments

Cet estimateur est consistant et asymptotiquement normal sous une condition de variation régulière du second ordre d'indices γ et ρ avec un terme de biais qui dépend de γ et du paramètre du second ordre ρ et un terme de variance qui ne dépend que de γ . Pour plus de détails, voir Dekkers et al. [25], ou encore de Haan and Ferreira [21, Théorème 3.5.4]

Théorème 2.3.13. (*Consistance faible, Dekkers al., 1989*). Supposons que $F \in \mathcal{MDA}(G_\gamma)$, $\gamma \in \mathbb{R}$ et k une suite intermédiaire, alors

$$\gamma_{k,n}^M \xrightarrow{\mathbb{P}} \gamma \quad n \rightarrow \infty$$

Théorème 2.3.14. (*Consistance forte, Dekkers al., 1989*). Supposons que $F \in \mathcal{MDA}(G_\gamma)$, $\gamma \in \mathbb{R}$ et k une suite intermédiaire telle que $k/(\log n)^\eta \rightarrow \infty$ quand $n \rightarrow \infty$, avec $\eta > 0$, alors

$$\gamma_{k,n}^M \xrightarrow{p.s.} \gamma \quad n \rightarrow \infty$$

Théorème 2.3.15. (*Normalité asymptotique*). (voir Théorème 3.1 et corollaire 3.2 de Dekkers et al. [25]) Si $F \in \mathcal{MDA}(G_\gamma)$, $\gamma \in \mathbb{R}$ satisfait la condition de variation régulière du second ordre (2.7) avec $\tau \neq \gamma$ (Théorème 2.3.2), et si (k_n) est une suite intermédiaire telle que

$$\lim_{n \rightarrow +\infty} \sqrt{k} A\left(\frac{n}{k}\right) = \lambda < \infty$$

alors

$$\sqrt{k}(\gamma_{k,n}^M - \gamma) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \quad n \rightarrow \infty$$

où

$$\begin{aligned} \sigma^2 &= (1 + \sigma^2) \mathbf{1}_{\{\gamma \geq 0\}} + (1 - \gamma^2)(1 - 2\gamma) \left(4 - 8 \frac{1 - 2\gamma}{1 - 3\gamma} + \frac{(5 - 11\gamma)(1 - 2\gamma)}{(1 - 3\gamma)(1 - 4\gamma)} \right) \mathbf{1}_{\{\gamma < 0\}} \\ &= (1 + \sigma^2) \mathbf{1}_{\{\gamma \geq 0\}} + \left(\frac{(1 - \gamma^2)(1 - 2\gamma)(1 - \gamma + 6\gamma^2)}{(1 - 3\gamma)(1 - 4\gamma)} \right) \mathbf{1}_{\{\gamma < 0\}} \end{aligned}$$

2.3.4 L'estimateur t-Hill

La consistance faible de l'estimateur t-Hill a été prouvée dans des conditions de régularité standard dans Stehlik et al. [80] ou dans Beran et al. [6, Théorème 1]. La normalité asymptotique est donné dans Beran et al. [6, Théorème 2]

Théorème 2.3.16. Supposons $F \in \mathcal{MDA}(\Phi_\gamma)$ et $k = k_n$ une suite intermédiaire alors

$$\hat{\gamma}_{k,n}^{(tH)} \xrightarrow{P} \gamma$$

Théorème 2.3.17. Soit $F \in \mathcal{MDA}(\Phi_\gamma)$ et supposons que la condition de second ordre (2.8) est satisfaite avec $\gamma > 1/2$. Alors pour toute suite intermédiaire $k = k_n$ tel que

$$\sqrt{k} A(n/k) \rightarrow \lambda, \quad n \rightarrow \infty.$$

on a

$$\sqrt{k} \left(\hat{\gamma}_{n,k}^{(tH)} - \gamma \right) \xrightarrow{d} \mathcal{N}(\lambda\mu, \sigma^2), \quad n \rightarrow \infty.$$

où $\mu = \mu(\gamma, \rho) = \frac{1+\gamma}{1-\rho+\gamma}$ et $\sigma^2 = \sigma^2(\gamma) = \frac{\gamma^2(1+\gamma)^2}{(1+2\gamma)}$

2.3.5 Comment choisir le nombre de statistiques d'ordre

La valeur de $k = k_n$ représente d'une part la quantité de données qu'on extrait d'un n -échantillon pour les estimateurs de l'indice. Le choix de k_n est d'une grande importance en pratique. Pour un n -échantillon donné, si $k = k_n$ est très grand (autrement dit, si on prend beaucoup de valeurs pour l'estimation), un biais apparaît. A l'inverse, si k est trop petit la variance d'estimation devient trop importante. Il s'agit donc d'effectuer d'un compromis entre biais et variance. Nous appelons k_{opt} la valeur de k optimale par rapport à un certain critère.

Choix automatique du seuil: Reiss et Thomas ont proposé une méthode heuristique pour choisir le nombre optimal de valeurs extrêmes k_{opt} utilisées pour estimer les paramètres de la loi (Reiss and Thomas, 2007). Celui-ci est donné par

$$k_{opt} = \arg \min_k \frac{1}{k} \sum_{i=1}^k i^\theta |\hat{\gamma}(i) - \text{median}\{\hat{\gamma}(1), \dots, \hat{\gamma}(k)\}|$$

où $\gamma(i)$ (pour $i \leq k$) est l'estimateur de l'indice extrême en utilisant les j plus grandes valeurs et med désigne la médiane. Cette algorithm est utilisé dans le chapitre 4 pour chercher le k_{opt} .

2.4 Estimation de quantiles extrêmes

Définissons à nouveau $Q(1-s) = F^{\leftarrow}(1-s) = U(1/s)$ comme quantile extrême d'ordre $(1-s)$ de la fdr sous-jacent F du v.a. X , avec s petit ($s \rightarrow 0$). On souhaite ici estimer un quantile extrême, comme dans l'estimation de γ , la seule hypothèse est que $F \in \mathcal{MDA}(G_\gamma)$. Les quantiles extrêmes jouent un rôle important dans le contexte de l'étude du risque où il est crucial d'évaluer adéquatement le risque pour des pertes lourdes d'occurrences rares voir par exemple les Chapitres 3 et 4. On peut trouver des estimateurs construits à partir des estimateurs de l'indice de queue présentés précédemment. En voici trois.

- à partir de l'estimateur de Pickands

Dekkers et al. [25] donnent l'estimateur de quantile extrême à partir de l'estimateur de Pickands:

$$\hat{Q}_n(1-s)^{(P)} = \frac{(k/sn)^{\hat{\gamma}_{k,n}^{(P)}-1}}{1-2^{-\hat{\gamma}_{k,n}^{(P)}}} (X_{n-k+1:n} - X_{n-2k+1:n}) + X_{n-k+1:n} \quad (2.12)$$

Le théorème de normalité asymptotique de cet estimateur est donné dans Dekkers and De Haan [26] (Théorème 3.3).

- à partir de l'estimateur de Hill

Weissman [90] a proposé un estimateur de quantile extrême à partir de l'estimateur de Hill:

$$\hat{Q}_n(1-s)^{(W)} = X_{n-k:n} \left(\frac{k}{ns} \right)^{\hat{\gamma}_{k,n}^{(H)}} \quad (2.13)$$

- à partir de l'estimateur des moments

Un estimateur de quantile extrême peut être formulé à partir de l'estimateur des moments de Dekkers et al. [25]:

$$\hat{Q}_n(1-s)^{(M)} = X_{n-k:n} + \frac{\hat{a}(k/n)}{\hat{\gamma}_{k,n}^{(M)}} \left[\left(\frac{k}{ns} \right)^{\hat{\gamma}_{k,n}^{(M)}} - 1 \right] \quad (2.14)$$

avec $\hat{a}(k/n) = X_{n-k,n} \hat{\gamma}_{k,n}^{(H)} \max(1 - \hat{\gamma}_{k,n}^{(M)}, 1)$.

La normalité asymptotique de cet estimateur de quantile est montrée dans Dekkers et al. [25] et dans de Haan and Rootzén [22].

Ces dernières années, une attention croissante a été portée aux méthodes permettant d'associer une mesure chiffrée à un risque, ce risque pouvant être issu d'une position financière ou d'aléas assurés par une compagnie d'assurance, ces derniers incluant des risques majeurs variés comme ceux des catastrophes naturelles, des pandémies, ou encore des risques industriels.

En général, un risque est un événement qui peut ou peut ne pas se produire (i.e, un événement aléatoire) et apporte quelques conséquences néfastes. Il est naturel de modéliser le risque en une variable aléatoire qui représente le montant aléatoire de perte qu'une société peut rencontrer. On peut supposer que les variables aléatoires modélisant les pertes de risque sont non négatives (semblable au risque d'assurance). En général, le risque peut être défini comme une variable aléatoire représentant une valeur future, mais nous nous concentrons sur les risques de pertes, pas de profits.

3.1 Notations et définitions

Considérons un espace de probabilité $(\Omega, \mathcal{A}, \Pr)$. Où

- Ω représente l'ensemble de tous les scénarios possibles;
- \mathcal{A} est une tribu;
- P est une mesure de probabilité.

La valeur future d'un scénario est incertaine et peut être représentée par une v.a. X . This is a fonction dans l'ensemble de tous les scénarios possibles vers les nombres réels, $X : \Omega \rightarrow \mathbb{R}$.

Définition 3.1.1. (*risque*) Soit (Ω, \mathcal{A}) un espace mesurable où Ω est l'espace de résultats et \mathcal{A} est la tribu définie ci-dessus. Un risque est une variable aléatoire définie sur (Ω, \mathcal{A}) .

Les éléments X de \mathcal{X} sont des variables aléatoires, ce qui représente des pertes. Notons \mathcal{X} l'ensemble de variables aléatoires de pertes réelles définies sur un espace mesurable (Ω, \mathcal{A}) . \mathcal{X} contient les constantes et est stable par l'addition et la multiplication par un scalaire.

Comme son nom l'indique, une mesure de risque permet de quantifier le danger inhérent à un risque représenté par une valeur aléatoire X (mesurer des risques signifie établir une correspondance entre la variable aléatoire représentant le risque et un nombre réel non-négatif).

Définition 3.1.2. Une mesure de risque $\rho(X)$ (éventuellement infinie) est une fonctionnelle ρ qui attribue une valeur réelle à la variable aléatoire X des pertes associées à un risque tel que

$$\rho : \mathcal{X} \rightarrow \mathbb{R} \quad (3.1)$$

Compte tenu de cette définition du risque, l'espérance, la variance et les moments d'ordres supérieures, les mesures basées sur le quantile comme la VaR, et les espérances conditionnelles telles que CVaR sont toutes des mesures de risque. Si $X > 0$ on a eu une perte et si $X < 0$ on a eu un gain.

Il existe de nombreuses mesures de risque introduites dans la littérature et la pratique, et le choix d'une mesure de risque peut être difficile. Une approche pour traiter la question de la mesure du risque consiste à commencer par une liste des propriétés qu'une mesure de risque doit satisfaire.

Pour répondre à la nécessité de principes théoriques et pratiques, il est d'usage qu'une mesure de risque vérifie un certain nombre de propriétés. Une liste de propriétés axiomatiques pour une bonne mesure de risque a été introduite dans l'article fondateur par Artzner et al. [2, 1999]. La vérification de ces propriétés amène à la notion de mesure de risque cohérente.

Définition 3.1.3. Deux risques X, Y sont dits comonotones si il existe une v.a. z et des fonctions non-décroissantes h_1, h_2 tels que $X = h_1(z), Y = h_2(z)$. En particulier, les risques comonotoniques peuvent être toujours représentés par $X = F_X^{-1}(U), Y = F_Y^{-1}(U)$, où U est une v.a. de fdr uniforme et $F_X(\cdot)$ et $F_Y(\cdot)$ est la fdr de X et Y respectivement.

Définition 3.1.4. (*Comonotonie*) Deux variables aléatoires à valeurs

réelles X et Y sur (Ω, \mathcal{F}) sont dites comonotones si

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0, \forall (\omega, \omega') \in \Omega \times \Omega.$$

Proposition 3.1.5. *Pour deux variables aléatoires à valeurs réelles X et Y sur (Ω, \mathcal{F}) les conditions suivantes sont équivalentes:*

- i. X et Y sont comonotones,
- ii. Ils existent une v.a. Z sur (Ω, \mathcal{F}) et deux fonctions non-décroissantes f et g sur \mathbb{R} tel que $X = f(Z)$ et $Y = g(Z)$.

Définition 3.1.6. (invariance en loi) Une mesure de risque $\rho : \mathcal{X} \rightarrow \mathbb{R}$ qui satisfait $\rho(X) = \rho(Y)$ pour tout $X, Y \in \mathcal{X}$ tels que X et Y ont la même distribution dessous P s'appelle une mesure de risque invariance en loi.

Définition 3.1.7. (Convexité et Concavité) Une fonction ψ , définie sur un intervalle I , est convexe sur I si la partie du plan située au-dessus de la courbe est convexe, i.e, si tout arc de sa courbe représentative est situé au-dessous de la corde correspondante. Cette définition se traduit par:

$$\psi(kx_1 + (1 - k)x_2) \leq k\psi(x_1) + (1 - k)\psi(x_2)$$

$\forall t \in [0, 1]$ et $\forall x_1, x_2 \in I$. Si $-\psi$ est convexe, ψ est dite concave.

Pour les définitions de tous les axiomes, X et Y sont des variables aléatoires représentant la perte, $c \in \mathbb{R}$ est un scalaire représentant la perte, et ρ est une mesure de risque.

Axiome 1. Invariance par translation: pour toute constante c , $\rho[X + c] = \rho[X] + c$;

Axiome 2. Monotonie: si $X < Y$ pour tous les résultats possibles, alors $\rho[X] \leq \rho[Y]$

Axiome 3. Homogénéité positive: pour toute réelle positive c , $\rho[cX] = c\rho[X]$.

Axiome 4. Sous-additivité: $\rho[X + Y] \leq \rho[X] + \rho[Y]$;

Axiome 5. Convexité: $\forall X, Y \in \mathcal{X}$

$$\rho[kX + (1 - k)Y] \leq k\rho[X] + (1 - k)\rho[Y] \quad (3.2)$$

Axiome 6. Invariance en loi:

$$X = Y \implies \rho[X] = \rho[Y] \quad (3.3)$$

Axiome 7. Additivité comonotone:

$$\rho[X + Y] = \rho[X] + \rho[Y] \quad \text{où } X, Y \text{ sont comonotons.} \quad (3.4)$$

Artzner et al. [2, 1999] analysé les mesures de risque et énoncé un ensemble des axiomes qui devraient être souhaitables pour toute mesure de risque. Toute mesure de risque qui satisfait ces axiomes est dite cohérente.

Définition 3.1.8. (*mesure du risque cohérente*) une mesure du risque $\rho[X]$ est dite cohérente si elle vérifie les propriétés

(1) *Invariance par translation,*

(2) *Monotonie,*

(3) *Homogénéité positive,*

(4) *Sous-additivité.*

Cette définition axiomatique est la pierre angulaire d'une théorie très riche qui tire ses modules de l'analyse fonctionnelle et a des interprétations économiques intéressantes.

Ininterprétation des 4 axiomes:

Axiome 1. Cet axiome signifie que l'ajout d'un montant fixe à au risque va changer la valeur de mesure du risque par le même montant. Cela garantit alors que $\rho[X - \rho[X]] = 0$.

Axiome 2. Cet axiome est intuitif et exprime une exigence minimale : si le risque Y est plus grand que le risque X , alors les mesures correspondantes doivent satisfaire à la même inégalité.

Axiome 3. Cet axiome signifie que mesurer une proportion c d'un risque revient à considérer la mesure du par la même proportion.

Axiome 4. "a merger does not create extra risk," disent Artzner et al. [2, 1999]. Le raisonnement derrière cet axiome est d'intégrer l'idée de diversification, i.e, la fusion de deux risques, X_1 et X_2 avec mesures associées $\rho[X_1]$ et $\rho[X_2]$ va créer le risque global $X = X_1 + X_2$ avec une mesure $\rho[X]$ inférieure ou égale à la somme des mesures associées $\rho[X_1] + \rho[X_2]$. Cet axiome est le plus connu parce que le VaR ne satisfait pas cette condition dans certaines situations.

