



République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES

Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : **Statistique**

Par

BACHIR NAOUAL

Titre :

Application des Valeurs Extrêmes

Membres du Comité d'Examen :

Dr. BENAMEUR SANA	UMKB	Présidente
Pr. Brahimi Brahim	UMKB	Encadreur
Dr. BERKANE HASSIBA	UMKB	Examinateur

Septembre 2020

DÉDICACE

Je dédie ce modeste travail :

A mon grand amour, ma mère qui a sacrifié sa vie pour notre bonheur et notre réussite.

A mon père à qui je témoigne de l'affection, et du respect.

À mon cher ami Melili Larbi, pour avoir toujours été un soutien pour moi dans ma vie et ma carrière universitaire.

À mes chers frères Amar, Lazhar, Chaouki, Smati.

À mes chères sœurs Nassira, Fatima, Elkhamsa, Jamila.

A toute ma famille.

À tous mes amis

A tout les enfants Boutchicha Merame , Alraouf, Lilyane, Lejane.

NAOUAL BACHIR

REMERCIEMENTS

Avant tout, je tiens à remercier **ALLAH** le tout puissant qui m'a donné la santé, le courage et la patience pour arriver à ce jour.

Je tiens tout d'abord à remercier sincèrement mon encadreur le professeur **Brahimi Brahim**, d'avoir bien assuré la direction et l'encadrement de ce mémoire. Merci pour votre gentillesse, votre patience, votre simplicité et vos précieux conseils. Vraiment merci beaucoup pour la deuxième chance que vous me donnez pour continuer mon parcours scientifique.

Mes remerciements vont également au membres de jury : **Dr.BENAMEUR SANA** et **Dr.BERKANE HASSIBA** pour avoir accepté de juger ce mémoire.

Je remercie aussi l'ensemble des enseignants et des professeurs du département de mathématiques.

Je souhaite particulièrement remercier mes parents qui m'ont stimulé et encouragé pendant mes études. qui étaient toujours prêts à fournir tous les moyens physiques et morales pour la réussite de ce projet.

J'adresse mes remerciement à tous mes amies.

MERCI

Table des matières

Table des matières	i
Table des matières	ii
Table des matières	iii
Liste des figures	v
Liste des tableaux	vi
Introduction	1
1 Théorie des valeurs extrêmes	3
1.1 Statistique d'ordre	3
1.1.1 Distribution d'une statistique d'ordre	4
1.2 Distributions des valeurs extrêmes	6
1.3 Distribution GEV	8
1.3.1 Domaine d'attraction	10
1.3.2 Les fonctions à variations régulières	10
1.3.3 Caractérisations des domaines d'attraction	12
1.4 Distribution GPD	15
2 Estimation de l'indice des valeurs extrêmes	20
2.1 Estimation paramétrique	21
2.1.1 Estimateur de maximum de vraisemblance	21

2.2	Estimation par des méthodes semi-paramétrique	23
2.2.1	Estimateur de Pickands $\gamma \in \mathbb{R}$	23
2.2.2	Estimateur de Hill	25
2.2.3	Estimateur du moment	29
2.2.4	Le choix du nombre k	30
3	Application des valeurs extrêmes	33
3.1	Données réelles : Danish fire	34
3.2	Modèle composite	35
3.3	Résultats et discussion	37
	Conclusion	41
	Bibliographie	42
	Annexe A : Logiciel <i>R</i>	45
	Annexe B : Abréviations et Notations	47

Table des figures

1.1	Densités et Distribution de Lois des Valeurs Extrêmes :Gumbel($\gamma = 0$), Fréchet ($\gamma = 1$) et Weibull ($\gamma = -1$)	10
1.2	Représentation graphique de la densité et distribution de lois Paréto généralisée	19
2.1	Représentation graphique de l'estimateur de Pickands	25
2.2	Représentation graphique de l'estimation de Hill	30
2.3	Représentation graphique de l'estimateur de Moment	31
2.4	Le graphe de $(k, \hat{\gamma}_{k(n),n}^H)$ pour les différents estimateur de γ	31
3.1	Données danoises : mensuelles (à gauche) et hebdomadaires (à droite).	35
3.2	Fonctions de densité de probabilité ajustées pour les données danoises de l'assurance incendie : noir pour le modèle composite lognormal-Pareto et rouge pour le modèle composite lognormal-Burr.	39
3.3	probabilité	39
3.4	Graphiques quantile-quantiles pour les données danoises de l'assurance-incendie : noir pour le modèle composite lognormal-Pareto et rouge pour le modèle composite lognormal-Burr.	40

Liste des tableaux

1.1	Exemple sur les lois classées selon leurs domaines d'attraction	15
3.1	Estimations des paramètres pour les modèles composites ajustés aux données danoises de l'assurance incendie	38

Introduction

La théorie des valeurs extrêmes (*TVE*) ou "Extreme Value Theory" en anglais, est une branche de la théorie des probabilités et des statistiques mathématiques qui se concentre sur l'analyse et l'inférence sur les événements extrêmes, c'est-à-dire les événements avec très faible probabilité d'occurrence. Les événements extrêmes revêtent une grande importance dans presque tous les domaines de la science et de la technologie, car ils peuvent s'avérer catastrophiques et donc très coûteux. Cette théorie a été développée dans le contexte d'observations indépendantes ; les auteurs Fisher et Tippett (1928) [13] ; montrent que sous certaines conditions, les seules distributions limites des extrêmes sont les lois de Fréchet, Gumbel et Weibull.

Ceci nous permet de classer la plupart des lois en trois domaines d'attraction où chaque domaine est déterminé par des caractérisations sur les fonctions de répartition. Von Mises (1954) [26] puis Jenkinson (1955) [19] ont rassemblé les distributions de ces trois domaines en une seule écriture (voir Embrechts et al (1997) [12], de Haan et Ferreira (2006) [9]). Cependant, on y retrouve deux modèles : loi des valeurs extrêmes généralisée (*GEV* : Generalized Extreme Value) et loi de Pareto généralisée (*GPD* : Generalized Pareto Distribution). Divers travaux ont été consacrés à l'estimation de l'indice des extrêmes dont l'objectif revient à construire des estimateurs et étudier leurs propriétés ; citons Hill (1975) [17] et Pickands (1975) [21]. La plupart des estimateurs reposent sur l'utilisation de la statistique d'ordre.

Les modèles des valeurs extrêmes sont appliquées à une grande variété de problèmes tels l'environnement (vitesse du vent, extrêmes pluviométriques et de températures,...), la pluviométrie, la finance et l'assurance (Mesure du risque, Valeur à risque VaR, modèle de volatilité

stochastique) et en hydrologie pour calculer la probabilité que la hauteur d'eau d'un fleuve dépasse un certain seuil, télécommunications, physique,...

Pour la littérature concernant la théorie des valeurs extrêmes nous référons aux ouvrages de Embrechts et al (1997) [12], Reiss et Thomas (2007) [22] et de Hann et Ferreira (2006) [9].

Les méthodes statistiques les plus avancées pour les événements extrêmes seront étudiées à la fois du côté théorique et du côté des applications. Pour ce dernier, le logiciel d'analyse statistique *R* sera utilisé.

Ce mémoire est subdivisé en trois chapitres comme suit :

Au premier chapitre, on présente une introduction sur les statistiques d'ordre, nous donnons un aperçu des définitions et des résultats théoriques essentiels de la *TVE*, nous présentons les résultats sur le comportement du maximum, un rappel sur les deux principaux outils servant à modéliser le comportement des valeurs extrêmes d'un échantillon : la loi des valeurs extrêmes généralisée (*GEV*) et la loi des excès (*GPD*). Nous nous intéressons ensuite à la caractérisation des domaines d'attraction et les fonctions à variation régulière.

Au deuxième chapitre, nous présentons les estimateurs de l'indice des valeurs extrêmes les plus utilisés avec une attention particulière à l'estimation paramétrique et semi-paramétrique (Hill, Moment, Pickands parmi autres). Dans l'estimation de l'IVE on distingue deux modèles d'estimation : la première est basée sur la distribution des valeurs extrêmes généralisée et la deuxième est basée sur la distribution de Pareto généralisée. Nous concentrons dans ce mémoire sur l'estimateur de Hill qui est l'estimateur le plus populaire pour l'indice de queue positif de plus dans la majorité des applications réelles les ensembles de données correspondent aux distributions appartenant au domaine d'attraction de Fréchet.

Au troisième chapitre, nous présentons les applications des valeurs extrêmes dans l'assurance en appliquant les données de danish fire, et définitions des modèles composites pour le danish fire data (voir S. Nadarajah et al [25]).

Chapitre 1

Théorie des valeurs extrêmes

Les valeurs extrêmes sont les valeurs plus fortes (ou plus faible) d'une série d'observations, l'étude de ces valeurs revient à l'analyse de la plus grande observation d'un échantillon. En ce sens nous pouvons considérer la théorie des valeurs extrêmes, qui admet d'extrapoler le comportement des queues de distribution à partir des plus grandes observations.

Dans ce chapitre nous rappelons quelques notions essentielles sur la théorie des valeurs extrêmes (*TVE*). Nous commençons par rappeler la définition des statistiques d'ordre, qui est très utile en *TVE* et nous étudions le comportement des valeurs extrêmes d'un échantillon de variables aléatoires, nous définissons rapidement les caractérisations des domaines d'attractions. Nous donnons des critères pour que la limite en loi du maximum suive une loi des valeurs extrêmes. Ces critères faisant appel à la notion de fonctions à variations régulières. Pour plus de détails sur ce chapitre, il se trouve des bonnes références sur la théorie et les applications des valeurs extrêmes suivantes : Embrechts et al (1997) [12], David et Nagaraja (2003) [6], Beirlant et al. (2006) [4].

1.1 Statistique d'ordre

Le point de départ de toutes les méthodes statistiques est la notion d'échantillon donné. Il est des fois nécessaires d'étudier la suite des valeurs observées par valeurs croissante ou

décroissante.

Définition 1.1 Soit (X_1, X_2, \dots, X_n) une suite de variables aléatoires réelles, indépendantes et identiquement distribuées (i.i.d), de fonction de répartition F , telle que :

$$F(x) = P(X \leq x), \text{ pour tout } x \in \mathbb{R}$$

On appelle statistique d'ordre le réarrangement croissant de (X_1, X_2, \dots, X_n) notée :

$$X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$$

Pour $1 \leq k \leq n$; la variable aléatoire (va) $X_{k,n}$ est appelée la $k^{\text{ème}}$ statistique d'ordre. En particulier, la statistique du minimum et du maximum sont respectivement données par :

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

et

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

En générale, les résultats des minimaux peuvent être tirés du maximum par la relation suivante :

$$\min(X_1, X_2, \dots, X_n) = -\max(-X_1, -X_2, \dots, -X_n).$$

1.1.1 Distribution d'une statistique d'ordre

Distribution du maximum et du minimum

La fonction de répartition du maximum $X_{n,n}$ et du minimum $X_{1,n}$ de l'échantillon (X_1, X_2, \dots, X_n) sont données respectivement par :

$$\begin{aligned}
 F_{X_{n,n}}(x) &= P(X_{n,n} \leq x) & (1.1) \\
 &= P(\max(X_1, X_2, \dots, X_n) \leq x) \\
 &= P(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \\
 &= P\left[\bigcap_{i=1}^n (X_i \leq x)\right] \\
 &= \prod_{i=1}^n P(X_i \leq x) \\
 &= [F(x)]^n
 \end{aligned}$$

et

$$\begin{aligned}
 F_{X_{1,n}}(x) &= P(X_{1,n} \leq x) & (1.2) \\
 &= 1 - P(X_{1,n} > x) \\
 &= 1 - P\left[\bigcap_{i=1}^n (X_i > x)\right] \\
 &= 1 - \prod_{i=1}^n [P(x_i > x)] \\
 &= 1 - [1 - F(x)]^n
 \end{aligned}$$

