

MA-510-55-1

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ de BLIDA 1

Faculté de Sciences

Département de Mathématiques

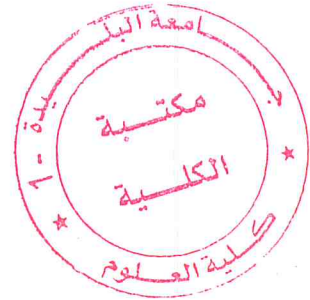


Mémoire de fin d'études

En vue de l'obtention du diplôme de

MASTER EN MATHEMATIQUE

Spécialité : Modélisation Stochastique et Statistique



Intitulé du mémoire

Estimation non-paramétrique des quantiles extrêmes conditionnels

Présenté par :

Benyoucef Mohamed Amine

Membres de jury :

ami Omar

Président de jury

USDB

MAA

aidi Mohamed

Examineur

ENST

MAA

rihi Redhouane

Examineur

ENSH

MAA

Rassoul Abdelazziz

Promoteur

ENSH

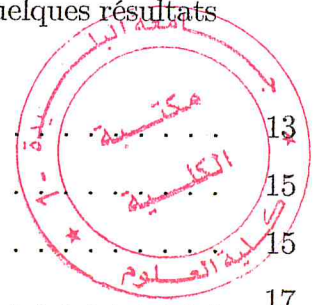
MCA

Année universitaire : 2016/2017

MA-510-55-1

Table des matières

0.1	Résumé	3
0.2	Abstract	3
0.3	Introduction Générale	5
1	Quelques éléments de la théorie des valeurs extrêmes	7
1.1	Introduction	7
1.2	Lois des valeurs extremes	8
1.3	Caractérisation des domaines d'attractions	11
1.3.1	Résultat sur les fonctions à variations régulières	11
1.3.2	Representation de Karamata	12
1.3.3	Définition d'une fonction à variation régulière et quelques résultats 12	
1.3.4	Domaine d'attraction de Fréchet	13
1.3.5	Domaine d'attraction de Gumbel	15
1.4	Estimation des quantiles extremes :	15
1.4.1	Quelques résultats sur les statistiques d'ordres	17
1.4.2	Approches des quantiles extremes par les lois de valeurs extremes :	19
1.4.3	Approches des valeurs extremes par la méthodes des excès	21
1.4.4	Approche des quantiles extremes par l'approche semi-paramétrique	24
1.5	Estimation du paramètre de l'indice de queue de la loi des valeurs extrêmes	25
1.5.1	Estimateur de Hill	27



1.5.2	Estimateur de Pickands	29
1.5.3	Estimateur de Zipf	30
2	Quantiles extrêmes conditionnelles	31
2.1	Introduction	31
2.2	Estimation paramétrique des quantiles conditionnels	32
2.3	Estimation non-paramétrique des quantiles conditionnels	33
2.3.1	La méthode d'estimation indirecte	33
2.4	Estimation des quantiles extrêmes conditionnels en design fixe	42
2.4.1	Contexte d'étude et présentation	43
2.4.2	Etude théorique des estimateurs	46
2.4.3	Exemple et discussion	51
2.5	Applications et discussion	53
2.5.1	Application du résultat de Gardes et Girard à nos estimateurs de quantiles extrêmes	53
3	Estimation non-paramétrique de mesures de risque extrêmes pour des lois conditionnelles à queues lourdes	57
3.1	Introduction	57
3.2	Mesures de risque	58
3.2.1	Mesure de risque cohérente	59
3.3	Le regression Conditional Tail Moment :définition et estimation	61
3.3.1	Pertes extrêmes et regression	64
3.4	Construction de l'estimateur de RCTE ET RVaR	65
3.5	Etude du comportement des deux estimateurs	66
3.6	Démonstration	71
3.6.1	Résultats préliminaires et leurs preuves	71
3.7	Validation sur simulation	77

0.1 Résumé

Cet mémoire appartient au contexte de la statistique des valeurs extrêmes.

La théorie des valeurs extrêmes a pour objectif l'étude du comportement asymptotique des grandes observations d'un échantillon de variables aléatoires indépendantes et identiquement distribuées. La théorie statistique "classique" permet de faire de l'inférence sur les valeurs centrales d'un échantillon mais ne donne que très peu d'information sur la queue de distribution. Le théorème fondamental de la théorie des valeurs extrêmes (connu sous le nom de théorème de Fisher-Tippett) donne quand à lui les lois limites possibles du maximum de l'échantillon et permet ainsi d'avoir une certaine connaissance sur la forme de la queue de distribution.

Le problème le plus fréquemment étudié en théorie des valeurs extrêmes est celui de l'estimation de quantiles extrêmes c'est à dire de quantiles plus grands que la valeur maximale observée. Pour estimer de tels quantiles, l'inversion de la fonction de répartition empirique est évidemment d'aucune utilité et c'est donc la théorie des valeurs extrêmes qui nous permet de répondre à ce problème. Alors que de nombreux travaux portent sur l'estimation des quantiles extrêmes d'une variable aléatoire réelle, très peu s'intéressent au cas où la variable d'intérêt Y est mesurée conjointement avec une covariable X . Le quantile extrême à estimer dépend alors de la covariable X et est appelé quantile extrême conditionnel.

0.2 Abstract

This memoir belongs to the context of statistical extremes values

The theory of extremes values has for objectif the studying of the asymptotic performance of big observations of a sample random variables independents and identically

distributed, the classical statistical theory allows to do the inference of central values but it gives us few informations of tail distribution. The fundamental theorem of the theory of extremes values give us the possible limits distribution of sample's maximum and allow us to have certain knowledge of tail's distribution form

The problem the most frequently studied in the theory of extremes values is the estimation of extremes values that is to say quantiles bigger than the maximal value observed. To estimate this quantiles, the inversion of the density is obviously useless so that the theory of extremes values allows us to reply this kinds of problems.

Many works take the estimation of extreme quantiles of a random variable, few of them are interested in case when the variable of interest Y measured jointly to a covariable X . So the estimated quantile depends to the covariable X called conditional quantile

lourde et qu'elle dépend d'une covariable. Les estimateurs proposés combinent des méthodes d'estimation non-paramétrique à noyau avec des méthodes issues de la statistique des valeurs extrêmes

Dès lors, la question que l'on pourrait se poser est de savoir ce que peuvent les statistiques face aux événements extrêmes? Autrement dit, peut-on réellement prévoir ou quantifier le risque des événements extrêmes?

Chapitre 1

Quelques éléments de la théorie des valeurs extrêmes

1.1 Introduction

La théorie des valeurs extrêmes a pour but d'étudier et de caractériser le comportement des valeurs extrêmes d'un échantillon de variables aléatoires. On souhaite estimer des petites probabilités ou des quantités dont la probabilité d'observation est très faible, c'est-à-dire proche de zéro. Ces quantités sont appelées quantiles, on parle de quantile extrême lorsque l'ordre du quantile (probabilité d'observation) converge vers zéro quand la taille de l'échantillon tend vers l'infini.

Plus précisément, on considère n variables aléatoires réelles $\{X_i, i = 1, 2, \dots, n\}$ indépendantes et identiquement distribuées de fonction de répartition F non nécessairement continue. A partir des observations de ces variables aléatoires, on souhaite estimer le quantile extrême d'ordre $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$ défini par :

$$q_{\alpha_n} \stackrel{\text{def}}{=} \bar{F}^{\leftarrow}(\alpha_n) = \inf \{x : \bar{F}(x) \leq \alpha_n\} \quad (1.1)$$

où \bar{F}^{\leftarrow} est l'inverse généralisé de $\bar{F} = 1 - F$. En particulier, pour n tendant vers ∞ , on

a :

$$\begin{aligned} P(\max(X_1, X_2, \dots, X_n) < q_{\alpha_n}) &= P(X_i < q_{\alpha_n}, i = 1, \dots, n) \\ &= (1 - \alpha_n)^n \\ &= \exp(n \log(1 - \alpha_n)^n) \\ &= \exp(-n\alpha_n(1 + o(1))) \end{aligned}$$

quand α_n tend vers zéro ce qui implique que

$$P(\max(X_1, X_2, \dots, X_n) < q_{\alpha_n}) \rightarrow 1$$

1.2 Lois des valeurs extremes

le principal résultat de la théorie des valeurs extremes repose sur le Théorème de Fisher et Tippett (1928) dont la première preuve rigoureuse est due à Gnedenko (1943) celui-ci fait appel à la notion de statistique d'ordre associée aux variables aléatoires.

Définition 1.1 *Notons par $X_{1,n}, \leq \dots \leq X_{n,n}$ le réarrangement croissant de l'échantillon $\{X_i, i = 1, 2, \dots, n\}$ pour tout $i = 1, 2, \dots, n$ la variable aléatoire $X_{i,n}$ s'appelle la i ème statistique d'ordre de l'échantillon.*

Celui-ci montre qu'il existe des suites (a_n) et $(b_n) \in R$ telle que la suites des variables aléatoires $(a_n^{-1}(X_{n,n} - b_n))$ converge vers une limite non dégénérée , ce résultat est la base de théorie des valeurs extrêmes. C'est l'équivalent du théorème central limite

Théorème 1.2 (Fisher Tippett (1928) Gnedenko (1943)) *Sous certaines conditions de régularité sur la fonction de répartition de F ,il existe un paramètre réel γ et deux*

suites (a_n) et $(b_n) \in \mathbb{R}$ tels que pour tout $x \in \mathbb{R}$

$$\lim P \left[\frac{X_{n,n} - b_n}{(a_n)} \leq x \right] = \mathcal{H}_\gamma(x)$$

avec

loi de Fréchet :

$$\Gamma_\gamma(x) = \left\{ \begin{array}{ll} 0 & \text{si } x < 0 \\ \exp(-(x)^{-\frac{1}{\gamma}}) & \text{si } x > 0 \end{array} \right\} \quad \text{et } \gamma > 0$$

loi de Weibull

$$\Gamma_\gamma(x) = \left\{ \begin{array}{ll} \exp(-(-x)^{-\frac{1}{\gamma}}) & \text{si } x < 0 \\ 1 & \text{si } x > 0 \end{array} \right\} \quad \text{et } \gamma < 0$$

loi de Gumbel

$$\Gamma_\gamma(x) = \exp[-\exp(-x)] \quad \text{si } x \in \mathbb{R} \quad \text{et } \gamma = 0$$

et $\Gamma_\gamma(x)$ et la fonction de répartition de la loi des valeurs extrêmes

Ce théorème est vrai pour la majorité des lois usuelles, si l'on fait un parallèle avec le théorème central limite, On considère la suite a_n comme un écart type de X et la suite b_n comme l'espérance. Les deux suites s'interprètent comme des paramètres d'échelle (resp. un paramètre de position ou de centrage), de plus les deux suites ne sont pas uniques .

La figure ci-dessus représente le cas de la loi normale centrée réduite la convergence de la suite de variables aléatoires $(a_n^{-1}(X_{n,n} - b_n))$ vers une limite non dégénérée Γ_0 , Dans cet exemple on va utiliser les suites de renormalisation théoriques associées à la loi normale standard (Embrechts et al 1997)

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

Grace aux travaux de Von Mises (1936) et de Jenkinson (1955) on a une forme unifiée de la fonction de répartition de la loi EVD à un facteur d'échelle de position près .

Définition 1.3 la représentation de Jenkinson_ Von Mises de la loi EVD qu'on appelle *loi des valeurs extremes generalisée* notée (**GEVD**) ou (**GEV**) a pour fonction de répartition

$$\Lambda_\gamma(x) = \left\{ \begin{array}{l} \exp[-1 + \gamma x]^{\frac{-1}{\gamma}} \text{ si } \gamma > 0 \\ \exp[-\exp(-x)] \text{ si } \gamma = 0 \end{array} \right\}$$

pour tout x tel que $1 + \gamma x > 0$

Définition 1.4 le paramètre γ du 1^{er} théorème et un paramètre de forme que l'on appelle *indice de valeurs extremes* ou *indice de queue*

si F vérifie le 1^{er} théorème on dit alors que F appartient au domaine d'attraction de Γ ou de Λ et selon le signe de γ on distingue trois domaines d'attraction dont quelques densités ont été représentées pour γ fixé tel que :

Si $\gamma > 0$. On dit que F appartient au domaine d'attraction de Fréchet et on notera $F \in D(\mathbf{Fréchet})$ ce domaine d'attraction est celui des distributions à queue lourdes ,i.equi on une fonction de survie à croissance polynomiale .comme exemple de lois appartenant à ce domaine d'attraction on a la loi de Fréchet, Pareto, khi-deux, Student, Cauchy etc .

Si $\gamma < 0$. On dit que F appartient au domaine d'attraction de Weibull et on notera $F \in D(\mathbf{Weibull})$. Ce domaine d'attraction est celui des fonctions de survie dont le support est borné supérieurement. Pour ce domaine on trouve les lois uniformes, Beta, Weibull

Si $\gamma = 0$. On dit que F appartient au domaine d'attraction de Gumbel et on notera $F \in D(\mathbf{Gumbel})$, ce domaine est celui des distributions à queues légères, i.e qui ont une fonction de survie à décroissance exponentielle. Dans ce domaine on trouve par exemple la loi normale, exponentielle, Log-normale ect

1.3 Caractérisation des domaines d'attractions

Dans cette partie, on se contentera de donner des conditions sur la fonction de répartition F pour qu'elle appartienne à l'un des domaines d'attractions définis précédemment. Ces conditions faisant appel à la notion de fonctions à variations régulières. On rappelle au préalable la définition de telles fonctions et on en donne quelques propriétés.

1.3.1 Résultat sur les fonctions à variations régulières

Avant qu'on commence avec les fonctions à variations régulières, il est nécessaire de donner une définition des fonctions à variations lentes.

Définition 1.5 On dit qu'une fonction l est à variation lente à l'infini si $l(x) > 0$ pour x assez grand et si pour tout $\lambda > 0$ on

$$\lim_{x \rightarrow \infty} \frac{l(\lambda x)}{l(x)} = 1.$$

Par exemple, la fonction constante, les fonctions qui tendent vers une constante $\log(x)$ sont des fonctions à variations lentes à l'infini. Les fonctions à variations lentes jouent un rôle prépondérant dans l'étude des lois de valeurs extrêmes

1.3.2 Représentation de Karamata

Théorème 1.6 *Resnick(1987)* Toute fonction à variations lentes l s'écrit sous la forme : ∞

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

ou c et ε sont deux fonctions mesurables

telles que

$$\lim_{x \rightarrow \infty} c(x) = c_0 \in]0; [\quad \lim_{x \rightarrow \infty} \varepsilon(x) = 0$$

si la fonction c est constante, alors on dit que l est normalisée, l'équation (1,2) implique que si l est normalisée alors elle est dérivable de dérivée l' avec pour tout $x > 0$ $l'(x) = \frac{\varepsilon(x)l(x)}{x}$, en particulier on a

$$\lim_{x \rightarrow \infty} x \frac{l'(x)}{l(x)} = 0$$

1.3.3 Définition d'une fonction à variation régulière et quelques résultats

Définition 1.7 On dit qu'une fonction G est à variations régulières d'indice $\rho \in \mathbb{R}$ à l'infini si $G > 0$ et pour tout $\lambda > 0$

$$\lim_{x \rightarrow \infty} \frac{G(\lambda x)}{G(x)} = \lambda^\rho$$

Dans le cas où $\rho = 0$ alors G est à variation lente, on peut écrire la fonction G sous la forme $x^\rho l(x)$ où l est une fonction à variations lentes à l'infini. Si G est une fonction à variations régulières d'indice ρ alors pour tout $0 < a < b$

$$\lim_{x \rightarrow \infty} \sup_{\lambda \in [a, b]} \left| \frac{G(\lambda x)}{G(x)} - \lambda^\rho \right| = 0$$

Lemme 1.8 (*inversion d'une fonction à variation régulière*)

Case 1.9 *Si G est à variation régulière d'indice $\rho > 0$, alors G^{\leftarrow} est à variation régulière d'indice $\frac{1}{\rho}$*

Si G est à variation régulière d'indice $\rho < 0$, alors G^{\leftarrow} est à variation régulière d'indice $\frac{-1}{\rho}$

Lemme 1.10 *soit l une fonction à variation lente et soient u_n et v_n deux suites positives telles que $u_n \rightarrow \infty$, $v_n \rightarrow \infty$ si $\frac{u_n}{v_n} = 1$ alors $l(u_n) \sim l(v_n)$*

Ce lemme qui est une conséquence du 1^{er} lemme, montre que les fonctions à variations lentes conservent les équivalents. Par conséquent. Ce résultat implique que si G est à variations régulières d'indice ρ et si $u_n \sim v_n$ alors $G(u_n) \sim G(v_n)$ et on dit que les fonctions à variations régulières conservent les équivalents

1.3.4 Domaine d'attraction de Fréchet

Avant de donner les expressions de la fonction de répartition de chaque domaine, il convient d'introduire une définition utile.

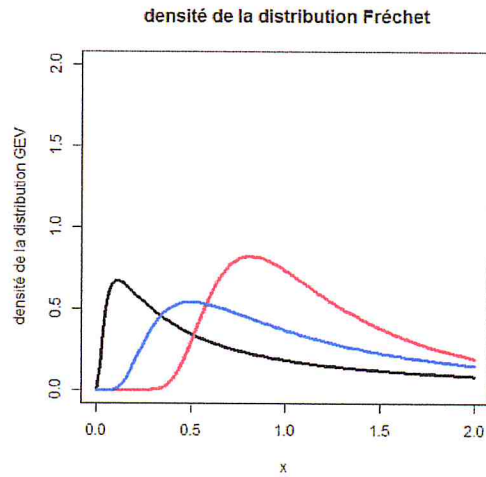
Définition 1.11 *On appelle point terminal de la fonction F le réel x_F défini par*

$$x_F = \sup \{x : F(x) < 1\}.$$

Théorème 1.12 *$F \in \mathcal{D}(\text{Fréchet})$ avec un indice de valeur extrême $\gamma > 0$ ssi $x_F = \infty$ et la fonction \bar{F} est à variations régulières d'indice $\frac{-1}{\gamma}$. Dans ce cas un choix possible pour les suites a_n et b_n est*

$$a_n = q_{\frac{1}{n}}, \quad b_n = 0$$

De ce théorème on en déduit que $F \in \mathcal{D}(\text{Fréchet})$ ssi le point terminal $x_F = \infty$ et $\bar{F}(x) = x^{\frac{-1}{\gamma}} l(x)$ où l est à variations régulières à l'infini et $\gamma > 0$



Domaine d'attraction de Weibull

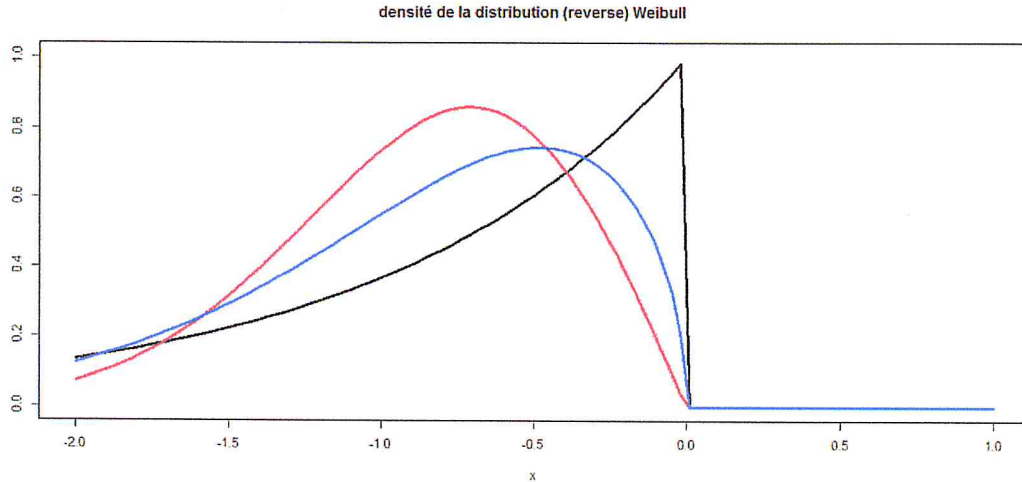
Théorème 1.13 $F \in \mathcal{D}(\text{Weibull})$ avec un indice de valeur extrême $\gamma < 0$ ssi $x_F < \infty$ et la fonction $(1 - F)$ est à variation régulière d'indice $\frac{1}{\gamma}$ avec :

$$\tilde{F}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ F(x_F - x^{-1}) & \text{si } x > 0 \end{cases}$$

Dans ce cas un choix possible pour a_n et b_n :

$$a_n = x_F - q_{\frac{1}{n}}; b_n = x_F$$

De ce théorème on déduit que $F \in \mathcal{D}(\text{Weibull})$ ssi le point terminal x_F est fini et $\bar{F}(x) = (x_F - x)^{\frac{-1}{\gamma}} l[(x_F - x)^{-1}]$ avec l fonction à variations lentes à l'infini et $\gamma < 0$



1.3.5 Domaine d'attraction de Gumbel

Théorème 1.14 $F \in \mathcal{D}(\text{Gumbel})$ ssi :

$$\bar{F}(x) = c(x) \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\}$$

où $\lim_{x \rightarrow x_F} c(x) = c$. Dans ce cas un choix possible pour les suites a_n et b_n :

$$a_n = q_{\frac{1}{n}}, b_n = \frac{1}{\bar{F}(a_n)} \int_{a_n}^{x_F} \bar{F}(y) dy = q_{\frac{1}{(ne)}} - q_{\frac{1}{n}}$$

