MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU

FACULTE DES SCIENCES

DEPARTEMENT : MATHEMATIQUES

THESE DE DOCTORAT

SPECIALITE: MATHEMATIQUES

Option : PROBABILITES ET STATISTIQUE

Présentée par:

Mme BOUALAM Karima

Intitulé:

Etude de l'estimateur de Hill sous dépendance faible

Devant le jury d'examen composé de:

M. FELLAG	Hocine	Professeur	U.M.M.T.O	Président.
M. BERKOUN	Youcef	Professeur	U.M.M.T.O	Rapporteur.
Mme DJABALLAH	Khadidja	Professeur	U.S.T.H.B	Examinatrice.
M. TATACHAK	Abdelkader	Professeur	U.S.T.H.B	Examinateur.
M. YOUSFATE	Abderrahmane	Professeur	U.Sidi Bel-Abbès	Examinateur.
Mme ATIL	Lynda	M.C.A	U.M.M.T.O	Examinatrice.

Soutenue le 17 / 05 /2017

Remerciements

Je souhaite exprimer ma plus grande gratitude envers tous celles et ceux qui m'ont aidé et encouragé dans la réalisation de ce travail.

En premier lieu, je tiens à remercier mon directeur de thèse, le professeur ***Berkoun Youcef****, pour la confiance qu'il m'a accordée en acceptant d'encadrer ce travail doctoral, pour ses multiples conseils et pour toutes les heures qu'il a consacrées à diriger cette recherche. Enfin, ses nombreuses relectures et corrections de cette thèse ont été très appréciables. Qu'il trouve ici l'expression de ma profonde gratitude.*

Je remercie sincèrement le Professeur ***Fellag Hocine*** *pour avoir accepté de juger ce travail et d'en présider le jury de soutenance.*

Mes remerciements vont également aux professeurs ***Yousfate Abderrahmane****,* ***Tatachak Abdelkader****,* ***Djaballah Khadidja*** *pour l'honneur qu'ils me font en acceptant de participer à ce jury de thèse et pour l'intérêt qu'ils ont porté à mon travail. Veuillez trouver ici le témoignage de toute mon estime et de ma grande reconnaissance.*

Je tiens aussi à exprimer mes remerciements à Mlle ***Atil Lynda*** *, Maitre de conférences au département de mathématiques, pour avoir accepté de participer à ce jury.*

Je remercie vivement le docteur ***Mamou Mohammed****, pour ses nombreuses remarques et suggestions qui ont contribué largement à l'amélioration de la qualité du document. Qu'il soit assuré de ma profonde reconnaissance.*

Je remercie mes amis du département de mathématiques, ainsi qu'à tous ceux qui m'ont consacré de leur temps. Particulièrement les enseignantes Nassima Moussouni, Karima Fahem et Fariza Arezki qui ont su m'apporter confiance et écoute à tous les moments.

Je dédie une mention spéciale et toute particulière à mes frères et soeurs, à mes parents surtout pour leur patience et leurs encouragements constants.

Merci enfin à mes deux poussins Fares et Ryma, à mon mari Younes, témoins de mes joies, de mes fatigues et de mes enthousiasmes. Vous étiez mon soutien quotidien.

Table des matières

Introduction générale

Ces dernières décennies, on observe une prolifération d'événements extrêmes (inondations, tremblements de terre de forte intensité, ouragans, krachs boursiers, etc.). L'impact de ces événements est considérable tant sur le plan humain qu'économique. La protection contre ces événements revêt donc un intérêt particulier car la prévention de ces risques peut efficacement contribuer à éviter des situations catastrophiques. Un des outils utilisés pour se prémunir contre ces risques est la théorie des valeurs extrêmes.
La théorie des valeurs extrêmes (TVE) est l'étude des événements rares ; c'est à dire des événements dont la probabilité de réalisation est très faible. Les valeurs extrêmes, aussi rares soient-elles, aux conséquences désastreuses, peuvent engendrer des dégâts considérables surtout dans le domaine de la finance et de l'assurance. C'est pourquoi, il est important de modéliser et de prévoir l'occurrence de ces phénomènes. Pour cette raison, les scientifiques portent un intérêt particulier à cette théorie qui représente un champ de recherche très actif aussi bien sur le plan théorique que sur le plan pratique ; elle a trouvé ses applications dans l'anticipation des risques (finance, actuariat, etc.) et vient en complément de la théorie statistique classique où on étudie le comportement d'une distribution autour de sa moyenne plutôt que dans le domaine des observations extrêmes. Dans ce cas, la plus grande et la plus petite valeur d'un échantillon apportent plus d'informations que ses valeurs centrales et permettent une meilleure compréhension et description du comportement de ces événements.
Les domaines d'application utilisant les modèles de la TVE n'ont cessé de se développer et de s'étendre en touchant des secteurs divers et variés comme :
- L'hydrologie, domaine dans lequel la prévision des crues et la connaissance de leurs débits est particulièrement important pour la conception des aménagements des cours d'eau et

la protection des zones urbaines inondables, souvent théâtre d'épisodes de crues violentes causant d'importants dommages économiques.
- La météorologie où il faut prévoir les grandes tempêtes, ouragans ou encore les inondations.
- La finance où la gestion de portefeuille et l'évaluation du risque sur les marchés (prédiction de crises monétaires) sont devenus des enjeux majeurs.
- L'industrie où l'on s'intéresse à la fiabilité des structures.

Cette théorie a d'abord été développée dans le contexte d'observations indépendantes ; les travaux précurseurs de Fréchet [45], Fisher et Tippett [44], montrent que sous certaines conditions, les seules distributions limites des extrêmes sont les lois de Fréchet, Gumbel et Weibull. Ceci nous permet de classer la plupart des lois en trois domaines d'attraction où chaque domaine est déterminé par des caractérisations sur les fonctions de répartition (voir Embrechts et al. [39], De Haan et Ferreira [26]). Une paramétrisation des trois comportements limites en une distribution unique, à savoir la GEV (generalized extreme value distribution) est due à Von Mises [85]. Un des paramètres importants est l'indice des valeurs extrêmes (indice de queue) qui décrit la lourdeur de la queue de la distribution. Le problème central de cette thèse est l'étude du comportement asymptotique de l'un des estimateurs de cet indice. Divers travaux ont été consacrés à l'estimation de l'indice des extrêmes dont l'objectif revient à construire des estimateurs et étudier leurs propriétés ; citons Smith [83], Pickands [71], Hill [56] ou encore Dekkers et al. [32]. La plupart des estimateurs reposent sur l'utilisation de la statistique d'ordre. En effet, seules les données extrêmes sont utilisées, ce qui assure un meilleur ajustement à la queue de la distribution. Dans la pratique, l'hypothèse d'indépendance n'est pas toujours vérifiée comme c'est le cas de chaînes de fabrication, des durées de vie en médecine et des temps d'attente. C'est ainsi que l'indépendance des variables a été remplacée par des conditions de dépendance. Différents concepts de dépendance sont apparus dans la littérature comme le mélange, l'association et la m-dépendance. Les premiers travaux sur le comportement asymptotique du maximum de variables aléatoires sous dépendance sont dus à Loynes [66] (pour la m-dépendance) et Leadbetter [60] (pour des processus stationnaires sous la condition D_n). Les résultats obtenus sont similaires à ceux montrés dans le cas $i.i.d.$

Pour mesurer les termes de covariance qui apparaissent de façon naturelle lorsqu'on travaille avec des processus dépendants, une nouvelle notion de dépendance faible a été introduite par Doukhan et Louhichi [36]. Cette notion est plus générale que le cadre classique de mélange et est suffisamment large pour inclure plus de modèles standards tels que les processus linéaires, bilinéaires et plus généralement les schémas de Bernoulli. De plus, elle rend explicite l'indépendance asymptotique entre des fonctions du futur et du passé d'un processus aléatoire.
Le but principal de cette thèse est d'étudier la normalité asymptotique de l'estimateur de Hill sous dépendance faible au sens de Doukhan.
Nous nous sommes focalisés sur le théorème de Resnick et Stărică [78] qui démontrent, en adoptant l'approche de Rootzén et al. [80], la normalité asymptotique de l'estimateur de Hill en supposant que les variables sont fortement mélangeantes. Dans ce contexte, nous généralisons tout d'abord les résultats démontrés par Rootzén et al. [80] dans le cas de mélange fort sur le comportement d'une version de l'estimateur de Hill pour des distributions décroissantes asymptotiquement de manière exponentielle. Par la suite, nous obtenons une extension du théorème de Resnick et Stărică [78] pour un processus linéaire faiblement dépendant au sens de Doukhan.

Cette thèse est composée de quatre chapitres.
Chapitre 1 : La première partie de ce chapitre est consacrée aux concepts fondamentaux de la théorie des valeurs extrêmes dans le cas univarié et essentiellement dans le cas *i.i.d.* On expose les résultats asymptotiques de base, notamment le théorème de convergence de Fisher-Tippett qui met en évidence trois lois limites possibles. Ceci conduira à donner des critères d'appartenance aux différents domaines d'attraction. La seconde partie est consacrée au comportement du maximum de variables aléatoires sous dépendance où on présentera les résulats obtenus pour des données m-dépendantes et aussi sous la condition D de Lindeberg .
Chapitre 2 : Dans cette partie, nous abordons l'aspect statistique. Nous avons choisi de présenter de façon synthétique les estimateurs semi-paramétriques de l'indice de queue : estimateur de Hill, estimateur de Pickands, estimateur des moments et d'autres variantes

de l'estimateur de Hill. Nous faisons une analyse des principales propriétés de consistance et de convergence asymptotique de ces estimateurs.
Chapitre 3 : Un des récents développements de la théorie des valeurs extrêmes consiste à lever l'hypothèse d'indépendance. Ce chapitre est complètement consacré aux propriétés asymptotiques de l'estimateur de Hill sous m-dépendance et mélange fort.
Chapitre 4 : Le dernier chapitre est consacré à la présentation de nos résultats concernant les propriétés asymptotiques de l'estimateur de Hill dans le cadre de dépendance faible au sens de Doukhan et Louhichi. Nous présenterons la définition, les propriétés de la dépendance faible ainsi que quelques processus vérifiant cette dépendance. Par la suite, nous énonçons nos résultats.

Nous terminons par une conclusion générale qui dresse une synthèse des principaux résultats obtenus dans le dernier chapitre et nous donnons quelques perspectives de recherche.

Chapitre 1

Théorie des valeurs extrêmes

1.1 Introduction

Dans ce chapitre, on s'intéresse à la théorie des valeurs extrêmes (TVE) dont le problème de base est de modéliser et prévoir l'occurrence d'événements extrêmes en s'intéressant principalement non pas au "corps" de la distribution mais à sa queue.

Cette théorie à la fois ancienne et moderne a donné lieu à de fructueux développements mathématiques ; l'ouvrage synthèse de Gumbel [52], résumé de ces travaux de 1954 et 1958, est souvent identifié à la théorie statistique des extrêmes. Il faut aussi citer les travaux de Fréchet [45], Fisher et Tippett [44] et De Finetti [25].

Les avancées théoriques sont de plus en plus nombreuses et touchent à différents domaines ; nous pouvons citer plusieurs monographies synthétisant ces enrichissements. Le livre de Resnick [76] présente une lecture complète et efficace des techniques d'étude des extrêmes. Embrechts, Klüppelberg, Mikosch [39] et Reiss, Thomas [75] proposent une analyse globale de la théorie des valeurs extrêmes pour des applications en assurance et en finance. Finkenstadt et Rootzén [43] développent les extrêmes en vu des applications en télécommunication et environnement. Beirlant et al. [7] traitent des problèmes statistiques liées aux extrêmes. Les fonctions à variation régulière ont été utilisées par De Haan et Ferreira [26] comme un outil analytique pour la théorie des extrêmes ; son travail a été d'une grande importance pour le développement de la théorie moderne des extrêmes. Enfin, pour une présentation assez complète nous renvoyons aussi aux ouvrages de David [21] et de Coles [17].

1.2 Lois des valeurs extrêmes sous indépendance

1.2.1 Distributions limites

On considère une suite $(X_n)_n$ de variables aléatoires indépendantes identiquement distribuées ($i.i.d$), de distribution F et posons

$$M_n = \max_{1 \leq i \leq n} X_i$$

Le but de la théorie des valeurs extrêmes est de déterminer la loi normalisée que suit le maximum (ou le minimum) en fonction de celle de la variable aléatoire X_1. La fonction de répartition de M_n est définie par :

$$P(M_n \leq x) = P(X_1 \leq x, ..., X_n \leq x) = F^n(x) \tag{1.1}$$

Il est généralement impossible de déterminer la distribution du maximum à partir de (1.1), vu qu'en pratique la fonction de répartition F est rarement connue avec précision. De plus, même si la loi de X_1 est connue, la distribution de M_n est quelquefois difficile à calculer ; c'est le cas d'une loi normale où la fonction de distribution n'a pas d'expression analytique.

Définition 1.1. *On appelle point terminal de la distribution F, noté x_F, la borne supérieure du support de F définie par :*

$$x_F = \sup\{x \in \mathbb{R}, F(x) < 1\} \leq \infty$$

Définition 1.2. *Soit*

$$\Delta(x) = \begin{cases} 0 & si\ x < 0 \\ 1 & si\ x \geq 0 \end{cases}$$

On dit qu'une fonction de répartition est dégénérée si elle est de la forme $\Delta(x - x_0)$.

Le lemme suivant montre que M_n tend vers le point terminal de la distribution F qui peut être fini ou infini.

Lemme 1.1. *(Embrechts et al. [39])*

$$M_n \xrightarrow{p.s} x_F$$

Comme

$$\lim_{n\to\infty} P(M_n \leq x) = \lim_{n\to\infty} F^n(x) = \begin{cases} 0 & \text{si} \quad x < x_F \\ 1 & \text{si} \quad x \geq x_F \end{cases}$$

On constate que la distribution asymptotique de la suite des maximums $(M_n)_n$ est une loi dégénérée. Ainsi, la recherche d'une loi limite non dégénérée revient à trouver des constantes a_n et b_n dites constantes de normalisation de telle sorte que la distribution limite de $a_n M_n + b_n$, $a_n > 0$ existe.

Posons

$$H(x) = \lim_{n\to\infty} F^n(a_n x + b_n) = \lim_{n\to\infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) \tag{1.2}$$

. Le principal résultat de la théorie des valeurs extrêmes repose sur le théorème de Fisher et Tippett [44] qui donne la loi de H dont la première démonstration rigoureuse est due à Gnedenko [48].

Théorème 1.1. *(Fisher-Tippett [44], Gnedenko [48])*
Soit $(X_n)_n$ une suite de variables aléatoires i.i.d. S'il existe des constantes de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$ et une distribution H non dégénérée telles que :

$$a_n^{-1}(M_n - b_n) \xrightarrow{d} H, \tag{1.3}$$

alors H appartient à l'une des familles de distributions suivantes : de Gumbel, Fréchet ou Weibull,
avec

Loi de Gumbel (type I) $\quad \Lambda(x) = \exp(-\exp(-x)) \quad \forall x \in \mathbb{R}$

Loi de Fréchet (type II) $\quad \Phi_\beta(x) = \begin{cases} 0 & si \;\; x \leq 0 \\ \exp(-(x)^{-\beta}) & si \;\; x > 0 \end{cases}, \quad \beta > 0$

Loi de Weibull (type III) $\quad \Psi_\beta(x) = \begin{cases} \exp(-(-x)^{-\beta}) & si \;\; x \leq 0 \\ 1 & si \;\; x > 0 \end{cases}, \quad \beta < 0$

Le passage de l'une des trois lois à l'autre est relativement facile. Si par exemple, la variable aléatoire X suit une distribution de Weibull de paramètre $\beta = -1$ alors, $Y = log(-X)$ suit une distribution de Gumbel, tandis que $Z = -X^{-1}$ suit une distribution de Fréchet de paramètre $\beta = 1$.

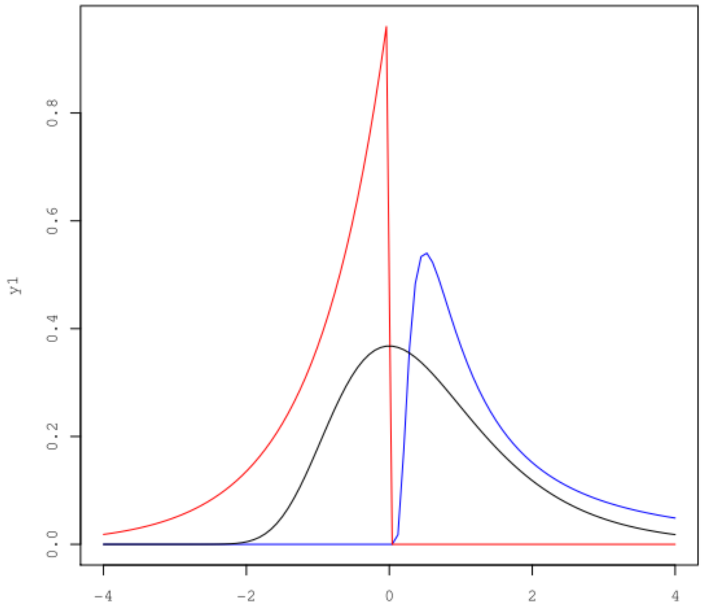

FIGURE 1.1 – Exemple de densités associées à la distribution des valeurs extrêmes avec $\beta = -1$ (rouge), $\beta = 0$ (noir) et $\beta = 1$ (bleu).

Remarque 1.1.
Soit $(X_n)_n$ une suite de variables aléatoires $i.i.d$, de même loi que X.
1. Si X suit la loi de Weibull de paramètre β, alors M_n a la même loi que $n^{\frac{1}{\beta}}X$.
2. Si X suit la loi de Gumbel, alors M_n a la même loi que $(X + \log n)$.
3. Si X suit la loi de Fréchet de paramètre β, alors M_n a la même loi que $n^{\frac{1}{\beta}}X$.

1.2.2 Distribution des valeurs extrêmes généralisée

Grâce aux travaux de Von Mises [85] et Jenkinson [59], il est possible de rassembler les trois familles de lois Weibull, Gumbel et Fréchet en une seule famille paramétrique $(H_\xi, \xi \in \mathbb{R})$ dite distribution généralisée des valeurs extrêmes (*Generalized Extreme Value* ou GEV).

Définition 1.3. *La fonction de répartition de H_ξ est :*

$$H_\xi(x) = \begin{cases} \exp\left(-(1+\xi x)^{-\frac{1}{\xi}}\right) & si \;\; 1+\xi x > 0 \\ \exp(-\exp(-x)) & si \;\; \xi = 0 \end{cases}$$

Le paramètre ξ est appelé l'indice de queue.

Plus l'indice ξ est élevé en valeur absolue, plus le poids des extrêmes dans la distribution initiale est important.

ξ est l'inverse au signe près du paramètre β qui apparaît dans les lois de Fréchet et de Weibull ; le cas intermédiaire $\xi = 0$ correspond à la loi de Gumbel qui peut être considérée comme une loi de transition entre les deux autres lois.

Remarque 1.2.

1. La loi de Weibull peut s'écrire $\Psi_\beta(x) = H_{\frac{1}{\beta}}(-\beta(x+1))$; si X suit une loi de Weibull de paramètre $\beta < 0$, alors $-\beta(X+1)$ suit la loi des valeurs extrêmes généralisée de paramètre $\xi = \frac{1}{\beta}$.
2. La loi de Gumbel correspond à H_0, i.e., la loi GEV de paramètre $\xi = 0$.
3. La loi de Fréchet peut s'écrire $\Phi_\beta(x) = H_{\frac{1}{\beta}}(\beta(x-1))$; si X suit une loi de Fréchet de paramètre $\beta > 0$, alors $\beta(X-1)$ suit la loi des valeurs extrêmes généralisée de paramètre $\xi = \frac{1}{\beta}$.

Si la loi des variables X_i est connue, Gnedenko [48] a donné des conditions nécessaires et suffisantes pour l'existence des constantes de normalisation qui peuvent être utilisées pour déterminer le type de la loi limite.

1.2.3 Exemples de comportements limites

a. Loi exponentielle : Considérons la loi exponentielle de paramètre $\lambda = 1$, de fonction de répartition $F(x) = 1 - \exp(-x), \forall x \in \mathbb{R}^+$.

En posant $b_n = \log n$ et $a_n = 1$, la fonction de répartition du maximum (1.2) s'écrit,

$$\begin{aligned} F^n(a_n x + b_n) &= (1 - \exp(-x - \log n))^n \\ &= \left(1 - \frac{\exp(-x)}{n}\right)^n \underset{n\to\infty}{\to} \exp(-\exp(-x)) = \Lambda(x) \end{aligned}$$

Le maximum normalisé de la loi exponentielle converge vers la loi de Gumbel.

b. Loi de Pareto : Considérons la loi de Pareto de fonction de répartition :

$$F(x) = 1 - cx^{-\beta} \; avec \; c > 0 \; et \; \beta > 0$$

On pose $b_n = 0$ et $a_n = (nc)^{\frac{1}{\beta}}$.
Pour $x > 0$, on écrit

$$F^n(a_n x + b_n) = \left(1 - \frac{x^{-\beta}}{n}\right)^n \underset{n\to\infty}{\to} \exp(-x^{-\beta}) = \Phi_\beta(x)$$

La loi limite est la loi de Fréchet.

c. Loi uniforme (loi à support borné à droite) : La distribution de la loi uniforme sur $[0, 1]$ est : $F(x) = x$ si $0 \leq x \leq 1$.
Pour $x < 0$ et $n > -x$, posons $a_n = \dfrac{1}{n}$ et $b_n = 1$, alors

$$F^n(a_n x + b_n) = F^n\left(\frac{x}{n} + 1\right) = \left(1 + \frac{x}{n}\right)^n \underset{n\to\infty}{\to} \exp(x) = \Psi_{-1}(x)$$

La dernière est la loi de Weibull avec $\beta = -1$.

d. Loi normale : La fonction de répartition de la loi normale centrée et réduite $N(0, 1)$ est : $F(x) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\dfrac{y^2}{2}} dy$, $x \in \mathbb{R}$.

$$\begin{aligned} F(-x) &= \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy, \quad x > 0 \\ &\leq \frac{1}{\sqrt{2\pi}} \int_x^\infty \frac{y}{x} e^{-\frac{y^2}{2}} dy = \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}} \end{aligned}$$

D'après l'inégalité de Gordon [49] on aura

$$\frac{x}{1+x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \leq F(-x) \leq \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x > 0$$

$$1 - F(x) \sim \frac{1}{x\sqrt{2\pi}} \exp(-\frac{x^2}{2}), \quad \text{pour } \ x \longrightarrow +\infty$$

On en déduit que,

$$\frac{1 - F(u + \frac{z}{u})}{1 - F(u)} \sim \left[\left(1 + \frac{z}{u^2}\right)^{-1} \exp\left(-\frac{1}{2}(u + \frac{z}{u})^2 + \frac{1}{2}u^2\right) \right] \sim \exp(-z), \quad \text{pour } \ u \longrightarrow +\infty$$

Soit b_n la solution de $F(b_n) = 1 - \frac{1}{n}$ et posons $a_n = \frac{1}{b_n}$, alors

$$n\left(1 - F(a_n x + b_n)\right) = \frac{1 - F(a_n x + b_n)}{1 - F(b_n)} \rightarrow \exp(-x)$$

d'où

$$\lim_{n\to\infty} F^n(a_n x + b_n) = \lim_{n\to\infty} \left(1 - \frac{\exp(-x)}{n}\right)^n = \exp(-\exp(-x)) = \Lambda(x)$$

La limite est la loi de Gumbel qui correspond à celle obtenue dans le cas de la loi exponentielle.

1.2.4 Convergence de M_n

Dans la partie précédente, on a considéré les probabilités de la forme $P(\frac{M_n - b_n}{a_n} \leq x)$ qui peuvent être réécrites sous la forme $P(M_n \leq u_n)$ avec $u_n = a_n x + b_n$; ces probabilités sont calculées avec la GEV. D'autre part, il est également intéressant d'examiner le cas général où $(u_n)_n$ ne dépend pas nécessairement de x.

