



N° Attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--



Année univ.: 2017/2018

Estimation non-paramétrique des quantiles extrêmes conditionnels

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saïda - Dr Moulay Tahar

Discipline : MATHÉMATIQUES

Spécialité : Analyse stochastique, statistique des processus et
applications (ASSPA)

par

Mlle. Bouazza Imene¹

Sous la direction de

Dr Mme F. Benziadi

Soutenu le 20/06/2018 devant le jury composé de

Dr. S. Idrissi	Université Dr Tahar Moulay - Saïda	Présidente
Dr. F. Benziadi	Université Dr Tahar Moulay - Saïda	Encadreur
Dr. R. Rouane	Université Dr Tahar Moulay - Saïda	Examinatrice
Dr. N. Hachemi	Université Dr Tahar Moulay - Saïda	Examinatrice

1. e-mail : imeneminou88@gmail.com

Remerciements

En tout premier lieu, je remercie le bon Dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés

Le moment est venu pour moi d'exprimer ma plus gratitude envers tous celles et ceux qui m'ont aidé et encouragé dans l'accomplissement de ce travail.

Je tiens tout d'abord à remercier Madame **F.BENZIADI** pour m'avoir encadré durant cette année de mémoire. Je suis vraiment très heureuse d'avoir eu cette chance de travailler avec vous. Au cours de cette année, vous avez toujours su me guider avec vos conseils et m'encourager avec beaucoup de patience. Votre gentillesse m'a infiniment aidé. Je ne pourrai jamais assez remercier.

Je désire aussi remercier les professeurs, qui m'ont fourni les outils nécessaires à la réussite de mes études universitaires. Je tiens à remercier spécialement **Mlle. IDRISSI Soumia** de m'avoir honorée en acceptant de présider le jury, je tiens aussi à remercier **Mme. ROUANE Rachida** et **Mlle. HACHEMI Nawel** qui me font le grand honneur d'évaluer ce travail.

Je profite aussi de l'occasion pour remercier tous les membres de laboratoire des Modèles Stochastiques, Statistique et Applications.

Enfin, mes remerciements les plus chaleureux vont à tous mes proches et collègues qui m'ont apporté leur support moral et intellectuel tout au long de ma démarche.

Dédicaces

*Au nom du dieu le clément et le miséricordieux louange à **ALLAH** le tout puissant.*

Je dédie ce modeste travail en signe de respect, reconnaissance et de remerciement à :

Mon père Kada et ma mère Mokhtaria qui m'ont indiqué le bon chemin à entreprendre et qui m'ont encouragé et soutenue tout au long de mon parcours quotidien.

Ma sœur Nadjat et mon frère Sid Ahmed.

Ma grande mère maternelle Fatna, que mon Dieu la préserve, mon oncle et ma tante .

Une spéciale dédicace à mes adorables amies BOUDAA Sarah et BEKHTI Naima qui m'ont encouragé le long de mon parcours universitaire.

Toutes mes chers amis(es) et mes camarades de la promotion de A.S.S.P.A 2017 /2018, chacun à son nom.

Tous mes enseignants et enseignantes de département de mathématique et informatique.

Table des matières

Introduction générale	5
1 Présentation de la théorie des valeurs extrêmes	8
1.1 Statistiques d'ordre	8
1.1.1 La loi du maximum	9
1.2 Lois des valeurs extrêmes	9
1.3 Domaines d'attraction	14
1.3.1 Fonctions à variations régulières	14
1.3.2 Domaine d'attraction de Fréchet	16
1.3.3 Domaine d'attraction de Weibull	17
1.3.4 Domaine d'attraction de Gumbel	18
1.4 Estimation de l'indice des valeurs extrêmes	19
1.4.1 Estimateur de Pickands	19
1.4.2 Estimateur de Hill	20
1.5 Le choix du nombre de statistiques d'ordre	22
1.5.1 Méthode graphique	22
1.5.2 Méthode analytique	23
2 Estimation de quantiles conditionnels	24
2.1 Quantile et Quantile Conditionnel	24
2.2 Estimation paramétrique des quantiles conditionnels	25
2.3 Estimation non-paramétrique des quantiles conditionnels	26
2.3.1 Méthode d'estimation indirecte	26
2.3.2 Méthode d'estimation directe	35

3 Estimation de quantiles extrêmes conditionnels	39
3.1 Estimation des quantiles extrêmes	39
3.1.1 L'estimation des quantiles extrêmes par la loi des valeurs extrêmes	42
3.1.2 L'estimation des quantiles extrêmes par la méthode des excès	44
3.1.3 L'estimation semi-paramétrique des quantiles extrêmes	47
3.2 Estimation de quantiles extrêmes conditionnels	48
3.2.1 Modèle	48
3.3 La construction des estimateurs	49
3.4 Quelques résultats asymptotiques	51
3.4.1 Hypothèses	51
3.4.2 Résultats préliminaires	52
3.4.3 La normalité asymptotique des estimateurs	53
3.5 L'estimation de l'indice de queue conditionnel	60
Conclusion	62

Introduction générale

La statistique non paramétrique connaît un grand essor chez de nombreux auteurs et dans différents domaines. En effet, celle-ci possède un champ d'application très large permettant, ainsi, l'explication de certains phénomènes mal modélisés jusqu'à présent, tels que les séries chronologiques, et prédire les réalisations futures.

Une branche principale dans la statistique est *l'estimation de quantiles conditionnels* qui fut traitée pour la première fois par **Roussas** (1969) [42] dans le cas de processus de Markov, en utilisant un estimateur de la fonction de répartition conditionnelle défini comme étant l'intégrale de l'estimateur de la densité conditionnelle. **Stone** (1977) [47] a obtenu la convergence en probabilité d'un estimateur basé sur l'estimation empirique de la fonction de répartition conditionnelle. **Samanta** (1989) [44] a établi la normalité asymptotique et la convergence uniforme de l'estimateur à noyau de quantiles conditionnels dans le cas i.i.d. Nous renvoyons à **Berlinet et al** (2001) [5], pour une étude générale de la normalité asymptotique de quantiles conditionnels.

En pratique, les domaines d'utilisation des quantiles conditionnels sont assez variés. En biologie, sont utilisés pour estimer des courbes de référence permettant d'analyser certaines propriétés biophysiques de la peau. En climatologie où l'on étudie comment les changements de climats au cours des années peuvent influencer les températures ou chutes de pluies. Les quantiles représentent également un moyen robuste de prévision. En pratique, ces quantiles sont calculés suivant un critère d'ordre sur les observations.

Au cours des dernières années, nous avons pu observer dans la recherche scientifique, une modélisation des événements rares (extrêmes). Ces derniers sont des événements dont la probabilité d'apparition est très faible c'est-à-dire : se trouve dans les queues des distributions. Ils sont dits extrêmes quand il s'agit de valeurs beaucoup plus grandes ou plus petites que celles observées habituellement, nous pouvons mentionner quelques catastrophes naturelles : inondations, séismes de forte intensité, vents violents, crues de rivière inhabituelles, ou bien les risques financiers (les kraks boursiers, les crises financières).

La théorie des valeurs extrêmes est une branche de la statistique qui essaie d'amener une solution face à ces phénomènes. Elle se repose principalement sur des distributions limites des extrêmes et leurs domaines d'attraction. Cependant, on y retrouve deux modèles : loi généralisée des extrêmes (GEV) et loi de Paréto généralisée (GPD). Ainsi, les auteurs **Fisher** et **Tippet** (1928) [26] puis plus tard **Gnedenko** (1943) [30] ont énoncé un théorème fondamental avec la création de trois domaines d'attraction : domaine d'attraction de Fréchet, Gumbel et Weibull. **Von Mises** (1936) [52] puis **Jenkinson** (1955) [36] ont rassemblé les distributions de ces trois domaines en une seule écriture. C'est en ce moment que plusieurs auteurs se sont focalisés aux estimations de l'indice des valeurs extrêmes. Ainsi que l'estimation de quantiles extrêmes qui est requise dans de nombreux domaines d'application parmi lesquels citons la fiabilité, la finance (**Embrechts et al** (1997) [21]), les assurances et la climatologie.

Dans la littérature, on dénombre plusieurs méthodes d'estimation des quantiles extrêmes conditionnels telles que : *l'approche paramétrique*, Les premiers travaux sont dus à **Smith** (1987) [46] qui propose de modéliser les maxima par une loi des valeurs extrêmes dont les paramètres sont des fonctions de la covariable. Une forme paramétrique est supposée sur ces fonctions et l'estimation est effectuée par maximum de vraisemblance ou par moindres carrés. Ce fut ensuite au tour de **Davison** et **Smith** (1990) [17] de proposer de modéliser les excès par une loi de Pareto généralisée dont les paramètres sont des fonctions de la covariable, les auteurs

estimant ces derniers par maximum de vraisemblance. *L'approche semi-paramétrique*, **Beirlant** et **Goegebeur** (2003) [4] ont été les premiers à utiliser une approche semi-paramétrique. Ils proposent de transformer tout d'abord les données dans le but d'obtenir des résidus suivant une loi de type Pareto puis de les utiliser dans un modèle de régression exponentielle où les paramètres de ce dernier sont estimés par la méthode du maximum de vraisemblance.

À la suite de ces travaux, les premiers à proposer *une approche non-paramétrique* sont **Davison** et **Ramesh** (2000) [16], qui modélisent les valeurs extrêmes conditionnelles où la covariable est le temps dont les paramètres sont estimés par la méthode non-paramétrique. Par contre, **Beirlant** et **al** (2001) [5] utilisent les statistiques d'ordre pour estimer les quantiles extrêmes conditionnels avec la méthode des polynômes locaux. Enfin, **Gardes** et **Girard** (2010) [28], **Gardes** et **al.** (2010) [29] ont aussi proposé des estimateurs de quantiles extrêmes conditionnels.

Ce travail a pour objectif d'estimer non-paramétriquement les quantiles extrêmes conditionnels.

Ce mémoire est partagé en trois chapitres. Dans le premier chapitre, Pour faciliter la lecture on regroupe des définitions et des résultats fondamentaux sur la théorie des valeurs extrêmes dans le cas univarié réel. Dans le deuxième chapitre, on présente quelques méthodes d'estimation paramétrique et non-paramétrique du quantile conditionnel et on cite quelques résultats asymptotiques de ces estimateurs. le troisième chapitre est composé de deux parties : dans la première partie, on présente trois approches pour estimer le quantile extrême : l'approche par la loi des valeurs extrêmes, par la méthode des excès et l'approche semi-paramétrique. dans la deuxième partie, nous nous intéressons à l'estimation de quantiles extrêmes conditionnels et de l'indice de queue conditionnel. Et on finalise ce travail par une conclusion.

Chapitre 1

Présentation de la théorie des valeurs extrêmes

Dans ce chapitre, on présente de façon très classique quelques concepts fondamentaux sur la statistique des valeurs extrêmes. Pour commencer, il faut avoir un grand bagage, alors notre point de départ sera les statistiques d'ordre.

Soit (X_1, \dots, X_n) une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d) de fonction de répartition F . Notons $\bar{F} = 1 - F$ la fonction de survie (ou de queue).

1.1 Statistiques d'ordre

Définition 1.1.1. *Notons par $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ le réarrangement croissant de l'échantillon $\{X_i, i = 1, \dots, n\}$.*

Pour tout $i = 1, \dots, n$ le vecteur $(X_{1,n}, \dots, X_{n,n})$ est appelé l'échantillon ordonné associé à l'échantillon (X_1, \dots, X_n) et la variable aléatoire $X_{i,n}$ étant la i -ième statistique d'ordre (ou statistique d'ordre i).

Dans un échantillon de taille n , deux statistiques d'ordre sont particulièrement intéressantes pour l'étude des événements extrêmes : ce sont les deux statistiques

extrêmes, autrement dit le minimum et le maximum de l'échantillon définis par :

$$X_{1,n} = \min(X_1, \dots, X_n) \quad \text{et} \quad X_{n,n} = \max(X_1, \dots, X_n),$$

et leur écart $D_n = X_{n,n} - X_{1,n}$ est dite déviation extrême.

1.1.1 La loi du maximum

Posons $M_n = X_{n,n} = \max(X_1, \dots, X_n)$. En utilisant la propriété d'indépendance des variables aléatoires X_1, \dots, X_n , nous en déduisons que la fonction de répartition du maximum est donnée par :

$$\begin{aligned} F_{M_n}(x) &= \mathbb{P}(M_n \leq x) \\ &= \mathbb{P} \left[\bigcap_{i=1}^n (X_i \leq x) \right] \\ &= \prod_{i=1}^n \mathbb{P}(X_i \leq x) \\ &= [F_X(x)]^n. \end{aligned}$$

La fonction de répartition de X n'étant pas toujours connue, il n'est généralement pas possible de déterminer la distribution du maximum à partir de ce résultat. On s'intéresse alors à la distribution asymptotique du maximum en faisant tendre n vers l'infini. On a :

$$\lim_{n \rightarrow \infty} F_{M_n}(x) = \lim_{n \rightarrow \infty} (F_X(x))^n = \begin{cases} 0 & \text{si } F(x) < 1, \\ 1 & \text{si } F(x) = 1. \end{cases} \quad (1.1)$$

Remarque 1.1.1. *Tous les résultats développés par le maximum peuvent être transposés pour le minimum en utilisant la formule mathématique suivante :*
 $\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n)$.

1.2 Lois des valeurs extrêmes

Le résultat fondamental de la théorie des valeurs extrêmes repose sur le théorème de **Fisher** et **Tippet** (1928) [26] avant que **Gnedenko** (1943) [30] n'obtienne

rigoureusement la convergence. Les travaux de **Von Mises** (1936) [52] et **Jenkinson** (1955) [36] ont permis de donner une forme unifiée à ce résultat. Les applications ont commencé suite aux travaux de **Gumbel** (1954) [32], en particulier en hydrologie.

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

$$F_1(ax + b) = F_2(x)$$

Le théorème suivant montre que s'il existe des suites normalisantes $(a_n)_{n \geq 1} > 0$, $(b_n)_{n \geq 1} \in \mathbb{R}$ et admettent une limite, alors la suite de variables aléatoires $(a_n^{-1}(X_{n,n} - b_n))_{n \geq 1}$ converge en loi vers une limite non dégénérée. Ce théorème est la base de la théorie des valeurs extrêmes.

Théorème 1.2.1. (*Fisher et Tippett* (1928) [26]; *Gnedenko* (1943) [30]) Sous certaines conditions de régularité sur la fonction de répartition F , il existe un paramètre réel γ et deux suites normalisantes réelles $(a_n)_{n \geq 1} > 0$ et $(b_n)_{n \geq 1} \in \mathbb{R}$ tels que pour tout $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{X_{n,n} - b_n}{a_n} \leq x \right] = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \mathcal{H}_\gamma(x) \quad (1.2)$$

Alors \mathcal{H}_γ est non dégénérée¹ et elle est soit de type :

$$\text{type 1 : loi de Fréchet } \mathcal{H}_\gamma(x) = \begin{cases} 0 & \text{si } x < 0 \\ \exp[-(x)^{-1/\gamma}] & \text{si } x \geq 0 \end{cases} \quad \gamma > 0$$

$$\text{type 2 : loi de Weibull } \mathcal{H}_\gamma(x) = \begin{cases} \exp[-(-x)^{1/\gamma}] & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases} \quad \gamma < 0$$

$$\text{type 3 : loi de Gumbel } \mathcal{H}_0(x) = \exp[-\exp(-x)] \quad \text{si } x \in \mathbb{R} \text{ et } \gamma = 0$$

et $\mathcal{H}_\gamma(\cdot)$ la fonction de répartition de la **loi des valeurs extrêmes (EVD)**. La densité associée est représentée Figure 1.2 pour différentes valeurs de γ .

1. Non centrée sur une valeur.

Ces trois lois caractérisent chacune un comportement spécifique des valeurs extrêmes. On les distingue au travers de deux caractéristiques :

- la vitesse de décroissance de la probabilité des grandes valeurs. Si la décroissance est rapide, de type exponentielle, alors la loi des extrêmes est la loi de Gumbel.
- l'étendue de variation des données. Si celle-ci est bornée, alors la loi des extrêmes est la loi de Weibull, si ce n'est pas le cas, c'est la loi de Fréchet.

Une démonstration détaillée du théorème 1.2.1 est trouvée dans **Resnick** (1987) [41].

