Comportement asymptotique presque sûre de la moyenne empirique pour les distributions à qeues lourdes

Djaber Ibtissem

Table des matières

0.1 Introduction

Parmi les estimateurs de position le plus classique est la moyenne empirique. La loi forte des grands nombres montre bien que cet estimateur est fortement consistant pour la moyenne théorique ou espérance mathématique.

Un autre résultat très classique en probabilité et en statistique c'est le Théorème Central limite. Ce fameux théorème nous permet en évidence de construire un intervalle de confiance pour la moyenne théorique. Néanmoins, cet estimateur n'est pas robuste contre les valeurs aberrantes ou les valeurs extrêmes, en particulier pour les distributions à queues lourdes.

Alors pour obtenir un estimateur robuste pour la moyenne on doit tronquer les plus grandes et les plus petites observations dans la sommation des données. On obtient alors un nouveau estimateur dit : la moyenne empirique tronquée. Malheureusement la normalité asymptotique n'est pas assurée dans ce cas. Ce problème nous mène alors à penser sur une autre version pour cet estimateur.

Il existe une manière pour tronquer les observations en conservant la normalité asymptotique. L'idée est la suivante : on remplace les valeurs extrêmes par un estimateur convenable pour les quantiles. Ce nouveau estimateur est appelé la moyenne Winsorisée.

Le présent mémoire alors et une synthèse sur les propriétés asymptotiques de cet estimateur : normalité asymptotique et convergence presque sûre. Nous nous basons ici sur la théorie des processus empiriques et le processus empiriques des quantiles.

Le présent mémoire est réparti comme suit :

Nous étudions au premier chapitre, les propriétés asymptotiques des valeurs extrêmes et les domaines d'attractions. Nous présentons aussi la loi du logarithme itéré classique pour la moyenne empirique sans troncature. Cette étude nous sert à développer des résultats analogues pour la moyenne Winsorisée.

Le deuxième chapitre est consacré à l'estimation de l'index des queues de distributions. Nous présentons ici les différents estimateurs de l'index ainsi que leurs propriétés asymptotiques.

Comme on s'intéresse aux sommations des données aux queues de distributions, nous avons alors réservé le troisième chapitre à une étude exhaustive pour les sommes des valeurs extrêmes.

Dans le dernier chapitre on étudie les comportements asymptotiques des moyennes tronquées et Winsorisées en se basant en particulier sur l'article de Peng (2001). La normalité asymptotique de l'estimateur de Peng (2001) a été établie. Nous suivons cette étude en présentant, de notre part, le comportement asymptotique presque sûre.

Notations et abrévations

$X_{i,n}$	La $i^{ème}$ statistique d'ordre
$\overset{\mathcal{D}}{\rightarrow}$	Convergence en distribution
$\overset{\mathcal{D}}{=}$	Egalité en distribution
$\overset{p.s}{\rightarrow}$	Convergence presque sûre
$\overset{\mathcal{P}}{\rightarrow}$	Convergence en probabilité
$\mathbb{P}$	Probabilité
$\mathbb{E}$	Esperence
$\mathbb{E}\,(X/Y)$	Espérence conditionnelle de X sachant Y
$\mathbf{B}(n, F(x))$	Loi binomiale
$\mathcal{N}\,(0,1)$	Loi normale standard
F_{DPG}	Distribution de Paréto Généralisée
Q_{DPG}	La fonction des quantiles de F_{DPG}
$\{W(t);\ 0 \leq t < \infty\}$	Processus de Wiener
$\{\mathcal{B}(t); 0 \leq t \leq 1\}$	Pont Brownien
$\mathbb{G}_n(t)$	La fonction de distribution empirique
$\alpha_n(t)$	Processus empirique uniforme
$\beta_n(t)$	Processus empirique des quantiles uniforme
$PC_+\,(0,1)$	Ensemble des fonctions continues sur $\left]u_{j-1}, u_j\right[$, $\forall j$
$\{\mathcal{B}_i(t)\}_{i=0}^{\infty}$	Suite de ponts Brownien
$x \vee y$	$\max\,(x,y)$
$x \wedge y$	$\min\,(x,y)$
$\mathbf{1}\,(x \leq y)$	La fonction indicatrice
[v.a]	Variable aléatoire
[iid]	Indépendante identiquement distribueés
p.s	Presque sûre

Chapitre 1

Préliminaires

1.1 Statistique d'ordre

Soit X une variable aléatoire réelle définie sur l'espace (Ω, B), d'une fonction de distribution F et d'une densité f.

On note par $(X_1, X_2, ..., X_n)$ l'échantillon indépendant de X.

Définition 1.1.1— *En classant les variables aléatoires $X_1, X_2, ..., X_n$ par ordre croissant, on obtient*

$$X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}. \tag{1.1.1}$$

On appelle $X_{1,n}, X_{2,n}, ..., X_{n,n}$ les statistiques d'ordre associées à $(X_1, ..., X_n)$, et on désigne par $X_{i,n}$ la $i^{ième}$ statistique d'ordre.

Intéressons nous plus particulièrement au comportement d'extrêmes correspondant à $i = 1$ et $i = n$, i.e,

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

et

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

1.2 Loi de la statistique d'ordre

Théorème 1.2.1 — *La loi de la statistique d'ordre est definie par la densité*

$$f\left(x_1, x_2, ..., x_n\right) = n! \prod_{i=1}^{n} f\left(x_i\right) \mathbf{1}_{C_n}\left(x_1, x_2, ..., x_n\right),$$

où

$$C_n = \left\{\left(x_1, x_2, ..., x_n\right) \in \mathbb{R}^n / x_1 < x_2 < ... < x_n\right\}.$$

Théorème 1.2.2 — *La loi marginal de* $\left(X_{1,n}, X_{2,n}, ..., X_{n,n}\right)$, $k \leq n$, *admet la densité*

$$\frac{n!}{(n-k)!} \mathbf{1}_{C_k}\left(x_1, x_2, ..., x_k\right) \left(\prod_{i=1}^{k} f\left(x_i\right)\right) \left[\int_{x_k}^{+\infty} f(x) dx\right]^{n-k}.$$

Loi de $X_{k,n}$

Soit

$$f_k(x) = \frac{n!}{(n-k)!\,(k-1)!} F^{k-1}(x) \left(1 - F(x)\right)^{n-k} f(x),$$

la densité marginale de $X_{k,n}$, $(k = 1, ..., n)$ obtenu par l'intégration de la densité de $\left(X_{1,n}, X_{2,n}, ..., X_{n,n}\right)$.

En considérant la variable Y comptant le nombre de points X tels que $X < x$, on a

$$F_k(x) = \mathbb{P}\left(X_{k,n} < x\right) = \mathbb{P}(Y > k).$$

La variable aléatoire Y suit une loi binomiale $\mathbf{B}(n,\ F(x))$. On obtient

$$F_k(x) = \sum_{i=k}^{n} C_n^i F^i(x) \left[1 - F(x)\right]^{n-i}.$$

Lois des valeurs extrêmes $X_{1,n}$ et $X_{n,n}$

L'un des rôles importants des statistiques d'ordre en statistique est la détection des points aberrants dans un échantillon, on s'intéresse aux deux lois suivantes :

Loi de $X_{1,n}$

Ona

$$F_1(x) = \mathbb{P}(X_{1,n} \leq x) = \sum_{i=1}^{n} C_n^i F^i(x)(1-F(x))^{n-i} = 1 - C_n^0 (1-F(x))^n,$$

d'où

$$F_1(x) = 1-(1-F(x))^n, \quad \text{et} \quad f_1(x) = nf(x)(1-F(x))^{n-1}.$$

Loi de $X_{n,n}$

Ona

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

d'où

$$f_n(x) = nf(x) F^{n-1}(x).$$

Loi jointe d'un couple$(X_{k,n}, X_{l,n})$

Soit un couple $(X_{k,n}, X_{l,n})$ avec $1 \leq k < i < n$, on obtient

$$\begin{aligned} f_{k,i}(u,v) &= L(n,i,k) f(u) f(v) F^{k-1}(u)(F(v)-F(u))^{i-k-1} \\ &\times [1-F(v)]^{n-1} 1_{[u<v]}(u,v), \end{aligned}$$

où la constante de normalisation est

$$L(n,i,k) = \frac{(k-1)!(i-k-1)!(n-i)!}{n!}.$$

Pour $k=1$, et $l=n$, on obtient la loi jointe de $(X_{1,n}, X_{n,n})$

$$f_{1,n}(u,v) = n(n-1) f(u) f(v) [F(u)-F(v)]^{n-2} 1_{[u<v]}(u,v).$$

1.3 Domaine d'attraction

Plusieurs conditions nécessaires et suffisantes, qu'on peut utiliser pour déterminer à quel domaine d'attraction une certaine fonction de distribution appartient. Ces conditions concernent le comportement du queue de cette fonction de distribution, c.à.d, le comportement de $1-F(x)$. Nous citons les conditions de *von Mises* en terme des trois distributions limites Λ, Φ_α et Ψ_α.

Notre premier théorème donne une caractérisation de la distribution limite du maximum de l'échantillon $X_{n,n}$.

Théorème 1.3.1 — *Soit $X_1, X_2, ..., X_n$ une suite de [v.a] [iid]. S'il existe deux constantes $a_n > 0$, $b_n \in \mathbb{R}$, et une distribution non dégénérée G telle que*

$$\lim_{n\to\infty} F^n(a_n x + b_n) = \lim_{n\to\infty} \left\{ \frac{\max(X_1, X_2, ..., X_n) - b_n}{a_n} \leq x \right\} = G(x),$$

alors G est l'une des trois distribution des valeurs extrêmes

$$\begin{array}{lclll} \textit{Gumbel} : \Lambda(x) & = & \{\exp(-e^{-x}), & & x \in \mathbb{R} \\ \textit{Fréchet} : \Phi_\alpha & = & \left\{ \begin{array}{cc} \exp(-x^{-\alpha}) & x > 0 \\ 0 & x \leq 0 \end{array} \right. & & \alpha > 0 \\ \textit{Weibull} : \Psi_\alpha & = & \left\{ \begin{array}{cc} \exp(-(-x)^\alpha) & x < 0 \\ 1 & x \geq 0 \end{array} \right. & & \alpha > 0 \end{array}$$

On note par :

$$F^{-1}(1) = \sup\{x : F(x) < 1\},$$

le point extrême superieur de la distribution F.

Théorème 1.3.2 — *$F \in D(\Lambda)$ si et seulement s'il existe une fonction strictement positive $q(t)$ telle que*

$$\lim_{t\uparrow F^{-1}(1)} \frac{1 - F(t + xq(t))}{1 - F(t)} = e^{-x} \quad (= -\log \Lambda). \tag{1.3.1}$$

On peut choisir

$$q(t) = \int_t^{F^{-1}(1)} (1 - F(u))\, du \big/ 1 - F(t),$$

pour $t < F^{-1}(1)$.

Théorème 1.3.3 — *Si $F^{-1}(1) = +\infty$, et supposons qu'il existe une constante $\alpha > 0$ telleque pour tout $x > 0$,*

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha} \quad (= -\log \Phi_\alpha). \tag{1.3.2}$$

Alors, $F \in D(\Phi_\alpha)$.

Théorème 1.3.4 — *$F \in D(\Psi_\alpha)$, pour $\alpha > 0$ si et seulement si $F^{-1}(1) < \infty$ et*

$$\lim_{h \downarrow 0} \frac{1 - F(F^{-1}(1) - xh)}{1 - F(F^{-1}(1) - h)} = x^\alpha \quad (= -\log \Psi_\alpha), \tag{1.3.3}$$

pour $x > 0$.

Les domaines d'attraction de la loi des valeurs extrêmes sont définis par le théorème suivant :

Théorème 1.3.5 — *Sous certaines conditions de régularité sur F, il existe $\alpha \in \mathbb{R}$ et deux suites réelles a_n et b_n $(b_n > 0)$ tels que $\forall x \in \mathbb{R}$,*

$$\lim_{n \to \infty} F^n(a_n + b_n x) = H_\alpha(x), \tag{1.3.4}$$

où H_α est la fonction de répartition de la loi des valeurs extrêmes

$$H_\alpha(x) = \exp\left[-(1 + \alpha x)^{-1/\alpha}\right].$$

On dit alors que F appartient au domaine d'attraction de H_α : $F \in D(\alpha)$.

Ce théorème nous donne, aprés normalisation, la loi asymptotique du maximum de l'échantillon (de fonction de répartition F^n).

Théorème 1.3.6 — *Soit F une fonction de distribution, d'une densité f, et*

$$h(x) = \frac{f(x)}{(1 - F(x))}. \tag{1.3.5}$$

Si $h(x) > 0$, $\forall x$ et pour $\alpha > 0$,

$$\lim_{x \to \infty} x h(x) = \alpha,$$

alors $F \in D(\Phi_\alpha)$.

Théorème 1.3.7 — *Supposons que $h(x) \neq 0$, et differentiable pour $x > F^{-1}(1)$. Alors, $F \in D(\Lambda)$ si*

$$\lim_{x \to F^{-1}(1)} \frac{d}{dx} \left\{ \frac{1}{h(x)} \right\} = 0. \tag{1.3.6}$$

Théorème 1.3.8 — *Si* $F^{-1}(1) < \infty$, *et pour* $\alpha > 0$,

$$\lim_{x \to F^{-1}(1)} \left(F^{-1}(1) - x\right) h(x) \quad = \quad \alpha, \tag{1.3.7}$$

alors $F \in D(\Psi_\alpha)$.

1.4 Loi du logarithme itéré

Nous donnerons seulement la preuve de la loi du logarithme itéré dans un cas spécial, où les vaiables aléatoires X_n sont [iid] d'une disribution normal stanard.

Définition 1.4.1— Une variable aléatoire $X \in \mathcal{L}$ est symétrique si sa loi μ_X est donnée par :

$$\mu([-b, -a)) = \mu([-a, -b)), \text{ pour tous } a < b.$$

Une variable aléatoire $X \in \mathcal{L}^1$ a une moyenne nulle. Nous utilisons la notation :

$$S_n = \sum_{k=1}^{n} X_k.$$

Lemme 1.4.1— Soit X_n, par symétrie et indépendance.Pour tous $\varepsilon > 0$.

$$\mathcal{P}\left(\max_{1 \leq k \leq n} S_k > \varepsilon\right) \leq 2\mathcal{P}(S_n > \varepsilon).$$

Preuve. C'est une conséquence directe de théorème de Levy, on peut prendre $m = 0$ comme mediane de la distribution symétrique. □

Définissons $\Lambda_n = \sqrt{2n \log\log n}$ (pour $n \geq 2$).

Théorème 1.4.1— *(Loi du logarithme itéré pour* $N(0,1)$*)*

Soit X_n une suite de variables aléatoires [iid] $\mathcal{N}(0,1)$-distribuées. Alors

$$\limsup_{n\to\infty} \frac{S_n}{\Lambda_n} = 1, \qquad \liminf_{n\to\infty} \frac{S_n}{\Lambda_n} = -1.$$

Preuve. Puisque le deuxiéme résultat s'obtient du premiére, en remplaçant X_n par $-X_n$, nous donnerons seulement la preuve du $\limsup\limits_{n\to\infty} \dfrac{S_n}{\Lambda_n} = 1$.

(i) $\mathcal{P}\left(S_n > (1+\varepsilon)\Lambda_n,\ i.s\right) = 0 \qquad$ pour tous $\varepsilon > 0$.

Soient $n_k = \left[(1+\varepsilon)^k\right] \in \mathbb{N}$ et les évènements

$$A_k = \left\{S_n > (1+\varepsilon)\Lambda_n,\ \text{pour quelque } n \in (n_k, n_{k+1})\right\}.$$

Evidement, on a

$$\limsup_k A_k = \left\{S_n > (1+\varepsilon)\Lambda_n,\ i.s\right\}.$$

Par le théorème de Borel Cantelli, nous montrons que $\sum\limits_k \mathcal{P}(A_k) < \infty$.

Pour chaque k, nous obtenons par le lemme au dessus

$$\begin{aligned}
\mathcal{P}(A_k) &= \mathcal{P}\left[\left\{S_n > (1+\varepsilon)\Lambda_n,\ n \in (n_k, n_{k+1})\right\}\right] \\
&\leq \mathcal{P}\left(\max_{n_k < n \leq n_{k+1}} S_n > (1+\varepsilon)\Lambda_k\right) \\
&\leq \mathcal{P}\left(\max_{1 < n \leq n_{k+1}} S_n > (1+\varepsilon)\Lambda_k\right) \\
&\leq 2\mathcal{P}\left(S_{n_{k+1}} > (1+\varepsilon)\Lambda_k\right).
\end{aligned}$$

En utilisons le fait que $S_{n_{k+1}} \big/ \sqrt{n_{k+1}}$ est $\mathcal{N}(0,1)$-distribuée et qu'une variable aléatoire $\mathcal{N}(0,1)$-distribuée

$$\mathcal{P}(X > t) \leq const \cdot e^{-t^2/2}, \text{ alors}$$

$$\begin{aligned}
2\mathcal{P}\left(S_{n_{k+1}} > (1+\varepsilon)\Lambda_k\right) &= 2\mathcal{P}\left(\frac{S_{n_{k+1}}}{\sqrt{n_{k+1}}} > (1+\varepsilon)\frac{\Lambda_{n_k}}{\sqrt{n_{k+1}}}\right) \\
&= 2\mathcal{P}\left(\frac{S_{n_{k+1}}}{\sqrt{n_{k+1}}} > (1+\varepsilon)\frac{\sqrt{2n_k \log\log n_k}}{\sqrt{n_{k+1}}}\right) \\
&\leq C\exp\left[-1/2\,(1+\varepsilon)^2 \frac{2n_k \log\log n_k}{n_{k+1}}\right] \\
&\leq C_1 \exp\left(-(1+\varepsilon)\log\log(n_k)\right) \\
&\leq C_1 \log(n_k)^{-(1+\varepsilon)} \\
&\leq C_2 k^{-(1+\varepsilon)}.
\end{aligned}$$

On a montré que $\mathcal{P}(A_k) \leq Const\cdot k^{-(1+\varepsilon)}$. Par la suite, on a $\sum\limits_k \mathcal{P}(A_k) < \infty$.

(ii) $\mathcal{P}\left(S_n > (1-\varepsilon)\Lambda_n,\ i.s \text{ (infiniment souvent)}\right) = 1$ pour tous $\varepsilon > 0$.

Il suffit de montrer que pour tous $\varepsilon > 0$, il existe une sous suite n_k.

$$\mathcal{P}\left(S_{n_k} > (1-\varepsilon)\Lambda_{n_k},\ i.s\right) = 1.$$

Donnons $\varepsilon > 0$. Choisissons $N > 1$ grand et $c < 1$ prés de 1 tel que

$$c\sqrt{1-1/N} - \frac{2}{\sqrt{n}} > 1-\varepsilon \tag{1.4.1}$$

Soit $n_k = N^k$ et $\Delta_{n_k} = n_k - n_{k-1}$. Les ensembles

$$A_k = \left\{S_{n_k} - S_{n_{k-1}} > c\sqrt{2\Delta_{n_k}\log\log\Delta_{n_k}}\right\},$$

sont indépendantes. Nous utilisons le fait que $\int\limits_t^\infty e^{-x^2/2}dx \geq Ce^{-t^2/2}$ pour quelque constante C.

$$\begin{aligned}
\mathcal{P}(A_k) &= \mathcal{P}\left(\left\{S_{n_k} - S_{n_{k-1}} > c\sqrt{2\Delta_{n_k}\log\log\Delta_{n_k}}\right\}\right) \\
&= \mathcal{P}\left(\left\{\frac{S_{n_k} - S_{n_{k-1}}}{\sqrt{\Delta_{n_k}}} > c\frac{\sqrt{2\Delta_{n_k}\log\log\Delta_{n_k}}}{\sqrt{\Delta_{n_k}}}\right\}\right) \\
&\geq C\exp\left(-c^2\log\log\Delta_{n_k}\right) \\
&= C\exp\left(-c^2\log\left(k\log N\right)\right) \\
&= C_1\exp\left(-c^2\log k\right) \\
&= C_1 k^{-c^2}.
\end{aligned}$$

Donc $\sum\limits_k \mathcal{P}(A_k) = \infty$. Nous avons par Borel-Cantelli l'ensemble A de mesure complet, alors pour $\omega \in A$

$$S_{n_k} - S_{n_{k-1}} > c\sqrt{2\Delta_{n_k}\log\log\Delta_{n_k}},$$

pour k infini. De (i), on sait que

$$S_{n_k} > -2\sqrt{2n_k\log\log n_k},$$

pour k suffisamment large. Les deux inégalités sont vérifiées pour quelques valeurs infinies de k. Nous avons pour un tel k

$$\begin{aligned}
S_{n_k}(\omega) &> S_{n_{k-1}}(\omega) + c\sqrt{2\Delta_{n_k}\log\log\Delta_{n_k}} \\
&\geq -2\sqrt{2n_{k-1}\log\log n_{k-1}} + c\sqrt{2\Delta_{n_k}\log\log\Delta_{n_k}} \\
&\geq \left(-2/\sqrt{N} + c\sqrt{1-1/N}\right)\sqrt{2n_k\log\log n_k} \\
&\geq (1-\varepsilon)\sqrt{2n_k\log\log n_k},
\end{aligned}$$

où, nous avons utilisé (1.4.1) dans la dernière inégalitè. $\square$

1.5 Comportement presque sûre du processus des quantiles uniforme pondérés

Nous proposons dans ce paragraphe une description complète de Comportement presque sûre du processus des quantiles uniforme pondérés.

Soit $U_1, U_2, ...$ une suite de variables aléatoires indépendantes uniforme sur $(0,1)$. Pour chaque entier $n \geq 1$, on définit

$$\mathbb{G}_n(s) \quad := \quad n^{-1} \sum_{i=1}^{n} \mathbf{1}(U_i \leq s), \quad 0 \leq s \leq 1,$$

la fonction de distribution empirique, où $\mathbf{1}(x \leq y)$ désigne la fonction indicatrice, et soit

$$\mathbb{V}_n(s) = \begin{cases} U_{k,n}, & k-1/n < s \leq k/n, \qquad k = 1, ..., n \\ \\ U_{1,n}, & s = 0, \end{cases}$$

où $U_{1,n} \leq U_{2,n} \leq ... \leq U_{n,n}$ la statistique d'ordre basée sur $U_1, U_2, ..., U_n$.
la fonction des quantiles.

