

وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR -ANNABA
UNIVERSITY
UNIVERSITE BADJI MOKHTAR
ANNABA



جامعة باجي مختار
- عنابة -

Faculté des Sciences

Année : 2021

Département de Mathématiques



THÈSE

Présentée en vue de l'obtention du diplôme de
Doctorat

Etude d'un Modèle dans la Théorie des Valeurs Extrêmes

Option :

Actuariat

Par

Dehaimi Samira

DIRECTEUR DE THÈSE : Dr Hadji Mohammed lakhdar U.B.M. ANNABA

Devant le jury

PRESIDENT :	Taallah Frekh	MCA	U.B.M. ANNABA
EXAMINATRICE :	Seddik Ameer Nacira	Prof	U.B.M. ANNABA
EXAMINATEUR :	Aissaoui Mohamed Zine	Prof	U. GUELMA
EXAMINATEUR :	Ezzebsa Abdelali	M.C.A	U. GUELMA

Remerciements

Mes sincères remerciements à mon directeur de thèse Dr Hadji mohammed Lakhar de l'université Badji Mokhtar-Annaba pour le sujet très intéressant qu'il m'a proposé, ce ci m'a ouvert un nouvel horizon de la recherche. Sans ses conseils, encouragements et un suivi continu je ne serais jamais arrivé à ce fruit.

Je remercie également tous les membres de jury pour le temps consacré à examiner et évaluer ce travail.

Je remercie toutes les personnes qui m'ont aidé de près ou de loin à toute réalisation.

Enfin, j'exprime ma gratitude et ma reconnaissance à mon mari qui grâce à son esprit très scientifique, sa patience, sans support moral et encouragements continus, sans oublier ma fille mayar.

Résumé

Dans cette thèse, on présente la statistique des distributions des valeurs extrêmes et ses applications. Le but de celle-ci est l'étude et la caractérisation du comportement des valeurs extrêmes d'un échantillon de variables aléatoires, issues de deux manières principales. La première, celle de Block Maxima (BM) qui modélise des échantillons du maximum par la distribution généralisée des extrêmes (GEV), l'autre méthode celle de Peaks Over Threshold (POT) modélise la distribution des excès au dessus d'un seuil. La théorie des valeurs extrêmes est un outil statistique largement utilisé pour la modélisation et la prévention d'une distribution des valeurs extrêmes d'un échantillon, pour cela, on introduit une nouvelle distribution (GE) qui dépend de la distribution exponentielle et de la distribution Gumbel. On a étudié aussi certaines propriétés mathématiques. A la fin, on a réalisé une application avec des données réelles sur les indices boursiers CAC40 et RTS50.

Mots clés : Théorie des valeurs extrêmes, distribution du maximum, méthode "BM", méthode "POT".

Abstract

In this work, we present the statistics of distributions of extreme values and their applications. The goal of this thesis is the study and characterization of the behavior of the extreme values of a sample of random variables, resulting from two main ways, the first, that of Block Maxima (BM) which models the sample of the maximum by the distribution Generalized Extreme Value (GEV), the other method of Peaks Over Threshold (POT) models the distribution of excess above a threshold. The theory of extreme values (EVT) is a statistical tool widely used for the modeling and the prevention of a distribution of extreme values of a sample. For that, we introduce a new distribution (GE) which depends on the exponential distribution and Gumbel distribution, we studied some of its mathematical properties. Finally, an application with real data on two stock market "CAC40, RTS50" is produced.

Key Words : Extreme value theory, maximum distribution, BM method, POT method.

الملخص

في هذا العمل نقدم إحصائيات توزيع القيم المتطرفة وقيمها التطبيقات. الهدف من هذه الأطروحة هو دراسة وتوصيف سلوك القيم القصوى لعينة من المتغيرات العشوائية مأخوذة من طريقتين رئيسيتين. الأول ، هو نموذج (BM) الذي يمثل العينة القصوى من خلال تعميم تحريم التطرف (GEV) ، وهي الطريقة الأخرى التي تستخدمها قمم الذروة (POT) نماذج توزيع الفائض فوق العتبة. نظرية القيم المتطرفة (EVT) هو أداة إحصائية تستخدم على نطاق واسع لنمذجة ومنع أ توزيع القيم المتطرفة لعينة لهذا نقدم جديدة التوزيع (GE) الذي يعتمد على التوزيع الأسّي وتوزيع اللثة بيل ، التي تمت دراسة بعض الخصائص الرياضية منها. في أون ، قمنا بعمل التطبيق مع بيانات حقيقية عن مؤشر سوق الأوراق المالية الروسية "RTS"

Introduction

Depuis quelques années, la théorie des valeurs extrêmes (TVE) a reçu beaucoup d'attention aussi bien sur le plan théorique que sur le plan pratique. Les domaines d'application sont en effet très variés: hydrologie, météorologie, biologie, ingénierie, gestion de l'environnement, finances, assurance, sciences sociales etc... La TVE est une branche de la théorie des probabilités et des statistiques mathématiques qui se concentre sur l'analyse et l'inférence des 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 être source d'énormes pertes. Cette motivation provient soit du domaine des finances soit de celui des assurances qui sont très fréquents dans la TVE. Pour la littérature concernant la théorie des valeurs extrêmes nous nous référons aux ouvrages de Beirlant et al. (8), Reiss et Thomas (68). La Théorie des valeurs extrêmes qui est une théorie de très grande utilité pour les catastrophes naturelles et industrielles, elles sont des événements rares qui s'éloignent fortement de la moyenne ou de sa tendance, elle ont des conséquences désastreuses pour l'homme et son environnement. Dans ce travail on s'intéresse à la distribution des valeurs extrêmes et son application. Le principal objectif est de donner les modèles probabilistes de la (TVE) qui permettent d'extrapoler le comportement de la queue de distribution des données à partir des plus grandes observations. Ce travail est consacré à l'étude des valeurs extrêmes sur un échantillon, d'un événement dont les probabilités d'occurrence sont faibles. la TVE complète la statistique classique, il s'agit dans cette théorie de caractériser le comportement des queues des distributions à l'aide de modèles permettant un bon ajustement au-delà du maximum, elle s'intéresse non pas à la modélisation totale d'une distribution mais seulement à la queue de cette distribution à partir de lois spécifiques propres que nous allons préciser dans ce travail. Pour plus de détails voir par exemple Embrechts et al (32) et Beirlant et al (9). La rédaction de cette thèse s'articule autour de quatre chapitres. Dans le premier nous rappelons quelques notions préliminaires de la statistique d'ordre, qui est très

utile en théorie des valeurs extrêmes et d'autres définitions liées à TVE. Pour une étude approfondie, sur les statistiques d'ordre nous suggérons l'ouvrage d'Arnold et al(2007). Dans le deuxième chapitre nous présentons les distributions des valeurs extrêmes, elles établissent deux types de comportement asymptotique pour l'étude de ces valeurs. D'une part, le théorème des valeurs extrêmes donne la loi asymptotique du maximum de d'échantillon. D'autre part, le théorème de Pickands donne la loi asymptotique des excès lorsque le seuil tend vers le point terminal. Les lois limites possibles dans les deux théorèmes sont indexées par un paramètre de queue pour contrôler la lourdeur de celle-ci. Pour cela, il est essentiel de trouver le meilleur estimateur de ce paramètre. Dans La (TVE), il y a deux méthodes essentielles de modélisation, celle de block maxima (BM) qui modélise l'échantillon du maximum par la distribution généralisée des extrêmes (GEV) sur le théorème de Fisher-Tippet(35) et celle de Peaks Over Threshold (POT) qui modélise la distribution des excès au dessus d'un seuil élevé par la distribution généralisée de Pareto (GPD) sur le théorème de Balkemade Haan-Pickandis (9). Nous nous sommes concentrés sur les outils statistiques, les quantile-quantile plot (QQ plot) (voir par exemple Kimball 1960) et le mean-excess plot (ME plot), qui sont souvent plus instructifs pour détecter l'existence des valeurs extrêmes. Ces outils peuvent être utilisés comme un moyen d'ajustement dans le cas où on réussit à mieux les adapter et ajuster, dans ce chapitre, on traite aussi l'estimation des statistiques extrêmes, qui est directement liée à l'estimation de l'indice de queue et la détermination du nombre optimal des valeurs extrêmes , nous avons présenté deux méthodes d'estimateurs, les premières sont les estimations paramétriques par exemple l'estimateur de maximum de vraisemblance et des moments pondérés, les deuxièmes semi paramétriques présentent l'estimateur de l'indice de queue de Hill (1975) qui est l'estimateur le plus populaire pour l'indice de queue et l'estimateur de Pickands (1975)(63), ainsi que l'estimateur des moments Dekkar (1989)(23). Notre contribution cette thèse, est contenu dans le troisième chapitre où nous introduisons une nouvelle distribution sur les lois des valeurs extrêmes, on a choisi la combinaison des deux distributions Gumbel et Exponentielle, en utilisant la méthode transformé-transformer T-X. Les nouvelles distributions bien connues est utilisées

sont la famille de distribution de probabilité de Weibull-G a été avancée par Bourguignon et al (10). La famille exponentielle-exponentielle introduite par Gupta et Kundu (2001). Nadarajah (2006) a énoncé la distribution de Gumbel exponentielle. Cordeiro et al (18) ont suggéré la distribution de Kumaraswamy Gumbel (2011). Pour générer de nouvelles familles, Alzaatreh et al(2013) ont introduit une méthode, dans laquelle ils ont proposé une famille de distributions exponentielles $T - X$ (3) et une distribution de Weibull-Pareto avec applications (3). Nous avons étudié en détail, les expressions de fonction générée, les fonctions de quantile, la variance, le kurtosis, la méthode du maximum de vraisemblance est proposée pour estimer les paramètres de cette distribution. Le quatrième chapitre est consacré à l'application des méthodes décrites ci-dessus en utilisant le logiciel R. On prend comme échantillon d'application les cours des indices boursier "CAC40, RTS50" et on modélise ses valeurs extrêmes avec les méthodes BM et POT. Pour valider nos résultats, on a utilisé le graphe de QQ-plot sur les trois distributions GEV, GPD et GED (distribution de Gumbel-Exponentielle). Il s'agit de présenter les méthodes retenues pour la sélection du seuil optimal, finalement nous pouvons estimer les paramètres à partir des données de l'indice. On termine cette thèse par une conclusion générale et quelques perspectives pour nos futurs travaux.

Chapter 1

Rappels

1.1 Notions fondamentales sur les probabilités

Au cours de cette thèse, l'ensemble des définitions de base suivantes seront utilisées.

1.1.1 Analyse de survie

Variable aléatoire

La variable aléatoire est la notion principale de la théorie des probabilités et de la statistique.

Definition 1 Soit $(\Omega; F; P)$ un espace de probabilité. Une variable aléatoire réelle est une application X dans \mathbb{R}

$$X : w \in \Omega \rightarrow X(w) \in \mathbb{R}$$

Lorsque l'ensemble des réalisations possibles de la variable aléatoire réelle X est fini ou dénombrable, on dit que la variable aléatoire réelle X est discrète. Sinon, on dit qu'elle est continue.

Definition 2 (*Fonction de répartition*)

La fonction de répartition d'une v.a X est l'application F définie de \mathbb{R} dans $[0; 1]$ par

$$F(t) = P(X \leq t) \tag{1.1}$$

F est aussi appelée fonction de distribution ou fonction de distribution cumulée.

Definition 3 (Fonction de densité)

On appelle densité de probabilité toute application

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad (1.2)$$

$$x \rightarrow f(x) \quad (1.3)$$

telle que :

1. pour tout $x \in \mathbb{R}$ $f(x) \geq 0$;

2. $\int_{-\infty}^{+\infty} f(x)dx = 1$ (en supposant que $\int_{-\infty}^{+\infty} f(x)dx$ existe).

Ainsi la relation entre la fonction de répartition F et la fonction de densité de probabilité $f(x)$ est la suivante :

$$F(t) = P(X \leq t) = \int_{-\infty}^t f(t)dt \quad (1.4)$$

Definition 4 (Fonction de survie)

La fonction de survie est la probabilité pour que l'événement d'intérêt (le décès par exemple) T intervienne après un délai supérieur à t , autrement dit, l'événement d'intérêt T ne survienne pas avant la date t . La fonction de survie, aussi appelée queue de distribution qu'on note par $\bar{F}(t)$ définie sur \mathbb{R}_+ par

$$\bar{F}(t) = 1 - F(t) = P(X > t) \quad (1.5)$$

Remark 5 La fonction de répartition F d'une v.a X est croissante, continue à droite et vérifie

$$\lim_{t \rightarrow 0^+} F(t) = 0 \text{ et } \lim_{t \rightarrow +\infty} F(t) = 1 \quad (1.6)$$

alors que \bar{F} est une fonction décroissante, continue à gauche telle que

$$\lim_{t \rightarrow 0^+} \bar{F}(t) = 1 \text{ et } \lim_{t \rightarrow +\infty} \bar{F}(t) = 0 \quad (1.7)$$

Definition 6 (*Risque instantané h ou taux de hasard*)

Le risque instantané (ou taux d'incidence), pour t fixé caractérise la probabilité de mourir dans un petit intervalle de temps après t , conditionnellement au fait d'avoir survécu jusqu'au temps t (c'est-à-dire le risque de mort instantané pour ceux qui ont survécu). Si X est une variable continue positive représentant une durée, la fonction de hasard, notée par $h(t)$, est définie par :

$$h(t) = \frac{f(x)}{1 - F(x)} = \lim_{dx \rightarrow 0} \frac{P(t < X < t + dx)}{dx} \quad (1.8)$$

Definition 7 (*Fonction génératrice des moments*)

La notion de fonction génératrice peut être utile parfois pour calculer plus facilement les moments de certaines lois de probabilité. Pour la variable aléatoire X la fonction génératrice des moments est définie par :

$$M_x(t) = \int_{\mu}^{\infty} e^{tx} f(x) dx \quad (1.9)$$

Entropie

L'entropie est la mesure du désordre dans les systèmes physiques ou une quantité d'information qui peut être acquise par les observations des systèmes désordonnés. Claude Shannon définit une mesure formelle de l'entropie appelée entropie de Shannon.

