HAL
open science

Quelques contributions à la Théorie univariée des Valeurs Extrêmes et Estimation des mesures de risque actuariel pour des pertes à queues lourdes

El Hadji Deme

To cite this version:

El Hadji Deme. Quelques contributions à la Théorie univariée des Valeurs Extrêmes et Estimation des mesures de risque actuariel pour des pertes à queues lourdes. Méthodologie [stat.ME]. Université Gaston Berger, Saint Louis, Sénégal, 2013. Français. NNT : . tel-01069382

HAL Id: tel-01069382

https://theses.hal.science/tel-01069382

Submitted on 29 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES

DE L'UNIVERSITÉ GASTON BERGER DE SAINT-LOUIS

Spécialité : MATHEMATIQUES APPLIQUÉES

Par **EL HADJI DEME**

Quelques contributions à la Théorie univariée des Valeurs Extrêmes et Estimation des mesures de risque actuariel pour des pertes à queues lourdes

Soutenance le 5 Juin 2013 devant le jury composé

Président:	Ali Souleyman DABYE	Professeur	UGB, Sénégal
Rapporteurs:	Maria Ivette GOMEZ	Professeur	Lisbonne, Portugal
	Diamel MERAGHNI	Professeur	Biskra, Algérie
	Papa NGOM	Maître de Conférence	UCAD, Sénégal
Examinateurs:	Abdou Kâ DIONGUE	Maître de Conférence	UGB, Sénégal
Directeurs de thèse:	Gane Samb LO	Professeur	UGB, Sénégal
	Aliou DIOP	Professeur	UGB, Sénégal
	Stéphane GIRARD	HDR	INRIA Rhône Alpes, France

Thèse préparée au sein du Laboratoire d'Études et de Recherches en Statistique et Développement de l'UFR SAT
et de l'Equipe Mistis de l'INRIA Rhône-Alpes de Grenoble.

Dédicace

A la mémoire de ma Grand mère Feu Séné Deme,

A mes parents,

A ma famille, à mes amis et à tous ceux qui me sont chers.

Remerciements

Je tiens tout d'abord à exprimer ma profonde reconnaissance à mon directeur de thèse le Professeur Gane Samb Lô et mes co-directeurs de thèse le Professeur Aliou Diop et Monsieur Stéphane Girard (HDR à l'inria de Grenoble). Je leur adresse mes plus sincères remerciements pour la confiance qu'ils m'ont accordée, pour leur disponibilité et leur précieux conseils et d'avoir permis à ce travail de voir le jour. Merci infiniment.

Je suis très honoré que les professeurs Maria Ivette Gomes (Université de Lisbonne, Portugal), Diamel Mereghni (Université Mohamed khider de Biskra, Algérie) et Papa Ngom (Université Cheikh Anta Diop de Dakar) aient accepté de rapporter ma thèse. Je voudrais exprimer ma gratitude au Professeur Ali Souleyman Dabye (Université de Gaston Berger) pour l'intérêt qu'il a bien voulu porter à ce travail et avoir accepté de présider mon jury. Mes remerciements vont également au Professeur Abdou Ka Diongue (Université de Gaston Berger) qui a accepté de faire partie de mon jury.

Je suis très reconnaissant aux professeurs Laurent Gardes (Université de Strasbourg) et Armelle Guillou (Université de Strasbourg) pour leurs disponibilités et avec qui j'ai eu l'opportinuité de travailler pendant mon séjour en France

Toute ma reconnaissance aux programmes AIRES-SUD (Appuis Intégrés pour le Renforcement des Equipes Scientifiques du Sud), FIRST (Fonds d'Impulsion pour la Recherche Scientifique et Technique) qui ont contribué financièrement à ma thèse. Je remercie également l'institut Nationale de Recherche en Informatique et Automatisme (INRIA) et le Service de Coopération et d'Action Culturelle (SCAC) de l'ambassade de france qui m'ont accordé trois fois de suite une bourse d'excellence pour des stages au sien de l'équipe Mistis de l'inria de Grenoble.

Un remerciement particulier à l'equipe Mistis de l'inria de Grenoble qui m'a acceuilli et dans la quelle j'ai pu travailler dans une ambillance trèes chaleureuse.

Il ne serait pas possible de travailler à l'université sans eux : merci aux personnels admisnistratifs et technique aux enseignants de l'UFR SAT. Je remercie particulièrement Monsieur Mamadou Sy (Professeur à l'UGB) pour son aide et tous ses conseils.

Je remercie tous les doctorants et jeunes docteurs du Laboratoire d'Etudes et de Recherches en Statistique et Développement (LERSTAD) avec qui j'ai passé de très bons moments et ainsi que ceux du Laboratoire d'Analyse Numérique et Informatique (LANI) pour leur soutien moral, pour les échanges fructueux lors des séminaires et discussions et

avec qui j'ai partagé mes angoisses, mes rêves et des moments inoubliables.

Merci à mes amis : Pape Djiby Mergane, Youssouf, Modou Lo, Christine, Jonathan, Cai, Lamia, Zildas, Abou Sene, Lamine Diagne, Ndeye Fatou Koné, Sara, Mbaye babacar Gueye, Junior, Sadou Dahaltou, Amadou Thiam, Mandiaye Fall, Bara Ndiaye, Adum, Rokhaya Dia, Waly Ngom, Babacar Fall, Mbargou Deme, Babacar Lèye, Amina Diop, Moustapha Kane, Donald et toute la cominauté Tchiadiène de l'UGB d'avoir passer de bons moments avec moi.

Je dédie une mention spéciale et toute particulière à mes parents surtout pour leur patience, à ma famille et à mes amis les plus proches.

Propos de Cont (2009)

« La loi des grands nombres et la distribution Gaussienne, fondements de l'étude statistique des grandeurs moyennes, échouent à rendre compte des événements rares ou extrêmes. Pour ce faire, des outils statistiques plus adaptés existent mais ne sont pas toujours utilisés ».

Résumé

Quelques contributions à la Théorie univariée des valeurs extrêmes et estimation des mesures de risque actuariel pour des pertes à queues lourdes

Cette thèse est divisée en cinq (5) chapitres auxquels s'ajoutent une introduction et une conclusion. Dans le premier chapitre, nous rappelons quelques notions de base sur la théorie des valeurs extrêmes. Dans le deuxième chapitre, nous considérons un processus statistique dépendant d'un paramétre continu $\boldsymbol{\tau} > 0$ et dont chaque marge peut être considérée comme un estimateur de Hill généralisé (Hill, 1975). Ce processus statistique a été introduit et étudié par Diop et Lo (2006), Diop et Lo (2009) pour estimer l'indice des valeurs extrêmes et permet de discriminer entièrement les domaines d'attraction des valeurs extrêmes. La normalité asymptotique de ce processus statistique a été seulement donnée pour $\boldsymbol{\tau} > 1/2$. Nous complétons cette étude pour $0 < \boldsymbol{\tau} \leq 1/2$, en donnant une approximation des domaines de Gumbel et de Fréchet. Des études de simulations effectuées avec le logiciel « R », permettent de montrer la performance de ces estimateurs. Comme illustration, nous proposons une application de notre méthodologie aux données hydrauliques. Dans le troisième chapitre, nous étendons l'étude du processus statistique précédent dans un cadre fonctionnel. Nous proposons donc un processus stochastique dépendant d'une fonctionnelle positive pour obtenir une grande classe d'estimateurs de l'indice des valeurs extrêmes dont chaque estimateur est une marge d'un seul processus stochastique. L'étude théorique de ces processus stochastiques que nous avions menée, est basée sur la théorie moderne de convergence vague fonctionnelle. Cette dernière permet de gérer des estimateurs plus complexes sous forme de processus stochastiques (voir par exemple van der Vaart et Wellner (1996)). Nous donnons les distributions asymptotiques fonctionnelles de ces processus et nous montrons que pour certaines classes de fonctions, nous avons un comportement asymptotique non Gaussien et qui sera entièrement caractérisé. Dans le quatrième chapitre, on s'intéresse à l'estimation du paramètre du second ordre. Notons que ce paramètre joue un rôle très important dans le choix adaptatif du nombre optimal de valeurs extrêmes utilisé lors de l'estimation de l'indice des valeurs extrêmes. L'estimation de ce paramètre est également utilisée pour la réduction du biais des estimateurs de l'indice de queue et a reçu une grande attention dans la littérature des valeurs extrêmes, on peut voir par exemple (Hall et Welsh, 1985; Gomes et Martins, 2002; Fraga Alves *et al.*, 2003b; Ciuperca

et Mercadier, 2010; Goegebeur *et al.*, 2010) Nous proposons une simple et générale approche pour estimer le paramètre du second ordre, permettant de regrouper un grand nombre d'estimateurs. Il est montré que les estimateurs cités précedemment peuvent être vus comme des cas particuliers de notre approche. Nous tirons également parti de notre formalisme pour proposer de nouveaux estimateurs asymptotiquement Gaussiens du paramètre du second ordre. Finalement, certains estimateurs sont comparés tant du point de vue asymptotique que performance sur des échantillons de tailles finies. Comme illustration, nous proposons une application sur des données d'assurance. Dans le dernier chapitre, on s'intéresse aux mesures de risque actuariel pour des phénomènes capables d'engendrer des pertes financières très importantes (ou phenomènes extrêmes c'est-à-dire à des risques dont on ne sait pas si le système d'assurance sera capable de les supporter). De nombreuses mesures de risque ou principes de calcul de la prime ont été proposés dans la littérature actuarielle. Nous nous concentrons sur la prime de risque-ajustée introduite par Wang (1996). Jones et Zitikis (2003) ont donné une estimation de cette dernière basée sur la distribution empirique et ont établi sa normalité asymptotique sous certaines conditions appropriées, et qui ne sont pas souvent remplies dans le cas des distributions à queues lourdes. Ainsi, nous regardons ce cadre là et nous considérons une famille d'estimateurs de la prime de risque-ajustée basée sur l'approche de la théorie des valeurs extrêmes. Nous établissons leur normalité asymptotique et nous proposons également une approche de réduction de biais pour ces estimateurs. Des études de simulation permettent d'apprécier la qualité de nos estimateurs. Comme illustration, nous proposons une application sur des données d'assurance.

Mots-clés : Théorie des valeurs extrêmes, processus statistique, estimation, queue lourde, indice de queue, paramètre du second ordre, estimateur de Hill, prime de risque, propriétés asymptotiques.

Abstract

Some contributions to the univariate Extreme Value Theory and estimation of actuarial risk measures for heavy-tailed losses

This thesis is divided into five (5) chapters with an additional introduction and a conclusion. In the first chapter, we recall some basics on extreme value theory. In the second chapter, we consider a statistical processes depending on a continuous time $\boldsymbol{\tau} > 0$ and whose any margin can arise as a Generalized Hill's estimator (Hill, 1975). This statistical processes was introduced and studied by Diop et Lo (2006), Diop et Lo (2009) to estimate the extreme value index and allows to discriminate completely the extreme values domain of attraction. The asymptotic normality of this statistical processes was only given for $\boldsymbol{\tau} > 1/2$. We complete this study for $0 < \boldsymbol{\tau} \leq 1/2$, and we give an approximation to the domain of attraction of Gumbel and Fréchet. Simulation studies carried out with the « R »software, can show the performance of these estimators. As an illustration, we propose an application of our methodology for hydraulic data. In the third chapter, we extend the study of the previous statistical processes to a functional framework. We propose a stochastic processes depending on a positive functional to obtain a large class of estimators of extreme values index which each estimator is a marginal one stochastic processes. The theoretical study of these stochastic process that we lead is based on the modern theory of functional weak convergence, which handle more complex estimators in the form of stochastic processes (see for example van der Vaart et Wellner (1996)). We give the asymptotic distributions functions of these processes and we show that for some class, we have a non Gaussian asymptotic behavior, which will be fully characterized. In the fourth chapter, we are interested in the estimation of the second order parameter. Note that this parameter is of primordial importance in the adaptive choice of the best number of upper order statistics to be considered in the estimation of the extreme-value index. The estimation of the second order parameter can also be used to propose bias reduced estimators of the extreme value index, and has received a lot of attention in the extreme-value literature (Hall et Welsh, 1985; Gomes et Martins, 2002; Fraga Alves *et al.*, 2003b; Ciuperca et Mercadier, 2010; Goegebeur *et al.*, 2010). We propose a simple and general approach to estimate the second order parameter. It is shown that many estimators in the extreme-value literature can be read as particular cases of our approach. We illustrate how a lot of new asymptotically Gaussian estimators can be derived from this framework. Finally, some estimators are compared both from the asymptotic

finite sample size performances points of view. As an illustration, we propose a case study in the field of insurance. In the last chapter, we are interested in the actuarial risk premiums for some phenomena generating very significant financial losses (or extreme phenomenas). Many Risk measures or premium calcutation principles have been proposed in the actuarial literature. We focus on the adjusted risk premium introduced by Wang (1996). Jones et Zitikis (2003) gave an estimate of the adjusted risk premium based on the empirical distribution and have established its asymptotic normality under some appropriate conditions, and which are not fulfilled in the case of heavy-tailed distributions (case of extreme phenomena). we look at this framework and we consider a family of estimators of the ajusted risk premium based on the extreme values theory approach. We establish their asymptotic normality and we propose also an approach to bias reduction for these estimators. A simulation study is used to evaluate the quality of our estimators. As an illustration, we propose an application on real data insurance.

Keywords: Extreme values theory, statistical process, estimation, heavy-tailed, tail index, second order parameter, Hill's estimator, risk premium, asymptotic properties.

Abréviations & Notations

Variables aléatoires et modes de convergence

$\mathbb{P}(A)$	: La probabilité de A.
$\mathbb{P}^*(A)$	: La probabilité extérieure de A.
$\mathbb{E}(X)$	: L'espérance mathématique de X.
$\mathbb{E}^*(T)$	: Intégrale extérieure de T.
$\overset{\mathcal{D}}{=}$	: Egalité en distribution
$\overset{\mathcal{D}}{\sim}$	: Equivalence en distribution
$\overset{\mathcal{D}}{\longrightarrow}$	: Convergence en loi
$\overset{\mathbb{P}}{\longrightarrow}$	: Convergence en probabilité
$\overset{p.s}{\longrightarrow}$	: Convergence presque sûre
$\overset{\mathcal{D}}{\rightsquigarrow}$	: Convergence vague
$o_{\mathbb{P}}(\cdot)$	: Tendre vers 0 en probabilité
$O_{\mathbb{P}}(\cdot)$	: Etre bornée en probabilité.
$o^*_{\mathbb{P}}(\cdot)$	: Tendre vers 0 en probabilité extérieure
$O^*_{\mathbb{P}}(\cdot)$	: Etre bornée en probabilité extérieure.
$p.s$	: Presque sûrement.

Notations d'ordre général

$\in$	: Appartient
$\mathbb{N}$	: Ensemble des entiers naturels.
$\mathbb{N}^*$	: Ensemble des entiers naturels non nuls.
$\mathbb{R}$	: Ensemble des réels et $\mathbb{R}^d = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{d \text{ fois}}$, $d \in \mathbb{N}^*$
${}^tX$	: Transposée du vecteur X
$\lvert x \rvert$	: Valeur absolue de la variable x
$\mathbb{1}_A$	: Fonction indicatrice de l' ensemble A
$[x]$	: Partie entière de x
i.e	: C'est-à-dire
ssi	: Si et seulement si.
$\mathcal{N}(0,1)$	: Loi normal standard
$\mathcal{N}(\mu, \sigma^2)$	: Loi normale de moyenne μ et de variance σ^2
$\mathcal{N}_d(m, \Sigma)$	: Loi normale multivariée de moyenne $m \in \mathbb{R}^d$ et de matrice de variance covariance Σ
$D(H)$	: Domaine d'attraction de la distribution H
x_F	: Point terminal de la distribution F
$\max(A)$	: Maximum de l'ensemble A.
$\min(A)$	: Minimum de l'ensemble A.
$\ell^\infty(\mathcal{F})$	: L'espace des fonctions réelles bornées définies dans $\mathcal{F}$.

Table des matières

Dédicace i

Remerciements ii

Propos de Cont (2009) iv

Résumé v

Abstract vii

Abréviations & Notations ix

Introduction générale xiv

1 Présentation de la Théorie des Valeurs Extrêmes 1

1.1 Introduction . . . 1

1.2 Comportement asymptotique des extrêmes . . . 2

1.3 Distribution conditionnelle des excès . . . 6

1.4 Caractérisation des Domaines d'attraction . . . 7

1.4.1 Fonctions à variation régulière . . . 7

1.4.2 Domaine d'attraction de Fréchet . . . 9

1.4.3 Domaine d'attraction de Weibull . . . 10

1.4.4 Domaine d'attraction de Gumbel . . . 11

1.4.5 Caractérisation générale . . . 11

1.5 Estimation des paramètres . . . 13

1.5.1 Estimation de l'indice de queue sous le modèle de type Paréto . . . 14

1.5.1.1 Le Paréto Quantile Plot . . . 15

1.5.1.2 Estimateur de Hill . . . 16

1.5.1.3 Correction de biais sur l'estimateur de Hill . . . 19

1.5.2 Estimation pour différents domaines d'attraction . . . 22

1.5.2.1 L'estimateur de Pickands . . . 22

1.5.2.2 Estimateur des moments . . . 23

1.5.3 Estimateurs du couple $(\boldsymbol{\gamma}, \boldsymbol{\sigma})$ . . . 26

1.5.3.1 Estimateurs par la méthode des moments . . . 26

1.5.3.2 Estimateurs par la méthode des moments pondérés . . . 27
1.6 Choix du nombre d'extrêmes k_n . . . 27
1.6.1 Méthode Graphique . . . 28
1.6.2 Méthode basée sur l'erreur moyenne quadratique . . . 29
1.6.3 Procédure numérique . . . 29
1.7 Estimation de quantiles extrêmes . . . 30
1.7.1 Approche basée sur un estimateur d'indices positifs ($\gamma > 0$) . . . 31
1.7.2 Approche basée sur un estimateur d'indices quelconques . . . 32
1.7.3 Approche basée sur la méthode des excès . . . 33

2 Processus statistique de paramètre continu généralisant l'estimateur de Hill 34
2.1 Contexte d'étude et définitions . . . 34
2.1.1 Introduction . . . 34
2.1.2 Représentation pour $F \in D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\gamma})$, $\gamma > 0$. . . . 36
2.2 Resultats asymptotiques . . . 38
2.3 Applications et Remarques . . . 41
2.3.1 Estimateurs asymptotiquement Gaussiens . . . 41
2.3.2 Illustration des Conditions (**C1**), (**C2**) et (**C3**) pour $\boldsymbol{\tau} = 1/2$ . . . 42
2.3.3 Test statistique sur le paramètre $\boldsymbol{\gamma}$ pour $\boldsymbol{\tau} = 1/2$ . . . 44
2.3.4 Estimateurs asymptotiquement non Gaussiens . . . 45
2.3.4.1 Illustration de la fonction de répartition de $\mathcal{L}(\boldsymbol{\tau})$ . . . 47
2.4 Etude de simulation . . . 48
2.5 Illustration sur des données hydrauliques . . . 49
2.6 Démonstrations . . . 52
2.6.1 Résultats préliminaires . . . 52
2.6.2 Démonstration des résultats théoriques . . . 55

3 Processus fonctionnel de Hill généralisé et théorie uniforme 62
3.1 Introduction . . . 62
3.2 Distributions finies et ponctuelles de $\boldsymbol{T}_{k_n}(f)$ . . . 65
3.3 Loi fonctionnelle de $\boldsymbol{T}_{k_n}(f)$ . . . 68
3.4 Quelques cas speciaux . . . 71
3.5 Démonstrations . . . 72
3.5.1 Résultats préliminaires . . . 72
3.5.2 Démonstration des résultats théoriques . . . 76

4 Famille d'estimateurs semi-paramétriques du paramètre du second ordre 81
4.1 Introduction 81
4.2 Propriétés asymptotiques 84
4.3 Exemples de variables aléatoires de type T_n 85
4.4 Applications 89
4.4.1 Estimateurs basés sur la statistique $R_{k,n}(\boldsymbol{\tau})$ 89
4.4.2 Estimateurs basés sur la statistique $S_{k,n}(\boldsymbol{\tau})$ 92
4.5 Comparaisons de quelques estimateurs 96
4.5.1 Estimateurs basés sur la variable aléatoire $R_{k,n}(\boldsymbol{\tau})$ 96
4.5.1.1 Borne supérieure de l' Erreur Quadratique Moyenne Asymptotique 96
4.5.1.2 Illustration sur une distribution de Burr 97
4.5.2 Estimateurs basés sur la variable aléatoire $S_{k_n}(\tau, \alpha)$ 98
4.5.2.1 Comparaison des termes de l'erreur quadratique moyenne asymptotique 98
4.5.2.2 Comparaison sur une distribution de Burr simulée 100
4.5.3 Tentative de conclusion 100
4.5.4 Illustrations sur des données d'assurance 101
4.6 Démonstrations des résultats théoriques 101

5 Estimation des mesures de risque actuariel pour des pertes à queues lourdes 105
5.1 Problématiques et définitions 105
5.2 Mesures de risque actuariel 107
5.2.1 Définitions et propriétés 107
5.2.2 Mesures de rique de Wang 108
5.2.3 Principe de la prime ajustée 109
5.2.4 Estimation de la prime ajustée 110
5.2.4.1 Approche du quantile empirique 110
5.2.4.2 Approche des quantiles extêmes 111
5.3 Résultats principaux 113
5.3.1 Resultats asymptotiques de $\widehat{\boldsymbol{\gamma}}_{k_n}^K$ 113
5.3.2 Résultats asymptotique de $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$ 115
5.3.3 Correction de Biais pour $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$ 116
5.4 Etude de simulations et de données réelles 118

5.4.1 Etude de simulations 118
5.4.2 Illustration sur des données d'assurance incendie 119
5.5 Démonstrations 120
5.5.1 Résultats préliminaires 120
5.5.2 Démonstrations des résultats théoriques 122
5.6 Démonstrations des résultats préliminaires 126

Conclusions et perspectives **135**

A Convergence de mesures **138**

B Démonstration du Théorème 3.2.1 **140**

C Démonstration du Corollaire 3.4.2 **144**

D Vérification de l' hypothèse (L3) pour $\tau > 1/2$ **147**

E Une identité utile **150**

Bibliographie **153**

Introduction générale

Généralité

La théorie des valeurs extrêmes (TVE) est une branche des statistiques qui a pour but de modéliser et de décrire la survenue et l'intensité d'évènements dits *rares* c'est-à-dire qui présentent des variations de très grandes amplitudes (ayant une faible probabilité d'apparition). Lorsque le comportement de ces événements est dû au hasard, on peut étudier leur loi. Ils sont dits *extrêmes* quand il s'agit de valeurs beaucoup plus grandes ou plus petites que celles observées habituellement.

Les événements rares et catastrophiques (tremblements de terre, inondations, accidents nucléaires, crises monétaires ou financières, Crachs boursiers, émergence d'un nouveau phénomène endémique, etc.) dominent l'actualité quotidienne par leur caractère imprévisible. Compte tenu de l'importance des enjeux sociaux et scientifiques, aucun débat sérieux sur le hasard ne saurait être mené sans une réflexion sur les événements rares et extrêmes.

La théorie statistique « classique » (et notamment le théorème bien connu de la limite centrale) permet de faire de l'inférence sur les valeurs centrales d'un échantillon mais ne donne que très peu d'informations sur la queue de distribution. La particularité de la théorie des valeurs extrêmes est qu'elle se concentre sur la queue de distribution engendrant les divers phénomènes extrêmes étudiés. Elle est développée pour l'estimation de la probabilité d'événements d'occurrence rares et permet d'obtenir des estimations fiables des valeurs extrêmes, pour lesquelles les observations sont peu nombreuses. Le millénaire passé et le début de ce millénaire ont connu beaucoup d'évènements extrêmes et c'est dans ce contexte que la théorie des valeurs extrêmes développée par Fisher et Tippet (1928) a trouvé sa place en

- *Climatologie*: étude des évènements climatiques extrêmes (précipitations, températures, chutes de neige), modélisation des grands feux de forêt (voir par exemple Alvarado *et al*, 1998);
- *Hydrologie*: crues consécutives à des pluies torrentielles : Au Sénégal (inondation de 1994 et celle de 2003 entraînant l'ouverture de la Brêche); aux Pays-Bas, digues menacées par l'effet conjoint des grandes marées et des conditions climatiques en Mer du Nord (Inondation de 1953);

- *Assurance*: survenue des sinistres d'intensité exceptionnelle (ouragan Katrina en 2005, importants incendies en risques industriels, sinistres graves en responsabilité civile automobile) qui peuvent avoir des conséquences négatives sur les résultats et la solvabilité des organismes d'assurance;
- *Finance*: fortes variations du cours d'actifs financiers, gestions du risque opérationnel des banques (crise des subprimes, fin des années 2000).

Parmi les précurseurs, on peut citer en particulier Gnedenko (1943), Leadbetter (1983), Resnick (1987), Embrechts *et al* (1997), de Haan et Ferreira (2006).

Présentation des travaux de recherche

Le thème principal de cette thèse repose sur l'étude des valeurs extrêmes et s'articule autour de trois axes principaux : estimation fonctionnelle de l'indice des valeurs extrêmes, esimation sémi-paramétrique du paramètre du second ordre et étude des mesures de risque acturiel pour des pertes extêmes.

Notre contribution

Cette thèse peut être considérée dans sa globabilité comme une contribution à la théorie univariée des valeurs extrêmes et ses applications statistiques.

Contribution du Chapitre 1

Le Chapitre 1 est surtout un chapitre de généralités qui permet de définir l'ensemble des notions et des résultats que nous utiliserons dans la suite de ce travail. Nous commençons par rappeler les résultats sur les distributions asymptotiques en théorie des valeurs extrêmes. Ces distributions dépendent d'un paramètre réel inconnu γ appelé l'*indice de queue* ou *l'indice des valeurs extrêmes* et dont sa connaisance est d'une grande importance en théorie des valeurs extrêmes. Nous verrons que ce paramètre contrôle le comportement de la queue de distribution. Plus γ est grand, plus la queue est lourde et donc son estimation a un grand intérêt. Nous rappelons également quelques méthodes d'estimation de ce paramètre, les procédures de selection du nombre de valeurs extrêmes utilisé pour l'estimation de ce paramètre. Et enfin, nous passons à l'estimation des quantiles extrêmes. Dans chacune de ces étapes, nous présentons les résultats les plus importants qui nous seront utiles dans les chapitres qui suivent.

Contribution du Chapitre 2

Dans le Chapitre 2, nous nous intéressons à une classe d'estimateurs de l'indice de queue γ, dépendant d'un paramètre continu $\tau > 0$, appelé processus statistique de Hill généralisé. Cette classe d'estimateurs de l'indice des valeurs extrêmes γ a été introduite et étudiée par Diop et Lo (2006) et Diop et Lo (2009) et par sa formulation, elle peut être vue comme l'estimateur à noyau de Csörgő *et al.* (1985) pour un choix du noyau $K(u) = u^{\tau-1}\mathbb{1}_{\{0<u<1\}}$. Lorsque le paramètre $\tau = 1$, alors l'estimateur obtenu est celui de Hill (1975), qui est l'estimateur le plus populaire en théorie des valeurs extrêmes. Pour $\tau > 0$, Diop et Lo (2006) ont étudié la consistance forte de ce type d'estimateur et ont donné sa variance asymptotique. Dans Csörgő *et al.* (1985) et Diop et Lo (2009), sous un certain nombre d'hypothèses la normalité de cette classe d'estimateurs a été etudiée seulement dans le cas $\tau > 1/2$. Le cas $0 < \tau \leq 1/2$ pose un problème du fait que les hypothèses utilisées par Csörgő *et al.* (1985) et Diop et Lo (2009) ne sont plus vérifiées, et donc leurs résultats ne sont plus valides pour ce cas. Ainsi, nous complétons les travaux précédents, en donnant la distribution asymptotique de cette classe d'estimateurs pour $0 < \tau \leq 1/2$. L'étude théorique de cette classe d'estimateurs montre que dans le cas $\tau = 1/2$ la distribution asymptotique de cette famille d'estimateurs est gaussienne. Par contre, elle est non gaussienne pour $0 < \tau < 1/2$. Cette contribution a donné lieu à la publication Lo *et al.* (2012) et a fait l'objet d'une présentation au Workshop:

- International Worksohp on Multivariate Risks and Copuas, April 12-15, 2010, Université Mohamed Khider Biskra, Algérie.

Toutefois, notons que la contribution présentée dans ce chapitre est plus générale par rapport à l'article cité précédemment.

Contribution du Chapitre 3

Dans le Chapitre 3, nous étendons les résultats obtenus dans le Chapitre 2 pour le processus de Hill Généralisé de l'indice des valeurs extrêmes γ dans un cadre fonctionnel. Nous considérons un processus dépendant d'une fonctionnelle $f : \mathbb{N}^* = \mathbb{N}\backslash\{0\} \longmapsto \mathbb{R}_+$. Nous appelerons ces processus, processus stochastiques des estimateurs de l'indice des valeurs extrêmes.

La motivation principale de ce travail est d'obtenir une grande classe d'estimateurs de l'indice des valeurs extrêmes dont chaque estimateur est un marginal d'un seul processus stochastique généralisant l'estimateur de Hill (1975). Pour la fonction uniforme $f(j) = j$, ce processus correspond à l'estimateur de Hill. Le chapitre 2 donne une première étape de l'estimateur de Hill fonctionnel associé à la fonction $f(j) = j^{\tau}$, $\tau > 0$ et respective-

ment étudié dans Diop et Lo (2006) et Diop et Lo (2009), pour $\boldsymbol{\tau} > 1/2$ et dans Lo *et al.* (2012) pour $0 < \boldsymbol{\tau} \leq 1/2$ en distributions finies. Groeneboom *et al.* (2003) ont également examiné une étude approfondie d'une famille de noyau d'estimateurs de l'indice des valeurs extrêmes. Cependant, ils ne considèrent pas une version avec les processus stochastiques. L'étude théorique de ces processus stochastiques est basée sur la théorie moderne de convergences vagues fonctionnelles qui permet de gérer des estimateurs plus complexes sous forme de processus stochastiques (on pourra se référrer à van der Vaart et Wellner (1996)). Ainsi, nous établissons la normalité asymptotique fonctionnelle de ces processus stochastiques des estimateurs de l'indice des valeurs extrêmes ou simplement leur distribution asymptotique pour des classes appropriées. Nous montrons que pour certaines classes, nous aurons un comportement asymptotique non Gaussien et qui sera entièrement caractérisé. Cette contribution à donné lieu à la publication Lo et Deme (2012).

Contribution du Chapitre 4

Dans le Chapitre 4, on s'intéresse uniquement à la classe des distributions à *queues lourdes* « *Heavy-tailed* ». Une caractérisation simple pour les distributions F appartenant à cette classe est donnée par $1 - F(x) = x^{-1/\boldsymbol{\gamma}}\ell(x)$, $\boldsymbol{\gamma} > 0$, où ℓ est une fonction vérifiant pour tout $\lambda > 0$, $\ell(\lambda x)/\ell(x) \to 1$, lorsque $x \to \infty$ et est appelée fonction à variations lentes. Ainsi, ce paramètre contrôle le comportement de la queue de distribution (plus γ est grand, plus la queue est lourde). Notons que l'estimation de γ est un thème central dans l'analyse de ces distributions et un grand nombre d'estimateurs est donné dans la littérature notamment dans le cas particulier où $\boldsymbol{\gamma} > 0$ et qui nous intéresse dans ce chapitre. Leurs distributions asymptotiques ont été établies à partir d'une condition appelée « *condition du second ordre* ». Cette condition fait appel à un autre paramètre inconnu appelé « *paramètre du second ordre* » et est noté ρ . Dans la littérature statistique des valeurs extrêmes, le paramètre du second ordre ρ joue un rôle très important. Il est de première importance dans le choix adaptatif du nombre optimal de valeurs extrêmes utilisé lors de l'estimation de l'indice de queue $\boldsymbol{\gamma}$, voir (Beirlant *et al.*, 2002; Cheng et Peng, 2001; Gomes et Oliveira, 2001; Hall et Welsh, 1985). Notons également que ce paramètre contrôle le biais des estimateurs de l'indice des valeurs extrêmes et donc la connaissance de ρ est alors d'un grand intérêt dans la pratique. Son estimation peut être utilisée pour la réduction du biais des estimateurs de l'indice des valeurs extrêmes, voir par exemple (Caeiro *et al.*, 2005; Gomes et Martins, 2002; Gomes *et al.*, 2007). Pour les raisons mentionnées ci-dessus, l'estimation du paramètre du second ordre ρ a reçu beaucoup

d'attention dans la theorie des valeurs extrêmes voir par exemple (Hall et Welsh, 1985; Gomes et Martins, 2002; Fraga Alves *et al.*, 2003b; Ciuperca et Mercadier, 2010; Goegebeur *et al.*, 2010). Nous proposons donc une famille d'estimateurs semi-paramétrique de ρ permettant de regrouper les travaux précèdents au sein d'un formalisme commun. Nous établissons la consistance et la normalité asymptotique de cette famille d'estimateurs du paramètre du second ordre ainsi proposé. Nous montrerons également que les estimateurs proposés dans Hall et Welsh (1985), Gomes et Martins (2002), Fraga Alves *et al.* (2003b), Ciuperca et Mercadier (2010), Goegebeur *et al.* (2010) peuvent être vus comme des cas particuliers de notre approche. En conséquence, cette remarque permet d'établir leurs propriétés asymptotiques d'une manière simple et unifiée. Nous tirons également parti de notre formalisme pour illustrer un grand nombre d'estimateurs nouveaux de ρ asymptotiquement Gaussiens. Cette contribution a donné lieu à la publication (Deme *et al.*, 2012) et a fait l'objet de présentations à des séminaires et conférences:

- Conférence: 43ièmes Journées de Statistique organisées par la Société Française de Statistique, Tunis, 2011.
- Conférence: 7th Conference on Extreme Value Analysis (EVA), Probabilistic and Statistical Models and their Applications June 27th-July 1st, 2011, Lyon, France.
- Séminaire LERSTAD, 7 Août 2011, Université Gaston Berger de Saint-Louis, Sénégal
- Séminaire équipe MISTIS, Janvier 2012, Inria rhônes Alpes, Grenoble, France.
- Séminaire Institut de Recherche de Mathématiques Avancées (IRMA), 6 Mars 2012, Université de Strasbourg, France.

Toutefois, notons que la contribution présentée dans ce chapitre est plus générale par rapport à l'article cité précédemment.

Contribution du Chapitre 5

Dans ce chapitre, nous nous interessons aux mesures de risque actuariel particulièrement aux primes de risque pour des phénomènes de très grande ampleur. C'est-à-dire des phénomènes capables d'engendrer des pertes financières très importantes sur la solvabilité des compagnies de réassurance et sur l'évaluation des pertes possibles lors d'une accumulation de sinistres. La modélisation de ces événements par la distribution normale a été longtemps utilisée dans la gestion du risque. Cependant, les études empiriques concluent que les données financières et actuarielles présentent des déviations systématiques de la normalité et leurs distributions ont généralement une queue plus lourde que celle du modèle Gaussien.

De nombreuses mesures de risque ou de principes de calcul de la prime de risque ont été proposés dans la littérature actuarielle, voir par exemple Buhlmann (1980); Goovaerts *et al.* (1984); Wang (1995) et Wang (1996); Artzner *et al.* (1999); Denuit et Charpentier (2005). Nous nous concentrons sur la prime de risque-ajustée introduite par Wang (1996). Cette prime de risque a été largement étudiée dans la littérature. Un estimateur naturel de cette mesure de risque peut être obtenu en utilisant une approche basée sur la distribution empirique. Cet estimateur est une combinaison linéaire des éléments de la statistique d'ordre communément appelée L-statistique, voir par exemple Shorack et Wellner (1986). Les principales propriétés asymptotiques de ce type d'estimateur ont été établies par Jones et Zitikis (2003) sous certaines conditions appropriées, et qui ne sont pas remplies dans le cas des pertes à queues lourdes (pertes extrêmes). C'est ainsi que Necir et Meraghni (2009) ont donné une estimation alternative de la prime de risque-ajustée dans le cas des pertes extrêmes et ont étudié sa normalité asymptotique. Leurs études sont basées sur l'approche de la théorie des valeurs extrêmes. Cette méthode produit un biais potentiel dans l'estimation de la prime de risque. Ainsi, nous regardons ce cadre là et nous considérons une grande classe d'estimateurs de la prime de risque-ajustée pour des pertes à queues lourdes et nous proposons une approche de réduction de biais pour ces estimateurs. Nous établissons également leurs normalités asymptotiques. Ce travaail a donné suite à la contribution scientifique (Deme *et al.*, 2013).

Organisation de la thèse

La thèse s'articulera essentiellement autour de cinq chapitres qui sont rédigés de manière indépendante.

Ainsi, le Chapitre 1 présente l'état de l'art en théorie des valeurs extrêmes. On se limite au cas univarié réel. Au cours de ce chapitre, on présente tout d'abord le résultat principal de la théorie des valeurs extrêmes. Celui-ci montre qu'à l'exception de certaines lois pathologiques, on peut regrouper les lois usuelles en des groupes appelés *domaines d'attractions*. Nous exposons les critères pour qu'une loi appartienne à l'un de ces groupes. Nous présentons ensuite quelques estimateurs classiques de l'indice de queue γ ainsi que leurs propriétés asymptotiques. Et nous illustrons ces estimateurs par simulation. 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.

Nous étudions, au Chapitre 2, une classe d'estimateurs de l'indice de queue γ, dépendant d'un paramètre continu $\tau > 0$, appelé processus continu de Hill généralisé. Après avoir mené l'étude théorique de cette classe d'estimateurs, on s'attardera sur des simulations dont le but est d'apprécier la qualité des estimateurs obtenus. Comme illustration, nous proposons une application de notre méthodologie aux données hydrauliques.

Le chapitre 3 sera consacré entièrement à l'estimation fonctionnelle de l'indice de queue γ et est une généralisation du Chapitre 2. L'étude théorique qui y sera mené est basée sur théorie moderne de convergences vagues fonctionelles.

Dans le chapitre 4, après avoir présenté les contributions existantes, on définit une famille d'estimateurs du paramètre du second ordre. En dehors de l'étude théorique qui y sera menée, on illustrera notre approche d'estimation par simulation dont le but est d'apprécier la qualité de nos estimateurs. Pour illustrer la performance de nos estimateurs, nous proposons une application sur des données d'assurance.

Le chapitre 5 sera éntièrement consacré à l'estimation d'une mesure de risque actuariel dans le cas des pertes à queues lourdes, après avoir présenté les contributions existantes, on expose clairement le cadre d'étude dans lequel on va investiguer. Ensuite, nous donnons les principaux résultats sur des estimateurs proposés. En dehors de l'étude théorique qui y sera menée, on s'attardera sur des simulations dont le but est d'apprécier la qualité de nos estimateurs. Nous proposons également une application sur des données d'assurance incendie.

Cette thèse offre de nombreuses perspectives, tant sur le plan pratique que théorique. Une liste non exhaustive sera présentée en conclusion.

CHAPITRE 1

Présentation de la Théorie des Valeurs Extrêmes

Résumé

Le but de ce chapitre est de décrire et de présenter les principaux résultats classiques sur la théorie des valeurs extrêmes qui permettent de faciliter la lecture de la thèse et auxquels on fera appel dans les autres chapitres.

1.1 Introduction

La théorie des valeurs extrêmes a pour objectif l'étude du comportement asymptotique des grandes observations d'un échantillon de variables aléatoires indépendantes et identiquement distribuées. L'approche standard en théorie des probabilités place l'accent sur le comportement en moyenne et la variabilité autour de la moyenne, par le biais d'outils probabilistes comme par exemple la loi des grands nombres ou le théorème central limite. Le théorème fondamental de la théorie des valeurs extrêmes (Connu sous le nom de Théorème de Fisher-Tippett) donne quant à lui les lois limites possibles du maximum de l'échantillon et permet ainsi d'avoir une certaine connaissance sur la queue de distribution. L'utilisation des lois des valeurs extrêmes repose sur des propriétés de statistiques d'ordre et sur des méthodes d'extrapolation. Plus précisément, elle repose principalement sur les distributions limites des extrêmes et leurs domaines d'attraction.

1.2 Comportement asymptotique des extrêmes

Nous présentons ici une version concise de la théorie des valeurs extrêmes. Pour plus de détails et d'éventuelles démonstrations on pourra se référer à Embrechts *et al.* (1997). Soit $(X_n)_{n\geq 1}$ une suite de copies indépendantes d'une variable aléatoire X et de fonction de distribution $F(x) = \mathbb{P}(X_1 \leq x)$. Dans la suite on note $X_{1,n} \leq \ldots \leq X_{n,n}$ l'échantillon ordonné. Pour tout $i = 1, \ldots, n$, la variable aléatoire $X_{i,n}$ s'appelle la *ième* statistique d'ordre de l'échantillon. Le résultat central de la théorie des valeurs extrêmes concerne la distribution asymptotique $\mathcal{H}$ du maximum [1]. Comme les variables aléatoires sont indépendantes et identiquement distribuées, alors la fonction de répartition $F_{X_{n,n}}$ du maximum $X_{n,n}$ est donnée par

$$\begin{aligned} F_{X_{n,n}}(x) = \mathbb{P}(X_{n,n} \leq x) &= \mathbb{P}(X_1 \leq x, ..., X_n \leq x) \\ &= \mathbb{P}(X_1 \leq x)...\mathbb{P}(X_n \leq x) \\ &= [F(x)]^n . \end{aligned} \tag{1.1}$$

La fonction de répartition F étant souvent inconnue, il n'est généralement pas possible de déterminer la distribution du maximum à partir de ce résultat. Comme les valeurs extrêmes se trouvent, à droite et à la fin du support de la distribution, intuitivement le comportement asymptotique du maximum $X_{n,n}$ doit permettre de rendre compte de la fin de la distribution. Précisons que dans toute la suite, nous noterons par $x_F = \sup\{x \in \mathbb{R},\ F(x) < 1\}$, le *point terminal à droite* (*right-end point*) de la distribution F. Ce point terminal peut être infini ou fini (voir Embrechts *et al.* (1997), Exemple 3.3.22, p.139). On s'intéresse alors à la distribution asymptotique du maximum en faisant tendre n vers l'infini,

$$\lim_{n\to\infty} F_{X_{n,n}}(x) = \lim_{n\to\infty} [F(x)]^n = \begin{cases} 1 & \text{si} \quad x \geq x_F \\ 0 & \text{si} \quad x < x_F. \end{cases} \tag{1.2}$$

On constate que la distribution asymptotique du maximum donne une loi dégénérée, une masse de Dirac en x_F, puisque pour certaines valeurs de x, la probabilité peut être égale à 1 dans le cas où x_F est fini et donc $X_{n,n}$ tend vers x_F presque sûrement, avec $x_F \leq \infty$. Ce fait ne fournit pas assez d'informations, d'où l'idée d'utiliser une transformation afin d'obtenir des résultats plus exploitables pour les lois limites des maxima. On s'intéresse par conséquent à une loi non dégénérée pour le maximum, la théorie des valeurs extrêmes permet de donner une réponse à cette problématique. Les premiers résultats sur la caractérisation du comportement asymptotique des maxima $X_{n,n}$ convenablement normalisés

[1] En utilisant la correspondance suivante $\min(X_1, \ldots, X_n) = -\max(-X_1, \ldots, -X_n)$, tous les résultats que nous allons présenter pour les maxima pourront être transposés pour les minima.

ont été obtenus par Fisher et Tippet (1928). Le théorème ci-dessous explicite ces résultats.

Définition 1.2.1 *: Soient F_1 et F_2 deux fonctions de distribution. On dit que F_1 et F_2 sont du même type si et seulement si, il existe $a \in \mathbb{R}_+^*$ et $b \in \mathbb{R}$ tels que $F_1(ax+b) = F_2(x)$.*

Théorème 1.2.2 *(Fisher et Tippet (1928); Gnedenko (1943))*
Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) avec $\mathbb{P}(X_1 \leq x) = F(x)$. S'il existe deux suites normalisantes réelles $(a_n > 0, n \geq 1)$ et $(b_n \in \mathbb{R}, n \geq 1)$ et une loi non-dégénérée $\mathcal{H}$ telle que:

$$\begin{aligned} \lim_{n\to\infty} \mathbb{P}(X_{n,n} \leq a_n x + b_n) &= \lim_{n\to\infty} F^n(a_n x + b_n) \\ &= \mathcal{H}(x), \ \forall\, x \in \mathbb{R}, \end{aligned} \tag{1.3}$$

alors $\mathcal{H}$ est du même type qu'une des trois lois suivantes:

$$\begin{aligned} &\boldsymbol{\phi}_\alpha(x) = \exp\left(-(x)^{-\alpha}\right) \mathbb{1}_{(x\geq 0)}, \ \alpha > 0 && (\textit{loi de Fréchet}), \\ &\boldsymbol{\Lambda}(x) = \exp\left(-e^{-x}\right), && (\textit{loi de Gumbel}), \\ &\boldsymbol{\psi}_\alpha(x) = \exp\left(-(-x)^{\alpha}\right) \mathbb{1}_{(x<0)} + \mathbb{1}_{(x\geq 0)}, \ \alpha > 0, && (\textit{loi de Weibull}), \end{aligned}$$

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

Une démonstration détaillée de ce thèoréme est donnée dans Resnick (1987) et avec des développements dans Embrechts *et al.* (1997), p. 152.
Le théorème établit que le comportement limite du maximum normalisé est décrit par l'une ou l'autre de ces trois distributions quelle que soit la loi F de la variable aléatoire originelle X. Il présente ainsi dans l'esprit des similarités avec le Théorème Central Limite (TCL) qui définit quant à lui le comportement limite de la somme (ou de la moyenne) d'une suite de variables aléatoires indépendantes et identiquement distribuées.

Les suites $(a_n > 0, n \geq 1)$ et $(b_n \in \mathbb{R}, n \geq 1)$ dépendent des paramètres de la loi de X. La Figure 1.1 illustre dans le cas d'une loi normale centrée réduite, la convergence de la suite de variables aléatoires $\left(a_n^{-1}\left(X_{n,n} - b_n\right)\right)_{n\geq 1}$ en loi vers une limite non dégénérée $\boldsymbol{\Lambda}$. Nous avons utilisé ici les suites de renormalisation théoriques associées à la loi normale standard dans Embrechts *et al.* (1997), page 145 à savoir :

$$a_n = (2\log n)^{-1/2} \ \text{ et } \ b_n = (2\log n)^{1/2} - \frac{\log\log n + \log 4\pi}{2\,(2\log n)^{1/2}}.$$

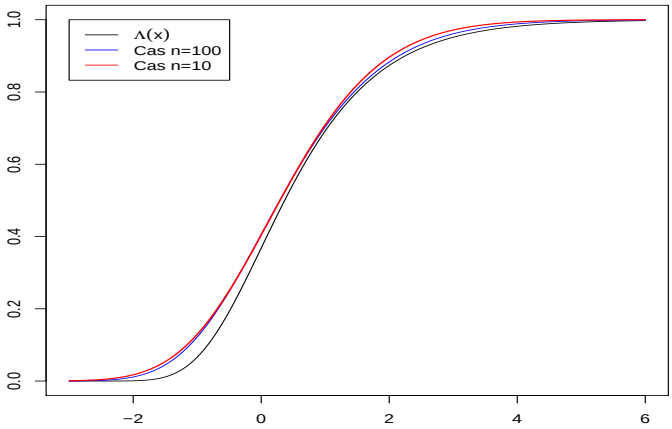

FIG. 1.1 – Illustration de la théorie des valeurs extrêmes sur une loi normale centrée réduite. Comparaison entre $\mathbf{\Lambda}(x)$ (noire), $\mathbb{P}\left[\frac{X_{n,n}-b_n}{a_n} \leq x\right]$ avec $n = 100$ (bleu) et $\mathbb{P}\left[\frac{X_{n,n}-b_n}{a_n} \leq x\right]$ avec $n = 10$ (rouge).

On dira que F appartient au domaine d'attraction de $\mathcal{H}$, et on note $F \in D(\mathcal{H})$ si la condition (1.3) est vérifiée. Notons que (1.3) peut aussi se réécrire sous la forme suivante

$$\lim_{n\to\infty} n\left(\overline{F}(a_n x + b_n)\right) = -\log\ \mathcal{H}(x), \quad x \in \mathbb{R}. \tag{1.4}$$

Lorsque $\mathcal{H}(x) = 0$, le terme de droite peut être interprété comme étant égal à ∞. En fonction de la loi limite $\boldsymbol{\phi}_\alpha$, $\mathbf{\Lambda}$ ou $\boldsymbol{\psi}_\alpha$, on dira que la distribution F de la variable aléatoire X appartient au domaine d'attraction de Fréchet ($D(\boldsymbol{\phi}_\alpha)$), de Gumbel ($D(\mathbf{\Lambda})$) ou de Weibull ($D(\boldsymbol{\psi}_\alpha)$).
Notons par $\mathcal{DA}$, la classe de tous les domaines d'attraction (l'ensemble de toutes les fonctions de distribution attirées par la fonction $\mathcal{H}$) on a $\mathcal{DA} = \cup_{\alpha>0} D(\boldsymbol{\phi}_\alpha) \cup D(\boldsymbol{\psi}_\alpha) \cup D(\mathbf{\Lambda})$.
D'après la Definition 1.2.1, on mentionne que cette fonction $\mathcal{H}$ représente une classe d'équivalence de la relation binaire $\mathcal{R}$ sur l'ensemble $\mathcal{DA}$:

$$\forall\ (\mathcal{H}_1, \mathcal{H}_2) \in \mathcal{DA}^2,\ \mathcal{H}_1\ \mathcal{R}\ \mathcal{H}_2 \Leftrightarrow \exists (a,b) \in \mathbb{R}_+^* \times \mathbb{R},\ \forall x \in \mathbb{R},\ \mathcal{H}_1(x) = \mathcal{H}_2(ax+b).$$

On peut voir facilement que si $F^n(a_n x + b_n) \to \mathcal{H}_1(x)$, alors $F^n(C_n x + d_n) \to \mathcal{H}_2(ax+b) = \mathcal{H}_1(x)$ lorsque

$$a_n/b_n \to a \text{ et } (b_n - d_n)/c_n \to b, \text{ quand } n \to \infty. \tag{1.5}$$

Dans ce cas, on peut prendre les suites de nombres réels $(a_n > 0,\ n \geq 1)$ et $(b_n,\ n \geq 1)$ tels que pour $\alpha > 0$, $a = \gamma = 1/\alpha$ et $b = 1$ (dans le cas de Fréchet), $a = -\gamma = -1/\alpha$ et $b = -1$ (dans le cas de Weibull). Finalement on peut interpréter $(1+\gamma x)^{-1/\gamma} = e^{-x}$ pour $\gamma = 0$ (dans le cas de Gumbel). Ceci conduit à une paramétrisation des trois comportements limites précédents en une distribution unique dite famille des lois de valeurs extrêmes généralisées (Generalized Extreme Value GEV), (on pourra se référer aux travaux de von Mises (1936) et Jenkinson (1955)) :

$$\mathcal{H}(x) = \mathcal{H}_\gamma(x) = \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right) & \forall x \in \mathbb{R},\ 1+\gamma x > 0, \quad \text{si } \gamma \neq 0 \\ \exp\left(-\exp\left(-x\right)\right) & \forall x \in \mathbb{R}, \qquad\qquad\qquad \text{si } \gamma = 0. \end{cases} \tag{1.6}$$

Ce résultat implique de façon évidente que le comportement de la queue de distribution dépend d'un unique paramètre, noté γ [2], et appelé *indice de queue* ou *indice des valeurs extrêmes*. Le signe de ce paramètre est un indicateur essentiel sur la forme de la queue:

• Le cas $\gamma > 0$ où $1 - \mathcal{H}_\gamma$ est à support borné à gauche et à queue lourde:

$$1 - \mathcal{H}_\gamma(x) = (\gamma x)^{-1/\gamma}(1 + o(1));$$

• Le cas $\gamma = 0$ où $1 - \mathcal{H}_\gamma$ est à décroissance exponentielle;

• Le cas $\gamma < 0$ où $1 - \mathcal{H}_\gamma$ est à support borné à droite. La Figure 1.2 illustre le comportement de différentes densités de la loi GEV pour certaines valeurs de γ.

On peut écrire $\mathcal{H}_\gamma(x)$ sous une forme plus générale dans la quelle on fait apparaître un paramètre de localisation [3] μ et un paramètre d'échelle [4] σ (Weissman, 1978), qui dépendent des suites (a_n) et (b_n). En introduisant les paramètres de localisation μ et de dispersion σ dans la paramétrisation (1.6), on obtient la forme la plus générale de la GEV

$$\mathcal{H}_{\gamma,\mu,\sigma}(x) = \begin{cases} \exp\left(-\left(1+\gamma(\frac{x-\mu}{\sigma})\right)^{-\frac{1}{\gamma}}\right) & \text{si} \quad \gamma \neq 0,\ 1+\gamma\frac{x-\mu}{\sigma} > 0, \\ \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) & \text{si} \quad \gamma = 0,\ \forall x \in \mathbb{R}. \end{cases} \tag{1.7}$$

[2] La connaissance de γ permet à elle seule, de caractériser à un changement d'échelle près, le comportement asymptotique du maximum normalisé.

[3] Un paramètre de localisation (ou de position) est comme son nom l'indique, un paramètre qui régit la position d'une densité de probabilité f. Si ce paramètre (scalaire ou vectoriel) est noté μ, la densité de X présente formellement comme: $f_\mu(x) = f(x-\mu)$.

[4] Un paramètre d'échelle (ou dispersion) est un paramètre qui régit l'applatissement d'une famille paramétrique de lois de probabilités. Il dirige l'échelle ou encore la dispersion de la distribution. La densité de X présente comme suit $f_\sigma(x) = \sigma^{-1} f(x/\sigma)$.

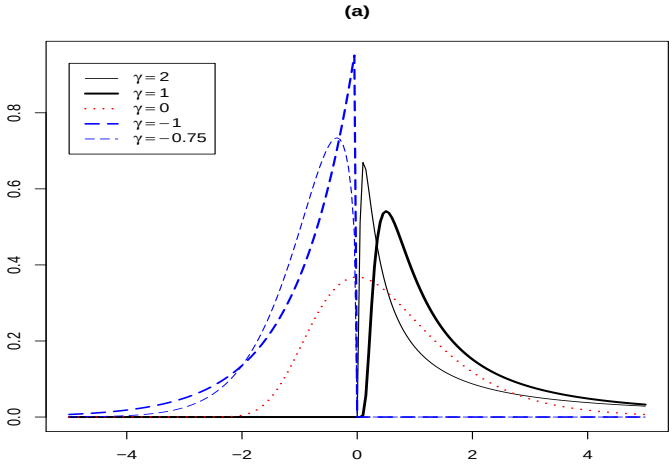

FIG. 1.2 – Densités de la distribution des valeurs extrêmes pour différentes valeurs de γ

1.3 Distribution conditionnelle des excès

L'idée d'utiliser un nombre croissant de statistiques d'ordre de l'échantillon a ensuite été plus largement développée dans le cadre de l'approche « *Peaks Over Threshold* »(POT), via l'approximation de la loi des excès au-delà d'un seuil par des GPD (Generalized Pareto Distributions). Plus précisément, soit $u < x_F$ et F_u la fonction de répartition des excès définie par

$$F_u(y) = \mathbb{P}(Y \leq y | X > u) = \mathbb{P}(X - u \leq y | X > u) = \frac{F(u+y) - F(u)}{\overline{F}(u)}.$$

Ce qui équivaut à $F_u(x) = \mathbb{P}(X \leq x | X > u) = \dfrac{F(x) - F(u)}{\overline{F}(u)}$ pour $x \geq u$. F_u déccrit la loi de X sachant $\{X > u\}$. On a alors le résultat suivant dû à Balkema et de Haan (1974) et Pickands (1975).

Théorème 1.3.1 *Si F appartient à l'un des trois domaines d'attraction de la loi des valeurs extrêmes (Fréchet, Gumbel ou Weibull), alors il existe une fonction $\sigma(u)$ strictement positive et un $\gamma \in \mathbb{R}$ tels que*

$$\lim_{u \uparrow x_F} \sup_{0 \leq y \leq x_F - u} |F_u(y) - \mathcal{G}_{\gamma,\sigma(u)}(y)| = 0 \tag{1.8}$$

où $\mathcal{G}_{\gamma,\sigma}$ est la fonction de répartition de la loi de Paréto Généralisée définie par

$$\mathcal{G}_{\gamma,\sigma}(y) = \begin{cases} 1 - (1 + \gamma y/\sigma)^{-1/\gamma} & si \quad \gamma \neq 0, \ \sigma > 0, \\ 1 - \exp(-y/\sigma) & si \quad \gamma = 0, \ \sigma > 0, \end{cases} \tag{1.9}$$

pour $y \in [0, (x_F - u)]$ *si* $\boldsymbol{\gamma} \geq 0$ *et* $y \in [0, \min(-\sigma/\boldsymbol{\gamma}, x_F - u)]$ *si* $\boldsymbol{\gamma} < 0$.

Ainsi, pour une valeur de u assez élevée, la loi des excès est approchée par une loi Paréto généralisée: $F_u \approx \mathcal{G}_{\boldsymbol{\gamma},\sigma(u)}$. Ce théorème établit le lien entre le paramètre de la loi du domaine d'attraction maximum et le comportement limite des excès au-delà d'un seuil assez grand [5]. En particulier, les paramètres $\boldsymbol{\gamma}$ et σ de la loi de Paréto généralisée sont les mêmes que ceux donnés dans l'étude du maximum. Il est intéressant de noter que le cas $\boldsymbol{\gamma} = 0$ correspond à la distribution exponentielle de moyenne σ et que le cas $\boldsymbol{\gamma} = 1$ correspond à la loi uniforme sur $[0, \sigma]$.

1.4 Caractérisation des Domaines d'attraction

Un problème important revient à définir les conditions (nécessaires et suffisantes) d'appartenance d'une distribution à un domaine d'attraction. La recherche de ce domaine d'attraction peut être considérée comme l'étude réciproque de la recherche de la distribution des valeurs extrêmes associée éventuellement à une distribution. Ceci consiste donc à répondre à la question suivante: étant donnée une loi $\mathcal{H}$ de type extrême (donc appartenant à l'une des trois familles Fréchet, Gumbel et Weibull) quels sont les critères à vérifier pour que la loi du maximum de la suite de variables aléatoires i.i.d. de loi F converge vers $\mathcal{H}$? Différentes caractérisations des trois domaines d'attraction de Fréchet, Gumbel et de Weibull ont été proposées dans de Haan (1985); Resnick (1987); Embrechts *et al.* (1997); de Haan et Ferreira (2006). Ces caractérisations font appel aux classes de fonctions à variation régulière. Dans la suite, on note $g^{\leftarrow}$, l'inverse généralisée d'une fonction monotonne g et définie par $g^{\leftarrow}(s) = \inf\{x \in \mathbb{R};\ g(x) \geq s\}$, $0 < s \leq 1$.

1.4.1 Fonctions à variation régulière

Le concept de la variation régulière est largement utilisé dans la théorie des valeurs extrêmes. Nous résumons ici quelques principaux résultats (définitions, extensions et propriétés) de la théorie de la variation régulière qui sont pertinents à notre portée, pour plus de details on pourra se référer à Bingham *et al.* (1987).

[5] Ceci permet de distinguer les lois à queue épaisse qui appartiennent au domaine d'attraction de Fréchet des lois à queue fine ou légère qui appartiennent au domaine d'attraction de Gumbel. Une queue de distribution est d'autant plus épaisse qu'elle se distingue de la loi normale en s'étalant plus lentement, et donc avec un coefficient d'aplatissement (kurtosis) supérieur à celui de la loi normale (égal à trois).

Définition 1.4.1 *Une fonction $g : \mathbb{R}_+ \to \mathbb{R}_+$ mesurable est à variation régulière à l'infini si et seulement si, il existe un réel α tel que pour tout $x > 0$*

$$\lim_{t\to\infty} \frac{g(tx)}{g(t)} = x^\alpha.$$

Et on note $g \in \mathcal{RV}_\alpha$, α est appelé indice (ou exposant) de la fonction à variation régulière g.

Remarque 1.4.2 *Dans le cas particulier où $\alpha = 0$, on dit que g est à variation lente à l'infini, c'est à dire*

$$\lim_{t\to\infty} \frac{g(tx)}{g(t)} = 1 \quad \forall x > 0.$$

Les fonctions à variation lente sont génériquement notées $\ell(x)$.

Conséquence 1.4.3 *Une fonction $g : \mathbb{R}_+ \to \mathbb{R}_+$ mesurable est à variation régulière d'indice α au voisinage de 0 notée ($g \in \mathcal{RV}_\alpha^0$), si pour tout $x > 0$*

$$\lim_{t\to 0} \frac{g(tx)}{g(t)} = x^\alpha.$$

c'est à dire $g(1/x)$ est à variation régulière d'indice $-\alpha$ à l'infini.

Lemme 1.4.4 *Inverse d'une fonction à variation régulière*
- *Si g est à variation régulière d'indice $\alpha > 0$ à l'infini, alors $g^{\leftarrow}(x)$ est à variation régulière d'indice $1/\alpha > 0$ à l'infini.*
- *Si g est à variation régulière d'indice $\alpha < 0$ à l' infini, alors $g^{\leftarrow}(1/x)$ est à variation régulière d'indice $-1/\alpha > 0$ à l'infini.*

Pour la preuve du Lemme 1.4.4, on pourra se référer à Bingham *et al.* (1987), Théorème 1.5.12 ou dans Resnick (1987). Il est clair que si ℓ est une fonction à variation lente et si $\alpha \in \mathbb{R}$, alors la fonction g définie par $g(x) = x^\alpha \ell(x)$ pour tout $x > 0$ appartient à $\mathcal{RV}_\alpha$.

Proposition 1.4.5 *Soient $\alpha \in \mathbb{R}$ et $g \in \mathcal{RV}_\alpha$. Alors il existe une fonction à variation lente ℓ à l'infini telle que:*

$$\forall\ x > 0,\ g(x) = x^\alpha \ell(x). \tag{1.10}$$

Ce résultat montre que l'étude des fonctions à variation régulière à l'infini se ramène à celle des fonctions à variation lente. Parmi les fonctions à variation lente, on peut citer:
- les fonctions possédant une limite strictement positive à l'infini;
- les fonctions de la forme $g : x \mapsto |\log x|^\beta$, $\beta \in \mathbb{R}$;

- les fonctions g telles que

$$\exists\ M > 0, \quad \forall x \geq M, \quad g(x) = c + dx^{-\beta}(1 + o(1)) \tag{1.11}$$

où $c, \beta > 0$ et $d \in \mathbb{R}$. L'ensemble de ces fonctions g est appelé classe de *Hall* .

Théorème 1.4.6 *Représentation de Karamata (Resnick, 1987)*
Toute fonction à variation lente ℓ à l'infini s'écrit sous la forme

$$\ell(x) = c(x) \exp\left(\int_1^x \Delta(t)t^{-1}\ dt\right), \tag{1.12}$$

où $c(\cdot) > 0$ et $\Delta(\cdot)$ sont deux fonctions mesurables telles que

$$\lim_{x\to\infty} c(x) = c_0 \in]0, \infty[\ \textit{et} \ \lim_{x\to\infty} \Delta(x) = 0.$$

Si la fonction $c(\cdot)$ est une constante, alors on dit que ℓ est *normalisée*. L'équation (1.12) implique que si ℓ est normalisée alors ℓ est dérivable de dérivée ℓ' avec pour tout $x > 0$, $\ell'(x) = \Delta(x)\ell(x)/x$. En particulier, on a $\lim_{x\to\infty} x\ell'(x)/\ell(x) = 0$.

Proposition 1.4.7 *Pour toute fonction à variation lente ℓ à l'infini, on a:*

$$\lim_{x\to\infty} \frac{\log(\ell(x))}{\log(x)} = 0. \tag{1.13}$$

Pour plus d'information sur la théorie des fonctions à variation régulière, on renvoie le lecteur à de Haan (1985), Bingham *et al.* (1987) and Lo (1986).

1.4.2 Domaine d'attraction de Fréchet

Précisons que dans toute la suite, nous noterons par $\overline{F}(\cdot) = 1 - F(\cdot)$, la fonction de survie de F, $Q(\cdot)$, l'inverse généralisée de F; définie par $Q(s) = F^{\leftarrow}(s) = \inf\{x \in \mathbb{R},\ F(x) \geq s\}$, $0 < s \leq 1$. $Q(\cdot)$ est aussi appelée fonction de *quantile* associée à la distribution F. On définit également la fonction de *quantile de queue* par $U(x) = Q(1 - 1/x)$, $x > 1$. Le résultat ci-dessous énoncé dans Gnedenko (1943) et dont on trouvera une démonstration dans Resnick (1987), Proposition 1.13, assure que toute fonction appartenant au domaine d'attraction de Fréchet est une fonction à variation régulière.

Théorème 1.4.8 *La fonction de distribution F ayant pour point terminal x_F appartient au domaine d'attraction de Fréchet $D(\mathcal{H}_\gamma)$, $\gamma > 0$ si et seulement si $x_F = +\infty$ et $\overline{F}$ est à variation régulière d'indice $-1/\gamma$ à l'infini ie.*

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(x)} = x^{-1/\gamma}. \tag{1.14}$$

Les suites de normalisation a_n et b_n sont données dans ce cas par:

$$a_n = F^{\leftarrow}(1 - 1/n) = U(n) \text{ et } b_n = 0, \ \forall n > 0.$$

Remarque 1.4.9 *On déduit de la Proposition 1.4.5 que F appartient au domaine $D(\mathcal{H}_\gamma)$, $\gamma > 0$ si et seulement si $x_F = \infty$ et $\overline{F}(x) = x^{-1/\gamma}\ell_F(x)$ où ℓ_F est une fonction à variation lente à l'infini.*

Remarque 1.4.10 *Il est facile de voir que pour tout $s \in (0,1)$, $Q(1-s) = \overline{F}(1/s)$, et dans ce cas en utilisant la conséquence 1.4.3, l'équation (1.14) est équivalente à $Q(1-s)$ est à variation régulière d'indice $-\gamma$ en 0 et on note $Q(1-\cdot) \in \mathcal{RV}^0_{-\gamma}$ (C'est à dire $Q(1-s) = s^{-\gamma}\ell(1/s)$, avec $\ell \in \mathcal{RV}_0$). Et donc la fonction de queue $U(\cdot)$ est à variation régulière d'indice γ à l'infini ($U \in \mathcal{RV}_\gamma$).*

Remarque 1.4.11 *D'après (1.14), $\overline{F} \in \mathcal{RV}_{-1/\gamma}$, $\gamma > 0$. Ainsi (1.10) et (1.12) permettent d'avoir la représentation suivante*

$$\overline{F}(x) = c(x)x^{-1/\gamma}\exp\left(\int_1^x \Delta(t)t^{-1}dt\right), \ x < x_F \tag{1.15}$$

avec $\lim_{x\to\infty} c(x) = c_0 \in]0,\infty[$ et $\lim_{x\to\infty}\Delta(x) = 0$.

1.4.3 Domaine d'attraction de Weibull

Le résultat suivant (voir Gnedenko (1943) ou Resnick (1987), Proposition 1.13), montre que l'on passe du domaine d'attraction de Fréchet à celui de Weibull par un simple changement de variable dans la fonction de répartition.

Théorème 1.4.12 *Une fonction de distribution F avec un point terminal x_F appartient au domaine d'attraction de Weibull $D(\mathcal{H}_\gamma)$, $\gamma < 0$ si et seulement si x_F est fini et la fonction de répartition F^* définie par*

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

appartient au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes $-\gamma > 0$, c'est-à-dire que $\overline{F}^*$ est une fonction à variation régulière d'indice $1/\gamma$ à l'infini et on note ($\overline{F}^* \in \mathcal{RV}_{1/\gamma}$). Dans ce cas un choix possible pour les suites a_n et b_n est:

$$a_n = x_F - F^{\leftarrow}(1 - 1/n) \text{ et } b_n = x_F.$$

Pour la preuve du Théorème 1.4.12, le lecteur pourra se référer à Resnick (1987), Proposition 1.13 ou à Embrechts *et al.* (1997), Théorème 3.3.12.

Remarque 1.4.13 *On déduit de ce Théorème 1.4.12 que F appartient au domaine de Weibull $D(\mathcal{H}_\gamma)$, $\gamma < 0$ si et seulement si $x_F = \infty$ et $\overline{F}(x) = (x_F - x)^{-1/\gamma}\ell((x_F - x)^{-1})$ où $\ell \in \mathcal{RV}_0$. De manière équivalente, la fonction quantile $Q(1-s)$ s'écrit:*

$$Q(1-s) = x_F - s^{-\gamma}\ell(1/s), \quad \ell \in \mathcal{RV}_0 \tag{1.16}$$

1.4.4 Domaine d'attraction de Gumbel

La caractérisation des fonctions de répartition du domaine de Gumbel est plus complexe. Le résultat ci-dessous qui est démontré notamment dans Resnick (1987), Proposition 1.4, permet de décrire que $F \in D(\mathcal{H}_\gamma)$, $\gamma = 0$.

Théorème 1.4.14 *La fonction de distribution F avec un point terminal x_F appartient au domaine d'attraction de Gumbel $D(\mathcal{H}_\gamma)$, $\gamma = 0$ si et seulement si il existe un réel z, $z < x_F \leq \infty$ tel que*

$$\overline{F}(x) = c(x)\exp\left(-\int_z^x \frac{g(t)}{a(t)}dt\right), \ z < x < x_F, \tag{1.17}$$

où $c(\cdot)$ et $g(\cdot)$ sont des fonctions mesurables positives satisfaisant $c(x) \to c > 0$, $g(x) \to 1$ lorsque $x \uparrow x_F$, et $a(\cdot)$ est une function positive absolument continue (par rapport à la mesure de Lebesgue), de densité $a'(\cdot)$ avec $\lim_{x\uparrow x_F} a'(x) = 0$. Dans ce cas, un choix possible pour les suites a_n et b_n est $a_n = F^{\leftarrow}(1 - 1/n) = U(n)$ et $b_n = a(a_n)$. Un choix possible de la fonction $a(\cdot)$ est

$$a(x) = \int_x^{x_F} \frac{\overline{F}(t)}{\overline{F}(x)}dt, \ x < x_F. \tag{1.18}$$

La fonction $a(\cdot)$ est appelée fonction auxiliaire.