Une autre façon de formaliser une mesure de risque cohérente est le théorème de représentation d'une mesure de risque cohérente:

Théorème 3.1.9. *Une mesure de risque est cohérente si et seulement si, il existe une famille \mathcal{P} de mesure de probabilité sur l'ensemble des états de la nature tel que:*

$$\rho(X) = \sup \{ \mathbb{E}_P(X) \mid P \in \mathcal{P} \}. \quad (3.5)$$

Voir Artzner et al. [2, 1999] pour une démonstration.

Remarque 3.1.10. *La Value-at-Risk, bien que très largement utilisée en finance, n'est pas une mesure de risque cohérente car elle n'est pas sous-additive ; de même pour la variance. Par contre la CVaR est une mesure de risque cohérente.*

Définition 3.1.11. *Les mesures de risque qui satisfont les propriétés (1) et (2) de la définition 3.1.8 sont appelées mesures de risque monétaires.*

Nous introduisons ici une relaxation de la propriété de cohérence. En effet, dans bien des cas, la propriété d'homogénéité positive n'est pas vérifiée autrement dit le risque ne croît pas de manière linéaire par rapport à la taille du portefeuille. Pour prendre en compte le risque de liquidité, Föllmer and Schied [37, 2002a] proposent de rejeter l'axiome d'homogénéité. Moyennant cela, ils définissent un nouvel ensemble de mesures de risque dites convexes, qui englobent les mesures de risque cohérentes.

Les critiques des axiomes de sous-additivité et d'homogénéité positive ont conduit à l'étude des mesures de risque convexe (voir la définition 3.1.7). Une mesure de risque convexe satisfait les mêmes axiomes qu'une mesure de risque cohérente, sauf que les axiomes de sous-additivité et d'homogénéité positive sont remplacés par l'axiome de convexité (Axiome 5.).

Si les axiomes (Axiome 3.) et (Axiome 4.) sont remplacés par l'axiome de convexité (Axiome 5.), la mesure de risque ρ est appelée mesure de risque convexe.

Définition 3.1.12. *Une mesure de risque ρ définie d'un ensemble de variables aléatoires convexe \mathcal{X} à valeur réelle est dite convexe (voir Föllmer and Schied [37, 2002a]) si elle vérifie les Axiomes suivantes:*

- (1) *Invariance par translation,*
- (2) *Monotonie,*
- (3) *Convexité.*

Interprétation. La propriété de convexité dit que la position $\rho[kX + (1 - k)Y]$ est moins risquée que les positions $k\rho[X]$ et $(1 - k)\rho[Y]$ prises individuelle-

ment. Les mesures de risque convexes ont été présentées et étudiées la première fois dans (Föllmer and Schied [38, 2002b]).

Remarque 3.1.13.

- D'après la définition 3.1.11, On peut donc dire aussi qu'une mesure de risque monétaire avec propriété de convexité comme dans 3.1.12 est appelée mesure de risque convexe.
- Lorsque la mesure de risque ρ est normalisée (i.e $\rho(0) = 0$) et qu'elle vérifie la propriété d'homogénéité positive alors l'axiome de convexité est équivalente à l'axiome de sous-additivité. Donc, la propriété de convexité possède certains aspects de la diminution du risque sous diversification comme pour l'axiome de sous-additivité.
- Une mesure du risque cohérente est une mesure de risque convexe. Cependant, une mesure de risque convexe n'est pas nécessairement cohérente car elle ne peut pas satisfaire les propriétés de sous-additivité et l'homogénéité positive de la définition 3.1.8. Si une mesure de risque convexe est positivement homogène, elle est également sous-additive.

Comme dans Artzner et al. [2, 1999] pour le cas d'une mesure du risque cohérente, nous donnons ici un théorème de représentation pour la mesure de risque convexe. Soient \mathcal{X} l'espace des fonctions à valeurs réelles sur un ensemble fini Ω , et \mathcal{Q} est l'ensemble de toutes les mesures de probabilités dans Ω .

Théorème 3.1.14. (Théorème de représentation) Toute mesure de risque convexe ρ définie de \mathcal{X} à valeurs dans \mathbb{R} peut s'écrire sous la forme:

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{Q}} (\mathbb{E}[-X] - \alpha(\mathbb{Q})), \quad (3.6)$$

où \mathcal{Q} est une famille de mesures de probabilité sur l'espace des états de la nature et α est une fonctionnelle sur \mathcal{Q} donnée par

$$\alpha(\mathbb{Q}) = \sup_{X \in \mathcal{A}_\rho} (\mathbb{E}[-X])$$

et

$$\mathcal{A}_\rho = \{X \in \mathcal{X} | \rho[X] \leq 0\}$$

Notons que le Théorème 3.1.14 comprend le théorème de représentation pour la mesure de risque cohérente (Théorème 3.1.9) comme un cas spécial. En effet, il est facile de voir que ρ satisfait la propriété d'homogénéité positive ,

c-à-d, ρ sera une mesure du risque cohérente, si et seulement si la fonction de pénalité ci-dessus $\alpha(\cdot)$ (voir Föllmer and Schied [37, 2002a]) ne prend que la valeur 0 et $+\infty$. Dans ce cas, Le Théorème 3.1.14 implique la représentation 3.5 en termes de l'ensemble $Q = \{Q \in \mathcal{P} | \alpha(Q) = 0\}$, et la représentation en 3.6 se simplifie en

$$\rho(X) = \sup_{Q \in \mathcal{P}} (\mathbb{E}[-X])$$

Nous présentons dans la suite des exemples de mesures couramment utilisées en gestion des risques.

3.2 Quelques mesures de risque usuelles

3.2.1 La variance

La variance est une mesure de dispersion d'une variable aléatoire autour de sa moyenne. Si X est une variable aléatoire de carré intégrable, sa variance est définie par

$$Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]$$

3.2.2 La Value-at-Risk

L'une de plus populaires mesures de risque est la Value-at-Risk (VaR), appelée également "Valeur à Risque".

Définition 3.2.1. On appelle Value-at-Risk de niveau $\alpha \in (0, 1)$ le quantile de niveau α ,

$$\rho[X] = VaR_\alpha[X] = x_\alpha \quad \text{où} \quad P(X \leq x_\alpha) = \alpha,$$

ou encore

$$\begin{aligned} VaR_\alpha[X] = F^{-1}(\alpha) &= \inf \{x : P[X > x] \leq 1 - \alpha\} \\ &= \inf \{x : P(X \leq x) \geq \alpha\} \\ &= \sup \{x : F(x) < \alpha\} \\ &= Q(\alpha) \end{aligned}$$

i.e, la VaR est le seuil minimum excédé par X avec la probabilité tout au plus $1 - \alpha$

La VaR a les propriétés suivantes:

- $VaR_\alpha[X] \leq \max[X]$ pour tout $\alpha \in (0, 1)$;

- La VaR est monotone: $VaR_\alpha [X] \leq VaR_\alpha [Y]$ si $X \leq Y$;
- La VaR est translation invariant, $VaR_\alpha [X + c] = VaR_\alpha [X] + c$;
- La VaR est positive homogène $VaR_\alpha [cX] = c \times VaR_\alpha [X]$.
- La VaR est comonotone additive (i.e, there is no diversification for perfectly dependent risks): Si les risques X_1, X_2, \dots, X_n sont comonotons, alors

$$VaR_\alpha [X_1 + X_2 + \dots + X_n] = VaR_\alpha [X_1] + VaR_\alpha [X_2] + \dots + VaR_\alpha [X_n]$$

Les deux dernières propriétés découlent également de la relation générale suivante.

Proposition 3.2.2. *(La VaR d'une fonction de v.a.) La VaR d'une v.a. $Y = g(X)$, où $g(\cdot)$ est une fonction non-décroissante d'une v.a. X , peut être calculée comme suivant*

$$VaR_\alpha [X] = g(VaR_\alpha [X])$$

Malgré sa simplicité, La mesure de risque de référence, Value-at-Risk (VaR), a été critiquée. Tout d'abord, la VaR ne fournit pas d'informations sur la gravité des pertes au-delà du seuil indiqué. Deuxièmement, la VaR ne satisfait pas la propriété de la subadditivité (donc ne pas cohérente) dans le sens d'Artzner et al. [2]. Il peut arriver que la diversification conduise à une augmentation de la VaR, notamment dans le cas où les v.a ont des distributions à queues lourdes (moyennes infinies). Ainsi, la VaR d'une somme de pertes n'est pas nécessairement inférieure ou égale à la somme des VaR de pertes individuelles. La VaR est subadditive pour les pertes distribuées elliptiquement (voir, par exemple, J. McNeil et al. [49]).¹ Alors que la mesure VaR indique un seul point de la queue de la distribution, il ne fournit aucune information sur l'ampleur des pertes qui peuvent survenir au-delà de la VaR. La VaR peut conduire à une mauvaise interprétation des risques: par exemple les portefeuilles qui affichent une même VaR ne sont pas nécessairement soumis aux mêmes risques.

Pour couvrir ces inconvénients, d'autres mesures de risque cohérentes ont été développées.

¹Une loi continue est dite elliptique de paramètre de position μ et de matrice de forme symétrique définie positive Σ si sa densité p peut s'écrire $p(x) = (\det \Sigma)^{-1/2} q((x - \mu)^t \Sigma^{-1} (x - \mu))$, où q est une fonction réelle à valeurs positives telle que $\int_{\mathbb{R}^n} q(\|y\|^2) dy = 1$. Cela peut être vu comme la généralisation d'une loi gaussienne.

3.2.3 La Tail Value-at-Risk

En plus de souligner le manque de cohérence de la VaR, Artzner et al. [2, 1999] a été aussi proposé un member de mesures de risque cohérentes comme une mesure de risque alternative: la Tail Conditional Expectation ou Tail Value-at-Risk (notée TCE ou $TVaR$). Une telle mesure de risque est étudiée intensivement par différents auteurs dans divers contextes, par conséquent, est parfois définie et différemment nommée. Wirch and Hardy [91, 1999] ont défini la même mesure de risque par Conditional Tail Expectation (CTE) et nous suivrons leur terminologie dans cette thèse. Rockafellar and Uryasev [75, 2000] ont nommé leur mesure de risque le Conditional Value-at-Risk (CVaR) (Valeur-à-Risque conditionnel (CVaR)) ; Bertsimas et al. [7] ont nommé une mesure de risque par semblable Mean Shortfall et une mesure de risque semblable avec un nom apparemment identique Expected Shortfall est étudiés par Acerbi et al. [1, 2001]. Nous suivrons la terminologie de Wirch and Hardy [91, 1999] (Conditional Tail Expectation (CTE)) dans ce qui suit.

La Tail Value-at-Risk(TVaR) ou Conditional Value-at-Risk (CVaR) ou encore Expected Shortfall(ES) à $\alpha\%$ représente les pertes moyennes attendues au-delà de la VaR pour un niveau de confiance de $(1 - \alpha\%)$. Elle traduit la moyenne équipondérée de toutes les VaR calculées à des niveaux de risques inférieurs ou égaux à α . Elle améliore la VaR en prenant en compte l'ampleur des pertes. La notoriété dont jouie cette mesure face à la VaR réside dans sa capacité à prendre en compte les pertes extrêmes et rares. Il s'agit d'une mesure cohérente au sens d'Artzner et al.(1999).

3.2.3.1 Définitions

La Conditional-Tail-Expectation (CTE) au niveau de probabilité α , appelée aussi la Tail-Conditional-Expectation, est une mesure de risque définie comme l'espérance de la v.a. X au delà de la VaR.

Définition 3.2.3. (*La Conditional-Tail-Expectation*) *La Conditional-Tail-Expectation d'une v.a. $X \sim F(x)$ au niveau de probabilité α , notée $CTE_\alpha [X]$, est*

$$CTE_\alpha [X] = \mathbb{E}[X|X > VaR_\alpha [X]] \quad (3.7)$$

ce qu'est l'espérance de la v.a. X dans les cas ou cette v.a. prend des valeurs supérieures à VaR, avec $CTE_0 [X] = \mathbb{E}[X|X > VaR_0 [X]] = \mathbb{E}[X]$

Cette mesure vérifie les propriétés 2 à 7. Dans le cas continu, la propriété de sous-additivité(Axiome S) est vérifiée et la mesure est cohérente. A noter

que dans ce cas, la CTE coïncide alors avec la TVaR et nous pouvons évaluer l'une ou l'autre indifféremment.

Nous considérons la perte au-dessus de VaR conditionnel étant excédé, i.e

$$X - VaR_\alpha[X] | X > VaR_\alpha[X]$$

Le moyen de cet excès conditionnel, appelé la mesure Conditional VaR, est noté par $CVaR_\alpha[X]$ et défini comme suit

Définition 3.2.4. (*La Conditional Value-at-Risk*) La Conditional Value-at-Risk d'une v.a. $X \sim F(x)$ au niveau de probabilité α , notée $CVaR_\alpha[X]$, est

$$\begin{aligned} CVaR_\alpha[X] &= \mathbb{E}[X - VaR_\alpha[X] | X > VaR_\alpha[X]] \\ &= CTE_\alpha[X] - VaR_\alpha[X] \end{aligned} \quad (3.8)$$

ce qu'est la valeur moyenne des pertes qui excèdent la VaR, i.e. il s'agit de l'excédent moyen de sinistre au-delà de la VaR.

Définition 3.2.5. (*L'Expected shortfall*) L'Expected shortfall au niveau α , notée $ES_\alpha[X]$ est

$$\begin{aligned} ES_\alpha[X] &= \mathbb{E}[(X - VaR_\alpha[X])_+] \\ &= \mathbb{E}[X - VaR_\alpha[X] | X > VaR_\alpha[X]] \mathbb{P}(X > VaR_\alpha[X]) \\ &= (1 - \alpha)CVaR_\alpha[X] \end{aligned} \quad (3.9)$$

ce qu'est la prime stop-loss dont la rétention est fixée à $VaR_\alpha[X]$. Où $(x - d)_+ = \max(x - d, 0)$, $d > 0$.

et nous avons, d'après les équations 3.8 et 3.9

$$\begin{aligned} CTE_\alpha[X] &= VaR_\alpha[X] + CVaR_\alpha[X] \\ &= VaR_\alpha[X] + \frac{1}{1 - \alpha} ES_\alpha[X] \end{aligned}$$

ce qui relie CTE_α au ES_α .

Une autre mesure de risque qui est de plus en plus considérée dans la littérature est la Tail-Value-at-Risk (TVaR). La TVaR est définie pour un niveau de confiance α , $0 \leq \alpha < 1$, comme

$$TVaR_\alpha[X] = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_s[X] ds \quad (3.10)$$

ce qui s'appelle parfois le Tail value-at-risk, noté par $TVaR_\alpha[X]$. et est la moyenne arithmétique du VaRs de X de s à 1, avec $TVaR_0[X] = \int_0^1 VaR_s[X] ds =$

$\mathbb{E}[X]$. Elle peut également s'écrire comme suit:

$$TVaR_\alpha[X] = \frac{1}{1-\alpha} (\mathbb{E}(X \mathbf{1}_{\{X > VaR_\alpha[X]\}}) + VaR_\alpha[X] (\mathbb{P}(X \leq VaR_\alpha[X]) - \alpha)) \quad (3.11)$$

Elle prend évidemment en compte les valeurs de la distribution au delà de la VaR au niveau de confiance α et nous informe donc sur l'épaisseur de la queue de distribution. si l'on considère une v.a. X continue, alors $\mathbb{P}(X \leq VaR_\alpha[X]) = \alpha$. L'équation (3.11) devient ainsi $TVaR_\alpha[X] = \mathbb{E}[X|X > VaR_\alpha[X]]$ qui est exactement la définition de la CTE_α , Donc CTE peut s'écrire comme suit

$$\begin{aligned} CTE_\alpha[X] &= \mathbb{E}[X|X > VaR_\alpha[X]] \\ &= \frac{1}{1-\alpha} \int_\alpha^1 VaR_s[X] ds \end{aligned} \quad (3.12)$$

3.2.3.2 Propriétés

Alors que la TVaR sera cohérente aussi bien pour des v.a. continues que non continues, plus, dans ce même contexte de non-continuité, la CTE ne respecte pas toutes les propriétés d'une mesure de risque cohérente. Quand X est continue, CTE satisfait les Axioms M, T, PH, and S, et est donc cohérente. La CTE est aussi satisfait toutes les propriétés désirables énoncées précédemment.

