UNIVERSITE SAAD DAHLAB DE BLIDA 1

Faculté des Sciences

Département de Mathématiques

THESE DE DOCTORAT

en Mathémathiques

MODELISATION DES DISTRIBUTIONS A QUEUES LOURDES : APPLICATION DANS LE DOMAINE D'ACTUARIAT

Par

Omar TAMI

Devant le jury composé de :

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

Blida, 10 Mars 2021

Remerciements

Toute ma gratitude et mes remerciements à Mr Rassoul Abdelaziz pour toute sa disponibilité, sa présence et ses précieux conseils ainsi qu'à Mr Ould Rouis Hamid.

Je voudrais aussi remercier les membres du jury qui ont accepté de juger cette thèse malgré la charge de travail qu'ils subissent.

A toute l'équipe du département de Maths de la faculté des Sciences :

un grand merci.

A toute ma famille aussi j'adresse mes remerciements.

Omar Tami

ملخص

تنقسم هذه الرسالة إلى أربعة فصول تضاف إليها مقدمة وخاتمة. في الفصل الأول ، نتذكر بعض المفاهيم الأساسية حول نظرية القيم المتطرفة. في الفصل الثاني ، نعتبر أن حالة التوزيعات تنتمي إلى مجال الجذب فريشت Fréchet أو توزيعات نوع Pareto ، وتعتمد هذه التوزيعات على معامل $xi > 0$ ، المقدر المعتاد لهذا تم تقديم التقدير بواسطة Hill 1975[107]، أحدها يعرض التقنيات المختلفة لبناء هذا المقدّر ، ودراسة السلوكيات المقاربة ، وتقنيات تقليل التحيز لـ xi على أساس مقدراتها بشرط الشرط الثاني وننتهي ببعض التوزيعات من هذا النوع. في الفصل الثالث ، نقدم لأنفسنا بعض مؤشرات عدم المساواة ، لورنز كيرف ، جيني ، ثيل ، أتكنسون ، بونفيروني ، زينغا. أخيرًا ، في الفصل الرابع ، نركز على مؤشر Zenga الذي قدمه Zenga 2007 [171] تقديرًا للأخير بناءً على التوزيع التجريبي وأثبت طبيعته المقاربة في ظل ظروف مناسبة ، والتي لا يتم الوفاء بها غالبًا في حالة التوزيعات ذات الذيل الثقيل. وهكذا ، فإننا ننظر إلى هذا الإطار ؛ نحن نعتبر عائلة مقدرة للمؤشر على أساس نهج نظرية القيم المتطرفة. نحن نؤسس طبيعتها المقاربة ونقترح أيضًا نهجًا لتقليل التحيز لهذه المقدرات. تتيح دراسات المحاكاة تقدير جودة مقدراتنا المقترحة.

الكلمات الرئيسية: نظرية القيمة المتطرفة ، الذيل الثقيل ، مؤشر الذيل ، معامل الدرجة الثانية ، مقدر التل ، مؤشر عدم المساواة ، منحنى لورنز ، مؤشر Zenga ، الخصائص المقاربة.

Résumé

Cette thèse est divisée en quatre 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 considèrons le cas des distributions appartenant au domaine d'attraction de Fréchet ou bien les distributions de type Pareto. Ces distributions dépendent d'un paramètre $\xi > 0$, l'estimateur usuel de ce paramètre est introduit par Hill, (1975) [107]. On présente les différentes techniques de construction de cet estimateur, l'étude des comportements asymptotiques, et les techniques de réduction de biais pour ξ, ses estimateurs basés sur la condition de régularité d'ordre deux, et on termine par quelques distributions de ce type. Dans le chapitre trois, nous présentons quelques indices des inégalités, Courbe de Lorenz, Gini, Theil, Atkinson, Bonferroni, Zenga. Finalement, au chapitre quatre, nous nous concentrons sur l'indice de Zenga introduit par Zenga (2007) [171]. Zenga 2007) a donné une estimation de ce dernier basé sur la distribution empirique et a établi sa normalité asymptotique sous certaines conditions appropriées, qui ne sont pas souvent remplies dans le cas des distributions à queues lourdes. Ainsi, dans ce carde la, nous considérons une une famille d'estimateurs de l'indice basé sur l'approche de la théorie des valeurs extrêmes. Nous établissons leurs 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 proposés.

Mots-clés : Théorie des valeurs extrêmes, queue lourde, indice de queue, paramètre du second ordre, estimateur de Hill, indice d'inégalité, courbe de Lorenz, indice de Zenga, propriétés asymptotiques.

Abstract

This thesis is divided into four chapters to which are added an introduction and a conclusion. In the first chapter, we recall some basic notions about the theory of extreme values. In the second chapter, we consider the case of the distributions belongs to the Fréchet domain of attraction or the Pareto type distributions, these distributions are dependent on a parameter $\xi > 0$, the usual estimator of this parameter is introduced by Hill, (1975) [107], one presents the different technics of construction of this estimator, and the study of the asymptotic behaviors, and the technics of reduction of bias for ξ its estimators based on the condition of the second oder condition, and we end with some distributions of this type. In chapter three, we introduce ourselves some index of inequality, Lorenz Curve, Gini, Theil, Atkinson, Bonferroni, Zenga. Finally, in chapter four, we focus on the Zenga index introduced by Zenga (2007)[171]. Zenga 2007) gave an estimate of the latter based on the empirical distribution and established its asymptotic normality under appropriate conditions, which are not often met in the case of heavy-tailed distributions. Thus, we look at this framework ; we consider a family of estimators of the index based on the approach of the theory of extreme values. We establish their asymptotic normality and we also propose an approach to reduce bias for these estimators. Simulation studies make it possible to appreciate the quality of our proposed estimators.

Key words : Extreme value theory, heavy tail, tail index, second order parameter, Hill estimator, inequality index, Lorenz curve, Zenga index, asymptotic properties.

Table des matières

Table des figures

Liste des tableaux

Introduction Générale

Les phénomènes rares et /ou catastrophiques dominent l'actualité quotidienne par leur caractère imprévisible. Ils sont variés et souvent de caractères physiques en particulier les catastrophes naturelles : les séismes, les éruptions volcaniques, les tsunamis, les mouvements de terrain, les inondations, les tempêtes, les cyclones, les orages etc faisant ainsi des ravages sur leur passage. Les exemples ci-dessous illustrent les dégâts matériels et humains que peut engendrer un événement catastrophique rare.

À l'échelle mondiale, on recense annuellement environ un millier de grandes catastrophes «naturelles» en majeure partie provoquées par les crues, événements naturels les plus fréquents et les plus de structeurs ; leurs causes initiales sont toujours météorologiques : moussons, cyclones, tempêtes.

En confondant les effets désastreux de certains de ces événements avec les causes des catastrophes qui en résultent, il a été longtemps considéré que les causes étaient des punitions et que les effets étaient inéluctables, fatals, prescrits... La science et la technique permettent maintenant de caractériser les événements, de prévoir leurs effets, d'établir et distinguer les causes naturelles d'avec les causes humaines des catastrophes pour améliorer la prévention et la gestion des secours.

Approche probabilistique

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, le théorème central limite ou encore l'analyse de la variance. Cette approche ne fournit pas d'informations fiables sur les évènements extrêmes c'est-à-dire sur les queues de distributions des événements. Pour caractériser et quantifier le comportement de ces événements extrêmes, une nouvelle théorie est nécessaire, la Théorie des Valeurs Extrêmes (TVE). Cette théorie englobe des modèles stochastiques extrêmes adéquats pour modéliser et décrire la survenue et l'intensité d'évènements dits rares c'est-à-dire qui présentent des variations à très grandes amplitudes ou à très faibles amplitudes et ayant une très faible probabilité d'apparition. L'étude des lois de ces événements extrêmes n'est possible que si le comportement de ces derniers est dû au hasard (notion de probabilité). Ils sont dits ex-

trêmes quand il s'agit de valeurs beaucoup plus grandes ou plus petites que celles observées habituellement. La modélisation de tels événements est de nos jours un champ de recherches particulièrement actif due notamment à l'importance de leurs impacts socio-économiques et à la longue collecte de données enregistrant de tels événements. L'analyse des valeurs extrêmes requiert l'estima-tion d'un indice de queue qui donne une indication essentielle sur la forme de la queue de distribution des événements.

Cette théorie développée par Fisher et Tippet (1928) [61] sur les lois limites possibles du maximum d'un échantillon a montré que la théorie des valeurs extrêmes était quel que chose de spécial et pas comme la théorie classique de la limite centrale. Depuis,ce résultat de Fisher et Tippet (1928) [61]. Cette théorie a été étudié par Gnedenko (1943) [88] qui obtient rigoureusement la convergence, dont la preuve fut simplifiée par de Haan (1976) [38]. L'unification de ce résultat est due aux travaux de Von Mises (1936) [165] et Jenkinson (1955) [124]. Cette analyse repose principalement sur des distributions limites des extrêmes et leurs domaines d'attraction. Cependant, on y retrouve deux modèles :

— La loi généralisée des valeurs extrêmes (GEV : «Generalized Extreme Value»),
— La loi de Paréto généralisée (GPD : «Generalized Pareto Distribution»).

Limites de cette approche

Par définition, les événements extrêmes sont peu nombreux de par leurs fréquences d'apparitions rares voire même inexistants rendant ainsi leurs modélisations diffciles. L'information la plus précise est celle contenue dans les valeurs observées les plus extrêmes. De ce fait, à partir de peu de données, on doit construire des modèles nous permettant d'extrapoler et de prédire un événement sans commune mesure ce qui conduit à deux problèmes en

— pratique Le premier est lié à la taille de l'échantillon qui est souvent faible remettant en question l'applicabilité des résultats asymptotiques (La taille minimale $n = 50$ a été re-commandée par Stedinger (2000) [158] pour avoir des estimations robustes). Néanmoins, cette taille ne peut pas fournir assez d'informations si on s'intéresse à la prédiction d'événements extrêmes sur de longues périodes.
— Le second est dû au fait qu'une loi de probabilité ne donne pas toujours un bon ajustement dans toutes les applications (cf.Bobée & Rasmussen(1995) [15]). Dans ce cas, il est nécessaire d'effectuer un classement des distributions en fonction du comportement de leurs queues et à partir de considérations physiques ou statistiques, d'établir des critères de discrimination entre les différentes classes dans le cas d'un échantillon de faible taille.

Domaines d'application

La théorie des valeurs extrêmes (TVE) fournit une base mathématique probabiliste rigoureuse sur laquelle il est possible de construire des modèles statistiques permettant de prévoir l'intensité et la fréquence de ces événements extrêmes. La TVE est très répandue ces dernières décennies dans la littérature car elle permet d'apporter des réponses à de nombreux problèmes pratiques et c'est dans ce contexte que la théorie des valeurs extrêmes développée par Fisher et Tippet (1928) [61] a trouvé toute sa place. Les domaines d'applications utilisant les modèles de la TVE n'ont cessé de se développer ces dernières années.

— En hydrologie, le domaine d'application historique dû notamment aux travaux de Gumbel & Lieblein (1954) [100], domaine dans lequel la prévision des crues par exemple est particulièrement importante (cf. Davison & Smith (1990) [34]; Katz et al.(2002) [113]).

— En climatologie avec l'étude et la prédiction des évènements climatiques extrêmes comme les précipitations extrêmes, les canicules, les chutes de neige, les avalanches (cf. Rootzén & Tajvidi (2001) [147]; Heneka et al. (2006) [106]; Brodin & Rootzén (2009) [19]).

— En météorologie où l'étude de la vitesse du vent, par exemple, permet d'évaluer le degré de résistance des matériaux face à la pression exercée par le vent (au cours d'une tempête par exemple) sur les bâtiments ou les structures de génie civil (cf. Coles & Walshaw (1994) [24]; Smith (2001) [157]; Klajnmic (2004) [115]; Khaliq et al. (2006) [114]; Davison et al. (2012) [35]).

— En environnement, avec la modélisation de grands feux de forêts comme des événements extrêmes (cf. Alvarado et al. (1998) [3] voir aussi Ferrez et al. (2011) [58]).

— Dans les domaines des sciences humaines et sociales, plus particulièrement dans le domaine de la démographie, tout un débat qui a été initié par Gumbel (1937) [98], auquel Fréchet a pris une part active, sur la notion de "durée extrême de la vie humaine" et sur sa mesure. Aarssen & Haan (1994) [1]; Han (2005) [105] proposèrent des résultats afin de calculer l'âge limite possible de l'être humain.

— En assurance, avec l'étude de la survenue des sinistres d'intensité exceptionnelle qui peuvent avoir des conséquences négatives sur les résultats et la solvabilité des organismes d'assurance (cf. Barrois (1834) [7]; McNeil & Saladin (1997) [128]; Rootzén &Tajvidi (1997) [146]).

— En finance, elle apporte une réponse immédiate à la remise en cause de l'hypothèse de normalité sur tout avec les observations à hautes amplitudes (cf. Embrechts et al. (1997) [56]; Danielsson & de Vries (1997) [31]; Embrechts et al.

(1999) [57] ; McNeil & Frey (2000) [129] ; Longin (2000) [121] ; Gençay & Selcuk (2004) [86]). Cependant, Bouleau (1991) [17] met en garde sur les mauvaises utilisations de la théorie des valeurs extrêmes.

Notre contribution

Cette thèse s'inscrit dans sa globalité comme étant une contribution à la théorie des valeurs extrêmes dans l'économie des revenus et ses applications statistiques.

En économie, un indicateur est une statistique construite afin de mesurer certaines dimensions de l'activité économique, ceci de façon aussi objective que possible. Leurs évolutions ainsi que leurs corrélations avec d'autres grandeurs sont fréquemment analysées à l'aide de méthodes économétriques.

Les indicateurs sont construits par l'agrégation d'indices qui figurent dans un document appelé « tableau de bord ». La construction des indicateurs découle d'un choix de conventions qui traduisent plus ou moins bien certaines priorités et valeurs éthiques et morales. Le « Tableau économique » de François Quesnay, l'un des premiers physiocrates qui a vécu au XVIIIe siècle, constitue l'un des premiers exemples d'un tel indicateur visant à mesurer la richesse d'un pays. Depuis les développements des comptes nationaux après la Seconde Guerre mondiale, le produit intérieur brut (PIB) et le produit national brut (PNB) sont les indicateurs les plus courants. Par ailleurs, il existe d'autres indicateurs qui prennent en compte d'autres facteurs ignorés par le PNB et le PIB afin de mesurer le bien-être des habitants d'un pays ; en incluant par exemple des indicateurs de santé, d'espérance de vie, de taux d'alphabétisation. Le Programme des Nations Unies pour le développement (PNUD) a ainsi créé l'indice de développement humain (IDH) dans les années 1990.

Des tentatives pour prendre en compte d'autres dimensions telles la sécurité ou pour inclure la « soutenabilité écologique » de l'activité économique dans des indicateurs ont aussi été menées plus récemment.

L'inégalité est un concept plus large que la pauvreté dans la mesure où elle est définie sur l'ensemble de la population et ne se concentre pas uniquement sur les pauvres. La mesure la plus simple de l'inégalité trie la population du plus pauvre au plus riche et montre le pourcentage de la dépense (ou du revenu) attribuable à chaque cinquième (quintile) ou dixième (décile) de la population. Le quintile le plus pauvre représente généralement de 6 à 10% de toutes les dépenses, le quintile supérieur étant de 35 à 50%. Une mesure populaire de l'inégalité est L'indice de Gini, basé sur le travail de Gini, est utilisé pour décrire l'inégalité de revenu dans une population.
Les pays publient leur propre indice de Gini. De nombreuses institutions, y compris la Banque mondiale et la CIA, calculent l'indice de Gini des pays du monde. L'inégalité des revenus dans les «limites optimales» favorise la croissance. Le taux d'inégalité

des revenus dans les pays riches du monde évite l'égalitarisme extrême et l'extrême inégalité. Il n'y a pas de corrélation entre Gini et la richesse pour les pays les plus pauvres, qui va de 0 (égalité parfaite) à 1 (inégalité parfaite), mais se situe généralement entre 0,3 et 0,5 pour les dépenses par habitant. L'indice de Gini est dérivé de la courbe de Lorenz, qui trie la population du plus pauvre au plus riche, et montre la proportion cumulative de la population sur l'axe horizontal et la proportion cumulée des dépenses (ou revenu) sur l'axe vertical. Bien que le coefficient de Gini ait de nombreuses propriétés souhaitables - indépendance moyenne, indépendance de la taille de la population, symétrie et sensibilité au transfert de Pigou-Dalton - il ne peut être facilement décomposé pour montrer les sources d'inégalité.

Pour mesurer et comparer la disparité, il faut incorporer la nature relative du «petit» et du «grand», et pour cette raison nous employons les indices de l'inégalité économique. Cela rend le développement de l'inférence statistique un défi, même pour les populations légères, laisser seuls ceux à queue lourde, comme c'est le cas avec les revenus du capital, nous utilisons l'indice d'inégalité économie Zenga nouvellement développé. Son estimateur non paramétrique ne fait partie d'aucune catégorie de statistiques bien connue. Cela rend le développement de l'inférence statistique un défi même pour les populations légères, et encore moins pour les populations à queue lourde, comme c'est le cas pour les revenus du capital.

L'augmentation observée de l'inégalité économique, où la préoccupation majeure est par rapport à l'énorme croissance des revenus les plus élevés, motive à revisiter les mesures classiques de l'inégalité et à offrir de nouvelles façons de synthétiser la variabilité de l'ensemble de la distribution des revenus.
L'idée et l'objectif sont de fournir aux études les inférences statistiques de l'indice de Zenga). Dans cette thèse, nous allons présenter les notions et les définitions de quelques vocabulaire dans l'économie et les revenus, et des outils pour analyser quelques indicateurs classiques d'inégalité : Courbe de Lorenz, Indice de Gini, Indicateur de Theil (1967) [162], Indice d'Atkinson (1970) [5], Indice Bonferroni (1930) [16], Indice d'inégalité de Zenga (2007) [171] et l'on va s'intéresser à une mesure de bien-être. De l'inégalité à la pauvreté, les indices de pauvreté et inégalité et la décomposition des indices.

Organisation de la thèse

— Le chapitre 1 présente un état de l'art de la théorie des valeurs extrêmes. La section 2 rappelle les principaux résultats et définitions de la théorie des valeurs extrêmes utiles dans nos travaux, plus particulièrement, l'estimation de l'indice de queue lourde et des quantiles extrêmes correspondants qui constituent la problématique de cette thèse. Dans cette partie, nous partons du résultat prin-

cipal de la TVE qui 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.

— Le deuxième chapitre est consacré pour l'étude des distributions de type Pareto ou les distributions à queues lourdes, nous étudierons les estimateurs de l'indice de queue et des quantiles extrêmes correspondants. D'autre part, on présente les différentes techniques pour réduire le biais de l'estimateur de Hill.

— Au troisième chapitre, nous allons présenter les notions et les définitions de quelques vocabulaire dans l'économie et les revenus, et desoutils pour analyser quelques indicateurs classiques d'inégalité : Courbe de Lorenz, Indice de Gini [87], Indicateur de Theil (1967) [162], Indice d'Atkinson (1970) [5], Indice Bonferroni (1930) [16], Indice d'inégalité de Zenga (2007) [171] et l'on va s'intéresser à une mesure de bien-être, De l'inégalité à la pauvreté, Les indices de pauvreté, Pauvreté et inégalité et la décomposition des indices.

— Le chapitre 4, nous avons étudié estimation paramétrique et semi paramétrique pour l'indice de Zenga, en particulier pour les distributions des inégalités à queues lourdes, leurs propriétés asymptotiques. Ensuite, nous avons simulé pour les deux modèles de Pareto et de Fréchet respectivement. Le document se termine par une conclusion et de perspectives.

Chapitre 1

Rappels sur la théorie des valeurs extrêmes

1.1 Introduction

La théorie des valeurs extrêmes communément appelée « **Extreme Value Theory** » (**EVT**) en anglais, est une vaste théorie dont le but est d'étudier les événements rares c'est-à-dire les événements dont la probabilité d'apparition est faible. Autrement dit elle essaie d'amener des éléments de réponses aux intempéries, aux inondations, aux catastrophes naturelles, aux problèmes financiers, etc. en prédisant leurs occurrences dans les années à venir. En d'autres termes on veut estimer des petites probabilités ou des quantités dont la probabilité d'observation est très faible c'est-à-dire proche de zéro. Ces quantités sont appelées quantiles extrêmes car l'ordre de ces quantiles tend vers zéro lorsque la taille de l'échantillon, n, tend vers l'infini.

On considère, pour cela, n variables aléatoires réelles $(X_i)_{1\leqslant i\leqslant n}$ indépendantes et identiquements distribuées (iid) de fonction de répartition F non nécessairement continue. Et soit $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ l'échantillon ordonné associé.

1.2 Fonction de répartition et quantile

Soit $X_1, ..., X_n$, un échantillon aléatoire tiré d'une loi dont la fonction de répartition est F. Si F est inconnue, on peut l'estimer par la fonction de répartition empirique

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \leq x\}}. \tag{1.1}$$

Pour $x \in \mathbb{R}$ fixé, $F_n(x)$ correspond au nombre d'observations qui sont inférieures à x. Nous allons nous intéresser à présent au comportement asymptotique de F_n.
Tout d'abord, il faut noter que la loi des grands nombres assure que, pour x fixé, $F_n(x)$

converge en probabilité vers $F(x)$.
Le Théorème de Glivenko-Cantelli étend largement ce résultat en démontrant la convergence uniforme de F_n vers F pour $x \in \mathbb{R}$. Ce résultat est énoncé dans la suite.

Théorème 1.1 *Soit $(X_n)_{n\geq 1}$, une suite de variables aléatoires de même fonction de répartition F. Alors*

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)|$$

converge en probabilité vers zéro.

Considérons maintenant le processus empirique

$$\mathbb{F}_n(x) = \sqrt{n}\{F_n(x) - F(x)\}.$$

Une application du Théorème central limite assure que pour un $x \in \mathbb{R}$ fixé, $\mathbb{F}_n(x)$ converge vers une variable aléatoire Normale de moyenne nulle et de variance

$$\sigma_x^2 = F(x)(1 - F(x)).$$

Le Théorème de Donsker va beaucoup plus loin en obtenant le comportement limite en loi de $F_n(x)$ en tant que fonction aléatoire définie pour tout $x \in \mathbb{R}$.

Définition 1.1 *Le quantile d'ordre α de la fonction de distribution F est défini par*

$$q(\alpha) := F^{-1}(\alpha) = \inf\{X : F(y) > \alpha\},$$

avec $\alpha \in]0,1[$, où F^{-1} est l'inverse généralisée de F. Par convention $\inf\emptyset = \infty$. Notons que l'inverse généralisé d'une fonction coïncide avec l'inverse classique lors que la fonction est continue. Ainsi, un quantile sera dit extrême il'on remplace son ordre α par une suite $\alpha_n \to 0$ quand $n \to \infty$. Concrètement, un quantile extrême d'ordre $1 - \alpha_n$ de la fonction de distribution F est définie par :

$$q(\alpha_n) := \overline{F}^{-1}(\alpha_n) = \inf\left\{y : \overline{F}(y) \leq \alpha_n\right\} \text{ avec } \alpha_n \to 0 \text{ quand } n \to \infty, \tag{1.2}$$

où $\overline{F} := 1 - F$ est la fonction de survie de F.

Il faut dire que le fait que l'ordre $\alpha_n \to 0$ quand $n \to \infty$ indique que l'information la plus importante pour estimer des quantiles extrêmes est contenue dans la queue de distribution.

Si la taille de l'échantillon tend vers l'infini, nous avons

$$\begin{aligned} P(X_{(n)} &< q(\alpha_n)) = P\left(X_i < q(\alpha_n), \forall i = 1, ..., n\right) \\ &= \mathbf{P}\left(\bigcap_{i=1}^{n} \{X_i < q(\alpha_n)\}\right) \\ &= \prod_{i=1}^{n} P\left(X_i < q(\alpha_n)\right) \\ &= F^n\left(q(\alpha_n)\right) \\ &= (1-\alpha_n)^n \\ &= \exp\left(n\log(1-\alpha_n)\right) \\ &= \exp\left(-n\alpha_n(1+o(1))\right) \quad \text{quand } \alpha_n \longrightarrow 0. \end{aligned}$$

Cette probabilité dépend donc du comportement asymptotique de $n\alpha_n$. Estimer les quantiles extrêmes revient à étudier la limite de $n\alpha_n$ lorsque n tend vers l'infini. Ainsi, on distingue trois situations selon la vitesse de convergence de α_n vers 0 :

— **Première situation :** Si $n\alpha_n \to \infty$ alors

$$\mathbb{P}\left(X_{(n)} < q(\alpha_n)\right) \to 0.$$

Dans cette situation, le quantile à estimer se trouve avec une grande probabilité dans l'échantillon disponible. Elle correspond au cas où α_n converge lentement vers 0, autrement dit le quantile à estimer converge lentement vers l'infini lorsque n tend vers l'infini. Dans ce cas estimer le quantile revient à interpoler à l'intérieur de l'échantillon et donc on obtient $X_{(n-\lfloor n\alpha_n \rfloor+1)}$ comme estimateur.

— **Deuxième situation :** Si $n\alpha_n \longrightarrow c \in [1, \infty[$ alors

$$\mathbb{P}\left(X_{(n)} < q(\alpha_n)\right) \longrightarrow e^{-c}.$$

Dans ce deuxième cas, nous avons une faible probabilité que le maximum de l'échantillon soit supérieur au quantile. Cela signifie que l'estimation du quantile extrême repose sur les grandes observations situées au voisinage de la frontière de l'échantillon et toujours dans l'ensemble des données. De ce fait, l'estimateur proposé dans la première situation est celui envisagé.

— **Troisième situation** : Si $n\alpha_n \longrightarrow c \in [0, 1[$ alors

$$\mathbb{P}\left(X_{(n)} < q(\alpha_n)\right) \to e^{-c}.$$

Dans ce cas, le quantile à estimer est supérieur au maximum des observations disponibles. Proposer un estimateur du quantile est impossible avec **la fonction**

de répartition empirique définie :

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq x\}} = \begin{cases} 0 & si\ x < X_{(1)}, \\ \frac{i-1}{n} & si\ X_{(i-1)} \leq x < X_{(i)}, 2 \leq i \leq n \\ 1 & si\ x \geq X_{(n)}. \end{cases} \tag{1.3}$$

Cependant pour estimer le quantile extrême il faut extrapoler au delà du maximum des observations dont on dispose car $F_n(y) = 1$ si $x \geq X_{(n)}$. Ce phénomène peut être expliqué en utilisant le schéma suivant :

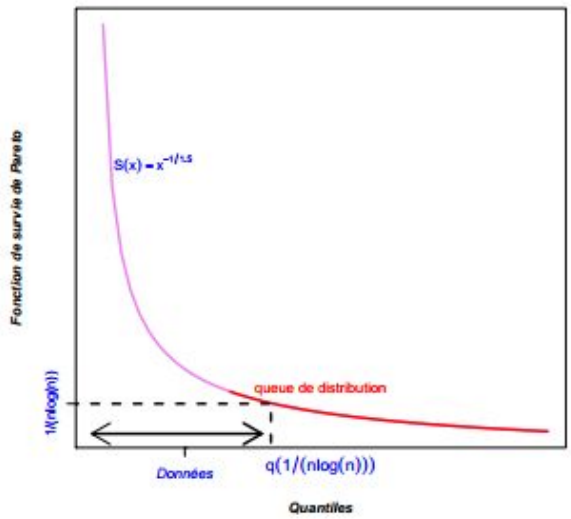

FIGURE 1.1 – Explication graphique de quantiles extrêmes et de queue de distribution.

La figure 1.1 révèle que le quantile d'ordre $1/(n\log(n))$, qui tend strictement vers 0 quand n tend vers l'infini, qu'on veut déterminer, se situe au delà du maximum des observations disponibles. Il faut donc une information sur la forme de la distribution en étudiant la loi du maximum. Remarquons d'abord que

$$\lim_{n\to\infty} F_{X(n)}(y) = \lim_{n\to\infty} [F(y)]^n = 1_{\{y \geq y_F\}} = \begin{cases} 1 & si \quad y \geq y_F \\ 0 & si \quad y < y_F \end{cases} \tag{1.4}$$

où

$$y_F = \sup\{y \in \mathbb{R}, F(y) < 1\}$$

est le point terminal de la loi F. L'équation (1.4) montre que la loi du maximum, $X_{(n)}$ est dégénérée, donc fournit peu d'information.

1.3 Approche des maxima par blocs : lois GEV

1.3.1 Valeurs extrêmes et lois α-stables

La Théorie des Valeurs Extrêmes **(EVT)** repose principalement sur deux résultats. Ces deux résultats nous donnent le comportement asymptotique de la variable aléatoire $X_{(n)}$ ou des dépassements d'un seuil u. L'intérêt de ces résultats provient du fait qu'il n'est pas nécessaire de connaître la loi F du processus X que nous souhaitons prédire. Cependant, l'**EVT** est analogue au Théorème Central Limite **(TCL)** mais pour les phénomènes extrêmes. Ainsi, là où le **TCL** montre que la moyenne empirique de la variable X converge vers une loi normale (indépendamment de la loi de la variable d'intérêt et lorsque les moments d'ordre 1 et 2 existent); l'**EVT** établit des résultats analogues mais pour les valeurs extrêmes de X.

Définition 1.2 *On dit qu'une variable aléatoire X suit une loi **stable** si pour tout $n \geq 2$, il existe un réel strictement positif C_n et réel D_n tel que :*

$$X_1 + X_2 + ... + X_n =^{loi} C_n X + D_n \tag{1.5}$$

Où $X_1, X_2, ..., X_n$ sont n copies indépendantes de X.
Le réel strictement positif C_n est nécessairement de la forme $n^{1/\alpha}$, pour un certaine $\alpha \in]0,2]$, d'où l'appelation loi $\alpha-$stable (Nikias et Shao (1995) [135].

Définition 1.3 *On dit qu'une variable aléatoire X suit une loi $\alpha-$**stable** si et seulement si sa fonction caractéristique est de la forme :*

$$\phi(t) = \exp\{iat - \gamma|t|^\alpha [1 + i\beta sign(t)\omega(t,\alpha)]\} \tag{1.6}$$

Avec

$$\omega(t,\alpha) = \begin{cases} -\tan\frac{a\pi}{2} & si\ \alpha \neq 1 \\ \frac{2}{\pi}\log|t| & si\ \alpha = 1 \end{cases} \tag{1.7}$$

Où $\alpha \in]0,2]$, $\beta \in]-1,1[$, $\gamma > 0$ et $a \in \mathbb{R}$. Les paramètres α, a, γ et β représentent respectivement l'exposant caractéristique de la loi stable, le paramètre de position, le paramètre d'échelle et le paramètre d'asymétrie.

Le théorème suivant montre que les lois stables peuvent être utilisées particulièrement, pour approximer la loi d'une somme de variables aléatoire i.i.d.

Théorème 1.2 *Soit $X_1, ..., X_n$ des variables aléaoires indépendantes, de même loi. Une variable aléatoire X est limite en loi de la suite :*

$$\frac{X_1 + ... + X_n}{a_n} - b_n$$

pour une suite de réels strictement positifs (a_n) *et une suite de réels* (b_n)*, si et seulement si, X suit une loi stable (pour la démonstration de ce théorème, (voir Nolan* 1996*) [136]).*

Les lois $\alpha-$stable présentent un grand intérêt dans la modélisation de nombreux problèmes physiques. La caractéristique de ces lois est leur index de stabilité α qui indique la vitesse de décroissance des queues. Ainsi, les lois peuvent être regroupées selon le comportement de leures queues. La classification suivante a été propsée recemment par El Adlouni et al.(2007) [55] :

— (E) : Les lois exponentielles dont les moments n'existent pas.
— (D) : Les lois subexponentielles .
— (C) : Les lois à variation régulières.
— (B) : Les lois avec un comportement de Pareto .
— (A) : Les lois $\alpha-$stables avec $\alpha < 2$.

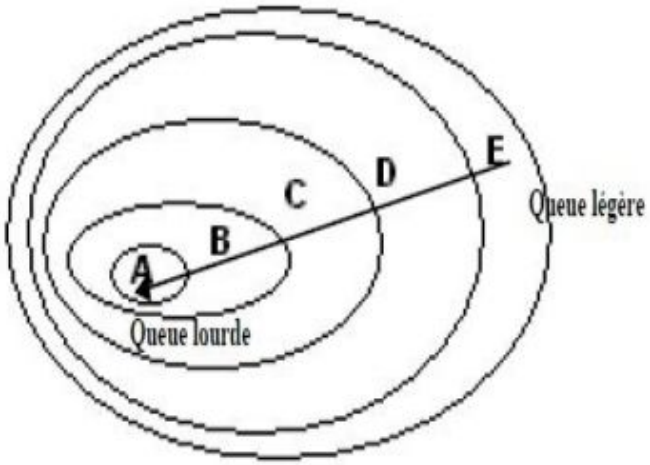

FIGURE 1.2 – Classification des lois selon les queues, (El Adlouni et al.(2007)).

— La classe (E) contient toutes les lois tel que $\mathbb{E}\left[e^X\right] = \infty$. La probabilité au dépassement, $\overline{F} = \Pr(X \geq x) = 1 - F(x)$, pour les extrêmes de cette classe, décroît moins rapidement que celle des lois ayant une queue plus lourde que la loi normale.
— La classe (D) quant à elle, contient les lois tel que $\overline{F}$ décroît plus lentement que n'importe quelle loi exponentielle.
— Pour la classe (C) dite classe des lois à variations régulières, la probabilité au dépassement des extrêmes décroît suivant une fonction puissance (appelée aussi décroissance géomètrique).
— La classe (B) est celle des lois de Pareto.
— Enfin, la classe (A) est celle qui regroupe des lois à différentes asymétries, avec des queues très lourdes.

1.4 Théorème fondamental des valeurs extrêmes

Le résultat de base de la TVE consiste à décrire la loi asymptotique du maximum de n variables indépendantes et identiquement distribuées (i.i.d). La forme de cette loi (notée GEV) dépend d'un seul paramètre qui permet de spécifier le comportement de la queue de la loi considérée. Ce paramètre est l'indice des valeurs extrêmes (noté ξ). En d'autres termes, le résultat de base de la TVE assure que la loi du maximum de n variables (i.i.d) est toujours dans le domaine d'attraction d'une loi GEV_ξ.

Définition 1.4 *Soit G une fonction de répartition définie sur $\mathbb{R}$. Une fonction de répartition F sera dite appartenir au domaine d'attraction $\mathcal{D}(G)$ de G s'il existe deux suites (a_n) $(a_n > 0)$ et (b_n) telles que :*

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \tag{1.8}$$

pour tout point de continuité x, de G.

Dans le cas où F est connue, G peut être identifiée à partir de la loi des statistiques d'ordre. En effet, si nous supposons que les données observées sont des réalisations d'une suite des variables aléatoires $X_1, X_2, ..., X_n$ (i.i.d) de loi F, alors la fonction de répartition de la statistique d'ordre $X_{i:n}$ est donnée par :

$$F_{i:n}(x) = P(X_{i:n} \leq x) = \sum_{k=1}^{n} C_n^k [F(x)]^k [1 - (F(x)]^{n-k} \tag{1.9}$$

En conséquence, les lois de

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

et de

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

notées respectivement F_M et F_W, sont données par :

$$\forall x \in \mathbb{R} : F_M(x) = [F(x)]^n ; F_W(x) = 1 - [1 - F(x)]^n \tag{1.10}$$

Dans la plupart des cas F n'est pas connue et parfois, l'utilisation de la loi des statistiques d'ordre aboutissent à des calculs compliqués. D'où l'utilité d'établir un résultat asymptotique (résultat de base de la TVE), connu sous le nom du théorème de Fisher-Tippett (1928) [61]. Par analogie avec le théorème central limite (TCL), le théorème de Fisher-Tippett établit la convergence en loi de M_n (le théorème est aussi applicable à W_n) :

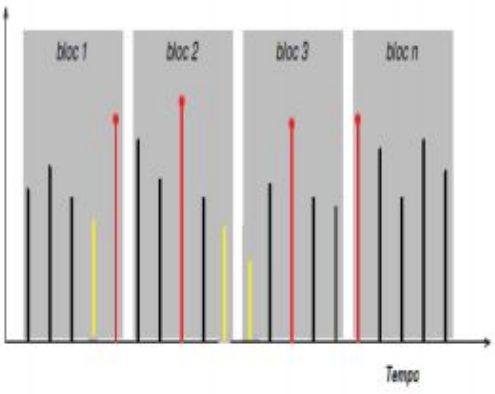

FIGURE 1.3 – Méthode des maxima (minima) par bloc.

Théorème 1.3 (Fisher-Tippett) *Soient $X_1,...,X_i,...,X_n$ n variables aléatoires indépendantes et de même loi de probabilité F. S'il existe des suites $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ et une loi G non dégénérée (c'est à dire différente d'une masse de Dirac) tel que :*

$$\lim_{n\to\infty} \Pr\left(\frac{M_n - b_n}{a_n} \leq x\right) = G(x), \forall x \in \mathbb{R} \tag{1.11}$$

Alors G appartient à l'un des trois types de loi :

— *Loi de Gumbel* ($\xi = 0$) *:*

$$\Lambda_{\mu,\sigma}(x) = \exp\left(-\exp-\frac{x-\mu}{\sigma}\right), x \in \mathbb{R} \tag{1.12}$$

— *Loi de Fréchet* ($\xi > 0$) :

$$\Phi_{\mu,\sigma,\xi}(x) = \begin{cases} \exp-\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}} & si\ x > \mu \\ 0 & si\ x \leq \mu \end{cases} \tag{1.13}$$

— *Loi de Weibull* ($\xi < 0$) :

$$\Psi_{\mu,\sigma,\xi}(x) = \begin{cases} \exp-\left(-\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}} & si\ x < \mu \\ 1 & si\ x \geq \mu \end{cases}. \tag{1.14}$$

Les trois lois de probabilité ci-dessus sont appelées lois des valeurs extrêmes.

Pour la preuve, nous renvoyons à Embrechts et al. (1997), chapitre 3, p.122 [56]. Il faut bien signaler que chacune des trois lois des valeurs extrêmes peut s'obtenir par une transformation fonctionnelle de l'autre. D'une façon similaire, on définit les lois des valeurs extrêmes associées au minimum.

S'il existe des Suites de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$ et une loi non dégénérée H^* tel que :

$$\lim_{n\to\infty} P\left[\frac{W_n - b_n}{a_n} \leq x\right] = \lim_{n\to\infty} (1 - F(a_n x + bn))^n = H^*(x) \tag{1.15}$$

alors H^* est une loi des valeurs extrêmes associée au minimum.

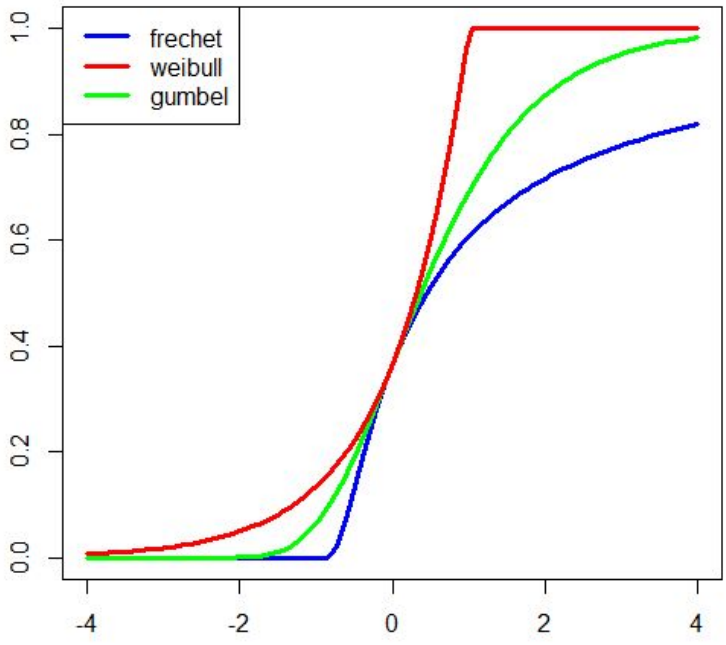

FIGURE 1.4 – Représentation de la fonction de répartition : Gumbel ($\xi = 0$), Fréchet ($\xi = 1$) et Weibull ($\xi = -1$).