Remarque 1.1 *Pour obtenir la densité du maximum et du minimum, on dérive respectivement 1.1 et 1.2, alors :*

$$f_{X_{n,n}}(x) = nf(x) \{F(x)\}^{n-1}$$

$$f_{X_{1,n}}(x) = nf(x) \{1 - F(x)\}^{n-1}$$

Proposition 1.1 *La fonction de répartition de la $k^{\text{ème}}$ statistique d'ordre $X_{k,n}$ pour $1 \leq k \leq$*

n , est donnée par :

$$F_{X_{k,n}}(x) = \sum_{i=k}^n C_n^i \{F(x)\}^i \{1 - F(x)\}^{n-i}, -\infty < x < +\infty.$$

De plus, si F admet une densité f , alors :

$$f_{X_{k,n}}(x) = \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1} (1 - F(x))^{n-k}, -\infty < x < +\infty.$$

Preuve. Voir David et Nagaraja (2003) [6] ■

Théorème 1.1 Si on suppose l'échantillon X_1, X_2, \dots, X_n de va's i.i.d solen une loi de densité f alors la densité jointe des n statistiques d'ordre $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ et donnée par :

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, x_3, \dots, x_n) = n! \prod_{i=1}^n f(x_i), -\infty \leq x \leq +\infty.$$

1.2 Distributions des valeurs extrêmes

La distribution du maximum devrait nous fournir des informations sur des événements extrêmes, et comme la fonction de répartition de X n'étant pas souvent connue, il n'est généralement pas possible de déterminer cette distribution à partir (1.1). On s'intéresse alors à la distribution asymptotique du maximum. Alors, on a :

$$\lim_{n \rightarrow +\infty} F_{X_{n,n}}(x) = \lim_{n \rightarrow +\infty} \{F(x)\}^n = \begin{cases} 1 & \text{si } F(x) = 1 \\ 0 & \text{si } F(x) < 1 \end{cases}$$

On constate que la distribution asymptotique du maximum donne une loi dégénérée (Elle prend les valeurs 0 et 1 seulement). On cherche alors une loi non dégénérée pour $X_{n,n}$.

De façon analogue au théorème centrale limite (*TCL*), Fisher et Tippett (1928)[13], Gnedenko (1943)[14] et de Hann (1970) ont trouvé une solution à ce problème qui précise les seules distributions limites non dégénérées H possibles sont les distributions de valeurs extrêmes. Ce théorème donne un résultat important pour l'existence d'une loi limite non dégénérée

pour le maximum.

Théorème 1.2 (Fisher et Tippett(1928), Gnedenko(1940)) Soit X_1, X_2, \dots, X_n n va's i.i.d de fonction de répartition F . S'il existe deux suites normalisantes réelles $(a_n)_{n>0}$, et $(b_n)_{n\geq 0} \in \mathbb{R}$, alors :

$$\lim_{n \rightarrow +\infty} P\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow +\infty} F^n(a_n x + b_n) = H(x), \forall x \in \mathbb{R}.$$

où H est une fonction de distribution non dégénérée, appelée *distribution des valeurs extrêmes* (EVD : *Extremes values Distribution*). La distribution H est du même type que l'une des trois distributions des valeurs extrêmes standard suivantes :

$$\text{Distribution de Gumbel : } \Lambda(x) := \exp\{-\exp(-x)\}; x \in \mathbb{R}.$$

$$\text{Distribution de Fréchet : } \Phi_\alpha(x) := \begin{cases} 0 & \text{si } x \leq 0, \\ \exp(-x^{-\alpha}) & \text{si } x > 0. \end{cases} ; \alpha > 0.$$

$$\text{Distribution de Weibull : } \Psi_\alpha(x) := \begin{cases} \exp\{-(-x)^\alpha\} & \text{si } x \leq 0, \\ 1 & \text{si } x > 0. \end{cases} ; \alpha < 0.$$

Preuve. Une preuve détaillée de ce théorème peut être trouvée dans Embrechts et al (1997) [12] ■

Remarque 1.2 (Relation entre Λ, Φ_γ , et Ψ_γ) On peut passer de l'une des trois lois à l'autre comme suit :

$$Y \sim \Phi_\gamma \Leftrightarrow -Y^{-1} \sim \Psi_\gamma \Rightarrow \ln Y^\gamma \sim \Lambda.$$

Exemple 1.1 (Loi exponentielle) Soit X une va suit la loi exponentielle standard de fonction de répartition $F(x) = 1 - \exp(-x)$. Prenons $a_n = 1$ et $b_n = \log(n)$, alors $Y = \frac{X_{n,n} - b_n}{a_n}$

tend asymptotiquement vers la loi de Gumbel. En effet :

$$\begin{aligned}
 F_y(x) &= P(X_{n,n} - \log(n) \leq x) = P(X_{n,n} \leq \log(n) + x) \\
 &= F_{X_{n,n}}(\log(n) + x) \\
 &= F^n(\log(n) + x) \\
 &= [1 - \{\exp(-(\log(n) + x))\}]^n \\
 &= \left[1 - \frac{e^{-x}}{n}\right]^n \\
 &= \exp\left[n \log\left(1 - \frac{e^{-x}}{n}\right)\right] \approx \exp[-\exp(-x)] \text{ quand } n \rightarrow +\infty \\
 &= \Lambda(x)
 \end{aligned}$$

1.3 Distribution GEV

Etant donnée qu'il est difficile de travailler avec les trois distributions limites, Jenkinson (1955)[19] et Von-Mises (1954)[26] ont proposés une famille paramétrique appelée distribution des valeurs extrêmes généralisées standard *GEV* (Generalized Extreme Value), qui permet d'unifier les trois types des lois extrêmes si dessus, la fonction H_γ devient :

$$H_\gamma(x) := \begin{cases} \exp(-(1 + \gamma x))^{\frac{-1}{\gamma}} & \text{si } \gamma \neq 0, 1 + \gamma x > 0, \\ \exp\{-\exp(x)\} & \text{si } \gamma = 0, x \in \mathbb{R}. \end{cases}$$

où γ s'appelle indice des valeurs extrêmes (E V I : Extreme Value Index).

La fonction de densité de probabilité h_γ associée est définie par :

$$h_\gamma(x) := \begin{cases} H_\gamma(x)(1 + \gamma x)^{\frac{-1}{\gamma-1}} & \text{si } \gamma \neq 0, 1 + \gamma x > 0, \\ \exp[-x - \exp(-x)] & \text{si } \gamma = 0, x \in \mathbb{R}. \end{cases}$$

pour $\gamma = 0$, il faut lire $H_0(x) = \exp[-\exp(-x)]$, $x \in \mathbb{R}$, qui s'obtient dans H_γ en faisant tendre γ vers 0. Les lois des valeurs extrêmes généralisées correspondent, à une translation et changement d'échelle près, au loi des valeurs extrêmes standard.

Nous avons alors les correspondances suivantes :

$$H_0(x) = \Lambda(x) \quad \text{si } \alpha = 0,$$

$$H_{\frac{1}{\gamma}}(\alpha(x-1)) = \Phi(x) \quad \text{si } \alpha > 0,$$

$$H_{\frac{-1}{\gamma}}(\alpha(x+1)) = \Psi(x) \quad \text{si } \alpha < 0.$$

Pour les variables non centrées et non réduites, on peut écrire $H_\gamma(x)$ sous une forme plus générale dite forme paramétrée de Von Mises, dans laquelle on fait apparaître un paramètre de localisation $\mu \in \mathbb{R}$ et un paramètre d'échelle $\sigma > 0$, telle que :

$$H_{\mu,\sigma,\gamma}(x) := \begin{cases} \exp \left\{ -\left(1 + \gamma \left(\frac{x-\mu}{\sigma}\right)^{\frac{-1}{\gamma}}\right) \right\} & \text{si } \gamma \neq 0, 1 + \gamma \left(\frac{x-\mu}{\sigma}\right) > 0 \\ \exp \left\{ -\exp\left(-\frac{x-\mu}{\sigma}\right) \right\} & \text{si } \gamma = 0, x \in \mathbb{R}. \end{cases}$$

Cette approche est apportée aux données qui consistent en un ensemble des maximums par blocs, dite "Block Maxima" : annuels journaliers, semestriels journaliers, trimestriels journaliers, etc.

La densité de la *GEV* s'écrit pour $\gamma \neq 0$ comme suit :

$$h_{\mu,\sigma,\gamma}(x) := \frac{1}{\sigma} \left[1 + \gamma \frac{x-\mu}{\sigma} \right]_{\mu,\sigma,\gamma}^{-(\frac{1+\gamma}{\gamma})} H(x).$$

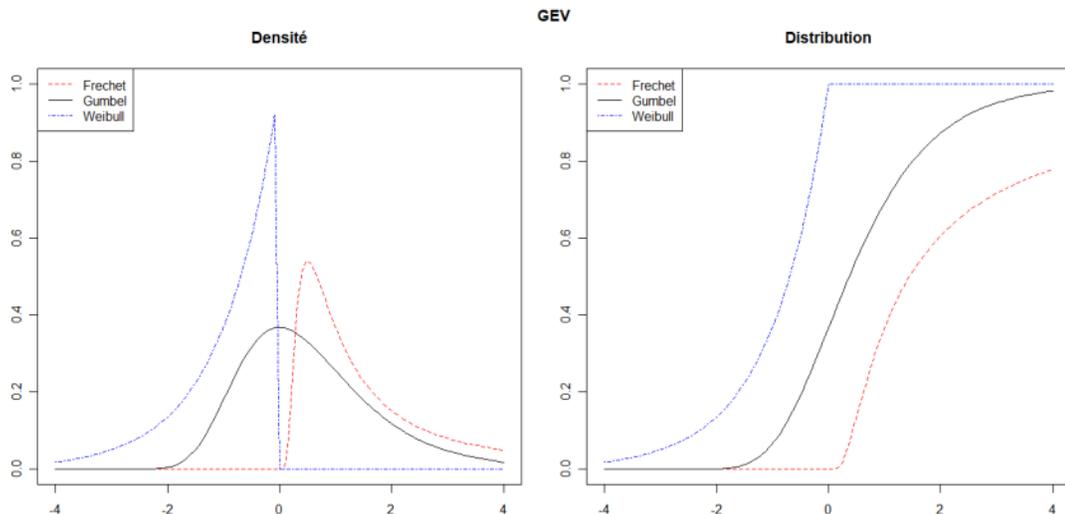


FIG. 1.1 – Densités et Distribution de Lois des Valeurs Extrêmes :Gumbel($\gamma = 0$), Fréchet ($\gamma = 1$) et Weibull ($\gamma = -1$)

1.3.1 Domaine d'attraction

Dans cette partie de la section, nous établirons les conditions nécessaires et suffisantes qui doivent être vérifiées par une fonction de distribution F , pour que la loi du maximum converge vers H_γ .

Définition 1.2 (Domaine d'attraction) *On dit qu'une distribution appartient au domaine d'attraction de H_γ (notée $F \in D(H_\gamma)$), si la distribution du maximum normalisé converge vers H_γ . Autrement dit, s'il existe des constantes réelles $a_n > 0$ et $b_n \in \mathbb{R}$ tels que*

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

1.3.2 Les fonctions à variations régulières

Dans cette partie, nous allons définir quelques notions essentielles de la théorie des valeurs extrêmes permettant de caractériser les domaines d'attraction. Cependant, ces fonctions à variations régulières permettent d'avoir une écriture unique pour chaque domaine d'attraction.

Définition 1.3 (Fonction à variation régulière) *Une fonction h positive à l'infini (c-à-d : s'il existe A tel que $x \geq A, h(x) > 0$) est dite à variation régulière à l'infini d'indice $\rho \in \mathbb{R}$*

et on note $h \in RV_\rho$ si et seulement si

$$\lim_{t \rightarrow +\infty} \frac{h(tx)}{h(t)} = x^\rho, \quad x > 0.$$

où : RV_ρ l'ensemble des fonctions à variations régulières d'indice ρ .