Théorème 1.2. *(Embrechts et al. [39])*
Soit $(X_n)_n$ une suite de variables aléatoires i.i.d de distribution F. Soit $0 \leq \tau \leq \infty$ et $(u_n)_n$ une suite de nombres réels telle que :

$$n\left(1 - F(u_n)\right) \rightarrow \tau \qquad \textit{quand } n \rightarrow \infty \tag{1.4}$$

Alors

$$P(M_n \leq u_n) \rightarrow \exp(-\tau) \qquad \textit{quand } n \rightarrow \infty \tag{1.5}$$

Inversement, si (1.5) *est vérifiée pour un certain $\tau, 0 \leq \tau \leq \infty$, alors* (1.4) *est vérifiée.*

Dans le cas de variables $i.i.d$, pour une suite arbitraire $(u_n)_n$, on a :

$$P(M_n \leq u_n) = F^n(u_n) = \exp(-n\,(1 - F(u_n))) + o(1)$$

Remarquons que (1.2) est un cas particulier du théorème précédent, en faisant les identifications suivantes : $\tau = -\log H(x),\ \ u_n = a_n x + b_n$.

$$\text{Si } \lim_{n\to\infty} n(1 - F(a_n x + b_n)) = -\log H(x)$$

On obtient

$$P(M_n \leq a_n x + b_n) \to H(x), \quad \text{quand } n \to \infty.$$

1.3 Caractérisation des domaines d'attraction

Nous rappelons les conditions nécessaires et suffisantes sur la fonction de répartition F pour qu'elle appartienne à l'un des domaines d'attraction de l'une des trois lois limites des valeurs extrêmes.

1.3.1 Définitions

Définition 1.4. *On dit qu'une variable aléatoire X (ou de sa fonction de répartition F) appartient au domaine d'attraction de la distribution des extrêmes généralisée (ou sa fonction de répartition H) s'il existe des constantes $a_n > 0$ et $b_n \in \mathbb{R}$ telles que* (1.3) *est vérifiée. On écrit $X \in D(H)$ (ou $F \in D(H)$).*

Avant de caractériser ces domaines d'attraction, on donne quelques définitions et outils nécessaires notamment la notion de fonctions à variation régulière.

Définition 1.5. *(Fonction quantile)*
On appelle fonction quantile ou inverse généralisée de la distribution F, la fonction Q définie par :

$$Q(p) = F^{\leftarrow}(p) = \inf\{x \in \mathbb{R}, F(x) \geq p\},\ \forall p \in]0,1[$$

Définition 1.6. *(Fonction quantile de queue)*
On appelle fonction quantile de queue de la distribution F, la fonction b :$]1,+\infty[\longrightarrow \mathbb{R}$ définie par :

$$b(t) = Q(1-\frac{1}{t}) = F^{\leftarrow}(1-\frac{1}{t}),\ \forall t > 1$$

Définition 1.7. *Une fonction L est dite à variation lente à l'infini si :*

$$\lim_{t\to+\infty} \frac{L(tx)}{L(t)} = 1,\ \ \forall x > 0$$

Proposition 1.1. *Pour toute fonction à variation lente L, on a :*

$$\lim_{t\to+\infty} \frac{\log(L(t))}{\log(t)} = 0 \tag{1.6}$$

Définition 1.8. *Une fonction L est dite à variation régulière (à l'infini) d'indice $\beta \in \mathbb{R}$ (on note $L \in RV_\beta$), si L est positive pour t assez grand et si :*

$$\lim_{t\to+\infty} \frac{L(tx)}{L(t)} = x^\beta,\ \ x > 0$$

On remarque que si L est à variation régulière d'indice β, alors $\dfrac{L(x)}{x^\beta}$ est à variation lente. Une fonction à variation régulière d'indice β peut toujours s'écrire sous la forme $x^\beta L(x)$, où L est à variation lente.

Proposition 1.2. *(Représentation de Karamata)(Resnick [76])*
Si L est une fonction à variation régulière d'ordre $\beta \in \mathbb{R}$, alors

$$L(x) = c(x).\exp\left\{\int_{x_0}^{x} \frac{\varepsilon(u)}{u} du\right\},\qquad x \geq x_0 > 0 \tag{1.7}$$

où c et ε sont des fonctions mesurables telles que : $\lim_{x\to\infty} c(x) = c_0 \in]0,+\infty[$ et $\lim_{x\to\infty} \varepsilon(x) = \beta$.

Définition 1.9. *(Fonction à variation rapide)*
Une fonction mesurable h définie sur $]0,+\infty[$ est dite à variation rapide d'indice $-\infty$, si

$$\lim_{t\to\infty} \frac{h(tx)}{h(x)} = \begin{cases} 0 & si\quad x > 1 \\ \infty & si\quad 0 < x < 1 \end{cases}$$

On écrit $h \in R_{-\infty}$.

1.3.2 Domaine d'attraction de H_ξ

Les résultats suivants caractérisent le domaine d'attraction de la GEV.

Proposition 1.3. *(Embrechts et al. [39])*
$F \in D(H_\xi)$ ssi pour une certaine suite $(a_n, b_n)_{n\geq 1}$ où $a_n > 0$ et $b_n \in \mathbb{R}$, on a

$$\lim_{n\to\infty} n\overline{F}(a_n x + b_n) = -\log H_\xi(x), \quad x \in \mathbb{R}$$

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

Le théorème suivant complète cette dernière proposition et est fondamental pour l'étude du $D(H_\xi)$.

Théorème 1.3. *(Embrechts et al. [39])*
Pour $\xi \in \mathbb{R}$, les affirmations suivantes sont équivalentes :

1. *$F \in D(H_\xi)$*
2. *Il existe une fonction mesurable a, telle que pour $1 + \xi x > 0$ on a*

$$\lim_{u\to x_F} \frac{\overline{F}(a(u)x + u)}{\overline{F}(u)} = \begin{cases} (1+\xi x)^{-\frac{1}{\xi}} & si \ \ \xi \neq 0 \\ \exp(-x) & si \ \ \xi = 0 \end{cases}$$

3. *Pour $x, y > 0$ et $y \neq 1$,*

$$\lim_{s\to\infty} \frac{b(sx) - b(s)}{b(sy) - b(s)} = \begin{cases} \dfrac{x^\xi - 1}{y^\xi - 1} & si \ \ \xi \neq 0 \\ \dfrac{\log x}{\log y} & si \ \ \xi = 0 \end{cases} \tag{1.8}$$

1.3.3 Domaine d'attraction de Fréchet

L'une des caractéristiques de la distribution de Fréchet est que sa queue est lourde à droite ; ce qui signifie que les événements extrêmes d'une distribution appartenant au domaine d'attraction de Fréchet ont une importante probabilité d'apparition.
Notons que par le développement de Taylor de Φ_β, quand $x \to \infty$:

$$1 - \Phi_\beta(x) = 1 - \exp(-x^{-\beta}) \sim x^{-\beta}$$

La queue de Φ_β décroît comme une fonction puissance.

Dans cette partie, nous allons caractériser le domaine d'attraction de la distribution de Fréchet en fonction des distributions à variation régulière.

Théorème 1.4. *(Embrechts et al. [39])*
La fonction de répartition F appartient au domaine d'attraction de la loi de Fréchet de paramètre $\beta > 0$ ssi $x_F = +\infty$ et

$$\overline{F}(x) = x^{-\beta} L(x) \tag{1.9}$$

où L est une fonction à variation lente. De plus, les constantes de normalisation peuvent être choisies de sorte que $a_n = F^{\leftarrow}(1 - \dfrac{1}{n})$ et $b_n = 0$.

Corollaire 1.1. *Pour $\beta > 0$, la fonction de répartition F appartient au domaine d'attraction de la loi de Fréchet ssi*

$$\lim_{t\to\infty} \frac{b(tx)}{b(t)} = x^{\frac{1}{\beta}} \quad \textit{pour} \;\; x > 0$$

Il suffit donc que la fonction quantile de queue soit à variation régulière d'ordre $\frac{1}{\beta}$, pour que F appartienne au domaine d'attraction de Fréchet.

Dans le cas où la variable aléatoire admet une densité, on caractérise ce domaine en utilisant le critère de Von Mises.

Proposition 1.4. *(Critère de Von Mises) [De Haan et Ferreira [26]]*
Soit F une fonction de répartition avec $x_F = \infty$. Si

$$\lim_{x\to\infty} \frac{x f(x)}{\overline{F}(x)} = \beta > 0, \quad f = dF \tag{1.10}$$

alors F appartient au domaine d'attraction de la loi de Fréchet de paramètre β.

La condition (1.10) peut être écrite en utilisant la fonction quantile de queue.
Soit $t = \dfrac{1}{1 - F(b(t))}$, alors

$$b'(t) = \frac{[\overline{F}(b(t))]^2}{f(b(t))},$$

par conséquent

$$\begin{aligned} t\frac{b'(t)}{b(t)} &= \frac{t\overline{F}(b(t)).\overline{F}(b(t))}{b(t)f(b(t))} \\ &= \frac{1}{\dfrac{b(t)f(b(t))}{\overline{F}(b(t))}} \end{aligned}$$

En utilisant (1.10), on déduit que

$$\lim_{x\to\infty} \frac{xb'(x)}{b(x)} = \frac{1}{\beta}$$

Exemples : Parmi les lois de probabilité appartenant $D(\Phi_\beta)$, on trouve les lois de Cauchy, Pareto et loggamma.

Remarque 1.3. Toutes les distributions appartenant au domaine d'attraction de Fréchet sont dites de type Pareto et leur distribution F s'écrit, pour x très grand,

$$\overline{F}(x) \sim Cx^{-\beta}, \;\; C, \beta > 0 \tag{1.11}$$

Les constantes de normalisation peuvent être choisies de la façon suivante : $a_n = (Cn)^{-\frac{1}{\beta}}$ et $b_n = 0$.

1.3.4 Domaine d'attraction de Weibull

Les distributions appartenant au domaine d'attraction de Weibull sont caractérisées par un point terminal fini ; ce qui les rend peu pratiques dans la modélisation de beaucoup de phénomènes où, en principe il n'y a pas de borne supérieure.

Théorème 1.5. *(Embrechts et al. [39])*
La fonction de répartition F appartient au domaine d'attraction de la loi de Weibull de paramètre $\beta < 0$ ssi $x_F < +\infty$ et

$$\overline{F}(x_F - \frac{1}{x}) = x^{-\beta} L(x)$$

où L est une fonction à variation lente.
Les constantes de normalisation sont données par :
$a_n = x_F - F^{\leftarrow}(1 - \frac{1}{n})$ *et* $b_n = x_F$.

Proposition 1.5. *(Critère de Von Mises)[De Haan et Ferreira [26]]*
Notons par F et f respectivement les fonctions de répartition et de densité d'une variable aléatoire X. On suppose que la fonction de densité f est strictement positive sur un intervalle (z, x_F), avec $x_F < \infty$ et que

$$\lim_{x\to x_F^-} \frac{(x_F - x) f(x)}{\overline{F}(x)} = -\beta$$

alors F appartient au domaine d'attraction de la loi de Weibull de paramètre $\beta < 0$.

Exemples : Dans le domaine d'attraction de la loi de Weibull, nous trouvons les distributions bornées à droite (Uniforme, Bêta, etc.).

1.3.5 Domaine d'attraction de Gumbel

La queue de la loi de Gumbel est une fonction à variation rapide à l'infini.
En effet

$$1 - \Lambda(x) = 1 - \exp(-\exp(-x)) \sim \exp(-x) \;\; \textit{pour} \;\; x \to \infty$$

Le domaine d'attraction de Gumbel contient une grande variété de distributions. Parmi elles, on trouve des distributions les plus utilisées, notamment l'exponentielle et la gaussienne.
La caractérisation du domaine d'attraction de Gumbel nécessite la définition d'un nouveau type de fonction.

Définition 1.10. *(Fonction de Von Mises)*
Soit F une fonction de répartition avec un point terminal $x_F \leq \infty$.
Si $\exists x_0 < x_F$, tel que $\overline{F}$ peut s'écrire :

$$\overline{F}(x) = c \exp\left\{ -\int_{x_0}^{x} \frac{1}{a(t)} dt \right\}, \quad x_0 < x < x_F, \; c > 0 \tag{1.12}$$

où a est une fonction positive absolument continue de derivée a' telle que $\lim_{x \to x_F} a'(x) = 0$, alors F est une fonction de Von Mises. a est dite fonction auxiliaire de F.

La relation (1.12) peut être comparée à la représentation de Karamata pour les fonctions à variation régulière.
Soit F une fonction de Von Mises avec une fonction auxiliaire a.
En remplaçant $\varepsilon(u)$ par $\dfrac{u}{a(u)}$ dans la relation (1.7), on remarque que $\lim_{x \to x_F = \infty} \dfrac{u}{a(u)} = \infty$.
Ceci montre que les fonctions de Von Mises ne sont pas à variation régulière mais à variation rapide à l'infini.

Proposition 1.6. *(Embrechts et al. [39])*
Toute fonction de Von Mises F de densité f est absolument continue sur (z, x_F) et la fonction auxiliaire est telle que $a(x) = \dfrac{\overline{F}(x)}{f(x)}$. De plus

1. *Si $x_F = \infty$, alors $\overline{F} \in R_{-\infty}$ et*

$$\lim_{x\to\infty} \frac{xf(x)}{\overline{F}(x)} = \infty$$

2. *Si $x_F < \infty$, alors $\overline{F}(x_F - \frac{1}{x}) \in R_{-\infty}$ et*

$$\lim_{x\to x_F} \frac{(x_F - x)f(x)}{\overline{F}(x)} = \infty$$

Proposition 1.7. *(Resnick [76])*
Si F est une fonction de Von Mises, alors $F \in D(\Lambda)$ et les constantes de normalisation sont données par : $b_n = F^{\leftarrow}(1 - \dfrac{1}{n})$, $a_n = a(b_n)$.

La réciproque de la proposition précédente n'est pas vraie, i.e., que les fonctions de Von Mises ne caractérisent pas totalement $D(\Lambda)$; une légère modification de (1.12) permet de le caractériser complètement.

Proposition 1.8. *La distribution F appartient à $D(\Lambda)$ ssi $\exists x_0 < x_F \leq \infty$, tel que :*

$$\overline{F}(x) = c(x) \exp\left\{ -\int_{x_0}^{x} \frac{\varepsilon(t)}{a(t)} dt \right\}, \quad x_0 < x < x_F \tag{1.13}$$

où c et ε sont des fonctions mesurables vérifiant : $\lim_{x\to x_F} c(x) = c > 0$, $\lim_{x\to x_F} \varepsilon(x) = 1$ et a est la fonction auxiliaire de F.

Un choix possible pour la fonction a est

$$a(x) = \int_{x}^{x_F} \frac{\overline{F}(t)}{\overline{F}(x)} dt, \quad x < x_F \tag{1.14}$$

On déduit de la proposition 1.8 qu'une distribution vérifiant la représentation (1.13) appartient à $D(\Lambda)$ et $b_n = F^{\leftarrow}(1 - \dfrac{1}{n})$, $a_n = a(b_n)$.

Une deuxième caractérisation du domaine d'attraction $D(\Lambda)$ est donnée par le théorème suivant.

Théorème 1.6. *Une distribution F avec un point terminal $x_F \leq \infty$ appartient à $D(\Lambda)$ ssi il existe une fonction positive $\tilde{a}$ telle que :*

$$\lim_{x \to x_F} \frac{\overline{F}(x + t\tilde{a}(x))}{\overline{F}(x)} = \exp(-t), \quad t \in \mathbb{R}$$

Un choix possible pour la fonction $\tilde{a}$ est donné par (1.14).

1.4 Lois des valeurs extrêmes des suites stationnaires

L'hypothèse d'indépendance des variables aléatoires n'est malheureusement pas souvent vérifiée ; ce qui a amené certains auteurs à travailler avec des données dépendantes. Un des récents développements de la théorie des valeurs extrêmes consiste à lever cette hypothèse et étudier les processus stationnaires.

Plusieurs généralisations du théorème 1.1 ont été proposées ; Watson [86] et Loynes [66] montrent respectivement que le résultat du théorème demeure valide si les variables sont m-dépendantes ou fortement mélangeantes. Sous des hypothèses relatives aux autocorrélations des observations, Berman [10] montre que le résultat reste valide si la série des coefficients de corrélation au carré est convergente. Leadbetter [60] introduit la condition D, plus faible que les autres notions de dépendance et établit le résultat du théorème 1.1. De Hann et al. [29] analyse le comportement des extrêmes provenant d'un processus ARCH(1) et montrent que la distribution limite est de type II. Sous certaines conditions sur l'autocovariance des observations, Leadbetter et al. [61] obtiennent la distribution des maximums de variables gaussiennes stationnaires converge vers une loi limite du type I. Ces extensions montrent que l'hypothèse d'indépendance et d'identité de la loi parente ne sont pas cruciales.

Dans cette partie, on s'intéresse aux lois limites des valeurs extrêmes pour des suites dépendantes strictement stationnaires ; la stationnarité caractérise un processus pour lequel les propriétés stochastiques sont homogènes à travers le temps.

Définition 1.11. *(Stationnarité stricte)*
On dit qu'un processus $(X_t)_{t\in\mathbb{Z}}$ est strictement stationnaire si la loi de $(X_{t_1}, X_{t_2}, ..., X_{t_m})$ est la même que la loi de $(X_{t_1+h}, ..., X_{t_m+h})$, pour tout $t_1 < t_2 < ... < t_m$ et h entier.

Définition 1.12. *(Stationnarité faible)*
On dit qu'un processus $(X_t)_{t\in\mathbb{Z}}$ est faiblement stationnaire s'il vérifie les trois propriétés suivantes :

- $\forall t \in \mathbb{Z}, \quad E(X_t) = m$ *l'espérance mathématique est constante,*
- $\forall t \in \mathbb{Z}, \quad V(X_t) = \sigma^2$ *la variance est constante,*
- $\forall (t, s) \in \mathbb{Z}^2, \quad cov(X_t, X_s) = \varphi(|t - s|)$, *où* $\varphi : \mathbb{Z} \to \mathbb{R}$.

Dans la suite de cette thèse, la stationnarité sera prise au sens strict.

1.4.1 Lois des extrêmes sous m-dépendance

On peut mesurer la dépendance entre le passé des variables aléatoires $(X_1, X_2, ..., X_r)$ et leur futur $(X_s, X_{s+1}, ..., X_n)$, suivant la vitesse avec laquelle leur séparation $s - r > m$ augmente.

Définition 1.13. *(m-dépendance)*
Une suite $(X_n)_n$ est dite m-dépendante si X_r et X_s sont indépendantes $\forall s, r$ tels que $|s - r| > m$.

Pour $i_1, i_2, ..., i_n$ on note :

$$F_{i_1,i_2,...,i_n}(x_1, ..., x_n) = P(X_{i_1} \leq x_1, ..., X_{i_n} \leq x_n)$$

La m-dépendance se traduit par :

$$F_{i_1,i_2,...,i_p,i_1,j_1,...,j_q}(x_1, ..., x_p, y_1, ..., y_q) = F_{i_1,i_2,...,i_p}(x_1, ..., x_p)F_{j_1,...,j_q}(y_1, ..., y_q) \qquad (1.15)$$

pour $i_1 < i_2 < ... < i_p < j_1 < ... < j_q$, avec $j_1 - i_p > m$. Sous la notion de m-dépendance, on a le résultat suivant :

Théorème 1.7. *(Watson [86])*
Soit $(X_t)_t$ une suite stationnaire de variables aléatoires, non bornée, m-dépendante ($m \geq 1$) de distribution marginale F. On suppose que

$$\lim_{c\to\infty} \max_{1\leq|s-r|\leq m} P\left(X_s > c \,|X_r > c\right) = 0$$

Alors, pour $\tau = \lim\limits_{n\to\infty} n(1-F(u_n))$, on a

$$\lim_{n\to\infty} P\left(M_n \leq u_n\right) = \exp(-\tau)$$

Ce résultat est une généralisation de (1.5) au cas de maximum de variables aléatoires stationnaires, m-dépendantes avec la condition restrictive $\tau = \lim\limits_{n\to\infty} n(1-F(u_n))$ qui peut être impossible à satisfaire.

Preuve succincte :
Pour démontrer que $\lim\limits_{n\to\infty} P(M_n \leq u_n) = \exp(-\tau)$, il suffit de majorer et de minorer le terme $P(M_n \leq u_n)$ par deux quantités T et L tendant vers $\exp(-\tau)$.
Pour

$$\begin{aligned} T &= 1 - \sum_{i=1}^{n} P(X_i > u_n) + ... + (-1)^{l-1} \sum_{i_{l-1}=1}^{n} ... \sum_{i_1=1}^{n} P(X_{i_1} > u_n, ..., X_{i_{l-1}} > u_n) \\ \text{et} \quad L &= 1 - \sum_{i=1}^{n} P(X_i > u_n) + ... + (-1)^{l} \sum_{i_l=1}^{n} ... \sum_{i_1=1}^{n} P(X_{i_1} > u_n, ..., X_{i_l} > u_n) \end{aligned}$$

où $l \leq n$.
On a

$$T \leq P(M_n \leq u_n) \leq L$$

Le terme $\sum_{i=1}^{n} P(X_i > u_n) = n(1-F(u_n)) \to \tau$.
D'autre part

$$\sum_{j=1}^{n}\sum_{i=1}^{n} P(X_i > u_n, X_j > u_n) = \sum_{i=1}^{m}(n-i)P(X_1 > u_n, X_{i+1} > u_n) + P(X_i > u_n)^2 \frac{1}{2}(n-m-1)(n-m).$$

Quand $u_n \to \infty$, on a

$$\sum_{i=1}^{m}(n-i)P(X_1 > u_n, X_{i+1} > u_n) \leq mn\left[1 - \frac{(m+1)}{2n}\right] \max_{|i-j|\leq m} \frac{P(X_i > u_n, X_j > u_n)}{P(X_i > u_n)} \to 0$$

On en déduit que

$$\lim_{n\to\infty} \sum P(X_i > u_n, X_j > u_n) = \frac{1}{2}\tau^2.$$

En calculant $\sum_{s=1}^{n}\sum_{j=1}^{n}\sum_{i=1}^{n} P(X_i > u_n, X_j > u_n, X_s > u_n)$ de la même manière, on obtient

$$\lim_{n\to\infty} \sum P(X_i > u_n, X_j > u_n, X_s > u_n) = -\frac{1}{6}\tau^3$$

De manière générale, pour tout entier $l \leq n$

$$\sum_{r=0}^{l-1} \frac{(-\tau)^r}{r!} \leq P(M_n \leq u_n) \leq \sum_{r=0}^{l} \frac{(-\tau)^r}{r!}.$$

1.4.2 Lois des extrêmes sous mélange fort

Dans ce qui suit, nous introduisons la notion de α-mélange appelé aussi mélange fort.

Définition 1.14. *On définit le coefficient de α- mélange entre deux tribus $\mathcal{A}$ et $\mathcal{B}$ d'un même espace de probabilité par :*

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A\in\mathcal{A}, B\in\mathcal{B}} |P(A \cap B) - P(A)P(B)|$$

Si $(X_n)_n$ est une suite de variables aléatoires définies sur le même espace de probabilité, on définit la suite des coefficients de mélange fort α_n par :

$$\alpha_n = \sup_{k\in\mathbb{Z}} \alpha(\mathcal{F}_{-\infty}^{k}, \mathcal{F}_{k+n}^{\infty}) \tag{1.16}$$

où $\mathcal{F}_a^b = \sigma(X_i, a \leq i \leq b)$. On dit que la suite $(X_n)_n$ est α-mélangeante si $\alpha_n \to 0$ quand $n \to \infty$.

Le mélange fort est donné en termes de coefficients de dépendance entre les tribus engendrées par les variables de la suite avant un instant t et les tribus engendrées par les variables après l'instant $t + n$; ces coefficients sont nuls quand les tribus sont indépendantes. Cette notion a été introduite par Rosenblatt [82] et a donné naissance à une abondante bibliographie, par exemple : Hall et Heyde [55], Bradley et Peligrad [12], Doukhan et al.[37], Doukhan [35].

Remarque 1.4.

1. Dans le cas d'un processus stationnaire, on peut simplifier le coefficient (1.16) par

$$\alpha_n = \alpha(\mathcal{F}_{-\infty}^{0}, \mathcal{F}_n^{\infty})$$

2. $0 \leq \alpha(\mathcal{A}, \mathcal{B}) \leq \frac{1}{4}$

Le résultat suivant donne la loi des extrêmes sous mélange fort.