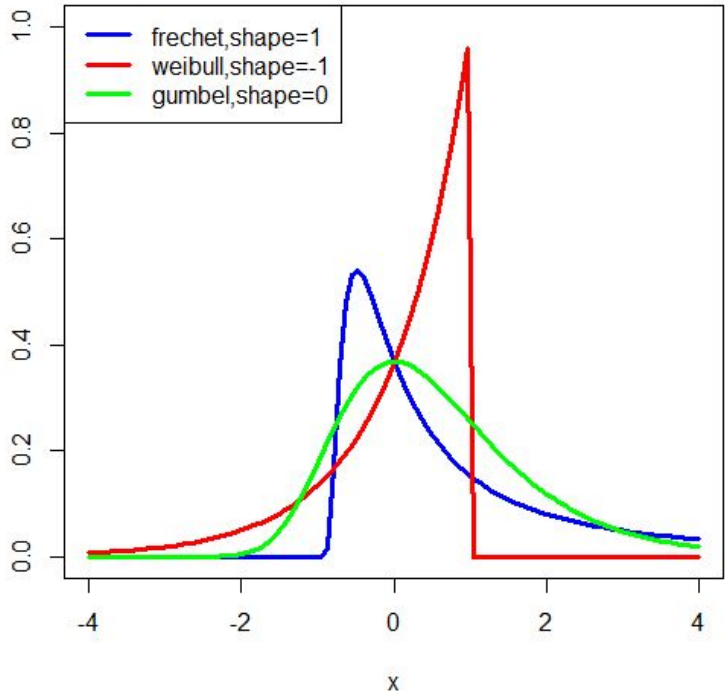

FIGURE 1.5 – Densité des lois des valeurs extremes.

En effet : $\min(X_i) = -\max(-X_i)$. Ainsi, $\forall x \in \mathbb{R}, \forall 1 \leq i \leq n$:

$$P[\min(X_i) > x] = P[\max(-X_i) \leq -x] \Longrightarrow 1 - F_W(x) = F_M(-x) \tag{1.16}$$

Von Mises (1954) [166] et Jenkinson (1955) [124] ont établi un résultat. théorique qui permet d'identifier une famille de lois GEV, qui regroupe les trois types de lois associés aux valeurs extrêmes. Le résultat théorique est formulé par le biais du théorème suivant :

Théorème 1.4 *S'il existe des suites de normalisations $a_n > 0$ et $b_n \in \mathbb{R}$ et une loi G^* tel que :*

$$\lim_{n\to\infty} \Pr\left[\frac{M_n - b_n}{a_n} \leq x\right] = G^*(x); \forall x \in \mathbb{R} \tag{1.17}$$

alors G^* est donnée par :

$$G_{\mu,\sigma,\xi}(x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \tag{1.18}$$

où x est tel que $1 + \xi\frac{(x-\mu)}{\sigma} > 0$, $-\infty < \mu < +\infty$, $\sigma > 0$, et $-\infty < \xi < +\infty$.

Les paramétres μ, σ et ξ sont respectivement, les paramètres de position, d'échelle

et de forme. On note que l'indice $\alpha = 1/\xi$ indice de la queue de la distribution (tail index).

Une démonstration détaillée du théorème 1.4 est donnée dans l'ouvrage de Resnick (1987) [144], avee des développements dans le livre d'Embrechts et al. (1997, p. 152) [56] et des illustrations dans le livre de Galambos (1987, pp.53- 54) [84].

1.5 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 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 H ? Différentes caractérisations des trois domaines d'attraction de Fréchet, Gumbel et de Weibull ont été proposées dans Resnick (1987) [144]; Embrechts et al. (1997) [56]; de Haan et Ferreira (2006) [39]. Ces caractérisations font appel aux classes de fonctions à variation régulière.

Il faut bien noter que le paramètre ξ conditionne le type de la loi des valeurs extrêmes. Nous présentons dans ce qui suit, les domaines d'attractions dans les trois cas correspondant au paramétre ξ.

On commence par définir les fonctions à variations régulières.

1.5.1 Fonctions à variations régulières

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) [14].

Définition 1.5 *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 $x > 0$*

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

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

Remarque 1.1 *Dans le cas particulier où $\alpha = 0$, on dit que g est à variation lente à l'infini, c'est à dire*

$$\lim_{t\to\infty} \frac{g(tx)}{g(t)} = 1. \tag{1.19}$$

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

Criterion 1.5 *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.

Proposition 1.1 *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 :*

$$\forall x > 0, g(x) = x^\alpha \ell(x).$$

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 $g : x \longmapsto |\log x|^\beta$, $\beta \in \mathbb{R}$.
— les fonctions g telles que

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

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

Théorème 1.6 (Représentation de Karamata) *Toute fonction à variation lente ℓ à l'infini s'écrit sous la forme*

$$\ell(x) = c(x) exp\left(\int_a^x \kappa(t) t^{-1} dt\right) \tag{1.20}$$

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

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

Vérifions que toute fonction de la forme (1.20) est à variation lente :

$$\begin{aligned}
\frac{\ell(xt)}{\ell(t)} &= \frac{c(xt)}{c(t)} \frac{\exp\left[\int_a^{xt} \frac{\kappa(u)}{u} du\right]}{\exp\left[\int_a^{t} \frac{\kappa(u)}{u} du\right]} \qquad , \ t \to \infty, \ x > 0 \\
&= \exp\left[\int_a^{xt} \frac{\kappa(u)}{u} du - \int_a^{t} \frac{\kappa(u)}{u} du\right] \\
&= \exp\left[\int_t^{xt} \frac{\kappa(u)}{u} du\right] \\
&\sim \exp\left[\varepsilon \int_t^{xt} \frac{1}{u} du\right] \\
&\sim x^{\varepsilon} \underset{\varepsilon \to 0}{\longrightarrow} 1
\end{aligned}$$

Proposition 1.2 ([42]) *Soit ℓ une fonction à variations lentes. Alors on a pour tout $\rho > 0$, $\ell(x) = o(x^{\rho})$ en $+\infty$ et*

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

Preuve. La démonstration repose sur la formule de représentation (1.20).
Soit $\rho > 0$. Il existe x_0, $M > 0$ tel que, pour $x \geq x_0$, on a $\kappa(x) \leq \rho/2$ et

$$c(x) \exp\left[\int_a^{x_0} \frac{\kappa(u) du}{u} du\right] \leq M.$$

On en déduit que pour $x \geq x_0$, on a

$$\mathcal{L}(x) \leq M \exp\left[\int_{x_0}^{x} \frac{\rho}{2u} du\right] \leq M' \left(\frac{x}{x_0}\right)^{\rho/2}$$

On obtient $\mathcal{L}(x) = o(x^{\rho})$ en $+\infty$.
Soit $u \geq 1$. La fonction $h_x(u) = \left(\frac{\ell(ux)}{\ell(x)} - 1\right) u^{-\rho-1}$ est majorée en valeur absolue par :

$$\left(1 + \frac{c(ux)}{c(x)} \exp\left[\int_x^{ux} \frac{\kappa(v)}{v} dv\right]\right) u^{-\rho-1}.$$

En utilisant les convergences de c et de κ, on en déduit que pour $x \geq x_0$, la fonction

$|h_x(u)|$ est majorée par la fonction

$$g(u) = \left(1 + A\exp\left[\int_x^{ux} \frac{\rho}{2v} dv\right]\right) u^{-\rho-1} \leq A' u^{-1-\rho/2}$$

oú A et A' sont des constantes qui ne dépendent pas de u. La fonction g est intégrable sur $[1, \infty[$. De plus on a

$$\lim_{x\to\infty} \left(\frac{\ell(ux)}{\ell(x)} - 1\right) u^{-\rho-1} = 0,$$

car ℓ est à variations lentes. Par le théorème de convergence dominée, on en déduit que

$$\lim_{x\to\infty} \int_1^{\infty} \left(\frac{\ell(ux)}{\ell(x)} - 1\right) u^{-\rho-1} = 0$$

Ce qui implique que

$$\lim_{x\to\infty} \int_1^{\infty} \frac{\ell(ux)}{\ell(x)} u^{-\rho-1} = \frac{1}{\rho}$$

et en posant le changement de variable $v = ux$,

$$\lim_{x\to\infty} \frac{1}{x^{-\rho}\ell(x)} \int_x^{\infty} v^{-\rho-1}\ell(v)dv = \frac{1}{\rho}$$

On obtient bien la dernière propriété du proposition. □

Proposition 1.3 ([69]) (Conservation de la variation régulière par integration - Théorème de Karamata) *Soit ℓ une fonction à variation lente et localement bornée sur $[x_0, +\infty[$ pour un certain $x_0 \geq 0$. Alors,*

1. *Pour $\alpha > -1$ et quand $x \to \infty$ on a :*

$$\int_{x_0}^{x} t^{\alpha}\ell(t)dt \sim \frac{x^{\alpha+1}}{\alpha+1}\ell(x)$$

2. *$\alpha < -1$ et quand $x \to \infty$ on a :*

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

Proposition 1.4 ([42]) (Inégalité de Potter). *Supposons que g est une fonction positive définie sur un voisinage de l'infni qui est à variation régulière d'indice $\alpha \in \mathbb{R}$, alors pour n'importe quel $0 < \varepsilon < 1$, il existe $t_0 = t_0(\varepsilon)$ tels que pour tout $t \geq t_0$ et $tx \geq t_0$ on a*

$$(1-\varepsilon)x^{\alpha}e^{-\varepsilon|\log x|} \leq \frac{g(tx)}{g(t)} \leq (1+\varepsilon)x^{\alpha}e^{\varepsilon|\log x|} \tag{1.21}$$

La preuve se fait en utilisant la représentation de Karamata d'une fonction à variation

régulière.

1.5.2 Domaine d'attraction de Gumbel $\mathcal{D}(\Lambda)$

La loi présente dans la queue une décroissance de type exponentielle, ce qui permet de caractériser dans ce cas, le domaine d'attraction de Gumbel $\mathcal{D}(\Lambda)$. Ce dernier est a une traitement délicate, car il n'y a pas de lien direct entre la queue de la loi et les fonctions à variations lentes définies par (Delmas et Jourdain (2006) [48]).

Von Mises (1936) [165] a donné une caractérisation simple pour le domaine d'attraction de Gumbel, formulée par le biais du théorème suivant :

Théorème 1.7 *S'il existe une fonction mesurable R, appelée fonction auxiliaire telle que :*

$$\lim_{x\to\omega(F)} \frac{1-F(t+xR(t))}{1-F(t)} = \exp(-x) \tag{1.22}$$

avec

$$\omega(F) = \sup\{x \in \mathbb{R} : F(x) < 1\} \tag{1.23}$$

est le point terminal de F, alors $F \in \mathcal{D}(\Lambda)$.

Dans ce cas les suites a_n et b_n sont ainsi définies :

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

Le tableau suivant fournit quelques exemples de lois qui appartiennent au domaine d'attraction de Gumbel.

Table 1.1 – Lois du domaine d'attraction de Gumbel.

Loi	$1-F(x)$
Benktender $I\,I$, $\alpha,\beta>0$	$x^{-(1-\beta)}\exp(-\frac{\alpha}{\beta}x^{\beta})$
Logistic	$\frac{1}{1+\exp(x)}$

1.5.3 Domaine d'attraction de Weibul $\mathcal{D}(\ominus_\xi)$

Le domaine d'attraction dans ce cas pour $\xi < 0$. Les lois de ce domaine sont bornées à droite, et par conséquent, le point terminal $\omega(F)$ est fini. Une caractérisation d'appartenance à ce domaine d'attraction est donnée par le théorème suivant (pour la démonstration, voir Gnedenko (1943) [88]) :

Théorème 1.8 *Une fonction de répartition F appartient au $\mathcal{D}(\Psi_\xi)$ si et seulement si $\omega(F) < +\infty$ et*

$$\bar{F}\left(\omega(F) - \frac{1}{x}\right) = x^{-\frac{1}{\xi}}\ell(x) \tag{1.24}$$

avec: $\bar{F}$ est la fonction de survie donnée par $\bar{F}(x) = 1 - F(x)$, ℓ est une fonction à variation lente.

Dans ce domaine d'attraction les suites de normalisation sont déterminées comme suit :

$$a_n = \omega(F) - F^{\leftarrow}\left(1 - \frac{1}{n}\right),\ b_n = \omega(F).$$

Le tableau ci-dessous présente quelques lois qui appartiennent au domaine d'attraction de Weibull.

Table 1.2 – Lois du domaine d'attraction de Weibull

Loi	$1 - F(\omega(F) - \frac{1}{x})$	ξ	$l(x)$
Uniforme	$\frac{1}{x}, x > 1$	-1	1
Weibull	$1 - exp(-x^{-\alpha})$	$-\frac{1}{\alpha}$	$1-\frac{x^{-\alpha}}{2} + o(x^{-\alpha})$
Reverse Burr (λ, β, τ)	$(\frac{\beta}{\beta+x^\tau})^\lambda avec\ x > 0$	$-\frac{1}{\lambda\tau}$	$\beta^\lambda(1 - \lambda\beta x^{-\tau} + o(x^{-\tau})$

1.5.4 Domaine d'attraction de Fréchet $\mathcal{D}(\Phi_\xi)$

Ce cas correspond au cas où $\xi > 0$. Les lois appartenant à ce domaine d'attraction sont caractérisées par une queue à décroissance lente (polynomiale) à l'infini, et un point terminal $\omega(F) = +\infty$. Elles sont dites aussi lois à queues lourdes (heavy-tailed). Une caractérisation de ce domaine d'attraction noté $\mathcal{D}(\Phi_\xi)$ est donnée par le théorème suivant (pour la démonstration, voir Gnedenko (1943) [88]) :

Théorème 1.9 *Une fonction de répartition F appartient au $\mathcal{D}(\Phi_\xi)$ si et seulement si sa fonction de survie est donnée par :*

$$\bar{F}(x) = x^{-\frac{1}{\xi}}\ell(x) \tag{1.25}$$

où ℓ est une fonction à variation lente.

Dans ce cas un choix possible des suites de normalisation a_n et b_n du théorème 1.3 est :

$$a_n = F^{\leftarrow}\left(1 - \frac{1}{n}\right) = \bar{F}^{\leftarrow}\left(\frac{1}{n}\right)\ b_n = 0\ et\ \ell \in \mathcal{RV}_0.$$

Le tableau ci-dessous présente quelques lois qui appartiennent au domaine d'attraction de Fréchet .

TABLE 1.3 – Lois du domaine d 'attraction de Fréchet

Loi	$1-F(x)$	ξ	$l(x)$
Pareto(α) avec $\alpha>0$	$x^{-\alpha}$	$\frac{1}{\alpha}$	1
Burr(η,τ,λ) avec $\eta,\tau,\lambda>0$	$(\frac{\eta}{\eta+x^{\tau}})^{\lambda}, x>0$	$\frac{1}{\lambda\tau}$	$(\frac{\eta}{1+\frac{\eta}{x^{\tau}}})^{\lambda}$
Pareto Généralisé(σ,ξ) avec $\sigma,\xi>0$	$(1+\frac{\xi x}{\sigma})^{-\frac{1}{\xi}}, x>0$	ξ	$(\frac{\sigma}{\xi})^{\frac{1}{\xi}}(1+\frac{\sigma}{\xi x})^{-\frac{1}{\xi}}$

1. Une fonction de répartition F peut appartenir à deux domaines d'attraction différents $\mathcal{D}_1$ et $\mathcal{D}_2$ (pour plus de détails, voir Resnick (1987) [144]).
2. Les variables aléatoires dans les trois domaines d'attraction de Gumbel, de Fréchet et de Weibull sont liées par la relation suivante (voir Embrecht et al. (1997) [56]) :

$$X \in \mathcal{D}(\Phi_\alpha) \Longleftrightarrow \log(X^{\frac{1}{\xi}}) \in \mathcal{D}(\Lambda) \Longleftrightarrow -X^{-1} \in \mathcal{D}(\Psi_\xi) \tag{1.26}$$

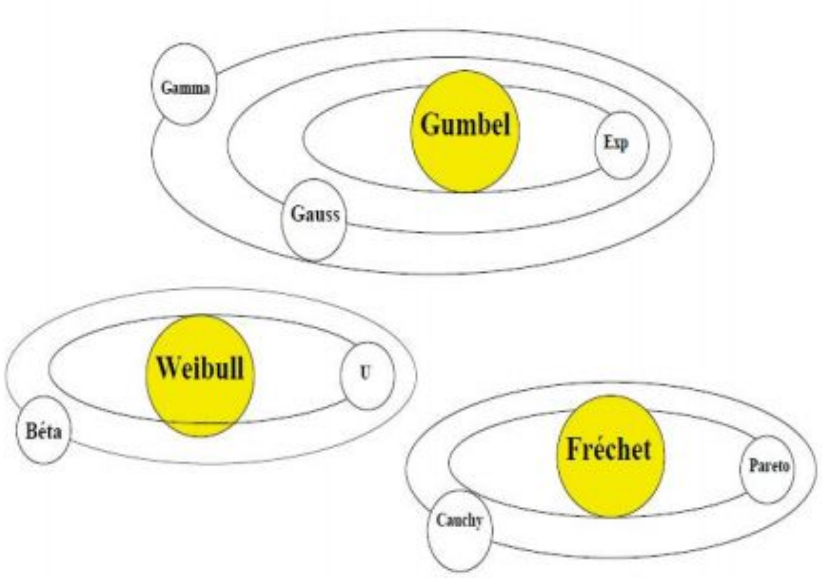

FIGURE 1.6 – Loi du domaine d'attraction $\mathcal{D}(\Lambda)$,$\mathcal{D}(\Phi_\xi)$ et $\mathcal{D}(\Psi_\xi)$.

3. Les trois domaiues d'attraction $\mathcal{D}(\Lambda)$, $\mathcal{D}(\Phi_\xi)$ et $\mathcal{D}(\Psi_\xi)$ ne sont pas restrictifs. Il se peut, qu'une loi n'appartienne à aucun de ces trois domaines d'attraction ; c'est le cas de la loi définie par :

$$F(x) = 1 - \frac{1}{\log(x)},\ x \geq e. \tag{1.27}$$

Outre les domaines d'attraction qui caractérisent. chacuue des lois GEV, les champs d'application conditionnent leurs choix. En effet. la loi de Gumbel par exemple, sert de

référence car elle permet de mesurer l'écart entre les lois empiriques et la loi normale. Elle est généralement utilisée pour prévoir le niveau des crues d'un fleuve. Elle peut aussi servir à prédire la probabilité d'un événement, critique comme un tremblement de terre. En ce qui concerne la loi de Fréchet, elle est couramment évoquée en finance et en macroéconomie, où les variatious sont géneralemeut non bornées. Par ailleurs, cette loi s'applique en hydrologie pour modéliser des évènements extrêmes tels que les débits maximaux des rivières. La loi de Weibull quant à elle, est largement utilisée grâce à sa grande flexibilité. Elle est préconisée eu particulier, dans les analyses de durée de vie ou pour l'estimation des potentiels éoliens.

1.6 Estimation des modéles GEV

1.6.1 Estimation pour la loi de Gumbel

Dans cette partie, nous notons la loi de Gumbel par $G_0(\mu_0, \sigma_0)$, où μ_0 et σ_0 désignent respectivement, les paramètres de position et d'échelle ($\xi = 0$ dans le cas de la loi de Gumbel). La fonction de densité de Gumbel est notée g_0.

1.6.1.1 Méthode de la regréssion linéaire

Soient$X_1, ..., X_n$ les maxima par bloc issus d'une loi de Gumbel et soient $X_{1,m} \leq ... \leq X_{m,m}$ les statistiques d'ordre associées à $X_1, ..., X_m$. La méthode d'estimation par régression linéaire consiste à ajuster le nuage de points :

$$(-\log(-\log(G_0(X_{i,m}))); X_{i,m}) \tag{1.28}$$

à une droite d'équation $X_{i,m} = aX + b$, avec $X = -\log(-\log(G_0(X_{i,m}))$, en calculant les coefficients a et b par la méthode des moindres carrés.

Le coefficient directeur (a) et l'ordonnée à l'origine (b) de la droite ajustée, permettent de calculer les estimateurs $\hat{\mu}_0^{reg}$ et $\hat{\sigma}_0^{reg}$:

$$\begin{cases} \hat{\mu}_0^{reg} = b \\ \hat{\sigma}_0^{reg} = a \end{cases}. \tag{1.29}$$

L'estimation des quantiles extrêmes dépassés avec une faible probabilité p, pour la loi de Gumbel G_0 et pour une période de retour

$$T = \frac{1}{p} = \frac{1}{1 - G_0(y)},$$

est obtenue en remplaçant μ et σ par leurs estimations dans la formule suivante :

$$q_0 = \mu_0 + \sigma_0(-\log(-\log(1-\frac{1}{T}))) \tag{1.30}$$

Ainsi, l'estimateur du quantile extrêmes $\hat{q}_0^{reg}$ est :

$$\hat{q}_0^{reg} = \hat{\mu}_0^{reg} + \hat{\sigma}_0^{reg}(-\log(-\log(1-\frac{1}{T}))). \tag{1.31}$$

1.6.1.2 Méthode du maximum de vraisemblance

Cette méthode consiste à chercher les paramètres $\Theta_0 = (\mu_0, \sigma_0)$ qui maximisent la fonction de vraisemblance $l(\mu_0, \sigma_0) = \prod_{i=1}^{m} g_0(y_i, \mu_0, \sigma_0)$ où log-vraisemblance $L(\mu_0, \sigma_0) = log\ l(\mu_0, \sigma_0)$, en résolvant le systéme suivant :

$$\begin{cases} \frac{\partial \log L(\Theta_0)}{\partial \mu_0} = 0 \\ \frac{\partial \log L(\Theta_0)}{\partial \sigma_0} = 0 \end{cases} \tag{1.32}$$

Cependant, la solution n'est pas explicite et il faut passer par une résolution numérique en utilisant par exemple, la méthode de **Newton-Raphson**.

1.6.1.3 Méthode des moments

La méthode consiste à égaliser respectivement, les deux premiers moments théoriques m_1 et m_2 aux deux premiers moments observés : moyenne ($\bar{y}_m$) et variance S_X^2 empiriques.

Dans le cas d'une variable aléatoire issue d'une loi de Gumbel, la méthode du maximum de vraisemblance abouti au système suivant :

$$\begin{cases} \mathbb{E}[X] = \mu_0 + \gamma^* \sigma_0 = \bar{y}_m \\ \mathbb{V}[X] = \frac{1}{6}\pi^2 \sigma_0^2 = S_X^2 \end{cases} \tag{1.33}$$

où $\gamma^* \approx 0.57721$ représente la constante d'Euler. La résolution de ce systéme d'équations fournit les valeurs des estimateurs $\hat{\mu}_0^{mom}$ et $\hat{\sigma}_0^{mom}$:

$$\begin{cases} \hat{\sigma}_0^{mom} = \frac{\sqrt{6} S_X}{\pi} \\ \hat{\mu}_0^{mom} = \bar{y}_m - \gamma^* \hat{\sigma}_0^{mom} \end{cases} . \tag{1.34}$$

1.6.2 Estimation pour les lois de Fréchet et de Weibull

Pour $\xi \neq 0$, les domaines d'attractions sont ceux de Fréchet et de Weibull. Ces lois seront notées $G_\xi(\mu, \sigma, \xi)$ ($\xi > 0$ pour une loi de Fréchet et $\xi < 0$ pour une loi de Weibull). La fonction de densité sera notée g_ξ.

Nous présentons pour les deux lois de Fréchet et de Weibull, les méthodes d'estimation du maximum de vraisemblance et des moments de probabilités pondérés.

1.6.2.1 Méthode du maximum de vraisemblance

La méthode consiste à chercher les paramètres $\Theta = (\mu, \sigma, \xi)$ qui maximise la fonction de log-vraisemblance. Pour l'obtention de $\hat{\mu}^{mv}, \hat{\sigma}^{mv}$ et $\hat{\xi}^{mv}$, la maximisation conduit à résoudre le système suivant :

$$\begin{cases} \frac{\partial \log L(\Theta)}{\partial \sigma} = 0 \\ \frac{\partial \log L(\Theta)}{\partial \xi} = 0 \\ \frac{\partial \log L(\Theta)}{\partial \mu} = 0 \end{cases} . \tag{1.35}$$

Smith (1985) [156] a étudié le problème de l'estimation par la méthode du maximum de vraisemblance. Il a obtenue les résultats suivant :

— Si $\xi > -0.5$, les estimateurs du maximum de vraisemblance possèdent des propriétés asymptotiques telle que la convergence vers la vraie valeur du paramètre inconnu, l'invariance par rapport à une transformation paramètrique et l'éfficacation asymptotique.
— Si $-1 < \xi < -0.5$, les estimateurs du maximum de vraisemblance ne possèdent pas des propriétés asymptotiques standards.
— Si $\xi < -1$, l'obtention des estimateurs du maximum de vraisemblance n'est pas garantie.

1.6.2.2 Méthode des moments de probabilités pondérés PWM :

Cette méthode a été introduite par Greenwood et al. (1979) [93] dans le but de surmonter les problèmes de convergence, surtout pour les petits échantillons ($n < 25$) (Lubes et al, 1991 [122]).

Les moments des probabilités pondérés sont définis par :

$$M_{p,r,s} = \mathbb{E}[X^p (F(X))^r (1 - F(X))^s] \tag{1.36}$$

où : p, r, s sont des réels et X est une variable aléatoire issue de la loi F.

Dans le cas où $\xi \neq 0$, un choix possible pour calculer les moments $M_{p,r,s}$ est de prendre $p = 1$, $s = 0$ et $r = 0, 1, 2$ (Beirlant et al. (2004) [12]). Ainsi :

$$M_{1,r,0} = \frac{1}{r+1} \left\{ \mu - \frac{\sigma}{\xi} \left[1 - (r+1)^{\xi} \Gamma(1 - \xi) \right] \right\}, \xi > 1 \tag{1.37}$$

Les estimateurs PWM déduit de la méthode des moments de probabilités pondérés sont notés $\hat{\mu}^{pwm}, \hat{\sigma}^{pwm}$ et $\hat{\xi}^{pwm}$. Ils sont obtenue suite à la résolution d'un système d'équations, pour $r = 0, 1, 2$:

$$\begin{cases} M_{1,0,0} = \mu - \frac{\sigma}{\xi}(1-\Gamma(1-\xi)) \\ 2M_{1,1,0} - M_{1,0,0} = \frac{\sigma}{\xi}\Gamma(1-\xi)(2^{\xi}-1) \\ \frac{3M_{1,2,0}-M_{1,0,0}}{2M_{1,1,0}-M_{1,0,0}} = \frac{3^{\xi}-1}{2^{\xi}-1} \end{cases} . \tag{1.38}$$

Dans la pratique, Hosking et al.(1985) [109] préconisent l'utilisation d'un estimateur asymptotiquement consistant, donné par :

$$\hat{M}_{1,r,0} = \frac{1}{m}\sum_{i=1}^{m}(\hat{F}(X_{i,m}))^r X_{j,m} \tag{1.39}$$

Avec :

$$\hat{F}(y_{i,m}) = \frac{i-a}{m}, 0 < a < 1 \tag{1.40}$$

$$\hat{F}(y_{i,m}) = \frac{i-a}{m+1-2a}, -\frac{1}{2} < a < \frac{1}{2} \tag{1.41}$$

1.6.3 Méthodes non paramétriques pour l'estimation de ξ

Une large catégorie d'estimateurs a été élaborée spécifiquement pour le paramètre ξ. Les estimateurs qui font partie de cette catégorie, tel que l'estimateur de Pickands (1975) [137], Hill (1975) [107] et Dekkers-Einmahl- De Haan (1989) [45], sont qualifiés d'estimateurs non paramétriques. Une bonne estimation de ξ peut être obtenue en identifiant une zone de stabilité sur le graphique qui représente différentes valeurs des estimateurs non paramétriques, en fonction du rang k correspondant à une statistique d'ordre $X_{(k)}$ (voir formules des estimateurs).

1.6.3.1 Estimateur de Pickands

L'estimateur de Pickands (1975) [137] combine les quatres statistiques d'ordre. Il est calculé pour un ensemble de rang k. Sa formule est donnée par :

$$\hat{\xi}^{p}_{k,n} = \frac{1}{\ln 2}\ln\left(\frac{X_{(n-k+1,n)} - X_{(n-2k+1,n)}}{X_{(n-2k+1,n)} - X_{(n-4k+1,n)}}\right. \tag{1.42}$$

où $X_1, ..., X_n$ un échantillon de v.a (i.i.d) et n est la taille de l'échantillon observé. L'estimateur de Pickands est un estimateur convergent. Plus précisément :

$$\sqrt{k}(\hat{\xi}^{p}_{k,n} - \xi) \rightarrow \mathcal{N}(0, \sigma^2(\xi)) \tag{1.43}$$

où

$$\sigma^2(\xi) = \frac{\xi^2(2^{2\xi+1}+1}{(2(2\xi-1)\ln 2)^2}. \tag{1.44}$$

1.6.3.2 Estimateur de Hill

L'estimateur de Hill (1975) [107] est l'un des estimateurs les plus répandus dans les applications de la théorie des valeurs extrêmes (TVE). Introduit par Hill, il est défini de la façon suivante (pour tout $\xi > 0$) :

$$\hat{\xi}^H_{k,n} = \frac{1}{k}\sum_{i=n-k+1}^{n} \ln X_{(i,n)} - \ln X_{(n-k+1,n)} \tag{1.45}$$

Le théorème suivant présente les propriétés de l'estimateur de Hill :

Théorème 1.10 *Soit $X_1,...,X_n$ est un échantillon de loi F tel que*

$$\bar{F}(X) = x^{-\frac{1}{\xi}}L(x), x > 0,$$

et L une fonction à variations régulières d'indice 0, alors:

Si $X_1,...,X_n$ un échantillon d'observation i.i.d, $k \to \infty$ et $k/n \to 0$ pour $n \to \infty$, alors l'estimateur de Hill $\hat{\xi}^H$ converge en probabilité vers ξ.

Si $k/n \to 0$, $k/\ln(\ln(n)) \to \infty$ et les $X_1,...,X_n$ sont i.i.d, alors l'estimateur de Hill converge presque sûrement vers ξ.

Si les $X_1,...,X_n$ sont i.i.d et si $\bar{F}$ est à variations régulières, alors :

$$\sqrt{k}(\hat{\xi}^H_{k,n} - \xi) \to \mathcal{N}(0, \xi^2). \tag{1.46}$$

Pour la preuve de ce théorème, voir Hill (1975) [107].

1.6.3.3 Estimateur des moments de Dekkers-Einmahl-De Haan

L'estimateur de Dekkers-Einmahl-De Haan présente une extension de l'estimateur de Hill pour $\xi \in \mathbb{R}$. Il est défini comme suit :

$$\hat{\xi}^D_{k,n} = 1 + H^{(1)}_n + \frac{1}{2}\left(\frac{(H^{(1)})^2}{H^{(2)}_n} - 1\right)^{-1}, \tag{1.47}$$

où :

$$H^{(1)}_n = \frac{1}{k}\sum_{i=n-k+1}^{n} \ln X_{(i,n)} - \ln X_{(n-k,n)} \tag{1.48}$$

et

$$H_n^{(2)} = \frac{1}{k} \sum_{i=n-k+1}^{n} (\ln X_{(i,n)} - \ln X_{(n-k,n)})^2 \tag{1.49}$$

Sous certaines conditions sur k, l'estimateur de Dekkers-Einmahl-De Haan converge asymptotiquement vers la loi normale.

1.7 Approche du loi GPD

1.7.1 Modélisation des excés

Alternativement à l'approche des maxima par bloc, le modéle de renouvellement pour des v.a (i.i.d), a donné naissance à un autre volet de la théorie des valeures extrêmes, illustré formellement par l'approche des dépassements de seuil (Peaks over Threshold (POT)). Contrairement à l'approche du maxima par bloc, la méthode POT consiste à utiliser toutes les observations, appeleés excès, qui dépassent un certain seuil suffisament élevé. L'objectif est d'analyser leur comportement asymptotique.

L'approche POT a été introduite pour remédier aux défauts de l'approche classique de la TVE. Les critiques sont essentiellements dues au fait que la méthode d'estimation de la loi GEV se base sur les "block component -wise ", ce qui implique systématiquement une perte d'information. En plus, certain blocs peuvent contenir plusieurs valeurs extrêmes issues de la loi initiale, alors que d'autres peuvent ne pas en contenir (Mc Neil et Frey (2002) [127], Katz (2002) [113]).

Historiquement, la méhode POT introduite par Pickands (1975) [137] et reprise par de Haan et Rootzen (1993) [40], a été initialement appliqué pour le traitement des données hydrologiques, notamment le volume et la durée des déficits en eau. Quant aux volets théoriques de la méthode, ils ont été abondamment développés par divers auteurs tels que Todorovic et Zelenhasie (1970) [163], Todorovic et Rousselle (1971) [164], Smith (1987) [155], Davison et Smith (1990) [34], Reiss et Thomas (2001) [143]. Le point de départ de l'approche POT consiste à choisir un seuil $u \in \mathbb{R}$ et de considérer les variables X_j définies, à partir des variables initiales $X_1, ..., X_n$, par :

$$\begin{cases} N_u = card\{i : i = 1, ..., n \mid X_i > u\} \\ X_j = X_i - u > 0, \qquad pour \;\; 1 \leq j \leq N_u \end{cases} \tag{1.50}$$

où : $X_1, ..., X_n$ un échantillon de v.a (i.i.d), N_u est le nombre des dépassements du seuil u et $Y_1, ..., Y_{N_u}$ les excés correspondants.

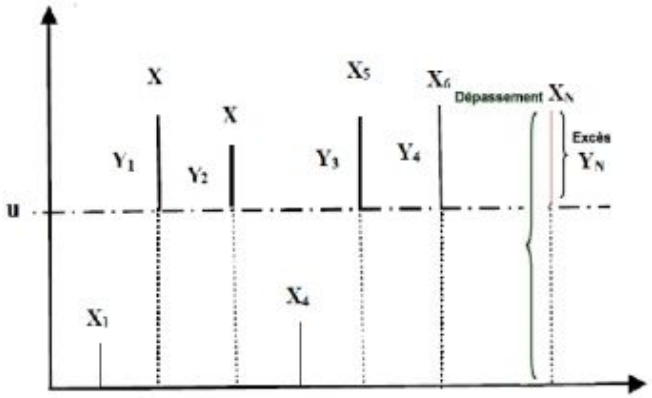

FIGURE 1.7 – Reprsentation des excs Y issue des dpassements X au-del d'un seuil u.

Le nombre de dépassements d'un seuil u suit une loi de Poisson de paramètre τ, notée $P(\tau)$. Ce résultat est justifié sur la base de l'approximation de la loi binomiale par la loi de Poisson. En effet, si on désigne par $\{X_i > u_n\}$ l'événement " Dépassement d'un seuil u_n " et sa fonction indicatrice $I_{\{X_i>u_n\}}$ définie par :

$$I_{\{X_i>u_n\}} = \begin{cases} 1, & X_i > u_n \\ 0, & \text{sinon} \end{cases} \tag{1.51}$$

alors le nombre de dépassements de u_n, égal à $N_{u_n} = \sum_{i=1}^{n} I_{\{X_i>u_n\}}$, suit une loi binomiale $Bin(n, 1 - F(u_n))$ qui peut être approchée par une loi de Poisson lorsque $n \to +\infty$, $1 - F(u_n) \to 0$ et $n(1 - F(u_n)) \to \tau$. Une fois les excés identifiés, l'approche POT consiste, en partant de la loi F de X , à déterminer une loi conditionnelle F_u, pour les variables aléatoires Y qui dépassent le seuil u.

La loi conditionnelle F_u est définie par :

$$F_u(y) = \mathbb{P}(X - u \leq y \mid X > u) = \begin{cases} \frac{F(u+y)-F(u)}{1-F(u)}, & si\ y \geq 0 \\ 0, & si\ y < 0 \end{cases} \tag{1.52}$$

La loi asymptotique associée à $F_u(y)$ est donnée par le théorème suivant :

Théorème 1.11 *(**Pickands, Balkema et de Haan**) Si F (la loi de X) appartient à l'un des trois domaines d'attraction : $\mathcal{D}$(Fréchet), $\mathcal{D}$(Gumbel) ou $\mathcal{D}$(Weibull), alors il existe une fonction $\sigma(u)$ positive, tel que :*

$$\lim_{u \to \omega(F)} \sup_{0<y<\omega(F)-u} | F_u(y) - H_{\xi,\sigma(u)(y)} | = 0, \tag{1.53}$$

où

$$\omega(F) = \sup\{x \in \mathbb{R} : F(x) < 1\} \tag{1.54}$$

est le point terminal de F et $H_{\xi,\sigma(u)(y)}$ est la fonction de répartition de la loi de Pareto Généralisée (GPD, Generalized Pareto Distribution), définie par :

$$H_{\xi,\sigma(u)}(y) = \begin{cases} 1-(1+\frac{\xi y}{\sigma(u)})^{-\frac{1}{\xi}}, & si\ \xi \neq 0 \\ 1-\exp(-\frac{y}{\sigma(u)}), & \xi = 0 \end{cases} . \tag{1.55}$$

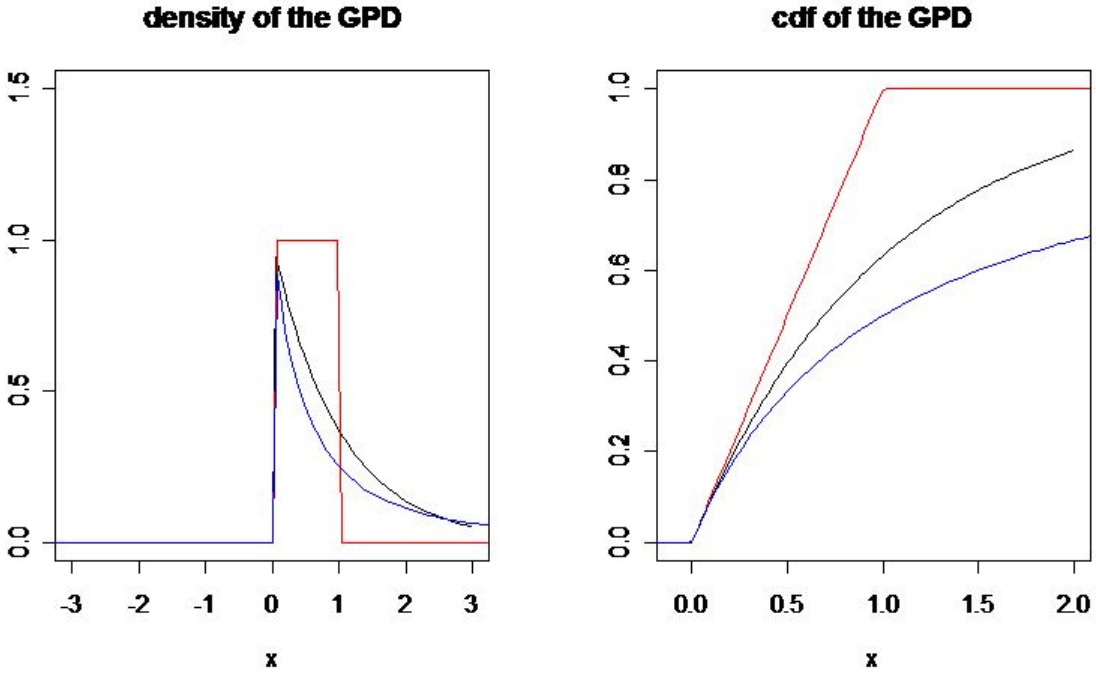

FIGURE 1.8 – Densité et fonction de répartition de la loi de Pareto Généralisée

Pour la démonstration de ce théorème, voir Pickands (1975) [137].

Nous pouvons estimer la période de retour par la relation suivante (Rosbjerg (1987)) :

$$T_u = \frac{1}{\tau p} = \frac{1}{\tau(1-H_{\xi,\sigma(u)}(y))} \tag{1.56}$$

Ainsi, si on s'intéresse particulièrement à un niveau de dépassement, une fois chaque N années, nous pouvons utiliser l'équation 1.56 pour déduire que ce niveau de dépassement est donné par :

$$Z_n = \begin{cases} \mu + \frac{\sigma}{\xi}\left[(N n_y \zeta_u)^{\xi} - 1\right], & si\ \xi \neq 0; \\ \mu + \sigma \log(N n_y \zeta_u), & \xi = 0. \end{cases} \tag{1.57}$$

où $\zeta_u = \mathbb{P}\{X > u\}$ et n_y est le nombre de dépassements.