Remarque 1.2.1. *Ce théorème présente un intérêt important, car si l'ensemble des distributions est 'large' alors celui des distributions des valeurs extrêmes est plutôt 'restreint'. S.Coles (2001) fait le parallèle entre le théorème centrale limite et celui de Fisher-Tippet, il trouve que la suite a_n joue le rôle de $n^{-1/2}\sigma(X)$ où $\sigma(X)$ désigne l'écart type de X et la suite b_n joue le rôle de l'espérance. La suite a_n (resp. b_n) s'interprète comme un paramètre d'échelle (resp. un paramètre de position ou de centrage). De plus, les suites a_n et b_n ne sont pas uniques.*

Exemple 1.2.1. *Donnons n variables aléatoires X_1, X_2, \dots, X_n i.i.d de fonction de répartition commune F supposée suit une loi exponentielle de paramètre 1, c'est à dire*

$$F(x) = 1 - \exp(-x).$$

Soit F_Y la fonction de répartition de la variable $Y = M_n - \log(n)$, alors :

$$\begin{aligned} F_Y(x) &= \mathbb{P}[M_n - \log(n) \leq x] \\ &= \mathbb{P}[M_n \leq x + \log(n)] \\ &= (F(x + \log(n)))^n \\ &= (\mathbb{P}[X \leq x + \log(n)])^n \\ &= (1 - \exp(-x - \log(n)))^n \\ &= \left(1 - \frac{1}{n} \exp(-x)\right)^n \\ &= \exp\left(n \log\left(1 - \frac{1}{n} \exp(-x)\right)\right) \\ &\simeq \exp(-\exp(-x)) \quad \text{quand } n \rightarrow \infty \\ &= \mathcal{H}_0(x). \end{aligned}$$

Prenons $a_n = 1$ et $b_n = \log(n)$, alors $\frac{(M_n - b_n)}{a_n}$ tend asymptotiquement vers la loi de Gumbel.

Exemple 1.2.2. La Figure suivante illustre dans le cas d'une loi normale centrée réduite, la convergence de la suite de variables aléatoires $(a_n^{-1}(X_{n,n} - b_n))_{n \geq 1}$ en loi vers une limite non dégénérée \mathcal{H}_0 . Dans cet exemple, nous avons utilisé les suites de normalisation théoriques associées à la loi normale standard :

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

Par suite et comme précédemment on déduit que le maximum normalisé tend asymptotiquement vers la loi de Gumbel.

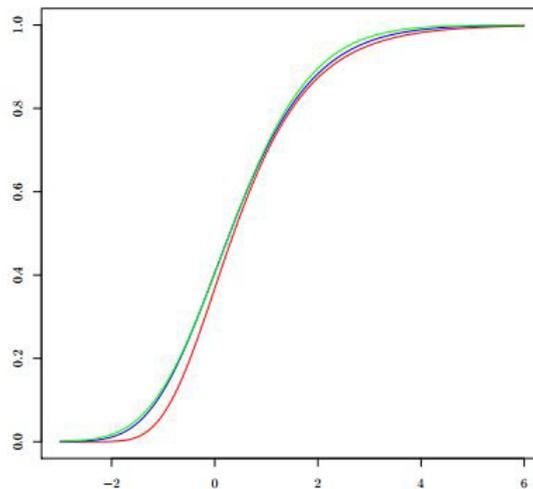


Figure 1.1- Illustration de la théorie des valeurs extrêmes sur une loi normale centrée réduite. Comparaison entre $\mathcal{H}_0(x)$ (rouge), $\mathbb{P} \left[\frac{X_{n,n} - b_n}{a_n} \leq x \right]$ avec $n = 100$ (bleu) et $\mathbb{P} \left[\frac{X_{n,n} - b_n}{a_n} \leq x \right]$ avec $n = 10$ (vert).

Pour plus d'exemples de suites de normalisation de chaque loi se référer à (Embrechts et al (1997) [21]).

Dans certains cas (en particulier pour des problèmes statistiques), il est préférable d'avoir une seule distribution qui unifie les trois précédentes. Pour cela **Von Mises** (1936) [52] et **Jenkinson** (1955) [36] ont proposé la distribution $\text{GEV}(\mu, \sigma, \gamma)$ suivante :

Définition 1.2.2. *La représentation de Jenkinson-von Mises de la loi EVD que l'on appelle **loi des valeurs extrêmes généralisée** notée (**GEVD**) ou (**GEV**) a pour tout x la fonction de répartition suivante :*

$$\Lambda_\gamma(x) = \begin{cases} \exp[-(1 + \gamma x)^{-1/\gamma}] & \text{si } \gamma \neq 0; \quad \text{avec } 1 + \gamma x > 0. \\ \exp[-\exp(-x)] & \text{si } \gamma = 0. \end{cases}$$

Ici, nous prenons le cas où le paramètre de position $\mu = 0$ et le paramètre de dispersion $\sigma = 1$ pour obtenir une forme standard.

Définition 1.2.3. *Le paramètre γ du Théorème 1.2.1 ou de la Définition 1.2.2 est un paramètre de forme que l'on appelle «indice des valeurs extrêmes» ou «indice de queue».*

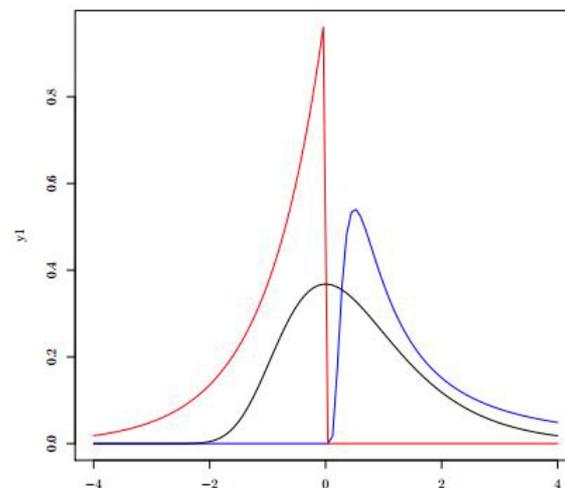


Figure 1.2- Exemple de densités associées à la loi des valeurs extrêmes avec $\gamma = -1$ (rouge), $\gamma = 0$ (noir) et $\gamma = 1$ (bleu).

Si F vérifie le Théorème 1.2.1, on dit alors que F appartient au «domaine d'attraction» de \mathcal{H}_γ ou Λ_γ et selon le signe de γ , on a trois domaines d'attraction.

1.3 Domaines d'attraction

La recherche du domaine d'attraction peut être considérée comme l'étude réciproque de la recherche de la distribution des valeurs extrêmes associée éventuellement à une distribution.

Définition 1.3.1. *On appelle domaine d'attraction de \mathcal{H}_γ (ou domaine d'attraction maximal) l'ensemble des lois F pour lesquelles le maximum normalisé suit la loi \mathcal{H}_γ .*

La caractérisation de ces domaines fait intervenir la notion de fonctions à variations régulières.

1.3.1 Fonctions à variations régulières

Définition 1.3.2. *Une fonction $U(\cdot)$ positive, mesurable est dite à variations régulières à l'infini d'indice $\delta \in \mathbb{R}$ (et on la note par $U(\cdot) \in \mathcal{RV}_\delta$) si pour tout $\lambda > 0$ et $x > 0$, on a :*

$$\lim_{x \rightarrow \infty} \frac{U(\lambda x)}{U(x)} = \lambda^\delta. \quad (1.3)$$

La notion de fonctions à variations régulières à l'infini est liée à celle de fonctions à variations lentes à l'infini.

Définition 1.3.3. *Une fonction $L(\cdot)$ mesurable est dite à variations lentes à l'infini si et seulement si, pour tout $\lambda > 0$, on a :*

$$\lim_{x \rightarrow \infty} \frac{L(\lambda x)}{L(x)} = 1. \quad (1.4)$$

Exemple 1.3.1. $f(x) = \log x$

$$\lim_{x \rightarrow \infty} \frac{\log(\lambda x)}{\log(x)} = \lim_{x \rightarrow \infty} \left(1 + \frac{\log \lambda}{\log x} \right) = 1.$$

La fonction \log est à variations lentes à l'infini.

- Remarque 1.3.1.**
1. Une fonction à variations régulières à l'infini d'indice $\delta = 0$ est une fonction à variations lentes à l'infini, et on notera par $L(\cdot) \in \mathcal{RV}_0$.
 2. Si $\delta \rightarrow \infty$, on dit que la fonction $U(\cdot)$ est à variation rapide.
 3. On montre facilement que toute fonction à variations régulières d'indice $\delta \in \mathbb{R}$ peut s'écrire sous la forme :

$$U(x) = x^\delta L(x), \quad L \in \mathcal{RV}_0. \quad (1.5)$$

Et ce résultat montre que l'étude des fonctions à variations régulières à l'infini se ramène à l'étude des fonctions à variations lentes à l'infini.

On peut caractériser les fonctions à variations lentes de manière plus précise à l'aide du théorème suivant :

Théorème 1.3.1. *Toute fonction à variations lentes à l'infini $L(\cdot)$ peut s'écrire sous la forme :*

$$L(x) = c(x) \exp \left(\int_1^x \frac{\Delta(u)}{u} du \right),$$

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

$$\lim_{x \rightarrow \infty} c(x) = c > 0 \quad \text{et} \quad \lim_{x \rightarrow \infty} \Delta(x) = 0.$$

Cette représentation des fonctions à variations lentes est connue sous le nom de *représentation de Karamata*. De plus, si la fonction $c(\cdot)$ est constante, la fonction $L(\cdot)$ est dite normalisée. De nombreux résultats sur les fonctions à variations régulières sont donnés dans le livre de **Bingham, Goldie et Teugels** (1987) [7].

Le lemme suivant nous renseigne sur l'inverse d'une fonction à variations régulières.

Lemme 1.3.1. *(Inverse d'une fonction à variations régulières)*

- Si U est à variations régulières d'indice $\delta > 0$, alors $U^\leftarrow(x)$ est à variations régulières d'indice $1/\delta$.

– Si U est à variations régulières d'indice $\delta < 0$, alors $U^{\leftarrow}(1/x)$ est à variations régulières d'indice $-1/\delta$.

Pour une preuve du Lemme 1.3.1, on pourra se référer au Théorème 1.5.12 de l'ouvrage de **Bingham, Goldie et Teugels** (1987) [7] ou à la Proposition 2.6 du livre de **Resnick** (1987) [41].

Avant de donner les conditions pour que F appartienne à l'un des trois domaines d'attraction ainsi que ses suites de normalisation (a_n) et (b_n) , il convient d'introduire une définition utile.

Définition 1.3.4. (Point terminal) On appelle point terminal (en anglais « upper endpoint ») de la fonction F , le réel x_F (i.e. la plus grande valeur possible pour $X_{i,n}$ peut prendre ∞) défini par

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

avec la convention $\sup\{\emptyset\} = \infty$.

1.3.2 Domaine d'attraction de Fréchet

Le domaine d'attraction de Fréchet contient les lois dont la fonction de survie est à décroissance polynomiale, (i.e) les lois à queues lourdes.

Le résultat ci-dessous énoncé par **Gnedenko** (1943) [30] et dont on trouvera une démonstration simple dans le livre de (**Resnick** (1987) [41], Proposition 1.11) assure que toute fonction appartenant au domaine d'attraction de Fréchet est une fonction à variations régulières et inversement.

Théorème 1.3.2. $F \in \mathcal{D}(\text{Fréchet})$ avec un indice des valeurs extrêmes $\gamma > 0$ si et seulement si $x_F = +\infty$ et la fonction \bar{F} est à variations régulières d'indice $-1/\gamma$ qui s'écrit $\bar{F} \in \mathcal{RV}_{-1/\gamma}$, (i.e) :

$$\forall \lambda > 0, \quad \lim_{x \rightarrow \infty} \frac{\bar{F}(\lambda x)}{\bar{F}(x)} = \lambda^{-1/\gamma}.$$

Autrement dit, une fonction de répartition F appartenant au domaine d'attraction de Fréchet s'écrit sous la forme :

$$F(x) = 1 - x^{-1/\gamma}L(x) \quad \text{avec} \quad L \in \mathcal{RV}_0.$$

Dans ce cas, un choix possible pour les suites de normalisation (a_n) et (b_n) est :

$$a_n = \overline{F}^{\leftarrow}(1/n) = q_{1/n} \quad \text{et} \quad b_n = 0.$$

De ce théorème, on en déduit que $F \in \mathcal{D}(\text{Fréchet})$ si et seulement si le point terminal x_F est infini et $\overline{F}(x) = x^{-1/\gamma}L(x)$, où $L(\cdot)$ est une fonction à variations lentes à l'infini et γ un réel strictement positif.

1.3.3 Domaine d'attraction de Weibull

Toutes les lois appartenant au domaine d'attraction de Weibull ont un point terminal x_F fini. Le résultat suivant (voir **Gnedenko** (1943) [30], (**Resnick** (1987) [41], Proposition 1.13)) montre que l'on passe du domaine d'attraction de Fréchet à celui de Weibull par un simple changement de variable dans la fonction de répartition.

Théorème 1.3.3. $F \in \mathcal{D}(\text{Weibull})$ avec un indice des valeurs extrêmes $\gamma < 0$ si et seulement si $x_F < +\infty$ et la fonction $(1 - \tilde{F})$ est à variations régulières d'indice $1/\gamma$ avec

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

Ainsi, une fonction de répartition F du domaine d'attraction de Weibull s'écrit pour $x \leq x_F$ sous la forme :

$$F(x) = 1 - (x_F - x)^{-1/\gamma}L((x_F - x)^{-1}).$$

Comme on peut facilement passer du domaine d'attraction de Fréchet à celui de Weibull on a les suites de normalisation (a_n) et (b_n) données pour tout $n > 0$ par :

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

Du Théorème 1.3.3, on en déduit que $F \in \mathcal{D}(\text{Weibull})$ si et seulement si le point terminal x_F est fini et $\overline{F}(x) = (x_F - x)^{-1/\gamma} L[(x_F - x)^{-1}]$ avec $L(\cdot)$ est une fonction à variations lentes à l'infini et γ un réel strictement négatif.

Ce domaine d'attraction a été considéré pour estimer le point terminal d'une distribution.

1.3.4 Domaine d'attraction de Gumbel

Les résultats concernant le domaine d'attraction de la loi de Gumbel sont plus délicats, ainsi que la caractérisation des fonctions de répartition de ce domaine est plus complexe. Le résultat ci-dessous est démontré notamment dans (**Resnick** (1987) [41], Proposition 1.4) donne une caractérisation plus précise.

Théorème 1.3.4. $F \in \mathcal{D}(\text{Gumbel})$ si et seulement si

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

où c est une fonction mesurable tel que $\lim_{x \rightarrow x_F} c(x) = c > 0$ et a est une fonction positive et dérivable de dérivée a' telle que : $\lim_{x \rightarrow x_F} a'(x) = 0$.

Le choix possible pour les suites (a_n) et (b_n) dans ce cas est le suivant :

$$a_n = q_{1/n} \quad \text{et} \quad b_n = \frac{1}{\overline{F}(a_n)} \int_{a_n}^{x_F} \overline{F}(y) dy.$$

On donne ici une liste non exhaustive de l'appartenance des lois à leur domaine d'attraction dans le tableau suivant :

Fréchet ($\gamma > 0$)	Gumbel ($\gamma = 0$)	Weibull ($\gamma < 0$)
Pareto	Normale	Uniforme
Student	Exponentielle	Beta
Burr	Log-normale	ReverseBurr
Chi-deux	Weibull	
Fréchet	Gamma	
Cauchy	Gumbel	
Log-gamma	Logistique	
Log-logistique	Benktander(type1, 2)	

Table 1.1- Quelques lois et leurs domaines d'attraction

1.4 Estimation de l'indice des valeurs extrêmes

L'estimateur de l'indice de queue joue un rôle important dans le comportement de la lois des extrêmes. Dans la littérature, plusieurs estimateurs de cet indice et de quantiles extrêmes ont été proposés. Mais, nous exposerons uniquement deux estimateurs de l'indice de queue, tous basés sur la statistique d'ordre. Notre choix s'explique par notre volonté d'introduire une adaptation des estimateurs ainsi proposés au cas conditionnel (chapitre 3).

1.4.1 Estimateur de Pickands

L'estimateur de Pickands est construit en utilisant trois statistique d'ordres. Cet estimateur a l'avantage d'être valable quel que soit le domaine d'attraction de la distribution et par conséquent, du domaine de définition de l'indice des valeurs extrêmes. De plus, cet estimateur est très sensible à la taille de l'échantillon sélectionné, ce qui le rend peu robuste. **Pickands** (1975) [40] démontre la consistance faible de son estimateur. La convergence forte ainsi que la normalité asymptotique ont été démontrées par **Dekkers** et de **Haan** (1989) [20].