On définit le processus empirique uniforme[PEU] comme suite :

$$\alpha_n(s) \quad := \quad n^{1/2} \{\mathbb{G}_n(s) - s\}, \quad 0 \leq s \leq 1$$

et le processus des quantiles uniforme par :

$$\beta_n(s) \quad := \quad n^{1/2} \{\mathbb{V}_n(s) - s\}, \quad 0 \leq s \leq 1.$$

Notons par :

$$\tilde{\beta}_n(s) = \begin{cases} \beta_n(s), & \text{si} \qquad 1/(n+1) \leq s \leq n/(n+1) \\ 0, & \text{si non} \end{cases}$$

le processus des quantiles uniforme tronqué.

1.5.1 Lois fonctionnelles du logarithme itéré pour le processus empirique des quantiles pondérés

On considère l'ensemble suivant :

$$\begin{aligned}\mathfrak{D} &= \{d : [0,1] \to [0,\infty) : \\ &\quad d \text{ est non décroissante strictement positive sur } [0,1]\},\end{aligned}$$

et on désigne par $\mathfrak{B}[0,1]$ l'espace des fonctions bornées, et définies sur $[0,1]$

avec la norme *"sup", et* $\mathfrak{F}[0,1]$ l'ensemble des fonctions absolument continues $f \in \mathfrak{B}[0,1]$, et

$$f(0) = 0, \quad \text{et} \quad \int_0^1 (f'(s))^2 \, ds \leq 1.$$

Définissons l'ensemble $\mathfrak{F}_d[0,1]$, Pour tous $d \in \mathfrak{D}$, comme suit

$$\mathfrak{F}_d[0,1] = \{f \diagup d : f \in \mathfrak{F}[0,1]\},$$

et posons

$$h_n = \log\log(n \vee 3), \quad \text{pour} \quad n \geq 1.$$

Théorème 2.1.1 — *Soit $d \in \mathfrak{D}$, et supposons que*

$$\lim_{s \downarrow 0} (s \log\log(1/s))^{1/2} \diagup d(s) = \tau \in [0,\infty]. \tag{1.5.1.1}$$

Alors, avec une probabilité 1

$$\limsup_{n\to\infty} \sup_{0<s\leq 1} \left|\tilde{\beta}_n(s)\right| \diagup \left(d(s)(2h_n)^{1/2}\right)$$

$$= \left(2^{-1/2}\tau\right) \vee \left(\sup_{0<s\leq 1/2} \left(s(1-s)^{1/2} \diagup d(s)\right)\right)$$

De plus, si $\tau = 0$, alors presque sûrement la suite

$$\left\{\tilde{\beta}_n(s) \diagup ((2h_n)^{1/2} d)\right\}_{n=1}^{\infty} \tag{1.5.1.2}$$

est relativement compacte dans $\mathfrak{B}[0,1]$, et l'ensemble des points limites égale à $\mathfrak{F}_d[0,1]$.

Si $\tau > 0$, presque sûrement la suite (1.5.1.1) ne peut pas être relativement compacte dans $\mathfrak{B}[0,1]$

Pour tout $0 < c < \infty$, soient $0 < \alpha_c^- < 1 < \alpha_c^+$ les deux solutions de l'équation

$$\lambda - \log \lambda - 1 = c^{-1}.$$

Théorème 2.1.2 — *Soit $0 \le \nu \le 1/2$ et $0 < a_n < 1/2$ avec $a_n \downarrow 0$.*

(I) Si $na_n/h_n \to 0$ et $a_n \ge a/n$, pour $a > 0$, alors presque sûrement,

$$\limsup_{n\to\infty} \sup_{a_n \le s \le 1/2} \sqrt{n} a_n^{1-\nu} \left|\beta_n(s)\right| \left(s^{1-\nu} h_n\right)^{-1} = 1 \tag{1.5.1.3}$$

(II) Si $na_n/h_n \to c \in (0,\infty)$, alors presque sûrement,

$$\limsup_{n\to\infty} \sup_{a_n \le s \le 1/2} a_n^{1-\nu} \left|\beta_n(s)\right| \left(s^{1-\nu} h_n^{1/2}\right)^{-1} = \gamma(c,\nu), \tag{1.5.1.4}$$

où

$$\begin{cases} \gamma(c,1/2) &= 2 \vee \left\{c^{1/2}\left(\alpha_c^+ - 1\right)\right\}, \\ \gamma(c,\nu) &= c^{1/2}\left(\alpha_c^+ - 1\right), \qquad \text{si} \quad 0 \le \nu \le 1/2. \end{cases}$$

(III) Si $na_n/h_n \uparrow (1/a_n) \diagup h_n \to c$, alors presque sûrement

$$\begin{aligned} \limsup_{n\to\infty} \sup_{a_n \le s \le 1/2} a_n^{1-\nu} \left|\beta_n(s)\right| \left(s^{1-\nu} h_n^{1/2}\right)^{-1} &= \left[2(1+c)\right]^{1/2}, \quad \nu = 1/2, \\ &= 2^{1/2}, \qquad 0 \le \nu < 1/2. \end{aligned}$$

Ce théorème nous donne une discription du comportement presque sûre du processus des quantiles uniforme pondéré pour une certaine classe de fonction $d \in \mathfrak{D}$ afin que la limite dans (1.5.1.1) est égale à ∞.

Soit

$$0 < k_n \le n, \quad k_n \uparrow \quad \text{et} \quad k_n/n \downarrow 0.$$

Définissons deux versions du processus empirique des quantiles des queues basés sur la suite $\{k_n\}_{n=1}^{\infty}$ comme suit :

$$\upsilon_n(s) \;:=\; (n/k_n)^{1/2} \beta_n(sk_n/n), \quad 0 \le s \le 1,$$

et

$$\tilde{\upsilon}_n(s) \ := \ (n/k_n)^{1/2}\,\tilde{\beta}_n(sk_n/n)\,, \quad 0 \le s \le 1.$$

Soit $\mathcal{K}[0,1]$ l'ensemble des fonctions absolument continues $f \in \mathfrak{B}[0,1]$ tel que

$$f(0) = 0 \quad \text{et} \quad \int_0^1 \left(f'(s)\right)^2 ds \le 1.$$

On pose pour tous $0 < \nu \le 1/2$,

$$\mathcal{K}_\nu[0,1] \ = \ \left\{ f I^{1/2-\nu} : f \in \mathcal{K}[0,1] \right\}$$

où I est la fonction identité.

Théorème 2.1.3 — *Soit $0 \le \nu \le 1/2$, $1 < k_n < n$, $k_n \uparrow$ et $k_n/n \downarrow 0$.*
(I) Si $k_n^{2\nu}/h_n \to 0$, alors presque sûrement,

$$\limsup_{n\to\infty} \sup_{0\le s\le 1} k_n^\nu \left|\tilde{\upsilon}_n(s)\right| \left(s^{1/2-\nu} h_n\right)^{-1} \ = \ 1. \tag{1.5.1.5}$$

(II) Si $k_n^{2\nu}/h_n \to c \in (0,\infty)$, alors presque sûrement,

$$\limsup_{n\to\infty} \sup_{0\le s\le 1} \left|\tilde{\upsilon}_n(s)\right| \left(s^{1/2-\nu} h_n^{1/2}\right)^{-1} \ = \ \varrho(c,\nu)\,, \tag{1.5.1.6}$$

où

$$\begin{array}{rcllll} \varrho(c,\nu) & = & c^{1/2}\left(\alpha_c^+ - 1\right), & \nu = 1/2, & c \in (0,\infty)\,, \\ & = & 2^{1/2}, & \nu = 1/2, & c = \infty, \\ & = & c^{-1/2}, & 0 < \nu < 1/2, & c \in (0,1/2)\,, \\ & = & 2^{1/2}, & 0 < \nu < 1/2, & c \in (1/2,\infty)\,. \end{array}$$

(III) Si $k_n^{2\nu}/h_n \to \infty$, alors presque sûrement, la suite

$$\left\{ \tilde{\upsilon}_n(s) \left[I^{1/2-\nu} (2h_n)^{1/2} \right]^{-1} \right\}_{n=1}^{\infty}, \tag{1.5.1.7}$$

est relativement compacte sur $\mathfrak{B}[0,1]$ et l'ensemble de points limites égale à $\mathcal{K}_\nu[0,1]$.

De plus, $(I) - (III)$ restent valables si on remplace $\tilde{\upsilon}_n$ par υ_n pour le cas $\nu = 1/2$.

Ensuite nous avons besoin des résultats suivants :

Fait1 [Kiefer (1972)]

(1) *Soit* $(n+1)^{-1} \leq a_n < 1$ *et* $na_n / h_n \to 0$. *Alors presque sûrement,*

$$\limsup_{n\to\infty} n\mathbb{V}_n(a_n)\,h_n^{-1} \;=\; 1.$$

(2) *Soit* $0 < a_n < 1$, $na_n / h_n \uparrow \infty$ *et* $a_n \downarrow$. *Alors presque sûrement,*

$$\limsup_{n\to\infty} \pm\beta_n(a_n)\,(a_n(1-a_n)\,h_n)^{-1/2} \;=\; 2^{1/2}.$$

(3) *Soit* $0 < a_n < 1$, $na_n / h_n \uparrow \infty$ *et* $a_n \downarrow$. *Alors presque sûrement,*

$$\limsup_{n\to\infty} \pm\alpha_n(a_n)\,(a_n(1-a_n)\,h_n)^{-1/2} \;=\; 2^{1/2}.$$

(4) *Si* $c \in (0,\infty)$, *alors presque sûrement,*

$$\limsup_{n\to\infty} n\mathbb{V}_n(ch_n/n)\,h_n^{-1} \;=\; c\alpha_c^+,$$

et

$$\liminf_{n\to\infty} n\mathbb{V}_n(ch_n/n)\,h_n^{-1} \;=\; c\alpha_c^-.$$

Fait 2 [Wellner (1978)]

Si $c \in (0,1)$, *alors presque sûrement,*

$$\limsup_{n\to\infty} \sup_{ch_n/n\leq s\leq 1} \mathbb{V}_n(s)/s \;=\; \alpha_c^+,$$

et

$$\limsup_{n\to\infty} \sup_{ch_n/n\leq s\leq 1} s/\mathbb{V}_n(s) \;=\; (\alpha_c^-)^{-1}.$$

Fait 3 [Kiefer (1970)]

On a, presque sûrement

$$\limsup_{n\to\infty} \sup_{0\leq s\leq 1} n^{1/4}\left|\alpha_n(s)+\beta_n(s)\right|\left(h_n^{1/2}\log_n\right)^{-1/2} \;=\; 2^{-1/4}.$$

Fait 4 [Csáki (1977)]

(1) *Soit* $0 < a_n \leq 1$, $na_n / h_n \to \infty$, $\log\log(1/a_n) / h_n \to c$ *et* $a_n \downarrow$. *Alors presque sûrement,*

$$\limsup_{n\to\infty} \sup_{a_n\leq s\leq 1/2} |\alpha_n(s)|\,(sh_n)^{-1/2} \;=\; (2(1+c))^{1/2}$$

(2) *Pour* $c \in (0,\infty)$, *presque sûrement*

$$\limsup_{n\to\infty} \sup_{ch_n/n\leq s\leq 1} |\alpha_n(s)| (sh_n)^{-1/2} = 2 \vee \left\{c^{1/2}\left(\beta_c^+ - 1\right)\right\},$$

où $\beta_c^+ > 1$ est la racine la plus grande de l'équation $\lambda(\log\lambda - 1) + 1 = c^{-1}$.

Le théorème suivant nous permet de construire une approximation forte du processus empirique des quantiles uniforme pondérés.

Théorème 2.2.1 —

(I) *Sur un espace probabilisé assez riche, il existe une suite de ponts Brownien indépendants* $B_1, B_2, ...,$ *et une suite de variables aléatoires uniformes sur* $(0,1)$*, tellque* $d \in \mathfrak{D}$ *satisfait* (1.5.1.1) avec $\tau = 0$, *alors on a presque sûrement,*

$$\sup_{0\leq s\leq 1} \left|\tilde{\beta}_n(s) - n^{1/2}\sum_{i=1}^{n} B_i(s)\right| \diagup d(s) = o\left(h_n^{1/2}\right). \tag{1.5.1.8}$$

(II) *Soit* $0 < \nu \leq 1/2$, $1 \leq k_n \leq n$, $k_n \uparrow$ *et* $k_n \diagup n \downarrow 0$. *Si* $k_n^{2\nu} \diagup h_n \to \infty$, *alors sur un espace de probabilité assez riche, il existe une suite de processus de Wiener standard* $W_1, W_2, ...,$ *et une suite* $U_1, U_2, ...,$ *de variables aléatoires uniformes et indépendantes sur* $(0,1)$*, telle que, presque sûrement*

$$\sup_{0\leq s\leq 1} \left|\tilde{v}_n(s) - k_n^{-1/2}\sum_{i=1}^{n} W_i(sk_n \diagup n)\right| s^{-1/2+\nu} = o\left(h_n^{1/2}\right). \tag{1.5.1.9}$$

Preuve. Voir [10]. □

Chapitre 2

Estimation de l'index des queues de distribution

2.1 Estimateur de Pickands

L'estimateur introduit par JAMES PICKANDS III dans Pickands (1975) est basé sur la condition suivante

$$\lim_{n\to\infty} F^n(a_n + b_n) = \lim_{n\to\infty} \mathbb{P}\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = G(x), \tag{2.1.1}$$

où

$$M_n = \max(X_1, X_2, ..., X_n),$$

et F est une distribution continue, d'une point limite supèrieur $F^{-1}(1)$.

On dit que F est dans le domaine d'attraction de G_α pour $\alpha \in \mathbb{R}$, si et seulement si

$$\lim_{t\uparrow F^{-1}(1)} \sup_{0\leq x\leq F^{-1}(1)-t} |F_t(x) - F_{DPG}(x;\alpha,\beta(t))| = 0,$$

où F_{DPG} est la distribution de Paréto généralisée, définie pour $\sigma > 0$ et $\alpha \in \mathbb{R}$ comme suit :

$$F_{DPG}(x;\alpha,\sigma) = 1 - (1 + \alpha x/\sigma)^{-1/\alpha},$$

avec

$$x > 0 \quad \text{et} \quad +\alpha x/\sigma > 0,$$

et F_t la probabilité conditionnelle de X qui ne dépasse pas $x+t$ donnée (il a déja dépassé t) c'est à dire que si le seuil t est grand, la distribution conditionnelle de X, telle que $X \geq tn$ peut être approximé par une distribution de Paréto généralisée.

La fonction des quantiles de la distribution de Paréto généralisée est donnée

$$Q_{DPG}(s;\alpha,\sigma) \;=\; \sigma \int\limits_0^{-\log(1-s)} e^{\alpha u}du \quad \begin{cases} \sigma\alpha^{-1}(1-s)^{-\alpha} & \alpha \neq 0, \\ -\sigma\log(1-s) & \alpha = 0, \end{cases}$$

pour $\alpha \in \mathbb{R}$ et $\sigma > 0$.

En utilisant cette formule, nous pouvons exprimer les paramètres α et σ en terme du quantiles $Q_{DPG}(1/2;\alpha,\sigma)$ et $Q_{DPG}(3/4;\alpha,\sigma)$

$$\alpha \;=\; \log\left(\frac{Q_{DPG}(3/4;\alpha,\sigma) - Q_{DPG}(1/2;\alpha,\sigma)}{Q_{DPG}(1/2;\alpha,\sigma)}\right) \diagup \log 2,$$

et

$$\sigma \;=\; (Q_{DPG}(1/2;\alpha,\sigma)) \diagup \int\limits_0^{\log 2} \exp(\alpha u)du.$$

2.1.1 Construction de l'estimateur

Pickands a proposé d'estimer les paramètres α et σ par la méthode du centile simple :

Soit M un entier plus petit que n, et on considère les espacements

$$E(M) := X_{(n-i+1)} - X_{(n-4M+1)},\ i = 1, ...,\ 4M-1,$$

sur le seuil $X_{(n-4M+1)}$ qui est un échantillon d'une distribution de Paréto généralisée. Alors, on utilise le quantile de l'échantillon de ces excés pour l'estimation de α et σ, c'est à dire,

$$\hat{\alpha}^P_{n,M} \;=\; \log\left(\frac{X_{(n-M+1)} - X_{(n-2M+1)}}{X_{(n-2M+1)} - X_{(n-4M+1)}}\right) \diagup \log 2,$$

et

$$\hat{\sigma}^P_{n,M} \;=\; (X_{(n-2M+1)} - X_{(n-4M+1)}) \diagup \int\limits_0^{\log 2} \exp\left(\hat{\alpha}^P_{n,M}u\right)du.$$

Les estimateurs dépendent du choix du l'entier M, et selon Pickands on a

$$M \xrightarrow{\mathcal{P}} \infty, \quad \text{et} \quad M/n \to 0, \quad \text{quand} \quad n \uparrow \infty.$$

2.1.2 Consistance de l'estimateur de Pickands

Consistance faible

Théoréme 2.1.2 — *Soit $m = m_n$ une suite d'entier telle que*

$$m_n \to \infty \quad et \quad m_n/n \to 0, \quad quand \quad n \to \infty.$$

Si (2.1.1) est vérifie, alors

$$\hat{\alpha}^P_{n,m} \xrightarrow{\mathcal{P}} \alpha, \quad quand \quad n \to \infty.$$

Consistance forte

Théoréme 2.1.3 — *Soit $m = m_n$ une suite d'entier telle que*

$$m_n/n \to 0 \quad et \quad m_n/\log\log n \to \infty, \quad quand \quad n \to \infty.$$

Si (2.1.1) est vérifie, alors

$$\hat{\alpha}^P_{n,m} \to \alpha, \quad presque\ sûr, \quad quand \quad n \to \infty.$$

2.1.3 Normalité asymptotique de $\hat{\alpha}^P_{n,m}$

Théorème 2.1.3 — *Soit F une fonction de distribution dans le domaine d'attraction d'une distribution des valeurs extrêmes G_α, pour $\alpha \in \mathbb{R}$. Supposons que la fonction des quantiles $Q = F^{-1}$ a un dérivé positif Q', et qu'il existe une fonction positive a, telle que la fonction*

$$t \mapsto t^{-\alpha-1} Q'(1 - 1/t),$$

est $\pi-$variation à l'infini, d'une fonction auxilière a. Alors pour une suite $m = m_n$ satisfait

$$m_n \to \infty, \quad m_n = o\left(n / g^{-1}(n)\right),$$

où g^{-1}est l'inverse de

$$g(t) \quad = \quad t^{-2\alpha-1}(Q'(1 - 1/t) / a(t))^2,$$

$$\sqrt{m}\left(\hat{\alpha}^P_{n,m} - \alpha\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_p^2\right) \quad quand \quad n \to \infty,$$

où $\mathcal{N}\left(0, \sigma_p^2\right)$ désigne la distribution normale de moyenne 0 et variance σ_p^2 donnée par

$$\sigma_p^2 \quad = \quad \alpha^2 \frac{\left(2^{2\alpha-1}\right)}{\left(2\left(2^\alpha - 1\right)\log 2\right)^2}$$

2.2 Estimateur de Hill

En 1975 un autre estimateur de l'index de la valeur extrême a été introduit par B. M. Hill dans (*Hill* 1975). Cet estimateur est basé sur les propriétés de variation régulière dans le cas où la fonction de distribution F dans le domaine d'attraction de G_α, où $\alpha > 0$. L'estimateur de *Hill* serait seulement applicable de la valeur extrême, au cas où l'index serait positif, c à d F a une queue lourde.

La construction de cet estimateur est basée sur l'estimation par la méthode du maximum de vraisemblance en supposant que $\alpha > 0$, et F vérifie la condition

$$\frac{1 - F(x)}{1 - F(u)} = \left(\frac{x}{u}\right)^{-1/\alpha}, \quad x \geq u, \quad u \in \mathbb{R}. \tag{2.2.1}$$

Alors, on utilise seulement les statistiques d'ordre qui depassent le seuil u.

On a, les statistiques d'ordre $X_{n,m}, ..., X_{n-k_n+1,n}$. La fonction de log-vraisem-blance est donné par

$$\begin{aligned} L(\alpha; X_{n-k_n+1,n,...}, X_{n,n}) &= -k_n \log(\alpha u) + \log(1 - F(u)) - \frac{\alpha + 1}{\alpha} \\ &\times \sum_{i=1}^{k_n} (\log X_{n-i+1,n} - \log u) \end{aligned} \tag{2.2.2}$$

En resolvant cette equation et remplaçant le seuil u par $X_{n-k,n}$, on obtient l'estimateur de Hill

$$\hat{\alpha}^H_{n,k_n} = \frac{1}{k_n} \sum_{i=1}^{k_n} (\log X_{n-i+1,n} - \log X_{n-k,n}). \tag{2.2.3}$$

Notons que cet estimateur peut être écrit en terme des espacements des logarithmes des observations.

$$\hat{\alpha}^H_{n,k_n} = \frac{1}{k_n} \sum_{i=1}^{k_n} i \left(\log X_{n-i-1,n} - \log X_{n-i,n}\right). \tag{2.2.4}$$

2.2.1 Consistance de l'estimateur de Hill

Mason (1982 a) a prouvé la consistance faible de l'estimateur de *Hill* pour toute suite k_n vérifiant

$$k_n \to \infty, \quad \text{et} \quad k_n/n \to 0 \quad \text{quand} \quad n \to \infty.$$

La consistance forte a été prouvée dans *Deheuvels, Haeusler et Mason* (1985) sous les conditions

$$k_n/\log\log n \to \infty, \quad \text{et} \quad k_n/n \to 0, \quad n \to \infty.$$

2.2.2 Normalité asymptotique de $\hat{\alpha}^H_{n,k_n}$

Dans l'article de *Hill* (1975), la normalité asymptotique n'a pas été établit pour cet estimateur. La normalité asymptotique peut être établie sous certaines conditions supplémentaires, (voir *Haeusler et Teugels* (1985))

Nous affirmerons le résultat de la normalité asymptotique qui est obtenu comme un cas spécial de la preuve dans *Dekkers, Einmahl et de Hann* (1989), de la normalité asymptotique de leur estimateur du moment (voir aussi la section 2.5 de la thèse de PhD de *Paul de Wolf*).