Remarque 1.4.15 *La représentation (1.17) n'est pas unique car elle dépend des fonctions $c(\cdot)$ et $g(\cdot)$. Et celle donnée avec une fonction $g = 1$ peut être alternativement employée (voir Resnick (1987), Proposition 1.4).*

1.4.5 Caractérisation générale

Dans la sous section précédente, nous avons présenté séparément les propriétés de caractérisation pour les trois domaines d'attraction. Après l'unification, une propriété « globale » de caractérisation peut être développée et qui comprend toutes les trois lois limites

pour les maxima. Notons également que le domaine d'attraction de $\mathcal{H}_\gamma$ est construit de la même manière que le domaine d'attraction des trois distributions des valeurs extrêmes. Les résultats ci-dessous donnent une caractérisation unifiée des trois domaines d'attraction précédents, voir par exemple de Haan et Ferreira (2006).

Théorème 1.4.16 *Soit $\gamma \in \mathbb{R}$, les assertions suivantes sont équivalentes*

1. *F appartient au domaine d'attraction de $\mathcal{H}_\gamma$.*
2. *Il existe une fonction positive $a(\cdot)$ telle que pour tout $x > 0$,*

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = D_\gamma(x) = \begin{cases} \dfrac{x^\gamma - 1}{\gamma} & si \;\; \gamma \neq 0, \\ \log x & si \;\; \gamma = 0. \end{cases} \tag{1.19}$$

3. *Il existe une fonction positive $f(\cdot)$ telle que pour tout x vérifiant $1 + \gamma x > 0$,*

$$\lim_{t\uparrow x_F} \frac{\overline{F}(t + xf(t))}{\overline{F}(t)} = \begin{cases} (1+\gamma x)^{-1/\gamma} & si \;\; \gamma \neq 0, \\ e^{-x} & si \;\; \gamma = 0. \end{cases} \tag{1.20}$$

Un choix possible de la fonction de $f(\cdot)$ est $f(t) = a(1/(\overline{F}(t)))$ où la fonction $a(\cdot)$ est celle utilisée dans (1.19).

En supposant que $U(\infty) > 0$, la condition (1.19) implique que

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)/U(t)} = \begin{cases} \log x & \text{si} \;\; \gamma \geq 0, \\ \dfrac{x^\gamma - 1}{\gamma} & \text{si} \;\; \gamma < 0. \end{cases} \tag{1.21}$$

Dans Dekkers *et al.* (1987), le résultat (1.21) a été utilisé pour proposer l'estimateur des moments du paramètre γ dans la sous-section 1.5.2.2.

Proposition 1.4.17 *Soit $\gamma \in \mathbb{R}$, $F \in D(\mathcal{H}_\gamma)$ si et seulement si pour tous $x > 0$, $y > 0$, $y \neq 1$,*

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{U(ty) - U(t)} = \begin{cases} \dfrac{x^\gamma - 1}{y^\gamma - 1} & si \;\; \gamma \neq 0, \\ \dfrac{\log x}{\log y} & si \;\; \gamma = 0. \end{cases} \tag{1.22}$$

La preuve sur les résultats de caractérisation générale peut être retrouvée dans Embrechts *et al.* (1997). La Proposition 1.4.17 est utilisée pour construire l'estimateur de Pickands de γ au paragraphe 1.5.2.1. Nous regroupons dans le tableau 1.1 quelques lois et leur domaine d'attraction.

Domaines d'attraction	**Gumbel** $\gamma = 0$	**Fréchet** $\gamma > 0$	**Weibull** $\gamma < 0$
Lois	Normale Exponentielle Lognormale Gamma Weibull	Cauchy Pareto Student Burr Loggamma	Uniforme Beta

TAB. 1.1 – Quelques exemples de lois classées selon leur domaine d'attraction.

1.5 Estimation des paramètres

Dans cette section nous nous intéressons aux différentes méthodes d'estimation du paramètre $\boldsymbol{\gamma}$ ou $\boldsymbol{\sigma}$ qui intervienent dans la distribution asymptotique des valeurs extrêmes (1.7). On trouve dans la littérature de la théorie des valeurs extrêmes plusieurs méthodes d'estimation. Parmi ces méthodes, on peut citer l'estimateur de Pickands (1975), l'estimateur de Hill (1975), l'estimateur des moments Dekkers *et al.* (1987), l'estimateur du rapport des moments Danielson *et al.* (1996), l'estimateur de Peng (1998), l'estimateur basé sur le QQ-plot, l'estimateur basé sur le graphique de la moyenne des excès Beirlant *et al.* (1996), l'estimateur construit par des méthodes de régression Beirlant *et al.* (2002) (optimisation avec contrainte (Csörgő et Mason, 1985) ou sans contrainte (Schultze et Steinebach, 1996; Kratz et Resnick (1996)), les estimateurs basés sur la méthodes des moments et des moments pondérés Hosking et Wallis (1987) et l'estimateur du maximum de vraisemblance. D'autres estimateurs sont également donnés : Lo (1986), Embrechts *et al.* (1997), Beirlant *et al.* (2004), de Haan et Ferreira (2006), Diop et Lo (2006).

On pourrait également ajouter les estimateurs avec correction du biais basés sur des méthodes de régression (Beirlant *et al.*, 1999; Beirlant *et al.*, 2002; Beirlant *et al.*, 2005; Diebolt *et al.*, 2008c) ou sur des méthodes de Bootstrap (Gomes et Oliveira, 2001) ou jackknife Gomes, 1999; Gomes *et al.*, 2005b) et les estimateurs à poids ou combinaisons linéaires (Viharos, 1995; Gomes *et al.*, 2005a) qui englobent d'une part les estimateurs construits par des méthodes d'optimisation avec et sans contrainte et d'autre part les estimateurs construits par des méthodes de régression. D'un point de vue théorique, toutes ces méthodes partagent les mêmes propriétés de consistance et

de normalité asymptotique. Cependant, les simulations montrent qu'il y a de grandes différences entre ces estimateurs. En général, il n'y a pas une meilleure méthode dans toutes les situations. Les méthodes les plus utilisées sont celles de Hill, Pickands et des moments. Certaines études de comparaison (théorique et par simulation) entre les différentes méthodes peuvent être trouvées dans Rosen et Weissman (1996), Peng (1998), de Haan et Peng (1998), Groeneboom *et al.* (2003), etc.
Tsourti et Panaretos (2001) et Tsourti et Panaretos (2003) pensent que la performance d'une méthode dépend de la distribution de la série étudiée. En d'autres termes, elle dépend de la vraie valeur de l'indice de queue. Ils recommandent l'utilisation de techniques pour déterminer le domaine d'attraction de la loi des valeurs extrêmes, et donc l'intervalle le plus probable pour l'indice des valeurs extrêmes. Les méthodes les plus utilisées à cette fin sont graphiques : le graphique log log, la moyenne empirique des excès, le graphique des rapports du maximum et de la somme, le graphique du rapport de Hill, la statistique de Jackson. Pour de plus amples explications sur ces méthodes voir El-Adlouni et al. (2007).

Dans les paragraphes ci-dessous, nous exposerons quelques estimateurs cités ci-dessus et que nous aurons à utiliser dans les autres chapitres.

1.5.1 Estimation de l'indice de queue sous le modèle de type Paréto

Le domaine de Fréchet ($\gamma > 0$) a été le plus largement étudié dans la littérature dans la mesure où il englobe un grand nombre d'applications pratiques. Les distributions F dans ce domaine sont dites de type Paréto et ont la propriété suivante:

$$\overline{F}(x) = x^{-1/\gamma}\ell_F(x) \text{ avec } \gamma > 0, \tag{1.23}$$

ou de manière équivalente la fonction de quantile de queue $U \in \mathcal{RV}_\gamma$ i.e

$$U(x) = Q(1 - 1/x) = x^\gamma \ell_U(x), \tag{1.24}$$

ℓ_F et ℓ_U sont deux fonctions à variation lente à l'infini. En utilisant la Proposition 1.4.7, nous remarquons que lorsque $x \to \infty$

$$\log U(x) = \gamma \log x + \log \ell_U(x) = \gamma \log x \left(1 + \frac{\log \ell_U(x)}{\gamma \log x}\right) \sim \gamma \log(x). \tag{1.25}$$

Un estimateur naturel de F est sa fonction de répartition empirique $F_n(x)$ définie par :

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\{X_j \leq x\}}.$$

Le quantile empirique associé à F est donné par :

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

qui est un estimateur de la fonction de quantile $Q(\cdot)$. Notons que $Q_n(\cdot)$ peut être exprimée comme étant une fonction de la statistique d'ordre associée à l'echantillon $(X_1, ..., X_n)$:

$$Q_n(s) = X_{n-j+1,n} \quad \text{si} \quad (n-j)/n < s \leq (n-j+1)/n,\ j = 1, ..., n \text{ et } Q_n(0) = X_{1,n}.$$

En remplaçant dans (1.25), la fonction de queue $U(\cdot)$ par sa version empirique

$$\widehat{U}_n(x) = Q_n\left(1 - 1/x\right) = X_{n-[n/x]+1,n}, \tag{1.26}$$

nous obtenons

$$\log \widehat{U}_n(x) = \log X_{n-[n/x]+1,n} \sim \boldsymbol{\gamma} \log x,\ x \to \infty. \tag{1.27}$$

Soit k_n, $n \geq 1$, une suite d'entiers correspondant au nombre de statistiques d'ordre supérieur (« nombre d'extrêmes ») utilisé tel que $1 \leq k_n < n$, $k_n \to \infty$ et $k_n/n \to 0$ quand $n \to \infty$. En prenant $x = (n+1)/j$, $j = 1, ..., k_n$ et en considérant (1.27) nous obtenons l'équivalence suivante:

$$\log X_{n-j+1,n} \sim \boldsymbol{\gamma} \log\left(\frac{n+1}{j}\right),\ n \to \infty,\ \left(\frac{n+1}{j}\right) \to \infty,\ j = 1, .., k_n. \tag{1.28}$$

1.5.1.1 Le Paréto Quantile Plot

Considérons la statistique d'ordre $X_{1,n} \leq ... \leq X_{n,n}$ associée à l'échantillon $(X_1, ..., X_n)$, le « Paréto Quantile plot » correspond au graphe de $(-\log\left(j/(n+1)\right), \log X_{n-j+1,n})$, $j = 1, ..., k_n$, et est une représentation très utile pour visualiser graphiquement si les observations sont distribuées selon une loi du domaine de Fréchet ou non. Dans ce domaine, on a d'après (1.28), le Paréto quantile plot serait approximativement linéaire avec une pente $\boldsymbol{\gamma}$, pour les petites valeurs de j, c'est-à-dire les points extrêmes. Cette linéarité apparaît au-delà du point $\left(-\log\left(\frac{k_n}{n+1}\right), \log X_{n-k_n+1,n}\right)$. Deux approches sont donc possible pour la construction de tels estimateurs: soit en forçant la droite à passer par ce point; soit simplement par la méthode des moindres carrés.

• Dans le premier cas, Csörgő *et al.* (1985) ont proposé une famille d'estimateurs à noyau $\widehat{\gamma}_{k_n}^K$ définie de la façon suivante:

$$\boxed{\widehat{\gamma}_{k_n}^K = k_n^{-1} \sum_{j=1}^{k_n} jK(j/k_n) \left(\log X_{n-j+1,n} - \log X_{n-j,n}\right) \Big/ k_n^{-1} \sum_{j=1}^{k_n} K(j/k_n),} \tag{1.29}$$

où K représente un noyau décroissant d'intégrale $\int K = 1$. Suivant le choix de ce noyau, différents estimateurs peuvent en résulter, le plus connu étant l'estimateur de Hill (1975), correspondant au cas particulier $K(x) = \mathbb{1}_{(0,1)}(x)$ et que nous précisons en detail dans la Section 1.5.1.2 ci-dessous.

• Dans le second cas, Schultze et Steinebach (1996) et indépendamment Kratz et Resnick (1996) ont introduit l'estimateur de *Zipf* ci-dessous en utilisant une approche par moindres carrés.

$$\boxed{\widehat{\gamma}_{k_n}^Z = \frac{\dfrac{1}{k_n} \displaystyle\sum_{j=1}^{k_n} \left(\log \frac{k_n+1}{j} - k_n^{-1} \sum_{j=1}^{k_n} \log \frac{k_n+1}{j} \right) \log X_{n-i+1,n}}{\dfrac{1}{k_n} \displaystyle\sum_{j=1}^{k_n} \left(\log \frac{k_n+1}{j} \right)^2 - \left(\dfrac{1}{k_n} \displaystyle\sum_{j=1}^{k_n} \log \frac{k_n+1}{j} \right)^2}.} \tag{1.30}$$

Remarque 1.5.1 *En remplaçant* $\log x$ *par* x *en ordonnée du graphe précédent, on obtient « L'exponential quantile plot » qui est similaire au « Paréto Quantile plot » mais concerne cette fois-ci le domaine de Gumbel* ($\boldsymbol{\gamma} = 0$)*. La pente asymptotique dans « L'exponential quantile plot » est donnée par le paramètre* $\boldsymbol{\sigma}$*.*

1.5.1.2 Estimateur de Hill

La définition de la fonction de distribution F de type Paréto peut être écrite comme

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma}, \ \forall\, x > 1.$$

Ce qui peut être interprété comme $\mathbb{P}(X/t > x | X > t) \to x^{-1/\gamma}$, quand $t \to \infty$ et $x > 1$. Et si on note $Y_j(t)$ les excès relatifs au-delà de t, i.e. $Y_j(t) = X_i/t$ avec $X_i > t$ tel que i est l'indice du j-*ème* excès dans l'échantillon originel et $j = 1, ..., N_t$, N_t représente le nombre d'excès.

En formant la log vraisemblance basée sur les excés $Y_1(t), ..., Y_{N_t}(t)$ conditionnellement à N_t on obtient

$$\log L\left(Y_1(t), ..., Y_{N_t}(t)\right) = -N_t \log \gamma - \left(1 + \frac{1}{\gamma}\right) \sum_{j=1}^{N_t} \log Y_j(t).$$

Alors

$$\frac{d \log L}{d\gamma} = -\frac{N_t}{\gamma} + \frac{1}{\gamma^2} \sum_{j=1}^{N_t} \log Y_j(t) = 0 \Leftrightarrow \gamma = \frac{1}{N_t} \sum_{j=1}^{N_t} \log Y_j(t).$$

Pour $1 \leq k_n < n$ où $k_n \to \infty$, choisissons pour le niveau d'excès t la valeur de la statistique d'ordre $X_{n-k_n,n}$ et remplaçons N_t par k_n, on obtient l'estimateur de Hill (1975) défini par:

$$\boxed{\widehat{\gamma}_{k_n}^H = \frac{1}{k_n} \sum_{j=1}^{k_n} \log X_{n-j+1,n} - \log X_{n-k_n,n} = \frac{1}{k_n} \sum_{j=1}^{k_n} j \left(\log X_{n-j+1,n} - \log X_{n-j,n}\right).} \quad (1.31)$$

D'autres méthodes pour la construction de l'estimateur de Hill sont données dans Beirlant *et al.* (2004) et dans de Haan et Ferreira (2006). Un grand nombre de travaux théoriques ont été consacré à l'étude des propriétés de l'estimateur de Hill. Mason (1982) a démontré la consistance faible et Deheuvels *et al.* (1988) ont établi la consistance forte. La normalité asymptotique est due entre autre à Davis et Resnick (1984), Csörgő *et al.* (1985), Haeusler et Teugels (1985) et Smith (1987).

Notons que la consistance (en probabilité ou presque sûrement) de l'estimateur $\widehat{\gamma}_{k_n}^H$ ne dépend que du comportement de k_n, alors que sa normalité asymptotique nécessite des conditions plus délicates sur la fonction de distribution F et donc sur la fonction de quantile de queue $U(\cdot)$ voir le Théorème 1.5.3.

Définition 1.5.2 *(Variation régulière du second ordre)*
Soit $F \in D(\mathcal{H}_\gamma)$, $\gamma > 0$, on dit que F est à variation régulière du second d'ordre à l'infini si elle satisfait les conditions équivalentes suivantes:

- *Il existe une constante réelle $\rho \leq 0$ et une fonction de signe constant $A^*(t) \to 0$ quand $t \to \infty$, telles que pour tout $x > 0$*

$$\lim_{t\to\infty} \frac{\left(\overline{F}(tx))/\overline{F}(t)\right) - x^{-1/\gamma}}{A^*(t)} = x^{-1/\gamma} \frac{x^\rho - 1}{\rho} \quad (1.32)$$

- *Il existe une constante réelle $\rho \leq 0$ et une fonction de signe constant $A(t) \to 0$ quand $t \to \infty$, telles que pour tout $x > 0$*

$$\lim_{t\to\infty} \frac{U(tx))/U(t) - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho}. \quad (1.33)$$

On peut interpréter $(x^\rho - 1)/\rho = \log x$ dans le cas où $\rho = 0$,

Remarquons que les relations (1.32) et (1.33) [6] impliquent que les fonctions $|A^*|$ et $|A|$ sont à variations regulières à l' infini d'indices respectifs ρ/γ et ρ (i.e $|A^*| \in \mathcal{RV}_{\rho/\gamma}$ et $|A| \in \mathcal{RV}_\rho$), pour plus de détails on peut se référer à Geluck et de Haan (1987). Le paramètre ρ (appelé *paramètre du second ordre*) contrôle donc la vitesse de convergence de $\big(\overline{F}(tx))/\overline{F}(t)\big)$ vers $x^{-1/\gamma}$ quand $t \to \infty$ ou de manière équivalente celle de $U(tx))/U(t)$ vers x^γ quand $t \to \infty$ (et donc la vitesse de convergence de $\ell_U(tx)/\ell_U(t)$ vers 1). Par exemple plus cette convergence est rapide, plus la linéarité dans le « Paréto quantile plot » apparaît vite. En d'autres termes, une valeur de ρ proche de 0 implique une faible vitesse de convergence, et plus l'estimation de l'indice de queue γ est difficile.
Les relations (1.32) et (1.33) ci-dessus peuvent être reformulées respectivement de la manière suivante:

$$\lim_{t\to\infty} \frac{\log(\overline{F}(tx)) - \log(\overline{F}(t)) + \gamma^{-1}\log x}{A^*(t)} = \frac{x^\rho - 1}{\rho}, \tag{1.34}$$

et

$$\lim_{t\to\infty} \frac{\log U(tx) - \log U(t) - \gamma \log x}{A(t)} = \frac{x^\rho - 1}{\rho}. \tag{1.35}$$

Ces conditions appelées « *conditions du second ordre* » sont satisfaites pour la plupart des lois appartenant au domaine de Fréchet (Cas des distribution à queues lourdes). Par exemple dans le cas des modèles de Hall (ou classe de Hall) ci-dessous, ces conditions ont été un peu plus spécifiées. La classe de Hall est composée de l'ensemble des distributions F telles que

$$F(x) = 1 - cx^{-1/\gamma}\left(1 + dx^{\rho/\gamma} + o(x^{\rho/\gamma}\right) \text{ quand } x \to \infty, \tag{1.36}$$

où $\gamma > 0$, $\rho \leq 0$, $c > 0$, et $d \in \mathbb{R}^*$. Cette sous classe des distributions à queues lourdes (de type Paréto) contient les lois de Paréto, Burr, Fréchet et t-Student fréquement utilisées en assurance mathématique, comme des modèles pour des grands risques (voir par exemple le Chapitre 5). La relation (1.36) peut être reformulée [7] en utilisant la fonction quantile de queue $U(\cdot)$ de la manière suivante $U(t) = c^\gamma t^\gamma\left(1 + \gamma d c^\rho t^\rho + o(t^\rho)\right)$ quand $t \to \infty$.

[6] Les équations (1.32) et (1.33) sont aussi équivalentes à:
Il existe une constante réelle $\rho \leq 0$ et une fonction $A^{**}(s) \to 0$ quand $s \to 0$, telles que pour tout $x > 0$

$$\lim_{s\to 0} \frac{(Q(1-sx))/Q(1-s)) - x^{-\gamma}}{A^{**}(s)} = x^{-\gamma}\frac{x^\rho - 1}{\rho}.$$

$|A^{**}|$ est à variation régulière d'indice $-\rho$ en zero et $A^{**}(s) = A(1/s)$.

[7] (1.36) peut être reformulée avec la fonction quantile $Q(\cdot)$

$$Q(1-s) = c^\gamma s^{-\gamma}\left(1 + \gamma d c^\rho s^{-\rho} + o(s^{-\rho})\right) \text{ quand } s \to 0.$$

Ainsi, pour le cas des modèles de Hall, les fonctions $A(\cdot)$ et $A^*(\cdot)$ sont respectivement équivalentes à $d\rho\gamma c^\rho t^\rho$ et $d\rho\gamma t^{\rho/\gamma}$ quand $t \to \infty$.

Théorème 1.5.3 *(Propriétés asymptotiques de l'estimateur de Hill)*
Soit k_n, $n \geq 1$ une suite d'entiers telle que $1 \leq k_n < n$, $k_n \to \infty$ et $k_n/n \to 0$ quand $n \to \infty$.
- *(Consistance faible): Alors, $\widehat{\gamma}^H_{k_n}$ converge en probabilité vers γ.*
- *(Consistance forte): Si de plus $k_n/\log\log n \to \infty$ quand $n \to \infty$, alors $\widehat{\gamma}^H_{k_n}$ converge presque sûrement vers γ.*
- *(Normalité asymptotique): Si la condition (1.33) est satisfaite avec $\sqrt{k_n}A(n/k_n) \to \lambda \in \mathbb{R}$ quand $n \to \infty$, alors $\sqrt{k_n}\left(\widehat{\gamma}^H_{k_n} - \gamma\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\lambda/(1-\rho), \gamma^2\right)$.*

En utilisant la normalité asymptotique de $\widehat{\gamma}^H_{k_n}$, Beirlant *et al.* (1996) trouvent un intervalle de confiance de $\widehat{\gamma}^H_{\widehat{k_{n\,opt}},n}$ où $\widehat{k_{n\,opt}}$ est une valeur optimale de k_n donnée par leur algorithme. Dans leurs travaux, ils mentionnent qu'il est nécessaire d'être prudent sur l'intervalle de confiance du fait que la variabilité du choix de k_n est ignorée et ce qui ne devait pas être le cas. Une méthode plus précise suggérée pour la construction d'un intervalle de confiance est le bootstrap (Danielsson *et al.*, 2001; Gomes et Oliveira, 2001). Des méthodes de sélection du meilleur choix de k_n sont données dans la Section 1.6.

Remarque 1.5.4 *D'après la normalité asymptotique de Hill, si $\sqrt{k_n}A(n/k_n) \to 0$, on peut écrire: $\sqrt{k_n}\left(\widehat{\gamma}^H_{k_n}/\gamma - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$. Ainsi, un intervalle de confiance au niveau $(1-\alpha)$ pour l'estimation de γ est donné par:*

$$\left(\widehat{\gamma}^H_{k_n}\left(1 + k_n^{-1/2} z_{1-\alpha/2}\right)^{-1}, \widehat{\gamma}^H_{k_n}\left(1 - k_n^{-1/2} z_{1-\alpha/2}\right)^{-1}\right),$$

où z_α est le quantile d'une loi normale standard au niveau α.

Par exemple pour une loi de Paréto standard on a $\overline{F}(x) = x^{-1/\gamma}$, $x \geq 1$ et $\gamma > 0$, la fonction à variations lentes associée est $\ell_F(x) = 1$ alors, il n'y a pas de biais asymptotique dans l'estimateur de Hill (la limite $\lambda = 0$). Le graphe de Hill et son intervalle de confiance au niveau de 95% pour cette loi sont représentés dans la Figure 1.3.

1.5.1.3 Correction de biais sur l'estimateur de Hill

Comme nous l'avions vu précédement, afin d'estimer systématiquement γ, nous avons considéré une suite croissante k_n de statistiques d'ordres supérieur. Lorsque k_n est trop petit, la variance de l'estimateur sera grande. D'autre part, si k_n est trop grand, alors le biais est introduit dans l'estimation. Le choix de k_n est donc une question d'équilibre

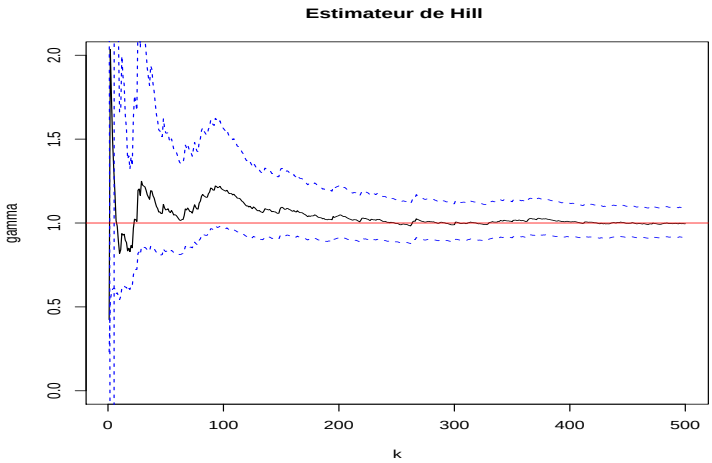

FIG. 1.3 – Graphique de l'estimateur de Hill en moyenne, en fonction du nombre d'extrêmes k_n (ligne pleine) avec 95% de borne de confiance (ligne pointillée), d'une loi de Paréto standard ($\gamma = 1$) basée sur 100 réalisations de 3000 observations. La ligne horizontale représente la vraie valeur de γ.

entre le biais et la variance et est donné dans la Section 1.6. Une des méthodes pour résoudre ce problème est de faire une correction de biais pour ces estimateurs quand un grand nombre k_n de statistiques d'ordre supérieur est utilisé dans l'estimation. Une telle formulation nécessite des hypothèses supplémentaires sur la fonction de distribution F. Pour plus de details sur ce type de mèthodes on pourra se reférer à Peng (1998), Beirlant *et al.* (1999), Feuerverger et Hall (1999), Beirlant *et al.* (2002), Gomes et Martins (2002), Matthys et Beirlant (2003), Caeiro *et al.* (2009) et Gomes *et al.* (2009). Nous allons seulement donner la méthode de correction de biais proposée dans Beirlant *et al.* (2002) car nous l'utiliserons dans la suite.

Cette méthode repose en partie sur le modèle de régression exponentielle des espacements entre les logarithmes des plus grandes statistiques d'ordre. Ce modèle, proposé par Beirlant *et al.* (1999), Feuerverger et Hall (1999) et Beirlant *et al.* (2002) pour des lois du domaine d'attraction de Fréchet, est défini de la manière suivante:

$$\boxed{Z_{\jmath k_n} = \jmath \log\left(\frac{X_{n-\jmath+k_n,n}}{X_{n-k_n,n}}\right) \overset{\mathcal{D}}{\sim} \left(\gamma + A(n/k_n)\left(\frac{\jmath}{k_n+1}\right)^{-\omega}\right) + \varepsilon_{\jmath k_n},\ 1 \leq \jmath \leq k_n,} \quad (1.37)$$

où $\varepsilon_{\jmath k_n}$ est un terme d'erreur centré. Si on ignore le terme $A(n/k_n)$ dans (1.37), on retrouve l'estimateur de Hill $\widehat{\gamma}_n^H(k_n)$. Par la méthode des moindres carrés, la représentation (1.37) a

été exploitée par Beirlant *et al.* (2002) pour déterminer un estimateur de $\boldsymbol{\gamma}$ et de $A(n/k_n)$ où ρ est remplacé par un estimateur consistant $\widehat{\boldsymbol{\rho}} = \widehat{\rho}_{\widetilde{k}_n,n}$ avec $\widetilde{k}_n$ donné de tel sorte que $\widetilde{k}_n \to \infty$, $\widetilde{k}_n/n \to 0$ et $\widetilde{k}_n^{1/2} A(n/\widetilde{k}_n) \to \infty$. On obtient les estimateurs suivants:

$$\boxed{\begin{cases} \widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) = \dfrac{1}{k_n}\displaystyle\sum_{\jmath=1}^{k_n} Z_{\jmath k_n} - \widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})/(1-\widehat{\boldsymbol{\rho}}), \\ \widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) = \dfrac{(1-2\widehat{\boldsymbol{\rho}})(1-\widehat{\boldsymbol{\rho}})^2}{\widehat{\boldsymbol{\rho}}^2}\dfrac{1}{k_n}\displaystyle\sum_{\jmath=1}^{k_n}\left(\left(\dfrac{\jmath}{k_n+1}\right)^{-\widehat{\boldsymbol{\rho}}} - \dfrac{1}{1-\widehat{\boldsymbol{\rho}}}\right) Z_{\jmath k_n}. \end{cases}} \tag{1.38}$$

Ici, $k_n^{-1}\sum_{\jmath=1}^{k_n}(\jmath/(k_n+1))^{-\widehat{\boldsymbol{\rho}}}$ peut être approximé par $1/(1-\widehat{\boldsymbol{\rho}})$ et $k_n^{-1}\sum_{\jmath=1}^{k_n}((\jmath/(k_n+1))^{-\widehat{\boldsymbol{\rho}}} - 1/(1-\widehat{\boldsymbol{\rho}}))^2$ par $\widehat{\boldsymbol{\rho}}^2/((1-2\widehat{\boldsymbol{\rho}})(1-\widehat{\boldsymbol{\rho}})^2)$. Notons que $\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})$ est un estimateur sans biais de $\boldsymbol{\gamma}$ et asymptotiquement normal. Les estimateurs donnés dans (1.38) sont utilisés pour la réduction du biais des estimateurs de primes d'assurance dans le Chapitre 5.

Théorème 1.5.5 *Supposons que $F \in D(\mathcal{H}_{\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$ et que la condition du second ordre (1.33) soit verifiée. Si on substitue ρ par un estimateur consistant $\widehat{\boldsymbol{\rho}}$ dans la procédure d'estimation, alors pour $k_n \to \infty$, $k_n/n \to 0$ et $\sqrt{k_n}A(n/k_n) = O(1)$ quand $n \to \infty$ on obtient*

$$\sqrt{k_n}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) - \boldsymbol{\gamma}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \boldsymbol{\gamma}^2(1-\rho)^2/\rho^2\right)$$

et

$$\widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) = A(n/k_n) + \frac{N_{k_n}}{\sqrt{k_n}} + o_{\mathbb{P}}(A(n/k_n)),$$

avec $N_{k_n} \xrightarrow{\mathcal{D}} \mathcal{N}(0,(1-2\rho)(1-\rho)^2/\rho^2)$ quand $k_n \to \infty$.

Pour la preuve du Théorème 1.5.5, le lecteur pourra se référer à Beirlant *et al.* (2002), Théorème 3.2. La question que l'on se pose sur (1.38) concerne l'estimation du paramètre du second ordre ρ. Comme estimateur on pourra se referer à(Hall et Welsh (1985); Gomes et Martins (2002); Fraga Alves *et al.* (2003b); Ciuperca et Mercadier (2010); Goegebeur *et al.* (2010).

On peut voir que les estimateurs de ρ basés sur le modèle de régression discuté ici ont les mêmes propriétés de consistance quand $\widetilde{k}_n^{1/2}A(n/\widetilde{k}_n) \to \infty$. Ces auteurs ont montré la normalité asymptotique de leurs estimateurs du paramètre en utilisant des conditions supplémentaires sur la distribution F, $A(n/\widetilde{k}_n)$ etc. On peut se referer au Chapitre 4 pour plus de détails.

Remarque 1.5.6 *L'estimation de ρ est souvent difficile en pratique. Certains auteurs ont montré par des simulations que la fixation de sa valeur n'a pas beaucoup d'influence*

sur les résultats. Ce problème a déjà été discuté dans Beirlant et al. (1999); Feuerverger et Hall (1999) et Beirlant et al. (2002) où des remarques similaires ont été faites dans le cas des distributions à queues lourdes. Le choix canonique $\widehat{\boldsymbol{\rho}} = \rho_0 = -1$ *est souvent utilisé comme critère d'estimateur de* ρ *car ayant de meilleures performances par rapport aux autres par un modèle de sélection.*

1.5.2 Estimation pour différents domaines d'attraction

Le cas général $\boldsymbol{\gamma} \in \mathbb{R}$ a été moins étudié. Nous recensons dans cette partie les estimateurs proposés dans Pickands (1975); Beirlant *et al.* (1999) et Dekkers *et al.* (1989).

1.5.2.1 L'estimateur de Pickands

Il est fondé sur le calcul des quantiles. Cet estimateur a été introduit par Pickands (1975) et revisité par Drees (1995) et Drees et Kaufmann (1998) d'une part, Dekkers et de Haan (1989) d'autre part, qui en ont établi sa consistance faible ainsi que sa normalité asymptotique. Cet estimateur est construit en utilisant trois statistiques d'ordres et a l'avantage d'être valable quelque soit le domaine d'attraction de la distribution. Il est défini de la manière suivante:

$$\boxed{\widehat{\boldsymbol{\gamma}}_{k_n}^{P} = \frac{1}{\log 2} \log \frac{X_{n-k_n+1,n} - X_{n-2k_n+1,n}}{X_{n-2k_n+1,n} - X_{n-4k_n+1,n}}, \ k_n = 1, ..., [n/4].}$$

Notons que l'estimateur de Pickands $\widehat{\boldsymbol{\gamma}}_{k_n}^{P}$, qui utilise l'information apportée par des distances entre deux statistiques ordonnées n'utilise pas le maximum de l'échantillon $X_{n,n}$, ce qui implique une perte d'information sur la queue de distribution. Une justification plus formelle de l'estimateur de Pickands a été fournie dans Embrechts *et al.* (1997)

Les propriétés de l'estimateur de Pickands ont été principalement explorées dans Dekkers et de Haan (1989). Plus précisément, ils ont prouvé les resultats suivants:

Théorème 1.5.7 *On suppose que* $F \in D(\mathcal{H}_{\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} \in \mathbb{R}$ *la condition du second ordre (1.33) soit verifiée. Soit* k_n, $n \geq 1$ *une suite d'entiers telle que* $1 \leq k_n < [n/4]$, $k_n \to \infty$ *et* $k_n/n \to 0$ *quand* $n \to \infty$.

- *(Consistance faible): Alors,* $\widehat{\boldsymbol{\gamma}}_{k_n}^{P} \xrightarrow{\mathbb{P}} \boldsymbol{\gamma}$ *quand* $n \to \infty$.
- *(Consistance forte): Si de plus* $k_n/\log\log n \to \infty$ *quand* $n \to \infty$, *alors* $\widehat{\boldsymbol{\gamma}}_{k_n}^{P} \xrightarrow{p.s} \boldsymbol{\gamma}$.
- *(Normalité asymptotique): on suppose que la fonction de queue* U *associée à* F *satisfait la condition du second ordre ci-dessous: ils existent une constante réelle* $\rho \leq 0$ *et*

une fonction $A(t) \to 0$ *quand* $t \to \infty$ *telles que pour tout* $x > 0$,

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - D_\gamma(x)}{A(t)} = \int_1^x s^{\gamma-1} \int_1^s u^{\rho-1} du ds, \tag{1.39}$$

avec $D_\gamma(x)$ *et* $a(\cdot)$ *donnés dans (1.5.7). Alors pour* $\sqrt{k_n}A(n/k_n) \to \lambda \in \mathbb{R}$ *quand* $n \to \infty$ *on a:*

$$\sqrt{k_n}\left(\widehat{\gamma}^P_{k_n} - \gamma\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\lambda b_{\rho,\gamma}, \sigma^2_P(\gamma)\right) \text{ quand } n \to \infty,$$

où

$$\sigma^2_P(\gamma) = \frac{\gamma^2(2^{2\gamma+1}+1)}{4(\log 2)^2(2^\gamma-1)^2}$$

et

$$b_{\rho,\gamma} = \begin{cases} \dfrac{4^{-\rho}\gamma(4^{\gamma+\rho}-1) - (2^\gamma+1)(2^{\gamma+\rho}-1)}{\rho 2^\gamma(\gamma+\rho)(2^\gamma-1)\log 2} & si \quad \rho < 0, \ \gamma \in \mathbb{R}, \\ 1 & si \quad \rho = 0. \end{cases}$$

Pour plus de détails on peut se référer dans de Haan et Ferreira (2006) p. 85-86. Une justification plus formelle sur l'estimateur de Pickands est fournie dans Embrechts *et al.* (1997).

Remarque 1.5.8 *En utilisant la normalité asymptotique de l'estimateur de Pickands du Théorème 1.5.7, nous avons pour* $\sqrt{k_n}A(n/k_n) \to 0$ *quand* $n \to \infty$, $\left(\sqrt{k_n}/\sigma_P(\gamma)\right)\left(\widehat{\gamma}^P_{k_n} - \gamma\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$. *En remplaçant* $\sigma_P(\gamma)$ *par son estimateur consistent* $\sigma_P(\widehat{\gamma}^P_{k_n})$, *on peut donner un intervalle de confiance au niveau* $(1-\alpha)$ *de* γ: $\left(\widehat{\gamma}^P_{k_n} - z_{\alpha-1/2}\frac{\sigma_P(\widehat{\gamma}^P_{k_n})}{\sqrt{k_n}}, \widehat{\gamma}^P_{k_n} + z_{\alpha-1/2}\frac{\sigma_P(\widehat{\gamma}^P_{k_n})}{\sqrt{k_n}}\right)$, z_α *est le quantile d'ordre* α *d'une loi normale standard.*

Pour une loi Uniforme standard de distribution F, on a $\overline{F}(x_F - 1/x) = 1/x$, $x > 1$ et $\gamma = -1$, la fonction à variation lente associée $\ell_F(x) = 1$ (voir Beirlant *et al.* (2005), p. 68). Alors, il n'y a pas de biais asymptotique dans l'estimateur de Pickands pour la loi uniforme (plus précisément $\sqrt{k_n}A(n/k_n) \to 0$ quand $n \to \infty$).
La figure 1.4 illustre l'estimateur de Pickands avec 95% de borne de confiance pour cette loi.

1.5.2.2 Estimateur des moments

En se basant sur la relation (1.21), Dekkers *et al.* (1987) ont proposé une extension de l'estimateur de Hill appelée estimateur des moments, qui lui est valable quel que soit

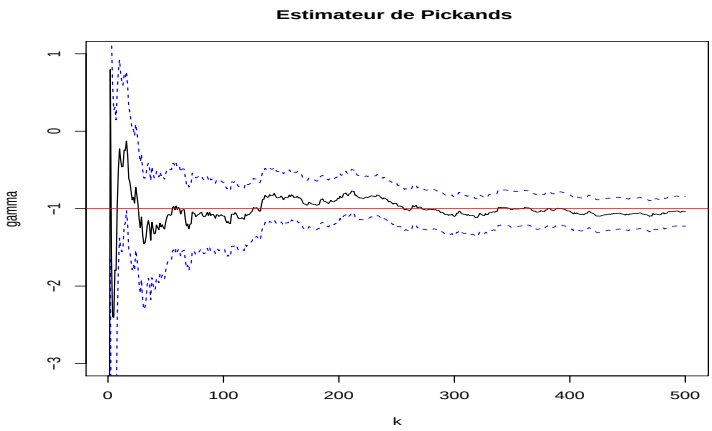

FIG. 1.4 – Graphique de l'estimateur de Pickands en moyenne (ligne pleine), en fonction du nombre d'extrêmes avec 95% de borne de confiance (ligne pointillée), d'une distribution Uniforme standard ($\gamma = -1$) basée sur 100 échantillons de 3000 observations. La ligne horizontale correspond à la vraie valeur de γ.

le signe de l'indice de queue (i.e $\gamma \in \mathbb{R}$) et qui est défini comme suit:

$$\boxed{\widehat{\gamma}_{k_n}^{\mathcal{M}} = \mathcal{M}_{k_n}^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(\mathcal{M}_{k_n}^{(1)}\right)^2}{\mathcal{M}_{k_n}^{(2)}}\right)^{-1}}$$

où $\mathcal{M}_{k_n}^{(r)} = \dfrac{1}{k_n}\displaystyle\sum_{j=1}^{k_n}(\log X_{n-j+1,n} - \log X_{n-k_n,n})^r$, $r = 1, 2$. Notons que $\mathcal{M}_{k_n}^{(1)}$ correspond à l'estimateur de Hill $\widehat{\gamma}_{k_n}^{H}$. L'estimateur des moments est également connu sous l'appellation d'estimateur de *Dekkers-Einmahl-de Haan*.
Les propriétés asymptotiques de l'estimateur des moments $\widehat{\gamma}_{k_n}^{\mathcal{M}}$ ont été établies par Dekkers *et al.* (1987), Théorème 3.1 et Corollaire 3.2 et que nous allons rappeler dans le théorème ci-dessous.

Théorème 1.5.9 *Supposons que $F \in D(\mathcal{H}_\gamma)$, $\gamma \in \mathbb{R}$, $k_n \to \infty$ et $k_n/n \to 0$ quand $n \to \infty$*

- *(Consistance faible): Alors, $\widehat{\gamma}_{k_n}^{\mathcal{M}} \xrightarrow{\mathbb{P}} \gamma$ quand $n \to \infty$.*
- *(Consistance forte): Si de plus $k_n/(\log n)^\delta \to \infty$ quand $n \to \infty$, pour un certains $\delta > 0$, alors $\widehat{\gamma}_{k_n}^{\mathcal{M}} \xrightarrow{p.s} \gamma$ quand $n \to \infty$.*

• *(Normalité asymptotique): Sous certaines conditions sur la loi F (voir Dekkers et al. (1987)),*

$$\sqrt{k_n}\left(\widehat{\gamma}_{k_n}^{\mathcal{M}} - \gamma\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_{\mathcal{M}}^2\right) \quad \textit{quand } n \to \infty.$$

où

$$\sigma_{\mathcal{M}}^2 = \begin{cases} 1+\gamma^2 & \textit{si} \quad \gamma \geq 0, \\ (1-\gamma^2)(1-2\gamma)\left(4 - 8\dfrac{1-2\gamma}{1-3\gamma} + \dfrac{(5-11\gamma)(1-2\gamma)}{(1-3\gamma)(1-4\gamma)}\right) & \textit{si} \quad \gamma < 0. \end{cases} \tag{1.40}$$

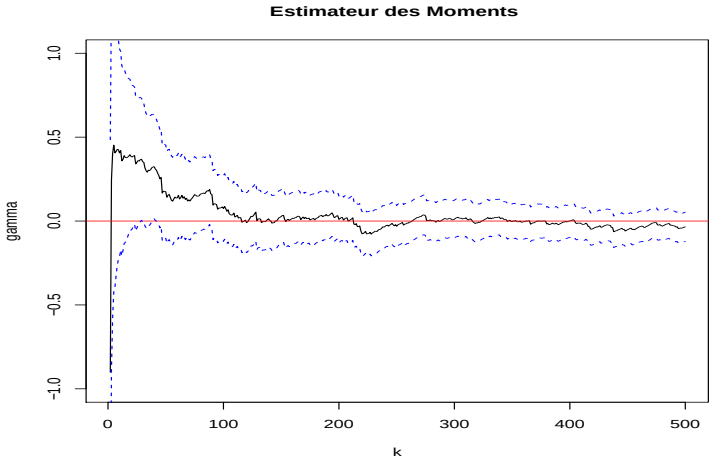

FIG. 1.5 – Graphique de l'estimateur des moments en moyenne (ligne pleine), en fonction du nombre d'extrêmes avec 95% de borne de confiance (lignes pointillées), d'une distribution de Gumbel ($\gamma = 0$) basée sur 100 réalisations de 3000 observations.

Notons que le problème de l'estimateur des moments est que l'on ne peut pas, contrairement à l'estimateur de Hill, l'interpréter graphiquement. Afin d'apporter une solution à ce problème, une généralisation du « Paréto quantile plot » a été proposée dans Beirlant *et al.* (1996).

Remarque 1.5.10 *Il n'est pas facile de choisir un estimateur de γ, mais l'estimateur de Hill présente une variance asymptotique plus faible que l'estimateur de Pickands et des moments (dans le cas où $\gamma > 0$), et est généralement le plus utilisé.*

Etant donné une multitude d'estimateurs pour l'indice des valeurs extrêmes, il est nécessaire de savoir celui qui marche le mieux. Malheureusement, il n'y a pas une meilleure

méthode dans toutes les situations pour qu'un estimateur soit meilleur qu'un autre. Tout dépend de l'indice $\boldsymbol{\gamma}$ et de la série étudiée. Nous représentons dans la Figure 1.6 les estimateurs de Hill, de Pickands et des Moments sur les distributions de Burr [8] et de Cauchy. Nous remarquons que l'estimateur de Pickands apparaît comme étant l'estimateur le moins stable.

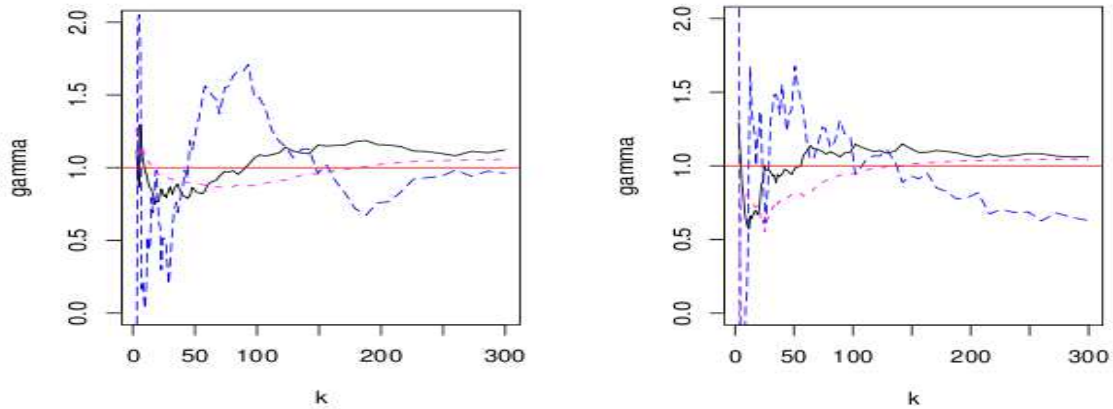

FIG. 1.6 – Estimateurs de Hill $\widehat{\gamma}^H_{k_n}$ (ligne pleine), de Pickands $\widehat{\gamma}^P_{k_n}$ (ligne tiret) et des moments (ligne en pointtillé), en fonction du nombre d'extrêmes k_n sur une distribution de Burr(1,1,1) (graphique de gauche) et sur une distribution de Cauchy (graphe de droite), basées sur 100 réalisations de 3000 observations. La ligne verticale correspond à la vraie valeur de $\boldsymbol{\gamma}$.

1.5.3 Estimateurs du couple $(\boldsymbol{\gamma}, \boldsymbol{\sigma})$

1.5.3.1 Estimateurs par la méthode des moments

Dans le cas d'une distribution de Paréto Généralisée $\mathcal{G}_{\boldsymbol{\gamma},\boldsymbol{\sigma}}$ (voir Equation 1.9), l'estimateur par la méthode des moments (voir Hosking et Wallis (1987), Section 3.2) est basé sur le fait que

$$\mathbb{E}\left[\left(1+\frac{\boldsymbol{\gamma}X}{\boldsymbol{\sigma}}\right)^r\right] = \frac{1}{1-r\boldsymbol{\gamma}} \quad \text{si} \quad 1-r\boldsymbol{\gamma}>0.$$

Il en découle de l'expression des paramètres en fonction des deux premiers moments μ_1 et μ_2

$$\boldsymbol{\gamma} = \frac{1}{2}\left(1-\frac{\mu_1^2}{\mu_2-\mu_1^2}\right) \quad \text{et} \quad \boldsymbol{\sigma} = \frac{1}{2}\mu_1\left(\frac{\mu_1^2}{\mu_2-\mu_1^2}+1\right).$$

[8] Distribution de Burr $Burr(\xi,\eta,\lambda)$ définie par $1-F(x) = (\xi/(\xi+x^\eta))^\lambda$, $x>0$, $\xi,\lambda,\eta>0$.

En remplaçant dans les expressions précédentes μ_1 et μ_2 par leurs versions empiriques, nous obtenons l'estimateur des moments $(\widehat{\boldsymbol{\gamma}}_{k_n}^{\mathcal{MOM}}, \widehat{\boldsymbol{\sigma}}_{k_n}^{\mathcal{MOM}})$ du couple $(\boldsymbol{\gamma}, \boldsymbol{\sigma})$ définis comme suit

$$\boxed{\widehat{\boldsymbol{\gamma}}_{k_n}^{\mathcal{MOM}} = \frac{1}{2}\left(1 - \frac{\overline{X}^2}{\overline{X^2} - \overline{X}^2}\right) \quad \text{et} \quad \widehat{\boldsymbol{\sigma}}_{k_n}^{\mathcal{MOM}} = \frac{1}{2}\overline{X}\left(\frac{\overline{X}^2}{\overline{X^2} - \overline{X}^2} + 1\right).}$$

où $\overline{X}$ et $\overline{X^2}$ sont les estimateurs empiriques des moments d'ordre 1 et 2 de léchantillon. La normalité asymptotique du couple $(\widehat{\boldsymbol{\gamma}}_{k_n}^{\mathcal{MOM}}, \widehat{\boldsymbol{\sigma}}_{k_n}^{\mathcal{MOM}})$ peut être établie sous la condition $\boldsymbol{\gamma} < 1/4$. Pour affaiblir cette condition d'autres alternatives ont été proposées, telle que la méthode des moments pondérés (éventuellement généralisé)

1.5.3.2 Estimateurs par la méthode des moments pondérés

Hosking et Wallis (1987) ont également proposé d'utiliser les deux moments pondérés ν_0 et ν_1 avec $\nu_s = \mathbb{E}[X(1 - \mathcal{G}_{\boldsymbol{\gamma},\boldsymbol{\sigma}}(X))^s]$. Nous avons

$$\forall\ s > \gamma - 1, \quad \nu_s = \frac{\boldsymbol{\sigma}}{(s+1)(s - \boldsymbol{\gamma} + 1)}.$$

En notant $P_n = \dfrac{1}{k_n}\displaystyle\sum_{j=0}^{k-1} X_{n-j,n} - X_{n-k_n,n}$ et $Q_n = \dfrac{1}{k_n}\displaystyle\sum_{j=0}^{k-1} \dfrac{j}{k_n}\left(X_{n-j,n} - X_{n-k_n,n}\right)$, P_n et Q_n sont respectivement les moments empiriques correspondant à ν_0 et ν_1; les estimateurs de $\boldsymbol{\gamma}$ et $\boldsymbol{\sigma}$ sont alors

$$\boxed{\widehat{\boldsymbol{\gamma}}_{k_n}^{\mathcal{PWM}} = \frac{P_n - 4Q_n}{P_n - 2Q_n} = 1 - \left(\frac{P_n}{2Q_n} - 1\right)^{-1} \quad \text{et} \quad \widehat{\boldsymbol{\sigma}}_{k_n}^{\mathcal{PWM}} = \frac{2P_nQ_n}{P_n - 2Q_n} = P_n\left(\frac{P_n}{2Q_n} - 1\right)^{-1}.}$$

La normalité asymptotique de $(\widehat{\boldsymbol{\gamma}}_{k_n}^{\mathcal{PWM}}, \widehat{\boldsymbol{\sigma}}_{k_n}^{\mathcal{PWM}})$ a été etablie dans Hosking et Wallis (1987). Le domaine de validité nécessaire de cette normalité asymptotique est assez restrictif : $\boldsymbol{\gamma} \in (-1, 1/2)$; néamoins, cette méthode présente un certain nombre d'avantages tant au niveau de la facilité du calcul que des performances à distance finie (voir Hosking et Wallis (1987)). Une généralisation de cette méthode permettant détendre le domaine de validité $\boldsymbol{\gamma} \in (-1, 3/2)$ a été proposée par Diebolt *et al.* (2004) et Diebolt *et al.* (2007).

1.6 Choix du nombre d'extrêmes k_n

Les résultats concernant les estimateurs de l'indice des valeurs extrêmes sont asymptotiques : ils sont obtenus lorsque $k_n \to \infty$ et $k_n/n \to 0$. La difficulté en pratique consiste à choisir le nombre d'extrêmes k_n utilisé dans les estimations. L'issue est importante:

l'extrême volatilité du graphe $\{(k_n, \widehat{\gamma}_{k_n}) : \ 1 \leq k_n < n\}$, où $\widehat{\gamma}_{k_n}$ représente n'importe quel estimateur introduit précédement, rend difficile l'utilisation de l'estimateur en pratique si aucune indication sur le choix de k_n n'est donnée. Des travaux ont montré qu'en utilisant trop d'observations, dans la procédure d'estimation de γ, on observe un biais substantiel tandis que l'utilisation de peu d'observations conduit à une variance considérable. Ce probème a été longuement abordé dans la littérature, voir par exemple Balkema et de Haan (1974), Hall et Welsh (1985), Dekkers et de Haan (1993), Reiss et Thomas (1997), de Haan et Peng (1998), Drees et Kaufmann (1998), Danielsson *et al.* (2001), Cheng et Peng (2001) Beirlant *et al.* (2002), Beirlant *et al.* (2004), etc.

1.6.1 Méthode Graphique

C'est la méthode la plus simple pour la détermination de k_n. Elle consiste à tracer le graphe $\{(k_n, \widehat{\gamma}_{k_n}) : \ 1 \leq k_n < n\}$ dans le but de trouver une valeur optimale de k_n. Ainsi le k_n optimal (notée k_{opt}) est choisi dans la première région où l'estimateur $\widehat{\gamma}_{k_n}$ devient stable. Nous illustrons cette méthode sur les données d'assurance de montants de sinistres incendies (en millions de couronnes Danoises), survenus au Danemark (2492 données). La Figure 1.6.1 suggère une valeur estimée de $0,7$ pour l'indice de queue sur ces données, et un nombre optimal d'extrêmes utilisé dans l'estimation de l'ordre de 300.

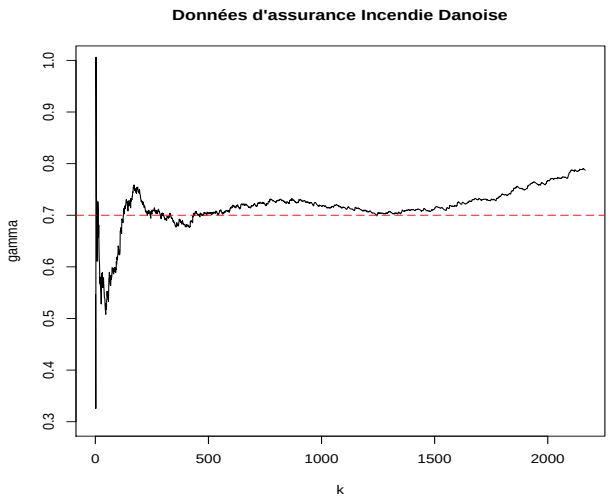

FIG. 1.7 – $\widehat{\gamma}^H_{k_n}$ en fonction du nombre d'extrêmes k_n sur les données d'assurance incendie Danoise pour la période allant de 01/01/1980 au 31/12/1990. La ligne horizontale correspond à la valeur estimée de l'indice de queue.

D'autres méthodes graphiques pour le choix optimal de k_n sont discutées dans Sousa (2002).

1.6.2 Méthode basée sur l'erreur moyenne quadratique

Dans un autre point de vue, une minimisation de l'*erreur moyenne quadratique asymptotique* (en anglais mean squared error: $\mathcal{AMSE}$) est souvent donnée comme critère. L'erreur quadratique moyenne de l'estimateur $\widehat{\gamma}_{k_n}$ de l'indice de queue $\boldsymbol{\gamma}$ est définie par $\mathcal{AMSE} = \mathbb{E}_\infty((\widehat{\gamma}_{k_n} - \boldsymbol{\gamma})^2)$, où $\mathbb{E}_\infty$ est l'espérance mathématique suivant la distribution asymptotique. Il est donc facile de voir que l'erreur quadratique moyenne de $\widehat{\gamma}_{k_n}$, qui est en fonction de k_n, n'est rien d'autre que le carré du biais plus la variance de l'estimateur considéré. Par conséquent pour une estimation précise de l'indice de queue $\boldsymbol{\gamma}$, il est nécessaire pour un estimateur classique de trouver un compromis entre le biais et la variance. Il semble raisonnable qu'une minimisation du $\mathcal{AMSE}$ permet de trouver une valeur intermédiaire entre les composantes du biais et de la variance pour ce compromis. Par exemple en utilisant le Théorème 1.5.3, l'érreur quadratique moyenne de l'estimateur de Hill est donnée par

$$\mathcal{AMSE}(\boldsymbol{\gamma}, A(n/k_n), \rho) = \frac{\boldsymbol{\gamma}^2}{k_n} + \left(\frac{A(n/k_n)}{1-\rho}\right)^2.$$

Il nous sera impossible de trouver une valeur de k_n minimisant $\mathcal{AMSE}(\boldsymbol{\gamma}, A(n/k_n), \rho)$, car cette dernière dépend des paramètres inconnus $\boldsymbol{\gamma}$, $A(n/k_n)$ et ρ. Pour résoudre ce problème une idée consiste à substituer respectivement $\boldsymbol{\gamma}$, $A(n/k_n)$ et ρ par leurs estimateurs consistants. Ainsi, pour un estimateur consistant $\widehat{\rho}$ donné, nous pouvons utiliser les estimateurs sans biais de $\boldsymbol{\gamma}$ et de $A(n/k_n)$ développés dans (1.38), et un estimateur de $\mathcal{AMSE}(\boldsymbol{\gamma}, A(n/k_n), \rho)$ est donné par

$$\widehat{\mathcal{AMSE}}(k_n) := \mathcal{AMSE}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}), \widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}), \widehat{\boldsymbol{\rho}}\right).$$

Il semble donc naturel de trouver une valeur $\widehat{k_n}$ qui minimise le graphe de l'erreur quadratique moyenne estimée $\left\{\left(k_n, \widehat{\mathcal{AMSE}}(k_n)\right),\ k_n = 1, ..., n-1\right\}$. La valeur optimale de k_n est donnée par

$$\widehat{k_n}_{opt} = \arg\min_{k_n} \widehat{\mathcal{AMSE}}(k_n).$$

1.6.3 Procédure numérique

Notons que la procédure du choix optimal de k_n est souvent difficile du fait que celui-ci ne dépend pas exclusivement de la taille de l'échantillon et de l'indice des valeurs extrêmes $\boldsymbol{\gamma}$, mais il dépend d'autres paramètres inconnus (le paramètre du second ordre ρ; etc..) intervenant dans la fonction de distribution F. Pour surmonter cet obstacle, une grande

variété d'algorithmes de $\widehat{k}_{nopt}$, est donnée dans le sens où l'estimateur $\widehat{k}_{nopt}$ de $\widehat{k}_{nopt}$ obtenu satisfait $\widehat{k}_{nopt}/\widehat{k}_{nopt} \xrightarrow{\mathbb{P}} 1$ quand $n \to \infty$. L'estimateur de l'indice de queue correspondant $\widehat{\gamma}_{\widehat{k}_{nopt},n}$ est aussi asymptotiquement efficace que $\widehat{\gamma}_{\widehat{k}_{nopt},n}$, on pourra se référer à Hall et Welsh (1985), Reiss et Thomas (1997), Cheng et Peng (2001)Beirlant *et al.* (2002), etc.

1.7 Estimation de quantiles extrêmes

Dans l'analyse des extrêmes, on s'intéresse souvent à l'estimation des quantités d'evénements d'occurrences rares. Dans de nombreux domaines d'application, comme l'assurance, la finance, l'hydrologie et les contrôles statistiques de la qualité, une exigence typique consiste à trouver les grandes valeurs, de sorte que les probabilités de les dépasser soient très faibles c'est-à-dire proches de zéro. Ces quantités sont appelées quantiles extrêmes [9] car l'ordre de ces quantiles tend vers zéro lorsque la taille n de l'échantillon tend vers l'infini. Nous nous sommes donc intéressés à l'estimation des quantiles extrêmes d'ordre $\alpha = \alpha_n \in (0,1)$ et $\alpha_n \to 0$ quand $n \to \infty$, défini par

$$q(\alpha_n) = F^{\leftarrow}(1-\alpha_n) = Q(1-\alpha_n) = U(1/\alpha_n).$$

Plus généralement si la taille de l'échantillon tend vers l'infini, nous avons

$$\begin{aligned}
\mathbb{P}\left(X_{n,n} < q(\alpha_n)\right) &= F^n(q(\alpha_n)) \\
&= (1-\alpha_n)^n \\
&= \exp\left(n \log\left(1-\alpha_n\right)\right) \\
&= \exp\left(-n\alpha_n(1+o(1))\right) \ \textit{quand } \alpha_n \to 0
\end{aligned}$$

Ainsi, au niveau du dernier terme pour estimer les quantiles extrêmes, deux cas sont envisagés pour $q(\alpha_n)$ (au sein et dehors de l'échantillon) en fonction de la vitesse de convergence de α_n vers zéro:

• Dans la première situation, la suite α_n converge lentement vers zéro ($\alpha_n \to 0$) en ce sens que $n\alpha_n \to c \in [1,\infty]$ lorsque $n \to \infty$. Autrement dit, ceci revient à supposer que le quantile $q(\alpha_n)$ ne tend pas trop vite vers l'infini quand $n \to \infty$ (par exemple le cas où $\alpha_n \geq 1/n$). Dans une telle situation, l'estimation du quantile extrême requiert une interpolation à l'intérieur de l'ensemble des données $(X_1, ..., X_n)$.

• Dans la deuxième situation, on autorise la suite α_n à converger vers zéro plus vite que

[9] Les quantiles extrêmes jouent un rôle important dans le contexte de l'étude du risque où il est crutial d'évaluer adéquatement le risque pour des pertes lourdes d'occurances rares voir par exemple Section 5.

dans la situation précédente $n\alpha_n \to c \in [0,1)$ lorsque $n \to \infty$. Dans une telle situation, l'estimation du quantile $q(\alpha_n)$ peut nécessiter d'extrapoler au-delà des observations $(X_1, ..., X_n)$, par exemple le cas où $\alpha_n < 1/n$. Ce dernier cas peut également être expliqué en utilisant la Figure 1.7 pour $\alpha_n = 1/(n \log n)$. Cette figure révèle que l'information sur la queue de distribution est cruciale pour estimer ces genres de quantiles. Ainsi, l' estimation de la queue de distribution proposée dans la Sous-Section 1.5 nous sera utile dans la suite.

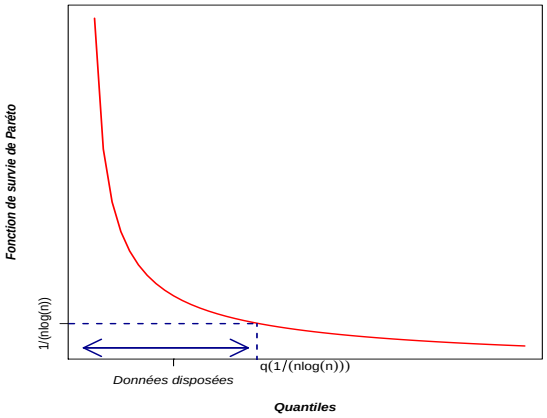

FIG. 1.8 – Explication graphique de quantile extrême.

Dans la sous-section 1.5, nous avons introduit un certain nombre d'estimateurs de l'indice des valeurs extrêmes γ, nous allons donner quelques méthodes d'estimation de quantiles extrêmes basées sur un estimateur d'indice et que nous utiliserons dans les autres chapitres.

1.7.1 Approche basée sur un estimateur d'indices positifs ($\gamma > 0$)

Notons que parmi les estimateurs d'indices des valeurs extrêmes le plus connu est sans conteste l'estimateur de Hill, qui peut être interprété comme un estimateur de la pente du « Paréto Quantile Plot ». Il apparaît donc naturel d'estimer un quantile extrême à l'aide de cet estimateur de γ en extrapolant le long de la droite du « Paréto quantile plot » d'équation

$$y = \log X_{n-k_n,n} + \widehat{\gamma}^H_{k_n} \left(x + \log \frac{k_n + 1}{n + 1} \right),$$

de point d'origine $(\log((n+1)/(k_n+1)), \log X_{n-k_n,n})$, voir par exemple Beirlant *et al.* (1996). Ceci conduit en prenant $x = -\log \alpha_n$ à un estimateur le plus connu d'un quantile

extrême $q(\alpha_n) = F^{\leftarrow}(1-\alpha_n)$ introduit par Weissman (1978) :

$$\boxed{\widehat{q}^{H}_{k_n}(\alpha_n) = X_{n-k_n,n}\left(\frac{k_n+1}{(n+1)\alpha_n}\right)^{\widehat{\gamma}^{H}_{k_n}}.} \tag{1.41}$$

Nous avons déjà vu que l'estimateur de Hill est valable uniquement dans le domaine de Fréchet (indices positifs). Autrement si on se restreint aux fonctions dans ce domaine pour lesquelles on a la caractérisation suivante $\overline{F}(x) = x^{-1/\gamma}\ell_F(x)$, avec ℓ_F étant une fonction à variation lente à l'infini et $\gamma > 0$. Conformément au Lemme 1.4.4, cette caractérisation implique que

$$\begin{aligned}\overline{F}^{\leftarrow}(1-\alpha_n) &= U(1/\alpha_n) = \alpha_n^{-\gamma}\ell_U(1/\alpha_n) \quad \text{avec} \quad \alpha_n \leq 1/n,\\ \overline{F}^{\leftarrow}(1-\beta_n) &= U(1/\beta_n) = \beta_n^{-\gamma}\ell_U(1/\beta_n) \quad \text{avec } \beta_n \geq 1/n,\end{aligned}$$

où $\ell_U(\cdot)$ est une fonction à variation lente à l'infini. Comme $\ell_U(tx)/\ell_U(t) \to 1$ quand $t \to \infty$, alors pour β_n suffisamment petit, on a

$$\overline{F}^{\leftarrow}(\alpha_n) \simeq \overline{F}^{\leftarrow}(\beta_n)\left(\frac{\beta_n}{\alpha_n}\right)^{\gamma}.$$

En remplaçant $\overline{F}^{\leftarrow}(\beta_n)$ par son estimateur naturel $X_{n-[n\beta_n],n}$ et γ par un estimateur quelconque d'indices positifs noté $\widehat{\gamma}^{+}_{k_n}$, on obtient aussi l'estimateur général de type Weissman (1978) défini par

$$\boxed{\widehat{q}^{W}_{k_n}(\alpha_n) = X_{n-[n\beta_n],n}\left(\frac{\beta_n}{\alpha_n}\right)^{\widehat{\gamma}^{+}_{k_n}},} \tag{1.42}$$

où $[x]$ désigne la partie entière de x. Remarquons que pour $\beta_n = (k_n+1)/(n+1)$ et $\widehat{\gamma}^{+}_{k_n} = \widehat{\gamma}^{H}_{k_n}$ on retrouve l'estimateur de (1.41). Les propriétés asymptotiques de l'estimateur de Weissman sont discutées et un intervalle de confiance a été construit sous certaines conditions sur la fonction de distribution F, k_n et α_n dans Embrechts *et al.* (1997), Beirlant et Matthys (2003), Ferreira *et al.* (2003) et Markovich (2005).

1.7.2 Approche basée sur un estimateur d'indices quelconques

Plus généralement, nous introduisons ici une méthode d'estimation du quantile extrême pour $\gamma \in \mathbb{R}$ que nous allons utiliser dans la suite. En utilisant l'estimateur des moments, Dekkers *et al.* (1987) ont proposé d'estimer $q(\alpha_n) = U(1/\alpha_n)$ pour $\alpha_n \to 0$

comme suit:

$$\boxed{\widehat{q}_{k_n}^{\mathcal{M}}(\alpha_n) = X_{n-k_n,n} + \widehat{a}\left(\frac{k_n}{n}\right)\frac{\left(\frac{n}{k_n}\alpha_n\right)^{-\widehat{\gamma}_{k_n}^{\mathcal{M}}} - 1}{\widehat{\gamma}_{k_n}^{\mathcal{M}}},} \tag{1.43}$$

où $\widehat{a}\,(k_n/n) = X_{n-k_n,n}\widehat{\gamma}_{k_n}^{H}\max(1-\widehat{\gamma}_{k_n}^{\mathcal{M}}, 1)$. La normalité asymptotique de cet estimateur a été etablie par Dekkers et de Haan (1989) et par de Haan et Rootzén (1993) sous diverses hypothèses, portant d'une part sur la queue et d'autre part sur α_n.

1.7.3 Approche basée sur la méthode des excès

Cette méthode s'appuie sur le fait que la distribution des excès au-delà d'un seuil élevé u peut être approchée par une loi GPD généralisée : $F_u(x) \approx \mathcal{G}_{\boldsymbol{\gamma},\sigma}(x)$. Soit N_u le nombre d'observations $X_{j_1}, ..., X_{j_{N_u}}$ dépassant u et notons $Y_i = X_{j_i} - u \geq 0$. On ajuste une loi GPD $\mathcal{G}_{\boldsymbol{\gamma},\sigma}$ aux excès $Y_1, ..., Y_{N_u}$ pour obtenir les estimateurs $\widehat{\boldsymbol{\gamma}}$ et $\widehat{\sigma}$ des paramètres de forme et d'échelle. Une application directe du Théorème 1.3.1 montre que $\overline{F}(x) = \overline{F}(u).\overline{F}_u(x-u)$ pour $x > u$, nous pouvons estimer la queue $\overline{F}(u) = 1 - F(u)$ par

$$\boxed{\widehat{\overline{F}}(x) = \frac{N_u}{n}\left(1 + \widehat{\boldsymbol{\gamma}}\frac{x-u}{\widehat{\sigma}}\right)^{-1/\widehat{\gamma}_u}, \quad x \geq u.} \tag{1.44}$$

Par inversion de (1.44), un estimateur du quantile $Q(\alpha_n) = F^{\leftarrow}(1-\alpha_n) > u$ est alors donné par

$$\widehat{q}_{N_u,n}(\alpha_n) = u + \frac{\widehat{\sigma}}{\widehat{\boldsymbol{\gamma}}}\left(\left(\frac{n}{N_u}\alpha_n\right)^{-\widehat{\gamma}_u} - 1\right). \tag{1.45}$$

En pratique, le seuil u est en général la (k_n+1)*ième* plus grande observation; c'est-à-dire $X_{n-k_n,n}$, qui correspond à la statistique d'ordre des observations situées au niveau de la queue de la distribution. On obtient dans ce cas $N_u = k$ et si on note $(\widehat{\boldsymbol{\gamma}}_{k_n}^{POT}, \widehat{\boldsymbol{\sigma}}_{k_n}^{POT})$ les estimateurs $(\widehat{\boldsymbol{\gamma}}_{k_n}^{\mathcal{MOM}}, \widehat{\boldsymbol{\sigma}}_{k_n}^{\mathcal{MOM}})$ où $(\widehat{\boldsymbol{\gamma}}_{k_n}^{\mathcal{PWM}}, \widehat{\boldsymbol{\sigma}}_{k_n}^{\mathcal{PWM}})$ du couple $(\boldsymbol{\gamma}, \boldsymbol{\sigma})$ donnés dans la Sous-Section 1.5.3, on obtient l'estimateur des quantiles extrêmes suivant:

$$\boxed{\widehat{q}_{k_n}^{POT}(\alpha_n) = X_{n-k_n,n} + \frac{\widehat{\boldsymbol{\sigma}}_{k_n}^{POT}}{\widehat{\boldsymbol{\gamma}}_{k_n}^{POT}}\left(\left(\frac{n}{k_n}\alpha_n\right)^{-\widehat{\gamma}_{k_n}^{POT}} - 1\right) \quad \text{pour } \alpha_n < \frac{k_n}{n}.} \tag{1.46}$$

CHAPITRE 2

Processus statistique de paramètre continu généralisant l'estimateur de Hill

Résumé

Dans ce chapitre, nous présentons un processus statistique de paramètre continu généralisant l'estimateur de Hill dont chaque marge caractérise entièrement les extrêmes. La normalité asymptotique de ce type d'estimateur est étudiée et une approximation des domaines de Gumbel et de Fréchet est donnée.