Définition 1.4.1. *On suppose que $\{X_i, i = 1, \dots, n\}$ est une suite de variables aléatoires indépendantes de loi F appartenant à 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 :

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

Théorème 1.4.1. (*Propriétés asymptotiques de $\widehat{\gamma}_{k_n}^P$*). Soit $(k_n)_{n \geq 1}$ une suite d'entiers telle que $1 \leq k_n \leq n$, $k_n \rightarrow \infty$ et $\frac{k_n}{n} \rightarrow 0$ quand $n \rightarrow \infty$. Alors :

- $\widehat{\gamma}_{k_n}^P \xrightarrow{\mathbb{P}} \gamma$.
- Si de plus $\frac{k_n}{\log \log n} \rightarrow \infty$ quand $n \rightarrow \infty$, alors $\widehat{\gamma}_{k_n}^P \xrightarrow{p.s.} \gamma$.
- 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) [20], on a :

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

1.4.2 Estimateur de Hill

Cet estimateur a été introduit par **Hill** (1975) [35] pour estimer d'une manière non-paramétrique le paramètre de queue des lois appartenant au \mathcal{D} (Fréchet), il fournit un estimateur de l'indice de queue plus efficace que l'estimateur de Pickands. Pour construire son estimateur, Hill utilise les k_n statistiques d'ordre supérieur. Un grand nombre de travaux théoriques ont été consacrés à l'étude des propriétés de l'estimateur de Hill. **Mason**(1982) [38] a démontré la consistance faible et **Deheuvels**, **Haeusler** et **Mason** (1988) [18] ont établi la consistance forte. La normalité asymptotique est due entre autres à **Davis** et **Resnick** (1984) [15], **Csörgö** et **Mason** (1985) [13] et **Smith** (1987) [46]. D'autres méthodes pour la construction de l'estimateur de Hill sont données dans le livre de **Haan** et **Ferreira** (2006) [19].

La construction de l'estimateur de Hill

Soient α_n et β_n deux suites de nombre positives, la construction de l'estimateur de Hill basée sur la relation suivante :

$$q_{\beta_n} \simeq q_{\alpha_n} \left(\frac{\alpha_n}{\beta_n} \right)^\gamma. \quad (1.6)$$

Passons au logarithme dans l'équation (1.6) ce qui donne :

$$\log(q_{\beta_n}) - \log(q_{\alpha_n}) \simeq \gamma \log \left(\frac{\alpha_n}{\beta_n} \right).$$

On choisit $\alpha_n = k_n/n$ et on considère plusieurs valeurs pour β_n . $\beta_n = i/n$ avec $i = 1, \dots, k_n - 1$ tout en ayant $\beta_n < \alpha_n$. On obtient alors :

$$\log(q_{i/n}) - \log(q_{k_n/n}) \simeq \gamma \log(k_n/i).$$

Ainsi en estimant les quantiles par leurs équivalents empiriques on obtient :

$$\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n}) \simeq \gamma \log(k_n/i).$$

En sommant de part et d'autre sur $i = 1, \dots, k_n - 1$, on obtient :

$$\gamma = \frac{\sum_{i=1}^{k_n-1} \log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n})}{\sum_{i=1}^{k_n-1} \log(k_n/i)}.$$

Le dénominateur se réécrit $\log(k_n^{k_n-1}/(k_n-1)!)$. En utilisant la formule de Stirling, il est équivalent à k_n au voisinage de l'infini. On obtient alors l'estimateur de Hill.

Définition 1.4.2. Soit $(k_n)_{n \geq 1}$ une suite d'entiers avec $1 \leq k_n \leq n$, 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 1.4.2. (*Propriétés asymptotiques de $\widehat{\gamma}_{k_n}^H$*). Soit $(k_n)_{n \geq 1}$ une suite d'entiers telle que $1 \leq k_n \leq n$, $k_n \rightarrow \infty$ et $\frac{k_n}{n} \rightarrow 0$ quand $n \rightarrow \infty$. Alors :

- $\widehat{\gamma}_{k_n}^H \xrightarrow{\mathbb{P}} \gamma$. (*consistance faible*)
- Si de plus $\frac{k_n}{\log \log n} \rightarrow \infty$ quand $n \rightarrow \infty$, alors $\widehat{\gamma}_{k_n}^H \xrightarrow{p.s.} \gamma$. (*consistance forte*)
- Et si $\sqrt{k_n} \Delta(\frac{n}{k_n}) \rightarrow 0$ quand $n \rightarrow \infty$, tel que : $\Delta(x)$ est une fonction mesurable à variations régulières d'indice $\rho < 0$ à l'infini qui converge vers 0 quand $x \rightarrow \infty$, on a alors :

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

Remarque 1.4.1. Une suite vérifiant les deux conditions $\lim_{n \rightarrow \infty} k_n = \infty$ et $\lim_{n \rightarrow \infty} \frac{k_n}{n} = 0$ sera appelée suite intermédiaire d'entiers.

- La première condition $k_n \rightarrow \infty$ nous assure que le nombre de statistiques d'ordre utilisé est assez grand afin d'obtenir des estimateurs stables.
- La seconde condition $\frac{k_n}{n} \rightarrow 0$ permet de rester dans la queue de distribution.

1.5 Le choix du nombre de statistiques d'ordre

En pratique, le choix du nombre k_n de statistiques d'ordre extrêmes à utiliser pose des problèmes. Il s'agit de choisir k_n de manière à ce que l'on dispose suffisamment de matériel statistique tout en restant dans la queue de distribution. Pour cela, on introduit quelques méthodes proposées pour sélectionner ce paramètre.

1.5.1 Méthode graphique

C'est la plus simple des méthodes pour la détermination de k_n . Elle consiste à tracer le graphe $\{(k_n, \widehat{\gamma}_{k_n}), 1 \leq k_n < n\}$, où $\widehat{\gamma}_{k_n}$ représente n'importe quel estimateur introduit précédemment et de prendre la valeur où l'estimateur $\widehat{\gamma}_{k_n}$ devient stable.

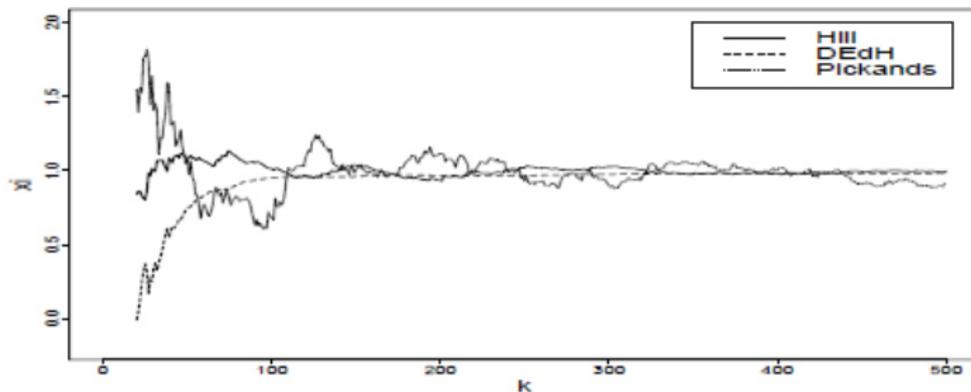


Figure 1.3- Le graphe de $(k_n, \hat{\gamma}_{k_n})$ pour les différents estimateurs de γ .

1.5.2 Méthode analytique

Il est nécessaire pour donner une précision à l'estimateur $\hat{\gamma}_{k_n}$, de calculer l'erreur en moyenne quadratique (EQM), qui est en fonction de k_n , tel que :

$$\begin{aligned} EQM(\hat{\gamma}_{k_n}) &= \mathbb{E}(\hat{\gamma}_{k_n} - \gamma)^2 \\ &= \text{Biais}^2(\hat{\gamma}_{k_n}) + \text{Var}(\hat{\gamma}_{k_n}). \end{aligned}$$

Donc le choix optimal de k_n (noté k_{opt}) correspond à minimiser EQM, (i.e)

$$k_{opt} = \arg \min_k EQM(\hat{\gamma}_{k_n}).$$

Par conséquent, pour une estimation exacte de l'indice de queue, il est nécessaire pour n'importe quel estimateur classique à faire un compromis entre le biais et la variance.

Remarque 1.5.1. *On distingue 2 cas :*

- Si k_n est trop petit, $\hat{\gamma}_{k_n}$ utilise peu d'observations, il a alors une grande variance dans ce cas.
- Si k_n est trop grand, $\hat{\gamma}_{k_n}$ a un grand biais.

Donc le nombre de statistiques d'ordre utilisées dans les procédés d'estimation doit être intermédiaire, (i.e) ni trop petit ni trop grand.

Chapitre 2

Estimation de quantiles conditionnels

L'étude *des fonctions quantiles conditionnelles* est intéressante pour un certain nombre de raisons, notamment car il fournit une image plus complète de la distribution conditionnelle d'une personne à charge variable que la régression classique qui restreint l'attention à la fonction moyenne conditionnelle seulement. La littérature sur l'estimation de quantile conditionnel, qui a considérablement augmenté ces dernières années, provient du travail fondateur de **Koenker** et **Bassett** (1978) [3] qui a introduit cette approche dans un cadre de régression paramétrique. La réalisation, cependant, que les praticiens peuvent exiger plutôt estimations non-paramétriques a donné lieu à une littérature grandissante qui inclut **Chaudhuri** (1991) [9], **Yu** et **Jones** (1998), (2007) [55] et **Daouia, Gardes** et **Girard** (2013) [14] d'illustration.

2.1 Quantile et Quantile Conditionnel

Soit \mathbb{E} un espace métrique (de dimension finie ou infinie) muni d'une distance d .

Définition 2.1.1. (*Quantile*) Soit Y une variable aléatoire réelle, F sa fonction de répartition et $\alpha \in]0, 1[$. Le quantile d'ordre α de la variable Y , noté $q_y(\alpha)$, est la plus petite valeur y vérifiant $F(y) \geq \alpha$ (i.e)

$$q_y(\alpha) := F^{-1}(\alpha) = \inf\{y \in \mathbb{R} : F(y) \geq \alpha\}. \quad (2.1)$$

Remarque 2.1.1. *Pour certaines valeurs de α , on donne un nom particulier aux quantiles ; par exemple, pour $\alpha = 0.5$ le quantile appelé médiane, pour $\alpha = 0.25, 0.75$ le quantile appelé quartile, . . . , etc.*

Définition 2.1.2. (Quantile conditionnel) *Soit X, Y deux variables aléatoires réelles (la variable Y appelée variable d'intérêt et la variable X appelée covariable). $F(y/x)$ la fonction de répartition conditionnelle de Y sachant $X = x$, et $\alpha \in]0, 1[$. D'une manière similaire à la définition du quantile, le quantile conditionnel $q_y(\alpha/x)$ (noté simplement $q_\alpha(x)$) est défini par*

$$q_\alpha(x) := F^{-1}(\alpha/x) = \inf\{y \in \mathbb{R} : F(y/x) \geq \alpha\}. \quad (2.2)$$

Plusieurs approches ont été développées pour l'estimation des quantiles conditionnels. L'approche **paramétrique** peut parfois être mal adaptée à la réalité des données en particulier biologiques. Une approche **non-paramétrique** du problème a alors été développée afin de pallier les problèmes d'hypothèses et de modélisation paramétriques. De nombreux travaux récents ont été menés pour l'estimation non paramétrique des quantiles conditionnels aussi bien dans un cadre théorique que sur le plan des applications. Ces méthodes ne nécessitent pas d'hypothèse sur la nature de la distribution.

2.2 Estimation paramétrique des quantiles conditionnels

Quand l'échantillon est de petite taille, des hypothèses paramétriques sont habituellement imposées afin de réduire le nombre de paramètres à estimer. En particulier, la fonction de répartition conditionnelle est supposée gaussienne. Ainsi, un estimateur du quantile conditionnel $q_\alpha(x)$ est donné par

$$\widehat{q}_{\alpha_n}(x) = \widehat{m}_n(x) + z_\alpha \widehat{\sigma}_n(x),$$

où $\widehat{m}_n(x)$ et $\widehat{\sigma}_n(x)$ désignent respectivement les estimateurs de l'espérance conditionnelle et de la variance conditionnelle de Y sachant $X = x$ et où z_α est le quantile

d'ordre α de la loi normal centrée réduite $\mathcal{N}(0, 1)$. Afin d'estimer $m(x)$ et $\sigma^2(x)$, **Royston et Wright** (1998) [43] ont utilisé un modèle polynomial qu'ils ont associé 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 (voir **Cole** (1988) [11] et **Tango** (1998) [49]). L'existence d'une telle transformation n'est nullement garantie. Il est aussi connu que les hypothèses paramétriques sont restrictives et peuvent rarement être faites avec certitude.

2.3 Estimation non-paramétrique des quantiles conditionnels

L'estimation non paramétrique des quantiles conditionnels a reçu un grand intérêt depuis (1969) quand **Roussas** (1969) [42] a montré la convergence et la normalité asymptotique des estimateurs à noyau. On peut citer les travaux de (**Stute** (1986) [48]; **Samanta** (1989) [44]) lorsque les données sont indépendantes et identiquement distribuées ou α -mélangeantes (**Berlinet et al** (1998) [5]). Dans le même esprit, Plusieurs méthodes pour estimer le quantile conditionnel de Y pour une valeur donnée de X ont été proposées. On les regroupe en deux grandes catégories : les méthodes indirectes, les méthodes directes.

2.3.1 Méthode d'estimation indirecte

Elle consiste à estimer au préalable la fonction de répartition conditionnelle puis à l'inverser pour obtenir un estimateur du quantile conditionnel. Et pour assurer l'existence et l'unicité de ce dernier on suppose que la fonction de répartition conditionnelle $F(\cdot/x)$ est strictement croissante et continue.

On dénombre donc deux techniques d'estimation indirectes de la fonction de répartition conditionnelle :

1. Les estimateurs à noyaux ("Kernel estimator").

2. Les estimateurs au sens des plus proches voisins.

Les estimateurs à noyaux

Il est naturel d'estimer le quantile conditionnel $q_\alpha(x)$ par $\hat{q}_{\alpha_n}(x)$ défini de la manière suivante :

$$\hat{q}_{\alpha_n}(x) = \hat{F}_n^{\leftarrow}(\alpha/x) = \inf\{y : \hat{F}_n(y/x) \geq \alpha\} \quad \text{avec} \quad \alpha \in]0, 1[.$$

(a) -Méthode d'estimation par simple noyau

On estime la fonction de répartition conditionnelle par :

$$\hat{F}_n(y/x) = \sum_{i=1}^n w_{n,i}(x) \mathbb{1}_{[Y_i \leq y]} \quad \text{avec} \quad \sum_{i=1}^n w_{n,i}(x) = 1. \quad (2.3)$$

Notons que si on prend $w_{n,i}(x) = 1/n$, on obtient l'expression classique de la fonction de répartition empirique. En posant $Y^* = \mathbb{1}_{[Y \leq y]}$, on a $F(y/x) = \mathbb{E}[Y^*/X = x]$ et $\hat{q}_{\alpha_n}(x)$ est appelé l'estimateur du quantile de régression.

Plusieurs auteurs s'intéressent à ce type d'estimation dans *le cas du design aléatoire*, tel que **Yu** et **Jones** (1998) [55]. Ainsi que la convergence ponctuelle en probabilité de l'estimateur (2.3) a été établie par **Stone** (1977) [47] lorsque les variables aléatoires $(X_i, Y_i) \in \mathbb{R}^p \times \mathbb{R}$ sont indépendantes et identiquement distribuées. Plus précisément, il donne des conditions sur les poids $w_{n,i}$ pour que l'estimateur converge ponctuellement en probabilité. **Collomb** (1980) [12] 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

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

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

Et alors l'estimateur de la fonction de répartition conditionnelle ainsi obtenu est défini par :

$$\hat{F}_n(y/x) = \sum_{i=1}^n \frac{K\left(\frac{x-X_i}{h_n}\right)}{\sum_{j=1}^n K\left(\frac{x-X_j}{h_n}\right)} \mathbb{1}_{[Y_i \leq y]}.$$

En utilisant un noyau de probabilité continu et borné sur $[-1, 1]$, **Stute** (1986) [48] posa :

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

avec $G_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{[X_j \leq x]}$. Où l'estimateur à noyau de la fonction de

répartition conditionnelle donné par :

$$\hat{F}_n(y/x) = \sum_{i=1}^n \frac{K\left(\frac{G_n(x) - G_n(X_i)}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{G_n(x) - G_n(X_i)}{h_n}\right)} \mathbb{1}_{[Y_i \leq y]}.$$

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.

Gannoun (1989) [27] étudie les propriétés de l'estimateur de **Collomb** (1980) [12] dans le cas des données indépendantes et identiquement distribuées puis α -mélangeantes. **Hart** (1991) [33] propose un estimateur plus général que l'estimateur de **Stute** (1986) [48], il prend un noyau continue et borné sur $[0, 1]$ et

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

Une modification de l'estimateur de **Collomb** (1980) [12] a été introduite par **Ferraty** et **Vieu** (2000) [24] dans le cas où la covariable X est de nature fonctionnelle ou de dimension infinie, ils posent

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

Où d est une distance semi-métrique sur \mathbb{E} qui mesure la proximité entre deux objets fonctionnels et K un noyau qui est positif et décroissant sur $[0, 1]$. **Ferraty** et **Vieu** (2006) [25] donnent la vitesse de convergence presque complète dans le cas où les variables X_i sont indépendantes et identiquement distribuées, et dans le cas dépendant.