Théorème 2.2.2 — *Soit F une fonction de distribution qui appartient au domaine d'attraction de la distribution des valeurs extrêmes G_α avec $\alpha > 0$. Soit Q sa fonction des quatiles. supposons qu'il existe une fonction positive b, telle que la fonction*

$$t \mapsto \pm t^{-\alpha} Q\left(1 - 1/t\right)$$

(avec l'un et l'autre de choix de signe) est π-variation à l'infini avec une fonction auxiliaire b. Alors, pour toute suite $k_n \to \infty$ satisfaisant $k_n = o\left(n/g^{-1}(n)\right)$, où g^{-1} est l'inverse de

$$g(t) \quad = \quad t^{1-2\alpha}\left(Q\left(1-1/t\right)/b\left(t\right)\right)^2,$$

$$\sqrt{k_n}\left(\hat{\alpha}_{k_n,n} - \alpha\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \alpha^2\right), \quad \textit{quand} \quad n \to \infty,$$

avec $\mathcal{N}\left(0, \alpha^2\right)$ la distribution normale de moyenne 0 et de variance α^2.

Preuve. Voir [20]. $\square$

2.3 Estimateur du moment

2.3.1 Construction de l'estimateur du moment

Une généralisation de l'estimateur de Hill était introduite dans Dakkers, Einmahl et de Hann (1989). Leur but principal est de trouver un estimateur dans certaine sens semblable à l'estimateur de Hill qui est consistant pour tout $\alpha \in \mathbb{R}$. Ils ont introduit les deux quantités suivantes

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

Notons que $M_{n,k}^{(1)}$ est l'estimateur de *Hill*. Alors ils ont proposé l'estimateur de l'index des valeurs extrêmes suivant :

$$\hat{\alpha}_{n,k}^{M} = M_{n,k}^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_{n,k}^{(1)}\right)^{2}}{M_{n,k}^{(2)}}\right)^{-1}. \tag{2.3.1.1}$$

Nous donnerons une dérivation heuristique de cet estimateur du moment, en utilisant le lemme suivant (voir. *Dekkers, Einmahl et de Hann*(1989), Lemme 2.5).

Lemme 2.3.1 — *Supposons que $F \in D(G_\alpha)$, pour $\alpha \in \mathbb{R}$, on note par Q la fonction des quantiles. Alors, pour quelque fonction positive a,*

$$\lim_{s \downarrow 0} \frac{\log Q(1 - sx) - \log Q(1 - s)}{a(s)/Q(1 - s)}$$

$$= \begin{cases} -\log x & \alpha \geq 0 \\ \dfrac{x^{-\alpha} - 1}{\alpha} & \alpha < 0 \end{cases} \qquad \text{pour } x > 0.$$

De plus, dans le cas $\alpha > 0$ nous pouvons prendre

$$a(s)/Q(1 - s) = \alpha,$$

et dans le cas $\alpha < 0$ nous pouvons prendre

$$a(s)/Q(1 - s) = -\alpha\left(\log Q(1) - \log Q(1 - s)\right).$$

Notons que $M_{n,k}^{(r)}$ peut être envisagé comme l'esperance conditionnel de $(\log (X/t))^r$ sachant que $X > t$. Dans ce cas le seuil t pres place la $(n-k)^{ème}$ statistique d'ordre de l'échantion. En effet

$$\begin{aligned} E\left((\log (X/t))^r / X > t\right) &= \int_t^{F^{-1}(1)} (\log (x/t))^r \, d\left(\frac{F(x) - F(t)}{1 - F(t)}\right) \\ &= -\int_t^{F^{-1}(1)} (\log (x/t))^r \, d\left(\frac{1 - F(x)}{1 - F(t)}\right). \end{aligned}$$

La substitution de $y = 1 - F(x)$ et $s = 1 - F(t)$ donne

$$\begin{aligned} &\int_0^s (\log Q(1-y) - \log Q(1-s))^r d(y/s) \\ &= \int_0^1 (\log Q(1-sx) - \log Q(1-s))^r dx. \end{aligned}$$

Notons le dernier intégral par $M_s^{(r)}$, $M_{n,k}^{(r)}$ pourait être considéré pour être un estimateur de $M_s^{(r)}$ avec $s = k/n$. Maintenant, on utilise le lemme (2.3.1) pour déduire que, dans le cas où $\alpha > 0$,

$$M_s^{(r)} \sim \alpha^r \int_0^1 (-\log x)^r \, dx = r\alpha^r, \quad r = 1, 2,$$

et dans le cas où $\alpha < 0$,

$$\begin{aligned} M_s^{(r)} &\sim (-\alpha)^r \left((\log Q(1) - \log Q(1-s))^r\right) \int_0^1 \left(\frac{x^{-\alpha} - 1}{\alpha}\right)^r dx \\ &= (\log Q(1) - \log Q(1-s))^r \frac{r(-\alpha)^r}{(1 - r\alpha)(1 - (r-1)\alpha)}, \quad r = 1, 2. \end{aligned}$$

Alors, la substitution de ces expressions dans (2.3.1.1) i.e,

$$\alpha_s := M_s^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_s^{(1)}\right)^2}{M_s^{(2)}}\right)^{-1},$$

satisfaite pour $\alpha > 0$.

$$\begin{aligned}\alpha_s &\sim \alpha + 1 - \frac{1}{2}\left(1 - \frac{\alpha^2}{2\alpha^2}\right)^{-1} \\ &= \alpha,\end{aligned}$$

dans le cas où $\alpha < 0$.

$$\begin{aligned}\alpha &\sim -\alpha\left(\log Q\left(1\right) - \log Q\left(1-s\right)\right)\frac{1}{1-\alpha} \\ &+ 1 - \frac{1}{2}\left(1 - \frac{1-2\alpha}{2(1-\alpha)}\right)^{-1} \\ &= \alpha + \frac{\alpha}{1-\alpha}\left(\log Q\left(1\right) - \log Q\left(1-s\right)\right) \\ &\sim \alpha,\end{aligned}$$

pour un petit s.

2.3.2 Normalité asymptotique de l'estimateur du moment

Théorème 2.4.1— *Soit F une fonction de distribution dans le domaine d'attraction d'une distribution de valeurs extrêmes G_α, pour $\alpha \in \mathbb{R}$. On note par Q la fonction des quantiles. Supposons qu'il existe des fonctions positives b_1, b_2, b_3* et *b_4 tel que pour tout $x > 0$ (avec le choix de signe).*

1. *dans le cas où $\alpha > 0$, la fonction $t \mapsto \pm t^{-\alpha} Q\left(1 - 1/t\right)$,est π-variation au voisinage de l'infini d'une fonction auxilière b_1.*
2. *Dans le cas où $\alpha = 0$,*
$$\lim_{t\to\infty} \frac{\log Q\left(1 - 1/\left(tx\right)\right) - \log Q\left(1 - 1/t\right) + b_2\left(t\right)\log x}{b_3\left(t\right)} = \pm\frac{\left(\log x\right)^2}{2}.$$
3. *Dans le cas où $\alpha < 0$, la fonction $t \mapsto \pm t^{-\alpha}\left(Q\left(1\right) - Q\left(1 - 1/t\right)\right)$,est π-variation au voisinage de l'infini d'une fonction auxilière b_4.*

Alors, pour la suite $k = k_n \to \infty$ *satisfaisant* $k_n = o\left(n/g^{-1}(n)\right)$, *où* g^{-1} *est l'inverse de*

$$g(t) = \begin{cases} t^{1-2\alpha}\left(Q(1-1/t)/b_1(t)\right)^2 & \alpha > 0, \\ \\ t\left(b_2(t)/b_3(t)\right)^2 & \alpha = 0, \\ \\ t^{1-2\alpha}\left(\log Q(t) - \log Q(1-1/t)/b_4(t)\right) & \alpha < 0, \end{cases}$$

et de plus, dans le cas où $\alpha = 0, k_n = o\left(n/g_t^{-1}(n)\right)$ *où* $g_t^{-1}(\cdot)$ *est l'inverse de*

$$g_t(t) = t\left(Q(1-1/t)/a(1/t)\right)^2,$$

avec la fonction a est définée dans le lemme(2.3.1)

$$\sqrt{k}\left(\hat{\alpha}_{n,k}^{M} - \alpha\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_M^2\right), \quad quand \quad n \to \infty,$$

avec $\mathcal{N}\left(0, \sigma_M^2\right)$ *désigne la distribution normale de moyenne 0 et variance* σ_M^2 *donné par :*

$$\sigma_M^2 = \begin{cases} 1+\alpha^2 & \alpha \geq 0, \\ \\ \left(1-\alpha^2\right)(1-2\alpha)\left(4 - 8\dfrac{1-2\alpha}{1-3\alpha} + \dfrac{(5-11\alpha)(1-2\alpha)}{(1-3\alpha)(1-4\alpha)}\right) & \alpha < 0. \end{cases}$$

Preuve. Voir [20]. □

2.4 Estimateur du moindre carée

Soit $X, X_1, X_2, ...$ une suite de variables aléatoires [v.a] indépendantes d'une fonction de distribution commune $F(x) = P(X \leq x)$, $x \in \mathbb{R}$ tel que $1 - F$ a variation régulière au voisinage de l'infinie, d'index $-1/\alpha$, $\alpha > 0$. Supposons que

$$1 - F(x) = x^{-1/\alpha}\mathbb{L}^*(x), \quad \text{quand} \quad x \to \infty,$$

où $\mathbb{L}^*$ est une fonction positive sur $[1, \infty)$, a variation lente au voisinage de l'infinie, alors

$$\frac{\mathbb{L}^*(tx)}{\mathbb{L}^*(x)} \to 1, \quad \text{quand} \quad x \to \infty,$$

pour tout $t > 0$.

Cette class de fonctions sera noté R_α. L'inverse ou la fonction des quantiles Q de F est définie par :

$$Q(s) = \inf\{x : F(x) \geq s\}, \quad 0 < s \leq 1, \quad Q(0) := Q(0+). \tag{2.4.1}$$

Il est bien connu (voir Csörgő et Mason (1985)) que $F \in R_\alpha$ si et seulement si

$$Q+(1-s) \;=\; \mathbb{U}(s) \;=\; s^{-\alpha}\mathbb{L}(s), \quad \text{quand} \quad s \downarrow 0, \tag{2.4.2}$$

pour la version continue à droite $Q+$ de Q, où $\mathbb{L}(\cdot)$ est une fonction positive a variation lente au voisinage de 0, i.e,

$$\frac{\mathbb{L}(ts)}{\mathbb{L}(s)} \to 1, \quad s \downarrow 0 \quad \text{pour tout} \quad t > 0.$$

Pour $n \geq 1$, soit $X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$ les statistiques d'ordre basé sur l'échantillon $X_1, X_2, ..., X_n$.

On note par $\log^+$ logarithm naturel log, où

$$\log^+ \;=\; \log[\max(x,1)], \quad x \in \mathbb{R}.$$

L'estimateur de Hill (Hill 1975) est définie par :

$$\hat{\alpha}^H(k_n) \;:=\; k_n^{-1}\sum_{i=1}^{k_n}\log\left(\frac{X_{n-1+1,n}}{X_{n-k_n,n}}\right),$$

où k_n est une suite d'entiers satisfaisants

$$1 \leq k_n < n, \quad k_n \to \infty \quad \text{et} \quad k_n/n \to 0. \tag{$\mathbb{K}$}$$

Csörgő, Deheuvels et Mason (1985) ont généralisé cet estimateur pour introduire une classe d'estimateurs à noyau.

En basant sur les considérations de moindre carées, Viharos (1999) a proposé un nouveau estimateur de l'index de la queue lequel est universellement asymptotiquement normal. Cet estimateur est défini par :

$$\hat{\alpha}^{(V)}(k_n) \;=\; \left(\sum_{i=1}^{k_n} a_{i,n}\log\left(\frac{k_n}{i}\right)\right)^{-1}\sum_{i=1}^{k_n} a_{i,n}\log X_{n-i+1,n}, \tag{2.4.3}$$

où $a_{i,n} := \int\limits_{i-1/k_n}^{i/k_n} \mathbb{J}(s)\,ds$, $i = 1, ..., k_n$, avec $\mathbb{J}(\cdot)$ est une fonction pondérée spécifie par les conditions (H1)-(H5), au dessous

(H1) $\int_0^1 \mathbb{J}(s)\, ds = 0;$

(H2) $\mathbb{J}(\cdot)$ est non croissante sur $(0,1]$ tel que $\lim_{x\downarrow 0} \mathbb{J}(x) > 0$ et $\mathbb{J}(1) < 0;$

(H3) $\mathbb{J}(\cdot)$ est differentiable d'une dérivé $\mathbb{J}'(\cdot)$ monotone sur $(0,1]$;

(H4) $\int_0^1 s^{-\upsilon-1/2} |\mathbb{J}(s)|\, ds < \infty$ pour $0 < \upsilon < 1/2;$

(H5) $\sup_{0<s<1} s\,|\mathbb{J}'(s)| < \infty.unesuited'entiertelleque\ m_n/n \to 0 \quad et \quad m_n/\log\log n \to \infty, \quad quand \quad n \to$ (2.1.1) est vrifie, alors

$$\hat{\alpha}^{P}_{n,m} \to \alpha, \quad presque\ s\hat{u}r, \quad quand \quad n \to \infty.$$

Normalité asymptotique de $\hat{\alpha}^{P}_{n,m}$

Théorème 2.1.3 —

Soit F une fonction de distribution dans le domaine d'attraction d'une distribution des valeurs extrêmes G_α, pour $\alpha \in \mathbb{R}$. Supposons que la fonction des quantiles $Q = F^{-1}$ a un dérivé positif Q', et qu'il existe une fonction positive a, telle que la fonction

$$t \mapsto t^{-\alpha-1} Q'(1 - 1/t),$$

est $\pi-$variation à l'infini, d'une fonction auxilière a. Alors pour une suite $m = m_n$ satisfait

$$m_n \to \infty, \quad m_n = o\left(n \big/ g^{-1}(n)\right),$$

où g^{-1}est l'inverse de

$$g(t) \quad = \quad t^{-2\alpha-1}(Q'(1-1/t) \big/ a(t))^2,$$

$$\sqrt{m}\left(\hat{\alpha}^{\hat{}P}_{n,m} - \alpha\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_p^2\right) \quad quand \quad n \to \infty,$$

où $\mathcal{N}\left(0, \sigma_p^2\right)$ désigne la distribution normale de moyenne 0 et variance σ_p^2 donnée par

$$\sigma_p^2 \quad = \quad \alpha^2 \frac{(2^{2\alpha-1})}{\left(2\left(2^\alpha - 1\right)\log 2\right)^2}$$

Estimateur de Hill

En 1975 un autre estimateur de l'index de la valeur extrême a été introduit par B. M. Hill dans (*Hill* 1975). Cet estimateur est basé sur les propriétés de variation régulière dans le cas où la fonction de distribution F dans le domaine d'attraction de G_α, où $\alpha > 0$. L'estimateur de *Hill* serait seulement applicable de la valeur extrême, au cas où l'index serait positif, c à d F a une queue lourde.

La construction de cet estimateur est basée sur l'estimation par la méthode du maximum de vraisemblance en supposant que $\alpha > 0$, et F vérifie la condition

$$\frac{1-F(x)}{1-F(u)} = \left(\frac{x}{u}\right)^{-1/\alpha}, \quad x \geq u, \quad u \in \mathbb{R}. \tag{2.2.1}$$

Alors, on utilise seulement les statistiques d'ordre qui depassent le seuil u.

On a, les statistiques d'ordre $X_{n,m}, ..., X_{n-k_n+1,n}$. La fonction de log-vraisem-blance est donné par

$$\begin{aligned} L(\alpha; X_{n-k_n+1,n,...}, X_{n,n}) &= -k_n \log(\alpha u) + \log(1-F(u)) - \frac{\alpha+1}{\alpha} \\ &\times \sum_{i=1}^{k_n} (\log X_{n-i+1,n} - \log u) \end{aligned} \tag{2.2.2}$$

En resolvant cette equation et remplaçant le seuil u par $X_{n-k,n}$, on obtient l'estimateur de Hill

$$\hat{\alpha}^H_{n,k_n} = \frac{1}{k_n} \sum_{i=1}^{k_n} (\log X_{n-i+1,n} - \log X_{n-k,n}). \tag{2.2.3}$$

Notons que cet estimateur peut être écrit en terme des espacements des logarithmes des observations.

$$\hat{\alpha}^H_{n,k_n} = \frac{1}{k_n} \sum_{i=1}^{k_n} i\left(\log X_{n-i-1,n} - \log X_{n-i,n}\right). \tag{2.2.4}$$

Consistance de l'estimateur de Hill

Mason (1982 a) a prouvé la consistance faible de l'estimateur de *Hill* pour toute suite k_n vérifiant

$$k_n \to \infty, \quad \text{et} \quad k_n/n \to 0 \quad \text{quand} \quad n \to \infty.$$

La consistance forte a été prouvée dans *Deheuvels, Haeusler et Mason* (1985) sous les conditions

$$k_n/\log\log n \to \infty, \quad \text{et} \quad k_n/n \to 0, \quad n \to \infty.$$

Normalité asymptotique de $\hat{\alpha}^H_{n,k_n}$

Dans l'article de *Hill* (1975), la normalité asymptotique n'a pas été établit pour cet estimateur. La normalité asymptotique peut être établie sous certaines conditions supplémentaires, (voir *Haeusler et Teugels* (1985))

Nous affirmerons le résultat de la normalité asymptotique qui est obtenu comme un cas spécial de la preuve dans *Dekkers, Einmahl et de Hann* (1989), de la normalité asymptotique de leur estimateur du moment (voir aussi la section 2.5 de la thèse de PhD de *Paul de Wolf*).

Théorème 2.2.2 —
Soit F une fonction de distribution qui appartient au domaine d'attraction de la distribution des valeurs extrêmes G_α avec $\alpha > 0$. Soit Q sa fonction des quatiles. supposons qu'il existe une fonction positive b, telle que la fonction

$$t \mapsto \pm t^{-\alpha} Q\left(1 - 1/t\right)$$

(avec l'un et l'autre de choix de signe) est π-variation à l'infini avec une fonction auxiliaire b. Alors, pour toute suite $k_n \to \infty$ satisfaisant $k_n = o\left(n \diagup g^{-1}(n)\right)$, où g^{-1} est l'inverse de

$$g(t) \;=\; t^{1-2\alpha}\left(Q\left(1-1/t\right) \diagup b(t)\right)^2,$$

$$\sqrt{k_n}\left(\hat{\alpha}_{k_n,n} - \alpha\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \alpha^2\right), \quad \text{quand} \quad n \to \infty,$$

avec $\mathcal{N}\left(0, \alpha^2\right)$ la distribution normale de moyenne 0 et de variance α^2.

Estimateur du moment

Construction de l'estimateur du moment

Une généralisation de l'estimateur de Hill était introduite dans Dakkers, Einmahl et de Hann (1989). Leur but principal est de trouver un estimateur dans certaine sens semblable à l'estimateur de Hill qui est consistant pour tout $\alpha \in \mathbb{R}$. Ils ont introduit les deux quantités suivantes

$$M^{(r)}_{n,k} \;=\; \frac{1}{k}\sum_{i=1}^{k}\left(\log X_{n-i+1,n} - \log X_{n-k,n}\right)^r, \quad r = 1, 2.$$

Notons que $M_{n,k}^{(1)}$ est l'estimateur de *Hill*. Alors ils ont proposé l'estimateur de l'index des valeurs extrêmes suivant :

$$\hat{\alpha}_{n,k}^{M} = M_{n,k}^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_{n,k}^{(1)}\right)^2}{M_{n,k}^{(2)}}\right)^{-1}. \tag{2.3.1.1}$$

Nous donnerons une dérivation heuristique de cet estimateur du moment, en utilisant le lemme suivant (voir. *Dekkers, Einmahl et de Hann*(1989), Lemme 2.5).