3.3 Principes de calcul de primes

Une prime d'assurance est le montant que le preneur d'assurance doit payer pour pouvoir bénéficier de la couverture d'assurance en cas de sinistre. Le montant de cette prime dépend de nombreux facteurs. Dans cette partie, nous discutons de certaines propriétés essentielles qui devraient être satisfaites par un principe de calcul de prime, pour être un principe idéal. Ensuite, nous définissons quelques principes de calcul de primes populaires existants.

Nous notons par $\Pi(X)$ la prime chargée par l'assureur pour couvrir le risque X . Toute règle qui permet d'associer à un risque X , une prime $\Pi(X)$ est appelée un principe de prime. Les propriétés suivantes sont des propriétés désirables pour un principe de prime et la liste suivante présente les propriétés les plus importantes en pratique.

1. Non chargement excessive. Pour une mesure de risque X donnée, si $X \leq x_{\max}$, alors $\Pi[X] \leq X_{\max}$.

Cette propriété spécifie que la prime ne dépasse pas ce le montant maximal. Dans le cas contraire, il n'y aurait aucune raison pour l'assuré de souscrire à la couverture d'assurance.

3. Chargement injustifié. Soit a une constante quelconque. Nous devons toujours avoir $\Pi[a] = a$.

Ce qui signifie que la mesure d'un montant certain est ce montant lui-même.

2. Chargement de sécurité: $\Pi(X) \geq \mathbb{E}(X)$

Cette propriété exige que la prime soit supérieure ou égale à la prime pure. Cette propriété permet à l'assureur de pouvoir résister à la volatilité naturelle des sinistres. Le capital requis doit ainsi excéder les pertes espérés (sous peine de ruine pour une compagnie d'assurance, par exemple). Ce qui signifie que la mesure d'un risque doit être plus grande que son espérance.

4. Additivité: $\Pi(X + Y) = \Pi(X) + \Pi(Y)$ Cette propriété garantit qu'il n'y a aucune différence entre assurer la totalité des risques ou répartir la couverture sur plusieurs contrats.

5. Homogénéité positive: pour toute constante $a \geq 0$, $\Pi(aX) = a\Pi(X)$

Si le montant des réclamations subit un changement d'échelle, à cause de l'inflation par exemple, cette propriété spécifie que la prime pour le nouveau risque aX est proportionnelle à la prime du risque de base $\Pi(X)$. Cette propriété garantit à l'assureur le même niveau de profitabilité en cas de ce changement d'échelle.

6. Invariance par translation: pour toute constante $a \geq 0$, $\Pi(X + a) = \Pi(X) + a$

Cette propriété garantit que le principe de prime donne un résultat logique face à de nouveaux coûts fixés par un contrat.

7. Sous-additivité. Soient X, Y deux risques. On doit avoir $\Pi(X + Y) \leq \Pi(X) + \Pi(Y)$,

Ce principe prévient qu'un individu ou une compagnie d'assurance partage un risque en deux et qu'il en tire avantage. De plus, la mutualisation des coûts de plusieurs contrats doit conduire à une réduction des primes demandées pour l'ensemble des contrats.

Les principes de prime sont les mesures de risque les plus importantes en actuariat. Les propriétés qu'il possède un principe de prime sont déterminantes pour le choisir. Nous donnons maintenant des principes de prime. Pour plus de détails, on se réfère au Young [94]:

A - Principe de la prime nette. Ce principe est le plus simple de principes de primes et est égale à la valeur espéré de pertes:

$$\Pi(X) = \mathbb{E}(X)$$

Ce principe n'inclut aucun chargement de sécurité et ne peut donc être utilisé pour établir la prime finale. Ce principe reste tout de même intéressant pour donner une idée de la prime par sa simplicité. En plus, il respecte toutes les autres propriétés.

B - Principe de la valeur espérée. Le principe de la valeur espérée est donnée par

$$\Pi(X) = (1 + \theta)\mathbb{E}(X), \quad \theta > 0$$

où θ et $\theta\mathbb{E}(X)$ sont les chargements de sécurité relative aet total, respectivement. Ici on a besoin de seulement un paramètre de la distribution du risque, à savoir le moyen. Ce principe contient un chargement de sécurité positive, est additif et proportionnel. Cependant il n'est pas cohérent, ni plafonné. Un inconvénient de ce principe est que tous les risques avec le même moyen ont la même prime. Intuitivement, un risque avec une grande dispersion est plus dangereux qu'un risque avec le même moyen et une petite dispersion et par conséquent devrait avoir une prime plus élevée. Ainsi il semble que une certaine mesure de dispersion devrait être incluse dans la formule de la prime. Cette idée est prise en compte par les deux principes suivantes.

C - Principe de la variance. Le principe de la variance est donné par

$$\Pi(X) = \mathbb{E}(X) + \theta Var(X), \quad \theta > 0$$

Ce principe modifie le principe de la valeur espérée en faisant dépendre la majoration de la variance du risque. Cette modification permet de donner une prime différente à des risques qui n'ont pas la même variance. Ce principe a aussi l'avantage d'être cohérent. Cependant, il n'est pas proportionnel, Il n'est pas plafonné non plus, car pour une certaine valeur de θ , il est possible d'avoir $\Pi(X) > X_m$.

D - Principe de l'écart-type. Le principe de l'écart-type est donné par

$$\Pi(X) = \mathbb{E}(X) + \theta\sqrt{\text{Var}(X)}, \quad \theta > 0$$

Ce principe est inspiré du principe de la valeur espérée et du principe de la variance. Ceci lui permet de bénéficier des avantages de chacun. Ce principe est proportionnel, Cependant, il n'est pas additif, car les écarts-types ne s'additionnent pas, il n'est pas plafonné.

E - Principe de l'utilité équivalente:

$$u(w) = \mathbb{E}(u(w - X + \Pi(X)))$$

Soit u la fonction d'utilité de l'assureur et w son surplus initial. Ce principe établi la prime comme le montant minimal que l'assureur doit charger, conformément à son aversion au risque, pour offrir la couverture. Ce principe contient un chargement de sécurité positive puisque, avec l'inégalité de Jensen

$$u(w) = \mathbb{E}(u(w - X + \Pi(X))) \leq u(w - X + \Pi(X)).$$

Le principe est cohérent et plafonné, cependant, il n'est pas additif ni proportionnel.

F - Principe exponentiel:

$$\Pi(X) = \frac{1}{\beta} \ln \mathbb{E}(\exp(\beta X)), \quad \beta > 0$$

Ce principe découle du principe de l'utilité équivalente lorsque la fonction d'utilité u est exponentielle, i.e. $u = -\exp(-\beta x), \beta > 0$. Ce principe possède la propriété de ne pas dépendre du surplus initial. Il est aussi proportionnel, cependant, comme le principe de l'utilité équivalente, il n'est pas additif.

G - Principe d'Esscher:

$$\Pi(X) = \mathbb{E}(Xe^{hX})/\mathbb{E}(e^{hX})$$

Ce principe découle du principe de l'utilité exponentiel. Il survient lorsque l'assureur vise à optimiser son utilité selon le principe de l'utilité équivalente

$$\max_{\Pi} u(w) - h\mathbb{E}(u(w - X + \Pi(X))), h > 0.$$

Ce principe peut aussi être vu comme une pondération du risque où plus de poids est donné aux événements extrêmes. Il possède l'avantage d'être additif. Cependant, il n'est pas proportionnel.

H - Principe de distortion.

$$\Pi(X) = \int_0^\infty \psi(S(t))dt$$

où $S = 1 - F$ est la fonction de survie du risque X et ψ est une fonction de distortion.

Ce principe est semblable au principe d'Esscher en ce sens qu'il produit aussi une pondération du risque où plus de poids est donné aux événements extrêmes. Contrairement au principe d'Esscher, il est proportionnel, mais pas additif, Ce principe vérifie toutes les autres propriétés. Ce principe est présenté dans la section suivante.

Les principes de primes sont aussi traités dans Rolski et al. [76, 1999] et Kaas et al. [56].

3.4 Mesure du risque de distorsion

Le risque de réassurance est à queue lourde et contient un risque systématique. Les primes basées sur la valeur actuarielle ou plus généralement sur les moments sous une mesure de probabilité ne reflètent pas suffisamment la queue de ce risque.

Les principes de primes présentées dans la subsection 3.3 définissent le risque basé sur un chargement de l'espérance de perte. L'espérance $\mathbb{E}(X)$ d'une perte aléatoire continue non négative X de fdr F et fonction de survie $S(x) = P(X > x) = 1 - F(x) = \bar{F}(x)$, peut être s'écrit comme

$$\mathbb{E}(X) = \int_0^\infty S(x)dx \quad (3.13)$$

Ainsi, au lieu d'ajouter un chargement à $\mathbb{E}(X)$ pour obtenir une prime nous pouvons redéfinir la distribution des pertes en décalant plus de probabilité pesant aux pertes élevées. Une façon d'atteindre cet objectif consiste à appliquer une distortion sur la fonction de survie. Ce procédure est introduit par Wang [86, 87, 1995,1996] et est étroitement liés à la théorie d'espérance distordue Yaari [93, 1987]. Ces principes de prime sont étudiés comme des mesures de risque dans Wang [88, 89, 2000,2001] et sont connus par mesures de risque de distorsion depuis lors. Le Tail Value at Risk (TVaR) ainsi que le Value at Risk (VaR) et PH-trasform sont des cas particuliers de mesures de risque de distorsion (ou principes de primes de distortion). Avant de définir la mesure du risque de distorsion (ou principes de primes de distortion), nous devons d'abord introduire quelques définitions utiles.

Définition 3.4.1. (fonction de distorsion) Une fonction $\psi : [0, 1] \rightarrow [0, 1]$ est une fonction de distorsion si:

1. $\psi(0) = 0$ et $\psi(1) = 1$.
2. $\psi(\cdot)$ est non-décroissante et continue.

Définition 3.4.2. (probabilité distordue) Soient (Ω, \mathcal{F}, P) un espace de probabilité et $\psi : [0, 1] \rightarrow [0, 1]$ une fonction de distorsion. Soit la fonction d'ensembles $\psi \circ P$ définie par $\psi \circ P(A) := \psi(P(A)), \forall A \in \mathcal{F}$. La fonction $\psi \circ P$ est appelée probabilité distordue.

Nous exposons maintenant la notion de mesures de risque de distorsion (ou principes de primes de distorsion). Soient X une v.a non-égative modélise une risque (ou perte) de fdr F et fonction de survie $S(x) = P(X > x) = 1 - F(x) = \bar{F}(x)$, la moyenne de perte ou la prime nette est

$$\mathbb{E}(X) = \int_0^{\infty} S(x)dx$$

La contribution de Wang est de devait suggérer une mesure de risque basée sur une fonction de distorsion $\psi(S(x))$. La mesure de risque de distorsion (ou principes de primes de distorsion) associée à la fonction de distorsion ψ notée par $\Pi_{\psi}[\cdot]$ est défini par

$$\Pi_{\psi} [X] = \int_0^{\infty} \psi(S(x))dx, \quad (3.14)$$

pour toute v.a. de perte X .

L'idée principale de mesures du risque de distorsion (ou principes de primes de distorsion) est de quantifier le risque ou de dériver la prime de risque de distorsion, ce qui contiennent une chargements de sécurité par distorsion. Ainsi, par le changement des probabilités originelles de sorte que les probabilités pour les événements défavorables soient artificiellement gonflées (il distord la distribution de la probabilité du risque, assignant une probabilité plus élevée aux événements défavorables). Plus précisément, un chargement de risque est impliquée en utilisant une mesure de probabilité distordue Q au lieu de la mesure objective P pour calculer une perte espérée gonflée comme une prime de risque brute tel que

$$\mathbb{E}(X) = \int_0^{\infty} S(x)dx \leq \int_0^{\infty} \psi(S(x))dx = \Pi_{\psi}(X)$$

Donc, la mesure du risque de distorsion peut être interprété comme un ajustement de la vraie mesure de probabilité pour donner plus de poids aux événements de risque plus élevé. L'expression 3.14 est appelée aussi **mesure du**

risque ajustée ou **mesure du risque de Wang**. Cette expression peut s'écrire sous des formes alternatives

$$\begin{aligned}\Pi_\psi[X] &= \int_0^\infty \psi(S(x))dx \\ &= \int_0^\infty x\psi'(S(x))f(x)dx\end{aligned}\quad (3.15)$$

$$= \int_0^1 \psi'(1-s)Q(s)ds \quad (3.16)$$

$$= \int_0^1 VaR_{1-s}[X]d\psi(s) \quad (3.17)$$

De (3.15), Nous pouvons considérer ainsi $\psi'(S(x))$ comme un mécanisme pesant qui escompte la probabilité des événements souhaitables (pour X élevé), tout en chargeant la probabilité d'événements défavorables (pour X bas). la représentation (3.16) est plus commode pour développer les estimateurs empiriques de la mesure du risque Π . la représentation (3.17) indique que les mesures de risque de distortion sont des moyennes pondérées de VaRs.

Dans le cas où la v.a. $X \in \mathbb{R}$, la mesure de risque de distortion est définie comme suit

$$\Pi_\psi[X] = - \int_{-\infty}^0 [1 - \psi(S(x))] dx + \int_0^\infty \psi(S(x))dx, \quad (3.18)$$

Notons que la deuxième intégrale dans l'expression 3.18 est égale à celui dans l'expression 3.14. Une d'interprétation simple d'une mesure du risque de distorsion est le suivant: d'abord, la fonction de survie de la variable aléatoire est distordue ($\psi(S)$); deuxièmement, l'espérance mathématique de la variable aléatoire distordue est calculée.

Remarque 3.4.3.

- Notons que l'espérance mathématique $\mathbb{E}[X]$ est une mesure de risque de distortion dont la fonction de distortion est la fonction d'identité $\psi(s) = id(s) = s$.
- Notons en outre que $\psi_1(s) \leq \psi_2(s)$ pour tout $s \in [0, 1]$ implique que $\rho_{\psi_1}[X] \leq \rho_{\psi_2}[X]$.
- On constate immédiatement que $\psi(S(x))$ est une fonction non-croissante de x avec des valeurs dans l'intervalle $[0, 1]$. Ainsi, la fonction de distortion $\psi(S(x))$ peut être considérée comme un fdr du risque ajustée. Cependant $\Pi_\psi[X]$ ne peut pas toujours être considéré comme espérance de X sous une nouvelle mesure de probabilité, parce que $\psi(S(x))$ ne sera pas nécessairement continue à droite.

- Pour une fonction de distorsion générale ψ , la mesure de risque $\Pi_\psi[X]$ peut être interprétée comme un intégral de Choquet (voir Denneberg [29, 1994]). La représentation de mesures de risque de distorsion comme un intégral de Choquet est utilisée pour explorer ses propriétés mathématiques, le calcul de mesures de risque de distorsion peut être facilement faite en prenant l'espérance de X sous la mesure de probabilité distordue $Q(A) = \psi(P(A))$.

Les propriétés et plusieurs exemples de fonctions de distorsion viables $\psi(\cdot)$ sont fournis dans les discussions de Wirch and Hardy [92, 2001].

3.4.1 Propriétés de mesures de risque de distorsion

Les mesures de risque de Wang ont beaucoup de propriétés commodes. Dans cette sous-section nous fournissons une liste de diverses propriétés des mesures de risque de distorsion.