Un autre point de vue d'estimation de la fonction de répartition conditionnelle dans *le cas du design fixe*, a été introduite par **Anthoch** et **Janssen**(1989)

[1], ils posent :

$$\left\{ \begin{array}{l} w_{n,i}(x) = \frac{1}{h_n} \int_{x_{i-1}}^{x_i} K\left(\frac{x-u}{h_n}\right) du \quad 2 \leq i \leq n-1 \\ w_{n,1}(x) = \frac{1}{h_n} \int_{-\infty}^{x_1} K\left(\frac{x-u}{h_n}\right) du \\ w_{n,n}(x) = \frac{1}{h_n} \int_{x_{n-1}}^{+\infty} K\left(\frac{x-u}{h_n}\right) du \end{array} \right.$$

où K est un noyau de probabilité.

(b) -Méthode d'estimation par noyau produit

Une version plus "lisse" et régulière des estimateurs du simple noyau de la fonction de répartition conditionnelle avait été introduite par **Roussas** (1969) [42] dans le cas d'un processus $(X_i, Y_i)_{i=1, \dots, n}$ à valeurs dans $\mathbb{E} \times \mathbb{R}$, supposé Markovien. Cet estimateur peut être également vu comme une primitive de l'estimateur à noyau de la densité conditionnelle. Les propriétés asymptotiques de cet estimateur sont établies par exemple dans l'article de **Berlinet, Gannoun et Matzner-Lober** (2001) [5]. Elle consiste tout d'abord à remplacer la fonction indicatrice de l'estimateur à noyau simple par une densité symétrique.

Ce qui permet d'écrire l'estimateur à double noyaux de la fonction de répartition conditionnelle comme suit :

$$\hat{F}_n(y/x) = \frac{\sum_{i=1}^n K_0\left(\frac{d(x, X_i)}{h_{n,0}}\right) K_1\left(\frac{y - Y_i}{h_{n,1}}\right)}{\sum_{i=1}^n K_0\left(\frac{d(x, X_i)}{h_{n,0}}\right)}.$$

avec $K_i, (i = 0, 1)$ est un noyau positif et décroissant sur $[0, 1]$.

Convergence presque complète : Cas i.i.d

Ferraty, Laksaci et Vieu (2005) [22] donnent la vitesse de convergence presque complète¹ dans le cas où les variables X_i sont indépendantes et identiquement distribuées de cet estimateur sous les conditions suivantes :

Hypothèses

(C1) $P(X \in B(x, r)) = \phi_x(r) > 0$,

(C2) F^x est k - fois continûment différentiable par rapport à y sur $(t_\alpha - \xi, t_\alpha + \xi)$, vérifiant $F^{x^{(l)}}(t_\alpha) = 0$, si $l \leq k$, et $0 < |F^{x^{(k)}}(t_\alpha)| < \infty$,

(C3) $\forall (y_1, y_2) \in \mathbb{R} \times \mathbb{R}, \forall (x_1, x_2) \in N_{x_1} \times N_{x_2}$,
 $|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C_x (d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2})$, où b_1 et b_2 deux réels strictement positifs.

(C4) $\left\{ \begin{array}{l} \forall (y_1, y_2) \in \mathbb{R}^2, |K_1(y_1) - K_1(y_2)| \leq C|y_1 - y_2|, \\ \int_{\mathbb{R}} |t|^{b_2} K_1^{(1)}(t) dt < \infty. \end{array} \right.$

(C5) K_0 est une fonction à support compact sur $[0, 1]$ vérifiant $0 < C_1 < K_0(t) < C_2 < \infty$, où C_1 et C_2 sont deux constantes strictement positives,

(C6) $\lim_{n \rightarrow \infty} h_{n,0} = 0$, $\lim_{n \rightarrow \infty} \frac{n\phi_x(h_{n,0})}{\log n} = \infty$.

(C7) $\lim_{n \rightarrow \infty} h_{n,1} = 0$, $\exists \alpha > 0 \lim_{n \rightarrow \infty} n^\alpha h_{n,1} = \infty$.

Théorème 2.3.1. (Ferraty, Laksaci et Vieu (2005) [22])

Sous les hypothèses (C1) – (C7) on a :

$$\widehat{q}_\alpha - q_\alpha = O\left(h_{n,0}^{\frac{b_1}{k}} + h_{n,1}^{\frac{b_2}{k}}\right) + O\left(\left(\frac{\log n}{nh_{n,1}\phi_x(h_{n,0})}\right)^{\frac{1}{2k}}\right), \quad p.co.$$

1. soit $(X_n)_n \in \mathbb{N}$ une suite de variables aléatoires. On dit que X_n converge presque complètement (p.co) vers 0 ssi : $\forall \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \epsilon) < \infty$. De plus, soit $(Z_n)_n \in \mathbb{N}$ une suite de nombres réels positifs. On dit que $X_n = \mathcal{O}(Z_n)$ p.co. ssi : $\forall \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \epsilon Z_n) < \infty$.

Lorsque X est de nature fonctionnelle, **Ferraty et al** (2005) proposent de modifier l'estimateur dans le cas des données non nécessairement dépendantes. L'estimateur ainsi défini est donné par

$$\widehat{F}_n(y/x) = \frac{\sum_{i=1}^n K_0\left(\frac{d(x, X_i)}{h_{n,0}}\right) \int_{-\infty}^y K_1\left(\frac{(u - Y_i)}{h_{n,1}}\right) du}{\sum_{i=1}^n K_0\left(\frac{d(x, X_i)}{h_{n,0}}\right)},$$

où d est une distance semi-métrique, K_1 un noyau positif et décroissant sur $[0, 1]$, K_0 un noyau symétrique, $h_{n,0}$ et $h_{n,1}$ deux suites positives.

Convergence presque complète : Cas α -mélangeant

Les auteurs (**Ferraty, Rabhi et Vieu** (2005) [23]) établissent la vitesse de convergence presque complète dans ce cas sous les hypothèses précédentes et les hypothèses suivantes :

Hypothèses

$$(C8) \sup_{i \neq j} P((X_i, X_j) \in B(x, r) \times B(x, r)) = \phi_x(r)\psi_x(r) > 0,$$

(C9) Le coefficient de mélange vérifie

$$\exists a > (5 + \sqrt{17})/2, \exists c > 0, \forall n, \alpha_n \leq cn^{-a},$$

$$(C10) \begin{cases} \lim_{n \rightarrow \infty} h_{n,0} = 0, \lim_{n \rightarrow \infty} \frac{\log n}{nh_{n,1} \chi_x(h_{n,0})} = 0, \\ \exists \beta_2 > 0, c_1 > 0, c_2 > 0, c_2 n^{\left(\frac{3-a}{a+1}\right)+\beta_2} \leq \chi_x(h_{n,0}) \leq c_1 n^{\frac{1}{1-a}-\beta_2}. \end{cases}$$

telle que, $\chi_x(h_{n,0}) = \max(\psi_x(h_{n,0}), \phi_x(h_{n,0}))$.

Cette vitesse est donnée par :

Théorème 2.3.2. (Ferraty, Rabhi et Vieu (2005) [23])

$$\widehat{q}_\alpha - q_\alpha = O\left(h_{n,0}^{\frac{b_1}{k}} + h_{n,1}^{\frac{b_2}{k}}\right) + O\left(\left(\frac{\log n}{nh_{n,1}\phi_x(h_{n,0})}\right)^{\frac{1}{2k}}\right), \quad p.co.$$

Convergence en norme L^p : Cas i.i.d

La convergence dans l'espace $L^p[0, 1]$ ² de cet estimateur a été étudiée par **Dabo-Niang** et **Laksaci** en (2007) dans le cas indépendant avec la vitesse suivante :

Théorème 2.3.3. (Dabo-Niang et Laksaci(2007))

$$\|\widehat{q}_\alpha - q_\alpha\|_p = O\left(h_{n,0}^{\frac{b_1}{k}} + h_{n,1}^{\frac{b_2}{k}}\right) + O\left(\left(\frac{1}{n\phi_x(h_{n,0})}\right)^{\frac{1}{2k}}\right)$$

sous les conditions (C1-C7)

Convergence en norme L^p : Cas α -mélangeant

Sous les hypothèses (C1-C7) et les conditions suivantes :

(C11) Le coefficient de mélange vérifie $\exists a > 0, \exists c > 0, \forall n \in \mathbb{N}, \alpha(n) \leq cn^{-a}$.

(C12) $\forall i \neq j, 0 < \sup_{i \neq j} P[(X_i, X_j) \in B(x, h) \times B(x, h)] = O\left(\frac{(\phi_x(h))^{a+1/a}}{n^{1/a}}\right)$.

(C13) $\forall (y_1, y_2) \in N_x \times N_x, |F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C(d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2}),$
 $b_1 > 0, b_2 > 0$.

On a :

Théorème 2.3.4. (Dabo-Niang et Laksaci(2007))

$$\|\widehat{q}_\alpha - q_\alpha\|_p = O\left(h_{n,0}^{\frac{b_1}{k}} + h_{n,1}^{\frac{b_2}{k}}\right) + O\left(\left(\frac{1}{n\phi_x(h_{n,0})}\right)^{\frac{1}{2k}}\right).$$

2. Soit X_n une suite de variables aléatoires définie sur le même espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, on dit que X_n converge vers X dans l'espace L^p ssi : $\mathbb{E}(|X_1|^p) < \infty$ et $\lim_{n \rightarrow \infty} \mathbb{E}(|X_n - X|^p) = 0$.

tel que $\|\cdot\|_p = (E^{1/p}|\cdot|^p)$.

Mentionnons que pour les deux méthodes, le choix du paramètre de lissage joue un rôle très important. Ce choix est un peu compliqué dans le cas du double noyau (noyau produit), où il faut faire un choix multiple de ce paramètre.

(c) - *Méthode de la fenêtre mobile*

Encore appelé médianogramme, l'estimateur de la fenêtre mobile introduit par **Tukey** (1961) [51] est valable quelque soit le design (fixe, aléatoire). 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 de centre x et de rayon h_n ((i.e) : $B(x, h_n) = \{y \in \mathbb{E}, d(y, x) \leq h_n\}$). En posant :

$$\left\{ \begin{array}{l} w_{n,i}(x) = \frac{\mathbb{1}_{\{X_i \in B(x, h_n)\}}}{\sum_{i=1}^n \mathbb{1}_{\{X_i \in B(x, h_n)\}}} \quad \text{designe aléatoire} \\ w_{n,i}(x) = \frac{\mathbb{1}_{\{x_i \in B(x, h_n)\}}}{\sum_{i=1}^n \mathbb{1}_{\{x_i \in B(x, h_n)\}}} \quad \text{designe fixe} \end{array} \right.$$

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 alors construit 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 $K(x) = \mathbb{1}_{\{x \in [-1, 1]\}}$, **Truong**(1989) [50] a étudié la vitesse de convergence optimale de l'estimateur de la fenêtre mobile de la médiane conditionnelle. La convergence uniforme presque complète de cet estimateur a été établie par **Gannoun**(1989) [27].

Les estimateurs au sens des plus proches voisins

La méthode de construction de ces estimateurs est analogue à ceux de la fenêtre mobile. **Bhattacharyya** et **Gangopadhyay** (1990) [6] 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 du point fixe x , (*i.e*)

$$\widehat{F}_k(y/x) = \frac{1}{k} \sum_{i=1}^k \mathbb{1}_{\{Y_{[n,i]} \leq y\}},$$

où $k \leq n$ et les $Y_{[n,i]}$ sont les concomitants des statistiques d'ordre $U_{n,1} < \dots < U_{n,n}$ associés aux $U_i = |X_i - x|$.

L'estimateur au sens des plus proches voisins de quantile conditionnel est alors défini par

$$\widehat{q}_\alpha(x) = \inf \left\{ y : \widehat{F}_k(y/x) \geq \frac{\lfloor k\alpha \rfloor}{k} \right\}.$$

Passant maintenant à la deuxième méthode d'estimation du quantiles conditionnels

2.3.2 Méthode d'estimation directe

Cette méthode basée sur la recherche d'un estimateur du quantile conditionnel comme étant une solution minimale d'un problème de la forme suivante :

$$\frac{1}{n} \sum_{i=1}^n l_\alpha(Y_i - r_\alpha(X_i)).$$

Koenker (2005) définit la fonction de perte ("check function") l_α par :

$$l_\alpha(u) = |u| + (2\alpha - 1)u.$$

Remarque 2.3.1. Pour $\alpha = 0.5$, on a la médiane conditionnelle.

Bassett et **Koenker** (1978) [3] étudient ainsi un estimateur de la médiane conditionnelle en minimisant un critère du type moindres valeurs absolues, et prouvent sa convergence et sa normalité asymptotique. Koenker a résolu ce problème de minimisation, il utilise une méthode algorithmique, notamment la méthode de simplexe.

Historiquement, il y a trois approches directes d'estimations :

1. La méthode de la constante locale

La technique consiste à approcher le quantile par une fonction linéaire

$$q_\alpha(z) \approx q_\alpha(x) + q'_\alpha(x)(z - x) := a + b(z - x),$$

pour z dans un voisinage de x . L'idée est de minimiser par rapport à a et b la quantité

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

en utilisant le principe de moindre carrés.

Si $b = 0$, on parle de la méthode dite de la constante locale et on estime le quantile conditionnel par :

$$\hat{q}_{\alpha_n}(x) = \arg \min_{a \in \mathbb{R}} \sum_{i=1}^n l_\alpha(Y_i - a)K\left(\frac{x - X_i}{h_n}\right).$$

Plusieurs auteurs s'intéressent à l'étude de ses propriétés asymptotiques, **Stone** (1977) [47], **Yu** et **Jones** (1998) [55], **Berlinet** et **al** (2001) [5], **Mint ElMouvit** (2000) [39].

2. La méthode de polynômes locaux

Cette méthode est plus générale que la méthode de la constante locale, si on prend un noyau de type fonction indicatrice d'intervalles. Pour $k = 1$, elle approxime le quantile conditionnel par un polynôme de degré k , la détermination de ce polynôme revient à résoudre le problème d'optimisation suivant :

$$\min_{(b_0, \dots, b_k)} \sum_{i=1}^n l_\alpha \left(Y_i - \sum_{j=0}^k b_j (X_i - x)^j \right) K \left(\frac{x - X_i}{h_n} \right).$$

L'estimateur de quantile conditionnel est alors défini par :

$$\hat{q}_{\alpha_n}(x) = \hat{b}_0.$$

Chauduri (1991) [9] a démontré la convergence forte ponctuelle. **He** et **Shi** (1994) [34] proposent un estimateur de $q_\alpha(x)$ basé sur des *splines de régression* et obtiennent des vitesses de convergence usuelles en statistique non-paramétrique. Dans la littérature, on trouve d'autres estimateurs, comme par exemple l'estimateur de **Koenker, Ng et Portnoy** (1994) basé sur des *splines de lissage*. D'un point de vue de l'estimation linéaire, **Hervé Cardot, Christophe Crambe** et **Pascal Sarda**(2005) [8], **Collomb** (1980) [12] estiment le quantile conditionnel lorsque la variable explicative X est fonctionnelle et Y est réelle, ils posent un estimateur spline du coefficients fonctionnels, ils suposent que pour tout $x \in L^2([0, 1])$, $\alpha \in [0, 1]$, la fonction quantile conditionnel $q_\alpha(x)$ est un opérateur linéaire continue sur $L^2([0, 1])$ dans \mathbb{R} , le théorème de Riesz³ assure l'existence d'une unique fonction $\Psi_\alpha \in L^2([0, 1])$ tel que :

$$q_\alpha(x) = \langle \Psi_\alpha, X \rangle = \int_0^1 \Psi_\alpha(t) X(t) dt. \quad (2.4)$$

Ce modèle est une généralisation de la régression sur quantiles introduite par **Koenker** et **Basset** (1978) [3]. Le but est de définir un estimateur spline de la fonction Ψ_α . Pour ce faire, soit $q \in \mathbb{N}$, $k = k_n \in \mathbb{N}^*$, on donne une subdivision en k sous-intervalle de $[0, 1]$, l'estimateur spline $\hat{\Psi}_\alpha$ de la fonction Ψ_α est défini comme une combinaison linéaire de vecteurs de bases de l'espace de fonction B-spline $B_l, l = 1, \dots, k + q$, qui sont des fonctions polynômiales de degré q par morceaux sur chaque sous-intervalle de $[0, 1]$, cette base est notée

3. H un sous espace de hilbert menu de son produit scaliare noté $\langle \cdot, \cdot \rangle$, $f \in H$ une forme linéaire continue sur H , alors il existe un unique $Y \in H$ tel que pour tout $x \in H$ on ait $f(x) = \langle Y, x \rangle$.

par $B_{k,q} = {}^t(B_1, \dots, B_{k+q})$. Le problème devient la recherche d'un vecteur $\hat{\theta} = {}^t(\hat{\theta}_1, \dots, \hat{\theta}_{k+q}) \in \mathbb{R}^{k+q}$ tel que :

$$\hat{\Psi}_\alpha = {}^t B_{k,q} \hat{\theta}$$

$\hat{\theta}$ est la solution du problème de minimisation suivant :

$$\min_{\theta \in \mathbb{R}^{k+q}} \left\{ \frac{1}{n} \sum_{i=1}^n l_\alpha(Y_i - \langle {}^t B_{k,q} \theta, X_i \rangle) + \rho \|({}^t B_{k,q} \theta)^{(m)}\|^2 \right\}$$

- $({}^t B_{k,q} \theta)^{(m)}$ est la dérivée d'ordre m de $({}^t B_{k,q} \theta)$.
- ρ le paramètre de pénalisation pour contrôler le degré de «régularité» de l'estimateur cherché.
- $\|\cdot\|$ est la norme associée au produit scalaire de $L^2([0, 1])$.