Lemme 2.3.1 —

Supposons que $F \in D(G_\alpha)$, pour $\alpha \in \mathbb{R}$, on note par Q la fonction des quantiles. Alors, pour quelque fonction positive a,

$$\lim_{s \downarrow 0} \frac{\log Q(1-sx) - \log Q(1-s)}{a(s)/Q(1-s)}$$

$$= \begin{cases} -\log x & \alpha \geq 0 \\ \dfrac{x^{-\alpha}-1}{\alpha} & \alpha < 0 \end{cases} \qquad \text{pour } x > 0.$$

De plus, dans le cas $\alpha > 0$ nous pouvons prendre

$$a(s)/Q(1-s) = \alpha,$$

et dans le cas $\alpha < 0$ nous pouvons prendre

$$a(s)/Q(1-s) = -\alpha\left(\log Q(1) - \log Q(1-s)\right).$$

Notons que $M_{n,k}^{(r)}$ peut être envisagé comme l'esperance conditionnel de $(\log(X/t))^r$ sachant que $X > t$. Dans ce cas le seuil t pres place la $(n-k)^{ème}$ statistique d'ordre de l'échantion. En effet

$$E\left((\log(X/t))^r / X > t\right) = \int_t^{F^{-1}(1)} (\log(x/t))^r \, d\left(\frac{F(x)-F(t)}{1-F(t)}\right)$$

$$= -\int_t^{F^{-1}(1)} (\log(x/t))^r \, d\left(\frac{1-F(x)}{1-F(t)}\right).$$

La substitution de $y = 1 - F(x)$ et $s = 1 - F(t)$ donne

$$\begin{aligned} & \int_0^s (\log Q(1-y) - \log Q(1-s))^r d(y/s) \\ &= \int_0^1 (\log Q(1-sx) - \log Q(1-s))^r dx. \end{aligned}$$

Notons le dernier intégral par $M_s^{(r)}$, $M_{n,k}^{(r)}$ pourait être considéré pour être un estimateur de $M_s^{(r)}$ avec $s = k/n$. Maintenant, on utilise le lemme (2.3.1) pour déduire que, dans le cas où $\alpha > 0$,

$$M_s^{(r)} \sim \alpha^r \int_0^1 (-\log x)^r \, dx = r\alpha^r, \quad r = 1, 2,$$

et dans le cas où $\alpha < 0$,

$$\begin{aligned} M_s^{(r)} &\sim (-\alpha)^r \left((\log Q(1) - \log Q(1-s))^r\right) \int_0^1 \left(\frac{x^{-\alpha} - 1}{\alpha}\right)^r dx \\ &= (\log Q(1) - \log Q(1-s))^r \frac{r(-\alpha)^r}{(1 - r\alpha)(1 - (r-1)\alpha)}, \ r = 1, 2. \end{aligned}$$

Alors, la substitution de ces expressions dans (2.3.1.1) i.e,

$$\alpha_s := M_s^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_s^{(1)}\right)^2}{M_s^{(2)}}\right)^{-1},$$

satisfaite pour $\alpha > 0$.

$$\begin{aligned} \alpha_s &\sim \alpha + 1 - \frac{1}{2}\left(1 - \frac{\alpha^2}{2\alpha^2}\right)^{-1} \\ &= \alpha, \end{aligned}$$

dans le cas où $\alpha < 0$.

$$\begin{aligned}
\alpha &\sim -\alpha\left(\log Q\left(1\right)-\log Q\left(1-s\right)\right)\frac{1}{1-\alpha} \\
&+ 1-\frac{1}{2}\left(1-\frac{1-2\alpha}{2(1-\alpha)}\right)^{-1} \\
&= \alpha+\frac{\alpha}{1-\alpha}\left(\log Q\left(1\right)-\log Q\left(1-s\right)\right) \\
&\sim \alpha,
\end{aligned}$$

pour un petit s.

Normalité asymptotique de l'estimateur du moment

Théorème 2.4.1—

Soit F une fonction de distribution dans le domaine d'attraction d'une distribution de valeurs extrêmes G_α, pour $\alpha \in \mathbb{R}$. On note par Q la fonction des quantiles. Supposons qu'il existe des fonctions positives b_1, b_2, b_3 et *b_4 tel que pour tout $x > 0$ (avec le choix de signe).*

1. *dans le cas où $\alpha > 0$, la fonction $t \mapsto \pm t^{-\alpha} Q\left(1-1/t\right)$,est π-variation au voisinage de l'infini d'une fonction auxilière b_1.*
2. *Dans le cas où $\alpha = 0$,*

$$\lim_{t\to\infty}\frac{\log Q\left(1-1/\left(tx\right)\right)-\log Q\left(1-1/t\right)+b_2\left(t\right)\log x}{b_3\left(t\right)}=\pm\frac{\left(\log x\right)^2}{2}.$$

3. *Dans le cas où $\alpha < 0$, la fonction $t \mapsto \pm t^{-\alpha}\left(Q\left(1\right)-Q\left(1-1/t\right)\right)$,est π-variation au voisinage de l'infini d'une fonction auxilière b_4.*

Alors, pour la suite $k = k_n \to \infty$ satisfaisant $k_n = o\left(n/g^{-1}\left(n\right)\right)$, où g^{-1} est l'inverse de

$$g\left(t\right)=\begin{cases} t^{1-2\alpha}\left(Q\left(1-1/t\right)/b_1\left(t\right)\right)^2 & \alpha>0, \\ t\left(b_2\left(t\right)/b_3\left(t\right)\right)^2 & \alpha=0, \\ t^{1-2\alpha}\left(\log Q\left(t\right)-\log Q\left(1-1/t\right)/b_4\left(t\right)\right) & \alpha<0, \end{cases}$$

et de plus, dans le cas où $\alpha = 0, k_n = o\left(n/g_t^{-1}\left(n\right)\right)$ où $g_t^{-1}\left(\cdot\right)$ est l'inverse de

$$g_t\left(t\right) = t\left(Q\left(1-1/t\right)/a\left(1/t\right)\right)^2,$$

avec la fonction a est définée dans le lemme(2.3.1)

$$\sqrt{k}\left(\hat{\alpha}_{n,k}^{M}-\alpha\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\sigma_{M}^{2}\right),\quad quand\quad n\rightarrow\infty,$$

avec $\mathcal{N}\left(0,\sigma_M^2\right)$ *désigne la distribution normale de moyenne 0 et variance* σ_M^2 *donné par :*

$$\sigma_M^2=\begin{cases}1+\alpha^2 & \alpha\geq 0,\\ \\ \left(1-\alpha^2\right)\left(1-2\alpha\right)\left(4-8\dfrac{1-2\alpha}{1-3\alpha}+\dfrac{\left(5-11\alpha\right)\left(1-2\alpha\right)}{\left(1-3\alpha\right)\left(1-4\alpha\right)}\right) & \alpha<0.\end{cases}$$

Estimateur du moindre carée

Soit $X, X_1, X_2, ...$ une suite de variables aléatoires [v.a] indépendantes d'une fonction de distribution commune $F\left(x\right)=P\left(X\leq x\right),\ x\in\mathbb{R}$ tel que $1-F$ a variation régulière au voisinage de l'infinie, d'index $-1/\alpha$, $\alpha>0$. Supposons que

$$1-F\left(x\right)\ =\ x^{-1/\alpha}\mathbb{L}^{*}\left(x\right),\quad\text{quand}\quad x\rightarrow\infty,$$

otion*où $\mathbb{L}^*$ est une fonction positive sur $[1,\infty)$, a variation lente au voisinage de l'infinie, alors

$$\frac{\mathbb{L}^{*}\left(tx\right)}{\mathbb{L}^{*}\left(x\right)}\rightarrow 1,\quad\text{quand}\quad x\rightarrow\infty,$$

pour tout $t>0$.

Cette class de fonctions sera noté R_α. L'inverse ou la fonction des quantiles Q de F est définie par :

$$Q\left(s\right)=\inf\left\{x:F\left(x\right)\geq s\right\},\quad 0<s\leq 1,\quad Q\left(0\right):=Q\left(0+\right).\qquad(2.4.1)$$

Il est bien connu (voir Csörgő et Mason (1985)) que $F\in R_\alpha$ si et seulement si

$$Q+\left(1-s\right)\ =\ \mathbb{U}\left(s\right)\ =\ s^{-\alpha}\mathbb{L}\left(s\right),\quad\text{quand}\quad s\downarrow 0,\qquad(2.4.2)$$

pour la version continue à droite $Q+$ de Q, où $\mathbb{L}\left(\cdot\right)$ est une fonction positive a variation lente au voisinage de 0, i.e,

$$\frac{\mathbb{L}\left(ts\right)}{\mathbb{L}\left(s\right)}\rightarrow 1,\quad s\downarrow 0\quad\text{pour tout}\quad t>0.$$

Pour $n\geq 1$, soit $X_{1,n}\leq X_{2,n}\leq...\leq X_{n,n}$ les statistiques d'ordre basé sur l'échantillon X_1, X_2

On note par $\log^{+}$logarithm naturel log, où

$$\log^{+} \quad = \quad \log[\max(x,1)], \quad x \in \mathbb{R}.$$

L'estimateur de Hill (Hill 1975) est définie par :

$$\hat{\alpha}^{H}(k_n) \quad := \quad k_n^{-1} \sum_{i=1}^{k_n} \log\left(\frac{X_{n-1+1,n}}{X_{n-k_n,n}}\right),$$

où k_n est une suite d'entiers satisfaisants

$$1 \leq k_n < n, \quad k_n \to \infty \quad \text{et} \quad k_n/n \to 0. \tag{$\mathbb{K}$}$$

Csörgő, Deheuvels et Mason (1985) ont généralisé cet estimateur pour introduire une classe d'estimateurs à noyau.

En basant sur les considérations de moindre carées, Viharos (1999) a proposé un nouveau estimateur de l'index de la queue lequel est universellement asymptotiquement normal. Cet estimateur est défini par :

$$\hat{\alpha}^{(V)}(k_n) \quad = \quad \left(\sum_{i=1}^{k_n} a_{i,n} \log\left(\frac{k_n}{i}\right)\right)^{-1} \sum_{i=1}^{k_n} a_{i,n} \log X_{n-i+1,n}, \tag{2.4.3}$$

où $a_{i,n} := \int\limits_{i-1/k_n}^{i/k_n} \mathbb{J}(s)\,ds,\ i = 1, ..., k_n,$ avec $\mathbb{J}(\cdot)$ est une fonction pondérée spécifie par les conditions (H1)-(H5), au dessous

(H1) $\int\limits_0^1 \mathbb{J}(s)\,ds = 0;$

(H2) $\mathbb{J}(\cdot)$ est non croissante sur $(0,1]$ tel que $\lim\limits_{x\downarrow 0} \mathbb{J}(x) > 0$ et $\mathbb{J}(1) < 0;$

(H3) $\mathbb{J}(\cdot)$ est differentiable d'une dérivé $\mathbb{J}'(\cdot)$ monotone sur $(0,1]\,;$

(H4) $\int\limits_0^1 s^{-\upsilon-1/2} |\mathbb{J}(s)|\,ds < \infty$ pour $0 < \upsilon < 1/2;$

(H5) $\sup\limits_{0<s<1} s\,|\mathbb{J}'(s)| < \infty.$

Normalité asymptotique de $\hat{\alpha}^{(V)}(k_n)$

La normalité asymptotique de l'estimateur $\hat{\alpha}^{(V)}(k_n)$ est donné par le théorème A au dessous. Rappelons (2.4.2).

On utilise les notations suivantes. Pour chaque fonction $\mathbb{J}$ satisfaisant (H1)-(H5) on pose

$$\lambda(\mathbb{J}) \quad := \quad -\int_0^1 \mathbb{J}(t) \log t dt, \tag{2.4.1.1}$$

$$\mu^{(V)}(k_n) \quad := \quad \{\lambda(\mathbb{J})\}^{-1} \int_0^1 \mathbb{J}(t) \log \mathbb{U}(k_n t / n) dt, \tag{2.4.1.2}$$

et

$$\sigma^2(\mathbb{J}) \quad := \quad \{\lambda(\mathbb{J})\}^{-2} \int_0^1 \int_0^1 \min(s,t)(st)^{-1} \mathbb{J}(s)\mathbb{J}(t) dsdt. \tag{2.4.1.3}$$

Théorème A —

Supposons que $F \in R_\alpha$, pour $\alpha > 0$ et soit J une fonction pondérée satisfaisant (H1)-(H5). Alors pour toute suite satisfaire ($\mathbb{K}$), *on a*

$$k_n^{1/2}\left\{\hat{\alpha}^{(V)}(k_n) - \mu^{(V)}(k_n)\right\} \overset{\mathcal{D}}{\rightarrow} \mathcal{N}\left(0, \alpha^2\sigma^2(\mathbb{J})\right), \quad quand \quad n \to \infty,$$

avec

$$\mu^{(V)}(k_n) \to \alpha, \quad quand \quad n \to \infty.$$

Ici$\overset{\mathcal{D}}{=}$ et $\overset{\mathcal{D}}{\rightarrow}$ désigne l'égalité et la convergence en distribution respectivement, et soit $\mathcal{N}(\mu, \sigma^2)$ la distribution normale de moyenne μ et de variance σ^2.

La convergence presque sûre de l'estimateur $\hat{\alpha}^{(V)}(k_n)$ a été établit par Necir[2002].

Loi fonctionnelle de logarithme itéré pour $\hat{\alpha}^{(V)}(k_n)$

Soit $\mathbb{B}[0,1]$ désigne l'espace des fonctions à valeurs réelles bornées définie sur $[0,1]$ muni de la topologie de convergence uniforme sur $[0,1]$ et

$$\mathbb{K}[0,1] \quad = \quad \left\{ f \in \mathbb{B}[0,1], f(0) = 0 \text{ et } \int_0^1 (f'(s))^2 ds \leq 1 \right\}, \tag{2.4.2.1}$$

l'ensemble de *Strassen* de fonctions absolument continues. Ici f' désigne le dérivé au sens de Lebesgue de f. On utilise la notation :

$$\ell_n \quad = \quad \log\log(\max(n,3)), \quad n = 1, 2, ... \tag{2.4.2.2}$$

Théorème 2.5.2 —

Supposons que $F \in R_\alpha$ pour tout $\alpha > 0$ et soit J une fonction pondérée satisfaisant (H1)-(H5). Si k_n vérifié (K) et $k_n/\ell_n \to \infty$, quand $n \to \infty$, presque sûrement la suite

$$\left\{(k_n)^{1/2} \ell_n^{-1/2} \left[\hat{\alpha}^{(V)}(k_n) - \mu^{(V)}(k_n)\right]\right\}_{n \geq 1},$$

est relativement compact dans $\mathbb{R}$. L'ensemmble de points limites est égale

$$\Theta(\mathbb{J};g) = \left\{ 2^{1/2} \{\lambda(\mathbb{J})\}^{-1} \alpha \int_0^1 s^{-1} \mathbb{J}(s) g(s) ds, \ g \in \mathbb{K}[0,1] \right\}. \quad (2.4.2.3)$$

De plus, presque sûrement

$$\begin{aligned} \limsup_{n\to\infty} \pm (k_n)^{1/2} \ell_n^{-1/2} \left[\hat{\alpha}^{(V)}(k_n) - \mu^{(V)}(k_n) \right] &= \sup_{g\in\mathbb{K}[0,1]} \Theta(\mathbb{J};g) \\ &= 2^{1/2} \alpha \sigma(\mathbb{J}), \end{aligned}$$

avec

$$\mu^{(V)}(k_n) \to \infty, \quad \text{quand} \quad n \to \infty.$$

Estimateur de type noyau

Soit $X_1, X_2, ...$ une suite de variables aléatoires indépendantes, d'une fonction de distribution commune $F(x) = P(X_1 \leq x)$, $x \in \mathbb{R}$. Pour tout entier $n \geq 1$, soit $X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$ les statistiques d'ordre basées sur l'échantillon $X_1, X_2, ..., X_n$.

La fonction des quantiles Q de F est définie par :

$$Q(s) = \mathbb{U}(1-s) = \inf\{x : F(x) \geq s\}, \quad 0 < s < 1. \quad (2.5.1)$$

La fonction de distribution des valeurs extêmes du paramètre α est donnée par :

$$G_\alpha = \begin{cases} \exp\left(-(1+\alpha x)^{-1/\alpha}\right), & \text{pour} \quad \alpha \neq 0, \quad 1+\alpha x > 0, \\ 0, & \text{pour} \quad \alpha \neq 0, \quad 1+\alpha x \leq 0, \\ \exp(-e^{-x}), & \text{pour} \quad \alpha = 0, \quad x \in \mathbb{R}. \end{cases}$$

Soit $I_\alpha = \{x \in \mathbb{R} : 1 + \alpha x \neq 0\}$, $\alpha \neq 0$, et $I_0 = \mathbb{R}$.

La condition pour que F appartient au domaine d'attraction faible de la distribution des valeurs extrêmes G_α, $\alpha \in \mathbb{R}$ notée $F \in D(G_\alpha)$ est l'existence des suites $a_n > 0$ et b_n telles que

$$\lim_{n\to\infty} \mathbb{P}\{a_n^{-1}(X_{n,n} - b_n) \leq x\} = G_\alpha(x), \quad \text{pour} \quad x \in I_\alpha.$$

L'estimateur de Hill (Hill (1975)) de l'index de la queue α est défini comme suit :

$$\hat{\alpha}^H(k_n) := k_n^{-1} \sum_{i=1}^{k_n} \log\left(\frac{X_{n-i+1,n}}{X_{n-k_n,n}}\right),$$

où les k_n sont des entiers positifes satisfaisant les conditions suivantes :

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

Sous ces conditions, on sait (Mason 1982) que $\hat{\alpha}^H(k_n)$ est consistant pour tout $\alpha > 0$, on supposons seulement que $F \in D(G_\alpha)$. Cet estimateur a été généralisé par Csörgő, Deheuvels et Mason $\Big(\bar{K}(s)$ *est celle dans* (2.5.2.6).

Rappellons (2.5.2.7) et (2.5.2.8).

Corollaire 3.7.3 —

supposons que $F \in D(G_\alpha)$pour tout $\alpha \in \mathbb{R}$,et supposons que (CP1)-(CP3) et (RV1)-(RV2) sont vérifiées. Fixons un $\gamma > 1/2$arbitraire, et soit Kla fonction satisfaisant les conditions (CK1)-(CK5). Alors, si $h = h_n$vérifiant $h \downarrow 0$ et $nh/\ell_n \to \infty$, quand $n \to \infty$, avec probabilité 1, avec les conditions supplémentaires

1. *Si $\alpha > 0$,*

$$(nh)^{1/2}\,\ell_n^{-1/2}\mathbb{U}_1(h) \to 0, \quad \textit{quand} \quad n \to \infty$$

2. *Si $\alpha < 0$, $(nh)^{1/2}\,\ell_n^{-1/2}\mathbb{U}_2(h) \to 0$,*
3. *et*

$$(nh)^{-\alpha+1/2}\,\ell_n^{-1/2} \to 0, \quad \textit{quand} \quad n \to \infty$$

alors avec une probabilité 1, nous avons

$$\begin{aligned} \limsup_{n\to\infty} \pm (nh)^{1/2}\,\ell_n^{-1/2}\left[\hat{\alpha}_{n,h}^{(GLW)} - \alpha\right] &= \sup_{f\in\mathbb{K}[0,1]} \bar{\Theta}(f) \\ &= 2^{1/2}\sigma\left(\bar{K}\right). \end{aligned}$$

où $\sigma\left(\bar{K}\right)$ est celle dans (2.5.2.8).

Corollaire 3.7.4 —

supposons que $F \in D(G_\alpha)$ avec $\alpha > 0$, et supposons que (CP3) et (RV1)sont vérifiées. Soit K un noyau satisfaisant les conditions (CK1), (CK2) et (CK4). Si $h = h_n$vérifiant $h \downarrow 0, nh/\ell_n \to \infty$, et $(nh)^{1/2}\,\ell_n^{-1/2}\mathbb{U}_1(h) \to 0$, quand $n \to \infty$, avec une probabilité 1

$$\begin{aligned} \limsup_{n\to\infty} \pm (nh)^{1/2}\,\ell_n^{-1/2}\left[\hat{\alpha}_{n,h}^{(CDM)} - \alpha\right] &= \sup_{f\in\mathbb{K}[0,1]} \Theta(f) \\ &= 2^{1/2}\sigma(K). \end{aligned}$$

PREUVE.

Voir [17]. $\square$

Sommes des valeurs extrêmes[SVE]

Normalité asymptotique des[SVE]

Soit $X_1, X_2, ...$ une suite de variables aléatoires indépendantes, d'une fonction de distribution commune F et pour chaque entier $n \geq 1$, soit $X_{1,n} \leq ... \leq X_{n,n}$ les statistiques d'ordre basées sur les premiers n de ces variables aléatoires. *Csörgő et Mason* (1985, 1986) ont montré que si

$$1 - F(x) = \mathbb{L}^*(x) x^{-\alpha} \quad \text{quand} \quad x \to \infty, \tag{3.1.1}$$

où $\mathbb{L}^*$ est une fonction à variation lente au voisinage de l'infini et $\alpha \geq 2$, alors pour toute suite des entiers satisfaisant

$$1 \leq k_n \leq n, \quad k_n \to \infty \quad \text{et} \quad k_n/n \to 0, \quad \text{quand} \quad n \to \infty, \tag{$\mathbb{K}$}$$

il existe deux suites des constantes $A_n > 0$ et C_n tel que

$$A_n \left(\sum_{i=1}^{k_n} X_{n-i+1,n} - C_n \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad n \to \infty. \tag{3.1.2}$$

Soit

$$Q(s) = \inf \{x : F(x) \geq s\}, \quad 0 < s \leq 1,$$

avec $Q(0) = Q(0+)$, désigne l'inverse de F.

On écrit

$$\sigma^2(s) = \int_{1-s}^{1} \int_{1-s}^{1} (\min(u,v) - uv)\, dQ(u)\, dQ(v).$$

Pour tous $0 < \beta < \infty$, posons

$$c(s,\beta) = s^{-\beta} \int_{1-s}^{1} (1-u)^{\beta}\, dQ(u), \quad 0 \leq s \leq 1.$$

Convention :

Si $\beta = 1$, on écrit $c(s) = c(s,1)$.

Soit $D^*(\Lambda)$ une sous classe de $D(\Lambda)$ de toutes les fonctions de distributions F, leur fonction de quantiles Q satisfaisant

$$Q(1-s) = a + \int_{s}^{1} u^{-1} r(u)\, du,$$

pour tout $s \geq 0$ suffisamment petit, a est une constante fixée et r est une fonction strictement positive à variation lente au voisinage de 0.