Definition 8 (*Entropie Shannon*)

L'entropie de Shannon est une fonction décroissante d'une dispersion de variable aléatoire, elle est maximale lorsque tous les résultats sont également probables. L'entropie de Shannon est :

$$\eta = E \{ -\ln(g(x)) \} \quad (1.10)$$

Definition 9 (*Entropie de Rènyi*)

Alfred Rènyi a proposé l'entropie généralisée qui pour $s = 1$, elle se réduit à l'entropie

de Shannon, elle est définie par :

$$I_R(s) = \frac{1}{1-s} \left\{ \ln \left(\int_1^\infty g^s(x) dx \right) \right\}, s > 0, s \neq 1 \quad (1.11)$$

1.1.2 Statistique d'ordre

Definition 10 (*Statistique d'ordre*)

La statistique d'ordre de l'échantillon (X_1, \dots, X_n) est le réarrangement croissant, notée par $(X_{1,n}, \dots, X_{n,n})$ de sorte que $X_{1,n} < X_{2,n} < \dots < X_{n,n}$. En particulier, la variable aléatoire $X_{i,n}$ où $i = 1, \dots, n$ désigne la i^{ieme} statistiques d'ordre.

Les statistiques d'ordre extrêmes sont définies comme des termes du maximum et du minimum de l'échantillon (X_1, \dots, X_n) . La variable $X_{1,n}$ est la plus petite statistique d'ordre et $X_{n,n}$ est la plus grande statistique d'ordre, elles sont définies par

$$X_{1,n} = \min(X_1, X_2, \dots, X_n) \quad (1.12)$$

$$X_{n,n} = \max(X_1, X_2, \dots, X_n) \quad (1.13)$$

Corollary 11 *Lois des statistiques d'ordre*

Soit (X_1, \dots, X_n) une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d) d'une densité commune f et d'une fonction de distribution F alors on a les lois de $X_{i,n}$:

$$F_{i,n} = P(X_{i,n} \leq x) = \sum_{r=1}^n C_n^r (F(x))^r (1 - F(x))^{n-r} \quad (1.14)$$

Nous en déduisons que la fonction de densité est

$$f_{i,n}(x) = \frac{n!}{(i-1)!(n-1)!} [F(x)]^{i-1} [1 - F(x)]^{i-1} f(x) \quad (1.15)$$

où $f(x)$ est la densité de probabilité de X_i et F sa fonction de répartition associée. En utilisant la propriété d'indépendance des variables aléatoires X_1, \dots, X_n , on obtient :

La fonction de répartition de $X_{1,n}$

$$F_{1,n}(x) = p[X_{1,n} \leq x] = 1 - (1 - F(x))^n \quad (1.16)$$

La fonction de densité de $X_{1,n}$

$$f_{1,n}(x) = n[1 - F(x)]^{n-1} f(x) \quad (1.17)$$

La fonction de répartition de $X_{n,n}$

$$F_{n,n}(x) = p[M_n \leq x] = [F(x)]^n \quad (1.18)$$

La fonction de densité de $X_{n,n}$

$$f_{n,n}(x) = n[F(x)]^{n-1} f(x), \quad x \in \mathbb{R} \quad (1.19)$$

Inverse généralisée

Definition 12 (*Inverse généralisée*)

Soit X_1, X_2, \dots, X_n un échantillon issu d'une fonction de répartition commune F telle que $F(x) = p(X_i < x), i = 1, \dots, n$, on note la fonction inverse généralisée de F l'application notée F^{\rightarrow} définie par

$$Q(p) = F^{\rightarrow}(p) = \inf \{x, F(x) \geq p\}, \quad 0 < p < 1 \quad (1.20)$$

l'inverse généralisé F^{\rightarrow} coïncide avec l'inverse F^{-1} lorsque F est strictement croissante et continue.

Fonction de répartition empirique

Definition 13 La fonction de répartition empirique associée à l'échantillon (X_1, X_2, \dots, X_n) notée F_n est évaluée à l'aide des statistiques d'ordre et définie par :

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n 1_{\{X_k \leq x\}} = \frac{1}{n} \text{card} \{k \in \{1, \dots, n\} : X_k \leq x\} \quad (1.21)$$

$$F_n(x) = \begin{cases} 0 & \text{si } x < X_{(1,n)} \\ \frac{1}{n} & \text{si } X_{(1,n)} \leq x < X_{(2,n)} \\ \frac{2}{n} & \text{si } X_{(2,n)} \leq x < X_{(3,n)} \\ \dots\dots\dots & \\ \dots\dots\dots & \\ \dots\dots\dots & \\ 1 & \text{si } x \geq X_{(n,n)} \end{cases} \quad (1.22)$$

Fonction des quantiles empiriques

Definition 14 La fonction quantile de la fonction de distribution F est la fonction inverse généralisée de F définie par :

$$Q_n(t) = \sup \{s, F(s) \geq t\}, \quad 0 < t < 1 \quad (1.23)$$

$$Q_n(t) = \begin{cases} X_{i,n} & \text{si } \frac{i-1}{n} \leq t \leq \frac{i}{n} \\ X_{n,n} & \text{si } \frac{i-1}{n} \leq t \leq \frac{n}{n} \end{cases} \quad (1.24)$$

Point terminal

Definition 15 On note par x_F (resp. x_F^*) le point extrême supérieur (resp. inférieur) de la distribution F i.e la plus grande valeur possible pour $X_{k,n}$, elle peut prendre la valeur $+\infty$ (resp. $-\infty$).

$$x_F = \sup \{x \in \mathbb{R} : F(x) < 1\} \leq \infty \quad (1.25)$$

$$x_F^* = \inf \{x \in \mathbb{R} : F(x) > 0\}$$

Lois à queue lourde

Les distributions à queues lourdes sont liées à la théorie des valeurs extrêmes et permettent de modéliser beaucoup de phénomènes que l'on trouve dans différentes disciplines telles que la finance, l'hydrologie, etc. Les définitions les plus simples basées sur la comparaison avec la distribution normale(El Adlouni et al (31)) ont été proposées pour ces distributions.

Definition 16 (*Distribution à queue lourde*)

On dit que la fonction F est à queue lourde si pour $a > 0$, $b > 0$

$$\bar{F}(x) > ae^{-b} \quad (1.26)$$

où $\bar{F} = 1 - F$ est la fonction de survie et sa fonction de répartition vérifie

$$E(e^{\lambda x}) = \int_{\mathbb{R}} e^{\lambda x} dF(x) = \infty \text{ pour tout } \lambda > 0 \quad (1.27)$$

Corollary 17 *Si le Kurtosis d'une distribution F vérifie*

$$Kurtosis = \frac{\mu_4}{\mu_2^2} > 3 \quad (1.28)$$

on peut dire que cette distribution est à queue lourde, où μ_4 et μ_2 sont les moments centrés d'ordres respectifs 4 et 2 de X . Ce qui est équivalent à dire qu'une distribution est à queue lourde si et seulement si son coefficient d'asymétrie est supérieur à celui de la distribution normale (Hubert and Bendjoudi (43)).

La classe de distributions à queue peut lourde effectivement être définie pour des distributions sur \mathbb{R} et elles sont alors dites à queue lourde sur \mathbb{R} .

Definition 18 (*Distribution à queue légère*)

Une distribution F est dite à queue légère ou à queue fine si et seulement si

$$E(e^{\lambda x}) = \int_{\mathbb{R}} e^{\lambda x} dF(x) < \infty \text{ pour tout } \lambda > 0 \quad (1.29)$$

C'est-à-dire si elle n'est pas à queue lourde.

1.1.3 Les quantiles

Les quantiles d'une variable aléatoire discrète (entière) ou continue (réelle) sont les valeurs que prend cette variable pour des valeurs de probabilité p ($0 < p < 1$). On les appelle aussi fractiles, elles sont les valeurs réciproques de la fonction de répartition de la loi de probabilité considérée.

Quantile d'ordre p

Definition 19 *On appelle quantile d'ordre p , le nombre x_p défini par*

$$x_p = \inf \{x \in \mathbb{R} : F(x) \geq p\} \text{ avec } p \in [0, 1] \quad (1.30)$$

Remark 20 *Si F est strictement croissante et continue, alors x_p est l'unique nombre réel tel que*

$$F(x_p) = p \quad (1.31)$$

Fonction quantile de queue

La fonction quantile de queue est définie par

$$U(t) = F^{-1}\left(1 - \frac{1}{t}\right) \text{ avec } 1 < t < \infty \quad (1.32)$$

F^{-1} étant l'inverse généralisée de F

Quantile extrême

On appelle quantile extrême le quantile d'ordre $(1 - p)$, défini par

$$\begin{aligned} x_p &= \inf \{x \in \mathbb{R} : F(x) \geq p\} \\ &= F^{-1}(1 - p) \end{aligned} \tag{1.33}$$

Où p est un nombre proche de zéro.

Quantile empirique

La fonction quantile empirique de l'échantillon (X_1, \dots, X_n) est donnée par :

$$Q_n(p) = F_n^{-1} = \inf \{x \in \mathbb{R} : F_n(x) \geq p\} \text{ avec } 0 < p < 1 \tag{1.34}$$

La fonction quantile empirique de la queue correspondante est définie par :

$$U(t) = F^{-1}\left(1 - \frac{1}{t}\right) \text{ avec } 0 < t < 1 \tag{1.35}$$

Definition 21 Soit X_1, \dots, X_n un échantillon issu d'une loi F , et $\{X_{(1,n)}, \dots, X_{(n,n)}\}$ l'échantillon ordonné. Soit $p \in]0, 1[$, on appelle quantile empirique d'ordre p , la variable aléatoire notée $\hat{X}_{(np,n)}$ définie par

$$\hat{X}_{(np,n)} = \begin{cases} \frac{X_{(np;n)} + X_{(np+1;n)}}{2} & \text{si } np \in \mathbb{N} \\ X_{([np]+1,n)} & \text{sinon} \end{cases} \tag{1.36}$$

où $[np]$ désigne la partie entière de np

En particulier $X_{([n/2]+1,n)}$ est la médiane empirique.

Conclusion

nous avons rappelé dans ce chapitre quelques concepts fondamentaux sur la statistique des valeurs extrêmes, ainsi que quelques notions fondamentales sur l'analyse de survie.

Chapter 2

Théorie des valeurs extrêmes

Lorsque l'on étudie un phénomène aléatoire, on s'intéresse principalement à la partie dite centrale de la loi modélisant au mieux le phénomène considéré (calcul de l'espérance, la médiane, la variance, l'utilisation du théorème central limite TCL, etc). Cependant, l'étude des "grandes" valeurs (ou de manière équivalente des "petites" valeurs) du phénomène est parfois essentielle lorsqu'il s'agit par exemple de quantifier le risque pour une compagnie d'assurance ou bien connaître la fréquence des crues d'une rivière. La théorie des valeurs extrêmes propose un cadre théorique solide pour l'étude de ces grandes (ou petites) valeurs dites extrêmes. En ce sens, on peut considérer que cette théorie est un complément à la théorie de la statistique classique, qui est principalement basée sur l'étude de la moyenne d'un échantillon plutôt que des observations extrêmes. La TVE est une vaste théorie dont le but est d'étudier les événements rares c-à-d les événements dont la probabilité d'apparition est faible. Elle a pour but de caractériser le comportement des valeurs extrêmes d'un échantillon de variables aléatoires. On peut citer par exemple comme une bonne référence classique sur la théorie et les applications des valeurs extrêmes les ouvrages de Embrechts et al(1997), qui donnent un aperçu sur la théorie mathématique avec une orientation vers des applications en assurance et en finance, le livre de Coles (2001)(16) se concentre sur l'analyse de données et l'inférence statistique pour les valeurs extrêmes, Reiss et Thomas (2007)(68) font le point sur les différentes techniques existantes.

2.1 Théorème de Fisher-Tippett (1928)

Le Théorème de Fisher-Tippett fournit, en quelque sorte, un complément au théorème TCL (théorème central limite) dans le cas d'événements extrêmes. Le fondement du théorème de Fisher-Tippett a été formulée plusieurs fois, au début, il a été annoncé en 1928(35), ils ont proposé une solution au problème suivant

$$\lim_{n \rightarrow \infty} F_{M_n}(x) = \lim_{n \rightarrow \infty} (F(x))^n = \begin{cases} 0 & \text{si } F(x) < 1 \\ 1 & \text{si } F(x) = 1 \end{cases} \quad (2.1)$$

et donc F_{M_n} converge vers 0 ou 1 selon la valeur de $F(x)$ et donc la loi asymptotique de M_n est dégénérée pour cela, on a cherché une loi non dégénérée pour le maximum de X fournie par le théorème des distributions extrêmes qui donne une condition nécessaire et suffisante pour l'existence d'une loi limite non dégénérée pour le maximum. Ce théorème a été démontré par Gendenko (1943)(46) qui donne la forme des lois limites, ensuite il a été repris et simplifié par Haan en (1970) puis Jenkison (1955), qui ont donné l'expression générale, ce théorème caractérise la loi des distributions extrêmes. le principal problème mathématique posé dans la théorie des valeurs extrêmes concerne la recherche de distributions des X_i , Fisher et Tippett ont démontré l'existence des suites de normalisation $a_n > 0$ et b_n et une loi non-dégénérée H_ξ dans le théorème suivant.

Theorem 22 "*Fisher-Tippett (1928)*" Soit M_n le maximum d'un échantillon aléatoire X_1, \dots, X_n de fonction de répartition F . S'il existe deux suites réelles $a_n > 0$ et $b_n \in \mathbb{R}$ appelées constantes de normalisation, telles que

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H_\xi(x), \quad x \in \mathbb{R} \quad (2.2)$$

a_n est un paramètre de position.

b_n est un paramètre d'échelle.

ξ est un paramètre de forme.

où $H_\xi(x)$ est non dégénérée, ce résultat est le fondement de la TVE.