Proposition 3.4.4. (*Propriétés de mesures de risque de distorsion*). Soient ψ une fonction de distorsion et. Une mesure de risque de distorsion Π_ψ a les propriétés suivantes

- (a) Homogénéité positive ,
- (b) Invariantes par translation,
- (c) Monotonéité,

Si de plus ψ est une fonction de distorsion concave:

- (d) Wirch and Hardy [91, 1999] ont montré qu'un mesure de risque de distorsion est subadditive, par conséquent cohérente dans le sens d'Artzner et al. 1999 si et seulement si sa fonction de distorsion ψ est concave, i.e, les fonctions de distorsion concave générer les mesures de risque cohérentes. La concavité de la fonction de distorsion est une condition nécessaire et suffisante pour que la mesure du risque soit cohérente.

- (b) Les mesures de risque de Wang Π_ψ sont comonotonement sous-additives.

Corollaire 3.4.5. (*Cohérence*). Une mesure de risque de distorsion est cohérente si et seulement si la fonction de distorsion ψ est concave.

En choisissant une fonction de distorsion convenable ψ , nous regroupons dans le Tableau 3.4.2 ci-dessous, certaines mesures de risque de distorsion très fréquentes dans la littérature actuarielle.

3.4.2 Exemples de mesures de risques de distorsion

Pour des choix appropriés de $\psi(\cdot)$, les mesures de risque bien connus tels que la VaR et l'ES sont des mesures du risque de distorsion. En particulier, pour un niveau de confiance $\alpha \in (0, 1)$:

- VaR correspond à

$$\psi_{VaR}(x, \alpha) = \begin{cases} 1, & \text{si } x > 1 - \alpha, \\ 0, & \text{sinon,} \end{cases}$$

Ce qui n'est pas une fonction concave la Value-at-risk n'est pas cohérente (car elle n'est pas sous-additive) ;

- ES correspond à

$$\psi_{TVaR}(x, \alpha) = \min\left(1, \frac{x}{1 - \alpha}\right) \quad (3.19)$$

ce qu'est une fonction de distorsion concave donc l'TVaR est une mesure de risque cohérente. La connexion entre TVaR et la fonction de distorsion (3.19) a été observée dans Wirch and Hardy [91, 1999].

De la même manière il est possible de représenter de nombreuses mesures de risque sous forme de mesure de Wang.

- Mesure de risque PHP. La mesure de risque PHP a été introduite par Wang [86, 1995]. La fonction de distorsion pour PHP est donnée par:

$$\psi_{PHP}(t, \xi) = t^{\frac{1}{\xi}}, \quad \forall \xi \geq 1$$

la mesure de risque de distorsion est définie par

$$\Pi_{PHP}[t] = \int_0^\infty (1 - F(t))^{\frac{1}{\xi}} dt$$

Le paramètre $\xi \geq 1$ représente le coefficient de distorsion ou indice d'aversion au risque. Et pour $\xi = 1$, $\Pi[X] = \mathbb{E}(X)$. Il est clair que ψ_{PHP} est concave pour $\xi \geq 1$ et par conséquent la mesure PHP est cohérente.

Pour un niveau de rétention $R > 0$, le PHP ou la prime de réassurance est définie comme suit

$$\Pi_{\xi, R}[X] = \int_R^\infty (\bar{F}(x))^{\frac{1}{\xi}} dx$$

- Mesure de risque Dual-power. La mesure de risque Dual-power est obtenue en considérant la fonction de distortion:

$$\psi_{DP}(t, \xi) = 1 - (1 - t)^\xi, \quad \forall \xi \geq 1$$

la mesure de Wang est définie par

$$\Pi [t] = \int_0^\infty (1 - (F(t))^\xi) dt$$

- Mesure de risque Wang Transform (WT). La mesure de risque Wang Transform a été présentée par Wang [89, 2000], dont la fonction de distortion est donnée par

$$\forall t \in [0, 1] : \quad \psi_{WT}(t, \beta) = (\Phi^{-1}(t) + \Phi^{-1}(\beta)), \quad \beta \in [0, 1]$$

ce qui s'appelle le "Transformation de Wang au niveau β ", la mesure de risque de distortion correspondante qui s'appelle "la mesure de risque Wang transform et est notée par $WT_p[X]$. Avec Φ est la fdr de la loi normale centrée réduite $\mathcal{N}(0, 1)$.

$$\Pi [t] = \int_0^\infty (\Phi^{-1}((1 - (F(t)))) + \Phi^{-1}(\beta)) dt$$

- La mesure de risque de LB-transform (Lookback(LB) distortion) a été présentée par Hürlimann [48] pour étudier des stratégies d'option et pour modéliser le risque d'échange. La fonction de distortion pour LB-transform est donnée par

$$\psi_{LB}(t, \xi) = t^\xi(1 - \xi \ln t), \quad \xi \in (0, 1]$$

ψ_{LB} est une fonction concave pour $\xi \in (0, 1]$, par conséquent LB_ξ est cohérente.

- La mesure de risque Beta distortion dont la fonction de distortion est donnée par la fonction beta

$$\psi(x) = \frac{1}{\beta(a, b)} \int_0^x t^{a-1}(1-t)^{b-1} dt, \quad a > 0, b > 0, \beta(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt.$$

Posons $b = 1$ ce qui donne power distortion $\psi(x) = x^a$. Posons $a = 1$ ce qui donne dual-power distortion $\psi(x) = 1 - (1 - x)^b$.

Pour plus de détails sur les mesures du risque de distorsion, voir Denuit et al. [30, 2005, sections 2.6.2 et 2.6.3].

Mesure de risque	Fonction de distortion	Restrictions pour Mesures de Risque Cohérentes
Beta-distortion	$\psi(t) = \int_0^t \frac{1}{\beta(a,b)} s^{a-1} (1-s)^{b-1} ds$	$\beta(a,b) = \int_0^1 s^{a-1} (1-s)^{b-1} ds$ $a \leq 1, b \geq 1$
Proportional hasard transform	$\psi(t) = t^a$	$a \leq 1$
Dual power function	$\psi(t) = 1 - (1-t)^a$	$a \geq 1$
Gini principle	$\psi(t) = (1+a)t - at^2$	$0 \leq a \leq 1$
Wang transform	$\psi(t) = (\Phi^{-1}(t) + \Phi^{-1}(p))$	
Lookback	$\psi(t) = t^a (1 - a \ln t)$	$0 < a < 1$
Conditional Tail Expectation	$\psi(x) = \min\left(1, \frac{x}{1-\alpha}\right)$	
Value at risk	$\psi(x) = \begin{cases} 1, & \text{si } x > 1 - \alpha, \\ 0, & \text{sinon,} \end{cases}$	

Tableau 3.1: Un Aperçu de Mesures de Risque de Distortion

3.5 Estimation de mesures de risque de distortion

L’objectif de cette partie est de donner différentes approches d’estimation pour les primes de risques de distortion de pertes à queue lourdes. Nous discutons respectivement des approches par les quantiles empirique et par les quantiles extrêmes (Approche semi-paramétrique). L’attention principale est sur les estimateurs semi-paramétrique. En particulier, nous donnons les définitions des estimateurs, des exemples ainsi que la normalité asymptotique. Dans l’approche par les quantiles empirique, nous estimons la fonction de distribution des pertes par la fonction de répartition empirique. Une discussion est donnée par Jones and Zitikis [52, 2003].

Considérons maintenant un ensemble des observations indépendants et identiquement distribuées (i.i.d) (X_1, X_2, \dots, X_n) d’une v.a X (le risque) avec fdr F et fonction quantile $Q = F^{-1}$ et soient $X_{(1:n)} \leq X_{(2:n)} \leq \dots \leq X_{(n:n)}$ les statistiques d’ordre correspondantes. La fonction quantile Q peut être estimée par le quantile empirique (d’échantillon) Q_n défini par:

$$Q_n(s) = \inf \{t > 0, F_n(t) \geq s\}, \quad s \in [0, 1].$$

où $F_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{X_j \leq x}$ est la distribution empirique de l’échantillon et $\mathbb{1}_A$ est la fonction indicatrice de l’ensemble A .

Un estimateur naturel pour $\Pi_\psi[X]$ par l’approche empirique est obtenu en remplaçant le vrai quantile Q par la fonction quantile empirique Q_n :

$$\Pi_{\psi,n}(X) = \sum_{i=1}^n \left(\psi\left(\frac{i}{n}\right) - \psi\left(\frac{i-1}{n}\right) \right) X_{n-i+1:n} \tag{3.20}$$

L'estimateur $\Pi_{\psi,n}()$ peut être vu comme une L -statistique qu'est une combinaison linéaire des statistiques d'ordre (voir par exemple Shorack and Wellner [78, 1986]. Pour une application récente d'une telle classe des statistiques dans l'estimation des mesures de risque, on se rapporte à Jones and Zitikis [52, 53, 54, 2003, 2005, 2007].

Pour certains inférence statistique sur les primes de risque de distorsion on pourra se référer aux Peng [67] Jones and Zitikis [52, 2003], Necir and Boukhetala [62], Centeno de Lourdes and Andrade e Silva [14], Necir et al. [65, 2007], Jones and Zitikis [54], Brazauskas et al. [12], Furman and Zitikis [40, 41], Greselin et al. [43], Necir and Meraghni [63, 64], Brahimy et al. [10, 11], Peng et al. [69], Rassoul [72], Deme et al. [27]

En utilisant la relation entre les mesures de risques de distorsion et les L -statistiques, Jones et Zitikis (voir ,Jones and Zitikis [52, 2003, Théoreme 3.2]) ont été établi la normalité asymptotique de l'estimateur de $\Pi_{\psi}(X)$. En effet, ils ont utilisé la théorie asymptotique pour les L -statistiques (par exemple, Shorack and Wellner [78, 1986]) et montrent que, pour des fdr F avec un nombre suffisant de moments finis et dans certaines conditions de régularité sur la fonction de distorsion ψ , le résultat de normalité asymptotique suivant:

$$\sqrt{n} \left(\widetilde{\Pi}_{\psi,n}(X) - \Pi_{\psi}(X) \right) \xrightarrow{d} \mathcal{N}(0, \sigma_{\psi}^2); \quad \text{quand } n \longrightarrow \infty$$

à condition que la variance

$$\sigma_{\widetilde{\Pi}_{\psi,n}}^2 = \int_0^1 \int_0^1 (\min(s, t) - st) \psi'(s) \psi'(t) dQ(1-s) dQ(1-t) \quad (3.21)$$

i.e, pour les mesures de risque définies dans (3.14), Jones and Zitikis [52] fournissent les conditions dans lesquelles $(\widetilde{\Pi}_{\psi,n}(X) - \Pi_{\psi}(X))$ est asymptotiquement normale de moyen zéro et de variance fini, où F_n est l'estimateur empirique de F . La consistance des estimateurs suit alors immédiatement de leur normalité asymptotique sous des conditions correspondantes.

Exemples

Dans la suite, nous présentons quelques mesures du risque de distorsion et leurs expressions en termes de la fonction quantile. En appliquant le même principe dans (3.20), nous en déduisons des estimateurs non paramétriques fonctionnelles de mesures du risque de distorsion et nous dérivons la forme exacte de la variance asymptotique pour les fonctions de distorsion spécifiques associées. Nous considérons ci-dessous les exemples de VaR, CTE ou TVaR,

PHT.

VaR

Si $g(x) = \mathbf{1}_{x>\alpha}$ pour $p \in [0, 1]$. Alors

$$\begin{aligned} VaR_\alpha [X] &= \int_0^\infty \mathbf{1}_{\{1-F(x)>1-\alpha\}} dx & (3.22) \\ &= \int_0^\infty \mathbf{1}_{\{x>F^{-1}(\alpha)\}} dx \\ &= \int_0^{F^{-1}(\alpha)} 1 dx + \int_{F^{-1}(\alpha)}^\infty 0 dx \\ &= F^{-1}(\alpha) = Q(\alpha) \end{aligned}$$

qui est la quantile empirique $Q(\alpha)$, donc, l'estimateur nonparamétrique de VaR_α peut être écrit de manière suivante

$$\begin{aligned} VaR_\alpha [X] &= \sum_{i=1}^n X_{i:n} [\mathbf{1}_{\{\frac{i-1}{n} \leq 1-\alpha\}} - \mathbf{1}_{\{\frac{i}{n} \leq 1-\alpha\}}] & (3.23) \\ &= \begin{cases} X_{n(1-\alpha)}, & \text{si } n(1-\alpha) \text{ est un entier} \\ X_{[n(1-\alpha)]+1}, & \text{sinon} \end{cases} \end{aligned}$$

où $[y]$ désigne la partie entière de y . La variance asymptotique de l'estimateur de VaR est

$$\sigma^2 = \frac{\alpha(1-\alpha)}{[f(Q(\alpha))]^2}$$

Remarque 3.5.1. On peut remplacer $[n\alpha]$ par $[n\alpha] + 1$ ou en général par n'importe quelle autre suite k_n pourvu que $k_n/n \rightarrow \alpha$. Dans ce cas, l'estimateur sera donné par $X_{(k_n:n)}$.

CTE

$$\Pi_{CTE,\alpha}[X] = \frac{1}{1-\alpha} \int_\alpha^1 Q(s) ds. \quad (3.24)$$

En remplaçant Q par sa version empirique Q_n définie par $Q_n(s) = X_{i:n}$ pour tout $(i-1)/n < s \leq i/n$ et $i = 1, \dots, n$ avec $Q_n(0) = X_{1:n}$, on obtient

$$\begin{aligned} \tilde{\Pi}_{CTE,n,\alpha}[X] &= \frac{1}{1-\alpha} \int_\alpha^1 Q_n(s) ds & (3.25) \\ &= \frac{1}{1-\alpha} \sum_{i=[n\alpha]+1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} ds X_{i:n} \\ &= \frac{1}{n(1-\alpha)} \sum_{i=[n\alpha]+1}^n X_{i:n} \end{aligned}$$

Le comportement asymptotique de l'estimateur $\widetilde{\Pi}_{CTE,n,\alpha}[X]$ a été étudié par Brazauskas et al. [12, 2008, Théorème 3.1], lorsque $\mathbb{E}[X^2] < \infty$. Ce théorème peut être considéré comme une corollaire de Théorème 3.2 dans Jones and Zitikis [52, 2003]. En effet, la variance asymptotique est obtenue en substituant la fonction de distortion dans l'expression (3.21). Donc, la variance asymptotique est

$$\sigma_{\widetilde{\Pi}_{CTE,n,\alpha}}^2 = \frac{1}{(1-\alpha)^2} \int_{\alpha}^1 \int_{\alpha}^1 (\min(s,t) - st) dQ(s) dQ(t)$$

PHT

En remplaçant Q par sa version empirique Q_n définie par $Q_n(s) = X_{i,n}$ pour tout $(i-1)/n < s \leq i/n$ et $i = 1, \dots, n$ avec $Q_n(0) = X_{1:n}$, on obtient

$$\begin{aligned} \widetilde{\Pi}_{PHT,n,\xi}[X] &= \int_0^1 \frac{1}{\xi} s^{\frac{1}{\xi}-1} Q_n(s) ds \\ &= \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{\xi} s^{\frac{1}{\xi}-1} ds X_{i:n} \\ &= \sum_{i=1}^n \left(\left(\frac{i}{n} \right)^{\frac{1}{\xi}} - \left(\frac{i-1}{n} \right)^{\frac{1}{\xi}} \right) X_{n-i+1:n} \end{aligned} \quad (3.26)$$

La variance asymptotique est obtenue en substituant la fonction de distortion dans l'expression (3.21). Donc, La variance asymptotique est

$$\sigma_{\widetilde{\Pi}_{PHT,n,\xi}}^2 = \frac{1}{\xi^2} \int_0^1 \int_0^1 (\min(s,t) - st) s^{\frac{1}{\xi}-1} t^{\frac{1}{\xi}-1} dQ(s) dQ(t);$$

à condition que $\mathbb{E}|X|^\eta < \infty$ pour certains $\eta > \frac{2\xi}{2-\xi}$.