1.4 Estimation des quantiles extremes :

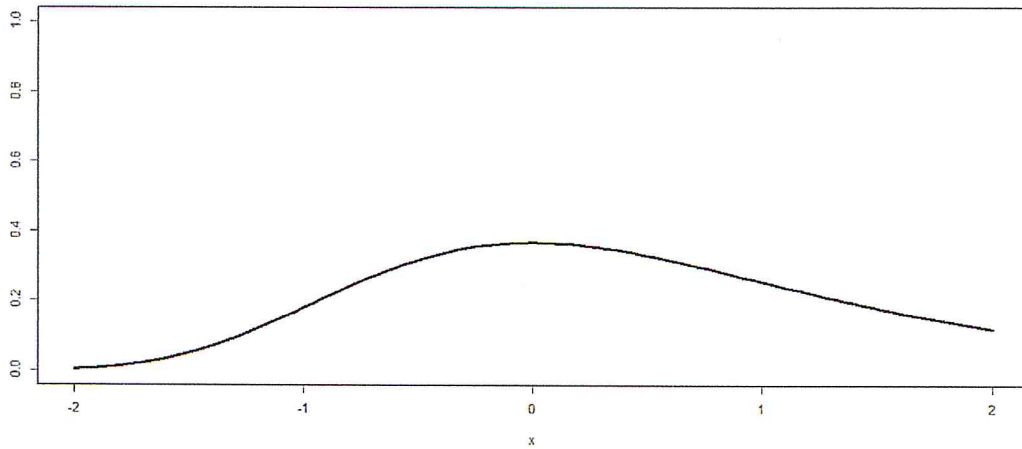
Dans tout ce qui suit. On fait l'hypothèse que F appartient à l'un des domaines d'attractions définis précédemment

Proposition 1.15 Si $\alpha_n \rightarrow 0$ et $n\alpha_n \rightarrow c$ quand n tend vers ∞ alors

$$P(X_{n,n} < q_{\alpha_n}) = e^{-c}$$

D'après le Lemme deux situations peuvent être distinguées

densité de la distribution Gumbel



si $c = \infty$ alors, $P(X_{n,n} < q_{\alpha_n}) = 0$. Dans un tel contexte, un estimateur naturel de q_{α_n} est le quantile empirique qui n'est rien d'autre que la $[n\alpha_n]$ ième plus grande observation de l'échantillon $\{X_i; i = 1, 2, \dots, n\}$ c'est à dire la statistique d'ordre $X_{n-[n\alpha_n]+1,n}$

si $c = 0$ alors, $P(X_{n,n} < q_{\alpha_n}) = 1$ par conséquent on ne peut pas estimer le quantile empiriquement, Pour résoudre ce problème on a répertorié trois catégories principales de méthodes

La théorie des valeurs extrêmes présentée par Guida et Longo (1988) et dont les premiers éléments bibliographiques remontent à Fisher et Tippet (1928) et Gnedenko (1943) consiste à diviser l'échantillon à m_0 sous groupes disjoints de taille $n_0 = \frac{n}{m_0}$ desquels on détermine les maxima. La loi de ces maxima est alors approchée, pour n_0 assez grand, par une loi des valeurs extrêmes. En utilisant la relation $P(\max(X_1, \dots, X_n) < q_{\alpha_n}) = F^n(q_{\alpha_n})$ on peut alors estimer le quantile extrême q_{α_n}

La méthode des excès initialement présentée par Pickands (1975). Elle préconise de ne retenir que les observations dépassant un seuil fixé u . La loi des k_n observations ainsi retenues que l'on note par $\{Y_i, i = 1, \dots, k_n\}$ peut-être approché, si u est assez grand par une loi de Pareto généralisée (GPD). Pour estimer le quantile extrême de q_{α_n} . Il suffit d'utiliser le résultat de (Balkema et de Haan (1974), Pickands (1975)) qui établit

l'équivalence entre la convergence en loi du maximum vers une loi des valeurs extrêmes et la convergence en loi d'un excès vers une GPD

Les méthodes semi-paramétriques où l'on suppose que pour tout $\gamma > 0$ on a $P(X > x) \approx x^{-\frac{1}{\gamma}}$ lorsque x tend vers ∞ , c'est à dire que la fonction de survie $\bar{F}(x)$ décroît en $x^{-\frac{1}{\gamma}}$. Cette hypothèse permet de construire des estimateurs nonparamétriques du paramètre γ dont le plus célèbre est l'estimateur de Hill (1975). Partant de ce résultat Weisman(1978) proposera trois ans plus-tard un estimateur de quantile extrême q_{α_n} . En effet comme nous venons de voir dans le lemme, supposer que $P(X > x) \approx x^{-\frac{1}{\gamma}}$ revient à supposer que le quantile q_{α_n} décroît en $\alpha_n^{-\gamma}$.

1.4.1 Quelques résultats sur les statistiques d'ordres

Lemme 1.16 (*transformation des quantiles*), soit $\{X_i, i = 1, \dots, n\}$ des variables aléatoires indépendantes de fonction de répartition F et soient $\{V_1, \dots, V_n\}$ des statistiques d'ordres associées aux variables aléatoires indépendantes de loi uniforme $\{V_i, i = 1, \dots, n\}$ alors

$$F^{-1}(V_1) \stackrel{L}{=} X_1$$

$$\{F^{-1}(V_1, \dots, V_n)\} \stackrel{L}{=} \{X_1, \dots, X_n\}$$

pour tout $n \in \mathbb{N}$

$F(X_1)$ suit la loi uniforme standard ssi F est continue

Ce lemme nous assure que l'étude des quantiles empirique associés à une loi quelconque peut se déduire de l'étude des statistiques d'ordre associées à la loi uniforme et exponentielle. Soit $\{E_i, 1, \dots, n\}$ une suite de variables aléatoires de loi exponentielle standard. On note $T_n = \sum_{i=1}^n E_i$. Soit $\{V_i, i = 1, \dots, n\}$ une suite de variables aléatoires de loi uniforme standard

Lemme 1.17 (représentation de Rényi (1953)). La suite de variables aléatoire $\{V_{1,n}, \dots, V_{n,n}\}$ a même loi que la suite $\left\{\frac{T_1}{T_{n+1}}, \dots, \frac{T_n}{T_{n+1}}\right\}$ i.e $V_{j,n} \stackrel{L}{=} \frac{T_j}{T_{n+1}} \quad j = 1, \dots, n$

Ce lemme nous permet d'établir une convergence faible d'une suite de statistiques d'ordres de loi exponentielle

Proposition 1.18 Soit $(k_n)_{n \geq 1}$ une suite d'entiers tel que $1 \leq k_n \leq n$ $k_n \rightarrow \infty$ et $\frac{k_n}{n} \rightarrow 0$ alors

$$\sqrt{k_n}(E_{n-k_n+1,n} - \log(\frac{n}{k_n})) \stackrel{L}{=} N(0, 1)$$

Définition 1.19 pour tout $t > 0$, la fonction queue est donnée par :

$$U(t) \stackrel{def}{=} \inf\{x : F(x) \geq 1 - \frac{1}{t}\} = q_{\frac{1}{t}}$$

L'utilité de cette définition est essentiellement d'ordre pratique . En effet, il peut être plus commode de travailler non pas sur la fonction de survie \bar{F} elle même , mais plutôt sur la fonction queue U .

Théorème 1.20 (Haan et Ferreira (2006)). Soit $(k_n)_{n \geq 1}$ une suite d'entiers tel que $1 \leq k_n \leq n$ $k_n \rightarrow \infty$ et $\frac{k_n}{n} \rightarrow 0$. Si $\lim_{x \rightarrow x_{\bar{F}}} \frac{\bar{F}(x)F''(x)}{(F'(x))^2} = -\gamma - 1$, alors $\sqrt{k_n}(\frac{X_{n-k_n+1,n} - U(\frac{n}{k_n})}{(n/k_n)U'(n/k_n)}) \stackrel{L}{=} N(0, 1)$

Ce théorème montre que l'estimateur de quantile extrême $X_{n-k_n+1,n}$ avec $k_n \stackrel{def}{=} \lfloor n\alpha \rfloor$ est asymptotiquement gaussien . En remarquant que la proposition est un cas particulier de ce théorème en posant $U = \log$

Proposition 1.21 Dans le cas où $\gamma > 0$ si la condition $\lim_{x \rightarrow \infty} \frac{xU'(x)}{U(x)} = \gamma$ est vérifiée alors on :

$$\sqrt{k_n}(\frac{X_{n-k_n+1,n}}{q_{k_n/n}} - 1) \stackrel{L}{=} N(0, \gamma^2)$$

La proposition est un cas particulier du théorème . Il parait que pour les lois dont la fonction de répartition $F \in \mathfrak{D}(\mathfrak{Frechet})$, l'estimateur de quantile extreme $X_{n-k_{n+1},n}$ est asymptotiquement normal avec une variance asymptotiquement proportionnelle à $\frac{\gamma^2}{k_n}$

1.4.2 Approches des quantiles extremes par les lois de valeurs extremes :

Pour estimer le quantile extreme Guida et Longo (1988) utilisent l'approximation

$$P(X_{n,n} \leq a_n x + b_n) = F^n(a_n x + b_n) \approx \Gamma_\gamma(x)$$

En effet, d'après le théorème de Fisher et Tippet on peut écrire

$$\lim_{x \rightarrow \infty} n \log F(a_n x + b_n) = \log \Gamma_\gamma(x)$$

Soit encore

$$\lim_{x \rightarrow \infty} n[\log(1 - \bar{F}(a_n x + b_n))] = \log \Gamma_\gamma(x)$$

Comme $n \rightarrow \infty$ on peut montrer que $a_n x + b_n \rightarrow x_F$ et par conséquent $\bar{F}(a_n x + b_n) \rightarrow 0$ avec un développement limité du 1^{er} ordre $\log(1 + u) = \bar{F}(a_n x + b_n)$ on obtient

$$\bar{F}(a_n x + b_n) \approx \frac{-1}{n} \log \Gamma_\gamma(x)$$

Pour tout γ on peut alors approcher le quantile q_{α_n} par :

$$q_{\alpha_n} \approx a_n x_{\alpha_n} + b_n \text{ où } x_{\alpha_n}$$

avec

$$-\log(\Gamma_\gamma(x_{\alpha_n})) = n\alpha_n$$

On a alors un estimateur de quantile extreme de type

$$\hat{q}_{\alpha_n} = \hat{a}_n x_{\alpha_n} + \hat{b}_n$$

$$\begin{aligned} \hat{a}_n (n\alpha_n)^{-\gamma} + \hat{b}_n & \quad \text{si } F \in D(\text{Fréchet}) \\ -\hat{a}_n (n\alpha_n)^{-\gamma} + \hat{b}_n & \quad \text{si } F \in D(\text{Gumbel}) \\ -\hat{a}_n \log(n\alpha_n)^{-\gamma} + \hat{b}_n & \quad \text{si } F \in D(\text{Weibull}) \end{aligned} \quad (1.2)$$

où (\hat{a}_n, \hat{b}_n) et $\hat{\gamma}$ sont respectivement des estimateurs de suites (a_n, b_n) et de l'indice de queue γ

Dans le cas où $\gamma = 0$, les auteurs proposent d'utiliser l'approche basée sur la loi GEV dont le résultat s'énonce ainsi

Théorème 1.22 (Weinstein(1973)) *soit $F \in D(\text{Gumbel})$, il existe deux suites $(a_n)_{n \geq 1} > 0$ et $(b_n)_{n \geq 1} \in \mathbb{R}$ telles que pour tout $x \in \mathbb{R}$ et $v > 0$ on a*

$$\lim_{x \rightarrow x_{\text{inf}}} n\bar{F}[(b_n^v + c_n x)^{\frac{1}{v}}] = \exp(-x)$$

où $c_n = a_n v b_n^{v-1}$

Dans une telle situation on approche le quantile par

$$q_{\alpha_n} \sim (b_n^v + c_n x_{\alpha_n})^{\frac{1}{v}} \quad \text{où } x_{\alpha_n} \text{ vérifie } \exp(-x_{\alpha_n}) = -n\alpha_n$$

est un estimateur du quantile extrême obtenu en remplaçant les suites b_n et c_n respectivement par leur estimateur \hat{b}_n et \hat{c}_n i.e :

$$\hat{q}_{\alpha_n} = (\hat{b}_n^v - \hat{c}_n \log(n\alpha_n))^{\frac{1}{v}}$$

L'avantage d'utiliser le résultat du théorème de Weinstein, provient du fait qu'il existe des valeurs de v pour lesquelles la convergence de $\lim_{x \rightarrow x_{\text{inf}}} n\bar{F}[(b_n^v + c_n x)^{\frac{1}{v}}] = \exp(-x)$ est plus rapide que dans le cas où $v = 1$. Dans ce cas, sur simulation, l'approximation du quantile q_{α_n} est de meilleure qualité que l'approximation basée sur l'approche EVD avec $v = 1$. Les auteurs fournissent la valeur optimale du paramètre v

les paramètres γ, a_n, b_n, c_n de ces lois peuvent être estimés par la méthode du maximum de vraisemblance (Prescott et Walden, 1980, 1983) ou par la méthode des moments pondérés (Hosking et al (1985)). Dans le cas de l'approche EVD, Smith (1985) fait une étude détaillée du comportement asymptotique des estimateurs des paramètres γ, a_n, b_n , obtenues par la méthode du maximum de vraisemblance. Toutes fois, il est conseillé d'utiliser les estimateurs des moments pondérés car ceux-ci sont non seulement explicites et facile à calculer mais aussi parce qu'ils donnent de meilleurs résultats que les estimateurs du maximum de vraisemblance quand on a des échantillons de petite ou de moyenne taille. La principale difficulté de l'estimation des paramètres γ, a_n, b_n est due au fait qu'il faut un échantillon de maxima, lequel est parfois difficile à extraire des données initiales

1.4.3 Approches des valeurs extremes par la méthodes des excès

Définition 1.23 (*Loi de Pareto Généralisée (GPD)*) : La loi de Pareto généralisée est la loi dont la fonction de répartition est donnée par :

$$G_{\gamma, \beta}(y) = \begin{cases} 1 - (1 + \frac{\gamma y}{\beta})^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0, \beta > 0 \\ 1 - \exp(\frac{-y}{\beta}) & \text{si } \gamma = 0, \beta > 0 \end{cases}$$

avec $y \in R_+$ si $\gamma > 0$ ou $y \in [0, \frac{-\beta}{\gamma}]$ si $\gamma < 0$

Dans l'expression précédente β représente le paramètre d'échelle et γ représente le paramètre de forme, il s'agit du même paramètre de forme que l'on appelle indice de valeurs extremes

La loi GDP représente quelques particularités, voici une liste non exhaustive :

si $\beta = 1$ on l'appelle GDP standard

si $\gamma = 0$ elle correspond à une loi Exponentielle d'espérance β

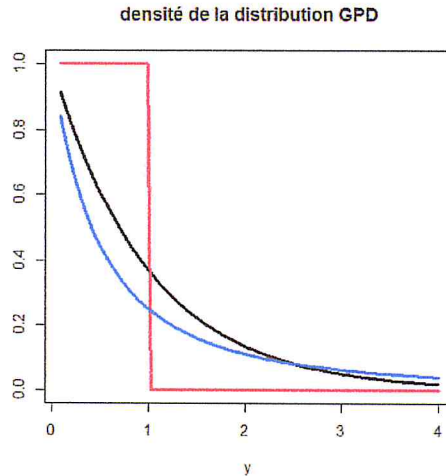


FIG. 1-1 – Graphe des densités associées à la loi de Pareto généralisée (noir $\gamma = 0$, rouge $\gamma = -1$, bleu $\gamma = 1$)

si $\gamma = -1$ elle correspond à une loi Uniforme dans l'intervalle $[0, \beta]$

si $\gamma > 0$ on a une loi de Pareto décentrée

Dans cette approche d'estimation des quantiles extrêmes, on ne retient que les observations dépassant un seuil fixé $u < x_F$. On définit alors l'excès Y de la variable X au dessus du seuil u par $X - u$ sachant que $X > u$. Si on note par F_u^+ la fonction de répartition d'un excès au dessus du seuil u on a pour tout $y > 0$:

$$\begin{aligned}
 1 - F_u^+(y) &= P(Y > y) \\
 &= P(X - u > y / X > u) \\
 &= \frac{P(X > y+u, X > u)}{P(X > u)} \\
 &= \frac{1 - F(y+u)}{1 - F(u)}
 \end{aligned}$$

Lorsque le seuil u est grand on peut approcher cette quantité par la fonction de survie d'une loi GPD . Afin d'approcher le quantile, il suffit alors d'utiliser le résultat de (Balkema et de Haan (1974)) et (Pickands (1975)) qui établit l'équivalence entre la convergence en loi du maximum vers une loi de valeurs extrêmes Γ_γ est la convergence en loi d'un excès vers une loi GPD

Théorème 1.24 (*Balkema et de Haan (1974)*) et (*Pickands (1975)*), Si F appartient au domaine d'attraction de $\Gamma_\gamma(x)$ alors :

$$\lim_{x \rightarrow x_F} \sup_{y \in]0, x_F - u[} |F_u(y) - G_{\gamma, \beta}(y)| = 0$$

d'après ce résultat, si pour une fonction de répartition F inconnue, l'échantillon des maximas normalisés convergent en loi vers une distribution non dégénérée, alors il s'en déduit que la distribution des excès au-dessus d'un seuil élevé converge vers une GPD lorsque le seuil tend vers la limite supérieure du support de F . Cette caractérisation est à la base des méthodes d'estimations Peaks Over Threshold (POT)

Comme $1 - F(u + y) = [1 - F(u)][1 - F_u(y)]$, si pour tout $y > 0$ on pose $q_{\alpha_n} = u + y$ alors

$$\begin{aligned} q_{\alpha_n} &= 1 - F(q_{\alpha_n}) = [1 - F(u)][1 - F_u(q_{\alpha_n} - u)] \\ &\simeq [1 - F(u)][1 - G_{\gamma, \beta}(q_{\alpha_n} - u)] \end{aligned}$$

Pour k_n excès au-dessus du seuil u , l'approximation $1 - F(u) \simeq k_n/n$ conduit à

$$\frac{k_n}{n} (1 - G_{\gamma, \beta}(q_{\alpha_n} - u)) \simeq q_n$$

si $\gamma \neq 0$ alors on approche le quantile par

$$q_{\alpha_n} = u + \frac{\beta}{\gamma} \left(\left(\frac{k_n}{n\alpha_n} \right)^\gamma - 1 \right)$$

On a alors un estimateur de type

$$\hat{q}_{\alpha_n} = \frac{\left(\frac{k_n}{n\alpha_n} \right)^{\hat{\gamma}} - 1}{\hat{\gamma}} \hat{\beta} + u \quad (1.3)$$

où $\hat{\gamma}, \hat{\beta}$ sont respectivement des estimateurs des paramètres de forme et d'échelle.

Les paramètres γ et β de la GPD peuvent être estimés par la méthode des moments . La méthode des moments pondérées (Hosking et Wallis,1987) ou la méthode du maximum de vraisemblance (Smith 1987,Davison et Smith 1990)

Cette méthode présente un avantage par rapport à la précédente en ce sens qu'il est plus facile d'avoir un échantillon d'excès que de maxima.Dans la pratique,on remplace u par $X_{n-k_n+1,n}$ c'est à dire la k_n plus grande observation de l'échantillon $\{X_i, i = 1, 2, \dots, n\}$

Deux variantes de cette méthode ont été présentées par Breiman et al.(1990) sous les applications Exponential Tail (ET) et Quadratic Tail (QT).

1.4.4 Approche des quantiles extrêmes par l'approche semi-paramétrique

On se restreint aux fonctions $F \in \mathcal{D}(\text{Fréchet})$ pour lesquelles on a la caractérisation suivante :

$$\bar{F} = x^{-\frac{1}{\gamma}} l(x)$$

avec l est une fonction à variations lentes à l'infini et $\gamma > 0$,Conformément au lemme de l'inverse d'une fonction à variation régulière cette caractérisation implique que :

$$\begin{aligned} q_{\alpha_n} & : = \bar{F}^{\leftarrow}(\alpha_n) = \alpha_n^{-\gamma} L\left(\frac{1}{\alpha_n}\right) \text{ avec } \alpha_n \leq 1/n \\ q_{\beta_n} & : = \bar{F}^{\leftarrow}(\beta_n) = \beta_n^{-\gamma} L\left(\frac{1}{\beta_n}\right) \text{ avec } \beta_n \leq 1/n \end{aligned}$$

où L est une fonction à variation lente à l'infini.En ce qui concerne les fonctions L et l , il apparaît important de signaler qu'il ne s'agit pas de la même fonction à variations lentes.