Définition 1.4 (Fonction à variation lente) Si une fonction $L : \mathbb{R} \mapsto [0, +\infty[$ est à variation régulière d'indice 0 ($L \in RV_0$), on dit que L est à variation lente, telle que :

$$\lim_{t \rightarrow +\infty} \frac{L(tx)}{L(t)} = 1, \quad x > 0.$$

Exemple 1.2 La fonction logarithme est une fonction à variation lente . En effet, soit $t > 0$

$$\lim_{t \rightarrow +\infty} \frac{\log(tx)}{\log(x)} = 1 + \lim_{t \rightarrow +\infty} \frac{\log(t)}{\log(x)} = 1.$$

Remarque 1.3 Tout fonction $h \in RV_\rho$ peut s'écrire sous la forme :

$$h(x) := x^\rho L(x), \quad \text{où } L \in RV_0$$

Proposition 1.2 (Représentation de Karamata) soit $h \in RV_\rho$, $\rho \in \mathbb{R}$. Il existe deux fonctions mesurable $c > 0$ et ε telles que :

$$\lim_{x \rightarrow +\infty} c(x) = c_0 \in]0, +\infty[\quad \text{et} \quad \lim_{x \rightarrow +\infty} \varepsilon(x) = \rho$$

et $x_0 > 0$ tels que pour tout $x \geq x_0$,

$$L(x) = c(x) \left\{ \int_{x_0}^x \frac{\varepsilon(u)}{u} du \right\}.$$

Remarque 1.4 Dans le cas où la fonction $c(\cdot)$ est constante, la fonction $L(\cdot)$ correspondante est dite normalisée.

1.3.3 Caractérisations des domaines d'attraction

Proposition 1.3 (Caractérisation de $D(H)$) $F \in D(H_\gamma)$ si et seulement si, pour une certaine suite $a_n > 0$ et $b_n \in \mathbb{R}$, on a

$$\lim_{n \rightarrow +\infty} n\bar{F}(a_n x + b_n) = -\log H_\gamma(x), \quad x \in \mathbb{R}.$$

avec \bar{F} est la fonction de survie (encore appelée queue de distribution), définie par :

$$\bar{F}(x) := 1 - F(x).$$

En pratique, il est souvent plus commode, non pas de travailler sur la fonction F elle-même, mais sur la fonction quantile de queue.

Définition 1.5 (Inverse généralisée) On appelle inverse généralisée (fonction des quantiles) de la fonction F , l'application notée F^{\leftarrow} définie par :

$$Q(p) = F^{\leftarrow}(p) := \inf \{x : F(x) \geq p\}, \quad p \in [0, 1].$$

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

$$U(t) := Q\left(1 - \frac{1}{t}\right) = F^{\leftarrow}\left(1 - \frac{1}{t}\right), \quad \forall t > 1.$$

où Q est la fonction des quantiles associée à F .

Définition 1.7 (Point terminal) Le point terminal de la fonction de répartition F est donné par

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

Théorème 1.3 (Caractérisation du $D(H_\gamma)$) Pour $\gamma \in \mathbb{R}$; les affirmations suivantes sont équivalentes :

– $F \in D(H_\gamma)$

– Il existe une fonction mesurable a telle que pour $1 + \gamma x > 0$, on a :

$$\lim_{u \rightarrow x_F} \frac{\overline{F}((a(u)x + u))}{\overline{F}(u)} = \begin{cases} (1 + \gamma x)^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0, \\ \exp(-x) & \text{si } \gamma = 0. \end{cases}$$

– Pour $x, y > 0$ et $y \neq 1$, on a

$$\lim_{s \rightarrow \infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^\gamma - 1}{y^\gamma - 1} & \text{si } \gamma \neq 0, \\ \frac{\log x}{\log y} & \text{si } \gamma = 0. \end{cases} \quad (1.3)$$

Preuve. Voir Embrechts et al.[12] ■

Théorème 1.4 (Caractérisation du $D(\Phi_\gamma)$) La fonction de répartition F appartient au domaine d'attraction de la loi de Fréchet de paramètre $\gamma > 0$ si et seulement si $x_F = +\infty$ et

$$\overline{F}(x) = x^{-\gamma} L(x),$$

où L est une fonction à variation lente. Dans ce cas en choisit $a_n = U(n) = F^{\leftarrow}(1 - \frac{1}{n})$ et $b_n = 0, \forall n > 0$. la suite $(a_n^{-1} X_{n,n})_{n \geq 1}$ converge en loi vers une variable aléatoire de fonction de répartition Φ_γ , quand $n \rightarrow \infty$.

Exemple 1.3 Soit X_1, X_2, \dots, X_n ; une suite de va's i.i.d de loi de Paréto de paramètre $\gamma > 0$, de fonction de répartition $F(x) := 1 - x^{-\frac{1}{\gamma}}$, pour $a_n = F^{\leftarrow}(1 - \frac{1}{n}) = n^\gamma$ et $b_n = 0$, on a :

$$\begin{aligned} \lim_{n \rightarrow +\infty} F^n(a_n x + b_n) &= \lim_{n \rightarrow +\infty} P\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) \\ &= \lim_{n \rightarrow +\infty} \left(1 - \frac{x^{-\frac{1}{\gamma}}}{n}\right)^n \\ &= \exp\left(-x^{-\frac{1}{\gamma}}\right) \\ &= \Phi_\gamma(x). \end{aligned}$$

Donc la loi limite est une loi de Fréchet $F \in D(\Phi_\gamma)$.

Remarque 1.5 Toutes les distributions appartenant au domaine d'attraction de Fréchet sont

dites de type Paréto et leur queue de distribution \bar{F} s'écrit, pour x très grand comme suit :

$$\bar{F}(x) = Cx^{-\gamma}, \quad C, \gamma > 0.$$

Théorème 1.5 (Caractérisation du $D(\Psi_\gamma)$) La fonction de répartition F appartient au domaine d'attraction de la loi de Weibull ($\gamma < 0$) si et seulement si $x_F < +\infty$ et

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

pour certains fonction à variation lente L .

Dans ce cas on peut choisir les constantes de normalisation $a_n = x_F - U(n) = x_F - F^{-1}(1 - \frac{1}{n})$ et $b_n = x_F, \forall n > 0$, alors : la suite $(a_n^{-1}(X_{n,n} - x_F))_{n \geq 1}$ converge en loi vers une va de fonction de répartition Ψ_γ , quand $n \rightarrow +\infty$.

Exemple 1.4 Soit X_1, X_2, \dots, X_n une suite de va's i.i.d de loi uniforme sur $[0, 1]$, de fonction de répartition $F(x) = x$, pour $a_n = \frac{1}{n}$ et $b_n = 1$, on a :

$$\begin{aligned} \lim_{n \rightarrow +\infty} F^n(a_n x + b_n) &= \lim_{n \rightarrow +\infty} (1 + \frac{1}{n}x)^n \\ &= \exp(x) \\ &= \exp(-(-x)) \end{aligned}$$

Donc la loi limite est une loi de Weibull $F \in D(\Psi_1)$.

Définition 1.8 (Fonction de von-Mises) Soit F une fonction de répartition de point terminale x_F , on suppose qu'il existe $z < x < x_F$, tel que :

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

où $c > 0$ et a est une fonction positive absolument continue avec la densité à vérifiant $\lim_{x \rightarrow x_F} a(x) = 0$.

Alors F est appelé fonction de von-Mises et a est une fonction auxiliaire de F .

Théorème 1.6 (Caractérisation du $D(\Lambda)$) *La fonction de répartition F appartient au domaine d'attraction de la loi de Gumbel avec le point terminale $x_F \leq +\infty$ si et seulement s'il existe $z < x_F$ tel que :*

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

où c et g sont deux fonctions mesurables telles que $c(x) \rightarrow c > 0$ et $g(x) \rightarrow 1$ quand $x \rightarrow x_F$ et a est une fonction auxiliaire. Dans ce cas on peut choisir les constantes de normalisation $b_n = \bar{F}(1 - \frac{1}{n})$ et $a_n = a(b_n)$. Un choix possible pour la fonction a est :

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt, x < x_F.$$

Dans ce tableau (1.1), nous avons regroupé quelques lois et leur domaine d'attraction

TAB. 1.1 – Exemple sur les lois classées selon leurs domaines d'attraction

<i>Domaine d'attraction</i>	<i>Gumbel</i> ($\gamma = 0$)	<i>Frechet</i> ($\gamma > 0$)	<i>Weibul</i> ($\gamma < 0$)
<i>Lois</i>	<i>Exponentielle</i>	<i>Pareto</i>	<i>Beta</i>
	<i>Normale</i>	<i>Cauchy</i>	<i>Uniforme</i>
	<i>Log – normale</i>	<i>Student</i>	<i>inverse de Burr</i>
	<i>Gamma</i>	<i>Log – gamma</i>	<i>inverse de Pareto</i>
	<i>Weibull</i>	<i>Burr</i>	

1.4 Distribution GPD

Cependant, l'approche des maximums par blocs a été critiquée dans la mesure où l'utilisation d'un seul maxima conduit à une perte d'information contenue dans les autres grandes valeurs de l'échantillon. Pour pallier ce problème, Pickands en (1975)[21] a introduit une nouvelle approche dans la *TVE* connue par la méthode *POT* : Peaks Over Threshold (ou des excès au-delà d'un seuil). Cette méthode consiste à observer non pas le maximum ou les plus grandes valeurs, mais toutes les valeurs des réalisations qui excèdent un certain seuil élevé

bien déterminé. Elle a été développée par divers auteurs tels que Smith en (1987)[24] et Reiss et Thomas en (2007)[22].

La *GPD* peut être utilisée pour modéliser des queues de distributions, c'est-à-dire pour des données dépassant un certain seuil pour être plus précis, on définit un nombre réel u suffisamment élevé, appelé seuil avec $N_u = \text{card} \{i : i = 1, \dots, n, X_i > u\}$ est le nombre de dépassements du seuil et $Y_i = X_i - u > 0$ pour $1 \leq i \leq N_u$ pour les $(X_i)_{1 \leq i \leq n}$ et Y_1, \dots, Y_{N_u} les excès correspondants.

Définition 1.9 (Distribution des excès et moyenne des excès) *La fonction de répartition des excès de X au-dessus du seuil u est définie par :*

$$F_u(y) = P(X - u \leq y \mid X > u)$$

De plus,

$$\begin{aligned} F_u(y) &= \frac{P(X - u \leq y, X > u)}{P(X > u)} = \frac{P(u < X \leq y + u)}{1 - P(X \leq u)} \\ &= \frac{P(X \leq y + u) - P(X \leq u)}{1 - P(X \leq u)} \\ &= \frac{F(u + y) - F(u)}{1 - F(u)} \\ &= \frac{F(u + y) - F(u)}{\bar{F}(u)} \end{aligned}$$

De même, si X est intégrable, la fonction de moyenne des excès de X est donnée par :

$$e_u(x) = E(X - u \mid X > u) = \int_u^\infty x dF_u(y), u < x_F,$$

qui peut s'écrire aussi comme suit :

$$e_u(x) = \frac{1}{\bar{F}(u)} \int_u^\infty \bar{F}(x) dx, u < x_F.$$

Définition 1.10 (Distribution de Paréto généralisée) *La loi de Pareto Généralisée (GPD),*

de paramètres $\gamma \in \mathbb{R}$ et $\sigma > 0$, est définie par sa fonction de répartition, donnée par :

$$G_{\gamma,\sigma,\mu}(x) := \begin{cases} 1 - (1 + \gamma \frac{x-\mu}{\sigma})^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0. \\ 1 - \exp(\frac{-x}{\sigma}) & \text{si } \gamma = 0. \end{cases},$$

où

$$\begin{aligned} x &\geq 0 && \text{si } \gamma \geq 0, \\ 0 \leq x &\leq -\frac{\sigma}{\gamma} && \text{si } \gamma < 0. \end{aligned}$$

On note que la GPD standard correspond au cas où $\mu = 0$ et $\sigma = 1$, telle que :

$$G_{\gamma}(x) := \begin{cases} 1 - (1 - \gamma x)^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \\ 1 - \exp(-x) & \text{si } \gamma = 0 \end{cases}.$$

avec

$$\begin{aligned} x &\geq 0 && \text{si } \gamma \geq 0, \\ 0 \leq x &\leq -\frac{1}{\gamma} && \text{si } \gamma < 0. \end{aligned}$$

Lorsque le paramètre de localisation est nul ($\mu = 0$) et le paramètre d'échelle est arbitraire ($\sigma > 0$). Balkema et de Haan [3] et Pickands [21] ont proposés le théorème suivant qui fait le lien entre le comportement asymptotique de la distribution des excès et la loi de Paréto généralisée.