Le cas $\xi \geq 1/2$ n'est pas couvert dans la formule ci-dessus. Jones and Zitikis [52, 2003] notent que, dans le cas du PHT, ces conditions sont satisfaites **si et seulement si** $\xi < 1/2$ et est consistant pour tout $\xi \geq 1$ pourvu que $\mathbb{E}|X|^\eta < \infty$ pour $\eta > \xi$ et est asymptotiquement normal si $\xi \in [1, 2]$ avec la même condition sur le moment X mais cette fois limitée à $\eta > \frac{2\xi}{2-\xi}$. L'utilisation de quantiles empiriques pour estimer les primes de risque de distortion Π_ψ ne garantit pas la normalité asymptotique lorsque les pertes ont une distribution à queue lourde. dans ce cas, les résultats de Jones and Zitikis [52, 2003] ne peuvent plus être appliqués, du fait que le moment d'ordre deux de X est infini.

Ainsi, Jones and Zitikis [52] concluent qu'il s'agit toujours d'une question ouverte pour ce cas. Donc, Une réponse ou un nouvel approche pour gérer cette situation est requise. Par une approche basée sur des quantiles extrêmes,

Necir and Meraghni [63] ont donné une estimation alternative de la prime de risque de distortion dans le cas des pertes extrêmes (distributions à queues lourdes) et ont étudié sa normalité asymptotique.

Donc, dans ce contexte, il est assez naturel de supposer que les fonctions de distorsion ψ sont telles que $t \mapsto \psi(1/t)$ est à variation régulière à l'infini avec indice de variation régulière $\beta \geq 1$, c'est-à-dire,

$$\psi(1/t) = t^{1/\beta} \ell_\psi(t) \quad (3.27)$$

où $\psi(t)$ est à variation lente à l'infini, c'est-à-dire que $\ell_\psi(tx)/\ell_\psi(t) \rightarrow 1$ quand $t \rightarrow \infty$, pour tout $x > 0$.

Définition de l'estimateur de Necir and Meraghni (2009)

L'idée principale est la suivante. La formule (3.16) peut être divisé comme suit

$$\Pi_\psi[X] = \int_0^{k/n} Q(1-s) d\psi(s) + \int_{k/n}^1 Q(1-s) d\psi(s) \quad (3.28)$$

Dans la formule (3.28), en remplaçant $Q(1-s)$ par l'estimateur de quantile extrême $Q_n^W(1-s)$ (Estimateur de Weissman [90, 1978]) dans la première intégrale et par le quantile empirique dans la deuxième intégrale, Necir and Meraghni [63, 2009] ont proposé un estimateur alternatif pour $\Pi_\psi(X)$ comme suit:

$$\begin{aligned} \widetilde{\Pi}_\psi[X] &= \int_0^{k/n} \psi(s) dQ_n^W(1-s) - \int_{k/n}^1 \psi(s) dQ_n(1-s) \\ &= \psi(k/n) \frac{X_{n-k,n}}{1 - \widehat{\gamma}\rho} + \sum_{i=k+1}^n c_{i,n}(\psi) X_{n-i+1:n} \end{aligned} \quad (3.29)$$

où $c_{i,n}(\psi) = (\psi(i/n) - \psi((i-1)/n))$ et $\rho\widehat{\gamma}^H \in]0, 1[$.

L'estimation de $\Pi_\psi[X]$ a été étudiée aussi par Centeno et Andrade e Silva (2005) en utilisant les techniques de bootstrap. Dans le cas des excès élevés de stratification de pertes $R \rightarrow \infty$, la technique d'estimation de $\Pi_{\xi,R}(X)$ est différente du cas où $R = 0$. Ce cas est traité par différents auteurs, par exemple, Necir and Boukhetala [62], Vandewalle and Beirlant [83, 2007] et Necir et al. [65, 2007] ont introduit et étudié différents estimateurs de $\Pi_{\xi,R}$ basés sur des échantillons de montant de revendications sur les couvertures de réassurance pour des pertes à queues lourdes.

En remplaçant Q par sa version empirique Q_n définie par $Q_n(s) = X_{i,n}$ pour tout $(i-1)/n < s \leq i/n$ et $i = 1, \dots, n$ avec $Q_n(0) = X_{1:n}$, Necir et al. [65] ont proposé un estimateur semi-paramétrique de $\Pi_{\xi,R_{opt}}(X)$ pour une valeur fixée $\xi \geq 1$ et un seuil optimal de la rétention $R_{opt} = F^{\leftarrow}(1-k/n)$ où $k = k_n$ est une

suite intermédiaire avec comme estimateur de R_{opt} la valeur $\hat{R}_{opt} = X_{n-k:n}$.

$$\tilde{\Pi}_{\xi, \hat{R}_{opt}}(X) = \left(\frac{k}{n}\right)^{\frac{1}{\xi}} \frac{\xi}{1/\hat{\gamma}_{k,n}^{(H)} - \xi} X_{n-k:n}, \text{ pour } \hat{\gamma}_{k,n}^{(H)} < 1/\xi$$

Deme and Lo [28] regroupent les estimateurs des primes de risque de distorsion Π_ψ sous une forme général comme suit

$$\tilde{\Pi}_n^*(\psi) = \tilde{\Pi}_n(\psi)\mathbf{1}_{\{\sigma_\psi^2 = \infty\}} + \tilde{\Pi}_n(\psi)\mathbf{1}_{\{\sigma_\psi^2 < \infty\}} \quad (3.30)$$

où $\tilde{\Pi}_n(\psi)$ est comme dans (3). Plus précisément

$$\tilde{\Pi}_n^*(\psi) = \tilde{\Pi}_n(\psi)\mathbf{1}_{\{S(\gamma, \beta)\}} + \tilde{\Pi}_n(\psi)\mathbf{1}_{\{\bar{S}(\gamma, \beta)\}} \quad (3.31)$$

où $S(\gamma, \beta) = \left\{(\gamma, \beta) \in (0, \infty) \times [1, \infty), \gamma \in (1/2, 1) \text{ et } \beta < \frac{1}{\gamma}\right\}$ et $\bar{S}(\gamma, \beta)$ est son complémentaire sur $(0, \infty) \times [1, \infty)$.

Un certain nombre de cas particuliers qui sont couverts par la théorie de la statistique inférentielle pour les primes de risque de distorsion ont été étudiés dans la littérature dans le cas de distributions à queues lourdes en utilisant la théorie des valeurs extrêmes. On peut se référer à Peng [67], Necir and Boukhetala [62], Necir et al. [65], Necir and Meraghni [63], Necir et al. [66], Brahimi et al. [10, 11], Necir et Zitikis (2012), Peng et al. [69], Rassoul [72] et Deme et al. [27].

Pour faire une bonne inférence, comme la recherche de \mathcal{CI} pour $\Pi_\psi(\cdot)$, par exemple, nous devons assumer une certaine caractérisation de la distribution de ces estimateurs. Cependant, l'existence d'une distribution asymptotique non dégénérée exige que la fdr de pertes F satisfait la condition de second ordre discutée

Théorème 3.5.2. *Supposons que F satisfait (2.10) avec $\gamma \in (\frac{1}{2}, 1)$ et son quantile correspondant La fonction $Q(\cdot)$ est continuellement différentiable sur $[0, 1]$. Pour toute fonction de distorsion différentiable ψ satisfaisant la condition (3.27) avec $\beta \in [1, \frac{1}{\gamma})$, et pour toute suite intermédiaire $k = k_n$ satisfaisant et $\sqrt{k}A(n/k) \rightarrow \lambda \in \mathbb{R}$, quand $n \rightarrow \infty$, on a*

$$\sqrt{k}\psi(k/n)Q(1 - k/n)(\tilde{\Pi}_\psi - \tilde{\Pi}_\psi) \rightarrow \mathcal{N}(\lambda\mu_\psi(\gamma, \xi, \beta), \sigma_\psi^2(\gamma, \beta))$$

où

$$\mu_\psi(\gamma, \xi, \beta) = \frac{\beta\xi(\gamma\beta + \beta - 1)}{(1 - \xi)(\gamma\beta + \xi\beta - 1)(1 - \gamma\beta)^2}$$

et

$$\sigma_\psi^2(\gamma, \beta) = \frac{\beta\gamma^2(\gamma\beta + \beta - 1)}{(2\gamma\beta + \beta - 2)(1 - \beta\gamma)^4}.$$

Quelques cas particuliers

Nous pouvons définir les expressions suivantes de σ^2 des quelques fonctions de distorsion ψ .

Dans le cas particulier où la fonction de distorsion ψ est une fonction de puissance, c'est-à-dire $\psi(t) = t^{1/\xi}$ (qui correspond au PHT, dans ce cas, $\beta = \xi$) son estimateur correspondant $\widetilde{\Pi}_\psi$ est celui proposé par Necir and Meraghni [63].

Théorème 3.5.3. *Soit F est une fdr satisfaisant (2.10) avec $\gamma > 1/2$ et de fonction des quantiles $Q(\cdot)$ continûment dérivable sur $[0, 1)$. Soit $k = k_n$ une suite intermédiaire et $\sqrt{k}A(n/k) \rightarrow 0$ quand $n \rightarrow \infty$. Pour $1 \leq \xi < 1/\gamma$, on a*

$$\frac{\sqrt{n}(\widetilde{\Pi}_{PHT,n}(X) - \Pi_{PHT}(X))}{(k/n)^{\xi-1/2}Q(1-k/n)} \xrightarrow{\mathcal{D}} \mathcal{N}(\mu_\psi(\gamma, \xi, \tau), \sigma_\psi^2(\gamma, \xi)), n \rightarrow \infty$$

où

$$\mu_\psi(\gamma, \xi, \tau) = \frac{\xi\tau(\gamma\xi + \xi - 1)}{(1-\tau)(\gamma\xi + \tau\xi - 1)(1-\gamma\xi)^2}$$

et

$$\sigma_\psi^2(\gamma, \xi) = \frac{\beta\gamma^2(\gamma\xi + \xi - 1)}{(2\gamma\xi + \xi - 2)(1-\xi\gamma)^4}.$$

Théorème 3.5.4. *(Normalité asymptotique de $\widetilde{\Pi}_{PHT, \hat{R}_{opt}}$)*

Si on suppose que (2.10) est vérifiée avec $t^{-1/\xi}U(t) \rightarrow 0$, pour une valeur fixée $\xi \geq 1$ et $0 < \gamma < 1/\xi$ quand $t \rightarrow \infty$. Alors pour une suite intermédiaire $k = k_n$ et $\sqrt{k}A(n/k) \rightarrow \lambda \in \mathbb{R}$. On a, quand $n \rightarrow \infty$

$$(\widetilde{\Pi}_{PHT, \hat{R}_{opt}} - \Pi_{PHT, R_{opt}}) \xrightarrow{\mathcal{D}} \mathcal{N}(\lambda\mu, \sigma_R^2(\xi, \gamma)) \quad (3.32)$$

où

$$\mu_R = \frac{\rho\gamma(\xi\gamma + \xi - 1)}{(1-\rho)(\xi\gamma + \xi\rho - 1)(1-\xi\gamma)^2}$$

et

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

Dans le cas particulier où la fonction de distorsion ψ est la fonction $t \mapsto \psi(t) = \min(\frac{t}{\alpha}, 1)$, $0 < \alpha \leq 1$ (qui correspond au CTE, dans ce cas, $\beta = 1$) Necir and Meraghni [64] ont introduit un estimateur de la CTE pour les pertes à queue lourde qui est un autre cas particulier de $\widetilde{\Pi}_\psi$.

Théorème 3.5.5. (*Normalité asymptotique de $\widetilde{\Pi}_{CTE,n}$*)

Soit F est une fdr satisfait la condition (2.10) avec $\gamma \in (1/2, 1)$. Alors pour une suite intermédiaire $k = k_n$ et $\sqrt{k}A(n/k) \rightarrow \lambda$ quand $n \rightarrow \infty$. On a, pour $\alpha \in (0, 1)$

$$\frac{\sqrt{n}(1 - \alpha)(\widetilde{\Pi}_{CTE,n}(X) - \Pi_{CTE}(X))}{(k/n)^{1/2}X_{n-k:n}} \xrightarrow{\mathcal{D}} \mathcal{N}(\lambda\mu_{CTE}, \sigma_{CTE}^2), \quad \text{quand } n \rightarrow \infty.$$

Où

$$\sigma_{CTE}^2 = \frac{\gamma^2}{(1 - \gamma)^2(2\gamma - 1)}$$

Toutes ces Théorèmes se trouvent dans Necir and Meraghni [63], pour l'estimateur de $\Pi_\psi(\cdot)$, et $\Pi_{PHT}(\cdot)$. Pour l'estimateur de $\Pi_{\xi, R_{opt}}(\cdot)$, nous nous référons à Necir et al. [65] et pour l'estimateur de CTE notre référence est Necir et al. [66].

La consistance des estimateurs suit immédiatement de leur normalité asymptotique dans des conditions correspondantes.

Ce chapitre a fait l'objet d'un article publié dans Journal Afrika Statistika (Brahimi and Kenioua, 2016).

Abstract

We use the so-called t-Hill tail index estimator proposed by Fabián [34], rather than Hill's one, to derive a robust estimator for the distortion risk premium of loss. Under the second-order condition of regular variation, we establish its asymptotic normality. By simulation study, we show that this new estimator is more robust than of Necir and Meraghni [63] both for small and large samples.

Keywords: Distortion risk premiums; Extreme values; Tail; Robustness.

AMS 2010 Subject Classification: 62G10; 62G32.

4.1 Introduction

IN many important applications in Finance, Actuarial Science, Hydrology, Insurance, one of most crucial topics is the determination of the amounts of losses of a heavy-tailed risks. In literature there are many possible definitions of risk according to the investment objectives, so in some sense risk itself is a subjective concept as well as the risk measure for an investor. Thus, the existence of a unique risk that solves the investor's problems is not confirmed. However, one must distinguish between risks. The concept of coherence due to the paper of Artzner et al. [2] which categorize risks by good or bad risks. For this reason and to improve the performance of

investor's strategies we may identify those risk measures and the appearance of the risk himself heavy tails, asymmetries,... Most of this risk measures, used are special cases of Wang's distortion premium Wang [87], defined as follows

$$\Pi[\psi; F] = \int_0^\infty \psi(1 - F(x)) dx, \quad (4.1)$$

where $\psi : [0, 1] \rightarrow [0, 1]$ is a non-decreasing function called distortion function, such that $\psi(0) = 0$ and $\psi(1) = 1$. The distortion functions ψ are concave, which makes the corresponding distortion premiums $\Pi[\psi; F]$ coherent (Artzner et al. [2]) as proved by Wirch and Hardy [91]. In this paper, we suppose that the distortion functions $t \mapsto \psi(t^{-1})$ is regularly varying at infinity with index of regular variation $\rho \geq 1$, such that

$$\psi(t^{-1}) = t^{-1/\rho} \mathcal{L}_\psi(t), \quad (4.2)$$

where $t \mapsto \mathcal{L}_\psi(t)$ is slowly varying as infinity, that is $\mathcal{L}_\psi(tx) / \mathcal{L}_\psi(t) \rightarrow 1$ as $t \rightarrow \infty$, for any $x > 0$. Note that a non negative random variable (rv) X with finite mean and a cumulative distribution function (cdf) F is called heavy-tailed, if $1 - F$ is regularly varying with index $-1/\gamma < 0$ (notation: $F \in \mathcal{RV}_{(-1/\gamma)}$), that is

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma}, \text{ for } x > 0. \quad (4.3)$$

In particular, the proportional-hazards premium (see, Rolski et al. [76, page 82]).

$$\Pi[\rho; F] = \int_0^\infty (1 - F(x))^{1/\rho} dx, \quad (4.4)$$

with the concave distortion function $\psi(t) = t^{1/\rho}$ for every $\rho \geq 1$. Since we are concerned with heavy-tailed losses with infinite second moment, then by following Brahimy et al. [11], we assume that $\gamma \in (1/2, 1)$ and $\rho\gamma \in (0, 1)$, thus we will work with

$$1/2 < \gamma < 1/\rho. \quad (4.5)$$

Suppose that we have an independent and identically distributed (iid) sample X_1, \dots, X_n of rv X of size n with a cdf F satisfying condition (4.3) and let denote by $X_{1:n} \leq \dots \leq X_{n:n}$ the corresponding order statistics. Also, let $1 < k = k_n$ be the number of extreme observations used in the computation of the tail index. We assume that k satisfies the conditions

$$1 < k < n, \quad k \rightarrow \infty \text{ and } k/n \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (4.6)$$