Vu la définition d'une fonction à variations lentes pour β_n suffisamment petit, on a

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

En remplaçant $\overline{F}^{\leftarrow}(\beta_n)$ et γ par des estimateurs, on obtient l'estimateur de Weissman(1978) défini par

$$\hat{q}_{\alpha_n}^W = X_{n-[n\beta_n]+1,n} \left(\frac{\beta_n}{\alpha_n}\right)^{\hat{\gamma}}$$

Pour les propriétés de l'estimateur de Weissman, On peut se référer à l'ouvrage de Embrechts et al (1997)

Comme autre estimateur de quantile extrême, on peut citer celui obtenu par l'approximation

$$\overline{F}^{\leftarrow}(\alpha_n) \simeq \frac{(\beta_n/\alpha_n)^{\gamma} - 1}{1 - 2^{-\gamma}} (\overline{F}^{\leftarrow}(\beta_n) - \overline{F}^{\leftarrow}(2\beta_n)) + \overline{F}^{\leftarrow}(\beta_n)$$

Cette approximation est valable quelque soit le domaine d'attraction de F . La normalité asymptotique de l'estimateur de quantile extrême qui en découle, i.e

$$\hat{q}_{\alpha_n}^{DH} = \frac{(\beta_n/\alpha_n)^{\hat{\gamma}} - 1}{1 - 2^{-\hat{\gamma}}} (X_{n-[n\beta_n]+1,n} - X_{n-[2n\beta_n]+1,n}) + X_{n-[n\beta_n]+1,n}$$

a été établi par Deckers et De Haan (1989), il apparait clairement que cet estimateur de quantile extrême peut se mettre sous la forme

$$\hat{a}_n = \frac{\hat{\gamma}}{1 - 2^{-\hat{\gamma}}} (X_{n-[n\beta_n]+1,n} - X_{n-[2n\beta_n]+1,n}) \text{ et } \hat{b}_n = X_{n-[n\beta_n]+1,n}$$

1.5 Estimation du paramètre de l'indice de queue de la loi des valeurs extrêmes

Dans la littérature de la théorie des valeurs extrêmes on trouve plusieurs techniques semi-paramétriques pour l'estimation de l'indice de queue. On peut citer l'estimateur de

Hill(1975) valable pour $\gamma > 0$. Il est considéré comme le plus simple estimateur de l'indice de queue. Pour pallier les limitations de l'estimateur de Hill, mais aussi pour l'étendre aux deux autres domaines d'attraction Dekkers et al (1989) ont proposé un estimateur des moments valable pour $\gamma \in \mathbb{R}$. Dans la littérature, certains auteurs l'appellent plutôt l'estimateur de Dekkers-Einmahl et Haan. Les auteurs ont établis la consistance forte, faible et la normalité asymptotique de leurs estimateurs. Soulignons que la méthode des estimateurs fut initialement utilisé par Hosking et Wallis(1987) qui proposèrent un estimateur des moments pondérées défini pour tout $\gamma < 1$. Cet estimateur est consistant si $\gamma < 1$ et asymptotiquement gaussien si $\gamma < \frac{1}{2}$. Toutefois, Pickands(1975) fût le premier à proposer un estimateur de l'indice de queue plus général que l'estimateur de Hill, i.e valable quelque soit γ Smith(1985) s'est quant à lui intéressé au comportement des estimateurs du maximum de vraisemblance dans le cas d'une loi GEV. Il a été démontré que l'estimateur du maximum de vraisemblance est consistant, asymptotiquement efficace et asymptotiquement normal pour tout $\gamma > -\frac{1}{2}$. Falk(1995) a proposé un estimateur du maximum de vraisemblance. Son estimateur, que l'on appelle aussi estimateur négatif de Hill est consistant si $\gamma < -\frac{1}{2}$ et asymptotiquement normal si $-1 < \gamma < -\frac{1}{2}$.

A côté de cette énumération on pourrait aussi ajouter les méthodes de corrections du biais basées sur des méthodes de regression (Beirlant et al(1995),(2005)), (Diebolt et al(2008) ou sur des méthodes de bootstrap(Gomes et Oliveira(2001)) et les estimateur à poids ou combinaison linéaire(Viharos(1993),(1995)), (Gomes et al(2005)) qui englobent d'un e part les estimateurs construits par des méthodes d'optimisation avec ou sans contrainte et d'autres part les estimateurs construits par des méthodes de regressions .

D'un point de vue théorique, toutes ses méthodes partagent les mêmes propriétés de consistance et de normalité asymptotique. Cependant, les simulations montrent qu'il y a une grande différence entre ces estimateurs. En général, il n'y a pas une meilleure méthode dans toutes les situations, les méthodes les plus utilisées sont celle de Hill, Pickands et celle des moments.

Tsourti et Panaretos(2001,2003) pensent que la performance d'une méthode dépend

de la distribution de la série étudiée. En d'autres termes, elle dépend de la vraie valeur de l'indice de queue. Ils recommandent l'utilisation de techniques pour déterminer le domaine d'attraction de la loi des valeurs extrême, et donc l'intervalle le plus probable pour l'indice des valeurs extrêmes. Les méthodes les plus utilisées à cette fin sont graphiques, le graphique loglog, la moyenne empirique des excès, le graphique des rapports du maximum et de la somme, le graphique du rapport de Hill, la statistique de Jackson

1.5.1 Estimateur de Hill

Cet estimateur a été introduit par Hill(1975) pour estimer d'une manière non-paramétrique le paramètre de queue des lois appartenant au $\mathcal{D}(\text{Fréchet})$. Pour construire son estimateur. Hill utilise la méthode du maximum de vraisemblance sur l'ensemble des k_n plus grandes observations d'un échantillon. Un grand nombre de travaux théoriques ont été consacrés à l'étude des propriétés de l'estimateur de Hill. Mason(1982) a démontré la consistance faible. Deheuvels, Haeusler et Mason ont établi la consistance forte dans Deheuvels et al(1988).

Définition 1.25 Soit $(k_n)_{n \geq 1}$ une suite d'entiers avec $1 < k_n \leq n$ l'estimateur de Hill est défini par

$$\hat{\gamma}_{k_n}^H = \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} \log X_{n-i+1,n} - \log X_{n-k_n+1,n}$$

Théorème 1.26 (propriétés de l'estimateur de Hill) soit $(k_n)_{n \geq 1}$ une suite d'entiers telle que $1 < k_n \leq n, k_n \rightarrow \infty$ et $k_n/n \rightarrow 0$ quand $n \rightarrow \infty$

alors $\hat{\gamma}_{k_n}^H$ converge en loi vers γ

si de plus $k_n/\log \log n \rightarrow \infty$ quand $n \rightarrow \infty$ alors $\hat{\gamma}_{k_n}^H$ converge presque sûrement vers γ

Pour établir la normalité asymptotique de $\hat{\gamma}_{k_n}^H$, on a besoin d'une hypothèse sur la fonction à variations lentes l . Il est en effet nécessaire d'imposer une condition qui spécifie la vitesse de convergence du rapport des fonctions à variations lentes vers l

Il existe une constante réelle $\rho < 0$ et une fonction $\varepsilon(x) \rightarrow 0$ quand $x \rightarrow \infty$ telles que pour tout $\lambda > 1$

$$\log \frac{l(\lambda x)}{l(x)} \sim \varepsilon(x) \frac{\lambda^\rho - 1}{\rho} \text{ quand } x \rightarrow \infty$$

Condition 1.27 Cette condition appelée "Condition du second ordre" est satisfaite pour la plupart des lois appartenant au $\mathcal{D}(\text{Fréchet})$. Plus la constante ρ est proche de 0, plus est difficile l'estimation de l'indice de queue γ

Remarque 1.28 La condition implique que $\forall \varepsilon > 0 \exists x_0$ telle que $\forall x \geq 0 \forall \lambda > 1$

$$\frac{(1 + \varepsilon)\lambda^{\varepsilon+\rho} - 1}{\rho} \leq \frac{1}{\varepsilon(x_n)} \log \frac{l(\lambda x)}{l(x)} \leq \frac{(1 - \varepsilon)\lambda^{\rho-\varepsilon} - 1}{\rho}$$

Théorème 1.29 (la normalité asymptotique de l'estimateur de Hill) Soit $(k_n)_{n \geq 1}$ une suite d'entiers tels que $1 \leq k_n \leq n$, $k_n \rightarrow \infty$ et $k_n/n \rightarrow 0$ quand $n \rightarrow \infty$ si la condition est satisfaite avec $\sqrt{k_n}\varepsilon(n/k_n) \rightarrow 0$ quand $n \rightarrow \infty$ alors

$$\sqrt{k_n}(\hat{\gamma}_{k_n}^H - \gamma) \xrightarrow{L} \mathcal{N}(0, \gamma^2)$$

En pratique le choix du paramètre k_n pose des problèmes. Si l'on trace le diagramme de Hill, on observe une extrême volatilité qui rend difficile l'utilisation de cet estimateur en pratique si l'on n'a aucune indication sur le choix de k_n . De plus, cet estimateur est biaisé ce biais est de l'ordre $\varepsilon(n/k_n)$ la condition $\sqrt{k_n}\varepsilon(n/k_n) \rightarrow 0$ impose au biais d'être négligeable devant l'écart type de l'estimateur qui est égal à $\sqrt{k_n}$. Une minimisation de l'erreur en moyenne quadratique peut-être utilisé comme critère. Cette méthode néanmoins inutilisable en pratique puisque l'erreur en moyenne quadratique reste inconnue !.

Le résultat sur la normalité asymptotique de l'estimation de Hill permet de donner un intervalle de confiance pour l'estimation. En pratique on se contentera de remplacer γ par sa valeur estimée. Par conséquent, si k_n est petit, on aura a fortiori, compte tenu des remarques faites précédemment, une estimation avec un intervalle de confiance large et à contrariori, si k_n est grand, on aura un intervalle de confiance plus étroit mais pas centré sur la vraie valeur .

1.5.2 Estimateur de Pickands

l'estimateur de Pickands est construit en utilisant trois statistiques d'ordres, cet estimateur a l'avantage d'être valable quelque soit le domaine d'attraction de la distribution et par conséquent, du domaine de définition de l'indice des valeurs extrêmes, Pickands(1975) démontre la consistance faible de son estimateur. La convergence forte ainsi que la normalité asymptotique ont été démontrées par Deekers et de Haan(1989)

Définition 1.30 *On suppose que $\{X_i : i = 1, \dots, n\}$ une suite de variables aléatoires indépendantes de loi F appartenant a l'un des domaines d'attractions, soit $(k_n)_{n \geq 1}$ une suite d'entiers avec $1 \leq k_n \leq n$, l'estimateur de Pickands est défini par*

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

Théorème 1.31 (propriétés de l'estimateur de Pickands) *Soit $(k_n)_{n \geq 1}$ une suite d'entiers tels que $1 \leq k_n \leq n, k_n \rightarrow \infty, \frac{k_n}{n} \rightarrow 0$ quand $n \rightarrow \infty$*

$\hat{\gamma}_{k_n}^P$ converge en loi vers γ

Si de plus $\frac{k_n}{\log \log n} \rightarrow \infty$ quand $n \rightarrow \infty$ alors $\hat{\gamma}_{k_n}^P$ converge presque surement vers γ

Sous des conditions additionnelles sur la suite k_n et la fonction de répartition F que l'on pourra consulter dans Dekkers et de Haan(1989)

$$\sqrt{k_n}(\hat{\gamma}_{k_n}^P - \gamma) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\gamma^2(2^{2\gamma+1} + 1)}{4(\log 2)^2(2^\gamma - 1)^2}\right)$$

Comme l'estimateur de Hill, cet estimateur est biaisé et le résultat sur sa normalité asymptotique permet de donner un intervalle de confiance pour l'estimation.

Compte tenu de la variance asymptotique de $\hat{\gamma}_{k_n}^P$ qui est assez importante comparativement à $\hat{\gamma}_{k_n}^H$ certains auteurs ont proposés des estimateurs à variance minimale construits à partir de combinaisons linéaires des logarithmes des accroissements des statistiques d'ordres. Par exemple Drees(1995) propose de faire la moyenne de plusieurs estimateurs de Pickands utilisant un nombre de plus grandes observations différentes dans le but d'obtenir un estimateur moins sensible aux choix des k_n

1.5.3 Estimateur de Zipf

Dans le but d'améliorer le biais asymptotique des estimateurs précédents, Kratz et Resnick (1996) et Schultze et Steinebach (1996) ont indépendamment proposés d'estimer l'indice de queue par la méthode des moindres carrés classique. Leur estimateur connu sous le nom de Zipf est asymptotiquement normal. Il est défini par

$$\hat{\gamma}_{k_n}^Z = \frac{\frac{1}{k_n} \sum_{j=1}^{k_n} \log \frac{k_n+1}{j} \log X_{n-j+1,n} - \frac{1}{k_n} \sum_{j=1}^{k_n} \log \frac{k_n+1}{j} \left(\frac{1}{k_n} \sum_{j=1}^{k_n} \log X_{n-j+1,n} \right)}{\frac{1}{k_n} \left(\sum_{j=1}^{k_n} \log \frac{k_n+1}{j} \right)^2 - \left(\frac{1}{k_n} \sum_{j=1}^{k_n} \log \frac{k_n+1}{j} \right)^2}$$

Chapitre 2

Quantiles extrêmes conditionnelles

2.1 Introduction

Dans cette partie, nous nous intéressons au cas où la variable aléatoire d'intérêt Y est mesurée conjointement avec une covariable X . Contrairement au problème d'estimation des quantiles de régression classique qui a été largement considéré, peu d'attention a été accordé aux quantiles conditionnels extrêmes.

Ce problème n'a été considéré que récemment , et ce malgré de nombreuses applications notamment en finance et en hydrologie. En effet, l'ajout d'une covariable permet de modéliser plus de phénomènes. Par exemple en hydrologie, la variable Y peut représenter le niveau de pluie journalier en millimètres tombée en un point géographique

caractérisé par sa position $x = (\textit{latitude}, \textit{longitude}, \textit{altitude})$. Un autre exemple d'application possible est l'astrophysique, où Y représente la quantité d'un certain paramètre physique et x une courbe hyperspectrale .

Nous trouvons dans la littérature deux types d'approches pour les estimations conditionnelles :

- l'approche paramétrique
- l'approche non-paramétrique

Lorsque l'on travaille en présence d'une covariable on distingue deux cas selon la

nature de cette dernière :

- Le modèle dit <à plan fixe> ou <design fixe> dont les données sont des couples $\{(x_i, Y_i), i = 1, \dots, n\}$ où les observations Y_i sont des variables aléatoires réelles non nécessairement indépendantes et les x_i sont des points d'observations non aléatoires
- Le modèle dit <à plan aléatoire> ou <design aléatoire> pour lequel les données sont des couples $\{(X_i, Y_i), i = 1, \dots, n\}$ de variables aléatoires non nécessairement indépendantes et identiquement distribuées de même loi. Dans ce cas, on peut définir la fonction de répartition conditionnelle $F(\cdot|x)$ par :

$$F(y|x) = \mathbb{P}(Y < y|X = x)$$

2.2 Estimation paramétrique des quantiles conditionnels

Elle est généralement utilisée quand on dispose d'un échantillon de petite taille. Afin d'estimer le quantile conditionnel, une façon de procéder consiste à supposer la fonction de répartition conditionnelle gaussienne. L'estimateur correspondant du quantile conditionnel $q(\alpha|x) = \hat{m}_n(x) + z_\alpha \hat{\sigma}_n(x)$

où $\hat{m}_n(x)$ (resp. $\hat{\sigma}_n(x)$) désigne l'estimateur de l'espérance conditionnelle (resp. l'écart type conditionnel) de Y sachant $X = x$ et z_α le quantile d'ordre α de la loi normale centrée réduite. Afin d'estimer $m(x)$ et $\sigma^2(x)$, Royston et Wright (1998) ont utilisé un modèle polynomial qu'ils ont associés à la méthode des moindres carrés. Cependant, dans la pratique, lorsque l'on dispose des données connues pour leurs valeurs aberrantes comme en biologie, il est parfois nécessaire de les transformer dans l'espoir d'obtenir des résidus normalement distribués. Hormis la restrictivité des hypothèses paramétriques, il semblerait même que l'existence d'une telle transformation ne soit pas toujours possible.

Pour pallier ces problèmes d'optimisations ou d'hypothèses une nouvelle approche dite non-paramétrique a été mise en oeuvre.

2.3 Estimation non-paramétrique des quantiles conditionnels

Dans cette approche on distingue deux méthodes d'estimation

- La première consiste à estimer au préalable la fonction de répartition conditionnelle puis à l'inverser pour en obtenir un estimateur du quantiles conditionnel.
- La seconde consiste quant à elle en une estimation directe basée sur le principe des moindres carrés

De nombreuses études on été menées dans le cas d'un processus Markovien (Roussas(1969)), lorsque les données sont indépendantes et identiquement distribuées(Stute(1986)) ou α -mélangeantes (Berlinet et al(1998))

2.3.1 La méthode d'estimation indirecte

On a deux techniques d'estimation indirecte de la fonction de répartition conditionnelle

- Les estimateurs à noyaux qui peuvent être construits suivant :
 - la méthode d'estimation du simple noyau
 - la méthode d'estimation par noyau produit
 - la méthode de médianogramme ou de la fenêtre mobile
- Les estimateurs au sens du plus proche voisin

Estimateur à noyau

L'expression de l'estimation du quantile conditionnel construit en inversant la fonction de répartition conditionnelle est donnée par

$$\hat{q}_n(\alpha|x) = \inf\{y|\hat{F}_n(y|x) \geq \alpha\} \quad (2.1)$$

(a)-L'estimateur par simple noyau de la fonction de répartition conditionnelle

On estime la fonction de répartition conditionnelle par

$$\hat{F}_n(y|x) = \sum_{i=1}^n w_{ni}(x) \mathbb{I}_{Y_i \leq y} \text{ avec } \sum_{i=1}^n w_{ni}(x) = 1 \quad (2.2)$$

Notons que si l'on ne conditionne pas, i.e $w_{ni}(x) = \frac{1}{n}$ on retrouve l'expression classique de la fonction de répartition empirique. En posant $Y^* = \mathbb{I}_{\{Y \leq y\}}$ on a $F(y|x) = \mathbb{E}(Y^*|X) = x$ et $\hat{q}_n(\alpha|x)$ est appelé l'estimateur du quantile de régression

Cas du design aléatoire : la convergence ponctuelle en probabilité de l'estimateur (2-2) a été établie par (Stone(1977)) lorsque les variables aléatoires $(X_i; Y_i) \in \mathbb{R}^p \times \mathbb{R}$ sont indépendantes et identiquement distribuées, (Collomb'(1980)) après avoir étudié les propriétés asymptotiques d'un estimateur à noyau de probabilité conditionnelle d'un couple de variables aléatoires indépendantes et identiquement distribuées à valeurs dans $\mathbb{R}^p \times \mathbb{R}$ a proposé d'estimer la fonction de répartition conditionnelle en posant

$$w_{ni}(x) = \frac{K\left(\frac{x-X_i}{h_n}\right)}{\sum_{j=1}^n K\left(\frac{x-X_j}{h_n}\right)}$$

où K est une densité de probabilité appelée noyau et h_n un paramètre qui converge vers zéro lorsque n tend vers l'infini. Il a démontré la convergence ponctuelle et uniforme en x de cet estimateur. Il n'a pas démontré la convergence uniforme en y . Il a donné également l'erreur quadratique moyenne optimale ainsi que la normalité asymptotique de ses estimateurs

En utilisant un noyau de probabilité continu et borné sur $[-1, 1]$ (Stute (1986)) propose :

$$w_{ni}(x) = \frac{K\left(\frac{G_n(x)-G_n(X_i)}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{G_n(x)-G_n(X_j)}{h_n}\right)} \quad \text{avec } G_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}_{\{X_j \leq x\}}$$

il a démontré la convergence uniforme de son estimateur de la fonction de répartition conditionnelle. Il a établi un résultat sur la normalité asymptotique de l'estimateur du quantile conditionnel associé

Horvath et Yandall (1988) donnent des conditions permettant d'approcher la suite $\varphi_n(y|x) = \sqrt{nh_n}(\hat{F}_n(y|x) - F(y|x))$ par un processus gaussien. Ils énoncent le résultat sur la convergence uniforme (en probabilité) en y de la suite $\varphi_n(y|x)$ des estimateurs de Collomb(1980) et Stute(1986). Toutefois, notons que Stute(1986) a établi un résultat sur la convergence faible de la suite $\varphi_n(y|x)$ vers un processus gaussien

Gannoun(1989) étudie les propriétés de l'estimateur de Collomb(1980) dans le cas de données indépendantes et identiquement distribuées puis α -mélangeantes. Hart(1991) propose quant à lui d'utiliser un noyau de probabilité de support $[0; 1]$ et de modifier l'estimateur de Stute(1986) en prenant une famille de poids définie par

$$w_{ni}(x) \frac{1}{h_n} \int_{\frac{i-1}{n}}^{\frac{i}{n}} K\left(\frac{G_n(x) - u}{h_n}\right) du$$

Lorsque la covariable X est de nature fonctionnelle ou de dimension infinie, Ferraty et Vieu (2000) proposent de modifier la famille de poids introduit par Collomb(1980) en posant

$$w_{ni}(x) = \frac{K\left(\frac{d(x, X_i)}{h_n}\right)}{\sum_{j=1}^n K\left(\frac{d(x, X_j)}{h_n}\right)}$$

où d est une distance semi-métrique qui mesure la proximité entre deux objets fonctionnels et K est un noyau positif et décroissant sur $[0; 1]$. Les vitesses de convergence presque complète ont été obtenues dans le cas d'un échantillon indépendant et dépendant, on pourra se référer à l'ouvrage (Ferraty et Vieu(2006), chapitre 6).