2.1 Contexte d'étude et définitions

2.1.1 Introduction

Soit $(X_1, \ldots, X_n)$, un échantillon de n variables aléatoires réelles indépendantes et de même loi que X avec une fonction de répartition $F(x) = \mathbb{P}(X \leq x)$. Considérons $X_{1,n} \leq \ldots \leq X_{n,n}$, les statistiques d'ordre associées aux observations $X_1, \ldots, X_n$. La théorie des valeurs extrêmes établit que le comportement asymptotique du maximum $X_{n,n} = \max(X_1, ..., X_n)$ normalisé est donné par

$$\mathcal{H}_{\boldsymbol{\gamma}}(x) = \exp\left(-(1+\boldsymbol{\gamma}x)_+\right)^{-1/\boldsymbol{\gamma}}$$

où $y_+ = \max(y, 0)$ et on dit que F appartient au domaine d'attraction de $\mathcal{H}_{\boldsymbol{\gamma}}$, noté $F \in D(\mathcal{H}_{\boldsymbol{\gamma}})$. Le paramètre $\boldsymbol{\gamma} \in \mathbb{R}$ est appelé *indice des valeurs extrêmes* ou *indice de queue*. Selon le signe de $\boldsymbol{\gamma}$, on dira que F appartient au domaine d'attraction de Fréchet

pour $\boldsymbol{\gamma} > 0$ ($F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$), au domaine d'attraction de Gumbel pour $\boldsymbol{\gamma} = 0$ ($F \in D(\boldsymbol{\Lambda})$) et le cas $\boldsymbol{\gamma} < 0$ correspond au domaine de Weibull ($F \in D(\boldsymbol{\psi}_{1/\boldsymbol{\gamma}})$). Un estimateur non-paramétrique très populaire de $\boldsymbol{\gamma}$ dans le cas des distributions de type Paréto (c'est-à-dire correspondant au cas d'indice $\boldsymbol{\gamma} > 0$), est l'estimateur de Hill (1975). Cet estimateur est le plus utilisé en théorie des valeurs extrêmes, et a été largement étudié voir par exemple Hall (1982), Mason (1982), Davis et Resnick (1984), Csörgő *et al.* (1985), Haeusler et Teugels (1985), Lo (1986), Mason et Turova (1994), Resnick et Starica (1995), Diop et Lo (2006), Diop et Lo (2009). On pourra aussi se référer à la section 1.5.1.2 du chapitre 1 pour d'autres estimateurs de l'indice de queue $\boldsymbol{\gamma}$.

Nous considérons dans ce chapitre un processus statistique indexé par un paramètre $\boldsymbol{\tau} > 0$ dont chaque marge peut être considérée comme un estimateur généralisé de Hill et est défini de la manière suivante:

$$\boxed{\boldsymbol{T}_{k_n}(\boldsymbol{\tau}) = k_n^{-\boldsymbol{\tau}} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}} \left(\log X_{n-j+1,n} - \log X_{n-j,n}\right).} \tag{2.1}$$

Notons que k_n est une suite d'entiers correspondant au nombre de statistiques d'ordre supérieur (nombre d'extrêmes) utilisés dans l'estimation. Le processus statistique de Hill généralisé (2.1) a été introduit et étudié par Diop et Lo (2006) et Diop et Lo (2009) pour estimer l'indice des valeurs extrêmes et permet de discriminer entièrement le domaine d'attraction des valeurs extrêmes. Dans toute la suite on suppose que

$$1 \leq k_n < n, \quad k_n/n \to 0, \quad \text{quand} \quad n \to \infty.$$

Si $\boldsymbol{\tau} = 1$, alors $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ correspond à l'estimateur classique de HIll (1975). Toutefois McCulloch (1997) observe que l'estimateur de Hill surestime beaucoup l'indice de queue $\boldsymbol{\gamma}$ quand la loi sous-jacente est stable d'exposant $1.5 < 1/\boldsymbol{\gamma} < 2$. Drees *et al.* (2000) constatent que le graphe de l'estimateur de Hill est un moyen efficace et rapide de trouver la valeur de l'indice de queue si la loi sous-jacente est une loi de Paréto ou ne s'écarte pas trop d'une loi de Paréto. Le paramètre continu $\boldsymbol{\tau}$ introduit par Diop et Lo (2006) a, en fait, permis d'obtenir non seulement la discrimination complète de tous les domaines d'attraction mais aussi d'avoir une souplesse remarquable dans les applications. Diop et Lo (2006) ont montré par simulation que le processus de Hill geńéralisé $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ est plus robuste que l'estimateur classique de Hill selon le choix du paramètre $\boldsymbol{\tau}$ quand la loi sous-jacente s'écarte d'une Paréto ou lorsqu'elle est stable d'exposant $1.5 < 1/\boldsymbol{\gamma} < 2$. de Haan et Resnick (1980) ont proposé un estimateur simple de l'indice de queue noté $DR_n = \boldsymbol{T}_{k_n}(0)/(\log k_n)$. Lo (1986) a montré que si F appartient au domaine d'attraction

de Gumbel $D(\boldsymbol{\Lambda})$ alors l'estimateur DR_n centré et normalisé converge en distribution vers un multiple de Gumbel. Cependant on s'intéresse seulement ici aux valeurs de $\boldsymbol{\tau} > 0$. Comme mentionné dans Diop et Lo (2006), on note dans ce cas que le processus (2.1) donne une famille d'estimateurs de l'indice des valeurs extrêmes $\boldsymbol{\gamma}$.

Cette classe d'estimateurs peut être vue dans sa formulation comme l'estimateur à noyau de Csörgő *et al.* (1985) pour un choix de $K(u) = u^{\boldsymbol{\tau}-1}1_{0<u<1}$. Pour $\boldsymbol{\gamma} > 0$, Csörgő *et al.* (1985) ont établi la consistance faible de cet estimateur sous des hypothèses générales sur le noyau, et la normalité asymptotique sous certaines hypothèses suplémentaires sur la distribution F (voir Csörgő *et al.* (1985)). L'estimateur $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ possède un noyau qui selon la valeur du paramètre $\boldsymbol{\tau}$ ne vérifie pas les hypothèses de Csörgő *et al.* (1985). Par exemple, si $\boldsymbol{\tau} \leq 1/2$ alors $\int u^{-1/2}K(u)du = \infty$. Pour $\boldsymbol{\tau} > 0$, Diop et Lo (2006) ont étudié la consistance forte de cet estimateur pour $\boldsymbol{\gamma} \geq 0$ et ont donné sa variance asymptotique. La normalité de cette classe d'estimateurs $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ a été étudiée seulement dans le cas $\boldsymbol{\tau} > 1/2$ (voir par exemple Csörgő *et al.* (1985) et Diop et Lo (2009)). Il reste maintenant à établir la normalité asymptotique $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ dans le cas où $0 < \boldsymbol{\tau} \leq 1/2$.

Dans ce chapitre, nous allons compléter cette étude, en donnant la normailité asymptotique de $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ pour $0 < \boldsymbol{\tau} \leq 1/2$ et sous l'hypothèse que $F \in D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. Nos résultats sont également un complément à ceux liés à $\boldsymbol{T}_{k_n}(0)$ dans Pickands (1975), de Haan et Resnick (1980) et Lo (1986). Nous nous restreignons seulement au cas $\boldsymbol{\gamma} \geq 0$, puisque le cas $\boldsymbol{\gamma} < 0$, peut être étudié à travers la transformation $F(x_F - 1/\cdot) \in D(\boldsymbol{\phi}_{-1/\boldsymbol{\gamma}})$ (voir la Sous-Section 1.4.3, Théorème 1.4.12), x_F représente le point terminal de F. Cela conduit à remplacer dans (2.1) $X_{n-j+1,n}$ par $x_F - 1/X_{j,n}$. Toutefois, une investigation directe de (2.1) pour $\boldsymbol{\gamma} < 0$ est possible. Cela requiert la théorie des sommes de variables aléatoires dépendantes tandis que le présent travail utilise les résultats sur les sommes de variables aléatoires indépendantes, comme nous le verrons plus loin. Nous considérons par conséquent une manipulation spéciale de cette affaire dans un travail distinct.

2.1.2 Représentation pour $F \in D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$.

Introduisons quelques notations et conventions auxquelles on fera appel dans ce chapitre. Comme nous nous intéressons à la queue de distribution de F et donc nous supposons sans perte de généralité que $X \geq 1$ et considérons l'échantillon $(Y_1, \ldots, Y_n)$ lié aux observations $X_1, \ldots, X_n$ par la transformation $Y_j = \log X_j$, $j = 1, .., n$ et de fonction de répartition $G(y) = \mathbb{P}(Y \leq y) = F(e^y)$, $y \geq 0$. Notons par $Y_{1,n} = \log X_{1,n} \leq ... \leq Y_{n,n} = \log X_{n,n}$, les statistiques d'ordre associées à l'échantillon $(Y_1, \ldots, Y_n)$.

Dans la suite, nous utiliserons la représentation de $Y_1, \ldots, Y_n$ par $G^{\leftarrow}(1-\xi_1), ..., G^{\leftarrow}(1-\xi_n)$ où $G^{\leftarrow}$ est l'inverse généralisée de G et $(\xi_1, ..., \xi_n)$ est un échantillon de n variables aléatoires indépendantes et uniformément distribuées sur $(0,1)$ et notons par $(\xi_{1,n}, ..., \xi_{n,n})$ la statistique d'ordre associée. Nous avons

$$\Big\{Y_j,\ 1 \leq j \leq n\Big\} \stackrel{\mathcal{D}}{=} \Big\{G^{\leftarrow}(1-\xi_j), 1 \leq j \leq n\Big\},$$

et donc

$$\Big\{Y_{j,n},\ j = 1, ..., n\Big\} \stackrel{\mathcal{D}}{=} \Big\{G^{\leftarrow}(1-\xi_{n-j+1,n})\ j = 1, ..., n\Big\}.$$

De plus, nous utiliserons aussi la représentation de Malmquist ci dessous (voir Shorack et Wellner (1986), p. 336)

$$\Big\{j\Big(\log \xi_{j+1,n} - \log \xi_{j,n}\Big), j = 1, ..., n\Big\} \stackrel{\mathcal{D}}{=} \Big\{E_1, ..., E_n\Big\}, \tag{2.2}$$

où $E_1, ..., E_n$ sont des variables aléatoires indépendantes de loi exponentielle standard.

A présent, nous allons rappeler quelques représentations classiques pour qu'une distribution F appartienne aux domaines d'attraction d'une distribution des valeurs extrêmes non dégénérée. Dans la Section 1.4, nous avons donné un grand nombre de propriétés permettant de caractériser l'appartenance aux domaines d'attraction. Nous allons utiliser dans ce chapitre les réprésentations de Karamata et celles proposées par de Haan Ferreira (2006). Pour toute fonction de répartition F, une représentation appropriée d'appartenance au domaine d'attraction d'une distribution des valeurs extrêmes est donnée par $G^{\leftarrow}(1-u) = \log Q(1-u)$, $u \in (0,1)$, où Q et $G^{\leftarrow}$ sont respectivement les inverses généralisées de F et G. Pour toute fonction de répartition H donnée, nous noterons dans toute la suite par (θ_H, x_H), le suport de H avec

$$\theta_H = \inf\{x, H(x) > 0\} \quad \text{et} \quad x_H = \sup\{x, H(x) < 1\}.$$

Nous avons les représentations ci-dessous

1. Représentation de Karamata (KARARE)

Si $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$, alors il est bien connu que $1-F$ est à variations régulières à l'infini d'indice $-1/\boldsymbol{\gamma}$, notée $1-F \in \mathcal{RV}_{-1/\boldsymbol{\gamma}}$ (voir Chapitre 1) et donc $Q(1-\cdot)$ est à variations régulières en zéro d'indice $-\boldsymbol{\gamma}$, notée $Q(1-\cdot) \in \mathcal{RV}^0_{-\boldsymbol{\gamma}}$. D'après la représentation de Karamata, on a

$$Q(1-u) = c_0(1+p(u))u^{-\boldsymbol{\gamma}} \exp\left(\int_u^1 \frac{\Delta(t)}{t} dt\right),\ 0 < u < 1, \tag{2.3}$$

où $p(u)$ et $\Delta(u)$, $u \in (0,1)$ sont deux fonctions mesurables avec $\sup(|p(u)|, |\Delta(u)|) \to 0$ quand $u \to 0$ et c_0 une constante positive. Par conséquent

$$G^{\leftarrow}(1-u) = \log c_0 + \log(1+p(u)) - \boldsymbol{\gamma} \log u + \int_u^1 \frac{\Delta(t)}{t} dt, \ 0 < u < 1, \tag{2.4}$$

avec $G^{\leftarrow}(0) = G^{\leftarrow}(0_+)$.

2. **Represéntation dans le domaine de Gumbel**

D'après Lo (1986), Lemmes 9 et 10, nous avons $F \in D(\boldsymbol{\Lambda})$ si et seulement si $G \in D(\boldsymbol{\Lambda})$ et $R(x,G) \to 0$ quand $x \to x_G$, où

$$R(x,G) = \int_x^{x_G} \frac{1-G(t)}{1-G(x)} dt, \quad \theta_G < x < x_G.$$

Comme $G \in D(\boldsymbol{\Lambda})$, nous avons d'aprés de Haan (1970), Théorème 2.4.1, $G^{\leftarrow}(1-u)$ admet la représenation suivante:

$$G^{\leftarrow}(1-u) = d_0 - s(u) + \int_u^1 s(t)t^{-1} dt, \ 0 < u < 1. \tag{2.5}$$

où d_0 est une constante réelle et $s(\cdot)$ est une fonction à variations lentes en zero ($s(\cdot) \in \mathcal{RV}_0^0$), i.e. $\forall x > 0$, $\lim_{u \to 0} s(ux)/s(u) = 1$. Et donc par la représentation de Karamata, on obtient

$$s(u) = c_0(1+p(u)) \exp\left(\int_u^1 \frac{\Delta(t)}{t} dt\right), \ 0 < u < 1, \tag{2.6}$$

où les fonctions $p(\cdot)$ et $\Delta(\cdot)$ sont les mêmes que celles données dans les représentations (2.3) et (2.4). Notons qu'il est important de remarquer que l'appartenance de la fonction de répartition F au domaine d'attraction $D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$ est associée à un couple de fonction (p, Δ) liant $G(x) = F(e^x)$ dans chaque représentation appropriée.

2.2 Resultats asymptotiques

Dans Diop et Lo (2006), Lemme 1, nous avons $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$ si et seulement si $G \in D(\boldsymbol{\Lambda})$. Ce résultat leur a permis d'établir la consistance forte de l'estimateur généralisé de Hill $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$, $\boldsymbol{\tau} > 0$ et est donnée dans le théorème ci-dessous.

Théorème 2.2.1 *Supposons $F \in D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. Soit $k_n = o(n)$ tel que*

$$\log\log n/k_n \to 0 \quad \textit{quand} \quad n \to \infty,$$

alors pour tout $\boldsymbol{\tau} > 0$, nous avons $\boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau}) \xrightarrow{p.s} \boldsymbol{\gamma}$ quand $n \to \infty$.

La normalité de l'estimateur $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ dans le cas $\boldsymbol{\tau} > 1/2$ a été étudiée dans Csörgő *et al.* (1985) et Diop et Lo (2009). Nous complétons cette étude dans le cas où $\boldsymbol{\tau} \geq 1/2$. Ainsi, pour établir nos résultats, nous allons utiliser quelques notations et hypothèses . Nous considérons les séquences $a_{k_n}(\boldsymbol{\tau})$ et $\sigma^2_{k_n}(\boldsymbol{\tau})$ définies pour $0 < \boldsymbol{\tau} \leq 1/2$ par:

$$a_{k_n}(\boldsymbol{\tau}) = k_n^{-\boldsymbol{\tau}} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1} \quad \text{et} \quad \sigma^2_{k_n}(\boldsymbol{\tau}) = k_n^{-2\boldsymbol{\tau}} \sum_{j=1}^{k_n} j^{2(\boldsymbol{\tau}-1)}. \tag{2.7}$$

Soit $\lambda > 1$ un nombre réel quelconque. Pour établir la normalité asymptotique de l'estimateur $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$, nous allons considérer les trois hypothèses ci-dessous:

$$\textbf{(C1)} \qquad g_{1,n}(p,\lambda)\Big(k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})\Big)^{-1} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}} \to 0 \text{ quand } \; n \to +\infty, \tag{2.8}$$

$$\text{avec } 0 \leq g_{1,n}(p,\lambda) = \sup_{0 \leq u \leq \lambda k_n/n} |p(u)|\,.$$

$$\textbf{(C2)} \qquad g_{2,n}(p,\lambda)\Big(k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})\Big)^{-1} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1} \to 0 \text{ quand } \; n \to +\infty, \tag{2.9}$$

$$\text{avec } g_{2,n}(\Delta,\lambda) = \sup_{0 \leq u \leq \lambda k_n/n} |\Delta(u)|\,.$$

$$\textbf{(C3)} \qquad d_n(p,\Delta,\lambda)\Big(k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})\Big)^{-1} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1} \to 0 \text{ quand } \; n \to +\infty, \tag{2.10}$$

$$\text{avec } d_n(p,\Delta,\lambda) = \max(g_{1,n}(p,\lambda), g_{2,n}(\Delta,\lambda)\log k_n) \to 0 \; \textit{quand } n \to \infty.$$

Introduisons la suite de fonction $\zeta_{k_n}(s)$, et la variable aléatoire centrée $\mathcal{L}_{k_n}(\boldsymbol{\tau})$, $0 < \boldsymbol{\tau} < 1/2$, définies par:

$$\zeta_{k_n}(s) = \sum_{j=1}^{k_n} j^{-s}, \; s > 1 \quad \text{et} \quad \mathcal{L}_{k_n}(\boldsymbol{\tau}) = \frac{1}{\sqrt{\zeta_{k_n}(2(1-\boldsymbol{\tau}))}} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(E_j - 1). \tag{2.11}$$

Considérons la fonction de Riemann $\zeta(s)$ et la variable aléatoire centrée réduite $\mathcal{L}(\boldsymbol{\tau})$ définies par :

$$\zeta(s) = \lim_{n\to\infty} \zeta_{k_n}(s) = \sum_{j=1}^{\infty} j^{-s}, \; s > 1 \quad \text{et} \quad \mathcal{L}(\boldsymbol{\tau}) = \frac{1}{\sqrt{\zeta(2(1-\boldsymbol{\tau}))}} \sum_{j=1}^{\infty} j^{\boldsymbol{\tau}-1}(E_j - 1).$$

Il est claire que $\mathcal{L}_{k_n}(\boldsymbol{\tau}) \xrightarrow{\mathbb{P}} \mathcal{L}(\boldsymbol{\tau})$, quand $n \to \infty$.

Dans toute la suite, sauf une mention partiulière les limites étudiées sont données lorsque $n \to \infty$. Nous sommes maintenant en mesure d'énoncer nos résultats sur la classe d'estimateurs étudiée.

Théorème 2.2.2

- *Soit $F \in D(\boldsymbol{\Lambda})$. Supposons que la condition* (**C3**) *soit vérifiée pour $\boldsymbol{\tau} = 1/2$, alors*

$$\frac{a_{k_n}(\boldsymbol{\tau})}{s(k_n/n)\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - s(k_n/n)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

et pour $0 < \boldsymbol{\tau} < 1/2$,

$$\frac{a_{k_n}(\boldsymbol{\tau})}{s(k_n/n)\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - s(k_n/n)\right) \xrightarrow{\mathcal{D}} \mathcal{L}(\boldsymbol{\tau}).$$

- *Soit $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. Supposons que les conditions* (**C1**) *et* (**C2**) *soient vérifiées pour $\boldsymbol{\tau} = 1/2$, alors*

$$\frac{a_{k_n}(\boldsymbol{\tau})}{\boldsymbol{\gamma}\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - \boldsymbol{\gamma}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

et pour $0 < \boldsymbol{\tau} < 1/2$

$$\frac{a_{k_n}(\boldsymbol{\tau})}{\boldsymbol{\gamma}\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - \boldsymbol{\gamma}\right) \xrightarrow{\mathcal{D}} \mathcal{L}(\boldsymbol{\tau}).$$

La preuve de ce théoréme est détaillée dans la Sous-section 2.6.2.

Commentaires: A présent, nous alllons donner quelques commentaires sur nos résultats. Ainsi, nous justifions que les vitesses de convergence utilisées dans le Théorème 2.2.2 tendent vers infini. Il suffit de remarquer que pour tout $\boldsymbol{\gamma} \geq 0$ et pour tout $\boldsymbol{\tau} \in]0, 1/2]$,

$$a_{k_n}(\boldsymbol{\tau})\Big(\max\left(s(k_n/n), \boldsymbol{\gamma}\right)\sigma_{k_n}(\boldsymbol{\tau})\Big)^{-1} \to \infty, \text{ quand } n \to \infty.$$

En effet, par les expréssions (**S2**) et (**S3**) dans le Lemme 2.6.1 de la Sous-section 2.6.2 ci-dessous, nous avons pour tout $\boldsymbol{\tau} \in]0, 1/2]$,

$$a_{k_n}(\boldsymbol{\tau}) = \frac{1}{\boldsymbol{\tau}}\left(1 + o(1)\right), \text{ quand } n \to \infty. \tag{2.12}$$

Et par les expressions (**S1**) et (**S4**) du Lemme 2.6.1, nous avons aussi pour $n \to \infty$,

$$\sigma^2_{k_n}(\boldsymbol{\tau}) = \begin{cases} \dfrac{\log k_n}{k}\left(1 + O(1/\log k_n)\right), & \text{pour} \quad \boldsymbol{\tau} = 1/2, \\ \zeta\left(2(1-\boldsymbol{\tau})\right) k_n^{-2\boldsymbol{\tau}}\left(1 + o(1)\right), & \text{pour} \quad 0 < \boldsymbol{\tau} < 1/2. \end{cases} \tag{2.13}$$

Il est facile de voir que $\zeta\left(2(1-\boldsymbol{\tau})\right) < \infty$, pour tout $0 < \boldsymbol{\tau} < 1/2$. Nous allons maintenant évaluer la limite de la suite $s(k_n/n)$. Deheuvels et al. (1989) ont montré que

la représentation 2.5 est équivalente à $G \in D(\boldsymbol{\Lambda})$. Et selon Lo (1986), Lemme 4, $s(u) \sim R(G^{\leftarrow}(1-u), G)$, $u \to 0$. Ce qui nous permet d'avoir $s(k_n/n) \sim R(G^{\leftarrow}(1-k_n/n), G)$, pour n assez grand et en remarquant que $x_G = G^{\leftarrow}(1)$, on voit facilement que $s(k_n/n) \to 0$ par les représentations ci-dessus. Et finalement pour tout $\boldsymbol{\tau} \in]0, 1/2]$, et $\boldsymbol{\gamma} \geq 0$, nous concluons que $a_{k_n}(\boldsymbol{\tau})\Big(\max\left(s(k_n/n), \boldsymbol{\gamma}\right)\sigma_{k_n}(\boldsymbol{\tau})\Big)^{-1} \to \infty$, quand $n \to \infty$.

Remarque 2.2.3 *Notons que la suite $s(k_n/n) \to 0$, joue un rôle très important dans le cas du domaine d'attraction de Gumbel ($D(\boldsymbol{\Lambda})$), on pourra se référer dans Lo (1986), elle nous permet d'établir la normalité asymptotique de $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})/a_{k_n}(\boldsymbol{\tau})$, lorsque $\boldsymbol{\gamma} = 0$.*

Remarque 2.2.4 *La méthode utilisée ici peut être appliquée pour $0 < \boldsymbol{\tau} < 1$. Par contre, on ne peut pas étendre ces résultats à des paramètres $\boldsymbol{\tau} \in]1/2, 1[$ puisque pour ce cas, $k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau}) \to 0$, alors que nos résultats sont basés sur l'hypothèse $k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau}) \to C \in]0, \infty]$.*

2.3 Applications et Remarques

Dans cette partie nous illustrons quelques applications statistiques sur nos résultats.

2.3.1 Estimateurs asymptotiquement Gaussiens

Soit $F \in D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$, rappelons que $\boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ est un estimateur de $\boldsymbol{\gamma}$ consistant pour $\boldsymbol{\tau} > 0$, et est asymptotiquement normal lorsque $\boldsymbol{\tau} > 1/2$ (voir Csörgő *et al.* (1985) pour $\boldsymbol{\gamma} > 0$, Diop et Lo (2006) et Diop et Lo (2009) pour $\boldsymbol{\gamma} \geq 0$). Notre présent résultat étend celui obtenu pour $\boldsymbol{\tau} = 1/2$ et montre que cette dernière valeur est cruciale puisque le corollaire 2.3.2 ci-dessous donne la normalité asymptotique de l'estimateur $\boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$, $\boldsymbol{\tau} = 1/2$. Maintenant, nous allons dans cette sous-section particulariser nos résultats pour ce cas puis nous illustrons les conditions (**C1**) et (**C2**) et (**C3**).

Corollaire 2.3.1

- *Soit $F \in D(\boldsymbol{\Lambda})$. Supposons que la condition (**C3**) soit vérifiée pour $\boldsymbol{\tau} = 1/2$, alors*

$$\sqrt{\frac{4k_n}{s^2(k_n/n)\log k_n}}\left(\frac{\boldsymbol{T}_{k_n}(1/2)}{a_{k_n}(1/2)} - s(k_n/n)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{2.14}$$

- *Soit $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. Supposons que (**C1**) et (**C2**) soient vérifiées pour $\boldsymbol{\tau} = 1/2$, alors*

$$\sqrt{\frac{4k_n}{\boldsymbol{\gamma}^2\log k_n}}\left(\frac{\boldsymbol{T}_{k_n}(1/2)}{a_{k_n}(1/2)} - \boldsymbol{\gamma}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{2.15}$$

Ainsi $\boldsymbol{T}_{k_n}(1/2)/a_{k_n}(1/2)$ est un nouvel estimateur de l'indice des valeurs extrêmes $\boldsymbol{\gamma} \geq 0$ et est asymptotiquement normal. Par (**S2**) dans le Lemme 2.6.1, $a_{k_n}(1/2) = 2(1 + O(k_n^{-1/2}))$ et donc $a_{k_n}(1/2) \to 2$, nous retrouvons le résultat général de Diop et Lo (2006): $\boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau}) \to \boldsymbol{\gamma}$, et nous avons le corollaire ci-dessous.

Corollaire 2.3.2

- *Soit $F \in D(\boldsymbol{\Lambda})$. Supposons que la condition (**C3**) soit vérifiée pour $\boldsymbol{\tau} = 1/2$, alors*

$$\sqrt{\frac{4k_n}{s^2(k_n/n)\log k_n}}\left(\frac{1}{2}\boldsymbol{T}_{k_n}(1/2) - s(k_n/n)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

- *Soit $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. supposons que (**C1**) et (**C2**) soient vérifiées pour $\boldsymbol{\tau} = 1/2$, alors*

$$\sqrt{\frac{4k_n}{\boldsymbol{\gamma}^2\log k_n}}\left(\frac{1}{2}\boldsymbol{T}_{k_n}(1/2) - \boldsymbol{\gamma}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

D'après les corollaires 2.3.1 et 2.3.2, nous pouvons formuler une famille d'estimateurs de $\boldsymbol{\gamma}$ par

$$\boldsymbol{T}_{k_n}^{(a)}(1/2) = \Big(2a + (1-a)/a_{k_n}(1/2)\Big)\boldsymbol{T}_{k_n}(1/2)\,, \quad \text{pour } 0 \leq a \leq 1 \tag{2.16}$$

et qui est asymptotiquement normal. Le paramètre « a », qui est choisi de manière arbitraire dans $[0,1]$, peut être utile dans les applications.

2.3.2 Illustration des Conditions (C1), (C2) et (C3) pour $\boldsymbol{\tau} = 1/2$

Pour $\boldsymbol{\tau} = 1/2$, par le Lemme 2.6.1 la condition (**C1**) devient

$$g_{1,n}(p,\lambda)(\log k_n)^{-1/2}\sum_{j=1}^{k_n} j^{1/2} \to 0.$$

Mais dans beaucoup de situations, on suppose que la fonction de répartition F admet une dérivée pour que nous puissions prendre $p(u) = 0$, ce qui rend (**C1**) inutile. En effet, supposons que $F(x)$ admet une dérivée $F'(x)$ pour de grandes valeurs de x. Une condition pour que F appartienne au domaine d'attraction $D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$, est:

$$\lim_{x\to\infty}\frac{1-F(x)}{xF'x)} = \frac{1}{\boldsymbol{\gamma}}. \tag{2.17}$$

(voir de Haan Ferreira (2006), p. 17, Theorem 1.1.11). Ceci implique que

$$\Delta(u) = u(G^{\leftarrow}(1-u))' + \boldsymbol{\gamma} = uG'(G^{\leftarrow}(1-u))^{-1} + \boldsymbol{\gamma} \to 0 \text{ quand } n \to \infty$$

et ensuite, pour une constante $c > 0$ et $u_0 \in (0,1)$,

$$G^{\leftarrow}(1-u) = c - \boldsymbol{\gamma} \log u + \int_u^{u_0} \frac{\Delta(t)}{t} dt.$$

Remarque 2.3.3 *Pour une telle fonction de répartition, la condition (**C1**) est effectivement inutile. La condition (**C2**) est plus consistante. Cependant, elle est quelque part liée avec la « condition du second ordre » de l'équation (1.32) (voir de Haan Ferreira (2006), p.43 pour plus de détails) et pour un modèle populaire utilisé dans la littérature.*

Par exemple pour la classe de Hall: $1 - F(x) = c_1 x^{-1/\boldsymbol{\gamma}}(1 + d_1 x^{-\beta} + o(x^{-\beta}))$, où c_1 et d_1 sont des constantes positives. Nous avons $F^{-1}(1-u) = c_2 u^{-\boldsymbol{\gamma}}(1 + d_2 u^{\beta\boldsymbol{\gamma}} + o(u^{\beta\boldsymbol{\gamma}}))$ avec c_2 et d_2 des constantes données. Et on obtient

$$G^{\leftarrow}(1-u) = \log c_2 - \boldsymbol{\gamma} \log u + O(u^{\beta\boldsymbol{\gamma}}),$$

qui peut être réécrite, avec $\Delta(t) = u^{\beta\boldsymbol{\gamma}}$ et $C(u) = O(1)$ quand $u \to 0$, par

$$G^{\leftarrow}(1-u) = d_2 - \boldsymbol{\gamma} \log u + C(u) \int_u^1 \frac{\Delta(t)}{t} dt.$$

Supposons maintenant que $C(u) = D$ une constante donnée ou bien $C(u) \displaystyle\int_u^1 \frac{\Delta(t)}{t} dt$ est obtenu à partir de $\displaystyle\int_u^1 \frac{C_1(t)\Delta(t)}{t} dt$ par le théorème de l'intégrale moyenne avec une fonction bornée $C_1(\cdot)$. Alors la condition (**C2**) devient

$$k_n^{\beta\boldsymbol{\gamma}+1/2} (\log k_n)^{-1/2} / n^{\beta\boldsymbol{\gamma}} \to 0,$$

et la condition (**C3**) devient

$$k_n^{\beta\boldsymbol{\gamma}+1/2} (\log k_n)^{1/2} / n^{\beta\boldsymbol{\gamma}} \to 0.$$

On conclut que pour $\boldsymbol{\tau} = 1/2$, $\boldsymbol{T}_{k_n}(1/2)/a_{k_n}(1/2)$ est un estimateur asymptotiquement normal de $\boldsymbol{\gamma}$ sous des conditions raisonnables basées sur la régularité de la fonction de répartition F. Des études de simulations et d'application sur des données réelles dans les sections 2.4 et 2.5 suggèrent une bonne performance de l'estimateur.

2.3.3 Test statistique sur le paramètre γ pour $\tau = 1/2$

Après avoir estimé le paramètre $\boldsymbol{\gamma}$, il est important de tester sa significativité. Ainsi, nous allons montrer comment établir un test d'hypothèse pour que $F \in D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$ sur l'estimateur $\boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau}) \to \boldsymbol{\gamma}$, $\boldsymbol{\tau} = 1/2$. Soit $F \in D(\boldsymbol{\Lambda})$ ($\boldsymbol{\gamma} = 0$), par le Corollaire 2.3.1, nous avons sous la condition **C3** que

$$\sqrt{\frac{4k_n}{s^2(k_n/n)\log k_n}}\left(\frac{\boldsymbol{T}_{k_n}(1/2)}{a_{k_n}(1/2)} - s(k_n/n)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Notons qu'en pratique on ne connait pas $s(k/n)$, puisque ce dernier dépend de la fonction de répartition de G liée à celle de F supposée iconnue. Nous allons donc trouver une approximation de $s(k/n)$ pour enfin pouvoir établir une procedure de test sur $\boldsymbol{\gamma}$. D'après Lo (1989), Lemme 4, nous avons:

$$\forall \lambda > 1,\ \left(G^{\leftarrow}(1-u) - G^{\leftarrow}(1-\lambda u)\right)/s(u) \to \log\lambda \ \text{ quand } u \to 0.$$

Comme $G^{\leftarrow}(1-u) = \log Q(1-u)$, posons $u = k_n/n$ en remplaçant $Q(1-k_n/n)$ et $Q(1-\lambda k_n/n)$ par leurs versions empiriques respectives $Q_n(1-k_n/n) = X_{n-k_n,n}$ et $Q_n(1-\lambda k_n/n) = X_{n-[\lambda k_n],n}$, où $[x]$ désigne la partie entière de x, nous pouvons approximer $s(k_n/n)$ par

$$\widehat{s}_\lambda(k_n/n) = \frac{1}{\log\lambda}\left(\log X_{n-k_n,n} - \log X_{n-[\lambda k_n],n}\right). \tag{2.18}$$

Nous sommes en mesure d'établir le corollaire ci-dessous.

Corollaire 2.3.4 *Soit $F \in D(\boldsymbol{\Lambda})$. Supposons que la condition (**C3**) soit vérifiée pour $\boldsymbol{\tau} = 1/2$, alors, $\forall\lambda > 0$ nous avons*

$$\sqrt{\frac{4k_n}{\widehat{s}^2_\lambda(k_n/n)\log k_n}}\left(\frac{\boldsymbol{T}_{k_n}(1/2)}{a_{k_n}(1/2)} - \widehat{s}_\lambda(k_n/n)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{2.19}$$

Nous allons maintenant décrire brièvement la procédure de test. On veut tester au niveau de significativité de 5% l'hypothèse selon la quelle $F \in D(\boldsymbol{\Lambda})$ (cas $\boldsymbol{\gamma} = 0$) contre celle $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. On pose donc des hypothèses de test unilatéral : $\boldsymbol{H}_0 : \boldsymbol{\gamma} = 0$ contre $\boldsymbol{H}_1 : \boldsymbol{\gamma} > 0$. La statistique de test est l'estimateur de $\boldsymbol{\gamma} = 0$ suivant:

$$\mathcal{ST}_{k_n,\lambda} = \boldsymbol{T}_{k_n}(1/2)/a_{k_n}(1/2) - \widehat{s}_\lambda(k_n/n).$$

Intuitivement, on comprend bien qu'on va rejeter $\boldsymbol{H}_0$ si $\mathcal{ST}_{k_n,\lambda}$ est trop grand en valeur absolue c'est-à-dire si $\boldsymbol{T}_{k_n}(1/2)/a_{k_n}(1/2)$ est trop éloigné de $\widehat{s}_\lambda(k_n/n)$ sous $\boldsymbol{H}_0$. D'autre

part, en utilisant la propriété asymptotique (2.19), et d'après la remarque faite plus haut, on rejette $\boldsymbol{H}_0$ si

$$\left|\sqrt{\frac{4k_n}{\widehat{s}_\lambda^2(k_n/n)\log k_n}}\mathcal{ST}_{k_n,\lambda}\right| > z_0.$$

Pour construire la région de rejet de $\boldsymbol{H}_0$, on cherche donc z_0 tel que

$$\mathbb{P}\left(\left|\sqrt{\frac{4k_n}{\widehat{s}_\lambda^2(k_n/n)\log k_n}}\,\mathcal{ST}_{k_n,\lambda}\right| > z_0\right) = \alpha.$$

Et par symétrie de la loi normale centrée réduite $F_{\mathcal{N}(0,1)}$, z_0 est tel que $1 - F_{\mathcal{N}(0,1)}(z_0) = \alpha/2$, ce qui s'écrit encore $F^{-1}_{\mathcal{N}(0,1)}(1-\alpha/2) = z_0$. D'après la table de la fonction de répartition inverse de la loi normale, on en déduit que $z_0 = 1.96$ pour $\alpha = 0.05$. Finalement, on rejette donc $\boldsymbol{H}_0$ si

$$|\mathcal{ST}_{k_n,\lambda}| > 1.96\sqrt{\frac{\widehat{s}_\lambda^2(k_n/n)\log k_n}{4k_n}}.$$

Dans la suite, nous utiliserons une valeur $\lambda = 1.5$ pour évaluer l'expression de $\widehat{s}_\lambda(k_n/n)$.

2.3.4 Estimateurs asymptotiquement non Gaussiens

Pour $0 < \boldsymbol{\tau} < 1/2$, nous pouvons aussi établir le corollaire suivant:

Corollaire 2.3.5

- *Soit $F \in D(\boldsymbol{\Lambda})$. Supposons que la condition* (**C3**) *soit vérifiée, alors*

$$\frac{k_n^{\boldsymbol{\tau}}}{\boldsymbol{\tau} s(k_n/n)\sqrt{\zeta(2(1-\boldsymbol{\tau}))}}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - s(k_n/n)\right) \xrightarrow{\mathcal{D}} \mathcal{L}(\boldsymbol{\tau}), \tag{2.20}$$

- *Soit $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. Supposons que les conditions* (**C1**) *et* (**C2**) *soient vérifiées, alors*

$$\frac{k_n^{\boldsymbol{\tau}}}{\boldsymbol{\tau}\boldsymbol{\gamma}\sqrt{\zeta(2(1-\boldsymbol{\tau}))}}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - \boldsymbol{\gamma}\right) \xrightarrow{\mathcal{D}} \mathcal{L}(\boldsymbol{\tau}). \tag{2.21}$$

$\boldsymbol{T}_{k_n}(\boldsymbol{\tau})/a_{k_n}(\boldsymbol{\tau}), 0 < \boldsymbol{\tau} < 1/2$ est une classe d'estimateurs de l'indice des valeurs extrêmes $\boldsymbol{\gamma} \geq 0$ et est asymptotiquement normal. Par (**S3**) dans le Lemme 2.6.1, $a_{k_n}(\boldsymbol{\tau}) = \frac{1}{\boldsymbol{\tau}}(1 + o(1))$, nous retrouvons le résultat général de Diop et Lo (2006): $\boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau}) \to \boldsymbol{\gamma}$, et nous pouvons simplifier les résultats du Corollaire 2.3.5 en remplaçant a_{k_n} par $1/\boldsymbol{\tau}$.

Ces résultats établissent une connexion entre la théorie des valeurs extrêmes et la théorie des nombres, du fait qu'il dépend fortement de la fonction de Riemann qui est en relation avec les nombres premiers. Soit $s > 1$, nous avons d'après Valiron (1990), p.506,

$$\zeta(s) = \prod_{p=2}^{\infty}(1-p^{-s})^{-1} = \frac{s}{s-1} + \frac{1}{s}\int_1^\infty \frac{x-[x]}{x^{s+1}}dx,$$

où le produit s'applique sur les nombres premiers $p \geq 2$ et $[x]$ est la valeur absolue de x. La loi limite $\mathcal{L}(\boldsymbol{\tau})$ est caraterisée par la fonction caractéristique suivante

$$\varphi(t) = \exp\left(\sum_{n=2}^{\infty} \frac{(it)^n}{n} \zeta\Big(n(1-\boldsymbol{\tau})\Big)\zeta\Big(2(1-\boldsymbol{\tau})\Big)^{-n/2}\right),$$

calculée et justifiée dans (2.23). La variable aléatoire $\mathcal{L}(\boldsymbol{\tau})$ ayant tous ses moments finis est liée avec la fonction de Riemann. En effet, posons $\varphi_\infty(t) = \exp(\widetilde{\varphi}(t))$ avec

$$\widetilde{\varphi}(t) = \sum_{n=2}^{\infty} \frac{(it)^n}{n} \zeta\Big(n(1-\boldsymbol{\tau})\Big)\zeta\Big(2(1-\boldsymbol{\tau})\Big)^{-n/2},$$

et soit $\widetilde{\varphi}^{(r)}$ la r^{ieme} dérivée de la fonction $\widetilde{\varphi}$. Nous avons facilement

$$\widetilde{\varphi}^{(1)}(t) = \sum_{n=2}^{\infty} i(it)^{n-1} \zeta\Big(n(1-\boldsymbol{\tau})\Big)\zeta\Big(2(1-\boldsymbol{\tau})\Big)^{-n/2},$$

et

$$\widetilde{\varphi}^{(2)}(t) = -\sum_{n=2}^{\infty} (n-1)(it)^{n-2} \zeta\Big(n(1-\boldsymbol{\tau})\Big)\zeta\Big(2(1-\boldsymbol{\tau})\Big)^{-n/2}.$$

Par suite,

$$\varphi^{(1)}(t) = \widetilde{\varphi}^{(1)}(t)\exp(\widetilde{\varphi}(t))$$

et

$$\varphi^{(2)}(t) = \Big(\big(\widetilde{\varphi}^{(1)}(t)\big)^2 + \widetilde{\varphi}^{(2)}(t)\Big)\exp(\widetilde{\varphi}(t)).$$

Ainsi,

$$\widetilde{\varphi}^{(1)}(0) = 0 \quad \text{et} \quad \widetilde{\varphi}^{(2)}(0) = -\left[(n-1)(it)^{n-2}\zeta(n(1-\boldsymbol{\tau}))\zeta(2(1-\boldsymbol{\tau}))^{-n/2})\right]_{n=2} = -1$$

et donc $\mathbb{E}\left(\mathcal{L}(\boldsymbol{\tau})\right) = 0$ et $\mathbb{E}\left(\mathcal{L}^2(\boldsymbol{\tau})\right) = 1$. On peut utiliser la même procédure pour obtenir $\mathbb{E}\left(\mathcal{L}^3(\boldsymbol{\tau})\right) = 2\zeta(3(1-\boldsymbol{\tau}))/\zeta(2(1-\boldsymbol{\tau}))^{3/2}$ et $\mathbb{E}\left(\mathcal{L}^4(\boldsymbol{\tau})\right) = 3 + 6\zeta(4(1-\boldsymbol{\tau}))/\zeta(2(1-\boldsymbol{\tau}))^2$, ainsi que les moments d'ordre élevé. Pour ces valeurs de $\boldsymbol{\tau}$, les conditions (**C2**) et (**C3**) deviennent respectivement,

$$g_{2,n}(p,\lambda) \to 0 \quad \text{et} \quad \max(g_{1,n}(p,\lambda), g_{2,n}(\Delta,\lambda)) \to 0.$$

Une chose intéressante est le fait que la condition (**C2**) est toujours vérifiée pour toutes les fonctions de distribution $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. Ceci est un résultat très rare en théorie des valeurs extrêmes. Pour $F \in D(\boldsymbol{\Lambda})$, la condition (**C3**) est aussi très faible puisque les expressions $g_{1,n}$ et $g_{2,n}$ sont seulement données de telles sorte qu'elles tendent plus vite vers zéro que $\log k_n$. Nous obtenons alors une loi asymptotique très générale qui semble être une compensation de l'absence de normalité asymptotique.

2.3.4.1 Illustration de la fonction de répartition de $\mathcal{L}(\boldsymbol{\tau})$

Etant donné que ces limites semblent nouvelles dans l'estimation de l'indice des valeurs extrêmes, il semble raisonnable de donner quelques commentaires sur la façon de les utiliser. Bien que nous ne pouvons pas donner une expression simple de la fonction de répartition de $\mathcal{L}(\boldsymbol{\tau})$ notée $F_{\mathcal{L}(\tau)}(u) = \mathbb{P}(\mathcal{L}(\boldsymbol{\tau}) \leq u)$, mais nous pouvons illustrer numériquement sa fonction de répartition empirique $\widehat{F}_{\mathcal{L}(\tau)}(x) = B^{-1} \sum_{m=1}^{B} \mathbb{1}_{\{\mathcal{L}(\boldsymbol{\tau},m) \leq x\}}$, où les $\mathcal{L}(\boldsymbol{\tau}, m),\ m = 1, ..., B$ représentent B observations de la variable aléatoire $\mathcal{L}(\boldsymbol{\tau})$. Ainsi, pour $\boldsymbol{\tau} \in]0, 1/2[$ fixé, nous allons montrer comment générer un échantillon de ces variables aléatoires puis établir sa distribution empirique. Considérons par exemple pour chaque $m = 1, ..., B = 1000$, un échantillon $E_1(m), ..., E_k(m)$ de loi exponentielle standard avec $k \in (100, 250, 500, 750, 1000, 2000, 2500, 5000)$. Et évaluons la variable aléatoire $\mathcal{L}_k(\boldsymbol{\tau})$ notée $\mathcal{L}_k(\boldsymbol{\tau}, m)$ suivante pour $\boldsymbol{\tau} \in]0, 1/2[$:

$$\mathcal{L}_k(\boldsymbol{\tau}, m) = \frac{1}{\sqrt{\zeta_k(2(1-\boldsymbol{\tau}))}} \sum_{j=1}^{k} j^{\boldsymbol{\tau}-1}(E_j(m) - 1),\ m = 1, ..., B$$

où $\zeta_k(s) = \sum_{j=1}^{k} j^{-s},\ s > 1$. Nous pouvons observer sur la Figure 2.1 que les fonctions de distributions empiriques $\widehat{F}_{\mathcal{L}_k(\boldsymbol{\tau})}$ sont très proches l'une de l' autre pour différentes valeurs de $\boldsymbol{\tau} \in \{0.1, 0.2, 0.3, 0.4\}$ et deviennent stables dans le sens où elles n'ont pas une différence significative pour $k \geq 500$. Et donc nous pouvons prendre $F_{\mathcal{L}_k(\tau)}(u) = \mathbb{P}(\mathcal{L}_k(\boldsymbol{\tau}) \leq u)$ comme étant une bonne approximation de $F_{\mathcal{L}(\tau)}(u) = \mathbb{P}(\mathcal{L}(\boldsymbol{\tau}) \leq u)$ pour k assez grand. De là, nous serons en mesure de calculer les probabilités et les quantiles de $\mathcal{L}(\boldsymbol{\tau})$, $0 < \boldsymbol{\tau} < 1/2$ et les résultats sont représentés par le Tableau 2.1. Des tests statistiques peuvent alors être basés sur ces résultats.

$\boldsymbol{\tau}$	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.48
$F_{\boldsymbol{\tau}}(-1.96)$	0,0040	0,0035	0,0060	0,0070	0,0090	0,0137	0,0160	0,0166	0,0180
$F_{\boldsymbol{\tau}}(0)$	0,5572	0,5570	0,5490	0,5440	0,5380	0,5400	0,5240	0,5010	0,5240
$F_{\boldsymbol{\tau}}(1.96)$	0,9530	0,9558	0,9545	0,9623	0,9610	0,9650	0,9710	0,9842	0,9660

TAB. 2.1 – Quelques probabilités et quantiles de la loi limite $\mathcal{L}(\boldsymbol{\tau})$ pour différentes valeurs de τ.

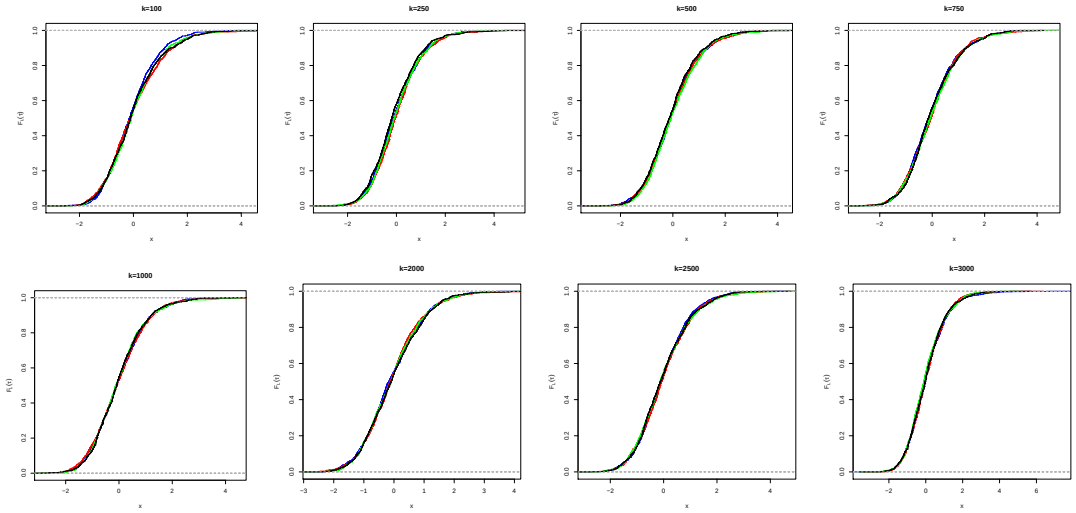

FIG. 2.1 – Fonction de distribution empirique de $\mathcal{L}(\boldsymbol{\tau})$ pour différentes valeurs de τ.

2.4 Etude de simulation

Etant données n réalisations $(x_1, ..., x_n)$ d'un échantillon $(X_1, ..., X_n)$ d'une variable aléatoire X vérifiant les hypothèses nécessaires données dans la Section 2.1.2, nous souhaitons dans cette partie illustrer la performance de l'estimateur $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})/a_{k_n}(\boldsymbol{\tau}) \approx \boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$, pour $\boldsymbol{\tau} \in (0, 1/2]$, en le comparant avec l'estimateur de Hill $\widehat{\boldsymbol{\gamma}}^H_{k_n}$ et celui de Pickands $\widehat{\boldsymbol{\gamma}}^P_{k_n}$. Pour simplifier, nous noterons dans toute la suite $\boldsymbol{T}^*_{k_n}(\boldsymbol{\tau}) := \boldsymbol{T}_{k_n}(\boldsymbol{\tau})/a_{k_n}(\boldsymbol{\tau}) \approx \boldsymbol{\tau}\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$, $0 < \boldsymbol{\tau} \leq 1/2$. La variance asymptotique de notre classe d'estimateurs $\boldsymbol{T}^*_{k_n}(\boldsymbol{\tau})$, $0 < \boldsymbol{\tau} \leq 1/2$ est donnée pour tout $\boldsymbol{\gamma} \geq 0$ par:

$$\mathcal{AV}(\omega^{(\boldsymbol{\gamma})}_{k_n}, \boldsymbol{\tau}) = \begin{cases} \dfrac{\omega^{(\boldsymbol{\gamma})}_{k_n} \log k_n}{4k_n} & \text{pour} \quad \boldsymbol{\tau} = 1/2, \\ \dfrac{\boldsymbol{\tau}^2 \omega^{(\boldsymbol{\gamma})}_{k_n} \zeta\left(2(1-\boldsymbol{\tau})\right)}{k_n^{2\boldsymbol{\tau}}} & \text{pour} \quad 0 < \boldsymbol{\tau} < 1/2, \end{cases} \tag{2.22}$$

avec $\omega^{(\boldsymbol{\gamma})}_{k_n}$ donné pour tout $\boldsymbol{\gamma} \geq 0$ par:

$$\omega^{(\boldsymbol{\gamma})}_{k_n} = \begin{cases} s^2(k_n/n) & \text{si} \quad \boldsymbol{\gamma} = 0, \\ \boldsymbol{\gamma}^2 & \text{si} \quad \boldsymbol{\gamma} > 0. \end{cases}$$

Nous représentons sur la Figure 2.4, l'expréssion de $\mathcal{AV}(\omega^{(\boldsymbol{\gamma})}_{k_n}, \boldsymbol{\tau})/\omega^{(\boldsymbol{\gamma})}_{k_n}$ en fonction de k_n pour différentes valeurs de $\boldsymbol{\tau} \in]0, 1/2]$, dans le but d'illustrer la variance asymptotique de notre classe d'estimateurs $\boldsymbol{T}^*_{k_n}(\boldsymbol{\tau})$, $\boldsymbol{\tau} \in]0, 1/2]$. Nous remarquons que l'estimateur $\boldsymbol{T}^*_{k_n}(1/2)$

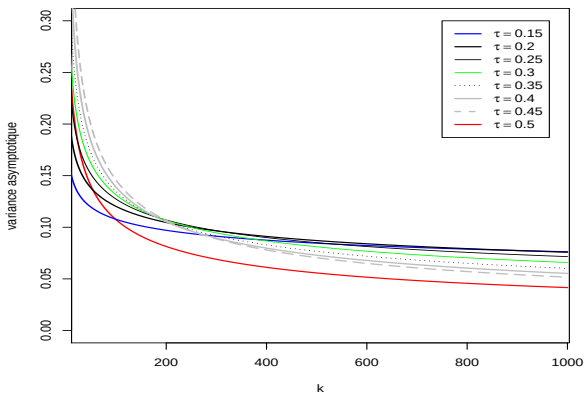

FIG. 2.2 – $\mathcal{AV}(\omega_{k_n}^{(\gamma)}, \boldsymbol{\tau})/\omega_{k_n}^{(\gamma)}$ en fonction de k_n pour différentes valeurs de $\boldsymbol{\tau} \in]0, 1/2]$.

a la plus petite variance pour k_n variant. Ce qui justifie son utilisation dans la suite. Notre étude de simulation est basée sur 500 échantillons de taille $n = 500$, 1000 et 1500 issus du modèle de Fréchet: $\overline{F}(x) = 1 - \exp(-x^{-1})$ et d'une distribution F vérifiant: $\overline{F}(x) = x^{-4/3}/\log(x)$. Tout d'abord, nous comparons, en termes de biais et d'erreur quadratique moyenne, la performance de l'estimateur $\boldsymbol{T}_{k_n}^*(1/2)$. Les résultats de cette partie sont illustrés par les graphes des Figures 2.4–2.7. En terme de biais et d'erreurs moyennes quadratiques, ces graphes montrent que l'estimateur généralisé de Hill $\boldsymbol{T}_{k_n}^*(1/2)$ a une meilleure performance que l'estimateur classique de Hill et celui de Pickands pour les modèles précédents. Finalement, nous étudions la qualité de l'approximation gaussienne sur des échantillons de tailles n assez grand basés sur les modèles précédents. Les résultats de cette partie sont représentés par les histrogrammes et les graphes du « Normal QQ-plot » des Figure 2.8 et 2.9. Ces résultats sont évalués la valeur $k_n = k_n^* = \arg\min_{k_n} \mathcal{MSE}(\boldsymbol{T}_{k_n}^*(1/2))$. A partir de ces graphes, nous obtenons une bonne approximation de la normalité de l'estimateur $\boldsymbol{T}_{k_n}^*(1/2)$ dans les Corollaires 2.3.1 et 2.3.2. Les tests d'ajustement de normalité sur l'estimateur $\boldsymbol{T}_{k_n}^*(1/2)$ au niveau de significativité de 5% de Cramer-von Mises (CvM), Lilliefors (Kolmogorov-Simirnov) (L), Anderson-Darling (AD) et Pearson (P) sur $\boldsymbol{T}_{k_n^*}^*(1/2)$ et basés sur les réalisations des modèles précédents sont donnés dans le Tableau 2.2.

2.5 Illustration sur des données hydrauliques

La performance de l'estimateur $T_{k_n}^*(1/2)$ est également illustrée à travers l'analyse des événements extrêmes sur les données de la rivière Nidd très utilisées dans les applications de la théorie des valeurs extrêmes (Hosking et Wallis, 1987 et Davison et Smith, 1990).

n	CvM	L	AD	P
500	0.552	0.517	0.600	0.292
1000	0.202	0.403	0.214	0.694
1500	0.515	0.492	0.502	0.111

n	CvM	L	AD	P
500	0.276	0.249	0.274	0226
1000	0.503	0.758	0.432	0.905
1500	0.112	0.234	0.103	0.857

TAB. 2.2 – Tests d'ajustement de normalité au niveau de significativité de 5%: Cramer-von Mises (CvM), Lilliefors (Kolmogorov-Simirnov) (L), Anderson-Darling (AD) et Pearson (P) pour l'estimateur $\boldsymbol{T}^*_{k_n^*}(1/2)$ évalué sur les réalisations du modèle de Fréchet: $F(x) = \exp\left(-x^{-1}\right)$ (taubleau de gauche) et d'une distribution F véfrifiant: $\overline{F}(x) = x^{-4/3}/\log(x)$ (tableau de droite).

Les données brutes sont en 154 dépassements de niveau 65 m^3/s par la rivière Nidd (Yorkshire, Angleterre) au cours de la période de 1934 à 1969 (soit une durée de 35 ans). Il n'y a pas de compromis général sur le domaine d'attraction de ces données. Diebolt et al. (2005) ont supposé que ces données appartiennent au domaine d'attraction de Fréchet. Cependant, dans Hosking et Wallis (1987), les données de la rivière Nidd semblent provenir d'une distribution appartenant au domaine d'attraction de Gumbel. Ainsi, notre famille d'estimateurs $\boldsymbol{T}^*_{k_n}(1/2)$, qui étudie à la fois le domaine de Fréchet et celui de Gumbel, sera d'un intérêt capital pour l'estimation de $\boldsymbol{\gamma}$, afin de trouver le domaine d'attraction. Sur la Figure 2.5, nous représentons l'estimateur $\boldsymbol{T}^*_{k_n}(1/2)$ pour plusieurs valeurs de k_n. Comparé aux autres estimateurs de Hill $\widehat{\gamma}^H_{k_n}$, des Moments $\widehat{\gamma}^{\mathcal{M}}_{k_n}$ et des Moments pondérés $\widehat{\gamma}^{\mathcal{PWM}}_{k_n}$, l'estimateur $\boldsymbol{T}^*_{k_n}(1/2)$ fournit l'estimation la plus stable pour k_n variant. Et en utilisant l'approche graphique donnée dans la Sous-Setion 1.6.1 sur le choix optimal de k_n, ce graphe montre que notre estimateur devient stable lorsque $k_n \geq 75$, ce qui correspond à une valeur estimée $\boldsymbol{T}^*_{k_n}(1/2) \simeq 0.29$. Un rejet de l'hypothèse $\mathbf{H_0}$ par le test de significativité du paramètre de l'indice de queue présenté dans la Sous-section 2.3.3 montre que les données de la rivière Nidd peuvent être supposées provenir d'une distribution appartenant au domaine d'attraction de Fréchet. Ce qui rejoint les résultats obtenus par El Methni et al. (2012), Figure 4 sur l'estimation du paramètre de queue de Weibull et Diebolt et al. (2005), Figure 7 pour l'estimation de l'indice de queue avec les méthodes bayésiennes, avec comme valeur estimée égale à 0.3 dans les deux cas cités. Ainsi, en utilisant l'équation (1.42) dans la Sous-Section 1.7.1, nous pouvons donner un estimateur de quantiles extrêmes ($q(\alpha_n)$) de type Weissman, basé sur $\boldsymbol{T}^*_{k_n}(1/2)$ et défini de la manière suivante:

$$\widehat{q}^{W,\boldsymbol{T}}_{k_n}(\alpha_n) = X_{n-k_n,n}\left(n\alpha_n/k_n\right)^{-\boldsymbol{T}^*_{k_n}(1/2)}.$$

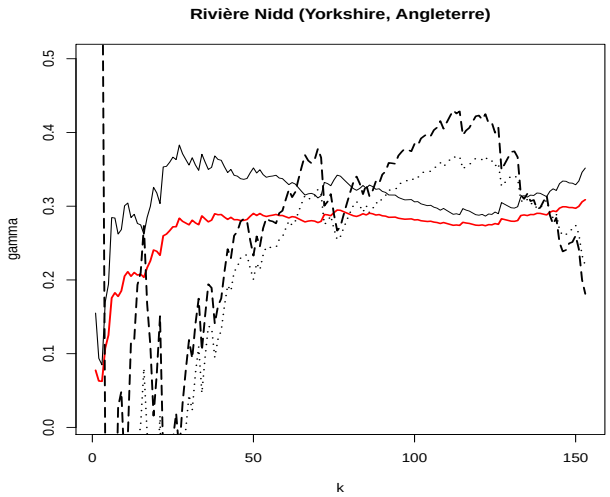

FIG. 2.3 – Estimateurs: $\boldsymbol{T}^*_{k_n}(1/2)$ (ligne pleine, rouge), Hill $\widehat{\gamma}^H_{k_n}$ (ligne pleine, noire), des Moments $\widehat{\gamma}^{\mathcal{M}}_{k_n}$ (ligne pointillée) et des Moments pondérés $\widehat{\gamma}^{\mathcal{PWM}}_{k_n}$ (ligne tiret) en fonction de k_n sur les données de la rivière Nidd.

La normalité asymptotique donnée dans de Haan Ferreira (2006), page 138, nous permet d'obtenir l'intervalle de confiance (I.C) de $q(\alpha)$ ci-dessous:

$$I.C = \left(\frac{\widehat{q}^{W,\boldsymbol{T}}_{k_n}(\alpha_n)}{1 + z_{1-\alpha/2}\log(k_n/(n\alpha_n))\frac{\boldsymbol{T}^*_{k_n}(1/2)}{\sqrt{k_n}}}, \frac{\widehat{q}^{W,\boldsymbol{T}}_{k_n}(\alpha_n)}{1 - z_{1-\alpha/2}\log(k_n/(n\alpha_n))\frac{\boldsymbol{T}^*_{k_n}(1/2)}{\sqrt{k_n}}} \right),$$

où z_α est le quantile dune loi normale standard au niveau α. Ce qui intéresse les hydrologues c'est la détermination des quantiles extrêmes dans le but de prédire le niveau de retour sur une longue durée. Rappelons que le niveau de retour sur N années noté NR_N (en anglais return level), correspond au niveau qui peut être dépassé en moyenne une fois au cours des N années. Selon Davison et Smith (1990) et Diebolt et al. (2005), nous avons $NR_N = q(k_n/(\lambda N n))$, avec n le nombre d'observations des données (pour les données de la rivière Nidd $n = 154$). Ces auteurs ont supposé que le processus des dépassements est modéliser par un Poisson de taux annuel λ. Ainsi l'estimation du niveau de retour est donnée par $\widehat{NR}_N = \widehat{q}_{k_n}\left(k_n/(\widehat{\lambda} N n)\right)$, $\widehat{\lambda}$ est un estimateur de λ. Si nous avons observé k_n dépassements au dessus de $X_{n-k_n,n}$ durant 35 ans, alors $\widehat{\lambda} = k_n/35$. Pour les données de la rivière Nidd nous avons $\alpha_n = k_n/(\widehat{\lambda} N n) \simeq 35/(Nn)$. Donc, une estimation d'un niveau de retour sur $N = 50$ et $N = 100$ ans correspond respectivement à l'estimation du quantile extrême $q(\alpha_n)$ avec $\alpha_n = 35/(50 \times 154) \simeq 4.54\ 10^{-3}$ et $\alpha_n = 35/(100 \times 154) \simeq 2.27\ 10^{-3}$. Notons par $\widehat{NR}^T_N$, le niveau de retour calculé à partir de l'estimateur de quantiles extrêmes $\widehat{q}^{W,\boldsymbol{T}}_{k_n}(\alpha_n)$. Nous regroupons dans le Tableau 2.3, les niveaux de retour $\widehat{NR}^T_{50}$ et $\widehat{NR}^T_{100}$ (en m^3/s) respectivement sur 50 et 100 ans pour ces données lorsque $k_n = 60, 70, 80$.

k_n	$\widehat{NR}_{50}^T$	I. Confiance	$\widehat{NR}_{100}^T$	I. Confiance
60	$317\ m^3/s$	$[239\ m^3/s, 467\ m^3/s]$	$386\ m^3/s$	$[292\ m^3/s, 569\ m^3/s]$
70	$310\ m^3/s$	$[239\ m^3/s, 443\ m^3/s]$	$376\ m^3/s$	$[289\ m^3/s, 537\ m^3/s]$
80	$319\ m^3/s$	$[248\ m^3/s, 449\ m^3/s]$	$390\ m^3/s$	$[303\ m^3/s, 549\ m^3/s]$

TAB. 2.3 – Niveaux de retour $\widehat{NR}_{50}$ et $\widehat{NR}_{100}$ respectivement sur 50 et 100 ans et leurs bornes de confiance au niveau de 95% calculés avec les nombres d'extrêmes $k_n = 60, 70, 80$ pour les données de la rivière Nidd.

2.6 Démonstrations

2.6.1 Résultats préliminaires

Notre premier résultat est basé sur le Lemme 2.6.1 ci dessous:

Lemme 2.6.1 *Définissons pour* $\boldsymbol{\nu} > 0$, $S_{k_n}(\boldsymbol{\nu}) = k_n^{-\boldsymbol{\nu}} \sum_{j=1}^{k_n} j^{-\boldsymbol{\nu}}$. *Alors*

(S1)	*pour* $\boldsymbol{\nu} > 1$,	$S_{k_n}(\boldsymbol{\nu}) \sim k_n^{-\boldsymbol{\nu}} \zeta(\boldsymbol{\nu})$,
(S2)	*pour* $\boldsymbol{\nu} = 1/2$,	$S_{k_n}(\boldsymbol{\nu}) = 2(1 + O(k_n^{-1/2}))$,
(S3)	*pour* $\boldsymbol{\nu} < 1$,	$S_{k_n}(\boldsymbol{\nu}) = k_n^{1-2\boldsymbol{\nu}}/(1-\boldsymbol{\nu})(1 + O(k_n^{-1-\boldsymbol{\nu}}))$,
(S4)	*pour* $\boldsymbol{\nu} = 1$,	$S_{k_n}(\boldsymbol{\nu}) = (\log k_n) k_n^{-1}(1 + O(1/\log k_n))$.

Démonstration du Lemme 2.6.1 Ces formules sont obtenues en comparant $\sum_{j=1}^{k_n} j^{-\boldsymbol{\nu}}$ et $\int_1^{k_n} x^{-\boldsymbol{\nu}} dx$, pour une analyse classique. ■

Notre second résultat préliminaire se résume au Lemme 2.6.2, qui est un outil permettant de déterminer la loi asymptotique de notre classe d'estimateurs $\boldsymbol{T}_{k_n}(\boldsymbol{\tau})$ pour $0 < \boldsymbol{\tau} \leq 1/2$. On considère les séquences de variables aléatoires indépendantes et uniformément distribuées $\xi_1, \xi_2, \ldots$ et données dans (2.1.2)

Lemme 2.6.2 *Soit* $V_{k_n}(\boldsymbol{\tau}) = k_n^{-\boldsymbol{\tau}} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}} (\log \xi_{j+1,n} - \log \xi_{j,n})$. *Alors nous avons:*

- *pour* $\boldsymbol{\tau} = 1/2$,

$$\frac{a_{k_n}(\boldsymbol{\tau})}{\sigma_{k_n}(\boldsymbol{\tau})} \left(\frac{V_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - 1 \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

- *et pour* $0 < \boldsymbol{\tau} < 1/2$,

$$\frac{a_{k_n}(\boldsymbol{\tau})}{\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{V_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{L}(\boldsymbol{\tau}),$$

où $a_{k_n}(\boldsymbol{\tau})$ *et* $\sigma_{k_n}(\boldsymbol{\tau})$ *sont définis dans (2.7) et la variable aléatoire* $\mathcal{L}(\boldsymbol{\tau})$ *donnée dans la Section 2.2 par:*

$$\mathcal{L}(\boldsymbol{\tau}) = \frac{1}{\sqrt{\zeta(2(1-\boldsymbol{\tau}))}} \sum_{j=1}^{\infty} j^{-(1-\boldsymbol{\tau})}(E_j - 1),$$

est centrée, réduite et ayant des moments finis.