Il faut bien signaler que les deux approches ; blocs des maxima et dépassements d'un seuil, sont équivalentes. Ainsi, pour un seuil adéquat, la loi des excès peut être approchée par une loi GPD d'indice extrême ξ, identique à celui de la loi GEV. En

effet :

$$\mathbb{P}\{X > x \mid X > u\} = \left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \mathbb{P}\{X > x\} = \zeta_u \left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}. \tag{1.58}$$

La fonction de répartition ($H_{\xi,\sigma(u)}$) met en exergue trois paramètres : μ, σ et ξ.

— Le paramètre ξ, appelé indice des valeurs extrêmes ou indice de queue, renseigne sur le type de la loi asymptotique (Fréchet pour $\xi > 0$, Gumbel pour $\xi = 0$ et Weibull pour $\xi < 0$). En outre, plus ξ est élevé en valeur absolue, plus le poids des extrêmes est important. Nous parlons dans ce cas d'une loi à queue épaisse.

L'identification du seuil est la phase la plus déterminante dans l'implémentation de l'approche POT, vu que la qualité du modèle en dépend. C'est à partir de la définition d'un bon seuil u que nous pouvons garantir la convergence des excès vers une GPD, estimer les autres paramètres et évaluer le temps de retour.

D'une manière similaire à la détermination des blocs dans la première approche de la TVE, le choix du seuil doit établir un compromis entre biais et variance. Concrètement, le seuil doit être suffisamment grand pour pouvoir utiliser les résultats asymptotiques, mais pas trop élevé afin d'obtenir des estimations précises. Par contre, le choix d'un seuil faible risque de déclarer abusivement des observations extrêmes, introduire un biais dans l'estimation et par conséquent, mal approximer la loi asymptotique. Dans ce sens, plusieurs méthodes de détection du seuil ont été proposées. Certaines ont le défaut d'être subjectives, et donc elles ne doivent être utilisées que pour proposer un intervalle de seuils candidats. D'autres par contre, ont l'avantage d'être objectives.

1.7.2 Sélection du seuil

L'estimation des paramètres de la distribution GPD pose le problème de la détermination du seuil u, le seuil ne doit pas être trop grand car il faut suffisamment de données pour avoir une bonne précision des estimateurs. La technique graphique est la plus utilisée pour estimer u.

1.7.2.1 Le Mean Excess Plot

Etant donnée un échantillon aléatoire $X_1, X_2, ..., X_n$ d'une distribution F et soit u un seuil donné. Le Mean Excess Plot est le graphe des points $(u, e(u))$ où $e(u)$ est la moyenne des excès au delà du seuil u, appelée aussi la Mean Excess Function définie par :

$$e(u) = E\left[X - u \mid X > u\right],\ 0 < u < \infty. \tag{1.59}$$

Elle correspond à une fonction de u que l'on peut exprimer à l'aide de $1-F$. On suppose que pour le modèle proposé, $E(X) < \infty$. En pratique, la fonction des excès moyens $e(u)$ est estimée par $\hat{e}_n(u)$:

$$\hat{e}_n(u) = \frac{\sum_{i=1}^{n}(x_i - u)1_{(x_i>u)}(x_i)}{\sum_{i=1}^{n} 1_{(x_i>u)}(x_i)}. \tag{1.60}$$

Un des principaux outils du choix du seuil est le graphe de la sample mean excess function $\hat{e}_n(u)$ ou le mean excess plot.

Définition 1.6 *Le mean excess plot ou ME-plot est défini par le nuage de points suivant :*

$$\{(u, \hat{e}_n(u)), x_{1,n} < u < x_{n,n}\},$$

où $x_{n,n}$ et $x_{1,n}$ sont respectivement le maximum et le minimum de l'échantillon.

Comment interprète-t-on ce graphe ? Il faut remarquer tout d'abord que la mean excess function d'une $GPD(\sigma, \xi)$, pour $\xi < 1$, est

$$E[X - u \mid X > u] = \frac{\xi}{1-\xi}u + \frac{\sigma}{1-\xi}. \tag{1.61}$$

Dans ce cas, le seuil à retenir u, est celui pour lequel la moyenne des excès est approximativement linéaire. Plus les queues des distributions sont épaisses, plus la fonction $e(u)$ tend rapidement vers l'infini. En effet, la fonction moyenne des excès relative à une distribution de loi Pareto de paramètre $\alpha > 0$, appartenant au domaine d'attraction de Fréchet d'indice des valeurs extrêmes $\xi = \frac{1}{\alpha}$, s'écrit comme suit :

$$e(u) = \frac{k+u}{\alpha - 1}. \tag{1.62}$$

Dans la pratique, si la ME-plot semble avoir un comportement linéaire au-dessus d'une certaine valeur de u, cela signifie que les excès au-dessus de ce seuil suivent une GPD. Cette valeur correspond à un arbitrage acceptable entre le souhait d'avoir un seuil élevé et celui d'obtenir un échantillon d'excès de taille suffisante.

1.7.2.2 Seuil aléatoire

McNeil et Frey choisissent un seuil aléatoire : $N_u = k$ (Il y a k observations excédentaires) où $k << N$ et le seuil est donc la $(k+1)^{ième}$ statistique d'ordre.

Soient $z_{1,n} \leq z_{2,n} \leq ... \leq z_{n,n}$, les excès ordonnés. La loi Pareto généralisée de paramètres ξ et σ est ajustée aux données :

$$(z_{1,n} - z_{k+1,n}, ..., z_{k,n} - z_{k+1,n}).$$

Ces auteurs font une étude simulatoire à partir d'une distribution t de Student pour déterminer une valeur de k appropriée.

1.7.2.3 Bootstrap

Caers, Beirlant et Maes utilisent une méthode basée sur le bootstrap semi-paramétrique afin de déterminer l'erreur moyenne carrée (**MSE**). Ceci permet de choisir le seuil qui donne la plus petite MSE. La méthodologie est la suivante :

1. Un seuil u est fixé pour lequel les paramètres de la Pareto généralisée, sont estimés par $\hat{\xi}_u$, $\hat{\sigma}_u$ avec la méthode de notre choix (par exemple, avec le maximum de vraisemblance). Ceci nous permet d'utiliser la densité semi-paramétrique suivante :

$$\hat{F}_s(x\backslash u) = \begin{cases} (1-\hat{F}(u))\left(1-\left(1+\frac{\hat{\xi}_u}{\hat{\sigma}_u}(x-u)\right)\right)^{\frac{-1}{\hat{\xi}_u}} + \hat{F}(u), & x > u, \\ \hat{F}(x), & x \leq u. \end{cases} . \tag{1.63}$$

2. Cette densité estimée permet de faire du bootstrap semi-paramétrique. Donc B échantillons sont tirés à partir $\hat{F}_s(x\backslash u)$:

$$X^{(b)} = \{X_1^{(b)}, ..., X_n^{(b)}\}, \tag{1.64}$$

   où $b = 1, ..., B$.

3. À partir de ces échantillons, un paramètre (par exemple l'indice de queue ξ) est estimé. De ces B estimés, le biais et la variance de l'estimateur sont calculés et combinés pour former un estimé de la **MSE**.
4. Le seuil ayant donné la plus petite **MSE** pour l'estimation du paramètre qui nous intéresse est retenu.

Il n'est pas clair que cette méthode produise un estimé raisonnable du biais du choix de seuil u, le biais correspondrait plutôt à la méthode d'estimation du paramètre d'intérêt (par exemple ξ) pour un seuil fixé.

L'estimation de ξ et σ pose le problème de la détermination du seuil u. Il doit être suffisamment grand pour que l'on puisse appliquer le résultat précèdent, mais ne doit pas être trop grand afin d'avoir suffisamment de données pour obtenir des estimateurs de bonne qualité.

1.8 Estimation du modèle GPD

La littérature propose une multitude de méthodes, consacrées principalement à l'estimation des paramètres des lois GEV. Nous pouvons citer par exemple, les mé-

thodes d'estimation empirique (Gumbel et Mustafi (1967) [99] ou Tiago de Oliveira (1975) [160]), la méthode du maximum de vraisemblance (Smith (1987 [155] et 1985 [156]), Prescott et Walden (1980 et 1983)([140], [139]),Tiago de Oliveira (1982) [161], Hougaard (1986) [108]), la méthode des moments (Christopeit, (1994) [23]), la méthode des moments de probabilité pondérés (Greenwood et al., (1979) [93]) ou encore, les méthodes bayésiennes (Lye et al., (1993) [123]).

La comparaison de ces méthodes d'estimation s'effectue dans la plupart du temps dans un cadre empirique (voir Benkhaled (2007) [13] et Lang et al. (2012) [118]).

Par ailleurs, d'autres approches principalement non paramétriques, ont été dédiées à l'estimation de l'indice de queue. Nous pouvons citer à titre d'exemple l'estimateur de Pickands (1975) [137], l'estimateur de Hill (1975) [107] (pour la loi de Fréchet uniquement) et l'estimateur de Dekkers-Einmahl-de Hann (Dekkers et al. (1989) [45]).

1.8.1 Méthode des moments

La Méthode consiste à égaliser les moments théoriques et les moments empiriques de façon à obtenir :

$$\begin{cases} \mathbb{E}(Y) = \frac{\sigma_u}{1-\xi} = \bar{Y} \\ \mathbb{V}(Y) = \frac{\sigma_u^2}{(1-\xi)^2(1-2\xi)} = S_Y^{*2} \end{cases}, \tag{1.65}$$

où $\bar{Y}$ et S_Y^{*2} représentent respectivement, la moyenne et la variance empirique des excès.

Deux méthodes d'estimation sont ici encore réalisables : l'Estimation par Maximum de Vraisemblance (EMV) et celle par les Moments Pondérés (EMP).

1.8.2 Méthode du maximum de vraisemblance (EMV)

Supposons que notre échantillon des excés $Y = (Y_1, ..., Y_{N_u})$ est i.i.d avec comme fonction de distribution la GPD. La fonction de densité $g_{\xi,\sigma}$ de GPD $G_{\xi,\sigma}$ est

$$g_{\xi,\sigma}(y) = \begin{cases} \frac{1}{\sigma}\left(1+\xi\frac{y}{\sigma}\right)^{-\frac{1}{\xi}-1} & si\ \xi \neq 0 \\ \exp\left(-\frac{y}{\sigma}\right) & si\ \xi = 0 \end{cases}, \ \sigma > 0. \tag{1.66}$$

La fonction log-vraisemblance est donc égale à

$$l((\xi,\sigma);X) = -N_u \ln\sigma - \left(\frac{1}{\sigma}+1\right)\sum_{i=1}^{N_u} \ln\left(1+\frac{\xi y_i}{\sigma}\right), \tag{1.67}$$

prendre des dérivées partielles de $l((\xi,\sigma);X)$ par rapport à ξ et σ, on obtient :

$$\begin{cases} \frac{1}{\sigma^2}\sum_{j=1}^{N_u}\ln\left(1+\frac{\xi y_i}{\sigma}\right)-\sum_{i=1}^{N_u}\frac{y_i}{\sigma+\xi x_i}=0 \\ -N_u+(1+\xi)\sum_{i=1}^{N_u}\frac{y_i}{\sigma+\xi y_i}=0 \end{cases}, \tag{1.68}$$

où $y_1,...,y_{N_u}$ est une réalisation de $Y_1,...,Y_{N_u}$. EMV $(\hat{\xi}_{N_u},\hat{\sigma}_{N_u})$ de (ξ,σ) comme une solution de se système d'équations. Notez que ce système n'a pas de solution explicite et que des méthodes numériques sont donc nécessaires pour calculer les valeurs estimées.

Smith 1987 [155] montre la normalité asymptotique de $(\hat{\xi}_{N_u},\hat{\sigma}_{N_u})$ fournie $\xi > -\frac{1}{2}$ plus précisément, nous avons

$$\sqrt{N_u}\begin{pmatrix}\hat{\xi}_{N_u}-\xi\\ \hat{\sigma}_{N_u}-\sigma\end{pmatrix}\xrightarrow{d}\mathcal{N}_2(0,\mathbb{Q}^{-1}) \text{ comme } N_u\to\infty, \tag{1.69}$$

où

$$\mathbb{Q}^{-1}=(1+\xi)\begin{pmatrix}1+\xi & -1\\ -1 & 2\end{pmatrix}, \tag{1.70}$$

où $\mathcal{N}_2(\varepsilon,\Sigma)$ représente la distribution normale bivariée avec le vecteur moyen ε et la matrice de covariance Σ, avec ce résultat, les intervalles de confiance pour les estimations de paramètres sont facilement construits.

1.8.3 Méthode des moments pondérés(EMP)

La définition des moments de probabilité pondérés pour une GPD, dans le cas $p = 1$, $r = 0$, $s = 0$ ou $s = 1$, est donnée par :

$$M_{1,0,s}=\frac{\sigma_u}{(s+1)(s+1-\xi)},\ \xi<1, \tag{1.71}$$

avec :

$$\hat{M}_{1,0,s}=\frac{1}{k}\sum_{i=1}^{k}\left(1-\frac{i}{k+1}\right)Y_{i,k} \tag{1.72}$$

La résolution de l'équation définissant $M_{1,0,s}$ en fonction de σ_u et ξ, conduit pour $s = 0$ et $s = 1$ aux estimateurs PWM suivants :

$$\hat{\xi}_u^{pwm}=2-\frac{\hat{M}_{1,0,0}}{\hat{M}_{1,0,0}-2\hat{M}_{1,0,1}}, \tag{1.73}$$

$$\hat{\sigma}_u^{pwm}=2\frac{\hat{M}_{1,0,0}\hat{M}_{1,0,1}}{\hat{M}_{1,0,0}-2\hat{M}_{1,0,1}}. \tag{1.74}$$

1.9 Estimation de la queue de la distribution

La distribution de Pareto généralisée est utilisée pour modéliser la queue (supérieure ou inférieure) d'une distribution présentant des valeurs extrêmes. Pour pouvoir faire de l'estimation des quantiles, il faut avoir une formulation qui combine la distribution estimée dans la queue et la distribution centrale. Une telle formulation est donnée par l'égalité suivante, $\forall u < x < x_F$

$$\bar{F}(x) = \bar{F}_u(x-u)\bar{F}(u), u < x < x_F. \tag{1.75}$$

où

$$F_u(y) = P(X - u \leq y \mid X > u). \tag{1.76}$$

Autrement dit, $\forall x > u$:

$$P(X > x) = P(X > u)P(X > x \mid X > u). \tag{1.77}$$

L'estimation se fait de la façon suivante : $\bar{F}(x)$ est estimée avec la probabilité de dépassement empirique N_u/n.

$$\widehat{\bar{F}}(u) = \bar{F}_n(u) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X>u\}} = \frac{N_u}{n},\ u < x_F. \tag{1.78}$$

La queue conditionnelle F_u de F est estimée par

$$\widehat{\bar{F}}(x-u) = 1 - G_{\hat{\xi},\hat{\sigma}}(x-u) = \left(1 + \hat{\xi}_u \frac{x-u}{\hat{\sigma}_u}\right)^{\frac{-1}{\hat{\xi}_u}}, u < x < x_F. \tag{1.79}$$

L'estimateur de la queue de la distribution est donc :

$$\widehat{\bar{F}}(u) = \frac{N_u}{n}\left(1 + \hat{\xi}_u \frac{x-u}{\hat{\sigma}_u}\right)^{\frac{-1}{\hat{\xi}_u}}, u < x < x_F. \tag{1.80}$$

1.10 Estimation des quantiles

1.10.1 Quantile d'ordre p

Définition 1.7 *Soit X une variable aléatoire d'une distribution F. On suppose que F est continue, Le quantile d'ordre p vérifie :*

$$F\left(x_p\right) = P\left(X < x_p\right) = p,\ où\ \ p \in [0,1]. \tag{1.81}$$

x_p : est un quantile d'ordre p. Par définition le quantile d ordre p c'est la fonction inverse de F c'est-à-dire $x_p = F^{-1}(p)$.

1.10.2 Fonction des quantiles

Si on a F est une fonction continue et monotone alors F est bjective donc F^{-1} existe.

Définition 1.8 *Soit X une variable aléatoire et $X_1, X_2, ..., X_n$ un échantillon de X .*

$$Q(p) = F^{-1}(p) = \inf\{x; F(x) \geq p\}, \text{ avec } \quad p \in [0,1]. \tag{1.82}$$

"la fonction des quantiles s'appelle la fonction inverse généralisée".

1.10.3 Estimation empirique

L'estimation quantile joue un rôle important dans le contexte de la gestion des risques, où il est crucial d'évaluer correctement le risque d'une grosse perte qui survient très rarement. La principale difficulté de cette estimation est due au fait que lorsque p est très petit, le point x_p est au-delà de la plage de l'échantillon $(X_1, ..., X_n)$ tiré d'un cdf F inconnu.

Comme nous utilisons la théorie asymptotique, p doit dépendre de la taille de l'échantillon n, c'est-à-dire, $p = p_n$.

Deux cas sont possibles pour x, à l'intérieur et à l'extérieur de l'échantillon.

— si $p_n \to 0$ avec $np_n \to c \in [1, \infty]$ comme $n \to \infty$, le $(1-p)$-quantile est dans l'échantillon,

— si $p_n \to 0$ avec $np_n \to c \in [0,1]$ comme $n \to \infty$, le $(1-p)$-quantile est en dehors de l'échantillon.

En d'autres termes, l'estimation intra-échantillon est possible jusqu'au quantième $(1/n)$ alors que, pour $p < 1/n$, les estimations quantiles sont au-delà de la plage des données. Ce dernier cas est le plus pertinent pour les applications réelles. Pour la première situation, nous avons $\mathbb{Q}_n(s) = X_{n-i+1,n}$, puis avec $s = 1 - p = 1 - (i-1)/n$ pour $i = 2, ..., n$, on a

$$\mathbb{Q}_n\left(1 - \frac{(i-1)}{n}\right) = X_{n-i+1,n}, i = 2, ..., n. \tag{1.83}$$

Ainsi, $X_{n-i+1,n}$ semble être un estimateur naturel pour le $(1 - \frac{(i-1)}{n})$-quantile.

Dans le second cas, nous devons déduire au-delà des limites de l'échantillon en extrapolant à partir des quantiles intermédiaires. Évidemment, cela ne peut pas être fait sans une sorte d'information sur les queues est alors nécessaire.

1.10.4 Intervalles de confiance des quantiles empirique

Proposition 1.5 (Intervalles de confiance 1) *Soit $p \in]0,1[$, supposons que la loi de X_1 possède une densité f continue en x_p et telle que $f\left(x_p\right) > 0$,*

on suppose de plus que $k(n) = np + o\left(\sqrt{n}\right)$ c-à-d

$$\lim_{n\to\infty} \frac{k(n)}{n} = p. \tag{1.84}$$

Alors, on a la convergence en loi suivante :

$$\frac{\sqrt{n}\left(X_{(k(n),n)} - x_p\right)\left(f\left(x_p\right)\right)}{\sqrt{p(1-p)}} \overset{n\to\infty}{\to} \mathcal{N}(0,1). \tag{1.85}$$

Alors, soit $\alpha > 0$, pour calculer l'intervalle de confiance $1-\alpha$ on procède de la façon suivante, on veut :

$$1 - \alpha = P\left(|\mathcal{N}(0,1)|\right). \tag{1.86}$$

Donc l'intervalle de confiance de x_p, de niveau asymptotique $1-\alpha$ est l'intervalle aléatoire :

$$\left[X_{(k(n),n)} \pm \delta \frac{\sqrt{p(1-p)}}{f\left(X_{(k(n),n)}\sqrt{n}\right)}\right]. \tag{1.87}$$

ou δ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0,1)$.

Proposition 1.6 (Intervalles de confiance 2) *Soit $p \in]0,1[$, soit δ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0,1)$. On considère par la suite que, les entiers $i_n = \left[np - \sqrt{n}\delta\sqrt{p(1-p)}\right]$ et $j_n = \left[np + \sqrt{n}\delta\sqrt{p(1-p)}\right]$. Pour n assez grand les entiers i_n et j_n sont compris entre 1 et n. De plus l'intervalle alèatoire :*

$$\left[X_{(i(n),n)}, X_{(j(n),n)}\right].$$

est un intervalle de confiance pour x_p de niveau asymptotique $1-\alpha$.

En d'autres termes, une bonne estimation de l'indice de queue est essentielle au processus d'estimation quantile extrême.

Puisque l'estimation des quantiles élevés est directement liée à l'estimation de l'IVE, on s'attendrait à trouver, dans la littérature, autant d'estimateurs quantiles que d'estimateurs de l'indice de queue.

1.10.5 Approche POT

La méthode POT s'appuie sur le théorème de Balkema-de Haan-Pickands pour estimer x_p et cet estimateur est obtenu en inversant $\widehat{\bar{F}}(x)$ dans l'équation (1.80).

L'estimateur obtenu par cette méthode s'écrit sous la forme :

$$\hat{x}_p = u + \frac{\hat{\sigma}_u}{\hat{\xi}_u}\left[\left(\frac{N_u}{np}\right)^{\hat{\xi}} - 1\right]. \tag{1.88}$$

où N_u désigne le nombre d'excès au-delà du seuil u, $\hat{\sigma}_n$ et $\hat{\xi}_n$ sont des estimateurs des paramètres de la loi GPD.

Quand $\xi > 0$, le point terminal est fini et il est estimé par :

$$\hat{x}_F = u - \frac{\hat{\sigma}_u}{\hat{\xi}_u}. \tag{1.89}$$

En pratique, u est choisi égal à l'une des statistiques d'ordre et en prenant $u = X_{n-k,n}$ la $(k+1)^{ieme}$ observation, donne $N_u = k$.

Les estimateurs résultants des paramètres ξ et σ sont respectivement dénotés par $\hat{\xi}^{(POT)}$ et $\hat{\sigma}^{(POT)}$. Dans ce cas, l'estimateur des quantiles est de la forme suivante :

$$\hat{x}_p^{(POT)} = X_{n-k,n} + \frac{\hat{\sigma}^{(POT)}}{\hat{\xi}^{(POT)}}\left(\left(\frac{k}{np}\right)^{\hat{\xi}^{(POT)}} - 1\right), \text{ pour } p < \frac{k}{n}, \tag{1.90}$$

le point terminal est estimé par :

$$\hat{x}_F = X_{n-k,n} - \frac{\hat{\sigma}^{(POT)}}{\hat{\xi}^{(POT)}}. \tag{1.91}$$

1.10.6 Estimation de Weissman

Dans l'analyse des extrêmes, une exigence typique consiste à trouver les grandes valeurs de sorte que les probabilités de les dépasser soient très faibles (c-à-d proches de zéro). Ces quantités sont appelées quantiles extrêmes 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 p.

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

On estime le quantile extrême, à partir de l'estimateur de Hill de γ en extrapolant le long de la droite du « Paréto quantile plot » d'équation :

$$y = \log X_{n-kn,n} + \widehat{\gamma}_{k_n}^{H}\left(x + \log\frac{k_n+1}{n+1}\right), \tag{1.92}$$

de point d'origine $(\log((n+1)/(kn+1)), \log X_{n-kn,n})$,Ceci conduit en prenant $x = -\log \alpha_n$ à l'estimateur le plus connu d'un quantile extrême :

$$q(\alpha_n) = F^{\leftarrow}(1-\alpha_n) = Q(1-\alpha_n) = U(1/\alpha_n), \tag{1.93}$$

introduit par Weismann (1978) [167] :

$$\widehat{q}^{H}_{k_n}(\alpha) = X_{n-k_n,n}\left(\frac{k_n+1}{(n+1)\,\alpha}\right)^{\widehat{\gamma}^{+}_{k_n}}, \tag{1.94}$$

et $\widehat{\gamma}^{+}_{k_n} = \widehat{\gamma}^{H}_{k_n}$.

1.10.8 Approche basée sur un estimateur d'indices quelconques

En utilisant l'estimateur des moments ont proposé d'estimer $q(\alpha_n) = U(1/\alpha_n)$ pour $\alpha_n \longrightarrow 0$ comme suit :

$$\widehat{q}^{M}_{k_n}(\alpha_n) = X_{n-K_n,n} + \widehat{a}\left(\frac{k_n}{n}\right)\frac{\left(\frac{n}{k_n}\alpha_n\right)^{\widehat{-q}^{M}_{k_n}} - 1}{\widehat{q}^{M}_{k_n}}, \tag{1.95}$$

avec

$$\widehat{a}(k_n/n) = X_{n-K_n,n}\,\widehat{\gamma}^{H}_{k_n}\max\left(1-\widehat{q}^{M}_{k_n}, 1\right).$$

Conclusion

La méthode des maxima par blocs permet d'utiliser l'approximation d'une fonction de répartition F que l'on suppose dans le domaine d'attraction d'une loi GEV de fonction de rrtition G, de paramètres σ, γ, μ), par $G^{1/n}$, en construisant à l'aide de l'échantillon initial un pseudo-échantillon de maxima par blocs. La mise en pratique est très simple, mais plusieurs mises en garde sont à faire cependant. Le choix de la taille n des blocs est essentiel, et souvent délicat. Une taille de blocs trop petite conduira à une mauvaise qualité de l'approximation de F par $G^{1/n}$, se concluant par un biais dans l'estimation, tandis qu'une taille de blocs trop grande introduira une grande variance dans l'estimation. Cette méthode semble "gaspiller" beaucoup de données, dans la mesure où seule est conservée la plus grande valeur de chaque bloc. Si des résultats existent quantifiant la distance entre la fonction de répartition F et son attracteur G.

Chapitre 2

Caractérisation des distributions de type de Pareto

2.1 Introduction

La loi de Pareto, appelée aussi principe de Pareto ou encore règle des 80/20, est un phénomène empirique révélant que 80% des effets sont le produit de 20% des causes. Pareto c'est un économiste italien Vilfredo Pareto (1848-1923) qui mit ce principe en évidence au début du 20ème siècle en analysant les données fiscales de différents pays en vue de connaître la répartition des richesses. Il constatera que **80 % des richesses étaient détenues par 20 % de la population.** Cette loi, issue du domaine économique, eut le succès que l'on sait et sera portée dans de très nombreux domaines. La plupart du temps, elle sert à illustrer un principe de répartition inégal des actions et des résultats. La conclusion de Pareto est que de façon générale "l'augmentation des richesses par rapport à la population produit soit :

— l'augmentation du revenu minimum,
— soit la diminution de l'inégalité des revenus,
— soit les deux effets simultanément".

L'un des problèmes les plus importants dans la Théorie des Valeurs Extrêmes est l'estimation de l'indice ξ, cet indice caractérise la forme la distribution. Il existe de nombreux estimateurs pour cet indice.

L'estimateur le plus célèbre pour un ξ strictement positif est l'estimateur de Hill (1975) [107] que nous allons introduire dans ce chapitre. Celui-ci est construit à partir des $k+1$ plus grandes valeurs de l'échantillon, autrement dit celles qui dépassent la statistique d'ordre $X_{n-k,n}$. On dit que celle-ci joue le rôle de seuil. Le choix du seuil est une question très épineuse car l'estimateur de Hill est très volatil en fonction de ce seuil.

2.2 Distributions de type de pareto

Dans ce chapitre, nous considérons l'estimation de l'indice des valeurs extrêmes, des quantiles extrêmes et des probabilités de dépassement faibles, dans le cas où la distribution est de type Pareto, c'est-à-dire :

$$\bar{F}(x) = x^{-1/\xi}\ell_F(x) \tag{2.1}$$

qui est équivalent à

$$Q(1-1/x) = U(x) = x^{\xi}\ell_U(x) \tag{2.2}$$

où $\ell_F(x)$ et $\ell_U(x)$ sont des fonctions à variations lentes.

2.3 Condition des fonctions régulières du premier ordre

La caractérisation des distributions de type de Pareto est présentée dans la proposition suivante, qui est connue en littérature par la condition des fonctions à varriations régulières de premier ordre.

Proposition 2.1 (de Haan et Ferreira (2006)) *Les assertions suivantes sont équivalentes :*

(i) *F est à queue lourde* $F \in D.A.(Fréchet)$, $\xi > 0$.

(ii) *F est une fonction à variations régulières à l'infini d'indice* $-1/\xi$

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\xi}, x > 0. \tag{2.3}$$

(iii) $Q(1-s)$ *est une fonction à variations régulières à zéro d'indice* $-\xi$

$$\lim_{s\to 0} \frac{Q(1-sx)}{Q(1-s)} = x^{-\xi}, x > 0. \tag{2.4}$$

(iv) *U est une fonction à variations régulières à l'infini d'indice* ξ

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\xi}, x > 0. \tag{2.5}$$

Depuis le début des années quatre-vingt du vingtième siècle, ce problème a été étudié en détail dans la littérature. L'estimateur de Hill (Hill (1975)), paru en 1975, continue d'avoir une grande importance et constitue le sujet principal de ce chapitre.

Cependant, pour avoir une meilleure idée du choix des estimateurs possibles, commençons par quelques exemples d'estimateurs naïfs. Ce qu'ils ont tous en commun, c'est une tentative d'éviter la partie qui varie lentement et qui n'est pas pertinente.

Nous supposons désormais que nous avons un échantillon de i.i.d. valeurs $\{X_i, 1 \le i \le n\}$ d'une queue de type Pareto $1 - F$.

Les queues de type Pareto sont systématiquement utilisées dans certaines branches de l'assurance non vie. En économie pour modéliser les distributions des revenus.

Également en finance (rendement des stocks) et en télécommunication (taille de fichier, temps d'attente), cette classe est appropriée. Dans d'autres domaines d'application des statistiques de valeurs extrêmes telles que l'hydrologie, l'utilisation des modèles de Pareto semble beaucoup moins systématique.

Cependant, les problèmes d'estimation considérés ici sont typiques de la méthodologie des valeurs extrêmes et, en même temps, le modèle de type Pareto est plus spécifique et plus simple à gérer. Donc, ce chapitre a aussi un but instructif ; les distributions à queue épaisse sont un «terrain de jeu» idéal pour développer des méthodes efficaces qui doivent être étendues dans le cas général $\xi \in \mathbb{R}$.

2.4 Une approche naïve

Essayons quelques moyens faciles de se débarrasser de la fonction U. D'après la relation (2.2), nous voyons que pour $x \to \infty$,

$$\log U(x) = \xi \log x + \log \ell_U(x) \sim \xi \log x.$$

Par conséquent, il semble naturel de remplacer dans l'expression ci-dessus la quantité déterministe U par une quantité aléatoire dont l'argument va à l'infini avec la taille de l'échantillon. Pour simplifier, l'argument x pourrait être pris pour être n ou plus généralement n/k. Dans la suite, nous avons mis $\hat{Q}_n(1 - 1/x) = \hat{U}_n(x)$ comme un estimateur empirique de U. Nous nous attendons alors à avoir une déclaration probabiliste du type

$$\log U\left(\frac{n}{k}\right) \sim \xi \log\left(\frac{n}{k}\right).$$

Alors, pour tout $r \in \{1, 2, ..., n\}$, on a

$$\hat{U}_n\left(\frac{n}{n-r}\right) = X_{r,n},$$

et nous nous attendons donc asymptotiquement à ce que pour $n \to \infty$, cette

$$\log X_{n-k+1,n} \sim \xi \log\left(\frac{n}{k}\right).$$

Si on note $\{E_i\}_{i=1}^{\infty}$ une suite i.i.d. des variables aléatoires exponentielles avec moyenne 1, alors

$$\frac{X_{n-k+1,n} - U(n)}{a(n)} \xrightarrow{\mathcal{D}} h_\xi\left(\left(\sum_{i=1}^{k} E_i\right)^{-1}\right). \tag{2.6}$$

En remplaçant a_n par $\xi U(n)$ dans (2.6), il en résulte que, lorsque k est fixé,

$$\log\left(\frac{X_{n-k+1,n}}{U(n)}\right) = O_p(1),$$

où

$$\log X_{n-k+1,n} - \xi \log(n) - \log \ell_U(x) = O_p(1),$$

on déduit en effet que si F satisfait (C_ξ) et $\xi > 0$

$$\log X_{n-k+1,n}/\log(n) \xrightarrow{P} \xi.$$

Ce simple résultat montre qu'une seule statistique d'ordre plus grand peut être utilisée pour estimer l'indice de valeur extrême ξ. Mais cette approche naïve présente de graves inconvénients. Par exemple, il ne semble en effet pas satisfaisant de n'utiliser qu'une seule statistique d'ordre dans la procédure d'estimation. Aussi, que veut dire garder k fixe. De plus, il en découle que le taux de convergence est lent sur le plan logarithmique.

Par intuition statistique de base, nous pouvons prévoir qu'un estimateur basé sur plus de statistiques d'ordre sera plus fiable. Une possibilité consiste à examiner les différences entre deux statistiques d'ordre extrême différentes, telles que (voir Bacro et Brito (1995) [6]) :

$$\frac{\log X_{n-k+1,n} - \log X_{n-2k+1,n}}{\log 2},$$

ou des généralisations avec des espacements d'ordre différent de k. En utilisant la variation régulière des queues de type Pareto, on peut facilement constater que cet estimateur est cohérent si $k \to \infty$ et $n/k \to \infty$. Il s'avère que cette statistique améliore considérablement le taux de cohérence par rapport au premier estimateur naïf, mais n'utilise encore que deux observations extrêmes. L'estimateur de Hill améliorera considérablement cet aspect.

Mais même dans ce cas, nous avons besoin de savoir quelles statistiques sur les gros ordres peuvent être utilisées dans la procédure. À partir des dérivations du chapitre précédent, nous pourrions déduire que, si la taille de l'échantillon tend à être égale à ∞, alors k devrait également être autorisé à faire de même, bien qu'à un certain taux.

2.5 Estimateur de Hill

Il existe au moins quatre manières naturelles d'introduire cet estimateur. Tous sont inspirés par l'analyse précédente. De plus, l'estimateur jouit d'une grande popularité grâce à quelques propriétés théoriques intéressantes, mais en dépit de sérieux inconvénients.

2.5.1 Construction

i) La vue quantile :

a) La première source d'inspiration provient des parcelles quantiles des distributions de type Pareto.

$$\frac{\log Q(1-p)}{-\log p} \to \xi, \text{ lorsque } p \to 0.$$

Il en résulte qu'un tracé quantile de Pareto, c'est-à-dire un tracé quantile exponentiel basé sur les données transformées par log, est finalement linéaire avec la pente ξ près des plus grandes observations.

b) En outre, la pente d'un tracé de quantile exponentiel finalement linéaire peut être estimée à l'aide des valeurs de dépassement moyen du type $E_{k,n}$, comme indiqué dans le chapitre 1. La combinaison de ces deux observations conduit à la valeur excédentaire moyenne des données transformées en log, connues sous le nom d'estimateur de Hill (Hill (1975) [107]) :

$$H_{k,n} := \frac{1}{k}\sum_{i=1}^{k} \log X_{n-i+1,n} - \log X_{n-k,n}.$$

Une question importante peut être posée concernant l'optimalité de l'estimateur de Hill en tant qu'estimateur de la pente d'un graphique QQ quantile-quantile. En fait, les coordonnées verticales $\log X_{n-i+1,n}$ ne sont pas indépendantes et ne possèdent pas la même variance et résument donc la partie supérieure du quantile de Pareto. plot

$$\left(\log = \frac{n+1}{j}, \log X_{n-i+1,n}, i = 1, ..., k\right),$$

en utilisant une ligne des moindres carrés

$$y = \log X_{n-k,n} + \xi(x - \log((n+1)/(k+1)))$$

ne semble pas être efficace, car les conditions classiques de Gauss-Markov ne

sont pas remplies.

En combinant les informations sur un ensemble de valeurs i possibles, nous pouvons rechercher la droite des moindres carrés qui correspond le mieux aux points

$$\left\{\left(-\log\frac{i}{n+1}, \log X_{n-i+1,n}\right), i = 1,...,k+1\right\}$$

où nous forçons la ligne droite à traverser le point le plus à gauche de ces points. Une telle ligne a la forme

$$y = \log X_{n-k,n} + \xi\left(-\log\frac{i}{n+1} - \log\left(\frac{n+1}{k+1}\right)\right).$$

Un peu de réflexion indique qu'il serait peut-être sage de donner des points à droite de l'ensemble de poids ci-dessus, à la lumière du problème de l'hétéroscédasticité. Afin de trouver la valeur des moindres carrés de ξ, nous minimisons donc la quantité

$$\sum_{i=1}^{k} w_{i,n}\left[\log X_{n-i+1,n} - \left(\log X_{n-k,n} + \xi\log\frac{k+1}{i}\right)\right]^2$$

où $\{w_{i,n}, i = 1,...,k\}$ sont des poids appropriés. Un simple calcul nous dit que la valeur résultante de ξ, disons $\hat{\xi}_k$, est donnée par

$$\hat{\xi}_k = \sum_{i=1}^{k} \alpha_{i,k}\log\left(\frac{X_{n-i+1,n}}{X_{n-k,n}}\right)$$

où

$$\alpha_{i,k} = \frac{w_{i,n}\log\frac{k+1}{i}}{\sum_{r=1}^{k} w_{r,n}\left(\log\frac{k+1}{r}\right)^2}.$$

Lorsque on choisir $\alpha_{i,k} = 1/k$ on arrive a l'estimateur de Hill.

ii) La vue de probabilité : La définition d'une queue de type Pareto peut être réécrite comme suit :

$$\frac{1-F(tx)}{1-F(t)} \to x^{-1/\xi} \text{ quand } t \to \infty \text{ pour tout } x > 1.$$

Ceci peut être exprimé comme suit

$$P(X/t > x \mid X > t) \approx x^{-1/\xi} \text{ quand } t \text{ grand et pour tout } x > 1.$$

iii) La représentation exponentielle de Rényi : Il existe une autre façon d'écrire

l'estimateur de Hill en introduisant les variables aléatoires

$$Z_i = i\,(\log X_{n-i+1,n} - \log X_{n-i,n}) := iT_i$$

cela jouera un rôle crucial plus tard. À travers une sommation partielle, on constate que

$$\sum_{i=1}^{k} Z_i = \sum_{i=1}^{k} iT_i = \sum_{i=1}^{k}\sum_{j=1}^{i} T = \sum_{j=1}^{k}\sum_{i=j}^{k} T_i$$

qui mène facilement à la relation cruciale

$$H_{k,n} := \frac{1}{k}\sum_{i=1}^{k} Z_i = \bar{Z}_k.$$

iv) Approche en excès moyenne : Encore une autre variante est basée sur la fonction d'excès moyen des données transformées par log. Si $1-F \in \mathcal{R}_{-1/\xi}$ avec $\xi > 0$, alors

$$E\,(\log X - \log x \mid X > x) = \int_x^{\infty} \frac{1-F(u)}{1-F(x)}\frac{du}{u} \to \xi \text{ lorsque } x \to \infty.$$

Remplacez la distribution F par sa contrepartie empirique $\hat{F}_n$, et x par la séquence aléatoire $X_{n-k,n}$ qui tend vers ∞. C'est alors un exercice agréable pour montrer que

$$H_{k,n} := \int_{X_{n-k,n}}^{\infty} \frac{1-\hat{F}_n(u)}{1-\hat{F}_n(X_{n-k,n})}\frac{du}{u}.$$

2.5.2 Propriétés de l'estimateur de Hill

Un grand nombre de travaux théoriques ont été consacré à l'étude des propriétés de l'estimateur de Hill.

2.5.2.1 Consistance faible

Mason (1982) [126] a montré que $H_{k,n}$ est un estimateur cohérent pour ξ (sous la condition $k,\ n \to \infty,\ k/n \to 0$) quelle que soit la fonction variant lentement ℓ_F (ou ℓ_U). Ceci est même vrai pour des données faiblement dépendantes (Hsing (1991) [111]) ou dans le cas d'un processus linéaire (Resnick et Stăricä (1995) [145]).

Théorème 2.1 *Soit $X_1, X_2, ..., X_n$ un échantillon i.i.d. des variables aléatoires avec une fonction de distribution F. Supposons que $F \in D\left(G_\xi\right)$ avec $\xi > 0$. Alors, pour une suite d'entiers $k = k(n) \to \infty$, $k(n)/n \to 0$ lorsque $n \to \infty$, avec*

$$H_{k,n} \xrightarrow{P} \xi.$$

Le théorème suivant est une réciproque du résultat précédent.