Théorème 1.7 (Balkema et de Haan(1974), Pickands(1975)) Si la fonction de répartition F de point terminal x_F appartient au domaine d'attraction de G_{γ} , alors il existe une fonction mesurable et positive $\sigma(u)$, tel que :

$$\lim_{u \rightarrow x_F} \sup_{0 < y < x_F - u} |F_u(y) - G_{\gamma,\sigma(u)}(y)| = 0. \quad (1.4)$$

L'interprétation de ce théorème est la suivante : On peut modéliser asymptotiquement la

distribution conditionnelle de dépassements au-dessus d'un seuil par la GPD.

Preuve. La preuve de ce théorème doit être trouvée dans Embrecht et al [12] ■

Remarque 1.6 – *On écrit la densité de la distribution GPD comme suit :*

$$g_{\gamma,\sigma}(x) := \begin{cases} \frac{1}{\sigma} (1 + \gamma \frac{x}{\sigma})^{\frac{-1}{\gamma} - 1} & \text{si } \gamma \neq 0, \\ \frac{1}{\sigma} \exp(-\frac{x}{\sigma}) & \text{si } \gamma = 0. \end{cases}$$

– *Il y'a un rapport simple entre la GPD standard $G_\gamma(x)$ et la GEV standard $H_\gamma(x)$, tels que :*

$$G_\gamma(x) = 1 + \log H_\gamma(x), \text{ si } \log H_\gamma(x) > -1.$$

Exemple 1.5 *pour la loi exponentielle standard, la fonction de répartition est $F(x) = 1 - \exp(-x)$, $x > 0$, alors :*

$$F_u(y) = \frac{\exp(-u) - \exp(-u - y)}{\exp(-u)} = 1 - \exp(-y).$$

ceci correspond à $\gamma = 0$ et $\sigma = 1$ dans (1.4).

Exemple 1.6 *Pour la loi de Fréchet standard, la fonction de répartition est $F(x) = \exp(-1/x)$ pour $x > 0$ alors :*

$$F_u(y) = \frac{1 - \exp(\frac{-1}{u+y})}{1 - \exp(\frac{-1}{u})} = 1 - (1 + \frac{x}{u})^{-1}$$

Ceci correspond à $\gamma = 1$ et $\sigma(u) = u$ dans (1.4).

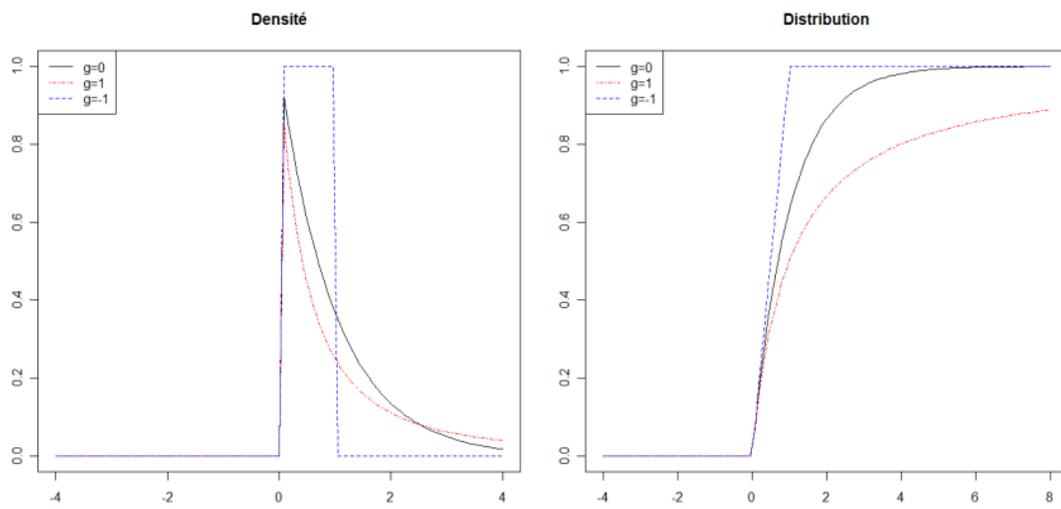


FIG. 1.2 – Représentation graphique de la densité et distribution de lois Paréto généralisée .

Chapitre 2

Estimation de l'indice des valeurs extrêmes

De nombreux auteurs se sont intéressés à l'estimation de l'indice des valeurs extrêmes γ pour des lois à queue lourde, voir Hill (1975) [17], Pickands (1975) [21], Mason (1982) [20], Csörgö et al (1985) [5], Dekkers et al (1989) [11], Davis et Resnick (1984) [7]. Il existe deux approches différentes pour cette estimation le premier c'est l'approche des maximums par blocs, est basés sur la *GEV* et le second connu comme approche de *POT*, utilise la *GPD*. Dans les deux cas, il existe plusieurs méthodes d'estimation : paramétrique, semi-paramétrique, etc., la méthode paramétrique la plus connue c'est la méthode du maximum de vraisemblance.

Dans ce chapitre, nous nous intéressons essentiellement à l'estimation de l'indice des valeurs extrêmes γ pour des lois à queue lourde, et nous concentrons sur certains des estimateurs avec leurs propriétés asymptotiques.

2.1 Estimation paramétrique

2.1.1 Estimateur de maximum de vraisemblance

La méthode la plus classique est celle du maximum de vraisemblance. L'estimation par le maximum de vraisemblance donne des résultats asymptotiques efficaces, et les estimateurs obtenus convergent sous certaines conditions vers les vraies valeurs des paramètres.

Pour la loi GEV

Soit X_1, X_2, \dots, X_n un échantillon, de n va's i.i.d, de densité h_θ , où $\theta = (\gamma, \sigma, \mu)$.

L'expression de la fonction de vraisemblance est donnée par :

$$l((X_1, X_2, \dots, X_n); (\gamma, \sigma, \mu)) = \prod_{i=1}^n h_\theta(x_i), \text{ avec } \theta = (\gamma, \sigma, \mu).$$

L'estimateur $\hat{\theta}$ est donné par la résolution du système suivant :

$$\begin{cases} \frac{\partial \log l(\theta)}{\partial \theta} = 0, \\ \frac{\partial^2 \log l(\theta)}{\partial^2(\theta)} < 0. \end{cases}$$

Dans le cas où $\gamma = 0$ (loi de Gumbel), la fonction logarithme de vraisemblance est égale à :

$$\log l((X_1, X_2, \dots, X_n); (\sigma, \mu, 0)) = -n \log \sigma - \sum_{i=1}^n \exp\left(-\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^n \frac{x_i - \mu}{\sigma}.$$

En dérivant cette fonction relativement aux deux paramètres, nous obtenons le système d'équations à résoudre suivant :

$$\begin{aligned} \frac{\partial \log l}{\partial \sigma} = 0 &\iff n + \sum_{i=1}^n \frac{x_i - \mu}{\sigma} \left[\exp\left(-\frac{x_i - \mu}{\sigma}\right) - 1 \right] = 0. \\ \frac{\partial \log l}{\partial \mu} = 0 &\iff n - \sum_{i=1}^n \exp\left(-\frac{x_i - \mu}{\sigma}\right) = 0. \end{aligned}$$

La résolution de ce système est relativement difficile et n'admet pas en général de solutions

explicites. Dans ce cas, on fait appel à des méthodes d'optimisation numériques (c'est ce qui est fait par exemple la fonction `fgev` du package R `evd`).

Pour la loi GPD

Soit X_1, X_2, \dots, X_n un échantillon, de n variables aléatoires supposées indépendantes identiquement distribuées i.i.d, de densité g_θ où $\theta = (\gamma, \sigma)$.

L'expression de la fonction de vraisemblance est donnée par :

$$l((X_1, X_2, \dots, X_n); (\gamma, \sigma)) = \prod_{i=1}^n g_\theta(x_i), \text{ avec } \theta = (\gamma, \sigma).$$

Dans le cas où $\gamma \neq 0$ la fonction logarithme de vraisemblance est égale à :

$$\log l((X_1, X_2, \dots, X_n); (\sigma, \gamma)) = -n \log \sigma - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^n \log\left(1 + \frac{\gamma}{\sigma} X_i\right).$$

En dérivant cette fonction relativement aux deux paramètres, nous obtenons le système d'équations à résoudre suivant :

$$\begin{aligned} \frac{\partial \log l}{\partial \sigma} = 0 &\iff -\frac{n}{\sigma} - \left(\frac{1+\gamma}{\gamma}\right) \sum_{i=1}^n \left(\frac{-\gamma x_i}{\sigma(\sigma + \gamma x_i)}\right) = 0. \\ \frac{\partial \log l}{\partial \gamma} = 0 &\iff \frac{1}{\gamma^2} \sum_{i=1}^n \log\left(1 + \gamma \frac{x_i}{\sigma}\right) - \left(\frac{1+\gamma}{\gamma}\right) \sum_{i=1}^n \left(\frac{x_i}{\sigma + \gamma x_i}\right) = 0. \end{aligned}$$

Et pour $\gamma = 0$, on obtient :

$$\log l((X_1, X_2, \dots, X_n); (\sigma, 0)) = -N_u \log \sigma - \frac{1}{\sigma} \sum_{i=1}^{N_u} x_i$$

Après on dérive $\log l$ par rapport à σ , on obtient :

$$\begin{aligned} \frac{\partial \log l}{\partial \sigma} = 0 &\iff -\frac{N_u}{\sigma} + \frac{1}{\sigma^2} \sum_{i=1}^{N_u} x_i = 0. \\ &\iff \hat{\sigma} = \frac{1}{N_u} \sum_{i=1}^{N_u} x_i = \bar{X}. \end{aligned}$$

2.2 Estimation par des méthodes semi-paramétrique

Le méthode que l'on a vue exige que les variables soient distribuées exactement selon H . La famille de lois H est de nature totalement paramétrique. En revanche, les estimateurs que l'on propose par la suite s'appliquent à des échantillons dont la distribution n'est plus H , mais qui appartient au domaine maximal d'attraction de H . La nature est par conséquent semi-paramétrique.

Si on se restreint au domaine d'attraction de Fréchet où l'on a la caractérisation

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

avec L une fonction à variation lente et $\gamma > 0$. Cette fonction comporte :

– une partie paramétrique $x^{-\gamma}$ ne dépendant que d'un paramètre réel.

Une partie non-paramétrique L sur laquelle on sait seulement que : $\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1$, pour $x > 1$.

2.2.1 Estimateur de Pickands $\gamma \in \mathbb{R}$

L'estimateur de Pickands est le plus simple et le plus ancien a été proposé par Pickands en (1975)[21] cet estimateur a l'avantage d'être valable quelque soit le domaine d'attraction de " $H_\gamma, \gamma \in \mathbb{R}$ ".

Définition 2.1 (Estimateur de Pickands) Soit X_1, X_2, \dots, X_n une suite de va's i.i.d de fonction de répartition $F \in D(H_\gamma)$, où $\gamma \in \mathbb{R}$. Soit une suite d'entier $k = k(n) \rightarrow \infty$ quand $n \rightarrow \infty$. L'estimateur de Pickands est donné par la statistique suivante :

$$\hat{\gamma}_{k(n)}^p := \frac{1}{\log 2} \log \frac{X_{n-k,n} - X_{n-2k,n}}{X_{n-2k,n} - X_{n-4k,n}}$$

Il a prouvé la convergence faible de son estimateur, la convergence forte ainsi que la normalité asymptotique ont été démontrées par Dekkers et de Haan (1989) [11].