Théorème 1.8. *(Loynes [66])*
Soit $(X_n)_n$ une suite stationnaire de variables aléatoires, vérifiant la condition de mélange fort. S'il existe deux suites de nombres réels $a_n > 0$ et $b_n \in \mathbb{R}$ telles que :

$$\lim_{n\to\infty} P\left(M_n \leq a_n x + b_n\right) = H(x)$$

Alors H appartient à l'une des familles de Gumbel, Weibull ou Fréchet.

Pour la démonstration de théorème précédent, on a besoin des lemmes suivantes.

Lemme 1.2. *(Leadbetter [60])*
Soient M_n une variable aléatoire et $a_n > 0$, $b_n \in \mathbb{R}$ deux suites de constantes telles que

$$P(M_n \leq a_n x + b_n) = H(x), \quad quand \ n \to \infty.$$

où H est une distribution non dégénérée.
Si de plus pour tout $r \geq 1$, on a

$$P^r(M_m \leq a_{mr} x + b_{mr}) = H(x), \quad quand \ m \to \infty.$$

Alors il existe deux suites de constantes $c_r > 0$ et d_r telles que

$$H^r(c_r x + d_r) = H(x) \tag{1.17}$$

Lemme 1.3. *(Leadbetter [60])*
Si H est une distribution non dégénérée telle que l'égalité (1.17) est vérifiée.
Alors H appartient à l'une des familles de Gumbel, Weibull ou Fréchet.

Preuve de Théorème 1.8 :

Soit n un entier tel que $n = rm$, on partage l'ensemble $\{X_1, X_2, ..., X_n\}$ en r groupes de m variables aléatoires. Notons M_{mi} le maximum de $i$ème groupe.i.e.,

$$M_{mi} = \max_{1 \leq j \leq m} X_{(i-1)m+j}, \quad 1 \leq i \leq r$$

Alors

$$P\left(M_n \leq a_n x + b_n\right) = P\left(\max_{1 \leq i \leq r} M_{mi} \leq a_n x + b_n\right)$$

Soit k un entier et notons

$$M'_{mr} = \max_{k \leq j \leq m} X_{(r-1)m+j}$$

et

$$M''_{mr} = \max_{1 \leq j \leq k} X_{(r-1)m+j}$$

Remarquons que $M_{mr} = \max(M'_{mr}, M''_{mr})$.
En considérant l'événement $B_m = \{M'_{mm} < M''_{mm}\}$ et en utilisant la loi des grands nombres pour les processus stationnaires, on peut montrer que $\lim\limits_{m \to \infty} P(B_m) = 0$.
Alors

$$P(\max_{1 \leq i \leq r} M_{mi} \leq a_n x + b_n) = P[(\max_{1 \leq i \leq r} M_{mi} \leq a_n x + b_n) \cap B_m] + P[(\max_{1 \leq i \leq r} M_{mi} \leq a_n x + b_n) \cap B_m^c]$$

Le premier terme tend vers zéro quand $m \to \infty$ et pour le deuxième terme on a :

$$\begin{aligned} P[(\max_{1 \leq i \leq r} M_{mi} \leq a_n x + b_n) \cap B_m^c] &= P[(\max_{1 \leq i \leq r-1} M_{mi} \leq a_n x + b_n) \cap (M'_{mr} \leq a_n x + b_n) \cap B_m^c] \\ &= P[(\max_{1 \leq i \leq r-1} M_{mi} \leq a_n x + b_n) \cap (M'_{mr} \leq a_n x + b_n)] - \delta_m, \end{aligned}$$

avec $0 \leq \delta_m \leq P(B_m)$. De plus

$$P[(\max_i M_{mi} \leq a_n x + b_n) \cap (M'_{mr} \leq a_n x + b_n)] = P[\max_i M_{mi} \leq a_n x + b_n] + P[M'_{mr} \leq a_n x + b_n] + \gamma_m,$$

où $|\gamma_m| \leq \alpha_k$, alors

$$\lim_{m \to \infty} \left(P(M_n \leq a_n x + b_n) - P(\max_{1 \leq i \leq r-1} M_{mi} \leq a_n x + b_n) P(M_{mr} \leq a_n x + b_n) \right) = 0.$$

On en déduit que

$$\lim_{m \to \infty} \left(P(M_n \leq a_n x + b_n) - P^r(M_{m1} \leq a_n x + b_n)\right) = 0 \tag{1.18}$$

Par hypothèse, on a

$$\lim_{m\to\infty} P(M_n \leq a_n x + b_n) = H(x)$$

.

En vertu de (1.18), on obtient

$$\lim_{m\to\infty} P^r(M_{m1} \leq a_n x + b_n) = H(x)$$

En appliquant le lemme 1.2, l'égalité (1.17) est donc vérifiée et d'après le lemme 1.3, on déduit que H appartient à l'une des familles de Gumbel, Weibull ou Fréchet.

Un autre théorème de Loynes exige une condition supplémentaire qui doit être considérée comme une condition suffisante.

Théorème 1.9. *(Loynes [66])*
Soit $(X_n)_n$ une suite stationnaire α-mélangeante. Si pour une suite $(X_n^*)_n$ i.i.d de même distribution que X_1, il existe deux suites de nombres réels $a_n > 0$ et $b_n \in \mathbb{R}$ telles que :

$$\lim_{n\to\infty} P\left(M_n^* \leq a_n x + b_n\right) = H(x)$$

avec $M_n^* = \max\limits_{1\leq i\leq n} X_i^*$, alors

$$\lim_{n\to\infty} P\left(M_n \leq a_n x + b_n\right) = H(x)$$

1.4.3 Lois des extrêmes sous condition D

La condition D est plus faible que la m-dépendance et se définit de la manière suivante :

Définition 1.15. *(Condition D)*
Une suite de variables aléatoires $(X_n)_n$ vérifie la condition D si pour tous $i_1, ..., i_p, j_1, ..., j_q$ tels que $i_1 < ... < i_p < j_1 < ... < j_q$ et $j_1 - i_p \geq l$, on a :

$$\left|F_{i_1,i_2,...,i_p,i_1,j_1,...,j_q}(u,...,u) - F_{i_1,i_2,...,i_p}(u,...,u)F_{j_1,...,j_q}(u,...,u)\right| \leq \beta_l$$

avec $\beta_l \to 0$ quand $l \to \infty$.

$F_{i_1,i_2,...,i_n}$ désigne la distribution jointe de $X_{i_1}, X_{i_2}, ..., X_{i_n}$.

On remarque que le mélange fort implique la condition D.
On définit maintenant des conditions plus faibles que D, appelées la condition $D(u_n)$ et la condition $D'(u_n)$ de Leadbetter [60].

Définition 1.16. *(Condition $D(u_n)$)*
Une suite de variables aléatoires $(X_n)_n$ vérifie la condition $D(u_n)$ si pour une suite de réels $(u_n)_n$ et $\forall p,q,n \in \mathbb{N}$ tels que : $1 \leq i_1 < ... < i_p < j_1 < ... < j_q \leq n$, $j_1 - i_p \geq l$, on a

$$\left|F_{i_1,i_2,...,i_p,i_1,j_1,...,j_q}(u_n,...,u_n) - F_{i_1,i_2,...,i_p}(u_n,...,u_n)F_{j_1,...,j_q}(u_n,...,u_n)\right| \leq \beta_{n,l_n}$$

avec $\beta_{n,l_n} \to 0$ quand $n \to \infty$ pour une suite $l = l_n = o(n)$.

Cette condition est peu restrictive car elle ne porte que sur les événements A de type $A = \bigcap_{r=1}^{p} \{X_{i_r} \leq u_n\}$ et pas nécessairement sur les lois jointes.

Définition 1.17. *(La condition $D'(u_n)$)*
Une suite stationnaire de variables aléatoires $(X_n)_n$ vérifie la condition $D'(u_n)$ si

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{\left[\frac{n}{k}\right]} P\left(X_1 > u_n, X_j > u_n\right) = 0$$

pour une suite réelle $(u_n)_n$ et $[.]$ désigne la partie entière.

La condition $D'(u_n)$ permet de borner la probabilité pour qu'il n'y ait pas plus d'un élément supérieur à u_n parmi $X_1,...,X_{\left[\frac{n}{k}\right]}$. Le théorème suivant généralise le théorème 1.1 au cas d'un processus stationnaire vérifiant la condition $D(u_n)$.

Théorème 1.10. *(Leadbetter et Rootzén [62])*
Soit $(X_n)_n$ une suite stationnaire de variables aléatoires et $a_n > 0$, b_n deux suites réelles telles que :

$$\lim_{n\to\infty} P\left(M_n \leq a_n x + b_n\right) = H(x)$$

Si $(X_n)_n$ vérifie la condition $D(u_n)$ pour $u_n = a_n x + b_n$, alors la distribution H appartient à l'une des familles de Gumbel, Fréchet ou Weibull.

Remarque 1.5.
Le résultat du théorème précédent reste vrai en remplaçant la condition $D(u_n)$ par la condition D.

La démonstration de ce théorème repose sur le lemme suivant

Lemme 1.4. *Soit une suite stationnaire de variables aléatoires $(X_n)_n$ vérifiant la condition $D(u_n)$, alors pour tout entier $k \geq 1$, on a*

$$\lim_{n\to\infty} \left[P(M_n \leq u_n) - P^k(M_{[\frac{n}{k}]} \leq u_n) \right] = 0$$

En divisant la partie $[1, n]$ en k intervalles de longueur $\left[\dfrac{n}{k}\right]$, on peut supposer que les maximums $M_{[\frac{n}{k}]}$ observés sur chaque intervalle sont approximativement indépendants.

Leadbetter a montré que le résultat du théorème 1.9 demeure vrai en remplaçant la condition de mélange fort par la condition $D'(u_n)$.

Théorème 1.11. *(Leadbetter et Rootzén [62])*
Soit $(X_n)_n$ une suite stationnaire de variables aléatoires, vérifiant les conditions $D(u_n)$ et $D'(u_n)$ avec $u_n = a_n x + b_n$, ($a_n > 0$, $b_n \in \mathbb{R}$).
Si pour une suite $(X_n^*)_n$ i.i.d de même distribution que X_1, on a :

$$\lim_{n\to\infty} P\left(M_n^* \leq u_n\right) = H(x)$$

avec $M_n^* = \max\limits_{1\leq i\leq n} X_i^*$, alors

$$\lim_{n\to\infty} P\left(M_n \leq u_n\right) = H(x)$$

Le résultat du théorème 1.2 a été établi par Leadbetter dans le cas d'une suite stationnaire vérifiant les conditions $D(u_n)$ et $D'(u_n)$ ($(u_n)_n$ est une suite de constantes). Une première preuve de ce résultat est proposée par Davis [22].

Théorème 1.12. *(Leadbetter et al.[61], Davis [22])*
Soit $(X_n)_n$ une suite stationnaire vérifiant les conditions $D(u_n)$ et $D'(u_n)$, alors

$$P(M_n \leq u_n) \to \exp(-\tau) \Longleftrightarrow n(1 - F(u_n)) \to \tau, \quad \textit{quand } n \to \infty \textit{ et } 0 \leq \tau < \infty$$

Si $n(1-F(u_n)) \to \infty$, la condition $D'(u_n)$ n'est pas toujours satisfaite. Le corollaire suivant donne une extension du théorème précédent au cas $\tau = \infty$.

Corollaire 1.2. *(Leadbetter et al.[61])*
On suppose que pour τ fini et suffisamment grand, il existe une suite $(v_n)_n$ telle que $(X_n)_n$ vérifie les conditions $D(v_n)$ et $D'(v_n)$ et si de plus $n(1-F(v_n)) \to \tau$, alors le résultat du théorème précédent reste vrai, i.e.,

$$P(M_n \leq u_n) \to 0 \Longleftrightarrow n(1-F(u_n)) \to \infty, \quad \textit{quand } n \to \infty$$

Chapitre 2

Estimation de l'indice des valeurs extrêmes

2.1 Introduction

Dans ce chapitre, nous nous intéressons aux différentes méthodes d'estimation du paramètre ξ intervenant dans la distribution des valeurs extrêmes généralisée présentée dans la partie précédente. En se référant à la littérature, diverses méthodes paramétriques et semi-paramétriques ont été proposées pour l'estimation de cet indice.
Pour les méthodes paramétriques, nous pouvons mentionner, par exemple, les méthodes d'estimation empirique (voir Gumbel et Mustafi [53] et Tiago de Oliveira [84]), la méthode du maximum de vraisemblance (voir Smith [83], Prescott et Walden [72]), la méthode des moments (voir Christopiet [15]) ou encore les méthodes bayésiennes (Lye et al.[67]).
D'autre part, des approches non-paramétriques ont été consacrées à l'estimation de l'indice de queue ; les estimateurs les plus utilisés sont les estimateurs de Hill [56] et de Pickands [71]. Le premier est appliqué seulement dans le domaine d'attraction de Fréchet ($\xi > 0$) et le deuxième est valable pour les différentes lois limites des extrêmes.
Hosking et Wallis [57] ont utilisé la méthode des moments et proposent un estimateur des moments pondérés défini seulement pour $\xi < 1$. Dekkers et al. [32] généralisent ce dernier estimateur pour $\xi \in \mathbb{R}$, qu'on appelle aussi l'estimateur de Dekkers-Einmahl-De Haan.

Falk[40] a proposé une variante de l'estimateur du maximum de vraisemblance appelé l'estimateur de Hill négatif. D'autres estimateurs ont été également donnés, voir Peng [70], Lo [64], Diop et Lo [34].

Dans ce chapitre, on considère une suite $(X_n)_n$ de variables aléatoires *i.i.d* de fonction de répartition F. Désignons par $X_{(1)} \geq X_{(2)} \geq ... \geq X_{(n)}$ les statistiques d'ordre correspondantes.

Définition 2.1. *On dit qu'une suite $(k_n)_n$ d'entiers est intermédiaire si :*

$$\lim_{n\to\infty} k_n = \infty \quad et \quad \lim_{n\to\infty} \frac{k_n}{n} = 0$$

.

Pour simplifier l'écriture, on notera $k = k_n$.

2.2 Estimateur de Pickands

Cet estimateur introduit par Pickands [71] est fondé sur le calcul des quantiles et peut être utilisé dans l'estimation du paramètre $\xi \in \mathbb{R}$ sans se limiter à un seul domaine d'attraction.

Définition 2.2. *L'estimateur de Pickands de $\xi \in \mathbb{R}$, noté $\widehat{\xi}_k^P$ est défini par :*

$$\widehat{\xi}_k^P = \frac{1}{\log 2} \log \left(\frac{X_{(k)} - X_{(2k)}}{X_{(2k)} - X_{(4k)}} \right)$$

Il faut noter que l'estimateur de Pickands utilise la statistique d'ordre $4k$ donc la suite k doit rester inférieure à $\frac{n}{4}$.

On peut justifier la définition de cet estimateur de la manière suivante :

En utilisant la relation (1.7), pour $\xi \in \mathbb{R}$, avec $t = 2s, x = 2$ et $y = \frac{1}{2}$, on a

$$\lim_{t\to\infty} \frac{b(t) - b(\frac{t}{2})}{b(\frac{t}{2}) - b(\frac{t}{4})} = 2^{\xi}$$

En remplaçant b par $\widehat{b}_n = \left(\frac{1}{1 - \widehat{F}_n} \right)^{\leftarrow}$ ($\widehat{F}_n$ étant la fonction de répartition empirique) et t par $\dfrac{n}{k}$, où k est une suite intermédiaire, on a :

$$\frac{\widehat{b}_n(\frac{n}{k}) - \widehat{b}_n(\frac{n}{2k})}{\widehat{b}_n(\frac{n}{2k}) - \widehat{b}_n(\frac{n}{4k})} = 2^{\xi} \quad \text{pour n suffisamment grand}$$

Du fait que $\widehat{b}_n(x) = X_{([\frac{n}{x}])}$, alors

$$\frac{X_{(k)} - X_{(2k)}}{X_{(2k)} - X_{(4k)}} = 2^{\xi}$$

L'estimateur de Pickands est alors la solution de cette dernière équation. La consistance faible a été obtenue par Pickands et la normalité asymptotique par Dekkers et al. [33].

2.2.1 Consistance

Théorème 2.1. *Supposons que $F \in D(H_\xi)$ et k est une suite intermédiaire, alors*

$$\widehat{\xi}_k^P \xrightarrow{P} \xi$$

En ajoutant une condition sur la suite k, on obtient la convergence presque sûre de l'estimateur de Pickands.

Théorème 2.2. *Supposons que $F \in D(H_\xi)$ et k est une suite intermédiaire vérifiant $\lim\limits_{n\to\infty} \dfrac{k}{\log\log n} = \infty$, alors*

$$\widehat{\xi}_k^P \xrightarrow{p.s} \xi$$

.

2.2.2 Normalité asymptotique

La normalité asymptotique nécessite une condition supplémentaire sur la fonction de distribution F.

Définition 2.3. *On dit qu'une fonction F est à variation régulière du second ordre d'indice (ξ, ρ), $\rho \leq 0$ (on note $\overline{F} \in 2RV(\xi, \rho)$), s'il existe une fonction g à signe constant avec $g \in RV_\rho$ et $g(t) \to 0$ quand $t \to \infty$, telle que :*

$$\lim_{t\to\infty} \frac{\dfrac{\overline{F}(tx)}{\overline{F}(t)} - x^{\xi}}{g(t)} = c x^{\xi} \frac{x^{\rho} - 1}{\rho} \quad , \ \forall x > 0, \ \ c \in \mathbb{R}^* \tag{2.1}$$

g est appelée la fonction auxiliaire de F.

Notons que la fonction à variation régulière g est aussi appelée la fonction de biais et le paramètre ρ, appelé paramètre du second ordre, contrôle la vitesse de convergence de $\dfrac{\overline{F}(tx)}{\overline{F}(t)}$ vers x^{ξ} ; une valeur de ρ proche de 0 implique une faible vitesse de convergence.

Théorème 2.3. *(Normalité asymptotique de l'estimateur de Pickands)*
Supposons que k est une suite intermédiaire et que $F \in D(H_\xi)$ satisfait la condition (2.1). *Si la fonction auxiliaire de F est telle que*

$$\lim_{n\to\infty} \sqrt{k}g(\frac{n}{k}) = \lambda, \quad \textit{avec } \lambda \textit{ fini} \tag{2.2}$$

alors

$$\sqrt{k}\left(\widehat{\xi}_k^P - \xi\right) \xrightarrow{d} N\left(\lambda b_{\xi,\rho},\ var_\xi\right)$$

où N est une loi normale, avec

$$b_{\xi,\rho} = \begin{cases} \dfrac{4^{-\rho}\xi\left[\left(4^{\xi+\rho}-1\right)-\left(2^{\xi}+1\right)\left(2^{\xi+\rho}-1\right)\right]}{\rho^{2\xi}\left(\xi+\rho\right)\left(2^{\xi}-1\right)\log 2} & si \quad \rho < 0 \neq \xi \\ \dfrac{1-2^{-\rho+1}+4^{-\rho}}{\rho^2\left(\log 2\right)^2} & si \quad \rho < 0 = \xi \\ 1 & si \quad \rho = 0 \end{cases}$$

et

$$var_\xi = \begin{cases} \dfrac{\xi^2\left(2^{2\xi+1}+1\right)}{4\left(\log 2\right)^2\left(2^{\xi}-1\right)^2} & si \quad \xi \neq 0 \\ \dfrac{3}{4\left(\log 2\right)^4} & si \quad \xi = 0 \end{cases}$$

Ce dernier théorème nous permet de donner un intervalle de confiance pour ξ. Notons que les simulations montrent que cet estimateur possède un biais important ; cela s'explique par le fait que $\widehat{\xi}_k^P$ n'utilise pas le maximum de l'échantillon $X_{(1)}$ mais l'information apportée par les distances entre deux statistiques d'ordre.

On remarque aussi que :

1- Si k est petit, on a une estimation avec un intervalle de confiance large.

2- Si k est grand, l'intervalle de confiance est plus étroit mais il faut tenir compte d'un biais.

De plus, l'estimateur de Pickands possède une grande variance, ce qui a amené de nombreux auteurs à proposer des variantes de cet estimateur avec des variances plus faibles (voir [7]).

2.3 Estimateur de Hill

Cet estimateur est le plus utilisé en théorie des valeurs extrêmes et a été largement étudié dans le cas de variables aléatoires *i.i.d.* Mason [68] et Deheuvels et al.[31] ont montré respectivement la consistance faible et forte qui ne dépend que du comportement de la suite k. Pour établir la normalité asymptotique, on a besoin de supposer que la fonction de répartition F est à variation régulière du second ordre. Plusieurs auteurs ont obtenu cette normalité, notamment Dekkers et al. [32], Csörgő et Mason [19], Davis et Resnick [24], Geluk et De Haan[46], Haeusler et Teugels [54].

2.3.1 Construction de l'estimateur de Hill

a. Première approche : Une première approche pour construire l'estimateur de Hill est basée sur le lemme suivant :

Lemme 2.1. *Si $F \in D(\Phi_\beta)$, alors*

$$\frac{1}{\overline{F}(t)} E\left[(\log X - \log t)1_{(X>t)}\right] \underset{t\to+\infty}{\longrightarrow} \frac{1}{\beta} = \xi$$

Pour construire l'estimateur de Hill, il faut trouver des estimateurs respectivement pour $\overline{F}$ et pour $E\left[(\log X - \log t)1_{(X>t)}\right]$.
Le meilleur estimateur de F est la fonction de répartition empirique

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(X_i \le x)} = \frac{1}{n}\sum_{i=1}^{n} 1_{\left(X_{(i)} \le x\right)}$$

D'après la loi forte des grands nombres, on a :

$$\frac{1}{n}\sum_{i=1}^{n} 1_{(X_i > t)} \xrightarrow{p.s} E(1_{(X>t)}) = P(X > t) = \overline{F}(t) \tag{2.3}$$

D'autre part, pour une suite intermédiaire k, on a $X_{(k)} \xrightarrow{p.s} x_F = \infty$ (voir De Haan et Ferreira [26]).