Ce théorème n'est valable que si les suites $\{a_n\}$ et $\{b_n\}$ existent et admettent des limites.

Ce théorème est un résultat très important car il n'est pas nécessaire de faire des hypothèses paramétriques sur la loi des X_i . Jenkinson (1955) a généralisé les résultats de Fréchet (1927) et Fisher et Tippett (1928) en montrant que la loi du maximum de l'échantillon tend vers la loi généralisée des valeurs extrêmes. Pour la preuve du théorème, nous renvoyons le lecteur à Embrechts et al, 1997, chapitre 3, p. 122.

2.2 Distribution généralisée des extrêmes "GEV"

Definition 23 (*Représentation de Jenkinson - Von Mises*)

Il est possible de rassembler les trois familles de lois (voir 2.2.1) (Jenkinson 1955) en une seule famille. Cette famille est dite loi des valeurs extrêmes GEV (distribution généralisée des extrêmes), elle est la seule et unique loi de probabilité qui modélise le comportement du maximum d'un échantillon

$$H_\xi(x) = \begin{cases} \exp(-(1 + \xi x)^{-1/\xi}), & \text{si } \xi \neq 0 \\ \exp(-e^{-x}), & \text{si } \xi = 0 \end{cases} \quad (2.3)$$

Le paramètre qui apparaît dans la formule (2.3) est appelé "indice de queue". Le comportement de la queue d'une suite de variables aléatoires sera complètement caractérisé par le paramètre ξ , son signe a une forte influence sur la distribution des extrêmes. Pour faciliter le travail avec les trois distributions limites, Jenkinson-Von Mises a donné une représentation, obtenue en introduisant les paramètres de localisation et de dispersion dans la paramétrisation des distributions extrêmes.

Alors pour les variables non centrées et non réduites, on peut écrire $H_\xi(x)$ sous une forme plus générale, notée par $H_{\mu,\delta,\xi}(x)$ dans laquelle on fait apparaître un paramètre de localisation $\mu \in \mathbb{R}$ et un paramètre d'échelle $\delta > 0$

La distribution $H_{\mu,\delta,\xi}(x)$ s'écrit comme suit

$$H_{\xi,\delta,\mu}(x) = \begin{cases} \exp\left(-\left(1 + \xi\left(\frac{x-\mu}{\delta}\right)\right)^{-1/\xi}\right), & \text{si } \xi \neq 0 \\ \exp\left(-e^{-\frac{x-\mu}{\delta}}\right), & \text{si } \xi = 0 \end{cases} \quad (2.4)$$

On peut facilement montrer que la fonction de densité de la distribution GEV correspondant à

$$h_{\xi,\delta,\mu}(x) = \begin{cases} \frac{1}{\delta} \left[1 + \xi\left(\frac{x-\mu}{\delta}\right)^{-\left(\frac{1+\xi}{\xi}\right)} H_{\mu,\delta,\xi}(x) \right] & \text{si } \xi \neq 0 \\ \frac{1}{\delta} \exp\left\{-\left(\frac{x-\mu}{\delta}\right) - \exp\left[-\left(\frac{x-\mu}{\delta}\right)\right]\right\} & \text{si } \xi = 0 \end{cases} \quad (2.5)$$

Le quantile $Q(p)$ de la distribution est donné par la formule suivante

$$Q(p) = H_{\xi,\delta,\mu}^{-1} = \begin{cases} \mu - \delta\xi^{-1} \left[1 - (-\log(p))^{-\xi} \right], & \text{si } \xi \neq 0 \\ \mu - \delta \log(-\log p), & \text{si } \xi = 0 \end{cases} \quad (2.6)$$

2.2.1 Les lois limites possibles

La loi GEV regroupe des distributions selon le signe du paramètre ξ , on distingue trois cas limites possibles.

Distribution de probabilité de Fréchet

Le mathématicien français Maurice René Fréchet a introduit une distribution considérée comme loi de valeurs extrêmes. Les travaux de Fréchet ont inspiré les recherches de Fisher-Tippett (1928).

Definition 24 *si $\xi > 0$, on dit que F appartient au domaine d'attraction de Fréchet (voir section de domaines d'attraction des lois de GEV). Il contient les lois dont la fonction de survie est à décroissance polynomiale comme une fonction puissance, i.e. les lois à queues lourdes. On dit qu'une variable aléatoire X est de distribution de Fréchet de paramètres μ ,*

$\delta > 0$ et $\xi > 0$, si sa fonction de distribution cumulative est

$$F(x) = \begin{cases} \exp - \left(1 + \xi \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}}, & \text{si } x > 0 \quad \xi > 0 \\ 0, & \text{si } x \leq 0 \end{cases} \quad (2.7)$$

La densité de probabilité de la distribution de Fréchet s'écrit

$$f(x) = \begin{cases} \left(\frac{\xi}{\delta} \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}-1} \exp - \left(1 + \xi \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}}, & \text{si } x > 0 \quad \xi > 0 \\ 0, & \text{si } x \leq 0 \end{cases} \quad (2.8)$$

et la fonction de survie est

$$\bar{F}(x) = 1 - F(x) = \begin{cases} 1 - \exp - \left(1 + \xi \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}}, & \text{si } x > 0 \quad \xi > 0 \\ 0, & \text{si } x \leq 0 \end{cases} \quad (2.9)$$

La moyenne et la variance

Le moment d'orde 1 de la distribution est

$$E(X) = \mu + \frac{\xi}{\delta} [1 - \Gamma(1 + \xi)] \quad (2.10)$$

sa variance

$$Var(X) = \left(\frac{\xi}{\delta}\right)^2 [\Gamma(1 + 2\xi) - \Gamma^2(1 + \xi)] \quad (2.11)$$

Les skewness et kurtosis de la distribution sont

$$skewness = - \frac{|\xi| \Gamma(1 + 3\xi) - 3\Gamma(1 + 2\xi) \Gamma(1 + \xi) + 2\Gamma^3(1 + \xi)}{[\Gamma(1 + 2\xi) - \Gamma^2(1 + \xi)]^{\frac{3}{2}}} \quad (2.12)$$

$$kurtosis = \frac{|\xi| \Gamma(1 + 4\xi) - 4\Gamma(1 + \xi) \Gamma(1 + 3\xi) + 6\Gamma^2(1 + \xi) \Gamma(1 + 2\xi) - 3\Gamma^4(1 + \xi)}{[\Gamma(1 + 2\xi) - \Gamma^2(1 + \xi)]^2} \quad (2.13)$$

Le tableau ci-dessous présente les valeurs skewness et kurtosis pour différentes valeurs du ξ

ξ	<i>skewness</i>	<i>kurtosis</i>
0.05	0.867951	4.267201
1	-2	87.72
3	1190.1132395-	1005.532

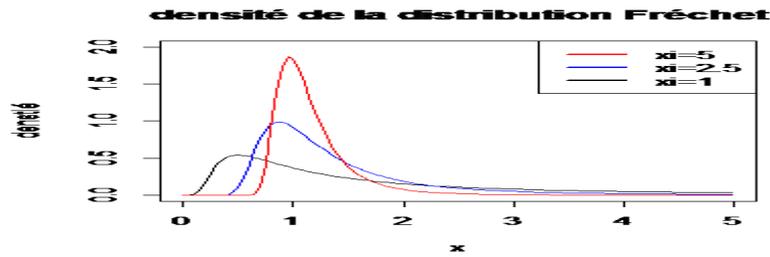


Figure 2.01: la fonction de densité de Frchet

La figure 2.01 représente la fonction de densité de Fréchet pour différentes valeurs de ξ . On remarque que plus le ξ augmente la distribution décroît rapidement, et devient plus fine.

Distribution de probabilité de Weibull

Le mathématicien suédois Waloddi Weibull (1951) a introduit la distribution de Weibull, cette dernière est couramment utilisée dans l'analyse de survie et en fiabilité des systèmes. La distribution de Weibull peut être également utilisée dans de nombreux domaines Rinne (69).

Definition 25 si $\xi < 0$, on dit que F appartient au domaine d'attraction de Weibull (voir section de domaines d'attraction des lois de GEV). Il contient des lois à queues légères. On dit qu'une variable aléatoire X est de distribution de Wiebull de paramètres $\mu, \delta > 0$, avec $\xi < 0$, si sa fonction de distribution cumulative est

$$F(x) = \begin{cases} \exp - \left(1 + \xi \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}}, & \text{si } x < 0 \quad \xi < 0 \\ 1, & \text{si } x \geq 0 \end{cases} \quad (2.14)$$

La fonction de densité de probabilité est donnée par la formule suivante

$$f(x) = \begin{cases} \left(\frac{\xi}{\delta} \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}-1} \exp - \left(1 + \xi \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}}, & \text{si } x < 0 \quad \xi < 0 \\ 0, & \text{si } x \geq 0 \end{cases} \quad (2.15)$$

La fonction de survie de Weibull est

$$\bar{F}(x) = 1 - F(x) = \begin{cases} 1 - \exp - \left(1 + \xi \left(\frac{x-\mu}{\delta}\right)\right)^{\frac{-1}{\xi}}, & \text{si } x < 0 \quad \xi < 0 \\ 1, & \text{si } x \geq 0 \end{cases} \quad (2.16)$$

La moyenne et la variance

Le moment d'ordre 1 de la distribution est

$$E(X) = \mu + \frac{\xi}{\delta} [1 - \Gamma(1 + \xi)] \quad (2.17)$$

Tel que la variance est

$$Var(X) = \left(\frac{\xi}{\delta}\right)^2 [\Gamma(1 + 2\xi) - \Gamma^2(1 + \xi)] \quad (2.18)$$

Les skewness et kurtosis de la distribution sont

$$skewness = -\frac{|\xi| \Gamma(1 + 3\xi) - 3\Gamma(1 + 2\xi) \Gamma(1 + \xi) + 2\Gamma^3(1 + \xi)}{[\Gamma(1 + 2\xi) - \Gamma^2(1 + \xi)]^{\frac{3}{2}}} \quad (2.19)$$

$$kurtosis = \frac{|\xi| \Gamma(1 + 4\xi) - 4\Gamma(1 + \xi) \Gamma(1 + 3\xi) + 6\Gamma^2(1 + \xi) \Gamma(1 + 2\xi) - 3\Gamma^4(1 + \xi)}{[\Gamma(1 + 2\xi) - \Gamma^2(1 + \xi)]^2} \quad (2.20)$$

Le tableau ci-dessous présente les valeurs des skewness et kurtosis pour différentes valeurs du ξ

ξ	skewness	kurtosis
-0.05	1.473884	7.333494
-1	1.910339	10.978566
-3	13.483552	-145.295346

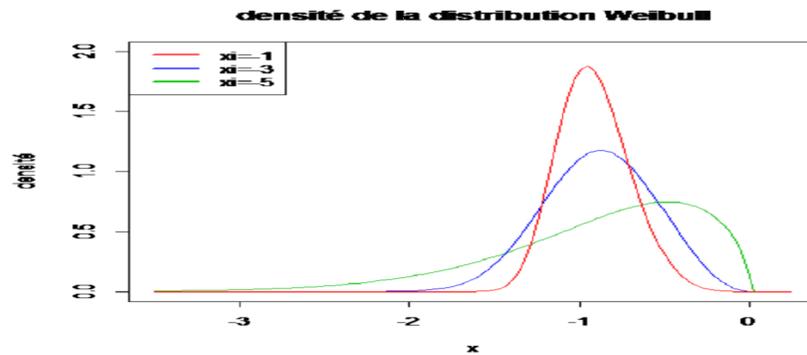


Figure 2.02: la fonction de densité de Weibull

La figure 2.02 représente la fonction de densité de Weibull pour différentes valeurs de ξ . Contrairement à la loi de Fréchet si ξ diminue, la courbe descend rapidement.

Distribution de probabilité de Gumbel

La distribution Gumbel a été introduite par le mathématicien Américain Emil Julius Gumbel (Gumbel (2012)) qui l'a développée comme loi de valeurs extrêmes. La distribution de Gumbel est l'une des distributions les plus populaires dans de nombreux domaines grâce à sa flexibilité.

Definition 26 si $\xi = 0$, on dit que F appartient au domaine d'attraction de Gumbel. Il contient les lois dont la fonction de survie est à décroissance exponentielle, i.e il n'y a pas

de queue de distribution. On dit qu'une variable aléatoire X est de distribution de Gumbel de paramètres μ et $\delta > 0$, si seulement si la fonction de distribution cumulative s'écrit

$$F(x) = \exp\left(-\exp\left(-\frac{x-\mu}{\delta}\right)\right), \text{ si } x \in \mathbb{R} \quad (2.21)$$

La fonction de densité de probabilité est donnée par la formule suivante

$$f(x) = \frac{1}{\delta} \left(\exp\left(-\frac{x-\mu}{\delta}\right) \exp\left(-\exp\left(-\frac{x-\mu}{\delta}\right)\right) \right), \text{ si } x \in \mathbb{R} \quad (2.22)$$

La fonction de survie est

$$\bar{F}(x) = 1 - F(x) = \begin{cases} 1 - \exp\left(-\exp\left(-\frac{x-\mu}{\delta}\right)\right), & \text{si } x \in \mathbb{R}, \text{ si } x < 0 \\ 1, & \text{si } x \geq 0 \end{cases} \quad (2.23)$$

La moyenne et la variance

Le moment d'ordre 1 de la distribution est

$$E(X) = \mu + \delta C \quad (2.24)$$

Où $C = 0.5772$ est la constante de Euler.

Tel que la variance

$$Var(X) = \frac{\delta^2 \pi^2}{6} \quad (2.25)$$

Les skewness et kurtosis de la distribution sont

$$skewness = 1.139 \quad (2.26)$$

$$kurtosis = 2.40 \quad (2.27)$$

Les lois des valeurs extrêmes généralisées correspondent à une translation et un change-

ment d'échelle proches des lois des valeurs extrêmes. Voir la proposition suivante.