Construction de l'estimateur de Pickands

D'après la relation (1.3), pour $\gamma \in \mathbb{R}$ avec le choix $t = 2s$, $x = 2$ et $y = 1/2$, on a

$$\lim_{t \rightarrow \infty} \frac{U(t) - U(t/2)}{U(t/2) - U(t/4)} = 2^\gamma \quad (2.1)$$

En remplaçant dans (2.1) U par $U_n = (1/1 - F_n)^\leftarrow$ (F_n étant la fonction de répartition empirique) et on choisit $t = \frac{n}{k}$, où k est une suite intermédiaire, on a :

$$\lim_{n \rightarrow \infty} \frac{U_n(\frac{n}{k}) - U_n(\frac{n}{2k})}{U_n(\frac{n}{2k}) - U_n(\frac{n}{4k})} = 2^\gamma$$

pour n assez grand du fait que $U_n(x) = X_{(\frac{n}{x})}$, alors :

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

L'estimateur de Pickands est alors la solution de cette dernière équation.

Propriétés asymptotiques de l'estimateur de Pickands

1. Consistance faible

Théorème 2.1 Soit X_1, X_2, \dots, X_n une suite de va's i.i.d de fonction de répartition $F \in$

$D(H_{\mu, \sigma, \gamma})$, et $\gamma > 0$. Soit $(k(n))_{n \geq 1}$ une suite d'entiers telles que $1 < k(n) \leq n$

$$\text{Si } \begin{cases} k(n) \rightarrow \infty \\ \text{et} \\ \frac{k(n)}{n} \rightarrow 0 \end{cases} \text{ alors } \hat{\gamma}_n^p \xrightarrow{p} \gamma, \text{ quand } n \rightarrow \infty.$$

2. Consistance forte

Théorème 2.2 Soit X_1, X_2, \dots, X_n une suite de va's i.i.d de fonction de répartition $F \in$

$D(H_{\mu, \sigma, \gamma})$, et $\gamma \in \mathbb{R}$. Soit $(k(n))_{n \geq 1}$ une suite d'entiers telles que $1 < k(n) \leq n$

$$Si \begin{cases} \frac{k(n)}{n} \rightarrow \infty \\ \text{et} \\ \frac{k(n)}{\log(\log(n))} \rightarrow \infty. \end{cases} \quad \text{alors } \hat{\gamma}_n^p \xrightarrow{p.s} \gamma, \text{ quand } n \rightarrow \infty.$$

3. Normalité asymptotique

Théorème 2.3 *Sous des conditions additionnelles sur la suite intermédiaire $k = k(n)$ et la fonction de répartition F , on a :*

$$\sqrt{k}(\hat{\gamma}_n^p - \gamma) \xrightarrow{d} N(0, \eta^2), \text{ quand } n \rightarrow \infty$$

où

$$\eta^2 := \frac{\gamma^2(2^{2\gamma+1} + 1)}{(2(2^\gamma - 1) \log 2)^2}.$$

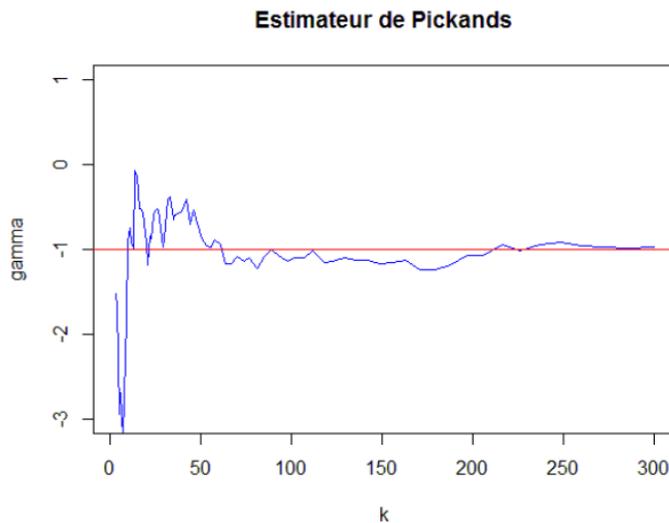


FIG. 2.1 – Représentation graphique de l'estimateur de Pickands .

2.2.2 Estimateur de Hill

L'estimateur de l'*EVI* le plus célèbre est l'estimateur de Hill qui a été introduit en 1975 et qu'il a été largement étudié dans le cas des v_a 's i.i.d. Pour construire cet estimateur on utilise deux approches, la première approche est basée sur la *GEV* et la seconde approche *POT* basée sur la *GPD* a été démontré par de Haan et Ferreira [9].

Définition 2.2 (Estimateur de Hill) Soit X_1, X_2, \dots, X_n une suite de va's i.i.d de fonction de répartition F appartenant au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes $\gamma > 0$, soit une suite d'entier $k = k(n) \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$. L'estimateur de Hill est défini par la statistique :

$$\hat{\gamma}_{(k(n),n)}^H := \frac{1}{k} \sum_{i=1}^k \log(X_{n-i+1,n}) - \log(X_{n-k,n}), 1 \leq k \leq n.$$

Mason (1982) [20] et Deheuvels et al.(1988) [10] ont montrés respectivement la consistance faible et forte qui ne dépend que du comportement de la suite $k(n)$. Pour établir la normalité asymptotique, on a besoin de supposer que la fonction de répartition F est à variation régulière du second ordre. Plusieurs auteurs ont obtenus cette normalité, notamment Davis et Resnick (1984) [7], Csörgö et Mason (1985) [5], Haeusler et Teugels (1985) [16].

Proposition 2.1 (Condition du premier ordre de Haan et Ferreira (2006)) Les assertions suivantes sont équivalentes :

– F à queue lourde

$$F \in D(\Phi_{1/\gamma}), \gamma > 0.$$

– F est à variation régulière à ∞ d'indice $-1/\gamma$

$$\lim_{t \rightarrow \infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma}, x > 0.$$

– $Q(1-s)$ est à variation régulière à 0 d'indice $-\gamma$

$$\lim_{s \rightarrow 0} \frac{Q(1-sx)}{Q(1-s)} = x^{-\gamma}, x > 0.$$

– U est à variation régulière à ∞ d'indice γ

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\gamma, x > 0.$$

Définition 2.3 (Fonction à variation régulière du second ordre) On dit que la queue

de $F \in D(\Phi_\alpha)$ avec $\alpha = 1/\gamma$, est à variation régulière du second ordre, d'indice (γ, ρ) avec le paramètre du second ordre $\rho \leq 0$ (on note $\bar{F} \in 2RV_{(\gamma, \rho)}$), à l'infinie si l'une des conditions équivalentes est satisfaite :

1. Il existe un paramètre $\rho \leq 0$ et une fonction A^* satisfaisant $\lim_{t \rightarrow \infty} A^*(t) = 0$ ne change pas son signe près de ∞ , telle que pour tout $x > 0$

$$\lim_{t \rightarrow \infty} \frac{(1 - F(tx)/1 - F(t)) - x^{-\alpha}}{A^*(t)} = x^{-\alpha} \frac{x^\rho - 1}{\rho}$$

2. Il existe un paramètre $\rho \leq 0$ et une fonction A^{**} satisfaisant $\lim_{t \rightarrow \infty} A^{**}(t) = 0$ et ne change pas son signe près de 0, telle que pour tout $x > 0$

$$\lim_{s \rightarrow \infty} \frac{(Q(1 - sx)/Q(1 - s)) - x^{-\gamma}}{A^{**}(t)} = x^{-\gamma} \frac{x^\rho - 1}{\rho}, \forall x > 0.$$

3. Il existe un paramètre $\rho \leq 0$ et une fonction A^* satisfaisant $\lim_{t \rightarrow \infty} A^*(t) = 0$ ne change pas son signe près de ∞ , telle que pour tout $x > 0$

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho} \quad (2.2)$$

Si $\rho = 0$, $x^\rho - 1/\rho$ s'interprète comme $\log x$. A , A^* et A^{**} sont des fonction à variation régulière avec $A^*(t) = A(1/\bar{F}(t))$ et $A^{**}(s) = A(1/s)$. Leurs rôle est de contrôler la vitesse de convergence.

Construction de l'estimateur de Hill

Il ya plusieurs approches pour construire l'estimateur de Hill

Approche par EMV On considère une suite de variables aléatoires i.i.d de loi de Pareto de paramètre $\alpha > 0$, de fonction de répartition est donnée par :

$$F(x) = 1 - x^{-\alpha} \text{ pour } x \succeq 1$$

alors

$$f(x) = \alpha x^{-\alpha-1} \text{ si } x \geq 1$$

On veut trouver l'estimateur du maximum de vraisemblance (*EMV*) de α pour cela, on donne $l(\cdot)$ la fonction de vraisemblance et les dérivées du logarithme de la fonction de vraisemblance par rapport à α

$$\begin{aligned} l(x_1, x_2, \dots, x_n, \alpha) &= \prod_{i=1}^n f(x_i) \\ &= \alpha^n \left(\prod_{i=1}^n x_i^{-\alpha-1} \right) \end{aligned}$$

$$\left\{ \begin{array}{l} \frac{\partial \log l(x_1, x_2, \dots, x_n, \alpha)}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^n \log x_i \\ \frac{\partial^2 \log l(x_1, x_2, \dots, x_n, \alpha)}{\partial \alpha^2} = -\frac{n}{\alpha^2} < 0 \end{array} \right.$$

L'estimateur du maximum de vraisemblance de $\frac{1}{\alpha}$ et donc la statistique :

$$\hat{T} = \frac{1}{n} \sum_{i=1}^n \log X_i = \frac{1}{n} \sum_{i=1}^n \log X_{(i,n)}$$

Une généralisation concern

$$F(x) = Cx^{-\alpha}, x > 0.$$

Si on pose $C = u^\alpha$ avec $x \geq u > 0$ on obtient de l'*EMV* de α :

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

Propriétés asymptotiques de l'estimateur de Hill

1. Consistance faible

Théorème 2.4 Soit (X_1, X_2, \dots, X_n) un échantillon de variables aléatoire *i.i.d* de fonction de répartition F avec $F \in D(\Phi_\gamma)$ et $\gamma > 0$. Soit $(k(n))_{n \geq 1}$ une suite d'entiers telles que

$$1 < k(n) \leq n$$

$$\text{Si } \begin{cases} k(n) \rightarrow \infty \\ \text{et} \\ \frac{k(n)}{n} \rightarrow 0 \end{cases} \quad \text{alors } \hat{\gamma}_{(k(n),n)}^H \xrightarrow{p} \gamma, \text{ quand } n \rightarrow \infty.$$

2. Consistance forte

Théorème 2.5 Soit (X_1, X_2, \dots, X_n) un échantillon de n variables aléatoires identiquement distribuées de fonction de répartition F avec $F \in D(H_{\mu,\sigma,\gamma})$ et $\gamma \in \mathbb{R}$. Soit $(k(n))_{n \geq 1}$ une suite d'entiers telles que $1 < k(n) \leq n$

$$\text{Si } \begin{cases} \frac{k(n)}{n} \rightarrow \infty. \\ \text{et} \\ \frac{k(n)}{\log(\log(n))} \rightarrow \infty. \end{cases} \quad \text{alors } \hat{\gamma}_{(k(n),n)}^H \xrightarrow{p.s} \gamma, \text{ quand } n \rightarrow \infty.$$

3. Normalité asymptotique

Théorème 2.6 Si la condition (2.2) est satisfaite avec $\lim_{n \rightarrow \infty} \sqrt{k} A(\frac{n}{K}) = \lambda$, alors

$$\sqrt{k}(\hat{\gamma}_{(K(n);n)}^H - \gamma) \xrightarrow{d} N\left(\frac{\lambda}{1-\rho}, \gamma^2\right), \text{ quand } n \rightarrow \infty.$$