Si l'on suppose que F est continue, la statistique d'ordre est strictement croissante presque sûrement. En remplaçant t par $X_{(k)}$ dans (2.3), on obtient une estimation de $\overline{F}(X_{(k)})$:

$$\frac{1}{n}\sum_{i=1}^{n} 1_{\left(X_i > X_{(k)}\right)} = \frac{1}{n}\sum_{i=1}^{n} 1_{\left(X_{(i)} > X_{(k)}\right)} = \frac{k-1}{n}.$$

En utilisant la loi des grands nombres, on obtient

$$\frac{1}{n}\sum_{i=1}^{n}(\log X_i - \log t)1_{(X_i>t)} \xrightarrow{p.s} g(t) = E\left[(\log X_i - \log t)1_{(X_i>t)}\right]$$

On remplace de nouveau t par $X_{(k)}$, on obtient comme estimateur de $g(X_{(k)})$:

$$\frac{1}{n}\sum_{i=1}^{n}(\log X_i - \log X_{(k)})1_{\left(X_i > X_{(k)}\right)} = \frac{1}{n}\sum_{i=1}^{k-1}\log X_{(i)} - (k-1)\log X_{(k)}$$

On déduit que l'estimateur de $\overline{F}(t)$ est $\dfrac{k-1}{n}$ et l'estimateur de $E\left[(\log X - \log t)1_{(X>t)}\right]$ est

$$\frac{1}{n}\sum_{i=1}^{k-1}\log X_{(i)} - (k-1)\log X_{(k)}$$

L'estimateur de $\dfrac{1}{\overline{F}(t)}E\left[(\log X\text{-} \log t)1_{(X>t)}\right]$ est donc

$$\frac{1}{k-1}\sum_{i=1}^{k-1}\log X_{(i)} - \log X_{(k)}$$

En remplaçant $k-1$ par k, on obtient

$$\frac{1}{k}\sum_{i=1}^{k}\log X_{(i)} - \log X_{(k+1)}$$

D'autre part (voir Embrechts et al.[39]), si la distribution F est à variation régulière et k est une suite intermédiaire on a

$$\frac{X_{(k)}}{X_{(k+1)}} \xrightarrow{P} 1$$

Donc

$$\frac{1}{k}\sum_{i=1}^{k}\log X_{(i)} - \log X_{(k)}$$

est un estimateur de $\dfrac{1}{\beta}$.

b- Deuxième approche : Une deuxième approche est celle où l'on utilise l'estimateur du maximum de vraisemblance conditionnel.
On considère une variable aléatoire X de fonction de répartition F telle que, pour $\beta > 0$

$$\overline{F}(x) = x^{-\beta}, \ x \geq 1$$

Si on pose $Y = \log X$, la fonction de répartition de Y est définie par :

$$\begin{aligned} P(Y > y) &= P(X > \exp y) = 1 - F(\exp y) \\ &= \exp(-\beta y), \ y \geq 0 \end{aligned}$$

Y suit donc une loi exponentielle de paramètre β de densité $f(y) = \beta \exp(-\beta y), \quad y \geq 0$.
La fonction de vraisemblance L est

$$\begin{aligned} L(y_1, y_2, ..., y_n, \beta) &= \prod_{i=1}^{n} \beta \exp(-\beta y_i) \\ &= \beta^n \exp\left(-\beta \sum_{i=1}^{n} y_i\right) \\ \frac{\partial \log L(y_1, y_2, ..., y_n, \beta)}{\partial \beta} &= \frac{n}{\beta} - \sum_{i=1}^{n} y_i \\ \frac{\partial^2 \log L(y_1, y_2, ..., y_n, \beta)}{\partial \beta^2} &= -\frac{n}{\beta^2} \end{aligned}$$

L'estimateur du maximum de vraisemblance de β est

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} \log X_i\right)^{-1} = \left(\frac{1}{n}\sum_{i=1}^{n} \log X_{(i)}\right)^{-1}$$

Une généralisation triviale concerne la fonction de répartition suivante :

$$F(x) = 1 - Cx^{-\beta}, \ x \geq u > 0, u \textit{ fixé}$$

On pose $C = u^{\beta}$, l'EMV de β est

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} \log \frac{X_{(i)}}{u}\right)^{-1} = \left(\frac{1}{n}\sum_{i=1}^{n} \log X_{(i)} - \log u\right)^{-1}$$

Dans le domaine d'attraction de Fréchet et pour x très grand, la fonction $\overline{F}$ suit la loi de Pareto (voir la remarque 1.3).
Posons $K = card\left\{i : X_{(i)} > u, i = 1, ..., n\right\}$
Conditionnellement à l'événement $\{K = k\}$, la méthode du maximum de vraisemblance pour estimer β et C consiste à maximiser la densité conjointe de $(X_{(k)}, ..., X_{(1)})$ définie par :

$$f_{x_{(k)},,...,x_{(1)}}(x_k, ..., x_1) = \frac{n!}{(n-k)!}(1 - Cx_k^{-\beta})^{n-k}C^k\beta^k\prod_{i=1}^{k}x_i^{-(\beta+1)}, \quad u < x_k < ... < x_1.$$

L'EMV conditionnel du paramètre β^{-1} est :

$$\frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k)}} = \frac{1}{k}\sum_{i=1}^{k}\log X_{(i)} - \log X_{(k)}$$

Définition 2.4. *L'estimateur de Hill de $\xi > 0$, noté $\widehat{\xi}_k^H$ est défini par :*

$$\widehat{\xi}_k^H = \frac{1}{k}\sum_{i=1}^{k}\log X_{(i)} - \log X_{(k)} \tag{2.4}$$

Une particularité de l'estimateur de Hill est qu'il est possible de l'interpréter graphiquement. Ceci est particulièrement important pour les praticiens qui préfèrent souvent des interprétations graphiques à des formulations mathématiques. Plus précisément, si l'on considère le graphe de coordonnées

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

appelé *Pareto quantile plot*, dans le cas d'une distribution de type Pareto ce graphe est approximativement linéaire, dans les points extrêmes, avec une pente $\xi = \dfrac{1}{\beta}$. L'estimateur de Hill n'est que l'estimateur de cette pente.

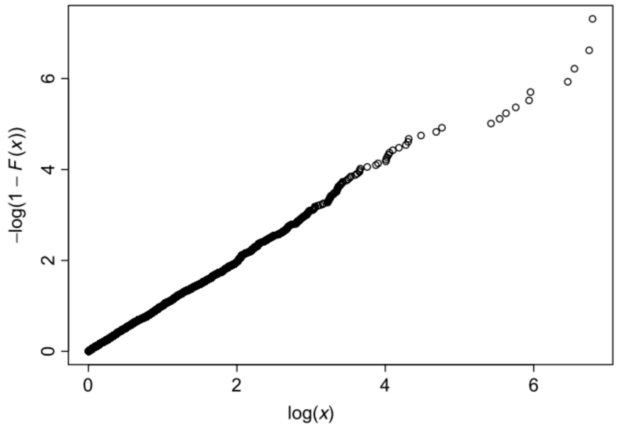

FIGURE 2.1 – Pareto quantile plot pour un échantillon de taille n, de distribution de Pa(1).

En effet, si la distribution F est dans le domaine d'attraction de Fréchet, la fonction de queue vérifie :

$$b(x) = x^{\frac{1}{\beta}} L(x)$$

où L est une fonction à variation lente ; il découle que

$$\log b(x) = \frac{1}{\beta} \log(x) + \log L(x) = \frac{1}{\beta} \log x \left(1 + \frac{\log L(x)}{\frac{1}{\beta} \log x} \right)$$

D'après la propriété (1.6), on a :

$$\log b(x) \sim \frac{1}{\beta} \log(x), \ x \to \infty$$

En remplaçant la fonction de queue par sa version empirique $\widehat{b}_n(t) = F_n^{\leftarrow}\left(1 - \frac{1}{t}\right)$ et en remarquant que $\widehat{b}_n\left(1 - \frac{i}{n+1}\right) = F_n^{\leftarrow}\left(\frac{i}{n+1}\right) = X_{(i)}$, on obtient

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

2.3.2 Consistance

La consistance faible est démontrée par plusieurs auteurs ; De Haan et Ferreira [26] proposent une démonstration s'appuyant sur la représentation de Rényi tandis que Mason [68] ou Resnick [76] utilisent la mesure empirique.

Théorème 2.4. *(Mason [68])*
Supposons que $F \in D(H_\xi)$ avec $\xi > 0$ et k est une suite intermédiaire, alors

$$\widehat{\xi}_k^H \xrightarrow{P} \xi$$

Le théorème suivant est une réciproque du résultat précédent.

Théorème 2.5. *(Mason [68]-De Haan et Ferreira [26])*
On suppose que k est une suite intermédiaire vérifiant $\lim_{n\to\infty} \dfrac{k_{n+1}}{k_n} = 1$ et

$$\widehat{\xi}_k^H \xrightarrow{P} \xi > 0$$

alors $F \in D(H_\xi)$.

En ajoutant une condition supplémentaire sur la suite k, Deheuvels et al.[31] obtiennent la consistant forte de cet estimateur.

Théorème 2.6. *(Deheuvels et al. [31])*
Si $F \in D(H_\xi)$ avec $\xi > 0$ et k est une suite intermédiaire vérifiant $\lim_{n\to\infty} \dfrac{k}{loglogn} = \infty$, alors

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

2.3.3 Normalité asymptotique

La normalité asymptotique de l'estimateur de Hill a été démontrée par Davis et Resnick [24], en utilisant les propriétés des statistiques d'ordre d'un échantillon issu d'une loi exponentielle et les conditions de Von Mises. Csörgő et Mason [19] l'ont montré en introduisant l'approximation des processus empiriques par des ponts browniens. De même, en supposant la variation régulière au second ordre de la distribution, De Haan et Resnick [28] ont aussi obtenu cette normalité asymptotique.

Théorème 2.7. *Supposons que $F \in D(H_\xi)$ $(\xi > 0)$ satisfait la condition* (2.1) *et sa fonction auxiliaire vérifie* (2.2), *alors*

$$\sqrt{k}\left(\widehat{\xi}_k^H - \xi\right) \xrightarrow{d} N(\frac{\lambda}{1-\rho}, \xi^2)$$

où k est une suite intermédiaire .

Dans le cas où l'on a $\lim_{t\to\infty} t^{\beta}\left(\frac{b(tx)}{b(t)} - x^{\xi}\right) = 0$, pour tout $x > 0$ et $\beta > 0$, l'assertion du théorème précédent reste valable et le bais $\frac{\lambda}{1-\rho}$ vaut zéro.

Si de plus $\sqrt{k}g\left(\frac{n}{k}\right) \to 0$, alors $\sqrt{k}\left(\frac{\overset{\wedge}{\xi}{}_k^H}{\xi} - 1\right) \xrightarrow{d} N(0,1)$.

On peut donc en déduire un intervalle de confiance pour ξ au niveau $(1-\beta)$ donné par :

$$\left]\frac{\widehat{\xi}_k^H}{1+\frac{Q_{1-\beta/2}}{\sqrt{k}}}, \frac{\overset{\wedge}{\xi}{}_k^H}{1-\frac{Q_{1-\beta/2}}{\sqrt{k}}}\right[$$

Q_{β} est le quantile d'une loi normale d'ordre β.

La précision de l'estimation dépend de la sélection de k, l'estimateur de Hill a un biais qui tend vers zéro quand k est relativement petit. La variance asymptotique de cet estimateur vaut $\frac{\xi^2}{k}$ qui tend vers zéro quand k est grand. Par conséquent, il faut trouver un compromis entre le biais et la variance asymptotique. Beirlant et al. [9] trouvent un intervalle de confiance $\overset{\wedge}{\xi}{}_{\hat{k}_{nopt}}^H$ où $\hat{k}_{nopt}$ est une valeur optimale de k donnée par leur algorithme. Danielsson et al. [20] proposent une autre méthode plus précise pour la construction d'un intervalle de confiance, basée sur le Bootstrap.

Pour la réduction du biais, d'autres estimateurs ont été donnés ; notamment celui proposé par Beirlant et al. [5, 6] qui utilisent un modèle de régression exponentielle. L'estimateur à noyau de Csörgő et al. [18] est une extension de estimateur de Hill pour $\xi \in \mathbb{R}$. Une autre variante meilleure que les précédentes est celle de Falk et Morohn [42].

2.4 Estimateur des moments (de Dekkers-Einmahl-De Haan)

Définition 2.5. *L'estimateur de Dekkers-Einmahl-De Haan (DEdH) du paramètre ξ, noté $\overset{\wedge}{\xi}{}_k^D$ est défini par :*

$$\overset{\wedge}{\xi}{}_k^D = 1 + H_n^{(1)} + \frac{1}{2}\left(\frac{(H_n^{(1)})^2}{H_n^{(2)}} - 1\right)^{-1}$$

où

$$\begin{cases} H_n^{(1)} = \dfrac{1}{k}\sum\limits_{i=1}^{k-1} \log X_{(i)} - \log X_{(k)} \\ H_n^{(2)} = \dfrac{1}{k}\sum\limits_{i=1}^{k-1} \left(\log X_{(i)} - \log X_{(k)}\right)^2 \end{cases}$$

Il est appelé l'estimateur des moments car $H_n^{(1)}$ et $H_n^{(2)}$ peuvent être considérés comme des moments empiriques.

Dekkers et al. [32] ont établi la consistance et la normalité asymptotique de cet estimateur.

2.4.1 Consistance

Théorème 2.8. *Supposons que $F \in D(H_\xi)$ avec $\xi \in \mathbb{R}$ et que k est une suite intermédiaire. Si $x_F > 0$, alors*

$$\overset{\wedge D}{\xi_k} \xrightarrow{P} \xi$$

Si de plus, $\lim\limits_{n\to\infty} \dfrac{k}{(logn)^\delta} = \infty$ avec $\delta > 0$, alors

$$\overset{\wedge D}{\xi_k} \xrightarrow{p.s} \xi$$

2.4.2 Normalité asymptotique

Pour la normalité asymptotique de l'estimateur des moments, on suppose que la fonction $\log b$ (b est la fonction quantile de queue) vérifie la condition (2.1).

Théorème 2.9. *(Dekkers et al. [32])*
Supposons que $F \in D(H_\xi)$ et $x_F > 0$. Si la condition (2.1) *est satisfaite pour la fonction* $\log b$ *et que sa fonction auxiliaire vérifie* (2.2),
alors pour une suite intermédiaire k, on a

$$\sqrt{k}\left(\overset{\wedge D}{\xi_k} - \xi\right) \xrightarrow{d} N(\lambda c_{\xi,\rho}, var_\xi)$$

avec

$$c_{\xi,\rho} = \begin{cases} \dfrac{(1-\xi)(1-2\xi)}{(1-\xi-\rho)(1-2\xi-\rho)} & si \quad \xi < \rho \leq 0 \\ \dfrac{\xi(1+\xi)}{(1-\xi)(1-3\xi)} & si \quad \rho < \xi \leq 0 \\ -\dfrac{\xi}{(1+\xi)^2} & si \quad 0 < \xi < -\rho \quad et\ l \neq 0 \\ \dfrac{\xi-\xi\rho+\rho}{\rho(1-\rho)^2} & si \quad (0 < \xi < -\rho \quad et\ l \neq 0)\ ou\ \xi \geq -\rho > 0 \\ 1 & si \quad \xi > \rho = 0 \end{cases}$$

et

$$var_\xi = \begin{cases} \xi^2+1 & si\ \ \xi \geq 0 \\ \dfrac{(1-\xi)^2(1-2\xi)(1-\xi+6\xi^2)}{(1-3\xi)(1-4\xi)} & si\ \ \xi < 0 \end{cases}$$

2.5 Estimateur de Hill adapté

Le problème de l'estimateur des moments, contrairement à celui de Hill, ne peut être interprété graphiquement. Afin d'apporter une solution à ce problème, une généralisation a été proposée par Beirlant et al. [8], ce nouveau estimateur est appelé l'estimateur de Hill adapté.

On considère la fonction bH définie par

$$bH(x) = b(x)H(x) = b(x)E\left(\log X - \log b(x) \mid X > b(x)\right) \tag{2.5}$$

où b est la fonction quantile de queue.

Proposition 2.1. *(Beirlant et al. [8])*
Si $F \in D(H_\xi)$, alors la fonction définie par (2.5) *est à variation régulière d'indice ξ.*

La fonction (2.5) peut être écrite sous la forme

$$bH(x) = x^\xi L(x)$$

où L est une fonction à variation lente.

En utilisant la représentation de Karamata pour L, on écrit

$$bH(x) = x^{\xi} c(x).\exp\left\{\int_1^x \frac{\varepsilon(u)}{u} du\right\}, \qquad x > 0 \tag{2.6}$$

où $\lim_{x\to\infty} c(x) = 1$ et $\lim_{x\to\infty} \varepsilon(x) = 0$

L'estimateur empirique de la fonction (2.5) évaluée en $x = \dfrac{n}{n-k}$ est

$$bH_k = X_{(k+1)}\left(\frac{1}{k}\sum_{j=1}^{k} \log X_{(j)} - \log X_{(k+1)}\right) = X_{(k+1)}\widehat{\xi}_k^H$$

Il est naturel de considérer le *quantile plot généralisé* dont les points ont pour coordonnées :

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

Beirlant et al. [8] ont proposé l'estimateur suivant pour $\xi \in \mathbb{R}$, basé sur une approximation de la pente de *quantile plot généralisé.*

Définition 2.6. *L'estimateur de Hill adapté de $\xi \in \mathbb{R}$ est défini par :*

$$\widehat{\xi}_k^{adH} = \frac{1}{k}\sum_{j=1}^{k} \log(bH_j) - \log(bH_{k+1})$$

Pour $\rho \leq 0$ et $t > 0$, on définit la fonction suivante

$$l_\rho(t) = \begin{cases} \dfrac{t^\rho - 1}{\rho} & \text{si} \quad \rho < 0 \\ \log t & \text{si} \quad \rho = 0 \end{cases}$$

Le théorème suivant nous donne la normalité de l'estimateur de Hill adapté.

Théorème 2.10. *(Beirlant et al. [8])*

Soit k une suite intermédiaire, on suppose que F est à variation régulière d'indice $\xi \in \mathbb{R}$ et la fonction ε définie dans (2.6) *est telle que* $\lim_{x\to\infty} \sqrt{k}\int_0^1 \varepsilon(\frac{n}{kt})dt = 0$.

Si de plus

1. *pour $\xi > 0$: il existe $\rho_1 \leq 0$ et une fonction positive b_1 vérifiant $\lim_{x\to\infty} b_1(x) = 0$ et $\lim_{n\to\infty} \sup_{\frac{1}{k}\leq t\leq 1} \sqrt{k} b_1\left(\frac{n}{tk}\right) = 0$, tels que*

$$\frac{(tx)^{-\xi} b(tx)}{x^{-\xi} b(x)} - 1 \sim b_1(x) l_{\rho_1}(t), \qquad t \to \infty$$

2. *pour* $\xi = 0$ *: il existe deux fonctions positives* a *et* b_2 *vérifiant* $\lim_{x\to\infty} b_2(x) = 0$ *et* $\lim_{n\to\infty} \sup_{\frac{1}{k}\leq t\leq 1} \sqrt{k}\frac{b_2\left(\frac{n}{tk}\right)}{a\left(\frac{n}{tk}\right)} = 0$*, telles que*

$$\lim_{x\to\infty} \frac{b(t)(\log b(tx) - \log b(t)) - a(x)\log t}{b_2(x)} = \pm\frac{(\log t)^2}{2}, \quad t \geq 1$$

3. *pour* $\xi < 0$ *: il existe* $\rho_2 \leq 0$ *et une fonction positive* b_3 *vérifiant* $\lim_{x\to\infty} b_3(x) = 0$ *et* $\lim_{n\to\infty} \sup_{\frac{1}{k}\leq t\leq 1} \sqrt{k} b_3\left(\frac{n}{tk}\right) = 0$ *tels que*

$$\frac{(tx)^{-\xi}(b(\infty) - b(tx))}{x^{-\xi}(b(\infty) - b(x))} - 1 \sim b_3(x) l_{\rho_2}(t), \quad t \to \infty$$

alors

$$\sqrt{k}\left(\widehat{\xi}_k^{adH} - \xi\right) \xrightarrow{d} N(0, c_\xi)$$

avec

$$c_\xi = \begin{cases} (1+\xi)^2 & si \quad \xi \geq 0 \\ \dfrac{(1-\xi)(1+\xi+2\xi^2)}{1-2\xi} & si \quad \xi < 0 \end{cases}$$

2.6 Estimateur de Hill négatif

L'estimateur de Hill négatif proposé par Falk [40] est un prolongement de l'estimateur du maximum de vraisemblance au cas où $\xi < -\dfrac{1}{2}$.

On remarque que si une distribution F appartient au domaine d'attraction de H_ξ avec $\xi < 0$, alors le point terminal x_F est fini ; de plus la variable $\widetilde{X} = \dfrac{1}{x_F - X}$ est dans le domain d'attraction de $H_{-\xi}$. Il suffit d'appliquer l'estimateur de Hill pour la variable $\widetilde{X}$. Comme x_F est inconnu, il est estimé par la plus grande statistique d'ordre $X_{(1)}$ (seulement dans le cas $\xi < -\dfrac{1}{2}$, voir De Haan et Ferreira [26]).

Définition 2.7. *Si* $\xi < -\dfrac{1}{2}$*, l'estimateur de Hill négatif est défini par :*

$$\widehat{\xi}_k^F = \frac{1}{k}\sum_{j=1}^{k-1} \log\left(X_{(1)} - X_{(j+1)}\right) - \log\left(X_{(1)} - X_{(k+1)}\right)$$

Théorème 2.11. *Soit k une suite intermédiaire vérifiant* $\lim_{n\to\infty} \frac{k^\tau}{\log n} = \infty$ *pour τ assez petit.*

1. *Si $F \in D(H_\xi)$ avec $\xi < -\frac{1}{2}$, alors*
$$\widehat{\xi}_k^F \xrightarrow{p} \xi$$
2. *Si $F \in D(H_\xi)$ avec $-1 < \xi < -\frac{1}{2}$, satisfait la condition* (2.1) *et sa fonction auxiliaire vérifie* (2.2), *alors*
$$\sqrt{k}(\widehat{\xi}_k^F - \xi) \xrightarrow{d} N(M, \xi^2)$$
avec
$$M = \begin{cases} \dfrac{\lambda\xi}{\rho(1+\xi)(1-\rho)} & si \;\; \rho < 0 \\ \lambda & si \;\; \rho = 0 \end{cases}$$

La principale difficulté en pratique consiste à choisir le nombre d'observations k. On remarque qu'en utilisant trop d'observations dans la procédure d'estimation de l'indice des valeurs extrêmes, on sort de la queue de la distribution et on observe un biais important. Tandis que si k est très faible, le peu d'observations rend l'estimateur instable et conduit à une variance considérable.

Rappelons que sous des conditions appropriées, l'estimateur de Hill est consistant seulement pour des valeurs positives de ξ ; les estimateurs de Pickands, des moments et Hill adapté sont définis et consistants pour tout réel ξ, tandis que l'estimateur de Hill négatif a cette propriété seulement pour $\xi < -\frac{1}{2}$.

D'un point de vue théorique, tous les estimateurs vérifient la même propriété de normalité asymptotique. Cependant, les estimateurs de Hill et Hill négatif ont la plus petite variance asymptotique, respectivement pour les valeurs $\xi > 0$ et $\xi < -\frac{1}{2}$.

La comparaison de biais de ces estimateurs est plus compliquée puisqu'en général le biais dépend des paramètres ξ et ρ, de la suite k et d'autres caractéristiques de la distribution considérée.

En général, il n'y a pas une meilleure méthode dans toutes les situations. Les estimateurs les plus utilisés sont ceux de Hill, Pickands et des moments. Certaines études de comparaison (théorique et par simulation) entre les différentes méthodes peuvent être trouvées dans Rosen et Weissman [81], Peng [70], de Haan et Peng [27], Groeneboom et al. [51].

Chapitre 3

Estimateur de Hill sous dépendance

3.1 Introduction

Dans certains cas, les données recueillies indiquent que la distribution de la variable est à queue lourde d'indice β qui peut être estimé en utilisant l'estimateur de Hill. Le comportement de cet estimateur est largement étudié dans le cas *i.i.d.* Cependant de nombreuses applications pratiques ne fournissent pas des données indépendantes mais plutôt dépendantes. Par conséquent, il est important de généraliser l'estimateur de Hill aux données non indépendantes stationnaires.

Plusieurs auteurs ont étudié le comportement de l'estimateur de Hill sous certaines formes de dépendance. Sous la condition de mélange fort, Hsing [58] et Rootzén et al. [80] ont étudié les propriétés de cet estimateur dans le cas d'observations provenant d'une suite stationnaire. Resnick et Stărică [79] ont donné une condition suffisante pour la consistance de l'estimateur dans le cas d'une suite m-dépendante et démontrent dans [78], la normalité asymptotique pour une suite stationnaire α- mélangeante, provenant d'un processus moyenne mobile d'ordre infini ($MA(\infty)$), ainsi que pour un processus autoregressif d'ordre p ($AR(p)$).

En utilisant la variation régulière des résidus et non des innovations, Ling et Peng [63] ont aussi obtenu la normalité asymptotique de l'estimateur de Hill.

3.2 Consistance

Sous mélange fort, Hsing [58] a obtenu la consistance de l'estimateur de Hill, les conditions imposées par Hsing ont été assouplies par Brito [13]. En utilisant la convergence des mesures empiriques, Resnick et Stărică [77] ont montré aussi la convergence de cet estimateur.

On rappelle que le coefficient de mélange fort est défini par :

$$\alpha_{n,l}(Y_i) = \sup_p \left\{ |P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_1^p, B \in \mathcal{F}_{p+l}^n, 1 \leq p \leq n-l \right\}$$

avec $\mathcal{F}_i^j$ est la $\sigma-$ algèbre générée par $(Y_p : i \leq p \leq j)$.

Soient k une suite intermédiaire et $(X_n)_n$ une suite stationnaire de variables aléatoires de distribution F.