Cas du design aléatoire : (Antoch et Janssen(1989)) étendent l'étude d'un modèle régression initialement introduit par (Gasser et Hüller(1984)) à l'estimation des quantiles conditionnels. Ils proposent alors d'estimer la fonction de répartition conditionnelle en posant :

$$\begin{aligned}
w_{ni}(x) &= \frac{1}{h_n} \int_{x_{i-1}}^{x_i} K\left(\frac{x_i - u}{h_n}\right) du \text{ pour } 2 \leq i \leq n-1 \\
w_{n1}(x) &= \frac{1}{h_n} \int_{-\infty}^{x_1} K\left(\frac{x - u}{h_n}\right) du \\
w_{nn}(x) &= \frac{1}{h_n} \int_{x_{n-1}}^{\infty} K\left(\frac{x_i - u}{h_n}\right) du
\end{aligned}$$

(b)-L'estimateur de Roussas

où K est un noyau de probabilité et les auteurs donnent une représentation du type de Bahadur de l'estimateur de quantile qui en découle.

une version plus lisse et régulière des estimateurs simple noyau de la fonction de répartition conditionnelle (2;2) avait été introduit par Roussas(1969) dans le cas d'un processus $(X_i; Y_i)$ à valeurs de $\mathbb{R} \times \mathbb{R}$, supposé Markovien. Elle consiste à remplacer la fonction indicatrice de l'estimateur (appel) par une densité symétrique, puis à effectuer préalablement à l'inversion un lissage par noyau de type Parzen(1979) de la fonction de densité marginale $g(x)$ de X de la fonction de densité conjointe $f(x, y)$ de (X, Y) . Cela revient à supposer que la variable aléatoire (X, Y) admet une densité de probabilité $f(., .)$ et que la densité conditionnelle de Y sachant $X = x$ admet une fonction régulière $f(., |x)$.

L'estimateur de la densité conditionnelle étant défini par le rapport des estimateurs de la densité du couple $f(x, y)$ et la densité marginale $g(x)$, il en découle que l'estimateur à noyau de la fonction de répartition conditionnelle est :

$$\hat{F}_n(y|x) = \int_{-\infty}^y f_n(u|x) du = \frac{\int_{-\infty}^y \hat{f}_n(x, u) du}{\hat{g}_n(x)} \stackrel{\text{def}}{=} \frac{\hat{\psi}_n(x, y)}{\hat{g}_n(x)}$$

$$\begin{aligned}
\text{où } \hat{f}_n(x, y) &= \frac{1}{nh_n^{d+1}} \sum_{i=0}^n K_0\left(\frac{x - X_i}{h_n}\right) K_1\left(\frac{y - Y_i}{h_n}\right) \\
\hat{g}_n(x) &= \int_{\mathbb{R}^p} \hat{f}_n(x, y) dy = \frac{1}{nh_n^d} \sum_{i=1}^n K_0\left(\frac{x - X_i}{h_n}\right) \\
\hat{\psi}_n(x, y) &= \frac{1}{nh_n^{d+1}} \sum_{i=0}^n K_0\left(\frac{x - X_i}{h_n}\right) \int_{-\infty}^y K_1\left(\frac{z - Y_i}{h_n}\right) dz
\end{aligned}$$

avec K_0 et K_1 des noyaux de probabilités.

De nombreuses approches ont été explorées suivant la structure de dépendance de données. Samanta(1989) énonce des résultats sur la consistance forte et de normalité asymptotique dans le cas des observations indépendantes et identiquement distribuées. Berline et al (1998) étendent ces résultats aux données non indépendantes. Ils considèrent le processus (X_i, Y_i) stationnaire et α -mélangeant. Ils énoncent des théorèmes qui établissent qu'un estimateur convergent du quantile $q(.|x)$ construit à partir d'un estimateur convenable de $F(.|x)$ et asymptotiquement normal, quelque soit la structure de dépendance de données.

Théorème 2.1 (Berline et al(1998)) : Soit $x \in \mathbb{R}^p$ et $(\mathcal{U}_n)_{n \geq 1}$ une suite de réels. On suppose que $F(.|x)$ admet un quantile conditionnel unique $q(.|x)$ d'ordre $\alpha \in [0; 1]$. On suppose que $f(.|x)$, $\hat{f}_n(.|x)$ et $F_n(.|x)$ existent, que $f(.|x)$ est continue et que $f(q(\alpha|x)|x) \neq 0$. Si pour n tend vers ∞

1. $\hat{q}(\alpha|x)$ converge presque sûrement vers $q(\alpha|x)$
2. $\mathcal{U}_n(\hat{F}_n(q(\alpha|x)|x) - F(q(\alpha|x)|x))$ converge en loi vers $\mathcal{N}(m_q(x), \sigma_q^2(x))$
3. $\hat{f}_n(.|x)$ converge uniformément en probabilité vers $f(.|x)$ sur un voisinage de $\mathcal{V}(q(.|x))$

alors

$$\mathcal{U}_n(\hat{q}_n(\alpha|x) - q(.|x)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\frac{m_q(x)}{f(q(\alpha|x)|x)}, \frac{\sigma_q^2(x)}{f^2(q(\alpha|x)|x)}\right)$$

Si l'on suppose que $\{(X_i, Y_i); i = 1, \dots, n\}$ sont des copies indépendantes du couple

aléatoire $(X, Y) \in \mathbb{R}^p \times \mathbb{R}$ alors le théorème 2.1 est vérifié avec

$$\mathcal{U}_n = (nh_n^p)^{\frac{1}{2}}m_q(x) = 0, \sigma_q^2 = \frac{\alpha(\alpha - 1)}{g(x)} \|K\|_2^2$$

voir (berlinet et al(2001)).

Signalons que les auteurs ont aussi établis la convergence presque complète de leurs estimateurs

Lorsque X est de nature fonctionnelle, (Ferraty et al(2005)) proposent de modifier l'estimateur dans le cas des données non nécessairements dépendantes l'estimateur ainsi défini est donnée par :

$$\hat{F}_n(y|x) = \frac{\sum_{i=1}^n K_0\left(\frac{d(x;X_i)}{h_{n,0}}\right) \int_{-\infty}^y K_1\left(\frac{z-Y_i}{h_{n,1}}\right) dz}{\sum_{i=1}^n K_0\left(\frac{d(x;X_i)}{h_{n,0}}\right)} \quad (2.3)$$

où d est une distance semi-métrique et K_1 un noyau positif et décroissant sur $[0; 1]$, K_0 est un noyau symétrique $h_{n,0}$ $h_{n,1}$ deux suites positives. Les auteurs établissent la vitesse de convergence presque complète de l'estimateur de quantile conditionnel qui en découle pour plus de détails, on pourra aussi consulter (Ferraty et Vieu(2005,2006), chapitre 5 et 6)

(c)-L'estimateur de la fenêtre mobile

Encore appelé médianogramme, son ancêtre le régressogramme fut introduit par Tukey(1961). Cet estimateur est valable quelque soit le design pour un réel fixé $h_n > 0$, on se place en un point fixé x et on sélectionne les seuls Y_i pour lesquels les points d'observations X_i (ou x_i) appartiennent à la boule centré en x de rayon h_n . Ceci revient à estimer la fonction de répartition conditionnelle par l'estimation paramétrique des quantiles conditionnels en posant

$$w_{ni}(x) = \frac{\mathbb{I}_{\{X_i \in B(x, h_n)\}}}{\sum_{i=1}^n \mathbb{I}_{\{X_i \in B(x, h_n)\}}} \quad \text{design aléatoire}$$

$$w_{ni}(x) = \frac{\mathbb{I}_{\{x_i \in B(x, h_n)\}}}{\sum_{i=1}^n \mathbb{I}_{\{x_i \in B(x, h_n)\}}} \quad \text{design fixe}$$

où $B(x, h_n)$ est une boule centrée en x et de rayon $h_n \rightarrow 0$ quand $n \rightarrow \infty$. L'estimateur du quantile conditionnel est construit alors en inversant l'estimateur de la fonction de répartition empirique du sous-échantillon des observations dont les covariables sont dans la boule, on peut alors remarquer que l'estimateur de la fenêtre mobile est un estimateur à noyau particulier correspondant au cas où le noyau est $K(x) = \mathbb{I}_{\{x \in [-1; 1]\}}$

L'estimateur de la médiane conditionnelle est construit suivant cette méthode. Truong(1989) a donné la vitesse optimale de la convergence d'un estimateur empirique de la médiane conditionnelle. Gannoun(1989) a établi la convergence uniforme presque complète de cet estimateur .

Estimateur au sens des plus proches voisins la méthode de construction de ces estimateurs est analogue à ceux de la fenêtre mobile. Bhattacharya et Gangopadhyay(1990) définissent l'estimateur de la fonction de répartition conditionnelle de Y sachant $X = x$ au sens des k plus proches voisins comme étant la fonction de répartition empirique des observations aux k points X_i (ou x_i) les plus proches des points x_i .

$$\hat{F}_k(y|x) = \frac{1}{k} \sum_{i=1}^n \mathbb{I}_{\{Y_{[ni]} \leq y\}}$$

où k est entier inférieur ou égal à n et les $Y_{[ni]}$ sont les concomitants des statistiques d'ordre $U_{n,1} < \dots < U_{n,n}$ associés aux $U_i = |X_i - x|$ (où $|x_i - x|$). Le quantile conditionnel d'ordre α au sens du plus proches voisins est alors défini par

$$\hat{q}_k(\alpha|x) = \inf\{y | \hat{F}_k(y|x) \geq \frac{\lfloor k\alpha \rfloor}{k}\}$$

les auteurs montrent que leur estimateur de quantile conditionnel est asymptotiquement gaussien pour $k = \lfloor n^{\frac{4}{5}} s \rfloor$ où $s \in [a; b]$ avec $a < 0 < b$

Méthode d'estimation directe

Elle consiste à ramener le problème d'estimation des quantiles conditionnels à un problème d'optimisation

$$q(\alpha|x) = \arg \min_{\theta \in \mathbb{R}} \mathbb{E}[\rho_{\alpha}(Y - \theta) | X = x]$$

où la fonction ρ est définie par $\rho_{\alpha}(y) = \alpha y \mathbb{I}_{[0, \infty[} - (1 - \alpha) y \mathbb{I}_{]-\infty, 0]}$ La théorie sous-jacente de cette technique d'estimation est décrite dans (Basset et Koenker(1982))

contrairement aux méthodes d'estimations indirecte on n'a pas besoin de déduire le comportement asymptotique du quantile conditionnel de la convergence de l'estimateur de la fonction de répartition. Dans la littérature on a répertorié trois méthodes d'estimation directe à savoir

- La méthode d'estimation par la constante locale qui est un cas particulier de la méthode d'estimation par les polynômes locaux.
- La méthode d'estimation par des fonctions splines de régression développée par He et Shi (1994). Ils proposent d'approcher la fonction quantile par une fonction spline. Cardot et al. (2004, 2005) généraliseront cet estimateur au cas où la variable explicative X est de type fonctionnel.
- La méthode d'estimation par des fonctions splines de lissage proposée par Koenker et al. (1994). Elle avait été spécialement développée pour le design fixe

Méthode de la constante locale la technique consiste à approcher le quantile par une fonction linéaire

$$q(\alpha|z) \approx q(\alpha|x) + q'(\alpha|x)(z - x) \stackrel{def}{=} a + b(z - x)$$

Pour z un voisinage de x . De façon naturelle, ceci revient à utiliser le principe de la méthode des moindres carrés

$$\min_{(a,b)} \sum_{i=1}^n \rho_{\alpha}(Y_i - a + b(X_i - x)) K\left(\frac{x - X_i}{h_n}\right)$$

Si $b = 0$ on parle de la méthode de la constante locale introduite par Tsybakov(1986) et on estime le quantile conditionnel

$$\hat{q}_n(\alpha|x) = \arg \min_{a \in \mathbb{R}} \rho_{\alpha}(Y_i - a) K\left(\frac{x - X_i}{h_n}\right)$$

Les résultats sur la consistance faible et la convergence en moyenne quadratique de cet estimateur sont dus à Stone(1977) et Yu et Jones(1998), Fan et al(1994) établissent la convergence en loi. La convergence uniforme sur un support compact peut être obtenue en adaptant le résultat de Berlinet et al(2001). Par ailleurs, cette méthode se comporte bien face aux effets de bord (Fan et al(1994))(Koenker et al (1992))(Mint El Mouvit, 2000)

Estimation par des polynômes locaux Plus générale que la méthode de la constante locale, elle consiste à adapter la méthode des polynômes locaux classique pour la moyenne conditionnelle aux quantiles conditionnels. Ainsi, si l'on approche le quantile conditionnel par un polynôme de degré k , cela revient à résoudre le problème d'optimisation

$$\min_{(b_0, \dots, b_k)} \sum_{i=0}^n \rho_{\alpha}(Y_i - \sum_{j=0}^k b_j (X_i - x)^j) K\left(\frac{x - X_i}{h_n}\right)$$

L'estimateur de quantile conditionnel est alors défini par

$$\hat{q}_n(\alpha|x) = \hat{b}_0$$

Chaduri(1991) a démontré la consistance forte ponctuelle dans le cas d'un noyau de type de fonction indicatrice d'intervalles. Pour $k = 1$, on retrouve l'estimateur de la

constante locale présenté précédemment

2.4 Estimation des quantiles extrêmes conditionnels en design fixe

Dans la littérature, on dénombre plusieurs méthodes d'estimations des quantiles extrêmes conditionnels. Smith (1989) et Davison et Smith (1990) proposent des familles de modèles paramétriques basées sur les excès au dessus du seuil. Beirlant et Goegebeur (2003) en adoptant une approche semi-paramétrique proposent de transformer tout d'abord les données, puis de les utiliser dans un modèle de régression exponentiel où les paramètres dudit modèle sont estimés par la méthode du maximum de vraisemblance. Hall et Tajvidi (2000) proposent dans le cas d'une série temporelle $\{(Y_i, t_i), i = 1, \dots, n\}$ où t_i est le temps, de combiner l'estimation non-paramétrique de la tendance temporelle avec une hypothèse sur la distribution conditionnelle de Y_i sachant t_i . L'estimation non-paramétrique des quantiles extrêmes conditionnels a été introduite à notre connaissance dans Davison et Ramesh (2000) où les auteurs

proposent des estimateurs par ajustement polynomial. Beirlant et Goegebeur (2004) étendent ces résultats aux covariables multidimensionnelles. En outre, ils donnent les propriétés asymptotiques des estimateurs qui en découlent. Dans le cas d'une covariable unidimensionnelle, Beirlant et al. (2004a) se proposent d'adapter les estimateurs de quantiles proposés dans Beirlant et Matthys (2001) et Beirlant et Matthys (2003) au cas conditionnel en remplaçant tout simplement les statistiques d'ordres par les quantiles estimés par la méthode des polynômes locaux (voir Koenker et Bassett, 1978). Plus récemment, Chavez-Demoulin et Davison (2005) en utilisant la méthode du maximum de vraisemblance pénalisé proposent des estimateurs splins de quantiles extrêmes conditionnels dans le cas d'une covariable unidimensionnelle. D'autres méthodes d'estimation ont été proposées par Chernozhukov (1998, 2001) et Hall et Ronde (2000).

Il existe de nombreuses applications en chimométrie ou en astrophysique où les cova-

riables sont des courbes. On parle alors de covariables fonctionnelles. A notre connaissance, aucun auteur ne s'est encore attardé sur l'estimation des quantiles conditionnels pour des jeux de données présentant cette particularité. Notons que dans un tel contexte, l'estimation de quantile requiert des techniques de lissage non-paramétrique adaptées aux données fonctionnelles afin de mieux prendre en compte la covariable (Ramsay et Silverman (1997)), (Bosq (2000)), (Ferraty et Vieu (2006)), (Ramsay et Silverman (2002)).

2.4.1 Contexte d'étude et présentation

Cette partie contient deux paragraphes, le premier est dédié au contexte d'étude et la présentation de notre méthode d'estimation des quantiles conditionnels, la deuxième est consacré à la présentation stricto sensu des estimateurs

Contexte d'étude et méthode d'estimation

Soit $Y \in \mathbb{R}$ une variable aléatoire associée à une covariable non-aléatoire $x \in E$ où E désigne un espace métrique, non nécessairement de dimension finie, muni d'une distance d . D'une manière précise, le problème est le suivant. Soient $\{(x_i, Y_i), i = 1, \dots, n\}$ des observations indépendantes du couple $(x, Y) \in E \times \mathbb{R}$, on veut estimer le quantile d'ordre $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$ dans le cas particulier où la fonction de répartition conditionnelle de Y sachant x est notée $F(y, x)$ est à queue lourde ceci signifie que pour tout $y > 0$

$$F(y, x) = 1 - y^{-\frac{1}{\gamma(x)}} l(y, x)$$

où $\gamma(x)$ est une fonction positive et inconnue de la covariable x que l'on appelle <indice de queue conditionnel> et pour tout x fixé $l(\cdot, x)$ est une fonction à variation lente à l'infini, alors on a pour tout $\lambda > 0$

$$\lim_{y \rightarrow \infty} \frac{l(\lambda y, x)}{l(y, x)} = 1$$

Ceci revient donc à supposer que pour tout x fixé, la fonction de répartition conditionnelle $F(\cdot, x) \in \mathcal{D}(\text{Fréchet})$. Dans ce cas, pour tout $x \in E$ le quantile conditionnel d'ordre $(1 - \alpha_n)$ et noté $q(\alpha_n, x)$ est à décroissance polynômiale d'indice $-\gamma(x)$.

Dans l'optique de combiner des techniques de lissage non-paramétrique avec des méthodes d'analyse de valeurs extrêmes dont le but est d'obtenir des estimateurs ayant de bonnes propriétés asymptotiques, nous utilisons une méthode d'estimation dite de la fenêtre mobile pour construire nos estimateurs. Pour cela, on introduit une boule centrée en x de rayon $h_n > 0$, notée $B(x, h_n)$ définie par

$$B(x, h_n) = \{t \in E, d(x, t) \leq h\}$$

Etant donné $h_{n,x}$ une suite positive convergeant vers 0 quand $n \rightarrow \infty$ on se propose de ne sélectionner que les observations Y_i dont les covariables x_i sont dans la boule $B(x, h_{n,x})$ la proposition de tels points donnés par

$$\varphi(h_{n,x}) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{x_i \in B(x, h_{n,x})\}}$$

Joue un rôle central dans cette étude. De façon similaire à la notion de probabilité de petites boules utilisées en analyse fonctionnelle dans (Ferraty et Vieu(2006)) $\varphi(h_{n,x})$ décrit comment cet ensemble de points se concentre dans un voisinage de x lorsque $h_{n,x} \rightarrow 0$. Dans la suite $m_n(x)$ désignera le nombre d'observations qui sont dans la boule $B(x, h_{n,x})$. Notons $\{Z_i(x), i = 1, \dots, m_{n,x}\}$ les observations de Y retenues par la procédure de sélection et nous désignerons par $Z_{1,m_{n,x}}(x) \leq \dots \leq Z_{m_{n,x},m_{n,x}}(x)$ les statistiques ordonnées correspondantes. Dans un tel contexte, il apparaît que dans la région $]0, \infty[\times B(x, h_{n,x})$, le nombre d'observations retenues par la méthode de sélection est lié à l'ensemble $\varphi(h_{n,x})$ par la relation $m_{n,x} = n\varphi(h_{n,x})$.

L'avantage de notre approche est que très peu d'hypothèses sont faites sur la régularité de la fonction $q(\cdot|x)$ ou $\gamma(x)$ et sur la nature de la covariable x . Des résultats sur le comportement asymptotique sur nos estimateurs des quantiles extrêmes conditionnels

sont établis sans faire d'hypothèses sur la dimension de E . D'un point de vue pratique nos estimateurs sont faciles à calculer car il ne requièrent aucune technique d'optimisation.