Démonstration du Lemme 2.6.2 D'après la représentation de Malmquist (2.2), nous avons $V_{k_n}(\boldsymbol{\tau}) = k_n^{-\boldsymbol{\tau}} \sum j^{\boldsymbol{\tau}-1} E_j$. Il s'en suit que $\mathbb{E}(V_{k_n}(\boldsymbol{\tau})) = a_{k_n}(\boldsymbol{\tau})$ et $\mathbb{V}ar(V_{k_n}(\boldsymbol{\tau})) = \sigma^2_{k_n}(\boldsymbol{\tau})$.
Posons

$$V^*_{k_n}(\boldsymbol{\tau}) = \frac{a_{k_n}(\boldsymbol{\tau})}{\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{V_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})} - 1\right)$$

alors on obtient

$$V^*_{k_n}(\boldsymbol{\tau}) = \frac{k_n^{\boldsymbol{\tau}}}{\sigma_{k_n}(\boldsymbol{\tau})} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(E_j - 1).$$

Pour $0 < \boldsymbol{\tau} < 1/2$, nous avons d'après (**S1**) ci dessus, $k_n^{\boldsymbol{\tau}}/\sigma_{k_n}(\boldsymbol{\tau}) \to \zeta(2(1-\boldsymbol{\tau}))^{-1/2}$ et donc

$$V^*_{k_n}(\boldsymbol{\tau}) \to \frac{1}{\sqrt{\zeta(2(1-\boldsymbol{\tau}))}} \sum_{j=1}^{\infty} j^{\boldsymbol{\tau}-1}(E_j - 1) = \mathcal{L}(\boldsymbol{\tau}).$$

Maintenant, nous allons montrer que $\mathcal{L}(\boldsymbol{\tau})$ est une variable aléatoire centrée et réduite de moments finis. La fonction caractéristique de $V^*_{k_n}(\boldsymbol{\tau})$ est

$$\Phi_{V^*_{k_n}(\boldsymbol{\tau})}(t) = \exp\left(-\frac{1}{\sqrt{\zeta(2(1-\boldsymbol{\tau}))}} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(it)\right) \prod_{j=1}^{k_n}\left(1 - \frac{itj^{-(1-\boldsymbol{\tau})}}{\sqrt{\zeta(2(1-\boldsymbol{\tau}))}}\right)^{-1}.$$

En utilisant un développement sur la fonction $\log(1-\cdot)$ et par le Théorème de Lebesgues, on montre que

$$\begin{aligned}\Phi_{V^*_{k_n}(\boldsymbol{\tau})}(t) &= \exp\left(\sum_{j=1}^{k_n}\sum_{n=2}^{\infty} \frac{(it)^n}{n} \times \frac{j^{-n(1-\boldsymbol{\tau})}}{\sqrt{\zeta\left(2(1-\boldsymbol{\tau})\right)}}\right) \\ &\to \Phi_\infty(t) = \exp\left(\sum_{n=2}^{\infty} \frac{(it)^n}{n} \times \frac{\zeta(n(1-\boldsymbol{\tau}))}{\sqrt{\zeta\left(2(1-\boldsymbol{\tau})\right)}}\right). \qquad (2.23)\end{aligned}$$

Dans Valiron (1990), p.506, nous avons pour $s > 1$,

$$\zeta(s) = \frac{s}{s-1} + \frac{1}{s}\int_1^\infty \frac{x-[x]}{x^{s+1}}dx,$$

où $[x]$ est la partie entière de x. On obtient

$$\frac{s}{s-1} \leq \zeta(s) \leq s\left(1+\frac{1}{s-1}\right).$$

En utilisant ceci, on voit que la valeur absolue du terme général de la série dans (2.23) est dominée pour de grandes valeurs de n, il s'en suit

$$\left|\frac{(it)^n}{n} \times \frac{\zeta(n(1-\boldsymbol{\tau}))}{\sqrt{\zeta\left(2(1-\boldsymbol{\tau})\right)}}\right| \leq 2(1-2\boldsymbol{\tau})\left(\frac{|t|}{2(1-\boldsymbol{\tau})}\right)^n.$$

Et donc $\Phi_\infty(t)$ caractérise la fonction de répartition de $\mathcal{L}(\boldsymbol{\tau})$. Il reste à étudier le cas $\boldsymbol{\tau} = 1/2$. Evaluons la fonction génératrice des moments de $V^*_{k_n}(1/2)$, notée $\Phi_{V^*_{k_n}(1/2)}$ et donnée par

$$\Phi_{V^*_{k_n}(1/2)}(t) = \prod_{j=1}^{k_n} \Phi_{(E_j-1)}\left(\frac{ti^{\boldsymbol{\tau}-1}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right),$$

où $\Phi_{(E_j-1)}$ est la fonction génératrice des moments de $E_j - 1$. Mais nous avons $\sigma^2_{k_n}(1/2) = S_{k_n}(1) \sim k_n^{-1}\log k_n$ et aussi $k_n^{1/2}\sigma_{k_n}(1/2) = (\log k_n)^{1/2} \to \infty$. Alors pour k_n fixé et assez grand,

$$\left|\frac{t\ i^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right| \leq \frac{2t}{\sqrt{\log k_n}} \tag{2.24}$$

uniformément en $i \geq 1$. Dans cette étape, nous utiliserons un développement de $\psi_{(E_j-1)}$ au voisinage de zéro $\Phi_{(E_j-1)}(u) = 1 + u^2/2 + u^3 g(u)$, où il existe u_0 et $C > 0$ tels que

$$0 \leq u \leq u_0 \Rightarrow |g(u)| \leq C.$$

En utilisant la bornitude uniforme dans (2.24), pour une valeur de k_n telle que $2t(\log k_n)^{-1/2} \leq u_0$, on obtient

$$\Phi_{(E_j-1)}\left(\frac{t\ j^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right) = 1 + \frac{1}{2}\left(\frac{t\ j^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right)^2 + \left(\frac{t\ j^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right)^3 g_{0,j,n}(t),$$

où $|g_{0,j,n}(t)| \leq 1$ pour tout $1 \leq j \leq k_n$. Par la bornitude uniforme du terme d'erreur $g_{0,j,n}(t)$, on obtient

$$\begin{aligned}\log\Phi_{(E_j-1)}\left(\frac{t\ j^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right) &= \frac{1}{2}\left(\frac{t\ j^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right)^2 + \left(\frac{t\ j^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right)^3 g_{0,j,n}(t) \\ &\quad + \left(\frac{t\ j^{-1/2}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right)^3 g_{1,j,n}(t),\end{aligned}$$

avec $|g_{1,j,n}(t)| \leq 1$ pour tout $1 \leq j \leq k_n$. Finalement,

$$\begin{aligned}\Phi_{V^*_{k_n}(1/2)}(t) &= \exp\left(\sum_{j=1}^{k_n} \log \psi_{(E_j-1)}\left(\frac{ti^{\boldsymbol{\tau}-1}}{k_n^{1/2}\sigma_{k_n}(1/2)}\right)\right) \\ &= \exp(t^2/2 + t^3 g_2),\end{aligned}$$

où pour une grande valeur de k_n, $|g_2| \leq (\log k_n)^{-2}$. Par conséquent $\Phi_{V^*_{k_n}(1/2)}(t) \to \exp(t^2/2)$ et $V^*_{k_n}(1/2) \to \mathcal{N}(0,1)$. Ce qui achève la preuve du lemme. ■

2.6.2 Démonstration des résultats théoriques

Démonstration du Théorème 2.2.2.
Considérons les répresentations de (2.1.2). nous avons, pour tout $n \geq 1$,

$$\{\log X_{n-j+1,n} = Y_{n-j+1,n}, 1 \leq j \leq n\} \stackrel{\mathcal{D}}{=} \{G^{\leftarrow}\{1-\xi_{j,n}), 1 \leq j \leq n\}.$$

Notre preuve s'appuiera sur le Lemme 2.6.1 ci-dessous qui établit que

$$\frac{a_{k_n}(\boldsymbol{\tau})}{\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{1}{k_n^{\boldsymbol{\tau}} a_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}} \log\left(\frac{\xi_{j+1,n}}{\xi_{j,n}}\right) - 1\right) \stackrel{\mathcal{D}}{=} \frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(E_j - 1)$$

converge en distribution vers une loi $\mathcal{N}(0,1)$ pour $\boldsymbol{\tau} = 1/2$ et vers la distribution finie

$$\mathcal{L}(\boldsymbol{\tau}) = \frac{1}{\sqrt{\zeta(2(1-\boldsymbol{\tau}))}}\sum_{j=1}^{\infty} j^{-(1-\boldsymbol{\tau})}(E_j - 1), \quad \text{pour} \quad 0 < \boldsymbol{\tau} < 1/2,$$

où $E_1, E_2, \ldots$ sont des variables aléatoires indépendantes de loi exponentielle standard définie dans (2.2).
Nous commençons par le cas où $F \in D(\boldsymbol{\Lambda})$. D'après la représentation (2.5), on obtient

$$\begin{aligned}\boldsymbol{T}_{k_n}(\boldsymbol{\tau}) &= \sum_{j=1}^{k_n}\left(\frac{j}{k_n}\right)^{\boldsymbol{\tau}}(s(\xi_{j,n}) - s(\xi_{j+1,n})) + \sum_{j=1}^{k_n}\left(\frac{j}{k_n}\right)^{\boldsymbol{\tau}}\int_{\xi_{j,n}}^{\xi_{j+1,n}} \frac{s(t)}{t}\, dt \\ &\equiv S_n(1) + S_n(2). \end{aligned} \tag{2.25}$$

En utilisant (2.6), nous obtenons pour $\xi_{1,n} \leq v, u \leq \xi_{k_n}$,

$$\frac{s(u)}{s(v)} = \frac{1+p(u)}{1+p(v)}\exp\left(-\int_u^v \Delta(t)\frac{dt}{t}\right).$$

En posant,

$$g_{1,n,0}(p) = \sup\{|p(u)|, 0 \leq u \leq \xi_{k_n+1,n}\} \text{ et } g_{2,n,0}(\Delta) = \sup\{|\Delta(u)|, 0 \leq u \leq \xi_{k_n+1,n}\},$$

et en remarquant que $\log(\xi_{k_n+1,n}/\xi_{1,n}) = O_{\mathbb{P}}(\log k_n)$, on obtient

$$\frac{s(u)}{s(v)} = (1 + O(g_{1,n,0})) \exp\left(-O_{\mathbb{P}}(g_{2,n,0} \log k_n)\right).$$

Si

$$g_{2,n,0} \log k_n \xrightarrow{\mathbb{P}} 0, \quad \text{alors} \tag{2.26}$$

$$\sup_{\xi_{1,n} \leq u,v \leq \xi_{k_n}} \left| \frac{s(u)}{s(v)} - 1 \right| = O_{\mathbb{P}}(\max(g_{1,n,0}, g_{2,n,0} \log k_n)) \tag{2.27}$$

et finalement

$$\sup_{\xi_{1,n} \leq u,v \leq \xi_{k_n}} \left| \frac{s(u) - s(v)}{s(k_n/n)} \right| = O_{\mathbb{P}}(\max(g_{1,n,0}, g_{2,n,0} \log k_n). \tag{2.28}$$

Nous allons donc montrer (2.26). Du fait que $nk_n^{-1}\xi_{k_n+1,n} \xrightarrow{\mathbb{P}} 1$, pour $\lambda > 1$ telle que la condition (C3) soit vérifiée, alors pour tout $\varepsilon > 0$, il existe $N_0 \in \mathbb{N}^*$ tel que pour tout $n \geq N_0$, on ait

$$\mathbb{P}\left(g_{2,n,0} \log k_n \geq g_{2,n} \log k_n\right) \equiv \mathbb{P}\left(A_n\right) < \varepsilon.$$

Maintenant fixons $\delta > 0$. En utilisant la décomposition d'ensembles ci-dessous

$$\Big\{ |g_{2,n,0} \log k_n| > \delta \Big\} = \Big\{ |g_{2,n,0} \log k_n| > \delta \Big\} \cap A_n + \Big\{ |g_{2,n,0} \log k_n| > \delta \Big\} \cap A_n^c,$$

on obtient, pour $n \geq N_0$, $\quad \mathbb{P}(|g_{2,n,0} \log k_n| > \delta) \leq \mathbb{P}(|g_{2,n} \log k_n| > \delta) + \varepsilon.$

En tendant dans un premier temps $n \to \infty$ et ensuite $\varepsilon \downarrow 0$, et en prenant en compte la condition $(C3)$, on voit facilement que (2.26) est vérifiée de même que (B.3). En combinant (2.25) et (B.3), et en utilisant la technique précédente, on obtient pour une valeur donnée $\lambda > 1$ comme dans (C3), pour tout $\varepsilon > 0$, il exite $N_0 \in \mathbb{N}^*$ tel que pour tout $n \geq N_0$ et pour tout $\delta > 0$,

$$\mathbb{P}\left(\left| \frac{S_n(1)}{\sigma_{k_n}(\boldsymbol{\tau}) s(k_n/n)} \right| > \delta \right) \leq \mathbb{P}\left(\frac{d_n(p, b, \lambda}{k_n^{\boldsymbol{\tau}} \sigma_{k_n}(\boldsymbol{\tau})} \sum_{j=1}^{k_n} j^{\boldsymbol{\tau}} > \delta \right) + \varepsilon,$$

où $d_n(p, \Delta, \lambda) = \max(g_{1,n}, g_{2,n} \log k_n)$.

En tendant d'abord $n \to \infty$ et ensuite $\varepsilon \downarrow 0$, et en choisissant $(C3)$, on obtient

$$\frac{S_n(1)}{\sigma_{k_n}(\boldsymbol{\tau}) s(k_n/n)} \xrightarrow{\mathbb{P}} 0.$$

Maintenant nous allons traiter le cas $S_n(2)$, nous remarquons que

$$\begin{aligned}
\frac{S_n(2)}{\sigma_{k_n}(\boldsymbol{\tau})s(k_n/n)} &= \frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}}\int_{\xi_{j,n}}^{\xi_{j+1,n}} \frac{s(t)}{s(k_n/n)}\,\frac{dt}{t} \\
&= \frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}}\int_{\xi_{j,n}}^{\xi_{j+1,n}} \frac{dt}{t} \\
&\quad + \frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}}\int_{\xi_{j,n}}^{\xi_{j+1,n}} \left(\frac{s(t)}{s(k_n/n)} - 1\right)\frac{dt}{t} \\
&= S_n(2,1) + S_n(2,2).
\end{aligned}$$

Nous avons, en utilisant (2.27) et la représentation de Malmquist (2.2),

$$\begin{aligned}
|S_n(2,2)| &\leq O_{\mathbb{P}}(1)\frac{d_n(p,\Delta,\lambda)}{k_n^{\boldsymbol{\tau}}\sigma_n(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}E_j \\
&\leq O_{\mathbb{P}}(1) \times \left\{\frac{d_{n,0}(p,\Delta,\lambda)}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(E_j-1) + \frac{d_{n,0}(p,\Delta,\lambda)}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}\right\}.
\end{aligned}$$

où $d_{n,0}(p,b,\lambda) = \max(g_{1,n,0}, g_{2,n,0}\log k_n)$.

Le premier terme tend vers zéro en probabilité du fait que $d_{n,0}(p,\Delta,\lambda) \xrightarrow{\mathbb{P}} 0$ par la condition (2.26), et $(k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau}))^{-1}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(E_j-1)$ converge en distribution vers une distribution finie par le Lemme 2.6.2 dans la section 2.6.1. On manipule aussi le second terme avec la technique précédente. On choisit une valeur $\lambda > 1$ comme dans $(C3)$. Pour tout $\varepsilon > 0$, il existe $N_0 \in \mathbb{N}^*$ tel que pour tout $n \geq N_0$ et $\delta > 0$,

$$\mathbb{P}\left(\left|\frac{d_{n,0}(p,\Delta,\lambda)}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}\right| > \delta\right) \leq \varepsilon + \mathbb{P}\left(\left|\frac{d_n(p,\Delta,\lambda)}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}\right| > \delta\right).$$

En tendant $n \uparrow \infty$ dans un premier temps et ensuite $\varepsilon \downarrow 0$, et en utilisant $(C3)$ on voit que ce terme tend vers 0 en probabilité. Finallement, par la représenation de Malmquist (2.2), on montre que

$$S_n(2,1) = \frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}E_j + o_{\mathbb{P}}(1).$$

Ce qui conduit à

$$S_n(2,1) - \frac{a_{k_n}(\boldsymbol{\tau})}{\sigma_{k_n}(\boldsymbol{\tau})} = \frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(E_j-1) \equiv V_{k_n}(\boldsymbol{\tau}).$$

Et finalement, par sommation on obtient,

$$\frac{a_{k_n}(\boldsymbol{\tau}}{\sigma_{k_n}(\boldsymbol{\tau})s(k_n/n)}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})}-s(k_n/n)\right)=V_{k_n}(\boldsymbol{\tau})+o_{\mathbb{P}}(1).$$

Ce qui achève la preuve de la partie du théorème, par le Lemme 2.6.2 dans la Section 2.6.1, $V_{k_n}(\boldsymbol{\tau})$ converge en distribution vers une loi $\mathcal{N}(0,1)$ pour $\boldsymbol{\tau}=1/2$ et vers $\mathcal{L}(\boldsymbol{\tau})$ pour $\boldsymbol{\tau}<1/2$.

Maintenant pour la seconde partie de la preuve nous considérons $F\in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma}>0$. Nous aurons ici une preuve similaire mais nous n'allons pas expliciter les convergences en distribution comme nous les avions developpé dans le cas précédent. Nous avons par (2.4) et par la représentation habituelle,

$$\begin{aligned}\boldsymbol{T}_{k_n}(\boldsymbol{\tau}) &= \frac{1}{k_n^{\boldsymbol{\tau}}}\sum_{j=1}^{k_n}j^{\boldsymbol{\tau}}\log\left(\frac{1+p(\xi_{j+1,n})}{1+p(\xi_{j,n})}\right)\\ &\quad+\boldsymbol{\gamma}\frac{1}{k_n^{\boldsymbol{\tau}}}\sum_{j=1}^{k_n}j^{\boldsymbol{\tau}}\log\left(\frac{\xi_{j+1,n}}{\xi_{j,n}}\right)\\ &\quad+k_n^{-\boldsymbol{\tau}}\sum_{j=1}^{k_n}j^{-\boldsymbol{\tau}}\int_{\xi_{j,n}}^{\xi_{j+1,n}}\Delta(t)\frac{dt}{t}\\ &\equiv S_n(1)+S_n(2)+S_n(3).\end{aligned}$$

On obtient, pour une valeur de k_n très grande,

$$\left|\frac{S_n(1)}{\sigma_{k_n}(\boldsymbol{\tau})}\right|\leq\frac{2g_{1,n,0}(p)}{k_n^{\boldsymbol{\tau}}\sigma_n(\boldsymbol{\tau})}\sum_{j=1}^{k_n}j^{\boldsymbol{\tau}},$$

où $g_{1,n,0}$ est défini dans (2.6.2), et tend vers zéro en probabilité par (**C1**) et (2.27).

$$\begin{aligned}\left|\frac{S_n(3)}{\sigma_{k_n}(\boldsymbol{\tau})}\right| &\leq \frac{g_{2,n,0}(\Delta)}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n}j^{\boldsymbol{\tau}}\log\left(\frac{\xi_{j+1,n}}{\xi_{j,n}}\right)\\ &= \frac{g_{2,n,0}(\Delta)}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n}j^{\boldsymbol{\tau}-1}(E_j-1)+\frac{g_{2,n,0}(\Delta)}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n}j^{\boldsymbol{\tau}-1},\end{aligned}$$

où $g_{2,n,0}$ est défini dans (2.6.2). alors $\dfrac{S_n(3)}{\sigma_{k_n}(\boldsymbol{\tau})}\xrightarrow{\mathbb{P}}0$ par $(C2)$, le Lemme 2.6.1 et les méthodes décrites ci-dessus. Finalement, par le Lemme 2.6.1 on obtient

$$\frac{a_{k_n}(\boldsymbol{\tau})}{\boldsymbol{\gamma}\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{S_n(3)}{a_{k_n}(\boldsymbol{\tau})}-\boldsymbol{\gamma}\right)=\frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n}j^{\boldsymbol{\tau}-1}(E_j-1)$$

et

$$\begin{aligned}\frac{a_{k_n}(\boldsymbol{\tau})}{\boldsymbol{\gamma}\sigma_{k_n}(\boldsymbol{\tau})}\left(\frac{\boldsymbol{T}_{k_n}(\boldsymbol{\tau})}{a_{k_n}(\boldsymbol{\tau})}-\boldsymbol{\gamma}\right) &= \frac{1}{k_n^{\boldsymbol{\tau}}\sigma_{k_n}(\boldsymbol{\tau})}\sum_{j=1}^{k_n} j^{\boldsymbol{\tau}-1}(E_j-1)+o_{\mathbb{P}}(1),\\ &= V_{k_n}(\boldsymbol{\tau})+o_{\mathbb{P}}(1)\end{aligned}$$

et ceci converge en distribution vers une loi $\mathcal{N}(0,1)$ pour $\boldsymbol{\tau}=1/2$ et vers $\mathcal{L}(\boldsymbol{\tau})$ pour $\boldsymbol{\tau}<1/2$. Ce qui complète la preuve en remarquant que $a_{k_n}(\boldsymbol{\tau})\to 1/\boldsymbol{\tau}$. ■

Démonstration des corollaires 2.3.1 et 2.3.2: Il suffit de remarquer par le Lemme 2.6.1, que $a_{k_n}(\frac{1}{2})=2(1+O(k_n^{-1/2}))$ et $a_{k_n}(\frac{1}{2})/\sigma_{k_n}(\frac{1}{2})\sim(4k_n/\log k_n)^{1/2}$. ■

Démonstration des corollaires 2.3.5: D'après le Lemme 2.6.1, nous avons pour $0<\boldsymbol{\tau}<1/2$, $a_{k_n}(\boldsymbol{\tau})/\sigma_{k_n}(\boldsymbol{\tau})\sim\boldsymbol{\tau}^{-1}k_n^{\boldsymbol{\tau}}/\sqrt{\zeta(2(1-\boldsymbol{\tau}))}$. ■

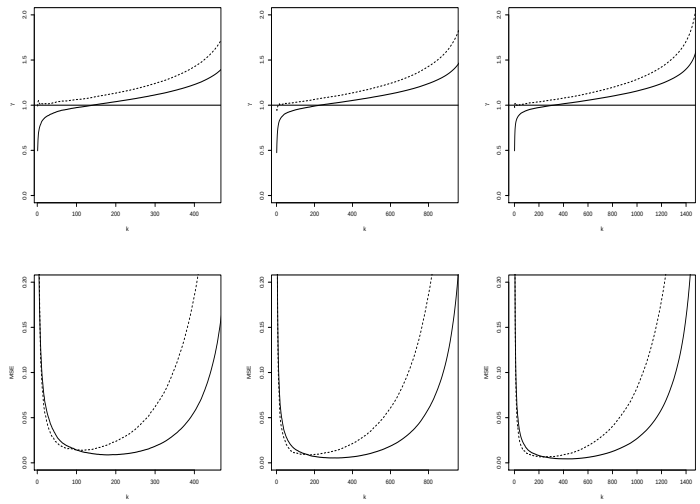

FIG. 2.4 – Estimateurs en moyenne (haut) et erreurs moyennes quadratiques empiriques (bas) des estimateurs $\boldsymbol{T}^*_{k_n}(1/2)$ (ligne pleine) et $\widehat{\boldsymbol{\gamma}}^H_{k_n}$ (ligne tiret) en fonction de k_n basés sur 500 réalisations du modèle de Fréchet: $F(x)=\exp\left(-x^{-1}\right)$, de tailles: $n=500$ (gauche), $n=1000$ (milieu) et $n=1500$ (droite).

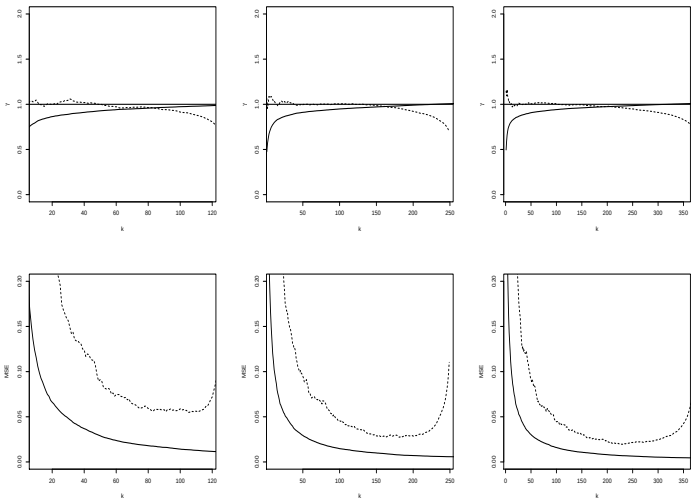

FIG. 2.5 – Estimateurs en moyenne (haut) et erreurs moyennes quadratiques empiriques (bas) des estimateurs $\boldsymbol{T}^*_{k_n}(1/2)$ (ligne pleine) et $\widehat{\gamma}^P_{k_n}$ (ligne tiret) en fonction de k_n basés sur 500 réalisations du modèle de Fréchet: $F(x) = \exp\left(-x^{-1}\right)$, de tailles: $n = 500$ (gauche), $n = 1000$ (milieu) et $n = 1500$ (droite).

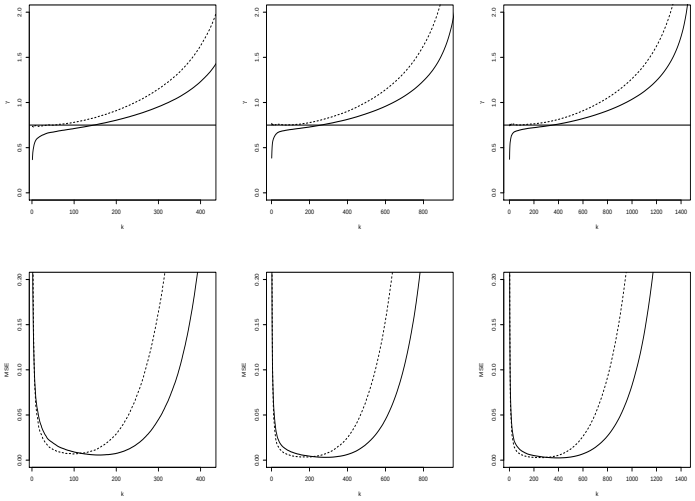

FIG. 2.6 – Estimateurs en moyenne (haut) et erreurs moyennes quadratiques empiriques (bas) des estimateurs $\boldsymbol{T}^*_{k_n}(1/2)$ (ligne pleine) et $\widehat{\gamma}^H_{k_n}$ (ligne tiret) en fonction de k_n basés sur 500 réalisations d'une distribution F véfrifiant: $\overline{F}(x) = x^{-4/3}/\log(x)$, de tailles: $n = 500$ (gauche), $n = 1000$ (milieu) et $n = 1500$ (droite).

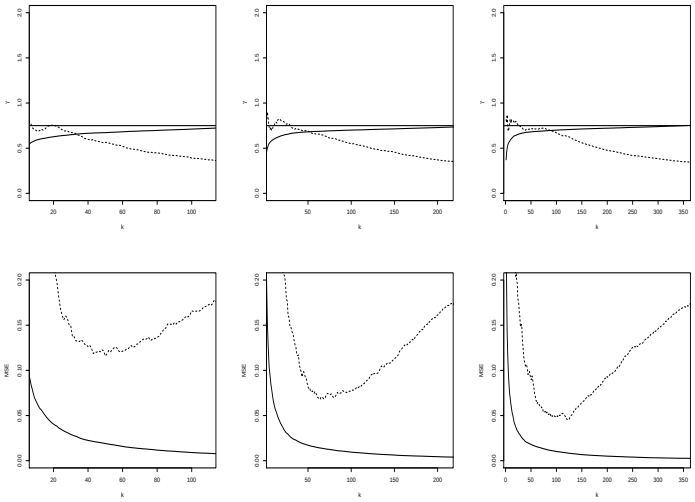

FIG. 2.7 – Estimateurs en moyenne (haut) et erreurs moyennes quadratiques empiriques (bas) des estimateurs $\boldsymbol{T}^*_{k_n}(1/2)$ (ligne pleine) et $\widehat{\gamma}^P_{k_n}$ (ligne tiret) en fonction de k_n basés sur 500 réalisations d'une distribution F véfrifiant: $\overline{F}(x) = x^{-4/3}/\log(x)$, de tailles: $n = 500$ (gauche), $n = 1000$ (milieu) et $n = 1500$ (droite).

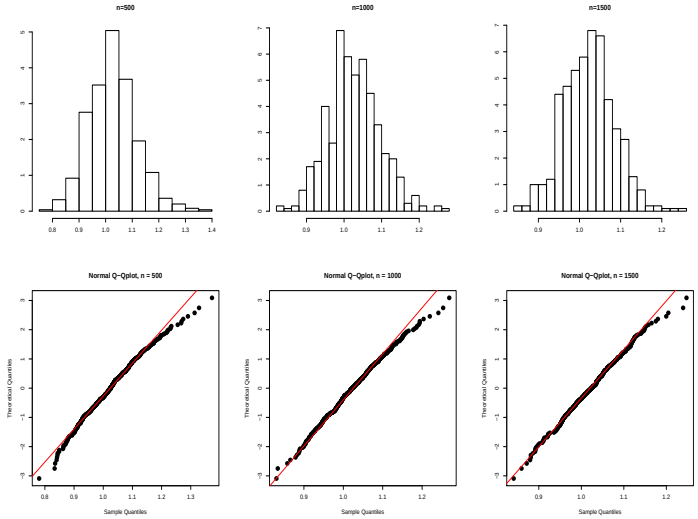

FIG. 2.8 – Histogrammes(haut) et graphes du Normal QQ-plot (bas) sur $\boldsymbol{T}^*_{k^*_n}(1/2)$ vasés sur des réalisations du modèle de Fréchet: $F(x) = \exp\left(-x^{-1}\right)$, de tailles: $n = 500,\ 1000,\ 1500$.

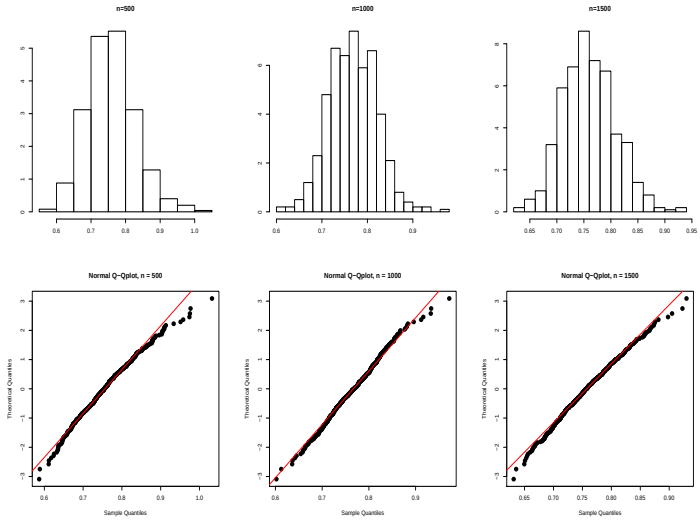

FIG. 2.9 – Histogrammes(haut) et graphes du Normal QQ-plot (bas) sur $\boldsymbol{T}^*_{k^*_n}(1/2)$ vasés sur des réalisations d'une distribution F véfrifiant: $\overline{F}(x) = x^{-4/3}/\log(x)$, de tailles: $n = 500,\ 1000,\ 1500$.

CHAPITRE 3

Processus fonctionnel de Hill généralisé et théorie uniforme

Résumé

Dans ce chapitre, nous nous proposons de développer un processus indexé par une classe de fonctions et permettant d'estimer l'indice des valeurs extrêmes. On obtient une grande classe d'estimateurs généralisant l'estimateur de Hill. Nous établissons la loi asymptotique de ce type de processus en distribution finie et ponctuelle. Et par une théorie moderne de convergence vague, nous donnons également sa loi asymptotique fonctionnelle.

3.1 Introduction

Soit $\{X_i,\ i = 1, \ldots, n\}$ un échantillon de n variables aléatoires réelles indépendantes et de même loi que $X \geq 1$ avec une fonction de répartition $F(x) = \mathbb{P}(X \leq x)$. On note par $X_{1,n} \leq \ldots \leq X_{n,n}$ les statistiques d'ordre associées à cet échantillon. Ensuite, on s'intérésse aux propriétés asymptotiques fonctionnelles de la suite de processus stochastique

$$\boxed{\boldsymbol{T}_{k_n}(f) = \sum_{j=1}^{k_n} f(j) \left(\log X_{n-j+1,n} - \log X_{n-j,n}\right),} \tag{3.1}$$

indexé par la classe $\mathcal{F}$ de fonctions $f : \mathbb{N}^* = \mathbb{N}\backslash\{0\} \longmapsto \mathbb{R}_+$ et k_n est une séquence d'entiers telle que $1 \leq k_n \leq n, \quad k_n/n \to 0$ quand $n \to \infty$.

La motivation principale de ce travail est d'obtenir une grande classe d'estimateurs de

l'indice des valeurs extrêmes lorsque F appartient au domaine d'attraction de la distribution asymptotique des extrêmes (on pourra se référer au chapitre 1). Chaque estimateur est une marge d'un seul processus stochastique. Pour la fonction uniforme $f(j) = j$, $k_n^{-1}\boldsymbol{T}_{k_n}(f)$ correspond à l'estimateur de Hill (1975) donné dans la section 1.5.1.2. Récemment, une premiere étape de l'estimateur de Hill fonctionel sous la forme $k_n^{-\boldsymbol{\tau}}\boldsymbol{T}_{k_n}(f)$, $f = j^{\boldsymbol{\tau}}$ a été respectivement étudiée en distributions finies dans Lo *et al.* (2012) pour $0 < \boldsymbol{\tau} \leq 1/2$ et dans Diop et Lo (2006), pour $\boldsymbol{\tau} > 1/2$. Groeneboom *et al.* (2003) ont également examiné une étude approfondie d'une famille d'estimateurs à noyau de l'indice des valeurs extrêmes. Ce pendant, ils n'ont pas considéré une version avec les processus stochastiques. Il existe un grand nombre d'estimateurs de l'indice des valeurs extrêmes. On peut citer Pickands (1975), de Haan et Resnick (1980), Csörgő *et al.* (1985), Dekkers *et al.* (1989), etc. Les propriétés asymptotiques en distribution finie de ces estimateurs sont fixées pour un nombre fini d'entre eux pour l'ensemble des domaines d'attraction. Le lecteur pourra se référer à Hall (1982), Mason (1982), Davis et Resnick (1984), Csörgő *et al.* (1985), Haeusler et Teugels (1985), Lo (1986), Mason et Turova (1994), Resnick et Starica (1995), Diop et Lo (2006), Diop et Lo (2009), etc.
Cependant la theorie moderne de la convergence vague fonctionnelle permet de gérer des estimateurs plus complexes sous forme de processus stochastiques. Ainsi, pour le processus fonctionel $\{\boldsymbol{T}_{k_n}(f),\ f \in \mathcal{F}\}$, nous souhaitons montrer que pour tout $f \in \mathcal{F}$, il existe une suite non aléatoire $a_{k_n}(f)$ telle que $T_{k_n}(f)/a_{k_n}(f)$ est un estimateur de l'indice des valeurs extrêmes $\boldsymbol{\gamma}$. Ces processus sont appelés processus stochastiques des estimateurs de l'indice des valeurs extrêmes.

Considérons maintenant le processus stochastique défini dans (3.1). Notre objectif est d'établir leur normalité asymptotique fonctionnelle lorsque cela est possible ou simplement leur distribution asymptotique pour des classes appropriées. Nous verrons pour certaines classes qu'il existe un comportement asymptotique non Gaussien et qui sera entièrement caractérisé. Nous considérons principalement deux classes de fonctions. La première est composée de fonctions f satisfaisant

$$\textbf{(K1)} \qquad A(2, f) = \sum_{j=1}^{\infty} f(j)^2 j^{-2} < \infty,$$

avec la notation générale $A(m, f) = \sum_{j=1}^{\infty} f(j)^m j^{-m}$. La seconde inclut les fonctions f telles que

$$\textbf{(K2)} \qquad \limsup_{n \to +\infty} B(n, f) = 0,$$

où

$$B(n,f) = \sigma_{k_n}(f)^{-1} \max\{f(j)j^{-1}, 1 \leq j \leq k_n\}$$

et $\sigma_{k_n}(f)$ est défini ci-dessous dans (3.2). Sous ces deux conditions, nous serons en mesure de trouver les distributions asymptotiques de $\boldsymbol{T}_{k_n}(f)$ pour une fonction f fixée, sous une hypothèse classique utilisée en théorie des valeurs extrêmes. Mais comme lois fonctionnelles, nous avons donc besoin des conditions uniformes ci-dessous .
Définissons $\mathcal{F}_1$ la sous-classe de $\mathcal{F}$ telle que

$$\textbf{(KU1)} \qquad 0 < \inf_{f\in\mathcal{F}_1} A(2,f) < \sup_{f\in\mathcal{F}_1} A(2,f) < +\infty,$$

et $\mathcal{F}_2$ la sous-classe de $\mathcal{F}$ telle que

$$\textbf{(KU2a)} \qquad \lim_{n\to\infty} \sup_{f\in F_2} B_n(f) = 0,$$

et telle que pour tout couple $(f_1, f_2) \in \mathcal{F}_2$,

$$\textbf{(KU2b)} \qquad \lim_{n\to\infty} \frac{1}{\sigma_{k_n}(f_1)\sigma_{k_n}(f_2)} \sum_{j=1}^{k_n} f_1(j)f_2(j)j^{-2} = \Gamma(f_1,f_2) < \infty,$$

où

$$\sigma^2_{k_n}(f) = \sum_{j=1}^{k_n} f(j)^2 j^{-2} \quad \text{et} \quad a_{k_n}(f) = \sum_{j=1}^{k_n} f(j)j^{-1}. \tag{3.2}$$

Nous supposerons dans toute la suite que $\mathcal{F}_h,\ h = 1,2$ est totalement bornée et est munie d'une semi-métrique ρ_h.

Notre plus grande réussite sera la description complète de la convergence vague de la suite

$$\{\boldsymbol{T}_{k_n}(f), f \in \mathcal{F}_h\},\ h = 1,2,$$

dans l'espace $\ell^\infty(\mathcal{F}_h)$ des fonctions réelles bornées définies dans $\mathcal{F}_h$. Ces résultats sont donnés par les Théorèmes 3.2.1 et 3.3.1 ci-dessous à la lumière de cette théorie moderne. En outre, nous développerons des cas d'études réels avec des classes explicites en application avec des résultats généraux, dans les corollaires de la Section 3.2. Cette approche donne un très grand nombre d'estimateurs de l'indice de queues $\boldsymbol{\gamma}$ pour les distributions F appartenant aux domaines d'attraction des valeurs extrêmes. Toutefois, ce chapitre portera essentiellement sur les lois fonctionnelles et uniformes du processus décrit ci-dessus.

On utilisera dans ce chapitre des résultats techniques de la théorie des valeurs extrêmes dont quelques notions de base ont été données dans le Chapitre 1. La Section 3.2 est

consacrée aux distributions asymptotiques ponctuelles de $\boldsymbol{T}_{k_n}(f)$, tandis que les résultats fonctionnels sont indiqués et établis dans la Section 3.3. Dans la Section 3.4, nous étudions certains cas particuliers, notamment la famille $\{f(j) = j^{\boldsymbol{\tau}}, \boldsymbol{\tau} > 0\}$, utilisé dans Lo *et al.* (2012).

Nous attirons l'attention du lecteur que nous allons traiter la convergence uniforme sur des familles de variables aléatoires indexées par des fonctions définies sur $\mathcal{F}_h$. Il n'est pas sûr que le supremum de ces familles de variables aléatoires soit mesurable. Nous sommes donc obligés d'utiliser la convergence en probabilité extérieure plutôt que la convergence en probabilité dans le cas d'une absence de mesurabilité. Rappelons qu'une application $(Z_n)_{n\geq 1}$ définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}$ converge vers zéro en probabilité extérieure quand $n \to \infty$ si et seulement si il existe une suite réelle et non négative de variables aléatoires $(u_n)_{n\geq 1}$ définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ telle que: $\forall\, n \geq 1,\ |Zn| \leq |u_n|$ et u_n tend vers zéro en probabilité quand $n \to \infty$. Ceci est également noté $Z_n = o^*_{\mathbb{P}}(1)$. Voir par exemple van der Vaart et Wellner (1996) pour plus de détails sur cette théorie.

3.2 Distributions finies et ponctuelles de $\boldsymbol{T}_{k_n}(f)$

Quelques représentations classiques d'une distribution F pour qu'elle appartienne au domaine d'attraction d'une distribution des valeurs extrêmes non dégénérée ont été données dans les chapitres 1 et 2. Dans la suite, on suppose que $F(1) = 0$ et nous considérons la suite de variables aléatoires $Y_1, Y_2, ...$, liée aux observations $X_1, X_2...$ par la transformation $Y_j = \log X_j$, $j \geq 1$ et de fonction de répartition $G(y) = F(e^y)$.
Dans le Chapitre 2, une représentation de $Y_1, Y_2, ...$ par $G^{-1}(1-\xi_1), G^{-1}(1-\xi_2), ...$ est donnée, où $\xi_1, \xi_2, ...$, est une suite de variables aléatoires indépendantes et uniformément distribuées sur $(0,1)$. Dans la Section 2.1.2, équations (2.4), et (2.5) une représentation appropriée d'appartenance au domaine d'attraction $D(\boldsymbol{\Lambda}) \cup D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$ pour les distributions F, est donnée par $G^{\leftarrow}(1-u) = \log F^{\leftarrow}(1-u)$, $u \in (0,1)$ où $F^{\leftarrow}$ et $G^{\leftarrow}$ sont respectivement les inverses généralisés de F et G. Maintenant introduisons certaines conditions sur la fonction de répartition G à travers les fonctions p et Δ données dans les représentations (2.4), (2.6) et (2.5) dans le chapitre 2. Soit $\lambda > 1$, considérons

$$0 \leq g_{1,n}(p, \lambda) = \sup_{0 \leq u \leq \lambda k_n/n} |p(u)|\,,$$

$$g_{2,n}(\Delta, \lambda) = \sup_{0 \leq u \leq \lambda k_n/n} |\Delta(u)|\,,$$

et

$$d_n(p, \Delta, \lambda) = \max(g_{1,n}(p, \lambda), g_{2,n}(\Delta, \lambda) \log k_n).$$

Nous aurons donc besoin de ces conditions ci-dessous pour $\lambda > 1$:

$$(\mathbf{C1}) \qquad \frac{g_{1,n}(p,\lambda)}{\sigma_{k_n}(f)} \sum_{j=1}^{k_n} f(j) \to 0, \text{ quand } n \to \infty, \tag{3.3}$$

$$(\mathbf{C2}) \qquad \frac{g_{2,n}(\Delta,\lambda)}{\sigma_{k_n}(f)} \sum_{j=1}^{k_n} f(j) \to 0, \text{ quand } n \to \infty, \tag{3.4}$$

et

$$(\mathbf{C3}) \qquad \sup_{f \in \mathcal{F}_h} \frac{d_n(p,\Delta,\lambda)}{\sigma_{k_n}(f)} \sum_{j=1}^{k_n} f(j) \to 0 \text{ quand } n \to \infty., \tag{3.5}$$

Dans tout ce qui suit sauf mention particulière toutes les limites sont considérées avec $n \to \infty$. Nous sommes en mesure d'établir le théoréme ci-dessous.

Théorème 3.2.1

• *Soit $F \in D(\boldsymbol{\Lambda})$ (cas $\boldsymbol{\gamma} = 0$). Supposons que la condition* (**C3**) *soit vérifiée, alors sous l'hypothèse* (**K2**) *on a*

$$\frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

et sous l'hypothèse (**K1**) *on a*

$$\frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right) \xrightarrow{\mathcal{D}} \mathcal{L}(f),$$

où $s(\cdot)$ est la fonction donnée dans le Chapitre 2, (2.5) et (2.6) et

$$\mathcal{L}(f) = \frac{1}{\sqrt{A(2,f)}} \sum_{j=1}^{\infty} \frac{f(j)}{j}(E_j - 1).$$

• *Soit $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$. Supposons que* (**C1**) *et* (**C2**) *soient vérifiées, alors sous l'hypothèse* (**K2**) *ona*

$$\frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

et sous l'hypothèse (**K1**) *on a*

$$\frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right) \xrightarrow{\mathcal{D}} \mathcal{L}(f)$$

Par le théorème précédent, nous obtenons deux résultats intermédiaires à l'égard des lois fonctionnelles. La première concerne la loi asymptotique en distributions finies. Quelques fondements mathématiques sur la théorie moderne de convergence en mesure sont donnés dans l'annexe A. Dans out ce qui suit $\overset{\mathcal{D}}{\rightsquigarrow}$ désigne la convergence vague.

Corollaire 3.2.2 *Supposons que les conditions* (**C1**), (**C2**) *et* (**C3**) *soient vérifiées et pour tout* $(f_1, f_2) \in \mathcal{F}_1^2 \cup \mathcal{F}_2^2$,

$$\lim_{n\to\infty} \frac{1}{\sigma_{k_n}(f_1)\sigma_{k_n}(f_2)} \sum_{j=1}^{k_n} f_1(j) f_2(j) j^{-2} = \Gamma(f_1, f_2) < \infty.$$

- *Si* $F \in D(\boldsymbol{\Lambda})$, *alors la distribution finie de*

$$\left\{ \frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right), f \in \mathcal{F}_1 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathcal{L}(f), f \in \mathcal{F}_1\}$$

et celle de

$$\left\{ \frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right), f \in \mathcal{F}_2 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathbb{G}(f), f \in \mathcal{F}_2\}$$

- *Si* $F \in D(\boldsymbol{\varphi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$, *alors la distribution finie de*

$$\left\{ \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right), f \in \mathcal{F}_1 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathcal{L}(f), f \in \mathcal{F}_1\}$$

et celle de

$$\left\{ \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right), f \in \mathcal{F}_2 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathbb{G}(f), f \in \mathcal{F}_2\}$$

avec $\mathbb{G}$ *un processus Gaussien centré de fonction de covariance*

$$\Gamma(f_1, f_2) = \lim_{n\to\infty} \frac{1}{\sigma_n(f_1)\sigma_{k_n}(f_2)} \sum_{j=1}^{k_n} f_1(j) f_2(j) j^{-2}.$$

Ce corollaire permet de faire une bonne transition vers la loi fonctionnelle. Afin d'aller plus loin, nous aurons besoin de ces conditions uniformes suivantes. Définissons pour tout $\lambda > 1$,

$$(\mathbf{CU1}) \qquad \sup_{f\in\mathcal{F}_h} \frac{g_{1,n}(p,\lambda)}{\sigma_{k_n}(f)} \sum_{j=1}^{k_n} f(j) \to 0, \tag{3.6}$$

$$(\mathbf{CU2}) \qquad \sup_{f\in\mathcal{F}_h} \frac{g_{2,n}(\Delta,\lambda)}{\sigma_{k_n}(f)} \sum_{j=1}^{k_n} f(j) \tag{3.7}$$

et

$$(\mathbf{CU3}) \qquad \sup_{f\in\mathcal{F}_h} \frac{d_n(p,\Delta,\lambda)}{\sigma_{k_n}(f)} \sum_{j=1}^{k_n} f(j) \to 0. \tag{3.8}$$

Maintenant, considérons les expressions suivantes:

$$V_{k_n}(f) = \sum_{j=1}^{k_n} f(j) \log\left(\frac{\xi_{j+1,n}}{\xi_{j,n}}\right) \quad \text{et} \quad \overline{V}_{k_n}(f) = \frac{a_{k_n}(f)}{\sigma_{k_n}(f)}\left(\frac{V_{k_n}(f)}{a_{k_n}(f)} - 1\right).$$

Corollaire 3.2.3 *Supposons que les conditions uniformes* (**KU1**), (**KU2a**), (**KU2b**), (**CU1**), (**CU2**) *et* (**CU3**) *soient vérifiées. Soit $\mathcal{F}_a$ l'ensemble non vide des fonctions satisfaisant* (**KU1**) *ou bien l'ensemble non vide des fonctions satisfaisant* (**KU2a**) *et* (**KU2b**). *Supposons que $\{\overline{V}_{k_n}(f), f \in \mathcal{F}_a\}$ converge vaguement dans $\ell^\infty(\mathcal{F}_a)$.*

- *Si $F \in D(\boldsymbol{\Lambda})$, alors uniformément dans $f \in \mathcal{F}_a$,*

$$\frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)}\left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n)\right) \stackrel{\mathcal{D}}{=} \overline{V}_{k_n}(f) + o^*_{\mathbb{P}}(1)$$

- *Si $F \in D(\boldsymbol{\varphi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$, alors uniformément dans $f \in \mathcal{F}_a$,*

$$\frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)}\left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma}\right) \stackrel{\mathcal{D}}{=} \overline{V}_{k_n}(f) + o^*_{\mathbb{P}}(1).$$

3.3 Loi fonctionnelle de $\boldsymbol{T}_{k_n}(f)$

Notons

$$Z_{j,n}(f) := \frac{f(j)j^{-1}}{\sigma_{k_n}(f)}(E_j - 1).$$

En utilisant la représentation (2.2) dans le chapitre 2, nous avons

$$\overline{V}_{k_n}(f) = \frac{a_{k_n}(f)}{\sigma_{k_n}(f)}\left(\frac{V_{k_n}(f)}{a_{k_n}(f)} - 1\right) \stackrel{\mathcal{D}}{=} \sum_{j=1}^{k_n} Z_{j,n}(f).$$

Nos principaux outils pour manipuler ces processus stochastiques se trouvent dans les Théorèmes 2.11.1 et 2.11.9 de van der Vaart et Wellner (1996) sur les lois uniformes de sommes de processus stochastiques indépendants. Nous aurons donc besoin des outils de bases de ces théorèmes.

Considérons $\mathcal{F}_h$ $(h = 1, 2)$, les sous-familles de $\mathcal{F}$ formées par les fonctions $f\colon \mathbb{N}^* \longmapsto \mathbb{R}^*_+$ satisfaisant l'hypothèse de mesurabilité ci-dessous:

Pour tout $h \in \{1,2\}$, $\forall \delta > 0$, $(e_1, \cdots, e_n) \in \{-1,0,1\}^n$, et $p = 1,2$,

$$\textbf{(MES)} \qquad \sup_{(f_1,f_2)\in\mathcal{F}_h, \rho(f_1,f_2)\leq\delta} \sum_{j=1}^{k_n} e_j \left|\sigma_{k_n}^{-1}(f_1)f_1(j)j^{-1} - \sigma_n^{-1}(f_2)f_2(j)j^{-1}\right|^p E_j^p, \tag{3.9}$$

est mesurable, avec ρ une semi-métrique sur l'espace $\mathcal{F}$. Plus précisément, nous introduisons dans cette classe, la sous famille $\mathcal{F}_{h,0}$ de $\mathcal{F}_h$ $(h = 1,2)$ telle que chacune d'entre elles soit munie d'une semi-métrique ρ_h telles que $(\mathcal{F}_{h,0}, \rho_h)$ soit totalement bornée et la condition de mesurabilité (**MES**) soit vérifiée. Nous aurons également besoin de la semi-metrique aléatoire suivante:

$$d_n^2(f_1, f_2) = \sum_{j=1}^{k_n} \left(\frac{f_1(j)}{j\sigma_{k_n}(f_1)} - \frac{f_2(j)}{j\sigma_{k_n}(f_2)} \right)^2 (E_j - 1)^2.$$

Maintenant, on supose que $(\mathcal{F}_{h,0}, \|\cdot\|)$ est un espace normé vérifiant la propriété de Riesz: pour tout couple de fonctions $(f,g) \in \mathcal{F}_{h,0}^2$, si $|f| \leq |g|$, alors $||f|| \leq ||g||$.
Définissons comme dans van der Vaart et Wellner (1996) (p. 211), le nombre de crochets $N_{[\,]}(\varepsilon, \mathcal{F}_{h,0}, L_2^n)$ correspondant au nombre minimal N_ε d'ensembles dans une partition $\mathcal{F}_{h,0} = \bigcup_{j=1}^{N_\varepsilon} \mathcal{F}_{\varepsilon j}^n$ de l'ensemble d'indices d'ensembles $\mathcal{F}_{\varepsilon j}^n$ tel que, pour toute partition de l'ensemble $\mathcal{F}_{\varepsilon j}^n$,

$$\sum_{i=1}^{k_n} \mathbb{E}^* \left(\sup_{f,g\in F_{\varepsilon j}^n} |Z_{ni}(f_1) - Z_{ni}(f_2)|^2 \right) \leq \varepsilon^2,$$

où $\mathbb{E}^*$ représente l'intégrale extérieure (voir Annexe A).

Le théorème ci-dessous établit notre première version de la loi fonctionnelle de $\boldsymbol{T}_{k_n}(f)$.

Théorème 3.3.1 *Supposons que pour tout $h = 1,2$, nous avons*

$$\textbf{(L1)} \qquad \sup_{(f_1,f_2)\in\mathcal{F}_{h,0}^2, \rho_h(f_1,f_2)\leq\delta_n} \sum_{j=1}^{k_n} \left(\frac{f_1(j)}{j\sigma_{k_n}(f_1)} - \frac{f_2(j)}{j\sigma_{k_n}(f_2)} \right)^2 \to 0 \tag{3.10}$$

avec $\delta_n \to 0$ quand $n \to \infty$, et

$$\textbf{(L2)} \qquad \int_0^{\delta_n} \sqrt{logN_{[\,]}(\varepsilon, \mathcal{F}_h, d_n)}d\varepsilon \to 0, \tag{3.11}$$

quand $n \to \infty$, où $N_{[\,]}(\varepsilon, \mathcal{F}_{h,0}, d_n)$ est le ε-nombre d'entropies de $\mathcal{F}_{h,0}$ munie de sa semi-métrique d_n, et qui représente le nombre minimal de d_n-boules de rayon d'au plus ε utilisé

pour couvrir $\mathcal{F}_h$.

• *Soit $F \in D(\boldsymbol{\Lambda})$ et supposons que* (**CU3**) *soit satisfaite, alors*

$$\left\{ \frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right), f \in \mathcal{F}_{2,0} \right\}$$

converge vers un processus Gaussien dans $\ell^\infty(\mathcal{F}_{2,0})$, avec une fonction de covariance

$$\Gamma(f_1, f_2) = \lim_{n\to\infty} \sum_{j=1}^{k_n} \{\sigma_{k_n}(f_1)\sigma_{k_n}(f_2)\}^{-1} f_1(j) f_2(j) j^{-2} \leq 1.$$

et

$$\left\{ \frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right), f \in \mathcal{F}_{1,0} \right\}$$

converge vers un processus stochastique $\{\mathcal{L}(f),\ f \in \mathcal{F}_{1,0}\}$ dans $\ell^\infty(\mathcal{F}_{1,0})$ avec une fonction de covariance $\Gamma(f_1, f_2)$. La distribution finie de $(\mathcal{L}(f_1), ..., \mathcal{L}(f_S))$, $S \in \mathbb{N}^*$, est caractérisée par la fonction génératrice des moments

$$(t_1, ..., t_S) \mapsto \prod_{j=1}^{+\infty} \exp\left(\sum_{q=1}^{S} t_q f_q(j) j^{-1} \right) \left(1 - \sum_{q=1}^{S} t_q f_q(j) j^{-1} \right), \tag{3.12}$$

pour $|t_q| \leq 1/S$, $q = 1, ..., S$.

• *Soit $F \in D(\phi_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$ et supposons que les conditions* (**CU1**) *et* (**CU2**) *soient verifiées. Alors*

$$\left\{ \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right), f \in \mathcal{F}_{2,0} \right\}$$

converge vers un processus Gaussien dans $\ell^\infty(\mathcal{F}_{2,0})$ avec une fonction de covariance Γ. Et

$$\left\{ \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right), f \in \mathcal{F}_{1,0} \right\}$$

converge vers un processus stochastique dans $\ell^\infty(\mathcal{F}_{1,0})$ avec une fonction de covariance $\Gamma(\cdot)$ et une distribution finie caracterisée par (3.12).

Maintenant, nous allons présenter la seconde version qui est plus générale puisqu'elle ne nécéssite pas l'hyposthèse de mesurabilité de telle sorte que l'on considère l'ensemble d'espaces $\mathcal{F}_h$ $(h = 1, 2)$.

Théorème 3.3.2 *Si* $(\mathcal{F}_h, \| \|)$ *est un espace normé verifiant la propriété de Riesz, alors les resultats du Théorème 3.3.1 sont vérifiés lorsque*

$$\textbf{(L3)} \qquad \int_0^{\delta_n} \sqrt{logN_{[\ ]}(\varepsilon, \mathcal{F}_h, L_2^n)}d\varepsilon \to 0, \textit{quand}\, n \to \infty, \tag{3.13}$$

à la place de (**L2**)*, et à condition que* $Z_{j,n}$ *admette des moments d'ordre deux finis.*

3.4 Quelques cas speciaux

Nous nous spécialisons ces résultats pour une classe spéciale de fonctions monotones $f_{\boldsymbol{\tau}}(j) = j^{\boldsymbol{\tau}}$, $\boldsymbol{\tau} > 0$. Comment on peut appliquer des lois particulières pour des classes spéciales sur nos résultats? Nous savons à partir de Lo *et al.* (2012), (voir par exemple le Chapitre 2) et Diop et Lo (2009) que $T_n(f_\tau)$ est asymptotiquement normal pour $\boldsymbol{\tau} \geq 1/2$, tandis qu' il suit asymptotiquement une loi de type $\mathcal{L}(f_{\boldsymbol{\tau}})$ pour $0 < \boldsymbol{\tau} < 1/2$, sous les conditions habituelles de la fonction de répartition G. Ilsera procédé ici au traitement du comportement asymptotique uniforme pour ces deux ensembles de valeurs de $\boldsymbol{\tau}$: $]0, 1/2[$ et $[1/2, +\infty[$. Pour le premier cas, on applique le Théorème 3.3.1 et pour le second, le Théorème 3.3.2. Soit $0 < a < b < 1/2$ et considérons $\mathcal{F}_0(a, b) = \{f(j) = j^{\boldsymbol{\tau}}, a \leq \boldsymbol{\tau} \leq b\}$. Nous avons:

Corollaire 3.4.1 *Soient* $0 < a < b < 1/2$ *et* $\mathcal{F}_0 = \mathcal{F}_0(a, b) = \{f(j) = j^{\boldsymbol{\tau}}, 0 < a \leq \boldsymbol{\tau} \leq b < 1/2\}$. *Alors,*

- *Pour* $F \in D(\boldsymbol{\Lambda})$*, si* (**CU2**) (**CU3**) *sont vérifiées, alors*

$$\left\{ \frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right), f \in \mathcal{F}_0 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathcal{L}(f), f \in \mathcal{F}_0\}.$$

- *Pour* $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$*,* $\boldsymbol{\gamma} > 0$*, si* (**CU1**) *est verifiée, alors*

$$\left\{ \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right), f \in \mathcal{F}_0 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathcal{L}(f), f \in \mathcal{F}_0\}.$$

Corollaire 3.4.2 *Soient* $b > 1$*,* $1/2 < a < b$ *et* $\mathcal{F}_1(a, b) = \{f(j) = j^{\boldsymbol{\tau}}, 0 < a \leq \boldsymbol{\tau} \leq b\}$.

- *Pour $F \in D(\boldsymbol{\Lambda})$, si* (**CU3**) *est vérifiée, alors*

$$\left\{ \frac{a_{k_n}(f)}{s(k_n/n)\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n) \right), f \in \mathcal{F}_0 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathbb{G}(f), f \in \mathcal{F}_0\}.$$

- *Pour $F \in D(\boldsymbol{\phi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$, si* (**CU1**) *est vérifiée alors*

$$\left\{ \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)} \left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma} \right), f \in \mathcal{F}_0 \right\} \overset{\mathcal{D}}{\rightsquigarrow} \{\mathbb{G}(f), f \in \mathcal{F}_0\}.$$

avec $\mathbb{G}$ un processus Gaussien centré de fonction de covariance $\Gamma(\cdot)$.

3.5 Démonstrations

3.5.1 Résultats préliminaires

Nous commençons par le lemme ci-dessous où on suppose qu'on dispose d'une séquence de variables indépendantes et uniformément distribuées $\xi_1, \xi_2, ...$ comme dans la Section 2.4.

Lemme 3.5.1 *Soit*

$$V_{k_n}(f) = \sum_{j=1}^{k_n} f(j) \left(\log \xi_{j+1,n} - \log \xi_{j,n} \right).$$

- *Si* (**K2**) *est verifiée, alors*

$$\frac{a_{k_n}(f)}{\sigma_{k_n}(f)} \left(\frac{V_{k_n}(f)}{a_{k_n}(f)} - 1 \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

- *Si* (**K1**) *est vérifiée, alors*

$$\frac{a_{k_n}(f)}{\sigma_{k_n}(f)} \left(\frac{V_{k_n}(f)}{a_{k_n}(f)} - 1 \right) \xrightarrow{\mathcal{D}} \mathcal{L}(f),$$

où $a_{k_n}(f)$ et $\sigma_{k_n}(f)$ sont définis dans (3.2) et

$$\mathcal{L}(f)) = \frac{1}{\sqrt{A(2,f)}} \sum_{j=1}^{\infty} f(j) j^{-1} (E_j - 1),$$

est une variable aléatoire centrée et reduite de moments finis.

Démonstration du Lemme 3.5.1: Par la représenation de Malmquist (2.2), nous avons $V_{k_n}(f) = \sum_{j=1}^{k_n} f(j)j^{-1}E_j$. Il s'en suit que $\mathbb{E}(V_{k_n}(f)) = a_{k_n}(f)$ et sa variance $\mathbb{V}ar(V_{k_n}(f)) = \sigma^2_{k_n}(f)$. Posons

$$\overline{V}_{k_n}(f) = \frac{a_{k_n}(f)}{\sigma_{k_n}(f)}\left(\frac{V_{k_n}(f)}{a_{k_n}(f)} - 1\right)$$

et

$$Z_{j,n}(f) := \frac{f(j)j^{-1}}{\sigma_{k_n}(f)}(E_j - 1).$$

Nous avons déjà mentionné que

$$\overline{V}_{k_n}(f) \stackrel{\mathcal{D}}{=} \sum_{j=1}^{k_n} Z_{j,n}(f).$$

On suppose dans un premier temps que (**K2**) est vérifiée. Evaluons la fonction génératrice des moments de $\overline{V}_{k_n}(f)$:

$$\phi_{\overline{V}_{k_n}(f)}(t) = \prod_{j=1}^{k_n} \phi_{(E_j-1)}\left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right). \tag{3.14}$$

Rappelons que dans ce cas $\sigma_{k_n}(f) \uparrow \infty$. Donc d'après (**K2**), pour tout $u_0 > 0$ et t fixé, et pour k assez grand,

$$\left|\frac{tf(j)}{j\sigma_{k_n}(f)}\right| \leq u_0 \tag{3.15}$$

uniformément en $j \geq 1$. A cette étape, on utilise une expansion de $\phi_{(E_j-1)}$ au voisinage de zéro:

$$\psi_{(E_j-1)}(u) = 1 + u^2/2 + u^3 g(u),$$

où il exite u_0 tel que $0 \leq u \leq u_0 \Rightarrow |g(u)| \leq 1$. En utilisant la bornitude uniforme dans (3.15), on obtient

$$\phi_{(E_j-1)}\left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right) = 1 + \frac{1}{2}\left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right)^2 + \left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right)^3 g_{0,j,n}(t),$$

où $|g_{0,j,n}(t)| \leq 1$ for all $1 \leq j \leq k$. Par la bornitude uniforme du terme d'erreur, nous avons

$$\log \phi_{(E_j-1)}\left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right) = \frac{1}{2}\left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right)^2 + \left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right)^3 g_{0,j,n}(t) + \left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right)^3 g_{1,j,n}(t),$$

où $|g_{1,j,n}(t)| \leq 1$ pour tout $1 \leq j \leq k$. Finalement

$$\begin{aligned}\phi_{\overline{V}_{k_n}(f)}(t) &= \exp\left(\sum_{j=1}^{k_n} \log \phi_{(E_j-1)}\left(\frac{tf(j)}{j\sigma_{k_n}(f)}\right)\right) \\ &= \exp\left(t^2/2 + g_{2,j,n}(t) \times \left(\frac{t}{\sigma_{k_n}(f)}\right)^3 \sum_{j=1}^{k_n}\left(\frac{f(j)}{j}\right)^3\right),\end{aligned}$$

avec $|g_{1,j,n}(t)| \leq 2$ pour tout $1 \leq j \leq k$. Comme

$$0 \leq \left(\frac{1}{\sigma_{k_n}(f)}\right)^3 \sum_{j=1}^{k_n}\left(\frac{f(j)}{j}\right)^3 \leq \frac{B(n,f)}{\sigma_{k_n}(f)^2} \times \sum_{j=1}^{k_n}\left(\frac{f(j)}{j}\right)^2 = B(n,f) \to 0.$$

On obtient alors $\phi_{\overline{V}_{k_n}(f)}(t) \to \exp(t^2/2)$ par conséquent $\overline{V}_{k_n}(f) \to \mathcal{N}(0,1)$. Ce qui conclut la première partie de la preuve du lemme.

Supposons maintenant que (**K1**) est vérifiée, donc $\sigma_{k_n}(f) \to \sqrt{A(2,f)} \in]0,1[$. Alors

$$\overline{V}_{k_n}(f) \xrightarrow{\mathcal{D}} \frac{1}{\sqrt{A(2,f)}} \sum_{j=1}^{\infty} \frac{f(j)}{j}(E_j - 1) = \mathcal{L}(f).$$

Maintenant nous allons montrer que $\mathcal{L}(f)$ définit bien une variable aléatoire ayant tous ses moments finis. La fonction caractéristique de $\overline{V}_{k_n}(f)$ est donnée par

$$\psi_{\overline{V}_{k_n}(f)}(t) = \exp\left(-\frac{1}{\sqrt{A(2,f)}}\sum_{j=1}^{k_n} it\frac{f(j)}{j}\right)\prod_{j=1}^{k_n}\left(1 - it\frac{f(j)}{j\sqrt{A(2,f)}}\right).$$

En utilisant un développement de $\log(1-\cdot)$, et le théorème de Lebesgues, on montre que

$$\psi_{\overline{V}_{k_n}(\tau)}(t) = \exp\left(\sum_{j=1}^{k_n}\sum_{n=2}^{\infty}\frac{(it)^n}{n} \times \left(\frac{f(j)}{j\sqrt{A(2,f)}}\right)^n\right) \tag{3.16}$$

$$\to \psi_\infty(t) = \exp\left(\sum_{n=2}^{\infty}\frac{(it)^n}{n} \times \frac{A(n,f)}{A(2,f)^{n/2}}\right). \tag{3.17}$$

Notons que si $A(2,f) < \infty$, alors on a aussi $A(n,f) < \infty$ pour tout $n \geq 2$, puisque le terme général (en j) de $A(n,f)$ est inférieur à celui de $A(2,f)$, pour de grandes valeurs de j. Ceci conclut la deuxième partie du lemme quand (**K1**) est vérifiée. ■

Lemme 3.5.2 *Soit $(a,b) \in \mathbb{R}^2$ et on suppose que pour tout couple $(f_1,f_2) \in \mathcal{F}_h^2$, $h = 1,2$,*

$$\lim_{n\to\infty} \frac{1}{\sigma_{k_n}(f_1)\sigma_{k_n}(f_2)} \sum_{j=1}^{k_n} f_1(j)f_2(j)j^{-2} = \Gamma(f_1,f_2) < \infty.$$

Alors pour $h = 1$, nous avons

$$a\overline{V}_{k_n}(f_1) + b\overline{V}_{k_n}(f_2) \overset{\mathcal{D}}{\rightsquigarrow} a\mathcal{L}(f_1) + b\mathcal{L}(f_2),$$

et pour $h = 2$, nous avons

$$a\overline{V}_{k_n}(f_1) + b\overline{V}_{k_n}(f_2) \overset{\mathcal{D}}{\rightsquigarrow} \mathcal{N}\left(0, v(a, b, f_1, f_2)\right)$$

où $v(a, b, f_1, f_2) = a^2 + b^2 + 2ab\Gamma(f_1, f_2)$. Dans les deux cas, les distributions finies de $\{\overline{V}_{k_n}(f), f \in \mathcal{F}_h\}$ convergent vaguement vers celles du processus $\{\mathcal{L}(f), f \in \mathcal{F}_h\}$ pour $h = 1$ et vers celle du processus Gaussien $\{\mathbb{G}(f), f \in \mathcal{F}_h\}$ pour $h = 2$, de fonction de covariance $\Gamma(f_1, f_2)$.

Démonstration du Lemme 3.5.2: Le cas $h = 1$ est simple. Pour $h = 2$, nous allons adapter la preuve du lemme précédent à partir de (3.14). Posons

$$V_{k_n}(a, b, f_1, f_2) = a\overline{V}_{k_n}(f_1) + b\overline{V}_{k_n}(f_2)$$

et

$$c(n, j) = a\frac{f_1(j)}{j\sigma_{k_n}(f_1)} + b\frac{f_2(j)}{j\sigma_{k_n}(f_2)}.$$

Nous avons toujours, pour $(f_1, f_2) \in \mathcal{F}_2^2$, $\sup_j |c(n, j)| \to 0$ et

$$\begin{aligned}\sum_{j=1}^{k_n} c(n, j)^2 &= a^2 + b^2 + 2ab\frac{1}{\sigma_{k_n}(f_1)\sigma_{k_n}(f_2)}\sum_{j=1}^{k_n} f_1(j)f_2(j)j^{-2} \\ &\longrightarrow v(a, b, f_1, f_2) = a^2 + b^2 + 2ab\Gamma(f_1, f_2).\end{aligned}$$

En utilisant les mêmes arguments que dans (3.14) et (3.16), nous obtenons

$$\begin{aligned}\phi_{V_{k_n}(a,b,f_1,f_2)}(t) &= \exp\left(\sum_{j=1}^{k_n} \log \psi_{(E_j-1)}(t\ c(n, j)\right) \\ &= \exp\left(\frac{t^2}{2}\left\{\sum_{j=1}^{k_n} c(n, j)^2\right\} + g_{3,j,n}(t) \times t^3\left(\sum_{j=1}^{k_n} c(n, j)^3\right)\right),\end{aligned}$$

avec $|g_{3,j,n}(t)| \leq 2$, pour tout $1 \leq j \leq k$. Puisque

$$0 \leq \sum_{j=1}^{k_n} c(n, j)^3 \leq \Big(aB(n, f_1) + bB(n, f_2)\Big) \times \sum_{j=1}^{k_n} c(n, j)^2 \to 0,$$

nous obtenons

$$\phi_{V_{k_n}(a,b,f_1,f_2)}(t) \to \exp\left(v(a,b,f_1,f_2)\frac{t^2}{2}\right).$$

Ce qui achève la preuve. Maintenant, ces méthodes peuvent être reproduites pour toute combinaison linéaire finie

$$a_1\overline{V}_{k_n}(f_1) + a_2\overline{V}_{k_n}(f_2) + ... + a_S V^*_{k_n}(f_S).$$

Dans tous ces cas, nous avons trouvé la même loi en distribution finie. Pour $h = 1$, $(\overline{V}_{k_n}(f_1), ..., V_{k_n}(f_S))$ converge en loi vers $(\mathcal{L}(f_1), ..., \mathcal{L}(f_S))$ et pour $h = 2$, $(V^*_{k_n}(f_1), ..., \overline{V}_{k_n}(f_S))$ converge vers un vecteur Gaussien de matrice de variance covariance $(\Gamma(f_i, f_j), 1 \leq i, j \leq S)$. ■

3.5.2 Démonstration des résultats théoriques

Démonstration du Théorème 3.2.1: Voir Annexe B.