Pour toute suite des entiers positive vérifie $(\mathbb{K})$ et $F \in D(\Lambda)$, posons pour $n = 1, 2, ...$

$$\mu_n(k_n) = n \int_{1-k_n/n}^{1} Q(s)\, ds.$$

Théorème 3.1.1 —

Sur un espace de probabilité assez riche, il existe une suite de variables aléatoires indépendantes $X_1, X_2, ...$ d'une fonction de distribution commune F et une suite de ponts Brownien $B_1, B_2, ...$ tel que pour toute suite k_n satisfaisant $(\mathbb{K})$, $F \in D(\Lambda)$,

$$\begin{aligned} & k_n^{-1/2} C(k_n/n)^{-1} \left\{ \sum_{i=1}^{k_n} X_{n-i+1,n} - \mu_n(k_n) \right\} \qquad (3.1.3) \\ = \; & -(n/k_n)^{1/2} C(k_n/n)^{-1} \int_{1-k_n/n}^{1} B_n(s)\, dQ(s) + o_p(1) \\ := \; & Z_n + o_p(1), \end{aligned}$$

et si $F \in D^*(\Lambda)$,

$$\begin{aligned} & k_n^{1/2} C(k_n/n)^{-1} \{ X_{n-k_n,n} - Q(1-k_n/n) \} \qquad (4.1.4) \\ = \; & -(n/k_n)^{1/2} B_n(1-k_n/n) + o_p(1) \\ := \; & Y_n + o_p(1), \end{aligned}$$

et

$$\begin{aligned} & k_n^{-1/2} C(k_n/n)^{-1} \left\{ \sum_{i=1}^{k_n} X_{n-i+1,n} - k_n X_{n-k_n,n} - n \int_{1-k_n/n}^{1} r(1-s)\, ds \right\} \\ & := Z_n - Y_n + o_p(1). \end{aligned} \qquad (3.1.5)$$

Les variables aléatoires de (3.1.3), (3.1.4) et (3.1.5) converge en distribution vers $\mathcal{N}(0,2)$, $\mathcal{N}(0,1)$ et $\mathcal{N}(0,1)$ respectivement, quand $n \to \infty$.

Preuve.

Voir [2]. $\square$

Distribution asymptotique des[SVE]d'une distribution à variation régulière

Soit $X_1, X_2, ..., X_n$ une suite de variables aléatoires[v.a] indépendantes et identiquement distribuées[iid], d'une fonction de distribution commune $F(x) = \mathbb{P}\{X \leq x\}$, $-\infty < x < \infty$, et la fonction des quantiles

$$Q(u) = \inf\{x : F(x) \geq u\}, \quad 0 < u \leq 1, \quad Q(0) = Q(0+).$$

On suppose que F à variation régulière d'index $-1/\alpha$, i.e,

$$1 - F(x) = x^{-1/\alpha}\tilde{\mathbb{L}}(x), \quad x > 0, \tag{3.2.1}$$

où α est un nombre positif arbitraire et $\tilde{\mathbb{L}}$ est une fonction à variation lente au voisinage de l'infini. Ceci est équivalant à

$$Q(1 - s) = s^{-\alpha}\mathbb{L}(s), \quad 0 < s < 1, \tag{3.2.2}$$

où $\mathbb{L}$ est une fonction à variation lente au voisinage de 0.

Soit $X_{1,n} \leq ... \leq X_{n,n}$ les statistiques d'ordre de $X_1, X_2, ..., X_n$. On s'intéresse à la distribution asymptotique des sommes des valeurs extrêmes

$$T_n(k_n) = \sum_{j=1}^{k} X_{n+1-j,n}.$$

Si $k \geq 1$ entier fixé alors,

$$(n^{\alpha}\mathbb{L}(1/n))^{-1}\sum_{j=1}^{k} X_{n+1-j,n} \overset{\mathcal{D}}{\to} \sum_{j=1}^{k}(S_j)^{-\alpha},$$

où $\overset{\mathcal{D}}{\to}$ désigne la convergence en distribution, quand $n \to \infty$, et $S_j = E_1 + ... + E_j$, $j = 1, ..., k$, où $E_1, ..., E_k$ sont des variables aléatoires exponentielle indépendantes de moyenne 1.

Les sommes de la forme $T_n(k_n)$ jouent un rôle important. Les premiers exemples sont l'estimation de l'index des queues de distributions (Hall (1982a), Mason (1982) et Csörgő, Deheuvels et Mason (1985)) et l'estimation des points limites de la distribution (Hall(1982b) et S. Csörgő et Mason (1984)).

Le cas $\alpha > 1/2$, quand F dans le domaine d'attraction d'une loi stable non normale a été traité dans S. Csörgő, Horvàth et Mason (1986).

Pour chaque $\alpha > 1/2$ soit $\Delta_{1/\alpha}$ désigne la variable aléatoire stable complètement asymptotique d'une fonction caractéristique

$$\Phi_{1/\alpha,\beta,\gamma,\theta}(t) = \exp\left\{i\theta t - \gamma\,|t|^{1/\alpha}\left[1 - i\beta\,\mathrm{sgn}(t)\,\omega(t,1/\alpha)\right]\right\},$$

où

$$\omega(t,1/\alpha) = \begin{cases} \tan\pi/2\alpha & \text{si} \quad \alpha \neq 1, \\ -2/\pi\log|t| & \text{si} \quad \alpha = 1 \end{cases}$$

donné par les paramètres

$$\beta = 1,$$

$$\gamma = \gamma(1/\alpha) = \begin{cases} \Gamma(1-1/\alpha)\cos\dfrac{\pi}{2\alpha}, & \text{si} \quad \alpha > 1, \\ \pi/2, & \text{si} \quad \alpha = 1, \\ 1/\alpha\,(1/\alpha - 1)^{-1}\,\Gamma(2-1/\alpha)\left|\cos\dfrac{\pi}{2\alpha}\right|, & \text{si} \quad 1/2 < \alpha < 1, \end{cases}$$

et

$$\theta = \theta(1/\alpha) = \begin{cases} 0, & \text{si} \quad \alpha \neq 1, \\ \displaystyle\int_0^\infty \left(\frac{\sin x}{x^2} - \frac{1}{x(1+x)}\right)dx, & \text{si} \quad \alpha = 1, \end{cases}$$

La description complète de la distribution asymptotique des sommes des valeurs extrêmes sous la variation régulière, est donnée par les résultats suivants :

Théorème 3.2.1 —
Soit k_n une suite des entiers positifes tel que

$$n + 1 - k_n \leq n, \quad k_n \to \infty, \quad k_n/n \to 0.$$

(i) *Si $\alpha > 1/2$, alors*

$$\left(n^\alpha \mathbb{L}(1/n)\right)^{-1}\left\{\sum_{j=1}^{k} X_{n+1-j,n} - C_n(k_n)\right\} \xrightarrow{\mathcal{D}} \Delta_{1/\alpha},$$

où

$$C_n(k_n) = \begin{cases} 0, & \alpha > 1, \\ n \int\limits_{1-k_n/n}^{1-1/n} Q(s)\,ds, & \alpha = 1, \\ n \int\limits_{1-k_n/n}^{1} Q(s)\,ds, & 1/2 < \alpha < 1. \end{cases}$$

Si $0 < \alpha \leq 1/2$, *alors*

$$\frac{1}{A_n(\alpha, k_n)} \left\{ \sum_{j=1}^{k} X_{n+1-j,n} - n \int\limits_{1-k_n/n}^{1} Q(s)\,ds \right\} \overset{\mathcal{D}}{\to} \mathcal{N}(0,1)$$

où $A_n(\alpha, k_n) = \begin{cases} n^{1/2}(k_n/n)^{1/2-\alpha}\,\mathbb{L}(k_n/n)\,W_\alpha, & 0 < \alpha < 1/2, \\ n^{1/2}\left(\int\limits_{1/n}^{k_n/n} s^{-1}\mathbb{L}^2(s)\,ds \right)^{1/2}, & \alpha = 1/2, \end{cases}$

et

$$W_\alpha = \left(\frac{2\alpha^2}{(1-2\alpha)(1-\alpha)} \right)^{1/2}.$$

Le cas (i) est un cas spécial dans le théorème 3 de S. Csörgő, Horvàth et Mason (1986).

La preuve du cas (ii) est basée sur l'approximation du processus empirique uniforme par une suite de ponts Brownien.

Preuve.
Voir [5]. □
Comportement presque sûre des[SVE]

Comportement presque sûre des[SVE]d'une distribution dans le domaine d'attraction de Gumbel

Nous étudions le comportement presque sûre des sommes de statistiques d'ordre dans le cas de lois attirées par la distribution de Gumbel.

Soit $X_1, X_2, ...$ une suite de variables aléatoires indépendantes et identiquement distribuées, d'une fonction de distribution $F(x) = \mathbb{P}(X_1 \leq x)$, et pour tout entier $n \geq 1$, on note par $X_{1,n} \leq ... \leq X_{n,n}$ les statistiques d'ordre basées sur $X_1, X_2, ..., X_n$.

Considérons les sommes de la forme

$$S_n(k_n) = \sum_{i=1}^{k_n} X_{n-i+1,n}.$$

Soit $F \in D(\Lambda)$. On définit la fonction des quantiles

$$Q(s) = \inf\{x : F(x) \geq s\} = U(1-s), \quad 0 < s < 1,$$

et soit $\{k_n, n \geq 1\}$ une suite des entiers non-décroissante satisfaisant

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

Le comportement asymptotique presque sûre de $S_n(k_n)$ est déterminé par

la vitesse de croissance de k_n dans les différents cas ; $k_n = o(\log_2 n)$, $k_n \sim \lambda \log_2 n$ pour $0 < \lambda < \infty$, ou $\log_2 n = o(k_n)$ quand $n \to \infty$.

Théorème 3.3.1.1 —
Si $F \in D(\Lambda)$ et $k_n = o(\log_2 n)$ quand $n \to \infty$, alors

$$\liminf_{n\to\infty} (k_n A_n(k_n))^{-1} \{S_n(k_n) - k_n U(k_n/n)\} = -1 \quad p.s, \qquad (3.3.1.1)$$

où

$$A_n(k_n) = U(k_n/n) - U(\log_2 n/n).$$

Preuve.
Voir [8]. □

Soit $V_1, V_2, ...$ une suite de variables aléatoires indépendantes et identiquement distribuées uniforme sur $(0,1)$, et soit G_n la fonction de distribution empirique basée sur $V_1, V_2, ..., V_n$, pour tout $n \geq 1$.

Théorème 3.3.1.2 —
Supposons que $F \in D(\Lambda)$, $\log_2 n = o(k_n)$ et $n^{-1}k_n \sim \beta_n \downarrow 0$ quand $n \uparrow \infty$, pour la suite des constantes β_n. Alors les deux résultats suivants sont équivalents

(i) *Pour* $0 \leq \alpha < \infty$ *dépendant de* F*et* k_n.

$$\limsup_{n\to\infty} c_n(k_n)^{-1}\{S_n(k_n) - \mu_n(k_n)\} = \alpha \quad p.s, \tag{3.3.1.2}$$

où

$$c_n(k_n) = (2k_n \log_2 n)^{1/2} c(k_n/n),$$

et

$$\mu_n(k_n) = n\int_{1-k_n/n}^{1} Q(s)\,ds = n\int_0^{k_n/n} U(s)\,ds.$$

(ii) *Pour* $0 < d < \infty$ *et* $0 < \alpha(d) < \infty$ *dépendant de* F*et* k_n.

$$\limsup_{n\to\infty} -c_n(k_n)^{-1}\int_0^{(d\log_2 n)/n} nG_n(s)\,dU(s) = \alpha(d) \quad p.s. \tag{3.3.1.3}$$

De plus, si (3.3.1.3) est vérifie pour tout $0 < d < d_0$*, avec* $\alpha(d) \to 0$ *quand* $d \downarrow 0$*, alors*

$$\limsup_{n\to\infty} \pm c_n(k_n)^{-1}\{S_n(k_n) - \mu_n(k_n)\} = \sqrt{2} \quad p.s.$$

Preuve.
Voir [8]. □
Loi de logarithme itéré pour les[SVE]d'une distribution à variation régulière

Soit F une fonction de distribution avec $F(0-) = 0$ et à variation régulière, i.e,

$$1 - F(x) = x^{-\alpha}\tilde{\mathbb{L}}(x), \quad x > 0, \tag{3.3.2.1}$$

pour un $0 < \alpha < \infty$ et une fonction $\tilde{\mathbb{L}}$ à variation lente au voisinage de l'infini.
Soit

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

désigne la fonction des quantiles correspondante. Alors (3.3.2.1) est équivalent à

$$Q(1-s) = s^{-1/\alpha}\mathbb{L}(s), \quad 0 < s < 1, \tag{3.3.2.2}$$

où $\mathbb{L}$ est une fonction à variation lente au voisinage de 0;

Considerons un $0 < \alpha < \infty$ fixé et une fonction $\mathbb{L}$ à variation lente. Soit $X_1, X_2, ...$ une suite de variables aléatoires [iid] d'une fonction de distribution commune F, et pour chaque $n \geq 1$, soit $X_{1,n}, X_{2,n}, ..., X_{n,n}$ les statistiques d'ordre basées sur $X_1, X_2, ..., X_n$.

S. Csörgő et Mason (1986) ont montré que si $2 \leq \alpha < \infty$, alors pour toute suite $(k_n)_{n\geq 1}$ des entiers positive, telle que $k_n \to \infty$ et $k_n/n \to 0$, on a

$$A_n(\alpha, k_n)^{-1} \left\{ \sum_{j=1}^{k_n} X_{n+1-j,n} - n\mu(k_n/n) \right\} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \qquad (3.3.2.3)$$

où

$$\begin{aligned} \mu(k_n/n) &= \int_{1-k_n/n}^{1} Q(s)\, ds, \\ A_n(\alpha, k_n) &= N_\alpha n^{1/2} (k_n/n)^{1/2-1/\alpha} \mathbb{L}(k_n/n), \end{aligned}$$

et

$$N_\alpha = \begin{cases} \left(\dfrac{2}{(\alpha-2)(\alpha-1)} \right)^{1/2}, & \text{pour} \quad \alpha > 2, \\ \left(n \displaystyle\int_{1/n}^{k_n/n} s^{-1} \mathbb{L}^2(s)\, ds \right)^{1/2}, & \text{pour} \quad \alpha = 2. \end{cases}$$

Pour $0 < \alpha < 2$ la somme de k_n superieure valeurs extrêmes centrée et normalisée converge en distribution vers une loi stable d'index α.

Soit

$$\begin{aligned} a_n(k_n) &= N_\alpha (2n \log\log n)^{1/2} (k_n/n)^{1/2-1/\alpha} \mathbb{L}(k_n/n) \\ &= N_\alpha (2k_n \log\log n)^{1/2} Q(1-k_n/n), \end{aligned}$$

et soit

$$k_n \sim \alpha_n \uparrow \infty, \qquad (3.3.2.4)$$

et

$$k_n/n \sim \beta_n \downarrow 0, \qquad (3.3.2.5)$$

pour des suites des nombres positives $(\alpha_n)_{n\geq 1}$ et $(\beta_n)_{n\geq 1}$. On ne peut pas supposer que la suite $(k_n)_{n\geq 1}$ satisfait $k_n \uparrow \infty$ et $k_n/n \downarrow 0$, mais une telle suite n'existe pas.

Théorème 3.3.2.1 —

sous (3.3.2.1) pour tout $\alpha > 2$ *et une suite des entiers positive* $(k_n)_{n\geq 1}$ *satisfaire (3.3.2.4) et (3.3.2.5), les trois énoncées sont équivalentes pour un entier* $k \geq 1$ *fixé.*

$$\sum_{n=1}^{\infty} n^{k-1} \left(1 - F\left(a_n\left(k_n\right)\right)\right)^k \quad < \quad \infty, \tag{3.3.2.6}$$

$$a_n\left(k_n\right)^{-1} X_{n+1-k,n} \quad \longrightarrow 0 \quad p.s, \tag{3.3.2.7}$$

$$\limsup_{n\to\infty} a_n\left(k_n\right)^{-1} \left\{ \sum_{j=k}^{k_n} X_{n+1-j,n} - n\mu\left(k_n/n\right) \right\} \quad = 1 \quad p.s, \tag{3.3.2.8}$$

de plus, si (3.3.2.6), (3.3.2.7), ou (3.3.2.8) vérifies, on a

$$\liminf_{n\to\infty} a_n\left(k_n\right)^{-1} \left\{ \sum_{j=k}^{k_n} X_{n+1-j,n} - n\mu\left(k_n/n\right) \right\} \quad = -1 \quad p.s. \tag{4.3.2.9}$$

Preuve.
Voir [11]. □
Sommes tronquées

Loi limite des sommes tronquées

Nous étudions dans ce paragraphe la loi limite des sommes des statistiques d'ordre des variables aléatoires [iid], d'une fonction de distribution commune dans le domaine d'attraction d'une loi stable d'index $0 < \alpha \leq 2$ (Fréchet). On s'intérèsse à l'étude de la loi limite des sommes de la forme

$$S_n\left(k,m\right) = \sum_{i=k+1}^{n-m} X_{i,n} \quad \text{et} \quad S_{n,k} = S_n\left(k,k\right) = \sum_{i=k+1}^{n-m} X_{i,n}. \tag{3.4.1.1}$$

Ce problème a été traité par M. Csörgő, S. Csörgő, Horvàth et Mason (1986),

dans le cas où k et m sont fixés, et ils ont trouvé des constantes convenables a_n et $b_n$pour que $\left(S_n\left(k,m\right) - b_n\right)/a_n$ converge en loi vers une distribution non dégénérée. Dans ce travail nous considérons le problème analogue pour $S_n\left(k_n,k_n\right)$ tel que $k_n \to \infty$, $n \to \infty$.

Avant de donner ces résultats, nous introduirons quelques notations.

Soit $X_1, X_2, ..., X_n$ une suite de variables aléatoires [iid], d'une fonction de distribution commune

$$F(x) = \mathbb{P}\{X \leq x\}, \quad -\infty < x < \infty,$$

et une fonction des quantiles correspondante

$$Q(u) = \inf\{x : F(x) \geq u\}.$$

Soit

$$G(x) = \mathbb{P}\{|X| \leq x\}, \quad -\infty < x < \infty,$$

et

$$K(u) = \inf\{x : G(x) \geq u\}, \quad 0 < u < 1$$

l'inverse de G.

Posons

$$Q_1(u) = (-Q(1-u)) \vee 0 \text{ et } Q_2(u) = Q(u) \vee 0, \ 0 < u < 1$$

tel que $a \vee b = \max(a, b)$.

On écrit

$$\sigma^2 = \int_s^{1-s}\int_s^{1-s} (u \wedge v - uv)\, dQ(u)\, dQ(v),$$

$$\sigma_1^2 = \int_0^{1-s}\int_0^{1-s} (u \wedge v - uv)\, dQ_1(u)\, dQ_1(v),$$

$$\sigma_2^2 = \int_0^{1-s}\int_0^{1-s} (u \wedge v - uv)\, dQ_2(u)\, dQ_2(v).$$

Nous donnons ici les résultats dus à M. Csörgő, S. Csörgő, Horvàth et Mason (1986).

1) Les conditions classiques necessaires et suffisantes pour que $F \in D(\alpha)$, $0 < \alpha < 1$ sont équivalentes aux trois conditions suivantes

(S.1) $K(1-u) = u^{-1/\alpha}\mathbb{L}(u)$, $\mathbb{L}$ est une fonction à variation lente au voisinage de 0,

(S.2) $\lim\limits_{u \downarrow 0} Q_1(1-u)/K(1-u) = q^{1/\alpha}$,

(S.3) $\lim_{u \downarrow 0} Q_2(1-u)/K(1-u) = p^{1/\alpha}$,

où $0 < \alpha < 2$, $0 \leq p \leq 1$ et $q = 1 - p$.

2) Les conditions classiques necessaires et suffisantes pour que $F \in D(2)$, équivalentes aux deux conditions suivantes

(N.1) σ^2 est à variation lente au voisinage de 0,

(N.2) $\lim_{u \downarrow 0} u \left\{Q^2(u) + Q^2(1-u)\right\} / \sigma^2(u) = 0$.

3) M. Csörgő, S. Csörgő, Horvàth et Mason (1986) ont construi un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ muni d'une suite $U_1, U_2, \ldots$ de variables aléatoires, d'une distribution commune uniforme sur $(0,1)$; et une suite des ponts Brownien $B_n(u)$, $0 \leq u \leq 1$, $n = 1, 2, \ldots$. Soit

$$\alpha_n(u) = n^{1/2}\left\{\mathbb{G}_n(u) - u\right\}, \quad 0 \leq u \leq 1$$

où

$$\mathbb{G}_n(u) = n^{-1}\left\{k : 1 \leq k \leq n, U_k \leq u\right\},$$

nous avons

$$\sup_{1/n \leq u \leq 1-1/n} n^{-\nu} \frac{|\alpha_n(u) - B_n(u)|}{(u(1-u))^{1/2-\nu}} = o(1), \quad n \to \infty$$

où ν nombre fixé, $\nu \in [0, 1/4]$.

Notre premier résultat est le suivant :

Théorème 3.4.1.1 —

Supposons que $F \in D(\alpha)$, où $0 < \alpha \leq 2$. Alors pour toute suite des entiers positifes k_n tel que

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

Alors

$$A_n(k_n)\left\{\sum_{i=k_n+1}^{n-k_n} X_{i,n} - C_n(k_n)\right\} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{3.4.1.2}$$

où

$$A_n(k_n) = \left[n^{1/2}\sigma(k_n/n)\right]^{-1},$$

et

$$C_n(k_n) = n \int_{k_n/n}^{1-k_n/n} Q(u)\, du.$$

Dans le cas où $F \in D(\alpha)$, et $0 < \alpha < 2$ le théorème suivant donne une information pricise sur la loi limite du partie centrale de l'échantillon.

Théorème 3.4.1.2 —

supposons que $F \in D(\alpha)$, où $0 < \alpha < 2$. Soit k_n et ℓ_n deux suites d'entiers tels que

$1 \leq k_n \leq \ell_n \leq n - \ell_n \leq n - k_n \leq n$ *et* $k_n/n \to 0$, $\ell_n/n \to 0$,

$k_n \to \infty$ *et* $\ell_n/k_n \to \infty$, $n \to \infty$.

Alors,

$$T_n^{(1)} = A_n(k_n) \left\{ \sum_{i=k_n+1}^{\ell_n} X_{i,n} - n \int_{k_n/n}^{\ell_n/n} Q(u)\, du \right\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_1),$$

$$T_n^{(2)} = A_n(k_n) \left\{ \sum_{i=n-\ell_n+1}^{n-\ell_n} X_{i,n} - n \int_{\ell_n/n}^{1-\ell_n/n} Q(u)\, du \right\} \xrightarrow{\mathcal{P}} 0,$$

$$T_n^{(3)} = A_n(k_n) \left\{ \sum_{i=n-\ell_n+1}^{n-k_n} X_{i,n} - n \int_{1-\ell_n/n}^{1-k_n/n} Q(u)\, du \right\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_2),$$

où

$$\sigma_1 = \frac{q^{2/\alpha}}{p^{2/\alpha} + q^{2/\alpha}}$$
$$\sigma_2 = \frac{p^{2/\alpha}}{p^{2/\alpha} + q^{2/\alpha}},$$

et $T_n^{(1)}$, $T_n^{(3)}$ sont des variables aléatoires asymptotiquement indépendantes.