Théorème 2.2 *Soit $X_1, X_2, ..., X_n$ un échantillon i.i.d. des variables aléatoires avec une fonction de distribution F. Supposons que pour une suite d'entiers $k = k(n) \to \infty$, $k(n)/n \to 0$ et $k(n+1)/k(n) \to 1$, lorsque $n \to \infty$, avec*

$$H_{k,n} \xrightarrow{P} \xi > 0.$$

Alors $F \in D\left(G_\xi\right)$.

2.5.2.2 Consistance forte

En ajoutant une condition supplémentaire sur la suite k, Deheuvels et al. 1988 [44] obtiennent la consistance forte de cet estimateur.

Théorème 2.3 *Soit $X_1, X_2, ..., X_n$ un échantillon i.i.d. des variables aléatoires avec une fonction de distribution F. Si $F \in D\left(G_\xi\right)$ avec $\xi > 0$ et k est une suite intermédiaire vérifiant $\lim_{n\to\infty} \frac{k}{\log\log n} = \infty$. Alors*

$$H_{k,n} \xrightarrow{p.s.} \xi.$$

2.5.2.3 Normalité asymptôtique

La normalité asymptotique est due entre autre à Davis et Resnick (1984) [33], Csörgő et al. (1985) [28], Haeusler et Teugels (1985) [101] et Smith (1987) [155].

Notons que la consistance (en probabilité ou presque sûrement) de l'estimateur $H_{k,n}$ 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 $U(\cdot)$.

La normalité asymptotique de $H_{k,n}$ a été abordée parmi d'autres dans Hall (1982) [102], Davis et Resnick (1984) [33], Csörgő et Mason (1985) [26], Haeusler et Teugels (1985) [101], Deheuvels et al. (1988) [44], Csörgő et Viharos (1998) [29], de Haan et Peng (1998) [41] et de Haan et Resnick (1998) [41]. Dans Drees (1998) [52] et Beirlant et al. (2002) [11], l'optimalité de la variance et du taux de l'estimateur de Hill a été dérivée pour les grands sous-modèles du modèle de type Pareto.

Définition 2.1 (Variation régulière du second ordre) *Soit $F \in D(H_\xi), \xi > 0$, 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* $x > 0$

$$\lim_{t\to\infty} \frac{\frac{\bar{F}(tx)}{\bar{F}(t)} - x^{-1/\xi}}{A(t)} = \begin{cases} x^{-1/\xi} \frac{x^{\rho}-1}{\rho}, \ si \ \rho < 0 \\ x^{-1/\xi} \log x, \ si \ \rho = 0 \end{cases} \tag{2.7}$$

ou encore

$$\lim_{t\to\infty} \frac{\frac{1}{\xi}\log x + \log\left(\bar{F}(tx)\right) - \log\left(\bar{F}(t)\right)}{A(t)} = \begin{cases} \frac{x^{\rho}-1}{\rho}, \ si \ \rho < 0 \\ \log x, \ si \ \rho = 0 \end{cases} \tag{2.8}$$

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

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^{\xi}}{b(t)} = \begin{cases} x^{\xi} \frac{x^{\rho}-1}{\rho}, \ si \ \rho < 0 \\ x^{\xi} \log x, \ si \ \rho = 0 \end{cases} \tag{2.9}$$

ou encore

$$\lim_{t\to\infty} \frac{\log\left(U(tx)\right) - \log\left(U(t)\right) - \xi \log x}{b(t)} = \begin{cases} \frac{x^{\rho}-1}{\rho}, \ si \ \rho < 0 \\ \log x, \ si \ \rho = 0 \end{cases}$$

Le paramètre ρ contrôle la vitesse de convergence de $\frac{\ell_F(tx)}{\ell_F(t)}$ vers 1 quand $t \to \infty$ ou de manière équivalente $\frac{\ell_U(tx)}{\ell_U(t)}$ vers 1 quand $t \to \infty$, pour plus de détails, voir Geluk & Haan (1987) [85]. Plus ρ est proche de 0, plus la convergence sera lente et donc l'estimation de ρ sera difficile. Beaucoup d'auteurs se sont intéressés à l'estimation de ce paramètre afin de réduire le biais asymptotique qui apparaît lors de l'estimation de l'indice extrême ξ (cf. Alves et al. (2003)[66],[67], Gomes et Martins (2001) [75], Gomes et al. (2000) [71], de Haan et Peng (1998) [41]).

Ces conditions permettent d'énoncer les propriétés asymptotiques de l'estimateur de Hill (1975) [107].

Théorème 2.4 *Supposons que F satisfait la condition régulière d'ordre deux (2.9) où B est une fonction positive avec* $\lim_{t\to\infty} b(t) = 0$, *avec* $k = k(n) \to \infty, k/n \to 0$ *lorsque* $n \to \infty$ *et* $\sqrt{k}\lim_{n\to\infty} b\left(\frac{n}{k}\right) \to \lambda (\lambda \ fini)$. *Alors*

$$\sqrt{k}\left(H_{n,k} - \xi\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \xi^2\right).$$

Cependant, plusieurs problèmes se posent.

— Pour chaque valeur de k, on obtient une valeur différente de ξ. Habituellement, on trace les estimations $H_{k,n}$ contre k, donnant le tracé de Hill : $\{(k, H_{k,n}), 1 \leq k \leq n-1\}$.

— Dans de nombreux cas, un biais grave peut apparaître. Cela se produit lorsque l'effet de la partie variant lentement dans le modèle disparaît lentement dans le graphique de quantiles de Pareto. En d'autres termes, dans l'hypothèse des probabilités, l'hypothèse selon laquelle les dépassements relatifs supérieurs à un certain seuil suivent une distribution de Pareto stricte est parfois trop optimiste.

— L'estimateur de Hill partage un grave défaut avec de nombreux autres estimateurs courants basés sur des données transformées en log : l'estimateur n'est pas invariant en ce qui concerne les décalages des données. Comme plusieurs auteurs l'ont mentionné, l'utilisation inadéquate de l'estimateur de Hill conjointement avec un décalage de données peut également conduire à des erreurs systématiques. Fraga Alves (2001) [65] propose une modification de l'estimateur de Hill non invariante par l'emplacement. A cette fin, une valeur k secondaire, notée k_0 ($k_0 < k$), est introduite, conduisant à

$$\hat{\xi}^{(H)}(k,k_0) = \frac{1}{k_0}\sum_{i=1}^{k_0} \log \frac{X_{n-i+1,n} - X_{n-k,n}}{X_{n-k_0,n} - X_{n-k,n}}.$$

Si on laisse $k = k_n$ et $k_0 = k_{0,n}$ tend vers l'infini avec $n \to \infty$, tel que $k/n \to 0$ et $k_0/k \to 0$, on peut montrer que $\hat{\xi}^{(H)}(k,k_0)$ est cohérent. Une version adaptative de l'estimateur proposé a été proposée à partir du meilleur k_0 théorique donné par

$$k_0 \sim \left[\frac{(1+\xi)}{\sqrt{2\xi}}\right]^{2/(1+2\xi)} k^{2\xi/(1+2\xi)}$$

départ avec une estimation initiale $\hat{\xi}^{(0)}$ dans un premier temps ; par exemple, obtenu en définissant $k_0^{(0)} = \left\lfloor 2k^{2/3} \right\rfloor$.

2.5.3 Autres estimateurs de régression

L'estimateur de Hill a été obtenu à partir du graphe de quantiles de Pareto à l'aide d'un estimateur assez naïf de la pente à l'extrémité droite du graphe de quantiles. Bien entendu, des méthodes de régression plus souples sur les k points les plus élevés du tracé du quantile de Pareto pourraient être appliquées. Ce programme a été exécuté en détail dans Schultze et Steinebach (1996) [152], Kratz et Resnick (1996) [116] et Csörgő et Viharos (1998) [29]. Nous nous référons à ces articles pour plus de détails mathématiques et nous nous limitons ici à la dérivation des estimateurs.

(i) Les moindres carrés pondérés du tracé du quantile de Pareto traité dans la sec-

tion 2.5 peuvent être réécrits sous la forme

$$\hat{\xi}_k = \frac{\sum_{j=1}^{k} T_j \sum_{i=1}^{j} w_{i,n} \log \frac{k+1}{i}}{\sum_{i=1}^{k} w_{i,k} \left(\log \frac{k+1}{i}\right)^2}.$$

posons

$$K\left(\frac{j}{n}\right) = \frac{1}{j} \sum_{i=1}^{j} w_{i,n} \log \frac{k+1}{i}$$

cette estimateur peut être exprimer comme suit

$$\hat{\xi}_{K,k} = \frac{\frac{1}{k} \sum_{i=1}^{k} K\left(\frac{i}{n}\right) i \left(\log X_{n-i+1,n} - \log X_{n-i,n}\right)}{\frac{1}{k} \sum_{i=1}^{k} K\left(\frac{i}{n}\right)},$$

montrant que l'estimation des moindres carrés pondérée conduit à la classe des estimateurs à noyau introduite par Csörgő et al. (1985) [28]. Ici, K désigne une fonction du noyau associant différentes pondérations aux différentes statistiques d'ordre. Csörgő et al. (1985) [28] considèrent également les fonctions du noyau avec support extérieur $(0,1]$. Un choix optimal de K est possible mais difficile à gérer en pratique.

La pondération des espacements Z_i a pour avantage que les graphiques des estimations en fonction de k sont plus lisses en comparaison, par exemple, avec l'estimateur de Hill, où les valeurs adjacentes de k peuvent conduire à des estimations très différentes.

(ii) Le problème de la régularité des estimations de Hill en fonction de k peut être résolu d'une autre manière : les moindres carrés simples non contraints avec estimation de la pente ξ ainsi que l'interception, disons δ, peuvent déjà fournir une régularité même sans le recours à une fonction du noyau. Cette procédure d'ajustement des lignes sur (des parties de) tracés QQ et en particulier des tracés à double logarithme peut être retracée jusqu'à Zipf à partir de la fin des années 1940 (voir Zipf (1949) [172]). Ce n'est que récemment que cette procédure a été étudiée plus en profondeur.

La procédure des moindres carrés classique minimisant

$$\sum_{i=1}^{k} \left(\log X_{n-i+1,n} - \left(\delta + \xi \log \frac{k+1}{i}\right)\right)^2$$

par rapport à δ et ξ conduit à

$$\hat{\xi}_{K,k} = \frac{\frac{1}{k}\sum_{i=1}^{k}\left(\log\frac{k+1}{i} - \frac{1}{k}\sum_{i=1}^{k}\log\frac{k+1}{i}\right)\log X_{n-i+1,n}}{\frac{1}{k}\sum_{i=1}^{k}\log^2\frac{k+1}{i} - \left(\frac{1}{k}\sum_{i=1}^{k}\log\frac{k+1}{i}\right)}.$$

C'est l'estimateur proposé dans Schultze et Steinebach (1996) [152] et Kratz et Resnick (1996) [116]. Dans Csörgő et Viharos (1998) [29], les propriétés asymptotiques de cet estimateur sont passées en revue. Ces auteurs proposent également une généralisation de cet estimateur qui peut à nouveau être motivée par un algorithme des moindres carrés pondéré :

$$\hat{\xi}_{WLS,k} = \frac{\frac{1}{k}\sum_{i=1}^{k}\left(\int_{(i-1)/k}^{i/k} J(s)ds\right)\log X_{n-i+1,n}}{\frac{1}{k}\sum_{i=1}^{k}\left(\int_{(i-1)/k}^{i/k} J(s)ds\right)\log\frac{k+1}{i}},$$

où J est une fonction non croissante définie sur $(0,1)$, qui intègrable en 0.

Csörgő et Viharos (1998) [29] proposent d'utiliser les fonctions de pondération J du type

$$J_\theta(s) = \frac{\theta+1}{\theta} - \frac{(\theta+1)^2}{\theta}s^\theta, s \in [0,1],$$

pour $\theta > 0$.

2.6 Représentation des espacements de log et des résultats asymptotiques

Dans cette section, nous étudions les propriétés mathématiques les plus importantes de l'estimateur de Hill et certaines généralisations sélectionnées, comme indiqué ci-dessus. En particulier, nous proposons des expressions pour le biais asymptotique et la variance asymptotique.

Ces résultats seront utiles plus tard lorsque nous discuterons du choix adaptatif de k. Les résultats obtenus peuvent également aider à trouver des solutions à certains des problèmes susmentionnés.

Dans la section 2.5, nous avons déduit que l'estimateur de Hill peut s'écrire comme une simple moyenne d'espacements de log à l'échelle :

$$H_{k,n} := \frac{1}{k}\sum_{i=1}^{k} Z_i, \text{ avec } Z_i = i\left(\log X_{n-i+1,n} - \log X_{n-i,n}\right).$$

Nous allons maintenant élaborer sur ces espacements. Nous poursuivons la discussion comme commencé dans la section 2.5, où nous avons constaté que dans le cas de

distributions de Pareto strictes

$$Z_i \overset{\mathcal{D}}{=} \xi E_i, i = 1, 2, ..., k,$$

avec $\{E_i, 1 \leq i \leq n\}$ un échantillon d'une distribution exponentielle de moyenne 1. Conformément à nos conventions, leurs statistiques d'ordre sont alors notées

$$E_{1,n} \leq E_{2,n} \leq ... \leq E_{n-k+1,n} \leq ... \leq E_{n,n}.$$

La double utilisation de la transformée intégrale de probabilité conduit aux égalités de liaison

$$X_{i,n} \overset{\mathcal{D}}{=} U\left(e^{E_{i,n}}\right), i = 1, 2, ..., n.$$

La principale raison d'utiliser un échantillon exponentiel réside dans une propriété remarquable concernant les statistiques d'ordre de cette dernière distribution, découverte par A. Rényi. Effectivement,

$$E_{n-j+1,n} - E_{n-k,n} \overset{\mathcal{D}}{=} \sum_{i=j}^{n} \frac{E_i}{i}, 1 \leq j \leq k < n$$

où $\{E_i, 1 \leq i \leq n\}$ un échantillon d'une distribution exponentielle de moyenne 1. À partir de cette équation, on peut par exemple déduire les attentes des statistiques d'ordre exponentielles dans

$$E\left(E_{n-j,n}\right) = \sum_{k=j}^{n-1} \frac{1}{k+1} \sim \log\left(\frac{n+1}{j+1}\right)$$

si n est grand.

Nous combinons maintenant ce qui précède avec les propriétés de second ordre de la fonction de quantile de queue U. À partir de là, nous supposons que $\log U$ vérifie $(\mathcal{C}_{-\beta}(b))$ pour certains $\beta > 0$ et $b \in \mathcal{R}_{-\beta}$. Cela signifie que nous pouvons écrire

$$\frac{U(tx)}{U(t)} = x^{\xi}\left(1 + h_{-\beta}(x)b(x) + o(b(x))\right) \tag{2.10}$$

En utilisant la condition de second ordre (2.10), nous développons la distribution des

espacements $Z_i = i\,(\log X_{n-i+1,n} - \log X_{n-i,n}), i = 1,...,k$

$$\begin{aligned} Z_i &= i \log \frac{X_{n-i+1,n}}{X_{n-i,n}} \\ &\overset{\mathcal{D}}{=} i \log \frac{U\left(e^{E_{n-i+n,n}-E_{n-i,n}} e^{E_{n-i,n}}\right)}{U\left(e^{E_{n-i,n}}\right)} \\ &\overset{\mathcal{D}}{=} i \log \frac{U\left(e^{E_j/j} e^{E_{n-i,n}}\right)}{U\left(e^{E_{n-i,n}}\right)} \\ &= i\left\{\xi \log e^{E_j/j} + \log 1 + h_{-\beta}(e^{E_j/j}) b(e^{E_j/j})(1+o(1))\right\} \\ &= \xi E_i + i \log(1 + W_{n,i}) \end{aligned}$$

où

$$W_{n,i} = h_{-\beta}(e^{E_j/j}) b(e^{E_j/j}) \tag{2.11}$$

On obtient la forme stochastique des espacement comme suit :

$$Z_i \overset{\mathcal{D}}{=} \xi E_i + i \log(1 + W_{n,i}).$$

Une façon d'utiliser ce résultat est de remplacer le long terme à droite par des inégalités telles que

$$\frac{\xi}{1+\xi} \leq \log(1+\xi) \leq \xi$$

qui produisent des inégalités universelles et stochastiques pour l'estimateur de Hill depuis

$$H_{k,n} := \frac{1}{k} \sum_{i=1}^{k} Z_i.$$

Tout d'abord, il est facile de voir que pour y petit,

$$h_{-\beta}(e^y) = y\,(1 + o(1)),$$

Ensuite, nous devons mieux comprendre le comportement de l'argument de $b(x)$ dans (2.11). aussi longtemps que $i/n \to 0$ lorsque $n \to \infty$, on a

$$E_{i-n,n}/\log n/i \xrightarrow{P} 1.$$

Mais

$$b(e^{E_{n-i,n}}) = b\left(\frac{n+1}{i+1}\right)(1 + o(1))$$

Ceci signifie qu'on peut approximé $i \log(1 + W_{n,i})$ en distribution par $E_i b\left(\frac{n+1}{i+1}\right)$. Nous

sommes donc amenés à la représentation approximative suivante :

$$Z_i \overset{\mathcal{D}}{\sim} \left(\xi + b\left(\frac{n+1}{i+1}\right)\right) E_i. \tag{2.12}$$

ou, en utilisant la variation régulière de b avec l'indice $-\beta$,

$$Z_i \overset{\mathcal{D}}{\sim} \left(\xi + \left(\frac{i}{k+1}\right)^{\beta} b\left(\frac{n+1}{k+1}\right)\right) E_i, i = 1, ..., k. \tag{2.13}$$

Dans la suite, en utilisant la notation $b_{n,k} = b\left(\frac{n+1}{k+1}\right)$.

Théorème 2.5 *Supposons (2.10) est satisfait. Ensuite, il existe des variables aléatoires $R_{i,n}$ et des variables aléatoires exponentielles standard E_j (indépendantes pour chaque n) telles que*

$$\sup_{1 \le i \le k} \left| Z_i - \left(\xi + \left(\frac{i}{k+1}\right)^{\beta} b_{n,k}\right) E_i - R_{i,n} \right| = o_P\left(b_{n,k}\right) \tag{2.14}$$

lorsque $k, n \to \infty$ avec $k/n \to 0$, où uniformément $i = 1, ..., k$

$$\left| \sum_{i=j}^{k} \frac{R_{i,n}}{i} \right| = o_P\left(b_{n,k} \max\left(\log\left(\frac{k+1}{i}\right), 1\right)\right).$$

Tirons quelques conclusions concernant l'estimateur de Hill.

(i) Le biais asymptotique de l'estimateur de Hill peut être retracé à l'aide de la représentation exponentielle. Effectivement,

$$ABias\left(H_{k,n}\right) \sim b_{n,k} \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k+1}\right)^{\beta} \sim \frac{b_{n,k}}{1+\beta}.$$

(ii) La variance asymptotique de l'estimateur de Hill est encore plus facile en ce que

$$AVar\left(H_{k,n}\right) \sim var\left(\frac{\xi}{k} \sum_{i=1}^{k} E_i\right) \sim \frac{\xi^2}{k}.$$

(iii) Enfin, la normalité asymptotique de l'estimateur de Hill peut être attendue lorsque $k, n \to \infty$ avec $k/n \to 0$. Alors, si $\sqrt{k} b_{n,k} \to 0$

$$\sqrt{k}\left(\frac{H_{k,n}}{\xi} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Ce résultat permet de construire des intervalles de confiance approximatifs

pour ξ. Au niveau $(1-\alpha)$, cet intervalle est donné par

$$\left(H_{k,n}\left(1+\frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{k}}\right)^{-1}, H_{k,n}\left(1-\frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{k}}\right)^{-1}\right) \tag{2.15}$$

ce qui est une approche acceptable si le biais n'est pas trop important, c'est-à-dire si $\beta \geq 1$. Typiquement, la condition $\sqrt{k}b_{n,k} \to 0$ restreint considérablement la plage de k-valeurs où l'intervalle de confiance fonctionne.

Nous terminons cette section en décrivant comment le résultat de la représentation exponentielle ci-dessus peut être utilisé pour dériver formellement les expressions de biais-variance ci-dessus et en déduire des résultats de normalité asymptotique pour les statistiques de type noyau

$$H_{k,n}^{K} := \frac{1}{k}\sum_{i=1}^{k} K\left(\frac{i}{n+1}\right) Z_i \tag{2.16}$$

comme discuté plus tôt dans ce chapitre. Ici, nous supposons que le noyau K peut s'écrire sous la forme $K(t) = \frac{1}{t}\int_0^t u(v)dv$, $0 < t < 1$ pour une fonction u définie sur $(0,1)$.

Selon le choix du noyau K, plusieurs estimateurs peuvent en résulter notamment celui de Hill avec $K(x) = 1_{]0,1]}(x)$.

Ce type de résultats peut être trouvé, par exemple, dans Csörgő et al. (1985) [28] et Csörgő et Viharos (1998) [29].

Théorème 2.6 *Supposons que (2.10) satisfait, soit* $K(t) = \frac{1}{t}\int_0^t u(v)dv$, *pour une fonction* u *satisfait* $\left|k\int_{(i-1)/n}^{i/n} u(v)dv\right| \leq f\left(\frac{i}{k+1}\right)$ *pour quelques fonction continue* f *définie sur* $(0,1)$ *telle que* $\int_0^1 \log^+(1/w) f(w)dw < \infty$ *et* $\int_0^1 |K|^{2+\delta}(w)dw < \infty$ *pour* $\delta > 0$. *Alors*

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k} K\left(\frac{i}{n+1}\right)\left(Z_i - \xi + \left(\frac{i}{k+1}\right)^{\beta} b_{n,k}\right)\right)$$

concerge en distribution vers $\mathcal{N}\left(0, \xi^2\int_0^1 K^2(w)dw\right)$.

2.7 Réduction de bias de l'estimateur de Hill

Dans de nombreux cas, l'estimateur de Hill surestime la valeur de ξ dans la population en raison de la convergence lente de $b_{n,k}$ à 0. Certaines propositions d'estimateurs à biais réduit ont récemment été introduites, par exemple dans de Haan et Peng (1998) [41], Feuerverger et Hall (1999) [59]. , Beirlant et al. (1999) [10], Gomes et al. (2000) [71] et Gomes et Martins (2002) [76]. Les dernières références utilisent la représentation exponentielle développée ci-dessus. Encore une fois, on peut aborder le problème à partir de la vue quantile ou de la vue de probabilité.

2.7.1 Approche quantile

La représentation (2.13) peut être considérée comme un modèle de régression généralisé à réponses distribuées de manière exponentielle. Pour chaque i fixé, les réponses Z_i sont distribuées approximativement de manière exponentielle avec $\xi + \left(\frac{i}{k+1}\right)^{\beta} b_{n,k}$. Si $b_{n,k} > 0$ alors les moyennes augmentent avec les valeurs croissantes de i tandis que l'interception est donnée par ξ.

Quelques variantes simples de (2.13) ont été proposées :

$$Z_i \overset{\mathcal{D}}{\sim} \xi \exp\left(d_{n,k}\left(\frac{i}{k+1}\right)^{\beta}\right) E_i, i = 1, ..., k. \tag{2.17}$$

où $d_{n,k} = b_{n,k}/\xi$, en utilisant l'approximation

$$1 + d_{n,k}\left(\frac{i}{k+1}\right)^{\beta} \sim \exp\left(d_{n,k}\left(\frac{i}{k+1}\right)^{\beta}\right).$$

Alternativement, en transformant le modèle linéaire généralisé (2.13) en un modèle de régression à bruit additif (en remplaçant les facteurs aléatoires E_i par leurs valeurs attendues dans le terme de biais), on obtient

$$Z_i \sim \xi + \left(\frac{i}{k+1}\right)^{\beta} b_{n,k} + \xi\,(E_i - 1), 1 \le i \le k. \tag{2.18}$$

Des estimations conjointes de ξ, $b_{n,k}$ (ou $d_{n,k}$) et β peuvent être obtenues pour chaque k à partir de (2.13) et (2.17) par maximum de vraisemblance, ou à partir de (2.18) par la méthode des moindres carrés

$$\sum_{i=1}^{k}\left(Z_i - \xi - \left(\frac{i}{k+1}\right)^{\beta} b_{n,k}\right)^2$$

en ce qui concerne ξ, $b_{n,k}$ et β. Nous désignons l'estimateur du maximum de vraisemblance de ξ basé sur (2.13) par $\hat{\xi}^{+}_{ML}$.

Dans chacun des trois modèles de régression considérés ci-dessus, on peut également résoudre pour ξ et $b_{n,k}$ ou ξ et $d_{n,k}$, après avoir substitué un estimateur cohérent $\hat{\beta} = \beta_{n,k}$ à β.

Par souci de brièveté, nous nous concentrons sur les estimateurs par les moindres carrés basés sur (2.18), conduisant à

$$\begin{aligned}
\hat{\xi}^{+}_{ML}\left(\hat{\beta}\right) &= \bar{Z}_k - \hat{b}^{+}_{ML}\left(\hat{\beta}\right)/\left(1+\hat{\beta}\right) \\
\hat{b}^{+}_{ML}\left(\hat{\beta}\right) &= \frac{\left(1+\hat{\beta}\right)^2\left(1+2\hat{\beta}\right)}{\hat{\beta}^2}\frac{1}{k}\sum_{i=1}^{k}\left(\left(\frac{i}{k+1}\right)^{\hat{\beta}} - \frac{1}{1+\hat{\beta}}\right)Z_i.
\end{aligned}$$

Ici, on va approximé $\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k+1}\right)^{\hat{\beta}}$ par $\frac{1}{1+\hat{\beta}}$ et $\frac{1}{k}\sum_{i=1}^{k}\left(\left(\frac{i}{k+1}\right)^{\hat{\beta}}-\frac{1}{1+\hat{\beta}}\right)^2$ par $\frac{\hat{\beta}^2}{(1+\hat{\beta})^2(1+2\hat{\beta})}$.

Sue ce base le théorème (2.6), Beirlant et al. (2002) [11] montre que la variance asymptotique de $\hat{\xi}^{+}_{ML}$ égale à $\frac{\xi}{k}\left(\frac{1+\beta}{\beta}\right)^2$.

Ici, l'augmentation de la variance par rapport à l'estimateur de Hill n'est pas aussi importante qu'avec $\hat{\xi}^{+}_{ML}$, mais la question qui se pose est celle d'un estimateur du paramètre de second ordre β. Drees et Kaufmann (1998) [52] ont proposé l'estimateur

$$\hat{\beta}=\frac{1}{\log\lambda}\log\frac{H_{\lfloor\lambda^2\tilde{k}\rfloor,n}-H_{\lfloor\lambda\tilde{k}\rfloor,n}}{H_{\lfloor\lambda\tilde{k}\rfloor,n}-H_{\tilde{k},n}}$$

pour $\lambda\in(0,1)$ et avec $\tilde{k}$ tel que $\sqrt{\tilde{k}}b_{n,\tilde{k}}\to\infty$. Un choix adapté de $\tilde{k}$ est choisi. Pour une discussion plus élaborée de l'estimation de β et de plusieurs autres estimateurs de β, nous renvoyons le lecteur à Gomes et al. (2002) [73], Gomes et Martins (2002) [76] et Fraga Alves et al. (2003) [66].

L'estimation de β est connue pour être difficile. Par conséquent, certains auteurs ont proposé de définir $\beta=1$ dans les procédures impliquant la connaissance de β. Les estimateurs résultants offrent un compromis entre la réduction du biais des estimateurs impliquant l'estimation de β et la variance plus faible lorsqu'on utilise, par exemple, l'estimateur de Hill, voir, par exemple, Gomes et Oliveira (2003) [74]. Guillou et Hall (2001) [97] utilisent l'estimateur de $b_{n,k}$ obtenu de (2.18) après avoir fixé $\beta=1$ dans le contexte de règles de sélection adaptative du nombre d'extrêmes k.

Enfin, nous mentionnons que $\hat{\xi}^{+}_{ML}$ et $\hat{\xi}^{+}_{ML}\left(\hat{\beta}\right)$, bien que n'étant pas invariants au sens mathématique du terme, sont déjà beaucoup plus stables sous les décalages que l'estimateur de Hill.

La modification de l'estimateur de Hill proposée par Fraga Alves (2001) [65], invariante par décalage, mentionnée ci-dessus, permet également de créer des tracés stables.

2.7.2 La vue de probabilité

Alternativement, Beirlant et al. (2004) [12] proposent d'utiliser un raffinement de second ordre de la vue de probabilité. En suivant l'approche décrite auparavant où l'estimateur de Hill découle de l'approximation de la distribution conditionnelle des excès relatifs $Y_j:=X_{n-j+1,n}/X_{n-k,n}, j=1,\ldots,k$, par une distribution de Pareto stricte, on peut affirmer que l'estimateur de Hill s'effondrera si cette approximation est mauvaise.

Afin de décrire le départ de

$$F_t(x)=P(X/t\le x\mid X>t)$$

d'une distribution de Pareto stricte, nous utilisons l'hypothèse :

$$\frac{1-F(tx)}{1-F(t)} = x^{-1/\xi}\left(1 + h_{-\tau}(x)\, B(x) + o\left(B(t)\right)\right) \tag{2.19}$$

où $\tau > 0$ et B est une fonction à variations régulière à l'infini avec indice $-\tau$. La condition (2.19) peut être reformuler sous la forme :

$$1 - F_t(x) = x^{-1/\xi}\left(1 - B(t)\,\tau^{-1}\left(x^{-\tau} - 1\right) + o\left(B(t)\right)\right), \text{ lorsque } t \to \infty.$$

En supprimant le terme d'erreur, cela affine l'approximation de Pareto d'origine en une approximation par un mélange de deux distributions de Pareto. L'idée est maintenant d'adapter une distribution de Pareto aussi perturbée aux excès multiplicatifs $Y_j, j = 1, ..., k$, visant une estimation plus précise de la queue inconnue.

Une telle distribution de Pareto perturbée est alors définie par la fonction de survie

$$1 - G(x; \xi, c, \tau) = (1-c)x^{-1/\xi} + cx^{-1/\xi - \tau}$$

pour $c \in (-1/\tau, 1)$ et $x > 1$. Notons que le cas $c = 0$, le mixte coincé avec la distribution de Pareto ordinaire.

Pour $c \downarrow 0$, on peut écrire

$$\begin{aligned} 1 - G(x; \xi, c, \tau) &= \left[x(1 + \xi c(1 - x^{-\tau}))\right]^{-1/\xi} + o(c) \\ &= \left[x\left[(1 + \xi c) - \xi c x^{-\tau}))\right]\right]^{-1/\xi} + o(c). \end{aligned}$$

En pratique, il s'avère que

$$\bar{G}_{PPD}(x) = x^{-1/\xi}\left[(1 + \xi c) - \xi c x^{-\tau}))\right]^{-1/\xi} \tag{2.20}$$

correspond bien à la méthode du maximum de vraisemblance conduisant aux estimateurs $\hat{\xi}^{+}_{PPD}$, $\hat{c}^{+}_{PPD}$ et $\hat{\tau}^{+}_{PPD}$. La surface de vraisemblance peut être considérée comme étant plutôt plate dans τ, de sorte que l'optimisation doit être traitée avec précaution, comparable à l'estimation de β dans le modèle linéaire généralisé (2.13).

Le modèle de Pareto généralisé perturbé (PPD) est une extension du modèle de pareto généralisé (PG), dans le sens suivant.

Dans les statistiques des extrêmes, il est courant d'approximer la distribution des dépassements absolus d'une variable aléatoire Y au-dessus d'un seuil suffisamment élevé u selon la distribution de Pareto généralisée :

$$P(Y - u > y \mid Y > u) = \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi}, y > 0, \sigma > 0 \tag{2.21}$$

remplaçant y par $ux-u$ avec $x \geq 1$ transforme le modèle (2.21) dans un modèle pour des excès relatifs

$$P(Y/u > x \mid Y > u) = \left[x\left(\frac{\xi y}{\sigma} - \left(\frac{\xi y}{\sigma} - 1\right)\frac{1}{x}\right)\right]^{-1/\xi}, y > 0, \sigma > 0$$

qui est équivalent à (2.20) avec $c = u/\sigma - 1/\xi$ et $\tau = 1$.

2.8 Quantiles extrêmes et probabilités de dépassement

2.8.1 Estimation de premier ordre des quantiles et des périodes de retour

Nous discutons d'abord de l'approche simple proposée par Weissman (1978) sur la base de l'estimateur de Hill.

Nous utilisons la méthode d'estimation de l'indice de Pareto basée sur la régression linéaire d'un tracé quantile de Pareto pour dériver un estimateur de $Q(1-p)$. En supposant que la linéarité finale du tracé du quantile de Pareto persiste à partir des k observations les plus grandes sur (jusqu'à l'infini), c'est-à-dire, en supposant que le modèle de Pareto strict persiste au-dessus de ce seuil, nous pouvons extrapoler le long de la ligne avec l'équation

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

approché au point $\left(-\log\frac{k+1}{n+1}, \log(X_{n-k,n})\right)$.

Prenons $x = -\log p$ pour obtenir un estimateur de $Q(1-p)$, noté $\hat{q}^{+}_{n,p}$ donné par

$$\begin{aligned}\hat{q}^{+}_{n,p} &= \exp\left(\log X_{n-k,n} + H_{k,n}\left(\log\frac{k+1}{(n+1)p}\right)\right)\\ &= X_{n-k,n}\left(\frac{k+1}{(n+1)p}\right)^{H_{k,n}}.\end{aligned}$$

Les caractéristiques asymptotiques de cette méthode se retrouvent dans l'extension suivante : depuis

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

et

$$X_{n-k,n} \overset{\mathcal{D}}{=} U^{-\xi}_{k+1,n}\ell_U\left(U^{-1}_{k+1,n}\right)$$

où $U_{j,n}$ est la statistique d'ordre d'un échantillon uniforme standard, on trouve que

$$\begin{aligned}\log\frac{\hat{q}^+_{n,p}}{Q(1-p)} &\overset{\mathcal{D}}{=} \log\left[\left(\frac{U_{k+1,n}}{p}\right)^{-\xi}\left(\frac{\ell_U\left(U^{-1}_{k+1,n}\right)}{\ell_U(1/p)}\right)\left(\frac{k+1}{(n+1)p}\right)^{H_{k,n}}\right]\\ = \log&\left[\left(\frac{U_{k+1,n}}{(k+1)/(n+1)}\right)^{-\xi}\left(\frac{\ell_U\left(U^{-1}_{k+1,n}\right)}{\ell_U(1/p)}\right)\left(\frac{k+1}{(n+1)p}\right)^{H_{k,n}-\xi}\right]\\ &\overset{\mathcal{D}}{=} \xi\left(E_{n-k,n}-\log\frac{n+1}{k+1}\right)+(H_{k,n}-\xi)\log\frac{k+1}{(n+1)p}+\log\frac{\ell_U\left(U^{-1}_{k+1,n}\right)}{\ell_U(1/p)}.\end{aligned}$$

Sous la condition (2.10), on peut approximé le dernier terme par

$$-b_{n,k}\frac{1-\left(\frac{(n+1)p}{k+1}\right)^{\beta}}{\beta}.$$

En utilisant la relation

$$\sqrt{k}\left(E_{n-k,n}-\log\frac{n}{k}\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\text{ lorsque }k,n\to\infty\text{ et }k/n\to 0,$$

avec la représentation exponentielle des espacements mis à l'échelle Z_j, nous trouvons les expressions de la variance asymptotique et du biais de l'estimateur de Weissman dans l'échelle logarithmique lorsque $p=p_n\to 0$ et $np_n\to c>0$ comme $n\to\infty$. On note l'espérance asymptotique de E_∞.

$$\begin{aligned}E_\infty\left(\log\frac{\hat{q}^+_{n,p}}{Q(1-p)}\right) &\sim ABias(H_{n,k})\log\left(\frac{k+1}{(n+1)p}\right)-b_{n,k}\frac{1-\left(\frac{(n+1)p}{k+1}\right)^{\beta}}{\beta}\\ &= \frac{b_{n,k}}{1+\beta}\log\left(\frac{k+1}{(n+1)p}\right)-b_{n,k}\frac{1-\left(\frac{(n+1)p}{k+1}\right)^{\beta}}{\beta}.\end{aligned}\tag{2.22}$$

$$AVar\left(\log\hat{q}^+_{n,p}\right)\sim\frac{\xi^2}{k}\left(1+\log^2\left(\frac{k+1}{(n+1)p}\right)\right)\tag{2.23}$$

De plus, on peut maintenant montrer que lorsque $k,n\to\infty$ et $k/n\to 0$ tels que

$$\sqrt{k}E_\infty\left(\log\frac{\hat{q}^+_{n,p}}{Q(1-p)}\right)\to 0,$$

$$\sqrt{k}\left(1+\log^2\left(\frac{k+1}{(n+1)p}\right)\right)^{-1/2}\left(\frac{\hat{q}^+_{n,p}}{Q(1-p)}-1\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\xi^2\right).\tag{2.24}$$

Alternativement, afin d'estimer $P(X>x)$ pour x large, on peut définir l'estimateur de

Weissman $\hat{q}^{+}_{k,p}$ égal à x et résoudre pour p :

$$\hat{p}^{+}_{k,p} = \left(\frac{k+1}{n+1}\right)\left(\frac{x}{X_{n-k,n}}\right)^{-1/H_{n,k}}.$$

2.8.2 Raffinements de deuxième ordre

2.8.2.1 La vue quantile

En utilisant la condition (2.10), on peut préciser $\hat{q}^{+}_{k,p}$ en exploitant les informations supplémentaires alors disponibles concernant la fonction U à variation lente. En utilisant encore

$$X_{n-k,n} \overset{\mathcal{D}}{=} U\left(1/U_{k+1,n}\right),$$

on trouve que

$$\begin{aligned}
\frac{Q(1-p)}{X_{n-k,n}} &\overset{\mathcal{D}}{=} \frac{p^{-\xi}}{U^{-\xi}_{k+1,n}} \frac{\ell_U(1/p)}{\ell_U(1/U_{k+1,n})} \\
&\sim \left(\frac{U_{k+1,n}}{p}\right)^{\xi} \exp\left(b\left(1/U_{k+1,n}\right) \frac{1-\left(\frac{U_{k+1,n}}{p}\right)^{-\beta}}{\beta}\right) \\
&\sim \left(\frac{k+1}{(n+1)p}\right)^{\xi} \exp\left(b_{n,k} \frac{1-\left(\frac{k+1}{(n+1)p}\right)^{-\beta}}{\beta}\right)
\end{aligned}$$

où dans la dernière étape, nous avons remplacé $U_{k+1,n}$ par sa valeur attendue $(k+1)/(n+1)$. On arrive donc à l'estimateur suivant pour les quantiles extrêmes avec $k = 3, \ldots, n-1$:

$$\hat{q}^{(1)}_{k,p} = X_{n-k,n}\left(\frac{k+1}{(n+1)p}\right)^{\hat{\xi}^{+}_{ML}} \exp\left(\hat{b}_{n,k} \frac{1-\left(\frac{k+1}{(n+1)p}\right)^{-\hat{\beta}}}{\hat{\beta}}\right), \tag{2.25}$$

où $\hat{\xi}^{+}_{ML}$, $\hat{\beta}$ et $\hat{b}_{n,k}$ représentent les estimateurs du maximum de vraisemblance basés sur (2.13).

Cet estimateur a été étudié plus en détail dans Matthys et Beirlant (2003) [125]. Entre autres, il a été prouvé que la distribution asymptotique de $\hat{\xi}^{+}_{ML}$ et de $\hat{q}^{(1)}_{k,p}$ est assez similaire.

En effet, par rapport à (2.24), la variance asymptotique devient maintenant $\xi^2\left(\frac{1+\beta}{\beta}\right)^4$ au lieu de ξ^2 in (2.24). Notez que l'équation (2.25) peut également être utilisée pour estimer les probabilités de petit dépassement. En effet, fixer $\hat{q}^{(1)}_{k,p}$ à un niveau élevé, (2.25) peut être résolu numériquement pour p. L'estimateur résultant pour p sera désigné par $\hat{p}^{(1)}_{k,x}$.

L'effet correcteur de biais obtenu en utilisant $\hat{\xi}^+_{ML}$ et le facteur $\exp\left(\hat{b}_{n,k}\frac{1-\left(\frac{k+1}{(n+1)p}\right)^{-\hat{\beta}}}{\hat{\beta}}\right)$ est illustré à la figure **??** où nous montrons les médianes calculées sur 100 échantillons de taille $n = 1000$ à partir de la distribution de $Burr(1, 0.5, 2)$ et $p = 0,0002$. Ensuite de $\hat{q}^+_{k,p}$ et $\hat{q}^{(1)}_{k,p}$, nous montrons également l'estimateur

$$\hat{q}^{(0)}_{k,p} = X_{n-k,n}\left(\frac{k+1}{(n+1)p}\right)^{\hat{\xi}^+_{ML}}$$

qui est en fait $\hat{q}^+_{k,p}$ avec $H_{k,n}$ remplacé par $\hat{\xi}^+_{ML}$.