Démonstration du Corollaire 3.2.2: Soit f fixée dans $\mathcal{F}_h, h = 1, 2$. Posons

$$V_{k_n}(0, f) = \frac{a_{k_n}(f)}{(s(k_n/n)\sigma_{k_n}(f)}\left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n)\right)$$

et

$$V_{k_n}(1, f) = \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)}\left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma}\right).$$

Pour une famille $(f_1, f_2, ..., f_S) \in \mathcal{F}_h^S, h = 1, 2,\ S \in \mathbb{N}^*$, nous avons par la preuve du Théorème 3.2.1 dans l'Annexe que pour tout $1 \leq i \leq S$,

- Pour $F \in D(\boldsymbol{\Lambda})$

$$V_{k_n}(0, f_i) = \overline{V}_{k_n}(f_i) + o_{\mathbb{P}}(1) \tag{3.18}$$

- Pour $F \in D(\boldsymbol{\varphi}_{1/\boldsymbol{\gamma}})$, $\boldsymbol{\gamma} > 0$

$$V_{k_n}(1, f_i) = \overline{V}_{k_n}(f_i) + o_{\mathbb{P}}(1). \tag{3.19}$$

Ainsi, nous pouvons conclure en appliquant le Lemme 3.5.2 qui établit les lois asymptotiques de $(V^*_{k_n}(f_1), ..., \overline{V}_{k_n}(f_S))$ sous les hypothèses du lemme. ■

Démonstration du Corollaire 3.2.3: Posons

$$V_{k_n}(0, f) = \frac{a_{k_n}(f)}{(s(k_n/n)\sigma_{k_n}(f)}\left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - s(k_n/n)\right)$$

et

$$V_{k_n}(1,f) = \frac{a_{k_n}(f)}{\boldsymbol{\gamma}\sigma_{k_n}(f)}\left(\frac{\boldsymbol{T}_{k_n}(f)}{a_{k_n}(f)} - \boldsymbol{\gamma}\right).$$

Sous les hypothèses uniformes (**KU1**) ou (**KU2a**) et (**KU2b**), nous avons sûrement

$$V_{k_n}(0,f) = \overline{V}_{k_n}(f)(1+o^*_{\mathbb{P}}(1)) + o^*_{\mathbb{P}}(1) \;\text{ et } V_{k_n}(1,f) = \overline{V}_{k_n}(f)(1+o^*_{\mathbb{P}}(1)) + o^*_{\mathbb{P}}(1),$$

uniformément dans $f \in \mathcal{F}_h$. Maintenant, soit $\mathcal{F}_a$, un sous ensemble de $\mathcal{F}_1$ tel que $\{\overline{V}_{k_n}(f), f \in \mathcal{F}_a\}$ converge vaguement vers $\{\mathbb{G}(f), f \in \mathcal{F}_a\}$ dans $\ell^\infty(\mathcal{F}_a)$, alors $\|\mathbb{G}\|^*_{\mathcal{F}_a} < \infty$, où $\|\cdot\|^*_{\mathcal{F}_a}$ représente la couverture minimale mesurable de la norme de convergence uniforme $\|\cdot\|_{\mathcal{F}_a}$ dans $\mathcal{F}_a$ (on pourra toujours se référer dans l'Annexe A). Puisque, par le théorème de continuité, $\left\|\overline{V}_{k_n}\right\|_{\mathcal{F}_a} \overset{\mathcal{D}}{\rightsquigarrow} \|\mathbb{G}\|^*_{\mathcal{F}_a}$, nous obtenons

$$V_{k_n}(0,f) = \overline{V}_{k_n}(f) + o^*_{\mathbb{P}}(1), \tag{3.20}$$

uniformément dans $f \in \mathcal{F}_a$, pour $F \in D(\boldsymbol{\Lambda})$. Les autres cas sont montrés de façon similaire. ■

Démonstration du Théorème 3.3.1 D'après la preuve du Corollaire 3.2.3, nous avons:

$$\boldsymbol{T}^*_{k_n}(f) = (s(k_n/n)\sigma_{k_n}(f))^{-1}(\boldsymbol{T}_{k_n}(f) - a_{k_n}(f)s(k_n/n)), \text{ pour } \; F \in D(\boldsymbol{\Lambda})$$

ou bien

$$T^*_n(f) = (a_{k_n}(f)/\sigma_n(f))(\boldsymbol{T}_{k_n}(f)/a_{k_n} - \boldsymbol{\gamma}), \text{ pour } F \in D(G_{1/\boldsymbol{\gamma}}),$$

et on a

$$T^*_n(f) = \overline{V}_{k_n}(f) + o^*_P(1),$$

uniforment dans $f \in \mathcal{F}_h$ et donc uniformément dans $f \in \mathcal{F}_{h,0}$ $(h = 1,2)$. De là nous appliquons le Théorème 2.11.1 de van der Vaart et Wellner (1996) sur le comportement uniforme du processus stochastique

$$\overline{V}_{k_n}(f) = \sigma_{k_n}(f)^{-1}(V_{k_n}(f) - a_{k_n}(f)) = \sum_{j=1}^{k_n} \frac{f(j)j^{-1}}{\sigma_{k_n}(f)} =: \sum_{j=1}^{k_n} Z_{j,n}(f),$$

indexé par $f \in \mathcal{F}_{h,0}$. Toutes les hypothèses du Théorème 2.11.1 de van der Vaart et Wellner (1996) sont prises en compte dans nos enoncés, sauf celle ci,

$$(L4) \qquad \mathbb{E}^*\left(\|Z_{j,n}\|^2_{\mathcal{F}_h}\, \mathbb{1}_{\left\{\|Z_{j,n}\|_{\mathcal{F}_h} > \eta\right\}}\right) \to 0, \quad \text{quand } n \to \infty, \text{ pour tout } \eta > 0 \tag{3.21}$$

et la convergence de la fonction de covariance. Mais $\|Z_{j,n}\|$ est mesurable et nous avons par l'hypothèse (**KU1**) que

$$\mathbb{E}\,\|Z_{j,n}\|_{\mathcal{F}_1} = \max\left\{\frac{f(j)}{j}, j \geq 1\right\} \Big/ \sigma_{k_n}(f) \to 0,$$

et par l'hypothèse (**KU2a**) que

$$\mathbb{E}\,\|Z_{j,n}\|_{\mathcal{F}_2} \leq B_n \to 0,$$

ce qui montre l'hypothèse (**L4**). En ce qui concerne les fonctions de covariance nous avons

$$\Gamma_n(f_1, f_2) = \sum_{j=1}^{k_n} \frac{f_1(j)}{j\sigma_{k_n}(f_1)} \times \frac{f_2(j)}{j\sigma_{k_n}(f_1)},$$

et nous remarquons par l'inegalité de Cauchy-Schwartz qu'elles sont bornées par 1. Pour $h = 1$, nous avons

$$\Gamma_n(f_1, f_2) = \frac{1}{\sqrt{A(2, f_1)A(2, f_2)}} \sum_{j=1}^{\infty} f_1(j)f_2(j)j^{-2} \leq 1,$$

tandis que pour $h = 2$, la condition (**KU2a**) assure le résultat attendu. Nous concluons que $\{T_n^*(f), f \in \mathcal{F}_{h,0}\}$ converge vaguement dans $\ell^\infty(\mathcal{F}_{h,0})$ pour chaque $h = 1, 2$. Maintenant par le Théorème 3.3.1 et le corollaire 3.2.2, nous savons que la limite obtenue est soit $\mathcal{L}$ définie par (3.12) soit un processus Gaussien centré $\mathbb{G}$ de fonction de covariance $\Gamma(\cdot)$.

Démonstration du Théorème 3.3.2: Il est obtenu en appliquant le Théorème 2.11.9 de van der Vaart et Wellner (1996), nous allons donc vérifier les hypothèses (**L1**), (**L2**), (**L3**) et que les $Z_{j,n}$ ont des moments finis. Mais (**L2**) et la dernière hypothèse sont vérifiées. Ainsi, (**L1**) et (**L3**) verifiées dans l'Annexe C et D, assurent les résultats de ce théorème.

Démonstration du Corollaire 3.4.1: Nous appliquons ici le Théorème 3.3.2. Nous avons $\mathcal{F}_{0,1} = \{g(j) = j^{-(b-\boldsymbol{\tau})}, a \leq \boldsymbol{\tau} \leq b\}$. Pour $f(j) = j^{\boldsymbol{\tau}}$, posons $g_f(j) = j^{-(b-\boldsymbol{\tau})}$, et $f(j)/j = g_f(j) \times j^{-(1-b)}$. Dans ce cas

$$V_{k_n}^{**}(f) = (V_{k_n}(f) - a_{k_n}(f)) = \sum_{j=1}^{k_n} g_f(j)(E_j - 1)j^{-(1-b)} =: \sum_{j=1}^{k_n} Z_{j,n}(f).$$

Nous avons également

$$\begin{aligned}
\rho_n^2(f_1,f_2) &= \mathbb{E}\sum_{j=1}^{k_n}(Z_{j,n}(f_1)-Z_{j,n}(f_2))^2 \\
&= \sum_{j=1}^{k_n}(f_1(j)-f_2(j))^2 j^{-2} \\
&\leq \rho^2(f_1,f_2) = \sum_{j=1}^{\infty}(f_1(j)-f_2(j))^2 j^{-2} \\
&= \sum_{j=1}^{\infty}(g_{f_1}(j)-g_{f_2}(j))^2 j^{-2(1-b)}.
\end{aligned}$$

Nous rappelons que $\rho^2(f_1,f_2)$ n'est rien d'autre que $\|g_{f_1}-g_{f_2}\|^2_{L_2(\mathcal{F}_{0,1},\mathbb{Q})}$ pour une mesure de probabilité sur $\mathbb{N}$,

$$\mathbb{Q} = A(2,b)^{-1}\sum_{j=1}^{\infty} j^{-2(1-b)}\delta_j$$

avec

$$A(2,b) = \sum_{j=1}^{\infty} j^{-2(1-b)} < \infty.$$

Pour ces fonctions monotones $g_f : \mathbb{N} \mapsto [0,1]$, nous avons en vertu du Théorème 2.7.5 de van der Vaart et Wellner (1996) que pour un certain $K > 0$, et pour tout $\varepsilon > 0$,

$$N_{[\,]}(\varepsilon, \mathcal{F}_{0,1}, L_2(\mathbb{Q})) \leq \exp(K\varepsilon^{-1}).$$

Ceci signifie que $(\mathcal{F}_0, \rho)$ est totalement borné et (**L1**) est reduit à

$$\sup_{\rho(f_1,f_2)\leq\delta_n} \rho_n(f_1,f_2) \leq \delta_n \to 0,$$

ce qui est trivial. Dans le même esprit

$$A(2,b,\omega) = \sum_{j=1}^{\infty} j^{-2(1-b)}(E_j-1)^2(\omega)$$

est fini presque sûrement et

$$\mathbb{Q}_0(\omega) = A(2,b,\omega)^{-1}\sum_{j=1}^{\infty} j^{-2(1-b)}(E_j-1)^2(\omega)\delta_j,$$

est une mesure de probabilité pour tout ω. Nous avons

$$0 \leq d_n^2(f_1,f_2) \to d^2(f_1,f_2) = \|g_{f_1}-g_{f_2}\|^2_{L_2(\mathcal{F}_{0,1},\mathbb{Q}_0)}.$$

Cette convergence est continue puisque $\boldsymbol{\tau}$ appartient à un ensemble compact $[a, b]$. Ainsi, de manière uniforme pour $\boldsymbol{\tau} \in [a, b]$, et pour de grandes valeurs de n, on obtient

$$0 \leq d_n^2(f_1, f_2) \geq 0.25\ d^2(f_1, f_2).$$

Nous pouvons utiliser les mêmes résultats de van der Vaart et Wellner (1996), et nous obtenons pour $K > 0$ et pour tout $\varepsilon > 0$,

$$\log N_{[\ ]}(\varepsilon, \mathcal{F}_0, d_n) \leq \log N_{[\ ]}(\varepsilon, \mathcal{F}_0, d/2) \leq K(\varepsilon/2)^{-1}.$$

Cela assure l'hypothèse (**L2**). La fonction de covariance, pour $f_1(j) = j^{\boldsymbol{\tau}_1}$ et $f_2(j) = j^{\boldsymbol{\tau}_2}$, est donnée par

$$\overline{\Gamma}(f_1, f_2) = \lim_{n\to\infty} \sum_{j=1}^{k_n} f_1(j) f_2(j) j^{-2} = \sum_{j=1}^{\infty} j^{-(2-\boldsymbol{\tau}_1-\boldsymbol{\tau}_2)} \leq \sum_{j=1}^{\infty} j^{-2(1-b)} = A(2, b).$$

Quant à l'hypthèse de mesurabilité, il est facile de voir que le suprémum suivant

$$\sup_{(f_1,f_2)\in\mathcal{F}_h,\ \rho(f_1,f_2)\leq\delta_n} \sum_{j=1}^{k_n} e_j \left| \frac{f_1(j)}{j\sigma_{k_n}(f_1)} - \frac{f_1(j)}{j\sigma_{k_n}(f_1)} \right|^p E_j^p \tag{3.22}$$

est atteint pour les valeurs rationnelles de $\boldsymbol{\tau}$ dans $[a, b]$, et est alors mesurable. Ce qui achéve la preuve. ■

Démonstration du Corollaire 3.4.2: Voir Annexe C. ■

CHAPITRE 4

Famille d'estimateurs semi-paramétriques du paramètre du second ordre

Résumé

Dans ce chapitre, nous proposons une méthode d'estimation du paramètre du second ordre noté ρ pour des distributions à queues lourdes. Cette méthode regroupe un grand nombre d'estimateurs existants dans la littérature. Nous montrons qu'il est possible d'établir la normalité asymptotique de ces estimateurs de façon simple et unifiée. Nous tirons également parti de notre formalisme pour proposer de nouveaux estimateurs de ρ asymptotiquement Gaussiens.

4.1 Introduction

Nous avons vu dans le Chapitre 1 que la théorie des valeurs extrêmes établit le comportement asymptotique des plus grandes observations dans un échantillon donné. Elle fournit des méthodes pour étendre la fonction de répartition empirique au-delà des données observées. Il est ainsi possible d'estimer les quantités relatives à la queue d'une distribution à savoir les probabilités faibles de dépassements ou des quantiles extrêmes. On pourra se référer au Chapitre 1 pour plus de détails sur la théorie des valeurs extrêmes ou également à Embrechts *et al.* (1997) et de Haan et Ferreira (2006) . Plus précisément, considérons $X_1, ..., X_n$, une suite de variables aléatoires, indépendantes et identiquement distribuées de fonction de répartition F. La théorie des valeurs extrêmes

établit que la loi asymptotique du maximum $X_{n,n} = \max(X_1, ..., X_n)$ normalisé est donné par

$$G_\gamma(x) = \exp\left(-(1+\gamma x)_+\right)^{-1/\gamma}$$

où $y_+ = \max(y, 0)$. Le paramètre inconnu $\gamma \in \mathbb{R}$ est appelé indice des valeurs extrêmes ou indice de queue et a été developpé dans les chapitres précédents. Notons que γ est le paramètre d'importance en théorie des valeurs extrêmes, il contrôle le comportement de la queue au premier ordre. Plus γ est grand, plus la queue est lourde. De nombreux estimateurs de ce paramètre ont été donnés dans les chapitres précédents notamment dans le cas particulier où la loi étudiée appartient au domaine d'attraction de Fréchet (cas $\gamma > 0$, et qui nous intéresse dans la suite de ce chapitre).

Nous avons déjà mentionné qu'une caratérisation simple des fonctions de répartition F appartenant au domaine de Fréchet est donnée par la fonction quantile de queue $\mathbb{U}(x) := F^{\leftarrow}(1 - 1/x)$ telle que,

$$U(x) = x^\gamma \ell(x), \tag{4.1}$$

où ℓ est une fonction à variation lente à l'infini c'est-à-dire pour tout $\lambda > 0$,

$$\lim_{x\to\infty} \frac{\ell(\lambda x)}{\ell(x)} = 1. \tag{4.2}$$

La distribution F est dite à queue lourde (Heavy-tailed) si (4.1) est vérifiée (voir aussi le Chapitre 1 pour plus de détails), et l'indice des valeurs extrêmes γ contrôle la lourdeur de la queue de cette distribution. Ainsi, l'estimation de γ est un thème central dans l'analyse de ces distributions. Plusieurs estimateurs ont ainsi été proposés dans la littérature statistique et leurs distributions asymptotiques ont été établies à partir d'une condition du second ordre donnée par (1.35) et qui est équivalente à la condition ci-dessous:

Il existe une fonction $A(x) \to 0$ de signe constant à l'infini et un paramètre du second ordre $\rho \leq 0$ tels que, pour tout $\lambda > 0$,

$$\boxed{\lim_{x\to\infty} \frac{1}{A(x)} \log\left(\frac{\ell(\lambda x)}{\ell(x)}\right) = K_\rho(\lambda) := \int_1^\lambda u^{\rho-1} du.} \tag{4.3}$$

Remarquons que la condition (4.3) implique que $|A|$ est une fonction à variations régulières à l'infini d'indice ρ, voir par exemple Geluck et de Haan (1987). Par conséquent, plus le paramètre du second ordre ρ est petit, plus la vitesse de convergence dans (4.3) est grande. La connaissance de ρ est alors d'un grand intérêt dans la pratique. Par exemple, le paramètre du second ordre est de première importance dans le choix adaptatif du

nombre optimal de statistiques d'ordre supérieur utilisé lors de l'estimation de l'indice des valeurs extrêmes (voir par exemple Hall et Welsh (1985), Cheng et Peng (2001), Gomes et Oliveira (2001), Beirlant *et al.* (2002), etc.

L'estimateur de ρ est aussi utilisé pour la réduction du biais des estimateurs de l'indice des valeurs extrêmes (voir par exemple , Gomes et Martins (2002), Caeiro *et al.* (2005), Gomes et al. (2007)), ou bien pour les coefficient de queue de Weibull, Diebolt *et al.* (2008a), Diebolt *et al.* (2008b) même si la réduction du biais est obtenue avec le choix canonique $\rho = -1$, comme cela est suggéré dans Feuerverger et Hall (1999) et Gomes et Martins (2004).
Pour les raisons mentionnées ci-dessus, l'estimation du paramètre du second ordre ρ a reçu beaucoup d'attention dans la théorie des valeurs extrêmes, voir par exemple Hall et Welsh (1985), Gomes *et al.* (2002), Fraga Alves *et al.* (2003a), Fraga Alves et al. (2003b), Ciuperca et Mercadier (2010), Goegebeur *et al.* (2010), Cai *et al.* (2011) et Worms et Worms (2011).

Nous proposons dans ce chapitre, une famille d'estimateurs semi-paramétriques de ρ permettant de regrouper les travaux précédents au sein d'un formalisme commun.
Soit $\mathbb{I} = {}^t(1, \ldots, 1) \in \mathbb{R}^d$, $d \geq 3$. L'outil principal de notre approche consiste à considérer la variable aléatoire $T_n = T_n(X_1, \ldots, X_n) \in \mathbb{R}^d$ et les trois hypothèses ci-dessous.
Il existe deux variables aléatoires ω_n, χ_n, deux fonctions $f : \mathbb{R}^- \to \mathbb{R}^d$ et $\psi : \mathbb{R}^d \to \mathbb{R}$ telles que:

(T1) $\omega_n^{-1}(T_n - \chi_n \mathbb{I}) \xrightarrow{\mathbb{P}} f(\rho)$.

(Ψ1) $\psi(x + \lambda \mathbb{I}) = \psi(x)$ pour tout $x \in \mathbb{R}^d$ et $\lambda \in \mathbb{R}$,

(Ψ2) $\psi(\lambda x) = \psi(x)$ pour tout $\lambda \in \mathbb{R} \setminus \{0\}$.

Notons que **(T1)** impose que T_n correctement normalisé, converge en probabilité vers une fonction de ρ, tandis que **(Ψ1)** et **(Ψ2)** signifient que ψ est à la fois invariante par translation et par homothétie. Suivant ces trois hypothèses, nous obtenons

$$\boxed{\psi(\omega_n^{-1}(T_n - \chi_n \mathbb{I}) = \psi(T_n) \xrightarrow{\mathbb{P}} \psi(f(\rho)),}$$

sous une hypothèse de continuité sur ψ. Posons $Z_n := \psi(T_n)$ et $\varphi := \psi \circ f : \mathbb{R}^- \to \mathbb{R}$, nous obtenons $Z_n \xrightarrow{\mathbb{P}} \varphi(\rho)$. Il est clair que, sous une hypothèse additionnelle de regularité et en supposant que Z_n et φ sont connues, ρ peut être estimé par $\varphi^{-1}(Z_n)$. Ce principe

d'estimation est décrit plus précisément dans la section 4.2. La consistance et la normalité asymptotique de l'estimateur proposé y sont établies. Des exemples de statistiques T_n sont présentés dans la section 4.3. Quelques exemples de fonctions ψ sont également proposés dans la section 4.4 et il est montré que les estimateurs proposés dans Gomes *et al.* (2002), Alves *et al.* (2003a), Alves et al. (2003b), Ciuperca et Mercadier (2010), Goegebeur *et al.* (2010) et peuvent être vus comme des cas particuliers de notre approche. En suite, cette remarque permet d'établir leurs propriétés asymptotiques d'une manière simple et unifiée. Nous tirons également parti de notre formalisme pour proposer un grand nombre d'estimateurs nouveaux de ρ asymptotiquement Gaussiens.

4.2 Propriétés asymptotiques

Rappelons que T_n est un vecteur aléatoire dans $\mathbb{R}^d$ vérifiant **(T1)** et ψ, une fonction de $\mathbb{R}^d \to \mathbb{R}$ vérifiant **(Ψ1)** et **(Ψ2)**. De plus nous supposons que:

(Ψ3) Il existe $J_0 \subseteq \mathbb{R}^-$ et un intervalle ouvert $J \subset \mathbb{R}$ tels que $\varphi = \psi \circ f$ est une bijection de $J_0 \to J$.

Sous cette hypothèse, on peut considérer l'estimateur de ρ suivant:

$$\hat{\boldsymbol{\rho}}_n = \begin{cases} \varphi^{-1}(Z_n) & \text{si } Z_n \in J \\ 0 & \text{sinon.} \end{cases} \tag{4.4}$$

Pour établir la consistance de $\hat{\boldsymbol{\rho}}_n$, nous introduisons une hypothèse additionnelle de régularité:

(Ψ4) ψ est continue au voisinage de $f(\rho)$ et f continue au voisinage de ρ.

La preuve du résultat ci dessous est basée sur les considérations heuristiques de la section 4.1 et est detaillée dans la section 4.6.

Théorème 4.2.1 *Si les hypothèses* **(T1)** *et* **(Ψ1)**–**(Ψ4)** *sont satisfaites alors*

$$\hat{\boldsymbol{\rho}}_n \xrightarrow{\mathbb{P}} \rho \quad \textit{quand} \quad n \to \infty.$$

La normalité asymptotique de $\hat{\boldsymbol{\rho}}_n$ peut être établie sous une version plus forte de **(Ψ4)**:

(Ψ5) ψ est continûment différentiable au voisinage de $f(\rho)$ et f est continûment différentiable au voisinage de ρ,

et l'hypothèse qu'une version normalisée de T_n est elle même asymptotiquement Gaussienne:

(T2) Il existe deux variables aléatoires ω_n, χ_n, une suite $v_n \to \infty$, deux fonctions f, $m : \mathbb{R}^- \to \mathbb{R}^d$ et une matrice Σ de dimension $d \times d$ telles que

$$v_n(\omega_n^{-1}(T_n - \chi_n \mathbb{I}) - f(\rho)) \xrightarrow{\mathcal{D}} \mathcal{N}_d(m(\rho), \boldsymbol{\gamma}^2 \Sigma).$$

Théorème 4.2.2 *Supposons que les hypothèses* **(T2)**, **(Ψ1)**–**(Ψ3)** *et* **(Ψ5)** *soient vérifiées. Si $\rho \in J_0$ et $\varphi'(\rho) \neq 0$, alors*

$$v_n(\hat{\boldsymbol{\rho}}_n - \rho) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{m_\psi(\rho)}{\varphi'(\rho)}, \frac{\boldsymbol{\gamma}^2 \sigma_\psi^2(\rho)}{(\varphi'(\rho))^2}\right),$$

avec $\varphi'(\rho) = {}^t f'(\rho) \nabla\psi(f(\rho))$ et où nous définissons

$$\begin{aligned} m_\psi(\rho) &:= {}^t m(\rho)\, \nabla\psi(f(\rho)), \\ \sigma_\psi^2(\rho) &:= {}^t \nabla\psi(f(\rho))\, \Sigma\, \nabla\psi(f(\rho)). \end{aligned}$$

La preuve du Théorème 4.2.2 est detaillée dans la section 4.6.

4.3 Exemples de variables aléatoires de type T_n

Soit $X_{1,n} \leq \ldots \leq X_{n,n}$ un échantillon de statistiques d'ordre croissant et k_n une suite intermédiaire telle que $k \to \infty$ et $k_n/n \to 0$ quand $n \to \infty$. Notons que beaucoup d'estimateurs en théorie des valeurs extrêmes sont basés soit sur les logarithmes des excès au delà k_n-*ième* plus grande observation d'un échantillon (*log-excesses*): $\log X_{n-j+1,n} - \log X_{n-k_n,n}$, $j = 1, \ldots, k_n$ ou bien sur les espacements des logarithmes entre les k_n plus grandes observations (*rescaled log-spacings*: $j(\log X_{n-j+1,n} - \log X_{n-j,n})$, $j = 1, \ldots, k_n$). Dans la suite, deux exemples de statistique T_n sont présentés et sont basés sur la moyenne pondérée des *log-excesses* et des *rescaled log-spacings*.

Le premier exemple est basé sur la statistique suivante

$$\boxed{R_{k_n}(\tau) = \frac{1}{k_n} \sum_{j=1}^{k_n} H_{\boldsymbol{\tau}}\left(\frac{j}{k_n+1}\right) j(\log X_{n-j+1,n} - \log X_{n-j,n}),} \tag{4.5}$$

où $H_{\boldsymbol{\tau}} : [0,1] \to \mathbb{R}$ est une fonction de poids indexée par $\boldsymbol{\tau} \in (0,\infty)$. Nous supposons sans perte de généralité que $H_{\boldsymbol{\tau}}$ est d'intégrale égale à l'unité c'est-à-dire $\int_0^1 H_{\boldsymbol{\tau}}(u)du = 1$. Cette statistique est utilisée par exemple dans Beirlant *et al.* (1999) pour estimer l'indice des valeurs extrêmes $\boldsymbol{\gamma}$ et dans Hall et Welsh (1985), Goegebeur *et al.* (2010) pour estimer le paramètre du second ordre ρ. Elle constitue un cas particulier de l'estimateur à noyau introduit dans Csörgő *et al.* (1985), voir la Sous-section 1.5.1.1.
Notons aussi que, dans le cas$H_{\boldsymbol{\tau}}(u) = 1$ pour tout $u \in [0,1]$, $R_{k_n}(\tau)$ correspond à l'estimateur de Hill (1975) développé dans la section 1.5.1.2. Les propriétés asymptotiques de $R_{k_n}(\tau)$ exigent une certaine condition technique (donnée par **(C1)**) sur la fonction de poids $H_{\boldsymbol{\tau}}$. Cette condition a été d'abord introduite dans Beirlant *et al.* (1999) et est rappelée ci-dessous. Introduisons l' opérateur

$$\mu : h \in L_2([0,1]) \longrightarrow \mu(h) = \int_0^1 h(u)du \in \mathbb{R}$$

et $I_t(u) = u^{-t}$ pour $t \leq 0$ and $u \in (0,1]$, la condition peut être écrite comme suit

(C1) $H_{\boldsymbol{\tau}} \in L_2([0,1])$, $\mu(|H_{\boldsymbol{\tau}}|I_{\rho+1+\varepsilon}) < \infty$ et $H_{\boldsymbol{\tau}}(t) = \frac{1}{t}\int_0^t u(\nu)d\nu$, pour un certain $\varepsilon > 0$ et pour une certaine fonction u vérifiant pour tout $j = 1,\ldots,k$

$$\left|(k_n+1)\int_{(j-1)/(k_n+1)}^{j/(k_n+1)} u(t)dt\right| \leq g\left(\frac{j}{k_n+1}\right),$$

où $g(\cdot)$ est une fonction positive continue et intégrable définie sur $(0,1)$. De plus, pour $\eta \in \{0,1\}$, et $k_n \to \infty$:

$$\begin{aligned}
\frac{1}{k_n}\sum_{j=1}^{k_n} H_{\boldsymbol{\tau}}\left(\frac{j}{k_n+1}\right)\left(\frac{j}{k_n+1}\right)^{-\eta\rho} &= \mu(H_{\boldsymbol{\tau}} I_{\eta\rho}) + o(k_n^{-1/2}),\\
\max_{j\in\{1,\ldots,k_n\}}\left|H_{\boldsymbol{\tau}}\left(\frac{j}{k_n+1}\right)\right| &= o(k_n^{1/2}).
\end{aligned}$$

Il est alors possible de définir la statistique $T_n^{(R)}$ à partir de $R_{k_n}(\tau)$

$$T_n^{(R)} = \left(T_{n,i}^{(R)} = (R_{k_n}(\tau_i)/\boldsymbol{\gamma})^{\boldsymbol{\theta}_i},\ i = 1,\ldots,d\right),$$

où $\boldsymbol{\theta}_i$, $i = 1,\ldots,d$ sont des paramètres arbitraires réels. Dans le lemme suivant, il est montré que $T_n^{(R)}$ satisfait l'hypothèse **(T2)** sous la condition du troisième ordre

ci-dessous, qui est un raffinement de (4.3):

(C2) Il existe deux fonctions $A(x) \to 0$ et $B(x) \to 0$ toutes les deux de signe constant à l'infini, il existe également un paramètre du second ordre $\rho \leq 0$ et un paramètre du troisième ordre $\beta \leq 0$ tels que, pour tout $\lambda > 0$,

$$\boxed{\lim_{x\to\infty} \frac{(\log \ell(\lambda x) - \log \ell(x))/A(x) - K_\rho(\lambda)}{B(x)} = L_{(\rho,\beta)}(\lambda) := \int_1^\lambda s^{\rho-1} \int_1^s u^{\beta-1} du ds,}$$

et les fonctions $|A|$ et $|B|$ sont à variations régulières à l'infini d'indices respectifs ρ et β.

Cette condition permet d'établir la normalité asymptotique des estimateurs de ρ. Soit $Y_{n-k_n,n}$ la $(n-k_n)$*ième* plus grande statistique d'ordre dans un échantillon de taille n d'une loi de Pareto standard avec une fonction de répartition $G(y) = 1 - 1/y$, $y > 1$.

Lemme 4.3.1 *Supposons que* **(C1)**, **(C2)** *soient vérifées et soit k_n une suite d'entiers tels que pour $\lambda_A \in \mathbb{R}$ et $\lambda_B \in \mathbb{R}$*

$$\begin{gathered} k_n \to \infty,\ n/k_n \to \infty,\ k_n^{1/2} A(n/k_n) \to \infty, \\ k_n^{1/2} A^2(n/k_n) \to \lambda_A \text{ et } k_n^{1/2} A(n/k_n) B(n/k_n) \to \lambda_B. \end{gathered} \tag{4.6}$$

Alors, le vecteur aléatoire $T_n^{(R)}$ satisfait l'hypothèse **(T2)** *avec*

$$\begin{aligned} \chi_n^{(R)} &= 1, \\ \omega_n^{(R)} &= A(Y_{n-k_n,n})/\boldsymbol{\gamma}, \\ v_n &= k_n^{1/2} A(n/k_n), \\ f^{(R)}(\rho) &= (\boldsymbol{\theta}_i \mu(H_{\boldsymbol{\tau}_i} I_\rho),\ i = 1, \ldots, d), \\ m^{(R)}(\rho) &= \left(\lambda_A \frac{\boldsymbol{\theta}_i(\boldsymbol{\theta}_i - 1)}{2\boldsymbol{\gamma}} \mu^2(H_{\boldsymbol{\tau}_i} I_\rho) - \lambda_B \boldsymbol{\theta}_i \mu(H_{\boldsymbol{\tau}_i} I_\rho K_{-\beta});\ i = 1, \ldots, d\right), \\ \Sigma_{i,j}^{(R)} &= \boldsymbol{\theta}_i \boldsymbol{\theta}_j \mu(H_{\boldsymbol{\tau}_i} H_{\boldsymbol{\tau}_j}),\ \textit{pour tout } (i,j) \in \{1, ..., d\}^2. \end{aligned}$$

Preuve: La preuve du Lemme 4.3.1 est une suite directe du Théorème 4.2.2 et de l'Appendice A.5 dans Goegebeur *et al.* (2010). ∎

Le second exemple nécessite quelques notations supplémentaires. Considérons l'opérateur

$$\vartheta : (h_1, h_2) \in L_2([0,1]) \times L_2([0,1]) \mapsto \vartheta(h_1, h_2) = \int_0^1 \int_0^1 \frac{h_1(u) h_2(v)}{u \wedge v} du dv - \mu(h_1)\mu(h_2) \in \mathbb{R}$$

et deux fonctions $\bar{I}_t(u) = (1-u)^{-t}$ et $J_t(u) = (-\log u)^{-t}$ définies pour tout $t \leq 0$ et $u \in [0,1)$. On s'intéresse maintenant à la statistique suivante

$$\boxed{S_{k_n}(\tau,\alpha) = \frac{1}{k_n}\sum_{j=1}^{k} G_{\boldsymbol{\tau},\boldsymbol{\alpha}}\left(\frac{j}{k_n+1}\right)(\log X_{n-j+1,n} - \log X_{n-k_n,n})^{\boldsymbol{\alpha}},} \tag{4.7}$$

où $G_{\boldsymbol{\tau},\boldsymbol{\alpha}}$ est une fonction positive indexée par deux paramètres positifs $\boldsymbol{\alpha}$ et $\boldsymbol{\tau}$. Sans perte de généralité nous allons supposer que $\mu(G_{\boldsymbol{\tau},\boldsymbol{\alpha}}J_{-\boldsymbol{\alpha}}) = 1$. Dans Dekkers *et al.* (1989) un estimateur de l'indice des valeurs extrêmes $\boldsymbol{\gamma}$ basé sur cette statistique a été introduit dans le cas particulier où $G_{\boldsymbol{\tau},\boldsymbol{\alpha}}$ est une constante. La statistique $S_{k_n}(\tau,\alpha)$ a été aussi utilisée dans Hall et Welsh (1985), Peng (1998), Fraga Alves *et al.* (2003b) et Ciuperca et Mercadier (2010) pour estimer les paramètres $\boldsymbol{\gamma}$ et ρ. Les distributions asymptotiques de ces estimateurs sont obtenues sous la condition ci-dessous sur la fonction $G_{\boldsymbol{\tau},\boldsymbol{\alpha}}$.

(C3) La fonction $G_{\boldsymbol{\tau},\boldsymbol{\alpha}}$ est positive, décroissante et intégrable sur $(0,1)$. De plus, il existe $\boldsymbol{\delta} > 1/2$ tel que $\mu(G_{\boldsymbol{\tau},\boldsymbol{\alpha}}I_{\boldsymbol{\delta}}) < \infty$ et $\mu(G_{\boldsymbol{\tau},\boldsymbol{\alpha}}\bar{I}_{\boldsymbol{\delta}}) < \infty$.

Il est alors possible de définir la variable aléatoire $T_n^{(S)}$ à partir de la statistique $S_{k_n}(\tau,\alpha)$ donnée dans (4.7)

$$T_n^{(S)} = \left(T_{n,i}^{(S)} = (S_{k_n}(\tau_i,\alpha_i)/\boldsymbol{\gamma}^{\boldsymbol{\alpha}_i})^{\boldsymbol{\theta}_i},\ i = 1,\ldots,d\right).$$

Le résultat suivant est analogue au Lemme 4.3.1 pour la statistique $S_{k_n}(\tau,\alpha)$.

Lemme 4.3.2 *Supposons que les conditions* **(C2)** *et* **(C3)** *soient verifiées. Si la suite d'entiers $k_n = k(n)$ satisfait (4.6) alors le vecteur aléatoire $T_n^{(S)}$ vérifie l'hypothèse* **(T2)** *avec*

$$\begin{aligned}
\chi_n^{(S)} &= 1,\\
\omega_n^{(S)} &= A(n/k_n)/\boldsymbol{\gamma},\\
v_n &= k_n^{1/2}A(n/k_n),\\
f^{(S)}(\rho) &= (-\boldsymbol{\theta}_i\boldsymbol{\alpha}_i\mu(G_{\boldsymbol{\tau}_i,\boldsymbol{\alpha}_i}J_{1-\boldsymbol{\alpha}_i}K_{-\rho});\ i = 1,\ldots,d),\\
m^{(S)}(\rho) &= \left(m_i^{(S)}(\rho);\ i = 1,\ldots,d\right),\\
\Sigma_{i,j}^{(S)} &= \boldsymbol{\theta}_i\boldsymbol{\theta}_j\boldsymbol{\alpha}_i\boldsymbol{\alpha}_j\vartheta(G_{\boldsymbol{\tau}_i,\boldsymbol{\alpha}_i}J_{1-\boldsymbol{\alpha}_i}, G_{\boldsymbol{\tau}_j,\boldsymbol{\alpha}_j}J_{1-\boldsymbol{\alpha}_j})\ \textit{pour tout}\ (i,j) \in \{1,...,d\}^2,
\end{aligned}$$

où

$$m_i^{(S)}(\rho) = \lambda_A\frac{\boldsymbol{\theta}_i\boldsymbol{\alpha}_i(\boldsymbol{\alpha}_i-1)}{2\boldsymbol{\gamma}}\mu(G_{\boldsymbol{\tau}_i\boldsymbol{\alpha}_i}J_{2-\boldsymbol{\alpha}_i}K_{-\rho}^2) + \lambda_B\boldsymbol{\alpha}_i\boldsymbol{\theta}_i\mu(G_{\boldsymbol{\tau}_i,\boldsymbol{\alpha}_i}J_{1-\boldsymbol{\alpha}_i}L_{(-\rho,-\beta)}).$$

Preuve: La preuve du Lemme 4.3.2 est une suite directe de la Proposition 2 et du Lemme 1 dans Ciuperca et Mercadier (2010). ■

Dans la section suivante, nous allons illustrer comment la combinaison des variables aléatoires $T_n^{(R)}$ et $T_n^{(S)}$ avec une certaine fonction ψ suivant (5.2.4.2) peut conduire à des estimateurs existants ou à de nouveaux estimateurs de ρ.

4.4 Applications

Dans cette section, nous allons proposer des estimateurs de ρ basés sur les variables aléatoires $T_n^{(R)}$ (Sous section 4.4.1) et $T_n^{(S)}$ (Sous section 4.4.2). Dans les deux cas, la dimension d est égale à 8 ($d = 8$) [1] et la fonction suivante $\psi_{\boldsymbol{\delta}} : \mathcal{D} \mapsto \mathbb{R} \setminus \{0\}$ est donnée par

$$\psi_{\boldsymbol{\delta}}(x_1, \ldots, x_8) = \widetilde{\psi}_{\boldsymbol{\delta}}(x_1 - x_2, x_3 - x_4, x_5 - x_6, x_7 - x_8),$$

où $\boldsymbol{\delta} \geq 0$,

$$\mathcal{D} = \{(x_1, \ldots, x_8) \in \mathbb{R}^8;\ x_1 \neq x_2,\ x_3 \neq x_4,\ \text{et } (x_5 - x_6)(x_7 - x_8) > 0\},$$

et $\widetilde{\psi}_{\boldsymbol{\delta}} : \mathbb{R}^4 \mapsto \mathbb{R}$ est donnée par:

$$\widetilde{\psi}_{\boldsymbol{\delta}}(y_1, \ldots, y_4) = \frac{y_1}{y_2}\left(\frac{y_4}{y_3}\right)^{\boldsymbol{\delta}}.$$

Soulignons que $\psi_{\boldsymbol{\delta}}$ vérifie les propriétés d'invariance **(Ψ1)** et **(Ψ2)**.

4.4.1 Estimateurs basés sur la statistique $R_{k,n}(\boldsymbol{\tau})$

Puisque $d = 8$, la statistique $T_n^{(R)}$ dépend de 16 paramètres: $\{(\boldsymbol{\theta}_i, \boldsymbol{\tau}_i) \in (0, \infty)^2,\ i = 1, \ldots, 8\}$. La condition ci-dessous sur ces paramètres est introduite. Soit $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_1, \ldots, \tilde{\boldsymbol{\theta}}_4) \in (0, \infty)^4$ avec $\tilde{\boldsymbol{\theta}}_3 \neq \tilde{\boldsymbol{\theta}}_4$.

(C4) $\{\boldsymbol{\theta}_i = \tilde{\boldsymbol{\theta}}_{\lceil i/2 \rceil},\ i = 1, \ldots, 8\}$ avec $\boldsymbol{\delta} = (\tilde{\boldsymbol{\theta}}_1 - \tilde{\boldsymbol{\theta}}_2)/(\tilde{\boldsymbol{\theta}}_3 - \tilde{\boldsymbol{\theta}}_4)$ où $\lceil x \rceil = \inf\{n \in \mathbb{N} | x \leq n\}$. De plus, $\boldsymbol{\tau}_1 < \boldsymbol{\tau}_2 \leq \boldsymbol{\tau}_3 < \boldsymbol{\tau}_4$, $\boldsymbol{\tau}_5 < \boldsymbol{\tau}_6 \leq \boldsymbol{\tau}_7 < \boldsymbol{\tau}_8$.

[1] Ce choix vient du fait que dans les estimations du paramètre du second ordre proposées dans Hall et Welsh (1985), Gomes *et al.* (2002), Fraga Alves *et al.* (2003a), Fraga Alves et al. (2003b), Ciuperca et Mercadier (2010), Goegebeur *et al.* (2010), Cai *et al.* (2011) et Worms et Worms (2011), les dimensions utilisées sont inférieures ou égales à 8.

Sous ces conditions, $T_n^{(R)}$ se retrouve avec 12 paramètres libres. Nous introduisons également les notations suivantes: $Z_n^{(R)} = \psi_{\boldsymbol{\delta}}(T_n^{(R)})$ et $\varphi_{\boldsymbol{\delta}}^{(R)} = \psi_{\boldsymbol{\delta}} \circ f^{(R)}$. D'après le choix de $\boldsymbol{\delta} = (\tilde{\boldsymbol{\theta}}_1 - \tilde{\boldsymbol{\theta}}_2)/(\tilde{\boldsymbol{\theta}}_3 - \tilde{\boldsymbol{\theta}}_4)$, il est facile de voir que $Z_n^{(R)}$ ne dépend pas du paramètre inconnu $\boldsymbol{\gamma}$. Nous allons maintenant établir la normalité asymptotique de l'estimateur $\hat{\boldsymbol{\rho}}_n^{(R)}$ défini par (5.2.4.2) lorsque la statistique $T_n^{(R)}$ et la fonction $\psi_{\boldsymbol{\delta}}$ sont utilisées:

$$\hat{\boldsymbol{\rho}}_n^{(R)} = \begin{cases} \left(\varphi_{\boldsymbol{\delta}}^{(R)}\right)^{-1}(Z_n^{(R)}) & \text{si } Z_n^{(R)} \in J \\ 0 & \text{sinon.} \end{cases} \tag{4.8}$$

La condition supplémentaire suivante est requise:

(C5) La fonction $\nu_\rho(\boldsymbol{\tau}) = \mu(H_{\boldsymbol{\tau}} I_\rho)$ est différentiable avec, pour tout $\rho < 0$ et pour tout $\boldsymbol{\tau} \in \mathbb{R}$, $\nu'_\rho(\boldsymbol{\tau}) > 0$.

Notons pour $i \in \{1, \ldots, 4\}$,

$$\begin{aligned} m_A^{(R,i)} &= \exp\left\{(\tilde{\boldsymbol{\theta}}_i - 1)\mu\left((H_{\boldsymbol{\tau}_{2i-1}} + H_{\boldsymbol{\tau}_{2i}})I_\rho\right)\right\}, \\ m_B^{(R,i)} &= \exp\left\{\frac{\mu\left((H_{\boldsymbol{\tau}_{2i-1}} - H_{\boldsymbol{\tau}_{2i}})I_\rho K_{-\beta}\right)}{\mu\left((H_{\boldsymbol{\tau}_{2i-1}} - H_{\boldsymbol{\tau}_{2i}})I_\rho\right)}\right\}, \\ v^{(R,i)}(u) &= \exp\left\{\frac{H_{\boldsymbol{\tau}_{2i-1}}(u) - H_{\boldsymbol{\tau}_{2i}}(u)}{\mu\left((H_{\boldsymbol{\tau}_{2i-1}} - H_{\boldsymbol{\tau}_{2i}})I_\rho\right)}\right\}, \quad \text{pour tout} \quad u \in [0,1]. \end{aligned}$$

Pour simplifier, nous introduisons aussi $m_A^{(R)} = (m_A^{(R,i)},\ i = 1, \ldots, 4)$, $m_B^{(R)} = (m_B^{(R,i)},\ i = 1, \ldots, 4)$ et $v^{(R)} = (v^{(R,i)},\ i = 1, \ldots, 4)$.

Corollaire 4.4.1 *Supposons que les conditions* **(C1)**, **(C2)**, **(C4)** *et* **(C5)** *soient vérifiées. S'il existe deux intervalles J et J_0 tels que pour tout $\rho \in J_0$ et une suite intermédiaire k_n vérifiant (4.6), alors*

$$k_n{}^{1/2} A(n/k_n)(\hat{\boldsymbol{\rho}}_n^{(R)} - \rho) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda_A}{2\boldsymbol{\gamma}}\mathcal{AB}_1^{(R)}(\boldsymbol{\delta}, \rho) - \lambda_B \mathcal{AB}_2^{(R)}(\boldsymbol{\delta}, \rho, \beta),\ \boldsymbol{\gamma}^2 \mathcal{AV}^{(R)}(\boldsymbol{\delta}, \rho)\right)$$

où

$$\begin{aligned} \mathcal{AB}_1^{(R)}(\boldsymbol{\delta}, \rho) &= \varphi_{\boldsymbol{\delta}}^{(R)}(\rho) \log \tilde{\psi}_{\boldsymbol{\delta}}(m_A^{(R)}) / [\varphi_{\boldsymbol{\delta}}^{(R)}]'(\rho), \\ \mathcal{AB}_2^{(R)}(\boldsymbol{\delta}, \rho, \beta) &= \varphi_{\boldsymbol{\delta}}^{(R)}(\rho) \log \tilde{\psi}_{\boldsymbol{\delta}}(m_B^{(R)}) / [\varphi_{\boldsymbol{\delta}}^{(R)}]'(\rho), \\ \mathcal{AV}^{(R)}(\boldsymbol{\delta}, \rho) &= \left(\varphi_{\boldsymbol{\delta}}^{(R)}(\rho)\right)^2 \mu\left(\log^2 \tilde{\psi}_{\boldsymbol{\delta}}(v^{(R)})\right) / \left([\varphi_{\boldsymbol{\delta}}^{(R)}]'(\rho)\right)^2. \end{aligned}$$

Notons que ce résultat peut être vu de deux manières comme une extension de la Proposition 3 de Goegebeur *et al.* (2010). D'abord, nous ne nous limitons pas au cas $\boldsymbol{\delta} = 1$. Ensuite nous ne supposons pas que la fonction $\varphi_{\boldsymbol{\delta}}^{(R)}$ est bijective, mais cela s'avère être une suite de **(C4)**. En outre, comme illustré ci-dessous, la preuve du Corollaire 4.4.1 est très simple et est basée sur le Théorème 4.2.2 et le Lemme 4.3.1.

Considérons par exemple la fonction $H_{\boldsymbol{\tau}} : u \in [0,1] \mapsto \boldsymbol{\tau} u^{\boldsymbol{\tau}-1}$, $\boldsymbol{\tau} \geq 1$. Il est facile de voir que $H_{\boldsymbol{\tau}}$ satisfait les conditions **(C1)** et **(C4)** puisque $\nu_\rho(\boldsymbol{\tau}) = \boldsymbol{\tau}/(\boldsymbol{\tau} - \rho)$. Soient $\boldsymbol{\tau}_1 \leq \boldsymbol{\tau}_5$, $\boldsymbol{\tau}_2 = \boldsymbol{\tau}_3$, $\boldsymbol{\tau}_4 = \boldsymbol{\tau}_8$ et $\boldsymbol{\tau}_6 = \boldsymbol{\tau}_7$, ce qui conduit à une expression simple de $\varphi_{\boldsymbol{\delta}}^{(R)}$:

$$\varphi_{\boldsymbol{\delta}}^{(R)}(\rho) = \omega(\boldsymbol{\delta}, \tilde{\boldsymbol{\theta}}) \left(\frac{\boldsymbol{\tau}_4 - \rho}{\boldsymbol{\tau}_1 - \rho}\right) \left(\frac{\boldsymbol{\tau}_5 - \rho}{\boldsymbol{\tau}_4 - \rho}\right)^{\boldsymbol{\delta}} \quad \text{où} \quad \omega(\boldsymbol{\delta}, \tilde{\boldsymbol{\theta}}) = \left(\frac{\tilde{\boldsymbol{\theta}}_1(\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2)}{\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\tau}_2 - \boldsymbol{\tau}_4)}\right) \left(\frac{\tilde{\boldsymbol{\theta}}_4(\boldsymbol{\tau}_6 - \boldsymbol{\tau}_4)}{\tilde{\boldsymbol{\theta}}_3(\boldsymbol{\tau}_5 - \boldsymbol{\tau}_6)}\right)^{\boldsymbol{\delta}}. \tag{4.9}$$

Nous avons également des formes explicites pour J_0 et J dans deux situations possibles:

(i) Si $0 \leq \boldsymbol{\delta} < \boldsymbol{\delta}_0 := (\boldsymbol{\tau}_4 - \boldsymbol{\tau}_1)/(\boldsymbol{\tau}_4 - \boldsymbol{\tau}_5)$ alors $\varphi_{\boldsymbol{\delta}}^{(R)}$ est strictement croissante de $J_0 = \mathbb{R}^-$ à valeur dans $J = \omega(\boldsymbol{\delta}, \tilde{\boldsymbol{\theta}}) \bullet (1, \widetilde{\psi}_{\boldsymbol{\delta}}(\boldsymbol{\tau}_4, \boldsymbol{\tau}_1, \boldsymbol{\tau}_4, \boldsymbol{\tau}_5))$.

(ii) Si $\boldsymbol{\delta} \geq \boldsymbol{\delta}_1 := \boldsymbol{\delta}_0 \boldsymbol{\tau}_5/\boldsymbol{\tau}_1$ alors $\varphi_{\boldsymbol{\delta}}^{(R)}$ est strictement de $J_0 = \mathbb{R}^-$ à valeur dans $J = \omega(\boldsymbol{\delta}, \tilde{\boldsymbol{\theta}}) \bullet (\widetilde{\psi}_{\boldsymbol{\delta}}(\boldsymbol{\tau}_4, \boldsymbol{\tau}_1, \boldsymbol{\tau}_4, \boldsymbol{\tau}_5), 1)$.

Ici, $\bullet$ désigne l'opérateur de multiplication par scalaire. Le cas $\boldsymbol{\delta} \in [\boldsymbol{\delta}_0, \boldsymbol{\delta}_1]$ n'est pas considéré dans la suite, puisque l'on peut montrer que dans cette situation, $J_0 \subsetneq \mathbb{R}^-$ et donc la condition $\rho \in J_0$ du Corollaire 4.4.1 n'est pas nécessairement satisfaite. Voyons maintenant quelques cas particuliers où la fonction inverse de $\varphi_{\boldsymbol{\delta}}^{(R)}$ est explicite.

Exemple 4.4.2 *Soit $\boldsymbol{\delta} = 1$ c'est-à-dire $\tilde{\boldsymbol{\theta}}_1 - \tilde{\boldsymbol{\theta}}_2 = \tilde{\boldsymbol{\theta}}_3 - \tilde{\boldsymbol{\theta}}_4$. Notons $Z_{n,1}^{(R)}$ la variable aléatoire $Z_n^{(R)}$ correspondant à ce choix. Puisque dans le cas $\boldsymbol{\delta}_0 > 1$, on se place dans la situation (i) et*

$$\hat{\boldsymbol{\rho}}_{n,1}^{(R)} = \frac{\boldsymbol{\tau}_1 Z_{n,1}^{(R)} - \boldsymbol{\tau}_5 \omega(1, \tilde{\boldsymbol{\theta}})}{Z_{n,1}^{(R)} - \omega(1, \tilde{\boldsymbol{\theta}})} \mathbb{1}_{\left\{Z_{n,1}^{(R)} \in \omega(1, \tilde{\boldsymbol{\theta}}) \bullet (1, \widetilde{\psi}_1(\boldsymbol{\tau}_4, \boldsymbol{\tau}_1, \boldsymbol{\tau}_4, \boldsymbol{\tau}_5))\right\}}.$$

Remarquons que cet estimateur coïncide avec celui proposé dans Goegebeur et al. (2010), Lemme 1.

Exemple 4.4.3 *Soit $\boldsymbol{\delta} = 0$ c'est-à-dire $\tilde{\boldsymbol{\theta}}_1 = \tilde{\boldsymbol{\theta}}_2$. Notons $Z_{n,2}^{(R)}$ la variable aléatoire $Z_n^{(R)}$ correspondant à ce choix de paramètre. Ici également on est dans la situation (i) et un*

nouvel estimateur de ρ est obtenu:

$$\hat{\boldsymbol{\rho}}_{n,2}^{(R)} = \frac{\boldsymbol{\tau}_1 Z_{n,2}^{(R)} - \boldsymbol{\tau}_4 \omega(0, \tilde{\boldsymbol{\theta}})}{Z_{n,2}^{(R)} - \omega(0, \tilde{\boldsymbol{\theta}})} 1\!\!1_{\left\{Z_{n,2}^{(R)} \in \omega(0,\tilde{\boldsymbol{\theta}}) \bullet (1, \widetilde{\psi}_0(\boldsymbol{\tau}_4, \boldsymbol{\tau}_1, \boldsymbol{\tau}_4, \boldsymbol{\tau}_5))\right\}}.$$

Exemple 4.4.4 *Soit $\boldsymbol{\tau}_1 = \boldsymbol{\tau}_5$. Dans ce cas $\boldsymbol{\delta}_0 = \boldsymbol{\delta}_1 = 1$ et on se place dans la situation (i) si $\boldsymbol{\delta} < 1$ et dans la situation (ii) sinon. Pour ce choix, si on note par $Z_{n,3}^{(R)}$ la variable aléatoire $Z_n^{(R)}$, une famille de nouveaux estimateurs de ρ est donnée par:*

$$\hat{\boldsymbol{\rho}}_{n,3}^{(R)} = \frac{\boldsymbol{\tau}_4 (Z_{n,3}^{(R)}/\omega(\boldsymbol{\delta}, \tilde{\boldsymbol{\theta}}))^{1/(\boldsymbol{\delta}-1)} - \boldsymbol{\tau}_1}{(Z_{n,3}^{(R)}/\omega(\boldsymbol{\delta}, \tilde{\boldsymbol{\theta}}))^{1/(\boldsymbol{\delta}-1)} - 1} 1\!\!1_{\{Z_{n,3}^{(R)} \in J\}}.$$

4.4.2 Estimateurs basés sur la statistique $S_{k,n}(\boldsymbol{\tau})$

La variable aléatoire $T_n^{(S)}$ dépend de 24 paramètres: $\{(\boldsymbol{\theta}_i, \boldsymbol{\tau}_i, \boldsymbol{\alpha}_i) \in (0, \infty)^3,\ i = 1, \ldots, 8\}$. Soit $(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_4) \in (0, \infty)^4$ avec $\boldsymbol{\zeta}_3 \neq \boldsymbol{\zeta}_4$. Dans la suite, nous allons supposer que

(C6) $\left\{\boldsymbol{\theta}_i \boldsymbol{\alpha}_i = \boldsymbol{\zeta}_{\lceil i/2 \rceil},\ i = 1, \ldots, 8\right\}$ avec $\boldsymbol{\delta} = (\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_2)/(\boldsymbol{\zeta}_3 - \boldsymbol{\zeta}_4)$. De plus, $(\boldsymbol{\tau}_{2i-1}, \boldsymbol{\alpha}_{2i-1}) \neq (\boldsymbol{\tau}_{2i}, \boldsymbol{\alpha}_{2i})$, pour $i = 1, \ldots, 4$ et $(\boldsymbol{\tau}_{2i-1}, \boldsymbol{\alpha}_{2i-1}) < (\boldsymbol{\tau}_{2i}, \boldsymbol{\alpha}_{2i})$, pour $i = 3, 4$, où

- $((x, y) = (s, t))$ signifie que $x = s$ et/ou $y = t$ et
- $(x, y) < (s, t)$ signifie que $x < s$ et $y \leq t$ ou $x = s$ et $y \leq t$.

Sous ces conditions, $T_n^{(S)}$ se retrouve avec 20 paramètres libres. Nous introduisons les notations suivantes: $Z_n^{(S)} = \psi_{\boldsymbol{\delta}}(T_n^{(S)})$ et $\varphi_{\boldsymbol{\delta}}^{(S)} = \psi_{\boldsymbol{\delta}} \circ f^{(S)}$. Puisque $\boldsymbol{\delta} = (\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_2)/(\boldsymbol{\zeta}_3 - \boldsymbol{\zeta}_4)$, il est donc facile de voir que $Z_n^{(S)}$ ne dépend pas du paramètre inconnu $\boldsymbol{\gamma}$. Nous allons également établir la normalité asymptotique de l'estimateur de $\hat{\boldsymbol{\rho}}_n^{(S)}$ défini ci-dessous, lorsque la statistique $T_n^{(S)}$ et la fonction $\psi_{\boldsymbol{\delta}}$ sont utilisées:

$$\hat{\boldsymbol{\rho}}_n^{(S)} = \begin{cases} \left(\varphi_{\boldsymbol{\delta}}^{(S)}\right)^{-1} (Z_n^{(S)}) & \text{si } Z_n^{(S)} \in J \\ 0 & \text{sinon.} \end{cases} \tag{4.10}$$

La condition supplémentaire suivante est réquise:

(C7) Pour tout $\rho < 0$, la fonction $\nu_\rho(\boldsymbol{\tau}, \boldsymbol{\alpha}) = \mu(G_{\boldsymbol{\tau},\boldsymbol{\alpha}} J_{1-\boldsymbol{\alpha}} K_\rho)$ est différentiable avec $\frac{\partial}{\partial \boldsymbol{\tau}} \nu_\rho(\boldsymbol{\tau}, \boldsymbol{\alpha}) > 0$ et $\frac{\partial}{\partial \boldsymbol{\alpha}} \nu_\rho(\boldsymbol{\tau}, \boldsymbol{\alpha}) > 0$ pour tout $\boldsymbol{\alpha} > 0$ et pour tout $\boldsymbol{\tau} \in \mathbb{R}$.

Définissons pour tout $i \in \{1, \ldots, 4\}$,

$$m_A^{(S,i)} = \exp\left\{\frac{(\boldsymbol{\alpha}_{2i-1}-1)\mu(G_{\boldsymbol{\tau}_{2i-1},\boldsymbol{\alpha}_{2i-1}}J_{2-\boldsymbol{\alpha}_{2i-1}}K_{-\rho}^2) - (\boldsymbol{\alpha}_{2i}-1)\mu(G_{\boldsymbol{\tau}_{2i},\boldsymbol{\alpha}_{2i}}J_{2-\boldsymbol{\alpha}_{2i}}K_{-\rho}^2)}{\mu(G_{\boldsymbol{\tau}_{2i},\boldsymbol{\alpha}_{2i}}J_{1-\boldsymbol{\alpha}_{2i}}K_{-\rho}) - \mu(G_{\boldsymbol{\tau}_{2i-1},\boldsymbol{\alpha}_{2i-1}}J_{1-\boldsymbol{\alpha}_{2i-1}}K_{-\rho})}\right\},$$

$$m_B^{(S,i)} = \exp\left\{\frac{\mu(G_{\boldsymbol{\tau}_{2i-1},\boldsymbol{\alpha}_{2i-1}}J_{1-\boldsymbol{\alpha}_{2i-1}}L_{(-\rho,-\beta)}) - \mu(G_{\boldsymbol{\tau}_{2i},\boldsymbol{\alpha}_{2i}}J_{1-\boldsymbol{\alpha}_{2i}}L_{(-\rho,-\beta)})}{\mu(G_{\boldsymbol{\tau}_{2i},\boldsymbol{\alpha}_{2i}}J_{1-\boldsymbol{\alpha}_{2i}}K_{-\rho}) - \mu(G_{\boldsymbol{\tau}_{2i-1},\boldsymbol{\alpha}_{2i-1}}J_{1-\boldsymbol{\alpha}_{2i-1}}K_{-\rho})}\right\},$$

et pour $u \in [0,1]$,

$$v^{(S,i)}(u) = \frac{G_{\boldsymbol{\tau}_{2i-1},\boldsymbol{\alpha}_{2i-1}}(u)J_{1-\boldsymbol{\alpha}_{2i-1}}(u) - G_{\boldsymbol{\tau}_{2i},\boldsymbol{\alpha}_{2i}}(u)J_{1-\boldsymbol{\alpha}_{2i}}(u)}{\mu(G_{\boldsymbol{\tau}_{2i},\boldsymbol{\alpha}_{2i}}J_{1-\boldsymbol{\alpha}_{2i}}K_{-\rho}) - \mu(G_{\boldsymbol{\tau}_{2i-1},\boldsymbol{\alpha}_{2i-1}}J_{1-\boldsymbol{\alpha}_{2i-1}}K_{-\rho})}.$$

Pour simplifier, nous allons considérer $m_A^{(S)} = (m_A^{(S,i)},\ i = 1, \ldots, 4)$, $m_B^{(S)} = (m_B^{(S,i)},\ i = 1, \ldots, 4)$ et $\widetilde{v}^S = v^{(S,1)} - v^{(S,2)} - \boldsymbol{\delta}(v^{(S,3)} - v^{(S,4)})$.
Le résultat suivant est une suite directe du Théorème 4.2.2 et du Lemme 4.3.2.

Corollaire 4.4.5 *Supposons que les conditions* **(C2)**, **(C3)**, **(C6)** *et* **(C7)** *soient satisfaites. S'il existe deux intervalles J et J_0 tels que pour tout $\rho \in J_0$ et une suite d'entiers k_n satisfaisant (4.6), alors*

$$k_n{}^{1/2}A(n/k_n)(\hat{\boldsymbol{\rho}}_n^{(S)} - \rho) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda_A}{2\boldsymbol{\gamma}}\mathcal{AB}_1^{(S)}(\boldsymbol{\delta},\rho) + \lambda_B\mathcal{AB}_2^{(S)}(\boldsymbol{\delta},\rho,\beta),\ \boldsymbol{\gamma}^2\mathcal{AV}^{(S)}(\boldsymbol{\delta},\rho)\right)$$

où

$$\mathcal{AB}_1^{(S)}(\boldsymbol{\delta},\rho) = \varphi_{\boldsymbol{\delta}}^{(S)}(\rho)\log\tilde{\psi}_{\boldsymbol{\delta}}(m_A^{(S)})/[\varphi_{\boldsymbol{\delta}}^{(S)}]'(\rho),$$

$$\mathcal{AB}_2^{(S)}(\boldsymbol{\delta},\rho,\beta) = \varphi_{\boldsymbol{\delta}}^{(S)}(\rho)\log\tilde{\psi}_{\boldsymbol{\delta}}(m_B^{(S)})/[\varphi_{\boldsymbol{\delta}}^{(S)}]'(\rho)$$

et

$$\mathcal{AV}^{(S)}(\boldsymbol{\delta},\rho) = \left(\varphi_{\boldsymbol{\delta}}^{(S)}(\rho)\right)^2\vartheta\left(\widetilde{v}^S,\widetilde{v}^S\right)/\left([\varphi_{\boldsymbol{\delta}}^{(S)}]'(\rho)\right)^2.$$

Nous allons maintenant montrer que les Propositions 5,7 et 9 de Ciuperca et Mercadier (2010) représentent des cas particuliers du Corollaire 4.4.5 pour trois valeurs différentes de $\boldsymbol{\delta}$ ($\boldsymbol{\delta} = 2$, $\boldsymbol{\delta} = 1$ et $\boldsymbol{\delta} = 0$ respectivement). La normalité asymptotique des estimateurs proposés dans Gomes *et al.* (2002) et dans Alves *et al.* (2003b) peut aussi être facilement etablie avec le Corollaire 4.4.5.

Comme exemple de fonction $G_{\boldsymbol{\tau},\boldsymbol{\alpha}}$, on peut considérer la fonction définie sur $[0,1]$ par:

$$G_{\boldsymbol{\tau},\boldsymbol{\alpha}}(u) = \frac{g_{\boldsymbol{\tau}-1}(u)}{\int_0^1 g_{\boldsymbol{\tau}-1}(x)J_{-\boldsymbol{\alpha}}(x)dx} \text{ pour } \boldsymbol{\tau} \geq 1 \text{ et } \boldsymbol{\alpha} > 0,$$

où la fonction $g_{\boldsymbol{\tau}}$ est donnée par

$$g_0(x) = 1, \quad g_{\boldsymbol{\tau}-1}(x) = \frac{\boldsymbol{\tau}}{\boldsymbol{\tau}-1}(1 - x^{\boldsymbol{\tau}-1}), \forall \boldsymbol{\tau} > 1.$$

Nous voyons clairement que $G_{\boldsymbol{\tau},\boldsymbol{\alpha}}$ satisfait les conditions **(C7)**, et sous la condition **(C6)**, l'expression de $\varphi_\rho^{(S)}$ est donnée par

$$\varphi_{\boldsymbol{\delta}}^{(S)}(\rho) = \frac{\boldsymbol{\zeta}_1}{\boldsymbol{\zeta}_2}\left(\frac{\boldsymbol{\zeta}_4}{\boldsymbol{\zeta}_3}\right)^{\boldsymbol{\delta}} \left[\frac{\nu_\rho(\boldsymbol{\tau}_1,\boldsymbol{\alpha}_1) - \nu_\rho(\boldsymbol{\tau}_2,\boldsymbol{\alpha}_2)}{\nu_\rho(\boldsymbol{\tau}_3,\boldsymbol{\alpha}_3) - \nu_\rho(\boldsymbol{\tau}_4,\boldsymbol{\alpha}_4)}\right] \left[\frac{\nu_\rho(\boldsymbol{\tau}_7,\boldsymbol{\alpha}_7) - \nu_\rho(\boldsymbol{\tau}_8,\boldsymbol{\alpha}_8)}{\nu_\rho(\boldsymbol{\tau}_5,\boldsymbol{\alpha}_5) - \nu_\rho(\boldsymbol{\tau}_6,\boldsymbol{\alpha}_6)}\right]^{\boldsymbol{\delta}}$$

avec

$$\nu_\rho(\boldsymbol{\tau},\boldsymbol{\alpha}) = \frac{1-(1-\rho)^{-\boldsymbol{\alpha}} + (\boldsymbol{\tau}-\rho)^{-\boldsymbol{\alpha}} - \boldsymbol{\tau}^{-\boldsymbol{\alpha}}}{\boldsymbol{\alpha}\rho(1-\boldsymbol{\tau}^{-\boldsymbol{\alpha}-1})} \quad \text{si} \quad \boldsymbol{\tau} \neq 1 \quad \text{et} \quad \nu_\rho(1,\boldsymbol{\alpha}) = \frac{(1-\rho)^{\boldsymbol{\alpha}} - 1}{\boldsymbol{\alpha}\rho(1-\rho)^{\boldsymbol{\alpha}}}.$$

Même si le Corollaire 4.4.5 assure l'existence des intervalles J_0 et J, il sera impossible de les spécifier dans le cas général. Dans la suite, nous considérons plusieurs combinaisons de paramètres pour lesquelles ces intervalles seront exhibés et la fonction inverse $\varphi_{\boldsymbol{\delta}}^{(S)}$ admettra une forme explicite. Ainsi, nous allons donc supposer que $\boldsymbol{\tau}_2 = \boldsymbol{\tau}_3 = \boldsymbol{\tau}_5 = \boldsymbol{\tau}_6 = \boldsymbol{\tau}_7 = \boldsymbol{\tau}_8 = \boldsymbol{\alpha}_7 = 1$, $\boldsymbol{\alpha}_6 = 3$, $\boldsymbol{\alpha}_8 = 2$ et nous introduisons la notation suivante:

$$\omega^*(\boldsymbol{\delta},\boldsymbol{\zeta}) = \frac{\boldsymbol{\zeta}_1}{\boldsymbol{\zeta}_2}\left(\frac{3\boldsymbol{\zeta}_4}{\boldsymbol{\zeta}_3}\right)^{\boldsymbol{\delta}}.$$

Dans tous les exemples ci-dessous, $J_0 = \mathbb{R}^-$ et donc la condition $\rho \in J_0$ est toujours vérifiée. Les trois premiers exemples correspondent à des estimateurs existants du paramètre de second ordre, tandis que les trois derniers donnent lieu à de nouveaux estimateurs.

Exemple 4.4.6 *Soit $\boldsymbol{\delta} = 0$ (c'est-à-dire $\boldsymbol{\zeta}_1 = \boldsymbol{\zeta}_2$), $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_3 = \boldsymbol{\alpha}_4 = 1$, $\boldsymbol{\tau}_1 = 2$ et $\boldsymbol{\tau}_4 = 3$. Notons $Z_{n,4}^{(S)}$ la variable aléatoire $Z_n^{(S)}$ pour ce choix de paramètres, un estimateur de ρ est donné par:*

$$\hat{\boldsymbol{\rho}}_{n,4}^{(S)} = \frac{6(Z_{n,4}^{(S)}+2)}{3Z_{n,4}^{(S)}+4} 1\!\!1_{\left\{Z_{n,4}^{(S)} \in (-2,-4/3)\right\}}.$$

Notons que cet estimateur correspond à l'estimateur $\hat{\boldsymbol{\rho}}_{n,k}^{[2]}$ défini dans la section 5.2 de Ciuperca et Mercadier (2010).