Necir and Meraghni [63] proposed an alternative estimator of (4.4) and establish its asymptotic normality by using the Weissman's estimator of the high quantile $q_t = F^{\leftarrow}(1 - t)$ defined by

$$\hat{q}_t = (k/n)^{\hat{\gamma}^H} X_{n-k:n} t^{-\hat{\gamma}^H}, \quad t \downarrow 0,$$

where F^{\leftarrow} denotes the generalized inverse of F and

$$\hat{\gamma}^H = \hat{\gamma}^H(k) := \frac{1}{k} \sum_{i=1}^k \log X_{n-i+1:n} - \log X_{n-k:n}, \quad (4.7)$$

is the well-known Hill estimator (Hill [47]) of the tail index γ . For a fixed aversion parameter ρ , their estimator is given by

$$\widehat{\Pi}_n(\hat{\gamma}^H, k) = (k/n)^{1/\rho} \frac{X_{n-k:n}}{1 - \hat{\gamma}^H \rho} + \sum_{i=k+1}^n \left((i/n)^{1/\rho} - ((i-1)/n)^{1/\rho} \right) X_{n-i+1:n} \quad (4.8)$$

The Hill estimator is a pseudo-maximum likelihood estimator based on the exponential approximation of normalized log-spacings

$$Y_j = j (\log X_{j:n} - \log X_{j+1:n}) \quad \text{for } j = 1, \dots, k.$$

In practice, the Hill estimator depends on the choice of the sample fraction k and is inherently not very robust to large values Y_j , which makes the estimator proposed by Necir and Meraghni [63] sensitive. This constitutes a serious problem in terms of bias and root mean squared error (RMSE). To improve the quality of $\widehat{\Pi}_n(\hat{\gamma}^H, k)$, instead of Hill's one, we propose to estimate the tail index γ by the so-called t-Hill estimator, proposed by Fabián [34], given by its harmonic mean

$$\hat{\gamma} = \hat{\gamma}(k) := \left(\frac{1}{k} \sum_{j=1}^k \frac{X_{n-k:n}}{X_{n-j+1:n}} \right)^{-1} - 1 \quad (4.9)$$

known as score moment estimation (t-score or t-estimation method). The latter is more robust than the classical Hill estimator $\hat{\gamma}^H$ defined in (4.7) (see Stehlík et al. [80]) and the asymptotic normality is given in Theorem 2 of Beran et al. [6]). For other robust estimators for γ we referred to Peng and Welsh [70], Juárez and Schucany [55], Vandewalle et al. [84] and Kim and Lee [57]. The rest of the Chapter is organized as follows, in Section 4.2 we present a construction of a robust estimator of $\Pi[\psi; F]$ in the case of heavy-tailed losses. In Section 4.3 we establish its asymptotic normality. In Section 4.4 we carry out a simulation study to illustrate empirical performance and

robustness of the estimator. Concluding notes are given in Section 4.5. Proofs are gathered in Section 4.6.

Throughout the chapter, we use the standard notation \xrightarrow{P} for the convergence in probability and $\mathcal{N}(\mu, \sigma)$ to denote a normal rv with mean μ and variance σ .

4.2 Defining the estimator

By using the generalized inverse F^{\leftarrow} and for a fixed distortion function ψ , we may rewrite (4.1) into

$$\Pi_{\psi}[X] := - \int_0^1 \psi(s) dF^{\leftarrow}(1-s). \quad (4.10)$$

The empirical estimator of the risk premium $\Pi_{\psi}[X]$ is obtained by substituting F^{\leftarrow} on the right-hand side of equation (4.10) by its empirical counterpart $F_n^{\leftarrow}(s) := \inf\{x \in \mathbb{R} : F_n(x) \geq s\}$, $0 < s \leq 1$, associated to the empirical cdf defined on the real line, defined by

$$F_n(x) = n^{-1} \# \{X_i \leq x, 1 \leq i \leq n\}$$

where $\#A$ denote the cardinality of a set A . After straightforward computations, we obtain the formula

$$\Pi_n[X] := \int_0^1 F_n^{\leftarrow}(1-s) d\psi(s)$$

which may be rewritten, in terms of $X_{1:n}, \dots, X_{n:n}$, as an L -statistic

$$\Pi_n[X] = \sum_{i=1}^n c_{i,n}(\psi) X_{n-i+1:n}, \quad (4.11)$$

where

$$c_{i,n}(\psi) \equiv \psi(i/n) - \psi((i-1)/n). \quad (4.12)$$

The form (4.11) is a linear combinations of the order statistics (see, Shorack and Wellner [78, page 260]). The limit behavior was discussed by many authors: Chernoff et al. [15], Stigler [82], Mason [59], Jones and Zitikis [52] (see its Theorem 3.2 in the case that X is not heavy-tailed) and in Brahimi et al. [11] (in heavy-tailed case).

4.2.1 Heavy-tailed losses case

Let X be a non-negative rv with cdf $F \in \mathcal{RV}_{(-1/\gamma)}$. The condition (4.3) is equivalent to

$$\lim_{t \rightarrow 0} \frac{F^{\leftarrow}(1 - tx)}{F^{\leftarrow}(1 - t)} = x^{-\gamma}, \text{ for every } x > 0. \quad (4.13)$$

We say that the function $s \rightarrow F^{\leftarrow}(1 - s)$ satisfying condition (4.13) is regularly varying at zero with the index $(-\gamma) < 0$. The parameter γ is called the tail index or extreme value index (EVI). A various tail index estimators have been suggested in the literature, based for instance of the conventional maximum likelihood method, moment estimation, ... (see, e.g. Hill [47], Dekkers et al. [25]. Csörgő et al. [18] and Drees [31]). For the robustness and bias reduction (see, i.e. Peng and Qi [68] and Stehlík et al. [80]). The regular-variation condition itself is not sufficient for establishing asymptotic distributions. To this end, we suppose that cdf F satisfy the well-known by the second-order condition of regular variation with second-order parameter $\tau \leq 0$, that is: there exists a function $t \rightarrow a(t)$ with constant sign at infinity and converges to 0 as $t \rightarrow \infty$ such that

$$\lim_{t \rightarrow \infty} \frac{\overline{F}(tx) / \overline{F}(t) - x^{-1/\gamma}}{a(t)} = x^{-1/\gamma} \frac{x^{\tau/\gamma} - 1}{\gamma\tau}, \quad (4.14)$$

for every $x > 0$. When $\tau = 0$, then the ratio $\frac{x^{\tau/\gamma} - 1}{\gamma\tau}$ should be interpreted as $\log x$. In terms of the quantile function F^{\leftarrow} , condition (4.14) is equivalent to the following one

$$\lim_{t \rightarrow 0} \frac{\frac{F^{\leftarrow}(1 - tx)}{F^{\leftarrow}(1 - t)} - x^{-\gamma}}{A(t)} = x^{-\gamma} \frac{x^{\tau} - 1}{\tau}, \quad (4.15)$$

for every $x > 0$, where $A(t) := \gamma^2 a(F^{\leftarrow}(1 - t))$, (see de Haan and Stadtmüller, 1996 or Theorem 3.2.9 in de Haan and Ferreira, 2006). The Weissman estimator (Weissman, 1978) of high quantiles F^{\leftarrow} is given by

$$F_n^{\leftarrow(W)}(1 - s) := (k/n)^{\widehat{\gamma}} X_{n-k:n} s^{-\widehat{\gamma}}, \quad s \downarrow 0. \quad (4.16)$$

The formula (4.10) can be split into

$$\Pi_{\psi}[X] = - \int_0^{k/n} \psi(s) dF^{\leftarrow}(1 - s) - \int_{k/n}^1 \psi(s) dF^{\leftarrow}(1 - s). \quad (4.17)$$

By using an integration by part to the second integral yields

$$\begin{aligned}\Pi_\psi &= \psi(k/n) F^{\leftarrow}(1 - k/n) - \int_0^{k/n} \psi(s) dF^{\leftarrow}(1 - s) + \int_{k/n}^1 F^{\leftarrow}(1 - s) d\psi(s) \\ &= \Pi_\psi^{(1)} + \Pi_\psi^{(2)} + \Pi_\psi^{(3)}.\end{aligned}$$

A simple estimator of $\Pi_\psi^{(1)}$ is

$$\Pi_{\psi,n}^{(1)} := \psi(k/n) X_{n-k,n}. \quad (4.18)$$

To estimate $\Pi_\psi^{(2)}$, we note that $\hat{\gamma}$ is a consistent estimator for γ (Stehlík et al. [80]) and since $\rho < 1/\gamma$, by substituting $F_n^{\leftarrow(W)}(1 - s)$ given in (4.16) instead of $F^{\leftarrow}(1 - s)$ and integrating yield the following estimator

$$\Pi_{\psi,n}^{(2)} := \hat{\gamma}(k/n)^{\hat{\gamma}} X_{n-k:n} \int_0^{k/n} s^{-\hat{\gamma}-1} \psi(s) ds. \quad (4.19)$$

Finally, by plugging F_n^{\leftarrow} instead of F^{\leftarrow} on second integral of Equation (4.17) we obtain the estimator

$$\Pi_{\psi,n}^{(3)} := \sum_{i=k+1}^n c_{i,n}(\psi) X_{n-i+1:n}, \quad (4.20)$$

of $\Pi_\psi^{(3)}$, where $F_n^{\leftarrow}(s) = \inf \{x \in \mathbb{R} : F_n(x) \geq s\}$, $0 < s \leq 1$, denote the sample quantile function associated to the empirical cdf defined on the real line by $F_n(x) := n^{-1} \sum_{i=1}^n \mathbf{1}_{(X_i \leq x)}$, with $\mathbf{1}_{(\cdot)}$ being the indicator function and the coefficients $c_{i,n}(\psi)$ are given in (4.12). The final form of our estimator

$$\begin{aligned}\tilde{\Pi}_{\psi,n} &= X_{n-k,n} \left(\psi(k/n) + \hat{\gamma}(k/n)^{\hat{\gamma}} \int_0^{k/n} s^{-\hat{\gamma}-1} \psi(s) ds \right) \\ &\quad + \sum_{i=k+1}^n c_{i,n}(\psi) X_{n-i+1:n}\end{aligned}$$

To establish the asymptotic normality of our estimator and compared with the estimator proposed by Necir and Meraghni (2009) given in Equation (4.8), we use the same function $\psi(t) = t^{1/\rho}$, for $\rho \geq 1$ used in Necir and Meraghni [63]. In this case our estimator have the following form

$$\tilde{\Pi}_n(\hat{\gamma}, k) := (k/n)^{1/\rho} \frac{X_{n-k,n}}{1 - \hat{\gamma}\rho} + \sum_{i=k+1}^n \left((i/n)^{1/\rho} - ((i-1)/n)^{1/\rho} \right) X_{n-i+1:n} \quad (4.21)$$

4.3 Asymptotic distribution

We will begin to expose our results as asymptotic representations theorems in the lines of Beran et al. [6]. For that purpose, we need to describe the

probability theory on which they hold. Indeed, we use the so-called Hungarian construction of Csörgő et al. [17]. For this we define $\{U_n(s), 0 \leq s \leq 1\}$, the uniform empirical distribution function and we consider the order statistics $U_{1:n} \leq \dots \leq U_{n:n}$ pertaining to the independent standard uniform rv's U_1, U_2, \dots , we introduce the uniform empirical quantile function $\{V_n(s), 0 \leq s \leq 1\}$ based on the $n \geq 1$ first observations of U_1, U_2, \dots , on $(0, 1)$ such that

$$V_n(s) = U_{i,n} \quad \text{for } (i-1)/n < s \leq i/n, \quad i = 1, \dots, n, \quad \text{and } V_n(0) = U_{1,n}.$$

Let $\beta_n(s) = \sqrt{n}(s - V_n(s)), 0 \leq s \leq 1$, be the corresponding quantile empirical process, for $n \geq 1$.

We use the well-known by Gaussian approximation given in Csörgő et al. [17] Corollary 2.1. It says that: on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, there exists a sequence of Brownian bridges $\{\mathbb{B}_n(s); 0 \leq s \leq 1\}$ has the representation

$$\{\mathbb{B}_n(s); 0 \leq s \leq 1\} \stackrel{D}{=} \{W_n(s) - sW_n(1); 0 \leq s \leq 1\},$$

where W_n is a standard Wiener process such that for every $0 \leq \zeta < 1/2$,

$$\sup_{1/n \leq s \leq 1-1/n} \frac{n^\zeta |\beta_n(s) - \mathbb{B}_n(s)|}{(s(1-s))^{1/2-\zeta}} = O_{\mathbb{P}}(n^{-\zeta}). \quad (4.22)$$

Our main result is the following

Théorème 4.3.1. *Let F be a df satisfying (4.15) with $\gamma > 1/2$ and suppose that $F^{\leftarrow}(\cdot)$ is continuously differentiable on $[0, 1)$. Let $k = k_n$ satisfying (4.6) such that $\sqrt{k}A(n/k) \rightarrow 0$ as $n \rightarrow \infty$. For any $1 \leq \rho < 1/\gamma$, we have*

$$\frac{\sqrt{n}(\widetilde{\Pi}_{\psi,n} - \Pi_{\psi})}{(k/n)^{1/\rho-1/2}F^{\leftarrow}(1-k/n)} \xrightarrow{d} \mathcal{N}(0, \sigma^2(\gamma, \rho)), \quad \text{as } n \rightarrow \infty,$$

where

$$\begin{aligned} \sigma^2(\gamma, \rho) &= \gamma^2 + \frac{\gamma^2 \rho (\rho - 2\rho\gamma^2 + 2\gamma)}{(\gamma\rho - 1)^2} + \frac{2\gamma^2}{(\rho + \gamma\rho - 1)(\rho + 2\gamma\rho - 2)} \\ &\quad + \frac{2\gamma}{2\gamma - 1} - \frac{2\gamma\rho(\rho\gamma^2 - \rho\gamma + 1)}{(\gamma\rho - 1)(\rho + \gamma\rho - 1)}. \end{aligned}$$

4.4 Simulation study

4.4.1 Performance and comparative study of $\widetilde{\Pi}_n$ and $\widehat{\Pi}_n$

In this simulation study we examine the performance of our estimator $\widetilde{\Pi}_n(\widehat{\gamma}, k)$ given in (4.21) and compare it with that of $\widehat{\Pi}_n(\widehat{\gamma}^H, k)$ given in (4.8). Thus we follow the steps below.

Step 1: We generate 1000 pseudorandom samples of size $n = 100, 200, 500$ and 1000 from Pareto cdf with $\gamma = 0.6$.

Step 2: We estimate the tail index parameter by Hill and t-Hill estimators $\hat{\gamma}^H(k_1^*)$ and $\hat{\gamma}(k_2^*)$, respectively given in (4.7) and (4.9). We adopt the **Reiss and Thomas** algorithm (see Reiss and Thomas [73, page 137]), for choosing the optimal numbers of upper extremes k_1 and k_2 . By this methodology, we define the optimal sample fraction of upper order statistics k_j^* by

$$k_j^* := \arg \min_k \frac{1}{k} \sum_{i=1}^k i^\theta |\hat{\gamma}_j(i) - \text{median} \{\hat{\gamma}_j(1), \dots, \hat{\gamma}_j(k)\}|, j = 1, 2$$

where $\hat{\gamma}_1 = \hat{\gamma}^H$ and $\hat{\gamma}_2 = \hat{\gamma}$. On the light of our simulation study, we obtained reasonable results by choosing $\theta = 0.3$. For an illustration of the behavior of $\tilde{\Pi}_n(\hat{\gamma}, k)$ as a function of n , see Figure 4.1.