Estimateurs des quantiles extrêmes conditionnels

Dans cette partie, on veut construire en tout point x de E , un estimateur du quantile extrême conditionnel d'ordre $1 - \alpha_{m_{n,x}}$. Ici on parle de quantile extrême si on a $\alpha_{m_{n,x}} \rightarrow 0$ quand $m_{n,x} \rightarrow \infty$. En fonction de la vitesse de convergence de $\alpha_{m_{n,x}} \rightarrow 0$, trois situations sont envisagées

1. $\alpha_{m_{n,x}}$ converge lentement vers 0, c'est à dire $\alpha_{m_{n,x}} \rightarrow 0$ alors $m_{n,x}\alpha_{m_{n,x}} \rightarrow \infty$ quand $m_{n,x} \rightarrow \infty$
2. $\alpha_{m_{n,x}}$ converge rapidement vers 0, c'est à dire $\alpha_{m_{n,x}} \rightarrow 0$ alors $m_{n,x}\alpha_{m_{n,x}} \rightarrow c \in [1; \infty[$ quand $m_{n,x} \rightarrow \infty$
3. $\alpha_{m_{n,x}}$ converge très rapidement vers 0, c'est à dire $\alpha_{m_{n,x}} \rightarrow 0$ alors $m_{n,x}\alpha_{m_{n,x}} \rightarrow c \in [0; 1]$ quand $m_{n,x} \rightarrow \infty$

Dans la 1^{ère} situation $\alpha_{m_{n,x}}$ converge moins vite vers 0 que le rapport $\frac{1}{m_{n,x}}$. Par conséquent, l'estimation du quantile extrême conditionnel requiert d'interpoler à l'intérieur de l'échantillon car $q(\alpha_{m_{n,x}}, x)$ est presque sûrement inférieur à l'observation maximale, on propose alors d'estimer $q(\alpha_{m_{n,x}}, x)$ par

$$\hat{q}_1(\alpha_{m_{n,x}}, x) = Z_{m_{n,x} - \lfloor m_{n,x}\alpha_{m_{n,x}} \rfloor + 1, m_{n,x}} \quad (2.4)$$

Dans la 2^{ème} situation, pour n assez grand on a $\lfloor m_{n,x}\alpha_{m_{n,x}} \rfloor = \lfloor c \rfloor > 0$. L'estimation du quantile extrême dans ce cas repose sur les plus grandes observations situées au voisinage de l'échantillon, mais toujours dans l'ensemble des données. Par conséquent, on peut réutiliser l'estimateur (2.4)

Dans la situation (S.3), $\alpha_{m_{n,x}}$ converge au moins aussi vite vers zéro que le rapport $1/m_{n,x}$. Estimer le quantile extrême conditionnel nécessite d'extrapoler au-delà des observations puisque $q(\alpha_{m_{n,x}}, x)$ est supérieur à l'observation maximale avec probabilité

$e^{-c} \geq e^{-1}$. Dans une telle situation, on propose d'adapter l'estimateur de Weissman (1978) au cas conditionnel. On estime alors $q(\alpha_{m_{n,x}}, x)$ par

$$\hat{q}_2(\alpha_{m_{n,x}}, x) = \hat{q}_1(\beta_{m_{n,x}}, x) \left(\frac{\beta_{m_{n,x}}}{\alpha_{m_{n,x}}} \right)^{\hat{\gamma}_n(x)}$$

où $\beta_{m_{n,x}}$ satisfait la 1^{ère} situation et $\hat{\gamma}_n(x)$ est un estimateur de l'indice des valeurs extrêmes conditionnels

Les trois situations ont déjà été étudiées dans le cas non conditionnel. de Haan (1984) énonce le premier résultat dans la situation (3) en posant $c = 0$. Dekkers et de Haan (1989) étudient les situations (1) et (3) avec $c \neq 0$. Leurs résultats peuvent être consultés dans (Embrechts et al., 1997, Théorèmes 6.4.14 et 6.4.15).

Dans la situation (2), si la constante c n'est pas entière alors, $m_{n,x} \alpha_{m_{n,x}} \rightarrow c$ implique que $\lfloor m_{n,x} \alpha_{m_{n,x}} \rfloor \rightarrow \lfloor c \rfloor$ quand $m_{n,x} \rightarrow \infty$. Sinon, si c est un entier, alors la condition $\lfloor m_{n,x} \alpha_{m_{n,x}} \rfloor \rightarrow \lfloor c \rfloor$ est nécessaire car elle empêche à la suite $(\lfloor m_{n,x} \alpha_{m_{n,x}} \rfloor)_{m_{n,x} > 0}$ d'avoir deux valeurs d'adhérence.

2.4.2 Etude théorique des estimateurs

Cette partie est subdivisée en deux paragraphes. Dans le premier on donne nos hypothèses de travail et dans le second on établit les lois asymptotiques. Dans tout ce qui suit, on suppose que $x \in E$

Hypothèses

Dans ce paragraphe, on énumère les hypothèses nous permettant d'étudier le comportement asymptotique de nos estimateurs de quantile extrêmes conditionnels

hypothèse sur le quantile conditionnel

(A.1) la fonction quantile conditionnel

$$\alpha \in]0; 1[\rightarrow q(\alpha, x) \in]0; +\infty[$$

est dérivable et la fonction biais

$$\alpha \in]0; 1[\rightarrow \Delta(\alpha, x) = \gamma(x) + \alpha \frac{d \log q}{d\alpha}(\alpha, x) \in]0; +\infty[$$

est continue et telle que $\lim_{\alpha \rightarrow 0} \Delta(\alpha, x) = 0$

L'hypothèse **(A.1)** a pour but de contrôler le comportement de la fonction log-quantile quant à sa première variable. C'est une condition suffisante pour que la fonction de répartition conditionnelle $F(\cdot, x)$ soit à queue lourde (Bingham et al (1987)), afin de simplifier les notations pour tout $a \in]0; 1[$, on notera

$$\bar{\Delta}(a, x) = \sup_{\alpha \in]0; a[} |\Delta(a, x)|$$

Il paraît important de préciser que dans **(A.1)**, on suppose implicitement que pour tout $x \in E$ ($\dim E \leq \infty$) la fonction à variations lentes $l(\cdot, x)$ est dérivable.

Hypothèse sur le paramètre de lissage

(A.2) Soit $(k_{n,x})_{n \geq 1}$ une suite d'entiers telle que $1 \leq k_{n,x} \leq m_{n,x}$, on suppose que

$$\frac{n\varphi(h_{n,x})}{k_{n,x}} \rightarrow \infty \text{ et } k_{n,x} \rightarrow \infty \text{ quand } n \rightarrow \infty$$

Cette hypothèse implique que $n\varphi(h_{n,x}) \rightarrow \infty$, c'est à dire le nombre d'observations retenues par la méthode de la fenêtre mobile tend vers l'infini lorsque la taille de l'échantillon tend vers l'infini.

Etude de comportement asymptotique des estimateurs

Cette partie est consacrée à l'étude théorique de nos estimateurs. Pour cela, il convient de donner quelques résultats auxiliaires utiles pour établir leur loi limite. Cependant il apparaît nécessaire de commencer par une définition de l'oscillation de la fonction log-quantile

Définition 2.2 Pour tout $a \in [0; 1/2]$, la plus grande oscillation de la fonction log-quantile par rapport à sa seconde variable est donnée par

$$\omega_n(a) = \sup \left\{ \left| \log \frac{q(\alpha, t)}{q(\alpha, t')} \right|, a \in [a, 1 - a], (t, t') \in B(x, h_{n,x})^2 \right\}$$

Notre premier résultat est dédié à l'étude de la présentation en loi des grandes variables aléatoires de l'échantillon

$\{Z_i(x); i = 1, \dots, m_{n,x}\}$ retenues par notre procédure de sélection

Proposition 2.3 Soit $J_{k_{n,x}} = \{1, \dots, k_{n,x}\}$ sous les hypothèses (A.1) et (A.2), $k_{n,x}^2 \omega_n(m_{n,x}^{-(1+\delta)}) \rightarrow 0$ pour certains $\delta > 0$, alors il existe un événement \mathcal{A}_n avec une probabilité convergant vers 1 quand $n \rightarrow \infty$ tel que

$$\{(\log Z_{m_{n,x}-i+1, m_{n,x}}(x), i \in J_{k_{n,x}} | \mathcal{A}_n)\} \stackrel{\mathcal{L}}{=} \{\log q(V_{i, m_{n,x}}, T_i), i \in J_{k_{n,x}} | \mathcal{A}_n\}$$

où $V_{1, m_{n,x}} \leq \dots \leq V_{m_{n,x}, m_{n,x}}$ sont les statistiques ordonnées associées à la suite des variables aléatoires de loi uniforme standard $\{V_i, i = 1, \dots, m_{n,x}\}$ et $\{T_i, i = 1, \dots, k_{n,x}\}$ sont des variables aléatoires appartenant à la boule $B(x, h_{n,x})$.

La condition $k_{n,x}^2 \omega_n(m_{n,x}^{-(1+\delta)}) \rightarrow 0$ montre qu'il est plus facile de contrôler les $k_{n,x}$ plus grandes observations dans la région $]0, \infty[\times B(x, h_{n,x})$ lorsque l'oscillation de la fonction log-quantile par rapport à sa second variable est petite. Un rapprochement peut être fait entre notre proposition et l'approximation utilisée par (Falk et al (2004), Théorème 3.5.2) dans l'étude de la loi de k-plus proches voisin en utilisant la distance de Hellinger

Donnons à présent un résultat sur la position du quantile extrême conditionnel

$q(\alpha_{m_{n,x}}, x)$ dans l'ensemble des données.

Proposition 2.4 Sous l'hypothèse (A.1) si $\omega_n(m_{n,x}^{-(1+\delta)}) \rightarrow 0$ pour un certain $\delta > 0$, alors

dans la première situation, $\mathbb{P}(Z_{m_{n,x}, m_{n,x}} < q(\alpha_{m_{n,x}}, x) \rightarrow 0$

dans les deux autres $\mathbb{P}(Z_{m_{n,x}, m_{n,x}} < q(\alpha_{m_{n,x}}, x) \rightarrow \exp(-c)$

Les résultats suivants établissent la loi limite d'un estimateur du quantile extrême conditionnel construit à partir de notre procédure d'estimation. Focalisons nous tout d'abord à l'estimation du quantile conditionnel dans la 1^{ère} situation

Théorème 2.5 *Soit $(\alpha_{m_{n,x}})_{n \geq 1}$ une suite satisfaisant la 1^{ère} situation sous (A.1), si pour $n \rightarrow \infty$, il existe $\delta > 0$ tel que $(m_{n,x} \alpha_{m_{n,x}})^2 \omega_n(m_{n,x}^{-(1+\delta)}) \rightarrow 0$ alors*

$$\sqrt{m_{n,x} \alpha_{m_{n,x}}} \left(\frac{\hat{q}_1(\alpha_{m_{n,x}}, x)}{q_1(\alpha_{m_{n,x}}, x)} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^2(x))$$

Dans cette situation, il apparaît que l'estimateur est asymptotiquement gaussien avec une variance asymptotique proportionnelle à $\frac{\gamma^2(x)}{(m_{n,x} \alpha_{m_{n,x}})}$. Ainsi, plus grande est la valeur de l'indice de queue conditionnel, c'est-à-dire plus lourde est la queue de la loi, et plus grande est la variance de notre estimateur de quantile extrême conditionnel. De plus, la variance asymptotique étant inversement proportionnelle à $\alpha_{m_{n,x}}$, l'estimation de quantile extrême est d'autant plus stable que l'on s'éloigne de la frontière de l'échantillon.

Considérons maintenant la 2^{ème} situation dont l'estimation du quantile repose sur les plus grandes observations situées au voisinage de la frontière de l'échantillon.

Théorème 2.6 *Soit $(\alpha_{m_{n,x}})_{n \geq 1}$ une suite satisfaisant la 2^{ème} situation, sous l'hypothèse (A.1), si pour $n \rightarrow \infty$, il existe $\delta > 0$ tel que $(m_{n,x} \alpha_{m_{n,x}})^2 \omega_n(m_{n,x}^{-(1+\delta)}) \rightarrow 0$ alors*

$$\left(\frac{\hat{q}_1(\alpha_{m_{n,x}}, x)}{q(\alpha_{m_{n,x}}, x)} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{E}(c, \gamma(x))$$

où $\mathcal{E}(c, \gamma(x))$ est une loi non dégénérée

Dans la situation (S.2), la loi asymptotique du quantile extrême conditionnel n'est pas gaussienne et son expression est assez compliquée. En outre, l'estimateur $\hat{q}_1(\cdot, x)$ n'est pas consistant

La dernière situation (S.3) est beaucoup plus intéressante mais aussi complexe puisqu'elle fait intervenir d'une part, l'estimateur de quantile $\hat{q}_1(\cdot, x)$ de la situation (S.1) et

d'autre part, l'estimateur de l'indice de queue conditionnel $\hat{\gamma}_n(x)$. Ainsi, sa loi asymptotique peut provenir de l'un ou de l'autre.

Théorème 2.7 Soit $(\beta_{m_{n,x}})_{n \geq 1}$ une suite satisfaisant la situation 1 et soit $(\alpha_{m_{n,x}})_{n \geq 1}$ une suite telle que $\alpha_{m_{n,x}} < \beta_{m_{n,x}}$. On pose $\zeta_{m_{n,x}} = (m_{n,x} \beta_{m_{n,x}})^{1/2} \log(\beta_{m_{n,x}} / \alpha_{m_{n,x}})$. Sous (A.1), si pour $n \rightarrow \infty$, il existe $\delta > 0$ tel que $(m_{n,x} \alpha_{m_{n,x}})^2 \omega_n(m_{n,x}^{-(1+\delta)}) \rightarrow 0$, et il existe une suite positive $\mathcal{V}_n(x) \rightarrow \infty$ et une loi \mathcal{D} telle que

$$\mathcal{V}_n(x)(\hat{\gamma}_n(x) - \gamma(x)) \xrightarrow{\mathcal{L}} \mathcal{D}$$

alors deux situations se présentent

1. Sous la condition additionnelle :

$$\zeta_{m_{n,x}} \max \left\{ \mathcal{V}_n^{-1}(x), \bar{\Delta}(\beta_{m_{n,x}}, x) \right\} \text{ nous avons } \sqrt{m_{n,x} \alpha_{m_{n,x}}} \left(\frac{\hat{q}_1(\alpha_{m_{n,x}}, x)}{q(\alpha_{m_{n,x}}, x)} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^2(x))$$

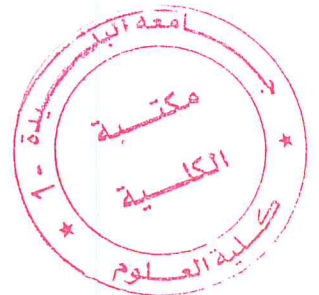
2. Sous la condition additionnelle :

$$\mathcal{V}_n(x) \max \left\{ \zeta_{m_{n,x}}^{-1}, \bar{\Delta}(\beta_{m_{n,x}}, x) \right\} \text{ nous avons } \frac{\mathcal{V}_n(x)}{\log(\beta_{m_{n,x}} / \alpha_{m_{n,x}})} \left(\frac{\hat{q}_2(\alpha_{m_{n,x}}, x)}{q(\alpha_{m_{n,x}}, x)} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{D}$$

De ce théorème, on remarque que la loi asymptotique de $\hat{q}_2(\cdot, x)$ dépend d'une part du comportement de $\hat{q}_1(\cdot, x)$ et d'une autre part du comportement de $\hat{\gamma}_n(x)$. Dans la situation 1, la normalité asymptotique de $\hat{q}_2(\cdot, x)$ est donnée par la statistique d'ordre $\hat{q}_1(\cdot, x)$ qui est ni rien d'autre $m_{n,x} \alpha_{m_{n,x}}$ ième plus grande observation parmi les $\{Z_i(x), i = 1, \dots, m_{n,x}\}$ observations contenues dans la boule $B(x, h_{n,x})$. À l'opposé, dans la 2^{ème} situation $\hat{q}_2(\cdot, x)$ hérite de la loi limite de l'estimateur de l'indice de queue conditionnel

La loi asymptotique de l'estimateur de quantile extrême conditionnel fait apparaître un biais asymptotique dont l'ordre en valeur absolue est donné par :

$$\left| \int_{\alpha_{m_{n,x}}}^{\beta_{m_{n,x}}} \frac{\Delta(u, x)}{u} du \right| \leq \log \left(\frac{\beta_{m_{n,x}}}{\alpha_{m_{n,x}}} \right) \bar{\Delta}(\beta_{m_{n,x}}, x)$$



Ainsi, dans la 1^{ère} situation la condition $\zeta_{m_{n,x}} \bar{\Delta}(\beta_{m_{n,x}}, x) \rightarrow 0$ impose au biais d'être négligeable devant l'écart type de $\hat{q}_2(\alpha_{m_{n,x}}, x)$ qui vaut quant à lui $(m_{n,x} \beta_{m_{n,x}})^{1/2}$. Dans la situation 2 la condition $\mathcal{V}_n(x) \max \left\{ \zeta_{m_{n,x}}^{-1}, \bar{\Delta}(\beta_{m_{n,x}}, x) \right\} \rightarrow 0$ impose au biais d'être négligeable devant l'écart type $\hat{q}_2(\alpha_{m_{n,x}}, x)$ qui est ici égal à $\frac{\mathcal{V}_n(x)}{\log(\beta_{m_{n,x}}/\alpha_{m_{n,x}})}$

Notons Toutefois que, même si l'intérêt principal du théorème (2.7) est d'établir le comportement asymptotique du quantile extrême conditionnel dans la situation où $\alpha_{m_{n,x}}$ converge très rapidement vers zéro, il peut également être appliqué dans toutes les situations moins restrictive que la situation (S.3), où $\alpha_{m_{n,x}}$ est plus petit que $\beta_{m_{n,x}}$. Par exemple dans la situation (S.2) il apparaît que $\hat{q}_2(\alpha_{m_{n,x}}, x)$ est un estimateur faiblement consistant i.e :

$$\frac{\hat{q}_2(\alpha_{m_{n,x}}, x)}{q_2(\alpha_{m_{n,x}}, x)} \xrightarrow{\mathbb{P}} 1$$

2.4.3 Exemple et discussion

La partie suivante comprend deux paragraphes. Tandis que le premier expose quelques méthodes d'estimation de l'indice de queue conditionnel, le deuxième s'articule quant-à lui autour de deux points, à savoir des applications et des illustrations

Quelques exemples d'estimateurs de $\gamma(x)$

Dans la littérature, on dénombre plusieurs approches d'estimation de l'indice de queue conditionnel. Celles-ci sont pour la majorité inhérentes aux méthodes d'estimations présentées dans l'introduction. Ainsi, on pourra se référer aux travaux de Davison et Smith (1990), Davison et Ramesh (2000), Hall et Tajvidi (2000), Beirlant et al. (2002), Beirlant et Goegebeur (2004) et Chavez-Demoulin et Davison (2005) pour de plus amples explications.

En ce qui nous concerne, nous allons présenter la famille d'estimateurs proposée par Gardes et Girard (2008) qui est une extension des estimateurs proposés par Beirlant et al. (2002) dans le cas univarié. Les raisons de ce choix sont essentiellement dûs à l'utilisation

de la méthode des fenêtres mobiles.

Les estimateurs de l'indice de queue conditionnel que l'on présente ici sont une somme pondérée des écarts de logarithmes entre les plus grandes observations retenues par la méthode de fenêtres mobiles.

Définition 2.8 Soient $Z_{i,m_{n,x}} \leq \dots \leq Z_{m_{n,x},m_{n,x}}$ les statistiques ordonnées correspondantes aux observations dans la boule $B(x, r_{n,x})$, La famille d'estimateurs défini dans Gardes et Girard(2008) est donnée par

$$\hat{\gamma}_n(x, W) = \frac{\sum_{i=1}^{k_{n,x}} i \log \left(\frac{Z_{m_{n,x}-i+1, m_{n,x}}(x)}{Z_{m_{n,x}-i, m_{n,x}}(x)} \right) W(i/k_{n,x}, x)}{\sum_{i=1}^{k_{n,x}} W(i/k_{n,x}, x)}$$

où $W(., x)$ est une fonction de poids défini sur $]0, 1[$ telle que $\int_0^1 W(s, x) ds \neq 0$

L'étude de cet estimateur repose sur le modèle de regression exponentielle pour des écarts pondérés de logarithme entre les $k_{n,x}$ plus grandes observations dans la boule $B(x, r_{n,x})$. Sous certaines hypothèses de régularité de la distribution conditionnelle de Y sachant x et sous certaines conditions sur la fonction des poids, les auteurs montrent que leur estimateur est asymptotiquement gaussien avec un biais dépendant à la fois de la distribution théorique de la fonction de poids $W(., x)$:

$$k_{n,x}^{1/2} (\hat{\gamma}_n(x, W) - \gamma(x) - \Delta(k_{n,x}/m_{n,x}, x) \mathcal{AB}(x, W)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^2(x) \mathcal{AV}(x, W))$$

où $\mathcal{AB}(x, W) = \int_0^1 W(s, x) s^{-\rho(x)} ds$ avec $\rho(x)$ fonction négative appelée paramètre du second ordre et $\mathcal{AV}(x, W) = \int_0^1 W^2(s, x) ds$

La variance asymptotique de leur estimateur est proportionnelle, à un facteur d'échelle près, au carré de l'indice de queue conditionnel. Ce facteur d'échelle ou terme multiplicatif donné par $\mathcal{AV}(x, W)$ fait intervenir la fonction de poids. Ceci montre qu'un mauvais choix

de la fonction de poids $W(\cdot, x)$ a forcément des répercussions sur la qualité des estimateurs et en particulier sur la construction des intervalles de confiance.

2.5 Applications et discussion

Dans cette partie on illustre le théorème sur les quantiles extrêmes conditionnels avec l'estimateur de l'indice de queue présenté précédemment. Ensuite, nous adapterons au cas conditionnel un autre estimateur non conditionnel défini par

$$q_{\alpha_n} = \left[\prod_{i=1}^{k_n} X_{n-i+1,n} \left(\frac{i g_{k_n}}{n \alpha_n} \right)^{\hat{\gamma}_i^H} \right]^{1/k_n}$$

où

$$g_{k_n} = \exp [\log(k_n + 1) - 1 - \log(k_n \text{fact})/k_n]$$

et $\hat{\gamma}_i^H$ est l'estimateur de Hill défini au premier chapitre. Enfin, on va donner quelques exemples de lois à queues lourdes

2.5.1 Application du résultat de Gardes et Girard à nos estimateurs de quantiles extrêmes

pour tout $\beta_{m_{n,x}} = \frac{k_{n,x}}{m_{n,x}} \rightarrow 0$ l'estimateur de quantile extrême de la situation (S.3) peut se réécrire sous la forme

$$\hat{q}_2(\alpha_{m_{n,x}}, x, W) = Z_{m_{n,x}-k_{n,x}+1, m_{n,x}}(x) \left(\frac{k_{n,x}}{\alpha_{m_{n,x}} m_{n,x}} \right)^{\hat{\gamma}_n(x,W)}$$

Compte tenu la remarque faite précédemment, seule la situation 2 du théorème 7 présente un intérêt certain au vu de la vitesse de convergence des estimateurs de l'indice de queue conditionnel introduit par Gardes et Girard(2008). Cette situation peut être réécrite de façon à mieux intégrer le comportement asymptotique de ces estimateur.