Preuve. Voir le livre de de Haan, L. et Ferreira, A. (2006) [9]. ■

2.2.3 Estimateur du moment

Cet estimateur introduit par Dekkers et al (1989) [11] est une généralisation de l'estimateur de Hill, valable pour tout $\gamma \in \mathbb{R}$:

$$\hat{\gamma}_{(n)}^{(M)} = \hat{\gamma}_{(k(n);n)}^H + 1 - \frac{1}{2} \left(1 - \frac{\hat{\gamma}_{(k(n);n)}^H}{M_n^{(2)}}\right)^{-1}$$

où

$$M_n^{(r)} := \frac{1}{k} \sum_{i=0}^K (\log X_{n-i+1,n} - \log X_{n-k,n})^r; \quad r = 1, 2.$$

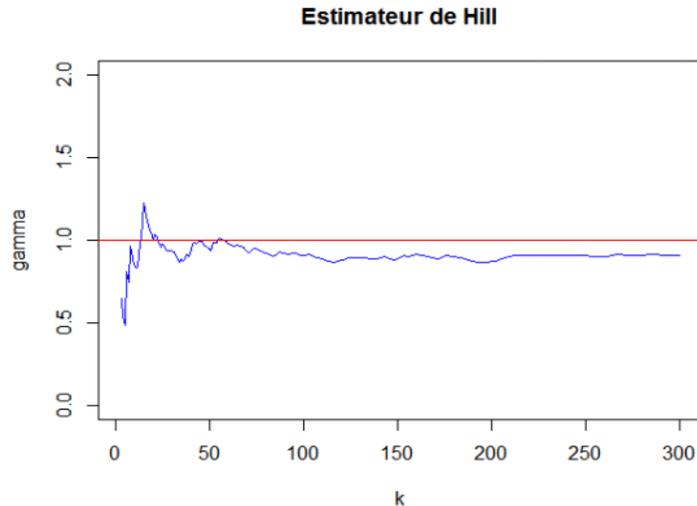


FIG. 2.2 – Représentation graphique de l'estimation de Hill .

Théorème 2.7 (Propriétés asymptotiques de $\hat{\gamma}_n^{(M)}$) Supposons que $F \in D(H_\gamma)$, $\gamma \in \mathbb{R}$, $k \rightarrow \infty$ et $\frac{k}{n} \rightarrow 0$ quand $n \rightarrow \infty$:

– *Consistance faible* :

$$\hat{\gamma}_n^{(M)} \xrightarrow{p} \gamma, \text{ quand } n \rightarrow \infty.$$

– *Consistance forte* :

$$\hat{\gamma}_n^{(M)} \xrightarrow{p.s} \gamma, \text{ quand } n \rightarrow \infty.$$

– *Normalité asymptotique, pour $\gamma \geq 0$* :

$$\sqrt{k}(\hat{\gamma}_n^{(M)} - \gamma) \xrightarrow{d} N(0, 1 + \gamma^2), \text{ quand } n \rightarrow \infty.$$

2.2.4 Le choix du nombre k

La difficulté consiste à choisir le nombre k de statistiques d'ordre extrême à utiliser. Les résultats concernant les estimateurs de l'indice des valeurs extrêmes sont asymptotiques : ils sont obtenus lorsque $k \rightarrow \infty$ et $\frac{k}{n} \rightarrow 0$. Comme en pratique, on ne dispose que d'un nombre d'observations n fini, il s'agit de choisir k de manière à ce que l'on dispose de suffisamment de matériel statistique tout en restant dans la queue de distribution ($k \ll n$).

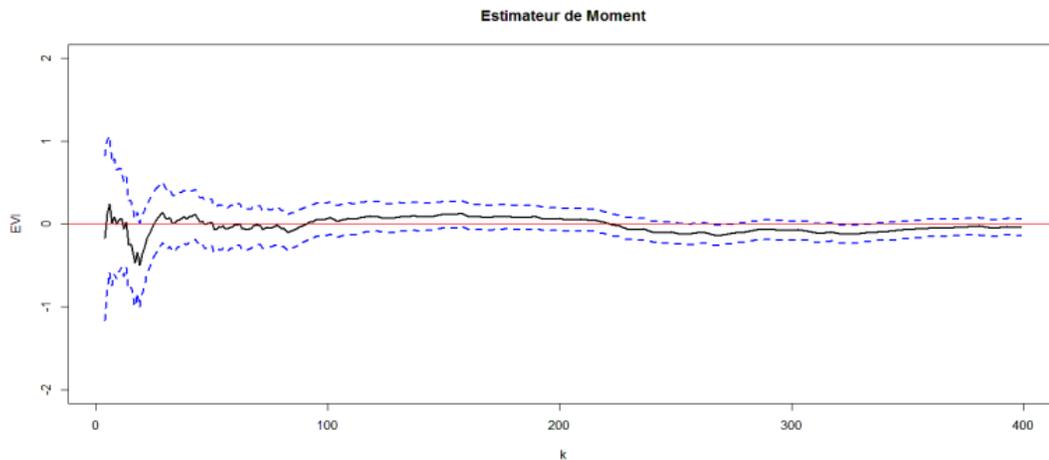


FIG. 2.3 – Représentation graphique de l'estimateur de Moment .

Méthode graphique

C'est une méthode la plus simple pour la détermination de k . Elle consiste à tracer le graphe $(k, \hat{\gamma}_{k(n),n}^H)$ et de prendre la valeur où $(k, \hat{\gamma}_{k(n),n}^H)$ devient horizontal. Cet estimateur est valable seulement dans le domaine d'attraction de Fréchet c'est à dire si $\gamma > 0$.

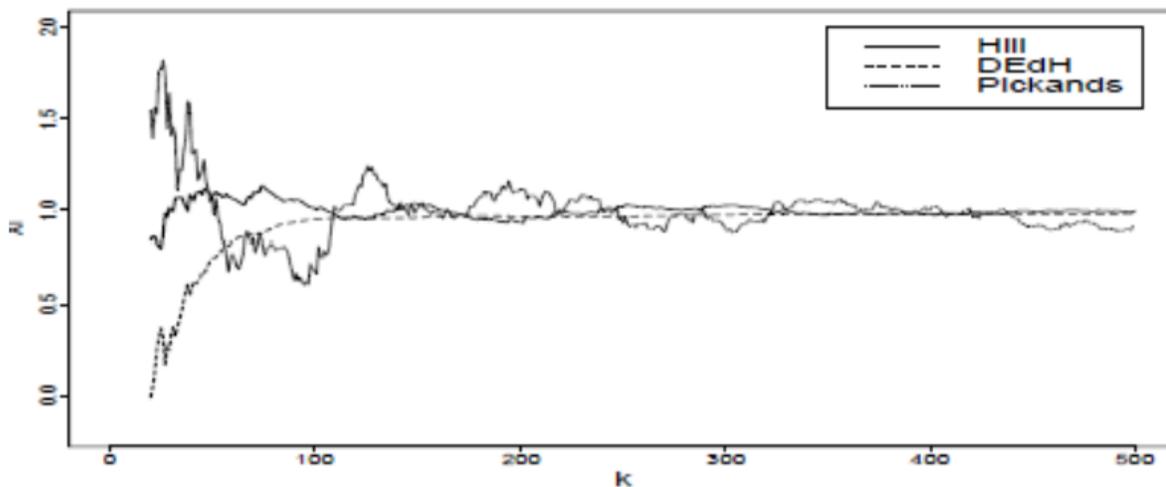


FIG. 2.4 – Le graphe de $(k, \hat{\gamma}_{k(n),n}^H)$ pour les différents estimateur de γ .

Méthode analytique

Il est nécessaire pour donner une précision à l'estimateur $\hat{\gamma}_{(k(n),n)}^H$, de calculer l'erreur moyenne quadratique (MSE), elle est en fonction de k :

$$\begin{aligned} \text{MSE}(\hat{\gamma}_{k(n),n}) &= \text{MSE}(\hat{\gamma}_{k(n),n} - \gamma)^2 \\ &= \text{biais}^2(\hat{\gamma}_{k(n),n}) + \text{var}(\hat{\gamma}_{k(n),n}). \end{aligned}$$

Le choix optimal de k , correspond à minimiser la MSE. Concernant l'estimateur de Hill pour des fonctions appartenant au domaine d'attraction maximale de Fréchet, de Haan et Peng en 1998 ont proposé de retenir le nombre d'observation k_{opt} qui minimise l'erreur moyenne quadratique de l'estimateur de Hill par

$$k_{opt} = \begin{cases} 1 + 2^{2\gamma/(2\gamma+1)} \left(\frac{\gamma+1}{2\gamma}\right)^{1/(2\gamma+1)} & \text{si } 0 < \gamma < 1. \\ 2\eta^{2/3} & \text{si } \gamma > 0. \end{cases}$$

Procédure numérique

Il existe plusieurs algorithmes de procédures adaptives pour trouver un estimateur \hat{k}_{opt} de k_{opt} :

$$\frac{\hat{k}_{opt}}{k_{opt}} \xrightarrow{p} 1 \text{ quand } n \rightarrow \infty.$$

alors $\hat{\gamma}_{(\hat{k}_{opt},n)}$ converge asymptotiquement vers $\hat{\gamma}_{(k_{opt},n)}$.

Remarque 2.1 – Si k est petit, $\hat{\gamma}_{k(n)}$ utilise peu d'observations et on a une grande variance.

– Si k est grand, le biais est grand, la variance est petite.

Pour une meilleure exactitude, le nombre de statistiques d'ordre utilisées dans les procédés d'estimation doit être intermédiaire, C'est à dire, ni trop petit ni trop grand.

Chapitre 3

Application des valeurs extrêmes

La modélisation des événements extrêmes (ouragan, tremblement de terre ou inondation, crues, crises financières, krachs, chocs pétroliers) est aujourd'hui un champ de recherches particulièrement actif, notamment par l'importance de leurs impacts économiques et sociaux. En particulier, depuis quelques années, on note un intérêt croissant pour l'application de la Théorie des Valeurs Extrêmes (*TVE*) pour la modélisation de tels événements. Les domaines d'applications utilisant les modèles de la *TVE* n'ont cessé de se développer ces dernières années touchant des domaines variés.

- En hydrologie : crues consécutives à des pluies torrentielles : aux Pays-Bas, digues menacées par l'effet conjoint des grandes marées et des conditions climatiques en Mer du Nord (Inondation de 1953).
- En assurance : survenue des sinistres d'intensité exceptionnelle (ouragan Katrina en 2005, importants incendies en risques industriels, sinistres graves en responsabilité civile automobile) qui peuvent avoir des conséquences négatives sur les résultats et la solvabilité des organismes d'assurance.
- En Finance : fortes variations du cours d'actifs financiers, gestions du risque opérationnel des banques (la crise fin des années 2000).
- En Climatologie : étude des évènements climatiques extrêmes (précipitations, températures, chutes de neige), modélisation des grands feux de forêt (voir par exemple Alvarado

et al. 1998).

- En météorologie [Coles et Walshaw, 1994 ; Smith, 2001] où l'étude de la vitesse du vent, par exemple, permet d'évaluer le degré de résistance des matériaux face à la pression exercée par le vent (au cours d'une tempête par exemple) sur les bâtiments ou les structures de génie civil.

3.1 Données réelles : Danish fire

Comme application réelle, considérons les catastrophes naturelles qui sont des exemples d'événement extrêmes qui conduisent à des pertes financières importantes. Parmi ces dernières on considère des données réelles décrivent les grandes demandes d'assurance d'incendie au Danemark (Danish fire - source : package evir - R), du jeudi 3 janvier 1980 au lundi 31 décembre 1990. Les données sont contenues dans un vecteur numérique de 2167 observations. Pour la modélisation de cet événement on utilise la *GEV* concernant les données relatives aux 132 blocs mensuelles des maximums et celles de 310 blocs hebdomadaire. Pour estimer les paramètres du modèle, considérant la méthode du maximum de vraisemblance, les valeurs estimées des paramètres de la *GEV* sont : $\hat{\gamma} = 0.623$, $\hat{\mu} = 5.971$, $\hat{\sigma} = 8.375$ avec des erreurs (écart-types) 0.103, 0.632 et 0.611 respectivement pour les maximums des blocs mensuels, et $\gamma = 0.689$, $\mu = 3.062$ et $\sigma = 4.612$ avec des erreurs 0.063, 0.217 et 0.198 respectivement pour les maximums des blocs hebdomadaires.