Cardot, Crambe et Sardat (2005) [8], **Collomb** (1980) [12] trouvent que cet estimateur converge en moyenne quadratique vers Ψ_α , sous les hypothèses suivantes :

Hypothèses

(H1) $\exists 0 < C_0 < \infty$ tel que $\|X_i\| \leq C_0 < +\infty$ *p.s.*

(H2) La fonction Ψ_α admet une dérivée d'ordre p' et $\exists 0 < C_1 < \infty$ tel que

$$|\Psi_\alpha^{(p')}(t) - \Psi_\alpha^{(p')}(s)| \leq C_1 |t - s|^r, \quad s, t, r \in [0, 1].$$

(H3) Les valeurs propres de Σ_X sont strictement positives.

(H4) $\forall x \in L^2([0, 1])$, la distribution conditionnelle de ε_i sachant $X = x$ admet une densité continue et strictement positive en zéro.

Chapitre 3

Estimation de quantiles extrêmes conditionnels

Dans ce chapitre, nous nous intéressons au cas où la variable aléatoire d'intérêt Y est mesurée conjointement avec une covariable $X = x$. Contrairement au problème d'estimation des quantiles de régression classique qui a été largement considéré, peu d'attention a été accordée 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 $X = x$.

3.1 Estimation des quantiles extrêmes

On désire 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* et 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.

On considère n variables aléatoires réelles $\{X_i, i = 1, \dots, n\}$ indépendantes et

identiquement distribuées de fonction de répartition F non nécessairement continue.

Définition 3.1.1. (*quantile extrême*) *Le quantile extrême d'ordre $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$ est défini pour tout x par*

$$q_{\alpha_n} = \overline{F}^{\leftarrow}(\alpha_n) = \inf\{x : \overline{F}(x) \leq \alpha_n\}, \quad (3.1)$$

avec la convention $\inf\{\emptyset\} = \infty$. Où α_n est une suite connue et $\overline{F}^{\leftarrow}$ est l'inverse généralisé¹ de \overline{F} . Notons que l'inverse généralisé d'une fonction coïncide avec l'inverse classique F^{-1} lorsque la fonction F est strictement croissante et continue.

En résumé, un quantile sera dit extrême si l'on remplace son ordre α par une suite $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$. Le fait que l'ordre $\alpha_n \rightarrow 0$ indique que l'information la plus importante pour estimer des quantiles extrêmes est contenue dans la queue de distribution.

Remarque 3.1.1. *On peut dire aussi que q_{α_n} est le quantile extrême d'ordre $1 - \alpha_n$ de la fonction de répartition F .*

En particulier, pour $n \rightarrow \infty$, on a

$$\begin{aligned} \mathbb{P}(\max(X_1, \dots, X_n) < q_{\alpha_n}) &= \mathbb{P}(X_i < q_{\alpha_n}, \forall i = 1, \dots, n) \\ &= (1 - \alpha_n)^n \\ &= \exp(n \log(1 - \alpha_n)) \\ &= \exp(-n\alpha_n(1 + o(1))) \quad \text{quand } \alpha_n \rightarrow 0. \end{aligned}$$

Donc, la probabilité que le quantile extrême soit plus grand que le maximum de l'échantillon dépend du comportement asymptotique de $n\alpha_n$. Ainsi, lorsque l'on souhaite estimer des quantiles extrêmes, on doit faire la distinction entre deux cas qui sont fonction de la vitesse de convergence de α_n vers 0. Plusieurs méthodes d'estimation du quantile extrême q_{α_n} ont été proposées dans la littérature.

1. l'inverse généralisé généralise celle d'inverse d'une application linéaire ou d'une matrice aux cas non inversibles en lui supprimant certaines des propriétés demandées aux inverses, ou en l'étendant aux espaces non algébriques plus larges.

Dans tout ce qui suit, on fait l'hypothèse que F appartient à l'un des domaines d'attraction défini précédemment. Pour résumer le problème d'estimation investigué dans ces travaux, on introduit le résultat suivant appelé *approximation de Poisson*.

Lemme 3.1.1. (*Embrechts et al (1997) [21], Proposition 3.1.1*) Si $\alpha_n \rightarrow 0$ et $n\alpha_n \rightarrow c$ (non nécessairement fini) quand $n \rightarrow \infty$, alors

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

D'après ce lemme, deux cas peuvent être distingués en fonction de c :

Premier cas, si $c = \infty$ alors, $\mathbb{P}(X_{n,n} < q_{\alpha_n}) \rightarrow 0$. Dans ce cas, on cherche à estimer un quantile qui se trouve avec une grande probabilité dans l'échantillon. 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, autrement dit la statistique d'ordre supérieur $X_{n-[n\alpha_n]+1,n}$ (où $[\cdot]$ est la fonction partie entière). Le résultat principal de cet estimateur de quantile extrême est donné par le Théorème 3.1.1.

Second cas, si $c = 0$ alors, $\mathbb{P}(X_{n,n} < q_{\alpha_n}) \rightarrow 1$. On cherche alors à estimer un quantile qui se trouve avec une grande probabilité en dehors de l'échantillon. Par conséquent, on ne peut pas estimer le quantile de manière empirique. Pour résoudre ce problème, on présente trois catégories principales de méthodes :

1. La théorie des valeurs extrêmes consiste à diviser l'échantillon en k sous-blocs disjoints choisis arbitrairement de même taille n assez grande desquels on détermine les maxima puis à approcher la loi des maxima de ces blocs par *une loi des valeurs extrêmes*. En utilisant la relation $\mathbb{P}(\max(X_1, \dots, X_n) < q_{\alpha_n}) = F^n(q_{\alpha_n})$, on peut alors estimer le quantile extrême q_{α_n} .
2. La méthode des excès consiste de ne retenir que les observations dépassant un seuil fixé u et approcher la loi des excès si u est assez grand par *une loi de Pareto généralisée (GPD)*. Pour estimer le quantile extrême q_{α_n} , il suffit alors d'utiliser le résultat de (**Balkema** et de **Haan** (1974) [2]; **Pickands** (1975) [40]).

3. Les méthodes semi-paramétriques où l'on suppose que pour tout $\gamma > 0$ on a $\mathbb{P}(X > x) \sim x^{-1/\gamma}$ lorsque $x \rightarrow \infty$, c'est-à-dire que la fonction de survie $\bar{F}(x)$ décroît en $x^{-1/\gamma}$. Partant de ce résultat, **Weissman** (1978) [54] proposa un estimateur semi-paramétrique du quantile extrême q_{α_n} dépendant de l'estimation du paramètre γ dont le plus célèbre est l'estimateur de **Hill** (1975) [35]. En effet, supposer que $\bar{F}(x)$ décroît en $x^{-1/\gamma}$ revient à supposer que le quantile q_{α_n} décroît en $\alpha_n^{-\gamma}$.

Remarque 3.1.2. *Ces trois méthodes d'estimation de quantiles extrêmes sont basées sur des estimateurs des paramètres des lois GEV et GPD et en particulier sur des estimateurs de l'indice des valeurs extrêmes.*

L'utilité de la définition suivante est essentiellement d'ordre pratique (i.e) pour le choix de α_n en pratique.

Définition 3.1.2. *Pour tout $t > 0$, la fonction queue (en anglais "tail quantile function") est donnée par :*

$$U(t) = \inf\{x : F(x) \geq 1 - 1/t\} = q_{1/t}.$$

Théorème 3.1.1. *(de Haan et Ferreira (2006) [19], Théorème 2.2.1). Soit $(k_n)_{n \geq 1}$ une suite d'entiers telle que $1 \leq k_n \leq n$, $k_n \rightarrow \infty$ et $k_n/n \rightarrow 0$ quand $n \rightarrow \infty$. Si*

$$\lim_{x \rightarrow x_{\bar{F}}} \frac{\bar{F}(x)F''(x)}{(F'(x))^2} = -\gamma - 1, \text{ alors}$$

$$\sqrt{k_n} \left(\frac{X_{n-k_n+1,n} - U(n/k_n)}{(n/k_n)U'(n/k_n)} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Le Théorème 3.1.1 montre que l'estimateur de quantile extrême $X_{n-k_n+1,n}$ avec $k_n \stackrel{\text{def}}{=} \lfloor n\alpha_n \rfloor$ est asymptotiquement gaussien.

3.1.1 L'estimation des quantiles extrêmes par la loi des valeurs extrêmes

Guida et Longo (1988) [31] se sont basés sur cette approche $\mathbb{P}(X_{n,n} \leq a_n x + b_n) = F^n(a_n x + b_n) \simeq \mathcal{H}_\gamma(x)$ pour estimer des quantiles extrêmes q_{α_n} . En effet,

d'après le Théorème 1.2.1, on peut écrire

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

Et on peut montrer que $a_n x + b_n \rightarrow x_F$ quand $n \rightarrow \infty$ et donc $\bar{F}(a_n x + b_n) \rightarrow 0$. Alors pour tout γ , on approche le quantile q_{α_n} par :

$$q_{\alpha_n} \simeq a_n x_{\alpha_n} + b_n$$

où x_{α_n} vérifie $(-\log(\mathcal{H}_\gamma(x_{\alpha_n}))) = n\alpha_n$.

Tel que par un développement limité au premier ordre de $\log(1 + u)$, on a

$$\bar{F}(a_n x + b_n) \simeq -\frac{1}{n} \log(\mathcal{H}_\gamma(x)).$$

Ce qui équivaut à :

$$\bar{F}(x) \simeq -\frac{1}{n} \log\left(\mathcal{H}_\gamma\left(\frac{x - b_n}{a_n}\right)\right) = -\frac{1}{n} \log(\mathcal{H}_{\gamma, a_n, b_n}(x)).$$

Et on estime le quantile extrême en remplaçant les suites a_n et b_n respectivement par leurs estimateurs \hat{a}_n et \hat{b}_n :

$$\hat{q}_{\alpha_n} = \hat{a}_n x_{\alpha_n} + \hat{b}_n \quad (3.2)$$

De même que précédemment, pour le cas $\gamma = 0$, les auteurs proposent d'utiliser l'approche basée sur la loi GEV dont le résultat s'énonce ainsi.

Théorème 3.1.2. (*Weinstein (1973) [53]*). Soit $F \in \mathcal{D}(\text{Gumbel})$, 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 $\nu > 0$,

$$\lim_{n \rightarrow \infty} n \bar{F}[(b_n^\nu + c_n x)^{1/\nu}] = \exp(-x), \quad (3.3)$$

où $c_n = a_n \nu b_n^{\nu-1}$.

On approche alors le quantile par

$$q_{\alpha_n} \simeq (b_n^\nu + c_n x_{\alpha_n})^{1/\nu}$$

où x_{α_n} vérifie $\exp(-x_{\alpha_n}) = n\alpha_n$.

L'estimateur du quantile extrême est donc obtenu en remplaçant les suites b_n et c_n respectivement par leurs estimateurs \hat{b}_n et \hat{c}_n , (i.e)

$$\hat{q}_{\alpha_n} = \left(\hat{b}_n^\nu - \hat{c}_n \log(n\alpha_n) \right)^{1/\nu}.$$

Les paramètres γ , a_n , b_n et c_n peuvent être estimés par l'une des méthodes paramétriques.

3.1.2 L'estimation des quantiles extrêmes par la méthode des excès

Il est irréaliste de croire que seul le maximum de l'échantillon permet de modéliser le comportement des valeurs extrêmes. Les autres grandes valeurs de l'échantillon contiennent elles aussi de l'information sur la queue de distribution. L'approche par dépassements de seuil, en anglais "Peaks-Over-Threshold approach" notée **POT** est une alternative à la loi **GEV** dans la modélisation du comportement du maximum d'un échantillon se basant sur l'utilisation des statistiques d'ordre supérieur de l'échantillon.

Cette méthode consiste à utiliser que les observations dépassant un certain seuil déterministe. L'excès au-delà du seuil est défini comme l'écart entre l'observation et le seuil (i.e) : $(Y_i = X_i - u \geq 0)$.

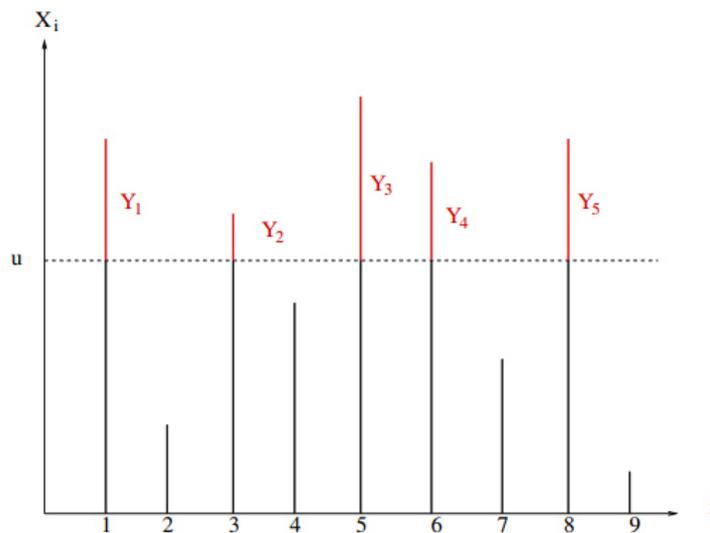


Figure 1.4- L'excès Y de la variable X au-delà du seuil u .

Nous supposons une suite d'observations X_1, \dots, X_n indépendantes et identiquement distribuées de fonction de répartition F et x_F un point terminal. Alors pour un seuil $u < x_F$ fixé (non aléatoire), on définit la fonction de distribution des excès de X au dessus du seuil u pour tout $0 \leq x \leq x_F - u$ par :

$$F_u(x) := \mathbb{P}(X - u \leq x / X > u) = \frac{F(x + u) - F(u)}{1 - F(u)} \quad (3.4)$$

ou de manière équivalente pour la fonction de survie :

$$\bar{F}_u(x) := \mathbb{P}(X - u > x / X > u) = 1 - F_u(x) = \frac{\bar{F}(x + u)}{\bar{F}(u)}. \quad (3.5)$$

La distribution de Pareto généralisée joue un rôle essentiel dans la modélisation des excès.

Définition 3.1.3. (*Loi de Pareto Généralisée (GPD)*).

La loi de Pareto généralisée est la loi dont la fonction de répartition est donnée par :

$$\mathcal{G}_{\gamma, \beta}(y) = \begin{cases} 1 - \left(1 + \gamma \frac{y}{\beta}\right)^{-1/\gamma} & \text{si } \gamma \neq 0 \text{ et } \beta > 0 \\ 1 - \exp\left(-\frac{y}{\beta}\right) & \text{si } \gamma = 0 \text{ et } \beta > 0 \end{cases}$$

Avec $y \in \mathbb{R}_+$ si $\gamma \geq 0$ ou $[0, -\beta/\gamma]$ si $\gamma < 0$. β représente le paramètre d'échelle et γ le paramètre de forme.

Exemple 3.1.1. "La loi exponentielle"

Soit $F(x) = (1 - e^{-x})\mathbb{1}_{\{x>0\}}(x)$ et

$$\mathbb{P}[X - u > y | X > u] = \frac{1 - F(u + y)}{1 - F(u)} = \frac{e^{-[u+y]}}{e^{-u}} = e^{-y}.$$

On retrouve la loi exponentielle qui est également la loi GPD de paramètre $\gamma = 0$ et $\beta_u = 1, \forall y > 0$.

Lorsque le seuil u est choisi grand, on approche $F_u(x)$ par la distribution d'une loi GPD. Afin d'approcher le quantile, il suffit alors d'utiliser le résultat de **Balkema** et de **Haan** (1974) [2] et **Pickands** (1975) [40] qui établit l'équivalence entre la convergence en loi du maximum (convenablement normalisé) vers une loi des valeurs extrêmes \mathcal{H}_γ et la convergence en loi d'un excès vers une GPD (loi de Pareto généralisée). Ce résultat s'énonce comme suit :

Théorème 3.1.3. (*Balkema et de Haan (1974) [2]; Pickands (1975) [40]*). Soit F_u la distribution des excès appartient au domaine d'attraction de \mathcal{H}_γ ; lorsque le seuil u tend vers le point terminal x_F , on a alors pour tout $\gamma \in \mathbb{R}$:

$$\lim_{u \rightarrow x_F} \sup_{0 \leq x \leq x_F - u} |F_u(x) - \mathcal{G}_{\gamma, \beta(u)}(x)| = 0. \quad (3.6)$$

Où $\beta(u)$ une fonction positive mesurable.