Notre dernier théorème prouve que la stabilité asymptotique est déterminée seulement par les statistiques d'ordre superieurs et inferieurs. Premièrement, on donne quelques notations et résultats de M. Csörgő, S. Csörgő, Horvàth et Mason (1986).

Soit $A_n^* = \left[\left(n^{1/\alpha}\mathbb{L}\left(1/n\right)\right)\right]^{-1}$, où $\mathbb{L}$ est de (S.1), et pour $i = 1, 2$, on a

$$C_n^{(i)} = \begin{cases} 0, & \text{si} \quad \alpha < 1, \\ (-1)^i \, n \displaystyle\int_{1/n}^{1} Q_i\left(1-u\right) du, & \text{si} \quad \alpha = 1, \\ (-1)^i \, n \displaystyle\int_{0}^{1} Q_i\left(1-u\right) du, & \text{si} \quad \alpha > 1. \end{cases}$$

Posons $X^- = X \wedge 0$ et $X^+ = X \vee 0$.

D'aprés le Corollaire (3.1) de M. Csörgő, S. Csörgő, Horvàth et Mason (1986), on a

Si $F \in D\left(\alpha\right)$, $0 < \alpha < 2$, alors

$$S_n^{(1)} = A_n^* \left\{ \sum_{i=1}^{n} X_i^- - C_n^{(1)} \right\} \xrightarrow{\mathcal{D}} -q^{1/\alpha}\Delta_\alpha, \tag{3.4.1.3}$$

$$S_n^{(2)} = A_n^* \left\{ \sum_{i=1}^{n} X_i^+ - C_n^{(2)} \right\} \xrightarrow{\mathcal{D}} p^{1/\alpha}\Delta_\alpha, \tag{3.4.1.4}$$

où $S_n^{(1)}$ et $S_n^{(2)}$ sont asymptotiquement indépendantes, et Δ_α est une variable aléatoire d'une fonction caracteristique stable.

Nous avons

$$A_n^* \left\{ \sum_{i=1}^{n} X_i - C_n \right\} \xrightarrow{\mathcal{D}} -q^{1/\alpha}\Delta_\alpha + p^{1/\alpha}\Delta_\alpha,$$

où $C_n = C_n^{(1)} + C_n^{(2)}$.

Théorème 3.4.1.3 —

Supposons que $F \in D\left(\alpha\right)$, où $0 < \alpha < 2$. Soit k_n une suite des entiers telle que

$$1 \leq k_n \leq n - k_n \leq n \quad et \quad k_n \to \infty \quad et \quad k_n/n \to 0.$$

Alors,

$$S_n^{(1)}\left(k_n\right) = A_n^* \left\{ \sum_{i=1}^{k_n} X_{i,n} - C_n^{(1)}\left(k_n\right) \right\} \xrightarrow{\mathcal{D}} -q^{1/\alpha}\Delta_\alpha, \tag{3.4.1.5}$$

$$S_n^{(2)}(k_n) = A_n^* \left\{ \sum_{i=n-k_n+1}^{n} X_{i,n} - C_n^{(2)}(k_n) \right\} \xrightarrow{\mathcal{D}} p^{1/\alpha}\Delta_\alpha, \qquad (3.4.1.6)$$

où $S_n^{(1)}$ et $S_n^{(2)}$ sont asymptotiquement indépendantes, pour $i = 1, 2$

$$C_n^{(i)}(k_n) = \begin{cases} 0, & si \quad \alpha < 1, \\ (-1)^i n \displaystyle\int_{1/n}^{k_n/n} Q_i(1-u)\, du, & si \quad \alpha = 1, \\ (-1)^i n \displaystyle\int_{0}^{k_n/n} Q_i(1-u)\, du, & si \quad \alpha > 1. \end{cases}$$

Conséquence :

$$A_n^* \left\{ \sum_{i=1}^{k_n} X_{i,n} + \sum_{i=n-k_n+1}^{n} X_{i,n} - C_n^*(k_n) \right\} \xrightarrow{\mathcal{D}} -q^{1/\alpha}\Delta_\alpha^{(1)} + p^{1/\alpha}\Delta_\alpha^{(2)}, \qquad (3.4.1.7)$$

tel que $C_n^*(k_n) = C_n^{(1)}(k_n) + C_n^{(2)}(k_n)$, et $\Delta_\alpha^{(1)}, \Delta_\alpha^{(2)}$ sont des copies indépendantes de Δ_α.

Pour prouver tous ces resultats, nous avons besion de quelques lemmes basés sur la théorie des fonctions à variation lente.

Lemme 3.4.1.1 —

Pour toute $F \in D(\alpha)$, où $0 < \alpha < 2$,

$$\lim_{s \downarrow 0} \sigma^2(s) \diagup \left\{ s^{1-2/\alpha} \mathbb{L}^2(s) \right\} = 2\left(p^{2/\alpha} + q^{2/\alpha}\right) \diagup (2-\alpha), \qquad (3.4.1.8)$$

$$\lim_{s \downarrow 0} \sigma_1^2(s) \diagup \left\{ s^{1-2/\alpha} \mathbb{L}^2(s) \right\} = 2q^{2/\alpha} \diagup (2-\alpha), \qquad (3.4.1.9)$$

$$\lim_{s \downarrow 0} \sigma_2^2(s) \diagup \left\{ s^{1-2/\alpha} \mathbb{L}^2(s) \right\} = 2p^{2/\alpha} \diagup (2-\alpha). \qquad (3.4.1.10)$$

Lemme 3.4.1.2 —

Soit L une fonction à valeurs réelles définie sur $(0,1)$*, et à variation lente au voisinage de 0.*

Soit k_n *et* l_n *deux suites d'entiers positives telles que*

$k_n/n \to 0, \quad l_n/n \to 0 \quad \text{et} \quad l_n/k_n \to \infty, \quad n \to \infty.$

Alors, pour tous $0 < \vartheta < \infty$*, nous avons*

$$\lim_{n\to\infty} \left\{(l_n/n)^{-\vartheta}\mathbb{L}\,(l_n/n)\right\} \diagup \left\{(k_n/n)^{-\vartheta}\,\mathbb{L}\,(k_n/n)\right\} \;=\; 0. \qquad (3.4.1.11)$$

Le lemme suivant simplifie les preuves des théorèmes (3.4.1.2) et (3.4.1.3).

Lemme 3.4.1.3 —

Supposons que $F \in D\,(\alpha)$*, où* $0 < \alpha < 2$*. Alors, pour toute suite des entiers positives,*

$1 \le m_n < n$ *tel que* $m_n \to \infty$ *et* $m_n/n \to 0, \quad n \to \infty,$

$$M_n^{(1)}\,(m_n) = \varphi\,(m_n, n) \left\{ \sum_{i=m_n+1}^{n} X_{i,n}^{-} - H_n^{(1)}\,(m_n) \right\} \xrightarrow{\mathcal{D}} \mathcal{N}\,(0, \delta_1) \quad (3.4.1.12)$$

et

$$M_n^{(2)}\,(m_n) = \varphi\,(m_n, n) \left\{ \sum_{i=1}^{n-m_n} X_{i,n}^{+} - H_n^{(2)}\,(m_n) \right\} \xrightarrow{\mathcal{D}} \mathcal{N}\,(0, \delta_2) \quad (3.4.1.13)$$

où

$$H_n^{(i)}\,(m_n) \;=\; (-1)^i\, n \int_{m_n/n}^{1} Q_i\,(1-u)\,du, \quad i = 1, 2,$$

$$\varphi\,(m_n, n) = \left\{ n^{1/2}\,(m_n/n)^{1/2-1/\alpha}\,\mathbb{L}\,(m_n/n)\right\}^{-1}$$

et

$$\delta_1 = 2q^{2/\alpha} \diagup 2 - \alpha \text{ et } \delta_2 = 2p^{2/\alpha} \diagup 2 - \alpha$$

Preuve.

Voir [4]. □

Loi de logarithme itéré pour les sommes tronquées

Nous avons vu que les sommes de la forme $\sum_{i=k_n+1}^{n-k_n} X_{i,n}$ où $(k_n)_{n\ge1}$ est une suite d'entier non-négative, tel que

$$k_n \to \infty \quad \text{et} \quad k_n/n \to 0 \quad \text{quand} \quad n \to \infty$$

peut être normalisées et centrées de sorte que la loi de logarithme itéré est vérifiée.

Soit $X_1, X_2, \ldots$ une suite de variables aléatoires indépendantes d'une fonction de distribution commune F. Pour tout n, on définit les statistiques d'ordre $X_{1,n} \leq \ldots \leq X_{n,n}$ basées sur $X_1, X_2, \ldots X_n$, et on désigne par $F \in D(\alpha)$ que F est dans le domaine d'attraction d'une loi stable d'index α.

Soit

$$Q(s) = \inf\{x : F(x) \geq s\}, \quad 0 < s \leq 1,$$

avec $Q(0) = Q(0+)$, désigne l'inverse ou la fonction des quantiles de F.

On écrit

$$\sigma^2(s) = \int_s^{1-s}\int_s^{1-s} (u \wedge v - uv)\, dQ(u)\, dQ(v), \quad 0 < s < 1/2,$$

où $x \wedge y = \min(x, y)$, et soit $(k_n)_{n \geq 1}$ une suite d'entier non-négative telle que $k_{n+1} \leq n - k_n$, pour $n \geq 1$, on a

$$T_n(k_n) = \sum_{i=k_n+1}^{n-k_n} X_{i,n}. \tag{3.4.2.1}$$

S. Csörgő, Horvàth et Mason (1986) ont montré la normalité asymptotique. Si

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

et $F \in D(\alpha)$, pour $0 < \alpha \leq 2$, alors

$$\left(n^{1/2}\sigma(k_n/n)\right)^{-1}\{T_n(k_n) - \mu_n(k_n)\} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{quand} \quad n \to \infty,$$

où

$$\mu_n(k_n) = n \int_{k_n/n}^{1-k_n/n} Q(s)\, ds.$$

et $\xrightarrow{\mathcal{D}}$ désigne la convergence en distribution et $\mathcal{N}(0,1)$ la loi normale standard.

On donne des conditions de régularité sur la suite $(k_n)_{n \geq 1}$

$$k_n \sim \alpha_n \uparrow \infty \qquad (n \to \infty), \tag{3.4.2.2}$$

$$k_n/n \sim \beta_n \downarrow 0 \qquad (n \to \infty). \tag{3.4.2.3}$$

Pour toute suite $(k_n)_{n\geq 1}$, on utilise la notation

$$\log_2 n \;=\; \log\log n \quad \text{pour} \quad n \geq 1.$$

Posons

$$a_n \;=\; (2n\log_2 n)^{1/2}\,\sigma\,(k_n/n)\,, \quad n \geq 3.$$

Théorème 3.4.2.1 —

Supposons que $(k_n)_{n\geq 1}$ *satisfait (3.4.2.2), et (3.4.2.3). Si* $F \in D(\alpha)$ *pour* $0 < \alpha < 2$, *et* $k_n/\log_2 n \to \infty$, *quand* $n \to \infty$ *ou si* $F \in D(2)$, *et* $\liminf_{n\to\infty} k_n/\log_2 n = \gamma$, *pour quelque* $0 < \gamma \leq \infty$, *alors*

$$\limsup_{n\to\infty} a_n^{-1}\left\{T_n(k_n) - \mu_n(k_n)\right\} \;=\; 1 \quad p.s, \tag{3.4.2.4}$$

et

$$\liminf_{n\to\infty} a_n^{-1}\left\{T_n(k_n) - \mu_n(k_n)\right\} \;=\; -1 \quad p.s. \tag{3.4.2.5}$$

Remarque :

Dans le cas où $F \in D(\alpha)$ pour $0 < \alpha < 2$, et

$$\liminf_{n\to\infty} k_n/\log_2 n \;=\; \gamma,$$

pour quelque $0 < \gamma < \infty$, et $0 \leq c < \infty$, on a

$$\limsup_{n\to\infty} a_n^{-1}\left\{T_n(k_n) - \mu_n(k_n)\right\} \;= c \quad p.s. \tag{3.4.2.6}$$

Preuve.

Voir [12]. $\square$

Estimation de la moyenne

Variation régulière du second ordre

Domaine d'attraction d'une loi stable

Soit $X, X_1, X_2, ...$ une suite de variables alèatoires indépendantes et identiquement distribuées, d'une fonction de distribution F. Il est connu que F est dans le domaine d'attraction d'une loi stable d'index $\alpha \in (0,2)$ si et seulement si la fonction

$$H(x) = 1 - F(x) + F(-x)\,,$$

à variation régulière d'index $-\alpha$, $\alpha > 0$, i.e,

$$\lim_{t\to\infty} H(tx) \diagup H(t) = x^{-\alpha}, \quad x > 0, \tag{4.1.1.1}$$

et

$$\lim_{t\to\infty} \{1 - F(t)\} \diagup H(t) \longrightarrow p \in [0,1]. \tag{4.1.1.2}$$

Dans ce cas il existe deux suites $a_n$et $b_n \in \mathbb{R}$ tels que

$$\lim_{t\to\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_i/a_n - b_n \leq x\right) = G_\alpha(x), \quad x \in \mathbb{R}, \tag{4.1.1.3}$$

où, G_α est une fonction de distribution d'une distribution stable. On écrit $H \in VR(-\alpha)$.

Une description de la domaine d'attraction d'une loi stable en terme de comportement asymptotique de la fonction caractéristique prés de zéro, est donné dans *Gnedenko et Klomogorov*([1]). Avec la notation

$$Ee^{itX} = U(t) + iV(t).$$

Ceci résulte de l'équivalence des conditions (4.1.1.1) et (4.1.1.2), avec les conditons

$$1 - U(1/t) \in RV(-\alpha) \tag{4.1.1.4}$$

et

$$\lim_{t\to\infty} \frac{txV(1/tx) - tV(1/t)}{t(1 - U(1/t))} = (2p-1)\left\{(1-\alpha)\tan\frac{\alpha\pi}{2}\right\} \times \frac{|x|^{1-\alpha} - 1}{1-\alpha}, \ x \in \mathbb{R}\diagdown\{0\}. \tag{4.1.1.5}$$

pour une constante $p \in [0,1]$..

La preuve de ces résultats est donnée dans *Geluk et deHann* ([2]).

La vitesse de covergence en (4.1.1.3) est étudie par plusieurs auteurs. *Gramèr* ([3]) considère le cas où $H(t)$ est une fonction puissance et un terme reste.

[1] B. V. Gnedenko, A. N. Kolmogorov, *Limit distributions for sums of independent random variables,* Addison-Wesley, Reading MA (1954).

[2] J.L. Geluk, L. de Haan, *Regular variation, extensions and Tauberian theorems,* CWI tract 40 Amsterdam (1987).

[3] H. Cramér, On the approximation to a stable probability distribution, D.Gilbarg and others (eds.), Studies in mathematical analysis and related topics. Essay in honor of George Pólya, Stanford University Press, Stanford, California (1971).

Au lieu des fonctions puissance avec un terme reste, c'est une extension naturelle de (4.1.1.1) pour considérer la variation de second ordre.

Variation régulière du second ordre

La fonction H est à variation régulière de second ordre avec le paramètre de premier ordre $-\alpha$, et une fonction auxiliaire A, s'il existe une fonction $A(t) \to 0$, $t \to \infty$ qui a un signe constant, telle que

$$\lim_{t\to\infty} \frac{\{H(tx) / H(t)\} - x^{-\alpha}}{A(t)} = M(x) \qquad (4.1.2.1)$$

avec

$$M(x) := cx^{-\alpha} \int_1^x u^{\rho-1} du,\ x > 0 \text{ pour } c \neq 0.$$

Notation :
$H \in 2RV(-\alpha, \rho, A)$.

Remarquons que pour $x > 0$

$$M(x) = \begin{cases} cx^{-\alpha} \log x, & \text{si} \quad \rho = 0, \\ cx^{-\alpha} \dfrac{x^\rho - 1}{\rho}, & \text{si} \quad \rho < 0. \end{cases}$$

Sous la condition de variation régulière de second ordre un raffinement naturel de (4.1.1.2) est la condition de second ordre d'équilibre de queue

$$\lim_{t\to\infty} \left[\frac{1 - F(t)}{H(t)} - p\right] A(t)^{-1} = r, \qquad (4.1.2.2)$$

pour un constant $p \in [0, 1]$ et $r \in \mathbb{R}$.

Avant que nous ne formulions les théorèmes nous définissons les constants suivants :

$$s_\alpha = \int_0^\infty x^{-\alpha} \sin x dx,$$

$$c_\alpha = \int_1^\infty x^{-\alpha} \cos x dx + \int_0^1 x^{-\alpha} (\cos x - 1) dx,$$

$$S_{\alpha,\rho} = \int_0^\infty x^{-\alpha} \frac{x^\rho - 1}{\rho} \sin x dx,$$

et

$$\begin{aligned}
C_{\alpha,\rho}(p,r) &= \int_1^\infty \left\{(2p-1)\,x^{-\alpha}\frac{x^\rho-1}{\rho}+2rx^{\rho-\alpha}\right\}\cos x dx \\
&+ \int_0^1 \left\{(2p-1)\,x^{-\alpha}\frac{x^\rho-1}{\rho}+2rx^{\rho-\alpha}\right\}(\cos x-1)\,dx.
\end{aligned}$$

Soit X une variable aléatoire d'une fonction de distribution F et esperence $\mathbb{E}X$. On utilise les notations suivantes :

$$\begin{aligned}
K(t) &= 1-F(t)-F(-t) \\
\lambda(t) &= Ee^{itX}=U(t)+iV(t)
\end{aligned}$$

Théorème 4.1.2.1 —
Soit $1<\alpha<2$, $\rho\le 0$, $\alpha-\rho<2$. Considérons les énoncées suivantes :

(i) $H(t)=1-F(t)+F(-t)\in 2RV(-\alpha,\rho,A_1)$ *et*

$$\lim_{t\to\infty}\left[\{1-F(t)\}\diagup H(t)-p\right]A_1(t)^{-1} \longrightarrow r, \tag{4.1.2.3}$$

quand $t\to\infty$ pour un constant $p\in[0,1]$ et $r\in\mathbb{R}$.

(ii) $1-U(1/t)\in 2RV(-\alpha,\rho,A_1)$ *et il existe*

$$c_\alpha^*\in\left[-\left(1+(1-\alpha)\,c_\alpha\right)\diagup s_\alpha,\left(1+(1-\alpha)\,c_\alpha\right)\diagup s_\alpha\right], \qquad d_\alpha^*(x)$$

tel que pour $x>0$

$$\left[\frac{txV(1/tx)-tV(1/t)}{t\,(1-U(1/t))}-c_\alpha^*\frac{x^{1-\alpha}-1}{1-\alpha}\right]A_1(t)\longrightarrow d_\alpha^*(x) \quad \text{quand } t\to\infty. \tag{4.1.2.4}$$

Alors (i) implique (ii). Inversement si $t^\alpha(1-F(t))$ et $t^\alpha H(t)$ sont monotone pour t suffisament large et $c_\alpha^*\neq 0$, alors (ii) implique (i).

Si (i) verifie, alors comme $t\to\infty$

$$1-U(1/t)-s_\alpha H(t) \sim S_{\alpha,\rho}A_1(t)\,H(t),$$

$$c_\alpha^*=(2p-1)\left[1+(1-\alpha)\,c_\alpha\right]\diagup s_\alpha,$$

et

$$
\begin{aligned}
d_\alpha^*(x) &= (2p-1)\int_1^x s^{-\alpha}\frac{s^\rho-1}{\rho}ds + (2p-1)^2 c_\alpha x^{1-\alpha}\frac{x^\rho-1}{\rho} \\
&+ \left\{C_{\alpha,\rho}(p,r) + \frac{2r}{\rho-\alpha+1}\right\}\left(x^{\rho+1-\alpha}-1\right) - S_{\alpha,\rho}c_\alpha^*\frac{x^{1-\alpha}-1}{1-\alpha}.
\end{aligned}
$$

Remarques :

1. L'expression $1-U(1/t) \in 2RV(-\alpha,\rho,A_1)$ vient de (4.1.2.4) dans le cas où $c_\alpha^* \neq 0$.
2. Les conditions de monotonie sur $t^\alpha(1-F(t))$ et $t^\alpha H(t)$ peut être remplacé d'autres conditions de monotonie sur les combinaisons linéaires de $t^\alpha(1-F(t))$ et $t^\alpha H(t)$ en adaptant la preuve.

Théorème 4.1.2.2 —

Supposons que $\mathbb{E}X = 0$ et $-1 < \rho \leq 0$. Les énoncées suivantes sont équivalentes.

(i)

$$H \in RV(\rho-2),$$

et

$$\{1-F(t)\} \diagup H(t) \longrightarrow p \in [0,1].$$

(ii)

$$1-U(1/t) \in 2RV(-2,\rho,A_1)$$

et quand $t \to \infty$

$$V(1/t) \diagup (1-U(1/t))\,A_1(t) \longrightarrow h_p \in \left[\zeta_\rho, -\zeta_\rho\right], \tag{4.1.2.5}$$

où

$$\zeta_\rho = \frac{\Gamma(1+\rho)\sin\frac{\rho\pi}{2}}{\rho(\rho-1)}.$$

De plus, étant donné (i) (ou (ii)), nous avons

$$\left(\frac{1-U(1/tx)}{1-U(1/t)} - x^{-2}\right)\frac{H_1(t)}{t^2H(t)} \longrightarrow (1+\rho g_\rho)\,x^{-2}\frac{x^\rho-1}{\rho},\ (t\to\infty), \tag{4.1.2.6}$$

où

$$g_{\rho} = \frac{1}{\rho(\rho-1)}\Gamma(1+\rho)\cos\frac{\rho\pi}{2}+\frac{1}{\rho}, \tag{4.1.2.7}$$

et le constant h_p dans (4.1.2.5) vérifiés

$$h_p = \frac{2p-1}{\rho(\rho-1)}\Gamma(1+\rho)\sin\frac{\rho\pi}{2}. \tag{4.1.2.8}$$

Estimation de la moyenne des distributions à queues lourdes

Ce paragraphe à pour objet d'étudier la distribution limite de la moyenne empirique pour les distributions appartenant au domaine d'attraction de la loi de Fréchét. On s'intérèsse à l'étude de la normalité asymptotique de la moyenne empirique des distributions à queues lourdes d'index $\alpha > 1$, car dans le cas où $\alpha < 2$ nous n'avons pas ce résultat. Pour traiter ce problème nous proposons un estimateur convenable d'une distribution limite sous des conditions de second ordre normale pour l'index $\alpha > 1$.