Step 3: We fix the distortion parameter with respect to Condition (4.5) by $\rho = 1.12$, then we compute the bias and RMSE of the four estimators $\hat{\gamma}^H(k_1^*)$, $\hat{\gamma}(k_2^*)$, $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ and $\hat{\Pi}_n(\hat{\gamma}^H, k_2^*)$. The results are summarized in Table 4.1. We see that when dealing with large samples our estimator performs better.

n	k_1^*	$\hat{\gamma}(k_1^*)$		$\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$		k_2^*	$\hat{\gamma}^H(k_2^*)$		$\hat{\Pi}_n(\hat{\gamma}^H, k_2^*)$	
		bias	RMSE	bias	RMSE		bias	RMSE	bias	RMSE
100	10	-0.0733	0.2511	0.3618	0.5199	17	-0.1641	0.2865	0.4096	0.7332
200	23	-0.0571	0.1821	0.3562	0.5147	34	-0.0993	0.2350	0.3918	0.7185
500	62	-0.0299	0.1142	0.3404	0.4820	86	-0.0301	0.0739	0.3639	0.6936
1000	129	-0.0147	0.0798	0.1966	0.2687	169	-0.0181	0.0545	0.2827	0.5279

Tableau 4.1: $\hat{\gamma}(k_1^*)$, $\hat{\gamma}^H(k_2^*)$, $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ and $\hat{\Pi}_n(\hat{\gamma}^H, k_2^*)$ estimators based on 1000 samples of Pareto-distributed claim amounts with tail index 0.6 and distortion parameter $\rho = 1.12$. The exact value of the premium is 2.0487.

4.4.2 Comparative robustness study

In this subsection we study the sensitivity to outliers of $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ and compare it with that of $\hat{\Pi}_n(\hat{\gamma}^H, k_2^*)$. We consider an ϵ -contaminated model known by mixture of Pareto distributions

$$F_{\gamma_1, \gamma_2, \epsilon}(x) = 1 - (1 - \epsilon)x^{-1/\gamma_1} + \epsilon x^{-1/\gamma_2}, \quad (4.23)$$

where $\gamma_1, \gamma_2 > 0$ and $0 < \epsilon < 0.5$ is the fraction of contamination. Note that for $\epsilon = 0$, $\hat{\gamma}^H$ and $\hat{\gamma}$ are asymptotically unbiased. Therefore, for $\epsilon > 0$, the effect of contamination becomes immediately apparent. If $\gamma_1 < \gamma_2$ and $\epsilon > 0$, (4.23) corresponds to a Pareto distribution contaminated by a longer tailed

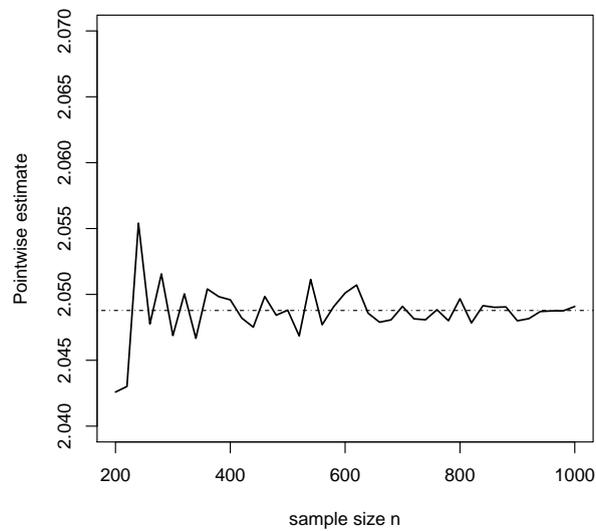


Fig 4.1: Plot of the premium estimator $\tilde{\Pi}_n(\hat{\gamma}, k_1)$ as a function of the sample size n based on 200 Pareto-distributed claims with tail index 0.6. The distortion parameter is 1.12. The horizontal line represents the true value of the premium 2.0487.

distribution. For the implementation of mixtures models to the outliers study one refers, for instance, to Barnett and Lewis [4, page 43]. In this context, we proceed our study as follows.

First, we consider $\gamma_1 = 0.6$, $\gamma_2 = 2$ to have the contaminated model and let $\rho = 1.12$. Then we consider four contamination scenarios according to $\epsilon = 5\%$, 10% , 15% , 25% .

For each value ϵ , we generate 1000 samples of size $n = 100$, 200 and 1000 from the model (4.23). Finally, we compare the $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ and $\widehat{\Pi}_n(\hat{\gamma}^H, k_2^*)$ estimators with this true value, by computing for each estimator, the appropriate bias and RMSE and summarize the results in Table (4.2).

As expected, the estimator $\widehat{\Pi}_n(\hat{\gamma}^H, k_2^*)$ as well as $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ turn out to be more sensitive to this type of contaminations. For example, in 0% contamination for $n = 200$, the couple (bias, RMSE) for $\widehat{\Pi}_n(\hat{\gamma}^H, k_2^*)$ take the values (0.3918, 0.7185), while for 15% contamination the bias and the RMSE are given by the couple $(-1.4355, 2.3107)$. We may conclude that the bias and RMSE of $\widehat{\Pi}_n(\hat{\gamma}^H, k_2^*)$ estimator are more sensitive (or note robust) to outliers. However, for 0% contamination the (bias, RMSE) of $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ is (0.3562, 0.5147), while for 15% contamination is (0.4508, 0.6870). Both the bias and the RMSE of $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ estimation are note sensitive to outliers. Then we may conclude that is the better estimator.

n	% contamination	$\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$		$\hat{\Pi}_n(\hat{\gamma}^H, k_2^*)$	
		bias	RMSE	bias	RMSE
100	5	0.4043	0.6664	-0.4286	1.4727
	10	0.4389	0.6862	-0.7291	1.9123
	15	0.4598	0.7464	-1.2786	2.1247
	25	1.0578	1.1305	-1.2103	2.1828
200	5	0.3831	0.5532	-0.4713	1.5118
	10	0.3964	0.5675	-1.2496	2.1907
	15	0.4508	0.6870	-1.4355	2.3107
	25	0.9470	1.0197	-1.7366	2.3274
1000	5	0.2124	0.3211	-0.3794	2.1222
	10	0.2329	0.3349	-1.0662	2.3978
	15	0.2931	0.3749	-1.2501	2.0355
	25	0.8124	0.9291	-1.5238	2.3596

Tableau 4.2: $\tilde{\Pi}_n(\hat{\gamma}, k_1^*)$ and $\hat{\Pi}_n(\hat{\gamma}^H, k_2^*)$ are based on 1000 samples of mixture of Pareto distributions with tail index 0.6, $\epsilon = 5\%, 10\%, 15\%, 25\%$ and distortion parameter $\rho = 1.12$. The exact value of the premium is 2.0487.

4.5 Concluding notes

We showed that the new estimator of premium based on t-Hill estimator is more robust and performs better than the one based on Hill estimator proposed by Necir and Meraghni [63]. Our estimator $\tilde{\Pi}_n(\hat{\gamma}, k)$ is based on Weissman's estimation of high quantiles, so we would lead to improve our result to use one of several bias-reduced estimators have been proposed (see for example Matthys and Beirlant [61]).

4.6 Proofs

To establish the asymptotic normality of $\tilde{\Pi}_{\psi,n}$ we need the asymptotic approximation of $\hat{\gamma}$ with the same sequence of Brownian bridges as $\tilde{\Pi}_{\psi,n}$, for this reason we give the following results.

Proposition 4.6.1. *Assume that the second order condition (4.15) holds with $\gamma > 1/2$ and let $k = k_n$ be an integer sequence satisfying (4.6) and $\sqrt{k}A(n/k) \rightarrow 0$. Then, there exists a sequence of Brownian bridges $\{\mathbb{B}_n(s), 0 \leq s \leq 1\}$ such that*

$$\sqrt{k}(\hat{\gamma} - \gamma) = \gamma(\gamma + 1)^2 \int_0^1 s^{\gamma-1} \mathbb{B}_n(s) ds + o_p(1),$$

leading to

$$\sqrt{k}(\hat{\gamma} - \gamma) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\gamma^2(1 + \gamma)^2}{(1 + 2\gamma)}\right), \text{ as } n \rightarrow \infty,$$

Proof. Our proofs are conducted in the probability space described in Section 4.3. Then we are entitled to write

$$S_k := \frac{1}{k} \sum_{j=1}^k \frac{X_{n-k:n}}{X_{n-j+1:n}}$$

then $\hat{\gamma} = S_k^{-1} - 1$. The asymptotic normality of $\hat{\gamma}$ is established in Beran et al. [6] by given their Wiener process representation. Here we suppose that $\sqrt{k}A(n/k) \rightarrow \lambda = 0$. So

$$\sqrt{k} \left(S_k - \frac{1}{\gamma + 1} \right) = \frac{\gamma}{\gamma + 1} W_n(1) - \gamma \int_0^1 t^{\gamma-1} W_n(t) dt + o_p(1).$$

Note that

$$\mathbb{B}_n(s) \stackrel{\mathcal{D}}{=} W_n(s) - sW_n(1),$$

it follows that

$$\sqrt{k} \left(S_k - \frac{1}{\gamma + 1} \right) = -\gamma \int_0^1 s^{\gamma-1} \mathbb{B}_n(s) ds + o_p(1).$$

Using the map $g(x) = 1/x - 1$, since $g(1/(\gamma + 1)) = \gamma$ and applying the delta method yields

$$\sqrt{k} (\hat{\gamma} - \gamma) = (\gamma + 1)^2 \gamma \int_0^1 s^{\gamma-1} \mathbb{B}_n(s) ds + o_p(1).$$

It is clear that $\sqrt{k} (\hat{\gamma} - \gamma)$ is a Gaussian rv with mean 0 and variance $\frac{\gamma^2(1+\gamma)^2}{(1+2\gamma)}$. This completes the proof of Proposition 4.6.1. \square

4.6.1 Proof of Theorem 4.3.1

Making use of Proposition 4.6.1 and from Necir et al. [65] we show that under the assumptions of Theorem 4.3.1, there exists a sequence of Brownian bridges $\{\mathbb{B}_n(s), 0 \leq s \leq 1\}$ such that, for all large n

$$\frac{\left(\Pi_{\psi,n}^{(1)} - \Pi_{\psi}^{(1)} \right)}{(k/n)^{1/\rho} F^{\leftarrow}(1 - k/n)} = -\gamma (n/k)^{1/2} \mathbb{B}_n(1 - k/n) + o_p(1).$$

Let \mathbb{U} be the left-continuous inverse of $1/(1 - F)$. Note that $\mathbb{U}(t)$ is defined for $t > 1$. Let Y_1, Y_2, \dots be independent and identically distributed rv's with cdf $1 - 1/y, y > 1$, and let $Y_{1,n} \leq Y_{2,n} \leq \dots \leq Y_{n,n}$ be the associated order statistics. Then, for $\rho \geq 1$, we may rewrite the statistic $\Pi_{\psi,n}^{(2)}$ as

$$\Pi_{\psi,n}^{(2)} = (k/n)^{1/\rho} \frac{\hat{\gamma} \rho}{1 - \hat{\gamma} \rho} X_{n-k:n}.$$

Then

$$\Pi_{\psi,n}^{(2)} = (k/n)^{1/\rho} \frac{\hat{\gamma}\rho}{1 - \hat{\gamma}\rho} \mathbb{U}(Y_{n-k:n}),$$

where $\hat{\gamma}$ is the t-Hill estimator of γ . So we have

$$\frac{\sqrt{k} \left(\Pi_{\psi,n}^{(2)} - \Pi_{\psi}^{(2)} \right)}{(k/n)^{1/\rho} \mathbb{U}(n/k)} := \sum_{j=1}^4 \Delta_{jn},$$

where

$$\begin{aligned} \Delta_{1n} &:= \sqrt{k} \frac{\hat{\gamma}\rho}{1 - \hat{\gamma}\rho} \left(\frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - \left(\frac{Y_{n-k:n}}{n/k} \right)^\gamma \right), \\ \Delta_{2n} &:= \sqrt{k} \frac{\hat{\gamma}\rho}{1 - \hat{\gamma}\rho} \left(\left(\frac{Y_{n-k:n}}{n/k} \right)^\gamma - 1 \right), \\ \Delta_{3n} &:= \sqrt{k} \left(\frac{\hat{\gamma}\rho}{1 - \hat{\gamma}\rho} - \frac{\gamma\rho}{1 - \gamma\rho} \right) \end{aligned}$$

and

$$\Delta_{4n} := \sqrt{k} (k/n)^{-1/\rho} \left(\frac{(k/n)^{1/\rho} \rho \mathbb{U}(n/k) - \Pi_{\psi}^{(2)}}{\mathbb{U}(n/k)} \right).$$

As showed in Necir et al. [65], we have : $\Delta_{1n} \rightarrow 0$ and $\Delta_{4n} \rightarrow 0$ as $n \rightarrow \infty$.

Next, we show that $\Delta_{2n} + \Delta_{3n}$ is asymptotically normal. Assume, without loss of generality, that the rv's $(Y_n)_{n \geq 1}$ are defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ which carries the sequence $(U_n)_{n \geq 1}$ in such a way that $Y_n = (1 - U_n)^{-1}$ for $n = 1, 2, \dots$ and $Y_{i,n} = (1 - U_{i,n})^{-1}$, $i = 1, \dots, n$. Then, this allows us to write $Y_{n-i+1,n} = (1 - V_n(1 - s))^{-1}$, for $\frac{i-1}{n} < s \leq \frac{i}{n}$, $i = 1, \dots, n$. From Necir et al. [65] we have For Δ_{2n}

$$\Delta_{2n} = - (n/k)^{1/2} \frac{\rho\gamma^2}{1 - \gamma\rho} (1 + o_p(1)) \mathbb{B}_n(1 - k/n).$$

For Δ_{3n} and by using the map $h(\theta) = \rho / (\frac{1}{\theta} - \rho)$ and applying the delta method yields:

$$\Delta_{3n} = \frac{\rho}{(\rho\gamma - 1)^2} \sqrt{k} (\gamma - \hat{\gamma}).$$

From Corollary 4.6.1 we get

$$\Delta_{3n} = \frac{\gamma\rho(\gamma + 1)^2}{(\rho\gamma - 1)^2} \int_0^1 s^{\gamma-1} \mathbb{B}_n(s) ds + o_p(1).$$

Finally we have

$$\frac{\sqrt{k} \left(\Pi_{\psi,n}^{(2)} - \Pi_{\psi}^{(2)} \right)}{(k/n)^{1/\rho} F^{\leftarrow}(1 - k/n)} = \frac{\gamma \rho (\gamma + 1)^2}{(\rho \gamma - 1)^2} \int_0^1 s^{\gamma-1} \mathbb{B}_n(s) ds - (n/k)^{1/2} \frac{\rho \gamma^2}{1 - \gamma \rho} \mathbb{B}_n(1 - k/n) + o_p(1)$$

From Necir et al. [65] we have

$$\frac{\sqrt{k} \left(\Pi_{\psi,n}^{(3)} - \Pi_{\psi}^{(3)} \right)}{(k/n)^{1/\rho} F^{\leftarrow}(1 - k/n)} = \frac{\int_{k/n}^1 s^{1/\rho-1} \mathbb{B}_n(1 - s) F^{\leftarrow'}(1 - s) ds}{\rho F^{\leftarrow}(1 - k/n) (k/n)^{1/\rho-1/2}} + o_p(1).$$

Then

$$\frac{\sqrt{n} \left(\widetilde{\Pi}_{\psi,n} - \Pi_{\psi} \right)}{(k/n)^{1/\rho-1/2} F^{\leftarrow}(1 - k/n)} = \Lambda(\gamma, \rho) + o_p(1)$$

where

$$\Lambda(\gamma, \rho) := W_{n1} + W_{n2} + W_{n3} + o_p(1)$$

and

$$\begin{aligned} W_{n1} &:= - (n/k)^{1/2} \gamma \mathbb{B}_n(1 - k/n), \\ W_{n2} &:= \frac{\gamma \rho (\gamma + 1)^2}{(\rho \gamma - 1)^2} \int_0^1 s^{\gamma-1} \mathbb{B}_n(s) ds - (n/k)^{1/2} \frac{\rho \gamma^2}{1 - \gamma \rho} \mathbb{B}_n(1 - k/n), \\ W_{n3} &:= \frac{\int_{k/n}^1 s^{1/\rho-1} \mathbb{B}_n(1 - s) F^{\leftarrow'}(1 - s) ds}{\rho F^{\leftarrow}(1 - k/n) (k/n)^{1/\rho-1/2}}. \end{aligned}$$