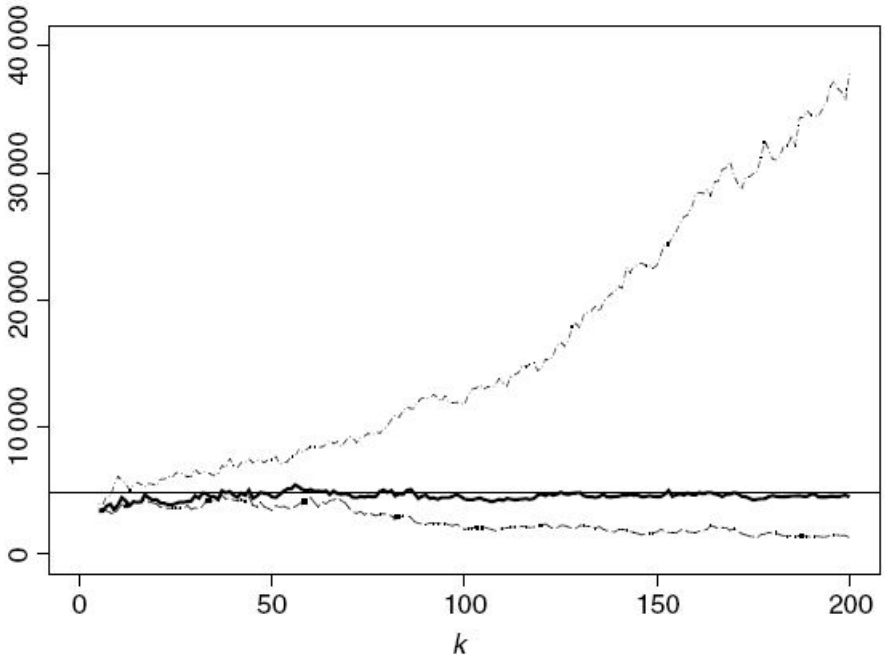

FIGURE 2.1 – Médiane de $\hat{q}^{(1)}_{k,p}$ (ligne continue), $\hat{q}^+_{k,p}$ (ligne pointillée) et $\hat{q}^{(0)}_{k,p}$ (ligne brisée) avec $p = 0,0002$ pour 100 échantillons simulés de taille $n = 1000$ de la distribution Burr (1,0.5,2), $k = 5,...,200$. La ligne horizontale indique la valeur vraie de $Q(1-p)$.

2.9 Sélection adaptative de la fraction d'échantillon de queue

Nous passons maintenant à l'estimation de la fraction optimale nécessaire pour appliquer un estimateur d'indice extrême tel que l'estimateur de Hill. Il devrait être intuitivement clair que les estimations de $b_{n,k}$, le paramètre qui domine le biais de l'estimateur de Hill, comme indiqué à la section précédent, devraient être utiles pour localiser les valeurs de k pour lesquelles le biais de l'estimateur de Hill est trop grand, ou pour lequel l'erreur quadratique moyenne de l'estimateur est minimale. Plusieurs méthodes ont été proposées récemment, que nous passons brièvement en revue. Voir

aussi Hall et Welsh (1985) [103] et Beirlant et al. (1996) [9].

1. Guillou et Hall (2001) [97] proposent de choisir $H_{\hat{k},n}$ où $\hat{k}$ est la plus petite valeur de k pour laquelle
$$\sqrt{\frac{k}{12}}\left|\frac{\hat{b}^{+}_{LS}(-1)}{H_{k,n}}\right| > c_{critique}$$
où $c_{critique}$ est une valeur critique prend 1.25 ou 1.5.
Pour comprendre cette normalisation, remarquons d'abord que sur la base du théorème 2.6, on peut montrer que si $\sqrt{k}b_{n,k} \to c \in \mathbb{R}$, alors
$$\sqrt{\frac{k}{12}}\frac{\hat{b}^{+}_{LS}(-1)}{\xi} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{c\beta\sqrt{3}}{\xi(1+\beta)(2+\beta)}, 1\right).$$
Ainsi, après normalisation appropriée de $\hat{b}^{+}_{LS}(-1)$, la procédure donnée dans Guillou et Hall (2001) [97] peut être considérée comme un test asymptotique pour une espérance nulle (asymptotique) de $\hat{b}^{+}_{LS}(-1)$: le biais dans l'estimateur de Hill est considéré comme trop grand et l'hypothèse du biais nul est donc rejetée lorsque la moyenne asymptotique du résultat limite apparaît significativement différente de zéro.

2. Une alternative importante, populaire parmi les statisticiens, consiste à minimiser l'erreur quadratique moyenne. Ensuite, nous essayons de minimiser l'erreur quadratique moyenne asymptotique de $H_{k,n}$, c'est-à-dire
$$AMSE(H_{k,n}) = AVar(H_{k,n}) + ABias^2(H_{k,n}) = \frac{\xi^2}{k} + \left(\frac{b_{n,k}}{1+\beta}\right)^2, \tag{2.26}$$
comme dérivé avant. Il semble donc naturel d'utiliser les estimateurs de maximum de vraisemblance décrits ci-dessus et de rechercher la valeur de $\hat{k}$, ce qui minimise cette courbe d'erreur quadratique moyenne estimée
$$\{(k, AMSE(H_{k,n})); k = 1, ..., n-1\}.$$

3. Limitons-nous à nouveau aux distributions de Hall-class où la distribution inconnue satisfait
$$U(x) = Cx^{\xi}\left(1 + Dx^{-\beta}(1+o(1))\right), x \to \infty$$
pour certaines constantes $C > 0$, $D \in \mathbb{R}$. Observons que, dans ce cas,
$$b(x) = -\beta D x^{\beta}(1+o(1)), \quad lorsque \quad x \to \infty.$$
Ensuite, l'erreur quadratique asymptotique moyenne de l'estimateur de Hill est

minimale pour

$$k_{n,opt} \sim \left(b^2(n)\right)^{-1/(1+2\beta)} \left(\frac{\xi^2(1+\beta)^2}{2\beta}\right)^{1/(1+2\beta)}, n \to \infty.$$

Ici, à cause de la forme particulière de b, on obtient

$$k_{n,opt} \sim \left(b^2\left(\frac{n}{k_0}\right)\right)^{-1/(1+2\beta)} k_0^{2\beta/(1+2\beta)} \left(\frac{\xi^2(1+\beta)^2}{2\beta}\right)^{1/(1+2\beta)}, n \to \infty. \tag{2.27}$$

pour toute valeur secondaire $k_0 \in \{1,...,n\}$ avec $k_0 = o(n)$. Nous intégrons des estimateurs cohérents de b_{n,k_0}, β et ξ dans cette expression, comme indiqué ci-dessus, tous basés sur les extrêmes k_0 supérieurs. De cette manière, on obtient pour chaque valeur de k_0 un estimateur de $k_{n,opt}$.
Alors, lorsque $k_0, n \to \infty$ et $k_0/n \to 0$ et $\frac{\sqrt{k_0} b_{n,k_0}}{\log k_0} \to \infty$, on a

$$\frac{\hat{k}_{n,k_0}}{k_{n,opt}} \xrightarrow{P} 1.$$

Bien sûr, un inconvénient de cette approche est qu'en pratique, il faut identifier la région k0 pour laquelle $\sqrt{k_0} b_{n,k_0} \to \infty$ afin d'obtenir une méthode cohérente. Cependant, les graphes de $\log \hat{k}_{n,k_0}$ en fonction de k_0 sont assez stables, sauf pour les k_0-régions correspondant à $\sqrt{k_0} b_{n,k_0} \to 0$.
Pour mettre en place une méthode automatique, d'un point de vue pratique, on peut utiliser la médiane des premières $\lfloor n/2 \rfloor \hat{k}$ comme estimation globale de $k_{n,opt}$

$$\hat{k}_{n,med} = mediane\left\{\hat{k}_{n,k_0}, k_0 = 3,...,\lfloor n/2 \rfloor\right\}$$

4. Dans Hall (1990) [104], une nouvelle technique de ré-échantillonnage est proposée pour estimer l'erreur quadratique moyenne de l'estimateur de Hill. À cet effet, le bootstrap habituel ne fonctionne pas correctement, notamment parce qu'il sous-estime gravement les biais.
   Ce problème peut être contourné en prenant des échantillons de taille inférieure à celle d'origine et en liant les estimations bootstrap de la fraction de sous-échantillon optimale à $k_{n,opt}$ pour l'échantillon complet. Cependant, pour établir ce lien, la méthode de Hall requiert que $\beta = 1$, ce qui impose une grave restriction au comportement des données en bout de chaîne. De plus, une estimation initiale est nécessaire pour estimer le biais. Comme l'ont souligné Gomes et Oliveira (2001) [72], l'ensemble de la procédure est très sensible au choix de cette valeur initiale. L'idée découpage de sous-échantillon est reprise dans une méthode plus large par Danielsson et al. (1997) [31]. Au lieu d'améliorer l'er-

reur quadratique moyenne de l'estimateur de Hill lui-même, ils utilisent une statistique auxiliaire, l'erreur quadratique moyenne qui converge au même taux et qui a une moyenne asymptotique connue, indépendante des paramètres ξ et β. Une telle statistique est

$$A_{n,k} = H_{n,k}^{(2)} - 2H_{n,k}^2$$

avec

$$H_{n,k}^{(2)} = \frac{1}{k}\sum_{i=1}^{k}\left(\log X_{n-i+1,n} - \log X_{n-k,n}\right)^2.$$

Comme les deux estimateurs $H_{n,k}^{(2)}/2H_{n,k}$ et $H_{n,k}$ sont consistentes pour ξ, $A_{n,k}$ converge vers 0 pour une valeur intermédiare de k lorsque $n \to \infty$. Ainsi

$$AMSE\left(A_{n,k}\right) = E_\infty\left(A_{n,k}^2\right),$$

et aucune estimation initiale des paramètres n'est nécessaire pour calculer la contrepartie de bootstrap. De plus, la valeur k qui minimise $AMSE\left(A_{n,k}\right)$, notée $\bar{k}_{n,opt}$, est du même ordre en n que $k_{n,opt}$:

$$\frac{\bar{k}_{n,k_0}}{k_{n,opt}} \to \left(1 + \frac{1}{\beta}\right)^{2/(1+2\beta)}, n \to \infty.$$

Malheureusement, l'estimation bootstrap habituelle pour $\bar{k}_{n,opt}$ ne converge pas en probabilité vers la valeur vraie ; elle converge simplement dans la distribution vers une séquence aléatoire en raison de l'équilibre caractéristique entre la variance et le biais carré au seuil optimal. Un bootstrap de sous-échantillon corrige ce problème. Prendre des sous-échantillons de taille $n_1 = O(n^{1-\varepsilon})$ pour quelque $0 < \varepsilon < 1$ fournit une estimation bootstrap cohérente $\widehat{\bar{k}}_{n_1,opt}$ pour $\bar{k}_{n_1,opt}$. De plus, le rapport des fractions optimales de l'échantillon et du sous-échantillon pour $A_{k,n}$ est de l'ordre

$$\frac{\bar{k}_{n,opt}}{\bar{k}_{n_1,opt}} \sim \left(\frac{n}{n_1}\right)^{\frac{2\beta}{1+2\beta}}.$$

Pour $n_1 = O(n^{1-\varepsilon})$ pour quelque $0 < \varepsilon < 0.5$, ce rapport peut être estimé via un deuxième bootstrap de sous-échantillon, désormais avec des sous-échantillons de taille $n_2 = n_1^2/n$, tels que

$$\frac{\bar{k}_{n,opt}}{\bar{k}_{n_1,opt}} \sim \frac{\bar{k}_{n_1,opt}}{\bar{k}_{n_2,opt}}.$$

La combinaison de ces résultats donne

$$k_{n,opt} = \frac{\left(\bar{k}_{n_1,opt}\right)^2}{\bar{k}_{n_2,opt}} \left(1 + \frac{1}{\beta}\right)^{\frac{-2}{1+2\beta}}$$

ce qui conduit à l'estimateur

$$\hat{k}_{n,opt} = \frac{\left(\widehat{\bar{k}}_{n_1,opt}\right)^2}{\widehat{\bar{k}}_{n_2,opt}} \left(1 + \frac{1}{\hat{\beta}}\right)^{\frac{-2}{1+2\hat{\beta}}}, \tag{2.28}$$

pour $k_{n,opt}$, où

$$\hat{\beta}_1 = \frac{\log \widehat{\bar{k}}_{n_1,opt}}{2\log\left(\widehat{\bar{k}}_{n_1,opt}/n_1\right)}$$

est un estimateur consistant de β. Sous la condition $(\mathcal{C}_{-\beta}(b))$ pour le $\log \ell_U$, il est possible de montrer que l'estimateur de Hill résultant $H_{\hat{k}_n,opt,n}$ a la même efficacité asymptotique que $H_{k_n,opt,n}$.
L'algorithme pour cette procédure de bootstrap est résumé comme suit

— Générer un sous-échantillons bootstrap B de taille $n_1 \in \left(\sqrt{n}, n\right)$ à partir de l'échantillon initial et déterminez la valeur $\widehat{\bar{k}}_{n_1,opt}$ qui minimise l'erreur quadratique moyenne bootstrap de A_{k,n_1}.
— Répétez cette opération pour les sous-échantillons de démarrage B de taille $n_2 = n_1^2/n$ et déterminez $\widehat{\bar{k}}_{n_2,opt}$, choisissez l'option où l'erreur quadratique moyenne d'amorçage de A_{k,n_2} est minimale.
— Calculer $\hat{k}_{n,opt}$ pour (2.28) et estimez ξ avec $H_{\hat{k}_{n,opt},n}$.

5. Drees et Kaufmann (1998) [52] présentent une procédure séquentielle pour sélectionner la fraction d'échantillon optimale $k_{n,opt}$. À partir d'une loi du logarithme itéré, ils construisent des 'temps d'arrêt' pour la séquence $H_{k,n}$ des estimateurs de Hill asymptotiquement équivalents à une séquence déterministe. Une combinaison ingénieuse de deux tels temps d'arrêt atteint alors le même taux de convergence que $k_{n,opt}$. Cependant, le facteur de conversion pour passer de cette combinaison de temps d'arrêt à $k_{n,opt}$ implique les paramètres inconnus ξ (qui nécessite une estimation initiale $\hat{\xi}_0$ et β. Nous nous référons au document original de Drees et Kaufmann (1998) [52] pour les principes théoriques à la base de cette procédure et décrivons immédiatement l'algorithme avec les choix de paramètres de nuisance proposés par ces auteurs.

— Obtenir une estimation initiale $\hat{\xi}_0 := H_{2\sqrt{n},n}$ pour ξ.

— Pour $r_n = 2.5\hat{\xi}_0 n^{0.25}$, calculez le «temps d'arrêt»

$$\hat{k}_n(r_n) = \min\left\{k \in \{1,...,n-1\}, \max_{1\le i\le k} \sqrt{i}\left(H_{i;n} - H_{k,n}\right) > r_n\right\}.$$

— De même, calcul $\hat{k}_n(r_n^\varepsilon)$ for $\varepsilon = 0.7$.
— Avec un estimateur cohérent $\hat{\beta}$ pour β, calculez

$$\hat{k}_{n,opt} = \left(\frac{\hat{k}_n(r_n^\varepsilon)}{\left[\hat{k}_n(r_n)\right]}\right)^{\frac{1}{1-\varepsilon}} \left(1+2\hat{\beta}\right)^{-\frac{1}{\hat{\beta}}} \left(2\hat{\beta}\hat{\xi}_0\right)^{\frac{1}{1+2\hat{\beta}}},$$

et estimé ξ par $H_{\hat{k}_{n,opt},n}$.
Dans les simulations, il a été constaté que la méthode fonctionnait généralement mieux si une valeur fixe β_0 était utilisée pour β dans (**??**), en particulier pour $\hat{\beta} = \beta_0 = 1$.

2.10 Exemple des distributions de type de Pareto (distributions des revenues)

2.10.1 Loi de Pareto

La loi de Pareto s'applique pour les distributions tronquées. Prenons un exemple de la vie courante, en France, la borne basse du salaire horaire est forcément le SMIG, il ne peut pas en être autrement. La loi de Pareto permet de tenir compte de cette contrainte en restreignant le domaine de définition de la v.a. X .

Définition 2.2 *La loi possède deux paramètres, $\alpha > 0$ et c qui introduit la contrainte $x > c$. Le domaine de définition de X est $]c, +\infty[$. La fonction de densité est monotone décroissante, elle s'écrit*

$$f(x) = \frac{\alpha}{c}\left(\frac{c}{x}\right)^{\alpha+1}, x > c. \tag{2.29}$$

La fonction de répartition est directement obtenue avec

$$F(x) = 1 - \left(\frac{c}{x}\right)^{\alpha} \tag{2.30}$$

Caractéristiques de la loi

$$E(X) = \frac{\alpha}{\alpha-1}, \text{ pour } \alpha > 1 \tag{2.31}$$

$$V(X) = \frac{\alpha}{(\alpha-1)^2(\alpha-2)}c^2, \text{ pour } \alpha > 2 \tag{2.32}$$

$$E[X^k] = \frac{\alpha}{\alpha-c}, \text{ pour } \alpha > c \tag{2.33}$$

2.10.2 Loi de Burr

En théorie des probabilités, en statistique et en économétrie, la **loi de Burr**, loi de Burr de type XII, loi de Singh-Maddala, ou encore loi log-logistisque généralisée est une loi de probabilité continue dépendant de deux paramètres réels positifs c et k. Elle est communément utilisée pour étudier les revenus des ménages

Si X suit une loi de Burr (ou Sigh-Maddala), on notera $X \rightsquigarrow SM(c,k)$.

Caractérisation

La densité de peobabilité de la loi de Burr est donnée par :

$$f(x;c,k) = \begin{cases} ck\frac{x^{c-1}}{(1+x^c)^{k+1}} & si\ x > 0 \\ 0 & sinon \end{cases} \tag{2.34}$$

et sa fonction de répartition est :

$$F(x;c,k) = \begin{cases} 1-(1+x^c)^{-k} & si\ x > 0 \\ 0 & sinon \end{cases} \tag{2.35}$$

Si $c = 1$, la loi de $Burr$ est la Distribution de $Pareto$.

2.10.3 Loi de Fréchet

La densité de probabilité de la distribution de paramètre $\alpha > 0$, peut être généralisée en introduisant un paramètre de position m du minimum et un paramètre d'échelle $s > 0$ prend la forme

$$f(x) = \begin{cases} \frac{\alpha}{s}\left(\frac{x-m}{s}\right)^{-1-\alpha} e^{-\left(\frac{x-m}{s}\right)^{-\alpha}} & si\ x > m \\ 0 & sinon \end{cases}$$

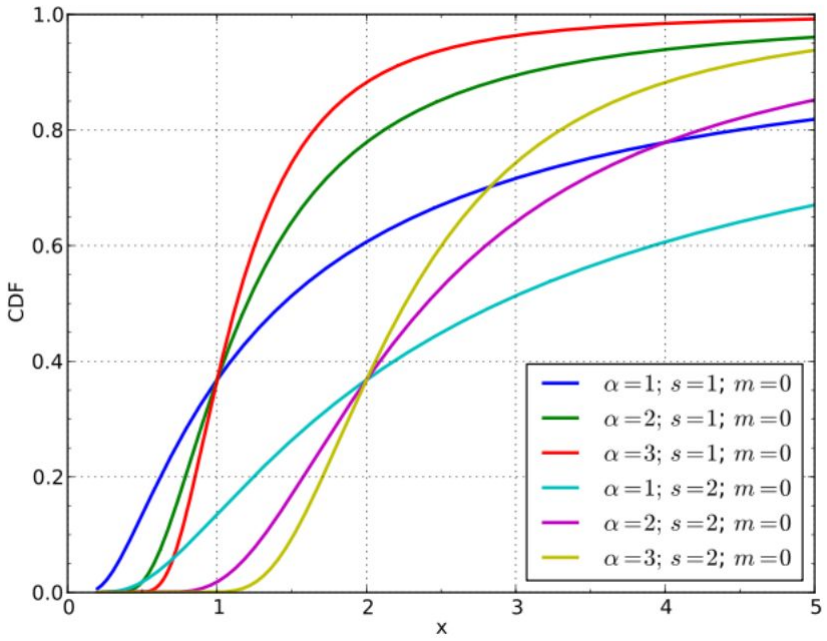

FIGURE 2.2 – Fonction de densité de la Loi de Fréchet.

Définition 2.3 *Sa fonction de répartition est est donnée par :*

$$\mathbb{P}(X \leq x) = \begin{cases} e^{-x^{-\alpha}} & si\ x > 0 \\ 0 & sinon \end{cases}$$

où $\alpha > 0$ est un paramètre de forme. Cette loi peut être généralisée en introduisant un paramètre de position m du minimum et un paramètre d'échelle $s > 0$. La fonction de répartition est alors :

$$\mathbb{P}(X \leq x) = \begin{cases} e^{-(\frac{x-m}{s})^{-\alpha}} & six > m \\ 0 & sinon \end{cases} \tag{2.36}$$

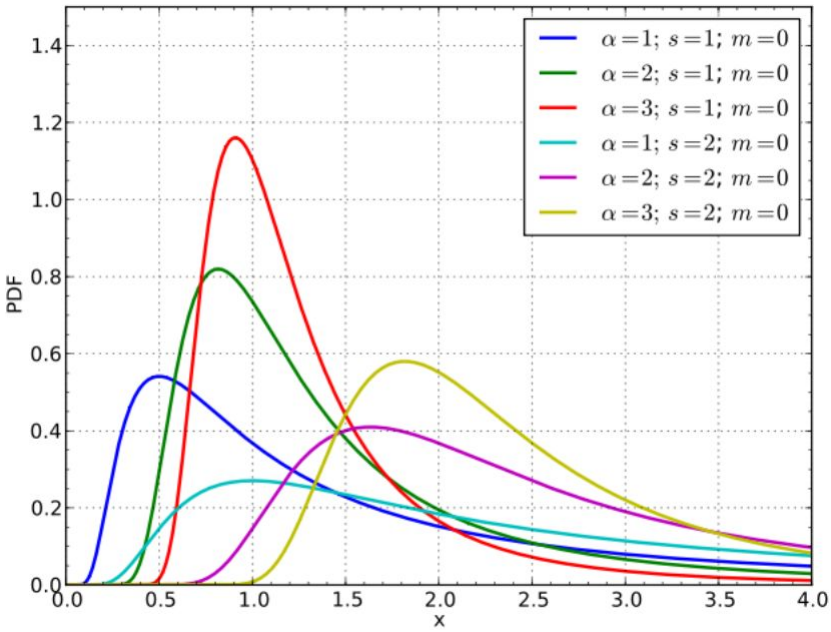

FIGURE 2.3 – Fonction de répartition de la Loi de Fréchet.

Propriétés :

Moments :

La loi de fréchet à un paramètre α a des moments standards :

$$\mu_k = \int_0^\infty x^k f(x)\,dx = \int_0^\infty t^{-\frac{k}{\alpha}} e^{-t} dt \tag{2.37}$$

(avec $t = x^{-\alpha}$) définis pour $k < \alpha$:

$$\mu_k = \Gamma\left(1 - \frac{k}{\alpha}\right) \tag{2.38}$$

où $\Gamma(z)$ est la fonction Gamma.

En particulier :

— Pour $\alpha > 1$ l'**espérance** est

$$\mathbb{E}[X] = \Gamma\left(1 - \frac{1}{\alpha}\right) \tag{2.39}$$

— Pour $\alpha > 2$ la **variance** est

$$Var(X) = \Gamma\left(1 - \frac{2}{\alpha}\right) - (\Gamma\left(1 - \frac{1}{\alpha}\right))^2. \tag{2.40}$$

2.10.3.1 Quantiles

Le **quantile** Q_Y d'ordre y peut être exprimé grâce à l'inverse de la fonction de répartition :

$$Q_Y = F^{-1}(y) = (-\log_e y)^{-\frac{1}{\alpha}}. \tag{2.41}$$

En particulier la **médiane** est :

$$Q_{1/2} = (\log_e 2)^{-\frac{1}{\alpha}}. \tag{2.42}$$

Le **mode** de la loi de *Fréchet* est :

$$\left(\frac{\alpha}{\alpha+1}\right)^{\frac{1}{\alpha}}. \tag{2.43}$$

Pour la loi de Fréchet à trois paramètres, le premier quartile est $Q_1 = m + \frac{s}{\sqrt[\alpha]{\log(4)}}$ est le troisième quartile est $Q_3 = m + \frac{s}{\sqrt[\alpha]{\lg(\frac{4}{3})}}$.

Chapitre 3

Indices des Inégalités

3.1 Introduction

Au cours des dernières décennies, les inégalités ont joué un rôle important dans de nombreuses branches des sciences sociales, principalement la sociologie et l'économie, étant l'une des clés les enjeux du discours sur le bien-être des sociétés et des individus . Ainsi, il apparaît une question de comment mesurer ces inégalités de manière appropriée. Il existe de nombreux indices d'inégalité dans la littérature le plus populaire, à savoir, l'indice de Gini, l'indice de Theil et la mesure d'Atkinson . De ceux-ci, l'indice de Gini est le plus souvent utilisé et également mieux connu des non-scientifiques. Récemment, un nouvel indice d'inégalité a été proposé par Zenga. Il a toutes les propriétés qui sont généralement requises pour les mesures d'inégalité. Afin de décider quelle mesure d'inégalité est la plus appropriée pour un sujet donné, il serait utile d'étudier et de comparer les propriétés de différents indices.

3.2 Définition de revenu

Définition 3.1 ***Etymologie*** *: de revenu, composé du préfixe, indiquant un retour à un état initial et du latin venire, aller, venir, arriver.*

- *En économie, un revenu est l'ensemble des ressources ou droits qu'un individu, une entreprise ou une collectivité publique, perçoit sur une période donnée, en nature ou en monnaie, sans prélever sur son patrimoine.*
- ***Synonymes :*** *allocation, gain, pension, produit, rente, rétribution, salaire.*
- *Contrairement au patrimoine qui est un stock de biens détenus à un instant donné, le revenu est un flux de biens et services dont on dispose pendant une période donnée. En outre, pour parler de revenus récurrents, par opposition aux revenus exceptionnels, ceux-ci doivent se répéter de période en période.*

On distingue les sommes perçues au titre de :

— la rémunération pour un travail. Ex : salaire.
— du patrimoine. Ex : loyers d'immeubles, produit d'un capital placé (intérêts, dividendes, redevance d'utilisition de brevet).
— l'activité : Services rendus et produits fournis par les professionnels et entreprises.
— des prestations et transferts sociaux. Ex : indememnités de chômage, allocations sociales...

Définition 3.2 *1. **Le revenu salarial** correspond à la somme de tous les salaires perçus par un individu au cours d'une année donnée, nets de toutes cotisations sociales, y compris Contribution Sociale Généralisée (CSG) et contribution au remboursement de la dette sociale (CRDS)".*

2. ***Le revenu net** est le revenu brut diminué des dépenses occasionnées pour sa perception (frais professionnels, entretretien d'un patrimoine, etc...).*

3. ***Le revenu disponible** d'un ménage est l'ensemble de ses revenus d'activité, de son patrimoine, et des prestations et transferts sociaux perçus, nets des impôts directs (impôt sur le revenu, taxe d'habitation, CSG, CRDS).*

4. ***Le revenu réel** correspond au pouvoir d'achat réel, c'est-à-dire en tenant compte des variations des prix des biens et des services.*

5. ***Le revenu national brut** (RNB) est la somme des revenus perçus, pendant une période donnée, par les agents économiques résidant sur le territoire national. Il est la somme du PIB et du solde des flux derevenus primaires avec le reste du monde.*

6. ***Le revenu par tête** (ou RNB par habitant) est le revenu national brut (RNB) annuel, divisé par le nombre total d'habitants, pour un pays ou une région donnée.*

3.3 Courbe de Lorenz

La courbe de Lorenz (économiste américain, 1880-1962) est une représentation graphique qui permet de visualiser graphiquement la répartition des concentrations entre individus et masses. On calcule les fréquences cumulées des effectifs (qu'on notera p_i) et celles des masses (qu'on notera q_i). On place sur graphe les points de coordonnées (p_i, q_i) et on les joint par une ligne polygonale. Cette ligne part du point $(0,0)$ et se termine au point $(1,1)$ puisque les fréquences cumulées varient toujours de 0 à 1. Elle est donc inscrite dans le carré de côté 1, parfois appelé le carré de Gini.

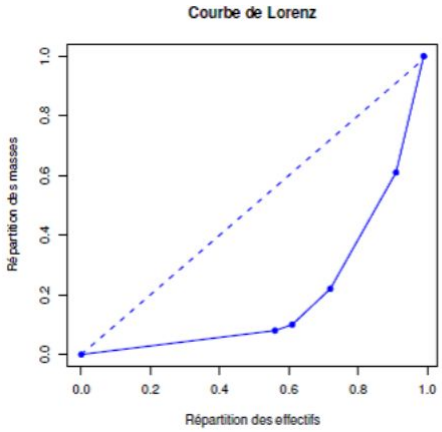

FIGURE 3.1 – Courbe de Lorenz.

Algèbriquement, on a la relation suivante pour la fréquence cumulée des effectifs :

$$p_i = \frac{1}{N}\sum_{j=1}^{i} n_j = \frac{1}{N}(n_1 + n_2 + ... + n_i) \tag{3.1}$$

avec $N = n_1 + n_2 + ... + n_k$.

De même, on a la relation suivante pour la fréquence cumulée des masses $n_i v_i$:

$$q_i = \frac{1}{T}\sum_{j=1}^{i} n_j v_j = \frac{1}{T}(n_1 v_1 + n_2 v_2 + ... + n_i v_i) \tag{3.2}$$

avec $T = n_1 v_1 + n_2 v_2 + ... + n_k v_k$.

Par convention, on pose $p_0 = q_0 = 0$.

Un point de coordonnées (p, q) sur la courbe de Lorenz indique que p% des individus se partagent q% de la masse.

La bissectrice du carré est la ligne d'équirépartition. C'est ce que serait la courbe de concentration s'il y avait équirépartition des masses. Sur cette diagonale, en tout point, p% des individus se partageraient exactement p% de la masse. Dans ce cas, la concentration est nulle.

Définition 3.3 *L'aire de concentration est la région comprise entre le diagonale et la courbe de Lorenz.*

Interprétation : plus cette aire est importante, c'est-à-dire plus la coube de concentration s'écarte de la bissectrice, plus la concentration est forte

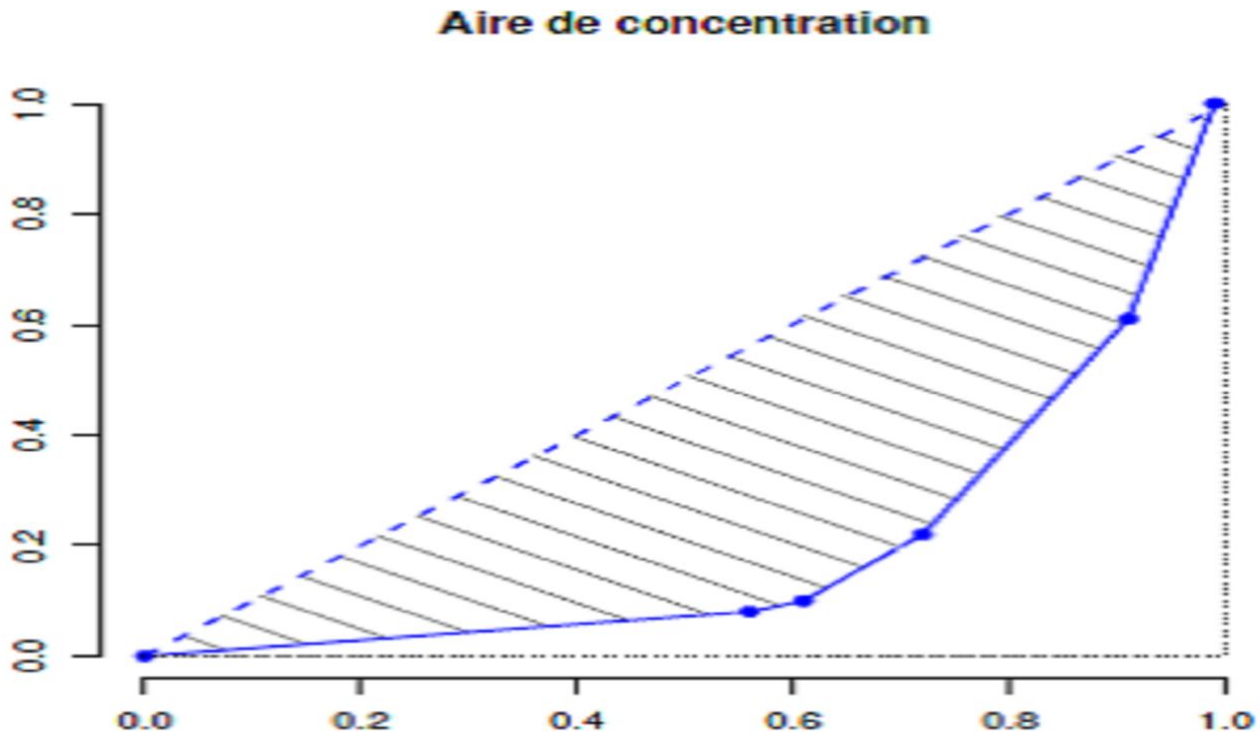

FIGURE 3.2 – Aire de Concentration.

$$L(p) = \frac{1}{\mu} \int_0^p Q(q)\, dq \tag{3.3}$$

avec $Q(q)$ représentant la répartition des revenus, μ le revenu moyen et pour des valeurs de p variant de 0 à 1. Lorsque $L(0,5) = 0,3$ on en déduira que 50 % des individus les plus modestes possèdent 30 % du revenu total.

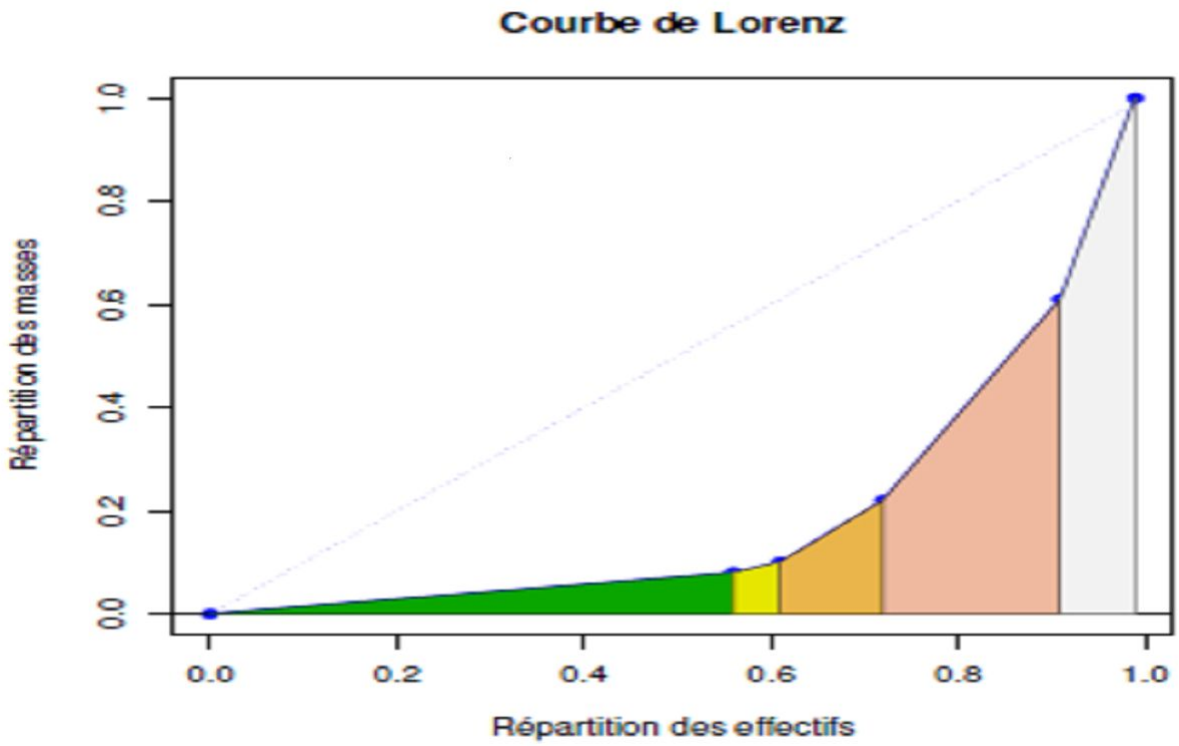

FIGURE 3.3 – Courbe de Lorenz et répartition des effectifs et des masses.

3.4 Indice de Gini

Le concept Gini ou la différence moyenne du Gini, initiée par Gini en 1912, est une caractéristique de dispersion très répandue dans le domaine de distribution des revenus. La spécificité de cet indicateur réside dans ses calculs simples. L'indice de Gini utilise la distance euclidienne entre toutes les paires de l'échantillon [Gini (1912,1914)]. Les interprétations qui en découlent sont faciles comparativement à la variance qui élève au carré les écarts entre les individus.

Il est à noter qu'il existe plusieurs approches qui dérivent du Gini, allant des statistiques de dispersion aux approches de régression, plus précisément la théorie des valeurs aberrantes. [Lerman et Yitzhaki (1989) [120], Yitzhaki et Schechtman (2013) [170]].

L'indice de Gini entretient un lien strict avec la représentation de l'inégalité des revenus à l'aide de la courbe de Lorenz. En particulier, il mesure le ratio entre l'aire située entre la courbe de Lorenz et la droite d'équidistribution (et donc l'aire de concentration) et l'aire de concentration maximale.

La figure (3.4) représente ces aires : elle trace trois courbes de Lorenz à partir de trois distributions de revenus hypothétiques O, P et Q.

La courbe basée sur la distribution des revenus O est la courbe standard que donne

l'analyse des distributions de revenus réelles.

Celle de la distribution P représente le cas extrême où tous les revenus sont égaux. Dans ce cas, elle prend aussi le nom de droite d'équidistribution.

Enfin, la courbe de la distribution Q illustre un autre cas extrême, celui où tous les revenus sont nuls, sauf le dernier.

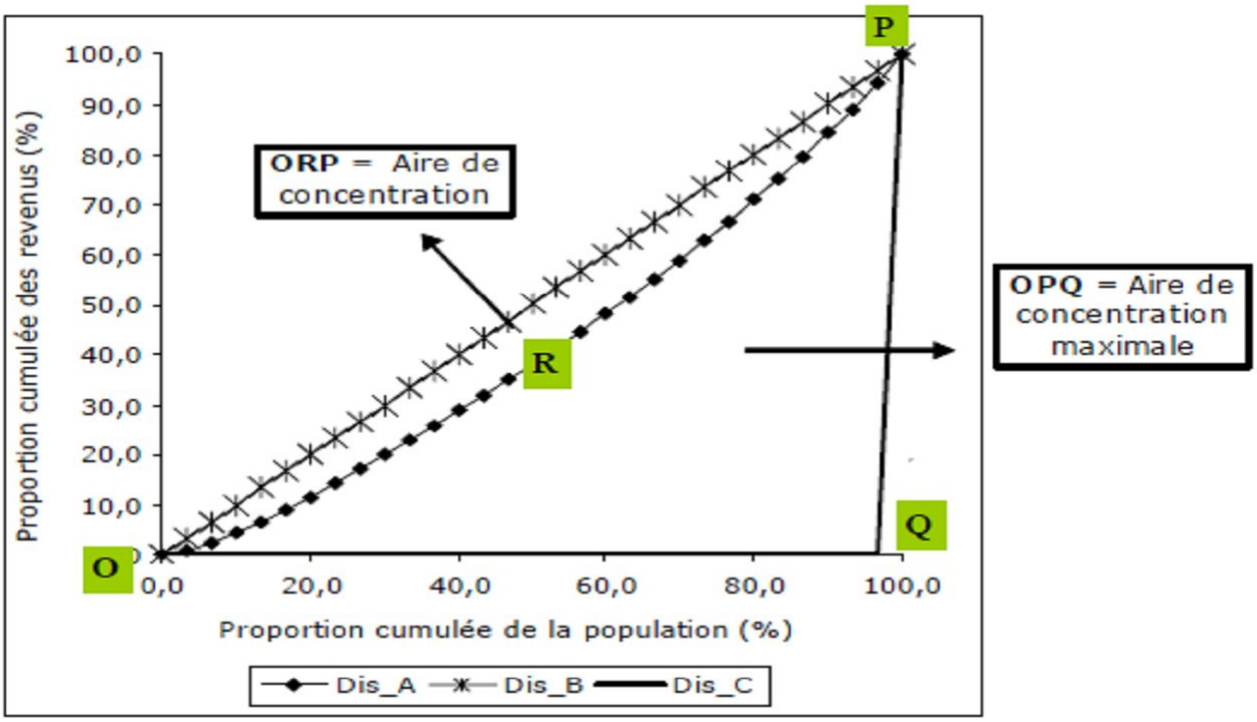

Figure 3.4 – Courbe de Lorenz et indice de Gini.