On définit Y_{ni} et I_{ni} deux fonctionnelles de X_i par :

$$Y_{ni} = \left(\log X_i - \log b(\frac{n}{k})\right)_+$$

et

$$I_{ni} = 1_{\left(\log X_i - \log b(\frac{n}{\rho k}) > \epsilon\right)}, \quad \epsilon \in \mathbb{R}$$

ρ est dans un intervalle contenant 1 et $b(t) = F^{\leftarrow}(1 - \frac{1}{t}), t > 1$.

Notons

$$\widehat{\xi}_k^{H+} = \frac{1}{k} \sum_{i=1}^{n} \left(\log X_i - \log b(\frac{n}{k})\right)_+$$

Théorème 3.1. *Hsing [58]*

Soit $(X_n)_n$ une suite stationnaire de variables aléatoires de distribution F vérifiant (1.9) *et supposons qu'il existe une suite d'entiers positifs $(p_n)_n$ telle que $p_n = o(n)$.*

Si

1. $r_n \alpha_{n,p_n}(Y_{ni}) \to 0$ *et* $r_n \alpha_{n,p_n}(I_{ni}) \to 0$ *avec* $r_n = [\frac{n}{p_n}]$
2. $$\left\{ \begin{array}{l} \frac{1}{k} \sum_{i=1}^{r_n} E(|S_{nj}| 1_{|S_{nj}| > k}) \to 0 \\ \frac{1}{k^2} \sum_{i=1}^{r_n} E(S_{nj}^2 1_{|S_{nj}| \leq k}) \to 0 \end{array} \right. \tag{3.1}$$

avec $S_{nj} = \sum_{i=(j-1)p_n+1}^{jp_n} Y_{ni}$ *et* $S_{nj} = \sum_{i=(j-1)p_n+1}^{jp_n} I_{ni}$

Alors

$$\widehat{\xi}_k^{H+} \xrightarrow{P} \frac{1}{\beta}$$

et

$$\widehat{\xi}_k^{H} \xrightarrow{P} \frac{1}{\beta}$$

Brito [13] montre qu'en choisissant la suite $(p_n)_n$ telle que $\frac{p_n}{k} \to 0$, les conditions du théorème précédent sont aussi satisfaites.

Corollaire 3.1. *(Brito [13])*
Soit $(X_n)_n$ *une suite stationnaire de variables aléatoires de distribution* F *vérifiant* (1.9) *et supposons qu'il existe une suite d'entiers positifs* $(p_n)_n$ *telle que :*

1. $p_n = o(n)$ *et* $\frac{p_n}{k} \to 0$
2. $r_n \alpha_{n,p_n}(I_{ni}) \to 0$ *et* $r_n \alpha_{n,p_n}(Y_{ni}) \to 0$, *avec* $r_n = [\frac{n}{p_n}]$

Alors

$$\widehat{\xi}_k^{H} \xrightarrow{P} \frac{1}{\beta}$$

Preuve : Pour montrer le corollaire précèdent, il suffit de vérifier la condition (3.1) du théorème de Hsing.
On pose $T_{ni} = Y_{ni}$ et $T_{ni} = I_{ni}$.
On a

$$\frac{1}{k}\sum_{i=1}^{r_n} E(|S_{nj}| 1_{|S_{nj}|>k}) = \frac{r_n}{k} E(|S_{n1}| 1_{|S_{n1}|>k}) \leq \frac{r_n}{k^2} E(S_{n1}^2)$$

$$\frac{1}{k^2}\sum_{i=1}^{r_n} E(S_{nj}^2 1_{|S_{nj}|\leq k}) = \frac{r_n}{k^2} E(S_{n1}^2 1_{|S_{n1}|\leq k}) \leq \frac{r_n}{k^2} E(S_{n1}^2)$$

D'autre part,

$$\frac{r_n}{k^2} E(S_{n1}^2) = \frac{r_n}{k^2} E\left(\left(\sum_{i=1}^{p_n} T_{ni}\right)^2\right) \leq \frac{r_n}{k^2} p_n^2 E(T_{n1}^2)$$

Pour terminer la preuve, il suffit de montrer que :

$$\frac{r_n}{k^2} p_n^2 E(T_{n1}^2) \to 0$$

pour $T_{ni} = I_{ni}$, on a

$$E(T_{n1}^2) = E(I_{n1}) = \mathcal{O}\left(\frac{k}{n}\right)$$

et pour $T_{ni} = Y_{ni}$, on obtient

$$E(T_{n1}^2) = E(Y_{n1}^2) = \mathcal{O}\left(\frac{k}{n}\right) \quad \text{(voir Hsing [58])}$$

Donc

$$\frac{r_n}{k^2} p_n^2 E(T_{n1}^2) \leq \frac{\left[\frac{n}{p_n}\right]}{k^2} p_n^2 \mathcal{O}\left(\frac{k}{n}\right) = \mathcal{O}\left(\frac{p_n}{k}\right),$$

Ce qui donne $\frac{r_n}{k^2} p_n^2 E(T_{n1}^2) \to 0$.

L'approche de Resnick est basée sur la notion de mesures empiriques, qu'on développe d'une façon succincte dans ce qui suit.
Soient $M_+(E)$ l'espace des mesures de radon positives sur $E = (0, \infty)$ et $C_K^+(E)$ l'espace des fonctions de $E = (0, \infty)$, continues, non négatives avec un support compact.

Définition 3.1. *On dit que $\mu_n \in M_+(E)$ converge vaguement vers $\mu_0 \in M_+(E)$ si*

$$\lim_{n \to \infty} \mu_n(f) = \mu_0(f), \ \forall f \in C_K^+(E).$$

On note $\mu_n \Rightarrow \mu_0$

Notons par δ la σ-algèbre Borélienne de $E = (0, \infty]$. Soit une mesure μ définie par :

$$\mu : \delta \to \mathbb{R}_+, \quad \mu((x, \infty]), \quad x > 0.$$

Pour une suite intermédiaire k, on définit la mesure empirique

$$\mu_{X,n} = \frac{1}{k} \sum_{i=1}^{n} \epsilon_{X_i / b(\frac{n}{k})}$$

où ϵ_x est la mesure de Dirac au point x ;

$$\epsilon_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \in A^c \end{cases}$$

On définit aussi

$$\widehat{\mu}_{X,n} = \frac{1}{k} \sum_{i=1}^{n} \epsilon_{X_i/X_{(k)}}$$

Pour démontrer la consistance de $\widehat{\xi}_k^H$ défini par (2.4), il suffit de montrer la convergence en probabilité de $\mu_{X,n}$ vers μ dans l'espace $M_+(E)$. Cette méthode a été utilisée aussi par Mason [69] dans le cas des variables *i.i.d.*

Théorème 3.2. *Soit* $(X_n)_n$ *une suite de variables aléatoires de même distribution* F *vérifiant* (1.9) *et* k *une suite intermédiaire. Si*

$$\mu_{X,n} \xrightarrow{P} \mu \text{ dans } M_+(E) \tag{3.2}$$

Alors

$$\widehat{\xi}_k^H \xrightarrow{P} \xi = \frac{1}{\beta}$$

et

$$\widehat{\xi}_k^{H+} \xrightarrow{P} \xi = \frac{1}{\beta}$$

L'idée de la démonstration repose sur le fait que si (3.2) est satisfaite, alors $\frac{X_{(k)}}{b(\frac{n}{k})} \xrightarrow{P} 1$, ce qui implique aussi que $\widehat{\mu}_{X,n} \xrightarrow{P} \mu$.

D'autre part, remarquons que

$$\widehat{\xi}_k^{H+} = \int_1^\infty \log(y) \mu_{X,n}(dy)$$

et

$$\widehat{\xi}_k^{H} = \int_1^\infty \log(y) \widehat{\mu}_{X,n}(dy)$$

En utilisant une intégration par parties, on obtient

$$\widehat{\xi}_k^{H} = \int_1^\infty \widehat{\mu}_{X,n}(u, \infty](\frac{du}{u}) \xrightarrow{P} \xi$$

De la même manière, on démontre que $\widehat{\xi}_k^{H+} \xrightarrow{P} \xi = \frac{1}{\beta}$.

Resnick et Stărică [77] ont appliqué ce dernier résultat aux processus linéaires et pour des processus autorégressifs d'ordre p avec des innovations à variation régulière.

Pour des processus stationnaires qui peuvent être tronqués et approximés par des suites m-dépendantes, Resnick et Stărică [79] ont établi le résultat suivant.

Proposition 3.1. *Soient* $\left(X_{n,i}^{(m)}\right)_{i\geq 1}$ *et* $(X_{n,i})_{i\geq 1}$ *deux suites stationnaires d'éléments de* E *(*$n \geq 1, m \geq 1$*), avec* $\left(X_{n,i}^{(m)}\right)_{i\geq 1}$ $m-$*dépendante.*
Si

1. *il existe une mesure de Radon* $v^{(m)}$ *dans* E*, vérifiant pour tout* $m \geq 1$
$$\frac{n}{k}P(X_{n,i}^{(m)} \in .) \Rightarrow v^{(m)}$$
2. $v^{(m)} \Rightarrow v$
3. $\forall \varepsilon > 0, \quad \lim_{m\to\infty} \limsup_{n\to\infty} \frac{n}{k} P\left(\left|X_{n,1}^{(m)} - X_{n,1}\right| > \varepsilon\right) = 0$

Alors
$$\frac{1}{k}\sum_{i=1}^{n} \epsilon_{X_{n,i}} \Rightarrow v, \quad \text{dans } M_+(E)$$
Si de plus, pour une suite stationnaire $(X_n)_n$ *on a :* $X_{n,i} = \dfrac{X_i}{b_n}$ *où* $b_n \to \infty$ *et* v *vérifie* $\int_1^\infty \log(u)v(du) < \infty$*, alors*
$$\widehat{\xi}_k^H \xrightarrow{P} \int_1^\infty \log(u)v(du)$$

La démonstration de la proposition précédente repose sur le lemme suivant.

Lemme 3.1. *Pour chaque* $n \geq 1$*, considérons une suite stationnaire* $(X_{n,i})_i$ *d'éléments de* E*. Pour une suite intermédiaire* k*, on suppose que :*

1. $\forall f \in C_K^+(E)$,
$$\lim_{n\to\infty} \frac{n}{k^2}\sum_{j=2}^{k} E(f(X_{n,1})f(X_{n,j})) = 0 \tag{3.3}$$
2. *Pour toute suite* $(l_n)_n$ *telle que* $l_n \to \infty$, $\dfrac{l_n}{k} \to 0$ *et pour* $f \in C_K^+(E)$*, on a*
$$\lim_{n\to\infty} E\left(\prod_{j=1}^{[\frac{n}{k}]} \exp\left(-\frac{1}{k}\sum_{i\in I_j} f(X_{n,i})\right)\right) - \prod_{j=1}^{[\frac{n}{k}]} E\left(\exp\left(-\frac{1}{k}\sum_{i\in I_j} f(X_{n,i})\right)\right) = 0 \tag{3.4}$$
où $I_1 = [1, k - l_n]$, $I_2 = [k+1, 2k - l_n], ..., I_{[\frac{n}{k}]} = [([\frac{n}{k}] - 1)k, [\frac{n}{k}]k - l_n]$

3. $\frac{n}{k}P(X_{n,1} \in .) \Longrightarrow v \qquad avec\ v(\{x\}) = 0, \forall x \in (0, \infty]$

Alors

$$\frac{1}{k}\sum_{i=1}^{n} \epsilon_{X_{n,i}} \Longrightarrow v \qquad dans\ M_+(E)$$

Preuve de la proposition 3.1 :

On montre d'abord qu'en utilisant le lemme 3.1, on a :

$$\frac{1}{k}\sum_{i=1}^{n} \epsilon_{X_{n,i}^{(m)}} \Longrightarrow v, \quad \forall m \geq 1 \tag{3.5}$$

Pour cela, il suffit de vérifier les conditions (3.3) et (3.4).

La condition (3.4) est satisfaite car pour $l(n) > m$, $(\sum_{i \in I_j} f(X_{n,i}^{(m)}))_{j=1,\ldots,p}$ sont des variables aléatoires indépendantes.

Il nous reste à montrer la limite (3.3).

On a

$$\frac{n}{k^2}\sum_{j=2}^{k} P\left(X_{n,1}^{(m)} > x, X_{n,j}^{(m)} > y\right) \leq$$

$$\begin{aligned}
&\leq \frac{n}{k^2}\left(\sum_{j=2}^{m} P\left(X_{n,1}^{(m)} > x, X_{n,j}^{(m)} > y\right) + \sum_{j=m+1}^{k} P\left(X_{n,1}^{(m)} > x, X_{n,j}^{(m)} > y\right)\right) \\
&\leq \frac{n}{k^2}\sum_{j=2}^{m} P\left(X_{n,j}^{(m)} > y\right) + \frac{n}{k}P\left(X_{n,1}^{(m)} > x\right)P\left(X_{n,1}^{(m)} > y\right) \\
&= \frac{(m-1)n}{k^2}P\left(X_{n,1}^{(m)} > y\right) + \frac{k}{n}\frac{n^2}{k^2}P\left(X_{n,1}^{(m)} > x\right)P\left(X_{n,1}^{(m)} > y\right) \\
&= \frac{(m-1)}{k}\left(v^{(m)}\left((y, \infty]\right) + o(1)\right) + \frac{k}{n}\left(v^{(m)}\left((x, \infty]\right) v^{(m)}\left((y, \infty]\right) + o(1)\right)
\end{aligned}$$

Donc

$$\lim_{n\to\infty} \frac{n}{k^2}\sum_{j=2}^{k} P\left(X_{n,1}^{(m)} > x, X_{n,j}^{(m)} > y\right) = 0$$

Ce qui termine la preuve de (3.5).

D'autre part, on a :

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left(d\left(\frac{1}{k}\sum_{i=1}^{n} \epsilon_{X_{n,i}}, \frac{1}{k}\sum_{i=1}^{n} \epsilon_{X_{n,i}^{(m)}}\right) > \varepsilon\right) = 0,$$

où ε est positif et d une distance sur $M_+(E)$.
D'après le théorème 4.2 de Billingsley [11], on déduit que

$$\frac{1}{k}\sum_{i=1}^{n}\epsilon_{X_{n,i}} \Longrightarrow v$$

Le résultat de la proposition 3.1 découle du théorème 3.2 qui démontre que la convergence de la mesure empirique implique la consistance de l'estimateur de Hill.

3.3 Normalité asymptotique

Sous la condition de mélange fort, la normalité asymptotique de l'estimateur de Hill a été établie par Hsing [58] et aussi par Rootzén [80] .

3.3.1 Résultats de Rootzén

Soit $(Y_n)_n$ une suite stationnaire de variables aléatoires α- mélangeante de fonction de répartition F vérifiant :

$$\lim_{t\to\infty}\frac{\overline{F}(t+x)}{\overline{F}(t)} = \exp(-\beta x) \quad \text{pour tout } x \geq 0 \tag{3.6}$$

et de densité f vérifiant la condition de Von Mises :

$$\lim_{t\to\infty}\frac{tf(t)}{\overline{F}(t)} = \beta \tag{3.7}$$

Si la condition (3.7) est vérifiée alors $\overline{F} \in RV_{-\beta}$. Rootzén et al. [80] ont étudié la normalité asymptotique de l'estimateur du paramètre β défini par :

$$\beta_n^* = \frac{1}{k}\sum_{i=1}^{k} Y_{(i)} - Y_{(k)}$$

Soit une suite d'entiers $(p_n)_n$ telle que $p_n \to \infty$ et

$$p_n\left(\alpha_{n,l_n} + \frac{l_n}{n}\right) \to 0 \tag{3.8}$$

où α_{n,l_n} est le coefficient de mélange fort.

Pour une fonction continue à droite $\psi(x) = a\psi_1(x) + b\psi_2(x)$, avec $\psi_1(x) = x1_{\{x\geq 0\}}$, $\psi_2(x) = 1_{\{x\geq 0\}}$, $a = 1$ et $b = -\dfrac{1}{\beta}$, on considère

$$\lambda_n = \frac{p_n}{k} var\left(\sum_{j=1}^{r_n} \psi((Y_j - u_n)_+)\right)$$

où

$$r_n = \left[\frac{n}{p_n}\right] \text{ avec } \frac{r_n}{n} \to 0 \tag{3.9}$$

$(u_n)_n$ est choisie de sorte que

$$\lim_{n\to\infty} \frac{\overline{F}(u_n)}{\frac{k}{n}} = 1 \tag{3.10}$$

Notons

$$\begin{aligned} \lambda_n^{(1)} &= \frac{n}{kr_n} var\left(\sum_{j=1}^{r_n} \psi_1((Y_j - u_n)_+)\right) \\ \lambda_n^{(2)} &= \frac{n}{kr_n} var\left(\sum_{j=1}^{r_n} \psi_2((Y_j - u_n)_+)\right) \end{aligned}$$

Soit la suite $(m_n)_n$, telle que $m_n = n - p_n r_n$.

On définit les intervalles $J_1, J_2, ..., J_{p_n}$ tels que :

$$J_i = \begin{cases} \{(i-1)r_n + i, (i-1)r_n + i + 1, ..., ir_n + i - 1\} & \text{pour } 1 \leq i \leq m_n \\ \{(i-1)r_n + m_n + 1, (i-1)r_n + m_n + 2, ..., ir_n + m_n\} & \text{pour } m_n \leq i \leq p_n \end{cases}$$

Notons J_i^* les premiers entiers $(r_n - l_n)$ dans J_i, $1 \leq i \leq p_n$ et posons

$$\begin{cases} Z_i = (\lambda_n k)^{-\frac{1}{2}} \sum\limits_{j\in J_i} \psi((Y_j - u_n)_+) \quad , \\ U_i = (\lambda_n k)^{-\frac{1}{2}} \sum\limits_{j\in J_i^*} \psi((Y_j - u_n)_+) \quad , \\ V_i = Z_i - U_i, \\ W_i = (\lambda_n k)^{-\frac{1}{2}} \left(\psi((Y_{i(r_n+1)} - u_n)_+)\right) \quad , \quad 1 \leq i \leq m_n \end{cases} \tag{3.11}$$

On partage $(\beta_n^*)_n$ en deux suites $(Z_i)_i$ et $(W_i)_i$ telles que $Z_1, Z_2, .., Z_n$ et $W_1, W_2, .., W_n$ soient indépendantes. On obtient :

$$\left(\frac{k}{\lambda_n}\right)^{\frac{1}{2}} \left(\hat{\beta}_n - E(\hat{\beta}_n)\right) = \sum_{i=1}^{p_n} (Z_i - E(Z_i)) + \sum_{i=1}^{m_n} (W_i - E(W_i))$$

avec $\hat{\beta}_n = \frac{1}{k}\sum_{i=1}^{n} \psi(Y_i - u_n)_+$ et $E(\hat{\beta}_n) = \frac{n}{k}E(\psi(Y_1 - u_n)_+)$

Avant d'énoncer le théorème de Rootzén, on donne le lemme suivant qui établit la convergence en probabilité du deuxième terme de la dernière équation.

Lemme 3.2. *Soit $(X_n)_n$ une suite stationnaire et α-mélangeante vérifiant les conditions* (3.8), (3.10) *et*

$$\lim_{n\to\infty} p_n(varV_1 + varW_1) = 0$$

Alors

$$\sum_{i=1}^{p_n} (V_i - E(V_i)) \xrightarrow{P} 0$$

$$\sum_{i=1}^{m_n} (W_i - E(W_i)) \xrightarrow{P} 0$$

Théorème 3.3. *(Rootzén et al. [80])*

Soit $(Y_n)_n$ une suite stationnaire, α- mélangeante, de distribution F vérifiant les conditions (3.6) *et* (3.7). *Si de plus les conditions suivantes sont vérifiées :*

1. $\lim_{n\to\infty} p_n(\alpha_{n,l_n} + \frac{l_n}{n}) = 0, \ \lim_{n\to\infty} \frac{k}{\lambda_n} = \infty$
2. $\lambda_n^{(1)} \leq \lambda_n \quad$ *et* $\quad \lambda_n^{(2)} \leq 2\lambda_n$
3. $E(\psi^2(Y_1 - u_n)_+) < \infty$ *et*
$$\forall \varepsilon > 0, \int_0^\infty e^{(\varepsilon-\beta)x} d\psi(x) < \infty$$
4. $\lim_{n\to\infty} p_n(varV_{n,1} + varW_{n,1}) = 0$
5. *Pour $0 < \varepsilon < 1$, soit $(w_n)_n$ telle que*
$$\lim_{n\to\infty} \frac{1}{\sqrt{k\lambda_n}} r_n \psi(w_n) = 0,$$
$$\lim_{n\to\infty} \frac{1}{\lambda_n} r_n \int_{w_n}^\infty e^{(\varepsilon-\beta)x} d\psi^2(x) = 0.$$
6. *Pour $I_n = [u_n, z_n)$ si ($z_n > u_n$ et $I_n = [z_n, u_n)$ dans l'autre cas)*
$$\frac{1}{\lambda_n}\frac{p_n}{k} var\left[\sum_{i=1}^{r_n} 1_{(Y_i \in I_n)}\right] \to 0 \ \text{ quand } n \to \infty$$
avec z_n non-aléatoire et $\sqrt{\frac{k}{\lambda_n}}\,(z_n - u_n)$ est bornée.

Alors

$$\left(\frac{k}{\lambda_n}\right)^{\frac{1}{2}} (\beta_n^* - \beta_n) \xrightarrow{d} N(0,1) \tag{3.12}$$

où $\beta_n = \frac{n}{k} E(Y_1 - u_n)_+$

3.3.2 Résultats de Resnick et Stărică

Pour des variables provenant d'un processus linéaire fortement mélangeant, l'approche de Rootzén a été utilisée par Resnick et Stărică [78] pour étudier la normalité asymptotique de l'estimateur de Hill issu d'une suite $Y_n = logX_n$ avec $(X_n)_n$ fortement mélangeante. Soit $(X_n)_n$ une suite de variables aléatoires fortement mélangeante de même distribution F vérifiant :

$$\overline{F}_{|X|}(x) = \overline{F}_X(x) \ + F_X(-x) \sim x^{-\beta} L_1(x)$$

On note

$$b_{|X|}(t) \quad = \quad \left(\frac{1}{1 - F_{|X|}} \right)^{\leftarrow} (t) = F_{|X|}^{\leftarrow}(1 - \frac{1}{t}), \quad t > 1$$

Avant d'énoncer les théorèmes de Resnick, on donne les conditions de mélange pour les processus linéaires $MA(\infty)$ et $AR(p)$.

3.3.2.1 Cas de processus moyenne mobile d'ordre infini

a. Conditions de mélange fort

Les conditions donnant le mélange fort d'un processus linéaire ont été données par Chanda [14]. Gorodetskii [50] a amélioré le résultat de Chanda en ajoutant des hypothèses supplémentaires.

Soit $(X_t)_t$ un processus réel défini par

$$X_t = \sum_{j=0}^{+\infty} c_j \varepsilon_{t-j}, \quad t \in \mathbb{Z} \tag{3.13}$$

avec $(\varepsilon_t)_t$ une suite de variables aléatoires réelles indépendantes de distribution G_i et de densité G_i' et $(c_n)_n$ une suite de réels telle que $c_0 \neq 0$.