Proposition 27 *Nous avons les correspondances suivantes (les relations suivantes entre Φ_ξ, Ψ_ξ et Δ en terme de $H_{\xi,\delta,\mu}$).*

On pose Φ_ξ la distribution de Fréchet, Ψ_ξ la distribution de Weibull et Δ la distribution de Gumbel

$$H_{1, \frac{1}{\xi}, \frac{1}{\xi}}(\xi(x-1)) = \Phi_\xi(x) \text{ si } \xi > 0 \quad (2.28)$$

$$H_{0, \frac{1}{\xi}, \frac{-1}{\xi}}(\xi(x+1)) = \Psi_\xi(x) \text{ si } \xi < 0$$

$$H_{0,1,0}(\xi(x)) = \Delta_\xi(x) \text{ si } \xi = 0$$

Soit Y une variable aléatoire positive, alors les affirmations suivantes sont équivalentes

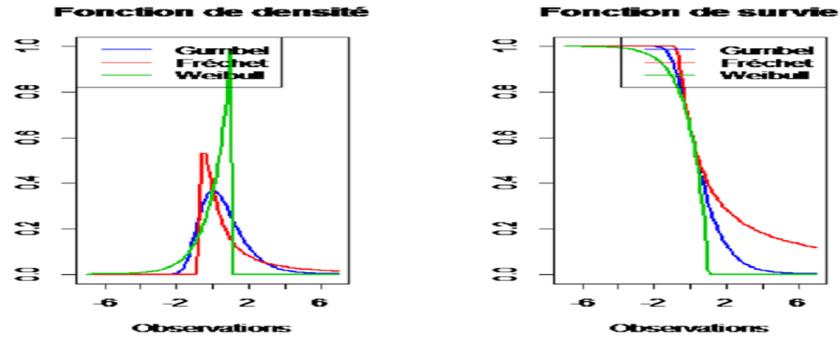
$$Y \sim \Phi_\xi \quad (2.29)$$

$$\ln Y^\xi \sim \Delta \quad (2.30)$$

$$-Y^{-1} \sim \Psi_\xi \quad (2.31)$$

Nous exprimons les trois distributions des valeurs extrêmes Φ_ξ, Ψ_ξ et Δ en termes de

la loi GEV. La figure (2.03) illustre le comportement de la distribution GEV.



la figure (2.03) les fonctions de survie et les fonctions de densité de GEV

On remarque d'après la figure (2.03) les fonctions de survie et les fonctions de densité que la distribution de Fréchet décroît moins vite que celles de Gumbel et de Weibull. C'est pour cela qu'on dit que la distribution de Fréchet est à queue lourde.

Gédenklo (1943)(46) donne une condition nécessaire et suffisante pour l'existence des constantes de normalisation quand la loi de la variable X est connue. Ces constantes peuvent être utilisées pour déterminer le type de la loi. Comme la GEV est très utile dans l'application de la TVE pour estimer le paramètre ξ . Les statisticiens ont souvent recours à une méthode qui s'appelle bloc maxima (BM).

Definition 28 "*Bloc maxima*"

Le Bloc maxima BM est une méthode principale de la TVE qui modélise la distribution extrême par la loi de la GEV, s'appuie sur le théorème de Fisher-Tippett(35).

Nous supposons savoir une suite d'observations X_1, \dots, X_n issues d'une fonction de distribution inconnue $F(\cdot)$ (c-à-d aucune hypothèse n'est présumée). Cet échantillon est séparé en k blocs disjoints de même longueur s (un bloc peut correspondre à un mois, un an, etc). Le bloc i est tel que $\{X_1^{(i)}, \dots, X_s^{(i)}\}$. Nous formons avec les maxima de ces k blocs un échantillon de maximum Y_1, Y_2, \dots, Y_k où $Y_i = \max(X_1^{(i)}, \dots, X_s^{(i)})$. Les variables

Y_1, \dots, Y_k sont supposées i.i.d et elles sont de loi GEV d'après le théorème de Fisher-Tippet.

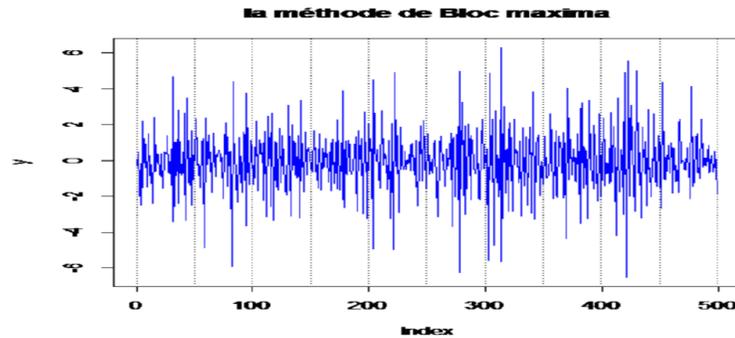


Figure 2.04: La mthode bloc maxima (BM)

/

Sélection de la taille des blocs

Il n'y a pas d'outils statistiques d'aide à la sélection de la taille, Il faut cependant que la taille des blocs soit suffisamment grande de sorte que nous ayons un nombre suffisant de maxima pour que l'estimation des paramètres de la GEV soit assez précise. Il est donc courant de prendre $s = 21$ (pour un mois) ou $s = 251$ (pour un an), valeurs qui représentent un bon compromis. Cette méthode à un inconvénient qui entraine une perte de certaines informations c-à-d certains blocs contiennent plusieurs valeurs extrêmes alors que d'autre bloc n'ont contiennent pas.

2.3 Domaines d'attraction des lois de GEV

Un domaine d'attraction est l'ensemble des distributions pour lesquelles la distribution limite du maximum converge vers la même limite de la GEV. Il faut bien noter que le paramètre conditionne le type de la loi des valeurs extrêmes. Nous présentons dans ce qui suit, les domaines d'attractions dans les trois types correspondant au paramètre. Les différentes caractérisations des trois domaines d'attraction de Fréchet, Gumbel et de Weibull ont été proposées dans (Resnick (1987), Haan and Ferreira (2007)(21), Embrechts et al. (2013)(33)).

Nous allons donner des conditions sur la fonction de répartition F pour qu'elle appartienne à l'un des trois domaines d'attraction.

Definition 29 On appelle fonction à variations régulières la caractérisation des domaines d'attraction des lois. Soit $U(x)$ une fonction positive à variation régulière (à l'infini) d'indice δ , on a

$$\lim_{x \rightarrow \infty} \frac{U(\lambda x)}{U(x)} = \lambda^\delta, \text{ pour tout } \lambda > 0 \quad (2.32)$$

Si $\delta = 0$: on dit que la fonction est à variation lente.

On notera dans la suite RV_δ l'ensemble des fonctions à variations régulières d'indice δ .

Remark 30 chaque fonction $U \in RV_\delta$ à variation régulières avec indice δ et pour $x > 0$ s'écrivent sous la forme

$$U(x) = x^\delta L(x) \quad (2.33)$$

Où $L(x) \in RV_0$

Theorem 31 (Karamata)

1- Théorème de la Convergence Uniforme

Si $U \in RV_0$, alors

$$\lim_{x \rightarrow \infty} \frac{U(x\lambda)}{U(x)} = 1 \quad (2.34)$$

est une convergence uniforme pour $t \in [a, b]$ où $0 < a < b < \infty$.

2 -(Représentation de Karamata)

Toute fonction à variations lentes s'écrit sous la forme

$$L(x) = c(x) \exp \left\{ \int_1^x \frac{\Delta(u)}{u} du \right\} \quad (2.35)$$

où c et Δ sont des fonctions mesurables avec $c(x) \rightarrow c > 0$ et $\Delta(u) \rightarrow 0$ quand $x \rightarrow \infty$.

La loi de Gumbel

Dans ce domaine le paramètre ξ est nul, la fonction de survie est à décroissance de type exponentielle, ce qui permet de caractériser dans ce cas, le domaine d'attraction de Gumbel qui est délicatement traitable, car il n'y a pas de lien direct entre la loi et les fonctions à variations lentes. Dans le théorème ci-dessous, on trouve une caractérisation de Von Mises (1936) pour le domaine d'attraction de Gumbel.

Definition 32 (Fonction de von Mises)

Soit F une fonction de répartition de point terminal x_F fini ou infini. S'il existe $z < x$ tel que $a(t)$ une fonction de répartition appartient au domaine d'attraction de Gumbel, s'il existe $z < x_f < \infty$ tel que

$$1 - F(x) = c \exp \left(- \int_z^x \frac{1}{a(t)} dt \right), z < x < x_F \quad (2.36)$$

Où $c > 0$ et a une fonction positive absolument continue de densité a' vérifiant $\lim_{x \rightarrow x_F} a'(x) = 0$: Alors F est une fonction de Von-Mises et a est sa fonction auxiliaire.

Theorem 33 F appartient au domaine d'attraction de Gumbel si et seulement si il existe une fonction de Von-Mises F^* telle que pour $z < x < x_F$ on ait

$$1 - F(x) = c(x)[1 - F^*(x)] = c(x) \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\} \quad (2.37)$$

$c(x) \rightarrow c > 0$ lorsque $x \rightarrow x_F$

Quelques exemples de lois qui appartiennent au domaine d'attraction de Gumbel

loi	$1 - F(x)$
Benktander, α, β	$x^{-(1-\beta)} \exp(-\frac{\alpha}{\beta}x\beta)$
logistic	$\frac{1}{1-\exp(x)}$

Loi de Fréchet

Les lois appartenant à ce domaine sont caractérisées par une queue lourde $\xi > 0$ à décroissance lente à l'infini, c-à-d que le point terminal $x_f = +\infty$ et $1 - F = x^{\frac{1}{\xi}}l(x)$. Les fonctions aux domaines d'attraction de Fréchet sont des fonctions à variation lente. La démonstration de ce résultat est énoncée par Géndenklo (1943) par le théorème suivant.

Theorem 34 *Si F la fonction de distribution appartient au domaine d'attraction de Fréchet, alors la variation régulière d'indice $\frac{-1}{\xi}$ est*

$$\lim \frac{\bar{F}(tx)}{\bar{F}(x)} = x^{\frac{-1}{\xi}} \quad (2.38)$$

Où les suites de normalisation sont pour tout n

$$\begin{aligned} a_n &= F^{\leftarrow} \left(1 - \frac{1}{n} \right) = U(n) \\ b_n &= 0 \end{aligned} \quad (2.39)$$

Corollary 35 *On peut écrire une fonction de répartition F au domaine d'attraction de Fréchet sous la forme*

$$F(x) = 1 - x^{-\frac{1}{\xi}}L(x) \quad (2.40)$$

Quelques lois qui appartiennent au domaine d'attraction de Fréchet

Loi	$1 - F(w(F))$	ξ	$L(x)$
Pareto (α) avec $\alpha > 0$	$\frac{1}{x}, x > 1$	$\frac{1}{\alpha}$	1
Burr (ζ, λ, η)	$\left(\frac{\eta}{\eta + x^\zeta} \right)^\lambda, x > 0$	$\frac{1}{\lambda x}$	$\left(\frac{\eta}{\eta + x^\zeta} \right)^\lambda$
reverse Burr	$\left(1 + \frac{\xi x}{\delta} \right)^{\frac{-1}{\xi}}, x > 0$	ξ	$\left(\frac{\delta}{\xi} \right)^{\frac{-1}{\xi}} \left(1 + \frac{\xi x}{\delta} \right)^{\frac{-1}{\xi}}$

Loi de Weibull

Dans ce cas, la queue $\xi < 0$, et les lois sont bornées à droite et par conséquent le point final x_f est fini $x_F < +\infty$ et $1 - F(x) = (x_f - x)^{-\frac{1}{\xi}} L[(x_f - x)^{-1}]$. En (1943) Gédenklo a donné une caractérisation pour le domaine d'attraction de Weibull formulée par le théorème ci-dessous.

Theorem 36 *Une fonction de répartition F au domaine d'attraction de Weibull est donnée par*

$$F(x_f - \frac{1}{x}) = x^{-\frac{1}{\xi}} L(x)$$

dans ce cas un choix possible pour les suites a_n et b_n est

$$\begin{aligned} a_n &= x_f - F^{-1}(1 - \frac{1}{n}) \\ b_n &= x_f \end{aligned} \tag{2.41}$$

Où \bar{F} est la fonction de survie.

Quelques lois qui appartiennent au domaine d'attraction de Weibull

Loi	$1 - F(w(F) - \frac{1}{x})$	ξ	$L(x)$
<i>Uniforme</i>	$\frac{1}{x}, x > 1$	-1	1
<i>weibull</i>	$1 - \exp(-x^{-\alpha})$	$-\frac{1}{\alpha}$	$1 - \frac{x^{-\alpha}}{2} + 0(x^{-\alpha})$
<i>revse Burr</i>	$(\frac{\beta}{\beta+x^\tau})^\lambda$ avec $x > 0$	$-\frac{1}{\lambda\tau}$	$\beta^\lambda (1 - \lambda\beta x^{-\tau}) + 0(x^{-\tau})$

2.4 La distribution des excès, approche POT

Dans les statistiques, la distribution généralisée de Pareto (*GPD*) est une famille de lois de probabilités continues, elle est employée souvent pour modéliser les queues d'une distribution. Cette distribution joue un rôle important dans l'analyse statistique des événements extrêmes, en fournissant une approximation appropriée pour l'excès au-delà d'un grand seuil.

Definition 37 On dit que la variable aléatoire X est distribuée suivant une loi GPD si sa fonction de répartition G est donnée par

$$G_{\xi,\delta}(x) = \begin{cases} 1 - \left(1 + \xi \frac{x}{\delta}\right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - e^{-\frac{x}{\delta}} & \text{si } \xi = 0 \end{cases} \quad (2.42)$$

où $\delta > 0$ est un paramètre d'échelle et $\xi \in R$ est un paramètre de forme.

Si $x \geq 0$ lorsque $\xi \geq 0$ il devient $0 \leq x \leq -\frac{\delta}{\xi}$ lorsque $\xi < 0$.

Pour les propriétés et la preuve de cette fonction ($G_{\xi,\delta}(x)$), voir exemple à l'ouvrage de Embrechts et al(32).