Corollaire 2.9 *Supposons les hypothèses de Gardes et Girard(2008, Théorème 2) satisfaites. Sous les deux hypothèses (A.1)-(A.2), si la suite $(k_{n,x})_{n \geq 1}$ est telle que :*

$$k_{n,x}^{1/2} \bar{\Delta}(k_{n,x}/m_{n,x}, x) \rightarrow 0 \quad n \rightarrow \infty \quad (2.5)$$

$$k_{n,x}^{1/2} \omega_n(m_{n,x}^{-(\delta+1)}) \rightarrow 0 \text{ pour } \delta > 0 \quad n \rightarrow \infty \quad (2.6)$$

et si $(\alpha_{m_{n,x}})$ est une suite satisfaisant la situation (S.2) ou (S.3) alors

$$\frac{k_{n,x}^{1/2}}{\log\left(\frac{k_{n,x}}{\alpha_{m_{n,x}} m_{n,x}}\right)} \left(\frac{\hat{q}_2(m_{n,x}, x, W)}{q(\alpha_{m_{n,x}}, x)} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^2(x) \mathcal{AV}(x, W))$$

Comme exemple de famille de poids nous pouvons citer :

La famille de poids constante qui consiste à poser pour tout $s \in [0, 1]$, $W^H(s, x) = 1$. Alors, l'estimateur de queue obtenu est une adaptation de l'estimateur de Hill (1975) au cas conditionnel. Dans ce cas, l'estimateur de quantile extrême conditionnel $\hat{q}_2(m_{n,x}, x, W^H)$ est asymptotiquement gaussien avec $\mathcal{AB}(x, W^H) = \frac{1}{1-\rho(x)}$ et $\mathcal{AV}(x, W^H) = 1$

La Famille de poids logarithmique définie pour tout $s \in]0, 1]$ par $W^Z(s, x) = -\log(s)$. L'estimateur de queue qui en découle est une adaptation de l'estimateur de Zipf (Schultze et Steinebach, 1996 ; Kratz et Resnick, 1996) au cas conditionnel. Dans ce cas $\hat{q}_2(m_{n,x}, x, W^Z)$ est asymptotiquement gaussien avec $\mathcal{AB}(x, W^Z) = \frac{1}{(1-\rho(x))^2}$ et $\mathcal{AV}(x, W^Z) = 2$

Pour d'autres exemples de familles de poids, le lecteur pourra se référer à l'article de Gardes et Girard (2008).

On peut estimer $q(\alpha_{m_{n,x}}, x)$ en adaptant l'estimateur de quantile extrême 1.1 introduit au cas conditionnel. Ainsi, on a donc

$$\hat{q}_3(\alpha_{m_{n,x}}, x) = \left[\prod_{i=1}^{k_{n,x}} Z_{m_{n,x}-i+1, m_{n,x}}(x) \left(\frac{ig_{k_{n,x}}}{m_{n,x} \alpha_{m_{n,x}}} \right)^{\hat{\gamma}_i(x, W^H)} \right]^{\frac{1}{k_{n,x}}}$$

avec $g_{k_n} = \exp[\log(k_{n,x} + 1) - 1 - \log(k_{n,x} \text{ fact})/k_{n,x}]$

et $\hat{\gamma}_i(x, W^H) \stackrel{\text{def}}{=} \frac{1}{i} \sum_{j=1}^i j \log \frac{Z_{m_{n,x}-i+1, m_{n,x}}(x)}{Z_{m_{n,x}-j, m_{n,x}}(x)}$

Corollaire 2.10 *Supposons les hypothèses de Gardes et Girard(2008, Théorème 2) satisfaites. Sous les deux hypothèses (A.1)-(A.2), si la suite $(k_{n,x})_{n \geq 1}$ vérifie les conditions et et $\frac{\log(k_{n,x})}{\log(\alpha_{m_{n,x}} m_{n,x})} \rightarrow 0$ et si $(\alpha_{m_{n,x}})_{n \geq 1}$ est une suite satisfaisant la situation (S.3) alors*

$$\frac{k_{n,x}^{1/2}}{\log\left(\frac{k_{n,x}}{\alpha_{m_{n,x}} m_{n,x}}\right)} \left(\hat{q}_3(m_{n,x}, x) - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2\gamma^2(x)).$$

Il apparait que les estimateurs $\hat{q}_2(\alpha_{m_{n,x}}, x, W^Z)$ et $\hat{q}_3(\alpha_{m_{n,x}}, x)$ ont la même loi asymptotiquement.

Illustration sur quelque lois à queue lourde La première illustration que nous proposons est la loi de Pareto. Cette loi est l'exemple le plus simple des lois à queue lourdes car sa fonction à variations lentes vaut 1. En effet son quantile extrême d'ordre $1 - \alpha_{m_{n,x}}$ a pour expression $q(\alpha_{m_{n,x}}, x) = \alpha_{m_{n,x}}^{-\gamma(x)}$. Par conséquent, sa fonction biais $\Delta(\alpha_{m_{n,x}}, x) = 0$ et la condition est trivialement satisfaite

La deuxième illustration est la loi de *Fréchet* dont le quantile extrême est défini par

$$q(\alpha_{m_{n,x}}, x) = \alpha_{m_{n,x}}^{-\gamma(x)} \left(\frac{1}{\alpha_{m_{n,x}}} \log \left(\frac{1}{1 - \alpha_{m_{n,x}}} \right) \right)^{-\gamma(x)} = \left(\log \left(\frac{1}{1 - \alpha_{m_{n,x}}} \right) \right)^{-\gamma(x)}$$

Ce quantile conditionnel décroît aussi comme un polynôme en $\alpha_{m_{n,x}}$ de puissance $-\gamma(x)$ i.e $q(\alpha_{m_{n,x}}, x) \simeq \alpha_{m_{n,x}}^{-\gamma(x)}$ et la qualité de cette estimation est contrôlée par

$$\Delta(\alpha_{m_{n,x}}, x) = \frac{-\gamma(x)}{2} \alpha_{m_{n,x}} (1 + O(\alpha_{m_{n,x}}))$$

Une autre illustration peut être donnée par la loi de Burr dont le quantile conditionnel et son erreur d'approximation sont respectivement donnés pour tout $\rho(x) < 0$ par :

$$q(\alpha_{m_{n,x}}, x) = \alpha_{m_{n,x}}^{-\gamma(x)} (1 - \alpha_{m_{n,x}}^{-\rho(x)})^{\frac{-\gamma(x)}{\rho(x)}}$$

et $\Delta(\alpha_{m_{n,x}}, x) = -\gamma(x) \alpha_{m_{n,x}}^{-\rho(x)} (1 + O(\alpha_{m_{n,x}}^{-\rho(x)}))$

On remarque que pour les lois de *Fréchet* ou de *Burr* la fonction de biais $\Delta(\alpha_{m_{n,x}}, x)$ est asymptotiquement proportionnelle à $\alpha_{m_{n,x}}^{-\rho(x)}$ quand $\alpha_{m_{n,x}} \rightarrow 0$ avec la convention $\rho(x) = -1$ pour la loi de *Fréchet*. Le paramètre du second ordre contrôle la qualité d'approximation du quantile conditionnel $q(\alpha_{m_{n,x}}, x)$ par le polynôme $\alpha_{m_{n,x}}^{-\gamma(x)}$. En outre, comme pour ces deux lois la valeur absolue de la fonction biais $|\Delta(\alpha_{m_{n,x}}, x)|$ est croissante alors condition peut se réécrire

$$m_{n,x}^{2\rho(x)} k_{n,x}^{1-\rho(x)} \rightarrow 0$$

Ce qui montre que l'on peut prendre $k_{n,x}$ grand lorsque la fonction $\rho(x)$ est éloignée de zéro, c'est-à-dire petite. Enfin, si on suppose que la fonction indice de queue conditionnelle $\gamma(\cdot)$ et le paramètre du second ordre $\rho(\cdot)$ sont lipschitziens, i.e s'il existe des constantes $c_\gamma > 0$ et $c_\rho > 0$ telles que

$$|\gamma(t) - \gamma(t')| \leq c_\gamma d(t, t') \text{ et } |\rho(t) - \rho(t')| \leq c_\rho d(t, t')$$

pour tout $(t, t') \in B(x, r_{n,x})^2$, alors la plus grande oscillation de la fonction log-quantile par rapport à sa seconde variable peut être bornée et on a

$$\omega_n(a) = O(h_{n,x} \log\left(\frac{1}{a}\right)) \text{ quand } a \rightarrow 0$$

Donc on peut simplifier la condition et on obtient :

$$k_{n,x}^2 r_{n,x} \log m_{n,x} \rightarrow 0$$

	$q(\alpha_{m_{n,x}}, x)$	$\Delta(\alpha_{m_{n,x}}, x)$
Pareto	$\alpha_{m_{n,x}}^{-\gamma(x)}$	0
Fréchet	$\alpha_{m_{n,x}}^{-\gamma(x)} \left(\frac{1}{\alpha_{m_{n,x}}} \log \left(\frac{1}{1 - \alpha_{m_{n,x}}} \right) \right)^{-\gamma(x)}$	$\frac{-\gamma(x)}{2} \alpha_{m_{n,x}} (1 + O(\alpha_{m_{n,x}}))$
Burr	$\alpha_{m_{n,x}}^{-\gamma(x)} (1 - \alpha_{m_{n,x}}^{-\rho(x)})^{\frac{-\gamma(x)}{\rho(x)}}$	$-\gamma(x) \alpha_{m_{n,x}}^{-\rho(x)} (1 + O(\alpha_{m_{n,x}}^{-\rho(x)}))$

Table : Quelques exemples de lois à queues lourdes. $\gamma(x) > 0$ est la fonction indice de queue conditionnel et $\rho(x) < 0$ est le paramètre du second-ordre

Chapitre 3

Estimation non-paramétrique de mesures de risque extrêmes pour des lois conditionnelles à queues lourdes

3.1 Introduction

La maîtrise des risques est un sujet de préoccupation aussi bien en hydrologie qu'en finance et en actuariat. Se prémunir contre les risques en finance et en actuariat est primordial afin d'anticiper respectivement des crises financières où d'importants sinistres coûteux pour les assurances. En hydrologie l'étude des pluies et débits extrêmes est d'un grand intérêt afin de se prémunir contre des catastrophes telles que des inondations et aussi permettre les aménagements du territoire.

Pour cela, il existe des outils permettant de quantifier et de prédire le risque : les mesures de risque. Elles permettent d'évaluer un niveau de dangerosité d'un risque, mais également de comparer différents risques entre eux. La quantification, la prévention, la comparaison et la prédiction du risque sont des éléments essentiels de notre société.

3.2 Mesures de risque

Dans la suite, la variable aléatoire d'intérêt dont on souhaite quantifier le risque sera notée Y . Elle désignera également la variable aléatoire du montant de perte. Donnons à présent la définition précise d'une mesure de risque.

Définition 3.1 *On appelle mesure de risque une fonction \mathcal{R} associant à Y une valeur positive ou nulle telle que :*

$$\mathcal{R} : Y \rightarrow \mathcal{R}^+ \in \mathbb{R}^+$$

De grandes valeurs de $\mathcal{R}(Y)$ indiqueront que Y est dangereux, $\mathcal{R}(Y)$ peut être vu comme le capital à détenir face aux pertes Y .

Une des mesures de risque les plus connues, si ce n'est la plus connue, est la Value-at-Risk (notée VaR) introduite dans les années 90 par Morgan (voir Jorion). C'est historiquement la plus utilisée des mesures de risque.

L'idée derrière le concept de la **VaR** est la suivante : on fixe un seuil probabiliste α et on calcule une valeur $VaR(\alpha)$ qui sera telle que la probabilité que la catastrophe survienne soit plus petite que α . On peut voir la $VaR(\alpha)$ comme le montant d'extra capital dont une entreprise a besoin afin de réduire à α la probabilité de faire faillite.

En termes statistiques, la VaR de niveau $\alpha \in]0, 1[$ correspond au quantile de niveau $1 - \alpha$ de la fonction de répartition de la loi des pertes.

La VaR présente cependant plusieurs défauts, elle ne donne qu'une information ponctuelle au quantile $q(\alpha)$ et aucune information au-delà de ce point ; elle ne prend ainsi pas en compte l'importance du sinistre lorsqu'il survient mais seulement sa fréquence. De ce fait comme elle ne prend pas en considération ce que serait la perte au-delà de ce quantile, alors elle ne fait pas la différence entre les sinistres. Donc la VaR sous estime les pertes.

Un des principaux reproche fait à la VaR et que des variables aléatoires à queues légères et à queues lourdes (voir Tasche) peuvent avoir la même **VaR**.

Pour répondre à la nécessité de principes théoriques et pratiques, Artzner et al ont cherché à caractériser ce qui ferait qu'une mesure de risque soit "performante". Pour cela ils ont introduit la notion de mesure de risque cohérente

3.2.1 Mesure de risque cohérente

Définition 3.2 Une mesure de risque \mathcal{R} est dite cohérente si, pour deux variables aléatoires T et Z , elle satisfait les 4 propriétés suivantes :

- 1) La Monotonie : si $\mathbb{P}(Z \leq T) = 1$, alors $\mathcal{R}(Z) \leq \mathcal{R}(T)$
- 2) Homogénéité positive : $\mathcal{R}(aZ) = a\mathcal{R}(Z)$ quelque soit $a > 0$
- 3) invariance par translation : $\mathcal{R}(Z + b) = \mathcal{R}(Z) + b$, quelque soit $b > 0$
- 4) Sous-additivité : $\mathcal{R}(Z + T) = \mathcal{R}(Z) + \mathcal{R}(T)$

La monotonie permet de s'assurer que si la variable aléatoire Z est presque sûrement plus petite que T alors le risque associé à Z est plus petit que celui associé à T .

L'homogénéité positive signifie que le fait de mesurer une proportion d'un risque revient à considérer la proportion de la mesure du risque seul.

De même, l'invariance par translation implique qu'ajouter une valeur positive au risque implique l'ajout de cette même valeur à la mesure du risque.

L'homogénéité et l'invariance par translation impliquent qu'une mesure de risque cohérente apprécie la nature aléatoire du risque.

La sous-additivité traduit le fait que considérer deux risques simultanément est moins risqué que traiter les risques séparément. Cela intègre donc l'idée de diversification du risque.

Le principal défaut de la Value-at-Risk est qu'elle n'est pas une mesure de risque cohérente car elle ne vérifie pas la condition de sous-additivité. Une mesure de risque cohérente est la Conditional Tail Expectation (notée CTE), aussi appelée dans la littérature sur les mesures de risque Tail-Value-at-Risk, Average-Value-at-Risk, Tail Conditional Expectation ou Expected Shortfall dans le cas d'une loi continue des pertes.

La *VaR* s'intéresse aux probabilités d'événements extrêmes alors que la *CTE* s'intéresse à ce qui se passe en moyenne lorsque ces événements surviennent. En effet, la *CTE* au niveau α est définie comme la perte moyenne attendue sachant que la $VaR(\alpha)$ est dépassée. Elle représente donc la moyenne des $(1 - \alpha)100\%$ sinistres les plus élevés et peut être vue comme le montant d'extra capital nécessaire afin de surmonter en moyenne les pires exercices.

Cette mesure de risque donne des informations sur la distribution de Y au-delà de la $VaR(\alpha)$. Ainsi, contrairement à la $VaR(\alpha)$, elle prend en compte l'information contenue dans la queue de la loi, ce qui est primordial aussi bien dans un contexte financier, actuariel ou hydrologique. L'épaisseur de la queue de la distribution des sinistres est un élément fondamental dans l'évaluation du niveau de danger d'un risque ayant une faible probabilité d'apparition. La Conditional Tail Expectation a été étudiée par Artzner et al., Cai et Tan, Cai et al., Tasche et est fréquemment utilisée dans les investissements financiers ou dans l'actuariat voir Brazaukas et al. et Landsman et Valdez.

A la suite de la *VaR* et de la *CTE*, Rockafellar et Uryasev ont introduit une nouvelle mesure de risque cohérente, la Conditional-Value-at-Risk (notée *CVaR*) définie comme la moyenne pondérée de la *VaR* et de la *CTE*.

Une autre mesure de risque basée sur la *VaR* et sur la *CTE* est la Stop-loss Premium réassurance (notée SP) avec un niveau de rétention égal à la *VaR*, pour une définition plus précise.

Il existe de nombreuses autres mesures de risque, citons également, la Conditional Tail Variance (notée *CTV*). Elle représente la variance de la loi des pertes sachant que les pertes sont supérieures à la *VaR*. De même, la Conditional Tail Skewness (notée *CTS*) est définie comme le coefficient d'asymétrie de la loi des pertes sachant que les pertes sont supérieures à la *VaR*. Le choix de la mesure la plus appropriée pour quantifier le risque en assurance comme en finance reste sujet à débat.

Dans ce chapitre, on commencera par introduire un outil permettant d'unifier l'estimation des mesures de risque mentionnées juste avant. Ce outil sera appelé le Conditional

Tail Moment (noté CTM). Il est défini comme le moment d'ordre $a \geq 0$ de la loi des pertes sachant que l'on se trouve au-dessus de la VaR de niveau α .

On montrera que l'estimation du CTM permet d'estimer toutes les mesures de risque basées sur les moments conditionnels d'ordre arbitraire sachant que les pertes sont supérieures à la VaR . Il est par exemple clair que le Conditional Tail Moment d'ordre un est la Conditional Tail Expectation.

Notre seconde contribution sera d'estimer le CTM dans le cas de pertes extrêmes, en se servant pour cela de la théorie des valeurs extrêmes. Bien que de nombreux liens entre la théorie des valeurs extrêmes et les mesures de risque aient fait l'objet de recherches citons notamment Embrechts, Embrechts et al, McNeil et al, l'estimation de mesures de risque est habituellement réalisée dans la littérature pour des valeurs fixées de α voir par exemple Deme et al. et Necir et al.

L'estimateur de mesures de risque extrêmes que l'on propose englobe également la présence d'une covariable. Dans ce contexte, notre nouvelle mesure de risque sera appelée le Regression Conditional Tail Moment (notée $RCTM$). En effet, en finance, la loi des pertes peut être affectée par de nombreux facteurs, tels que les taux d'intérêts ou l'inflation. De même les météorologues s'intéressent aux pluies extrêmes comme fonction de la position géographique.

3.3 Le regression Conditional Tail Moment :définition et estimation

Soit $Y \in \mathbb{R}$ une variable aléatoire de pertes. Pour $\alpha \in [0, 1]$, la Value-at-Risk de niveau α et la quantité $VaR(\alpha)$ satisfaisant :

$$\mathbb{P}(Y > VaR(\alpha)) = \alpha$$

Comme on a expliqué dans l'introduction, la Value-at-Risk est la plus connue des

mesures de risque mais de nombreuses autres existent dans la littérature :

- La Conditional Tail Expectation est définie par

$$CTE(\alpha) = \mathbb{E}(Y|Y > VaR(\alpha))$$

- La Conditional Tail Variance a été introduite par Valdez et elle est définie par

$$CTV(\alpha) = \mathbb{E}((Y - CTE(\alpha))^2|Y > VaR(\alpha))$$

Elle mesure la variabilité conditionnelle de Y sachant $Y > VaR(\alpha)$ et nous dit à quel point les évènements s'éloignent de $CTE(\alpha)$

- Conditional Tail Skewness a été définie par Hong, J. et Elshahat(2010) :

$$CTS(\alpha) = \frac{\mathbb{E}(Y^3|Y > VaR(\alpha))}{CTV(\alpha)^{\frac{3}{2}}}$$

- La Conditional Value-at-Risk est définie par :

$$CVaR_\lambda(\alpha) = \lambda VaR(\alpha) + (1 - \alpha)CTE(\alpha)$$

avec $0 \leq \lambda \leq 1$, il est clair que dans le cas où $\lambda = 0$ (resp 1), la $CVaR_\lambda(\alpha) = CTE(\alpha)$ (resp $VaR(\alpha)$). Cette dernière permet de quantifier des risques au-delà de la $VaR(\alpha)$ et elle est de plus cohérente. Si $0 \leq \lambda < 1$ d'autres propriétés fondamentales de cette mesure se retrouvent dans Rockafellar et Uryasev.

- La mesure de risque Stop-loss Premium réassurance avec un niveau de rétention égal à $VaR(\alpha)$ est définie par :

$$SP = \mathbb{E}(Y - Var(\alpha)_+) = \alpha(CTE(\alpha) - Var(\alpha))$$

Elle est proportionnelle à la différence entre la $CTE(\alpha)$ et la $VaR(\alpha)$. Cette mesure de risque permet ainsi de mettre en évidence les cas dangereux.