It is clear that $\Lambda(\gamma, \rho)$ is a Gaussian rv with mean 0 and variance

$$\begin{aligned} E(\Lambda(\gamma, \rho))^2 &= E(W_{n1}^2) + E(W_{n2}^2) + E(W_{n3}^2) + 2E(W_{n1}W_{n2}) \\ &\quad + 2E(W_{n1}W_{n3}) + 2E(W_{n2}W_{n3}). \end{aligned}$$

An elementary calculation gives, we get

$$\begin{aligned} E(W_{n1}^2) &= \gamma^2 + o(1), \\ E(W_{n2}^2) &= \frac{\gamma^2 \rho^2}{(1 - \gamma \rho)^2} + \frac{2\gamma}{2\gamma - 1} + o(1), \\ E(W_{n3}^2) &= \frac{2\gamma^2}{(\rho + \gamma \rho - 1)(\rho + 2\gamma \rho - 2)} + o(1), \\ E(W_{n1}W_{n2}) &= \frac{\rho \gamma^3}{1 - \gamma \rho} + o(1), \\ E(W_{n1}W_{n3}) &= \frac{\gamma \rho}{\rho + \gamma \rho - 1} + o(1) \end{aligned}$$

and

$$E(W_{n2}W_{n3}) = -\frac{\gamma^3 \rho^2}{(\gamma\rho - 1)(\rho + \gamma\rho - 1)} + o(1).$$

The proof of Theorem 4.3.1 is completed by combining all the preceding results. \square

- [1] Acerbi, C., Nardio, C., and Sirtori, C. (2001). Expected shortfall as a tool for financial risk management. *arXiv preprint cond-mat/0102304*. 53
- [2] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3):203–228. 1, 46, 48, 49, 50, 52, 53, 62, 73, 74
- [3] Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, 2(5):792–804. 28
- [4] Barnett, V. and Lewis, T. (1995). *Outliers in Statistical Data*. John Wiley & Sons, New York, 3rd edition. 81
- [5] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, volume 558. John Wiley & Sons. 2, 6, 33, 35
- [6] Beran, J., Schell, D., and Stehlík, M. (2014). The harmonic moment tail index estimator: asymptotic distribution and robustness. *Annals of the Institute of Statistical Mathematics*, 66(1):193–220. 42, 75, 78, 83
- [7] Bertsimas, D., Lauprete, G. J., and Samarov, A. (2004). Shortfall as a risk measure: properties, optimization and applications. *Journal of Economic Dynamics and Control*, 28(7):1353–1381. 53
- [8] Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987). *Regular variation*. Encyclopedia of mathematics and its applications 27. Cambridge University Press. 19, 20, 21
- [9] Brahim, B. and Kenioua, Z. (2016). Robust estimator of distortion risk premiums for heavy-tailed losses. *Afrika Statistika*, 11(1):869–882. 4, 73

- [10] Brahim, B., Meddi, F., and Necir, A. (2012). Bias-corrected estimation in distortion risk premiums for heavy-tailed losses. *Afrika Statistika*, 7(1):474–490. 66, 70
- [11] Brahim, B., Meraghni, D., Necir, A., and Zitikis, R. (2011). Estimating the distortion parameter of the proportional-hazard premium for heavy-tailed losses. *Insurance: Mathematics and economics*, 49(3):325–334. 66, 70, 74, 76
- [12] Brazauskas, V., L. Jones, B., L. Puri, M., and Zitikis, R. (2008). Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference*, 138(11):3590–3604. 66, 68
- [13] Brazauskas, V. and Serfling, R. (2000). Robust and efficient estimation of the tail index of a single-parameter pareto distribution. *North American Actuarial Journal*, 4(4):12–27. 35
- [14] Centeno de Lourdes, M. and Andrade e Silva, J. (2005). Applying the proportional hazard premium calculation principle. *Astin Bulletin*, 35(2):409. 66
- [15] Chernoff, H., Gastwirth, J. L., and Johns, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *The Annals of Mathematical Statistics*, 38(1):52–72. 76
- [16] Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Lecture Notes in Control and Information Sciences. Springer. 6
- [17] Csörgő, M., Csörgő, S., Horváth, L., and Mason, D. M. (1986). Weighted empirical and quantile processes. *The Annals of Probability*, 14(1):31–85. 79
- [18] Csörgő, S., Deheuvels, P., and Mason, D. M. (1985). Kernel estimates of the tail index of a distribution. *The Annals of Statistics*, 13(3):1050–1077. 41, 77
- [19] Davis, R. and Resnick, S. (1984). Tail estimates motivated by extreme value theory. *The Annals of Statistics*, 12(4):1467–1487. 41
- [20] de Haan, L. (1976). Sample extremes: an elementary introduction. *Statistica Neerlandica*, 30(4):161–172. 5

- [21] de Haan, L. and Ferreira, A. (2006). *Extreme value theory*. Springer Series in Operations Research and Financial Engineering, New York Inc. Springer. 2, 6, 15, 20, 21, 22, 24, 25, 27, 31, 33, 37, 39, 41, 77
- [22] de Haan, L. and Rootzén, H. (1993). On the estimation of high quantiles. *Journal of Statistical Planning and Inference*, 35(1):1–13. 44
- [23] de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *Journal of the Australian Mathematical Society (Series A)*, 61(03):381–395. 39, 77
- [24] Deheuvels, P., Haeusler, E., and Mason, D. M. (1988). Almost sure convergence of the hill estimator. *Mathematical Proceedings of the Cambridge Philosophical Society*, 104(2):371–. 41
- [25] Dekkers, A., Einmahl, J., and De Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, 17(4):1833–1855. 34, 41, 42, 43, 44, 77
- [26] Dekkers, A. L. M. and De Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation. *The Annals of Statistics*, 17(4):1795–1832. 40, 44
- [27] Deme, E. H., Girard, S., and Guillou, A. (2013). Reduced-bias estimator of the proportional hazard premium for heavy-tailed distributions. *Insurance: Mathematics and Economics*, 52(3):550–559. 66, 70
- [28] Deme, E. H. and Lo, G. S. (2013). Reduced-bias estimators for the distortion risk premiums for heavy-tailed distributions. 70
- [29] Denneberg, D. (1994). *Non-additive measure and integral*. Kluwer Academic Publishers, Dordrecht. 62
- [30] Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2005). *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. 1 edition. 64
- [31] Drees, H. (1995). Refined pickands estimators of the extreme value index. *The Annals of Statistics*, 23(6):2059–2080. 33, 77
- [32] Drees, H. and Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and their Applications*, 75(2):149–172. 31

- [33] Embrechts, P., Mikosch, T., and Klüppelberg, C. (1997). *Modelling extremal events: for insurance and finance*. Springer-Verlag. 2, 6, 12, 23, 24, 26, 28, 29, 33
- [34] Fabián, Z. (2001). Induced cores and their use in robust parametric estimation. *Communication in Statistics- Theory and Methods*, 30(3):537–555. iv, v, 4, 35, 36, 73, 75
- [35] Fabián, Z. and Stehlík, M. (2008). A note on favorable estimation when data is contaminated. *Communications in Dependability and Quality Management*, 11(4):36–43. 36
- [36] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2):180–190. 5, 11
- [37] Föllmer, H. and Schied, A. (2002a). Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447. 49, 51
- [38] Föllmer, H. and Schied, A. (2002b). Robust preferences and convex measures of risk. In *Advances in finance and stochastics*, pages 39–56. Springer. 50
- [39] Fréchet, M. (1927). Sur la loi de probabilité de l' écart maximum. *Annales de la Société Polonaise de Mathématique*, 6:93–116. 5
- [40] Furman, E. and Zitikis, R. (2008a). Weighted premium calculation principles. *Insurance: Mathematics and Economics*, 42(1):459–465. 66
- [41] Furman, E. and Zitikis, R. (2008b). Weighted risk capital allocations. *Insurance: Mathematics and Economics*, 43(2):263–269. 66
- [42] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44(3):423–453. 5, 11
- [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(2):227–245. 66
- [44] Gumbel, E. J. (1958). Statistics of extremes. *Columbia University Press, Columbia*. 5

- [45] Haeusler, E. and Teugels, J. L. (1985). On asymptotic normality of hill's estimator for the exponent of regular variation. *The Annals of Statistics*, 13(2):743–756. [41](#)
- [46] Henry, J. (2009). A harmonic moment tail index estimator. *Journal of Statistical Theory and Applications*, 8(2):141–162. [36](#)
- [47] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174. [iv](#), [v](#), [31](#), [33](#), [75](#), [77](#)
- [48] Hürlimann, W. (1998). Inequalities for lookback option strategies and exchange risk modelling. In *In Proceedings of the 1st Euro-Japanese Workshop on Stochastic Modelling for Finance, Insurance, Production and Reliability, Brussels*. [64](#)
- [49] J. McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management - Concepts, Techniques and Tools*. Princeton Series in Finance. Princeton University Press. [52](#)
- [50] Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171. [5](#), [15](#)
- [51] Johan, S. (2005). Generalized pickands estimators for the extreme value index. *Journal of Statistical Planning and Inference*, 128(2):381–396. [33](#)
- [52] Jones, B. L. and Zitikis, R. (2003). Empirical estimation of risk measures and related quantities. *North American Actuarial Journal*, 7(4):44–54. [65](#), [66](#), [68](#), [76](#)
- [53] Jones, B. L. and Zitikis, R. (2005). Testing for the order of risk measures: an application of l-statistics in actuarial science. *Metron*, 63(2):193–211. [66](#)
- [54] Jones, B. L. and Zitikis, R. (2007). Risk measures, distortion parameters, and their empirical estimation. *Insurance: Mathematics and Economics*, 41(2):279–297. [66](#)
- [55] Juárez, S. F. and Schucany, W. R. (2004). Robust and efficient estimation for the generalized pareto distribution. *Extremes*, 7(3):237–251. [35](#), [75](#)
- [56] Kaas, R., Goovaerts, M., Dhaene, J., and Denuit, M. (2008). *Modern Actuarial Risk Theory: Using R*. Springer-Verlag Berlin Heidelberg, 2 edition. [59](#)

- [57] Kim, M. and Lee, S. (2008). Estimation of a tail index based on minimum density power divergence. *Journal of Multivariate Analysis*, 99(10):2453–2471. [35](#), [75](#)
- [58] Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer-Verlag New York, 1st edition. [2](#), [6](#), [11](#)
- [59] Mason, D. M. (1981). Asymptotic normality of linear combinations of order statistics with a smooth score function. *The Annals of Statistics*, 9(4):899–904. [76](#)
- [60] Mason, D. M. (1982). Laws of large numbers for sums of extreme values. *The Annals of Probability*, 10(3):756–764. [41](#)
- [61] Matthys, G. and Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica*, 13(3):853–880. [82](#)
- [62] Necir, A. and Boukhetala, K. (2004). Estimating the riskadjusted premium for the largest claims reinsurance covers. In *COMPSTAT 2004—Proceedings in Computational Statistics*, pages 1577–1584. [66](#), [69](#), [70](#)
- [63] Necir, A. and Meraghni, D. (2009). Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance: Mathematics and Economics*, 45(1):49–58. [iv](#), [v](#), [4](#), [66](#), [69](#), [70](#), [71](#), [72](#), [73](#), [75](#), [78](#), [82](#)
- [64] Necir, A. and Meraghni, D. (2010). Estimating l-functionals for heavy-tailed distributions and application. *Journal of Probability and Statistics*, 2010:34. [66](#), [71](#)
- [65] Necir, A., Meraghni, D., and Meddi, F. (2007). Statistical estimate of the proportional hazard premium of loss. *Scandinavian Actuarial Journal*, 3(3):147–161. [66](#), [69](#), [70](#), [72](#), [83](#), [84](#), [85](#)
- [66] Necir, A., Rassoul, A., and Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics*, 2010(ID 596839):1–17. [70](#), [72](#)
- [67] Peng, L. (2001). Estimating the mean of a heavy tailed distribution. *Statistics & Probability Letters*, 52(3):255–264. [66](#), [70](#)

- [68] Peng, L. and Qi, Y. (2004). Estimating the first- and second-order parameters of a heavy-tailed distribution. *Australian & New Zealand Journal of Statistics*, 46(2):305–312. 77
- [69] Peng, L., Qi, Y., Wang, R., and Yang, J. (2012). Jackknife empirical likelihood method for some risk measures and related quantities. *Insurance: Mathematics and Economics*, 51(1):142–150. 66, 70
- [70] Peng, L. and Welsh, A. (2001). Robust estimation of the generalized pareto distribution. *Extremes*, 4(1):53–65. 35, 75
- [71] Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3. 28, 31, 40
- [72] Rassoul, A. (2012). Reduced bias estimation of the reinsurance premium of loss distribution. *Journal of Statistics Applications & Probability*, 1(2):147. 66, 70
- [73] Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields*. Basel: Birkhäuser Verlag AG, 3rd edition. 2, 6, 43, 80
- [74] Resnick, S. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag New York, 1st edition. 2, 6, 11, 12, 21, 22, 24, 25
- [75] Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of risk*, 2(3):21–42. 53
- [76] Rolski, T., Schmidli, H., Schmidt, V., and Teugels, J. (1999). *Stochastic processes for insurance and finance*, volume 505. John Wiley & Sons. 59, 74
- [77] Seneta, E. (1976). *Regularly Varying Functions*. Lecture Notes in Mathematics 508. Springer-Verlag Berlin Heidelberg, 1st edition. 21
- [78] Shorack, G. R. and Wellner, J. A. (1986). *Empirical processes with applications to statistics*. Wiley. 66, 76
- [79] Smith, R. L. (1987). Estimating tails of probability distributions. *The Annals of Statistics*, 15(3):1174–1207. 41
- [80] Stehlík, M., Fabián, Z., and Střelec, L. (2012). Small sample robust testing for normality against pareto tails. *Communications in Statistics - Simulation and Computation*, 41. 35, 36, 42, 75, 77, 78

- [81] Stehlík, M., Potocký, R., Waldl, H., and Fabián, Z. (2010). On the favorable estimation for fitting heavy-tailed data. *Computational Statistics*, 25(3):485–503. 36
- [82] Stigler, S. M. (1974). Linear functions of order statistics with smooth weight functions. *The Annals of Statistics*, 2(4):676–693. 76
- [83] Vandewalle, B. and Beirlant, J. (2006). On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance: Mathematics and Economics*, 38(3):441–459. 69
- [84] Vandewalle, B., Beirlant, J., Christmann, A., and Hubert, M. (2007). A robust estimator for the tail index of pareto-type distributions. *Computational Statistics & Data Analysis*, 51(12):6252–6268. 35, 75
- [85] von Mises, R. (1936). La distribution de la plus grande de n valeurs. *Rev. math. Union interbalcanique*, 1:141–160. Reproduced. Selected papers of Richard von Mises, American Mathematical Society, 1964, 2: 271–294. 5, 15
- [86] Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics & Economics*, 17(1):43–54. 59, 63
- [87] Wang, S. (1996). Premium calculation by transforming the layer premium density. *Astin Bulletin*, 26(01):71–92. 59, 74
- [88] Wang, S. (2001). A risk measure that goes beyond coherence. *Research Report*, pages 1–18. 59
- [89] Wang, S. S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance*, 67(1):15–36. 59, 64
- [90] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815. 16, 44, 69, 77
- [91] Wirch, J. L. and Hardy, M. R. (1999). A synthesis of risk measures for capital adequacy. *Insurance: Mathematics and Economics*, 25(3):337–347. 53, 62, 63, 74
- [92] Wirch, J. L. and Hardy, M. R. (2001). Distortion risk measures: coherence and stochastic dominance. In *International Congress on Insurance: Mathematics and Economics*, pages 15–17. 62

-
- [93] Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica: Journal of the Econometric Society*, pages 95–115. 59
- [94] Young, V. R. (2004). Premium principles. *Encyclopedia of actuarial science*. 57