Dans la (3.4), OP est la droite d'équidistribution et ORP l'aire définie par la courbe de Lorenz de la distribution des revenus standard et la courbe d'équidistribution, baptisée aire de concentration. OPQ est l'aire de concentration maximale, c'est-à-dire la zone entre la courbe de Lorenz de la distribution de revenus Q et la droite d'équidistribution.

La droite d'équidistribution OP et l'aire OPQ représentent les valeurs extrêmes de l'aire de concentration dans une courbe de Lorenz. Soit cette aire est nulle (comme dans le cas de la droite d'équidistribution de la distribution P), soit elle est maximale (cas de la distribution Q). Pour une distribution des revenus standard, l'aire de concentration se situe quelque part entre zéro et l'aire de concentration maximale, comme dans la figure (3.4).

L'indice de Gini mesure le ratio entre l'aire de concentration et l'aire de concentra-

tion maximale. Par conséquent, dans la figure (3.4) :

$$\mathbf{G} = \frac{\text{aire de concentration}}{\text{aire de concentration maximale}} = \frac{ORP}{OPQ}. \tag{3.4}$$

Comme l'aire de concentration maximale correspond à une distribution où un seul individu détient la totalité des revenus, l'indice de Gini **G** mesure en général la distance entre l'aire définie par une quelconque distribution de revenus standard et l'aire de concentration maximale.

Il faut maintenant comprendre comment s'applique la formule de la figure (3.4) dans la pratique. Commençons par le dénominateur de **G**. Que les coordonnées maximales de la courbe de Lorenz se situent au point $(1,1)$. Par conséquent, l'aire OPQ doit être un triangle possédant une longueur de base de 1 et une hauteur de 1. Son aire est donc égale à $1/2$. Le dénominateur de **G** est donc $1/2$.

Mais qu'en est-il du numérateur ? Au lieu de calculer directement l'aire de concentration, nous pouvons exploiter le fait que cette aire représente la différence entre l'aire de concentration maximale et l'aire sous la courbe de Lorenz (cette dernière étant donnée par $ORPQ$). Le mode de calcul le plus facile de l'aire sous la courbe de Lorenz est décrit ci-après.

Commençons par rappeler la définition des coordonnées de la courbe de Lorenz. Si $y_1 \leq y_2 \leq \ldots \leq y_n$:

$$q_i = \frac{y_1 + y_2 + \ldots + y_i}{y_1 + y_2 + \ldots + y_n} = \frac{y_1 + y_2 + \ldots + y_i}{Y} \rightarrow \textit{proportion cumulée des revenus} \tag{3.5}$$

$$p_i = \frac{i}{n} \rightarrow \textit{proportion cumulée de la population} \tag{3.6}$$

où $q_0 = p_0 = 0$ et $q_n = p_n = 1$.

L'aire sous la courbe de Lorenz $ORPQ$ est la somme des aires d'une série de polygones. Regardons la figure (3.5), où une courbe de Lorenz simplifiée a été crée pour une population de quatre individus. Le premier polygone est un triangle (PQO) et les trois autres sont des trapèzes isocèles pivotés. On peut donc calculer chaque aire séparément et ajouter les résultats obtenus pour obtenir la valeur de l'aire globale. Symbolisons l'aire du ième polygone par Z_i et l'aire totale obtenue de cette manière par Z.

L'aire du triangle est donnée par :

$$Z_1 = \frac{\overbrace{p_1}^{base} \quad \overbrace{q_1}^{hauteur}}{2} \tag{3.7}$$

tandis que l'aire de chaque trapèze est donnée par :

$$Z_i = \frac{\overset{\text{(base longue + base coute) } \times \text{hauteur}}{(q_i + q_{i-1})(\overline{p_i} - \overline{p_{i-1}})}}{2} \tag{3.8}$$

Comme $q_0 = p_0 = 0$, la somme de toutes ces aires donne :

$$Z = \sum_{i=1}^{n} Z_i = \frac{1}{2} \sum_{i} [(q_i + q_{i-1})(p_i - p_{i-1})] \quad \text{pour } n = 4. \tag{3.9}$$

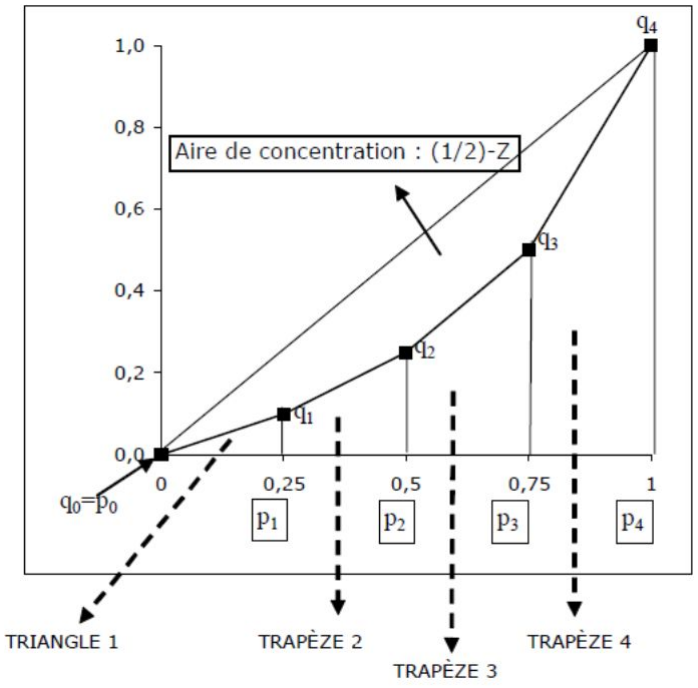

FIGURE 3.5 – Mode de calcul de l'aire de concentration.

Cependant, Z n'est pas l'aire de concentration, mais l'aire sous la courbe de Lorenz. Pour calculer l'aire de concentration (numérateur de l'indice de Gini), il suffit maintenant de soustraire Z de l'aire de concentration maximale (1/2) comme suit :

$$Aire\ de\ concentration = \frac{1}{2} - Z = \frac{1}{2} - \frac{1}{2} \sum_{i} [(q_i + q_{i-1})(p_i - p_{i-1})] \tag{3.10}$$

Selon (4), l'indice de Gini G est donc égal à :

$$G = \frac{\frac{1}{2} - \frac{1}{2} \sum_i [(q_i + q_{i-1})(p_i - p_{i-1})]}{\frac{1}{2}} = 1 - \sum_{i} [(q_i + q_{i-1})(p_i - p_{i-1})] \tag{3.11}$$

que l'on peut également écrire :

$$G = 1 - 2Z. \tag{3.12}$$

La formule ci-dessus indique seulement que l'indice de Gini est égal à 1 moins deux fois l'aire sous la courbe de Lorenz.

Cette interprétation géométrique basée sur la courbe de Lorenz ne constitue que l'un des modes de calcul possibles de l'indice de Gini. Une autre approche, qui va s'avérer particulièrement utile ci-après, consiste à exprimer directement l'indice de Gini en termes de covariance entre les niveaux de revenus et la distribution cumulée des revenus.

En particulier :

$$G = Cov(y, F(y))\frac{2}{y} \tag{3.13}$$

où Cov représente la covariance entre des niveaux de revenus y et la distribution cumulée des mêmes revenus $F(y)$ et où $\overline{y}$ est le revenu moyen. Il est utile de rappeler ici que la covariance est la valeur attendue E des produits des écarts sur la moyenne de chaque variable. Soit dans ce cas précis :

$$Cov\,[y, F(y)] = E([y - \overline{y}].\left|F(y) - \overline{F(y)}\right|). \tag{3.14}$$

En général, l'index de Gini est une fonction $G : \mathbb{R}^n_+ \to [0,1]$ qui attribue à chaque vecteur de revenu non négatif un nombre réel compris entre 0 et 1, ce qui représente le niveau d'inégalité de la société. Cette mesure est 0 en égalité maximale et 1 en parfaite inégalité. La dénotation attrayante de l'indice de Gini est le double de la surface entre la ligne d'égalité et la courbe de Lorenz dans la boîte de l'unité. La ligne à 45° représente l'égalité parfaite des revenus et la zone située entre cette ligne et la courbe de Lorenz est appelée zone de concentration. Par conséquent, l'indice de Gini peut être exprimé comme

$$\mathbf{G} = 2\int_0^1 (p - L(p))\,dp, \tag{3.15}$$

tel que $p = F(x)$ est une fonction de distribution cumulative (fdc) non-négatif revenu avec espérance positive et négative μ, $L(p)$ la fonction de Lorenz donnée par

$$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(t)\,dt, \text{ où } F^{-1}(t) = \inf\{x \mid F(x) \geq p : p \in [0,1]\}. \tag{3.16}$$

En utilisant la définition de l'indice de Gini dans l'équation (3.15), soit le double de la surface entre la droite d'égalité et la courbe de Lorenz, et en appliquant un changement de variable $p = F(x)$, on peut constater que :

$$\mathbf{G} = \frac{2}{\mu}\int_0^\infty xF(x)\,dF(x) - 1. \tag{3.17}$$

Supposons qu'un échantillon i.i.d de taille n est tiré au hasard de la population, et $\hat{F}$

désigne la fonction de distribution empirique correspondante. Soit $X_1,...,X_n$ être un échantillon aléatoire et $X_{1:n} \leq ... \leq X_{n:n}$ les statistiques d'ordre obtenues à partir de l'échantillon. En suite, un estimateur alternatif du coeffcient de Gini peut être obtenu par le fdc empirique $\left(\hat{F}\right)$ du revenu au lieu de sa fonction distribution correspondante F dans (3.17) comme :

$$\hat{\mathbf{G}} = \frac{2}{\hat{\mu}} \int_0^{\infty} x\hat{F}(x)\, d\hat{F}(x) - 1 \tag{3.18}$$

à cet égard, l'exemple de l'indice de Gini peut être exprimé comme

$$\begin{aligned} \tilde{\mathbf{G}} &= \frac{2}{\hat{\mu}} \int_0^{\infty} x d\hat{F}(x))^2 - 1, \\ &= \frac{2\sum_{i=1}^{n} X_{i:n}\left(i - \frac{1}{2}\right)}{n\sum_{i=1}^{n} X_i} - 1. \end{aligned} \tag{3.19}$$

Davidson (2009) a trouvé une expression approchée du biais de $\hat{\mathbf{G}}$ à partir du quel il a dérivé l'estimateur à correction de biais du coefficient de Gini noté $\tilde{\mathbf{G}}$, lequel est donné par :

$$\tilde{\mathbf{G}} = \frac{n}{n-1}\hat{\mathbf{G}}, \tag{3.20}$$

alors que l'estimateur (3.20) est toujours biaisé mais son biais est d'ordre n^{-1}, il est parfois recommandé d'utiliser cet estimateur parceque l'estimateur correctement corrigé du biais est non seulement plus facile à calculer que les autres estimateurs mais aussi son biais converge vers 0 plus vite que $n \to \infty$.

L'indice de Gini généralisé a été popularisé par les travaux de Donaldson et Weymark (1980) [50], et de Yitzhaki (1983) [169] :

$$I_\rho = \frac{\mu - \zeta_\rho}{\mu} \tag{3.21}$$

avec

$$\zeta_\rho = \sum_{i=1}^{L} \left[\frac{(R_i)^\rho - (R_{i+1})^\rho}{(R_i)^\rho}\right] y_i \tag{3.22}$$

et

$$R_i = \sum_{i=1}^{L} \omega_i \tag{3.23}$$

où μ représente la moyenne des revenus, ω_i et y_i le poids et le niveau de revenu de

l'individu, et ρ le paramètre d'aversion à l'inégalité.

3.5 Indicateur de Theil

L'indice de Theil (1967) [162] mesure l'écart entre le poids d'un individu (ou d'un groupe) dans la population et le poids de son revenu dans le revenu total . L'indice de Theil repose sur le concept physique d'entropie [Figini, 1998]. Cet indice correspond à la variation d'entropie entre la situation parfaitement égalitaire et la situation réelle. « En thermodynamique, l'entropie définit l'état de désordre d'un système, croissant lorsque celui-ci évolue vers un état de désordre accru » (Petit Robert, 1986). Sa valeur varie entre 0, la situation d'égalité et $logN$, dans le cas où tous les revenus sont nuls, sauf un.
Soit y_i le revenu de l'individu i appartenant à une population de N individus et μ le revenu moyen, l'indice s'écrit :

$$\mathbf{T} = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{\mu}(\log(\frac{y_i}{\mu})). \tag{3.24}$$

Cet indice accorde un peu plus d'importance à l'inégalité dans le bas de la distribution qu'à l'inégalité parmi les riches. Moins couramment utilisé que l'indice de Gini, l'indice de Theil présente néanmoins des atouts pratiques incontestables. Son principal intérêt est de pouvoir se décomposer à l'infini en partitionnant la population puis en redécomposant chacun des groupes en différents sous-groupes, cela afin d'analyser l'évolution des inégalités dans et entre différentes sous-populations. Cependant son expression mathématique, qui utilise la forme logarithmique, limite son usage à des valeurs non nulles.

3.6 Indice d'Atkinson (1970)

L'indice d'Atkinson est un indice de l'inégalité des revenus basé sur la théorie économique. Afin de définir son indice d'inégalité, Atkinson (1970) [5] suppose que le bien-être dans la société peut être évalué à partir de l'équation suivante :

$$W = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}\frac{x_i^{1-\varepsilon}}{1-\varepsilon}, & pour \quad \varepsilon \neq 1 \\ \frac{1}{n}\sum_{i=1}^{n}\ln x_i & \text{pour} \quad \varepsilon = 1 \end{cases} \tag{3.25}$$

Le paramètre ε décrit l'aversion de la société pour l'inégalité.

— Si $\varepsilon = 0$ il n'y a aucune aversion à l'inégalité. Qu'il soit distribué à un riche ou à un pauvre, un euro supplémentaire augmente pareillement le bien-être social

W.

— Si $\varepsilon \to \infty$ l'aversion à l'inégalité est extrême et le bien-être social est confondu avec le bien-être de l'individu le plus pauvre (Rawls)

Les indices d'inégalité associés à la fonction de bien-être social d'Atkinson s'écrivent :

$$I_A = \begin{cases} 1 - \left(\frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i}{\mu} \right)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}} & si \quad \varepsilon \neq 1 \\ 1 - \prod_{i=1}^{n} \left(\frac{x_i}{\mu} \right)^{\frac{1}{n}} & si \quad \varepsilon = 1 \end{cases} \tag{3.26}$$

où μ représente la moyenne des revenus observés. A l'aide de l'utilisation de la moyenne généralisée

$$M^{\varepsilon} = \begin{cases} \sqrt[\varepsilon]{\frac{1}{n} \sum_{i=1}^{n} x_i^{\varepsilon}}, & si \quad \varepsilon \neq 0 \\ \prod_{i=1}^{n} x_i, & si \quad \varepsilon = 0 \end{cases}, \tag{3.27}$$

l'indice d'Atkinson peut s'exprimer par :

$$A_{\varepsilon} = \frac{M^1 - M^{1-\varepsilon}}{M^1} = 1 - \frac{M^{1-\varepsilon}}{M^1}. \tag{3.28}$$

L'indice d'Atkinson est donc fonction du paramètre ε. Pour cette étude, nous avons utilisé l'indice pour deux valeurs du paramètre : $\varepsilon = 0.5$ et $\varepsilon = 1$:

$$\begin{aligned} A_{1/2} &= 1 - \frac{\left(\frac{1}{n} \sum_{i=1}^{n} \sqrt{x_i} \right)^2}{\mu}, \\ A_1 &= 1 - \prod_{i=1}^{n} \left(\frac{x_i}{\mu} \right)^{\frac{1}{n}} = 1 - \frac{G}{\mu}, \end{aligned} \tag{3.29}$$

avec G la moyenne géometrique :

$$G = \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}}. \tag{3.30}$$

En pratique on interprète le coefficient ε en remarquant que plus ce paramètre décroît, plus on attache d'importance aux transferts concernent les revenus les plus faibles.

3.7 Indice Bonferroni

C.E. Bonferroni (1930) [16] a proposé une mesure de l'inégalité des revenus, basée sur des moyens partiels, ce qui est souhaitable lorsque la principale source d'inégalité de revenu est la présence d'unités dont le revenu est très inférieur à celui des autres.

Le but est de passer en revue les propriétés théoriques et statistiques de l'indice de Bonferroni et de les relier aux caractéristiques de la distribution des revenus. Les

principaux résultats obtenus peuvent être résumés comme suit :

a) L'indice se concentre sur les bas revenus.

b) L'indice satisfait au principe de transfert décroissant introduit par Mehran (1976) [130].

c) L'indice n'est pas additionnellement décomposable.

3.7.1 Iindice de Bonferroni pour les distributions continues.

Définition 3.4 *Soit Y une variable aléatoire continue non négative avec fonction de distribution cumulative F. La moyenne partielle (ou conditionnelle) de Y sur l'intervalle* $[0, y]$ *est donnée par*

$$m(y) = \frac{\int_0^y u dF(u)}{F(y)}$$

Pour un niveau donné y du revenu Y

$$r(y) = \frac{\mu - m(y)}{\mu}; 0 < \mu < \infty \tag{3.31}$$

est une fonction bornée, monotone décroissante et non négative dans $[0, \infty[$ *et mesure la différence relative entre le revenu moyen total* μ *et la moyenne des revenus inférieurs ou égale à y. La moyenne (3.31) de tous les revenus donne l'indice de Bonferroni*

$$B = \int_0^\infty r(y) dF(y) \tag{3.32}$$

Notez que comme suggéré Pizzetti (1955) l'indice de Gini **G** *peut être exprimé comme suit*

$$\mathbf{G} = \int_0^\infty r(y) \left[\frac{F(y)}{\int_0^\infty F(y) dF(y)} \right] dF(Y), \tag{3.33}$$

par conséquent **G** *est la moyenne pondérée des* $r(Y)$ *alors que B est leur moyenne simple. Puisque* $r'(Y)$ *et* $F'(Y)$ *ont un signe opposé, alors* $B \geq \mathbf{G}$ *(voir De Vergottini, 1940). La moyenne progressive* $m(Y)$ *est égale au rapport* $\mu F_1(Y)/F(Y)$*, où* $F_1(Y)$ *est la distribution incomplète de premier instant correspondant à F. Par conséquent, la formule (3.32) peut aussi être écrit comme*

$$\begin{aligned} B &= \int_0^\infty \left[\frac{F(y) - F_1(y)}{F(y)} \right] dF(y) = 1 - \int_0^\infty \left[\frac{F_1(y)}{F(y)} \right] dF(y) \\ &= 1 - \int_0^\infty F_1(y) d \ln [F(y)]. \end{aligned} \tag{3.34}$$

L'indice de Bonferroni peut également être considéré comme la statistique réca-

pitulative de la courbe de Bonferroni de la distribution des revenus. Une telle courbe notée $B(p) : [0,1] (\subset \mathbb{R}) \rightarrow [0,1]$, est définie (voir Zenga, 1984a) comme la relation entre la proportion cumulée $p = F(y)$ des unités de revenu (IRU) et le ratio de la part cumulée du revenu $q = F(y)$ et p . Soit :

$$F^{-1}(t) = \inf\{y : F(y) > t\} \tag{3.35}$$

est la fonction inverse de $F(y)$ et $F_1(y)$ respectivement. Alors :

$$B(p) = \int_0^p \left[\frac{F_1^{-1}(t)}{F^{-1}(t)}\right] dt. \tag{3.36}$$

La courbe de Bonferroni est représentée dans un carré unitaire. Il est facile de vérifier que $B(0) = 0$, $B(1) = 1$, et que $B(p)$ est une fonction non décroissante pour $p \in [0,1]$. Clairement l'égalité parfaite aboutirait à des points le long de la ligne $B(p) = 1$, et si une IRU avait tous les revenus, la courbe de Bonferroni coïnciderait avec les cathètes OW et WZ.

D'un point de vue géométrique, l'indice de Bonferroni est la zone située entre la courbe de Bonferroni et la ligne d'égalité parfaite

$$B = 1 - \int_0^1 B(p)\, dp. \tag{3.37}$$

3.7.2 Indice de Bonferroni pour les distributions discrètes.

La population est supposée être composée de n IRU qui sont étiquetées dans l'ordre non-croissant du revenu de sorte que l'indice i indique le rang de y_i parmi $y_1, y_2, ..., y_n$. Soit μ le revenu moyen arithmétique, P_i soit la part cumulée de la population et Q_i la part cumulée des revenus correspondant aux premières IRU. Ainsi

$$P_i = \frac{i}{n}; Q_i = \frac{1}{n\mu}\sum_{j=1}^{i} y_j \quad (i = 1, 2, ..., n) \tag{3.38}$$

La moyenne des revenus inférieurs ou égaux à y_j est

$$M_i = \frac{1}{i}\sum_{j=1}^{i} y_j \ ;$$

par conséquent pour une distribution discrète, nous avons

$$B_i = \frac{1}{n-1}\sum_{i=1}^{n-1}\left[\frac{\mu - M_i}{\mu}\right] = \frac{1}{n-1}\sum_{i=1}^{n-1}\left[\frac{P_i - Q_i}{P_i}\right], \tag{3.39}$$

ces expressions montrent que B est facilement estimable à partir de sources de données existantes. Alternativement B_n peut être écrit comme un rapport de combinaisons linéaires de statistiques d'ordre

$$B_n = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} y_i}, \tag{3.40}$$

avec

$$w_i = 1 - \sum_{j=1}^{n} \frac{1}{j};\ w_{i+1} = w_i + \frac{1}{i};\ \sum_{i=1}^{n} w_i = 0. \tag{3.41}$$

Cette spécification aide à caractériser le système de pondération des revenus dans la fonction de bien-être derrière l'indice de Bonferroni. Si un transfert d'une unité de revenu préservant le rang a lieu de la s-ème à la r-ème IRU (avec $s > r$) B_n changera d'un montant

$$\Delta B_n = -\frac{1}{(n-1)\,\mu} \sum_{j=r}^{s-1} \frac{1}{j}, \tag{3.42}$$

qui est proportionnelle au nombre d'IRU dont le revenu tombe en $[y_r, y_{s-1}]$. En outre la formule (3.42) montre que pour une différence fixe $(s-r)$ entre les deux rangs, le plus bas est r le plus haut est ΔB_n pour que le Bonferroni satisfasse le principe de transfert décroissant : un petit transfert positif d'une unité plus riche vers une unité plus pauvre diminue l'inégalité et la diminution est d'autant plus importante que l'unité est pauvre. Il faut noter, cependant, que l'effet du transfert dépend uniquement des rangs r et s et non de la taille des niveaux de revenu (voir aussi Salvaterra, 1986).

De Vergottini (1940) a interprété (3.40) comme la fraction du revenu total qui devrait être transférée pour atteindre un état d'égalité parfaite au moyen d'un nivellement graduel des revenus à partir de l'IRU la plus pauvre. En fait, pour égaler y_1 et y_2, la quantité M_2 doit être soustraite de y_2 et $(M_2 - M_1)$ ajoutée à y_1 . De même, pour éliminer les différences entre $y_1 = M_2, y_2 = M_2$ et y_3, la quantité M_3 doit être soustraite de y_3 et $(M_3 - M_2)$ ajoutée à chaque M_2. En général, $i\,(M_{i+1} - M_i)$ est transféré à la i-ème redistribution de sorte que le revenu transféré pendant tout le processus de péréquation est

$$\sum_{i=1}^{n-1} i\,(M_{i+1} - M_i) = \sum_{i=1}^{n-1} (\mu - M_i) \tag{3.43}$$

en divisant (3.43) par le revenu qui doit être transféré dans le cas d'une inégalité complète, c'est-à-dire $(n-1)\,\mu$, on obtient B_n.

Il est facile de vérifier que l'indice de Bonferroni vérifie les propriétés suivantes :

1. $0 \leq B \leq 1$.
2. $B = 0$ si et seulement si tous les revenus sont égaux.

3. $B = 1$ si et seulement si un seul revenu est positif.
4. B est indépendant de l'échelle.
5. B est symétrique (ne dépend pas de l'affectation des étiquettes à l'IRU)
6. Les additions égales (soustractions) diminuent (augmentent) B
7. B ne satisfait pas la propriété d'invariance de la réplication de la population (Zenga, 1986) a prouvé qu'il existe une incompatibilité logique entre la troisième propriété et la propriété d'invariance à la réplication de la population)
8. B appartient à la classe des mesures linéaires de l'inégalité des revenus

$$J = \frac{\int_0^1 \left[F^{-1}(p) - \mu\right] W(p)\, dp}{\mu}, \int_0^1 W(p)\, dp = 0$$

défini par Mehran (1976) [130] avec $W(p) = 1 + Log(p)$. Mehran (1976) [130] a examiné la fonction de score linéaire $2(p-1)$ correspondant à l'indice de Gini et la fonction de score quadratique $W(p) = 1 - 3(1-p)$ qui ne correspond pas à une mesure d'inégalité bien connue. "Peu d'autres choix de $W(p)$ semblent avoir été explorés" (Arnold, 1983 [4]). Les deux auteurs ont omis de mentionner l'indice de Bonferroni.

3.8 Indice d'inégalité de Zenga

Zenga (2007) [171] propose une mesure d'inégalité basée sur la courbe d'inégalité $I(p)$, définie en termes de moyennes arithmétiques inférieure et supérieure d'une distribution. Cette idée de mesure de l'inégalité des revenus consiste à comparer les moyennes arithmétiques des revenus de deux groupes, appelés groupes inférieurs et supérieurs. La division des données ordonnées en deux groupes est faite en choisissant un point de division. un extrême, le groupe inférieur consiste seulement en l'observation la plus basse. l'autre extrême, le groupe supérieur ne comprend que les revenus les plus élevés. Soit

3.8.1 Indice de Zenga cas discrèt

$$\left\{\left(x_j, n_j\right) : j = 1, ..., s; \quad 0 \leq x_1 \leq x_2 \leq ... \leq x_s; \quad \sum_{j=1}^{s} n_j = N\right\} \tag{3.44}$$

dénote la distribution de fréquence d'une variable aléatoire non négative X. En suite, divisons cette distribution en deux parties, respectivement le groupe inférieur et le groupe supérieur :

$$\Big\{(x_1,n_1),(x_2,n_2),...,\big(x_j,n_j\big)\Big\},\Big\{\big(x_j,n_j\big),\big(x_{j+1},n_{j+1}\big),..,(x_s,n_s)\Big\}. \tag{3.45}$$

Pour chaque point de division, il est possible de définir la moyenne inférieure $\overline{M}\left(p_j\right)$et la moyenne supérieure $M^+\left(p_j\right)$ de la distribution divisée comme suit :

$$\begin{aligned} \overline{M}\left(p_j\right) &= \frac{1}{N_j}\sum_{i=1}^{j} x_i n_i,\ j=1,...,s \\ M^+\left(p_j\right) &= \frac{1}{N-N_{j-1}}\sum_{i=j}^{s} x_i n_i,\ j=1,...,s \end{aligned} \tag{3.46}$$

où $N_j=\sum_{i=1}^{j} n_i$ et $p_j=\frac{N_j}{N}$.

Comparons $\overline{M}\left(p_j\right)$ et $M^+\left(p_j\right)$ en utilisant l'indice, nous obtenons une mesure ponctuelle de l'uniformité de la distribution. $\left(U\left(p_j\right)\right)\times 100$ donne la moyenne inférieure en pourcentage de la moyenne supérieure. L'indice d'inégalité de point est défini en terme de $U\left(p_j\right)$ comme :

$$I\left(p_j\right)=1-U\left(p_j\right) \tag{3.47}$$

La mesure d'inégalité synthétique proposée par Zenga est la moyenne arithmétique pondérée suivante du point mesures $I\left(p_j\right)$:

$$Z=\sum_{j=1}^{s} I\left(p_j\right)\frac{n_j}{N} \tag{3.48}$$

$U\left(p_j\right)$ et $I\left(p_j\right)$ prennent des valeurs comprises entre 0 et 1 inclusivement. En particulier :

$$U\left(p_1\right)=\frac{x_1}{M},\quad U\left(p_s\right)=\frac{M}{x_s},\quad I\left(p_1\right)=1-\frac{x_1}{M},\quad I\left(p_s\right)=1-\frac{M}{x_s} \tag{3.49}$$

où M est la moyenne de toutes les observations.
L'indice de Zenga prend la valeur 0 dans le cas d'aucune inégalité.
La forme de la courbe $I\left(p_j\right)$ en fonction de p_j n'est pas contrainte par des points fixes $(0,0)$ et $(1,1)$, comme dans le cas de la courbe de Lorenz .

3.8.2 Indice de Zenga cas continu

Dans le cas continu, l'indice de Zenga est donné par

$$Z = 1 - \int_0^1 \frac{L(\alpha)}{\alpha} . \frac{1-\alpha}{1-L(\alpha)} d\alpha. \tag{3.50}$$

— Nouvel mesure d'inégalité proposée par Zenga (2007).

— Comme l'ndice de Gini, l'ndice de Zenga prend une valeur entre 0 et 1.

— $L(\alpha)$ est la courbe de Lorenz.

Il a été prouvé que l'indice de Zenga est caractérisé par toutes les propriétés principales que toute mesure d'inégalité devrait satisfaire.

3.9 Fonction de bien-être

On considère que la société est formée par une collection de n individus et l'on va s'intéresser à une mesure de bien-être pour l'ensemble de ces n éléments pris comme une entité. La mesure se fera à partir d'une quantité uni-dimensionnelle que l'on prendra égale soit au revenu, soit à la dépense de consommation que l'on va noter x_i pour l'individu i. On a donc la première donnée de

$$X = (x_1, x_2, ..., x_n) \tag{3.51}$$

qui représente la distribution des revenus (ou de toute autre caractéristique) au niveau de la population. On définit en suite la fonction de bien-être comme une fonction à n arguments :

$$W(x) = V(x_1, ..., x_n). \tag{3.52}$$

Cette fonction a un aspect très normatif et sa construction répond à une série d'axiomes qui précisent les comparaisons que l'on s'autorise à faire entre les individus.

1. **Axiome de Pareto** : la fonction est croissante en chacun de ses termes. On peut affaiblir cette axiome en demandant à la fonction d'être simplement non-décroissante en ses termes. Alors, on peut construire une fonction de bien-être qui restera constante si le revenu des plus riches augmente et ne croîtra que si le revenu des plus pauvres augmente.
2. **Axiome de symétrie ou anonymat :** On doit pouvoir intervertir les individus sans que la valeur de la fonction change. Mais, il existe des problèmes soulevés par la composition des ménages. Les données concernent les ménages, alors que la notion de bien-être s'intéresse aux individus. Il y aura donc une incidence non triviale de la composition des ménages que l'on essayera de gommer au moyen des échelles d'équivalence.
3. **Principe du transfert :** La quasi concavité de la fonction de bien-être implique que si l'on transfère d'un riche vers un pauvre, le bien-être augmente, à condi-

tion que le transfert ne change pas l'ordre du classement des individus. Il s'agit du principe dit de Pigou-Dalton.

4. **Autres axiomes :** la littérature économique sur la construction des fonctions de bien-être et des indices d'inégalité est importante. Certains axiomes se recoupent. On peut chercher le nombre minimal d'axiomes qui conduise à la construction de la fonction de bien-être. On consultera à ce propos l'ouvrage de Sen (1997) [153].

La conséquence de ces axiomes, est qu'une fonction de bien-être exprime l'aversion d'une société pour l'inégalité et que cette fonction sera maximale quand tous les ménages auront le même revenu.

3.10 De l'inégalité à la pauvreté

La consultation de la forme de la fonction de bien-être permet de voir que la croissance économique, c'est à dire l'augmentation conjointe de μ et de W peut s'accompagner d'une augmentation des inégalités : certains vont s'enrichir plus vite que les autres. C'est ce que l'on a constaté par exemple au Royaume Uni pendant la période Thatcher. Atkinson (2003) [**?**] montre comment dans les années 1980, le revenu réel des plus pauvres est resté constant alors que la croissance des revenus a concerné les groupes moyens et surtout les groupes les plus riches. Malgré cela la mesure du bien-être a augmenté.

La pauvreté est ressentie comme un échec et cela justifie que l'on s'y intéresse plus particulièrement. La fonction de bien-être transforme une distribution complète en un nombre permettant d'analyser les effets de mesures de politique économique sur l'ensemble de la distribution des revenus. Si l'on veut concentrer son attention sur les plus pauvres, on va s'intéresser plus particulièrement à une partie de la distribution des revenus, celle qui concerne les plus pauvres, ne serait-ce que pour les compter. On va donc passer de l'analyse des inégalités à l'analyse de la pauvreté en concentrant son attention sur la queue gauche de la distribution des revenus.

Pour concenter son attention sur les plus pauvres, il faut définir ce que l'on appelle une ligne de pauvreté, c'est à dire un seuil en delà duquel une personne (ou un ménage) sera considéré comme pauvre et au delà duquel il basculera dans la catégorie des non-pauvres. On mesure combien ce qu'un tel seuil a d'arbitraire. On peut le définir de deux façons

1. un seuil de pauvreté absolu se définit par rapport à un niveau minimum de subsistance. Le gouvernement Indien par exemple a défini un nombre minimum de calories nécessaires en ville et qui est différent de celui nécessaire à la campagne. En passant par un indice de prix, il arrive à un seuil monétaire de pauvreté en

ville et à la campagne. Sur la même base alimentaire, le gouvernement américain a défini un seuil absolu de pauvreté, mais en divisant celui-ci par la part de la nourriture dans le budget d'un ménage moyen. Le RMI(revenu minimum d'insertion) peut également se situer dans ce cadre.

2. Dans les pays développés et plus particulièrement au sein de l'Union Européenne, on préfère définir un seuil relatif de pauvreté. L'Union Européenne a lancé un programme de mesure de la pauvreté où le seuil de pauvreté y est défini par rapport à la moyenne ou la médiane de la distribution des revenus. Sera considéré comme pauvre tout individu touchant un revenu inférieur à 50% ou 60% de la moyenne des revenus de son pays. Il s'agit alors de pauvreté relative, ce qui nous rapproche de la notion de pauvreté ressentie.

3.11 Les indices de pauvreté

Il existe toute une série d'indices de pauvreté, mais d'une certaine façon les indices les plus simples à comprendre et à manipuler sont les indices linéaires de Foster, Greer, and Thorbecke (1984). Ces indices sont basés sur des moyennes partielles construites à partir de la distribution des revenus. Si $F(.)$ est la distribution des revenus et z le seuil de pauvreté, alors pour un α donné cet indice s'écrit

$$P_\alpha = \int_0^z \left(\frac{z-x}{z}\right)^\alpha dF(x). \tag{3.53}$$

En faisant varier le paramètre α entre 0 et 2, on retrouve un certain nombre des mesures classiques de pauvreté.

— **Pour** $\alpha = 0$, on tombe sur la mesure traditionnelle dite de headcount qui est une mesure de dénombrement :

$$P_0 = \int_0^z f(x)\,dx = F(z). \tag{3.54}$$

Il s'agit d'une première mesure, intéressante en soi, elle permet de connaître le nombre de pauvres en multipliant simplement P_0 par la taille de la population. Toutefois cette mesure est insuffisante car elle ne distingue pas entre les pauvres et ne tient pas compte de leur niveau de pauvreté, c'est à dire qu'elle ne distingue pas entre les individus qui sont près de la ligne de pauvreté et ceux qui en sont loin.

— **Pour** $\alpha = 1$, on tombe sur une mesure faisant intervenir le déficit de pauvreté ou poverty gap $z - x_i$ qui affecte chaque individu en dessous du seuil de pauvreté :

$$P_1 = \int_0^z (1 - x/z)\, f(x)\,dx$$

Cet indice respecte le principe de transfert à l'inverse de la mesure de comptage P_0 qui ne le respecte pas. Cet indice est continu alors que le head count ne l'est pas. Mais il est insensible à certains types de transferts entre les pauvres.

— **Pour** $\alpha = 2$, on arrive à une mesure qui est sensible à la distribution parmi les pauvres :

$$P_2 = \int_0^z (1 - x/z)^2 f(x)\, dx, \tag{3.55}$$

mais qui n'est pas très souvent employée. Atkinson (1987) examine simplement les propriétés d'une généralisation des deux premiers indices et leur relation avec la dominance stochastique restreinte, notion que l'on explicitera plus bas.

L'indice de Foster, Greer, and Thorbecke (1984) répond à la propriété de décomposabilité car il a une structure linéaire. Considérons par exemple une partition de la population entre urbaine et rurale. Si X représente l'ensemble des revenus, la partition de X se définira comme $X = X_U + X_R$. Appelons p la proportion de X_U dans X. Alors l'indice total de pauvreté se décomposera en

$$\begin{aligned} P_\alpha &= p \int_0^z \left(\frac{z-x}{z}\right)^\alpha dF(x_U) + (1-p) \int_0^z \left(\frac{z-x}{z}\right)^\alpha dF(x_R) \\ &= p P_\alpha^U + (1-p) P_\alpha^R. \end{aligned} \tag{3.56}$$

Une autre classe d'indices de pauvreté a été proposée à la suite de Sen (1976) pour tenir compte de l'inégalité entre les pauvres. L'indice original de Sen, P_s, combine une mesure de headcount P_0, une mesure du poverty gap,

$$z P_0 I_p = \int_0^z (z - x) f(x)\, dx$$

et un indice de Gini G_p calculé sur le segment $x < z$. Cet indice se note

$$P_s = P_0 \Big(I_P + (1 - I_P) G_p \Big) = P_0 \left(1 - (1 - I_P) \frac{\mu_p}{z} \right) \tag{3.57}$$

où μ_p est la moyenne des revenus parmi les pauvres :

$$\mu_p = \int_0^z x f(x)\, dx / F(z) \quad et \quad I_p = 1 - \frac{\mu_p}{z}. \tag{3.58}$$

Quand il n'y a pas d'inégalité entre les pauvres $\Big(G_p = 0\Big)$, alors $P_s = P_1$. Quand l'inégalité devient extrême $\Big(G_p = 1\Big)$, on retombe sur la mesure de headcount, ce que traduit bien la factorisation

$$P_s = P_0 G_p + P_1 (1 - G_p). \tag{3.59}$$

Mais tout comme l'indice de Gini, cet indice n'est pas décomposable. Il viole également l'axiome de transfert et de plus n'est pas continu. Shorrocks (1995) a proposé une modification de cet indice appelé aussi l'indice de Sen-Schorrocks-Thon qui résoud une partie de ces difficultés. Cet indice s'écrit par analogie avec l'indice de Sen comme

$$P_{SST} = (2 - P_0) P_0 I_p + P_0^2 \left(1 - I_p\right) G_p. \tag{3.60}$$

3.12 Pauvreté et inégalité

La formulation initiale de la fonction de bien-être (3.52) implique qu'une augmentation du bien-être peut tout à fait s'accompagner d'un accroissement des inégalités. Comment inclure dans cette décomposition une attention plus particulière à la pauvreté ? En d'autres termes, quelle forme doit-on considérer pour $W(x)$ si l'on veut maximiser le bien-être tout en insistant sur la pauvreté. Atkinson (1987) traite de cette question dans la section 3 de son papier en distinguant quatre options possibles.

— La première option consiste à ne pas se soucier particulièrement de la pauvreté. On va simplement maximiser

$$W(x) = \mu(1 - I), \tag{3.61}$$

où I est un indice d'inégalité et μI mesure le coût de l'inégalité. Si l'on a choisi de manière adéquate la fonction de bien-être, on peut décomposer cet indice en distinguant le groupe des pauvres du reste de la population. On pourra donc mesurer l'évolution de la pauvreté sans avoir la réduction de la pauvreté comme objectif principal.