Notre but est d'unifier les définitions des mesures de risques énoncées juste auparavant. Pour cela, on introduit une nouvelle mesure de risque, la Conditional Tail Moment notée CTM définie par :

$$CTM_a(\alpha) = \mathbb{E}(Y^a | Y > Var(\alpha))$$

où $a > 0$ et tel que le moment d'ordre a de Y existe. Il est facile de voir que toutes les mesures de risque précédentes de niveau α peuvent être réécrites sous la forme suivante :

$$\Phi(Var(\alpha), CTM_1(\alpha), CTM_2(\alpha), CTM_3(\alpha))$$

où la fonction $\Phi : \mathbb{R}^4 \rightarrow \mathbb{R}$ est prise dans le tableau ci-dessous :

Mesure de risque	$\Phi(t_0, t_1, t_2, t_3)$
CTE	t_1
CTV	$t_2 - t_1^2$
CTS	$t_3 / (t_2 - t_1^2)^{3/2}$
$CVaR$	$\lambda t_0 + (1 - \lambda)t_1 \quad \lambda \in [0, 1]$
SP	$\alpha(t_1 - t_0)$

Liens entre le CTM et les mesures de risque classiques.

De manière plus générale, le CTM peut être utilisé dans le but de définir n'importe quelle mesure de risque basée sur les moments conditionnels de la variable aléatoire des pertes sachant que les pertes sont supérieures à la VaR de niveau α . En effet, on peut par exemple introduire le Conditional Tail Kurtosis à l'aide de la fonction :

$$\Phi(t_0, t_1, t_2, t_3, t_4) = \frac{t_4}{(t_2 - t_1^2)^2}$$

3.3.1 Pertes extrêmes et regression

Comme évoqué dans l'introduction, notre second but est d'adapter les mesures de risque classiques dans le cas de pertes extrêmes en présence d'une covariable $X \in \mathbb{R}^p$. Pour cela, le niveau fixé $\alpha \in]0, 1[$ est remplacé par une suite $(\alpha_n) \in]0, 1[$, telle que $\alpha_n \rightarrow 0$. De plus, si l'on note par $\bar{F}(\cdot|x)$ la fonction de survie conditionnelle de Y sachant $X = x$, on peut alors définir :

- la Regression Value-at Risk par :

$$RVaR(\alpha_n|x) = \inf \{t, \bar{F}(t|x)\} \leq \alpha\}$$

- La regression Conditional Tail Moment par :

$$RCTM_a(\alpha_n|x) = \mathbb{E}(Y^a|Y > RVaR(\alpha_n|x), X = x)$$

où $a \geq 0$ et tel que le moment d'ordre a de Y existe. Notons que dans ce cadre d'étude, la $RVaR(\alpha_n|x)$ est le quantile extrême conditionnel de niveau $\alpha_n \in]0, 1[$.

Il est alors assez facile d'adapter les mesures de risque classiques dans le cas de pertes extrêmes en présence d'une covariable en appliquant la fonction adéquate au vecteur :

$$(RVaR(\alpha_n|x), RCTM_1(\alpha_n|x), RCTM_2(\alpha_n|x), RCTM_3(\alpha_n|x))$$

Cela nous donne les mesures de risque suivantes :

$$RCTE(\alpha_n|x), RCTV(\alpha_n|x), RCTS(\alpha_n|x), RCVaR(\alpha_n|x) \text{ et } RSP(\alpha_n|x)$$

3.4 Construction de l'estimateur de RCTE ET RVaR

Nous rappelons que les deux mesures de risque sont données par les formules suivantes :

$$\begin{aligned}\mathbb{P}(Y > VaR(\alpha)) &= \alpha \\ CTE(\alpha) &= \mathbb{E}(Y|Y > VaR(\alpha))\end{aligned}$$

Soient (X_i, Y_i) $i = 1, \dots, n$, des copies indépendantes du couple de variables aléatoires (X, Y) . Afin d'estimer le RCTM on commence par donnée l'égalité suivante :

$$RCTM_1(\alpha_n|x) = \frac{1}{\alpha_n} \varphi_a(\varphi_0^-(\alpha_n|x)|x) \quad (3.1)$$

où pour $y > 0$ le moment conditionnel d'ordre $a \geq 0$ est donné par :

$$\varphi_a(y|x) = \mathbb{E}(Y^a \mathbb{I}\{Y > y\} | X = x) \quad (3.2)$$

L'estimation du RCTM dépend ainsi de l'estimation du moment conditionnel. On propose d'utiliser un estimateur à noyau classique donné par

$$\hat{\varphi}_{a,n}(y|x) = \frac{\sum_{i=1}^n K_h(x - X_i) Y_i^a \mathbb{I}\{Y_i > y\}}{\sum_{i=1}^n K_h(x - X_i)} \quad (3.3)$$

où $\mathbb{I}(\cdot)$ est la fonction indicatrice et $h = h_n$ est une suite non aléatoire telle que $h \rightarrow 0$ quand $n \rightarrow \infty$. Dans le cas où $a = 0$ dans l'équation 3 on trouve l'estimateur à noyau de la fonction de survie conditionnelle introduit par Collomb. On a aussi introduit $K_h(t) = K(t/h)/h^p$ où K est une densité sur \mathbb{R}^p . Dans ce contexte d'étude, h est appelée paramètre de lissage

Puisque $\hat{\varphi}_{a,n}(y|x)$ est une fonction décroissante, on peut définir un estimateur de $\varphi_a^-(\alpha|x)$ pour $\alpha \in]0, 1[$ par

$$\varphi_{a,n}^{\leftarrow}(\alpha|x) = \inf \{t, \hat{\varphi}_{a,n}(y|x) < \alpha\} \quad (3.4)$$

En remarquant que $\hat{\varphi}_0(y|x) = \overline{F}(y|x)$, la *RVaR* de niveau α_n est alors estimée par

$$\widehat{RVaR}_{\alpha_n}(\alpha_n|x) = \hat{\varphi}_{0,n}^{\leftarrow}(y|x) \quad (3.5)$$

On trouve aussi l'estimateur du quantile extrême conditionnel étudié par Daouia et al. Le *RTCM* est estimé par

$$\widehat{RTCM}_{1,n}(\alpha_n|x) = \frac{1}{\alpha_n} \hat{\varphi}_{1,n}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) \quad (3.6)$$

Un estimateur de chaque mesure de risque mentionnée auparavant est alors donné par :

$$\phi \left(\widehat{RVaR}_n(\alpha_n|x), \widehat{RTCM}_{1,n}(\alpha_n|x) \right) \quad (3.7)$$

Où la fonction ϕ est choisi dans le tableau qui montre le lien entre le *CTM* et les mesures de risque

3.5 Etude du comportement des deux estimateurs

Notre hypothèse principale est la suivante :

(F.1) on suppose que la fonction de survie conditionnelle de Y sachant $X = x$ est à queue lourde et admet une fonction densité

En résumé (F.1) revient à supposer que la loi conditionnelle de Y sachant $X = x$ appartient au domaine d'attraction de *Fréchet*. L'hypothèse (F.1) équivaut à supposer que pour tout $y > 0$, $\overline{F}(x) = \mathbb{P}(y > x|X = x)$ est à variations régulières à l'infini avec un indice $-\frac{1}{\gamma(x)}$ que l'on notera $F(\cdot|x) \in \mathcal{RV}_{\frac{-1}{\gamma(x)}}$ i.e pour tout $\lambda > 0$ on a

$$\lim_{y \rightarrow \infty} \frac{\bar{F}(\lambda y|x)}{\bar{F}(y|x)} = \lambda^{-\frac{1}{\gamma(x)}}$$

Dans ce contexte, $\gamma(\cdot)$ est une fonction positive de la covariable x est sera appelée indice de queue conditionnel puisqu'elle contrôle la lourdeur de la queue de la loi conditionnelle de Y sachant $X = x$.

On peut alors remarquer que sous l'hypothèse (F.1), une condition suffisante de l'existence de du $RTCM_a(1/\cdot|x)$ et $a < \frac{1}{\gamma(x)}$. Comme établi dans le Lemme ..., la condition (F.1) implique également que pour tout $a \in [0, \frac{1}{\gamma(x)}[$, $\varphi_a(\cdot|x) \in \mathcal{RV}_{a-\frac{1}{\gamma(x)}}$, puisque de plus $\bar{F}(\cdot|x) \in \mathcal{RV}_{\frac{-1}{\gamma(x)}}$, on a alors le $RTCM_a(1/\cdot|x) \in \mathcal{RV}_{a\gamma(x)}$. Cela équivaut à supposer que pour $a \in [0, 1/\gamma(x)[$ et pour tout $y > 0$:

$$RTCM_a(1/y|x) = y^{a\gamma(x)} l_a(y|x) \quad (3.8)$$

où pour x fixé, $a(\cdot|x)$ est une fonction à variations lentes à l'infini, i.e pour tout $\lambda > 0$

$$\lim_{y \rightarrow \infty} \frac{l_a(\lambda y|x)}{l_a(y|x)} = 1 \quad (3.9)$$

Afin d'établir la normalité asymptotique de l'équation 5, les conditions suivantes sont nécessaires. Premièrement, puisque les fonctions à variations lentes sont d'un intérêt seulement asymptotique on peut ainsi supposer que dans l'équation 7 :

(F.2) $a(\cdot|x)$ est normalisée pour tout $a \in [0, 1/\gamma(x)[$.

Dans ce cas, la représentation de Karamata d'une fonction à variations lentes nous donne :

$$l_a(y|x) = c_a(x) \exp \left(\int_1^y \frac{\varepsilon_a(u|x)}{u} du \right) \quad (3.10)$$

où $c_a(\cdot)$ est une fonction positive et $\lim_{y \rightarrow \infty} \varepsilon_a(y|x) \rightarrow 0$. Ainsi, $l_a(\cdot|x)$ est différentiable et la fonction auxiliaire est donnée par $\varepsilon_a(y|x) = y \frac{l'_a(y|x)}{l_a(y|x)}$. Cette fonction joue un rôle important dans la théorie des valeurs extrêmes car elle contrôle la vitesse de convergence

dans l'équation 8, et plus généralement le biais des estimateurs en théorie des valeurs extrêmes. Comme estimateur à noyau de cette fonction, on pourra citer les travaux de Goegebeur et al. Goegebeur et de Wet supposent que la fonction auxiliaire est à variations régulières à l'infini et s'intéressent à l'estimation de l'indice conditionnel de variations régulières. Ici, on se limite à faire l'hypothèse que pour tout $a \in [0, 1/\gamma(x)[$.

(F.3) $|\varepsilon_a(\cdot|x)|$ est continue et décroissante à l'infini.

Une condition de type Lipschitz sur la fonction de densité g de X est aussi nécessaire. Pour tout $(x, x') \in \mathbb{R}^p \times \mathbb{R}^p$ la distance entre x et x' sera notée $d(x, x')$ on peut introduire les hypothèses suivantes

(L) il existe une constante $c_g > 0$ telle que $|g(x) - g(x')| \leq c_g d(x, x')$

L'hypothèse suivante est classique lorsque l'on fait de l'estimation par noyau.

(K) K est une densité bornée sur \mathbb{R}^p , de support S inclu dans la boule unité de \mathbb{R}^p .

Pour $\xi > 0$ la plus grande oscillation au point $(x, y) \in \mathbb{R}^p \times \mathbb{R}_*^+$ du moment conditionnel d'ordre $a \in [0, 1/\gamma(x)[$ est donnée par $:\omega(x, y, a, \xi, r_{n,x}) = \sup \left\{ \left| \frac{\varphi_a(z|x)}{\varphi_a(z|x')} - 1 \right|, z \in [(1 - \xi)y, (1 + \xi)y] \right\}$

Théorème 3.3 *Supposons que (F.1) (F.2) (L) et (K) soient vérifiées. Introduisons $0 \leq a_1 < a_2 < \dots < a_J$ où J est un entier positif. Pour tout $x \in \mathbb{R}^p$ tel que $g(x) > 0$ et $\gamma(x) < \frac{1}{2a_J}$, on introduit une suite $\alpha_n \rightarrow 0$ et $nh_n^p \alpha_n \rightarrow \infty$ quand $n \rightarrow \infty$ s'il existe $\xi > 0$ tel que*

$$nh_{n,x} \alpha_n \left(h \vee \max_{a \in \mathcal{L}(\{0, a_1, \dots, a_J\})} \omega(x, y, a, \xi, h_{n,x}) \right)^2 \rightarrow 0$$

alors le vecteur aléatoire

$$\sqrt[nh_{n,x} \alpha_n] \left\{ \left(\frac{\widehat{RTCE}_{J,n}(\alpha_n|x)}{\widehat{RTCE}_J(\alpha_n|x)} - 1 \right)_{J=1, \dots, J}, \left(\frac{\widehat{RVaR}_n(\alpha_n|x)}{\widehat{RVaR}(\alpha_n|x)} - 1 \right) \right\}$$

est asymptotiquement Gaussien, centré, avec matrice de covariance de taille $(J + 1) \times$

$(J + 1)$ donnée par $\|K\|_2^2 \gamma^2(x) \sum \frac{(x)}{g(x)}$ où pour $(i, j) \in \{1, \dots, J\}^2$

$$\begin{aligned} \sum_{i,j} (x) &= \frac{a_i a_j (2 - (a_i + a_j) \gamma(x))}{(1 - (a_i + a_j) \gamma(x))} \\ \sum_{J+1,j} (x) &= a_j, \quad \sum_{i,J+1} (x) = a_j, \quad \sum_{J+1,J+1} (x) = 1 \end{aligned}$$

Autrement écrit, la matrice de variance-covariance asymptotique est

$$\frac{\|K\|_2^2 \gamma^2(x)}{g(x)} \begin{vmatrix} & & & & & a_1 \\ & & & & & a_2 \\ & & & & & \cdot \\ & & & & & \cdot \\ & & & & & \cdot \\ & & & & & \cdot \\ & & & & & \cdot \\ & & & & & \cdot \\ \frac{a_i a_j (2 - (a_i + a_j) \gamma(x))}{(1 - (a_i + a_j) \gamma(x))} & & & & & a_J \\ a_1 & a_2 & \dots & a_J & & 1 \end{vmatrix}$$

Elle est proportionnelle à $\gamma^2(x)$ ainsi si $\gamma(x)$ augmente (i.e. plus la queue est lourde) plus la variance de nos estimateurs augmente.

La densité $g(x)$ de la covariable est au dénominateur de la matrice de variance-covariance asymptotique ainsi moins il y a de point (i.e. plus la densité est faible) plus la variance des estimateurs sera grande.

La condition $nh_{n,x}\alpha_n \rightarrow \infty$ est nécessaire et suffisante pour qu'il y ait presque sûrement au moins une observation dans la région $B(x, r_{n,x}) \times [\varphi_0^-(\alpha_n|x), +\infty[$ de $\mathbb{R}^p \times \mathbb{R}$ la condition

$$nr_{n,x}\alpha_n \left(h \vee \max_{a \in \mathcal{L}(\{0, a_1, \dots, a_J\})} \omega(x, y, a, \xi, h_{n,x}) \right)^2 \rightarrow 0$$

implique que le biais induit par le lissage est négligeable en comparaison de l'écart type.

Le théorème ci-dessous permet d'établir la normalité asymptotique de n'importe quel estimateur d'une mesure de risque basée sur un moment arbitraire au-dessus d'un quantile

conditionnel extrême.

Si $\alpha_n = \alpha$ fixé on retrouve la condition de normalité asymptotique classique : $nh_{n,x} \rightarrow \infty$. En l'absence de covariable on retrouve la condition de normalité asymptotique classique en théorie des valeurs extrêmes non conditionnelles : $n\alpha_n \rightarrow \infty$

Corollaire 3.4 *supposons que (F.1) (F.2) (K) et (L) soient vérifiées. Pour tout $x \in \mathbb{R}^p$ tel que $g(x) > 0$ et $\gamma(x) < \frac{1}{2}$, introduisons une suite (α_n) avec $\alpha_n \rightarrow 0$ et $nr_{n,x}\alpha_n \rightarrow \infty$ quand $n \rightarrow \infty$. S'il existe $\xi > 0$ tel que*

$$nh_{n,x}\alpha_n \left(h \vee \max_{a \in \{0,1,2\}} \omega(x, y, a, \xi, h_{n,x}) \right)^2 \rightarrow 0$$

alors on a

$$\sqrt[2]{nh_{n,x}\alpha_n} \left(\frac{\widehat{RTCE}_{J,n}(\alpha_n|x)}{RTCE_J(\alpha_n|x)} - 1 \right) \xrightarrow{d} \mathcal{N} \left(0, \frac{2\gamma^2(x)(1-\gamma(x))\|K\|_2^2}{1-2\gamma(x)g(x)} \right)$$

Théorème 3.5 *Supposons que les trois conditions (F.1) (F.2) et (F.3) sont vérifiées. Considérons $\hat{\gamma}_n(x)$ un estimateur de l'indice de queue tel que*

$$\sqrt[2]{nh^p\alpha_n} (\hat{\gamma}_n(x) - \gamma(x)) \xrightarrow{d} \mathcal{N}(0, v^2(x))$$

avec $v(x) > 0$. Si de plus, $(\beta_n)_{n \geq 1}$ est une suite positive telle que $\beta_n \rightarrow 0$ et $\frac{\beta_n}{\alpha_n} \rightarrow 0$ et $\sqrt[2]{nh^p\alpha_n}\varepsilon_a\left(\frac{1}{\alpha_n}|x\right) \rightarrow 0$ quand $n \rightarrow \infty$, on a alors :

$$\frac{\sqrt[2]{nh^p\alpha_n}}{\log\left(\frac{\alpha_n}{\beta_n}\right)} \left(\frac{\widehat{RCTE}_n^W(\beta_n|x)}{RCTE(\beta_n|x)} - 1 \right) \xrightarrow{d} \mathcal{N}(0, (av(x))^2)$$

La condition $\sqrt[2]{nh^p\alpha_n}\varepsilon_a\left(\frac{1}{\alpha_n}|x\right) \rightarrow 0$ implique que le biais induit par la fonction à variations lentes normalisées et négligeable en comparaison de l'écart type .

La condition $\frac{\beta_n}{\alpha_n} \rightarrow 0$ permet d'extrapoler et donc de choisir un ordre β_n arbitrairement petit

Notons que la normalité asymptotique de

$$\widehat{RVaR}_n^W(\beta_n|x) = \widehat{RVaR}_n(\beta_n|x) \left(\frac{\alpha_n}{\int \beta_n} \right)^{\hat{\gamma}_n(x)}$$

a été établie par Daouia et al. Remarquons de plus que l'on retrouve exactement l'estimateur de Weissman adapté au cas conditionnel. Par conséquent, en remplaçant \widehat{RVaR}_n par \widehat{RVaR}_n^W et \widehat{RCTM}_1 par $\widehat{RCTM}_{1,n}^W$ dans cela nous fournit des estimateurs pour les deux mesures de risque considérées dans ce chapitre pour des niveaux arbitrairement petits. Leurs normalités asymptotiques sont une simple conséquence de ce Théorème.