Remark 38 La densité de la distribution GPD s'écrit comme suit

$$g_{\xi,\delta}(x) = \begin{cases} \delta^{-1} \left(1 + \xi \frac{x}{\delta}\right)^{-\frac{1}{\xi}-1} & \text{si } \xi \neq 0 \\ \delta^{-1} \exp\left(-\frac{x}{\delta}\right), & \text{si } \xi = 0 \end{cases} \quad (2.43)$$

Le quantile $Q(s)$ de la distribution $G_{\xi,\delta}$, est donné par

$$Q(s) = u + \frac{\delta}{\xi} \left\{ \left(\frac{n}{N_u} s\right)^{-\xi} - 1 \right\} \quad (2.44)$$

Remark 39 Il y a un rapport simple entre la GPD standard G_ξ et la GEV standard H_ξ tels que

$$G_\xi(x) = 1 + \log H_\xi(x) \quad (2.45)$$

2.4.1 Les lois limites

La loi GPD regroupe trois distributions selon les valeurs du paramètre

Lorsque $\xi > 0$, c'est la distribution Pareto usuelle

$$G_{\mu,\sigma,\xi}(x) = 1 - \left(1 + \xi \frac{(x - \mu)}{\delta}\right)^{-\frac{1}{\xi}}, \text{ si } x \geq 1 \quad (2.46)$$

Lorsque $\xi < 0$, nous avons la distribution de Pareto type II

$$G_{\mu,\sigma,\xi}(x) = 1 - \left(1 + \xi \frac{(x - \mu)}{\delta}\right)^{-\frac{1}{\xi}}, \text{ si } -1 \leq x \leq 0 \quad (2.47)$$

Et si $\xi = 0$ on a la distribution exponentielle

$$G_{\mu,\sigma,0}(x) = 1 - e^{-\frac{x - \mu}{\delta}}, \text{ si } x \geq 0 \quad (2.48)$$

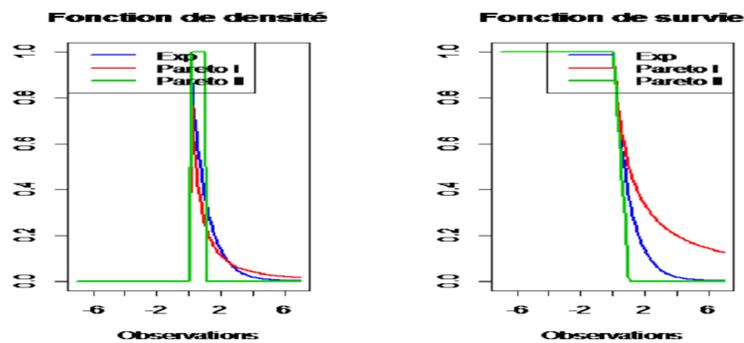


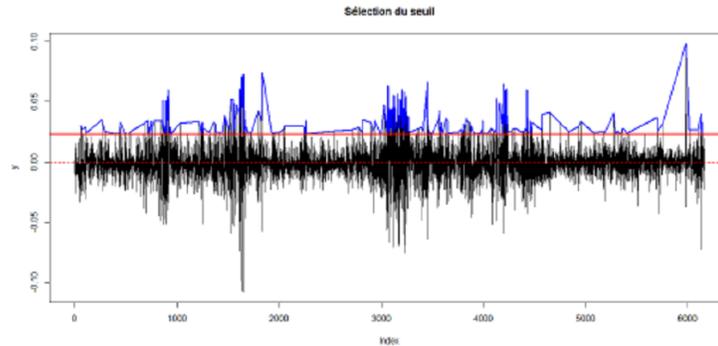
Figure 2.05: La fonction de densité et la fonction de survie de la loi
GPD

On voit d'après la figure 2.05 présente les fonctions de survie et les fonctions de densité de la GPD, la distribution de Pareto *I* décroît moins vite que celles de l'Exponentielle et de Pareto *II*. C'est pour cela qu'on dit que la distribution de Pareto *I* est à queue lourde.

2.5 Distribution conditionnelle des excès

Pour remédier à l'inconvénient de la méthode de BM les statisticiens (63) utilisent une autre méthode qui permet de prendre en compte beaucoup plus de données pour assurer la précision dans l'estimation des paramètres, cette méthode qu'on appelle "Peaks Over Threshold" (POT) consiste à utiliser les observations qui dépassent un certain seuil et plus particulièrement la différence entre le seuil et ces observations appelée excès (ou le seuil est un réel ni trop faible pour ne pas prendre en considération des valeurs non extrêmes,

ni trop élevé pour avoir suffisamment d'observations).



La figure 2.06 : la méthode de POT

La méthode POT est une méthode qui modélise la distribution des excès au dessus d'un seuil élevé par la loi de GPD qui s'appuie sur le théorème de Belkema-de Haan-Pickands. Le principe de POT est d'estimer la distribution des excès par une classe de GPD dépendant de paramètres. Cette méthode initialement développée par Pickands (1975) et abondamment étudiée par divers auteurs tels que Smith (70) et Davison et Smit(1990).

2.5.1 la loi des excès

On définit les excès au dessus d'un seuil u comme un ensemble des variables aléatoires conditionnelles $Y = [X - u | X > u]$ on cherche à partir de la distribution de F de X la distribution conditionnelle F_u par rapport à u pour les variables aléatoires dépassant ce seuil.

On définit un seuil u par

$$N_u = \text{card} \{i, i = 1, \dots, n, : X_i > u\} \quad (2.49)$$

N_u est le nombre de dépassements du seuil u par les $(X_i)_{1 \leq i \leq n}$.

Definition 40 " La fonction de répartition des excès " On appelle la distribution des excès

de la variable aléatoire X au dessus d'un seuil u la distribution de probabilité de la variable aléatoire $Y = [X - u | X > u]$, donnée par sa fonction de répartition F_u suivante

$$F_u(x) = P(X - u \leq y | X > u) = P(X \leq y + u | X > u) \quad (2.50)$$

$$= \frac{P(\{X \leq y + u\} \cap X > u)}{p(X > u)} = \frac{p(u < X \leq y + u)}{p(X > u)} \quad (2.51)$$

$$= \frac{F(y + u) - F(u)}{1 - F(u)} \quad (2.52)$$

Donc pour tout $y \in R$ on a

$$F_u(x) = P(X - u \leq y | X > u) = \begin{cases} 0 & \text{si } y \leq 0 \\ \frac{F(y+u)-F(u)}{1-F(u)} & 0 \leq y \leq x_0 - u \\ 1 & y \geq x - u \end{cases}$$

La fonction de survie de Y au dessus d'un seuil u est définie par

$$P(Y > y) = \bar{F}_u(y) = 1 - F_u(y) = \frac{1 - F(y + u)}{1 - F(u)} \quad (2.53)$$

2.5.2 choix de seuil

Le choix de seuil est très important car il induit une grande variabilité dans l'estimateur extrême, dans le paramètre de loi des excès si le seuil est très élevé: les écarts types des estimateurs seront très importants et aussi la taille de l'échantillon est insuffisamment grande pour les estimateurs des paramètres et induit une grande variance des estimateurs. Si le seuil est très bas: on va prendre en considération des valeurs non extrêmes et les estimateurs seront biaisés, en effet le seuil doit être suffisamment grand pour satisfaire l'asymptotique du modèle mais pas élevé et non bas, dans ce cas la manière la plus naturelle de modéliser la fonction des excès au delà d'un seuil suffisamment élevé est l'utilisation de la loi de GPD. Il suffit alors de la définir par le théorème suivant.

2.6 Théorème de Balkema-De Haan-Pickands

Balkema- De Haan et Pickands(9) ont proposé le théorème ci-après, qui va être le résultat théorique central de la TVE. Ce théorème précise la distribution conditionnelle des excès lorsque le seuil déterministe tend vers le point terminal.

La distribution limite d'un excès est donnée par le théorème de Balkema-De Haan -Pickands (Balkema and De Haan (1974)(9), Pickands et al (1975)).

Le théorème énonce que si F appartient à l'un des trois domaines d'attraction de la loi limite des extrêmes (Fréchet, Gumbel ou Weibull), alors il existe une fonction de répartition des excès au-delà de u , noté F_u qui peut être approchée par une loi de Pareto généralisée (GPD).

Theorem 41 *Soit F_u la distribution des excès, lorsque le seuil u tend vers la limite supérieure finie x_0 , alors, il existe une fonction $\sigma(u)$ strictement positive et un réel ξ et on a*

$$\lim_{u \rightarrow x_0} \sup_{0 \leq x < x_0 - u} |F_u(x) - G_{\xi, \delta(u)}(x)| = 0 \quad (2.54)$$

si et seulement si $F \in MDA(H_\xi)$ où $\sigma(\mu)$ est une fonction positive mesurable.

- La preuve de ce théorème se trouve dans Embrechts et al(32). Cette considération théorique suggère que, lorsque nous avons des données issues d'une distribution inconnue, il est possible d'approximer la distribution au-delà d'un certain seuil (assez grand) par une distribution de Pareto généralisée.

Ce théorème montre que la distribution des excès au delà d'un certain seuil peut être approchée par une loi de GPD, autrement dit pour une certaine distribution (cette classe est large, elle regroupe toutes les distributions continues classiques). La loi de GPD est la distribution limite des excès lorsque le seuil tend vers le point final, en générale on prend $x_f = \infty$, ce théorème signifie que pour une grande classe de la fonction de distribution F

quand le seuil augmente la fonction de distribution des excès converge vers la loi de GPD, alors on peut dire que la grande classe de la

fonction de distribution contient toutes les fonctions de distributions continues classiques des statistiques de l'actuariat (log normal, gamma, uniforme,..., etc).

Ce Théorème est très utile lorsqu'on travaille avec des observations qui dépassent un seuil fixé, puisqu'il assure que la loi des excès peut-être approchée par une loi de Pareto généralisée.

2.7 Les outils graphiques pour l'identification de la distribution

Les outils graphiques jouent un rôle important dans l'analyse statistique. La courbe quantile quantile (QQ) et la courbe mean excess (ME) sont parmi les graphiques les plus populaires et utilisées dans l'analyse des données utilisés pour obtenir une indication de la lourdeur des queues.

2.7.1 Quantile Quantile Plot "QQ plot"

C'est une technique graphique statistique apparue dans les années (1960)(voir par exemple Kimball, 1960) utilisée pour tester l'adéquation des modèles des valeurs extrêmes. Le quantile-quantile plot permet de comparer graphiquement les distributions de deux échantillons pour savoir si elles sont identiques c-à-d faire la comparaison entre la fonction de quantile d'un échantillon et la fonction quantile théorique envisagée.

Si l'échantillon provient bien de cette distribution théorique, alors le QQ plot sera linéaire autrement dit, si le modèle paramétrique correspond bien aux données, ce graphique doit avoir une forme linéaire, la linéarité dans un graphique peut être facilement vérifiée par l'observation. Ainsi, le graphique permet de comparer différents modèles estimés et de choisir le meilleur. Le Graphique QQ plot est plus approprié en termes de qualité d'ajustement. De plus, si l'origine de la distribution des données est plus ou moins connu, les graphiques QQ plot peuvent aider à détecter les valeurs aberrantes, voir Embrechts, Kluppelberg et Mikosch (1997).

Definition 42 *Le QQ plot est le nuage des points suivants*

$$\left\{ X_{i,n}, F^{\leftarrow} \left(\frac{i}{n+1} \right), i = 1, \dots, n \right\} \quad (2.55)$$

La même méthodologie peut alors être explorée et prolongée afin de fournir des comparaisons de la preuve empirique disponible dans les données en adaptant des modèles tels que Pareto, le log normal, Weibull ou d'autres. On se limite d'abord sur le modèle exponentielle.

QQ plot d'une distribution exponentielle

Le QQ plot basé sur la distribution exponentielle joue un rôle très important dans la théorie des valeurs extrêmes, il est la représentation des quantiles de la distribution empirique sur l'axe des Y contre les quantiles de la fonction de distribution exponentielle sur l'axe des X .

La distribution exponentielle standard est

$$1 - F_1(x) = \exp(-x), \quad x > 0 \quad (2.56)$$

L'exemple de la classe des distributions avec la fonction de survie générale est

$$1 - F_\lambda(x) = \exp(-\lambda x) \quad (2.57)$$

Nous voulons savoir si la population réelle de distribution F appartient à cette classe, paramétrée par $\lambda > 0$. La réponse doit se baser sur les données x_1, \dots, x_n que nous avons à notre disposition.

La fonction quantile pour la distribution exponentielle a la forme suivante

$$Q_\lambda(p) = -\frac{1}{\lambda} \log(1-p) \text{ pour } p \in [0, 1] \quad (2.58)$$

On a de la même façon

$$Q_1(p) = -\log(1-p) \text{ pour } p \in [0, 1] \quad (2.59)$$

Par conséquent, il existe une relation linéaire entre les quantiles d'une distribution exponentielle et les quantiles d'une distribution exponentielle standard

$$Q_\lambda(p) = \frac{1}{\lambda} Q_1(p) \text{ pour } p \in [0, 1] \quad (2.60)$$

A partir de notre échantillon x_1, \dots, x_n , on estime la fonction quantile Q par le quantile empirique \widehat{Q}_λ et dans un système de coordonnées orthogonales, les points avec des valeurs sont tracés pour différentes valeurs de $p \in [0, 1]$

$$(-\log(1-p), \widehat{Q}_\lambda(p)) \quad (2.61)$$

Donc, nous présumons que le graphe QQ plot aura la forme d'une ligne droite si le modèle exponentiel fournit un ajustement statistique donné. lorsque nous obtenons une ligne droite, la pente de cette droite peut être utilisée comme un estimateur du paramètre λ^{-1} . En effet, si le modèle est correct, alors l'équation

$$Q_\lambda(p) = -\frac{1}{\lambda} \log(1-p) \quad (2.62)$$

est valable.