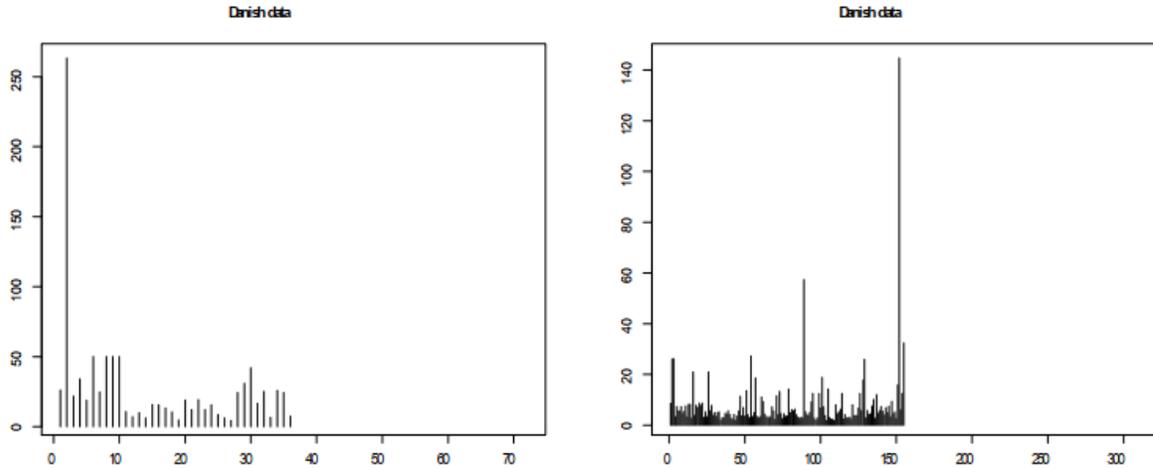


FIG. 3.1 – Données danoises : mensuelles (à gauche) et hebdomadaires (à droite).

3.2 Modèle composite

Dans cette section, de nouveaux modèles composites sont construits avec une valeur de seuil, θ , et des poids de mélange flexibles, a_1 et a_2 , où $a_1 + a_2 = 1$. Nous dérivons les poids de mélange en fonction des paramètres du modèle.

À des fins de comparaison, nous reproduisons le modèle composite lognormal-Pareto introduit par Scollnik (2007). Ce modèle a donné l'ajustement le plus remarquable parmi tous les modèles examinés par Scollnik (2007).

Soit X une variable aléatoire ayant la fonction de densité de probabilité :

$$f(x) = \begin{cases} a_1 f_1^*(x) & 0 < x \leq \theta, \\ a_2 f_2^*(x) & \theta \leq x < \infty, \end{cases} \quad (3.1)$$

tel que $a_1 = \frac{1}{1+\Phi}$ et $a_2 = \frac{\Phi}{1+\Phi}$, où f_1^* et f_2^* sont des fonctions de probabilité tronquées. Scollnik (2007) a pris les distributions tronquées pour être log-normales et Pareto du deuxième type. Autrement dit, $f_1(x)$ et $f_2(x)$ sont donnés respectivement par :

$$f_1(x) = \frac{1}{x\sigma} \psi\left(\frac{\ln x - \mu}{\sigma}\right) \quad (3.2)$$

et

$$f_2(x) = \frac{\alpha \lambda^\alpha}{(x + \lambda)^{\alpha+1}}$$

pour $x > 0, \sigma > 0, \lambda > 0$, et $-\infty < \mu < \infty$, où $\psi(\cdot)$ désigne la fonction de densité de probabilité normale standard. L'application de continuité en θ ($f'(\theta-) = f'(\theta+)$) donne :

$$\Phi = \frac{(\theta + \lambda)\psi\left(\frac{\ln\theta - \mu}{\sigma}\right)}{\alpha\theta\sigma\Phi\left(\frac{\ln\theta - \mu}{\sigma}\right)},$$

et

$$\mu = \ln(\theta) - \left(\frac{\alpha\theta - \lambda}{\theta + \lambda}\right)\sigma^2.$$

Ici, on prend $f_2(\cdot)$ pour correspondre à la distribution de Burr, résultant en le modèle composite lognormal-Burr, une fonction continue par morceaux composée d'une fonction de densité de probabilité lognormale tronquée et d'une probabilité de bavure tronquée et d'une fonction de densité de probabilité de bavure tronquée reliée par un seuil, θ .

Par conséquent, le modèle composite lognormal-Burr a une forme comme dans (3.1) avec $f_1(x)$ comme dans (3.2) et $f_2(x)$ donnée par :

$$f_2(x) = \frac{\alpha\beta\left(\frac{x}{s}\right)^\beta}{x\left[1 + \left(\frac{x}{s}\right)^\beta\right]^{\alpha+1}} \quad (3.3)$$

pour $x > 0, \alpha > 0, \beta > 0$, et $s > 0$. Notez également que (3.3) se réduit à une fonction de densité de probabilité de Pareto lorsque $\beta = 1$.

L'application de Continuité en θ conduit à l'expression suivante pour Φ :

$$\Phi = \frac{(\theta^\beta + s^\beta)\psi\left[\frac{\ln\theta - \mu}{\sigma}\right]}{\sigma\alpha\beta\theta^\beta\Phi\left[\frac{\ln\theta - \mu}{\sigma}\right]}.$$

L'application de la condition de différentiabilité ($f'(\theta-) = f'(\theta+)$) conduit à la simplification suivante :

$$\mu = \ln \theta - \sigma^2 \left\{ \frac{(\alpha + 1)\beta\theta^\beta}{\theta^\beta + s^\beta} - \beta \right\}.$$

Par conséquent, le modèle composite lognormal-Burr a cinq paramètres inconnus $\sigma > 0$,

$\theta > 0$, $\alpha > 0$, $\beta > 0$ et $s > 0$.

3.3 Résultats et discussion

Ici, nous illustrons une application aux données danoises de l'assurance incendie (en millions de DKK) de 2492 mesures. Ces données ont été précédemment appliquées par Cooray et Ananda (2005) pour leur modèle composite lognormal-Pareto. Scollnik (2007) a amélioré le modèle de Cooray et Ananda (2005) en utilisant la distribution de Pareto du second type pour remplacer la distribution de Pareto classique dans la deuxième partie du modèle. Scollnik (2007) a également autorisé pour des poids de mélange flexibles.

Certaines statistiques récapitulatives des données sont : la valeur minimale est 0.3134, le premier quartile est 1.1570, la médiane est 1.6340, la moyenne est 3.0630, le troisième quartile est 2.6450, la valeur maximale est 263.3000 et l'écart type est 7.976703.

Nous avons ajusté les modèles composite lognormal-Pareto et composite lognormal-Burr de la section 2 aux données. La méthode du maximum de vraisemblance a été utilisée. Le tableau 3.1 donne les estimations des paramètres, les valeurs de log-vraisemblance maximisées et les valeurs des critères d'information Akaike (AIC) (Akaike, 1974)

TAB. 3.1 – Estimations des paramètres pour les modèles composites ajustés aux données danoises de l'assurance incendie

<i>Des modèles</i>	<i>Estimations des paramètres</i>	<i>Log – vraisemblance</i>	<i>AIC</i>
<i>Lognormal – Pareto</i>	$\hat{\sigma} = 0.182(0.012)$	-3860.471	7728.943
	$\hat{\theta} = 1.145(0.030)$		
	$\hat{\alpha} = 1.563(0.088)$		
	$\hat{s} = 0.363(0.125)$		
<i>Lognormal – Burr</i>	$\hat{\sigma} = 0.178(0.011)$	-3857.827	7725.654
	$\hat{\theta} = 1.093(0.038)$		
	$\hat{\alpha} = 0.347(0.161)$		
	$\hat{\beta} = 4.111(1.808)$		
	$\hat{s} = 0.841(0.107)$		

Les nombres entre parenthèses sont les erreurs types calculées en inversant les matrices d'information observées.

Le modèle composite lognormal-Burr a un paramètre de plus que le mode composite lognormal-Pareto. Sur la base des valeurs AIC, nous voyons que le premier donne le meilleur modèle avec un *AIC* minimal.

La conclusion basée sur l'AIC peut être vérifiée au moyen de graphiques de probabilité-probabilité, de graphiques quantiles-quantiles et de graphiques de densité. Un graphique de probabilité-probabilité consiste en des graphiques des probabilités observées par rapport aux probabilités prédites par le modèle ajusté. Le graphique quantile-quantile consiste en des graphiques des quantiles observés par rapport aux quantiles prédits par le modèle ajusté. Un graphique de densité compare les fonctions de densité de probabilité ajustées des modèles avec l'histogramme empirique des données observées.

Le graphique de densité pour les deux modèles ajustés est présenté à la figure 3.2. Les graphiques de probabilité-probabilité et quantile-quantile sont présentés aux figures 3.3 et 3.4. Les fonctions de densité de probabilité ajustées semblent presque impossibles à distinguer.

Mais ils semblent bien capturer le schéma général de l'histogramme empirique. Les graphiques de probabilité-probabilité suggèrent que les deux modèles fournissent un ajustement adéquat aux données. Les graphiques quantiles-quantiles suggèrent que le modèle composite lognormal-Burr a des points plus proches de la ligne diagonale dans la queue supérieure.

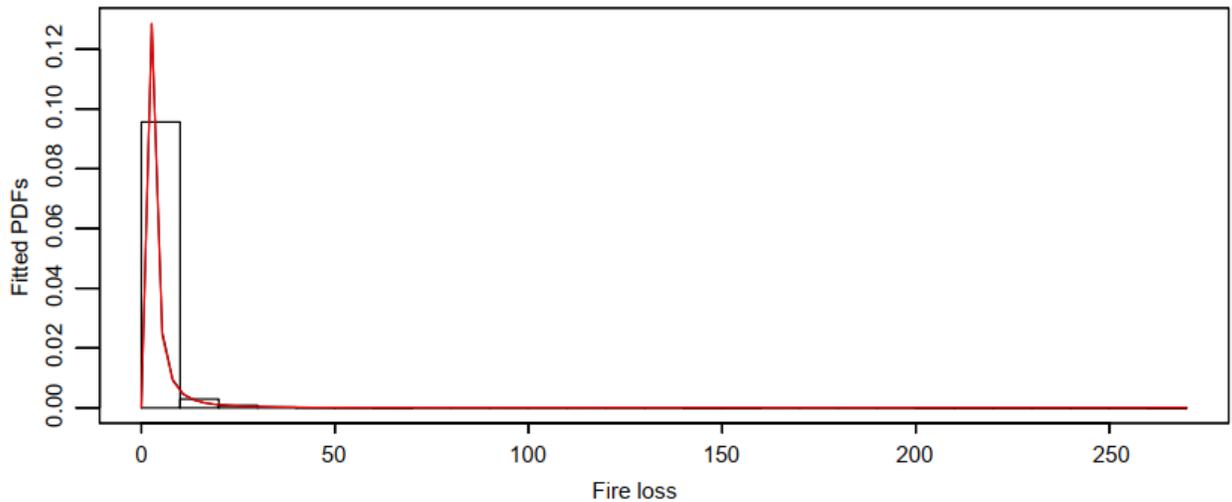


FIG. 3.2 – Fonctions de densité de probabilité ajustées pour les données danoises de l'assurance incendie : noir pour le modèle composite lognormal-Pareto et rouge pour le modèle composite lognormal-Burr.