3.6 Démonstration

3.6.1 Résultats préliminaires et leurs preuves

Proposition 3.6 *Supposons que les hypothèses (F.1) (F.2) (K) et (L) sont vérifiées. Soit $x \in \mathbb{R}^p$ tel que $g(x) > 0$ et soient $0 \leq a_1 \leq \dots \leq a_{J+1} \leq \frac{1}{2\gamma(x)}$ où J est un entier positif. Considérons $\alpha_n \rightarrow 0$ tel que $nh^p\alpha_n \rightarrow \infty$ soit $(\alpha_{n,j})$ des suites tel que*

$$\max_{j \in \{1, \dots, J+1\}} \left| \frac{\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)}{\varphi_0^{\leftarrow}(\alpha_n|x)} - 1 \right| \rightarrow 0$$

S'il existe $\xi > 0$ tel que

$$nh_{n,x}\alpha_n \left(h \vee \max_{a \in \mathcal{L}(\{a_1, \dots, a_{J+1}\})} \omega(x, \varphi_0^{\leftarrow}(\alpha_n|x), a, \xi, h_{n,x}) \right)^2 \rightarrow 0$$

alors le vecteur aléatoire

$$\left\{ \sqrt{nh_{n,x}\alpha_n} \left(\frac{\hat{\varphi}_{a_j,n}^{\leftarrow}(\alpha_{n,j}|x)}{\varphi_{a_j}^{\leftarrow}(\alpha_n|x)} - 1 \right) \right\}_{j \in \{1, \dots, J+1\}}$$

est asymptotiquement gaussien, centré avec une matrice de variance-covariance

$$\frac{\|k\|_2^2 \sum_{i,j}^{(2)}}{g(x)}, \text{ où } \sum_{i,j}^{(2)} = \frac{\gamma^{2(x)}}{1 - (a_i + a_j)\gamma(x)} \text{ avec } (i, j) \in \{1, \dots, J+1\}$$

Lemme 3.7 Ce Lemme nous donne un équivalent de $\varphi_a(y|x)$ quand $y \rightarrow \infty$. Sous l'hypothèse (F.1) pour $a \in [0, \frac{1}{\gamma(x)}[$ on a

$$\varphi_a(y|x) = \frac{1}{1 - a\gamma(x)} y^a \bar{F}(y|x) (1 + o(1))$$

De plus, si l'on ajoute la condition (F.2), la dérivée $\phi'_a(\cdot|x)$ de la fonction $\phi_a(\cdot|x)$ existe et elle est à variations régulières telle que :

$$\phi'_a(y|x) = \frac{a\gamma(x) - 1}{\gamma(x)} \frac{\phi_a(y|x)}{y} (1 + o(1)).$$

preuve du Lemme . Premièrement, une intégration par partie nous donne :

$$a \int_y^\infty z^{a-1} \bar{F}(y|x) dz = \varphi_a(y|x) - y^a \bar{F}(y|x) \quad (3.11)$$

En utilisant le résultat de Resnick avec :

$$y \rightarrow y^{a-1}\overline{F}(y|x) \in \mathcal{RV}_{a-\frac{1}{\gamma(x)}-1} \text{ où } a - \frac{1}{\gamma(x)} - 1 < -1 \text{ et } y \rightarrow \infty$$

on obtient

$$\int_y^\infty z^{a-1}\overline{F}(y|x)dz = \frac{\gamma(x)}{1-a\gamma(x)}y^a\overline{F}(y|x)(1+o(1))$$

En remplaçant dans l'équation # (4) et en divisant des deux cotés par : $\frac{1}{1-a\gamma(x)}y^a\overline{F}(y|x)$

on a

$$\frac{\varphi_a(y|x)}{\frac{1}{1-a\gamma(x)}y^a\overline{F}(y|x)} - 1 + a\gamma(x) = a\gamma(x)(1+o(1))$$

ce qui conclut la première partie de la preuve. Deuxièmement, sous la condition (F.2), en dérivant des deux cotés l'équation # (4) on obtient :

$$\varphi'_a(y|x) = y^a\overline{F}'(y|x) = y^{a-1}\overline{F}(y|x)\frac{y\overline{F}'(y|x)}{\overline{F}(y|x)}$$

En utilisant le resultat du théorème de Resnick on obtient :

$$\frac{y\overline{F}'(y|x)}{\overline{F}(y|x)} = \frac{-1}{\gamma(x)}(1+o(1))$$

ce qui conclut la preuve

Comme conséquence de ce Lemme et de l'équation (3.1) on obtient

$$RCTE(\alpha_n|x)/RVaR(\alpha_n|x) \sim \frac{1}{1-\gamma(x)}$$

qui est une extension du résultat sans covariable de Hua et Joe et de Zhu et Li. Le second lemme est aussi d'une nature analytique. Il nous donne un développement au second ordre du RCTM. ■

Lemme 3.8 *Supposons que les hypothèses (F.1), (F.2) et (F.3) soient vérifiées. Soient*

$0 < \beta_n < \alpha_n$ deux suites telles que $\alpha_n \rightarrow 0$ lorsque $n \rightarrow \infty$. Alors

$$|\log RTCM_a(\alpha_n|x) - \log RTCM_a(\beta_n|x) + a\gamma(x) \log(\alpha_n/\beta_n)| = O(\log(\alpha_n/\beta_n) \varepsilon_a \left(\frac{1}{\alpha_n}\right))$$

En se servant de (3.7) et de l'hypothèse (F.2), on a :

Preuve.

$$\log RTCM_a(\alpha_n|x) = -a\gamma(x) \log(\alpha_n) + \log c_a(x) + \int_1^{1/\alpha_n} \frac{\varepsilon_a(u|x)}{u} du$$

Par Conséquent

$$\Delta_n := \log RTCM_a(\alpha_n|x) - \log RTCM_a(\beta_n|x) + a\gamma(x) \log(\alpha_n/\beta_n) = \int_{1/\beta_n}^{1/\alpha_n} \frac{\varepsilon_a(u|x)}{u} du$$

à l'aide de (F.3), on obtient $|\Delta_n| \leq |\varepsilon_a \left(\frac{1}{\alpha_n}\right)| \log(\alpha_n/\beta_n)$, la conclusion en découle

Remarquons que l'estimateur à noyau de l'équation 3 du moment conditionnel peut être se réécrit sous la forme

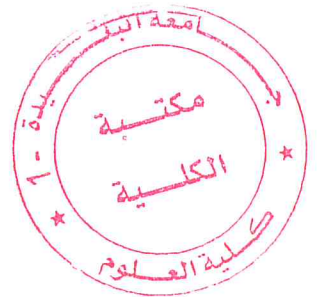
$$\hat{\varphi}_{a,n}(y_n|x) = \hat{\psi}_{a,n}(y_n|x) / \hat{g}_n(x)$$

Où

$$\hat{\psi}_{a,n}(y_n|x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) Y_i^a \mathbb{I}\{Y_i > y\}$$

est un estimateur de $\psi_a(y|x) = \varphi_a(y|x) g(x)$ et $\hat{g}_n(x)$ est l'estimateur à noyau classique de la densité $g(x)$

$$\hat{g}_n(x) = \sum_{i=1}^n K_h(x - X_i)$$



■

Théorème 3. Considérons pour $j \in \{1, \dots, J\}$ les termes suivantes :

$$v_{n,j}(x) = \frac{(1-a_j\gamma(x))(nh^p\alpha_n)^{1/2}}{\gamma(x)\varphi_0^{\leftarrow}(\alpha_n|x)}$$

$$\sigma_{n,j}(x) = \varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x)(nh^p\alpha_n)^{-1/2}$$

$$\sigma_{n,0}(x) = \varphi_0^{\leftarrow}(\alpha_n|x)(nh^p\alpha_n)^{-1/2}$$

$$t_{n,j} = v_{n,j}(x) \left(\varphi_0^{\leftarrow}(\alpha_n|x) - \varphi_{a_j}^{\leftarrow}(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x) \right)$$

$$W_{n,j}(x) = v_{n,j}(x) \left(\left(\hat{\varphi}_{a_j,n}^{\leftarrow} \varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x \right) - \varphi_{a_j}^{\leftarrow}(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x) \right)$$

$$W_{n,j}^{(0)}(x) = v_{n,j}(x) \left(\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x) - \varphi_0^{\leftarrow}(\alpha_n|x) \right)$$

$$W_{n,0}^{(0)}(x) = \sigma_{n,0}^{-1}(x) \left(\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x) - \varphi_0^{\leftarrow}(\alpha_n|x) \right)$$

où $(z_0, z_1, \dots, z_J) \in \mathbb{R}^{J+1}$ Regardons le comportement asymptotique de la fonction de répartition définie par

$$\begin{aligned} & \phi_n(z_0, z_1, \dots, z_J) \\ &= \mathbb{P} \left(\left\{ \bigcap_{j=1}^J \left\{ \sigma_{n,j}^{-1}(x) \hat{\varphi}_{a_j,n}^{\leftarrow} \left(\left(\varphi_{0,n}^{\leftarrow}(\alpha_n|x)|x \right) - \varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) \leq z_j \right) \right\} \right\} \cap \left\{ W_{n,0}^{(0)}(x) \leq z_0 \right\} \right) \\ &= \mathbb{P} \left(\left\{ \bigcap_{j=1}^J \left\{ W_{n,j}(x) - W_{n,j}^{(0)}(x) \leq t_{n,j} \right\} \right\} \cap \left\{ W_{n,0}^{(0)}(x) \leq z_0 \right\} \right) \end{aligned}$$

Concentrons-nous d'abord sur les termes déterministes $t_{n,j}(x)$. D'après le lemme pour $a \in [0, 1/2\gamma(x)[, \varphi_a^{\leftarrow}(\cdot|x)$ est une fonction à variations régulières dérivable telle que

$$(\varphi_a^{\leftarrow})'(y_n|x) = \frac{1}{\varphi_a'(\varphi_a^{\leftarrow}(y_n|x)|x)} = \frac{\gamma(x)\varphi_a^{\leftarrow'}(y_n|x)}{(a\gamma(x)-1)y_n} (1+o(1)) \quad (3.12)$$

lorsque $n \rightarrow \infty$ pour $j \in \{1, \dots, J\}$ un développement de Taylor à l'ordre 1 donne

$$\varphi_{a_j}^{\leftarrow} \left(\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) | x \right) - \varphi_{a_j}^{\leftarrow} \left(\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) + \sigma_{n,j} (x) z_j | x \right) =$$

où

$$q_{n,j} (x) = \left(\varphi_{a_j}^{\leftarrow} \right)' \left(\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) + \theta_{n,j} \sigma_{n,j} (x) z_j | x \right)$$

avec

$(\theta_{n,1}, \dots, \theta_{n,J}) \in]0, 1[^J$. Puisque $\frac{\sigma_{n,j}(x)}{\varphi_{a_j}^{\leftarrow}(\varphi_0^{\leftarrow}(\alpha_n|x)|x)} = (nh^p \alpha_n)^{-1/2} \rightarrow 0$ lorsque $n \rightarrow \infty$, l'équation implique

$$q_{n,j} (x) = \frac{\gamma (x) \varphi_0^{\leftarrow} (\alpha_n | x)}{(a_j \gamma (x) - 1) \varphi_{a_j}^{\leftarrow} (\varphi_0^{\leftarrow} (\alpha_n | x) | x)} (1 + o(1))$$

Ainsi

$$\varphi_{a_j}^{\leftarrow} \left(\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) \right) - \varphi_{a_j}^{\leftarrow} \left(\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) - \sigma_{n,j} (x) z_j | x \right) = \frac{z_j}{v_{n,j} (x)} (1 + o(1)) \quad (3.13)$$

ce qui montre que pour tout $j \in \{1, \dots, J\}$ $t_{n,j} (x) \rightarrow z_j$ lorsque $n \rightarrow \infty$. Concentrons nous sur les termes aléatoires $W_{n,j} (x)$ pour tout $j \in \{1, \dots, J\}$. Il est clair que

$$W_{n,j} (x) = \frac{1 - a_j \gamma (x)}{\gamma (x)} \left(\frac{\hat{\varphi}_{a_j,n}^{\leftarrow} (\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) + \sigma_{n,j} (x) z_j | x)}{\varphi_{a_j}^{\leftarrow} (\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) + \sigma_{n,j} (x) z_j | x)} - 1 \right) (nh^p \alpha_n)^{-1/2} (1 + o(1))$$

D'après l'équation :

$$\frac{\varphi_{a_j}^{\leftarrow} (\varphi_{a_j} \left(\varphi_0^{\leftarrow} (\alpha_n | x) \right) + \sigma_{n,j} (x) z_j | x)}{\varphi_0^{\leftarrow} (\alpha_n | x)} = 1 + \frac{z_j}{\varphi_0^{\leftarrow} (\alpha_n | x) v_{n,j} (x)} (1 + o(1)) = 1 + o(1)$$

De plus on a

$$W_{n,j}^0 (x) = \frac{1 - a_j \gamma (x)}{\gamma (x)} \left(\frac{\hat{\varphi}_{0,n}^{\leftarrow} (\alpha_n | x)}{\varphi_0^{\leftarrow} (\alpha_n | x)} - 1 \right) (nh^p \alpha_n)^{-1/2} (1 + o(1))$$

Comme conséquence il'en résulte en appliquant la proposition (3.6) avec $a_{J+1} = 0$, $\alpha_{n,j} = \varphi_{a_j} (\varphi_0^* (\alpha_n|x) |x) + \sigma_{n,j} (x) z_j$ et $\alpha_{n,J+1} = \alpha_n$ ■

3.7 Validation sur simulation

Dans cette partie, on teste la procédure précédente sur deux lois à queues lourdes conditionnelles : une loi de Fréchet et une loi de Burr.

La fonction de survie conditionnelle de la loi de Fréchet est donnée par :

$$\bar{F}(y|x) = 1 - \exp(-y^{\frac{-1}{\gamma(x)}}) \quad \text{pour } y > 0$$

La *RCTE* associée peut s'écrire :

$$RCTE(\beta|x) = \frac{1}{\beta} \int_0^{RVaR(\beta|x)^{\frac{-1}{\gamma(x)}}} t^{\gamma(x)} \exp(t) dt$$

avec

$$RVaR(\beta|x) = (-\log(1 - \beta))^{-\gamma(x)}$$

La fonction de survie conditionnelle de la loi de Burr choisie est donnée par :

$$\bar{F}(y|x) = \frac{1}{1 + y^{\frac{1}{\gamma(x)}}} \quad \text{pour } y > 0$$

La *RCTE* associée peut s'écrire :

$$RCTE(\beta|x) = \mathcal{I} \left(\frac{RVaR(\beta|x)^{\frac{-1}{\gamma(x)}}}{RVaR(\beta|x)^{\frac{-1}{\gamma(x)}} + 1}, 1 - \gamma, 1 + \gamma \right)$$

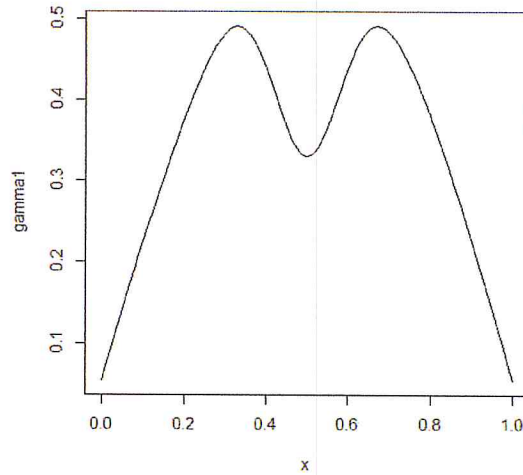


FIG. 3-1 – Graphe de la fonction de l'indice de queue conditionnel

avec

$$RVaR(\beta|x) = \beta^{-\gamma(x)} (1 - \beta)^{\gamma(x)}$$

où

$$\mathcal{I}(r, p, q) = \frac{\beta(r, p, q)}{\beta(p, q)} \quad \text{avec} \quad \beta(r, p, q) = \int_0^r w^{p-1} (1-w)^{q-1} dw$$

étant la fonction Beta incomplète et $B(p, q)$ la fonction Beta.

Pour notre validation sur simulation, on a choisi l'indice de queue conditionnel suivant :

$$\gamma : x \in [0, 1] \rightarrow \gamma(x) = 1/2 \left(\frac{1}{10} + \sin \pi x \right) \left(\frac{11}{10} - \exp \left(-64 \left(x - \frac{1}{2} \right)^2 \right) \right)$$

Si l'on garde en mémoire les restrictions sur $\gamma(x)$ dues à nos résultats asymptotiques (voir Théorème 3), ce choix de fonction rend l'estimation plus difficile. En effet, on peut remarquer que $\gamma(x)$ est proche de $1/2$ pour $x = 0.3$ et $x = 0.7$.

La fonction $\gamma(x)$ choisie étant définie sur l'intervalle unité $[0, 1]$ il nous faut normaliser les vecteurs que l'on utilisera comme covariable. La normalisation d'un vecteur positif $v = (v_i)_{i \in \{1, \dots, n\}}$ et un vecteur $w = (w_i)_{i \in \{1, \dots, n\}} \in [0, 1]$, est obtenue à l'aide de la transformation

$$w_i = \frac{v_i - \min(v_1, \dots, v_n)}{\max(v_1, \dots, v_n) - \min(v_1, \dots, v_n)} \text{ pour } i \in \{1, \dots, n\}$$

3.8 Conclusion

Notre travail a porté sur l'estimation non-paramétrique des quantiles extrêmes conditionnels. Tout d'abord le premier chapitre a présenté une synthèse sur les connaissances actuelles des lois à queues lourdes ainsi que des notions sur les estimateurs des paramètres des lois de valeurs extrêmes. Dans une deuxième partie nous avons élargi notre étude en donnant des définitions sur les quantiles extrêmes conditionnels et leur méthodes d'estimation et à la fin quelques applications et discussions ont été exposées.

à la fin de notre travail, on a donné une définition au mesure de risque et on a étudié leur comportement asymptotique en proposant deux lois à queues lourdes

Bibliographie

- [1] Resnick(1987) Generalized jackknife moment estimator of the tail index. Bulletin of the International Statistical Institute, 58(1) :401–402
- [2] embrechts et al.(1997) Modelling Extremal Events for Insurance and Finance. Springer Verlag
- [3] Pickands (1975) Statistical inference using extreme order statistics. Annals of Statistics, 3(1) :119–131.
- [4] Hill (1975)A simple general approach to inference about the tail of a distribution. Annals of Statistics, 3(5) :1163–1174
- [5] Weisman(1978) Estimation of parameters and large quantiles based on the klargest observations. Journal of the American Statistical Association, 73(364) :812–815.
- [6] Rényi (1953) On the theory of order statistics. Acta Mathematica Hungarica, 4(3–4) :191–231.
- [7] Guida et Longo (1988) Estimation of probability tails based on generalized extreme value distributions. Reliability Engineering and System Safety, 20(3) :219–242.
- [8] weinstein(1973) Theory and application of some classical and generalized asymptotic distributions of extreme values. IEEE Transactions on Information Theory, 19(2) :148–154.
- [9] Prescott et Walden(1980) Maximum likelihood estimation of the parameters of generalized extreme-value distribution. Biometrika, 67(3) :723–724.

- [10] Prescott et Walden(1983) Maximum likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution for censored samples. *Journal of Statistical Computation and Simulation*, 16(3–4) :241–250.
- [11] Hosking et el (1985) Estimation of the generalized extreme-value distribution by the method of probability-weighted comments. *Technometrics*, 27 :251–261.
- [12] Smith(1985) Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1) :67–92.
- [13] Balkema et de Haan (1974) Residual life time at a great age. *Annals of Probability*, 2(5) :792–804.
- [14] Pickands (1975) Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1) :119–131.
- [15] Davison et Smith (1990) Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, B*, 52(3) :393–442.
- [16] Breiman et al.(1990) Robust confidence bounds for extreme upper quantiles. *Journal of Statistical Computation and Simulation*, 37(3–4) :127–149.
- [17] Deckers et De Haan (1989) On testing the extreme value index via the Pot-method. *Annals of Statistics*, 23(6) :2013–2035.
- [18] Beirlant et al(2005) Estimation of the extreme value index and regression on generalized quantile plots. *Annals of Statistics*, 11(6) :949–970.
- [19] dieblot et al(2008) Bias-reduced estimators of the weibull tail-coefficient. *Test*, 17(2) :311–331.
- [20] Gomes et oliveira(2001) The bootstrap methodology in statistics of extremes theory and applications - choice of the optimal sample fraction. *Extremes*, 4(4) :331–358.
- [21] Viharos(1993) Asymptotic distributions of linear combinations of extreme values. *Acta Scientiarum Mathematicarum*, 58 :211–231.

- [22] Viharos(1995) Limit theorems for linear combinations of extreme values with applications to inference about the tail of a distribution. *Acta Scientiarum Mathematicarum*, 60 :761–777.
- [23] Gomes et al(2005) Revisiting the role of the jackknife methodology in the estimation of a positive tail index. *Communications in Statistics– Theory and Methods*, 34(2) :319–335.
- [24] Tsourti et Panaretos(2001) simulation study on the performance of extreme value index estimators and proposed robustifying modifications. 5th Hellenic European Conference on Computer Mathematics and its Applications, Athens, Greece, 2 :847–852.
- [25] Tsourti et Panaretos(2003) Stochastic musings : Perspectives from the pioneers of the late 20th century. pages 141–160. Laurence Erlbaum, John Panaretos édition.
- [26] Mason(1982) Laws of large numbers for sums of extreme values. *Annals of Probability*, 10(3) :754–764.
- [27] deheuvels et al(1988) Almost sure convergence of the Hill estimator. *Mathematical Proceedings of the Cambridge Philosophical Society*, 104(2) : 371–381.
- [28] Drees(1995) Refined Pickands estimator of the extreme value index. *Annals of Statistics*, 23(6) :2059–2080
- [29] Kratz et Resnick (1996) The qq-estimator and heavy tails. *Stochastic Models*, 12(4) :699–724
- [30] Schultze et Steinebach(1996) On least squares estimates of an exponential tail coefficient. *Statistics and Decisions*, 14(3) :353–372.
- [31] Royston et Wright(1998) How to construct normal ranges for fetal variables. *Ultrasound in Obstetrics and Gynecology*, 11(1) :30–38.
- [32] Roussas(1969) Nonparametric estimation of the transition distribution function of Markov process. *Annals of Mathematical Statistics*, 40(4) :1386–1400.
- [33] stute(1986) Conditional empirical processes. *Annals of Statistics*, 14(2) :638–647.

- [34] berlinet et al(1998) Propriétés asymptotiques des estimateurs des quantiles. Comptes Rendus de l'Académie des Sciences, 326(5) :611–614.
- [35] Stone(1977) Consistent nonparametric regression. Annals of Statistics, 5(4) :595–645.
- [36] Collomb (1980) Estimation non paramétrique de probabilités conditionnelles. Comptes Rendus de l'Académie des Sciences, 291 :427–430.
- [37] Horvath et Yandall (1988) Asymptotics of conditional empirical processes. Journal of Multivariate Analysis, 26(2) :184–206.
- [38] Gannoun(1989) Estimation de la médiane conditionnelle. Thèse de doctorat, Université de Paris VI.
- [39] Hart(1991) Comment to “choosing a kernel regression estimator”. Statistical Sciences, 6(4) :425–427.
- [40] Ferraty et Vieu (2000) Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. Comptes Rendus de l'Académie des Sciences, 330 :403–406.
- [41] Ferraty et Vieu (2006) Nonparametric Functional Data Analysis : Theory and Practice. Springer Series in Statistics, Springer.
- [42] Antoch et Janssen(1989) Nonparametric regression m-quantiles. Statistics and Probability Letters, 8(4) :355–362.
- [43] Gasser et Huller(1984) Robust nonparametric function fitting. Journal of the Royal Statistical Society, B, 46(1) :42–51.
- [44] Parzen(1979) Nonparametric estimation statistical data modeling. Journal of the American Statistical Association, 74(365) :105–131.