Exemple 4.4.7 *Soit $\boldsymbol{\delta} = 0$ ($\boldsymbol{\zeta}_1 = \boldsymbol{\zeta}_2$), $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_3 = \boldsymbol{\alpha}_4 = 1$, $\boldsymbol{\tau}_1 = \boldsymbol{\tau}_4 = \boldsymbol{\alpha}_2 = 2$. Notons $Z_{n,5}^{(S)}$ la variable aléatoire pour ce choix de paramètres, nous retrouvons l'estimateur $\hat{\boldsymbol{\rho}}_{n,k}^{[3]}$ proposé dans la section 5.2 de Ciuperca et Mercadier (2010):*

$$\hat{\boldsymbol{\rho}}_{n,5}^{(S)} = \frac{2(Z_{n,5}^{(S)}-2)}{2Z_{n,5}^{(S)}-1} 1\!\!1_{\left\{Z_{n,5}^{(S)} \in (1/2,2)\right\}}.$$

Exemple 4.4.8 *Soit* $\boldsymbol{\alpha}_1 = \boldsymbol{\zeta}_1 = 4$, $\boldsymbol{\alpha}_3 = \boldsymbol{\alpha}_8 = \boldsymbol{\zeta}_2 = \boldsymbol{\zeta}_4 = 2$, $\boldsymbol{\alpha}_6 = \boldsymbol{\zeta}_3 = 3$ *et* $\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_4 = \boldsymbol{\alpha}_7 = \boldsymbol{\tau}_1 = \boldsymbol{\tau}_4 = 1$. *Notons* $Z_{n,6}^{(S)}$ *la variable aléatoire* $Z_n^{(S)}$ *pour ce choix de paramètres, l'estimateur de* ρ *est donné par:*

$$\hat{\boldsymbol{\rho}}_{n,6}^{(S)} = \frac{6(Z_{n,6}^{(S)} - 4 + 3(Z_{n,6}^{(S)} - 2)^{1/2}}{4Z_{n,6}^{(S)} - 3} \mathbb{1}_{\left\{Z_{n,6}^{(S)} \in (2/3, 3/4)\right\}},$$

et correspond à celui proposé par Gomes et al. (2002), Equation (12).

Exemple 4.4.9 *Considérons le cas* $\boldsymbol{\delta} = 1$ *(i.e.* $\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_2 = \boldsymbol{\zeta}_3 - \boldsymbol{\zeta}_4$*),* $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_3 = \boldsymbol{\alpha}_4 = 1$, $\boldsymbol{\tau}_1 = \boldsymbol{\alpha}_5 = \boldsymbol{\alpha}_8 = 2$ *et* $\boldsymbol{\tau}_4 = \boldsymbol{\alpha}_6 = 3$. *Notons* $Z_{n,7}^{(S)}$ *la variable aléatoire* $Z_n^{(S)}$*, un nouvel estimateur de* ρ *est donné par:*

$$\hat{\boldsymbol{\rho}}_{n,7}^{(S)} = \frac{6Z_{n,7}^{(S)} + 4\omega^*(1, \boldsymbol{\zeta})}{3Z_{n,7}^{(S)} + 4\omega^*(1, \boldsymbol{\zeta})} \mathbb{1}_{\left\{Z_{n,7}^{(S)} \in \omega^*(1, \boldsymbol{\zeta}) \bullet (-4/3, -2/3)\right\}}.$$

Exemple 4.4.10 *Soit* $\boldsymbol{\delta} = 1$, $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_3 = \boldsymbol{\alpha}_4 = 1$, $\boldsymbol{\tau}_1 = \boldsymbol{\tau}_4 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_5 = \boldsymbol{\alpha}_8 = 2$ *et* $\boldsymbol{\alpha}_6 = 3$. *Notons pour ce choix* $Z_{n,8}^{(S)}$ *la variable aléatoire* $Z_n^{(S)}$*, un nouvel estimateur de* ρ *est donné par:*

$$\hat{\boldsymbol{\rho}}_{n,8}^{(S)} = \frac{6Z_{n,8}^{(S)} - 4\omega^*(1, \boldsymbol{\zeta})}{2Z_{n,8}^{(S)} - \omega^*(1, \boldsymbol{\zeta})} \mathbb{1}_{\left\{Z_{n,8}^{(S)} \in \omega^*(1, \boldsymbol{\zeta}) \bullet (1/2, 2/3)\right\}}.$$

Exemple 4.4.11 *Soit* $\boldsymbol{\alpha}_1 = \boldsymbol{\tau}_1 = \boldsymbol{\tau}_4 = 1$, $\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_3 = 2$ *et* $\boldsymbol{\alpha}_4 = 3$. *Notons pour ce choix* $Z_{n,9}^{(S)}$ *la variable aléatoire* $Z_n^{(S)}$*, un nouvel estimateur de* ρ *est donné par:*

$$\hat{\boldsymbol{\rho}}_{n,9}^{(S)} = \frac{3\left(Z_{n,9}^{(S)}/(3\omega^*(\boldsymbol{\delta}, \boldsymbol{\zeta}))\right)^{1/(\boldsymbol{\delta}+1)} - 1}{\left(Z_{n,9}^{(S)}/(3\omega^*(\boldsymbol{\delta}, \boldsymbol{\zeta}))\right)^{1/(\boldsymbol{\delta}+1)} - 1} \mathbb{1}_{\left\{Z_{n,9}^{(S)} \in \omega^*(\boldsymbol{\delta}, \boldsymbol{\zeta}) \bullet (3^{-\boldsymbol{\delta}}, 3)\right\}}.$$

Pour résumer, nous avons illustré comment le Théorème 4.2.2 peut être utilisé pour montrer la normalité asymptotique de ces estimateurs obtenus à partir des variables aléatoires $T_n^{(R)}$ et $T_n^{(S)}$: le Corollaire 4.4.1 et le Corollaire 4.4.5 traitent un grand nombre d'estimateurs proposés dans la littérature. Cinq (5) nouveaux estimateurs de ρ ont été introduits: $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$, $\hat{\boldsymbol{\rho}}_{n,3}^{(R)}$, $\hat{\boldsymbol{\rho}}_{n,7}^{(S)}$, $\hat{\boldsymbol{\rho}}_{n,8}^{(S)}$ et $\hat{\boldsymbol{\rho}}_{n,9}^{(S)}$. Tous ces estimateurs sont explicites et asymptotiquement Gaussiens. La comparaison de leurs propriétés pour des échantillons de tailles finies est une lourde tâche, car elles dépendent des paramètres ($\boldsymbol{\theta}_i$, $\boldsymbol{\tau}_i$, $\boldsymbol{\alpha}_i$) et de la distribution simulée. Ce point est au-delà du champ d'application de ce chapitre. Nous concluons cette étude en proposant une méthode pour sélectionner certains paramètres asymptotiques optimaux dans une familles d'estimateurs. Les performances et les limites de cette technique sont illustrées en comparant certains estimateurs sur des données simulées. Nous proposerons également une application sur des données d'assurance.

4.5 Comparaisons de quelques estimateurs

Nous comparons maintenant quelques estimateurs de ρ pour un modèle spécifique de type Pareto, à savoir la distribution de Burr avec comme fonction de répartition

$$F(x) = 1 - (\xi/(\xi + x^{\eta}))^{\lambda}, \ x > 0, \ \xi, \lambda, \eta > 0,$$

et qui considérée par exemple dans Beirlant *et al.* (1998), Equation (3). L'indice des valeurs extrêmes associé est $\boldsymbol{\gamma} = 1/(\lambda\eta)$ et ce modèle satisfait la condition du troisième ordre **(C2)** avec $\rho = \beta = -1/\lambda$, $A(x) = \boldsymbol{\gamma} x^{\rho}/(1 - x^{\rho})$ et $B(x) = \rho x^{\rho}/(1 - x^{\rho})$. Nous nous limitons seulement au cas $\xi = 1$ et $\lambda = 1/\eta$ c'est-à-dire $\boldsymbol{\gamma} = 1$.

4.5.1 Estimateurs basés sur la variable aléatoire $R_{k,n}(\boldsymbol{\tau})$

Nous nous concentrons dans un premier temps sur les estimateurs de ρ basés sur la statistique $R_{k_n}(\boldsymbol{\tau}_i)$ considerée dans la sous-section 4.4.1 avec le noyau $H_{\boldsymbol{\tau}_i}(u) = \boldsymbol{\tau}_i u^{\boldsymbol{\tau}_i - 1}$, pour $i = 1, \ldots, 8$. Les valeurs des paramètres $\boldsymbol{\tau}_1, ..., \boldsymbol{\tau}_8$, $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_3$ et $\boldsymbol{\theta}_4$ sont prises comme dans Ì Ì dans Goegebeur *et al.* (2010), Goegebeur et de Wet (2012): $\boldsymbol{\tau}_i = \tilde{\boldsymbol{\tau}}_i + 1$ avec $\tilde{\boldsymbol{\tau}}_1 = 0.25$, $\tilde{\boldsymbol{\tau}}_2 = \tilde{\boldsymbol{\tau}}_3 = 0.75$, $\tilde{\boldsymbol{\tau}}_4 = \tilde{\boldsymbol{\tau}}_8 = 1$, $\tilde{\boldsymbol{\tau}}_5 = 0.5$, $\tilde{\boldsymbol{\tau}}_6 = \tilde{\boldsymbol{\tau}}_7 = 0.75$, $\tilde{\boldsymbol{\theta}}_1 = 0.01$, $\tilde{\boldsymbol{\theta}}_3 = 0.02$ et $\tilde{\boldsymbol{\theta}}_4 = 0.04$. Selon ces auteurs, ces valeurs donnent de bons résultats pour les distributions satisfaisant la condition du troisième ordre (**C2**) avec $\rho = \beta$. Pour ces paramètres, nous obtenons l'expression de $\varphi_{\boldsymbol{\delta}}^{(R)}$ de (4.9) et nous avons pris également $\boldsymbol{\delta}_0 = 1.5$ et $\boldsymbol{\delta}_1 = 1.8$. Rappelons que $\tilde{\boldsymbol{\theta}}_2 = \tilde{\boldsymbol{\theta}}_1 + \boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}_4 - \tilde{\boldsymbol{\theta}}_3)$ pour $\boldsymbol{\delta} \geq 0$. Dans la suite, nous proposons de choisir le paramètre $\boldsymbol{\delta}$ en utilisant une méthode similaire à celle proposée dans Gardes et Girard (2010). Elle consiste à minimiser une borne supérieure de l'erreur quadratique moyenne asymptotique en fonction de $\boldsymbol{\delta}$. Cette méthode est décrite dans la Sous-section 4.5.1.1 et un exemple d'application est présenté dans la Sous-section 4.5.1.2.

4.5.1.1 Borne supérieure de l' Erreur Quadratique Moyenne Asymptotique

Comme dans Goegebeur *et al.* (2010), nous supposons que $\rho = \beta$. Suivant le Corollaire 4.4.1, les composantes du biais asymptotique de $\hat{\boldsymbol{\rho}}_n^{(R)}$ sont respectivement proportionnelles à $\mathcal{AB}_1^{(R)}(\boldsymbol{\delta}, \rho)$ et $\mathcal{AB}_2^{(R)}(\boldsymbol{\delta}, \rho, \rho)$ tandis que leur variance asymptotique est proportionnelle à $\mathcal{AV}^{(R)}(\boldsymbol{\delta}, \rho)$. L'erreur quadratique moyenne asymptotique ($\mathcal{AMSE}$) de $\hat{\boldsymbol{\rho}}_n^{(R)}$ peut être définie par

$$\mathcal{AMSE}(\boldsymbol{\delta}, \boldsymbol{\gamma}, \rho) = \frac{1}{kA^2(n/k_n)} \left(\left(\frac{\lambda_A}{2\boldsymbol{\gamma}} \mathcal{AB}_1^{(R)}(\boldsymbol{\delta}, \rho) - \lambda_B \mathcal{AB}_2^{(R)}(\boldsymbol{\delta}, \rho, \rho) \right)^2 + \boldsymbol{\gamma}^2 \mathcal{AV}^{(R)}(\boldsymbol{\delta}, \rho) \right). \tag{4.11}$$

Une façon de choisir le paramètre $\boldsymbol{\delta}$ est de minimiser l'erreur quadratique moyenne asymptotique ci dessus. En pratique, les paramètres $\boldsymbol{\gamma}$, ρ ainsi que les fonctions A et B sont inconnus et donc il sera impossible d'évaluer l'erreur quadratique moyenne asymptotique. Pour résoudre ce problème, il est possible d'introduire une borne supérieure sur $\mathcal{AMSE}(\boldsymbol{\delta}, \boldsymbol{\gamma}, \rho)$. Supposons que $0 \leq \boldsymbol{\delta} < \boldsymbol{\delta}_0$ ou $\boldsymbol{\delta} > \boldsymbol{\delta}_1$ et $\rho_{\mathbf{min}} \leq \rho \leq \rho_{\mathbf{max}}$, il est facile de voir que $|\mathcal{AB}_1^{(R)}(\boldsymbol{\delta}, \rho)| \geq |\mathcal{AB}_1^{(R)}(\boldsymbol{\delta}_1, \rho_{\mathbf{max}})|$, $|\mathcal{AB}_2^{(R)}(\boldsymbol{\delta}, \rho, \rho)| \geq |\mathcal{AB}_2^{(R)}(\boldsymbol{\delta}_0, \rho_{\mathbf{min}}, \rho_{\mathbf{min}})|$ et numériquement nous avons aussi $\mathcal{AV}^{(R)}(\boldsymbol{\delta}, \rho) \geq \mathcal{AV}^{(R)}(1.32, -0.46)$. Nous avons donc:

$$\mathcal{AMSE}(\boldsymbol{\delta}, \boldsymbol{\gamma}, \rho) \leq \frac{C\pi(\boldsymbol{\delta}, \rho)}{kA^2(n/k_n)},$$

avec $\pi(\boldsymbol{\delta}, \rho) = (\mathcal{AB}_1^{(R)}(\boldsymbol{\delta}, \rho)\mathcal{AB}_2^{(R)}(\boldsymbol{\delta}, \rho, \rho))^2\mathcal{AV}^{(R)}(\boldsymbol{\delta}, \rho)$ et où la constante C ne dépend ni de $\boldsymbol{\delta}$ ni de ρ. Considérons pour $\rho < 0$ le paramètre $\boldsymbol{\delta}$ minimisant la fonction $\pi(\boldsymbol{\delta}, \rho)$. Par exemple, quand ρ est au voisinage de 0, on peut montrer que la valeur optimale est $\boldsymbol{\delta} = \boldsymbol{\delta}_0 = 1.5$.

4.5.1.2 Illustration sur une distribution de Burr

Ici, trois estimateurs sont à comparer:

➠ L'estimateur $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$ proposé dans Goegebeur *et al.* (2010) ($\boldsymbol{\delta} = 1$, voir Exemple 4.4.2),

➠ Le nouvel estimateur explicite $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ introduit dans l'Exemple 4.4.3 ($\boldsymbol{\delta} = 0$),

➠ Le nouvel estimateur explicite défini par $\hat{\boldsymbol{\rho}}_{n,0}^{(R)} = \rho_n^{(R)}$ avec $\boldsymbol{\delta} = \boldsymbol{\delta}_0 = 1.5$, voir l'équation (4.8).

Nous considérons trois ensembles de paramètres: $\xi = 1$, $\lambda = 1/\eta$, $\rho = -\eta$ avec $\eta \in \{0.25, 1\}$. D'abord, les estimateurs sont comparés sur la base de leurs erreurs quadratiques moyennes asymptotiques. En prenant $n = 5000$ et $k \in \{1500, ..., 4999\}$, nous évaluons $\lambda_A = {k_n}^{1/2}A^2(n/k_n)$ et $\lambda_B = \rho\lambda_A/\boldsymbol{\gamma}$ puis les erreurs quadratiques moyennes asymptotiques des $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$, $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ et $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ en fonction de k_n, correspondant respectivement à $\mathcal{AMSE}(1, 1, \rho)$, $\mathcal{AMSE}(0, 1, \rho)$ et $\mathcal{AMSE}(1.5, 1, \rho)$ avec $\rho \in \{-0.25, -1\}$. Les résultats obtenus sont représentés sur les graphes de gauche de la Figure 4.1. Il apparait que $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ donne le meilleur résultat pour $\rho \geq -1$. Ce qui est en conformité avec le résultat du paragraphe précédent: $\boldsymbol{\delta} = 1.5$ est « *optimal* » quand ρ est proche de 0. Comme conclusion préliminaire, le critère de la borne supérieure $\pi(\cdot)$ semble être bien adapté

pour le choix des paramètres de l'estimateur. Contrairement au cas $\rho < -1$, le meilleur estimateur du point de vue erreur quadratique moyenne asymptotique est donné par $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$.
Ensuite, nous comparons la performance de ces estimateurs sur un échantillon de taille finie. Considérons $k \in \{1500, ..., 4999\}$, l'erreur quadratique moyenne empirique des estimateurs $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$, $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ et $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ est évaluée sur 500 réplications d'un échantillon de taille $n = 5000$ par le modèle précédent. Les résultats obtenus sont représentés sur les graphes de droite de la Figure 4.1. Les conclusions sont qualitativement les mêmes: $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ donne les meilleurs résultats lorsque $\rho \geq 1$ et $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ donne les meilleurs résultats dans le cas $\rho < -1$. Par conséquent $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$, n'est jamais le meilleur estimateur pour la situation considérée dans cette partie. En pratique, le cas $\rho \geq -1$ est plus intéréssant, puisqu'il correspond à un biais élevé. Pour cette raison, il nous semble que $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ doit être préféré.

4.5.2 Estimateurs basés sur la variable aléatoire $S_{k_n}(\tau, \alpha)$

Considérons maintenant les estimateurs de ρ basés sur les variables aléatoires $S_{k_n}(\boldsymbol{\tau}_i, \boldsymbol{\alpha}_i)$, $i = 1, \ldots, 8$ données dans la section 4.4.2 dans le cas$(\boldsymbol{\tau}_1, \boldsymbol{\alpha}_1) = (\boldsymbol{\tau}_7, \boldsymbol{\alpha}_7)$, $(\boldsymbol{\tau}_2, \boldsymbol{\alpha}_2) = (\boldsymbol{\tau}_8, \boldsymbol{\alpha}_8)$, $(\boldsymbol{\tau}_3, \boldsymbol{\alpha}_3) = (\boldsymbol{\tau}_5, \boldsymbol{\alpha}_5)$ et $(\boldsymbol{\tau}_4, \boldsymbol{\alpha}_4) = (\boldsymbol{\tau}_6, \boldsymbol{\alpha}_6)$.
Dans le Paragraphe 4.5.2.1, nous montrons que l'erreur quadratique moyenne asymptotique est indépendante des paramètres $\boldsymbol{\theta}_i$, $i = 1, \ldots, 8$ et $\boldsymbol{\delta}$.

4.5.2.1 Comparaison des termes de l'erreur quadratique moyenne asymptotique

D'après le Corollaire 4.4.5, les composantes du biais et de la variance de $\hat{\rho}_n^{(S)}$ sont respectivement proportionnelles à

$$\begin{aligned}
\mathcal{AB}_1^{(S)}(\boldsymbol{\delta}, \rho) &= \frac{g^{(S)}(\rho)}{(g^{(S)})'(\rho)} \left(\log m_A^{(S,1)} - \log m_A^{(S,2)} \right), \\
\mathcal{AB}_2^{(S)}(\boldsymbol{\delta}, \rho) &= \frac{g^{(S)}(\rho)}{(g^{(S)})'(\rho)} \left(\log m_B^{(S,1)} - \log m_B^{(S,2)} \right), \\
\mathcal{AV}^{(S)}(\boldsymbol{\delta}, \rho) &= \left(\frac{g^{(S)}(\rho)}{(g^{(S)})'(\rho)} \right)^2 \vartheta(v^{(S,1)} - v^{(S,2)}, v^{(S,1)} - v^{(S,2)}),
\end{aligned}$$

où

$$g^{(S)}(\rho) = \frac{\boldsymbol{\zeta}_1}{\boldsymbol{\zeta}_2} \frac{\mu(G_{\boldsymbol{\tau}_1,\boldsymbol{\alpha}_1} J_{1-\boldsymbol{\alpha}_1} K_\rho) - \mu(G_{\boldsymbol{\tau}_2,\boldsymbol{\alpha}_2} J_{1-\boldsymbol{\alpha}_2} K_\rho)}{\mu(G_{\boldsymbol{\tau}_3,\boldsymbol{\alpha}_3} J_{1-\boldsymbol{\alpha}_3} K_\rho) - \mu(G_{\boldsymbol{\tau}_4,\boldsymbol{\alpha}_4} J_{1-\boldsymbol{\alpha}_4} K_\rho)}.$$

Il apparaît ainsi que l'erreur quadratique moyenne asymptotique (définie de manière similaire à (4.11)) ne dépend pas des $\boldsymbol{\theta}_i$, $i = 1, \ldots, 8$ et $\boldsymbol{\delta}$. Du point de vue asymptotique,

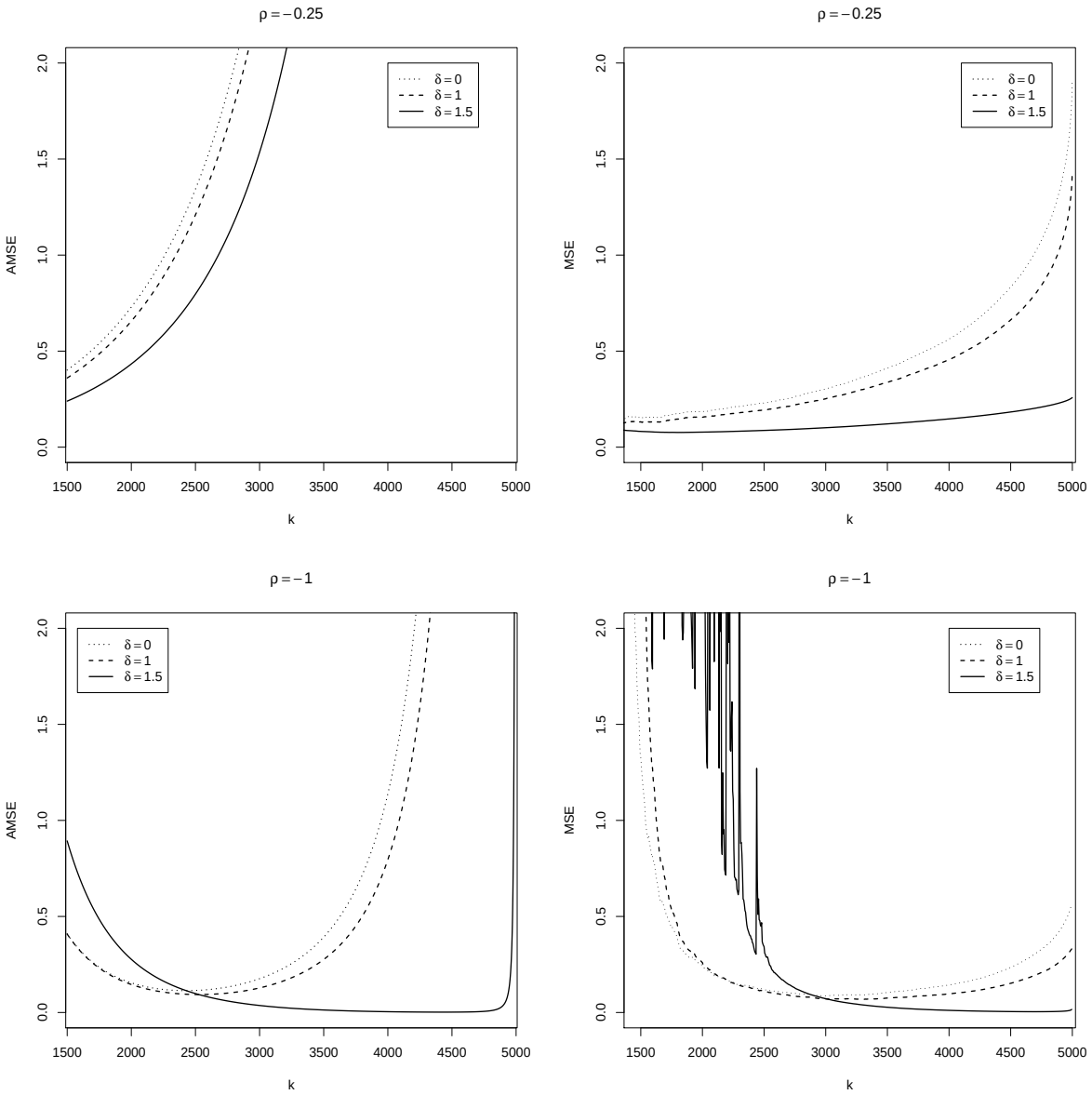

FIG. 4.1 – Erreur quadratique moyenne asymptotique des estimateurs (gauche) et erreur quadratique moyenne empirique (droite) des estimateurs $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$ ($\boldsymbol{\delta} = 1$), $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ ($\boldsymbol{\delta} = 0$) et $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ ($\boldsymbol{\delta} = \boldsymbol{\delta}_0 = 1.5$), en fonction de k_n pour le modèle de Burr, $\gamma = 1$, $\xi = 1$, $\eta = -\rho$, $\rho \in \{-0.25, -1\}$ et $\lambda = 1/\eta$.

tous les estimateurs $\hat{\rho}_n^{(S)}$ tels que $(\boldsymbol{\tau}_1, \boldsymbol{\alpha}_1) = (\boldsymbol{\tau}_7, \boldsymbol{\alpha}_7)$, $(\boldsymbol{\tau}_2, \boldsymbol{\alpha}_2) = (\boldsymbol{\tau}_8, \boldsymbol{\alpha}_8)$, $(\boldsymbol{\tau}_3, \boldsymbol{\alpha}_3) = (\boldsymbol{\tau}_5, \boldsymbol{\alpha}_5)$ et $(\boldsymbol{\tau}_4, \boldsymbol{\alpha}_4) = (\boldsymbol{\tau}_6, \boldsymbol{\alpha}_6)$ sont ainsi équivalents.

4.5.2.2 Comparaison sur une distribution de Burr simulée

Par soucis de simplicité, nous fixons $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_7 = \boldsymbol{\theta}_5 = \boldsymbol{\theta}_7 = \boldsymbol{\tau}_1 = \cdots = \boldsymbol{\tau}_8 = 1$, $\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_3 = \boldsymbol{\alpha}_5 = \boldsymbol{\alpha}_8 = 2$, $\boldsymbol{\alpha}_4 = \boldsymbol{\alpha}_6 = 3$, $\boldsymbol{\theta}_3 = \boldsymbol{\theta}_8 = 1/2$, $\boldsymbol{\theta}_4 = 1/3$, $\boldsymbol{\theta}_6 = 2/3$, $\boldsymbol{\theta}_1 = \boldsymbol{\delta} + 1$ et $\boldsymbol{\theta}_2 = (\boldsymbol{\delta} + 1)/2$ et donc $\boldsymbol{\delta}$ est le seul paramètre libre. L'estimateur résultant est $\hat{\boldsymbol{\rho}}_{n,9}^{(S)}$, il coincïde avec celui proposé dans Alves *et al.* (2003b) quand $\boldsymbol{\delta} = 0$. Pour une valeur de $k \in \{500, \ldots, 4999\}$, l'erreur quadratique moyenne empirique de $\hat{\boldsymbol{\rho}}_{n,9}^{(S)}$ est évaluée sur 500 réplications d'un échantillon de taille $n = 5000$ pour $\boldsymbol{\delta} \in \{0, 1, 2\}$ et pour $\rho \in \{-0.25, -1\}$. Les résultats sont donnés par la Figure 4.2. Il apparaît que $\boldsymbol{\delta} = 0$ donne les meilleurs résultats pour les deux valeurs de ρ utilisées: l'erreur quadratique moyenne empirique est plus faible que celle donnée pour le cas $\boldsymbol{\delta} = 1$ ou $\boldsymbol{\delta} = 2$. Cette hiérarchie ne peut être observée que pour l'erreur quadratique moyenne empirique.

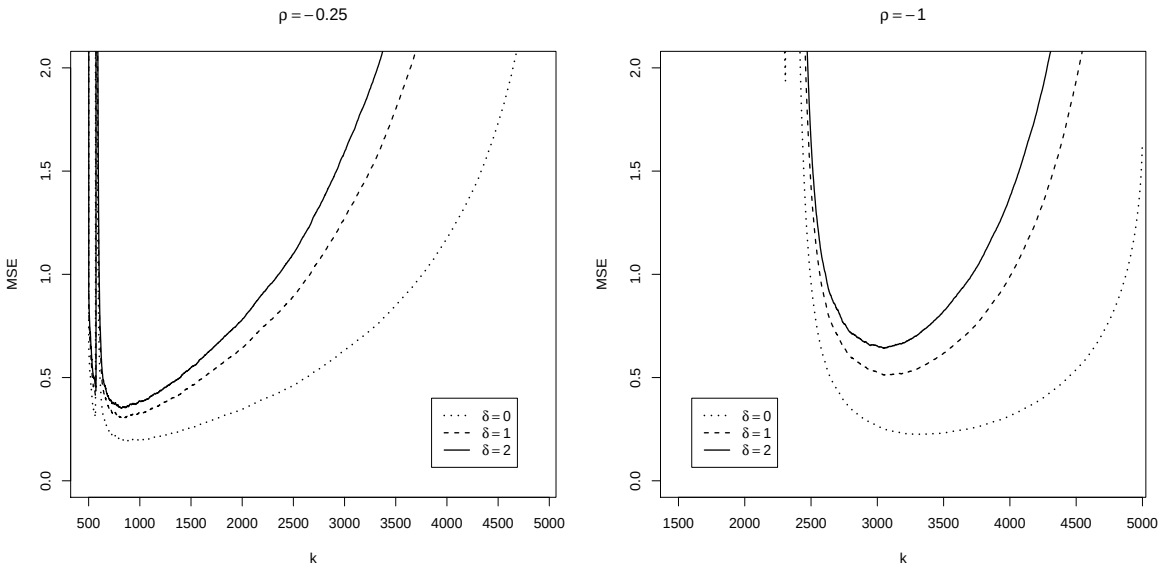

FIG. 4.2 – Erreur quadratique moyenne empirique de $\hat{\boldsymbol{\rho}}_{n,9}^{(S)}$ en fonction de k_n pour une distribution de Burr. Gauche: $\rho = -0.25$, droite: $\rho = -1$.

4.5.3 Tentative de conclusion

Les familles d'estimateurs du paramètre du second ordre dépendent habituellement d'une large classe de paramètres notée $\boldsymbol{\Theta}$ (12 paramètres pour les estimateurs basés sur les

variables aléatoires $R_{k_n}(\boldsymbol{\tau})$ et 20 paramètres pour ceux basés sur $S_{k_n}(\boldsymbol{\tau}, \boldsymbol{\alpha})$). La méthodologie proposée dans le Paragraphe 4.5.1.1 permet d'évaluer une borne supérieure $\pi(\cdot)$ de l'erreur quadratique moyenne asymptotique associée à ces estimateurs. Ceci requiert de montrer que les quantités $\mathcal{AB}_1$, $\mathcal{AB}_2$ et $\mathcal{AV}$ sont bornées inférieurement quand $\boldsymbol{\Theta}$ varie dans une région $R_{\boldsymbol{\Theta}}$. Ainsi, il est possible, pour un bon choix de la région $R_{\boldsymbol{\Theta}}$, de trouver un ensemble « optimal » de paramètres minimisant $\pi(\cdot)$. Malheureusement, l' erreur quadratique moyenne asymptotique peut ne pas dépendre de tous les paramètres dans $\boldsymbol{\Theta}$ (voir Paragraphe 4.5.2.1) où les performances sur échantillon fini de ces estimateurs sont données (voir Paragraphe 4.5.2.2). Dans ce cas, la définition du critère pour la sélection optimale de $\boldsymbol{\Theta}$ est une question ouverte.

4.5.4 Illustrations sur des données d'assurance

Dans cette partie nous souhaitons donner une application de nos estimateurs sur des données d'assurance. Notre illustration se fera à travers l'analyse des montants de sinistres automobiles dépassant 1.200.000 euros sur la période de 1988 à 2001, recueillies auprès de plusieurs sociétés d'assurance européennes et réassurées par la compagnie (« *Secura Belgian Re* »), en Belgique (371 observations). Ce jeu de données est disponible dans Beirlant *et al.* (2004). Nous représentons dans le graphe de gauche de la Figure 4.3, les montants sinistres automobiles par année pour cette compagnie de réassurance. Selon Beirlant *et al.* (2004), section 6.2.2, les données de la compagnie de réassurance « *Secura Belgian Re* » proviennent d'une distribution de type Pareto (distribution à queue lourde). Nous pouvons donc appliquer nos estimateurs sur ces données. Nous représentons dans le graphe de droite de la Figure 4.3, les estimateur $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$, $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ et $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ en fonction de $k_n = 100, ..., 370$. Nous remarquons que pour de grandes valeurs de k_n, l'estimateur $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ correspondant au cas $\boldsymbol{\delta} = \boldsymbol{\delta}_0 = 1.5$ est plus stable que $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$ ($\boldsymbol{\delta} = 1$) et $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ ($\boldsymbol{\delta} = 0$). Nous concluons que l'estimateur $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ a une meilleure performance pour ce jeu de données. Ce qui est bien en conformité avec les conclusions données dans les Sous-sections 4.5.1.1 et 4.5.1.2.

4.6 Démonstrations des résultats théoriques

Demonstration du Théorème 4.2.1. Nous avons clairement que, **(Ψ1)** et **(Ψ2)** entraînent que $Z_n = \psi(\omega_n^{-1}(T_n - \chi_n \mathbb{I})$. De plus, **(T1)** et **(Ψ4)** donnent $Z_n \xrightarrow{\mathbb{P}} \psi(f(\rho)) =$

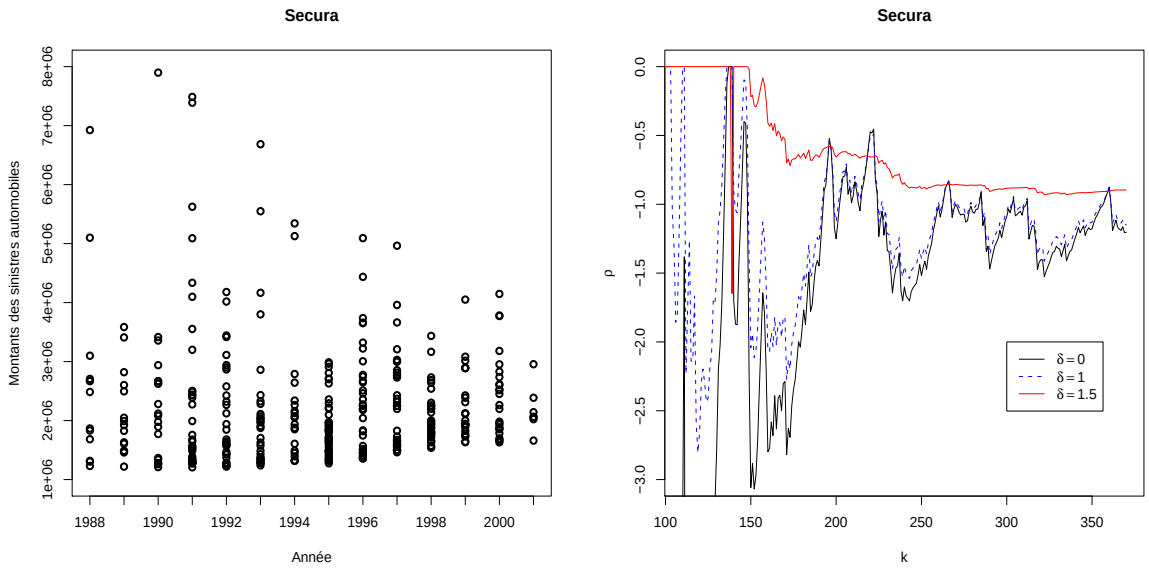

FIG. 4.3 – Graphe de gauche: montants des sinistres automobiles de la compagnie de réassurance *Secura Belgian Re* . Graphe de droite: estimateurs $\hat{\boldsymbol{\rho}}_{n,1}^{(R)}$ ($\boldsymbol{\delta} = 1$), $\hat{\boldsymbol{\rho}}_{n,2}^{(R)}$ ($\boldsymbol{\delta} = 0$) et $\hat{\boldsymbol{\rho}}_{n,0}^{(R)}$ ($\boldsymbol{\delta} = \boldsymbol{\delta}_0 = 1.5$) du paramètre du second ordre ρ en fonction de k_n.

$\varphi(\rho)$. Pour tout $\varepsilon > 0$, nous avons

$$\begin{aligned}
\mathbb{P}(|\hat{\boldsymbol{\rho}}_n - \rho| > \varepsilon) &= \mathbb{P}(\{|\hat{\boldsymbol{\rho}}_n - \rho| > \varepsilon\} \cap \{Z_n \in J\}) + \mathbb{P}(\{|\hat{\boldsymbol{\rho}}_n - \rho| > \varepsilon\} \cap \{Z_n \notin J\}) \\
&\leq \mathbb{P}(\{|\hat{\boldsymbol{\rho}}_n - \rho| > \varepsilon\} \cap \{Z_n \in J\}) + \mathbb{P}(\{Z_n \notin J\}) \\
&= \mathbb{P}(\{|\varphi^{-1}(Z_n) - \rho| > \varepsilon\} \cap \{Z_n \in J\}) + \mathbb{P}(\{Z_n \notin J\}).
\end{aligned}$$

D'après **(Ψ3)** et **(Ψ4)**, φ^{-1} est aussi continue au voisinage de $\varphi(\rho)$. Puisque $Z_n \xrightarrow{\mathbb{P}} \varphi(\rho)$, il s'en suit que

$$\mathbb{P}(\{|\varphi^{-1}(Z_n) - \rho| > \varepsilon\} \cap \{Z_n \in J\}) \to 0 \text{ quand } \ n \to \infty.$$

Comme, $\rho \in J_0$ alors $\varphi(\rho) \in J$ et donc

$$\mathbb{P}(\{Z_n \notin J\}) \to 0 \text{ quand } n \to \infty. \tag{4.12}$$

Par conséquent, $\mathbb{P}(|\hat{\boldsymbol{\rho}}_n - \rho| > \varepsilon) \to 0$ quand $n \to \infty$ et le résultat est prouvé. ■

Demonstration du Théorème 4.2.2. Rappelons que $Z_n = \psi(\omega_n^{-1}(T_n - \chi_n \mathbb{I})$, un développement de Taylor du premier ordre montre qu'il existe $\varepsilon \in (0, 1)$ tel que

$$v_n(Z_n - \varphi(\rho)) = {}^t(v_n \xi_n)\, \nabla\psi(f(\rho) + \varepsilon\xi_n),$$

où nous avons défini $\xi_n = \omega_n^{-1}(T_n - \chi_n \mathbb{I} - f(\rho)$. Donc $\xi_n \xrightarrow{\mathbb{P}} 0$ et **(Ψ5)** entraînent que

$$\nabla\psi(f(\rho) + \varepsilon\xi_n) \xrightarrow{\mathbb{P}} \nabla\psi(f(\rho)).$$

Ainsi, en tenant compte de la condition **(T2)**, on obtient

$$v_n(Z_n - \varphi(\rho)) \xrightarrow{\mathcal{D}} \mathcal{N}(m_\psi(\rho), \sigma_\psi^2(\rho)). \tag{4.13}$$

Maintenant, $P_n(x) := \mathbb{P}(\{v_n(\hat{\boldsymbol{\rho}}_n - \rho) \leq x\})$ peut être réécrite de la manière suivante

$$\begin{aligned} P_n(x) &= \mathbb{P}(\{v_n(\hat{\boldsymbol{\rho}}_n - \rho) \leq x\} \cap \{Z_n \in J\}) + \mathbb{P}(\{v_n(\hat{\boldsymbol{\rho}}_n - \rho) \leq x\} \cap \{Z_n \notin J\}) \\ &= \mathbb{P}(\{v_n(\varphi^{-1}(Z_n) - \rho) \leq x\} \cap \{Z_n \in J\}) + \mathbb{P}(\{v_n(\hat{\boldsymbol{\rho}}_n - \rho) \leq x\} \cap \{Z_n \notin J\}) \\ &=: P_{1,n}(x) + P_{2,n}(x). \end{aligned}$$

Notons d'abord que

$$0 \leq P_{2,n}(x) \leq \mathbb{P}(\{Z_n \notin J\}) \to 0 \text{ quand } n \to \infty, \tag{4.14}$$

d'après (4.12) dans la preuve du Théorèmes 4.2.1. Concentrons maintenant sur $P_{1,n}(x)$, puisque φ est continûment différentiable au voisinage ρ et $\varphi'(\rho) \neq 0$, alors φ est monotone au voisinage de ρ. Considérons le casφ est décroissante, le cas où φ est croissante étant similaire. Soit $J = (a, b)$, alors

$$\begin{aligned} P_{1,n}(x) &= \mathbb{P}(\{a \vee \varphi(\rho + x/v_n) \leq Z_n \leq b\}) \\ &= \mathbb{P}(\{v_n(a \vee \varphi(\rho + x/v_n) - \varphi(\rho)) < v_n(Z_n - \varphi(\rho)) \leq v_n(b - \varphi(\rho))\}). \end{aligned}$$

Introduisons G_n la fonction de répartition de $v_n(Z_n - \varphi(\rho))$, nous avons

$$\begin{aligned} 1 - P_{1,n}(x) &= 1 - G_n(v_n(b - \varphi(\rho))) + G_n(v_n(a \vee \varphi(\rho + x/v_n) - \varphi(\rho))) \\ &= 1 - G_n(v_n(b - \varphi(\rho))) + G_n(v_n(a - \varphi(\rho))) \vee G_n(v_n(\varphi(\rho + x/v_n) - \varphi(\rho))) \\ &=: P_{1,1,n} + P_{1,2,n} \vee P_{1,3,n}(x). \end{aligned}$$

Soit G la fonction de répartition la loi $\mathcal{N}(m_\psi(\rho), \sigma_\psi^2(\rho))$. Il est clair que

$$P_{1,1,n} \leq 1 - G(v_n(b - \varphi(\rho))) + \sup_{t \in \mathbb{R}} |G_n(t) - G(t)|.$$

Puisque $\rho \in J_0$, nous avons $\varphi(\rho) \in J = (a, b)$. En particulier, $b > \varphi(\rho)$ donne $1 - G(v_n(b - \varphi(\rho))) \to 0$ quand $n \to \infty$. En outre, (4.13) montre que $G_n(t) \to G(t)$ pour tout $t \in \mathbb{R}$ et donc $G_n(t) \to G(t)$ uniformément, voir par exemple Embrechts *et al.* (1997), p.552.

Comme conclusion préliminaire $P_{1,1,n} \to 0$ et de manière similaire, $P_{1,2,n} \to 0$ quand $n \to \infty$. Finalement,

$$|P_{1,3,n}(x) - G(x\varphi'(\rho))| \leq |G(v_n(\varphi(\rho + x/v_n) - \varphi(\rho)) - G(x\varphi'(\rho))| + \sup_{t\in\mathbb{R}} |G_n(t) - G(t)|$$

et d'après **(Ψ5)**, $v_n(\varphi(\rho+x/v_n)-\varphi(\rho)) \to x\varphi'(\rho)$ quand $n \to \infty$, ce qui donne $P_{1,3,n}(x) \to G(x\varphi'(\rho))$ quand $n \to \infty$. Nous allons donc montrer que

$$P_{1,n}(x) \to 1 - G(x\varphi'(\rho)) = G(x|\varphi'(\rho)|) \text{ quand } n \to \infty. \tag{4.15}$$

En combinant (4.14) et (4.15) on obtient

$$\mathbb{P}(\{v_n(\hat{\boldsymbol{\rho}}_n - \rho) \leq x\}) \to G(x|\varphi'(\rho)|) \text{ quand } n \to \infty$$

Ce qui complète la preuve du théorème. ■

Demonstration du Corollaire 4.4.1. Nous avons clairement que $\psi_{\boldsymbol{\delta}}$ satisfait **(Ψ1)** et **(Ψ2)**. Le Lemme 4.3.1 montre que **(T2)** est verifiée. Pour appliquer le Théorème 4.2.2, il ne nous reste qu'à prouver que les hypothèses **(Ψ3)** et **(Ψ5)** sont vérifiées. Dans un premier temps nous remarquons que les conditions **(C4)** et **(C5)** sont vérifiées et $\varphi_{\boldsymbol{\delta}}^{(R)}(\rho)$ est bien définie pour tout $\rho \leq 0$ puisque $f^{(R)}(\rho) \in \mathcal{D}$. De plus d'après le Lemme 4.3.1, nous avons pour $i = 1, \ldots, 4$,

$$T_{n,2i-1}^{(R)} - T_{n,2i}^{(R)} = \frac{\tilde{\boldsymbol{\theta}}_i A(Y_{n-k_n,n})}{\boldsymbol{\gamma}}(\mu(H_{\boldsymbol{\tau}_{2i-1}} I_\rho) - \mu(H_{\boldsymbol{\tau}_{2i}} I_\rho))(1 + o_{\mathbb{P}}(1)),$$

quand n tend vers infini. D'où les conditions **(C4)** et **(C5)** impliquent que $T_n^{(R)} \in \mathcal{D}$. Finalement, en utilisant le Théorème de Lerch (voir Carslaw et Jaeger (1948), page 345), la condition **(C4)** entraîne qu'il existe $\rho_0 < 0$ telle que la dérivée première de $\varphi_\rho^{(R)}$ est non nulle au point ρ_0. Alors, le théorème de la fonction inverse assure l'existence d'intervalles J_0 et J pour lesquels la fonction $\varphi_{\boldsymbol{\delta}}^{(R)}$ est une bijection de J_0 à valeurs dans J continûment différentiable. En conclusion, les conditions **(Ψ3)** et **(Ψ5)** sont satisfaites et le résultat est obtenu en appliquant le Théorème 4.2.2. ■

Demonstration du Corollaire 4.4.5. La preuve suit les mêmes lignes que le Corollaire 4.4.1. Elle consiste à remarquer que, sous les conditions **(C6)** et **(C7)**, on obtient $f^{(S)}(\rho) \in \mathcal{D}$ et $T_n^{(S)} \in \mathcal{D}$, puisque

$$T_{n,2i-1}^{(S)} - T_{n,2i}^{(S)} = \frac{\boldsymbol{\zeta}_i A(n/k_n)}{\boldsymbol{\gamma}} \left(\mu(G_{\boldsymbol{\tau}_{2i-1},\boldsymbol{\alpha}_{2i-1}} J_{1-\boldsymbol{\alpha}_{2i-1}} K_\rho) - \mu(G_{\boldsymbol{\tau}_{2i},\boldsymbol{\alpha}_{2i}} J_{1-\boldsymbol{\alpha}_{2i}} K_\rho)\right)(1 + o_{\mathbb{P}}(1)).$$

■

CHAPITRE 5

Estimation des mesures de risque actuariel pour des pertes à queues lourdes

Résumé

Dans ce chapitre, nous nous interéssons à l'estimation des mesures de risques actuariels pour des pertes extrêmes. Nous nous focalisons sur la prime de risque ajustée qui est la mesure de risque la plus étudiée dans la littérature actuarielle. Et nous proposons d'étudier une grande classe d'estimateurs de cette dernière dans le cas des pertes à queues lourdes en utilisant l'approche de la théorie des valeurs extrêmes puis nous établissons leur normalité asymptotique. Nous verrons que cette approche peut conduire à un biais potentiel dans l'estimation. Nous proposons donc une technique de réduction de biais pour ce type d'estimateurs.

5.1 Problématiques et définitions

Les compagnies de réassurance assurent un rôle économique essentiel par la fourniture d'une couverture de risques assumés pour leurs clients. Les risques extrêmes se définissent comme des risques dont on ne sait pas si le système d'assurance sera capable de les supporter, par exemple les ouragans, les incendies, les tremblements de terre, les attentats terroristes etc. et sont généralement responsables de la grande partie des indemnités versées par les compagnies de réassurance. Ainsi, l'attention des actuaires se porte en priorité sur ces risques qualifiés de majeurs et caractérisés par une faible fré-

quence et une extrême gravité sur les dommages qu'ils provoquent. Ceux pour lesquels la survenance d'un sinistre, que l'on peut qualifier d'exceptionnel, aurait une incidence sur le taux de prime pure moyen des sous portefeuilles et mettrait à mal les résultats de la société d'assurance. On pourra se référer à Griselda et Guillaume (2006) et Centeno et Guerra (2010). Un exemple de tel problème a été discuté dans Ceberián *et al.* (2003) sur les montants de revendications d'assurance médicale.

Pour cela, les statistiques traditionnelles notamment le théorème central limite, la loi des grands nombres qui étudient un phénomène à travers son comportement moyen, caractérisent mal ces extrêmes. Elles fournissent uniquement des indications sur la distribution générale d'un phénomène, et donc elles ne donnent que très peu d'informations sur la queue de distribution. Cette méthode peut ne pas fonctionner correctement si le portefeuille contient des risques hétérogènes. C'est ainsi que les actuaires trouvent que le secteur entier de la réassurance est une application très importante de la théorie des valeurs extrêmes, puisqu'on doit toujours sauvegarder une compagnie d'assurance contre les revendication excessives qui peuvent mettre en danger la solvabilité du portefeuille (voir par exemple Ceberián *et al.* (2003)).
McNeil (1997) et Rootzen et Tajvidin (1997) montrent que l'application de la théorie des valeurs extrêmes donne un cadre plus systématique, rigoureux et cohérent que les méthodes classiques. En effet, cette méthode permet de déduire la distribution des extrêmes à partir de la distribution totale, sans traiter séparément les extrêmes des risques récurrents. Ces événements appartiennent à ce que l'on appelle la « queue de distribution », ce qui veut dire que l'on dispose de très peu d'observations pour caractériser les distributions de probabilité. Donc, il est important de bien modéliser les queues de distributions pour éviter les erreurs grossières de tarification.

La modélisation de ces observations par la distribution normale a été longtemps utilisée dans la gestion du risque. Cependant, les études empiriques concluent que les données financières et actuarielles présentent des déviations systématiques de la normalité et leurs distributions ont généralement une queue plus lourde que celle du modèle Gaussien.
Notons qu'une fonction de distrubtion F est dite à queue lourde ou épaisse (en anglais heavy-tailed) si pour tout $\lambda > 0$

$$\lim_{x\to\infty} e^{\lambda x}\overline{F}(x) = \infty.$$

La classe des fonctions de répartitions à variations régulières à l'infini est un bon exemple pour les modèles à queues lourdes. Elle comprend les distributions populaires (comme Pareto, Burr, Student, Lévy-stable et log-gamma) très connues et sont des modèles appro-

priés, adaptés pour les grandes revendication de primes d'assurance, de larges fluctuations de prix, etc.(voir par exemple Beirlant *et al.* (2001)). Ces distributions appartiennent au domaine d'attration de Fréchet, pour plus de détails on pourra se référer au Chapitre 1. Rappelons qu'une caractérisation simple des fonctions de distribution F appartenant à ce domaine est donnée par

$$\lim_{x\to\infty} \frac{\overline{F}(\lambda x)}{\overline{F}(x)} = x^{-1/\gamma} \Leftrightarrow \overline{F}(x) = x^{-1/\gamma}\ell_F(x),\ \forall\lambda > 0 \tag{5.1}$$

où $\gamma > 0$, est l'indice des valeurs extrêmes et ℓ_F est une fonction à variations lentes à l'infini satisfaisant $\ell_F(\lambda x)/\ell_F(x) \to 1$ quand $x \to \infty$, $\lambda > 0$. Dans la suite de ce chapitre, nous serons amenés à utiliser les définitions précédentes.

5.2 Mesures de risque actuariel

Nous donnons dans cette section quelques mesures de risque utilisées dans la littérature actuarielle puis nous établissons quelques méthodes d'estimation de la prime de réassurance.

5.2.1 Définitions et propriétés

Nous reprenons ici la définition d'une mesure de risque telle qu'elle est formalisée dans Denuit et Charpentier (2005).

Définition 5.2.1 *: Soit X une variable aléatoire désignant un montant de perte. Une mesure de risque (Prime) est une fonctionnelle π faisant correspondre à un risque X (perte) une valeur positive $\pi[X]$ éventuellement infinie. Pour autant que π soit normalisée (i.e. $\pi[0] = 0$), $\pi[X]$ est le montant minimum qui, additionné à la perte X en début de période rend la couverture de X « acceptable ». Elle représente donc le capital à détenir pour faire face à une perte financière de montant X . De grandes valeurs de $\pi[X]$ indiqueront que X est « dangereux ».*

En particulier, cette définition nous permet d'établir que l'espérance, la variance ou l'écart-type sont des mesures de risque.
Dans tout ce qui suit, le montant de perte X est modélisé comme une variable aléatoire positive et nous supposerons que la mesure de risque $\pi[X]$ est finie. IL est généralement admis qu'une mesure de risque (ou prime) doit vérifier certaines conditions pour être utile dans les applications. Ceci mène à la notion de mesure de risque *Cohérente* selon la termologie de Artzner *et al.* (1999).

Propriété 5.2.2 *: Cohérence*

Une mesure de risque $\pi[X]$ est dite cohérente lorsqu'elle posséde les propriétés suivantes:

A1. *$\pi[X]$ est inférieure ou égale au montant muximum de perte: $\pi[X] \leq \max(X)$.*

A2. *$\pi[X]$ est supérieure ou égale à la perte moyenne: $\pi[X] \geq \mathbb{E}(X)$.*

A3. *Invariante par translation: $\pi[X+a] = \pi[X] + a$ pour toute constante a.*

A4. *Homogéneité positive: $\pi[bX] = b\pi[X]$ pour toute constante $b > 0$.*

A4. *Sous-additivité: $\pi[X+Y] \leq \pi[X] + \pi[Y]$ quels que soient les risques X et Y.*

Remarque 5.2.3 *La propriété* **A2.** *signifie que les mesures de risque contiennent un chargement de sécurité, c'est-à-dire que le capital minimal doit excéder la perte attendue, sans quoi la ruine devient certaine.*

5.2.2 Mesures de rique de Wang

Il existe plusieurs variantes de mesures de risque (ou principes de calcul de la prime), voir par exemple Buhlmann (1980), Wang (1995) et Wang (1996). Nous nous référons à Goovaerts *et al.* (1984) pour des exemples divers et leurs propriétés. Pour une tarification raisonnable des risques extrêmes,Wang (1996) propose un nouveau procédé de calcul de primes en appliquant une transformation à la fonction de survie. Cette approche de tarification des contrats d'assurance considère les montants de pertes comme des variables aléatoires positives X et engendre des primes qui sont plus hautes que la valeur attendue de la perte. Parmi les plus utilisées de ces mesures de risque, nous avons la *moyenne du risque* ou *prime nette* définie par :

$$\boldsymbol{\pi} := \mathbb{E}(X) = \int_0^\infty \overline{F}(x)dx,$$

où $\overline{F} = 1 - F$ est la fonction de survie de la variable aléatoire positive X. En général, les primes doivent être supérieures ou égales à la prime nette $\boldsymbol{\pi}$ (c'est le critère de chargement de sécurité, condition **A2** de la Propiété 5.2.2), afin d'éviter que l'assureur perde de l'argent en moyenne. Une façon d'atteindre cet objectif consiste à introduire une fonction croissante $g : [0,1] \mapsto [0,1]$ telle que $g(0) = 0$ et $g(1) = 1$, appellée *fonction de distorsion*, et de définir

$$\boldsymbol{\pi}(g, F) := \int_0^\infty g(\overline{F}(x))dx.$$

Cette mesure de risque a été introduite par Wang (1996) et est cohérente si et seulement si la fonction de distorsion correspondante g est concave. En choisissant une fonction de distorsion convenable g, nous regroupons dans le Tableau 5.1 ci-dessous, certaines

mesures de risque très fréquentes dans la littérature actuarielle. Pour une présentation détaillée de ces mesures de risque ainsi que leurs propriétes, on pourra se référer à Denuit et Charpentier (2005). Notons que la Value-at-risk n'est pas *cohérente* car elle n'est pas sous-additive. Les paramètres $\boldsymbol{\varrho}$, α, ς et κ sont appelés *paramètres de distorsion*. Ils contrôlent le montant de la charge des risques inclu dans la prime pour le degré de risque donné par la variable de pertes X. Les cas $\boldsymbol{\varrho} = 1$ et $\alpha = \kappa = 0$ correspondent à la prime nette $\boldsymbol{\pi} = \mathbb{E}(X)$. Parmi ces mesures de risque, la prime ajustée (en anglais Risk-Adjusted-Premium) $PRA_{\boldsymbol{\varrho}}$ est la plus populaire. Dans la suite, nous nous intéressons à cette dernière pour des pertes de très grandes ampleurs.

Mesure de risque	Paramètre		Fonction de distorsion
Value-at-Risk	VaR_α,	$0 \leq \alpha < 1$	$g(x) = \mathbb{1}_{[\alpha,\infty]}(x)$
Conditionnal Tail Expectation	$CTE\alpha$,	$0 \leq \alpha < 1$	$g(x) = \min(x/\alpha, 1)$
La prime ajustée	$PRA_{\boldsymbol{\varrho}}$,	$1 \leq \boldsymbol{\varrho} < \infty$	$g(x) = x^{1/\boldsymbol{\varrho}}$
Mesure de risque Look-back	LB_ς,	$0 < \varsigma \leq 1$	$g(x) = x^\varsigma(1 - \varsigma \ln x)$
Wang Transform	WT_κ,	$0 \leq \kappa < \infty$	$g(x) = \phi(\phi^{-1}(x) + \kappa)$

TAB. 5.1 – Exemples de mesures de risque avec différentes fonctions de distorsion g. $\phi(\cdot)$ est la fonction de répartition d'une loi normale standard.

5.2.3 Principe de la prime ajustée

Nous avons vu dans le Tableau 5.1, que la prime ajustée est la mesure de risque $RAP_{\boldsymbol{\varrho}}$ associée à la fonction de distorsion $g(x) = x^{\frac{1}{\boldsymbol{\varrho}}}$, $\boldsymbol{\varrho} \geq 1$. Pour simplifier notons $\boldsymbol{\pi}(\boldsymbol{\varrho}) = RAP_{\boldsymbol{\varrho}}$, la prime de risque-ajustée est définie comme l'espérance de distorsion (voir par exemple Wang (1995) et Wang (1996)) par:

$$\boldsymbol{\pi}(\boldsymbol{\varrho}) := \int_0^\infty (\overline{F}(x))^{\frac{1}{\boldsymbol{\varrho}}} dx.$$

Le paramètre $\boldsymbol{\varrho} \geq 1$ représente le *coéfficient de distorsion* ou *indice de l'aversion au risque*. Lorsque ce paramètre est à sa valeur minimale $\boldsymbol{\varrho} = 1$, alors $\boldsymbol{\pi}(\boldsymbol{\varrho})$ est la prime nette $\boldsymbol{\pi}$, et donc il n'y a pas de chargement. La charge du risque augmente lorsque $\boldsymbol{\varrho}$ augmente. Cette mesure de risque a été largement étudiée dans la littérature en particulier en Economie sur la théorie duale de l'espérance d'utilité dans l'analyse des comportements de demande d'assurance par l'intermédiaire de Yaari (1987). Pour une stratification élevée

avec un niveau de rétention $R > 0$, la prime ajustée peut être définie comme,

$$\boldsymbol{\pi}_R(\boldsymbol{\varrho}) := \int_R^\infty (\overline{F}(x))^{\frac{1}{\varrho}} dx.$$

Dans le cas des excès élevés de stratification de pertes ($R \to \infty$), Necir et Boukhetala (2004) Vandewalle et Beirlant (2006) et Necir *et al.* (2007) ont introduit et étudié différents estimateurs de $\boldsymbol{\pi}_R(\boldsymbol{\varrho})$ basés sur des échantillons de montant de revendications sur les couvertures de réassurance pour des pertes à queues lourdes. Associé au système d'information statistique de l'assurance automobile, le principe proposé par Necir et Boukhetala (2004) est mis en œuvre sous forme d'un logiciel permettant un calcul automatique d'estimation de primes de réassurance par intervalle de confiance.
Cependant, dans le cas de sans rétention ($R = 0$), la technique d'estimation de $\boldsymbol{\pi}(\boldsymbol{\varrho})$ est différente du cas où R est élevée et a été récemment étudiée dans la littérature par Centeno et Andrade e Silva (2005) en utilisant les techniques de bootstrap, et par Necir et Meraghni (2009) par l'intermédiaire des processus empiriques et de la théorie des valeurs extrêmes. Dans ce chapitre, cette dernière approche est également utilisée, mais contrairement à Necir et Meraghni (2009), nous traitons le problème du biais des estimateurs proposés.

5.2.4 Estimation de la prime ajustée

Nous présentons dans les deux sections suivantes deux approches pour estimer la prime de risque-ajustée $\boldsymbol{\pi}(\boldsymbol{\varrho})$.

5.2.4.1 Approche du quantile empirique

En utilisant la fonction de quantile associée à la distribution F,

$$Q(s) = \inf\{x,\ F(x) \geq s\},\ 0 < s \leq 1,$$

on peut réécrire $\boldsymbol{\pi}(\boldsymbol{\varrho})$ comme

$$\boldsymbol{\pi}(\boldsymbol{\varrho}) = -\int_0^1 s^{\frac{1}{\varrho}} d\left(Q(1-s)\right). \tag{5.2}$$

Soit $X_1, ..., X_n$, un échantillon de variables aléatoires indépendantes de même loi F que X et considérons $X_{1,n} \leq ... \leq X_{n,n}$, la statistique d'ordre associée à cet échantillon. En remplaçant dans (5.2) Q par sa version empirique Q_n définie par $Q_n(s) = X_{\jmath,n}$ pour tout

$(\jmath-1)/n < s \leq \jmath/n$ et $\jmath = 1, ..., n$ avec $Q_n(0) = X_{1,n}$, on obtient par une intégration par parties l'estimateur de la prime $\boldsymbol{\pi}(\boldsymbol{\varrho})$ suivant :

$$\widehat{\pi}_n(\boldsymbol{\varrho}) = \sum_{\jmath=1}^{n} \left(\left(\frac{\jmath}{n}\right)^{1/\boldsymbol{\varrho}} - \left(\frac{\jmath-1}{n}\right)^{1/\boldsymbol{\varrho}} \right) X_{n-\jmath+1,n}. \tag{5.3}$$

L'estimateur $\widehat{\pi}_n(\boldsymbol{\varrho})$ peut être vu comme une combinaison linéaire des éléments de la statistique d'ordre communément appelée *L-statsitique* (voir par exemple Shorack et Wellner (1986)). Les principales propriétés asymptotiques de $\widehat{\boldsymbol{\pi}}_n(\boldsymbol{\varrho})$ ont été établies par Jones et Zitikis (2003): $\widehat{\boldsymbol{\pi}}_n(\boldsymbol{\varrho})$ est consistant pour tout $\boldsymbol{\varrho} \geq 1$ pourvu que $\mathbb{E}|X|^{\eta} < \infty$ pour $\eta > \boldsymbol{\varrho}$ et est asymptotiquement normal si $\boldsymbol{\varrho} \in [1, 2)$ avec la même condition sur le moment X mais cette fois limitée à $\eta > 2\boldsymbol{\varrho}/(2-\boldsymbol{\varrho})$.
Nous allons maintenant traiter un problème d'estimation de la prime ajustée pour la classe des fonctions de répartitions à queues lourdes, i.e nous allons supposer que F vérifie (5.1). Cependant, nous nous concentrons dans ce chapitre sur le cas $\boldsymbol{\gamma} \in \left(\frac{1}{2}, 1\right)$ et $\boldsymbol{\varrho} \in [1, \frac{1}{\boldsymbol{\gamma}})$, afin de s'assurer que la prime de risque-ajustée $\boldsymbol{\pi}(\boldsymbol{\varrho})$ est finie et puisque dans ce cas, les résultats de Jones et Zitikis (2003) ne peuvent plus être appliqués, du fait que le moment d'ordre deux de X est infini. On aura donc recours à une autre approche basée sur la Théorie des Valeurs Extrêmes (TVE).

5.2.4.2 Approche des quantiles extêmes

D'après (5.1), les distributions à queues lourdes dépendent d'un paramètre réel positsi inconnu $\boldsymbol{\gamma}$. L'estimation de ce paramètre a été largement étudiée dans la littérature (on pourra toujours se référer au chapitre 1) et l'estimateur le plus célèbre est celui de Hill (1975), donné dans la sous-section 1.5.1.2 et que nous allons rappeler ci-dessous

$$\widehat{\boldsymbol{\gamma}}_{k_n}^{H} = \frac{1}{k_n} \sum_{\jmath=1}^{k_n} \jmath \left(\log X_{n-\jmath+1,n} - \log X_{n-\jmath,n}\right) \tag{5.4}$$

où k_n est une suite d'entiers telle que $k_n \to \infty$ et $k_n/n \to 0$ quand $n \to \infty$. Plus généralement, Csörgő *et al.* (1985) étend cet estimateur à une classe d'estimateurs à noyau ci-dessous (voir la sous-section 1.5.1.1)

$$\widehat{\boldsymbol{\gamma}}_{k_n}^{K} = \frac{1}{k_n} \sum_{\jmath=1}^{k_n} K\left(\frac{\jmath}{k_n+1}\right) Z_{\jmath,k_n}, \tag{5.5}$$

où K est un noyau tel que $\int K = 1$ et $Z_{\jmath,k_n} = \jmath\left(\log X_{n-\jmath+1,n} - \log X_{n-\jmath,n}\right)$. Notons que l'estimateur de Hill correspond au cas particulier $K := \underline{K} = \mathbb{1}_{(0,1)}$.

Dans cet esprit, nous proposons un type d'estimateur à noyau pour la prime de risque-ajustée. En effet, rappelons qu'à partir de (5.2), $\boldsymbol{\pi}(\boldsymbol{\varrho})$ peut être réécrite comme

$$\begin{aligned}\boldsymbol{\pi}(\boldsymbol{\varrho}) &= \underbrace{\left(\frac{k_n}{n}\right)^{\frac{1}{\varrho}} Q\left(1-\frac{k_n}{n}\right) - \int_0^{k_n/n} s^{\frac{1}{\varrho}} d(Q(1-s))}_{\boldsymbol{\pi}_{k_n}^{(1)}(\boldsymbol{\varrho})} + \underbrace{\frac{1}{\boldsymbol{\varrho}} \int_{k_n/n}^1 s^{\frac{1}{\varrho}-1} Q(1-s)ds}_{\boldsymbol{\pi}_{k_n}^{(2)}(\boldsymbol{\varrho})} \\ &=: \boldsymbol{\pi}_{k_n}^{(1)}(\boldsymbol{\varrho}) + \boldsymbol{\pi}_{k_n}^{(2)}(\boldsymbol{\varrho}).\end{aligned}$$

Ainsi, une estimation de $\boldsymbol{\pi}(\boldsymbol{\varrho})$ est basée sur celle du quantile $Q(1-s)$, $0 < s \leq 1$. Nous proposons donc d'estimer $Q(1-s)$ par:

$$\widehat{Q}_n(1-s) = \begin{cases} \widehat{q}_n(s) & \text{pour } 0 < s \leq k_n/n \\ Q_n(1-s) & \text{pour } k_n/n < s \leq 1, \end{cases}$$

où $Q_n(1-s)$ est l'estimateur empirique du quantile $Q(1-s)$ et $\widehat{q}_n(s)$, $s \to 0$ correspond à l'un des estimateurs de quantiles extrêmes proposés dans la Section 1.7. Comme on s'intéresse à des indices $\boldsymbol{\gamma} > 0$, on peut donc utiliser un estimateur de type Weissman (Weissman, 1978) défini dans (1.42) et basé sur l'estimateur à noyau $\widehat{\boldsymbol{\gamma}}_{k_n}^K$ et que nous allons rappeler ci-dessous

$$\widehat{q}_{k_n}^W(s) = X_{n-k_n,n}\left(\frac{k_n}{ns}\right)^{\widehat{\boldsymbol{\gamma}}_{k_n}^K}.$$

Ainsi, on obtient l'estimateur de $\boldsymbol{\pi}(\boldsymbol{\varrho})$ suivant

$$\begin{aligned}\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho}) &= \frac{(k_n/n)^{\frac{1}{\varrho}}}{1-\boldsymbol{\varrho}\widehat{\boldsymbol{\gamma}}_{k_n}^K} X_{n-k_n,n} + \sum_{\jmath=k_n+1}^{n} \left(\left(\frac{\jmath}{n}\right)^{1/\boldsymbol{\varrho}} - \left(\frac{\jmath-1}{n}\right)^{1/\boldsymbol{\varrho}}\right) X_{n-\jmath+1,n} \qquad (5.6) \\ &=: \widetilde{\pi}_{k_n}^{(K,1)}(\boldsymbol{\varrho}) + \widetilde{\pi}_{k_n}^{(2)}(\boldsymbol{\varrho}).\end{aligned}$$

Notons que pour estimer $\pi_{k_n}^{(2)}(\boldsymbol{\varrho})$ nous avons utilisé la même astuce que pour (5.3), par contre pour $\pi_{k_n}^{(1)}(\boldsymbol{\varrho})$ nous avons utilisé un estimateur de quantiles extrêmes de type Weissman: $\widehat{q}_{k_n}^W(s)$, $s \to 0$.

Remarque 5.2.4 *Notons également que les estimateurs des quantiles extrêmes donnés dans la Section 1.7 sont aussi utilisés pour estimer les mesures de risques telles que la Value-at-Risk, la Conditionnal Tail Expectation (CTE) dans le cas des pertes extrêmes.*

La normalité asymptotique de $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$ est évidement liée à celle de $\widehat{\gamma}_{k_n}^K$. D'habitude en théorie des valeurs extrêmes, pour montrer ce type de résultat, nous avons besoin de la

condition du second ordre (1.35) sur la fonction $\mathbb{U}(x) = Q(1-1/x)$ (pour plus de détails, on pourra se référer à la Section 1.5).

Le reste de ce chapitre est organisé de la manière suivante: la Section 5.3 établit les principaux résultats sur les propriétés asymptotiques de l'estimateur à noyau $\widehat{\gamma}_{k_n}^K$ pour en déduire ceux de $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$. Ces résultats illustrent le fait que cet estimateur produit un biais potentiel dans beaucoup de situations (voir par exemple la Figure 5.2 ci-dessous). Pour résoudre ce problème, une approche de réduction de biais est aussi proposée. L'éfficacité de notre méthode est montrée par une étude de simulation et de données réelles dans la Sections 5.4. Enfin la Section 5.5 est alouée aux démonstrastions de nos principaux résultats.