Généralement

$$\widehat{Q}_\lambda(p) = X_{i,n} \text{ pour } \frac{i-1}{n} < p < \frac{i}{n} \quad (2.63)$$

Un choix très pratique de valeurs $p_{i,n}$ est donné par

$$p_{i,n} \in \left\{ \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right\} \quad (2.64)$$

Le choix alternatif

$$p \in \left\{ \frac{1 - 0.5}{n}, \frac{2 - 0.5}{n}, \dots, \frac{n - 1 - 0.5}{n}, 1 \right\} \quad (2.65)$$

Finalement, les coordonnées sont

$$\{-\log(1 - p_{i,n}), X_{i,n}\} \quad (2.66)$$

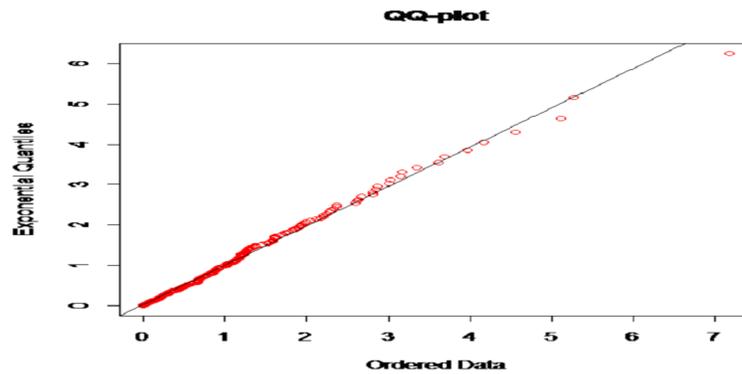


Figure 2.07: QQ plot d'une distribution exponentielle

L'intérêt du graphique 2.07 est de nous permettre d'obtenir la forme de la queue de distribution. Trois cas sont possibles :

-Les données suivent la loi exponentielle : la distribution présente une queue très légère, les points du graphique présente une forme linéaire.

- Les données suivent une distribution à queue lourde : le graphique QQ plot est concave.

- Les données suivent une distribution à queue légère : le graphique QQ plot a une forme convexe.

Exemple 43 *Le quantile $Q(p)$ de la distribution H_ξ est donné par la formule suivante*

$$Q(p) = H_\xi^{-1}(p) = \begin{cases} -\xi^{-1} [1 - (-\log p)^{-\xi}] & \text{si } \xi \neq 0 \\ -\log(-\log p) & \text{si } \xi = 0 \end{cases} \quad (2.67)$$

Ce quantile est donc fortement influencé par le paramètre. Intuitivement, on comprend que plus le paramètre est grand, plus le quantile est élevé.

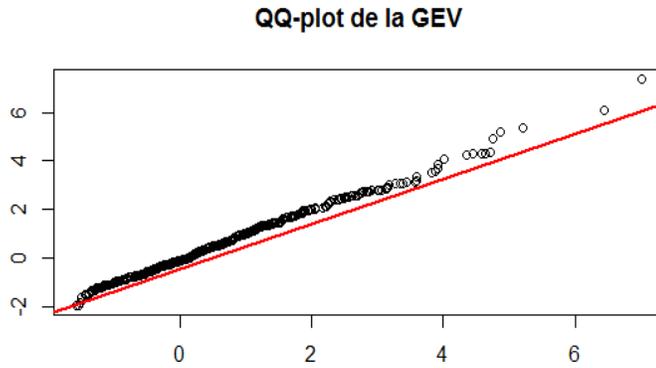


Figure 2.08: QQ plot d'une distribution GEV

Exemple 44 *Le quantile $Q(p)$ de la distribution G_ξ qui est également la VaR de niveau de confiance élevé p , est donné par*

$$Q(p) = VAR(p) = \frac{1}{\xi} \left\{ \left(\frac{n}{N_u} \right)^{-\xi} - 1 \right\} \quad (2.68)$$

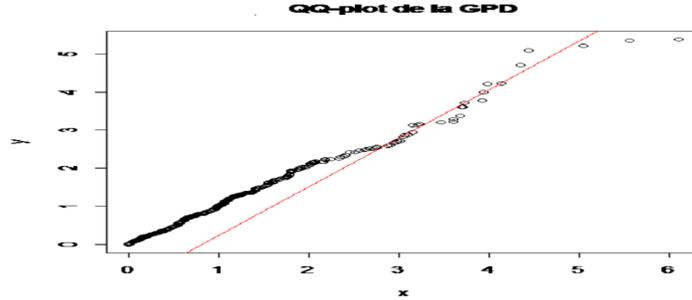


Figure:2.09: QQ Plot de la GPD

QQ plot d'une distribution de Pareto

Nous présentons maintenant le QQ plot d'une distribution de Pareto de paramètre $\frac{1}{\xi}$ qui pourra servir comme un test d'ajustement. Pour le même échantillon de données réelles X_1, \dots, X_n , on veut savoir si ces dernières proviennent d'une distribution de Pareto telle que pour tout $\xi > 0$

$$F_\gamma(x) = 1 - x^{-\frac{1}{\xi}} \quad (2.69)$$

Ainsi pour $\xi = 1$

$$F(x) = 1 - x^{-1} \quad (2.70)$$

La fonction quantile de la distribution de Pareto a la forme simple

$$\log(Q_\xi(p)) = -\xi \log(1 - p), \text{ pour } p \in [0, 1] \quad (2.71)$$

Et de la même manière on a

$$\log(Q_1(p)) = -\log(1 - p), \text{ pour } p \in [0, 1] \quad (2.72)$$

Par conséquent, il existe une relation linéaire simple entre les quantiles d'une distribu-

tion de Pareto et les quantiles d'une distribution de Pareto standard.

$$\log(Q_\xi(p)) = -\xi \log(Q_1(p)), \text{ pour } p \in [0, 1] \quad (2.73)$$

De la même façon qu'on a fait avec la distribution exponentielle on estime Q par \hat{Q}_n . Dans un système de coordonnées orthogonales, les points de valeurs

$$\{-\log(1-p), \log(Q_n(p))\} \quad (2.74)$$

Sont tracés pour différentes valeurs de $p \in [0, 1]$. Nous attendons alors qu'un motif de ligne droite apparait dans le nuage de points, si le modèle de Pareto permet un ajustement statistique plausible pour la population statistique donnée. Quand un motif de ligne droite est obtenu, la pente de la droite de régression peut être utilisée comme une estimation du paramètre $\frac{1}{\gamma}$. En effet, si le modèle est correct, alors l'équation

$$\log(Q_\gamma(p)) = -\gamma(\log(1-p)) \quad (2.75)$$

est valable

Enfinement, les coordonnées sont données par :

$$\{-\log(1-p_{i,n}), \log(X_{i,n})\} \quad (2.76)$$

2.7.2 Mean excess plot

L'estimation des paramètres de la distribution GPD pose le problème de la détermination du seuil, ce dernier ne doit pas être trop bas ni trop haut, car il faut suffisamment de données pour avoir une bonne précision des estimateurs. La technique la plus utilisée pour estimer μ est le graphique mean excess plot note " ME plot", voir Embrechts, Kluppelberg et Mikosch(33).

Le graphique mean excess plot est largement utilisé dans les articles d'applications ou

de simulations, il est à la fois un graphique qui aide à déterminer le nombre de valeur extrême "le choix du seuil" et pour distinguer les modèles à queue lourde et les modèles à queue légère.

Definition 45 Soit une variable aléatoire X , le graphe de la fonction moyenne des excès appelé *mean excess plot* est le graphe des points

$$\{(u, e_n(u)), X_{1,n} < u < X_{n,n}\} \quad (2.77)$$

avec $e_n(u)$ la fonction moyenne des excès théorique au delà du seuil u , définie par une estimation naturelle. Il est l'espérance conditionnelle de $X - u$ sachant que $X > u$

$$e_n(u) = E(X - u | X > u) = \frac{\int_u^{x_0} \bar{F}(s) ds}{\bar{F}(u)}, x < x_0 \quad (2.78)$$

Definition 46 "**Fonction moyenne des excès empiriques**" on définit la fonction moyenne des excès empiriques $\hat{e}_n(u)$ par la somme des excès dépassant un certain seuil élevé noté u , divisé par le nombre de points des données qui dépassent ce seuil .

$$\hat{e}_n(u) = \frac{\sum_{i=1}^n (X_i - u)^+ 1_{(X_i > u)}}{\sum_{i=1}^n 1_{(X_i > u)}} = \frac{1}{N_u} \sum_{i=1}^n (X_i - u)^+ \quad (2.79)$$

avec $u \geq 0$ et

$$1_{(X_i > u)} = \begin{cases} 1 & X_i > u \\ 0 & \text{sinon} \end{cases}$$

Le ME plot est un test graphique pour établir le comportement de la queue, il peut être effectué sur la base de la forme de la distribution de l'excès moyen, soit X_1, \dots, X_n i i d avec F_n la distribution empirique correspondante et $\Delta_n(u) = \{i, i = 1, \dots, n, X_i > u\}$.Alors

$$\hat{e}_n(u) = \frac{1}{\text{card}(\Delta_n(u))} \sum_{i \in \Delta_n(u)} (X_i - u), u \geq 0$$

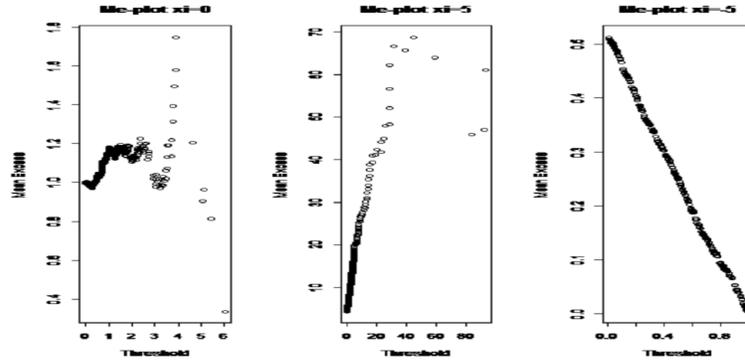
Où card fait référence au nombre de points dans l'ensemble $\Delta_n(u)$.

L'ensemble $\{X_{K,n} \cdot \hat{e}_n(X_{K,n}), k = 1, \dots, n\}$ forme le graphique d'excès moyen.

Les distributions à queue lourde donnent une fonction $e(u)$ qui tend vers l'infini pour un seuil haut u (forme linéaire avec pente positive).

La fonction moyenne des excès empiriques de la GPD s'écrit :

$$\hat{e}_n(u) = \frac{\hat{\delta} + \hat{\xi}u}{1 - \hat{\xi}}, \quad \hat{\delta} + \hat{\xi}u > 0 \quad (2.80)$$



ME plot de la loi GPD

Trois cas peuvent alors se présenter

-Si à un certain seuil, la fonction moyenne des excès empiriques est marquée par une pente positive, alors les données suivent une distribution de Pareto généralisée avec un paramètre positif.

- Si la fonction moyenne des excès empiriques présente une pente horizontale, les données suivent une distribution exponentielle.

-Si la fonction moyenne des excès empiriques est décroissante, on a une distribution bornée à droite.

2.7.3 Mean Residual Life plot (MRL plot)

La moyenne des excès peut être utilisée pour le choix d'un seuil adéquat "optimal" u . la loi de GPD est une bonne approximation de la loi des excès, alors l'espérance conditionnelle de $X - u$ sachant que $X > u$ est

$$E[X - u | X > u] \approx \frac{\delta_u}{1 - \xi} \quad (2.81)$$

2.8 Estimation de l'indice des valeurs extrêmes

La modélisation de la théorie des valeurs extrêmes est liée à l'étude de la queue de distribution. Les lois de cette dernière étant indexées par un paramètre ξ appelé l'indice de valeur extrême, il peut être estimé par plusieurs méthodes. Toutes ces dernières partagent les mêmes propriétés de consistance et de normalité asymptotique. Les méthodes les plus utilisées sont celles du maximum de vraisemblance, de Hill, de Pickands et des moments pondérés.

2.8.1 Estimation par des méthodes paramétriques

Il existe quatre propriétés pour les estimateurs de paramètres: convergent, non-biaisé, efficace et suffisance. Si on trouve qu'un certain estimateur vérifie les deux propriétés "non-biaisé et efficace", on dit qu'il est le meilleur estimateur.

1 Estimateurs convergents

L'estimateur $\hat{\theta}$ est convergent lorsque celui-ci tend en probabilité vers la vraie valeur θ

$$p[|\hat{\theta} - \theta| > \xi] \rightarrow 0, \text{ lorsque } n \rightarrow \infty \text{ et pour tout } \xi > 0 \quad (2.82)$$

2 Estimateurs non-biaisés

L'estimateur non-biaisé est un estimateur où l'espérance mathématique est égale à la vraie valeur du paramètre.

Estimateur de Maximum de vraisemblance

Une fois le seuil fixé, plusieurs approches existent pour l'estimation des paramètres, on peut calculer la vraisemblance (Reiss and Thomas)(68) associée aux dépassements de seuil, car la loi des excès suit asymptotiquement une loi Pareto généralisée. Il est ainsi possible de construire l'estimateur du maximum de vraisemblance pour estimer le paramètre, ce dernier étant le même que lorsque l'on s'intéresse au bloc maxima.).

L'estimation par maximum de vraisemblance donne des résultats asymptotiques efficaces, les estimateurs obtenus convergent sous certaines conditions vers les vrais valeurs des paramètres.