On pose $S_i(\delta) = \sum_{j=i}^{+\infty} |c_j|^\delta$

$$\psi(k) = \begin{cases} \sum\limits_{i=k}^{+\infty} (S_i(\delta))^{\dfrac{1}{1+\delta}} & \text{si } \delta < 2 \\ \sum\limits_{i=k}^{+\infty} \max \left\{ (S_i(\delta))^{\dfrac{1}{1+\delta}}, \sqrt{S_i(2) \left| \log S_i(2) \right|} \right\} & \text{si } \delta \geq 2 \end{cases}$$

Théorème 3.4. *(Gorodetskii [50])*
Sous les notations précédentes supposons que :
C1 :
$$\int_{-\infty}^{+\infty} |G_i'(x) - G_i'(x+y)|\, dx \leq c_1\, |y| \tag{3.14}$$
C2 :
$$\begin{cases} E(|\varepsilon_i|^{\delta}) \leq C_2 < +\infty, & \textit{pour un } \delta > 0 \\ E(\varepsilon_i) = 0 \quad\quad \textit{si } 1 \leq \delta < 2 \\ E(\varepsilon_i) = 0 \textit{ et } var(\varepsilon_i) = 1 \;\; \textit{si } \delta \geq 2, \end{cases}$$
C3 :
$$0 \neq \sum_{j=0}^{+\infty} c_j z^j < \infty,\ \forall\, |z| \leq 1$$
C4 :
$$\psi(0) < \infty$$
Alors le processus défini par (3.13) *satisfait la condition du mélange fort.*

Dans la partie suivante, on donne le résultat sur la normalité asymptotique de l'estimateur de Hill appliqué au processus défini par (3.13).

b. Normalité asymptotique

Soit le processus défini par (3.13), avec ε_i de même distribution G vérifiant :
$$\bar{G}(x) \sim r x^{-\beta} L_2(x), \quad G(-x) \sim s x^{-\beta} L_2(x) \tag{3.15}$$
avec $\beta > 0, r, s \geq 0, r + s = 1$ et L_2 est une fonction à variation lente.
On suppose que la suite $(c_j) \in \mathbb{R}^{\infty}$ contient au moins un élément non nul et $\exists A > 0$ et $u > 1$ tels que :
$$|c_j| < A u^{-j}, \quad j \in \mathbb{N} \tag{3.16}$$
La condition (3.16) qui est généralement vraie pour un processus linéaire causal, est vérifiée s'il existe z_0, $|z_0| > 1$ telle que
$$\sum_{j=0}^{\infty} c_j z^j < \infty, |z| < |z_0|$$
On rajoute la condition
$$\sum_{j=0}^{\infty} |\ c_j\ |^{\delta} < \infty,\ \textit{pour } 0 < \delta < \beta \wedge 1 \tag{3.17}$$

Ce qui assure la finitude du processus linéaire (3.13), i.e., $\sum_{j=0}^{\infty} c_j \varepsilon_{t-j} < \infty$ et que

$$\lim_{x\to\infty} \frac{P(|X_1| > x)}{P(|\varepsilon_1| > x)} = \sum_{j=0}^{\infty} |c_j|^\beta$$

Ceci implique que la queue de la distribution X_t est aussi à variation régulière (voir Cline [16]).

On note

$$\lambda_n = \frac{n}{kr_n} var\left(\sum_{i=1}^{r_n}\left(\left(\log\frac{|X_i|}{b(\frac{n}{k})}\right)_+ - \frac{1}{\beta}1_{\left(|X_i|>b(\frac{n}{k})\right)}\right)\right)$$

avec $r_n = o\left(\dfrac{n}{k}\right)$.

Lemme 3.3. *On suppose que* $(\varepsilon_t)_t$ *satisfait la condition* (3.15) *et que*

$$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} |c_k|^\beta \wedge |c_{j+k}|^\beta \log\left(\frac{|c_{k+j}| \vee |c_k|}{|c_{k+j}| \wedge |c_k|}\right) < \infty \tag{3.18}$$

Alors

1.

$$\frac{n}{k}\sum_{j=1}^{r_n} P\left(|X_1| > b\left(\tfrac{n}{k}\right)x, |X_{j+1}| > b\left(\tfrac{n}{k}\right)y\right) \to \frac{\sum\limits_{j=1}^{\infty}\sum\limits_{k=0}^{\infty} |c_k|^\beta x^{-\beta} \wedge |c_{j+k}|^\beta y^{-\beta}}{\sum\limits_{k=0}^{\infty} |c_k|^\beta}$$

localement uniforme pour x *et* y *dans* $(0,\infty)$

2.

$$\frac{n}{k}\sum_{j=1}^{r_n}\int_1^\infty\int_1^\infty P\left(|X_1| > b\left(\tfrac{n}{k}\right)x, |X_{j+1}| > b\left(\tfrac{n}{k}\right)y\right)\frac{dx}{x}\frac{dy}{y} \to$$

$$\frac{1}{\beta}\frac{\sum\limits_{j=1}^{\infty}\sum\limits_{k=0}^{\infty} |c_k|^\beta \wedge |c_{j+k}|^\beta \left(2+\beta\log\left(\frac{|c_{k+j}|\vee|c_k|}{|c_{j+k}|\wedge|c_k|}\right)\right)}{\sum\limits_{k=0}^{\infty} |c_k|^\beta}$$

3.

$$\frac{n}{k}\sum_{j=1}^{r_n}\int_1^\infty \left(P\left(|X_1| > b\left(\tfrac{n}{k}\right)x, |X_{j+1}| > b\left(\tfrac{n}{k}\right)y\right) + P\left(|X_1| > b\left(\tfrac{n}{k}\right)y, |X_{j+1}| > b\left(\tfrac{n}{k}\right)x\right)\right)\frac{dy}{y}$$

$$\to \frac{1}{\beta^2}\frac{\sum\limits_{j=1}^{\infty}\sum\limits_{k=0}^{\infty} |c_k|^\beta \wedge |c_{j+k}|^\beta \left(2+\beta\log\left(\frac{|c_{k+j}|\vee|c_k|}{|c_{j+k}|\wedge|c_k|}\right)\right)}{\sum\limits_{k=0}^{\infty} |c_k|^\beta}$$

localement uniforme sur $(0,\infty)$.

La limite de λ_n est une conséquence directe du lemme précédent (voir Resnick et Stărică [78]). On en déduit que :

$$\lambda_n \to \lambda = \frac{1}{\beta^2}\left(1 + 2\frac{\sum_{j=1}^{\infty}\sum_{k=0}^{\infty}|c_k|^{\beta}\wedge|c_{k+j}|^{\beta}}{\sum_{j=0}^{\infty}|c_k|^{\beta}}\right), \text{ quand } n \to \infty \tag{3.19}$$

la limite de λ_n est finie et ne dépend seulement des coefficients c_j.

Pour vérifier les conditions du mélange fort, on suppose que la densité de G' vérifie la condition (3.14) du théorème 3.4 et $\exists d < 1$ tel que

$$E(\varepsilon_1^d) < \infty \tag{3.20}$$

Si de plus (3.16) est satisfaite alors le processus (3.13) est fortement mélangeant.

Théorème 3.5. *Soit* $(X_t)_t$ *un processus linéaire défini par* (3.13)*, vérifiant les conditions* (3.7)*,* (3.14)*,* (3.15)*,*(3.16)*,* (3.18) *et* (3.20)*. Si la suite intermédiaire k vérifie*

$$\begin{cases} \liminf\limits_{n\to\infty} \dfrac{n}{k^{\frac{3}{2}}} > 0 \\ \quad ou \\ \limsup\limits_{n\to\infty} \dfrac{n}{k^{\frac{3}{2}}} < \infty \end{cases} \tag{3.21}$$

alors

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log\frac{|X|_{(i)}}{|X|_{(k+1)}} - \int_1^{\infty}\frac{n}{k}P\left(\frac{|X_1|}{b\left(\frac{n}{k}\right)} > x\right)\frac{dx}{x}\right) \to N(0,\lambda).$$

Si de plus $\overline{F}_{|X|} \in 2RV(-\beta,\rho)$ *et sa fonction auxiliaire* g *est telle que*

$$\sqrt{n}g\left(b_{|X|}\left(\frac{n}{k}\right)\right) \to 0, \quad \textit{quand } n \to \infty \tag{3.22}$$

alors

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log\frac{|X|_{(i)}}{|X|_{(k+1)}} - \frac{1}{\beta}\right) \to N(0,\lambda)$$

La condition de régularité de second ordre sur F implique la condition (3.6) imposée par Rootzén (voir Rootzén et al. [80], Appendix. p44) et les conditions (3.21) sur la suite k nous permettent de montrer l'existence des suites $(r_n)_n$ et $(w_n)_n$ définies dans le théorème 3.3.

3.3.2.2 Cas de processus autorégressif d'ordre p

a. Conditions de mélange fort

Athreya et Pantula [2, 3] ont étudié les propriétés de mélange des processus autorégressifs. Ils établissent d'abord la propriété de mélange fort pour une classe de chaines récurrentes au sens de Harris et, en utilisant ce résultat, ils donnent un ensemble de conditions suffisantes pour garantir la propriété du mélange fort pour les processus autorégressifs. Athreya et Pantula commencent par démontrer le théorème suivant.

Théorème 3.6. *Soit $(X_n)_n$ un processus autorégressif d'ordre 1 donné par :*

$$X_n = \phi X_{n-1} + \varepsilon_n, \quad n \in \mathbb{N}^*$$

avec $|\phi| \leq 1$ et $(\varepsilon_n)_n$ sont des variables aléatoires i.i.d indépendantes de X_0 ; on suppose que :

1. $E(\log|\varepsilon_1|^+) < \infty$.
2. *Pour un certain $n_0 \geq 1$, $\phi_{n_0} = \sum_{j=1}^{n_0} \phi^j \varepsilon_j$ a une composante absolument continue non triviale.*

Alors pour toute distribution initiale de X_0, $(X_n)_n$ est fortement mélangeant.

Ils ont généralisé ce dernier théorème à un processus $AR(p)$ défini par :

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t \ , \quad t \in \mathbb{N} \tag{3.23}$$

Théorème 3.7. *Un processus défini par* (3.23) *est fortement mélangeant s'il vérifie,*

1. $E(\log|\varepsilon_1|^+) < \infty$.
2. *la distribution de ε_1 a une composante absolument continue non dégénérée.*
3. *X_0 est indépendante des ε_i.*
4. *ε_i sont des variables aléatoires i.i.d.*
5. *les racines de l'équation caractéristique $z^p - \phi_1 z^{p-1} - ... - \phi_p = 0$ sont de module strictement inférieure à 1.*

Dans ce qui suit, on donne la normalité asymptotique de l'estimateur de Hill pour des observations issues d'un processus $AR(p)$.

b. Normalité asymptotique

Soit un processus défini par (3.23), tel que les variables aléatoires ε_i sont de même distribution vérifiant la condition (3.15). On suppose que

$$\phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i \neq 0 \ , \ |z| \leq 1$$

où z est l'opérateur de translation retard, alors X_t défini par (3.23) admet une solution stationnaire, de la forme

$$X_t = \sum_{j=1}^{\infty} c_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}$$

avec

$$\sum_{j=1}^{\infty} c_j z^i = \frac{1}{\phi(z)}, \quad |z| \leq 1$$

On suppose l'existence de la suite $\hat{\phi}^{(n)} = (\hat{\phi}_1^{(n)}, ..., \hat{\phi}_p^{(n)}), n \geq 1$, des estimateurs des coefficients de la régression vérifiant :

$$d(n)\left(\hat{\phi}^{(n)} - \phi\right) \to S \tag{3.24}$$

où S est une variable aléatoire non dégénérée et $d(n) \to \infty$.

Pour cette suite d'estimateurs on définit la suite des résiduels estimés $\hat{\varepsilon}_t^{(n)}$ par

$$\hat{\varepsilon}_t^{(n)} = X_t - \sum_{j=1}^{p} \hat{\phi}_j^{(n)} X_{t-j}, \qquad t = p+1, ..., n$$

Alors

$$\varepsilon_t - \hat{\varepsilon}_t^{(n)} = \sum_{i=1}^{p} (\hat{\phi}_i^{(n)} - \phi_i) X_{t-i}$$

Resnick et Stărică [78] ont appliqué l'estimateur de Hill pour l'échantillon $\left|\hat{\varepsilon}_1^{(n)}\right|, \left|\hat{\varepsilon}_2^{(n)}\right|, ..., \left|\hat{\varepsilon}_n^{(n)}\right|$ et ont obtenu le résultat suivant.

Théorème 3.8. *Supposons que*
- la distribution $\overline{G}_{|\varepsilon|} \in 2RV(-\beta, \rho)$ *et vérifie* (3.15).
- il existe une suite $d(n) \to +\infty$ *et une variable aléatoire non dégénérée* S*, telles que* (3.24) *est satisfaite.*
- la suite intermédiaire k *est choisie de sorte que* (3.22) *soit vérifiée et que*

$$\frac{\sqrt{k} b\left(\frac{n}{\sqrt{k}}\right)}{b\left(\frac{n}{k}\right)} = o(d(n)), \quad pour\ n \to \infty$$

Alors

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k} \log \frac{\left|\hat{\varepsilon}^{(n)}\right|_{(i)}}{\left|\hat{\varepsilon}^{(n)}\right|_{(k+1)}} - \frac{1}{\beta}\right) \xrightarrow{d} N\left(0, \frac{1}{\beta^2}\right)$$

3.3.3 Résultat de Ling et Ping (sous le modèle ARMA)

Une deuxième approche pour l'estimation du paramètre β est celle de Ling et Peng [63] où l'estimateur est appliqué pour les résidus, en supposant la condition de variation régulière de second ordre seulement pour la distribution des innovations. Cette méthode donne les mêmes résultats que ceux de Resnick et Stărică mais avec une variance plus petite.

On considère le processus linéaire $(X_t)_t$ de type $ARMA(p, q)$ défini par :

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_i \ , \quad t \in \mathbb{Z} \tag{3.25}$$

où p, q sont des entiers positifs et $(\varepsilon_t)_t$ une suite de variables aléatoires $i.i.d$ de distribution continue G vérifiant ;

Condition1 :

$$\begin{cases} \lim\limits_{t\to\infty} \dfrac{\overline{G}(tx)}{\overline{G}(t)} = x^{-\beta}, \quad x > 0, \beta > 0 \\ \lim\limits_{t\to\infty} \dfrac{\overline{G}(x)}{\overline{G}(t) + G(-t)} = c \in [0, 1] \end{cases} \tag{3.26}$$

On note $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p$, $\psi(z) = 1 - \psi_1 z - ... - \psi_q z^q$, où z est l'opérateur de translation retard.

Condition 2 : On suppose que les racines de $\phi(z)$ et $\psi(z)$ sont à l'extérieur du cercle unité et qu'il n'y a pas de racines communes. On peut donc représenter (3.25) par

$$X_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j} \,, \qquad t \in \mathbb{Z}$$

avec $\phi^{-1}(z)\psi(z) = \sum_{j=0}^{\infty} c_j z^j$ et $c_j = \mathcal{O}(\rho^j)$ où $\rho \in (0,1)$. Soient $(n+p)$ observations $X_{-p+1}, X_{-p+2}, ..., X_0, X_1, ..., X_n$ et notons par $(\widehat{\phi}_1, ..., \widehat{\phi}_p)$ et $(\widehat{\psi}_1, ..., \widehat{\psi}_q)$ les estimateurs respectifs de $(\phi_1, ..., \phi_p)$ et $(\psi_1, ..., \psi_q)$ avec

$$d_n(\widehat{\phi}_i - \phi_i) = \mathcal{O}_p(1) \;\text{ et }\; d_n(\widehat{\psi}_j - \psi_j) = \mathcal{O}_q(1), \;\; i = 1, ..., p; j = 1, ..., q \tag{3.27}$$

et

$$d_n = \begin{cases} \sqrt{n} & \text{si} \quad \beta > 2 \\ \left(\dfrac{n}{\log n}\right)^{\dfrac{1}{\beta}} & \text{si} \quad]0,2] \end{cases}$$

Les résidus sont donnés par :

$$\widehat{\varepsilon}_t = X_t - \sum_{i=1}^{p} \widehat{\phi}_i X_{t-i} - \sum_{i=1}^{q} \widehat{\psi}_i \widehat{\varepsilon}_{t-i}$$

L'estimateur de Hill appliqué aux résidus est défini par :

$$\widehat{\beta} = \left(\frac{1}{k} \sum_{i=1}^{k} \log \frac{\widehat{\varepsilon}_{(k+1)}}{\widehat{\varepsilon}_{(i)}} \right)^{-1}$$

où $\widehat{\varepsilon}_{(1)} \leq ... \leq \widehat{\varepsilon}_{(n)}$ sont les statistiques d'ordres de $\widehat{\varepsilon}_1, ..., \widehat{\varepsilon}_n$.

Théorème 3.9. *(Ling et Peng [63])*

Supposons que :

1. $\dfrac{k}{n} \to 0$ *et qu'il existe* $l_0, l_1 \in [0,1]$ *tels que* $k \in (n^{l_0}, n^{l_1})$ *pour* n *assez grand.*
2. *les racines de* $\phi(z)$ *et* $\psi(z)$ *sont à l'extérieur du cercle unité et qu'il n'y a pas de racines communes.*
3. *les conditions* (3.26), (3.27) *sont vérifiées.*

4. $\overline{G} \in 2RV(-\beta, \rho)$ *et sa fonction auxiliaire* g *est telle que*

$$\lim_{n\to\infty} \sqrt{k}g(\frac{n}{k}) = \sigma \in (-\infty, \infty).$$

Alors

$$\sqrt{k}(\widehat{\beta} - \beta) \xrightarrow{d} N(-\frac{\sigma\beta^2}{1-\rho}, \beta^2)$$

Dans le cas où $\beta > 2$, la condition (3.27) est vérifiée pour les estimateurs de ϕ et ψ obtenus par la méthode des moindres carrés. Si $\beta \in]0, 2]$, cette condition est vérifiée pour les estimateurs donnés par Davis [23] basés sur les méthodes de Gauss-Newton et de M-estimateurs.

Chapitre 4

Estimateur de Hill sous dépendance faible

4.1 Introduction

Les coefficients de mélange sont souvent difficiles à calculer et certains processus ne vérifient pas la condition de mélange, a titre d'exemple le processus $AR(1)$ défini par :

$$X_n = \varphi X_{n-1} + \varepsilon_n,$$

où $(\varepsilon_n)_n$ sont des innovations indépendantes de même distribution de Bernouilli de paramètre $\frac{1}{2}$. Andrews [1] a montré que si $0 < |\varphi| \leq \frac{1}{2}$, alors le processus $(X_n)_n$ n'est pas fortement mélangeant.

Ce qui a motivé Doukhan et Louhichi [36] à introduire un nouveau concept de dépendance, plus général et qui est moins restrictif que les différentes notions du mélange. De nombreux articles et ouvrages traitent de cette dépendance et de ses applications (voir Dedecker et al. [30], Louhichi [65], Prieur [74, 73], Bardet et al. [4]).

Dans ce chapitre, nous étudions le comportement asymptotique de l'estimateur de Hill, sous cette nouvelle notion de dépendance avec des observations provenant d'un processus linéaire avec des innovations *i.i.d.* L'approche qu'on utilise est identique à celle de Rootzén [80].

4.2 Dépendance faible

Pour un processus stationnaire, la dépendance faible au sens de Doukhan et Louhichi (voir Doukhan et Louhichi [36]) est mesurée en termes de covariance de fonctions du passé et du futur, i.e., $cov(h(\text{"}passé\text{"}), g(\text{"}futur\text{"}))$. Sous certaines conditions sur h et g, cette covariance est assez petite quand la distance entre le "passé" et le "futur" est suffisamment grande.

4.2.1 Définition

Définition 4.1. *(Doukhan et Louhichi [36])*
Soient $\pounds = \bigcup_{p=1}^{+\infty} \pounds^p$, *où* $\pounds^p$ *est la classe de fonctions de* $\mathbb{R}^p$ *dans* $\mathbb{R}$ *et* $\Psi : \pounds \times \pounds \times \mathbb{N}^2 \to \mathbb{R}_+$.
La suite $(X_n)_n$ *est dite* $(\epsilon, \pounds, \Psi)-$*faiblement dépendante si*
$\forall r \in \mathbb{N}, \forall u, v \in \mathbb{N}^*, \forall (h, g) \in \pounds^u \times \pounds^v, \forall i_1 < .... < i_u \leq i_u + r \leq j_1 < .... < j_v.$

$$Cov(h(X_{i_1}, ..., X_{i_u}), g(X_{j_1}, ..., X_{j_v})) \leq \Psi(h, g, u, v)\epsilon_r$$

où $\epsilon = (\epsilon_r)_{r\in\mathbb{N}}$ une suite décroissante vers zéro.

La fonction Ψ est définie par

$$\Psi(h, g, u, v) = c(u, v)\mu\left(Lip(h), Lip(g)\right)$$

où μ est une fonction localement bornée sur $\mathbb{R}_+^2$ et c est une fonction définie sur $\mathbb{N}^{*^2}$. La constante de Lipschitz de h est donnée par :

$$Lip(h) = \sup_{x\neq y} \frac{|h(x) - h(y)|}{\|x - y\|_1}$$

où $\delta_1 = \|x - y\|_1 = \sum_{i=1}^n |x_i - y_i|$.
Selon le choix de la fonction Ψ, on obtient les différents coefficients de dépendance η, κ, λ et ζ, pour plus de détails nous renvoyons à Dedecker et al. [30].
Pour h et g bornées et Lipschitziennes par rapport à la distance δ_1, on définit :
- le coefficient η correspondant à la fonction :

$$\Psi\left(h, g, u, v\right) = u\left\|g\right\|_\infty Lip(h) + v\left\|h\right\|_\infty Lip(g)$$

- le coefficient λ correspondant à la fonction :

$$\Psi(h, g, u, v) = u \|g\|_\infty Lip(h) + v \|h\|_\infty Lip(g) + uvLip(h)Lip(g)$$

Pour la définition des coefficients κ et ζ, on a besoin de la classe de fonctions Lipschitiziennes non nécessairement bornées ; dans ce cas, on suppose que les variables sont L^1-intégrables.

- Le coefficient κ correspond à la fonction :

$$\Psi(h, g, u, v) = uvLip(h)Lip(g)$$

- Le coefficient ζ correspond à la fonction :

$$\Psi(h, g, u, v) = \min(u, v)Lip(h)Lip(g)$$

Pour le coefficient de dépendance θ, on suppose que les fonctions h et g sont respectivement bornées et Lipschitziennes et les variables L^1-intégrables.

-Le coefficient θ correspond à la fonction :

$$\Psi(h, g, u, v) = v \|h\|_\infty Lip(g)$$

Remarque 4.1. (Doukhan et Louhichi [36])
Remplacer $\pounds$ par la classe des fonctions bornées donne lieu aux suites fortement mélangeantes avec $\Psi(h, g, u, v) = 4 \|h\|_\infty \|g\|_\infty$ et $\epsilon_r = \alpha_r$.

4.2.2 Exemples de processus faiblement dépendant

a. Schéma de Bernoulli avec des innovations indépendantes

Doukhan et Louhichi ont donné les propriétés de dépendance faible des schémas de Bernoulli avec des innovations indépendantes.

Définition 4.2. *Soient $(\zeta_t)_t$ une suite i.i.d de variables aléatoires réelles et $H : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}$ une fonction mesurable. La suite $(X_t)_t$ définie par :*

$$X_t = H\left((\zeta_{t-j})_{j \in \mathbb{Z}}\right), \quad t \in \mathbb{Z}$$

est appelée schéma de Bernoulli (Bernoulli shift).

Si

$$\sup_{t\in\mathbb{Z}} E|H\left((\zeta_{t-j})_{j\in\mathbb{Z}}\right) - H\left((\zeta_{t-j}1_{(|j|\leq r)})_{j\in\mathbb{Z}}\right)| \leq \delta_r$$

où δ_r est une suite positive tendant vers zéro, alors le schéma de Bernoulli $(X_t)_t$ est un processus stationnaire, η-faiblement dépendant avec $\eta_r \leq 2\delta_{[\frac{r}{2}]}$.

b. Association

Définition 4.3. *On dit que $(X_n)_n$ est une suite de variables aléatoires associée si*

$$0 \leq cov\left(h(X_1, ..., X_n), g(X_1, ..., X_n)\right) < \infty,$$

pour toute paire de fonctions h, g définies de $\mathbb{R}^n \to \mathbb{R}$, croissantes coordonnée par coordonnée.

- Si $(X_n)_n$ est une suite stationnaire de variables aléatoires réelles associées et centrées, alors elle est $\lambda-$ faiblement dépendante avec $\lambda_r = \mathcal{O}\left(\sup_{t\geq r}|cov(X_0, X_t)|\right)$.
- Si un processus est gaussien associé avec $\lim_{t\to\infty}|cov(X_0, X_t)| = 0$, alors il est $\kappa-$ faiblement dépendant avec : $\kappa_r = \mathcal{O}(\sup_{t\geq r}|cov(X_0, X_t)|)$.

c. Processus moyenne mobile

Soit un processus linéaire causal ou non causal $(X_t)_t$ défini par l'équation

$$X_t = \sum_{j=-\infty}^{\infty} c_j\varepsilon_{t-j}, \quad t \in \mathbb{Z}$$

avec $(c_j)_j \in \mathbb{R}^{\mathbb{Z}}$ et $(\varepsilon_t)_t$ est une suite de variables aléatoires $i.i.d$ centrées.