Soit $X_1, X_2, ..., X_n$ une suite de variables aléatoires [v.a] indépendantes, identiquement distribuées [iid], d'une fonction de distribution [f.d.] commune $F(x) = \mathbb{P}(X_1 \leq x)$, $x \in \mathbb{R}$. Soit

$$G(x) = \mathbb{P}(|X_1| \leq x) = F(x) - F(-x), \ x \in \mathbb{R}. \tag{4.2.1}$$

Supposons que G a une queue à variation régulière d'index $\alpha > 1$ et satifait la condition de balence de queue (voir, Csörgő, Horvàth et Mason, 1986).

$$\lim_{t\to\infty} \{1-G(tx)\} \diagup \{1-G(x)\} = x^{-\alpha}, \ x > 0, \tag{4.2.2}$$

et

$$\lim_{t\to\infty} \{1-F(x)\} \diagup \{1-G(t)\} = p, \tag{4.2.3}$$

où $0 < p < 1$.Si $0 < \alpha < 2$, alors (4.2.2) équivalent à dire que F dans le domaine d'attraction d'une loi stable. Supposons que G satisfait une condition de second ordre (voir de Haan et Stadmüller, 1996), ainsi que, pour une fonction A, qui ne change pas de signe prés de l'infini avec $A(t) \to 0$ quand $t \to \infty$, nous avons

$$\lim_{t\to\infty} \frac{1-G(tx)/1-G(t)-x^{-\alpha}}{A(t)} = x^{-\alpha}\frac{x^{-\beta}}{-\beta}, \quad x > 0, \tag{4.2.4}$$

et

$$\lim_{t\to\infty} \frac{1 - F(x)/1 - G(t) - p}{A(t)} = q, \tag{4.2.5}$$

où $\beta > 0$ et $q > 0$. La fonction A à variation régulière d'index $-\beta$.

Pour chaque entier $n \geq 1$, soit $X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$ les statistiques d'ordre basées sur l'échantillon $X_1, X_2, ..., X_n$.

Soit $\mu := \mathbb{E}X_1$ la valeur attendue de la distribution F. Si nous voulons évaluer une hypothèse de μ, ou construire un intervalle de confiance pour μ, donc la théorie de statistique classique suggère que nous basions notre procédure sur l'approximation normale au moyenne de l'échantillon $\hat{\mu}_n = n^{-1} \sum_{i=1}^{n} X_i$. Pour faire cette procédure plus robuste, nous pourrions légèrement couper la moyenne en enlevant des extrêmes de l'échantillon. C'est montré que cette procédure peut très améliorer la vitesse de convergence dans le théorème central limite (voir, Hall (1984)). Parmi les estimateurs robustes de la moyenne, on suggère, la *Moyenne Tronquée* $\hat{\mu}_n^{(T)}(k)$ et la *Moyenne Wonsorisée* $\hat{\mu}_n^{(W)}(k)$ qui sont définis comme suit :

$$\hat{\mu}_n^{(T)}(k) = \frac{1}{n - 2k} \sum_{i=k+1}^{n-k} X_{i,n}, \ k = 1, 2, .., \tag{4.2.6}$$

et

$$\hat{\mu}_n^{(W)}(k) = \frac{k}{n} X_{k,n} + \frac{1}{n} \sum_{i=k+1}^{n-k} X_{i,n} + \frac{k}{n} X_{n-k+1,n}, \ k = 1, 2, .. \tag{4.2.7}$$

Supposons que $k = k_n$, $n \geq 1$. En basant sur l'estimation des extrêmes quantiles, Peng (2001) a propsé une nouvelle version de la *Moyenne Wonsorisée* $\hat{\mu}_n^{(W)}(k)$ comme suit :

$$\mu_n^{(P)}(k) = \hat{\mu}_n^{-}(k) + \hat{\mu}_n(k) + \hat{\mu}_n^{+}(k),$$

où

$$\hat{\mu}_n(k) := \frac{1}{n} \sum_{i=k+1}^{n-k} X_{i,n}, \tag{4.2.8}$$

$$\hat{\mu}_n^{-}(k) := \frac{k}{n} X_{n-k+1,n} \frac{\hat{\alpha}_n^{(2)}}{\hat{\alpha}_n^{(2)} - 1} \quad \text{et} \quad \hat{\mu}_n^{+}(k) := \frac{k}{n} X_{k,n} \frac{\hat{\alpha}_n^{(1)}}{\hat{\alpha}_n^{(1)} - 1}, \tag{4.2.9}$$

avec

$$\hat{\alpha}_n^{(1)} := \left\{ \frac{1}{k} \sum_{i=1}^{k} \log^+ \left(-X_{n-i+1,n}\right) - \log^+ \left(-X_{n-k+1,n}\right) \right\}^{-1}, \tag{4.2.10}$$

et

$$\hat{\alpha}_n^{(2)} := \left\{ \frac{1}{k} \sum_{i=1}^{k} \log^+ \left(-X_{i,n}\right) - \log^+ \left(-X_{k,n}\right) \right\}^{-1}, \tag{4.2.11}$$

sont les deux versions de l'estimateur de Hill (Hill (1975)) de l'index de queue. Ici $\log^+ x = \log(x \vee 1)$.

Pour obtenir la normalité asymptotique de l'estimateur $\mu_n^{(P)}(k)$, *Peng* a considéré un cas spécial de la variation régulière de second ordre (4.2.4) - (4.2.5), appelé le modèle de *Hall*, c'est-à-dire,

$$1 - G(x) = cx^{-\alpha} \left\{ 1 + bx^{-\beta} + o\left(x^{-\beta}\right) \right\}, \tag{4.2.12}$$

où $c > 0$, $\beta > 0$ et $b \neq 0$ (voir, *Cheng* et *Peng,* 2001). Nous exposons maintenant le résultat concernant la normalité asymptotique de $\hat{\mu}_n^{(P)}$ dû à *Peng* (2001).

Normalité asymptotique de $\hat{\mu}_n^{(P)}$

Théorème 4.2.1.1 —

Supposons que (4.2.12) est vérifiée pour $c > 0$, $\beta > 0$ et $b \neq 0$. *De plus* $k = k_n$ *est tel que* $k \to \infty$, $k/n \downarrow 0$ *et* $k = o\left(n^{2\beta/(\alpha+2\beta)}\right)$ *quand* $n \to \infty$, *nous avons*

$$n^{1/2} \left\{ \hat{\mu}_n^{(P)}(k) - \mu \right\} / \sigma(k/n) \overset{\mathcal{D}}{\to} \mathcal{N}\left(0, \sigma_0^2\right), \tag{4.2.1.1}$$

où

$$\sigma^2(k/n) := \int_{k/n}^{1-k/n} \int_{k/n}^{1-k/n} \left(\min(u, v) - uv\right) dQ(u)\, dQ(v),$$

et

$$\sigma_0^2 := 1 + \left\{ \frac{(2-\alpha)\left(2\alpha^2 - 2\alpha + 1\right)}{2(\alpha-1)^4} + \frac{2-\alpha}{\alpha-1} \right\} \mathbf{1}(\alpha < 2) \tag{4.2.1.2}$$

Ici $\mathbf{1}(x < y)$ est la fonction indicatrice et $\mathcal{N}\left(\eta, \delta^2\right)$ la distribution normale de moyenne η et variance δ^2.

Loi du logarithme itéré pour la moyenne des distributions à queues lourdes

Étant donné la normalité asymptotique de l'estimateur $\hat{\mu}_n^{(P)}$, il est seulement naturel de supposer que, avec des conditions de régularité complémentaires sur la suite $(k_n)_{n\geq 1}$, la loi du logarithme itéré doit être aussi se tiennent. Nous devrons montrer que c'est, en effet, le cas.

Loi du logarithme itéré pour $\hat{\mu}_n^{(P)}$ [Necir (2003)]

Nous utilisons la notation,$\log_2 n = \log(\max(n,3))\,,\ n = 1, 2, ...$

Notre résultat est le suivant :

Théorème 4.3.1.1 —

Supposons que (4.2.12) est vérifiée pour $c > 0$, $\beta > 0$ *et* $b \neq 0$.Soit $k = k_n$ *est tel que* $k \to \infty$, $k/n \downarrow 0$ *et* $k = o\left(n^{2\beta/(\alpha+2\beta)}\right)$ *et* $\log_2 n = o(k)$. *Alors nous avons presque sûrement*

$$\limsup_{n\to\infty} \pm \frac{n^{1/2}\left(\hat{\mu}_n^{(P)}(k) - \mu\right)}{\left(2\sigma^2 (k_n/n) \log_2 n\right)^{1/2}} = \sigma_0,\ \textit{quand}\ n \to \infty. \tag{4.3.1.1}$$

En outre

$$\hat{\mu}_n^{(P)}(k) \overset{p.s}{\to} \mu,\ \textit{quand}\ n \to \infty.$$

Preuve.

Soit $U_1, U_2, ...$ une suite de [v.a] [iid], d'une fonction de distribution commune uniforme sur $(0,1)$. Pour tout entier $n \geq 1$, soit

$$\mathbb{G}_n(s) \ := \ n^{-1} \sum_{i=1}^{n} \mathbf{1}(U_i \leq u)\,, \quad 0 \leq s \leq 1,$$

la fonction de distribution empirique basée sur le premier n de ces variables aléatoires, et soit $\mathbb{V}_n(t) = U_{i,n}$, pour $(i-1)/n < t < i/n$, $i = 1, ..., n$, avec $\mathbb{V}_n(0) = U_{1,n}$, où $U_{1,n} \leq U_{2,n} \leq ... \leq U_{n,n}$ désigne les statistiques d'ordre basées sur $U_1, U_2, ..., U_n$. Nous définissons le processus empirique uniforme

$$\alpha_n(s) \ := \ n^{1/2}(\mathbb{G}_n(s) - s)\,, \quad 0 \leq s \leq 1,$$

et le processus des quantiles uniforme

$$\beta_n(s) \ := \ n^{1/2}(\mathbb{V}_n(s) - s)\,, \quad 0 \leq s \leq 1.$$

Le lemme suivant donne une approximation pour le processus $\alpha_n(.)$ en terme d'une suite de ponts Brownien.

Lemme 4.2.1.1 —

Il existe une suite de ponts Brownien $(B_i)_{i\geq 1}$, *tel que, pour tout* $\nu \in [0, 1/4]$, *nous avons*

$$\sup_{1/n\leq s\leq 1-1/n} n^{\nu}\frac{\left|\sqrt{n}\left(\mathbb{G}_n\left(s\right)-s\right)-B_n\left(s\right)\right|}{\left(s\left(1-s\right)\right)^{1/2-\nu}}=o_p\left(1\right), \quad \textit{quand,} \quad n\to\infty,$$

Lemme 4.2.1.2 —
Supposons que (4.2.12) est vérifiée, et

$$k\to\infty, \quad k/n\to 0, \quad \textit{quand} \quad n\to\infty.$$

Alors

$$\lim_{n\to\infty}\left(k_n/n\right)^{1/2}\int\limits_{k/n}^{1-k/n}\left(1-t\right)dQ\left(t\right)\big/\sigma\left(k/n\right) \quad = \quad S_{p,\alpha}\left(1-p\right)^{1/\alpha}\mathbf{1}\left(\alpha<2\right), \tag{4.3.1.2}$$

et

$$\lim_{n\to\infty}\left(k/n\right)^{1/2}\int\limits_{k/n}^{1-k/n} tdQ\left(t\right)\big/\sigma\left(k/n\right) \quad = \quad S_{p,\alpha}p^{1/\alpha}\mathbf{1}\left(\alpha<2\right). \tag{4.3.1.3}$$

où

$$S_{p,\alpha} \quad = \quad \left[\frac{2-\alpha}{2\left(p^{2/\alpha}+\left(1-p\right)^{2/\alpha}\right)}\right]^{1/2}.$$

Preuve.
Voir Csörgő et al. (1986a, b). □

Ensuite, nous utilisons les deux statistiques suivantes

$$\theta_n^-\left(k\right)=k^{-1}\sum_{i=1}^{k}-\log^+\left(U_{i,n}/U_{k,n}\right), \tag{4.3.1.4}$$

et

$$\theta_n^+\left(k\right)=k^{-1}\sum_{i=1}^{k}-\log^+\left\{\left(1-U_{n-i+1,n}\right)\big/\left(1-U_{n-k,n}\right)\right\}. \tag{4.3.1.5}$$

Preuve du théorème 4.2.1.1.

Pour tout entier $n\geq 1$, nous écrivons

$$\mu_n^-\left(k\right):=\int\limits_0^{k/n}Q\left(s\right)ds, \quad \mu_n^+\left(k\right):=\int\limits_{1-k/n}^{1}Q\left(s\right)ds,$$

et

$$\mu_n(k) := \int_{k/n}^{1-k/n} Q(s)\,ds.$$

Evidemment, nous avons

$$\mu = \mathbb{E}X_1 = \int_0^1 Q(s)\,ds = \mu_n^-(k) + \mu_n(k) + \mu_n^+(k).$$

Alors nous écrivons

$$\hat{\mu}_n^{(P)}(k) - \mu = \left\{\hat{\mu}_n^-(k) - \mu_n^-(k)\right\} + \left\{\hat{\mu}_n(k) - \mu_n(k)\right\} + \left\{\hat{\mu}_n^+(k) - \mu_n^+(k)\right\}.$$

Supposons que $\hat{\alpha}_n^{(1)}$ et $\hat{\alpha}_n^{(2)}$ sont exprimé en termes de $X_{j,n} = Q(U_{j,n})$, $j = 1, ..., n$. De la décomposition de *Peng* (2001), nous avons

$$\begin{aligned} \hat{\mu}_n^-(k) - \mu_n^-(k) &= \frac{k}{n} X_{k,n} \frac{\hat{\alpha}_n^{(1)}}{\hat{\alpha}_n^{(1)} - 1} - \int_0^{k_n/n} Q(s)\,ds \\ : &= I_{n,1} + I_{n,2} + I_{n,3} + I_{n,4} + I_{n,5}, \end{aligned}$$

où

$$\begin{aligned} I_{n,1} &:= \frac{k}{n} \frac{\alpha \hat{\alpha}_n^{(1)} Q(U_{k,n})}{\left(\hat{\alpha}_n^{(1)} - 1\right)(\alpha - 1)} \left\{ \frac{1}{k} \sum_{i=1}^{k} \left(\log \frac{Q(U_{i,n})}{Q(U_{k,n})} - \log \left(\frac{U_{i,n}}{U_{k,n}} \right)^{-1/\alpha} \right) \right\}, \\ I_{n,2} &:= \frac{k}{n} \frac{\alpha \hat{\alpha}_n^{(1)} Q(U_{k,n})}{\left(\hat{\alpha}_n^{(1)} - 1\right)(\alpha - 1)} \left\{ \frac{1}{k} \sum_{i=1}^{k} \left(\log \left(\frac{U_{i,n}}{U_{k,n}} \right)^{-1/\alpha} - 1/\alpha \right) \right\}, \\ I_{n,3} &:= \frac{k}{n} \frac{\alpha Q(k/n)}{\alpha - 1} \left\{ \frac{Q(U_{k,n})}{Q(k/n)} - \left(\frac{n}{k} U_{k,n} \right)^{-1/\alpha} \right\}, \\ I_{n,4} &:= \frac{k}{n} \frac{\alpha Q(k/n)}{\alpha - 1} \left\{ \left(\frac{n}{k} U_{k,n} \right)^{-1/\alpha} - 1 \right\}, \end{aligned}$$

et

$$I_{n,5} := \frac{k}{n} \frac{\alpha Q(k/n)}{\alpha - 1} - \int_0^{k_n/n} Q(s)\,ds.$$

De même, nous avons

$$\hat{\mu}_n^+(k) - \mu_n^+(k) = \frac{k}{n} X_{n-k+1,n} \frac{\hat{\alpha}_n^{(2)}}{\hat{\alpha}_n^{(2)} - 1} - \int_{1-k_n/n}^{1} Q(s)\, ds$$
$$: = III_{n,1} + III_{n,2} + III_{n,3} + III_{n,4} + III_{n,5},$$

où

$$III_{n,1} := \frac{k}{n} \frac{\alpha \hat{\alpha}_n^{(2)} Q(U_{n-k+1,n})}{\left(\hat{\alpha}_n^{(2)} - 1\right)(\alpha - 1)} \times \left\{ \frac{1}{k} \sum_{i=1}^{k} \left(\log \frac{Q(U_{n-i+1,n})}{Q(U_{n-k+1,n})} - \log \left(\frac{U_{n-i+1,n}}{U_{n-k+1,n}} \right)^{-1/\alpha} \right) \right\},$$

$$III_{n,2} := \frac{k}{n} \frac{\alpha \hat{\alpha}_n^{(2)} Q(U_{n-k+1,n})}{\left(\hat{\alpha}_n^{(2)} - 1\right)(\alpha - 1)} \times \left\{ \frac{1}{k} \sum_{i=1}^{k} \left(\log \left(\frac{U_{n-i+1,n}}{U_{n-k+1,n}} \right)^{-1/\alpha} - 1/\alpha \right) \right\},$$

$$III_{n,3} := \frac{k}{n} \frac{\alpha Q(1 - k/n)}{\alpha - 1} \left(\frac{Q(U_{n-k+1,n})}{Q(1 - k/n)} - \left(\frac{n}{k} U_{n-k+1,n} \right)^{-1/\alpha} \right),$$

$$III_{n,4} := \frac{k}{n} \frac{\alpha Q(1 - k/n)}{\alpha - 1} \left(\left(\frac{n}{k} U_{n-k+1,n} \right)^{-1/\alpha} - 1 \right),$$

et

$$III_{n,5} := \frac{k}{n} \frac{\alpha Q(1 - k/n)}{\alpha - 1} - \int_{1-k_n/n}^{1} Q(s)\, ds.$$

D'autre part, de Csörgő et al. (1986), nous avons

$$\hat{\mu}_n(k) - \mu_n(k) = II_{n,1} + II_{n,2} + II_{n,3},$$

où

$$II_{n,1} \ := \ -\int\limits_{k/n}^{1-k/n} \{G_n(s) - s\}\, dQ(s),$$

$$II_{n,2} \ := \ -\int\limits_{U_{n-k,n}}^{1-k/n} \{G_n(s) - k/n\}\, dQ(s),$$

et

$$II_{n,3} := -\int\limits_{1-k/n}^{U_{n-k,n}} \{G_n(s) - (1-k/n)\}\, dQ(s).$$

Le lemme suivant donne une approximation de $\hat{\mu}_n(k) - \mu_n(k)$ en terme d'une suite des ponts Brownien $(B_i)_{i \geq 1}$.

Lemme 4.2.1.3 —

Supposons que (4.2.12) est vérifiée pour $\alpha > 1$, *et* $k \to \infty$, $k/n \to 0$, *quand* $n \to \infty$. *Alors, nous avons*

$$n^{1/2}\{\hat{\mu}_n(k) - \mu_n(k)\} \diagup \sigma(k/n) = -\int\limits_{k/n}^{1-k/n} B_i(s)\, dQ(s) \diagup \sigma(k/n) + o(1).$$

Preuve.

De la preuve du théorème 1 de Csörgő et al. (1986a,b), nous avons

$$\frac{n^{1/2}}{\sigma(k/n)} II_{2,n} \xrightarrow{\mathcal{P}} 0, \quad \frac{n^{1/2}}{\sigma(k/n)} II_{3,n} \xrightarrow{\mathcal{P}} 0,$$

et

$$\frac{n^{1/2}}{\sigma(k/n)} II_{1,n} = -\int\limits_{k/n}^{1-k/n} B_n(s)\, dQ(s) \diagup \sigma(k/n) + o(1).$$

D'où la preuve du Lemme 4.2.1.3. $\square$

Ensuite, nous présentons le comportement au limite de $\hat{\mu}_n^-(k)$ et $\hat{\mu}_n^+(k)$.

Lemme 4.2.1.4 —

Sous les suppositions du théorème 4.2.1.1, nous avons

$$\sqrt{n}\sigma(k/n)^{-1} I_1 \xrightarrow{\mathcal{P}} 0, \quad \sqrt{n}\sigma(k/n)^{-1} I_3 \xrightarrow{\mathcal{P}} 0, \quad \sqrt{n}\sigma(k/n)^{-1} I_5 \xrightarrow{\mathcal{P}} 0,$$

et

$$\sqrt{n}\sigma\left(k/n\right)^{-1}III_1\xrightarrow{\mathcal{P}}0,\quad \sqrt{n}\sigma\left(k/n\right)^{-1}III_3\xrightarrow{P}0,\quad \sqrt{n}\sigma\left(k/n\right)^{-1}III_5\xrightarrow{\mathcal{P}}0,$$

$$\begin{aligned}\sqrt{n}\sigma\left(k/n\right)^{-1}II_1 &= -\int_{k/n}^{1-k/n}B_i\left(s\right)dQ\left(s\right)\diagup\sigma\left(k/n\right)+o\left(1\right)\\ &= L_n+o\left(1\right).\end{aligned}$$

Le Lemme suivant donne une représentation de θ_n^- et θ_n^+ en terme d'une suite de ponts Brownien $\left(B_i\right)_{i\geq 1}$.