Alors on peut approcher le quantile pour tout $\gamma \neq 0$ par :

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

Et on a un estimateur du quantile extrême de type :

$$\hat{q}_{\alpha_n} = \frac{\left(\frac{k_n}{n\alpha_n} \right)^{\hat{\gamma}} - 1}{\hat{\gamma}} \hat{\beta} + \hat{u}, \quad (3.7)$$

on peut noter alors $\hat{\beta} = \hat{a}_n$ et $\hat{u} = \hat{b}_n$, d'après l'expression du quantile (3.2).

Les paramètres γ et β de la GPD peuvent être estimés par la méthode des moments ou la méthode du maximum de vraisemblance (**Smith** (1987) [46]; **Davison** et **Smith** (1990) [17]).

Remarque 3.1.3. 1. *Cette méthode présente un avantage par rapport à la méthode des blocs, en ce sens qu'il est plus facile d'avoir un échantillon d'excès que de maximums.*

2. *L'estimation des paramètres de la distribution **GPD** pose le problème de la détermination du seuil u . Ce dernier ne doit pas être trop grand mais suffisamment grand pour avoir le Théorème 3.1.3. En pratique, on remplace u par $X_{n-k_n+1,n}$ (k_n : désigne le nombre d'excès).*

3.1.3 L'estimation semi-paramétrique des quantiles extrêmes

Le plus connu des estimateurs de quantiles extrêmes basé sur une approche semi-paramétrique est l'estimateur de **Weissman** (1978) [54]. On verra qu'il repose sur l'estimation de $\gamma > 0$. Et comme estimateur de l'indice des valeurs extrêmes $\gamma > 0$ on cite déjà le plus connu d'entre eux qui est l'estimateur de Hill.

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

$$\bar{F}(x) = x^{-1/\gamma} L(x),$$

avec L une fonction à variations lentes à l'infini et $\gamma > 0$. Et d'après le Lemme 1.3.1, on a :

$$q_{\alpha_n} := \bar{F}^{\leftarrow}(\alpha_n) = \alpha_n^{-\gamma} L(1/\alpha_n) \quad \text{avec} \quad \alpha_n \leq 1/n, \quad (3.8)$$

$$q_{\beta_n} := \bar{F}^{\leftarrow}(\beta_n) = \beta_n^{-\gamma} L(1/\beta_n) \quad \text{avec} \quad \beta_n \geq 1/n. \quad (3.9)$$

En divisant (3.8) par (3.9) et à l'aide de la définition d'une fonction à variations lentes, si β_n suffisamment petit et $\alpha_n < \beta_n$ on obtient l'approximation suivante :

$$\bar{F}^{\leftarrow}(\alpha_n) = \bar{F}^{\leftarrow}(\beta_n) \left(\frac{\beta_n}{\alpha_n} \right)^{\gamma}. \quad (3.10)$$

En remplaçant $\overline{F}^{\leftarrow}(\beta_n)$ par son estimateur naturel $X_{n-\lfloor n\beta_n \rfloor+1,n}$ et γ par l'estimateur de Hill $\widehat{\gamma}_{k_n}^H$, on obtient l'estimateur de **Weissman** (1978) [54] défini par :

$$\widehat{q}_{\alpha_n}^W = X_{n-\lfloor n\beta_n \rfloor+1,n} \left(\frac{\beta_n}{\alpha_n} \right)^{\widehat{\gamma}_{k_n}^H}. \quad (3.11)$$

Le terme $(\beta_n/\alpha_n)^{\widehat{\gamma}_{k_n}^H}$ est le terme servant à l'extrapolation. Pour les propriétés de cet estimateur, on peut se référer à l'ouvrage de **Weissman** (1978) [54] et de **Haan et Ferreira** (2006) [19].

3.2 Estimation de quantiles extrêmes conditionnels

3.2.1 Modèle

Dans le but d'estimer des courbes de niveaux extrêmes dans le cas des lois à queues lourdes qui est équivalent à l'étude des quantiles conditionnels quand l'ordre du quantile converge vers un, on donne au préalable une définition utile dans cette partie.

Définition 3.2.1. (*courbes de niveaux extrêmes*) Soient $\{(X_i, Y_i), i = 1, \dots, n\}$ des copies indépendantes du couple aléatoire $(X, Y) \in \mathbb{R}^p \times \mathbb{R}$ où Y est une variable d'intérêt associée à une covariable X . Pour tout $x \in \mathbb{R}^p$ et pour toute suite réelle $\alpha_n \rightarrow 0$, les courbes de niveaux extrêmes sont définies comme les graphes de fonctions $x \mapsto q_{\alpha_n}(x) \in \mathbb{R}$ vérifiant

$$\mathbb{P}[(Y > q_{\alpha_n}(x))/X = x] = \alpha_n,$$

lorsque la fonction de survie conditionnelle de Y sachant $X = x$ est à variations régulières d'indice $-1/\gamma(x)$ à l'infini. Ceci signifie que pour tout $y > 0$,

$$\overline{F}(y/x) \stackrel{\text{def}}{=} 1 - F(y/x) = y^{-1/\gamma(x)} L(y/x), \quad (3.12)$$

avec $\gamma(\cdot)$ une fonction inconnue et positive de la covariable x que l'on appelle «indice de queue conditionnel» ou «indice des valeurs extrêmes conditionnel» et $L(\cdot/x)$ une fonction à variations lentes à l'infini (*i.e*) pour tout $x \in \mathbb{R}^p$ et $\lambda > 0$, on a :

$$\lim_{y \rightarrow \infty} \frac{L(\lambda y/x)}{L(y/x)} = 1, \quad (3.13)$$

dans le cas particulier où la loi conditionnelle de Y sachant $X = x$ est à queue lourde, (i.e) $F(\cdot/x) \in \mathcal{D}(\text{Fréchet})$.

3.3 La construction des estimateurs

L'estimateur naturel du quantile extrême conditionnel d'ordre α_n défini par :

$$\hat{q}_{\alpha_n}(x) \stackrel{\text{def}}{=} \hat{F}_n^{\leftarrow}(\alpha_n/x) = \inf\{t, \hat{F}_n(t/x) \leq \alpha_n\}. \quad (3.14)$$

Avec $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$, nécessite d'estimer la probabilité

$$\bar{F}(y/x) \rightarrow 0 \text{ lorsque } y \rightarrow \infty.$$

On peut utiliser la méthode de *la fenêtre mobile* introduite au chapitre 2, pour estimer la fonction de survie conditionnelle de Y sachant $X = x$. Tel que pour tout $(x, y) \in \mathbb{R}^p \times \mathbb{R}$, on a :

$$\hat{F}_n(y/x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{X_i \in B(x, h_n)\}}}{\sum_{i=1}^n \mathbb{1}_{\{X_i \in B(x, h_n)\}}} \mathbb{1}_{\{Y_i > y\}},$$

où $B(x, h_n)$ est une boule de centre x et de rayon h_n strictement positif. Cet estimateur a un inconvénient d'être discontinu par nature. Pour cela on utilise l'estimateur à noyau à cause de ses bonnes propriétés asymptotiques. Ce dernier est introduit par **Collomb** (1980) [12] et est défini pour tout $(x, y) \in \mathbb{R}^p \times \mathbb{R}$ par :

$$\hat{F}_n(y/x) = \frac{\sum_{i=1}^n \frac{K\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)} \mathbb{1}_{\{Y_i > y\}}, \quad (3.15)$$

où la fonction $K(\cdot)$ appelée *noyau* est une densité de probabilité sur \mathbb{R}^p et h_n est une suite non aléatoire telle que $h_n \rightarrow 0$ quand $n \rightarrow \infty$ appelée *paramètre de lissage*.

L'estimateur (3.15) peut aussi s'écrire sous cette forme :

$$\widehat{F}_n(y/x) = \frac{\widehat{\psi}_n(x, y)}{\widehat{g}_n(x)}, \quad (3.16)$$

où $\widehat{g}_n(x)$ est l'estimateur à noyau classique de la densité $g(x)$ et $\widehat{\psi}_n(x, y)$ est un estimateur de $\psi(x, y)$.

Dans le cadre d'estimation du réel $q_{\alpha_n}(x)$ en fonction de la vitesse de convergence de α_n vers zéro, on distingue deux situations :

(S.1) La suite α_n *converge lentement* vers zéro en ce sens que $nh_n^p \alpha_n \rightarrow \infty$ quand $n \rightarrow \infty$.

(S.2) La suite α_n *converge rapidement* vers zéro.

Dans la situation **(S.1)**, la condition $nh_n^p \alpha_n \rightarrow \infty$ quand $n \rightarrow \infty$ revient à supposer que le quantile $q_{\alpha_n}(x)$ ne tend pas trop vite vers l'infini quand $n \rightarrow \infty$. Dans une telle situation, on se propose d'estimer le quantile extrême conditionnel par (3.14).

Dans la situation **(S.2)** où il est supposé que le quantile $q_{\alpha_n}(x)$ tend vite vers l'infini quand $n \rightarrow \infty$, on retrouve l'estimateur de **Weissman** (1978) [54] (voir (3.11)) adapté au cas conditionnel donné par :

$$\widehat{q}_{\alpha_n}^W(x) = \widehat{q}_{\beta_n}(x) \left(\frac{\beta_n}{\alpha_n} \right)^{\widehat{\gamma}_n(x)}, \quad (3.17)$$

où β_n satisfait la situation **(S.1)** et $\widehat{\gamma}_n(x)$ est un estimateur de l'indice de queue conditionnel (indice des valeurs extrêmes conditionnel) et $\widehat{q}_{\beta_n}(x)$ c'est l'estimateur (3.14).

Un autre exemple des estimateurs du quantile extrême conditionnel est donné pour la loi **GPD** par analogie avec (3.7) comme suit :

$$\widehat{q}_{\alpha_n}(x) = \frac{\left(\frac{k_n}{n\alpha_n} \right)^{\widehat{\gamma}_n(x)} - 1}{\widehat{\gamma}_n(x)} \widehat{\beta}_n(x) + \widehat{u}_n(x).$$

3.4 Quelques résultats asymptotiques

3.4.1 Hypothèses

On présente ici toutes les conditions utiles pour établir la loi asymptotique de tous nos estimateurs. Dans la suite, on désigne par $d(x, x')$ la distance entre deux points $(x, x') \in \mathbb{R}^p \times \mathbb{R}^p$.

Hypothèses sur la fonction à variations lentes

(F.1) $L(\cdot/x)$ est normalisée.

(F.2) $|\Delta(\cdot/x)|$ est continue et asymptotiquement décroissante vers 0.

Une généralisation du Théorème 1.3.1 permet de réécrire la fonction à variations lentes à l'infini au cas conditionnel sous la forme suivante :

$$L(y/x) = c(x) \exp \left(\int_1^y \frac{\Delta(u/x)}{u} du \right), \quad (3.18)$$

avec $c(\cdot)$ une fonction positive et $\Delta(y/x) \rightarrow 0$ quand $y \rightarrow \infty$. Ceci implique que la fonction $L(\cdot/x)$ est dérivable et pour tout $x \in \mathbb{R}^p$, la fonction $\Delta(y/x)$ est définie par

$$\Delta(y/x) = y \frac{L'(y/x)}{L(y/x)}.$$

D'après la Définition 1.3.3 du Chapitre 1, la fonction à variations lentes au cas conditionnel est définie par (3.13).

Certains auteurs supposent que la fonction biais $\Delta(y/x)$ est une fonction à variations régulières d'indice $\rho < 0$ à l'infini.

L'hypothèse **(F.2)** connue sous le nom de *condition du second ordre* dont le but est de contrôler le comportement de la fonction de survie conditionnelle de Y sachant $X = x$ par rapport à sa première variable.

Hypothèses de régularité lipschitzienne

(L.1) Il existe $c_\gamma > 0$ tel que $\left| \frac{1}{\gamma(x)} - \frac{1}{\gamma(x')} \right| \leq c_\gamma d(x, x')$.

(L.2) Il existe $c_l > 0$ et $y_0 > 1$ tels que $\sup_{y \geq y_0} \left| \frac{\log L(y/x)}{\log y} - \frac{\log L(y/x')}{\log y} \right| \leq c_l d(x, x')$.

(L.3) Il existe $c_g > 0$ tel que $|g(x) - g(x')| \leq c_g d(x, x')$.

Les hypothèses **(L.1)** et **(L.2)** nous assurent la régularité de la loi conditionnelle de Y sachant $X = x$.

Hypothèse sur le noyau

(K.1) K est une fonction positive, bornée, intégrable et à support compact $S \subseteq \mathbb{R}^p$.

Les hypothèses **(L.3)** et **(K.1)** sont des hypothèses classiques sur l'estimation non-paramétrique par la méthode du noyau.

3.4.2 Résultats préliminaires

Lemme 3.4.1. *Supposons que les hypothèses **(L.1)** et **(L.2)** satisfaites. Si $y_n \rightarrow \infty$ et $h_n \log y_n \rightarrow 0$ quand $n \rightarrow \infty$ alors*

$$\sup_{d(x, x') \leq h_n} \left| \frac{\overline{F}(y_n/x)}{\overline{F}(y_n/x')} - 1 \right| = O(h_n \log y_n).$$

Le Lemme 3.4.1 montre que sous les conditions de Lipschitz, le rapport de deux fonctions de survies conditionnelles évaluées aux points $(x, x') \in \mathbb{R}^p \times \mathbb{R}^p$ converge uniformément vers un lorsque n tend vers l'infini.

Lemme 3.4.2. *Supposons que les hypothèses **(F.1)** et **(F.2)** satisfaites.*

(i) *Soient $(\alpha_n)_{n \geq 1}$ et $(\beta_n)_{n \geq 1}$ deux suites positives telles que $0 < \beta_n < \alpha_n$ avec $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$. Alors*

$$\left| \log q_{\beta_n}(x) - \log q_{\alpha_n}(x) + \gamma(x) \log \left(\frac{\beta_n}{\alpha_n} \right) \right| = O \left(\log \left(\frac{\beta_n}{\alpha_n} \right) \Delta(q_{\alpha_n}(x)/x) \right).$$

(ii) Si de plus $\liminf \left(\frac{\beta_n}{\alpha_n} \right) > 0$, alors

$$\frac{\beta_n^{\gamma(x)} q_{\beta_n}(x)}{\alpha_n^{\gamma(x)} q_{\alpha_n}(x)} = 1 + O(\Delta(q_{\alpha_n}(x)/x)).$$

Le Lemme 3.4.2 montre que le quantile extrême conditionnel est à décroissance polynomiale d'indice $-\gamma(x)$ puisque $\Delta(q_{\alpha_n}(x)/x) \rightarrow 0$ quand $n \rightarrow \infty$.

Lemme 3.4.3. *Supposons que les hypothèses (L.1), (L.2) et (L.3) satisfaites. Considérons la région de \mathbb{R}^{p+1} définie par $R_n(x) = B(x, h_n) \times (q_{\alpha_n}(x), \infty)$ où $x \in \mathbb{R}^p$ est tel que $g(x) > 0$. Si $h_n \log q_{\alpha_n}(x) \rightarrow 0$ quand $n \rightarrow \infty$, alors $\mathbb{P}(\exists i \in \{1, \dots, n\}, (X_i, Y_i) \in R_n(x)) \rightarrow 1$ quand $n \rightarrow \infty$ si et seulement si, $nh_n^p \alpha_n \rightarrow \infty$.*

Le Lemme 3.4.3 nous montre qu'il n'est pas logique d'utiliser (3.14) pour estimer le quantile extrême conditionnel dans la situation (S.2) où $nh_n^p \alpha_n$ peut converger vers une constante.

Remarque 3.4.1. – *On peut donc remplacer $R_n(x) = B(x, h_n) \times (q_{\alpha_n}(x), \infty)$ par $R_n^*(x) = B(x, h_n) \times (y_n, \infty)$ et $nh_n^p \alpha_n \rightarrow \infty$ par $nh_n^p \bar{F}(y_n/x) \rightarrow \infty$, puisque α_n converge vers zéro et alors la condition $nh_n^p \alpha_n \rightarrow \infty$ est équivalente à la condition $nh_n^p \bar{F}(y_n/x) \rightarrow \infty$ où la quantité d'intérêt $y_n \rightarrow \infty$.*
– *La condition $nh_n^p \bar{F}(y_n/x) \rightarrow \infty$ est une condition nécessaire et suffisante pour qu'il y ait presque sûrement au moins un point dans la région $R_n^*(x)$ de \mathbb{R}^{p+1} .*

3.4.3 La normalité asymptotique des estimateurs

On a déjà vu que l'estimateur (3.15) peut se réécrire sous la forme (3.16) où

$$\hat{\psi}_n(x, y) = \frac{1}{nh_n^p} \sum_{i=1}^n K \left(\frac{x - X_i}{h_n} \right) \mathbb{1}_{\{Y_i > y\}} \quad (3.19)$$

et

$$\hat{g}_n(x) = \frac{1}{nh_n^p} \sum_{i=1}^n K \left(\frac{x - X_i}{h_n} \right). \quad (3.20)$$

Le comportement asymptotique de l'estimateur $\widehat{g}_n(x)$ a été étudié par **Collomb** (1980) [12]. Il montre que $\widehat{g}_n(x) \xrightarrow{\mathbb{P}} g(x)$.