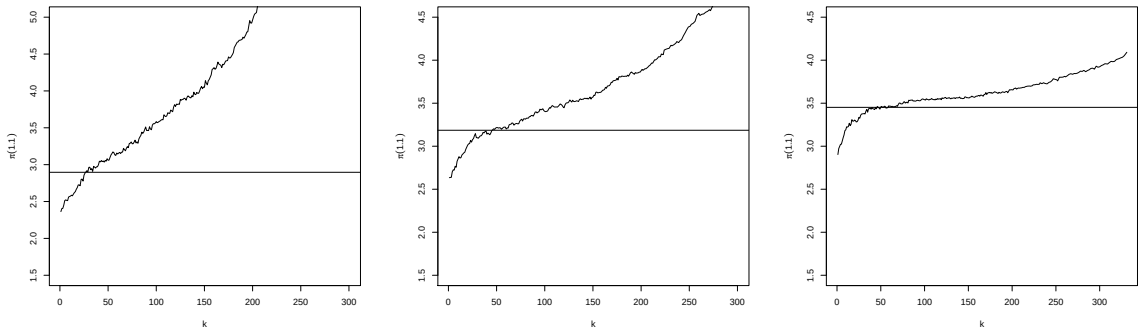

FIG. 5.1 – Médiane de l'estimateur $\widetilde{\pi}_{k_n}^K(1.1)$ en fonction de k_n basées sur 500 échantillons de taille 1000 d'une distribution de Burr définie par $\overline{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$. De gauche à droite $\rho = -0.75$, $\rho = -1$ et $\rho = -1.5$. La ligne horizontale correspond à la vraie valeur de la prime.

5.3 Résultats principaux

Nous commençons par établir quelques résultats sur l'estimateur à noyau $\widehat{\boldsymbol{\gamma}}_{k_n}^K$ dans le Théorème 5.3.1 et ensuite on en déduit la normalité asymptotique de l'estimateur $\widetilde{\boldsymbol{\pi}}_{k_n}^K(\boldsymbol{\varrho})$ de $\boldsymbol{\pi}(\boldsymbol{\varrho})$. Comme ce dernier présente un certain biais, nous proposons donc une technique de réduction de biais.

5.3.1 Resultats asymptotiques de $\widehat{\boldsymbol{\gamma}}_{k_n}^K$

Pour établir la normalité asymptotique de l'estimateur $\widehat{\boldsymbol{\gamma}}_{k_n}^K$, nous donnons quelques hypothèses classiques sur le noyau:

Condition $(\mathcal{K})$. Soit K une fonction définie sur $(0,1]$, telle que

(i) $K(s) \geq 0$, pour tout $0 < s \leq 1$ et $K(1) = 0$;

(ii) $K(\cdot)$ est différentiable, décroissante et continue à droite sur $(0,1]$;

(iii) K etK' sont bornées;

(iv) $\int_0^1 K(u)du = 1$;

(v) $\int_0^1 u^{-1/2}K(u)du < \infty$.

Théorème 5.3.1 *Soit K un noyau satisfaisant $(\mathcal{K})$ et supposons que la distribution F satisfait (1.35). Alors, pour $k_n \to \infty$, $k_n/n \to 0$ et $\sqrt{k_n}A(n/k_n) = O(1)$ quand $n \to \infty$, nous avons*

$$\sqrt{k_n}\left(\widehat{\gamma}_{k_n}^K - \gamma - A(n/k_n)\int_0^1 s^{-\rho}K(s)ds\right) \stackrel{\mathcal{D}}{=} \gamma\sqrt{\frac{n}{k_n}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-\frac{sk_n}{n}\right)d(sK(s)) + o_{\mathbb{P}}(1),$$

où $\{\mathbb{B}_n(s)\ n \geq 1\}$ est une suite de ponts Browniens standards pour $s \in (0,1)$.

Notons ce résultats est similaire à ceux donnés dans Csörgő *et al.* (1985) et Groeneboom *et al.* (2003) mais sous des hypothèses légèrement différentes: par exemple, une représentation de Karamata n'est pas nécessaire pour $Q(\cdot)$ et aucune hypothèse sur la dérivée seconde du noyau.
Une manipulation directe du Théorème 5.3.1 entraîne la normalité asymptotique suivante.

Corollaire 5.3.2 *Sous les hypothèses du Théorème 5.3.1, nous avons*

$$\sqrt{k_n}\left(\widehat{\gamma}_{k_n}^K - \gamma - A(n/k_n)\int_0^1 s^{-\rho}K(s)ds\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \gamma^2\int_0^1 K^2(s)ds\right).$$

Corollaire 5.3.3 *Sous les hypothèses du Théorème 5.3.1, et dans le cas particulier où $K = \underline{K}$, nous avons*

$$\sqrt{k_n}\left(\widehat{\gamma}_{k_n}^{\underline{K}} - \gamma - \frac{A(n/k_n)}{1-\rho}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \gamma^2\right).$$

Comme conséquence du Corollaire 5.3.2, l'erreur quadratique moyenne asymptotique de l'estimateur à noyau $\widehat{\boldsymbol{\gamma}}_{k_n}^K$ est donnée par:

$$\boxed{\mathcal{MSE}(\boldsymbol{\gamma}, A(n/k_n), \rho) = \frac{\boldsymbol{\gamma}^2}{k_n}\int_0^1 K^2(s)ds + A^2\left(\frac{n}{k_n}\right)\left(\int_0^1 s^{-\rho}K(s)ds\right)^2.}$$

Ainsi, connaissant les estimateurs de $\boldsymbol{\gamma}$, $A(n/k_n)$ et ρ on peut obtenir un critère de sélection de k_n comme par exemple dans le cas de la Sous-section 1.6.2:

$$k_n = \arg\min \mathcal{MSE}(\widehat{\boldsymbol{\gamma}}, \widehat{A}(n/k_n), \widehat{\boldsymbol{\rho}}).$$

Voir aussi Beirlant *et al.* (2002), Section 4.

5.3.2 Résultats asymptotique de $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$

Le théorème ci-dessous établit la normalité asymptotique de l'estimateur de $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$.

Théorème 5.3.4 *Supposons que F satisfait (1.35) avec $\boldsymbol{\gamma} \in (1/2, 1)$ et que la fonction de quantile $Q(\cdot)$ associée à F est continûment différentiable sur [0,1). De plus, si la condition $(\mathcal{K})$ est satisfaite et que la suite k_n vérifie $k_n \to \infty$, $k_n/n \to 0$ et $\sqrt{k_n}A(n/k_n) = O(1)$ quand $n \to \infty$, alors pour tout $1 \leq \boldsymbol{\varrho} < 1/\boldsymbol{\gamma}$ nous avons*

$$\frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\varrho}}\mathbb{U}(n/k_n)}\left(\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho}) - \pi(\boldsymbol{\varrho})\right) \stackrel{\mathcal{D}}{=} \sqrt{k}A\left(\frac{n}{k_n}\right)\mathcal{AB}_K(\boldsymbol{\varrho}, \boldsymbol{\gamma}, \rho) + \mathbb{W}_{n,1} + \mathbb{W}_{n,2}(K) + \mathbb{W}_{n,3} + o_{\mathbb{P}}(1)$$

où

$$\mathcal{AB}_K(\boldsymbol{\varrho}, \boldsymbol{\gamma}, \rho) := \frac{\boldsymbol{\varrho}}{1-\boldsymbol{\varrho}\boldsymbol{\gamma}}\left(\frac{1}{\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}\rho-1} + \frac{1}{1-\boldsymbol{\varrho}\boldsymbol{\gamma}}\int_0^1 s^{-\rho}K(s)ds\right)$$

et

$$\begin{cases} \mathbb{W}_{n,1} := -\dfrac{\boldsymbol{\gamma}}{1-\boldsymbol{\varrho}\boldsymbol{\gamma}}\sqrt{\dfrac{n}{k_n}}\,\mathbb{B}_n(1-k_n/n) \\ \mathbb{W}_{n,2}(K) := \dfrac{\boldsymbol{\varrho}\boldsymbol{\gamma}}{(1-\boldsymbol{\varrho}\boldsymbol{\gamma})^2}\sqrt{\dfrac{n}{k_n}}\displaystyle\int_0^1 s^{-1}\mathbb{B}_n(1-sk_n/n)d(sK(s)) \\ \mathbb{W}_{n,3} := -\dfrac{\boldsymbol{\varrho}^{-1}\int_{k_n/n}^1 s^{\frac{1}{\varrho}-1}\mathbb{B}_n(1-s)Q'(1-s)ds}{(k_n/n)^{\frac{1}{\varrho}-1/2}\mathbb{U}(n/k_n)}. \end{cases}$$

Corollaire 5.3.5 *Sous les hypothèses du Théorème 5.3.4, si $\sqrt{k_n}A(n/k_n) \to \lambda \in \mathbb{R}$, nous avons*

$$\frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\varrho}}\mathbb{U}(n/k_n)}\left(\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho}) - \pi(\boldsymbol{\varrho})\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\lambda\mathcal{AB}_K(\boldsymbol{\varrho}, \boldsymbol{\gamma}, \rho), \mathcal{AV}_K(\boldsymbol{\varrho}, \boldsymbol{\gamma})\right).$$

où

$$\mathcal{AV}_K(\boldsymbol{\varrho},\boldsymbol{\gamma}) = \frac{\boldsymbol{\gamma}^2\boldsymbol{\varrho}}{(1-\boldsymbol{\varrho}\boldsymbol{\gamma})^2(2\boldsymbol{\gamma}\boldsymbol{\varrho}+\boldsymbol{\varrho}-2)} + \frac{\boldsymbol{\varrho}^2\boldsymbol{\gamma}^2}{(1-\boldsymbol{\varrho}\boldsymbol{\gamma})^4}\int_0^1 K^2(s)ds.$$

Le Corollaire 5.3.5 généralise le Théorème 2 dans Necir et Meraghni (2009) dans le cas $\lambda \neq 0$ et quand et quand nous utilisons un noyau gén'eral au lieu de $\underline{K}$.
Compte tenu de ces résultats, $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$ est un estimateur de $\pi(\boldsymbol{\varrho})$ avec un biais asymptotique donné par

$$(k_n/n)^{\frac{1}{\varrho}}\,\mathbb{U}(n/k_n)A(n/k_n)\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho).$$

Cette quantité dépend de la distribution originale à travers les fonctions $\mathbb{U}(\cdot)$ et $A(\cdot)$, les paramètres des valeurs extrêmes $\boldsymbol{\gamma}$ et ρ, le paramètre de distorsion $\boldsymbol{\varrho}$, le noyau K et le rapport k_n/n de statistiques d'ordres supérieur. Pour un noyau spécifique, nous pouvons calculer le biais et la variance asymptotique. Par exemple, nous avons le corollaire suivant si $K = \underline{K}$. Pour simplifier les notations, nous allons poser $\boldsymbol{a_0} = (\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}-1)/(1-\boldsymbol{\varrho}\boldsymbol{\gamma})^2$.

Corollaire 5.3.6 *Sous les hypothèses du Corollaire 5.3.5 et dans le cas particulier où $K = \underline{K}$, nous avons*

$$\frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\varrho}}\mathbb{U}(n/k_n)}\left(\widetilde{\pi}_{k_n}^{\underline{K}}(\boldsymbol{\varrho})-\pi(\boldsymbol{\varrho})\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda\boldsymbol{\varrho}\rho\boldsymbol{a_0}}{(1-\boldsymbol{\varrho})(\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}\rho-1)}, \frac{\boldsymbol{\varrho}\boldsymbol{\gamma}^2\boldsymbol{a_0}^2}{(2\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}-2)}\right).$$

Notre objectif dans la prochaine section est de proposer un estimateur sans biais de la prime de risque-ajustée.

5.3.3 Correction de Biais pour $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$

Rappelons qu'à partir du Théorème 5.3.4,

$$\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho}) - (k_n/n)^{\frac{1}{\varrho}}\,\mathbb{U}(n/k_n)A(n/k_n)\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho) \tag{5.7}$$

est un estimateur sans biais pour $\pi(\boldsymbol{\varrho})$. Notons également que $\boldsymbol{\gamma}$, ρ, $\mathbb{U}(n/k_n)$ and $A(n/k_n)$ sont des quantités inconnues que nous devons estimer. Pour un estimateur consistant $\widehat{\boldsymbol{\rho}} = \widehat{\boldsymbol{\rho}}_{\widetilde{k}_n,n}$ avec $\widetilde{k}_n$ donné de telle sorte que $\widetilde{k}_n \to \infty$, $\widetilde{k}_n/n \to 0$ et $\widetilde{k}_n^{1/2}A(n/\widetilde{k}_n) \to \infty$, nous pouvons utiliser les estimateurs sans biais $\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})$ et $\widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})$ respectivement de $\boldsymbol{\gamma}$ et de $A(n/k_n)$ donnés dans la sous-section 1.5.1.2, équation (1.38) et proposés par Beirlant *et al.* (2002) et basés sur l'approche des moindres carrés. On remarque facilement que $\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})$ peut être vu comme un estimateur à noyau $\widehat{\boldsymbol{\gamma}}_{k_n}^{(K_\rho)}$, où pour $0 < u \leq 1$:

$$K_\rho(u) = \frac{1-\rho}{\rho}\underline{K}(u) + \left(1-\frac{1-\rho}{\rho}\right)\underline{K}_\rho(u) \tag{5.8}$$

avec $\underline{K}(u) = 1\!\!1_{\{0<u<1\}}$ et $\underline{K}_\rho(u) = ((1-\rho)/\rho)(u^{-\rho}-1)1\!\!1_{\{0<u<1\}}$. Les noyaux $\underline{K}(u)$ et $\underline{K}_\rho$ satisfont la condition $(\mathcal{K})$ et contrairement à K_ρ, qui ne vérifie pas l'hypothèse (i) dans cette condition. Nous référons à Gomes et Martins (2004) et Gomes *et al.* (2007) pour d'autres techniques de réduction de biais basées sur l'estimation du paramètre du second ordre.

Nous sommes maintenant en mesure d'obtenir un estimateur avec réduction du biais pour la prime de risque-ajustée $\pi(\boldsymbol{\varrho})$ à partir de (5.7) et en utilisant les estimateurs donnés dans (1.38).

$$\begin{aligned}\widehat{\pi}_{k_n}^{K,\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho}) &= \widetilde{\pi}_{k_n}^{K}(\boldsymbol{\varrho}) - \left(\frac{k_n}{n}\right)^{\frac{1}{\boldsymbol{\varrho}}} X_{n-k_n,n}\widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})\mathcal{AB}_K\left(\boldsymbol{\varrho},\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}),\widehat{\boldsymbol{\rho}}\right) \\ &= \left(\frac{k_n}{n}\right)^{\frac{1}{\boldsymbol{\varrho}}} X_{n-k_n,n}\left(\frac{1}{1-\boldsymbol{\varrho}\widehat{\boldsymbol{\gamma}}_{k_n}^{K}} - \widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})\mathcal{AB}_K\left(\boldsymbol{\varrho},\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}),\widehat{\boldsymbol{\rho}}\right)\right) \\ &\quad + \sum_{\jmath=k_n+1}^{n}\left(\left(\frac{\jmath}{n}\right)^{1/\boldsymbol{\varrho}} - \left(\frac{\jmath-1}{n}\right)^{1/\boldsymbol{\varrho}}\right) X_{n-\jmath+1,n}.\end{aligned} \tag{5.9}$$

Notre prochain objectif est d'établir sous des hypothèses convenables la normalité asymptotique de $\widehat{\pi}_{k_n}^{K,\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho})$.

Théorème 5.3.7 *Supposons que F satisfait (1.35) avec $\boldsymbol{\gamma} \in (1/2,1)$ et que la fonction de quantile $Q(\cdot)$ associée à F est continûment différentiable sur [0, 1). Supposons également que $(\mathcal{K})$ est vérifié et que la suite k_n satisfait $k_n \to \infty$, $k_n/n \to 0$ et $\sqrt{k_n}A(n/k_n) = O(1)$ quand $n \to \infty$. Si $\widehat{\boldsymbol{\rho}}$ est un estimateur consistant de ρ, alors pour tout $1 \le \boldsymbol{\varrho} < 1/\boldsymbol{\gamma}$ nous avons*

$$\frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\boldsymbol{\varrho}}}\mathbb{U}(n/k_n)}\left(\widehat{\pi}_{k_n}^{K,\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho}) - \pi(\boldsymbol{\varrho})\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \widetilde{\mathcal{AV}}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho)\right)$$

avec

$$\begin{aligned}\widetilde{\mathcal{AV}}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho) &= \mathcal{AV}_K(\boldsymbol{\varrho},\boldsymbol{\gamma}) + \rho^{-2}\boldsymbol{\gamma}^2(1-2\rho)(1-\rho)^2\mathcal{AB}_K^2(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho) \\ &\quad + \frac{2\boldsymbol{\varrho}\boldsymbol{\gamma}^2(1-\rho)(1-2\rho)}{\rho^2(1-\boldsymbol{\varrho}\boldsymbol{\gamma})^2}\left(1-(1-\rho)\int_0^1 s^{-\rho}K(s)ds\right)\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho).\end{aligned}$$

Corollaire 5.3.8 *Sous les hypothèses du Théorème 5.3.7 et dans le cas particulier où $K = \underline{K}$, nous avons*

$$\frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\boldsymbol{\varrho}}}\mathbb{U}(n/k_n)}\left(\widehat{\pi}_{k_n}^{\underline{K},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho}) - \pi(\boldsymbol{\varrho})\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\boldsymbol{\varrho}\boldsymbol{\gamma}^2\boldsymbol{a_0}^2(\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}-\boldsymbol{\varrho}\rho-1)^2}{(2\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}-2)(\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}\rho-1)^2}\right),$$

avec $\boldsymbol{a_0} = (\boldsymbol{\varrho\gamma} + \boldsymbol{\varrho} - 1)/(1 - \boldsymbol{\varrho\gamma})^2$.

Maintenant, dans le cas particulier où $K = K_\rho$, et comme nous l'avons déjà mentionné, l'estimateur $\widehat{\boldsymbol{\gamma}}_{k_n}^{(K_\rho)}$ coincide avec $\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})$. L'objectif du corollaire ci-dessous est détablir la normalité asymptotique de l'estimateur résultant $\widehat{\pi}_{k_n}^{K_\rho,\widehat{\boldsymbol{\rho}}}$ de la prime de risque $\pi(\boldsymbol{\varrho})$. Notons par $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}$ l'estimateur $\widehat{\pi}_{k_n}^{K_\rho,\widehat{\boldsymbol{\rho}}}$.

Corollaire 5.3.9 *Sous les hypothèses du Théorème 5.3.7 et dans le cas particulier où* $K = K_\rho$*, nous avons*

$$\frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\varrho}}\mathbb{U}(n/k_n)}\left(\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho}) - \pi(\boldsymbol{\varrho})\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \widetilde{\mathcal{AV}}_{K_\rho}(\boldsymbol{\varrho}, \boldsymbol{\gamma}, \rho)\right)$$

avec

$$\begin{aligned}\widetilde{\mathcal{AV}}_{K_\rho}(\boldsymbol{\varrho}, \boldsymbol{\gamma}, \rho) &= \frac{\boldsymbol{\varrho}^2\boldsymbol{\gamma}^2(1-\rho)^2}{\rho^2(1-\boldsymbol{\varrho\gamma})^4} + \frac{\boldsymbol{\varrho\gamma}^2}{(2\boldsymbol{\varrho\gamma} + \boldsymbol{\varrho} - 2)(1-\boldsymbol{\varrho\gamma})^2} \\ &\quad + \frac{\boldsymbol{\varrho}^2\boldsymbol{\gamma}^2(1-2\rho)(1-\rho)\left(\boldsymbol{\varrho\gamma}\rho + 2\boldsymbol{\varrho}\rho + \boldsymbol{\varrho\gamma} - \rho - 1\right)}{\rho^2(1-\boldsymbol{\varrho\gamma})^3(\boldsymbol{\varrho\gamma} + \boldsymbol{\varrho}\rho - 1)^2}.\end{aligned}$$

5.4 Etude de simulations et de données réelles

Dans cette Section, nous souhaitons illustrer la performance de nos estimateurs sur des échantillons de tailles finies et sur des données réelles d'assurance. Notons que dans toute la suite de cette partie, nous utiliserons le choix canonique $\widehat{\boldsymbol{\rho}} = \rho_0 = -1$, tel que cela a été suggéré par Feuerverger et Hall (1999) et Beirlant et al. (1999).

5.4.1 Etude de simulations

Dans cette section, nous comparons l'estimateur biaisé $\widetilde{\pi}_{k_n}^{K}(\boldsymbol{\varrho})$ avec à biais réduit $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho})$ sur simulations. Dans ce but, nous simulons 500 échantillons de tailles 500, 1 000 et 1 500 puis nous évaluons la médiane et l'erreur médiane quadratique ($\mathcal{MSE}$) de ces estimateurs en fonction de k_n à partir d'une distribution de Burr définie par: $\overline{F}(x) = (1 + x^{-\frac{3}{2}\rho})^{1/\rho}$. L'aversion au risque $\boldsymbol{\varrho}$ choisie est 1.1 et nous utilisons différentes valeurs de $\rho \in (-1.5, -1, -0.75)$ pour voir son impact. La médiane et l'erreur médiane quadratique de ces estimateurs sont évaluées sur 500 réplications. Les résultats sont illustrés par les Figures 5.2, 5.3 et 5.4. Il apparaît sur les graphes de haut des Figures 5.2, 5.3 et 5.4 que plus ρ est proche de 0, plus le biais de de l'estimateur $\widetilde{\pi}_{k_n}^{K}(\boldsymbol{\varrho})$ est important. A titre de comparaison, le biais de $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho})$ est beaucoup plus faible et presque indépendant de

ρ. Sans surprise, tous les biais diminuent lorsque n augmente. L'effet de la correction de biais sur l'erreur médiane quadratique est illustré sur les graphes de bas des Figures 5.2, 5.3 et 5.4. Il semble que cela peut conduire à une légère augmentation de l'erreur médiane quadratique lorsque le biais est faible, i.e lorsque ρ est faible. Cependant, l'erreur médiane quadratique de l'estimateur à biais réduit $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho})$ est à peu près constante par rapport à k_n, en particulier lorsque le biais de de $\widetilde{\pi}_{k_n}^{K}(\boldsymbol{\varrho})$ est forte, c'est-à-dire lorsque ρ est proche de 0. Cette propriété rend le choix de k_n plus facile dans la pratique. Des résultats similaires peuvent être obtenus avec d'autres distributions.

5.4.2 Illustration sur des données d'assurance incendie

Dans cette partie nous souhaitons illustrer également l'efficacité de la procédure de réduction de biais sur des montants de revendications de sinistres incendies.

Description des échantillons de données: Nous considérons dans cette partie une base de données provenant d'une société d'assurance où les variables étudiées correspondent aux coûts des sinistres. Le jeu de données étudié concerne un portefeuille d'assurance de sinistres incendies survenus en Norvège de 1972 à 1992. Ces données contiennent le jour de l'apparition des montants revandiqués ($\times$1000 couronnes norvégiennes) et sont disponibles dans Beirlant *et al.* (2004) . Elles ont été un sujet de préoccupation en ce sens que le nombre de demandes a augmenté systématiquement avec un maximum en 1988. Durant cette année, la taille moyenne annuelle des sinistres a atteint un pic, comme ce fut le cas en 1976 (voir le graphe (a) de la Figure 5.5). Ces données ont été utilisées par Reiss et Thomas (1999) et Reiss et Thomas (2001) sur la période allant de 1983 à 1992 pour estimer la prime nette. Dans la suite, nous nous concentrons sur les donnnées de l'année 1976 (soit $n = 207$ observations) sur lesquelles le graphe (b) de la Figure 5.5 montre l'histogramme.

Ajustement de queue de distribution: Nous allons étudier le comportement de queue de distribution pour ces données. Le graphe (c) de la Figure 5.5 montre que les points de l'exponential QQ-plot sont courbés vers le haut et présentent une forme convexe. Cela indique que les distributions pour ces deux séries possèdent une queue plus lourde que la loi exponentielle. Le graphe du Pareto QQ-plot de la Figure 5.5 montre une allure plus ou moins linéaire indiquant un ajustement raisonnable de la distribution de Pareto pour les montants de sinistres incendies norvégiens de l'année 1976. Encore une fois, l'indice de queue $\boldsymbol{\gamma}$ peut être estimé par la mesure de la pente de cette partie linéaire ultime (voir par exemple la Sous-section 1.5.1.1). A partir du graphe (e) de la Figure 5.5, on peut observer

la difficulté pour trouver la partie stable du graphe de Hill $\left\{(k_n, \widehat{\gamma}_{k_n}^H),\ k_n = 1....,n-1\right\}$ en raison du bais de cet estimateur. Nous pouvons donc appliquer notre méthodologie à ce jeu de données en tant que l'indice de valeur extrême (ou au moins son estimateur) est dans l'intervalle (1/2,1) quelle que soit la valeur de k_n.

Résultats: On étudie l'évaluation du montant de la prime détecté par la théorie des valeurs extrêmes qui reflète le mieux la charge de sinistres extrêmes. Sur les graphes de la Figure 5.7,on représente le montant de la prime d'assurance de ces sinistres incendies par les estimateurs $\widetilde{\pi}_{k_n}^K(\boldsymbol{\varrho})$ et $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho})$ en fonction du nombre d'extrêmes k_n. L'indice d'aversion utilisé est $\boldsymbol{\varrho} \in \{1, 1.1\}$. Rappelons que le cas $\boldsymbol{\varrho} = 1$ correspond à la prime nette. On remarque que l'estimateur $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho})$ est plus stable, ce qui justifie bien l'efficacité de la procédure de réduction de biais. On voit également à partir de ces graphes que cette méthode permet de diminuer le montant de la primes de sinistres extrêmes. Elle sera donc d'une très grande utilité pour les compagnies de réassurance, car leur permettant de proposer des contrats plus adaptés à un moindre coût.

5.5 Démonstrations

Soit $Y_1, ..., Y_n$ une suite de variables aléatoires indépendantes et identiquement distribuées selon la distribution de Pareto standard G, définie par $G(y) = 1 - y^{-1}$, $y \geq 1$. Pour tout $n \geq 1$, soit $Y_{1,n} \leq ... \leq Y_{n,n}$ la statistique d'ordre associée à $Y_1, ..., Y_n$. Nous avons clairement $X_{\jmath,n} \stackrel{\mathcal{D}}{=} \mathbb{U}(Y_{\jmath,n})$, $\jmath = 1, ..., n$. Dans le but d'utiliser les résultats de Csörgő *et al.* (1986), un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ est construit portant la suite $\xi_1, \xi_2, ...$ de variables aléatoires indépendantes unsiormément distribuées sur $(0, 1)$ et une suite de ponts Browniens $\mathbb{B}_n(s)$, $0 \leq s \leq 1, n \geq 1$. Le processus de quantile résultant est donné par

$$\beta_n(t) = \sqrt{n}\ (t - \mathbb{V}_n(t))$$

où $\mathbb{V}_n(s),\ 0 < s \leq 1$ est le quantile empirique de ξ_j $j = 1, , 2, ...$ et donné par

$$\mathbb{V}_n(s) = \xi_{j,n}, \frac{j-1}{n} < s \leq \frac{j}{n},\ j = 1, ..., n \text{ et } \mathbb{V}_n(0) = 0.$$

5.5.1 Résultats préliminaires

Pour établir les preuves des résultats principaux, nous avons donc besoin des lemmes ci-dessous :

Lemme 5.5.1 *Soit λ_n, une suite positive de nombres réels telle que $\lambda_n \to 0$ et $n\lambda_n \to \infty$ quand $n \to \infty$. Alors*

$$\int_0^{1/[\lambda_n(n+1)]} [\log(1 - \mathbb{V}_n(1 - s\lambda_n)) - \log(s\lambda_n)]\, d(sK(s)) = o_{\mathbb{P}}\left(\frac{1}{\sqrt{n\lambda_n}}\right).$$

Lemme 5.5.2 *Sous les hypthèses du Théorème 5.3.4, nous avons*

$$\frac{\sqrt{k_n}\left(\widetilde{\pi}_{k_n}^{(K,1)}(\rho) - \pi_{k_n}^{(1)}(\rho)\right)}{(k_n/n)^{1/\rho}\mathbb{U}(n/k_n)} \stackrel{\mathcal{D}}{=} \sqrt{k_n}A(n/k_n)\mathcal{AB}_K(\rho, \boldsymbol{\gamma}, \rho) + \mathbb{W}_{n,1} + \mathbb{W}_{n,2}(K) + o_{\mathbb{P}}(1).$$

Lemme 5.5.3 *Sous les hypthèses du Théorème 5.3.4,*

$$\frac{\sqrt{k_n}}{(k_n/n)^{1/\rho}\mathbb{U}(n/k_n)}\left(\widetilde{\pi}_{k_n}^{(2)}(\rho) - \pi_{k_n}^{(2)}(\rho)\right) \stackrel{\mathcal{D}}{=} \mathbb{W}_{n,3} + o_{\mathbb{P}}(1).$$

Lemme 5.5.4 *Soit $\mathbb{L}(\cdot)$ une fonction intégrable, bornée et positive sur $(0,1)$ et k_n une suite d'entiers telle que $k_n \to \infty$ et $\frac{k_n}{n} \to 0$ quand $n \to \infty$. Posons $\lambda_n = \frac{k_n}{n+1}$. Alors pour tout $\beta > 0$, nous avons*

$$\int_0^1 \left[\left(\frac{\mathbb{V}_n(s\lambda_n)}{\xi_{k_n+1,n}}\right)^{\beta} - s^{\beta}\right]\mathbb{L}(s)ds \xrightarrow{\mathbb{P}} 0, \text{ as } n \to \infty.$$

Pour établir le Théorème 5.3.7, nous avons besoin d'introduire le lemme ci-dessous.

Lemme 5.5.5 *Supposons que la distribution F satisfait (1.35). Si $k_n \to \infty$, $k_n/n \to 0$ et $\sqrt{k_n}A(n/k_n) = O(1)$ quand $n \to \infty$, alors pour tout estimateur consistant $\widehat{\boldsymbol{\rho}}$ de ρ nous avons*

$$\sqrt{k_n}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) - \boldsymbol{\gamma}\right) \stackrel{\mathcal{D}}{=} \boldsymbol{\gamma}\sqrt{\frac{n}{k_n}}\int_0^1 s^{-1}\mathbb{B}\left(1 - \frac{sk_n}{n}\right)d(sK_\rho(s)) + o_{\mathbb{P}}(1),$$

et

$$\sqrt{k_n}\left(\widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) - A(n/k_n)\right) \stackrel{\mathcal{D}}{=} \boldsymbol{\gamma}(1-\rho)\sqrt{\frac{n}{k_n}}\int_0^1 s^{-1}\mathbb{B}\left(1 - \frac{sk_n}{n}\right)d(s(\underline{K}(s) - K_\rho(s)) + o_{\mathbb{P}}(1).$$

Le dernier lemme est une conéquence directe du Théorème de Karamata (voir Proposition 1.5.8 dans Bingham *et al.* (1987)).

Lemme 5.5.6 *Soit ℓ une fonction à variations lentes en 0. Alors*

$$\lim_{s\to 0}\frac{1}{s^{1-\alpha}\ell(s)}\int_s^1 t^{-\alpha}\ell(t)dt = \frac{1}{\alpha - 1} \quad si\ \alpha > 1.$$

5.5.2 Démonstrations des résultats théoriques

Démonstration du Théorème 5.3.1: Nous avons déjà mentionné que la condition $(\mathcal{K})$ est similaire à celle donnée dans Csörgő *et al.* (1985). L'estimateur à noyau $\widehat{\gamma}_{k_n}^K$ est un exemple particulier de leur classe d'estimateurs définie ci-dessous:

$$a_n^{-1} = \sum_{j=1}^{n} \frac{1}{n\lambda_n} \widetilde{K}\left(\frac{j}{n\lambda_n}\right) Z_{j,k_n} \left(\int_0^{\frac{1}{\lambda_n}} \widetilde{K}(u)du\right)^{-1},$$

où $\lambda_n = \frac{k_n+1}{n}$, $\widetilde{K} = K$ sur $(0,1]$ et 0 sinon. Ainsi,

$$\begin{aligned}
\widehat{\gamma}_{k_n}^K &= \sum_{j=1}^{n} \frac{1}{n\lambda_n} \widetilde{K}\left(\frac{j}{n\lambda_n}\right) Z_{j,k_n} + O_{\mathbb{P}}\left(\frac{1}{k_n}\right) \\
&= \int_0^{1/\lambda_n} \log Q_n(1-s\lambda_n) d(s\widetilde{K}(s)) + O_{\mathbb{P}}\left(\frac{1}{k_n}\right) \\
&= \int_0^{1} \log Q_n(1-s\lambda_n) d(sK(s)) + O_{\mathbb{P}}\left(\frac{1}{k_n}\right),
\end{aligned}$$

par le Lemme 2(i) dans Csörgő *et al.* (1985), nous avons

$$\begin{aligned}
\widehat{\gamma}_{k_n}^K &\stackrel{\mathcal{D}}{=} \int_0^1 \left[\log Q_n(1-s\lambda_n) - \log Q(1-\xi_{k_n+1,n})\right] d(sK(s)) + O_{\mathbb{P}}\left(\frac{1}{k_n}\right), \\
&\stackrel{\mathcal{D}}{=} \int_0^1 \left[\log Q(1-\mathbb{V}_n(s\lambda_n)) - \log Q(1-\xi_{k_n+1,n})\right] d(sK(s)) + O_{\mathbb{P}}\left(\frac{1}{k_n}\right).
\end{aligned}$$

Selon de Haan et Ferreira (2006), page 74, la condition du second ordre (1.35) implique que pour pour une fonctionéventuellement différente A_0, avec $A_0(t) \sim A(t)$, $t \to \infty$, et pour un certain $\delta > 0$, il existe un $t_0(\delta)$ tel que pour $t \geq t_0(\delta)$, $x \geq 1$,

$$\left|\frac{\log \mathbb{U}(tx) - \log \mathbb{U}(t) - \gamma \log x}{A_0(t)} - \frac{x^\rho - 1}{\rho}\right| \leq \delta x^{\rho+\delta}. \tag{5.10}$$

Ainsi

$$\begin{aligned}
\widehat{\gamma}_{k_n}^K &= \gamma \int_0^1 \log\left(\frac{\xi_{k_n+1,n}}{\mathbb{V}_n(s\lambda_n)}\right) d(sK(s)) + A_0\left(\frac{1}{\xi_{k_n+1,n}}\right) \int_0^1 \frac{\left(\frac{\xi_{k_n+1,n}}{\mathbb{V}_n(s\lambda_n)}\right)^\rho - 1}{\rho} d(sK(s)) \\
&\quad + o_{\mathbb{P}}\left(A_0\left(\frac{1}{\xi_{k_n+1,n}}\right)\right) \int_0^1 \left(\frac{\xi_{k_n+1,n}}{\mathbb{V}_n(s\lambda_n)}\right)^{\rho+\delta} d(sK(s)) + O_{\mathbb{P}}\left(\frac{1}{k_n}\right) \\
&=: A_{n1} + A_{n2} + A_{n3} + O_{\mathbb{P}}\left(\frac{1}{k_n}\right).
\end{aligned}$$

Nous allons donc étudier séparément les termes ci-dessus.
Terme A_{n1}: par le Lemme 2(i) dans Csörgő *et al.* (1985), nous avons

$$A_{n1} = -\gamma \int_0^{1/\lambda_n} \log \mathbb{V}_n(s\lambda_n) d(s\widetilde{K}(s)),$$

par le Lemme 2(ii) dans Csörgő *et al.* (1985), nous avons

$$\begin{aligned} A_{n1} &\stackrel{\mathcal{D}}{=} \gamma - \gamma \int_0^1 \left[\log \mathbb{V}_n(s\lambda_n) - \log(s\lambda_n)\right] d(sK(s)) \\ &\stackrel{\mathcal{D}}{=} \gamma - \gamma \int_0^1 \left[\log\left(1 - \mathbb{V}_n(1 - s\lambda_n)\right) - \log(s\lambda_n)\right] d(sK(s)) \end{aligned}$$

et par le Lemme 5.5.1

$$\begin{aligned} A_{n1} &\stackrel{\mathcal{D}}{=} \gamma - \gamma \int_{1/[\lambda_n(n+1)]}^1 \left[\log\left(1 - \mathbb{V}_n(1 - s\lambda_n)\right) - \log(s\lambda_n)\right] d(sK(s)) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n\lambda_n}}\right) \\ &\stackrel{\mathcal{D}}{=} \gamma - \gamma \int_{1/[\lambda_n(n+1)]}^{1/\lambda_n} \left[\log\left(1 - \mathbb{V}_n(1 - s\lambda_n)\right) - \log(s\lambda_n) + \frac{\mathbb{B}_n(1 - s\lambda_n)}{s\lambda_n\sqrt{n}}\right] d(s\widetilde{K}(s)) \\ &\quad + \gamma \int_{1/[\lambda_n(n+1)]}^{1/\lambda_n} \frac{\mathbb{B}_n(1 - s\lambda_n)}{s\lambda_n\sqrt{n}} d(s\widetilde{K}(s)) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n\lambda_n}}\right). \end{aligned}$$

En utilisant le Lemme 12 dans Csörgő *et al.* (1985), on obtient

$$A_{n1} \stackrel{\mathcal{D}}{=} \gamma + \gamma \int_{1/[\lambda_n(n+1)]}^{1/\lambda_n} \frac{\mathbb{B}_n(1 - s\lambda_n)}{s\lambda_n\sqrt{n}} d(s\widetilde{K}(s)) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n\lambda_n}}\right),$$

et finalement par le Lemme 8 dans Csörgő *et al.* (1985), on a

$$A_{n1} \stackrel{\mathcal{D}}{=} \gamma + \gamma \int_0^1 \frac{\mathbb{B}_n(1 - s\lambda_n)}{s\lambda_n\sqrt{n}} d(sK(s)) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n\lambda_n}}\right) \tag{5.11}$$

sous les hypothèses sur k_n et $(\mathcal{K})$.
Terme A_{n2}: soit $\mathbb{L}(s) = d(sK(s))/ds = \mathbb{L}^+(s) - \mathbb{L}^-(s)$ où $\mathbb{L}^\pm(s)$ est une fonction positive

bornée. Nous avons

$$\begin{aligned}
A_{n2} &= \frac{1}{\rho}A_0\left(\frac{1}{\xi_{k_n+1,n}}\right)\int_0^1\left(\frac{\mathbb{V}_n(s\lambda_n)}{\xi_{k_n+1,n}}\right)^{-\rho}\mathbb{L}(s)ds \\
&= \frac{1}{\rho}A_0\left(\frac{1}{\xi_{k_n+1,n}}\right)\left\{\int_0^1\left(\frac{\mathbb{V}_n(s\lambda_n)}{\xi_{k_n+1,n}}\right)^{-\rho}\mathbb{L}^+(s)ds-\int_0^1\left(\frac{\mathbb{V}_n(s\lambda_n)}{\xi_{k_n+1,n}}\right)^{-\rho}\mathbb{L}^-(s)ds\right\} \\
&= \frac{1}{\rho}A_0\left(\frac{1}{\xi_{k_n+1,n}}\right)\left(\int_0^1 s^{-\rho}\mathbb{L}^+(s)ds-\int_0^1 s^{-\rho}\mathbb{L}^-(s)ds\right)(1+o_{\mathbb{P}}(1)) \text{ par le Lemme 5.5.4} \\
&= \frac{1}{\rho}A_0\left(\frac{1}{\xi_{k_n+1,n}}\right)\int_0^1 s^{-\rho}\mathbb{L}(s)ds(1+o_{\mathbb{P}}(1)) \\
&= A_0\left(\frac{1}{\xi_{k_n+1,n}}\right)\int_0^1 s^{-\rho}\,K(s)\,ds\,(1+o_{\mathbb{P}}(1)) \quad \text{en utilisant une intégration par parties} \\
&= A_0\left(\frac{n}{k_n}\right)\int_0^1 s^{-\rho}\,K(s)\,ds\,(1+o_{\mathbb{P}}(1)), \qquad (5.12)
\end{aligned}$$

puisque A_0 est à variations régulières à l'infini.

Terme A_{n3}: de manière similaire aux termes précédents, par le Lemme 5.5.4 avec $\beta = -\rho-\delta>0$, et sous nos hypothèses, nous obtenons

$$A_{n3} = o_{\mathbb{P}}\left(A\left(\frac{n}{k_n}\right)\right). \qquad (5.13)$$

Finalement, en combinant (5.11), (5.12) et (5.13), on établit le Thórème 5.3.1. ■

Démonstration du Théorème 5.3.4. En combinant les Lemmes 5.5.2 et 5.5.3, nous obtenons

$$\frac{\sqrt{k_n}\left(\widetilde{\pi}^K_{k_n}(\boldsymbol{\varrho})-\pi(\boldsymbol{\varrho})\right)}{(k_n/n)^{\frac{1}{\varrho}}\mathbb{U}(n/k_n)} \stackrel{\mathcal{D}}{=} \sqrt{k_n}A\left(\frac{n}{k_n}\right)\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho)+\mathbb{W}_{n,1}+\mathbb{W}_{n,2}(K)+\mathbb{W}_{n,3}+o_{\mathbb{P}}(1). \qquad (5.14)$$

Ainsi le Théorème 5.3.4 est établi. ■

Démonstration du corollaire 5.3.5. D'après (5.14), nous allons seulement calculer la variance asymptotique du processus $\mathbb{W}_{n,1}+\mathbb{W}_{n,2}(K)+\mathbb{W}_{n,3}$. Les calculs sont fastidieux, mais très directs. Nous donnons seulement ci-dessous les principaux arguments. Remarquons que

$$\begin{aligned}
\frac{J_\alpha(s)}{s^\alpha Q(1-s)} := \frac{\int_s^1 t^\alpha Q'(1-t)dt}{s^\alpha Q(1-s)} &= 1+\alpha\frac{\int_s^1 t^{\alpha-1}Q(1-t)\,dt}{s^\alpha Q(1-s)} \\
&= 1+\alpha\frac{\int_s^1 t^{\alpha-\boldsymbol{\gamma}-1}\ell(t)\,dt}{s^{\alpha-\boldsymbol{\gamma}}\ell(s)} \longrightarrow \frac{\boldsymbol{\gamma}}{\boldsymbol{\gamma}-\alpha} \quad \text{si } \boldsymbol{\gamma}>\alpha
\end{aligned}$$

d'après le Lemme 5.5.6 et en utilisant le fait que $Q(\cdot)$ est à variation régulière en 0 d'indice $-\boldsymbol{\gamma}$ sous (5.1), c'est-à-dire $Q(1-s) = s^{-\boldsymbol{\gamma}}\ell(s)$ avec ℓ une fonction à variations lentes en 0. Cela nous permet de prouver que

$$\frac{\int_s^1 J^2_{\frac{1}{\boldsymbol{\varrho}}-1}(t)dt}{sJ^2_{\frac{1}{\boldsymbol{\varrho}}-1}(s)} \longrightarrow \frac{\boldsymbol{\varrho}}{2\boldsymbol{\gamma}\boldsymbol{\varrho}-2+\boldsymbol{\varrho}} \text{ quand } s \to 0.$$

D'après la Proposition 1.3.6 dans Bingham *et al.* (1987), nous avons

$$\begin{aligned} \forall \varepsilon > 0 \quad x^{-\varepsilon}\ell(x) &\longrightarrow \infty \text{ quand } x \to 0 \\ \forall \delta > 0 \quad x^{\delta}\ell(x) &\longrightarrow 0 \text{ quand } x \to 0. \end{aligned}$$

Ainsi, en choisissant $\delta = \varepsilon$ et $0 < \varepsilon < \min\left(1-\frac{1}{\boldsymbol{\varrho}}, -\boldsymbol{\gamma}+\frac{1}{\boldsymbol{\varrho}}\right)$, nous avons

$$\begin{aligned} 0 \le s\left(\frac{J_{\frac{1}{\boldsymbol{\varrho}}}(s)}{s^{\frac{1}{\boldsymbol{\varrho}}}Q(1-s)}\right)^2 &= s\left(1+\frac{1}{\boldsymbol{\varrho}}\frac{\int_s^1 t^{\frac{1}{\boldsymbol{\varrho}}-\boldsymbol{\gamma}-1}\ell(t)dt}{s^{\frac{1}{\boldsymbol{\varrho}}-\boldsymbol{\gamma}}\ell(s)}\right)^2 \\ &\le s\left(1+Cs^{\boldsymbol{\gamma}-\frac{1}{\boldsymbol{\varrho}}-\varepsilon}\right)^2 = O\left(s^{1+2[\boldsymbol{\gamma}-\frac{1}{\boldsymbol{\varrho}}-\varepsilon]}\right) = o(1) \end{aligned}$$

sous nos hypothèses, où C est une constante convenable. ∎

Démonstration du Théorème 5.3.7. A partir du Théorème 5.3.4 et (5.9), nous avons

$$\frac{\sqrt{k_n}\left(\widehat{\pi}^{K,\widehat{\boldsymbol{\rho}}}_{k_n}(\boldsymbol{\varrho})-\pi(\boldsymbol{\varrho})\right)}{(k_n/n)^{1/\boldsymbol{\varrho}}\mathbb{U}(n/k_n)} \stackrel{\mathcal{D}}{=} \mathbb{W}_{n,1}+\mathbb{W}_{n,2}(K)+\mathbb{W}_{n,3}+\mathbb{W}_{n,4}(K)+o_{\mathbb{P}}(1), \quad (5.15)$$

où

$$\begin{aligned} \mathbb{W}_{n,4}(K) &:= \sqrt{k_n}\left(A(n/k_n)\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho) - \widehat{A}^{\boldsymbol{LS}}_{k_n}(\widehat{\boldsymbol{\rho}})\mathcal{AB}_K\left(\boldsymbol{\varrho},\widehat{\boldsymbol{\gamma}}^{\boldsymbol{LS}}_{k_n}(\widehat{\boldsymbol{\rho}}),\widehat{\boldsymbol{\rho}}\right)\frac{X_{n-k_n,n}}{\mathbb{U}(n/k_n)}\right) \\ &= -\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho)\sqrt{k_n}\left(\widehat{A}^{\boldsymbol{LS}}_{k_n}(\widehat{\boldsymbol{\rho}})-A(n/k_n)\right) \\ &\quad -\sqrt{k_n}\widehat{A}^{\boldsymbol{LS}}_{k_n}(\widehat{\boldsymbol{\rho}})\left(\mathcal{AB}_K\left(\boldsymbol{\varrho},\widehat{\boldsymbol{\gamma}}^{\boldsymbol{LS}}_{k_n}(\widehat{\boldsymbol{\rho}}),\widehat{\boldsymbol{\rho}}\right)-\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho)\right) \\ &\quad -\sqrt{k_n}\widehat{A}^{\boldsymbol{LS}}_{k_n}(\widehat{\boldsymbol{\rho}})\mathcal{AB}_K\left(\boldsymbol{\varrho},\widehat{\boldsymbol{\gamma}}^{\boldsymbol{LS}}_{k_n}(\widehat{\boldsymbol{\rho}}),\widehat{\boldsymbol{\rho}}\right)\left(\frac{X_{n-k_n,n}}{\mathbb{U}(n/k_n)}-1\right). \end{aligned}$$

Par le Lemme 5.5.5, nous avons

$$\mathbb{W}_{n,4}(K) \stackrel{\mathcal{D}}{=} -\mathcal{AB}_K(\boldsymbol{\varrho},\boldsymbol{\gamma},\rho)\boldsymbol{\gamma}(1-\rho)\sqrt{\frac{n}{k_n}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-\frac{sk_n}{n}\right)d(s(\underline{K}(s)-K_\rho(s))+o_{\mathbb{P}}(1),$$

en utilisant la consistance et l'inégalité $|\frac{e^x-1}{x}-1| \le e^{|x|}-1$ pour tout $x \in \mathbb{R}$. Par ailleurs, un calcul direct mène à la variance asymptotique désirée, ce qui complète la preuve du Théorème 5.3.4. ∎

Démonstration du Corollaire 5.3.9: Rappelons que K_ρ ne satisfait pas la condition $(\mathcal{K})$, mais nous pouvons le réécrire comme (5.8) avec les noyaux $\underline{K}$ et $\underline{K}_\rho$ satisfaisant tous les deux $(\mathcal{K})$. Donc, en suivant les lignes de la démonstration du Théorème 5.3.7 et du Lemme 5.5.5, le Corollaire 5.3.9 est établi. ∎

5.6 Démonstrations des résultats préliminaires

Démonstration du Lemme 5.5.1: Nous allons étudier

$$\begin{aligned} Q_n &:= \int_0^{1/[\lambda_n(n+1)]} [\log(1-\xi_{n,n}) + \log(n+1)]\, d(sK(s)) \\ &\quad - \int_0^{1/[\lambda_n(n+1)]} \log(s\lambda_n(n+1)) d(sK(s)) \\ &=: Q_n^{(1)} + Q_n^{(2)}. \end{aligned}$$

Rappelons que $\lim_{n\to\infty} \mathbb{P}\left(-\log(1-\xi_{n,n}) - \log(n+1) \le x\right) = e^{-e^{-x}}$.
Ainsi, d'après le Lemme 1 dans Csörgő *et al.* (1985), nous avons

$$\sqrt{n\lambda_n} Q_n^{(1)} = \sqrt{n\lambda_n}\, O_{\mathbb{P}}(1) \int_0^{1/[\lambda_n(n+1)]} d(sK(s)) = o_{\mathbb{P}}(1).$$

De manière similaire, par les Lemmes 1 et 2 dans Csörgő *et al.* (1985)

$$\begin{aligned} \sqrt{n\lambda_n} Q_n^{(2)} &= -\sqrt{n\lambda_n} \int_0^{1/[\lambda_n(n+1)]} \log(s\lambda_n(n+1)) d(sK(s)) \\ &= \sqrt{n\lambda_n} \int_0^{1/[\lambda_n(n+1)]} K(s) ds \\ &= o_{\mathbb{P}}(1). \end{aligned}$$

Ce qui achève la preuve du Lemme 5.5.1. ∎

Démonstrations du Lemme 5.5.2: Notons que $\widetilde{\pi}_{k_n}^{(K,1)}(\boldsymbol{\varrho})$ peut être réécrit comme suit

$$\widetilde{\pi}_{k_n}^{(K,1)}(\boldsymbol{\varrho}) = \frac{(k_n/n)^{\frac{1}{\varrho}}}{1-\boldsymbol{\varrho}\widehat{\boldsymbol{\gamma}}_{k_n}^K} \mathbb{U}\left(Y_{n-k_n,n}\right).$$

Ainsi,

$$\frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\varrho}} \mathbb{U}(n/k_n)} \left(\widetilde{\pi}_{k_n}^{(K,1)}(\boldsymbol{\varrho}) - \pi_{k_n}^{(1)}(\boldsymbol{\varrho})\right) = \sum_{j=1}^{4} T_{n,j},$$

où

$$\begin{aligned}
T_{n,1} &= \frac{\sqrt{k_n}}{1-\boldsymbol{\varrho}\widehat{\boldsymbol{\gamma}}_{k_n}^K}\left[\frac{\mathbb{U}\left(Y_{n-k_n,n}\right)}{\mathbb{U}(n/k_n)}-\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\boldsymbol{\gamma}}\right],\\
T_{n,2} &= \frac{\sqrt{k_n}}{1-\boldsymbol{\varrho}\widehat{\boldsymbol{\gamma}}_{k_n}^K}\left[\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\boldsymbol{\gamma}}-1\right],\\
T_{n,3} &= \frac{\rho}{(1-\boldsymbol{\varrho}\widehat{\boldsymbol{\gamma}}_{k_n}^K)(1-\rho\boldsymbol{\gamma})}\sqrt{k_n}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^K-\boldsymbol{\gamma}\right),\\
T_{n,4} &= \frac{\sqrt{k_n}}{(k_n/n)^{\frac{1}{\boldsymbol{\varrho}}}\mathbb{U}(n/k_n)}\left[\frac{(k_n/n)^{\frac{1}{\boldsymbol{\varrho}}}}{1-\boldsymbol{\varrho}\boldsymbol{\gamma}}\mathbb{U}(n/k_n)-\pi_{k_n}^{(1)}(\boldsymbol{\varrho})\right].
\end{aligned}$$

Nous allons étudier séparément chaque terme.

Terme $T_{n,1}$. Selon de Haan et Ferreira (2006), p. 60 et Théorème 2.3.9, p. 48, pour tout $\delta>0$, nous avons

$$\begin{aligned}
\frac{\mathbb{U}\left(Y_{n-k_n,n}\right)}{\mathbb{U}(n/k_n)}-\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\boldsymbol{\gamma}} &= A_0\left(\frac{n}{k_n}\right)\left\{\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\boldsymbol{\gamma}}\frac{\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\rho}-1}{\rho}\right.\\
&\qquad\left.+o_{\mathbb{P}}(1)\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\boldsymbol{\gamma}+\rho\pm\delta}\right\},
\end{aligned}$$

où $A_0(t)\sim A(t)$ quand $t\to\infty$. Ainsi, comme $kY_{n-k_n,n}/n=1+o_{\mathbb{P}}(1)$, nous avons

$$T_{n,1}=o_{\mathbb{P}}(1). \tag{5.16}$$

Terme $T_{n,2}$. En utilisant l'égalité $Y_{n-k_n,n}\overset{\mathcal{D}}{=}(1-\xi_{n-k_n,n})^{-1}$, nous avons

$$\sqrt{k_n}\left[\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\boldsymbol{\gamma}}-1\right]\overset{\mathcal{D}}{=}\sqrt{k_n}\left(\left(\frac{n}{k_n}(1-\xi_{n-k_n,n})\right)^{-\boldsymbol{\gamma}}-1\right),$$

et par un développement de Taylor, nous obtenons

$$\begin{aligned}
\sqrt{k_n}\left[\left(\frac{k_n}{n}Y_{n-k_n,n}\right)^{\boldsymbol{\gamma}}-1\right] &\overset{\mathcal{D}}{=} -\boldsymbol{\gamma}\sqrt{k}\left(\frac{n}{k_n}(1-\xi_{n-k_n,n})-1\right)(1+o_{\mathbb{P}}(1))\\
&= -\boldsymbol{\gamma}\sqrt{\frac{n}{k_n}}\beta_n\left(1-\frac{k_n}{n}\right)(1+o_{\mathbb{P}}(1))\\
&= -\boldsymbol{\gamma}\sqrt{\frac{n}{k_n}}\left(\mathbb{B}_n\left(1-\frac{k_n}{n}\right)+O_{\mathbb{P}}(n^{-\nu})\left(\frac{k_n}{n}\right)^{1/2-\nu}\right)(1+o_{\mathbb{P}}(1)),
\end{aligned}$$

pour $0\le\nu<1/2$, d'après Csörgő *et al.* (1986). Et donc

$$T_{n,2}\overset{\mathcal{D}}{=}-\frac{\boldsymbol{\gamma}}{1-\boldsymbol{\varrho}\boldsymbol{\gamma}}\sqrt{\frac{n}{k_n}}\mathbb{B}_n\left(1-\frac{k_n}{n}\right)(1+o_{\mathbb{P}}(1))=\mathbb{W}_{n,1}+o_{\mathbb{P}}(1). \tag{5.17}$$

Terme $T_{n,3}$. D'après le Théorème 5.3.1, nous avons

$$\begin{aligned} T_{n,3} &\overset{\mathcal{D}}{=} \frac{\boldsymbol{\varrho}}{(1-\boldsymbol{\varrho}\boldsymbol{\gamma})^2}\left\{\sqrt{k_n}A\left(\frac{n}{k_n}\right)\int_0^1 s^{-\rho}K(s)ds+\boldsymbol{\gamma}\sqrt{\frac{n}{k_n}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-\frac{k_n}{n}\right)d(sK(s))\right\}+o_{\mathbb{P}}(1) \\ &= \frac{\rho}{(1-\boldsymbol{\varrho}\boldsymbol{\gamma})^2}\sqrt{k_n}A\left(\frac{n}{k_n}\right)\int_0^1 s^{-\rho}K(s)ds+\mathbb{W}_{n,2}(K)+o_{\mathbb{P}}(1). \end{aligned} \tag{5.18}$$

Terme $T_{n,4}$. Une intégration par parties entraîne que

$$\begin{aligned} T_{n,4} &= \sqrt{k_n}\left\{\frac{1}{1-\boldsymbol{\varrho}\boldsymbol{\gamma}}-\frac{1}{\boldsymbol{\varrho}}\int_1^\infty x^{-1-\frac{1}{\boldsymbol{\varrho}}}\frac{\mathbb{U}(nx/k_n)}{\mathbb{U}(n/k_n)}dx\right\} \\ &= -\frac{1}{\boldsymbol{\varrho}}\sqrt{k_n}\int_1^\infty x^{-1-\frac{1}{\boldsymbol{\varrho}}}\left(\frac{\mathbb{U}(nx/k_n)}{\mathbb{U}(n/k_n)}-x^{\boldsymbol{\gamma}}\right)dx. \end{aligned}$$

D'après le Théorème 2.3.9 dans de Haan et Ferreira (2006), nous avons pour $\boldsymbol{\gamma}\in(1/2,1)$ et $1\leq\boldsymbol{\varrho}<1/\boldsymbol{\gamma}$,

$$\begin{aligned} T_{n,4} &= -\frac{1}{\boldsymbol{\varrho}}\sqrt{k_n}A_0\left(\frac{n}{k_n}\right)\int_1^\infty x^{\boldsymbol{\gamma}-1-\frac{1}{\boldsymbol{\varrho}}}\frac{x^\rho-1}{\rho}dx\,(1+o(1)) \\ &= \sqrt{k_n}A\left(\frac{n}{k_n}\right)\frac{\boldsymbol{\varrho}}{(1-\boldsymbol{\varrho}\boldsymbol{\gamma})(\boldsymbol{\varrho}\boldsymbol{\gamma}+\boldsymbol{\varrho}\rho-1)}(1+o_{\mathbb{P}}(1)). \end{aligned} \tag{5.19}$$

En combinant (5.16)-(5.19), le Lemme 5.5.2 est établi. ■

Démonstrations du Lemme 5.5.3. La preuve du Lemme 5.5.3 découle de celle du Théorème 2 dans Necir et Meraghni (2009). ■

Démonstrations du Lemme 5.5.4. Le Lemme 5.5.4 est une conséquence directe du Lemme 3.2 dans Groeneboom *et al.* (2003). ■

Démonstrations du Lemme 5.5.5. Notons

$$\begin{aligned} \sqrt{k}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})-\boldsymbol{\gamma}\right) &= \sqrt{k}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})-\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\rho)\right)+\sqrt{k}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\rho)-\boldsymbol{\gamma}\right) \\ &= \sqrt{k_n}\frac{1}{k_n}\sum_{\jmath=1}^{k_n}\left\{K_{\widehat{\boldsymbol{\rho}}}^*\left(\frac{\jmath}{k_n+1}\right)-K_\rho^*\left(\frac{\jmath}{k_n+1}\right)\right\}Z_{\jmath,k_n}+\sqrt{k}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\rho)-\boldsymbol{\gamma}\right) \end{aligned}$$

où

$$K_\rho^*(x)=\frac{(1-\rho)^2}{\rho^2}-\frac{(1-\rho)(1-2\rho)}{\rho^2}x^{-\rho}.$$

Ainsi, à partir de la preuve du Théorème 3.2 de Beirlant *et al.* (2002), nous avons

$$\sqrt{k}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})-\boldsymbol{\gamma}\right) = \sqrt{k}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\rho)-\boldsymbol{\gamma}\right)+o_{\mathbb{P}}(1).$$

Rappelons maintenant que $\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\rho)$ peut être vu comme l'estimateur à noyau $\widehat{\boldsymbol{\gamma}}_{k_n}^{(K_\rho)}$. Ainsi,

$$\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\rho) = \frac{1-\rho}{\rho}\widehat{\boldsymbol{\gamma}}_{k_n}^{K} + \left(1 - \frac{1-\rho}{\rho}\right)\widehat{\boldsymbol{\gamma}}_{k_n}^{K_\rho},$$

d'après le Théorème 5.3.1, nous avons

$$\begin{aligned}
\sqrt{k_n}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{K} - \boldsymbol{\gamma} - \frac{A(n/k_n)}{1-\rho}\right) &\overset{\mathcal{D}}{=} \boldsymbol{\gamma}\sqrt{\frac{n}{k_n}}\int_0^1 s^{-1}\mathbb{B}\left(1-\frac{sk_n}{n}\right)d(s\underline{K}(s)) + o_{\mathbb{P}}(1),\\
\sqrt{k_n}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{K_\rho} - \boldsymbol{\gamma} - \frac{A(n/k_n)}{1-2\rho}\right) &\overset{\mathcal{D}}{=} \boldsymbol{\gamma}\sqrt{\frac{n}{k_n}}\int_0^1 s^{-1}\mathbb{B}\left(1-\frac{sk_n}{n}\right)d(s\underline{K}_\rho(s)) + o_{\mathbb{P}}(1).
\end{aligned}$$

En combinant les deux dernières égalités, la première partie du Lemme 5.5.5 est démontrée.
Et pour la seconde partie, en utilisant (1.38), nous avons

$$\widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) = (1-\widehat{\boldsymbol{\rho}})\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{K} - \widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}})\right).$$

Nous avons,

$$\begin{aligned}
\sqrt{k_n}\left(\widehat{A}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) - A(n/k_n)\right) &= (1-\widehat{\boldsymbol{\rho}})\sqrt{k_n}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{K} - \boldsymbol{\gamma} - \frac{A(n/k_n)}{1-\rho}\right)\\
&\quad -(1-\widehat{\boldsymbol{\rho}})\sqrt{k_n}\left(\widehat{\boldsymbol{\gamma}}_{k_n}^{\boldsymbol{LS}}(\widehat{\boldsymbol{\rho}}) - \boldsymbol{\gamma}\right)\sqrt{k}A\left(\frac{n}{k_n}\right)\left(\frac{1-\widehat{\boldsymbol{\rho}}}{1-\rho} - 1\right).
\end{aligned}$$

Ce qui aboutira au résultat attendu. ■

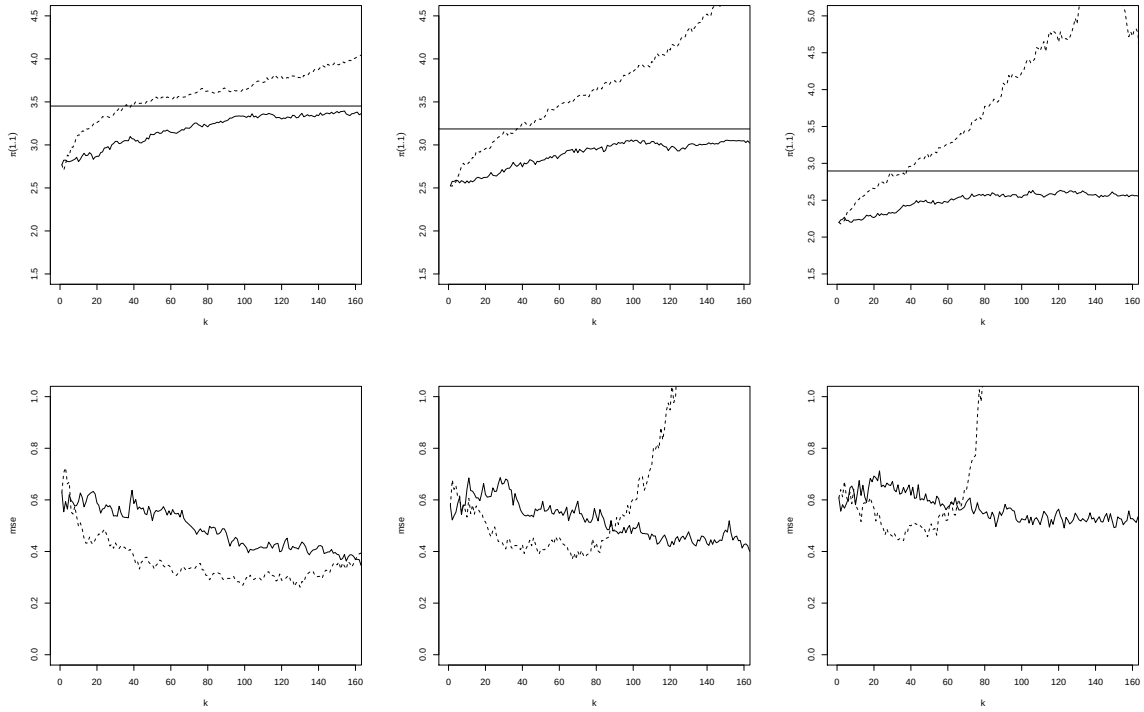

FIG. 5.2 – Médianes (haut) eterreurs médianes quadratiques empiriques (bas) des estimateurs $\widetilde{\pi}_{k_n}^{K}(1.1)$ (ligne tiret) et $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(1.1)$ (ligne pleine) en fonction de k_n basées sur 500 échantillons de taille 500 d'une distribution de Burr définie par $\overline{F}(x) = (1+x^{-\frac{3\rho}{2}})^{1/\rho}$. Gauche: $\rho = -1.5$, milieu $\rho = -1$, et droite: $\rho = -0.75$. La ligne horizontale correspond à la vraie valeur de la prime.

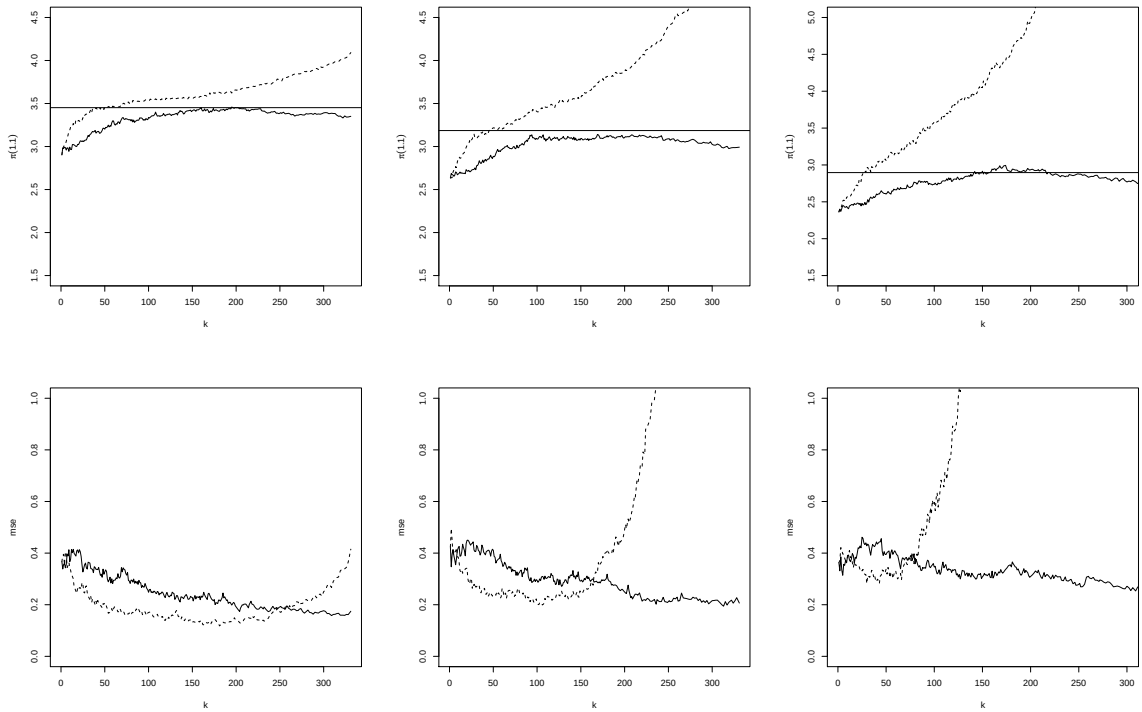

FIG. 5.3 – Médianes (haut) eterreurs médianes quadratiques empiriques (bas) des estimateurs $\widetilde{\pi}^{K}_{k_n}(1.1)$ (ligne tiret) et $\widehat{\pi}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}_{k_n}(1.1)$ (ligne pleine) en fonction de k_n basées sur 500 échantillons de taille 1000 d'une distribution de Burr définie par $\overline{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$. Gauche: $\rho = -1.5$, milieu $\rho = -1$, et droite: $\rho = -0.75$. La ligne horizontale correspond à la vraie valeur de la prime.

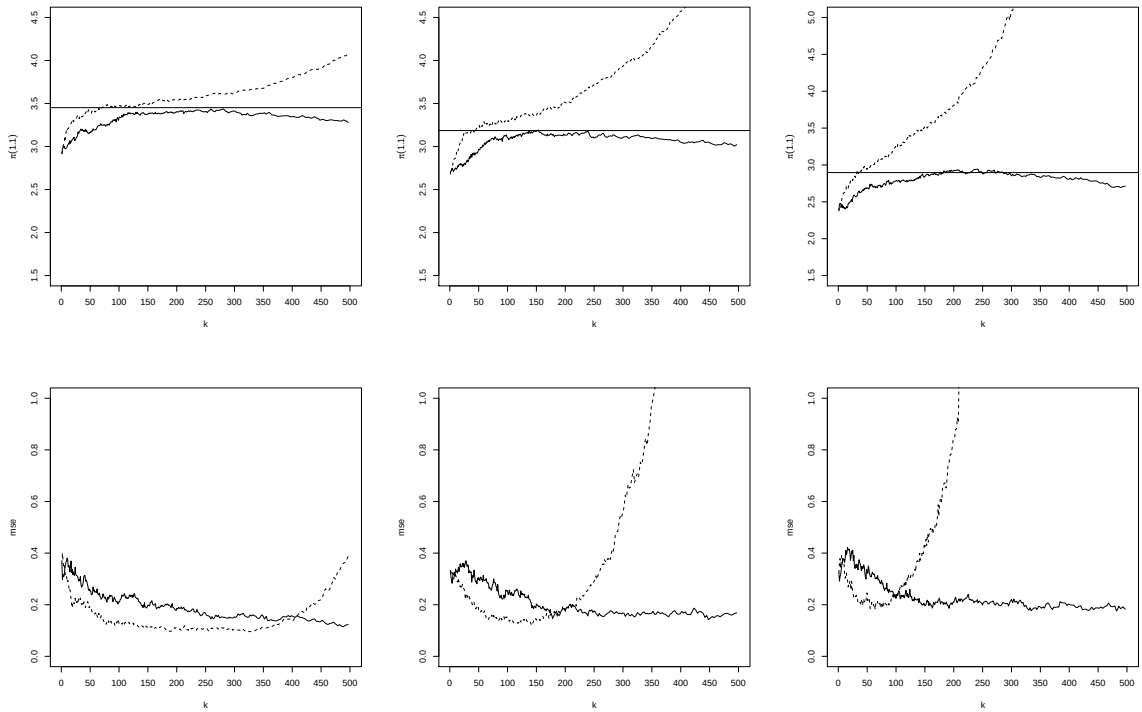

FIG. 5.4 – Médianes (haut) eterreurs médianes quadratiques empiriques (bas) des estimateurs $\widetilde{\pi}^{K}_{k_n}(1.1)$ (ligne tiret) et $\widehat{\pi}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}_{k_n}(1.1)$ (ligne pleine) en fonction de k_n basées sur 500 échantillons de taille 1500 d'une distribution de Burr définie par $\overline{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$. Gauche: $\rho = -1.5$, milieu $\rho = -1$, et droite: $\rho = -0.75$. La ligne horizontale correspond à la vraie valeur de la prime.