Soient n variables aléatoires $X_1, X_2 \dots X_n$ indépendantes identiquement distribuées (i i d) la fonction de log-vraisemblance est donnée par

Pour $\xi \neq 0$

$$\ln L(\delta, u, \xi) = -n \ln(\delta) - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \left(1 + \xi \frac{x_i - u}{\delta}\right) - \sum_{i=1}^n \left(1 + \xi \frac{x_i - u}{\delta}\right)^{-\frac{1}{\xi}} \quad (2.83)$$

et pour $\xi = 0$

$$\ln L(\delta, u) = -n \ln(\delta) - \sum \frac{x_i - u}{\delta} - \sum \exp\left[\frac{x_i - u}{\delta}\right] \quad (2.84)$$

En prenant la dérivée de $\ln L(\alpha, u)$ par rapport à δ , μ et ξ , on obtient le système d'équations suivant :

Pour $\xi \neq 0$

$$\frac{\partial \log l}{\partial \mu} = \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \frac{\xi}{\delta + \xi (x_i - u)} - \frac{1}{\delta} \sum_{i=1}^n \left(1 + \xi \frac{x_i - u}{\delta}\right)^{-\frac{1}{\xi}-1} = 0 \quad (2.85)$$

$$\frac{\partial \log l}{\partial \delta} = -\frac{n}{\delta} + \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \frac{\xi (x_i - u)}{\delta^2 + \xi \delta (x_i - u)} - \sum_{i=1}^n \left(\frac{(x_i - u)}{\delta^2}\right) \left(1 + \xi \frac{x_i - u}{\delta}\right)^{-\frac{1}{\xi}-1} = 0 \quad (2.86)$$

$$\begin{aligned} \frac{\partial \log l}{\partial \xi} &= -(1 - \frac{1}{\xi^2}) \sum_{i=1}^n \log \left(1 + \xi \frac{x_i - u}{\delta} \right) - (1 - \frac{1}{\xi}) \sum_{i=1}^n \frac{\xi (x_i - u)}{\delta + \xi \delta (x_i - u)} + \\ &\sum_{i=1}^n \left(1 + \xi \frac{x_i - u}{\delta} \right)^{\frac{-1}{\xi}} \left\{ \sum_{i=1}^n \frac{(x_i - u)}{\delta + \xi^2 \delta (x_i - u)} - \frac{1}{\xi^2} \log \left(1 + \xi \frac{x_i - u}{\delta} \right) \right\} = 0 \end{aligned} \quad (2.87)$$

et pour $\xi = 0$

$$\frac{\partial \log l}{\partial \delta} = 0 = n + \sum \frac{x_i - u}{\delta} [\sum \exp[\frac{x_i - u}{\delta}] - 1] \quad (2.88)$$

$$\frac{\partial \log l}{\partial \mu} = 0 = n - \sum \exp(\frac{x_i - u}{\delta}) \quad (2.89)$$

La résolution de ce système est relativement difficile et n'admet pas généralement de solutions explicites, dans ce cas on fait appel à des méthodes d'optimisation numérique.

Méthode des moments pondérés

Soit Y_1, Y_2, \dots, Y_n un échantillon de fonctions de répartition connue H_θ où $\theta = (\xi, \delta, \mu)$. Les moments pondérés d'ordre r , notés PWM en anglais (Probability Weighted Moments (Hosking and Wallis)[23] sont définis par

$$\mu_r = E(Y H_\theta^r(Y)), \quad r \in N \quad (2.90)$$

Les moments empiriques pondérés d'ordre r sont :

$$w_r = \frac{1}{n} \sum Y_{i,n} H_{\xi, \mu, \delta}^r(y_{i,n}), \quad r = 0, 1, 2 \quad (2.91)$$

En utilisant la distribution des extrêmes généralisés H_θ , on obtient

$$\mu_r = \frac{1}{r+1} \left[\mu - \frac{\delta}{\xi} (1 - (r+1)^\xi \Gamma(1 - \xi)) \right], \quad \xi < 1 \quad (2.92)$$

Où Γ est définie par

$$\Gamma(t) = \int_0^\infty x^{t-1} \exp(-x) dx \quad (2.93)$$

L'estimateur des moments pondérés $\theta = (\xi, \mu, \delta)$ est la solution des systèmes d'équations suivant :

$$\mu_0 = \mu - \frac{\delta}{\xi}(1 - \Gamma(1 - \xi)) \quad (2.94)$$

$$2\mu_1 - \mu_0 = \frac{\delta}{\xi}(1 - 2\xi)\Gamma(1 - \xi) \quad (2.95)$$

$$\frac{3\mu_2 - \mu_0}{3\mu_1 - \mu_0} = \frac{1 - 3\xi}{1 - 2\xi} \quad (2.96)$$

En inversant ces formules, on obtient (ξ, μ, δ) en fonction de (μ_0, μ_1, μ_2) , il reste à estimer ces trois moments.

On remplace l'espérance par une moyenne empirique.

$$\mu_r \simeq \frac{1}{n} \sum_{i=1}^n Y_i H_{\xi, \mu, \delta}^r(Y_i) = \frac{1}{n} \sum_{k=1}^n Y_{i,k} H_{\xi, \mu, \delta}^r(Y_{i,k}) \quad (2.97)$$

En ordonnant les observations, on remplace $H_{\xi, \mu, \delta}$ par la fonction de répartition empirique

$$\mu_r \simeq \frac{1}{n} \sum_{i=1}^n Y_i F_{\xi, \mu, \delta}^r(Y_i) = \frac{1}{n} \sum_{i=1}^n Y_{i,k} \left(\frac{i-1}{k}\right)^r \quad (2.98)$$

On obtient alors un estimateur sous forme d'une combinaison linéaire

$$\mu_r = \frac{1}{n} \sum_{i=1}^n Y_{i,k} \left(\frac{i-1}{n}\right)^r \quad (2.99)$$

2.8.2 Estimation par des méthodes semi-paramétriques

Estimateur de Hill

Cet estimateur a été introduit par Hill (1975) de forme semi paramétrique, il est basé sur les statistiques d'ordre qui n'est utilisable que pour les distributions appartenant au domaine de Fréchet, l'estimateur de Hill est probablement l'estimateur le plus étudié dans la

littérature. Un grand nombre de travaux a été consacré à l'étude de la normalité asymptotique de l'estimateur de Hill, Davis et Resnick (1984) et Haeusler et Teugels (1985) et pour les propriétés de l'estimateur, Mason(1982), Deheuvels et al(1988).

Soit $(k_n)_{n \geq 1}$ une suite d'entiers avec $1 < k_n < n$, l'estimateur de Hill est donné par la forme empirique suivante

$$\bar{\xi}(k) = \frac{1}{k-1} \sum_{i=1}^{K-1} \log X_{n-i+1,n} - \log X_{n-k+1,n} \quad (2.100)$$

On peut justifier la démonstration par la caractérisation

$$\bar{F}(x) = x^{-\frac{1}{\xi}} l(x) \quad (2.101)$$

avec l une fonction à variation lente.

Nous avons deux parties : $x^{-\frac{1}{\xi}}$ la partie paramétrique et $l(x)$ la partie non paramétrique.

Pour $t > 0$

$$\lim_{u \rightarrow \infty} \frac{l(tu)}{l(u)} = 1 \quad (2.102)$$

Avec $t > 0$

$$\lim_{u \rightarrow \infty} \frac{\bar{F}(tu)}{\bar{F}(u)} = t^{-\frac{1}{\xi}} \left(\lim_{u \rightarrow \infty} \frac{l(tu)}{l(u)} \right) = t^{-\frac{1}{\xi}} \quad (2.103)$$

On déduit l'approximation

$$\bar{F}(tu) \simeq t^{-\frac{1}{\xi}} \bar{F}(u) \quad (2.104)$$

En posant $x = t$ et $u\alpha = \bar{F}(u)$, on a :

$$F(x) = \alpha \left(\frac{x}{\bar{F}^{-1}(\alpha)} \right)^{-\frac{1}{\xi}} \quad (2.105)$$

si on prend $p = \bar{F}(u)$, on en déduit

$$\left(\frac{p}{\alpha} \right)^{-\xi} = \frac{p^{-1}(p)}{p^{-1}(\alpha)} \quad (2.106)$$

Pour $x > u$ ou de façon équivalente $p < \alpha$, on a :

$$F^{-1}(P) = F^{-1}(\alpha)\left(\frac{P}{\alpha}\right)^{-\xi} \quad (2.107)$$

Ces approximations sont des cas particuliers de l'approche GPD avec $\delta = \xi F^{-1}(\alpha)$

Il reste à estimer ξ , en se basant encore sur

$$F^{-1}(p) = F^{-1}(\alpha)\left(\frac{p}{\alpha}\right)^{-\xi} \quad (2.108)$$

On peut réécrire

$$\log \bar{F}^{-1}(p) - \log \bar{F}^{-1}(\alpha) = \xi \log\left(\frac{p}{\alpha}\right) \quad (2.109)$$

On choisit $\alpha = \frac{k}{n}$ et on considère plusieurs valeur de $p = \frac{i}{n}, i = 1, \dots, k-1$ (on doit avoir $p < \alpha$), On obtient

$$\log \bar{F}^{-1}\left(\frac{i}{n}\right) - \log \bar{F}^{-1}\left(\frac{k}{n}\right) = \xi \log\left(\frac{k}{i}\right) \quad (2.110)$$

Et en estimant les fonctions de survie par leurs équivalentes empiriques

$$\log X_{n-i+1,n} - \log X_{n-k+1,n} = \xi \log\left(\frac{k}{i}\right) \quad (2.111)$$

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

$$\xi = \frac{\sum_{i=1}^{k-1} (\log X_{n-i+1,n} - \log X_{n-k+1,n})}{\sum_{i=1}^{K-1} \log\left(\frac{k}{i}\right)} \quad (2.112)$$

Il est équivalent à $k-1$ au voisinage de l'infini, on obtient alors l'estimateur de Hill.

Propriétés de l'estimateur de Hill

1 : Convergence en probabilité

Theorem 47 Soit X_1, \dots, X_n un échantillon de variables aléatoires i.i.d de fonction de réparation $F \in MDA(H_\xi), \xi > 0$. Soit $(K(n))_{n>1}$ une suite d'entiers telle que $1 < K(n) < n$, si $\frac{K(n)}{n} \rightarrow 0$ et $K(n) \rightarrow \infty$, alors

$$\hat{\xi}_{((k(n)),n)}^H \xrightarrow{p} \xi, n \rightarrow \infty \quad (2.113)$$

2 : Convergence forte (convergence presque sûre) Si $\frac{k(n)}{n} \rightarrow 0$ et $\frac{k(n)}{\log \log(n)} \rightarrow \infty$ alors

$$\hat{\xi}_{((k(n)),n)}^H \xrightarrow{ps} \xi, n \rightarrow \infty \quad (2.114)$$

3 : Normalité asymptotique

Theorem 48 Soit $(K(n))_{n>1}$ une suite d'entiers telle que $1 < K(n) < n$, si $\frac{K(n)}{n} \rightarrow 0$ et $K(n) \rightarrow \infty$, Si la condition de l'ordre est satisfaite avec $K(n)A(\frac{n}{k(n)}) \rightarrow 0$ quand $n \rightarrow \infty$ alors

$$K(n)^{\frac{1}{2}}(\hat{\xi}_{(k(n),n)}^H - \xi) \xrightarrow{L} N(0, \xi^2) \quad (2.115)$$

ces propriétés nous permettent de calculer l'intervalle de confiance pour $\hat{\xi}_{(k(n),n)}^H$ qui donné par

$$[\xi(1 \mp \frac{\sigma}{k^{\frac{1}{2}}})] \quad (2.116)$$

Estimateur de Pickands

L'estimateur de Pickands est construit en utilisant les statistiques d'ordres, il 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. Cet estimateur a été introduit par (Pickands et al (1975)) et revisité (Drees (1995), Drees and Kaufmann (1998)) d'une part, (Dekkers et al (1989)) d'autre part, ils ont établi sa consistance faible ainsi que sa normalité asymptotique.

Definition 49 On suppose que $\{X_i, i = 1, \dots, n\}$ 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 et $1 < k_n < n$ un estimateur de Pickands est défini par :

$$\xi_{k,n}^P = \frac{1}{\ln 2} \ln \left(\frac{X_{n-k+1,n} - X_{n-2k+1,n}}{X_{n-2k+1,n} - X_{n-4k+1,n}} \right) \quad (2.117)$$

Soit $U(\cdot)$ la fonction définie par

$$U(\cdot) = \left(\frac{1}{1 - F(x)} \right)^{\leftarrow} \quad (2.118)$$

Où \leftarrow représente l'inverse généralisée

Soit $\varphi_t(\cdot)$ la fonction définie par

$$\varphi_t(x) = \int_1^{\infty} u^{t-1} du, \quad x > 0, t \in \mathbb{R} \quad (2.119)$$

De Haan (1984) montre que la relation est vérifiée si et seulement si pour tout $x > 0$ et $y = 1$

$$\lim \frac{U(tx) - U(x)}{U(ty) - U(y)} = \frac{\varphi(x)}{\varphi(y)} \quad (2.120)$$

En remplaçant dans cette relation

$$U \text{ par } \widehat{U} = \left[\frac{1}{1 - \widehat{F}_n} \right]^{\leftarrow}$$

$$t \text{ par } \frac{n}{2k},$$

$$y \text{ par } 2,$$

x par $\frac{1}{2}$, on a pour n suffisamment grand.

$$2^{-\xi} \frac{X_{n-k+1,n} - X_{n-2k+1,n}}{X_{n-2k+1,n} - X_{n-4k+1,n}} = 1 \quad (2.121)$$

L'estimateur de pickands est la solution de l'équation précédente.