Le processus $(X_t)_t$ est $\eta-$faiblement dépendant, si

$$c_j = \mathcal{O}\left(|j|^{-\mu}\right) \text{ avec } \mu > \frac{1}{2}, \tag{4.1}$$

de plus, on a $\eta_r = \mathcal{O}\left(\frac{1}{r^{u-\frac{1}{2}}}\right)$ (voir Doukhan et Neumann [38]).

d. Processus $GARCH(p,q)$ et $ARCH(\infty)$

Définition 4.4. *Soient $(\zeta_t)_t$ une suite de variables aléatoires i.i.d et $(X_t)_t$ un processus $GARCH(p,q)$ donné par les relations suivantes*

$$X_t = \rho_t\zeta_t \quad avec \quad \rho_t^2 = a_0 + \sum_{j=1}^{p} a_j X_{t-j}^2 + \sum_{j=1}^{q} b_j \rho_{t-j}^2, \quad t \in \mathbb{Z} \tag{4.2}$$

où $(p,q) \in \mathbb{N}^2, a_0 > 0, a_j \geq 0$ et $b_j \geq 0 \quad$ pour $j \in \mathbb{N}$.

Définition 4.5. *Soient $(\zeta_t)_t$ une suite de variables aléatoires i.i.d et $(X_t)_t$ un processus $ARCH(\infty)$ défini par :*

$$X_t = \rho_t\zeta_t \quad avec \quad \rho_t^2 = b_0 + \sum_{j=1}^{\infty} b_j X_{t-j}^2 \tag{4.3}$$

où $b_j \geq 0$.

On suppose qu'il existe $m > 2$ tel que $E(|\zeta_0|^m) < \infty$ et que la condition de stationnarité $\|\zeta_0\|_m^2 \cdot \sum_{k=1}^{\infty} |b_k| < 1$ est vérifiée.
- S'il existe $C > 0$ et $\mu \in]0,1[$ tels que $\forall j \in \mathbb{N}, 0 \leq b_j \leq C\mu^{-j}$, alors le processus (4.2) est $\eta-$faiblement dépendant avec $\eta_r = O(e^{-c\sqrt{r}}), c > 0$.
- S'il existe $C > 0$ et $\nu > 1$ tels que $\forall k \in \mathbb{N}, 0 \leq b_k \leq Ck^{-\nu}$, alors le processus (4.3) est $\eta-$ faiblement dépendant avec $\eta_r = \mathcal{O}\left(r^{-\nu+1}\right)$.

e. Processus de Volterra non causal

Définition 4.6. *Soit $(\zeta_t)_t$ une suite de variables aléatoires i.i.d. Le processus de Volterra est défini par :*

$$X_t = \sum_{j=1}^{\infty} Y_t^{(p)} \quad avec \quad Y_t^{(p)} = \sum_{j_1 < j_2 < \ldots < j_p} a_{j_1,\ldots,j_p} \zeta_{k-j_1} \ldots \zeta_{k-j_p}, \quad t \in \mathbb{Z} \tag{4.4}$$

où $(a_{j_1,\ldots,j_p}) \in \mathbb{R}$ pour $p \in \mathbb{N}^*$.

Supposons que $\sum_{p=0}^{\infty} \sum_{j_1<j_2<\ldots<j_p} \left|a_{j_1,\ldots,j_p}\right|^m E\left|\zeta_0\right|^{mp} < \infty$ avec $m > 0$ et qu'il existe $p_0 \in \mathbb{N}^*$ tel que $a_{j_1,\ldots,j_p} = 0$ pour $p > p_0$.
Si $a_{j_1,\ldots,j_p} = \mathcal{O}\left(\max_{1\leq i\leq p} \mid j_i \mid^{-\mu}\right)$ avec $\mu > 0$, alors le processus (4.4) est $\eta-$ faiblement dépendant avec $\eta_r = \mathcal{O}\left(\frac{1}{r^{\mu+1}}\right)$ (voir Doukhan et Neumann. [38]).

4.2.3 Propriété d'hérédité

Les coefficients de dépendance ont une propriété d'hérédité,i.e., si la suite $(X_t)_t$ est η, κ, λ ou θ faiblement dépendante, alors $\forall h : \mathcal{X} \longrightarrow \mathbb{R}$ ($\mathcal{X}$ un espace polonais) une fonction Lipschitizienne, $(h(X_t))_t$ est aussi faiblement dépendante. On peut avoir aussi cette propriété avec des fonctions qui ne sont pas Lipschitiziennes mais vérifiant certaines conditions spécifiques données par la proposition suivante.

Proposition 4.1. *(Dedecker et al. [30])*
Soit $(X_n)_n$ une suite de variables aléatoires à valeurs dans $\mathbb{R}^s$. On suppose qu'il existe une constante $C > 0$ telle que $\max_{1\leq i\leq s} \|X_i\|_p \leq C$ avec $p > 1$. Soit h une fonction de $\mathbb{R}^s \to \mathbb{R}$ vérifiant :

$$\begin{cases} h(0) = 0 \\ \forall x, y \in \mathbb{R}^s, \exists a \in [1, p[\text{ et } c > 0,\ |h(x) - h(y)| \leq c\,|x - y|\left(1 + |x|^{a-1} + |y|^{a-1}\right) \end{cases} \tag{4.5}$$

Posons $Y_n = h(X_n)$, alors
Si $(X_n)_n$ est η- faiblement dépendante (respectivement λ, θ-faiblement dépendante), alors $(Y_n)_n$ l'est aussi, avec

$$\eta_r(Y) = \mathcal{O}\left(\eta_r^{\frac{p-a}{p-1}}\right),\ \lambda_r(Y) = \mathcal{O}\left(\lambda_r^{\frac{p-a}{p+a-2}}\right) \text{ et } \theta_r(Y) = \mathcal{O}\left(\theta_r^{\frac{p-a}{p-1}}\right).$$

4.3 Normalité de l'estimateur de Hill sous dépendance faible

Soient $X_1, X_2, \ldots, X_n$ des observations provenant d'un processus linéaire stationnaire défini par (3.13) et notons $X_{(1)} \geq X_{(2)} \geq \ldots \geq X_{(n)}$ les statistiques d'ordre correspondantes. Rappelons que l'estimateur de Hill du paramètre $\xi = 1/\beta$ basé sur les k plus grandes observations est donné par :

$$\overset{\wedge}{\xi}{}_k^H = \frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k+1)}} \tag{4.6}$$

Supposons que la condition (4.1) est satisfaite, alors $(X_t)_t$ est $\eta-$ faiblement dépendante, i.e., pour une suite monotone $\eta = (\eta_{l_n})_{l_n \in \mathbb{N}}$ décroissante vers zero et pour $(h, g, u, v) \in \mathcal{L}^2 \times \mathbb{N}^2$:

$$Cov(h(X_{i_1}, ..., X_{i_u}), g(X_{j_1}, ..., X_{j_v})) \leq (u\left\|g\right\|_\infty Lip(h) + v\left\|h\right\|_\infty Lip(g))\eta_{l_n} \tag{4.7}$$

pour tous $i_1 < .... < i_u \leq i_u + l_n \leq j_1 < .... < j_v$ et l_n une suite d'entiers. $\mathcal{L}$ est la classe des fonctions définies sur $\mathbb{R}^n$, Lipschitziennes et bornées.
Dans ce cas le coefficient η_{l_n} vérifie $\eta_{l_n} = \mathcal{O}(\frac{1}{l_n^{\mu-1/2}}), \ \ \mu > \frac{1}{2}$.

L'estimateur de Hill s'écrit en fonction du logarithme des observations initiales X_t, provenant d'un processus η -faiblement dépendant. Pour montrer la normalité asymptotique de cet estimateur, on s'intéresse tout d'abord à la dépendance faible de la suite $(Y_t)_t$ définie par $Y_t = \log X_t$. Sous certaines conditions et malgré le fait que la fonction logarithmique ne satisfait pas la condition (4.5), on montre la propriété d'hérédité de $(Y_t)_t$ relativement à la η-dépendance.

Lemme 4.1. *Soit $(X_t)_t$ une suite stationnaire de variables aléatoires η-faiblement dépendante telle que $\exists C > 0$ telle que $\|X_1\|_p \leq C$ avec $p > 1$, alors le processus $(Y_t)_t$ défini par $Y_t = \log X_t$ est aussi η-faiblement dépendant, avec*

$$\eta_r(Y) = \mathcal{O}(\eta_r^{\frac{p}{p-1}})$$

Preuve :

Soient $x=(x_1,...,x_k)\in\mathbb{R}_+^k$ et $x^{(M)}=(x_1^{(M)},...,x_k^{(M)})$ deux vecteurs où $x^{(M)}=\min(x,M)$, pour $M>0$. Soient $(f,g)\in\pounds_u\times\pounds_v$, $X_i=(X_{i_1},...,X_{i_u})$ et $X_j=(X_{j_1},...,X_{j_v})$, $(u,v)\in\mathbb{N}^2$. On considère deux fonctions F et G définies respectivement par
$F:\mathbb{R}^{uk}\to\mathbb{R}$ et $G:\mathbb{R}^{vk}\to\mathbb{R}$ avec

$$F(X_i)=f(h(X_{i_1}),...,h(X_{i_u}))\ ,\ F^{(M)}(X_i)=f(h(X_{i_1}^{(M)}),...,h(X_{i_u}^{(M)}))$$

$$G(X_j)=g(h(X_{j_1}),...,h(X_{j_v}))\ ,\ G^{(M)}(X_j)=g(h(X_{j_1}^{(M)}),...,h(X_{j_v}^{(M)}))$$

où $h(x)=\log x$.

$$\begin{aligned}|cov(F(X_i),G(X_j))| &= \left|cov(F(X_i),G(X_j)-G^{(M)}(X_j)+G^{(M)}(X_j))\right|\\ &\leq \left|cov(F(X_i),G(X_j)-G^{(M)}(X_j))\right|+\left|cov(F(X_i)-F^{(M)}(X_i),G^{(M)}(X_j))\right|\\ &\quad+\left|cov(F^{(M)}(X_i),G^{(M)}(X_j))\right|\end{aligned}$$

On peut majorer le premier terme de la façon suivante

$$\begin{aligned}\left|cov(F(X_i),G(X_j)-G^{(M)}(X_j))\right| &\leq \left|E(F(X_i)(G(X_j)-G^{(M)}(X_j)))\right|\\ &\quad+\left|E(F(X_i))E(G(X_j)-G^{(M)}(X_j))\right|\\ &\leq 2\left\|f\right\|_\infty E\left|(G(X_j)-G^{(M)}(X_j))\right|\end{aligned}$$

avec $\left\|f\right\|_\infty=\sup_t|f(t)|$.
De la même manière, on a

$$\left|cov(F(X_i)-F^{(M)}(X_i),G(X_j))\right|\leq 2\left\|g\right\|_\infty E\left|(F(X_i)-F^{(M)}(X_i))\right|$$

On obtient

$$\begin{aligned}|cov(F(X_i),G(X_j))| &\leq 2\left\|f\right\|_\infty E\left|(G(X_j)-G^{(M)}(X_j))\right|+2\left\|g\right\|_\infty E\left|(F(X_i)-F^{(M)}(X_i))\right|\\ &\quad+\left|cov(F^{(M)}(X_i),G^{(M)}(X_j))\right|\end{aligned}$$

$$\begin{aligned}E\left|(G(X_j)-G^{(M)}(X_j))\right| &= E\left|g(\log X_{j_1},...,\log X_{j_v})-g(\log X_{j_1}^{(M)},...,\log X_{j_v}^{(M)})\right|\\ &\leq (Lipg)\sum_{l=1}^{v}E(\left|\log X_{j_l}-\log X_{j_l}^{(M)}\right|)\\ &\leq (Lipg)vE\left|(\log X_{j_l}-\log M)\,1_{(X_{j_l}>M)}\right|\end{aligned}$$

En utilisant le fait que : $\log x - \log y \leq \frac{1}{y}(x-y)\ \forall (x,y) \in \mathbb{R}_+^*$ avec $x \geq y$, on a
On a :

$$\begin{aligned} E\left|(G(X_j) - G^{(M)}(X_j))\right| &\leq (Lipg)\frac{v}{M}E\left|(X_{j_l} - M)\,1_{(X_{j_l}>M)}\right| \\ &\leq (Lipg)\frac{v}{M}E\left(\left|X_{j_l}\right|1_{(X_{j_l}>M)}\right) \\ &\leq (Lipg)\frac{v}{M}E\left(\left|X_{j_l}\right|^{1-p}\left|X_{j_l}\right|^{p}1_{(X_{j_l}>M)}\right),\quad p>1 \\ &\leq (Lipg)\frac{v}{M}M^{1-P}C^p \end{aligned}$$

De la même manière,

$$E\left|(F(X_i) - F^{(M)}(X_i))\right| \leq (Lipf)\frac{u}{M}M^{1-P}C^p$$

Comme $(X_t)_t$ est $\eta-$ faiblement dépendante, on a

$$\left|cov(F^{(M)}(X_i), G^{(M)}(X_j))\right| \leq \left(u\left\|F^{(M)}\right\|_\infty \left(LipG^{(M)}\right) + v\left\|G^{(M)}\right\|_\infty\left(LipF^{(M)}\right)\right)\eta_r$$

La fonction $h(x) = \log x$ est différentiable dans $[M, +\infty[$, $(M>0)$, donc

$$Liph = \sup_{x\in[M,+\infty[} h'(x) = \sup_{x\in[M,+\infty[}\frac{1}{x} = \frac{1}{M}$$

On en déduit que

$$LipF^{(M)} = \sup_{x\in[M,+\infty[} F'^{(M)}(x) = \sup_{x_i\in[M,+\infty[}\left(f(h(x_1),h(x_2),...,h(x_u))\right)' = \frac{1}{M}Lipf$$

De même $LipG^{(M)} \leq \dfrac{1}{M}Lipg$.
De plus, on a $\left\|F^{(M)}\right\|_\infty \leq \|f\|_\infty$ et $\left\|G^{(M)}\right\|_\infty \leq \|g\|_\infty$

$$\begin{aligned} |cov(F(X_i), G(X_j))| &\leq 2\|f\|_\infty (Lipg)\frac{v}{M}M^{1-P}C^p + 2\|g\|_\infty (Lipf)\frac{u}{M}M^{1-P}C^p \\ &\quad + \left(u\|f\|_\infty \frac{1}{M}(Lipg) + v\|g\|_\infty\frac{1}{M}(Lipf)\right)\eta_r \\ &\leq \left(u\|f\|_\infty (Lipg) + v\|g\|_\infty (Lipf)\right)\left(\frac{2C^p}{M^p} + \frac{1}{M}\eta_r\right) \end{aligned}$$

En choisissant $M = \eta_r^{\frac{1}{1-p}}$, on obtient

$$|cov(F(X_i), G(X_j))| \leq \left(u\|f\|_\infty (Lipg) + v\|g\|_\infty (Lipf)\right)(2C^p+1)\eta_r^{\frac{p}{p-1}}$$

$(Y_t)_t$ est donc $\eta-$faiblement dépendante avec $\eta_r(Y) = \mathcal{O}\left(\eta_r^{\frac{p}{p-1}}\right)$.

4.3.1 Généralisation des résultats de Rootzén

Nous allons maintenant généraliser le lemme 3.2 de Rootzén au cas de variables aléatoires η-faiblement dépendantes.

Soient $(p_n)_n$ et $(r_n)_n$ deux suites d'entiers vérifiant (3.9) et supposons que la suite $(l_n)_n$ introduite dans (4.7) vérifie

$$\lim_{n\to\infty} \frac{l_n}{n} = 0, \lim_{n\to\infty} \frac{l_n}{r_n} = 0 \text{ et } \lim_{n\to\infty} \frac{n l_n^{\frac{1}{2}-\mu}}{r_n} = 0, \quad \mu > \frac{1}{2} \tag{4.8}$$

Lemme 4.2. *Soit $(X_t)_t$ un processus $\eta-$faiblement dépendant vérifiant les conditions* (3.10), (4.8). *Si*

$$\lim_{n\to\infty} p_n(varV_1 + varW_1) = 0 \tag{4.9}$$

Alors

$$\sum_{i=1}^{p_n} (V_i - E(V_i)) \xrightarrow{P} 0$$
$$\sum_{i=1}^{m_n} (W_i - E(W_i)) \xrightarrow{P} 0$$

Preuve :

On note $L_u = V_u - E(V_u)$

$$\begin{aligned} |E(e^{it\sum_{u=1}^{p_n} L_i}) - \prod_{u=1}^{p_n} E(e^{itL_u})| &\leq \sum_{s=2}^{p_n} |E(e^{it\sum_{u=1}^{s} L_u}) - E(e^{it\sum_{u=1}^{s-1} L_u})E(e^{itL_u})| \\ &\leq \sum_{s=2}^{p_n} |cov(h(L_1, L_2, ..., L_{s-1}), g(L_s))| \end{aligned}$$

avec

$$\begin{aligned} h(L_1, L_2, ..., L_{s-1}) &= e^{\left(it\sum_{u=1}^{s-1} L_u\right)} \\ g(L_s) &= e^{itL_u} \end{aligned}$$

En utilisant l'inégalité $|e^{ix} - e^{iy}| \leq \left| sin\left(\frac{x-y}{2}\right)\right| \leq \frac{|x-y|}{2}$, alors

$$Liph = \sup_{x\neq y} \frac{|h(x)-h(y)|}{\|x-y\|_1} = \sup_{L_u} \frac{|e^{it\sum_{u=1}^{s_1-1} L_u} - e^{it\sum_{u=1}^{s_2-1} L_u}|}{\left|\sum_{u=1}^{s_1-1} L_u - \sum_{u=1}^{s_2-1} L_u\right|} \leq \sup_{L_u} \frac{\frac{|t|}{2}\left|\sum_{u=1}^{s_1-1} L_u - \sum_{u=1}^{s_2-1} L_u\right|}{\left|\sum_{u=1}^{s_1-1} L_u - \sum_{u=1}^{s_2-1} L_u\right|} = \frac{|t|}{2}$$

De la même manière, $Lipg \leq \frac{|t|}{2}$.
En utilisant la dépendance faible (4.7) et du fait que V_u est définie par un groupe de variables aléatoires séparées par l_n avec n très grand, alors

$$\begin{aligned}\sum_{s=2}^{p_n} |cov(h(L_1, L_2, ..., L_{s-1}), g(L_s))| &\leq (p_n - 1)((s-1)Liph + Lipg)\eta_{l_n} \\ &\leq (p_n - 1)s\frac{|t|}{2}\eta_{l_n} \to 0, \text{ quand } n \to \infty\end{aligned}$$

Soient $\phi_L et \phi_{L_1}$ respectivement les fonctions caractéristiques de $L = \sum_{u=1}^{p_n} L_u$ et L_1.
D'après le dernier résultat, on déduit que $\lim_{n\to\infty} \phi_L(t) = \lim_{n\to\infty} (\phi_{L_1}(t))^{p_n}$.
En utilisant le développement de Taylor de ϕ_{L_1}, on obtient

$$\phi_{L_1}(t) = \phi_{L_1}(0) + t\phi'_{L_1}(0) + \frac{t^2}{2}\phi''_{L_1}(0) + o(t^2)$$

avec $\phi_{L_1}(0) = 1$, $\phi'_{L_1}(0) = 0$, $\phi''_{L_1}(0) = -varV_1$.

$$\lim_{n\to\infty} \phi_L(t) = \lim_{n\to\infty} \exp\{ip_n \ln(1 - \frac{t^2}{2}varV_1 + o(t^2))\}$$

En vertu de l'hypothèse (4.9), on a : $\lim_{n\to\infty} p_n varV_1 = 0$ et donc $\lim_{n\to\infty} \phi_L(t) = 1$.
Ceci implique que $\sum_{u=1}^{p_n} L_u \xrightarrow{d} 0$. Par conséquent $\sum\limits_{i=1}^{p_n} (V_i - E(V_i)) \xrightarrow{P} 0$.
On procède de la même manière pour montrer $\sum\limits_{i=1}^{m_n} (W_i - E(W_i)) \xrightarrow{P} 0$.

En utilisant l'approche de Rootzén, nous allons étudier la normalité asymptotique de l'estimateur de Hill lorsque les observations sont issues d'un processus linéaire faiblement dépendant au sens de Doukhan. Pour cela, nous commençons tout d'abord à étendre le théorème 3.3 pour des variables η-faiblement dépendantes. Ensuite on l'appliquera par la suite au processus $(Y_t)_t$ défini par $Y_t = \log X_t$.

On rappelle tout d'abord le théorème de Lindeberg (Voir Billingsley, th 7.2 [11]) qui est essentiel pour la démonstration de notre résultat.

Théorème 4.1. *(Lindeberg)*
Supposons que $X_1, ..., X_n$ *sont des variables aléatoires indépendantes telles que :*
$E(X_n) = \mu_n < \infty$ *et* $Var(X_n) = \sigma_n^2 < \infty$ *. On pose*

$$T_n = \sum_{i=1}^{n}(X_i - \mu_i), \quad s_n^2 = VarT_n = \sum_{i=1}^{n}\sigma_i^2$$

Si, $\forall \epsilon > 0$

$$\lim_{n\to\infty} \frac{1}{s_n^2}\sum_{i=1}^{n} E\{(X_i - \mu_i)^2 1_{(|X_i-\mu_i|\geq\ \epsilon S_n)}\} = 0$$

Alors

$$\frac{T_n}{s_n} \xrightarrow{d} N(0,1)$$

Théorème 4.2.
Soit $(Y_t)_t$ *une suite stationnaire* η*-faiblement dépendante, sous les conditions du théorème3.3 et si la condition* $\lim_{n\to\infty} p_n(\alpha_{n,l_n} + \frac{l_n}{n}) = 0$ *est remplacée par* (4.8)*, alors la convergence* (3.12) *reste vraie.*

Preuve : Soient

$$\hat{\beta}_n = \frac{1}{k}\sum_{i=1}^{n}\psi(Y_i - u_n)_+$$

$$E(\hat{\beta}_n) = \frac{n}{k}E(\psi(Y_1 - u_n)_+)$$

On partage $(\hat{\beta}_n)_n$ en deux suites $(Z_i)_i$ et $(W_i)_i$ définies dans (3.11), les variables $Z_1, Z_2, .., Z_n$ (respectivement $W_1, W_2, .., W_n$) sont indépendantes.

$$\left(\frac{k}{\lambda_n}\right)^{\frac{1}{2}}\left(\hat{\beta}_n - E(\hat{\beta}_n)\right) = \sum_{i=1}^{p_n}(Z_i - E(Z_i)) + \sum_{i=1}^{m_n}(W_i - E(W_i))$$

La suite $(Y_t)_t$ est η-faiblement dépendante et vérifie la condition (4.8). En utilisant le lemme 4.2, on en déduit que $\sum_{i=1}^{m_n}(W_i - E(W_i)) \xrightarrow{P} 0$.
Afin de prouver que

$$\left(\frac{k}{\lambda_n}\right)^{\frac{1}{2}}\left(\hat{\beta}_n - E(\hat{\beta}_n)\right) \xrightarrow{d} N(0,1)$$

Il suffit de montrer que

$$\sum_{i=1}^{p_n}(Z_i - E(Z_i)) \xrightarrow{d} N(0,1)$$

Pour appliquer le théorème de Lindeberg à la suite $(Z_n)_n$, nous devons vérifier la condition suivante :

$$\forall \epsilon > 0,\ \lim_{n\to\infty} p_n E\left\{(Z_1 - E(Z_1))^2 1_{(|Z_1-E(Z_1)|>\epsilon)}\right\} = 0$$

Pour ceci, on a besoin de tronquer $\psi(Y_i - u_n)$ comme suit :

Pour une suite donnée $(w_n)_n$ tendant vers l'infini, on pose :

$$\begin{aligned} Y_i^{'} &= Y_i 1_{(Y_i\leq w_n)} + w_n 1_{(Y_i>w_n)}\ , 1\leq j\leq n \\ \text{et } Z_1^{'} &= (\lambda_n k)^{-\frac{1}{2}} \sum_{i=1}^{r_n} \psi\left(Y_i^{'}\right) \end{aligned}$$

On a

$$p_n E\{(Z_1 - E(Z_1))^2 1_{(Z_1-E(Z_1)>\epsilon)}\} \leq 4p_n(E\{(Z_1' - E(Z_1'))^2 1_{(Z_1'-E(Z_1')>\frac{\epsilon}{2})}\} + E(Z_1 - Z_1')^2)$$

Les conditions 2, 3 et 5 du théorème 3.3 étant satisfaites, alors le premier terme tend vers zéro car

$$0 \leq Z_1' \leq \frac{1}{\sqrt{k\lambda_n}} r_n \psi(w_n) \to 0$$

De plus, en utilisant le lemme 3.3 de Rootzén et al. [80], on a :

$$\lim_{n\to\infty} p_n E\left(Z_1 - Z_1^{'}\right)^2 = 0$$

La condition de Lindeberg étant vérifiée, alors

$$\left(\frac{k}{\lambda_n}\right)^{\frac{1}{2}} \left(\frac{1}{k}\sum_{i=1}^{n} \psi(Y_i - u_n)_+ - \frac{n}{k}E(\psi(Y_1 - u_n)_+)\right) \xrightarrow{d} N(0,1)$$

La condition 6 du théorème 3.3 étant satisfaite, alors

$$\left(\frac{k}{\lambda_n}\right)^{\frac{1}{2}} \left(\frac{1}{k}\sum_{i=1}^{k} \left(Y_{(i)} - Y_{(k)}\right) - \frac{n}{k}E\left(Y_1 - u_n\right)_+\right) \xrightarrow{d} N(0,1)$$

(Voir le théorème 4.3 de [80]).