— Dans une deuxième option, on va chercher à introduire un coût prioritaire pour la pauvreté $C_p = \mu P$ et laissant un rôle secondaire au coût de l'inégalité. Ceci peut se faire en adoptant une fonction de bien-être du type

$$W(x) = \mu - \mu P - \mu I.$$

Atkinson (1987) indique que dans ce cas, il est logique d'utiliser une mesure de comptage pour P et une mesure satisfaisant le principe de transfert pour I.

— La troisième option consiste à se focaliser uniquement sur la pauvreté. La fonction de bien-être à maximiser sera alors de la forme

$$W(x) = \mu - \mu P. \tag{3.62}$$

— Enfin, la dernière option consiste à utiliser un arbitrage entre inégalité et pau-

vreté. On aura toujours la fonction de bien-être donnée en

$$W(x) = \mu - \mu I - \mu P. \tag{3.63}$$

Mais cette fois-ci des considérations de justice vont conduire à utiliser pour I un indice de Gini calculé sur toute la population et pour P un indice de pauvreté de Sen (1976) modifié.

Ces considérations montrent que la construction de la fonction de bien-être peut être relativement complexe quant à ses propriétés et à la façon dont sont agrégés les individus. La forme simple était peut-être un peu trop simple.

3.13 La décomposition des indices

Certains indices d'inégalité et de pauvreté peuvent facilement se décomposer comme on la vu avec les indices de Foster, Greer, et Thorbecke (1984). La décomposition se fait alors par groupes de la population. Mais dans cette procédure aucune explication n'est donnée en fonction des caractéristiques des sous groupes. Oaxaca (1973) a le premier tenté une explication des inégalités en utilisant une technique de régression.

Oaxaca (1973) s'intéresse aux inégalités de salaire entre hommes et femmes. Pour chaque groupe, on estime une équation de salaire

$$\log(W_i) = X_i\beta + u_i, \quad i = h, f. \tag{3.64}$$

On regarde ensuite les différences salariales moyennes entre hommes et femmes. Une partie de cette différence s'explique par des différences de caractéristiques objectives mesurées par X_h, X_f, l'autre partie s'explique par des différences de rendements de ces mêmes caractéristiques, c'est à dire par la discrimination que le marché opère entre hommes et femmes. Comme dans (3.64) une régression $\log(Wi) = \overline{X}_i\widehat{\beta_i}$, on aura la décomposition suivante appelée déomposition de Oaxaca :

$$\log(W_h) - \log\left(W_f\right) = \left(X_h - X_f\right)\widehat{\beta_h} + \overline{X}_f\left(\widehat{\beta_h} - \overline{\beta}_f\right). \tag{3.65}$$

Ce type de décomposition a donné lieu à des développements importants dans la littérature. Par exemple Juhn, Murphy, and Pierce (1993) généralisent le résultat précédent aux différents quantiles de la distribution des résidus. Radchenko et Yun (2003) donnent une approche Bayésienne qui permet d'implémenter facilement des tests de significativité.

La décomposition de Oaxaca (1973) est basée sur une hypothèse de régression linéaire. Yun (2004) en donne une généralisation non-linéaire qui permet de proposer une décomposition des mesures de head count. Celles-ci étant assimilables à des pro-

portions, on peut les relier à des variables explicatives au moyen d'un modèle probit. Une des applications dans Yun (2004) concerne justement les modèles probit. Une équation de régression par groupe permet d'expliquer le ratio entre la dépense y et le seuil de pauvreté z :

$$\log\left(\frac{y}{z}\right) = X\beta + e, \tag{3.66}$$

où X est une matrice de caractéristiques personnelles à k composantes pour n observations. Sous une hypothèse de normalité pour e, la probabilité d'être pauvre pour le groupe des n individus est égale au vecteur $\Phi(-X\beta/\sigma)$ où σ^2 est la variance des résidus. Quand n tend vers l'infini, la moyenne des $\Phi(-X\beta/\sigma)$tend vers le head count ratio, c'est à dire P_0 dans nos notations. Alors, on aura la décomposition suivante de la différence entre deux indices de pauvreté correspondant à deux groupes distincts A et B :

$$P_A^0 - P_B^0 = \left[\overline{\Phi\left(-X_A\beta_A/\sigma_A\right)} - \overline{\Phi\left(-X_B\beta_A/\sigma_A\right)}\right] + \left[\overline{\Phi\left(-X_B\beta_A/\sigma_A\right)} - \overline{\Phi\left(X_B\beta_B/\sigma_B\right)}\right] \tag{3.67}$$

ce qui correspond à la différence entre les caractéristiques et la différence entre les coefficients.

L'application de la procédure de Yun (2004) permet de pondérer cette décomposition en fonction du poids de chacune des k caractéristiques individuelles

$$\begin{aligned} P_A^0 - P_B^0 \quad = \quad & \sum_{i=1}^{k} W_{\Delta x}^i \left[\overline{\Phi\left(-X_A\beta_A/\sigma_A\right)} - \overline{\Phi\left(-X_B\beta_A/\sigma_A\right)}\right] \\ & + \sum_{i=1}^{k} W_{\Delta\beta}^i \left[\overline{\Phi\left(-X_B\beta_A/\sigma_A\right)} - \overline{\Phi\left(X_B\beta_B/\sigma_B\right)}\right] \end{aligned} \tag{3.68}$$

où les poids $\sum_{i=1}^{k} W_{\Delta x}^i$ et $\sum_{i=1}^{k} W_{\Delta\beta}^i$ sont donnés dans Bhaumik, Gang, et Yun (2006) sur un argument de linéarisation de Yun (2004).

3.14 Conclusion

Dans ce chapitre, nous avons présenté des indicateurs statistiques dans l'économie qui ont un impact significatif dans le monde d'aujourd'hui. Mesurer les changements dans la répartition des revenus et l'inference économique est un objectif majeur de la recherche sur l'inégalité. Le développement de l'inégalité économique, une préoccupation majeure pour une croissance énorme avec des niveaux élevés de revenu, la raison est de considérer les mesures traditionnelles de l'inégalité et de fournir de nouvelles façons de collecter la distribution des fluctuations du revenu complet. L'objectif principal de notre approche utiliser est d'obtenir des résultats pour certaine inférences statistiques des quelques indices en économie (indice de Gini et indice de Zenga).

Nous avons identifié quelques variables aléatoires et des indicateurs généraux, et nous avons montré que les indicateurs économiques sont une classe de mesures connexes, et nous avons mené une étude les estimations pour indice Gini et indice de Zenga et proposer un nouvelle estimateur de l'indice Gini, les procédures détaillées et les exemples numériques montrent également comment utiliser l'indice de Gini par les lois de Pareto et Fréchet. Cette proposition est peut-être concluante ils sont truqués lors de distributions supplémentaires pour des considérations théoriques dans le compte.

Chapitre 4

Inférence statistique sur l'indice de Zenga

4.1 Introduction

Ce chapitre est consacré pour présenter des estimateurs statistiques de l'indice de Zenga,

- Estimation paramétrique basé sur un échantillon i.i.d. avec des distributions de variance fini,
- Un estimation semi paramétrique basé sur un échantillon i.i.d. pur des distribution de revenus de type de Paréto ou bien des distribution à queue lourdes, cet estimateur basé sur l'estimateur de Hill et l'estimateur du quantile de Weissman.
- Un estimateur à Noyau qui généralise l'estimateur semi paramétrique basé sur la généralisation de l'estimateur de Hill à un estimateur de type de Noyau.
- Un estimateur avec un biais réduit par l'utilisation un modèle de régression introduit par Beirlant 2004 pour réduire le bias de l'estimateur de Hill

4.1.1 Rappel sur l'indice de Zenga

Soit $F(x) = P[X \leq x]$ la fonction de distribution cumulative (fdc) de la variable aléatoire X, et

$$Q(t) = \inf\{x : F(x) \geq t\}, \quad t \in [0\ ;\ 1]$$

dénote la fonction quantile correspondante.

L'indice Z_F de l'inégalité de Zenga est défini par la formule

$$Z_F = \int_0^1 z_F(p)dp \tag{4.1}$$

où $z_F(p)$ est la courbe de Zenga, donnée par :

$$z_F(t) = 1 - \frac{L_F(p)}{p} \cdot \frac{1-p}{1-L_F(p)} \tag{4.2}$$

$L(p)$ est la fonction de Lorenz donnée par

$$L(p) = \frac{1}{\mu} \int_0^P Q(s)ds. \tag{4.3}$$

4.1.2 Estimateur traditionnel de l'indice Zenga

Soit $X_1, X_2, ..., X_n$ observations indépendants identiquement distribuées suivant de la fonction de distribution F. On considère l'estimation empirique de F par

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \leq x\}} \tag{4.4}$$

avec 1_A désigne la fonction d'indicateur sur l'ensembe A. Par le remplacement de l'estimateur (4.4) dans la formule (4.1), nous arrivons à un estimateur empirique de l'indice de Zenga traditionnel (par exemple, Greselin et Pasquazzi 2009 [94]; Greselin et al. 2013 [95])

$$\hat{Z}_n = 1 - \frac{1}{n} \sum_{i=1}^{n-1} \frac{i^{-1} \sum_{k=1}^{i} X_{k:n}}{(n-i)^{-1} \sum_{k=i+1}^{n} X_{k:n}} \tag{4.5}$$

où $X_{1:n} < X_{2:n} < ... < X_{n:n}$ sont les statistiques d'ordre basées sur la série $X_1, X_2, ..., X_n$.

Théorème 4.1 *Si X est une v.a. de CDF F avec le moment $E[X^{2+\varepsilon}]$ est fini pour tout $\varepsilon > 0$. Alors on a la représentation asymptotique*

$$\sqrt{n}(\hat{Z}_n - Z_F) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h(X_i) + o_p(1) \tag{4.6}$$

ici $o_p(1)$ désigne une variable aléatoire qui converge vers 0 en probabilité lorsque $n \to \infty$, et

$$h(X_i) = \int_0^{\infty} \left(1_{\{X_i \leq x\}} - F(x)\right) \omega_F(F(x))dx$$

avec la fonction de poids

$$\omega_F(t) = \frac{1}{\mu_F} \int_0^t \left(\frac{1}{p} - 1\right) \frac{L_F(p)}{(1-L_F(p))^2} dp + \frac{1}{\mu_F} \int_t^0 \left(\frac{1}{p} - 1\right) \frac{L_F(p)}{1-L_F(p)}. \tag{4.7}$$

Proof.[Voir Greselin et al. 2013] ■

4.2 Estimation semi paramétrique

Puisque nous sommes concernés par les distributions de revenus de type de Paréto ou une populations à queue lourde, nous travaillons inévitablement avec les fdc qui varient régulièrement à l'infini. Par conséquent, nous supposons que F satisfait

$$\lim_{t\longrightarrow\infty}\frac{1-F(tx)}{1-F(t)}=x^{-1/\gamma} \tag{4.8}$$

pour certains $\gamma>0$, ce que l'on appelle l'indice de la queue de la distribution. L'estimateur semi paramétrique pour les distributions à queue lourde pour l'indice Z_F est basé sur l'estimateur de Weissman (1978) [167] et l'estimateur de Hill's (1975) [107] de l'indice de queue γ. A savoir, avec une suite d'entiers $k:=k_n\to\infty$ tels que $k/n\to 0$ quand $n\to\infty$, le Hill's l'estimateur :

$$\hat{\gamma}_n=\frac{1}{k}\sum_{i=1}^{k}\log\left(\frac{X_{n-i+1:n}}{X_{n-k:n}}\right). \tag{4.9}$$

De Haan (1994) a étudié les propriétés asymptotiques de l'estimateur. Avec la fonction quantile empirique définie par

$$Q_n(t)=\inf\{x:F_n(x)\geq t\},0<t<1.$$

Z est estimé par :

$$\bar{Z}_{n,k}=1-\int_0^1\left(\frac{1-t}{t}\frac{\int_0^t Q_n(s)ds}{\int_t^{1-k/n}Q_n(s)ds+\frac{kX_{n-k:n}}{n(1-\hat{\gamma}_n)}}\right)dt. \tag{4.10}$$

4.2.1 Propriétés asymptôtiques

Établir les distributions asymptotiques des estimateurs dans les situations à queue lourde nécessite une hypothèse plus forte que celle de l'équation (4.8), ce qui est suffisant pour la cohérence de résultats. Par conséquent, à partir de maintenant, nous supposons que F satisfait le condition généralisé de variation régulière de second ordre avec le paramètre de second ordre $\rho\leq 0$, qui signifie qu'il existe une fonction α sur $[0,\infty[$ telle que

— $\alpha(t)\to 0$ lorsque $t\to\infty$;
— $\alpha(t)$ ne change pas de signe pour tout t suffisamment grand ; et
— l'équation

$$\lim_{t\to+\infty}\frac{1}{\alpha(t)}\left(\frac{1-F(tx)}{1-F(t)}\right)-x^{-1/\gamma}=x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\rho/\gamma} \tag{4.11}$$

est vérifiée pour tout $x > 0$, avec le côté droit interprété comme $x^{-1/\gamma} \log x$ quand $\rho = 0$.

Théorème 4.2 *Supposons que le fdc F vérifie la condition (4.11) avec un $\gamma \in (1/2, 1)$ et $\rho \leq 0$, et soit $k = k_n \to \infty$ quand $n \to \infty$ est tel que $k/n \to 0$ et $\sqrt{k}\alpha(Q(1-k/n)) \to 0$. Puis sur un espace de probabilité approprié, et avec des ponts Brownien construits de manière appropriée $\mathcal{B}_n$, nous avons*

$$\begin{aligned} \frac{\sqrt{n}(\bar{Z}_{n,k} - Z)}{\sqrt{k/n}Q(1-k/n)} &= -\int_0^{1-k/n} \frac{\mathcal{B}_n(s)v(s)}{\sqrt{k/n}Q(1-k/n)} dQ(s) \\ &\quad + \frac{\gamma^2 v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\mathcal{B}_n\left(1-\frac{k}{n}\right) \\ &\quad - \frac{\gamma v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\int_{1-k/n}^1 \frac{\mathcal{B}_n(s)}{1-s} ds + o_p(1) \end{aligned} \tag{4.12}$$

lorsque $n \to \infty$, où, pour $0 \leq s \leq 1$,

$$v(s) = \int_0^s \frac{1-t}{t} \frac{\int_0^t Q(s)ds}{\int_t^1 Q(s)ds} dt. \tag{4.13}$$

Corollaire 4.1 *Sous les hypothèses du théorème 4.2 , nous avons*

$$\frac{\sqrt{n}(\bar{Z}_{n,k} - Z)}{\sqrt{k/n}Q(1-k/n)} \longrightarrow \mathcal{N}(0, \sigma_Z^2). \tag{4.14}$$

avec

$$\sigma_Z^2 = \frac{\gamma^4}{(1-\gamma)^4(2\gamma-1)} v^2(1). \tag{4.15}$$

Proof. Notons $U_i = F(X_i)$ pour$i = 1, 2, ..., n$. Alors $U_1, U_2, ..., U_n$ est une séquence de i.i.d. variables aléatoires suivant la distribution uniforme sur $[0, 1]$. Ce qui suit théorème montre que est les processus empiriques et quantiles basés sur la séquence $U_1, U_2, ..., U_n$ peut être approchés par une série de ponts browniens; (voir Csörgő et Horváth 1993 [**?**])

$$\begin{aligned} \bar{G}_{n,k} - G &= \left(1 - \frac{2}{\hat{\mu}_n}\int_0^1\int_0^t Q_n(s)dsdt\right) - \left(1 - \frac{2}{\mu}\int_0^1\int_0^t Q(s)dsdt\right) \\ &= -\frac{2}{\hat{\mu}_n}\int_0^1\int_0^t Q_n(s)dsdt + \frac{2}{\mu}\int_0^1\int_0^t Q(s)dsdt \\ &= -\frac{2}{\hat{\mu}_n}\int_0^1\int_0^t Q_n(s)dsdt + \frac{2}{\hat{\mu}_n}\int_0^1\int_0^t Q(s)dsdt \\ &\quad + \frac{2}{\mu}\int_0^1\int_0^t Q(s)dsdt - \frac{2}{\hat{\mu}_n}\int_0^1\int_0^t Q(s)dsdt \end{aligned} \tag{4.16}$$

$$\bar{G}_{n,k} - G = A_n + B_n \tag{4.17}$$

où

$$A_n = -\frac{2}{\hat{\mu}_n}\left(\int_0^1 \int_0^t [Q_n(s) - Q(s)]\, ds dt\right) \tag{4.18}$$

et

$$\begin{aligned} B_n &= \left(\frac{2}{\mu} - \frac{2}{\hat{\mu}_n}\right)\int_0^1 \int_0^t Q(s) ds dt \\ &= 2\left(\frac{(\hat{\mu}_n - \mu)}{\mu\hat{\mu}_n}\right)\int_0^1 \int_0^t Q(s) ds dt \end{aligned} \tag{4.19}$$

A_n qui fait partie intégrante du processus quantile général $Q_n - Q$. Pour le réduire à une intégrale de l'empirique générale processus $F_n - F$, nous utilisons le procédé Vervaat (général)

$$V_n(t) = \int_0^t (Q_n(s) - Q(s)) ds + \int_{-\infty}^{Q(t)} (F_n(x) - F(x)) dx \tag{4.20}$$

Le processus $V_n(t)$ vérifie les conditions aux limites $V_n(0) = 0$ et $V_n(1) = 0$, est non négatif pour tout $t \in [0,1]$, et tel que

$$\sqrt{n} V_n(t) \leq |e_n(t)||Q_n(t) - Q(t)|. \tag{4.21}$$

Par conséquent, en rappelant que $e_n(t) = \sqrt{n}(F_n(Q(t)) - t)$ nous concluons à partir de l'équation(4.20) que la différence entre les quantités

$$\sqrt{n}\int_0^t (Q_n(s) - Q(s)) ds \tag{4.22}$$

et

$$-\sqrt{n}\int_{-\infty}^{Q(t)} (F_n(x) - F(x)) dx \tag{4.23}$$

tend vers zéro quand $n \to \infty$ quand $Q_n(t)$ converge vers $Q(t)$, ce qui est vrai F est continu et strictement croissant. C'est l'idée même d'employer le Vervaat processus dans la présente preuve, car il nous permet de remplacer la quantité (4.22) par (4.23) ce qui est beaucoup plus facile à aborder. Nous avons l'équation suivante

$$A_n(t) = \int_0^{1-k/n} (Q_n(s) - Q(s)) ds - \int_0^t (Q_n(s) - Q(s)) ds \tag{4.24}$$

$$= -\int_{Q(t)}^{Q(1-k/n)} (F_n(x) - F(x)) dx + V_n(1 - k/n) - V_n(t). \tag{4.25}$$

que nous appliquons sur Eq.(4.17). En changeant la variable de l'intégration, nous

obtenons

$$A_n(t) = -\int_t^{1-k/n} \frac{e_n(s)}{\sqrt{n}} dQ(s) + V_n(1-k/n) - V_n(t)$$

et

$$A_n(0) - A_n(t) = -\int_0^t \frac{e_n(s)}{\sqrt{n}} dQ(s) + V_n(t). \tag{4.26}$$

Alors

$$\frac{\sqrt{n}A_n(t)}{\sqrt{k/n}Q(1-k/n)} = -\frac{\int_{Q(t)}^{Q(1-k/n)} e_n(F(x))dx}{\sqrt{k/n}Q(1-k/n)} + O_p\frac{|\, e_n(1-k/n)\,\|}{\sqrt{k/n}Q(1-k/n)}.$$

D'après le résultat de Peng L. 2001 [**?**], Necir, Rassoul and Zitikis (2010) [**?**], il existe une suite des ponts Brownian $\{B_n(s), 0 \le s \le 1\}_{n\ge 1}$ telle que, pour tout n assez grand, on a :

$$\begin{aligned}\frac{\sqrt{n}(\hat{\mu}_n-\mu)}{\sqrt{k/n}Q(1-k/n)} \stackrel{d}{=} & -\int_0^{1-k/n} \frac{B_n(s)}{\sqrt{k/n}Q(1-k/n)} dQ(s) \\ & + \left\{\sqrt{n/k}\mathcal{B}_n(s)B_n(1-k/n)\right\} \\ & -\sqrt{n/k}\int_{1-k/n}^1 \frac{B_n(s)}{1-s} ds + o_P(1),\end{aligned}$$

Alors

$$\frac{\sqrt{n}(\bar{G}_{n,k} - G)}{\sqrt{k/n}Q(1-k/n)} = \sum_{i=1}^3 T_{n,i} + o_p(1)$$

où

$$T_{n,1} = -\int_0^{1-k/n} \frac{\mathcal{B}_n(s)v}{\sqrt{k/n}X_{n-k:n}} dQ(s)$$

$$T_{n,2} = \frac{\gamma^2 v}{(1-\gamma)^2}\sqrt{\frac{k}{n}}\mathcal{B}_n(s)(1-\frac{k}{n})$$

$$T_{n,3} = -\frac{\gamma v}{(1-\gamma)^2}\sqrt{\frac{k}{n}}\int_{1-k/n}^1 \frac{\mathcal{B}_n(s)}{1-s} ds$$

la somme $T_{n,1} + T_{n,2} + T_{n,3}$ est une variable aléatoire Gaussienne centrée. Pour calculer sa variance asymptotique nous établissons la limite suivante

$$\begin{aligned} E[T_{n,1}^2] &\to \frac{2\gamma}{2\gamma-1}, \; E[T_{n,2}^2] \to \frac{\gamma^4}{(1-\gamma)^4} \\ E[T_{n,3}^2] &\to \frac{\gamma^2}{(1-\gamma)^4}, \; E[T_{n,1}T_{n,2}] \to \frac{\gamma^2}{(1-\gamma)^2} \\ E[T_{n,1}T_{n,3}] &\to \frac{\gamma}{(1-\gamma)^2}, E[T_{n,2}T_{n,3}] \to \frac{\gamma^3}{(1-\gamma)^4}. \end{aligned}$$

■

4.3 Kernel type estimateur

More generally, Csörgő et al. (1985) [28] extended the Hill estimator (2.5.1) into a kernel class of estimators as follows :

$$\widehat{\gamma}_{n,k}^{K} = \frac{1}{k}\sum_{i=1}^{k} K\left(\frac{i}{k+1}\right) Y_i, \tag{4.27}$$

where K is a kernel integrating to one and $Y_i = i\,(\log X_{n-i+1,n} - \log X_{n-i,n})$.
Note that the Hill estimator corresponds to the particular case where $K = \underline{K} := 1_{(0,1)}$.

In this spirit, we can construct a new estimator of Z_F for a heavy tailed distribution satisfying the second order condition as follow :

$$\widetilde{Z}_{n,k}^{K} = 1 - \int_0^1 \left(\frac{\widetilde{TVaR}_{n,k}^{*}(t)}{\widetilde{TVaR}_{n,k}(t)}\right) dt, \tag{4.28}$$

where $\widetilde{TVaR}_{n,k}(t)$ is a semi parametric estimator of $TVaR_{n,k}(t)$, we can rewrite $TVaR_{n,k}(t)$ as follow :

$$\begin{aligned} TVaR_{n,k}(t) &= \frac{1}{(1-t)}\int_t^{1-k/n} \mathbb{Q}(s)ds + \frac{1}{(1-t)}\int_0^{k/n} \mathbb{Q}(1-s)ds \\ &= TVaR_{n,k}^{(1)}(t) + TVaR_{n,k}^{(2)}(t), \end{aligned}$$

then, an estimator of $TVaR_{n,k}(t)$ is defined by the formula :

$$\widetilde{TVaR}_{n,k}(t) = \widetilde{TVaR}_{n,k}^{(1)}(t) + \widetilde{TVaR}_{n,k}^{(2)}(t) \tag{4.29}$$

$$= \frac{1}{(1-t)}\int_t^{1-k/n} \mathbb{Q}_n(s)ds + \frac{(k/n)\,X_{n-k,n}}{(1-t)\left(1-\widehat{\gamma}_{n,k}^{K}\right)}, \tag{4.30}$$

where $\mathbb{Q}_n(s)$ is a trimmed empirical estimaor of the quantile for $s \in (t, 1-k/n)$. To estimate $TVaR_{n,k}^{(2)}(t)$ we use a extreme quantile, known by Weissman-type estimator [167] for $\mathbb{Q}$, such that :

$$\hat{\mathbb{Q}}_n^{W}(1-s) := X_{n-k,n}\ (k/n)^{\widehat{\gamma}_{n,k}^{K}}\ s^{-\widehat{\gamma}_{n,k}^{K}},\ s \to 0. \tag{4.31}$$

where $\widehat{\gamma}_{n,k}^{K}$ is the kernel class of Hill estimator of the tail index γ.

Also, we define an estimator for $TVaR_F^{*}$ as follow

$$\widetilde{TVaR}_{n,k}^{*} = \frac{1}{t}\left(\widetilde{TVaR}_{n,k}(0) + (1-t)\,\widetilde{TVaR}_{n,k}(t)\right).$$

Thus, the estimator (4.28) already proposed by Greselin et al., 2014 [92] in the particular case where $K = \underline{K} := 1_{(0,1)}$.
Asymptotic normality for $\widetilde{Z}^K_{n,k}$ is obviously related to the one $\widehat{\gamma}^K_{n,k}$ estimator. As usual in the extreme value framework, to prove such type of results, we need a second-order condition on the tail quantile function $\mathbb{U}$, defined as

$$\mathbb{U}(z) = \inf\{y : F(y) \geq 1 - 1/z\}, \; z > 1. \tag{4.32}$$

We say that the function $\mathbb{U}$ satisfies the second-order regular variation condition with second-order parameter $\rho \leq 0$ if there exists a function $A(t)$ which does not change its sign in a neighbourhood of infinity and such that, for every $x > 0$,

$$\lim_{t\to\infty} \frac{\log \mathbb{U}(tx) - \log \mathbb{U}(t) - \gamma \log(x)}{A(t)} = \frac{x^\rho - 1}{\rho}, \tag{4.33}$$

when $\rho = 0$, then the ratio on the right-hand side of equation (4.33) should be interpreted as $\log x$. As an example of heavy-tailed distributions satisfying the second order condition, we have the so called and frequently used Hall's model which is a class of cdf's, such that

$$\mathbb{U}(t) = ct^\gamma (1 + dA(t)/\rho + o(t^\rho)) \text{ as } t \to \infty. \tag{4.34}$$

where $\gamma > 0$, $\rho \leq 0$, $c > 0$, and $d \in \mathbb{R}^*$.
This sub-class of heavy-tailed distributions contains the Pareto, Burr, Fréchet and t–Student, cdf's usually used, in economic and insurance mathematics, as models for dangerous risks. For statistical inference concerning the second-order parameter ρ we refer, for example, to Peng and Qi (2004) [138], Gomes *et al.*, (2005) [78], Gomes and Pestana (2007) [79].
The study of the asymptotic distributions of the estimator given by equation (4.28) for the uniform kernel $K = \underline{K} := 1_{(0,1)}$, is established by Greselin et al., (2014) [92] under the second order framework (4.33), is asymptotically normal with null mean value whenever $\sqrt{k}A(n/k) \to 0$, but there appears a non-null asymptotic bias, whenever $\sqrt{k}A(n/k) \to \lambda \neq 0$, finite.
In this paper we are going to base on reduced bias estimators of the Zenga index, even when $\sqrt{k}A(n/k) \to \lambda$ finite, non-necessarily null, our procedure based on the exponential regression model.

4.3.1 Propriétés asymptôtiques

To study the asymptotic normality of the estimator of $\widetilde{Z}^K_{n,k}$, we need some results and classical assumptions about the kernel :

Condition ($\mathcal{K}$) : Let K be a function defined on $(0, 1]$

CK1. $K(s) \geq 0$ whenever $0 < s \leq 1$ and $K(1) = K'(1) = 0$;

CK2. $K(\cdot)$ is differentiable, nonincreasing and right continuous on $(0,1]$;

CK3. K and K' are bounded (K' is the derivative function of K);

CK4. $\int_0^1 K(u)du = 1$;

CK5. $\int_0^1 u^{-1/2}K(u)du < \infty$.

4.3.2 Asymptotic result for the $\widetilde{Z}^K_{n,k}$ estimator

Théorème 4.3 *Assume that F satisfies the condition (4.33) for $\gamma \in (1/2, 1)$. If further $(\mathcal{K})$ holds and the sequence k satisfies $k \to \infty$ and $k/n \to 0$ as $n \to \infty$, we have*

$$\frac{\sqrt{k}}{(k/n)\,\mathbb{U}(n/k)}\left(\widetilde{Z}^K_{n,k} - Z_F\right) \overset{d}{=} \sqrt{k}A\left(\frac{n}{k}\right)\mathcal{AB}_K(\gamma,\rho) + \mathbb{W}_{1,n} + \mathbb{W}_{2,n}(K) + \mathbb{W}_{3,n} + o_{\mathbb{P}}(1),$$

where

$$\begin{aligned}
\mathcal{AB}_K(\gamma,\rho) &= \left(\frac{v(1)}{(1-\gamma)(\gamma+\rho-1)} + \frac{v(1)}{(1-\gamma)^2}\int_0^1 s^{-\rho}K(s)ds\right);\\
\mathbb{W}_{1,n} &= \frac{\gamma v(1-k/n)}{(1-\gamma)}\sqrt{\frac{n}{k}}\mathbf{B}_n\left(1-\frac{k}{n}\right);\\
\mathbb{W}_{2,n}(K) &= -\frac{\gamma v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\left\{\int_0^1 \frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(sK(s)\right)\right\};\\
\mathbb{W}_{3,n} &= -\frac{\int_0^{1-k/n} v(s)\mathbf{B}_n(s)}{(k/n)^{1/2}\mathbb{Q}(1-k/n)}d\mathbb{Q}(s) + o_{\mathbf{P}}(1),
\end{aligned}$$

and

$$v(s) = \int_0^s \frac{TVaR^*_F(t)}{(1-t)\,TVaR^2_F(t)}dt.$$

Now, by computing the asymptotic variances of the different processes appearing in Theorem 4.3, we deduce the following corollary.

Corollaire 4.2 *Under the same assumptions of Theorem 4.3, if $\sqrt{k}A(n/k) \to \lambda \in \mathbb{R}$, we have*

$$\frac{\sqrt{k}}{(k/n)\,\mathbb{U}(n/k)}\left(\widetilde{Z}^K_{n,k} - Z_F\right) \xrightarrow{d} \mathcal{N}\left(\lambda\mathcal{AB}_K(\gamma,\rho), \mathcal{AC}_K(\gamma,\rho)\right),$$

where

$$\mathcal{AC}_K(\gamma,\rho) = \frac{v^2(1)\gamma^2}{(1-\gamma)^2(2\gamma-1)} + \frac{v^2(1)\gamma^2}{(1-\gamma)^4}\int_0^1 K^2(s)\,ds.$$

The Corollary 4.2 generalizes the result of the Theorem 1 in Greaselin, F. et al. (2014)[92] in case $\lambda \neq 0$ and when we use a general kernel instead of K.

In view of these results, $\widetilde{Z}^K_{n,k}$ is an estimator of Z_F with an asymptotic bias given by

$$(k/n)\,\mathbb{U}(n/k)A(n/k)\mathcal{AB}_K(\gamma,\rho).$$

For a specific kernel, the asymptotic bias and variance can be computed. For instance, we have the following corollary 4.3 if $K = \underline{K}$.

Corollaire 4.3 *Under the assumptions of Corollary 4.2 and in the special case where $K = \underline{K}$, we have*

$$\frac{\sqrt{k}\left(\widetilde{Z}^K_{n,k} - Z_F\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(\lambda \frac{\gamma\rho v(1)}{(1-\rho)(\gamma+\rho-1)(1-\gamma)^2}, \sigma^2_\gamma\right) \tag{4.35}$$

for any fixed $t \in (0,1)$, where the asymptotic variance σ^2_γ is given by the formula

$$\sigma^2_\gamma = \frac{\gamma^4 v^2(1)}{(1-\gamma)^4(2\gamma-1)}.$$

The next step is to propose a reduced-bias estimator of Z_F.

4.3.3 Bias-correction for the $\widetilde{Z}^K_{n,k}$ estimator [159]

The problem of reduced bias of inequality measures estimation is a well known in the literature, has been addressed recently by several authors, among whom we mention Fichtenbaum and Shahidi, 1988 [60], Breunig, 2002 [18], Deltas 2003 [47], all these researchers consider the possibility of dealing with the bias term in an appropriate way, building different new reduced bias estimators.
For the kernel-type estimator $\widetilde{Z}^K_{n,k}$, we recall that, from Theorem 4.3,

$$\widetilde{Z}^K_{n,k} - (k/n)\,\mathbb{U}(n/k)A(n/k)\mathcal{AB}_K(\gamma,\rho),$$

is an asymptotically unbiased estimator for Z_F. Note that γ, ρ, $\mathbb{U}(n/k)$ and $A(n/k)$ are unknown quantities that we have to estimate. Under the condition (4.33), Feuerverger and Hall (1999) [59] and Beirlant et al. (1999, 2002) [10, 11] propose the following exponential regression model for the log-spacings of order statistics :

$$Y_i \sim \left(\gamma + A\left(\frac{n}{k}\right)\left(\frac{i}{k+1}\right)^{-\rho}\right) + \epsilon_{i,k},\ 1 \le i \le k, \tag{4.36}$$

where $\epsilon_{i,k}$ are zero-centered error terms. If we ignore the term $A(n/k)$ in equation (4.36), we obtain the Hill estimator $\widehat{\gamma}^H_{n,k}$ by taking the mean of the left-hand side of (4.36). By using a least-squares approach, (4.36) can be further exploited to propose a reduced-bias estimator for γ in which ρ is substituted by a consistent estimator $\widehat{\rho} = \widehat{\rho}(n,k)$ (see for instance Beirlant et al., 2002 [11] and Fraga Alves et al., 2003)

[66] or by a canonical choice, such as $\rho = -1$(see e.g. Feuerverger and Hall (1999) [59] or Beirlant et al., (1999)[10]). The least squares estimators for γ and $A(n/k)$ are then given by

$$\begin{aligned}\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}) &= \frac{1}{k}\sum_{i=1}^{k} Y_i - \frac{\widehat{A}_{n,k}^{L.S}(\widehat{\rho})}{1-\widehat{\rho}};\\ \widehat{A}_{n,k}^{L.S}(\widehat{\rho}) &= \frac{(1-2\widehat{\rho})(1-\widehat{\rho})^2}{\widehat{\rho}^2}\frac{1}{k}\sum_{i=1}^{k}\left(\left(\frac{i}{k+1}\right)^{-\widehat{\rho}} - \frac{1}{1-\widehat{\rho}}\right)Y_i.\end{aligned}$$

Note that $\widehat{\gamma}_{n,k}^{L.S}(\rho)$ can be viewed as the kernel estimator $\widehat{\gamma}_{n,k}^{K_\rho}$, where for $0 < u \leq 1$:

$$K_\rho(u) := \frac{1-\rho}{\rho}\underline{K}(u) + \left(1 - \frac{1-\rho}{\rho}\right)\underline{K}_\rho(u) \tag{4.37}$$

with $\underline{K}(u) = 1_{\{0<u<1\}}$ and $\underline{K}_\rho(u) = \left(\frac{1-\rho}{\rho}\right)(u^{-\rho}-1)1_{\{0<u<1\}}$, both kernels satisfying condition $(\mathcal{K})$. On the contrary $\underline{K}_\rho$ does not satisfy statement (CK1) in $(\mathcal{K})$. We refer to Gomes and Martins (2004) [77] and Gomes et al., (2007) [80] for other techniques of bias reduction based on the estimation of the second order parameter.

We are now able to obtain a reduced-bias estimator for the Zenga index Z_F from condition (4.33) and using the above estimators for the different unknown quantities :

$$\widetilde{Z}_{n,k}^{K,\widehat{\rho}} = \widetilde{Z}_{n,k}^{K} - \left(\frac{k}{n}\right)X_{n-k,n}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right). \tag{4.38}$$

The asymptotic normality of $\widetilde{Z}_{n,k}^{K,\widehat{\rho}}$ is established in the theorem 4.4.

Théorème 4.4 *Under the assumptions of Theorem 4.3, if $\widehat{\rho}$ is a consistent estimator for ρ, then we have*

$$\frac{\sqrt{k}\left(\widetilde{Z}_{n,k}^{K,\widehat{\rho}} - Z_F\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \widetilde{\mathcal{AC}}_K(\gamma,\rho)\right),$$

where

$$\begin{aligned}\widetilde{\mathcal{AC}}_K(\gamma,\rho) &= \mathcal{AC}_K(\gamma,\rho) + \frac{\gamma^2}{\rho^2}(1-2\rho)(1-\rho)^2\mathcal{AB}_K^2(\gamma,\rho)\\ &\quad + \frac{2\gamma^2(1-2\rho)(1-\rho)}{\rho^2(1-\gamma)^2}\left(1-(1-\rho)\int_0^1 \frac{K(s)}{s^\rho}ds\right)\mathcal{AB}_K(\gamma,\rho).\end{aligned}$$

Let us observe that $\widetilde{Z}_{n,k}^{K,\widehat{\rho}}$ has a null asymptotic bias, which was not the case for $\widetilde{Z}_{n,k}^{K}$ (Corollary 4.2).

Corollaire 4.4 *Under the same assumptions as in Theorem 4.4 and in the special case where* $K = \underline{K}$, *we have*

$$\frac{\sqrt{k}\left(\widetilde{Z}_{n,k}^{K,\widehat{\rho}} - Z_F\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \frac{\gamma^4(\gamma-\rho)^2 v^2(1)}{(2\gamma-1)(\gamma+\rho-1)^2(1-\gamma)^4}\right).$$

Now, in the special case where $K = K_\rho$, as already mentioned, the estimator $\widehat{\gamma}_{n,k}^{K_\rho}$ coincides with $\widehat{\gamma}_{n,k}^{L.S}(\rho)$.
The aim of the next corollary is to establish the asymptotic normality of the resulting Zenga estimator $\widetilde{Z}_{n,k}^{K_\rho,\widehat{\rho}}$, denoted by $\widetilde{Z}_{n,k}^{L.S,\widehat{\rho}}$, when the least squares approach is adopted.

Corollaire 4.5 *Under the same assumptions as in Theorem 4.4 and in the special case where* $K = K_\rho$, *we have*

$$\frac{\sqrt{k}\left(\widetilde{Z}_{n,k}^{L.S,\widehat{\rho}} - Z_F\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \widetilde{\mathcal{AC}}_{K_\rho}(\gamma,\rho)\right),$$

where

$$\begin{aligned}\widetilde{\mathcal{AC}}_{K_\rho}(\gamma,\rho) &= \frac{\gamma^2(1-\rho)^2}{\rho^2(1-\gamma)^4}v^2(1) + \frac{v^2(1)\gamma^2}{(2\gamma-1)(1-\gamma)^2} \\ &\quad + \frac{\gamma^2(1-2\rho)(1-\rho)(\gamma\rho+2\rho+\gamma-\rho-1)}{\rho^2(1-\gamma)^3(\gamma+\rho-1)^2}v^2(1).\end{aligned}$$

4.4 Simulation study

In this section we examine via Monte Carlo simulations a finite-sample performance of the proposed (asymptotic) Zenga index measure, with particular emphasis on comparisons and illustrations of the performance of the biased estimator $\widetilde{Z}_{n,k}^{K}$ and the reduced-bias estimator $\widetilde{Z}_{n,k}^{L.S,\widehat{\rho}}$, we compare the two estimators in terms of the bias and Root Mean Square Error (RMSE), we note that bias1 and RMSE1 are for the estimator $\widetilde{Z}_{n,k}^{K}$ and bias2 and RMSE2 are for the estimator $\widetilde{Z}_{n,k}^{L.S,\widehat{\rho}}$, also we compare between the coverage probability (cov prob) of the bias and the RMSE for a significance level $\zeta = 0.95$, through its application to sets of samples taken from Pareto distribution with tail of distribution

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

and Fréchet distributions with tail of distribution

$$\overline{F}(x) = 1 - \exp\left(-x^{-1/\gamma}\right), x > 0,$$

in this study, we consider two values of tail index ($\gamma = 2/3$ and $\gamma = 3/4$), and the second order parameter $\rho = -1$, we generate 1000 independent replicates of samples

sizes 1000, 2000 and 4000 from the selected parent distribution. With the optimal values of k, we estimate the following minimal Asymptotic Mean Squared Error (AMSE) values of $\widehat{\gamma}^{K}_{n,k_{opt}}$, several procedures have been suggested in the literature, and we refer to, e.g., (Dekkers and de Haan, 1993 [46], Drees and Kaufmann, 1998 [51], Danielsson et al., 2001 [32], Cheng and Peng, 2001 [22], Neves and Fraga Alves, 2004 [134] and references therein.