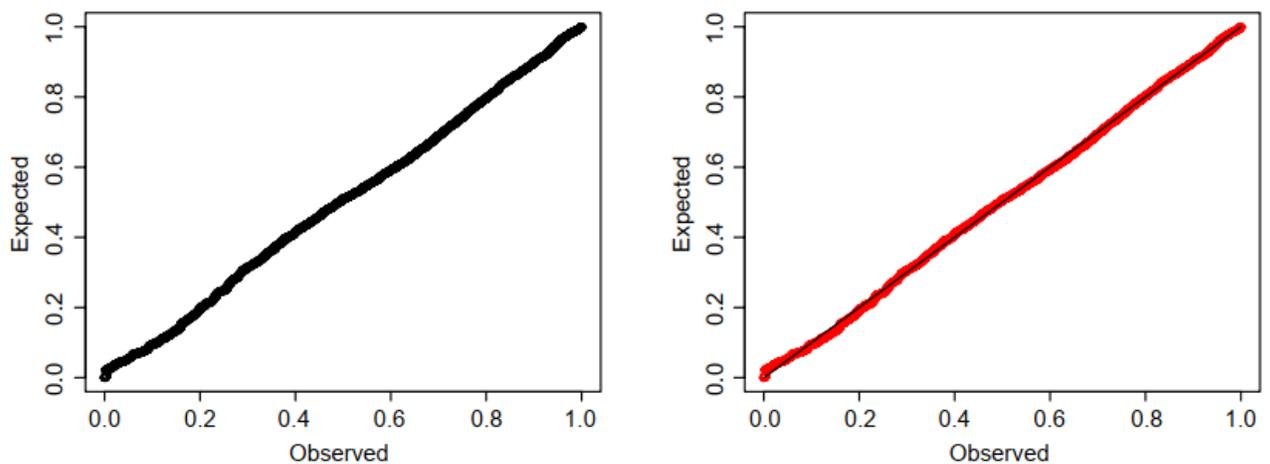


FIG. 3.3 – Diagrammes de probabilité pour les données danoises d'assurance incendie : noir pour le modèle composite lognormal-Pareto et rouge pour le modèle composite lognormal-Burr.

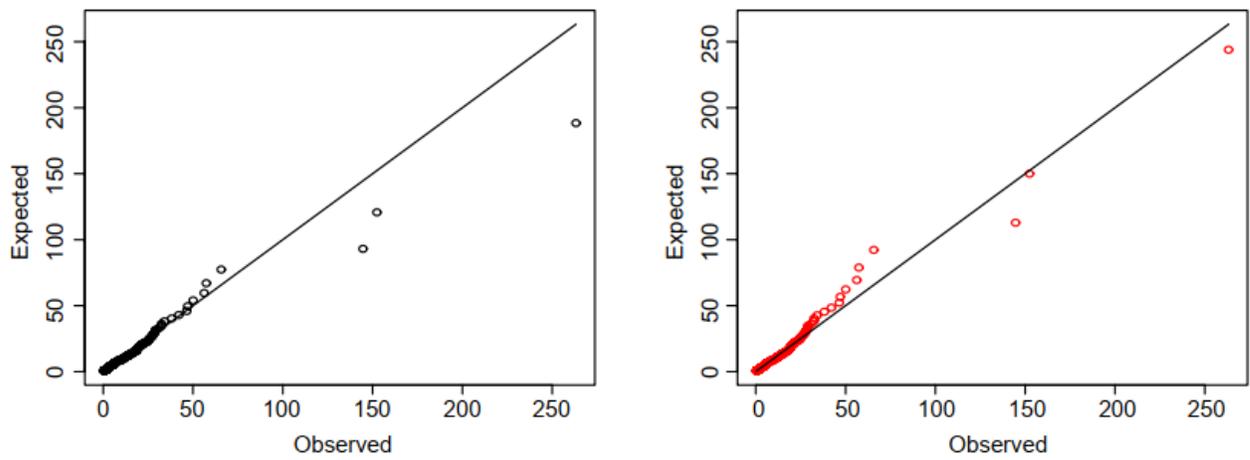


FIG. 3.4 – Graphiques quantile-quantiles pour les données danoises de l’assurance-incendie : noir pour le modèle composite lognormal-Pareto et rouge pour le modèle composite lognormal-Burr.

Conclusion

Depuis quelques années, la théorie des valeurs extrêmes a reçu beaucoup d'attention aussi bien sur le plan théorique que sur le plan pratique.

Dans ce mémoire, nous avons consacré à la présentation de la théorie des valeurs extrêmes.

Dans un premier temps, nous avons énoncés les principaux résultats et définitions de la théorie des valeurs extrêmes utiles dans nos travaux. Nous avons ensuite présentés une adaptation de l'estimation semi-paramétrique de l'indice des valeurs extrêmes proposées par Einmahl et al (2008).

Dans un second temps, nous présentons des applications des valeurs extrêmes en assurance (danish fire) sur des données aussi bien réelles que simulées permettront d'illustrer les résultats obtenus.

Bibliographie

- [1] Anis, Borchani, (2011). Statistiques des valeurs extrêmes dans le cas de lois discrètes
HAL Id : hal-00572559.
- [2] Brahim, B., Meraghni, D., Necir, A., and Yahia, D. (2013). A bias-reduced estimator for the mean of a heavy-tailed distribution with an infinite second moment. *J. Statist. Plann. Inference*, 143(6), 1064 – 1081
- [3] Belkema, A. A. ,and De Haan, L. (1974). Residual life time at great age. *Ann. Probab.*, 792 – 804.
- [4] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., (2006). *Statistics of Extremes – Theory and Applications*. Wiley.
- [5] Csörgő, S., and Mason, D. (1985). Central limit theorems for sums of extreme values. *Mathematical Proceedings of the Cambridge Philosophical Society*. 98(3). 547 – 558.
- [6] David, H. A., and Nagaraja, H. N. (2003). *Order Statistics, Third Edition*. John Wiley.
- [7] Davis, R., and Resnick, S. (1984). Tail estimates motivated by extreme value theory. *The Annals of Statistics*. 12(4). 1467 – 1487.
- [8] Davis, R.A., Resnick, S.I (1989). Basic properties and prediction of max-arma processes. *Adv. in Appl. Probab.*, 21 : 781 – 803.
- [9] De Haan, L. and Ferreira, A. (2006). *Extreme Value Theory : An Introduction*. Springer-Verlag, New York.
- [10] Deheuvels, P., Haeusler, E., and Mason, D. (1988) Almost sure convergence of the hill estimator. *Mathematical Proceedings of the Cambridge Philosophical Society*. 104(2). 371 – 381.

- [11] Dekkers.A.L.M, and Haan.L, D. (1989). On the estimation of the extreme-value index and largequantile estimation. *The Annals of Statistics*. 17(4). 1795 – 1832.
- [12] Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Event for Insurance and Finance*. Springer, Berlin.
- [13] Fisher, R., Tippet, L.(1928) Limitting forms the frequency distribution of the largestor smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*,24 : 180 – 190.
- [14] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d’une serie aleatoire. *Ann. Math.*, 423453.
- [15] Gilbert Saporta (2006) *Probabilité analyses des données et statistique* . Edition thechnip 27 rue Ginoux, 75737 Paris Cedex 15, France.
- [16] Haeusler, E., and Teugels, J. (1985) On asymptotic normality of hill’s estimator for the exponent of regular variation. *The Annals of Statistics*. 13(2). 743 – 756.
- [17] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3(5), 1163 1174.
- [18] [http ://www.statelem.com/statistique _d _ordre.php](http://www.statelem.com/statistique_d_ordre.php).
- [19] Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly J. R. Methodol.Soc.*, 81(348), 158 171
- [20] Mason, D. (1982), Laws of large numbers for sums of extreme values. *The Annals of Probability*.10(3), 754 – 764.
- [21] Pickands III, J. (1975). Statistical inference using extreme order statistics.*Ann. Statist.*, 119 131.
- [22] Reiss, R.D., and Thomas, M. (2007). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*.Birkhäuser, Basel.
- [23] Samah, Bateka (2010). *Determination du nombre de statistiques d’ordre extrême*. Thèse de doctorat de université Mohamed khider, Biskra, Algeria.

- [24] Smith, R.L., (1987). Estimating tails of probability lois. *The Annals of Statistics*. 3, 1174-1207
- [25] S, Nadarajah and S.A.A. Bakar (2012) : New composite models for the Danish fire insurance data, *Scandinavian Actuarial Journal*, DOI :10.1080/03461238.2012.695748
- [26] Von Mises R. (1954), La distribution de la plus grande de n valeurs. *Amer. Math. Soc.*, 2 : 271 – 294.

Annexe A : Code sous logiciel *R*

```
# Density of the GEV distribution "Densité des Lois de Valeurs Extrêmes"
op= par(mfrow=c(1,2), mar=c(3,2,4,2)+.1)
library(evd)
curve(dfrechets(x,shape=1),xlim=c(-4,4),ylim=c(0,1),ylab="",col='red',main="Densité",lty=2)
curve(dgumbel(x),add=T,lty=1)
curve(drweibull(x,shape=1),col="blue",add=T, lty=4)
legend("topleft", c("Frechet", "Gumbel", "Weibull"),pt.bg="white", lty=c(2,1,4),col = c('red','black','bl

# Distribution of the GEV "La fonction de Distribution des Lois de Valeurs Extremes"
curve(pfrechet(x,shape=1),xlim=c(-4,4),ylim=c(0,1),ylab="",col='red',main="Distribution",lty=2)
curve(pgumbel(x),add=T,lty=1)
curve(prweibull(x,shape=1),col="blue",add=T, lty=4)
legend("topleft", c("Frechet", "Gumbel", "Weibull"), pt.bg="white", lty=c(2,1,4), col = c('red','black','bl

par(op)
mtext("GEV", side=3, line=3, font=2, cex=1.2)

# Density of the GPD distribution "Densité de la loi de Paréto généralisée"
op= par(mfrow=c(1,2), mar=c(3,2,4,2)+.1 )
library(evd)
curve(dgpd(x, shape=0), xlim=c(-4,4),ylim=c(0,1),ylab="",main="Densité",lty=1)
curve(dgpd(x, shape=1), add=T, col='red', lty=4)
curve(dgpd(x, shape=-1), add=T, col='blue',lty=2)
```

```
legend("topleft", c("g=0", "g=1", "g=-1"),pt.bg="white", lty=c(1,4,2),col = c('black','red','blue'))

# Distribution of the GPD "Distribution de la loi de Paréto généralisée"
curve(pgpd(x, shape=0), xlim=c(-4,8),ylim=c(0,1),ylab="",main="Distribution",lty=1)
curve(pgpd(x, shape=1), add=T, col='red', lty=4)
curve(pgpd(x, shape=-1), add=T, col='blue',lty=2)
legend("topleft", c("g=0", "g=1", "g=-1"),pt.bg="white", lty=c(1,4,2),col = c('black','red','blue'))
par(op)
mtext("GPD", side=3, line=3, font=2, cex=1.2)
```

Annexe B : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous.

F	Fonction de répartition
D	Domaine d'attraction
EVI	Indice des valeurs extrêmes
F^{\leftarrow}	Inverse généralisée de F
TEV	Théorie des valeurs extrêmes.
GEV	Distribution des valeurs extrêmes généralisés.
GPD	Distribution de Pareto généralisée.
H_γ	Distribution des valeur extrême.
i.i.d	Indépendente et identiquement distribuées.
va	variable aléatoire.
$X_{n,n}$	Maximum de X_1, \dots, X_n .
RV_ρ	Variation régulière au ∞ avec l'indice ρ .
L	Fonction à variation lente.
POT	Pics au-delà d'un seuil.
TCL	Théorème Centrale Limite.
\bar{F}	Fonction de survie.
Λ	Loi Gumbel.
Φ	Loi de Fréchet.
Ψ	Loi de Weibull.
u	Seuil

x_F	Point terminale .
$:=$	Egalité en définition .
\xrightarrow{p}	convergence en probabilité .
$\xrightarrow{p.s}$	Convergence presque sur .
$\hat{\gamma}^H$	Estimateur de Hill
$\hat{\gamma}^p$	Estimateur de Pickands
$\hat{\gamma}^M$	Estimateur de Moment
$l((X_1, X_2, \dots, X_n); (\gamma, \sigma))$	Fonction de vraisemblance .
<i>al</i>	Autres