Lemme 4.2.1.5 —

Soit $k=k_n$ *est tel que* $k\to\infty$, $k/n\downarrow 0$.*Alors, nous avons*

$$\begin{aligned}\sqrt{k}\left\{\theta_n^+-1\right\} &= (n/k)^{1/2}B_i\left(1-k/n\right)-(n/k)^{1/2}\\ &\times\int_{1-k/n}^{1}\frac{B_i\left(s\right)}{1-s}ds+o\left(1\right).\\ &= C_n-H_n+o\left(1\right),\\ \sqrt{k}\left\{\theta_n^--1\right\} &= -\left(n/k\right)^{1/2}B_i\left(k/n\right)+\left(n/k\right)^{1/2}\\ &\times\int_0^{k/n}\frac{B_i\left(s\right)}{s}ds+o\left(1\right).\\ &= -Z_n+Y_n+o\left(1\right).\end{aligned}$$

et

$$\begin{aligned}\sqrt{k}\left\{-\log\left(1-U_{n-k,n}\right)+\log k/n\right\} &= -(n/k)^{1/2}B_i\left(1-k/n\right)+o\left(1\right).\\ &= -C_n+o\left(1\right),\end{aligned}$$

$$\begin{aligned}\sqrt{k}\left\{-\log\left(U_{k,n}\right)+\log k/n\right\} &= \left(n/k\right)^{1/2}B_i\left(k/n\right)+o\left(1\right).\\ &= Z_n+o\left(1\right).\end{aligned}$$

Preuve.

Voir les Théorèmes (2.3) et (2.4) de Csörgő et Mason (1985). □

Ensuite, nous présentons un Lemme Dans lequel nous donnons des propriétés de $Q(.)$.

Lemme 4.2.1.6 —

Sous les supposition du théorème 4.2.1.1, nous avons

$$\frac{\sqrt{k/n}Q(U_{k,n})}{\sigma(k/n)} \xrightarrow{\mathcal{P}} -S_{p,\alpha}(1-p)^{1/\alpha} 1_{(\alpha<2)},$$

$$\frac{\sqrt{k/n}Q(U_{n-k+1,n})}{\sigma(k/n)} \xrightarrow{\mathcal{P}} S_{p,\alpha}p^{1/\alpha}1_{(\alpha<2)}.$$

où

$$S_{p,\alpha} = \left[\frac{2-\alpha}{2\left(p^{2/\alpha}+(1-p)^{2/\alpha}\right)}\right]^{1/2}.$$

Posons

$$a_0 = \left[\frac{2-\alpha}{2\left(p^{2/\alpha}+(1-p)^{2/\alpha}\right)}\right]^{1/2} \mathbf{1}(\alpha<2).$$

Ensuite nous utilisons la notation suivante :

Soit $n^{1/2}\left\{\hat{\mu}_n^{(P)}(k)-\mu\right\}/\sigma(k/n) := \Psi_{i,n}$, où

$$\begin{aligned}
\Psi_{i,n} &:= -a_0(1-p)^{1/\alpha}\frac{\alpha}{(\alpha-1)^2}(Y_n-Z_n) - a_0(1-p)^{1/\alpha}\frac{1}{\alpha-1}(Z_n-L_n)\\
&\quad +a_0p^{1/\alpha}\frac{\alpha}{(\alpha-1)^2}(C_n-H_n) - a_0p^{1/\alpha}\frac{\alpha}{\alpha-1}C_n + o_p(1).\\
&= -a_0(1-p)^{1/\alpha}\frac{\alpha}{(\alpha-1)^2}Y_n + a_0(1-p)^{1/\alpha}\frac{1}{(\alpha-1)^2}Z_n - L_n\\
&\quad +a_0p^{1/\alpha}\frac{1}{(\alpha-1)^2}C_n - a_0(1-p)^{1/\alpha}\frac{\alpha}{(\alpha-1)^2}H_n + o_p(1).\\
&\xrightarrow{d} \mathcal{N}(0,\sigma_0^2).
\end{aligned}$$

où $\sigma_0^2 := \lim_{n\to\infty}[var\Psi_{i,n}]$.

En utilisant le Lemme 4.2.1.1, un calcule directe donne

$$\sigma_0^2 = 1 + \left\{\frac{(2-\alpha)(2\alpha^2-2\alpha+1)}{2(\alpha-1)^4} + \frac{2-\alpha}{\alpha-1}\right\}\mathbf{1}_{(\alpha<2)}.$$

Preuve du théorème 4.3.2.1.
Nous utilisons la même décomposition (*Peng* (2001)).

Le Lemme suivant donne une approximation forte pour les processus $\alpha_n(.)$ et $\beta_n(.)$ en termes des sommes des ponts Brownien dû à Einmahl et Mason (1988).

Lemme 4.3.2.1 —

Il exist une suite de ponts Brownien $(B_i)_{i\geq 1}$ *tel que, pour chaque* $0 < \eta < 1/2$, *et pour tout* n *grand, nous avons, presque sûrement*

$$\sup_{1/(n+1)\leq s\leq 1-1/(n+1)} s^{-\eta}\left|\alpha_n(s) - n^{-1/2}\sum_{i=1}^{n} B_i(s)\right| = o\left((\log_2 n)^{1/2}\right),$$

et

$$\sup_{1/(n+1)\leq s\leq 1-1/(n+1)} s^{-\eta}\left|\beta_n(s) - n^{-1/2}\sum_{i=1}^{n} B_i(s)\right| = o\left((\log_2 n)^{1/2}\right).$$

Preuve.
Voir Théorème 4 de Einmahl et Mason (1986a). □

Le Lemme suivant donne les représentations de $\theta_n^-(k)$ et $\theta_n^+(k)$ en termes de processus $\beta_n(.)$

Lemme 4.3.2.2 —

Soit $k = k_n$ *est tel que* $k \to \infty$, $k/n \downarrow 0$ *et* $\log_2 n = o(k)$. *Alors, pour chaque* $0 < \tau < 1/2$ *fixé, pour tout* n *grand, nous avons, presque sûrement*

$$\begin{aligned} k^{1/2}\left(\theta_n^-(k) - 1\right) &= (n/k)^{1/2}\int_\tau^1 s^{-1}\beta_n(ks/n)\,ds \\ &\quad - (n/k)^{1/2}\beta_n(k/n) + O\left(\tau^{1/2}(\log_2 n)^{1/2}\right), \end{aligned}$$

et

$$\begin{aligned} k^{1/2}\left(\theta_n^+(k) - 1\right) &= (n/k)^{1/2}\int_\tau^1 s^{-1}\beta_n(1 - ks/n)\,ds \\ &\quad - (n/k)^{1/2}\beta_n(1 - k/n) + O\left(\tau^{1/2}(\log_2 n)^{1/2}\right). \end{aligned}$$

Preuve.
Notons que $\theta_n^+(k)$ et $\theta_n^-(k)$ sont des cas spéciaux de statistique donnée dans le Théorème 1 de Necir (2003). Nous omettons les détails. □

Ensuite, nous présentons deux Lemmes dans lequel nous donnons les propriétés de la fonction $Q(.)$.

Lemme 4.3.2.3 —

Supposons que (4.2.1) vérifié. Alors, les deux résultats suivants vérifiés

i) Les deux fonctions $-Q(.)$ *et* $Q(1-.)$ *sont à variation régulière près du zéro avec l'index* $-1/\alpha$.

ii) $-Q(s) \sim Q(1-s)$ *quand* $s \downarrow 0$.

Preuve.

Voir les conditions S1-S2 de Csörgő et al. (1986). □

Lemme 4.3.2.4 —

Supposons que (4.2.1) vérifié pour $0 < p < 1$ *et* $0 < \alpha < 2$. *Alors,*

$$\lim_{n\to\infty} \sqrt{\frac{k}{n}} \frac{Q(1-k/n)}{\sigma(k/n)} = \lim_{n\to\infty} -\sqrt{\frac{k}{n}} \frac{Q(k/n)}{\sigma(k/n)} = C(\alpha, p) \tag{4.3.1.6}$$

où

$$C(\alpha, p) = \sqrt{\frac{2-\alpha}{2\left(p^{2/\alpha} + (1-p)^{2/\alpha}\right)}} (1-p)^{1/\alpha}. \tag{4.3.1.7}$$

Preuve.

Voir Lemme 1 de Csörgő, Horvàth et Mason (1986). □

Lemme 4.3.2.5 —

Supposons que (4.2.12) vérifié pour $c > 0$, $\beta > 0$ *et* $b \neq 0$. *Alors*

$$\lim_{t\to\infty} \frac{\dfrac{Q(1-tx)}{Q(1-t)} - x^{-1/\alpha}}{-\alpha^{-2} b\beta c^{-\beta/\alpha} t^{\beta/\alpha}} = \lim_{t\to\infty} \frac{\dfrac{Q(tx)}{Q(t)} - x^{-1/\alpha}}{-\alpha^{-2} b\beta c^{-\beta/\alpha} t^{\beta/\alpha}} = x^{-\alpha} \frac{x^{\beta/\alpha} - 1}{-\beta}, \; x > 0 \tag{4.3.1.8}$$

et

$$\lim_{t\to\infty} \frac{\log \dfrac{Q(1-tx)}{Q(1-t)} - \log x^{-1/\alpha}}{-\alpha^{-2} b\beta c^{-\beta/\alpha} t^{\beta/\alpha}} = \lim_{t\to\infty} \frac{\log \dfrac{Q(tx)}{Q(t)} - \log x^{-1/\alpha}}{-\alpha^{-2} b\beta c^{-\beta/\alpha} t^{\beta/\alpha}} = x^{-\alpha} \frac{x^{\beta/\alpha} - 1}{-\beta}, \; x > 0 \tag{4.3.1.9}$$

Le Lemme suivant donne une approximation forte de $\hat{\mu}_n(k) - \mu_n(k)$ en terme d'une suite des ponts Brownien $(B_i)_{i\geq 1}$.

Lemme 4.3.2.6 —

Supposons que (4.2.1) vérifié pour $\alpha > 1$. *Soit* $k = k_n$ *est tel que* $k \to \infty$, $k/n \downarrow 0$ *et* $\log_2 n = o(k)$. *Alors, nous avons presque sûrement*

$$\frac{n^{1/2}(\hat{\mu}_n(k) - \mu_n(k))}{(\log_2 n)^{1/2}\sigma(k/n)} = \frac{-\sum_{i=1}^{n}\int_{k/n}^{1-k/n} B_i(s)\,dQ(s)}{(\log_2 n)^{1/2}\sigma(k/n)} + o(1).$$

Preuve.

Notons que du Lemme 1 de Haeusler et Mason (1987), nous avons, presque sûrement

$$\frac{n^{1/2}II_{n,2}}{(\log_2 n)^{1/2}\sigma(k/n)} = \frac{n^{1/2}II_{n,3}}{(\log_2 n)^{1/2}\sigma(k/n)} = o(1), \text{ quand } n \to \infty.$$

D'autre part, La représentation de $II_{n,1}$ en terme du processus empirique $\alpha_n(.)$ donne

$$n^{1/2}II_{n,1} = -\int_{k/n}^{1-k/n} \alpha_n(s)\,dQ(s).$$

Par le Lemme 4.3.2.1, nous avons

$$n^{1/2}II_{n,1} = \Omega_n + \Delta_n,$$

où

$$\Omega_n = -n^{1/2}\sum_{i=1}^{n}\int_{k/n}^{1-k/n} B_i(s)\,dQ(s),$$

et

$$\Delta_n = -\int_{k/n}^{1-k/n}\left\{\alpha_n(s) - -n^{1/2}\sum_{i=1}^{n} B_i(s)\right\}dQ(s).$$

Alors pour compléter la preuve de Lemme 4.3.2.6 nous montrons que, presque sûrement

$$\Delta_n / \sigma(k/n) = o\left((\log_2 n)^{1/2}\right), \text{ quand } n \to \infty.$$

En fait, observons que, en vue de Lemme 4.3.2.6, pour tous $0 < \eta < 1/2$, nous avons, presque sûrement

$$\Delta_n = o\left((\log_2 n)^{1/2}\right)\int_{k/n}^{1-k/n} s^{\eta}\,dQ(s), \text{ quand } n \to \infty.$$

D'autre part, du la condition (4.3) de Csörgő et al. (1986a,b), pour tout n grand, et pour chaque $0 < \eta < 1/2$, nous avons

$$\int_{k/n}^{1-k/n} s^{\eta} dQ(s) \diagup \sigma(k/n) = \begin{cases} O\left(n^{-\eta}k^{\eta-1/2}\right) & \text{si} \quad 1 < \alpha < 2, \\ o\left(n^{-\eta}k^{\eta-1/2}\right) & \text{si} \quad \alpha \geq 2. \end{cases}$$

Cela implique aisément, pour $0 < \eta < 1/2$ et $k \to \infty$ quand $n \to \infty$, que pour tout $\alpha > 1$, nous avons

$$\int_{k/n}^{1-k/n} s^{\eta} dQ(s) \diagup \sigma(k/n) = o(1), \text{ quand } n \to \infty,$$

Qui achève la preuve de Lemme 4.3.2.6. □

Ensuite, Nous présentons seulement des détails pour la preuve concernant le comportement de limite de $\hat{\mu}_n^-$, la preuve pour $\hat{\mu}_n^+$ étant très semblable.

Lemma 4.3.2.7 —

Sous les suppositions du Théorème 4.3.2.1, nous avons presque sûrement

$$\frac{n^{1/2} I_{n,1}}{(\log_2 n)^{1/2} \sigma(k/n)} = \frac{n^{1/2} I_{n,2}}{(\log_2 n)^{1/2} \sigma(k/n)} = o(1), \text{ quand } n \to \infty,$$

et

$$\frac{n^{1/2} III_{n,1}}{(\log_2 n)^{1/2} \sigma(k/n)} = \frac{n^{1/2} III_{n,2}}{(\log_2 n)^{1/2} \sigma(k/n)} = o(1), \text{ quand } n \to \infty.$$

Preuve.

Nous écrivons

$$H_n := k^{1/2} \left\{ \frac{1}{k} \sum_{i=1}^{k} \left(\log \frac{Q(U_{i,n})}{Q(U_{k,n})} - \log \left(\frac{U_{i,n}}{U_{k,n}} \right)^{-1/\alpha} \right) \right\}.$$

Alors nous pouvons récrire $I_{n,1}$ comme suit :

$$I_{n,1} = \frac{k^{1/2}}{n} \frac{\alpha \hat{\alpha}_n^{(1)} Q(U_{k,n})}{\left(\hat{\alpha}_n^{(1)} - 1\right)(\alpha - 1)} H_n.$$

Nous montrons d'abord que, $H_n = o(1)$, quand $n \to \infty$, presque sûrement.

Soit

$$D_{n,i} = \log \frac{Q\left(U_{i,n}\right)}{Q\left(U_{k,n}\right)} - \log \left(\frac{U_{i,n}}{U_{k,n}}\right)^{-1/\alpha}, \ i = 1, ..., k.$$

De Kiefer (1972), pour tous n suffisament large, pour chaque suite $k = k_n$ tel que $k \to \infty$, $k/n \downarrow 0$ et $k \diagup \log_2 n \to \infty$, nous avons, presque sûrement

$$nU_{i,n} \diagup > (1 - \varepsilon),$$

et

$$1 - \varepsilon < nU_{k,n} \diagup k < (1 + \varepsilon).$$

Est alors aisément vérifié que, en vue de(3.2.1), nous avons

$$\frac{(1-\varepsilon)\,\ell_n}{(1+\varepsilon)\,k} < \frac{U_{i,n}}{U_{k,n}} < 1, \text{ pour } i = 1, ..., k.$$

Annexe

Définitions et notations

Définition 1 (Fonctions à variations lentes) —

Une fonction l, positive, Lebesgue mesurable sur $(0,+\infty)$, est à variations lentes en $+\infty$ *(noté $l \in \mathcal{R}_0$) si*

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

Définition 2 (Fonctions à variations régulières) —

Une fonction f, positive, Lebesgue mesurable sur $(0,+\infty)$, *est à variations réguliéres d'indice $\rho \in \mathbb{R}$ en $+\infty$ (noté $f \in \mathcal{R}_\rho$) si*

$$\forall t > 0, \ \lim_{x\to\infty} \frac{f\,(tx)}{f\,(x)} = t^\rho.$$

Définition 3 —

Un processus stochastique $\{\mathcal{B}(t); 0 \leq t \leq 1\}$ est appelé pont Brownien si

(i) *La distribution jointe de $(\mathcal{B}(t_1), \mathcal{B}(t_2), ..., \mathcal{B}(t_k))$ $(0 \leq t_1 < t_2 < t_k \leq 1\,;\, n = 1, 2, ...)$ est Gaussienne, avec $EB(t) \equiv 0$,*

(ii) *La fonction de covariance de $B(t)$ est*

$$R(r,t) = \mathbb{E}\mathcal{B}(s)\mathcal{B}(t) = s \wedge t - st,$$

(iii) *La trajectoire de de la fonction de l'échantillon de $B(t;\omega)$ est continue dans t avec probabilité 1.*

Définition 4 —

Un processus stochastique $\{W\,(t;\omega) = W\,(t)\,; 0 \leq t < \infty\}$, où $\omega \in \Omega$, et $\{\Omega,\ \mathcal{A},\ \mathbb{P}\}$ *est un espace de probabilité, est appelé processus de Wiener si*

(i) $W\,(t) - W\,(s) \in \mathcal{N}\,(0,\ t-s)$ *pour tous* $0 \leq s < t < \infty$ *et* $W\,(0) = 0$,

(ii) $W\,(t)$ *est un processus d'incrément indépendant,i.e,* $W\,(t_2) - W\,(t_1)\,,$ $W\,(t_4) - W\,(t_3)\,, ..., W\,(t_{2i}) - W\,(t_{2i-1})$ *sont des variables aléatoires indépendents pour tous* $0 \leq t_1 < t_2 \leq t_3 < t_4 \leq ... \leq t_{2i-1} < t_{2i} < \infty$ $(i = 2, 3, ...)\,,$

(iii) *La trajectoire de de la fonction de l'échantillon de $W(t;\omega)$ est continue dans t avec probabilité 1.*

Notons que (i) et (ii) implique que La fonction de covariance de $W(t)$ est

$$R(r,t) = \mathbb{E}W(s)W(t) = s \wedge t.$$

Définition 5 —

On dit que F *de queue lourde si sa fonction de queue* $\bar{F} := 1 - F$, *est donnée par :*

$$\bar{F} = x^{-\alpha} \ell_F(x),$$

où α *est l'index de queue et* ℓ_F *est une fonction à variation lente.*

D'une manière équivalente, la fonction de quantile

$$Q(s) := \inf \{x : F(x) \geq s\}$$

satisfait

$$Q(1-p) = p^{-1/\alpha} \ell(1/p),$$

avec ℓ *à variation lente.*

Bibliographie

[1] **Barry C, A., Balakrishnan, N. et Nagaraja, H.N. (1992).** A first cours in order statistics.

[2] **Csörgő, M., Csörgő, S., Horváth, L., D.M.** Weighted empirical and quantile processus. *Ann. Probability. To appear*

[3] **Csörgő, M., Csörgő, S., Horváth, L., D.M.** Normal and stable converence of integral functionals of the empirical distribution function. *Ann. Probability. To appear*

[4] **S. Csörgő, L. Horváth and D. M. Mason (1986).** What portion of the sample makes a partial sum asymptotically stable or normal ? Proba. Th. Rel. Fields 72, 1-16.

[5] **S. Csörgő, and D. M. Mason (1986).** The asymptotic distribution of sums of extreme values from a regulary varying distribution. *The Annals of Probability.* Vol. 14. No, 3. 974-983.

[6] **Deheuvels P. and Mason, D. M. (1991).** A tail-empirical process approach to some non-standard laws of the iterated logarithm. *Ann. Statist.,* ***17,*** 4, 1833-1855.

[7] **Deheuvels P. and Mason, D. M. (1990).** Nonstandard functional laws of the iterated logarithm for tail empirical and quantile processes. *The Annals of Probability,* Vol. 18, No. 4, 1693-1722.

[8] **Deheuvels, P. and Haeusler, E. and Mason, D. (1990).** On the almost sure behavior of sums of extreme values from a distribution in the domain of attraction of a Gumbel law. *Bull. Sci. Math.* **114,** 61-95.

[9] **Dekkers, A. L. M., Einmahl J. H. J., and de Haan (1989).** A moment estimator for the index of an extrem-value distribution, *Ann. Statist.,* ***17***, 4, 1833-1855.

[10] **Einmahl, J. H., and Mason, D. M. (1988).** Strong limit theorems for weighted quantile process. Ann. Prob., Vol. 16, No. 4, 1623-1643.

[11] **Erich Haeusler and David M. Mason (1987).** A law of iterated logarithm for sums of extreme values from a distribution with a regulary varying upper tail. The *Annals of Probability,*Vol. 15, N0. 3, 932-953.

[12] **Erich Haeusler and David M. Mason (1987).** Laws of iterated logarithm for sums of the middle portion of the sample. *Math. Proc. Camb. phil. Soc.*, 101-301. *Printed in Great Britain.*

[13] **Gane Samb LO (1989).** A note on the asymptotic normality of sums of extreme values. Journal of Statistical Planning and Inference 22, 127-136. North-Holland.

[14] **J. L. Geluk and L. Peng (2000).** Second-order regular variation and the domain of attraction of stable distribution.

[15] **Liang Peng (2001).** Estimating the mean of a heavy tailed distribution. Statistics & Probability Letters 52, 255-264.

[16] **Necir A. (2002).** A functional laws of the iterated logarithm for weighted least-squares estimators of tail index, (submitted).

[17] **Necir A. (2002).** A functional laws of the iterated logarithm for kernel-type estimator of tail indices, *J. Satist. Plann. Inference.* (submitted).

[18] **Necir A. (2003).** A law of the iterated logarithm for mean of heavy tailed distribution, (submited).

[19] **Pickand J. (1975).** Statistical inferrence using extrem order statistics. *Annals of statistics,* **3**, 119-131.

[20] **De Wolf, P.P. (1999).** Estimating the extreme value index, PhD thesis, TU Delft, 1999.