Lemme 3.4.4. (*Collomb (1980) [12], Propositions 2.1 et 2.2*). *Supposons que les hypothèses (L.3) et (K.1) sont satisfaites. Si $nh_n^p \rightarrow \infty$, alors, pour tout $x \in \mathbb{R}^p$,*

(i) $\mathbb{E}(\widehat{g}_n(x) - g(x)) = O(h_n),$

(ii) $Var(\widehat{g}_n(x)) = \frac{g(x)\|K\|_2^2}{nh_n^p}(1 + o(1)).$

Le lemme suivant montre que pour tout $x \in \mathbb{R}^p$, l'estimateur $\widehat{\psi}_n(y_n/x)$ est asymptotiquement gaussien et sans biais lorsque la suite y_n ne tend pas trop vite vers l'infini.

Lemme 3.4.5. *Supposons que les hypothèses (L.1), (L.2), (L.3) et (K.1) soient vérifiées. Soit $(y_n)_{n \geq 1}$ une suite positive satisfaisant les conditions $y_n \rightarrow \infty$, $h_n \log y_n \rightarrow 0$ et $nh_n^p \overline{F}(y_n/x) \rightarrow \infty$ quand $n \rightarrow \infty$. Et soit $\{a_j, j = 1, \dots, J\}$ une suite strictement positive et croissante. Alors, pour tout $x \in \mathbb{R}^p$ tel que $g(x) > 0$, on a :*

(i) $\left(\mathbb{E}(\widehat{\psi}_n(a_j y_n/x)) \right)_{\{j=1, \dots, J\}} = (\psi(a_j y_n/x)(1 + O(h_n \log y_n)))_{\{j=1, \dots, J\}}.$

(ii) $\left(\sqrt{nh_n^p \psi(y_n/x)} \left(\frac{\widehat{\psi}_n(a_j y_n/x) - \mathbb{E}(\widehat{\psi}_n(a_j y_n/x))}{\psi(a_j y_n/x)} \right) \right)_{\{j=1, \dots, J\}} \xrightarrow{\mathcal{L}} \mathcal{N}(0_{\mathbb{R}^J}, \|K\|_2^2 C(x)),$

où $C_{j,j'}(x) = a_{j \wedge j'}^{1/\gamma(x)}$ pour $(j, j') \in \{1, \dots, J\}^2$.

De plus, sous les conditions de Lipschitz énoncées précédemment, on a

$$\frac{\widehat{\psi}_n(y_n/x)}{\psi(y_n/x)} \xrightarrow{\mathbb{P}} 1.$$

La normalité asymptotique de l'estimateur $\widehat{F}_n(\cdot/x)$ est donnée dans le théorème suivant :

Théorème 3.4.1. *Supposons que les hypothèses (L.1), (L.2), (L.3) et (K.1) soient vérifiées. Soit $(y_n)_{n \geq 1}$ une suite positive satisfaisant les conditions $y_n \rightarrow \infty$, $nh_n^p \overline{F}(y_n/x) \rightarrow \infty$ et $nh_n^{p+2} \log^2(y_n) \overline{F}(y_n/x) \rightarrow 0$ quand $n \rightarrow \infty$.*

Et soit $\{a_j, j = 1, \dots, J\}$ une suite strictement positive et croissante (où J est un entier positif). Alors, pour tout $x \in \mathbb{R}^p$ tel que $g(x) > 0$, on a :

$$\left(\sqrt{nh_n^p \bar{F}(y_n/x)} \left(\frac{\widehat{F}_n(a_j y_n/x)}{\bar{F}(a_j y_n/x)} - 1 \right) \right)_{\{j=1, \dots, J\}} \xrightarrow{\mathcal{L}} \mathcal{N} \left(0_{\mathbb{R}^J}, \frac{\|K\|_2^2}{g(x)} C(x) \right),$$

où $C_{j,j'}(x) = a_{j \wedge j'}^{1/\gamma(x)}$ pour $(j, j') \in \{1, \dots, J\}^2$.

Remarque 3.4.2. – Si y_n est borné alors on retrouve la condition de normalité asymptotique classique $nh_n^p \rightarrow \infty$.

– La condition $nh_n^{p+2} \log^2(y_n) \bar{F}(y_n/x) \rightarrow 0$ est une condition pour que le carré du biais asymptotique de l'ordre de $(h_n \log y_n)^2$ soit négligeable devant la variance asymptotique de l'ordre de $\frac{1}{nh_n^p \bar{F}(y_n/x)}$.

Preuve du Théorème 3.4.1

En gardant à l'esprit les notations du Lemme 3.4.5, il s'ensuit les développements suivants :

$$\Lambda_n^{-1}(x) \sum_{j=1}^J \beta_j \left(\frac{\widehat{F}_n(y_{n,j}/x)}{\bar{F}(y_{n,j}/x)} - 1 \right) = \frac{\Delta_{1,n} + \Delta_{2,n} - \Delta_{3,n}}{\widehat{g}_n(x)}, \quad (3.21)$$

avec $\Lambda_n(x) = (nh_n^p \bar{F}(y_n/x))^{-1/2}$, et

$$\Delta_{1,n} = g(x) \Lambda_n^{-1}(x) \sum_{j=1}^J \beta_j \left(\frac{\widehat{\psi}_n(y_{n,j}/x) - \mathbb{E}(\widehat{\psi}_n(y_{n,j}/x))}{\psi(y_{n,j}/x)} \right),$$

$$\Delta_{2,n} = g(x) \Lambda_n^{-1}(x) \sum_{j=1}^J \beta_j \left(\frac{\mathbb{E}(\widehat{\psi}_n(y_{n,j}/x)) - \psi(y_{n,j}/x)}{\psi(y_{n,j}/x)} \right),$$

$$\Delta_{3,n} = \left(\sum_{j=1}^J \beta_j \right) \Lambda_n^{-1}(x) (\widehat{g}_n(x) - g(x)),$$

et les suites $(y_{n,j})$ vérifiant : $\max_{j \in \{1, \dots, J\}} \left| \frac{y_{n,j}}{y_n} - 1 \right| \rightarrow 0$. Commençons par remarquer que les hypothèses $nh_n^{p+2} \log^2(y_n) \bar{F}(y_n/x) \rightarrow 0$ et $nh_n^p \bar{F}(y_n/x) \rightarrow \infty$ impliquent que

$h_n \log y_n \rightarrow 0$ quand $n \rightarrow \infty$. Ainsi, à l'aide de point (ii) du Lemme 3.4.5, le terme aléatoire $\Delta_{1,n}$ peut se réécrire comme

$$\Delta_{1,n} = g(x) \|K\|_2 \sqrt{\beta^t C(x) \beta} \xi_n, \quad (3.22)$$

où ξ_n converge vers une loi normale centrée réduite. Le terme déterministe $\Delta_{2,n}$ est contrôlé par le point (i) du Lemme 3.4.5 tel que :

$$\Delta_{2,n} = O(\Lambda_n^{-1}(x) h_n \log y_n) = O(nh_n^{p+2} \log^2(y_n) \bar{F}(y_n/x))^{1/2} = o(1). \quad (3.23)$$

Finalement, $\Delta_{3,n}$ est un terme classique lorsque l'on estime une densité par la méthode du noyau. Il peut être borné par le Lemme 3.4.4 :

$$\begin{aligned} \Delta_{3,n} &= O(h_n \Lambda_n^{-1}(x)) + O_p(\Lambda_n^{-1}(x) (nh_n^p)^{-1/2}) \\ &= O(nh_n^{p+2} \bar{F}(y_n/x))^{1/2} + O_p(\bar{F}(y_n/x))^{1/2} \\ &= o_p(1). \end{aligned} \quad (3.24)$$

En combinant (3.21)-(3.24), on obtient :

$$\hat{g}_n(x) \Lambda_n^{-1}(x) \sum_{j=1}^J \beta_j \left(\frac{\widehat{F}_n(y_{n,j}/x)}{\bar{F}(y_{n,j}/x)} - 1 \right) = g(x) \|K\|_2 \sqrt{\beta^t C(x) \beta} \xi_n + o_p(1).$$

D'après le Lemme 3.4.4, on a $\hat{g}_n(x) \xrightarrow{\mathbb{P}} g(x)$, il en découle que

$$\sqrt{nh_n^p \bar{F}(y_n/x)} \sum_{j=1}^J \beta_j \left(\frac{\widehat{F}_n(y_{n,j}/x)}{\bar{F}(y_{n,j}/x)} - 1 \right) = \|K\|_2 \sqrt{\frac{\beta^t C(x) \beta}{g(x)}} \xi_n + o_p(1)$$

ce qui prouve le théorème énoncé.

Maintenant, on s'intéresse à la normalité asymptotique de l'estimateur du quantile extrême conditionnel dans la situation (S.1).

Théorème 3.4.2. *Supposons que les hypothèses (F.1), (L.1), (L.2), (L.3) et (K.1) soient vérifiées. Soit $(\alpha_n)_{n \geq 1}$ une suite positive satisfaisant les conditions $\alpha_n \rightarrow 0$, $nh_n^p \alpha_n \rightarrow \infty$ et $nh_n^{p+2} \alpha_n \log^2(\alpha_n) \rightarrow 0$ quand $n \rightarrow \infty$ et soit $\{\tau_j, j = 1, \dots, J\}$*

une suite strictement positive et décroissante. Alors, pour tout $x \in \mathbb{R}^p$ tel que $g(x) > 0$, on a :

$$\left(\sqrt{nh_n^p \alpha_n} \left(\frac{\widehat{q}_n(\tau_j \alpha_n / x)}{q(\tau_j \alpha_n / x)} - 1 \right) \right)_{\{j=1, \dots, J\}} \xrightarrow{\mathcal{L}} \mathcal{N} \left(0_{\mathbb{R}^J}, \gamma^2(x) \frac{\|K\|_2^2}{g(x)} \Sigma \right),$$

où $\Sigma_{j,j'} = \frac{1}{\tau_{j \wedge j'}}$ pour $(j, j') \in \{1, \dots, J\}^2$.

Dans la situation **(S.1)**, la variance asymptotique étant inversement proportionnelle à $nh_n^p \alpha_n$. Aussi, cette variance est proportionnelle à $\gamma^2(x)$, ceci implique que l'estimation de $q(\alpha_n/x)$ est plus difficile pour des grandes valeurs de l'indice de queue conditionnel (i.e) Si $\gamma(x)$ augmente (queue "plus lourde"), la variance des estimateurs augmente.

Preuve du Théorème 3.4.2

Introduisons pour $j = 1, \dots, J$, les termes suivants :

$$\sigma_{n,j}(x) = q(\alpha_{n,j}/x)(nh_n^p \alpha_n)^{-1/2}.$$

$$\nu_{n,j}(x) = \alpha_{n,j}^{-1} \gamma(x)(nh_n^p \alpha_n)^{1/2}.$$

$$W_{n,j}(x) = \nu_{n,j}(x) \left(\widehat{F}_n(q(\alpha_{n,j}/x) + \sigma_{n,j}(x)z_j/x) - \overline{F}(q(\alpha_{n,j}/x) + \sigma_{n,j}(x)z_j/x) \right).$$

$$a_{n,j}(x) = \nu_{n,j}(x) \left(\alpha_{n,j} - \overline{F}(q(\alpha_{n,j}/x) + \sigma_{n,j}(x)z_j/x) \right).$$

Et $z_j \in \mathbb{R}^J$. On veut établir la loi asymptotique de la fonction J-variée définie par

$$\begin{aligned} \Phi_n(z_1, \dots, z_J) &= \mathbb{P} \left(\bigcap_{j=1}^J \{ \sigma_{n,j}^{-1}(x) (\widehat{q}_n(\alpha_{n,j}/x) - q(\alpha_{n,j}/x)) \leq z_j \} \right) \\ &= \mathbb{P} \left(\bigcap_{j=1}^J \{ \widehat{F}_n(q(\alpha_{n,j}/x) + \sigma_{n,j}(x)z_j/x) \leq \alpha_{n,j} \} \right) \\ &= \mathbb{P} \left(\bigcap_{j=1}^J \{ W_{n,j}(x) \leq a_{n,j}(x) \} \right). \end{aligned}$$

Concentrons nous d'abord sur le terme déterministe $a_{n,j}(x)$. D'après la propriété (3.12) et l'hypothèse (F.2), la représentation de Karamata (3.18) montre que $\overline{F}(\cdot/x)$ est dérivable. Ainsi, pour chaque $j \in \{1, \dots, J\}$ il existe $\theta_{n,j} \in]0, 1[$ tel que

$$\overline{F}(q(\alpha_{n,j}/x)/x) - \overline{F}(q(\alpha_{n,j}/x) + \sigma_{n,j}(x)z_j/x) = -\sigma_{n,j}(x)z_j\overline{F}'(q_{n,j}/x), \quad (3.25)$$

où $q_{n,j} = q(\alpha_{n,j}/x) + \theta_{n,j}\sigma_{n,j}(x)z_j$. Il est ainsi évident que $q(\alpha_{n,j}/x) \rightarrow \infty$ et $\frac{\sigma_{n,j}(x)}{q(\alpha_{n,j}/x)} \rightarrow 0$ quand $n \rightarrow \infty$. Par conséquent, $q_{n,j} \rightarrow \infty$ et la représentation de Karamata entraînent que

$$\lim_{n \rightarrow \infty} \frac{q_{n,j}\overline{F}'(q_{n,j}/x)}{\overline{F}(q_{n,j}/x)} = -\frac{1}{\gamma(x)}. \quad (3.26)$$

De plus, puisque $q_{n,j} \sim q(\alpha_{n,j}/x)$ quand $n \rightarrow \infty$ et $\overline{F}(\cdot/x)$ est à variations régulières, il s'en suit que $\overline{F}(q_{n,j}/x) \sim \overline{F}(q(\alpha_{n,j}/x)/x) = \alpha_{n,j}$. Au vu des deux équations (3.25) et (3.26), il en découle que

$$a_{n,j}(x) = \frac{\nu_{n,j}(x)\sigma_{n,j}(x)\alpha_{n,j}z_j}{\gamma(x)q(\alpha_{n,j}/x)}(1 + o(1)) = z_j(1 + o(1)). \quad (3.27)$$

Concentrons nous maintenant à la variable aléatoire $W_{n,j}(x)$. En définissant $a_j = \tau_j^{-\gamma(x)}$, $y_{n,j} = q(\alpha_{n,j}/x) + \sigma_{n,j}(x)z_j$ pour $j = 1, \dots, J$ et $y_n = q(\alpha_n/x)$, nous avons $y_{n,j} \sim q(\alpha_{n,j}/x) \sim a_j y_n$ puisque $q(\cdot/x)$ est une fonction à variations régulières d'indice $-\gamma(x)$. En utilisant le même raisonnement, il est facile de montrer que $\log y_n \sim -\gamma(x) \log \alpha_n$. Par conséquent, le Théorème 3.4.1 s'applique et le vecteur aléatoire

$$\left(\frac{\sqrt{nh_n^p \overline{F}(y_n/x)}}{\nu_{n,j}(x)\overline{F}(y_{n,j}/x)} W_{n,j} \right)_{\{j=1, \dots, J\}} = (1 + o(1)) \left(\frac{W_{n,j}}{\gamma(x)} \right)_{\{j=1, \dots, J\}}$$

converge vers une variable aléatoire gaussienne centrée de matrice de variance-covariance $\frac{\|K\|_2^2}{g(x)} C(x)$. En tenant compte de l'équation (3.27), on obtient que

$\Phi_n(z_1, \dots, z_J)$ converge vers une fonction de répartition d'une loi normale centrée de matrice de variance-covariance $\frac{\|K\|_2^2 \gamma^2(x)}{g(x)} C(x)$ évaluée en (z_1, \dots, z_J) , qui est le résultat attendu.

Théorème 3.4.3. *Supposons que les hypothèses (F.1), (L.1), (L.2), (L.3) et (K.1) soient vérifiées. Soit $(\beta_n)_{n \geq 1}$ une suite positive satisfaisant les conditions $\beta_n \rightarrow 0$, $nh_n^p \beta_n \rightarrow \infty$ et $nh_n^{p+2} \beta_n \log^2(\beta_n) \rightarrow 0$ quand $n \rightarrow \infty$. Considérons $\hat{\gamma}_n(x)$ un estimateur de l'indice de queue tel que*

$$\sqrt{nh_n^p \beta_n} (\hat{\gamma}_n(x) - \gamma(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \nu^2(x)) \quad \text{avec } \nu(x) > 0. \quad (3.28)$$

Si de plus $(\alpha_n)_{n \geq 1}$ est une suite positive tendant vers zéro quand $n \rightarrow \infty$ et $\frac{\alpha_n}{\beta_n} \rightarrow 0$ alors, pour tout $x \in \mathbb{R}^p$, on a :

$$\frac{\sqrt{nh_n^p \beta_n}}{\log\left(\frac{\beta_n}{\alpha_n}\right)} \left(\frac{\hat{q}_{\alpha_n}^W(x)}{q_{\alpha_n}(x)} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \nu^2(x)).$$

Et alors la loi asymptotique de $\hat{q}_{\alpha_n}^W(x)$ dépend du comportement de $\hat{\gamma}_n(x)$.