Ce qui complète la démonstration.

4.3.2 Généralisation du résultat de Resnick et Stărică

Dans cette partie, considérons $(X_t)_t$ un processus linéaire défini par (3.13) et notons par F la distribution de X_t et G la distribution des innovations ε_t.

Pour généraliser le théorème 3.5 de Resnick Stărică, on a besoin de la proposition suivante donnée par Geluk et al. [47] qu'est une version du théorème de Karamata pour des fonctions à variation régulière de second ordre.

Proposition 4.2. *Si F est une distribution à support sur $[0,\infty)$, alors*

$$\overline{F} \in 2RV(-\beta,\rho)$$

ssi il existe une fonction g vérifiant $g>0, g(t)\to 0$ quand $t\to\infty$ et $g\in RV_\rho$, $\rho\leq 0$ telle que :

$$\lim_{t\to\infty}\frac{\int\limits_t^\infty \frac{\overline{F}(x)}{\overline{F}(t)}\frac{dx}{x}-\frac{1}{\beta}}{g(t)}=c$$

où c est une constante non nulle.

Le résultat suivant est une conséquence du théorème 4.2 qui nous permettra de déduire la normalité asymptotique de l'estimateur de Hill.

Corollaire 4.1. *Soit $(X_t)_t$ un processus linéaire défini par (3.13) vérifiant (3.16), (3.17), (3.18) et (4.1). Supposons que F vérifie (3.7), $G\in RV_{-\beta}$ et que les conditions (4.8), (3.21) sont satisfaites, alors*

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k+1)}}-\int\limits_1^\infty \frac{n}{k}P\left(\frac{X_1}{b\left(\frac{n}{k}\right)}>x\right)\frac{dx}{x}\right)\xrightarrow{d} N(0,\lambda)$$

avec

$$\lambda=\frac{1}{\beta^2}\left(1+2\frac{\sum_{j=1}^\infty\sum_{k=0}^\infty |c_k|^\beta\wedge|c_{k+j}|^\beta}{\sum_{j=0}^\infty |c_k|^\beta}\right)$$

Si de plus, $\overline{F}\in 2RV(-\beta,\rho)$ et que sa fonction auxiliaire g vérifie (3.22), alors

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k+1)}}-\frac{1}{\beta}\right)\xrightarrow{d} N(0,\lambda)$$

Preuve :

Il suffit de montrer que les conditions du théorème 4.2 sont satisfaites.

Remarquons que si $X \in RV_{-\beta}$, alors $E(X^p) < \infty$, $\forall p < \beta$.

Si de plus, X est positive, alors $E(\|X\|_p) = E(X^p)^{\frac{1}{p}} \leq C$, $\forall\, p$ où $1 < p < \beta$.

En utilisant (3.16),(3.17) et (4.1), $(X_t)_t$ est η-faiblement dépendant et à variation régulière.

En appliquant le lemme 4.1, $(Y_t)_t$ est aussi η-faiblement dépendant avec $Y_t = \log X_t$ et $1 < p < \beta$.

Notons que si on choisit la suite $(u_n)_n$ telle que $u_n = \log b\left(\frac{n}{k}\right)$, alors (3.10) est vérifiée et d'après (3.21), il existe deux suites $(r_n)_n$ et $(w_n)_n$ vérifiant les conditions du théorème 4.2 (voir [78]).

En appliquant le théorème 4.2 au processus $Y_n = \log X_n$ et en prenant $u_n = \log b(\frac{n}{k})$, on en déduit que

$$\sqrt{\frac{k}{\lambda_n}}\left(\frac{1}{k}\sum_{i=1}^{k}\left(\log\frac{X_{(i)}}{X_{(k+1)}}\right) - \frac{n}{k}E\left(\log\frac{X_1}{b\left(\frac{n}{k}\right)}\right)_+\right) \xrightarrow{d} N(0,1)$$

D'après le lemme 3.3 de Resnick et Stărică [78],

$$\lambda_n \to \lambda = \frac{1}{\beta^2}\left(1 + 2\frac{\sum\limits_{j=1}^{\infty}\sum\limits_{k=0}^{\infty}|c_k|^\beta \wedge |c_{j+k}|^\beta)}{\sum\limits_{k=0}^{\infty}|c_k|^\beta}\right), \quad \text{quand} \quad n \to \infty$$

et donc

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\left(\log\frac{X_{(i)}}{X_{(k+1)}}\right) - \frac{n}{k}E\left(\log\frac{X_1}{b\left(\frac{n}{k}\right)}\right)_+\right) \xrightarrow{d} N(0,\lambda)$$

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k+1)}} - \int_1^\infty \frac{n}{k}P\left(\frac{X_1}{b\left(\frac{n}{k}\right)} > x\right)\frac{dx}{x}\right) \xrightarrow{d} N(0,\lambda) \tag{4.10}$$

Comme $\overline{F} \in 2RV(-\beta,\rho)$ et en vertu de la proposition 4.2

$$\lim_{t\to\infty}\frac{\int_t^\infty \frac{\overline{F}(x)}{\overline{F}(t)}\frac{dx}{x} - \frac{1}{\beta}}{g(t)} = c \neq 0$$

où g est la fonction auxiliaire de F.

Par conséquent,

$$\int_t^{\infty} \frac{\overline{F}(x)}{\overline{F}(t)} \frac{dx}{x} \frac{1}{\beta} \sim cg(t) \qquad \text{au } v(\infty)$$

$$\sqrt{k}\left(\frac{1}{\overline{F}(t)} \int_t^{\infty} \overline{F}(x) \frac{dx}{x} - \frac{1}{\beta}\right) \sim \sqrt{k} cg(t)$$

$$\sqrt{k}\left(\frac{1}{\overline{F}(b(\frac{n}{k}))} \int_{b\left(\left(\frac{n}{k}\right)\right)}^{\infty} \overline{F}(x) \frac{dx}{x} - \frac{1}{\beta}\right) \sim \sqrt{k} cg(b(\frac{n}{k}))$$

En remplaçant x par $s.b\left(\frac{n}{k}\right)$, la dernière intégrale devient

$$\sqrt{k}\left(\frac{1}{\overline{F}(b(\frac{n}{k}))} \int_1^{\infty} \overline{F}\left(b\left(\frac{n}{k}\right) s\right) \frac{ds}{s} - \frac{1}{\beta}\right) \sim \sqrt{k} cg(b(\frac{n}{k}))$$

$$\sqrt{k}\left(\frac{n}{k} \int_1^{\infty} P\left(\frac{X_1}{b\left(\frac{n}{k}\right)} > s\right) \frac{ds}{s} - \frac{1}{\beta}\right) \sim \sqrt{k} cg(b(\frac{n}{k}))$$

Comme

$$\lim_{n \to \infty} \sqrt{k} g(b(\frac{n}{k})) = 0$$

(4.10) s'écrit donc comme suit

$$\sqrt{k}\left[\frac{1}{k} \sum_{i=1}^{k} \log \frac{X_{(i)}}{X_{(k+1)}} - \frac{1}{\beta}\right] + \sqrt{k}\left[\frac{1}{\beta} - \int_1^{\infty} \frac{n}{k} P\left(\frac{X_1}{b\left(\frac{n}{k}\right)} > x\right) \frac{dx}{x}\right] \xrightarrow{d} N(0, \lambda)$$

Alors,

$$\sqrt{k}\left[\frac{1}{k} \sum_{i=1}^{k} \log \frac{X_{(i)}}{X_{(k+1)}} - \frac{1}{\beta}\right] \xrightarrow{d} N(0, \lambda)$$

Ce qui termine la démonstration.

Conclusion et perspectives

La dernière partie de ce mémoire est pour nous l'occasion de revenir sur les principales étapes de la recherche que nous avons présentées, d'en discuter quelques points et perspectives importantes et de décrire les développements en cours. Peut être cette thèse a t-elle atteint, au moins, les objectifs qu'elle visait à l'origine. Mais il convient de conclure également, qu'on n'a pas épuisé le champ des possibilités et sur certains points, elle a l'avantage d'ouvrir des pistes d'exploitation.

La problématique fondamentale de cette thèse était de montrer que les résultats existant sur le comportement asymptotique de l'estimateur de Hill obtenus dans le cas dépendant du type mélange intense ou provenant d'un processus stationnaire pouvaient être étendus à la dépendance faible au sens de Doukhan. Nous avons montré la normalité de cet estimateur, sous dépendance faible et sous des hypothèses moins restrictives. Le résultat obtenu est une généralisation des travaux existant dans la littérature.

Aussi, à long et court terme, il serait intéressant d'étendre ce travail à de nouvelles perspectives que nous listons ci-après :

1- Étudier les propriétés des autres estimateurs cités dans le deuxième chapitre (Pickands, Dekkers-Einmahl-De Haan, estimateur de Hill négatif,...) sous dépendance faible.

2- Se pencher sur la consistance forte et faible de l'estimateur de Hill.

3- Étudier le comportement asymptotique de l'estimateur de Hill sous données censurées dans un premier temps dans le domaine de Fréchet et ensuite dans les autres domaines.

4- Au cours du dernier chapitre nous avons travaillé avec le processus $MA(\infty)$ mais il serait aussi possible d'appliquer l'estimateur de Hill à d'autres processus faiblement dépendants.

5- Réfléchir sur la réduction de la variance asymptotique de l'estimateur en adoptant la méthode de Ling et Peng dans le cas de processus autorégressif.

6- Élargir l'étude au cas multivarié.

Bibliographie

[1] Andrews, D. Non-strong mixing autoregressive processes. *Journal of Applied Probability. 21(4).* (1984), 930–934.

[2] Athreya, K., and Pantula, S. Mixing properties of harris chains and autoregressive processes. *Journal of Applied Probability. 23(4).* (1986), 880–892.

[3] Athreya, K., and Pantula, S. A note on strong mixing of arma processes. *Statistics and Probability Letters. 4(4).* (1986), 187–190.

[4] Bardet, J., Doukhan, P., and León, J. A functional limit theorem for η-weakly dependent processes and its applications. *Statistical Inference for Stochastic Processes. 11(3).* (2008), 265–280.

[5] Beirlant, J., Dierckx, G., Goegebeur, Y., and Matthys, G. Tail index estimation and an exponential regression model. *Extremes. 2(2).* (1999), 177–200.

[6] Beirlant, J., Dierckx, G., Guillou, A., and Stărică, C. On exponential representations of log-spacings of extreme order statistics. *Extremes. 5(2).* (2002), 157–180.

[7] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. *Statistics of Extremes : Theory and Applications.* Wiley Series in Probability and Statistics. Chichester., 2004.

[8] Beirlant, J., Vynckier, P., and Teugels, J. Excess functions and estimation of the extreme-value index. *Bernoulli. 2(4)* (1996), 293–318.

[9] Beirlant, J., Vynckier, P., and Teugels, J. Tail index estimation, pareto quantile plots regression diagnostics. *Journal of the American Statistical Association. 91(436).* (1996), 1659–1667.

[10] Berman, S. Limit theorems for the maximum term in stationary sequences. *The Annals of Mathematical Statistics. 35(2).* (1964), 502–516.

[11] Billingsley, P. Convergence of probability measures. *Wiley Serie in Probability and Statistics, New York.* (1968).

[12] Bradley, R., and Peligrad, M. Invariance principles under a tow-part mixing assumption. *Stochastic Processes and their Applications. 22(2).* (1986), 271–289.

[13] Brito, M., and Freitas, A. Consistent estimation of the tail index for dependent data. *Statistics and Probability Letters. 80(23-24).* (2010), 1835–1843.

[14] Chanda, K. Strong mixing properties of linear stochastic processes. *Journal of Applied Probability. 11(2).* (1974), 401–408.

[15] Christopeit, N. Estimating parameters of an extreme value distribution by the method of moments. *Journal of Statistical Planning and inference. 41(2).* (1994), 173–186.

[16] Cline, D. Estimation and linear prediction for regression, autoregression and arma with infinite variance data. *Thesis, Dept of Statistics, Colorado State University, Ft. Collins CO 80521 USA.* (1983).

[17] Coles, S. An introduction to statistical modeling of extremes values. *Springer Series in Statistics. Springer-Verlag. London.* (2001).

[18] Csörgő, S., Deheuvels, P., and Mason, D. Kernel estimates of the tail index of a distribution. *The Annals of Statistics. 13(3).* (1985), 1050–1077.

[19] Csörgő, S., and Mason, D. Central limit theorems for sums of extreme values. *Mathematical Proceedings of the Cambridge Philosophical Society. 98(3).* (1985), 547–558.

[20] Danielsson, J., De Haan, L., Peng, L., and de Vries, C. Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis. 76(2).* (2001), 226–248.

[21] David, H. Order statistics. *Wiley Series in Probability and Mathematical Statistics. second édition. John Wiley and Sons. New York.* (1981).

[22] Davis, R. Maxima and minima of stationary sequences. *The Annals of Probability. 7(3).* (1979), 453–460.

[23] Davis, R. Gauss-newton and m-estimation for arma processes with infinite variance. *Stochastic Processes and their Applications. 63(1).* (1996), 75–95.

[24] Davis, R., and Resnick, S. Tail estimates motivated by extreme value theory. *The Annals of Statistics. 12(4).* (1984), 1467–1487.

[25] De Finetti, B. Sulla legge di probabilita degli estremi. *Metron. Roma. vol IX. n.3/4.* (1932), 125.

[26] De Haan, L., and Ferreira, A. Extreme value theory : An introduction. *Springer Series in Operations Research and Financial Engineering. Springer Science and Business Media. LLC.* (2006).

[27] De Haan, L., and Peng, L. Comparison of tail index estimators. *Statistica Neerlandica. 52(1).* (1998), 60–70.

[28] De Haan, L., and Resnick, S. On asymptotic normality of the hill estimator. *Communications in Statistics. Stochastic Models. 4.* (1997), 849–866.

[29] De Haan, L., Resnick, S., Rootzén, H., and De Vries, C. Extremal behavior of solutions to a stochastic difference equation with applications to arch processes. *Stochastic Processes and their Applications. 32(2).* (1989), 213–224.

[30] Dedecker, J., Doukhan, P., Lang, G., León, J. R., Louhichi, S., and Prieur, C. *Weak dependence : With Examples and applications.* 2007.

[31] Deheuvels, P., Haeusler, E., and Mason, D. Almost sure convergence of the hill estimator. *Mathematical Proceedings of the Cambridge Philosophical Society. 104(2).* (1988), 371–381.

[32] Dekkers, A., Einmalh, J., and De Haan, L. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics. 17(4).* (1989), 1833–1855.

[33] Dekkers.A.L.M, and Haan.L, D. On the estimation of the extreme-value index and large quantile estimation. *The Annals of Statistics. 17(4).* (1989), 1795–1832.

[34] Diop, A., and Lo, G. Generalized hill's estimator. *Far East Journal of Theoretical Statistics. 20(2).* (2006), 129–149.

[35] Doukhan, P. Mixing : Properties and examples. *Lecture notes in statistics. Springer. New York. 85.* (1994).

[36] Doukhan, P., and Louhichi, S. A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications. 84(2).* (1999), 313–342.

[37] Doukhan, P., Massart, P., and Rio, E. The functional central limit theorem for strongly mixing processes. *Annales de l'I.H.P Probabilités et statistiques. 30(1).* (1994), 63–82.

[38] Doukhan, P., and Neumann, M. The notion of $\psi-$ weak dependence and its applications to bootstrapping time series. *Probability Surveys. 5.* (2008), 146–168.

[39] Embrechts, P., Klüppelberg, C., and Mikosch, T. Modelling extremal events for insurance and finance. *Springer-Verlag. Berlin.* (1997).

[40] Falk, M. Some best parameter estimates for distributions with finite endpoint. *Statistics : A Journal of Theoretical and Applied Statistics. 27(1-2).* (1995), 115–125.

[41] Falk, M., Hüsler, J., and Reiss, R. Laws of small numbers : Extremes and rare events. *DMV Seminar. Birkhäuser Verlag. Basel. 23.* (1994).

[42] Falk, M., and Marohn, F. Efficient estimation of the shape parameter in pareto models with partially known scale. *Statistics and Decisions. 15(3).* (1997), 229–239.

[43] Finkenstädt, B., and Rootzén, H. Extreme values in finance, telecommunications and the environment. *Monographs on Statistics and Applied Probability 99. Chapman & Hall/CRC. CRC Press LLC. Boca Raton. Florida.* (2004).

[44] Fisher, R., and Tippett, L. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society. 24.* (1928), 180–190.

[45] Fréchet, M. Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique. 6.* (1927), 93–116.

[46] Geluk, J., and De Haan, L. Regular variation, extensions and tauberian theorems. *Centrum voor Wiskunde en Informatica. Amsterdam.* (1987).

[47] Geluk, J., De Haan, L., Resnick, S., and Stărică, C. Second-order regular variation, convolution and central limit theorem. *Stochastic processes and their applications. 69(2).* (1997), 139–159.

[48] Gnedenko, B. Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics. 44(3).* (1943), 423–453.

[49] Gordon, R. Values of mill's ratio area to bounding ordinate and of the normal probability integral for large values of the argument. *The Annals of Mathematical Statistics. 12(3).* (1941), 364–366.

[50] Gorodetskii, V. On the strong mixing property for linear sequences. *Theory Probability and Its Applications. 22(2).* (1978), 411–413.

[51] Groeneboom, P., Lopuhaä, H. P., and de Wolf, P. Kernel-type estimators for the extreme value index. *The Annals of Statistics. 31(6).* (2003), 1956–1995.

[52] Gumbel, E. Statistics of extremes. *Columbia University Press. New York.* (1958).

[53] Gumbel, E., and Mustafi, C. Some analytical properties of bivariate extremal distributions. *Journal of the American Statistical Association. 62(318).* (1967), 569–588.

[54] 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.

[55] Hall, P., and Heyde, C. Martingale limit theory and its application. *Academic Press, INC. New York.* (1980).

[56] Hill, B. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics. 3(5).* (1975), 1163–1174.

[57] Hosking, J., and Wallis, J. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics. 29(3).* (1987), 339–349.

[58] Hsing, T. On tail index estimation using dependent data. *The Annals of Statistics. 19(3).* (1991), 1547–1569.

[59] Jenkinson, A. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *The Quarterly Journal of the Royal Meteorological Society. 81(348).* (1955), 158–171.

[60] Leadbetter, M. On extreme values in stationary sequences. *Z. Wahrscheinlichkeitstheorie und Verwandte Gebiete. 28(4).* (1974), 289–303.

[61] Leadbetter, M., Lindgren, G., and Rootzén, H. Extremes and related properties of random sequences and processes. *Springer Verlag, New York.* (1983).

[62] Leadbetter, M., and Rootzén, H. Extreme value theory for continuous parameter stationary processes. *Z.Wahrscheinlichkeitstheorie und Verwandte Gebiete. 60.* (1982), 1–20.

[63] Ling, S., and Peng, L. Hill's estimator for the tail index of an arma model. *Journal of Statistical Planning and Inference. 123(2).* (2004), 279–293.

[64] Lo, G. *Sur quelques estimateurs de l'index d'une loi de Pareto : Estimateur de Hill, de Csörgő-Deheuvels-Mason, de De Haan-Resnick et lois limites pour de sommes de valeurs extrêmes pour une variable dans le domaine d'attraction de Gumbel.* Thèse de doctorat, Université Paris VI. France, 1986. Directeur de thèse : Deheuvels, P.

[65] Louhichi, S. Independence via uncorrelatedness, examples and moment inequalities. *Préprint de l'Université Paris Sud. Orsay.* (2000).

[66] Loynes, R. Extreme values in uniformly mixing stationary stochastic processes. *The Annals Mathematical Statistics. 36(3)* (1965), 993–999.

[67] Lye, L., Hapuarachchi, K., and Ryan, S. Bayes estimation of the extreme-value reliability function. *IEEE Transactions on Reliability. 42(4).* (1993), 641–644.

[68] Mason, D. Laws of large numbers for sums of extreme values. *The Annals of Probability. 10(3).* (1982), 754–764.

[69] Mason, D. A strong invariance theorem for the tail empirical process. *Annales de l' I.H.P Probabilités et Statistiques. 24(4).* (1988), 491–506.

[70] Peng, L. Asymptotically unbiased estimators for the extreme-value index. *Statistics and Probability Letters. 38(2).* (1998), 107–115.

[71] Pickands, J. Statistical inference using extreme order statistics. *The Annals of Statistics. 3(1).* (1975), 119–131.

[72] Prescott, P., and Walden, A. Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika. 67(3).* (1980), 723–724.

[73] Prieur, C. Density estimation for one-dimensional dynamical systems. *ESAIM : Probability and Statistics. 5.* (2001), 51–76.

[74] Prieur, C. Estimation de la densité invariante de systèmes dynamiques en dimension 1. *Comptes Rendus de l'Académie des Sciences. Séries 1. Mathematics. 332(8).* (2001), 761–764.

[75] Reiss, R., and Thomas, M. Statistical analysis of extreme values with applications to assurance, finance, hydrology and other fields. *Basel. Birkhauser. Verlag.* (2001).

[76] Resnick, S. Extreme values regular variation and point processes. *Springer Series in Operations Research and Financial Engineering. Springer Verlag, New-York.* (1987).

[77] Resnick, S., and Stărică, C. Consistency of hill's estimator for dependent data. *Journal of Applied Probability. 32(1).* (1995), 139–167.

[78] Resnick, S., and Stărică, C. Asymptotic behavior of hill's estimator for autoregressive data. *Communications in statistics. Stochastic models 13(4).* (1997), 703–721.

[79] Resnick, S., and Stărică, C. Tail index estimation for dependent data. *The Annals of applied probability. 8(4).* (1998), 1156–1183.

[80] Rootzén, H., Leadbetter, M., and De Haan, L. Tail and quantile estimation for strongly mixing stationary sequences. *Techical report 292, Center for Stochastic Processes, Departement of Statistics, University of North Carolina, Chapel Hill. NC 27599-3260.* (1990).

[81] Rosen, O., and Weissman, I. Comparison of estimation methods in extreme value theory. *Communication in Statistics-Theory and Methods. 25(4).* (1996), 759–773.

[82] Rosenblatt, M. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the USA. 42(1).* (1956), 43–47.

[83] Smith, R. Maximum likelihood estimation in a class of nonregular cases. *Biometrika. 72(1).* (1985), 67–90.

[84] Tiago de Oliveira, J. Bivariate extremes : decisions. *Bull. Internat. Statist. Inst. XLVI.* (1975), 241–251.

[85] Von Mises, R. La distribution de la plus grande de n valeurs. *Revue de Mathématique Union Interbalcanique. 1.* (1936), 141–160.

[86] Watson, G. Extreme values in samples from m-dependent stationary stochastic processes. *The Annals of Mathematical Statistics. 25(4).* (1954), 798–800.