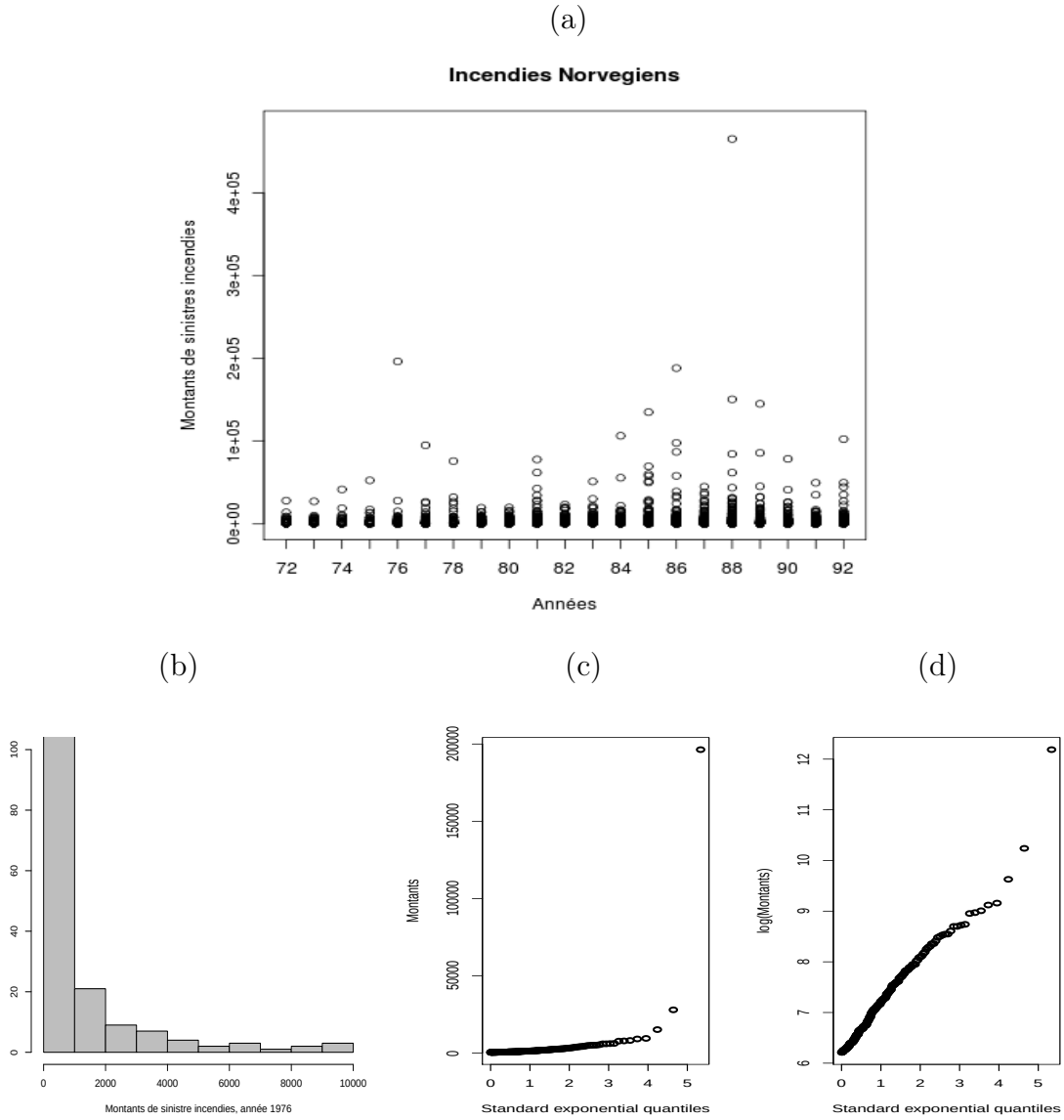

FIG. 5.5 – Montants de sinistres incendies norvégiens sur la période allant de 1972 à 1992 (graphe (a)) et son histogramme pour l'année 1976 (graphe (b)). Les graphes (c) et (d) correspondent respectivement à l'exponential QQ-plot et au Pareto QQ-plot pour les données de l'année 1976.

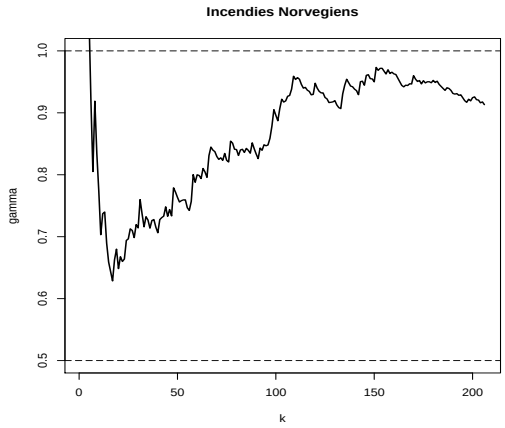

FIG. 5.6 – Estimateurs de Hill en fonction de k_n pour les montants de sinistres incendies Norvégiens en 1976.

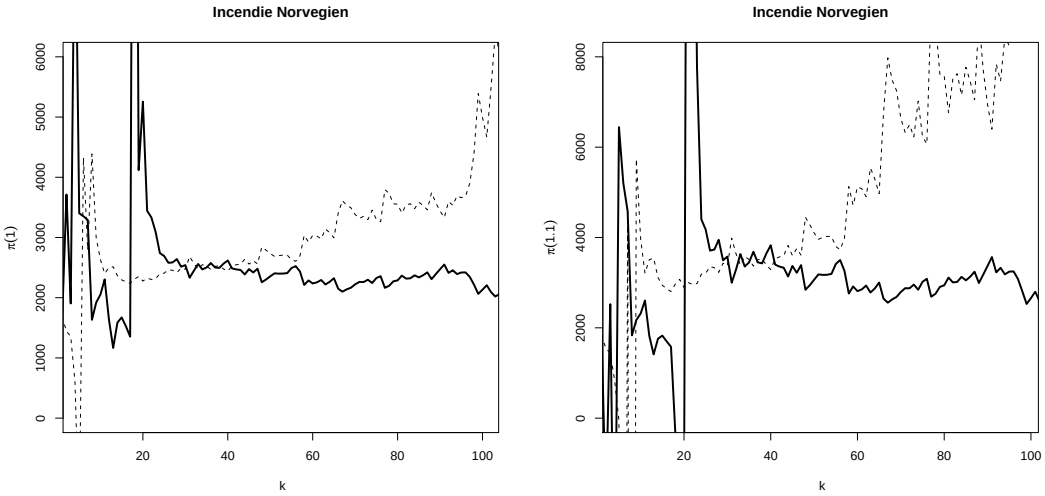

FIG. 5.7 – Estimateurs $\widetilde{\pi}_{k_n}^{K}(\boldsymbol{\varrho})$ (ligne tiret) et $\widehat{\pi}_{k_n}^{\boldsymbol{LS},\widehat{\boldsymbol{\rho}}}(\boldsymbol{\varrho})$ (ligne pleine) en fonction de k_n pour les montants de sinistres incendies Norvégiens en 1976. L'aversion de risque utilisée $\boldsymbol{\varrho} \in \{1, 1.1\}$.

Conclusions et perspectives

Dans cette thèse nous avons abordé différents aspects de la théorie univariée des valeurs extrêmes dans le cas indépendant.

Après avoir rappelé certaines des notions clés en théorie des valeurs extrêmes dans le premier chapitre, nous nous sommes intéressés dans le deuxième chapitre à une méthode d'estimation basée sur la généralisation de l'estimateur classique de Hill. Nous avons considéré un processus statistique dépendant d'un paramètre continu $\boldsymbol{\tau} > 0$ dont chaque marge peut être considérée comme un estimateur généralisé de Hill. Dans ce chapitre, nous avons étudié les propriétés asymptotiques de ce type d'estimateur pour $0 < \boldsymbol{\tau} \leq 1/2$ pour les distributions appartenant aux domaines d'atraction de Fréchet et de Gumbel. Le domaine de Weibull, présente une approche radicalement différente car comportant une somme de variables aléatoires dépendantes, et mérite une étude séparée. En outre, un travail futur consistera à développer une approche bayésienne en considérant un choix aléatoire du paramètre $\boldsymbol{\tau}$ menant à un choix optimal parmi une large classe de distributions adéquates. Il serait également intéressant d'étudier les propriétés asymptotiques de l'estimateur généralisé de Hill dans le cas de processus fortement dépendants.

Dans le troisième chapitre, nous nous sommes intéressés à un processus indexé par une classe de fonctions et permettant d'estimer l'indice des valeurs extrêmes. Nous avons développé une grande classe d'estimateurs généralisant l'estimateur de Hill. Par une théorie moderne de convergence vague, nous avons établi leur loi asymptotique en distribution finie, ainsi que leur loi asymptotique fonctionnelle. Ce travail qui est seulement basé sur une caractérisation des domaines d'attraction de Gumbel et de Fréchet est une généralisation des résultats du deuxieme chapitre précédent. Toutefois, une investigation dans le cas du domaine d'attraction de Weibull est possible. Cela requiert la théorie des sommes de variables aléatoires dépendantes tandis que nous avons utilisé dans ce travail des résultats sur des sommes de variables aléatoires indépendantes. Nous considérons par conséquent une manipulation spéciale de cette affaire dans un travail futur.

Nous avons mentionné dans les chapitres précédents que le paramètre d'importance en théorie des valeurs extrêmes est l'indce de queue. Il contrôle le comportement de la queue de distribution du premier ordre et de nombreuses méthodes d'estimation de ce paramètre ont été données dans la littérature. Les propriétés asymptotiqiues de certains de ces estimateurs sont basées sur une condition du second ordre faisant

intervenir un paramètre du second ordre supposé inconnu. Ce dernier est de première importance dans le choix adaptatif du nombre optimal de statistiques d'ordre supérieur utilisé lors de l'estimation de l'indice de queue et utilisé pour la réduction du biais de ces estimateurs. Dans le quatrième chapitre, nous avons développé une famille d'estimateurs semi-paramétrique du paramètre du second permettant de regrouper un grand nombre d'estimateurs au sein d'un formalisme commun. Nous avons étudié leur propriété asymptotique. Par notre formalisme, nous avons proposé de nouveaux estimateurs asymptotiquement Gaussiens du paramètre du second ordre. Cette famille d'estimateurs du paramètre du second ordre dépend habituellement d'une large classe de paramètres notée $\boldsymbol{\Theta}$. Pour certains cas nous avons proposé un critère de choix optimal de ces paramères, basé sur la minimisation d'une borne supérieure de l'erreur quadratique moyenne asymptotique associée à ces estimateurs quand $\boldsymbol{\Theta}$ varie dans une région $R_{\boldsymbol{\Theta}}$. Nous avons montré qu'il est possible, pour un bon choix de la région $R_{\boldsymbol{\Theta}}$, de trouver un ensemble « optimal » de paramètres minimisant cette borne. Malheureusement, l' erreur quadratique moyenne asymptotique peut ne pas dépendre de tous les paramètres dans $\boldsymbol{\Theta}$. Dans ce cas, la définition du critère pour la sélection optimale de $\boldsymbol{\Theta}$ est une question ouverte.
Nous envisageons dans un futur travail, de proposer une méthode d'estimation du paramètre du second en présence de covariables, afin de pouvoir établir une réduction de biais de certains d'indices de queues où de quantiles extrêmes dépendant de ces mêmes covariables.

Dans le dernier chapitre, nous nous sommes intéressés à l'estimation des mesures de risque actuariel pour des pertes extrêmes, et plus précisément à l'estimation de la prime de risque ajustée dans le cas des pertes à queues lourdes. Nous avons proposé dans ce chapitre une grande classe d'estimateurs de la prime de risque basée sur l'approche des quantiles extrêmes, en étudiant leur normalité asymptotique. Comme cette méthde induit à un biais potentiel dans l'estimation, nous avons proposé une technique de réduction de biais pour ce type d'estimateurs. L'intérêt de cette réduction de biais est de diminuer le montant des primes d'assurance; ce qui permettra aux assureurs de proposer des contrats plus adaptés à un moindre coût. Elle fournit des jumelles à travers lesquelles les assureurs peuvent observer les événements extrêmes avec une certaine objectivité afin de contrôler l'adéquation des primes pures avec la sinistralité et la hiérarchisation des classes de risque d'une part et de constituer des provisions pour faire face à ces risques extrêmes d'autre part. Nous pouvons donc conclure que la théorie des valeurs extrêmes peut figurer parmi les outils d'analyse pour les assureurs.

Nous envisageons dans un futur travail, détendre cette méthode pour les autres mesures de risque actuariel. Notons que la méthode utilisée dans ce travail est basée sur des variables aléatoires indépendantes de pertes et il serait intéréssant dans un travail futur de regarder l'influence de la dépendance sur l'estimation d'un quantile extrême et plus particulièrement l'influence de la dépendance sur l'estimation de la prime de risque. Cette étude se fera par rapport au biais à la variance asymptotique.

ANNEXE A

Convergence de mesures

Nous rappelons dans cette partie quelques éléments essentiels sur la théorie moderne de convergence de mesures. Nous introduisons les notions de mesures en l'absence de la mesurabilté utilisées dans les résultats du Chapitre 3. Pour plus de détails on poura consulter van der Vaart et Wellner (1996).

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, $T : \Omega \to \overline{\mathbb{R}}$, une application non nécessairement mesurable et B une partie de Ω.

Intégrale extérieure: On appelle intégrale extérieure de T, le réel noté $\mathbb{E}^*(T)$, et défini par:

$$\mathbb{E}^*(T) = \inf \Big\{ \mathbb{E}(U) : \ U \geq T, \ U : \Omega \to \overline{\mathbb{R}} \Big\},$$

avec U mesurable et $\mathbb{E}(U)$ existe.

Probabilité extérieure: On appelle Probabilité extérieure de B, le réel noté $\mathbb{P}^*(B)$, et défini par:

$$\mathbb{P}^*(B) = \inf \Big\{ \mathbb{P}(A) : \ B \subset A, \ A \in \mathcal{A} \Big\}.$$

Nous avons $\mathbb{P}^*(B) = \mathbb{E}^*(\mathbb{1}_B)$.

Couverture minimale mesurable: Soit $T : \Omega \to \overline{\mathbb{R}}$, une application. On appelle couverture minimale mesurable de T l'application mesurable $T^* : \Omega \to \overline{\mathbb{R}}$, telle que $T^* \geq T$, et pour toute aapplication mesurable $U : \Omega \to \overline{\mathbb{R}}$, vérifiant $U \geq T$, on a $U \geq T^*$. Si T^* existe, alors nous avons $\mathbb{E}^*(T) = \mathbb{E}(T^*)$.
Notons que T est mesurable si et seulement si $T^* = T$.

Convergence vague: La convergence vague est à la base des tests asymptotiques en statistique. Elle concerne la distribution limite d'une suite d'applications X_n définie

sur $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$, des espaces probabilisés, à valeurs dans un même espace probabilisable $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$. Ici, $\mathbb{D}$ est un espace métrique, et $\mathcal{B}(\mathbb{D})$ est la σ-algébre boréliènne associée.

• La suite d'applications X_n: $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n) \to (\mathbb{D}, \mathcal{B}(\mathbb{D}))$ converge vaguement vers X : $(\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{D}, \mathcal{B}(\mathbb{D}))$ mesurable lorsque $n \to \infty$ (notée $X_n \overset{\mathcal{D}}{\rightsquigarrow} X$), ssi pour toute fonction $f : \mathbb{D} \to \overline{\mathbb{R}}$ continue et bornée (notée $f \in C_b(\mathbb{D})$),

$$\mathbb{E}^*(f(X_n)) \to \mathbb{E}(X), n \to \infty. \tag{A.1}$$

• Notons L la loi de X définie par: $\forall B \in \mathcal{B}(\mathbb{D}),\ L(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(X \in B)$.
Nous avons:

$$\mathbb{E}(f(X)) = \int_{\mathbb{D}} f(x)dL(x) = \int f(X(\omega))d\mathbb{P}(\omega).$$

On peut donc remplacer (A.1) par $\mathbb{E}^*(f(X_n)) \to \displaystyle\int_{\mathbb{D}} f dL$.

Continuité: Si la suite d'aplications X_n: $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n) \to (\mathbb{D}, \mathcal{B}(\mathbb{D}))$ converge vaguement vers $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{D}, \mathcal{B}(\mathbb{D}))$, et si g est une application continue de : $\mathbb{D} \to \mathbb{S}$ (un autre espace métrique), alors $g(X_n) \overset{\mathcal{D}}{\rightsquigarrow} X$.
Espace des fonctions bornées: Soit $\mathbb{S}$. Un espace arbitraire. L'espace des fonctions bornées noté $\ell^\infty(\mathbb{S})$ est défini par l'ensemble de toutes les fonctions réelles uniformément bornées sur $\mathbb{S}$, c'est à dire les fonctions $z : \mathbb{S} \to \mathbb{R}$ telle que

$$||z||_{\mathbb{S}} = \sup_{t \in \mathbb{S}} |z(t)| < \infty.$$

C'est un espace métrique muni de la distance uniforme $d(z_1, z_2) = ||z_1 - z_2||_{\mathbb{S}}$.

Convergence en distribution finie: Soit X_n une suite d'applications définie dans $\ell^\infty(\mathbb{S})$. On dit que les distributions finies de X_n convergent, si pour toute partie $(t_1, ..., t_S) \in \mathbb{R}^S$, $S \in \mathbb{N}^*$, les vecteurs $(X_n(t_1), ..., X_n(t_S))$ convergent vaguement.

Si de plus les distributions finies convergent vers celles d'une même variable aléatoire, c'est à dire: il existe une variable aléatoire X mesurable dans $\ell^\infty(\mathbb{S})$ telle que $(X_n(t_1), ..., X_n(t_S)) \xrightarrow{\mathcal{D}} (X(t_1), ..., X(t_S))$, pour toute partie $(t_1, ..., t_S) \in \mathbb{R}^S$, alors on dit que X_n convergent en distribution finie vers X.

Si X_n convergent vaguement vers X, alors X_n converge en distribution finie.

ANNEXE B Démonstration du Théorème 3.2.1

Cette demonstration est une généralisation de celle du Théorème 2.2.2. Nous utiliserons donc le même procédé dans la suite. Rappelons la représentation de la Sous-section 2.1.2. Nous avons, pour tout $n \geq 1$,

$$\{\log X_{n-j+1,n} = Y_{n-j+1,n}, 1 \leq j \leq n\} =_d \left\{G^{-1}\{1 - U_{j,n}), 1 \leq j \leq n\right\}.$$

Soit $F \in D(\Lambda)$. Par la représentation (2.5), nous avons

$$\begin{aligned} T_n(f) &= \sum_{j=1}^{k} f(j)(s(U_{j,n}) - s(U_{j+1,n})) + \sum_{j=1}^{k} f(j) \int_{U_{j,n}}^{U_{j+1,n}} s(t)/t \, dt \\ &\equiv S_n(1) + S_n(2). \end{aligned}$$

En utilisant (2.6), nous avons pour $U_{1,n} \leq v, u \leq U_{k,n}$,

$$s(u)/s(v) = (1 + p(u))/(1 + p(v)) \exp\left(-\int_{U_{1,n}}^{U_{k+1,n}} t^{-1}\Delta(t)dt\right).$$

Posons

$$g_{1,n,0}(p) = \sup\{|p(u)|, 0 \leq u \leq U_{k+1,n}\} \text{ et } g_{2,n,0}(\Delta) = \sup\{|\Delta(u)|, 0 \leq u \leq U_{k+1,n}\}, \tag{B.1}$$

Puisque $\log(U_{k+1,n}/U_{1,n}) = O_p(\log k)$ quand $n \to \infty$, on obtient,

$$s(u)/s(v) = (1 + O(g_{1,n,0})) \exp(-O_p(g_{2,n,0} \log k)).$$

Ce qui implique que

$$\sup_{U_{1,n} \leq u,v \leq U_{k,n}} |s(u)/s(v) - 1| = O_p(\max(g_{1,n,0}, g_{2,n,0} \log k)) \tag{B.2}$$

quand $n \to \infty$ et

$$\sup_{U_{1,n} \leq u,v \leq U_{k,n}} \left| \frac{s(u) - s(v)}{s(k/n)} \right| = O_p(\max(g_{1,n,0}, g_{2,n,0} \log k)). \tag{B.3}$$

Puisque $nk^{-1}U_{k+1,n} \xrightarrow{p.s} 1$ quand $n \to \infty$, on peut trouver pour tout $\varepsilon > 0$ et pour tout $\lambda > 1$, un entier N_0 tel que pour tout $n \geq N_0$,

$$\mathbb{P}(g_{1,n,0} \leq g_{1,n}(p, \lambda),\ g_{2,n,0} \leq g_{2,n}(\Delta, \lambda)) \geq 1 - \varepsilon. \tag{B.4}$$

D'où (**C3**) implique que

$$\frac{S_n(1)}{\sigma_{k,n}(f)s(k/n)} \leq d_n(p, \Delta, \lambda)\ (\sigma_{k,n}(f))^{-1} \sum_{j=1}^{k} f(j) \xrightarrow{\mathbb{P}} 0,$$

où $d_n(p, \Delta, \lambda) = \max(g_{1,n}, g_{2,n} \log k)$. Ensuite

$$\begin{aligned}
\frac{S_n(2)}{\sigma_{k,n}(f)s(k/n)} &= \sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) \int_{U_{j,n}}^{U_{j+1,n}} \{s(t)/s(k/n)\}/t\ dt \\
&= \sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) \int_{U_{j,n}}^{U_{j+1,n}} t^{-1}\ dt \\
&\quad + \sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) \int_{U_{j,n}}^{U_{j+1,n}} \{s(t)/s(k/n) - 1\}/t\ dt \\
&= S_n(2,1) + S_n(2,2).
\end{aligned}$$

D'après (B.2) et la représenation de Malmquist (2.2), on obtient

$$\begin{aligned}
|S_n(2,2)| &\leq O_p(1) d_n(p, \Delta, \lambda) \times \sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1} E_j \\
&\leq O_p(1)\ \times \Bigg\{ d_n(p, \Delta, \lambda) \times \sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1} (E_j - 1) \\
&\qquad + d_n(p, \Delta, \lambda) \sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1} \Bigg\}.
\end{aligned}$$

Le premier terme tends vers zero car $\sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1} (E_j - 1)$ converge en distribution vers une variable aléatoire finie par le Lemme 3.5.1 et $d_n(p, \Delta, \lambda) \xrightarrow{\mathbb{P}} 0$ d'après

(B.2). Le second terme tend vers zero d'après (**C3**). Finallement, par la représentation de Malmquist (2.2), on obtient

$$S_n(2,1) = \sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1} E_j.$$

Et cela conduit à

$$S_n(2,1) - \{a_{k,n}(f)/\sigma_{k,n}(f)\} = \sigma_{k,n}(\boldsymbol{\tau})^{-1} \sum_{j=1}^{k} j^{\boldsymbol{\tau}-1}(E_j - 1),$$

qui converge en distribution vers une loi normale $\mathcal{N}(0,1)$ sous la condition (**K2**) et vers $\mathcal{L}_\infty(f)$ sous (**K1**) et par le Lemme 3.5.1. En combinant tous ces éléments, on montre que

$$(\sigma_{k,n}(f)s(k/n))^{-1}(T_n(f) - a_{k,n}(f)s(k/n))$$

converge en distribution vers une variable aléatoire normale $\mathcal{N}(0,1)$ sous (**K2**) et vers $\mathcal{L}_\infty(f)$ sous (**K1**).
Maintenant pour $F \in D(\phi_{1/\alpha})$, nous avons d'après (2.4) et la représentation usuelle que,

$$\begin{aligned} T_n(f) &= \sum_{j=1}^{k} f(j) \left\{\log(1 + p(U_{j+1,n})) - \log(1 + p(U_{j,n}))\right\} \\ &\quad + \boldsymbol{\gamma} \sum_{j=1}^{k} f(j) \log(U_{j+1,n}/U_{j,n}) + \sum_{j=1}^{k} f(j) \int_{U_{j,n}}^{U_{j+1,n}} \Delta(t)/t \, dt \\ &\equiv S_n(1) + S_n(2) + S_n(3). \end{aligned}$$

On obtient, pour k assez grand

$$|S_n(1)/\sigma_{k,n}(f)| \leq 2 g_{1,n,0}(f)(\sigma_{k,n}(f))^{-1} \sum_{j=1}^{k} f(j),$$

où $g_{1,n,0}$ est défini dans (B.1), qui tend vers zero en probabilité d'aprés (**C1**) et (B.2). Ensuite

$$\begin{aligned} |S_n(3)/\sigma_{k,n}(\boldsymbol{\tau})| &\leq g_{2,n,0}(b)(\sigma_{k,n}(f))^{-1} \sum_{j=1}^{k} f(j) \log(U_{j+1,n}/U_{j,n}) \\ &= g_{2,n,0}(b)\sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1}(E_j - 1) + g_{2,n,0}(b)\sigma_{k,n}(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1}, \end{aligned}$$

où $g_{2,n,0}$ est défini dans (B.1). Alors $S_n(3)/\sigma_n(f) \to 0$ d'après (**C3**), le Lemme 3.5.1 et les méthodes décrites ci-dessus. Finalement, en utilisant à nouveau le Lemme 3.5.1, on obtient

$$\{(S_n(3) - \boldsymbol{\gamma} a_{k,n}(f)\} / \sigma_{k,n}(f) = \boldsymbol{\gamma} \sigma_n(f)^{-1} \sum_{j=1}^{k} f(j) j^{-1} (E_j - 1)$$

et ceci converge en distribution vers une variable aléatoire $\mathcal{N}(0, \boldsymbol{\gamma}^2)$ sous la condition (**K2**) et vers $\boldsymbol{\gamma}\mathcal{L}_\infty(f)$ sous (**K1**). ■

ANNEXE C Démonstration du Corollaire 3.4.2

Nous appliquons ici le Théorème 3.3.2. Fixons d'abord $f \in \mathcal{F}_1(a,b)$. Soit $f(j) = j^{\boldsymbol{\tau}}$. Nous avons

$$\begin{aligned} B(n,f) &= O((\log k)^{-1}), \text{ pour } \quad \boldsymbol{\tau} = 1/2, \\ B(n,f) &= O(k^{-(2\boldsymbol{\tau}-1)}), \text{ pour } \quad 0 < \boldsymbol{\tau} < 1/2, \\ B(n,f) &= O(k^{-\boldsymbol{\tau}}), \qquad \text{pour } \quad \boldsymbol{\tau} > 1. \end{aligned}$$

Alors (**K1**) est vérifié et ceci conduit à la normalité dans le cas du Théoréme 3.2.1 pour $\boldsymbol{\tau} \geq 1/2$. Ensuite, pour $f_1(j) = j^{\boldsymbol{\tau}_1}$ et $f_2(j) = j^{\boldsymbol{\tau}_2}$, $\boldsymbol{\tau}_1 > 1/2$, $\boldsymbol{\tau}_2 > 1/2$,

$$\lim_{n\to\infty} \sum_{j=1}^{k(n)} \frac{f_1(j)}{j\sigma_{k,n}(f_1)} \times \frac{f_2(j)}{j\sigma_{k,n}(f_2)} = \frac{\sqrt{(2\boldsymbol{\tau}_1 - 1)(2\boldsymbol{\tau}_2 - 1)}}{\boldsymbol{\tau}_1 + \boldsymbol{\tau}_2 - 1} = \Gamma(f_1, f_2) < \infty.$$

Alors pour $\boldsymbol{\tau}_1 = 1/2$ et $\boldsymbol{\tau}_2 > 1/2$

$$\sum_{j=1}^{k(n)} \frac{f_1(j)}{j\sigma_{k,n}(f_1)} \times \frac{f_2(j)}{j\sigma_{k,n}(f_2)} \sim \sqrt{2\boldsymbol{\tau}_2 - 1}(\log k) k^{(2\boldsymbol{\tau}_2 - 1)/2},$$

et donc, $\mathcal{F}_1(a,b)$ satisfait (**KU2a**). Ce qui implique aussi que nous avons la normalité en distribution finie (au sens vague) pour $\boldsymbol{\tau} > 1/2$. Et également l'hypothèse (**KU2b**) est vérifiée puisque

$$B(n) = \sup_{f \in \mathcal{F}_1(a,b)} B(n,f) \leq k^{-b} \to 0.$$

Considérons ici

$$V_n^{**}(f) = \frac{1}{\sigma_n(f)}(V_n(f) - a_{k,n}(f)) = \sum_{j=1}^{k} \frac{f(j)}{j\sigma_n(f)}(E_j - 1) =: \sum_{j=1}^{k} Z_{j,n}(f)$$

Nous avons

$$\begin{aligned}
\rho_n^2(f_1,f_2) &= \mathbb{E}\sum_{j=1}^{k}(Z_{j,n}(f_1)-Z_{j,n}(f_2))^2 \\
&= 2-\frac{2}{\sigma_{k,n}(f1)\sigma_{k,n}(f_2)}\sum_{j=1}^{k} f_1(j)f_2(j)j^{-2} \\
&\to 2(1-\Gamma(f_1,f_2)) = \sum_{j=1}^{\infty}\left(\frac{f_1(j)}{j\sigma_{k,n}(f_1)}-\frac{f_2(j)}{j\sigma_{k,n}(f_2)}\right)^2.
\end{aligned}$$

D'après les calculs de routine et par un argument de continuité basé sur la remarque que les paramètres de $\boldsymbol{\tau}$ sont dans le compact $[a,b]$, nous pouvons montrer que

$$\rho_n^2(f_1,f_2) < 3(1-\Gamma(f_1,f_2)) \tag{C.1}$$

pour de grandes valeurs de n. Puisque nous avons utiliser plusieurs fois ce même type d'argument, nous donnerons dans la partie convergence continue ci dessous, la preuve exacte de (C.1). Mais

$$1-\Gamma(f_1,f_2) = (1-\frac{\sqrt{(2\boldsymbol{\tau}_1-1)(2\boldsymbol{\tau}_2-1)}}{\boldsymbol{\tau}_1+\boldsymbol{\tau}_2-1})$$

Pour $\boldsymbol{\tau}_1-\boldsymbol{\tau}_2=\delta>0$,

$$1-\Gamma(f_1,f_2) = \frac{(2\boldsymbol{\tau}_2-1)+\delta-\sqrt{(2\boldsymbol{\tau}_2-1)^2+2\delta(2\boldsymbol{\tau}_2-1)}}{\boldsymbol{\tau}_1+\boldsymbol{\tau}_2-1}.$$

Nous utilisons un développement de Taylor au premier ordre, afin d'obtenir pour $1/2 < a < \boldsymbol{\tau}_1,\boldsymbol{\tau}_2 < b$, $3(1-\Gamma(f_1,f_2)) \le B(a,b)\delta$, où $B(a,b)$ dépend seulement de a et b. Ainsi, pour δ, et k assez grand,

$$\rho_n^2(f_1,f_2) \le B(a,b)\delta.$$

Nous pouvons maintenant prendre une métrique $\rho(f_1,f_2) = \|\boldsymbol{\tau}_1-\boldsymbol{\tau}_2\|$, pour la quelle $(\mathcal{F}_1(a,b),\rho_2)$ est un espace de Riesz totalement borné et nous obtenons

$$\sup_{\rho_2(f_1,f_2)\le\delta_n}\rho_n(f_1,f_2) \le 4B(a,b)\delta_n \to 0.$$

Ce ci donne l'hypothèse (**L1**). Maintenant, pour conclure la preuve en établissant la loi fonctionnelle déjà donnée dans le Corollaire 3.2.2, nous allons montrer que l'hypothèse (**L3**) pour $(\mathcal{F}_1(a,b),\rho_2)$ avec des partitions ne dépendant pas de n, de sorte que l'hypothèse (**L2**) soit inutile. La preuve concernant l'hypothèse (**L3**) est très technique,

nous l'affirmons dans l'Annexe D.

Convergence Continue.
Dans le but de montrer (C.1), ous supposons que pour $N > 0$, il existe $n \geq N$ et un couple (f_1, f_2) tels que

$$\rho_n^2(f_1, f_2) < (1 - \Gamma(f_1, f_2)) \tag{H}$$

Nous pouvons réduire les notations et mettre $\rho_n^2(f_1, f_2) = \rho_n^2(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$ pour $f_1(j) = j^{\boldsymbol{\tau}_1}$ and $f_2(j) = j^{\boldsymbol{\tau}_2} \in (a, b)^2$. Il est facile de montrer que

$$\rho_n^2(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2) \to 2(1 - \Gamma(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)),$$

continuement, ceci est,

$$\rho_n^2(\boldsymbol{\tau}_{1,n}, \boldsymbol{\tau}_{2,n}) \to 2(1 - \Gamma(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)).$$

si $(\boldsymbol{\tau}_{1,n}, \boldsymbol{\tau}_{2,n}) \to (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$, quand $n \to +\infty$. Mais, avec l'hypothèse (H), nous pouvons trouver une suite d'entiers $n_1 < n_2 < ... < n_k < n_{k+1} < ..$ et une suite de couples $(\boldsymbol{\tau}_{1,n_k}, \boldsymbol{\tau}_{2,n_k}) \in (a, b)^2$, $k = 1, 2, ...$ tel que pour tout k,

$$\rho_n^2(\boldsymbol{\tau}_{1,n_k}, \boldsymbol{\tau}_{2,n_k}) < (1 - \Gamma(\boldsymbol{\tau}_{1,n_k}, \boldsymbol{\tau}_{2,n_k}). \tag{H1}$$

Par le théorème de Bolzano-Weierstrass, nous pouvons extraire de $(\boldsymbol{\tau}_{1,n_k}, \boldsymbol{\tau}_{2,n_k})$, une sous suite notée $(\boldsymbol{\tau}_{1,n_k^*}, \boldsymbol{\tau}_{2,n_k^*})$ qui converge vers $(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2) \in (a, b)^2$ et par le résultat de continuité

$$\rho_{n^*}^2(\boldsymbol{\tau}_{1,n_k^*}, \boldsymbol{\tau}_{2,n_k^*}) \to 2(1 - \Gamma(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)).$$

Ce ci contredit (H1) et (C.1) est montré. ■

ANNEXE D

Vérification de l' hypothèse (L3) pour $\boldsymbol{\tau} > 1/2$

Nous allons considérer que $\mathcal{F}_2(a,b)$ a un vecteur normé en identifiant $f_{\boldsymbol{\tau}}(j) = j^{\boldsymbol{\tau}}$ avec $\boldsymbol{\tau}$ et en posant $f_{\boldsymbol{\tau}_1} + f_{\boldsymbol{\tau}_2} = f_{\boldsymbol{\tau}_1+\boldsymbol{\tau}_2}$ et $f_{\boldsymbol{\tau}_1} \leq f_{\boldsymbol{\tau}_2}$ si et seulement si $\boldsymbol{\tau}_1 \leq \boldsymbol{\tau}_2$. Nous méttons sans perte de généralité que $b-a=1$ et que $\varepsilon^2 = 1/p$. Ensuite, nous divisons $\mathcal{F}_2(a,b)$ en p intervalles $[0,\boldsymbol{\tau}_1[, [\boldsymbol{\tau}_1,\boldsymbol{\tau}_2[, ..., [\boldsymbol{\tau}_{p-1},\boldsymbol{\tau}_p]$. Posons maintenant,

$$\sigma^2_{k,n}(1,\boldsymbol{\tau}) = k^{-2\boldsymbol{\tau}+1}\sigma^2_{k,n}(f_{\boldsymbol{\tau}})$$

et considérons $\boldsymbol{\tau}_i < \nu < \boldsymbol{\tau} < \boldsymbol{\tau}_{i+1}$ et soit $i=1$, $\boldsymbol{\tau} - \nu = \delta$ et

$$\begin{aligned} d^2_n(f_\nu, f_{\boldsymbol{\tau}}, j) &= \left(\frac{f_\nu(j)}{j\sigma_{k,n}(f_\nu)} - \frac{f_{\boldsymbol{\tau}}(j)}{j\sigma_{k,n}(f_{\boldsymbol{\tau}})}\right)^2 (E_j-1)^2 \qquad \text{(D.1)} \\ &= \left\{\frac{k^{-2\boldsymbol{\tau}+1}}{\sigma^2_{k,n}(1,\boldsymbol{\tau})} j^{2\boldsymbol{\tau}-2} + \frac{k^{-2\nu+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\nu-2} - \frac{2k^{-\nu-\boldsymbol{\tau}+1}}{\sigma_{k,n}(1,\nu)\sigma_{k,n}(1,\boldsymbol{\tau})} j^{\boldsymbol{\tau}+\nu-2}\right\}(E_j-1)^2 \\ &= \Big\{T_0(n,\boldsymbol{\tau},j) + T_0(n,\nu,j) - 2T_1(n,\nu,\boldsymbol{\tau})\Big\}(E_j-1)^2. \end{aligned}$$

Evaluons $T_0(n,\boldsymbol{\tau},j)$. En y ajoutant cette expréssion nulle

$$0 = \left\{\frac{k^{-2\boldsymbol{\tau}+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\boldsymbol{\tau}-2} - \frac{k^{-2\boldsymbol{\tau}+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\boldsymbol{\tau}-2}\right\} + \left\{\frac{k^{-2\nu+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\nu-2} - \frac{k^{-2\nu+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\nu-2}\right\}.$$

On obtient

$$\begin{aligned} T_0(n,\boldsymbol{\tau},j) &= \left(\frac{1}{\sigma^2_{k,n}(1,\boldsymbol{\tau})} - \frac{1}{\sigma^2_n(1,\nu)}\right) k^{-2\boldsymbol{\tau}+1} j^{2\boldsymbol{\tau}-2} \\ &\quad + \frac{k^{-2\nu+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\nu-2}(1-(j/k)^{2\delta}) + \frac{k^{-2\nu+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\nu-2}, \qquad \text{(D.2)} \end{aligned}$$

où nous avons utilisé dans l'expression précédente (D.2) l'identité suivante

$$k^{-2\boldsymbol{\tau}+1} j^{2\boldsymbol{\tau}-2} = k^{-2\nu+1} j^{2\nu-2} + k^{-2\nu+1} j^{2\nu-2}(1-(j/k)^{2\delta}). \qquad \text{(D.3)}$$

En utilisant encore (D.3) in (D.2), nous arrivons à

$$\begin{aligned} T_0(n,\boldsymbol{\tau},j) &= T_0(n,\nu,j) + \left(\frac{1}{\sigma^2_{k,n}(1,\boldsymbol{\tau})} - \frac{1}{\sigma^2_n(1,\nu)}\right) k^{-2\nu+1} j^{2\nu-2} \\ &\quad + \left(\frac{1}{\sigma^2_{k,n}(1,\boldsymbol{\tau})} - \frac{1}{\sigma^2_{k,n}(1,\nu)}\right) k^{-2\nu+1} j^{2\nu-2}(1-(j/k)^{2\delta}) \\ &\quad + \frac{k^{-2\nu+1}}{\sigma^2_n(1,\nu)} j^{2\nu-2}(1-(j/k)^{2\delta}). \end{aligned} \tag{D.4}$$

Nous utiliserons les même techniques pour obtenir

$$\begin{aligned} T_1(n,\nu,\boldsymbol{\tau}) &= T_0(n,\nu,j) + \left(\frac{1}{\sigma_{k,n}(1,\boldsymbol{\tau})\sigma_{k,n}(1,\nu)} - \frac{1}{\sigma^2_{k,n}(1,\nu)}\right) k^{-2\nu+1} j^{2\nu-2} \\ &\quad + \left(\frac{1}{\sigma_{k,n}(1,\boldsymbol{\tau})\sigma_{k,n}(1,\nu)} - \frac{1}{\sigma^2_n(1,\nu)}\right) k^{-2\nu+1} j^{2\nu-2}(1-(j/k)^{\delta}) \\ &\quad + \frac{k^{-2\nu+1}}{\sigma^2_{k,n}(1,\nu)} j^{2\nu-2}(1-(j/k)^{\delta}) \end{aligned} \tag{D.5}$$

Regardons $\sigma^{-2}_{k,n}(1,\boldsymbol{\tau}) - \sigma^{-2}_{k,n}(1,\nu)$ dans (D.4). Nous avons, en utilisant les même méthodes,

$$\sigma^{-2}_{k,n}(1,\boldsymbol{\tau}) - \sigma^{-2}_{k,n}(1,\nu) = \sigma^{-2}_{k,n}(1,\boldsymbol{\tau})\sigma^{-2}_{k,n}(1,\boldsymbol{\tau}) k^{-2\nu+1} \sum_{j=1}^{k} j^{2\nu-2}(1-(j/k)^{2\delta}).$$

Nous notons toujours que $\sigma^2_{k,n}(1,\boldsymbol{\tau}) \to (2\boldsymbol{\tau}-1)^{-1} \in [B=(2b-1)^{-1}, A=(2a-1)^{-1}]$ uniformément dans $\boldsymbol{\tau} \in [a,b]$ par un argument de la convergence continue. Pour ε choisis tel que $2a-1-2\varepsilon > 0$, nous obtenons

$$k^{-2\nu+1} \sum_{j=1}^{k} j^{2\nu-2}(1-(j/k)^{2\delta}) \leq \sum_{j=1}^{k} j^{-1}(j/k)^{2a-1}(1-(j/k)^{2\varepsilon^2})$$

$$= k^{2a-1} \sum_{j=1}^{k} j^{2a-2} - k^{2\varepsilon^2} j^{2a-2+2\varepsilon^2}.$$

En utilisant une approximation asymptotique à l'intégrale correspondante, nous obtenons, quand $n \to \infty$,

$$k^{2a-1} \sum_{j=1}^{k} j^{2a-2} - k^{2\varepsilon^2} j^{2a-2+2\varepsilon^2}$$

$$= k^{-2a+1} K\{(1+o(1))\frac{k^{2a-1}-1}{2a-1}\} - (1+o(1)k^{-2\varepsilon^2}\frac{k^{2a-1+2\varepsilon^2}}{2a-1+2\varepsilon^2})$$

$$\rightarrow \{(2a-1)^{-1} - (2a-1+2\varepsilon^2)^{-1}\} \leq (2\varepsilon^2)/(2a-1)^2.$$

Nous concluons que, pour de grandes valeurs de n

$$k^{-2\nu+1} \sum_{j=1}^{k} j^{2\nu-2}(1-(j/k)^{2\delta}) \leq K\varepsilon^2 \tag{D.6}$$

et donc

$$\left|\sigma_{k,n}^{-2}(1,\boldsymbol{\tau}) - \sigma_{k,n}^{-2}\right| \leq K\varepsilon^2,$$

pour une constante universelle positive K. Cette constante est générique et peut être changée de ligne en ligne. Dans le même esprit, nous avons, pour ce terme dans (D.5),

$$\begin{aligned}\{\sigma_{k,n}(1,\boldsymbol{\tau})\sigma_{k,n}(1,\nu)\}^{-1} - \sigma_n^{-2}(1,\nu) &= \sigma_{k,n}(1,\nu)^{-1}\sigma_{k,n}(1,\boldsymbol{\tau})^{-1} - \sigma_n^{-1}(1,\nu) \\ &= -\frac{\sigma_{k,n}^{-2}(1,\boldsymbol{\tau}) - \sigma_{k,n}^{-2}(1,\nu)}{\sigma_n(1,\nu)^2\sigma_{k,n}(1,\boldsymbol{\tau})(\sigma_{k,n}(1,\boldsymbol{\tau}) + \sigma_{k,n}(1,\nu))}.\end{aligned}$$

La méthode utilisée ci dessus conduit à

$$\left|\{\sigma_{k,n}(1,\boldsymbol{\tau})\sigma_{k,n}(1,\nu)\}^{-1} - \sigma_{k,n}^{-2}(1,\nu)\right| \leq K\varepsilon^2. \tag{D.7}$$

Maintenant, retournons à la formule (D.1), (D.2) et (D.4), nous voyons que les termes $T_0(n,\nu,j)$ disparaît comme $T_0(n,\nu,j) - 2T_0(n,\nu,j) + T_0(n,\nu,j) = 0$. En calculant l'espérance des termes restants, nous obtenons

$$\begin{aligned}Ed_n^2(f_\nu, f_{\boldsymbol{\tau}}, j) \quad \leq \quad & 2\left|\frac{1}{\sigma_n^2(1,\boldsymbol{\tau})} - \frac{1}{\sigma_{k,n}^2(1,\nu)}\right| j^{-1}(j/k)^{2a-1} \\ & +2\left|\frac{1}{\sigma_{k,n}(1,\boldsymbol{\tau})\sigma_{k,n}(1,\nu)} - \frac{1}{\sigma_{k,n}^2(1,\nu)}\right| j^{-1}(j/k)^{2a-1} \\ & +2j^{-1}(j/k)^{2a-1}(1-(j/k)^{2\delta}).\end{aligned}$$

Le terme de droite ne dépend pas de $(\nu,\boldsymbol{\tau}) \in [\boldsymbol{\tau}_{i-1},\boldsymbol{\tau}_i]$ de sorte que, avec l'application de (D.6) and (D.7), nous obtenons

$$\max_{(\nu,\boldsymbol{\tau})\in[\boldsymbol{\tau}_{i-1},\boldsymbol{\tau}_i]} d_n^2(f_\nu, f_{\boldsymbol{\tau}}, j) \leq K\left\{(j/k)^{2a-1}\varepsilon^2 + 2j^{-1}(j/k)^{2a-1}(1-(j/k)^{2\varepsilon^2})\right\}(E_j-1)^2.$$

Finallement, par (D.6),

$$\sum_{j=1}^{k} \mathbb{E} \sup_{(\nu,\boldsymbol{\tau})\in[\boldsymbol{\tau}_{i-1},\boldsymbol{\tau}_i]} d_n^2(f_\nu, f_{\boldsymbol{\tau}}, j) \leq K\varepsilon^2,$$

pour de grande valeurs de n. Alors

$$logN_{[\,]}(\varepsilon, \mathcal{F}_h, L_2^n) \leq p \leq K\varepsilon^{-2}.$$

Ce qui assure l'hypothèse (**L3**) du Théorème 3.3.2. ■

ANNEXE E Une identité utile

Considérons le lemme ci-dessous

Lemme E.0.1

Soit $1 \leq k \leq n$, une suite d'entiers et $x_k, ..., x_n$ $(n-k+1)$ des nombres réels. Nous avons

$$\sum_{i=1}^{k}(x_{n-i+1} - x_{n-k})^2 = 2\sum_{i=1}^{k}\sum_{j=1}^{i} j(1-\delta_{ij}/2)(x_{n-i+1} - x_{n-i})(x_{n-j+1} - x_{n-j}); \tag{E.1}$$

où $\delta_{ij} = 1$ si $i = j$ et 0 sinon, est le symbole de Kronecker.

Démonstrastion du Lemme E.0.1
Nous utiliserons ces notation $S_r = \sum_{1\leq i\leq k} x^r_{n-j+1}$, $r = 1, 2$. Nous avons deux formules:

$$\sum_{i=1}^{h} j(x_{n-j+1} - x_{n-j}) = x_n + ... + x_{n-h+1} - hx_{n-h} \tag{E.2}$$

et

$$\sum_{i=1}^{k}\sum_{j=1}^{i} j(x_{n-i+1} - x_{n-i})(x_{n-j+1} - x_{n-j}) = \frac{1}{2}\left\{(\sum_{j=1}^{k} x_j)^2 - \sum_{j=1}^{k} x_j^2\right\}$$

$$= (S_1^2 - S_2)/2. \tag{E.3}$$

La premiere formule est obtenue par induction et est montrée pour h=2, 3, etc.. et alors, la déduction est imédiate. La seconde is simplement déduite par un développement du carré de la somme des $n-k+1$ nombres. Maintenant, le second terme est (E.1)

$$\sum_{i=1}^{k}\sum_{j=1}^{i-1} j(x_{n-i+1} - x_{n-i})(x_{n-j+1} - x_{n-j}) + \frac{1}{2}\sum_{i=1}^{k} i(x_{n-i+1} - x_{n-i})^2 \equiv A + B.$$

En utilisant (E.2), on obtient

$$A = \sum_{i=1}^{k}(x_{n-i+1} - x_{n-i}) \sum_{j=1}^{i-1} j(x_{n-j+1} - x_{n-j})$$

$$= \sum_{i=1}^{k}(x_{n-i+1} - x_{n-i})(x_n + ... + x_{n-i+2} - (i-1)x_{n-i+1})$$

$$= \sum_{i=1}^{k}(x_{n-i+1} - x_{n-i})(x_n + ... + x_{n-i+1} - ix_{n-i+1})$$

$$= \sum_{i=1}^{k}\sum_{j=}^{i} x_{n-i+1}x_{n-j+1} - \sum_{i=1}^{k} ix_{n-i+1}^2 - \sum_{i=1}^{k}\sum_{j=1}^{i} x_{n-i}x_{n-j+1} + \sum_{i=1}^{k} ix_{n-i}x_{n-i+1}$$

$$\equiv A_{11} + A_{12} + A_{21} + A_{22}.$$

Par changement de variables $i = h - 1$

$$-A_{21} = \sum_{h=2}^{k+1}\sum_{j=1}^{h-1} x_{n-h+1}x_{n-j+1} = \sum_{h=1}^{k}\sum_{j=1}^{h-1} x_{n-h+1}x_{n-j+1} + \sum_{j=1}^{k} x_{n-j+1}x_{n-k}$$

$$= \sum_{h=1}^{k}\sum_{j=1}^{h} x_{n-h+1}x_{n-j+1} - \sum_{h=1}^{k} x_{n-j+1}^2 + \sum_{j=1}^{k} x_{n-j+1}x_{n-k}$$

$$= A_{11} - S_2 + x_{n-k}S_1.$$

De plus

$$2B = \sum_{i=1}^{k} ix_{n-i+1}^2 + \sum_{i=1}^{k} ix_{n-i}^2 - 2\sum_{i=1}^{k} ix_{n-i}x_{n-i+1} = B_1 + B_2 + B_3$$

avec, par changement de variables $i = h - 1$,

$$B_2 = \sum_{h=2}^{k+1}(h-1)x_{n-h+1}^2 = \sum_{h=1}^{k+1}(h-1)x_{n-h+1}^2 = \sum_{h=1}^{k}(h-1)x_{n-h+1}^2 + kx_{n-k}^2$$

$$= \sum_{h=1}^{k} hx_{n-h+1}^2 - \sum_{h=1}^{k} x_{n-h+1}^2 + kx_{n-k}^2 = \sum_{h=1}^{k} hx_{n-h+1}^2 - S_2 + kx_{n-k}^2.$$

Finallement

$$B = \frac{1}{2}(-A_{12} - A_{21}) - \frac{1}{2}S_2 + \frac{1}{2}kx_{n-k}^2 - A_{22}$$

et le second terme E.1 est

$$A_{11} + A_{12} - A_{11} + S_2 - x_{n-k}S_1 + A_{22} - A_{12} - \frac{1}{2}S_2 + \frac{1}{2}kx_{n-k}^2 - A_{22}$$

$$= \frac{S_2 - 2x_{n-k}S_1 + kx_{n-k}^2}{2}.$$

$$\sum_{i=1}^{k}(x_{n-i+1} - x_{n-k})^2 = \sum_{i=1}^{k} x_{n-i+1}^2 - 2x_{n-k}\sum_{i=1}^{k} x_{n-i+1} + kx_{n-k}^1 = S_2 - 2x_{n-k}S_1 + kx_{n-k^2}.$$

Ce qui compléte la preuve du lemme ■

Bibliographie

Alvarado, E., D. Sand Berg, and S. Picford (1998). Modelling large forest fires as extreme events. *Northwest Sci. 72*, 66–75.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *.Mathematical Finance* (9), 203–228.

Balkema, A. and L. de Haan (1974). Residual life time at a great age. *Annals of Probability 2*(5), 792–804.

Beirlant, J., G. Dierckx, A. Guilllou, and C. Stărică (2002). On exponential representations of log-spacings of extreme order statistics. *Extremes 5*(2), 157–180.

Beirlant, J., G. Dierckx, and A. Guillou (2005). Estimation of the extreme value index and regression on generalized quantile plots. *Annals of Statistics 11*(6), 949–970.

Beirlant, J., G. Dierckx, G. Y., and G. Matthys (1999). Tail index estimation and an exponential regression model. *Extremes 2*(2), 177–200.

Beirlant, J., Y. Goegebeur, J. Teugels, and J. Segers (2004). *Statistics of extremes: theory and applications*. Wiley Series in Probability and Statistics. John Wiley and Sons Ltd.

Beirlant, J., Y. Goegebeur, R. Verlaack, and P. Vynckier (1998). Burr regression and portfolio seg- mentation. *Insurance: Mathematics and Economics* (23), 231–250.

Beirlant, J. and G. Matthys (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica 13*(3), 853–880.

Beirlant, J., G. Matthys, and G. Dierckx (2001). Heavy-tailed distributions and rating. *Astin Bulletin* (31), 37–58.

Beirlant, J., P. Vynckier, and J. Teugels (1996). Excess functions and estimation of the extreme value index. *Bernoulli 2*(4), 293–318.

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1987). *Regular Variation*, Volume 27. Encyclopedia of Mathematics and its Applications, Cambridge University Press.

Buhlmann, H. (1980). An economic premium principle. . *Astin Bulletin Ì* (11), 11, 52.

Caeiro, F., M. Gomes, and D. Pestana (2005). Direct reduction of bias of the classical hill estimator. *Statistical Journal* (3(2)), 113–136.

Caeiro, F., M. I. Gomes, and L. H. Rodrigues (2009, January). Reduced-bias tail index estimators under a third-order framework. *Communications in Statistics: Theory and Methods 38*(7), 1019–1040.

Cai, J., L. de Haan, and C. Zhou (2011). Bias correction in extreme value statistics with index around zero. *Submitted*.

Carslaw, H. and J. Jaeger (1948). *Operational Methods in Applied Mathematics, Oxford University Press*.

Ceberián, A. C., M. Denuit, and P. Lambert (2003). Generalzed pareto fit to the society of actuariesâlarge claims database. *North American Actuarial Journal* (7), 18–36.

Centeno, M. and J. Andrade e Silva (2005). Applying the proportional hazard premium calculation principle. *Astin Bulletin* (35), 409–425.

Centeno, M. and M. Guerra (2010). The optimal reinsurance strategy: the individual claim case. *Insurance: Mathematics and Economics 46*, 450–460.

Cheng, S. and L. Peng (2001). Confidence intervals for the tail index. *Bernoulli 7*, 751–760.

Ciuperca, G. and C. Mercadier (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes 13*, 55–87.

Cont, R. (2009). La statistique face aux événements rares = statistics dealing with rare events. *Pour la science* (385), 116–123.

Csörgö, M., S. Csörgő, L. Horváth, and D. Mason (1986). Weighted empirical and quantile processes. *Annals of Probability* (14), 31–85.

Csörgö, S., P. Deheuvels, and D. Mason (1985). Kernel estimates of the tail index of a distribution. *Annals of Statistics 13*(3), 1050–1077.

Csörgö, S. and D. Mason (1985). Central limit theorems for sums of extreme values. *Mathematical Proceedings of the Cambridge Philosophical Society 98*(3), 547–558.

Danielsson, J., L. de Haan, L. Peng, and C. de Vries (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis 76*, 226–248.

Danielsson, J., D. W. J., and C. G. de Vries (1996). The method of moments ratio estimator for the tail shape parameter. *Communication in Statistics, Theory and Methods 4*(25), 711–720.

Davis, R. and S. Resnick (1984). Tail estimates motivated by extreme value theory. *Annals of Statistics 12*(4), 1467–1487.

Davison, A. C. and R. L. Smith (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, B 52*(3), 393–442.

de Haan, L. (1970). On regular variation and its application to the weak convergence of sample extremes. *MMathematical Centre Tracts, 32, Amsterdam. (MR0286156).*

de Haan, L. and A. Ferreira (2006). *Extreme Value Theory: An Introduction.* Springer Series in Operations Research and Financial Engineering, New York Inc.

de Haan, L. and L. Peng (1998). Comparison of tail index estimators. *Statistica Neerlandica 52*(1), 60–70.

de Haan, L. and S. Resnick (1980). A simple asymptotic estimate for the index of a stable law. *Journal of the Royal Statistical Society* (B), 83–87.

de Haan L. and H. Rootzén (1993). On the estimation of high quantiles. *Journal of Statistical Planning and Inference 35*, 1–13.

Deheuvels, P., E. Haeusler, and D. Mason (1988). Almost sure convergence of the Hill estimator. *Mathematical Proceedings of the Cambridge Philosophical Society 104*(2), 371–381.

Deheuvels, P., E. Haeusler, and M. Mason (1989). Laws of the iterated logarithm for sums of extreme values in the domain of attraction of a gumbel law. *Bull. Sci. Math..*

Dekkers, A. and L. de Haan (1989). On the estimation of the extreme value index and large quantile estimation. *Annals of Statistics 17*(4), 1795–1832.

Dekkers, A., J. H. J. Einmalh, and L. De Hann (1989). A moment estimator for the index of an extreme-value distribution. *Annals of Statistics 17*(4), 1833–1855.

Dekkers, A. L. M. and L. de Haan (1993). Optimal choice of sample fraction in extreme-value estimation. *Journal of Multivariate Analysis 47*, 173–195.

Deme, E. H., L. Gardes, and S. Girard (2012). On the estimation of the second order parameter for heavy-tailed distributions. *To appear in REVSTAT, available on http://hal.inria.fr/hal-00634573*.

Deme, E. H., S. Girard, and A. Guillou (2013). Reduced-bias estimator of the proportional hazard premium for heavy-tailed distributions. *Insurance: Mathematics and Economics 52*, 550–559.

Denuit, M. and A. Charpentier (2005). *Mathématiques de lâassurance non-vie : Tarification et provisionnement . Economica, Paris.*

Diebolt, J., M. El-Aroui, M. Garrido, and S. Girard (2005). Quasi-conjugate bayes estimates for gpd parameters and application to heavy tails modeling. *Extremes* (8), 57–78.

Diebolt, J., L. Gardes, S. Girard, and A. Guillou (2008a, August). Bias-reduced estimators of the weibull tail-coefficient. *Test 17*(2), 311–331.

Diebolt, J., L. Gardes, S. Girard, and A. Guillou (2008b). Bias-reduced extreme quantiles estimators of weibull distributions. *Journal of Statistical Planning and Inference 138*, 1389–1401.

Diebolt, J., L. Gardes, and A. Guillou (2008). Bias-reduced estimators of the weibull tail-coefficient. *Test* (17(2)), 311–331.

Diebolt, J., A. Guillou, and I. Rached (2004). A new look at probability-weighted moments estimators. *C. R. Acad. Sci. de Paris,* (338), Serie I, 8, 629–634.

Diebolt, J., A. Guillou, and I. Rached (2007). Approximation of the distribution of excesses through a generalized probability-weighted moments method. *Journal of Statistical Planning and Inference 3*(137), 841–857.

Diop, A. and G. Lo (2006). Generalized hill's estimator. *Journal of Theorical Statistic 20(2)*, 129–149.

Diop, A. and G. S. Lo (2009). Ratio of generalized hillâs estimator and its asymptotic normality theory. *Mathematical Methods of Statistics* (18(2)), 117–133.

Drees, H. (1995). Refined Pickands estimator of the extreme value index. *Annals of Statistics 23*(6), 2059–2080.

Drees, H. and E. Kaufmann (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and their Applications 75*(2), 149–172.

Drees, H. L.

El-Adlouni, S., B. Bobée, and T. B. Ouarda (2007). Caracterisation des distributions à queues lourdes pour l'analyse des crues. Technical Report no r-929, INRS-ETE, Université du Québec.

El Methni, J., L. Gardes, S. Girard, and A. Guillou (2012). Estimation of extreme quantiles from heavy and light tailed distributions. *Journal of Statistical Planning and Inference* (142(10)), 2735–2747.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Springer Verlag.

Ferreira, A., L. de Haan, and L. Peng (2003). On optimizing the estimation of high quantiles of a probability distribution. *Statistics 37*, 401–434.

Feuerverger, A. and P. Hall (1999). Estimating a tail exponent by modeling departure from a pareto distribution. *Annals of Statistics 27*, 760–781.

Fisher, R. and L. Tippet (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society 24*, 180–190.

Fraga Alves, M., L. de Haan, and T. Lin (2003). Estimation of the parameter controlling the speed of convergence in extreme value theory. *Mathematical Methods of Statistics 12*, 155–176.

Fraga Alves, M., M. Gomes, and L. de Haan (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica 60*, 193–214.

Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique 6*, 93–116.

Gardes, L. and S. Girard (2010). Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes* (13(2)), 177–204,.

Geluck, J. and L. de Haan (1987). Regular variation, extensions and tauberian theorems, cwi tract 40, center for mathematics and computer science, amsterdam, netherlands.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics 44*(3), 423–453.

Goegebeur, Y., J. Beirlant, and d. T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference 140*, 2632–2652.

Goegebeur, Y. and T. de Wet (2012). Estimation of the third-order parameter in extreme value statistics. *Test* (21(2)), 330–354,.

Gomes, M., L. de Haan, and L. Peng (2002). Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes 5*, 387–414.

Gomes, M. and M. Martins (2004). Bias reduction and explicit semi-parametric estimation of the tail index. *ournal of Statistical Planning and Inference* (24), 361–378.

Gomes, M. I. (1999). Generalized jackknife moment estimator of the tail index. *Bulletin of the International Statistical Institute 58*(1), 401–402.

Gomes, M. I., F. Figueiredo, and S. Mendonça (2005a). Asymptotically best linear unbiased tail estimators under a second-order regular variation condition. *Journal of Statistical Planning and Inference 134*(2), 409–433.

Gomes, M. I. and M. Martins (2002). Asymptotically unbiased estimators of the tail index based on external estimation of the second order parameter. *Extremes 5*, 5–31.

Gomes, M. I., M. Martins, and M. Neves (2007). Improving second order reduced bias extreme value index estimator. *REVSTAT - Statistical Journal* (5(2)), 177–207.

Gomes, M. I. and O. Oliveira (2001). The bootstrap methodology in statistics of extremes: theory and applications - choice of the optimal sample fraction. *Extremes 4*(4), 331–358.

Gomes, M. I., H. Pereira, and M. Miranda (2005b). Revisiting the role of the jackknife methodology in the estimation of a positive tail index. *Communications in Statistics–Theory and Methods 34*(2), 319–335.

Gomes, M. I., D. Pestana, and C. Frederico (2009, febuary). A note on the asymptotic variance at optimal levels of a bias-corrected hill estimator. *Statistics and Probability Letters 79*(3), 295–303.

Goovaerts, M., F. de Vylder, and J. Haezendonck (1984). Insurance premiums. *North-Holland,* (VII +398).

Griselda, D. and P. Guillaume (2006). *Théorie du risque et réassurance. Economica, Paris.*

Groeneboom, P., H. P. Lopuhaa, and de Wolf P.-P. (2003). Kernel-type estimators for the extreme value index. *Annals of Statistics 31*, 1956–1995.

Gumbel, E. J. (1958). *Statistics of extremes: theory and applications.* Columbia University Press.

H. Drees, a. d. and S. Resnick (2000). How to make a hill plot. *Annals of Statistics 28*, 254–274.

Haeusler, E. and J. Teugels (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. *Annals of Statistics 13*(2), 743–756.

Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society 44*, 37–42.

Hall, P. and A. Welsh (1985). Adaptive estimates of parameters of regular variation. *Annals of Statistics 13*, 331–341.

Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics 3*(5), 1163–1174.

Hosking, J. R. M. and J. R. Wallis (1987, August). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics 29*(3), 339–1349.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *The Quarterly Journal of the Royal Meteorological Society 81*(384), 158–171.

Jones, B. L. and R. Zitikis (2003). Empirical estimation of risk measures and related quantities. *North American Actuarial Journal* (7), 44–54.

Kratz, M. and S. Resnick (1996). The qq-estimator and heavy tails. *Stochastic Models 12*(4), 699–724.

Leadbetter, M. R. (1983). Extremes and local dependence of stationnary sequences. *Z. Wahr. verw. Gebiete* (65), 291–306.

Lo, G. (1986). *Sur quelques estimateurs de lâindex dâune loi de Pareto : Estimation de Deheuvels Csorgo, Mason, de De Haan-Resnick et lois limites pour des sommes de valeurs extrêmes pour une variable dans le domaine de Gumbel. Thèse de Doctorat. Universit´ Paris VI. France.* Ph. D. thesis.

Lo, G. (1989). A note on the asymptotic normality of sums of extreme values. *127 ournal of Statistical Planning and Inference* (22), 127–136.

Lo, G. and E. H. Deme (2012). A functional generalized hill process and applications. *International Journal of Statistics and Probability 1*(2), 250–268.

Lo, G., E. H. Deme, and A. Diop (2012). On the generalized hill process for small parameters and applications. *Journal statistical theory applications 11*(4), 397–418.

Lo, G. S. (1992). Sur une caract´risation empirique des extrêmes. *Mathematical Reports of the Academy of Science. Canada* (XIV,2,3), 89–94.

Markovich, N. (2005). High quantile estimation for heavy-tailed distributions. *Performance Evaluation 62*, 178–192.

Mason, D. (1982). Laws of large numbers for sums of extreme values. *Annals of Probability 10*(3), 754–764.

Mason, D. and T. Turova. Weak convergence of the hill estimator pro- cess. in j. galambos, j. lechner and simiu, editors, extreme value theory and applications. kluwer academic publishers, dordrecht, holland.

Matthys, G. and J. Beirlant (2003). Estimating the extreme value in- dex and high quantiles with exponential regression models. *Statistica Sinica 13*, 853–880.

McCulloch, J. (1997). Measuring tail thickness to estimate the stable index α : A critique. *Journal of Business & Economic Statistics* (15), 74–81.

McNeil, A. (1997). Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin* (27), 117–137.

Necir, A. and K. Boukhetala (2004). Estimating the risk adjusted premium of the largest reinsurance covers. *Proceeding in Compu- tational Statistics (Edited by Jaromir Antoch)*, 1577–1584. Physica–Verlag, Springer, ISBN: 3–7908–1554–3.

Necir, A. and D. Meraghni (2009). Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance: Mathematics and Economics* (45), 49–58.

Necir, A., D. Meraghni, and F. Meddi (2007). Statistical estimate of the proportional hazard premium of loss. *Scandinavian Actuarial Journal* (3), 147–161.

P., H. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis 32*, 177–203.

Peng, L. (1998). Asymptotically unbiased estimators for the extreme-value index. *Statistics and Probability Letters 38*(2), 107–115.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics 3*(1), 119–131.

Reiss, R. and M. Thomas (2001). Statistical analysis of extreme values. *Birkhanser Verlag*.

Reiss, R.-D. and M. Thomas (1997). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. Birkh Ìuser, Basel.*

Reiss, R.-D. and M. Thomas (1999). A new class of bayesian estimators in paretian excess-of-loss reinssurance. *Astin Journal 29*(2), 339–349.

Resnick, S. (1987). *Extreme Values, regular Variation, and Point Process.* Springer Verlag, New-York.

Resnick, S. and C. Starica (1995). Consistency of hill's estimator for dependent data. *Journal of Applied Probability* (32), 139–167.

Rootzen, H. and N. Tajvidin (1997). Extreme value statistics and wind storm losses: a case study. *Scandinavian Actuarial Journal 1*, 70–94.

Rosen, O. and I. Weissman (1996). Comparison of estimation methods in extreme value theory. *Communication in Statistics-Theory and Methods 24*(4), 759–773.

Schultze, J. and J. Steinebach (1996). On least squares estimates of an exponential tail coefficient. *Statistics and Decisions 14*, 353–372.

Shorack, G. and J. Wellner (1986). *Empirical Processes with Applications to Statistics. John Wiley*
Sons, New York.

Smith, R. L. (1987). Estimating tails of probability distributions. *Annals of Statistics 15*, 1174–1207.

Sousa, B. (2002). A contribution to the estimation of the tail index of heavy-tailed distributions. phd. thesis, university of michigan.

Tsourti, Z. and J. Panaretos (2001). A simulation study on the performance of extreme-value index estimators and proposed robustifying modifications. *5th Hellenic Eu- ropean Conference on Computer Mathematics and its Applications, Athens, Greece, 2*, 847–852.

Tsourti, Z. and J. Panaretos (2003). Stochastic musings : Perspectives from the pioneers of the late 20th century. *Laurence Erlbaum, john panaretos édition*, 141–160.

Valiron, G. (1990). *Cours dâAnalyse Math´matique : Th´orie des fonctions. 3ième Edition. Masson, Paris.*

van der Vaart , A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics . Springer, New-York. (MR1385671).*

Vandewalle, B. and J. Beirlant (2006). On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance: Mathematics and Economics* (38), 441–459.

Viharos, L. (1995). Limit theorems for linear combinations of extreme values with applications to inference about the tail of a distribution. *Acta Scientiarum Mathematicarum 60*, 761–777.

von Mises, R. (1936). La distribution de la plus grande de n valeurs. *Revue de Mathématique Union Interbalcanique* (1), 141–160.

Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazard transforms. *nsurance: Mathematics and Economics* (17), 43–54.

Wang, S. (1996). Premium calculation by transforming the layer premium density. *Astin Bulletin* (26), 71–92.

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.

Worms, J. and R. Worms (2011). Estimation of second order parameters using probability weighted moments. *ESAIM: Probability and Statistics, to appear*.

Yaari, M. (1987). The dual theory of choice under risk. *Econometrica 55*, 95–115.