Remarque 3.4.3. *La condition $\frac{\alpha_n}{\beta_n} \rightarrow 0$ permet d'extrapoler et donc de choisir un ordre α_n arbitrairement petit.*

Preuve du Théorème 3.4.3

La preuve de ce résultat est basée sur la décomposition suivante :

$$\frac{\sqrt{nh_n^p \beta_n}}{\log\left(\frac{\beta_n}{\alpha_n}\right)} (\log(\hat{q}_{\alpha_n}^W(x)) - \log(q_{\alpha_n}(x))) = \frac{\sqrt{nh_n^p \beta_n}}{\log\left(\frac{\beta_n}{\alpha_n}\right)} (Q_{n,1} + Q_{n,2} + Q_{n,3}),$$

avec :

$$\begin{aligned}
Q_{n,1} &= \sqrt{nh_n^p \beta_n} (\widehat{\gamma}_n(x) - \gamma(x)), \\
Q_{n,2} &= \frac{\sqrt{nh_n^p \beta_n}}{\log\left(\frac{\beta_n}{\alpha_n}\right)} \log\left(\frac{\widehat{q}_{\beta_n}(x)}{q_{\beta_n}(x)}\right), \\
Q_{n,3} &= \frac{\sqrt{nh_n^p \beta_n}}{\log\left(\frac{\beta_n}{\alpha_n}\right)} \left(\log q_{\beta_n}(x) - \log q_{\alpha_n}(x) + \gamma(x) \log\left(\frac{\beta_n}{\alpha_n}\right) \right).
\end{aligned}$$

Traisons les trois termes séparément. *Premièrement*, sous les hypothèses du Théorème 3.4.3, on a $Q_{n,1} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \nu^2(x))$. *Deuxièmement*, le Théorème 3.4.2 implique que $\frac{\widehat{q}_{\beta_n}(x)}{q_{\beta_n}(x)} \xrightarrow{\mathbb{P}} 1$ quand $n \rightarrow \infty$ et ainsi :

$$Q_{n,2} = \frac{\sqrt{nh_n^p \beta_n}}{\log\left(\frac{\beta_n}{\alpha_n}\right)} \left(\frac{\widehat{q}_{\beta_n}(x)}{q_{\beta_n}(x)} - 1 \right) (1 + o_p(1)) = \frac{O_p(1)}{\log\left(\frac{\beta_n}{\alpha_n}\right)}.$$

Comme conséquence, $Q_{n,2} \xrightarrow{\mathbb{P}} 0$ quand $n \rightarrow \infty$. *Troisièmement*, d'après le point (i) du Lemme 3.4.2, on a $Q_{n,3} = O\left(\sqrt{nh_n^p \beta_n} \Delta(q_{\beta_n}(x)/x)\right)$ qui converge vers 0 puisque par hypothèse $\frac{\Delta(q_{\beta_n}(x)/x)}{\sigma_n} \stackrel{\text{def}}{=} \sqrt{nh_n^p \beta_n} \Delta(q_{\beta_n}(x)/x) \rightarrow 0$. Ce qui achève la preuve.

3.5 L'estimation de l'indice de queue conditionnel

La construction des estimateurs pour l'indice de queue conditionnel nous permet de donner des intervalles de confiance du quantile extrême conditionnel $q_{\alpha_n}(x)$ et de pouvoir utiliser en pratique l'estimateur $\widehat{q}_{\alpha_n}^W(x)$. Le premier estimateur de $\gamma(x)$ que nous proposons est basé sur trois statistiques d'ordre.

Définition 3.5.1. Soit $(\beta_n)_{n \geq 1}$ une suite positive telle que $\beta_n \rightarrow 0$ et $nh_n^p \beta_n \rightarrow \infty$ quand $n \rightarrow \infty$. Un estimateur à noyau de type **Pickands** (1975) [40] est donné par :

$$\widehat{\gamma}_n^P(x) = \frac{1}{\log 2} \log\left(\frac{\widehat{q}_{\beta_n}(x) - \widehat{q}_{2\beta_n}(x)}{\widehat{q}_{2\beta_n}(x) - \widehat{q}_{4\beta_n}(x)}\right).$$

Lemme 3.5.1. *Supposons la condition (F.2) vérifiée. Sous les hypothèses du Théorème 3.4.2, si $\sqrt{nh_n^p \beta_n} \Delta(q_{2\beta_n}(x)/x) \rightarrow 0$ quand $n \rightarrow \infty$, alors*

$$\sqrt{nh_n^p \beta_n} (\hat{\gamma}_n^P(x) - \gamma(x)) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \frac{\|K\|_2^2}{g(x)} \frac{\gamma^2(x) (2^{2\gamma(x)+1} + 1)}{4(\log 2)^2 (2^{\gamma(x)} - 1)^2} \right).$$

Puisque la variance asymptotique de l'estimateur $\hat{\gamma}_n^P(x)$ est assez importante pour des grandes valeurs de l'indice de queue conditionnel, il nous paraît préférable de proposer un estimateur à noyau de variance inférieure.

Définition 3.5.2. *Soit $(\beta_n)_{n \geq 1}$ une suite positive telle que $\beta_n \rightarrow 0$ et $nh_n^p \beta_n \rightarrow \infty$ quand $n \rightarrow \infty$. Un estimateur à noyau de type **Hill** (1975) [35] est donné pour tout $J > 1$ par :*

$$\hat{\gamma}_n^H(x) = \frac{\sum_{j=1}^J [\log \hat{q}_n(\tau_j \beta_n / x) - \log \hat{q}_n(\tau_1 \beta_n / x)]}{\sum_{j=1}^J \log \left(\frac{\tau_1}{\tau_j} \right)},$$

où $(\tau_j)_{j \geq 1}$ est une suite de poids strictement positive et décroissante (i.e) :

$$1 = \tau_1 > \tau_2 > \dots > \tau_J > 0.$$

Lemme 3.5.2. *Supposons la condition (F.2) vérifiée. Sous les hypothèses du Théorème 3.4.2, si $\sqrt{nh_n^p \beta_n} \Delta(q(\tau_1 \beta_n / x)/x) \rightarrow 0$ quand $n \rightarrow \infty$, alors*

$$\sqrt{nh_n^p \beta_n} (\hat{\gamma}_n^H(x) - \gamma(x)) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \gamma^2(x) \frac{\|K\|_2^2}{g(x)} V_J \right),$$

$$\text{où } V_J = \frac{\left(\sum_{j=1}^J \frac{2(J-j)+1}{\tau_j} - J^2 \right)}{\left(\sum_{j=1}^J \log \left(\frac{\tau_1}{\tau_j} \right) \right)^2}.$$

Pour construire un intervalle de confiance du quantile $\hat{q}_{\alpha_n}^W(x)$ ou $\hat{q}_{\alpha_n}(x)$, il suffit de remplacer $\gamma(x)$ et $g(x)$ par leur estimateur respectif.

Conclusion

Dans ce mémoire nous avons proposé des estimateurs de quantiles extrêmes dans le cadre conditionnel c'est-à-dire dans la situation où la variable d'intérêt Y est mesurée simultanément avec une covariable X . Néanmoins, le cas sans covariable a été considéré dans la première partie du troisième chapitre.

Après avoir rappelé certaines des notions clés en théorie des valeurs extrêmes d'un échantillon de variables aléatoires réelles indépendantes et identiquement distribuées et proposé des estimateurs de l'indice de queue dans le premier chapitre, nous nous sommes intéressés dans le deuxième chapitre à l'estimation non-paramétrique des quantiles conditionnels pour les deux méthodes : indirecte et directe

Dans le dernier chapitre, nous nous sommes focalisés sur l'étude des valeurs extrêmes d'un échantillon d'observations indépendantes dont la loi conditionnelle de Y en un point x de la covariable $X \in \mathbb{R}^p$ appartient au \mathcal{D} (Fréchet). La variable d'intérêt Y a toujours été supposée aléatoire et réelle. Concernant la nature des covariables X , nous prenons seulement le cas aléatoire, nous avons introduit un estimateur de petites probabilités conditionnelles, deux estimateurs des quantiles extrêmes conditionnels selon la vitesse de convergence de α_n vers zéro et deux estimateurs de l'indice de queue conditionnel. Et nous avons établi la normalité asymptotique de tous ces estimateurs.

Bibliographie

- [1] Antoch, J. et Janssen, P. (1989). Nonparametric regression m-quantiles. *Statistics and Probability Letters*, 8(4) :355-362. 48
- [2] Balkema, A. et de Haan, L. (1974). Residual life time at a great age. *Annals of Probability*, 2(5) :792-804. 17,21
- [3] Bassett, G. et Koenker, R. (1978). Regression quantiles. *Econometrica*, pages 33-50. 54
- [4] Beirlant, J, Goegebeur, Y. (2003) Regression with response distributions of paratype. *Computational Statistics and Data Analysis*, 42 :595-619. (Cit  en page 22.)
- [5] Berlinet, A. Gannoun, A. et Matzner-Lober, E. (2001). Asymptotic normality of convergent estimates of conditional quantiles. *Statistics*, 18 :1400-1415. 50, 52, 95, 97
- [6] Bhattacharyya, P. K. et Gangopadhyay, A. K. (1990). Kernel and nearest-neighbor estimation of a conditional quantile. *Annals of Statistics*, 8(3) :1400-1415.51
- [7] Bingham, N. H., Goldie, C. M. et Teugels, J. L. (1987). *Regular Variation*, volume 27. *Encyclopedia of Mathematics and its Applications*, Cambridge University Press.
- [8] Cardot, H., Crambes, C. and Sarda, P. (2005). Quantiles regression when the covariates are functions. *Journal of Nonparametric Statistics*, 17, 841-856.
- [9] Chauduri, P. (1991). Nonparametric estimates of regression quantiles and their local Bahadur representation. *Annals of Statistics*, 19(2) :760-777. 52

-
- [10] Chavez-Demoulin, V. et Davison, A. C. (2005). Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society, C*, 54(1) :207-222. 54, 63, 88
- [11] Cole, T. J. (1988). Fitting smoothed centile curves to reference data. *Journal of Royal Statistical Society, A*, 151(3) :385-418.
- [12] Collomb, G. (1980). Estimation non paramétrique de probabilités conditionnelles. *Comptes Rendus de l'Académie des Sciences*, 291 : 427-430. 47, 48, 90, 94, 116
- [13] Csörgö, S. et Mason, D. (1985). Central limit theorems for sums of extreme values. *Mathematical Proceedings of the Cambridge Philosophical Society*, 98(3) :547-558. 25
- [14] A. Daouia, L. Gardes, and S. Girard (2013). On kernel smoothing for extremal quantile regression, *Bernoulli*, 19, 2557-2589.
- [15] Davis, R. et Resnick, S. (1984). Tail estimates motivated by extreme value theory. *Annals of Statistics*, 12(4) :1467-1487. 25
- [16] Davison, A. et Ramesh, N. I. (2000). Local likelihood smoothing of sample extremes. *Journal of the Royal Statistical Society, B*, 62(1) :191-208. 54, 63, 88
- [17] Davison, A. C. et Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, B*, 52(3) :393-442. 22, 54, 63, 88
- [18] Deheuvels, P., Haeusler, E. et Mason, D. (1988). Almost sure convergence of the Hill estimator. *Mathematical Proceedings of the Cambridge Philosophical Society*, 104(2) : 371-381. 25
- [19] De Haan, L. et Ferreira, A. (2006). *Extreme Value Theory : An Introduction*. Springer Series in Operations Research and Financial Engineering, New York Inc. 16, 19, 24
- [20] Dekkers, A. et de Haan, L. (1989). On the estimation of the extreme value index and large quantile estimation. *Annals of Statistics*, 17(4) :1795-1832. 23, 26, 28, 58
- [21] Embrechts, P., Klüppelberg, C. et Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer Verlag. 11, 13, 17, 23, 24, 58, 84, 113

- [22] Ferraty, F., Laksaci, A., Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statist. Infer. Stoch. Process* 9 :47-76.
- [23] Ferraty, F., Rabhi, A. et Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Nio phenomenon. *Indian Journal of Statistics*, 67(2) :378-398. 50, 126
- [24] Ferraty, F. et Vieu, P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *Comptes Rendus de l'Académie des Sciences*, 330 :403-406. 48
- [25] Ferraty, F. et Vieu, P. (2006). *Nonparametric Functional Data Analysis : Theory and Practice*. Springer Series in Statistics, Springer. 48, 50, 54, 56, 68
- [26] Fisher, R. et Tippett, L. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24 :180-190. 10, 11, 17
- [27] Gannoun, A. (1989). Estimation de la médiane conditionnelle. Thèse de doctorat, Université de Paris VI. 48, 51
- [28] Gardes, L. et Girard, S. (2010). Conditional extremes from heavy-tailed distributions : an application to the estimation of extreme rainfall return levels. *Extremes*, 13(2) :177-204. 56, 74
- [29] Gardes, L. Girard, S. Lekina, A. (2010) Functional nonparametric estimation of conditional extreme quantiles. *Journal of Multivariate Analysis*, 101 :419-433. (Cité en pages xix, 23, 81 et 134.)
- [30] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44(3) :423-453. 10, 11, 17
- [31] Guida, M. et Longo, M. (1988). Estimation of probability tails based on generalized extreme value distributions. *Reliability Engineering and System Safety*, 20(3) :219-242.
- [32] Gumbel, E.J. (1954), *Statistical Theory of Extreme Values and Some Practical Applications*, Applied Mathematics Series, 33, National Bureau of Standards, Washington.

- [33] Hart, J. D. (1991). Comment to "choosing a kernel regression estimator". *Statistical Sciences*, 6(4) :425-427. 48
- [34] He, X. et Shi, P. (1994). Convergence rate of B-spline estimators of non parametric conditional quantiles functions. *Journal of Nonparametric Statistics*, 3(3-4) :299-308.51
- [35] Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5) :1163-1174. 17, 23, 25, 32, 64, 98, 99, 129
- [36] Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *The Quarterly Journal of the Royal Meteorological Society*, 81(384) :158-171.
- [37] Kratz, M. et Resnick, S. (1996). The qq-estimator and heavy tails. *Stochastic Models*, 12(4) :699-724. 24, 28, 65
- [38] Mason, D. (1982). Laws of large numbers for sums of extreme values. *Annals of Probability*, 10(3) :754-764. 25
- [39] Mint El Mouvit, L. (2000). Sur l'estimateur linéaire local de la fonction de répartition conditionnelle. Thèse de doctorat, Université de Montpellier 2.
- [40] Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1) :119-131. 17, 21, 24, 26, 31, 98
- [41] Resnick, S. (1987). *Extreme Values, regular Variation, and Point Process*. Springer Verlag, New-York. 13, 14, 15
- [42] Roussas, G. G. (1969). Nonparametric estimation of the transition distribution function of Markov process. *Annals of Mathematical Statistics*, 40(4) :1386-1400. 46, 48
- [43] Royston, P. et Wright, E. M. (1998). How to construct normal ranges for fetal variables. *Ultrasound in Obstetrics and Gynecology*, 11(1) :30-38
- [44] Samanta, T. (1989) Non-parametric estimation of conditional quantiles. *Statistics and Probability Letters*, 7, 407-412.
- [45] Schultze, J. et Steinebach, J. (1996). On least squares estimates of an exponential tail coefficient. *Statistics and Decisions*, 14(3) :353-372. 24, 28, 65

- [46] Smith, R. (1987). Estimating tails of probability distributions. *Annals of Statistics*, 15(3) : 1174-1207. 22, 25
- [47] Stone, C. J. (1977). Consistent nonparametric regression. *Annals of Statistics*, 5(4) :595-645. 47, 52
- [48] Stute, W. (1986). Conditional empirical processes. *Annals of Statistics*, 14(2) :638-647.
- [49] Tango, T. (1998). Estimation of age-specific reference ranges via smoother avas. *Statistics in Medicine*, 17(11) :1231-1243. 46
- [50] Truong, Y. K. (1989). Asymptotic properties of kernel estimators based on local medians. *Annals of Statistics*, 17(2) :606-617.
- [51] Tukey, J. W. (1961). Curves as parameters, and touch estimation. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1 :681-694.
- [52] Von Mises, R. (1936). La distribution de la plus grande de n valeurs. *Revue de Mathématique Union Interbalcanique*, 1 :141-160.
- [53] Weinstein, S. B. (1973). Theory and application of some classical and generalized asymptotic distributions of extreme values. *IEEE Transactions on Information Theory*, 19(2) :148-154.
- [54] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364) :812-815.
- [55] Yu, K. et Jones, M. (1998). Local linear quantile regression. *Journal of the American Statistical Association*, 93 :228-237. 52