In our current study we employ the method of Cheng and Peng (2001) [22] for deciding on an appropriate value k_{opt} of k. For each simulated sample, we obtain an approximation of the estimators of Zenga index Z_F. The overall estimated Z_F is then taken as the empirical mean of the values in the 1000 repetitions. To this end, we summarize the results in Table 4.1 and Table 4.2

Table 4.1 – Simulations results based on Pareto distribution with $\gamma = 3/4$ and $\gamma = 2/3$, the corresponding Zenga index values 0.8247 and 0.7590, respectively

γ	2/3				3/4			
	bias1	RMSE1	bias2	RMSE2	bias1	RMSE1	bias2	RMSE2
n=1000	-0.0024	0.0075	-0.0012	0.0031	-0.0026	0.0069	-0.0016	0.0038
n=2000	-0.0018	0.0049	-0.0013	0.0023	-0.0019	0.0055	-0.0012	0.0029
n=4000	-0.0013	0.0036	-0.0010	0.0019	-0.0013	0.0038	-0.0010	0.0015
cov prob								
n=1000	0.627	0.665	0.714	0.755	0.605	0.685	0.669	0.775
n=2000	0.635	0.676	0.726	0.762	0.619	0.697	0.682	0.782
n=4000	0.639	0.681	0.744	0.785	0.632	0.701	0.705	0.796

Table 4.2 – Simulations results based on Fréchet distribution with $\gamma = 3/4$ and $\gamma = 2/3$, the corresponding Zenga index values 0.8652 and 0.8229, respectively

γ	2/3				3/4			
	bias1	RMSE1	bias2	RMSE2	bias1	RMSE1	bias2	RMSE2
n=1000	0.0019	0.0073	0.0013	0.0055	0.0021	0.0055	0.0012	0.0024
n=2000	0.0010	0.0048	0.0009	0.0026	0.0015	0.0034	0.0011	0.0019
n=4000	0.0008	0.0022	0.0006	0.0012	0.0012	0.0023	0.0009	0.0010
cov prob								
n=1000	0.625	0.701	0.734	0.755	0.596	0.659	0.682	0.748
n=2000	0.633	0.713	0.747	0.774	0.613	0.665	0.699	0.768
n=4000	0.639	0.721	0.762	0.781	0.629	0.671	0.718	0.782

4.5 Proofs

Let $Y_1, ..., Y_n$ be i.i.d. r.v.'s from the unit Pareto distribution G, defined as $G(y) = 1 - 1/y, y > 1$. For each $n \geq 1$, let $Y_{1,n} \leq ... \leq Y_{n,n}$ be the order statistics pertaining to $Y_1, ..., Y_n$. Clearly

$$X_{j,n} \stackrel{d}{=} \mathbb{U}(Y_{j,n}), j = 1, ..., n.$$

In order to use results from Csörgö et al., (1986)[27], a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is constructed carrying a sequence $1, 2, ...$ of independent random variables uniformly distributed on $(0,1)$ and a sequence of Brownian bridges $\{\mathbf{B}_n(s), 0 \leq s \leq 1, n = 1, 2....\}$. The resulting empirical quantile is denoted by

$$\beta_n(r) = \sqrt{n}(r - \mathbb{V}(r)),$$

where

$$\mathbb{V}(r) = \zeta_{j,n}, \frac{j-1}{n} \leq r \leq \frac{j}{n}, j = 1, 2, ..., n \quad \text{and} \quad \mathbb{V}(0) = 0.$$

Proof of Theorem 4.3. Theorem 4.3 follows from the asymptotic expansion

$$\widetilde{Z}^K_{n,k} = 1 - \int_0^1 \left(\frac{\widetilde{TVaR}^{*K}_{n,k}(t)}{\widetilde{TVaR}^{K}_{n,k}(t)} \right) dt,$$

we have

$$\begin{aligned}
\frac{\sqrt{k}}{\sqrt{k/n}\mathbb{U}(n/k)} \left(\widetilde{Z_{n,k}}^K - Z_F \right) &= -\int_0^1 \frac{1}{TVaR_F(t)} \sqrt{n} \frac{\widetilde{TVaR}^{*K}_{n,k}(t) - TVaR^*_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&\quad + \int_0^1 \frac{TVaR^*_F(t)}{TVaR^2_F(t)} \sqrt{n} \frac{\widetilde{TVaR}^{K}_{n,k}(t) - TVaR_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&\quad + \int_0^1 \left(\frac{1}{TVaR^K_{n,k}(t)} - \frac{1}{TVaR_F(t)} \right) \\
&\quad \times \sqrt{n} \frac{\widetilde{TVaR}^{*K}_{n,k}(t) - TVaR^*_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&\quad + \int_0^1 \frac{TVaR^*_F(t)}{TVaR_F(t)} \left(\frac{1}{\widetilde{TVaR}^{K}_{n,k}(t)} - \frac{1}{TVaR_F(t)} \right) \\
&\quad \times \sqrt{n} \frac{\widetilde{TVaR}^{K}_{n,k}(t) - TVaR_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Greselin et al. 2014[92] shows that $I_3 = o_P(1)$ and $I_4 = o_P(1)$ when $n \to \infty$.
On the other hand, we have

$$(1-t)\left(\widetilde{TVaR}^{K}_{n,k}(t) - TVaR_F(t) \right) = A_{n,1} + A_{n,2}(t),$$

and

$$t\left(\widetilde{TVaR}^{K*}_{n,k}(t) - TVaR^*_F(t) \right) = A_{n,2}(0) - A_{n,2}(t),$$

for all t such that $0 \leq t \leq 1 - k/n$, where

$$\begin{aligned} A_{n,1} &= \left(\frac{k}{n}\right)\frac{X_{n,k}}{1-\widehat{\gamma}_{n,k}^{K}} - \int_{1-k/n}^{1} Q(s)ds \\ A_{n,2}(t) &= \int_{t}^{1-k/n} (Q_n(s) - Q(s))\, ds \end{aligned}$$

We have, from Rassoul, 2013 [142], that

$$\frac{\sqrt{k}A_{n,1}(t)}{(k/n)\mathbb{Q}(1-k/n)} = -\frac{\int_0^{1-k/n}\mathbf{B}_n(s)d\mathbb{Q}(s)}{(k/n)^{1/2}\mathbb{Q}(1-k/n)} + o_{\mathbf{P}}(1) = \mathbb{W}_{n,3} \tag{4.39}$$

and

$$\frac{\sqrt{k}A_{n,2}(t)}{(k/n)\,\mathbb{Q}(1-k/n)} = \sqrt{k}A(n/k)\mathcal{AB}_K(\gamma,\rho) + \mathbb{W}_{1,n} + \mathbb{W}_{2,n}. \tag{4.40}$$

The proof of statement (4.39) is similar to that of Theorem 2 in Necir *et al.*, (2010) [133], though some adjustments are needed since, we are now concerned with the Zenga index. Therefore, we present the main blocks of the proof together with pinpointed references to Necir *et al.*, (2010) [133] for specific technical details.
Next, from theorem 1 in Rassoul (2013) [142], we have

$$\begin{aligned} &\frac{(1-t)\sqrt{k}}{(k/n)\,\mathbb{U}(n/k)}\left(\widetilde{TVaR}_{n,k}^{K}(t) - TVaR_F(t)\right) \\ &\stackrel{d}{=} \sqrt{k}A\left(\frac{n}{k}\right)\mathcal{AB}_K(\gamma,\rho) + \mathbb{W}_{1,n} + \mathbb{W}_{2,n}(K) + \mathbb{W}_{3,n} + o_{\mathbb{P}}(1), \end{aligned}$$

where

$$\begin{aligned} \mathcal{AB}_K(\gamma,\rho) &= \left(\frac{v(1)}{(1-\gamma)(\gamma+\rho-1)} + \frac{v(1)}{(1-\gamma)^2}\int_0^1 s^{-\rho}K(s)ds\right); \\ \mathbb{W}_{1,n} &= \frac{\gamma v(1-k/n)}{(1-\gamma)}\sqrt{\frac{n}{k}}\mathbf{B}_n\left(1-\frac{k}{n}\right)(1+o_{\mathbb{P}}(1)); \\ \mathbb{W}_{2,n}(K) &= -\frac{\gamma v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\left\{\int_0^1 \frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d(sK(s))\right\}; \\ \mathbb{W}_{3,n} &= -\frac{\int_0^{1-k/n} v(s)\,\mathbf{B}_n(s)}{(k/n)^{1/2}\mathbb{Q}(1-k/n)}d\mathbb{Q}(s) + o_{\mathbf{P}}(1). \end{aligned}$$

Finally

$$\frac{\sqrt{k}}{\sqrt{k/n}\mathbb{U}(n/k)}\left(\widetilde{Z}_{n,k}^{K} - Z_F\right) \stackrel{d}{=} \sqrt{k}A\left(\frac{n}{k}\right)\mathcal{AB}_K(\gamma,\rho) + \mathbb{W}_{1,n} + \mathbb{W}_{2,n}(K) + \mathbb{W}_{3,n} + o_{\mathbb{P}}(1).$$

■

Proof of corollary 4.2. From Theorem 4.3, the sum $\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}$ is a centered Gaussian random variable. To calculate its asymptotic variance, the computations are tedious but quite direct. the classical Sultsky ?s lemma completes the proof of corollary 4.2. ■

Proof of corollary 4.3. The proof of the corollary 4.3 is a direct result of the corollary 4.2 with the kernel $K=\underline{K}=1_{(0,1)}$. ■

Proof of Theorem 4.4.

We have

$$\frac{\sqrt{k}\left(\widetilde{Z}_{n,k}^{K,\widehat{\rho}}-Z_F\right)}{(k/n)\mathbb{U}(n/k)}\overset{d}{=}\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}+\mathbb{W}_{4,n}+o_{\mathbb{P}}(1),$$

where

$$\begin{aligned}
\mathbb{W}_{4,n} &= \sqrt{k}\left(A(n/k)\mathcal{AB}_K(\gamma,\rho)-\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\frac{X_{n-k,n}}{\mathbb{U}(n/k)}\right)\\
&= -\mathcal{AB}_K(\gamma,\rho)\sqrt{k}\left(\widehat{A}_{n,k}^{L.S}(\widehat{\rho})-A(n/k)\right)\\
&\quad -\sqrt{k}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\left(\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)-\mathcal{AB}_K(\gamma,\rho)\right)\\
&\quad -\sqrt{k}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\left(\frac{X_{n-k,n}}{\mathbb{U}(n/k)}-1\right)\\
&\quad \overset{d}{=} -\mathcal{AB}_K(\gamma,\rho)\gamma(1-\rho)\sqrt{\frac{n}{k}}\left\{\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(s\left(\underline{K}(s)-K_\rho(s)\right)\right)\right\}\\
&\quad +o_{\mathbb{P}}(1).
\end{aligned}$$

By the result of Lemma 5 of Girard and Guillou, (2013) [70], for any consistent estimator $\widehat{\rho}$ of ρ, we have

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho})-\gamma\right)\overset{d}{=}\gamma\sqrt{\frac{n}{k}}\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(sK_\rho(s)\right)+o_{\mathbb{P}}(1),\tag{4.41}$$

and

$$\begin{aligned}
&\sqrt{k}\left(\widehat{A}_{n,k}^{L.S}(\widehat{\rho})-A\left(\frac{n}{k}\right)\right)\\
&\overset{d}{=}\gamma(1-\rho)\sqrt{\frac{n}{k}}\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(s\left(\underline{K}(s)-K_\rho(s)\right)\right)+o_{\mathbb{P}}(1),
\end{aligned}\tag{4.42}$$

and by using the consistency and the inequality $|\frac{e^x-1}{x}-1|\leq e^{|x|}-1$ for all $x\in\mathbb{R}$. Moreover, direct computations lead to the desired asymptotic variance which ends the proof of Theorem 4.4. ■

Proof of corollary 4.4. The proof of the corollary 4.4 is a direct result of the Theorem 4.4 with the kernel $K = \underline{K} = 1_{(0,1)}$. ■

Prooof of corollary 4.5. Recall that K_ρ does not satisfy condition $(\mathcal{K})$ but it can be rewritten as (4.37) with both K and K_ρ satisfying $(\mathcal{K})$. So, following the lines of the proof of Theorem 4.4, Corollary 4.5 follows. ■

Conclusion générale

La théorie des valeurs extrêmes offre un outil utile pour modéliser la distribution des valeurs extrêmes. La théorie fournit des méthodes pour modéliser les queues spécifiquement sans faire attention au centre de la distribution. Cette approche incorpore autant d'informations que possible sur les queues et exploite au mieux le petit nombre d'extrêmes précédemment observés. Cette thèse était axée sur les indices des inégalités et traitait donc de la gestion des risques d'inégalité dans un cadre de distribution univarié.

Pour identifier les extrêmes dans un échantillon observé, la méthode des maxima dans des blocs de temps a été choisie par rapport à la sélection des dépassements par rapport à un seuil donné.

Par conséquent, l'inférence statistique était liée à la distribution de type de Pareto. Pour s'adapter au modèle de distribution, une approche semi-paramétrique a été utilisée.

Ici, le paramètre d'intérêt principal était l'indice de queue et sa valeur réciproque (paramètre de forme) car il caractérise avec précision le comportement de la queue.

Plus le paramètre de forme est élevé, plus la queue est lourde. La méthode de Hill a été utilisée pour l'estimation des paramètres car elle a fait ses preuves dans la pratique. Il a également été choisi en raison de son utilisation générale et de sa simplicité de calcul.

Les lois GEV se distinguent, essentiellement selon le paramètre ξ. Ainsi, l'estimation sera établie selon deux cas : $\xi = 0$ (loi de Gumbel) et $\xi \neq 0$ (lois de Fréchet et de Weibull). Nous avons mentionné dans les deux premiers chapitres que le paramètre d'importance en théorie des valeurs extrs est l ?indice 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 paramètre est de grand 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 dernier chapitre, nous nous sommes intéressés à l'estimation d'un indice des inégalité pour des revenus extrêmes, et plus précisement

de l'estimation de l'indice de Zenga dans le cas des distribution des revenus 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 méthode 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 l'inégalité des revenus dans la population. Nous pouvons donc conclure que la théorie des valeurs extrêmes peut figurer parmi les outils d'analyse pour l'économie d'un pays.

Bibliographie

[1] Aarssen, K., & De Haan, L. (1994). On the maximal life span of humans. Mathematical Population Studies, 4(4), 259-281.

[2] Abel, A.B. (2007). Optimal capital income taxation. NBER Working Paper No. 13354. http://www.nber.org/papers/w13354.

[3] Alvarado, E., Sandberg, D. V., & Pickford, S. G. (1998). Modeling large forest fires as extreme events. National Emergency Training Center.

[4] Arnold, B. C. (1983). Pareto distributions(International Cooperative Publishing House, Fairland, MD).

[5] Atkinson, A. B. (1970). On the measurement of inequality. Journal of economic theory, 2(3), 244-263.

[6] Bacro, J. N., & Brito, M. (1995). Weak limiting behaviour of a simple tail Pareto-index estimator. Journal of Statistical Planning and inference, 45(1-2), 7-19.

[7] Barrois, M. T. (1834). Essai sur l'application du calcul des probabilités aux assurances contre l'incendie.

[8] Batana, Y. M. (2007). Dominance stochastique et pauvreté multidimensionnelle dans les pays de l'UEMOA. CIPREE, Université Laval, Canada. 38pages.

[9] Beirlant, J., Vynckier, P., & Teugels, J. L. (1996). Tail index estimation, pareto quantile plots regression diagnostics. Journal of the American statistical Association, 91(436), 1659-1667.

[10] Beirlant, J., Dierckx, G., Goegebeur, M., Matthys, G. (1999). Tail index estimation and an exponential regression model, Extremes, 2, 177-200.

[11] Beirlant, J., Dierckx, G., Guillou, A., Starica, C. (2002). On exponential representations of log-spacings of extreme order statistics, Extremes, 5, 157-180.

[12] Beirlant J., Goegebeur Y., Teugels J. and Segers J., Statistics of extremes. Wiley Series in Probability and Statistics. John Wiley and Sons Ltd, Chichester, 2004.

[13] Benkhaled, A. (2007). Distributions statistiques des pluies maximales annuelles dans la région du Cheliff : comparaison des techniques et des résultats. Courrier du Savoir, 8, 83-91.

[14] Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation, Cambridge University Press, Cambridge.

[15] Bobée, Bernard, and Peter F. Rasmussen. "Recent advances in flood frequency analysis." Reviews of Geophysics 33.S2 (1995) : 1111-1116.

[16] Bonferroni C.E. (1930). Elmenti di statistica generale. Libreria Seber, Firenze.

[17] Bouleau, N. (1991). Splendeurs et misères des lois de valeurs extêmes. Revue Risques-Les cahiers de l'assurance, 4, 85-92.

[18] Breunig R., (2002), Bias correction for inequality measures : an application to China and Kenya, Applied Economics Letters, 9, 12, 783-786.

[19] Brodin, E., & Rootzén, H. (2009). Univariate and bivariate GPD methods for predicting extreme wind storm losses. Insurance : Mathematics and Economics, 44(3), 345-356.

[20] Castillo, E., Hadi, A.S., 1997. Fitting the generalized Pareto distribution to data. J. Amer. Statist. Assoc., 92, 1604-1620.

[21] Chamley, C.(1986). Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54, 607-622.

[22] Cheng, S. and Peng,L. (2001). Confidence intervals for the tail index, Bernoulli, 7(5), 751–760.

[23] Christopeit, N. (1994). Estimating parameters of an extreme value distribution by the method of moments. Journal of statistical planning and inference, 41(2), 173-186.

[24] Coles, S. G., & Walshaw, D. (1994). Directional modelling of extreme wind speeds. Journal of the Royal Statistical Society : Series C (Applied Statistics), 43(1), 139-157.

[25] Coles, S., 2001. An introduction to statistical modeling of extreme values, Springer Series in Statistics, Springer-Verlag, London.

[26] Csörgő, M., Mason, D. M. (1985). Central limit theorems for sums of extreme values. Math. Proc. Camb. Phil. Soc., 98, 547-558.

[27] Csörgő, M., Csörgo, S., Horvärth, L., Mason, D. M. (1986). Weighted empirical and quantile processes. Ann. Probab., 14, 31-85.

[28] Csörgő, S., Deheuvels, P., Mason, D.M. (1985). Kernel estimates of the tail index of a distribution, Annals of Statistics, 13, 1050- 1077.

[29] Csörgő, S., & Viharos, L. (1998). Estimating the tail index. In Asymptotic Methods in Probability and Statistics (pp. 833-881). North-Holland.

[30] Csörgő, M.,Horváth, L., & Shao, Q. M. (1993). Convergence of integrals of uniform empirical and quantile processes. Stochastic processes and their applications, 45(2), 283-294.

[31] Danielsson, J., & De Vries, C. G. (1997). Tail index and quantile estimation with very high frequency data. Journal of empirical Finance, 4(2-3), 241-257.

[32] Danielsson, J., de Haan, L., Peng, L., de Vries, C. G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. Journal of Multivariate analysis, 76(2), 226-248.

[33] Davis, R., & Resnick, S. (1984). Tail estimates motivated by extreme value theory. The Annals of Statistics, 12(4), 1467-1487.

[34] Davison, A. et Smith, R. (1990). Models for exceedances over high thresholds. Journal of the Royal Statistical Society Series B, 52(3) :393–442.

[35] Davison, A. C., Padoan, S. A., & Ribatet, M. (2012). Statistical modeling of spatial extremes. Statistical science, 27(2), 161-186.

[36] Davydov, Youri, and Ričardas Zitikis. "Convex rearrangements of random elements." Asymptotic Methods in Stochastics 44 (2004) : 141-171.

[37] Davydov, Youri, and Ričardas Zitikis. "Generalized Lorenz curves and convexifications of stochastic processes." Journal of Applied Probability 40.4 (2003) : 906-925.

[38] de Haan, L. d. (1976). Sample extremes : an elementary introduction. Statistica Neerlandica, 30 , 161–172.

[39] de Haan, L. et Ferreira, A. (2006). Extreme Value Theory : An Introduction. Springer Series in Operations Research and Financial Engineering, New York Inc.

[40] de Haan L. and Rootzen H., On the estimation of high quartiles J. of Statistical Planning and Inference,35, n.1, 1-13, 1993.

[41] de Haan, L., Peng, L. (1998). Comparison of tail index estimators, Statistica Neerlandica, 52, 60-70.

[42] de Haan, L., Resnick, S. (1996). Second order regular variation and rates of convergence in extreme value theory. Annals of Probability 24, 97-124.

[43] 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.

[44] Deheuvels, P., Haeusler, E., Mason, D.M., 1988. Almost sure convergence of the Hill estimator. Math. Proc. Cambridge Philos. Soc., 104, 2, 371-381.

[45] Dekkers A., Einmahl J. and De Haan L., A moment estimator for the index of an extrerne-value loi, Ann. Statist. 17, p. 1833-1855, 1989.

[46] Dekkers, A. L., de Haan, L. (1993). Optimal choice of sample fraction in extreme value estimation. Journal of Multivariate Analysis, 47(2), 173-195.

[47] Deltas G., (2003), The Small-Sample Bias of the Gini Coefficient : Results and Implications for Empirical Research, The Review of Economics and Statistics, 85, 1, 226-234.

[48] Delmas, Jean-François, and Benjamin Jourdain. Modèles Aléatoires. Vol. 57. Springer-Verlag Berlin Heidelberg, 2006.

[49] Denuit, M., Dhaene, J., Goovaerts,M.J., Kaas, R. (2005). Actuarial Theory for Dependent Risk :Measures, Orders and Models. Wiley, New York.

[50] Donaldson, D., & Weymark, J. A. (1980). A single-parameter generalization of the Gini indices of inequality. Journal of economic Theory, 22(1), 67-86.

[51] Drees, H., Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. Stochastic Processes and their Applications, 75(2), 149-172.

[52] Drees, H. (1998). A general class of estimators of the extreme value index. Journal of Statistical Planning and Inference, 66(1), 95-112.

[53] Drees, H. "Extreme quantile estimation for dependent data, with applications to finance." Bernoulli 9.4 (2003) : 617-65

[54] Drees, H., Ferreira, A., de Haan, L., (2004). On maximum likelihood estimation of the extreme value index. Ann. Applied Probab., 14, 1179-1201.

[55] 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.

[56] Embrechts P., Kluppelberg C. and Mikosch T., Modelling extremal events, Springer Verlagt, Berlin, 1997.

[57] Embrechts, P., Resnick, S. I., & Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. North American Actuarial Journal, 3(2), 30-41.

[58] Ferrez, J., Davison, A. C., & Rebetez, M. (2011). Extreme temperature analysis under forest cover compared to an open field. Agricultural and Forest Meteorology, 151(7), 992-1001.

[59] Feuerverger, A., Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, Annals of Statistics, 27, 760-781.

[60] Fichtenbaum R. and Shahidi H., 1988, Truncation Bias and the Measurement of Income Inequality, Journal of Business.

[61] Fisher, R. A., Tippett, L. (1928). Limiting Forms of the Frequency Distribution of the Largest of Smallest Member of a Sample. volume 24.

[62] Fougère, D., & Kramarz, F. (2001). La mobilité salariale en France de 1967 à 1999. Inégalités économiques, 333-354.

[63] Foster, J., Greer, J., & Thorbecke, E. (1984). A class of decomposable poverty measures. Econometrica : journal of the econometric society, 761-766.

[64] Frädorf, A., Grabka, M. M., Schwarze, J. (2011). The impact of household capital income on income inequality : a factor decomposition analysis for the UK, Germany and the USA. J. Econ. Inequal. 9, 35-56.

[65] Alves, M. F. (2001). A location invariant Hill-type estimator. Extremes, 4(3), 199-217.

[66] Fraga Alves, M.I., de Haan, L. and Lin, T. (2003). Estimation of the parameter controlling the speed of convergence in extreme value theory. Math. Methods Statist., 12, 155-176.

[67] 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.

[68] Gastwirth, J. L. (1972). The estimation of the Lorenz curve and Gini index. The Review of Economics and Statistics, 306-316.

[69] Geluk, J. L., Haan, L., de (1987). Regular variation, Extension and Tauberian Theorems. CWI Tract 40, Amsterdam.

[70] Girard, S., Guillou, A. (2013). Reduced-bias estimator of the Proportional Hazard Premium for heavy-tailed distributions. Insurance : Mathematics and Economics, 52(3), 550-559.

[71] Gomes, M. I., Martins, M. J., & Neves, M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events-the jackknife methodology. Extremes, 3 , 207–229.

[72] Gomes, M. I., & Oliveira, O. (2001). The bootstrap methodology in statistics of extremes—choice of the optimal sample fraction. Extremes, 4(4), 331-358.

[73] 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.

[74] Gomes MI and Oliveira O 2003 Maximum likelihood revisited under a semi-parametric context—estimation of the tail index. Journal of Statistical Planning and Inference 73, 285–301.

[75] Gomes, M. I., & Martins, M. J. (2001). Generalizations of the Hill estimator–asymptotic versus finite sample behaviour. Journal of Statistical Planning and Inference, 93(1-2), 161-180.

[76] Gomes, M. I., & Martins, M. J. (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. Extremes, 5(1), 5-31.

[77] Gomes, M.I., Martins, M.J. (2004). Bias reduction and explicit semi-parametric estimation of the tail index, Journal of Statistical Planning and Inference, 124, 361-378.

[78] Gomes, M.I., Figueiredo, F. and Mendona, S. (2005) Asymptotically best linear unbiased tail estimators under a second-order regular variation condition. J. Stat. Plann. Inference 134, No. 2, 409-433.

[79] Gomes, M.I., Pestana, D., 2007. A sturdy reduced-bias extreme quantile (VaR) estimator. Journal of the American Statistical Association102 (477), 280-292.

[80] Gomes, M.I., Martins, M.J., Neves, M. (2007). Improving second order reduced bias extreme value index estimator, REVSTAT-Statistical Journal, 5(2), 177-207.

[81] Golosov, M., Kocherlakota, N., Tsyvinski, A. (2003). Optimal indirect and capital taxation. Rev. Econ. Studies 70, 569-587.

[82] Goovaerts, M.J., de Vlyder, F., Haezendonck, J. (1984). Insurance premiums, theory and applications, North Holland, Amsterdam.

[83] Gourieroux, C. (1984). Econométrie des variables qualitatives. Paris : Economica.

[84] Galambos, Janos. The asymptotic theory of extreme order statistics. No. 04; QA274, G3.. 1978.

[85] Geluk, J. L., & de Haan, L. (1987). Regular Variation, Extensions and Tauberian Theorems. CWI Tract. Stichting Math. Centrum, Amsterdam. Mathematical Reviews (MathSciNet) : MR89a, 26002.

[86] Gençay, R., & Selçuk, F. (2004). Extreme value theory and Value-at-Risk : Relative performance in emerging markets. International Journal of Forecasting, 20(2), 287-303.

[87] Gini C., 1912, Variabilità e mutabilità, Bologna, Italy.

[88] Gnedenko B., Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math., 44(3), 423-453, 1943.

[89] Greselin, F., Pasquazzi, L. (2009). Asymptotic confidence intervals for a new inequality measure. Comm. Stat. Comput. Simul 38, 17-42.

[90] Greselin, F., Pasquazzi, L., Zitikis, R. (2010) Zenga's new index of economic inequality, its estimation, and an analysis of incomes in Italy. J. Prob. Stat. (Spec. Issue on Actuarial and Financial Risks : Models, Statistical Inference, and Case Studies). Article ID 718905, p. 26.

[91] Greselin, F., Pasquazzi, L., Zitikis, R. (2013). Contrasting the Gini and Zenga indices of economic inequality. J. Appl. Stat. 40, 282-297.

[92] Greselin, F., Pasquazzi, L. and Zitikis, R. (2014) Heavy tailed capital incomes : Zenga index, statistical inference, and ECHP data analysis, Extremes, 17 :127-155.

[93] Greenwood J., Landweher J., Matalas N. and Wallis J., Probability weighted moments : Definition and relation to parameters of several lois expressible in inverse form, Water Resources Research 15, p. 1049-1054, 1979.

[94] Greselin, F., & Pasquazzi, L. (2009). Asymptotic confidence intervals for a new inequality measure. Communications in Statistics-Simulation and Computation, 38(8), 1742-1756.

[95] Greselin, F., Pasquazzi, L., & Zitikis, R. (2013). Contrasting the Gini and Zenga indices of economic inequality. Journal of Applied Statistics, 40(2), 282-297.

[96] Groeneboom, P., LopuhaÄ , H. P., de Wolf, P. P. (2003). Kernel-type estimators for the extreme value index, Annals of Statistics, 31, 1956-1995.

[97] Guillou, A., & Hall, P. (2001). A diagnostic for selecting the threshold in extreme value analysis. Journal of the Royal Statistical Society : Series B (Statistical Methodology), 63(2), 293-305.

[98] Gumbel, E. J. (1937). La durée extrême de la vie humaine (No. 520). Hermann et cie.

[99] Gumbel E.J. and Mustafi C.K., Some analytical properties of bivariate extremal distributions, Am. Stat. Assoc. J., 62, p.569–589, 1967.

[100] Gumbel, E. J., & Lieblein, J. (1954). Some applications of extreme-value methods. The American Statistician, 8(5), 14-17.

[101] Haeusler, E., and Teugels, J. On asymptotic normality of hill's estimator for the exponent of regular variation. The Annals of Statistics. 13(2). (1985), 743–756.

[102] Hall, P. On some simple estimates of an exponent of regular variation. Stat. Scoc., B, 44 :37–42, 1982. (Cité en page 13.)

[103] Hall, P., & Welsh, A. H. (1985). Adaptive estimates of parameters of regular variation. The Annals of Statistics, 13(1), 331-341.

[104] Hall, P. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. Journal of multivariate analysis, 32(2), 177-203.

[105] Han, S. H., & Lee, J. H. (2005). An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. IEEE wireless communications, 12(2), 56-65.

[106] Heneka, P., Hofherr, T., Ruck, B., & Kottmeier, C. (2006). Winter storm risk of residential structures ? model development and application to the German state

of Baden-Württemberg. Natural Hazards and Earth System Science, 6(5), 721-733.

[107] Hill B. M., A simple general approach to inference about the tail of a Loi, Annals of Statistics 5, p. 1163-1174, 1975.

[108] Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. Biometrika, 73(2), 387-396.

[109] Hosking, J., Wallis, J. et Wood, E. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics, 27(3) :251–261.

[110] Hourriez, J. M., & Roux, V. (2001). Une vue d'ensemble des inégalités de revenu et de patrimoine. Inégalités économiques, Rapport pour le Conseil d'Analyse économique, La Documentation Française, Paris, 269.

[111] Hsing, Tailen. "On tail index estimation using dependent data." The Annals of Statistics (1991) : 1547-1569.

[112] Judd, K.L. (2002). Capital-income taxation with imperfect competition. Am. Econ. Rev. 92, 417-421.

[113] Katz R.W, Techniques for estimating Uncertainty in climate change Scenarios and impact studies Climate research, 20, 167-185, 2002.

[114] Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations : A review. Journal of hydrology, 329(3-4), 534-552.

[115] Klajnmic, H. (2004). Estimation et comparaison des niveaux de retour des vitesses extrêmes des vents. XXXVIemes Journées de Statistiques, Montpellier, France.

[116] Kratz, M., & Resnick, S. (1996). The qq-estimator of the index of regular variation. Communications in Statistics : Stochastic Models, 12, 699-724.

[117] Lachaud, J. P. (1996). Croissance économique, pauvreté et inégalité des revenus en Afrique subsaharienne : analyse comparative. Centre d'économie du développement, Université Montesquieu-Bordeaux IV.7.

[118] Lang, M., Renard, B., Kochanek, K., Arnaud, P., Aubert, Y., (2012). Project Extra-Flo : Comparaison de différents cadres d'analyse (locale, régionale, locale-régionale) pour l'estimation de quantiles de crue.

[119] Lerman, R.I., Yitzhaki, S.(1985). Income inequality effects by income source : a new approach and applications to the United States. Rev. Econ. Stat. 67, 151-156.

[120] Lerman, R. I., Yitzhaki, S. (1989). Improving the accuracy of estimates of Gini coefficients. Journal of econometrics, 42(1), 43-47.

[121] Longin, F. M. (2000). From value at risk to stress testing : The extreme value approach. Journal of Banking & Finance, 24(7), 1097-1130.

[122] Lubes H. and Mason J., Méthode des moments de probabilité pondérés. Application à la loi de Jenkinson, Hydrol. continent., vol. 6, n 1, p. 67-84, 1991.

[123] Lye, L., Hapuarachchi, K., and Ryan, S. Bayes estimation of the extreme-value reliability function. IEEE Transactions on Reliability. 42(4). (1993), 641{644.

[124] Jenkinson A.F., The frequency distribution of the annual maximum (or minimum) values of meteorological events, Journal of the Royal Meteorological Society, 81, p. 158-272, 1955.

[125] Matthys, G., & Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models. Statistica Sinica, 853-880.

[126] Mason, D.M., 1982. Laws of large numbers for sums of extreme values. Ann. Probab., 10, 3, 754-764.

[127] Mc Neil A.J. and Frey R, Estimation of tail-related risk measures for heteroscedastic financial time series : An extreme value approach Journal of Empirical Finance, 7, p. 271-300, 2002.

[128] McNeil, A. J., & Saladin, T. (1997, April). The peaks over thresholds method for estimating high quantiles of loss distributions. In Proceedings of 28th International ASTIN Colloquium (pp. 23-43).

[129] McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series : an extreme value approach. Journal of empirical finance, 7(3-4), 271-300.

[130] Mehran, F. (1976). Linear measures of income inequality. Econometrica : Journal of the Econometric Society, 805-809.

[131] Mussard, S. (2006). La décomposition des mesures d'inégalité en sources de revenu : l'indice de Gini et les généralisations. Cahier de recherche/Working Paper, 6, 05.

[132] Necir, A., Rassoul, A. et Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. Journal of Probability and Statistics, ID 596839 :17 pages.

[133] Necir, A., Rassoul, A. and Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses, JPS, doi :10.1155/596839.

[134] Neves, C., Alves, M. F. (2004). Reiss and Thomas' automatic selection of the number of extremes. Computational statistics data analysis, 47(4), 689-704.

[135] Nikias, C. L., & Shao, M. (1995). Signal processing with alpha-stable distributions and applications. Wiley-Interscience.

[136] Nolan, J. P. (2001). Lévy processes : Theory and applications. Barndorff-Nielsen, Ole E. ; Mikosch, Thomas ; Resnick, Sidney I. Chapter maximum likelihood estimation and diagnostics for stable distributions, 379-400.

[137] Pickands J., Statistical inference using extreme order statistics, Annals of Statistics 3, p. 119-131, 1975.

[138] Peng, L. and Qi, Y. (2004). Estimating the first- and second-order parameters of a heavy-tailed distribution. Aust. N. Z. J. Stat. 46, 305-312.

[139] Prescott, P. and WaldenA.T., Maximum likelihood estimation of the parameters of the three-parameter generalized extrerne-value distribution for censored samples, Journal of Statistical Computation and Simulation, 16(3-4) :241-250, 1983.

[140] Prescott, P. and WaldenA.T., Maximum likelihood estimation of the parameters of the generalized extreme-value distribution, Biometrika ,67, 723-724, 1980.

[141] Raggad B., Fondements de la théorie des valeurs extrêmes, ses principales applications et son apport à la gestion des risques du marché pétrolier, Mathematics and Social Sciences. p. 29-63, 2009.

[142] Rassoul, A. (2013). Kernel-type estimator of the conditional tail expectation for a heavy-tailed distribution. Insurance : Mathematics and Statistics

[143] Reiss R.D. and Thomas M., Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, Birkhauser Basel. 2nd edition, 2001.

[144] Resnick S., Extreme Values, Regular variation and Point Processes, Springer-Verlag, New-York, 1987.

[145] Resnick, S., and Stărică, C. (1995). Consistency of Hill's estimator for dependent data. Journal of Applied Probability, 32(1), 139-167.

[146] Rootzén, H., & Tajvidi, N. (1997). Extreme value statistics and wind storm losses : a case study. Scandinavian Actuarial Journal, 1997(1), 70-94.

[147] Rootzén, H., & Tajvidi, N. (2001). Can losses caused by wind storms be predicted from meteorological observations ?. Scandinavian Actuarial Journal, 2001(2), 162-175.

[148] Saez, E. (2005). Top incomes in the United States and Canada over the twentieth century. J. Eur. Econ. Assoc. 3, 402-411.

[149] Sen, A. (1976). Poverty : an ordinal approach to measurement. Econometrica : Journal of the Econometric Society, 219-231.

[150] Shorack, G. R., and J. A. Wellner. "Empirical Processes with Applications to Statistics Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics." New York (1986).

[151] Shorrocks, A. F. (1995). Revisiting the Sen poverty index. Econometrica : Journal of the Econometric Society, 1225-1230.

[152] Schultze, J., & Steinebach, J. (1996). On least squares estimates of an exponential tail coefficient. Statistics & Risk Modeling, 14(4), 353-372.

[153] Sen, A. (1997). Choice, welfare and measurement. Harvard University Press.

[154] Sörensen, P.B. (2007). Can capital income taxes survive? And should they? CESifo Econ. Stud. 53, 172-228.

[155] Smith R.L., Estimating tails of probability lois. The Annals of Statistics 3, p. 1174-1207, 1987.

[156] Smith R.L., Maximum likelihood estimation in a class of nonregular cases. Biometrika 72, p. 67-90, 1985. Economics, 53(3), 698-703.

[157] Smith, R. C., & Stammerjohn, S. E. (2001). Variations of surface air temperature and sea-ice extent in the western Antarctic Peninsula region. Annals of Glaciology, 33, 493-500.

[158] Stedinger, J. R., & Crainiceanu, C. M. (2000). Climate variability and flood-risk management. In 9th United Engineering Foundation Conference on Risk-Based Decisionmaking in Water Resources-Risk-Based Decisionmaking in Water Resources IX (pp. 77-86).

[159] Tami, O., Rassoul, A. and Ould-Rouis, H. (2019) An Improved Estimator of the Zenga Index for Heavy-Tailed Distributions. J. Stat. Appl. Pro. 8, No. 2, 1-12.

[160] Tiago de Oliveira, J. Bivariate extremes : decisions. Bull. Internat. Statist. Inst. XLVI. (1975), 241{251.

[161] Tiago de Oliveira, J. (1982). Decision and modelling for extremes. Some Recent Advances in Statistics, 101-110.

[162] Theil, H. (1967). Economics and information theory (No. 04; HB74. M3, T4.).

[163] Todorovic, P., & Zelenhasic, E. (1970). A stochastic model for flood analysis. Water Resources Research, 6(6), 1641-1648.

[164] Todorovic, P., & Rousselle, J. (1971). Some problems of flood analysis. Water Resources Research, 7(5), 1144-1150.

[165] Von Mises R., La distribution de la plus grande de n valeurs, Revue de Mathématique, Union Interbalcanique, p. 141-160, 1936.

[166] Von Mises R., La distribution de la plus grande de n valeurs, American Mathematical Society, RI, USA, vol. II, pp. 271-294, 1954.

[167] Weissman I., Estimation of Parameters and Large Quantiles Based on the k-Largest Observations, J. Amer. Statist. Assoc, 73, 812-815, 1978.

[168] Wellner, Jon A. "Limit theorems for the ratio of the empirical distribution function to the true distribution function." Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 45.1 (1978) : 73-88.

[169] Yitzhaki, S. (1983). On an extension of the Gini inequality index. International economic review, 617-628.

[170] Yitzhaki, S., Schechtman, E. (2013). More than a dozen alternative ways of spelling Gini. In The Gini Methodology (pp. 11-31). Springer, New York, NY.

[171] Zenga, M. (2007). Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. Stat. Appl. 5, 3-27.

[172] Zipf, G. K. (1949). Human behavior and the principle of least effort.

[173] Zitikis, Ričardas. "-The Vervaat process." Asymptotic methods in probability and statistics. 1998. 667-694.