Propriétés de l'estimateur de Pickands

1 : Convergence en probabilité

Theorem 50 Soit X_1, \dots, X_n un échantillon de variables aléatoires i.i.d de fonction de répartition $F \in MDA(H_\xi)$, si $n \rightarrow \infty$, $K(n) \rightarrow \infty$ et $\frac{K(n)}{n} \rightarrow \infty$ alors

$$\xi_{((k(n)),n)}^P \rightarrow \xi \quad (2.122)$$

2 : Convergence forte (convergence presque sûre)

Theorem 51 Si $\frac{k}{n} \rightarrow 0$ et $\frac{k(n)}{\log \log(n)} \rightarrow \infty$ $n \rightarrow \infty$, alors

$$\xi_{((k(n)),n)}^P \rightarrow \xi, n \rightarrow \infty \quad (2.123)$$

3 : Normalité asymptotique

Theorem 52 Sous les conditions additionnelles sur la suite $K(n)$ et la fonction de répartition F on a

$$K(\xi_{(k,n)}^P) - \xi \rightarrow N(0, v(\xi)) \quad (2.124)$$

avec $K(n)A(\frac{n}{k(n)}) \rightarrow 0$ quand $n \rightarrow \infty$ où

$$v(\xi) = \frac{\xi^2(2^{2\xi+1} + 1)}{(2(2^\xi - 1) \log 2)^2} \quad (2.125)$$

Pour plus de détails (De Haan and Ferreira (2007)) et (Embrechts et al. (2013))

Estimateur des moments

Cet estimateur est proposé par Dekkers et al (30), il présente une extension de l'estimateur de Hill, il est valable quel que soit le signe de ξ . Il est défini par la statistique.

$$\hat{\xi}_{X,k,n}^{(M)} = \hat{\xi}_{X,k,n}^H + 1 - \frac{1}{2} \left(1 - \frac{\hat{\xi}_{X,k,n}^H}{S_{X,k,n}}\right)^{-1} \quad (2.126)$$

où

$$S_{X,k,n} = \frac{1}{k} \sum_{j=1}^k (\log X_{n-j+1,n} - \log X_{n-k,n})^2 \quad (2.127)$$

2.8.3 Le choix du nombre k

Les résultats concernant les estimateurs de l'indice des valeurs extrêmes sont asymptotiques : ils sont obtenus lorsque $k \rightarrow \infty$ et $\frac{k}{n} \rightarrow 0$, le choix du nombre k de valeurs extrêmes n'est pas facile à gérer car il doit être intermédiaire, il ne doit être ni trop grand, ni trop petit car : si k est petit on ne garde pas beaucoup d'observations et on a une grande variance et l'utilisation de k introduit un grand biais, si k est grand la variance dans ce cas est petite. L'équilibrage de ces composants (la variance et le biais) est important dans les applications de la théorie des valeurs extrêmes, parce que cela réduit l'erreur moyenne quadratique MSE. Il existe certaines méthodes pour choisir ce nombre optimal de statistiques d'ordre par exemple k_{opt} , on a la méthode de Hall et Welsh (1985) et la méthode de Bootstrap Hall (1990) pour plus de détails sur le choix de k Balkema and De Haan (1974), Dekkers et de Haan (1993), De Haan and Peng (1998) Drees et Kaufmann (1998), Danielsson et al. (2001), Cheng et Peng (2001), Neves et Fraga Alves (2004) et Beirlant et al. (2006).

2.9 Mesure de risque

Le risque de marché correspond à un risque de perte des portefeuilles des investisseurs du aux variations des marchés financiers, les marchés des instruments de base (actions, obligations, devises, matières premières) et les marchés des produits dérivés (contrats à terme, options). Pour gérer le risque de marché, il faut donc mesurer de manière précise ce risque pour mieux l'appréhender, puis développer des outils de réduction et de couverture des risques, on a besoin d'une mesure plus synthétique du risque encouru par son portefeuille d'investissement.

Nous pouvons faire ressortir deux mesures de risque à partir de la distribution des rentabilités des actifs, la volatilité et la valeur de risque (Value-at-Risk, VaR).

La **volatilité** mesure la dispersion moyenne des rentabilités autour de leur moyenne, elle mesure les variations d'un actif autour de la tendance centrale. Elle prend donc en compte toutes les rentabilités, positives ou non, extrêmes ou non, la volatilité contribue à calculer les gains espérés et les pertes potentielles et ne donne pas de l'importance aux pertes seulement, elle peut être exprimée en fonction des rentabilités comme suit

$$\delta = E((R - E(R))^2) \quad (2.128)$$

La **Value-at-Risk** est un quantile de perte qui détermine la plus petite perte subie par un actif avec une probabilité d'occurrence faible sur un horizon donné, elle mesure ainsi le risque extrême. Elle est une méthode qui permet de matérialiser le risque de perte d'une position ou de l'ensemble d'un portefeuille, en un seul montant.

Nous venons d'introduire le fait que la volatilité n'est certainement pas la meilleure mesure de risque extrême car à partir de cette dernière on peut enregistrer un taux élevé et capturer des risques moyens, certes significatifs, mais pas extrêmes. C'est pour cela que la Value-at-Risk est utilisée par la plupart de managers.

Definition 53 (*Mesure de risque*)

On appelle mesure de risque toute application ρ associant à un risque X un réel

$$\rho(X) \in \mathbb{R}_+ \quad (2.129)$$

Definition 54 "*Chargement de sécurité*"

Une mesure de risque ρ contient un chargement de sécurité si pour tout risque X , on a

$$\rho(X) \geq E(X) \quad (2.130)$$

Remarque D'après la définition, on peut considérer l'espérance, la variance et l'écart-type comme des mesures de risque, d'une façon générale toute fonction réelle positive d'une v.a peut être considérée comme étant une mesure de risque.

2.9.1 Mesure de risque cohérente

Une mesure de risque est une fonction ρ faisant correspondre à un risque X un nombre positive noté $\rho(X)$ éventuellement infini, elle possède une partie des caractéristiques suivantes :

Monotonie

Soient X et Y deux risques tels que $\rho(X \leq Y) = 1$. On dit qu'une mesure de risque est monotone si

$$X \leq Y \implies \rho(X) \leq \rho(Y) \quad (2.131)$$

Si la perte de l'investisseur X est plus petite que celle de l'investisseur Y , alors le risque de l'investisseur X est également plus petit.

Sous-additivité

Soient X et Y , deux risques. On dit qu'une mesure de risque est sous-additive si

$$\rho(X + Y) \leq \rho(x) + \rho(y) \quad (2.132)$$

Le risque d'un portefeuille comprenant des investissements en X et Y est aussi plus grand que la somme des risques individuels.

Homogénéité Soit X un risque et λ un scalaire positif. On dit qu'une mesure de risque est homogène si

$$\rho(\lambda X) = \lambda \rho(X), \text{ pour } \lambda > 0 \quad (2.133)$$

Lorsque l'investissement est multiplié par le nombre λ par conséquent le risque est également multiplié par cette même valeur.

Invariance par translation Soit X un risque, on dit qu'une mesure de risque est invariante à la translation pour toute constante α si :

$$\rho(X + \alpha) = \alpha + \rho(X) \quad (2.134)$$

Lorsqu'on ajoute des pertes connues α au perte de l'investissement X , le risque change par une valeur égale à α

2.10 Value at Risk

La notion de Value-at-Risk (VaR) est apparue pour la première fois dans le secteur de l'assurance à la fin des années 1980, la banque Bankers Trust l'une des premières institutions à utiliser cette notion sur les marchés financiers aux Etats-Unis, mais c'est principalement la banque JP Morgan qui dans les années 90 a popularisé ce concept notamment grâce à son système RiskMetrics.

D'après A. Louis Calvet (en 2000), la VaR d'un portefeuille d'actifs financiers correspond au montant des pertes maximum sur un horizon de temps donné, si l'on exclut un ensemble d'évènements défavorables ayant une faible probabilité de se produire, Berdin et Hyde (2001) apportent une définition plus exhaustive au concept de VaR. En effet, ils définissent la VaR comme étant la mesure qui fournit une estimation de la perte potentielle sur un actif ou un portefeuille qui peut survenir avec une probabilité donnée suite à des mouvements de prix ou de taux relativement adverses, sous l'hypothèse que pendant une période de temps. (l'horizon de la VaR) la composition du portefeuille reste inchangée. Selon Jorion 2002, la VAR d'un portefeuille ou d'un actif pour une durée T et un niveau de probabilité α , se définit comme le montant de perte attendu de façon que ce montant sur la période $[0, T]$, ne devrait pas être plus important que la VAR et ceci avec une probabilité

de $1 - \alpha$. Selon Hervé 2004, la VAR est la perte maximale observable sur un portefeuille avec une probabilité sur un horizon de temps donnée. Donc de façon générale la VaR d'un actif financier ou d'un portefeuille d'actifs représente la perte potentielle qui ne devrait pas être dépassée avec une probabilité donnée sur un horizon de temps donné $T \in N$ par un certain niveau de confiance $\alpha \in (0, 1)$. Elle permet de répondre à la question suivante.

Combien l'établissement financier peut-il perdre avec une probabilité α pour un horizon de temps T fixé ?

La VAR n'est rien d'autre qu'un quantile du risque considéré comme une fonction perte, ainsi la VaR dépend de trois éléments: (i) **la distribution des pertes et profits** du portefeuille valable pour la période de détention. (ii) **le niveau de confiance** (ou de façon équivalente le taux de couverture égal à au moins le niveau de confiance) et (iii) **la période de détention de l'actif**.

la période de détention de l'actif.

L'horizon temporel ou holding period qui correspond à la période sur laquelle la variation de valeur du portefeuille est mesuré, ce paramètre est très important car plus l'horizon est long plus les pertes peuvent être importantes. Il n'existe aucune règle quant au choix de la période de détention dans le calcul de la VaR puisque ce choix dépend fondamentalement de l'horizon de reporting ou d'investissement des opérateurs. Toutefois les autorités de régulation peuvent spécifier des horizons de détention spécifiques notamment dans le cadre des procédures de validation de la VaR. L'horizon correspond au temps pendant lequel le portefeuille va subir les fluctuations du marché donnant lieu à des pertes ou des profits.

Taux de couverture et Niveau de Confiance

Le niveau de confiance est un paramètre compris entre 0 et 1 s'élevant à 95% ou 99% en général qui permet de contrôler la probabilité que l'on obtienne un rendement supérieur ou égal à la VaR. Le seuil de confiance représente le pourcentage de chances pour que le montant de la perte dépasse la valeur définie par la VaR.

Nous donnons ci-dessous sa définition mathématique

La VaR représente une perte potentielle dont il faut définir la probabilité d'occurrence ainsi que l'horizon h de durée d'investissement.

Soit V_t^p la valeur en t du portefeuille et V_{t+h}^p sa valeur à l'horizon h , La rentabilité du portefeuille $R_{t,t+h}^p$ investi à la date t et déteu sur un horizon de durée h a pour expression, avec $R_{t,t+h}^p$ le vecteur aléatoire des rentabilités des n actifs.

$$R_{t,t+h}^p = \ln \frac{V_{t+h}^p}{V_t^p} \approx \frac{V_{t+h}^p}{V_t^p} - 1 \quad (2.135)$$

Supposons que $R_{t,t+h}^p$ est issu d'une variable aléatoire positive ou négative notée R^P . alors $\text{VAR}(h, \alpha)$ est définie par

$$P[R^P \leq \text{VAR}(h, \alpha)] = 1 - \alpha \quad (2.136)$$

C'est-à-dire la probabilité que la rentabilité du portefeuille soit inférieure à la VaR sur un horizon h est inférieur ou égal à $1 - \alpha$.

Les praticiens parlent aussi du niveau de confiance $(1 - \alpha)$ qui correspond au quantile de probabilité α . Ainsi, une VaR à 99% par exemple correspond au quantile de probabilité $\alpha = 1 - 99\% = 1\%$.

Definition 55 Notons F la fonction de répartition de la variable aléatoire R^P des rentabilités du portefeuil de périodicité h . La VaR de probabilité d'occurrence α et d'horizon d'investissement h est alors le quantile défini par

$$\text{VAR}(h, \alpha) = F^{-1}(1 - \alpha) \quad (2.137)$$

Pour déterminer explicitement la VaR, il nous reste à caractériser la forme de F . La première solution consiste à estimer F par sa distribution empirique, nous avons alors la VaR historique, une seconde solution est obtenue en utilisant un modèle paramétrique, les praticiens utilisent classiquement la VaR Gaussienne.

Les Value-at-Risk ont un certain nombre de "bonnes" propriétés mathématiques parmi lesquelles le fait que pour toute fonction g croissante et continue à gauche, on a

$$VAR(g(X)) = g(VAR(X)) \quad (2.138)$$

Il découle de cette propriété en prenant $g = F_1^{-1} + F_2^{-1}$ que les VAR sont comonotones additives puisque pour tout $\alpha \in]0, 1[$, on a

$$VAR_\alpha((F_1^{-1} + F_2^{-1})(X)) = (F_1^{-1} + F_2^{-1})(VAR_\alpha(X)) \quad (2.139)$$

La fonction de répartition de la variable aléatoire X est l'application

$$\begin{aligned} F & : \quad \mathbb{R} \rightarrow [0, 1] \\ x & \longmapsto F(x) = P(X < x) \end{aligned} \quad (2.140)$$

2.10.1 Méthode historique

Comme également son nom l'indique, on utilise les gains et les pertes passés de notre portefeuille d'instruments financiers. La méthode historique est une méthode non-paramétrique, la plus intuitive dans sa conception et sa mise œuvre et très simple à utiliser pour les estimations des mesures de risque, puisqu'elle ne pose aucune hypothèse sur la forme de la distribution des rendements du portefeuille (la loi de distribution), elle nécessite l'hypothèse de stationnarité de ces rendements observés à partir des données historiques pouvant se reproduire à l'avenir, car elle utilise uniquement la distribution empirique.

Cette approche utilise des données de marché plus ou moins récentes pour déterminer la distribution des variations des facteurs de risque qui peuvent simplement être les prix des actifs dans le cas des actions. La méthode historique reste toujours la plus couramment utilisée car c'est la méthode la plus facile à mettre en place malgré la nécessité de maintenir des historiques nombreux et longs selon les nombres d'actifs pris en compte. Le principe

général de cette approche consiste à estimer la distribution des variations des facteurs de risque par la distribution observée à partir des historiques.

Pour la VaR historique, nous estimons F dans l'équation $VAR(h, \alpha) = F^{-1}(1 - \alpha)$ par sa distribution empirique \hat{F} par une méthode non-paramétrique. Supposons que nous disposons d'un historique de N observations des rentabilités, la distribution empirique \hat{F} des rentabilités du portefeuille est alors caractérisée à partir du vecteur $(\hat{R}_{t|t+h}, \dots, \hat{R}_{N|N+h})'$ et $\hat{F}^{-1}(\alpha)$ correspond au quantile de probabilité α de $(\hat{R}_{t|t+h}, \dots, \hat{R}_{N|N+h})'$. L'avantage de cette méthode est donc de ne pas imposer d'hypothèse sur la loi de distribution des rentabilités. Mais le problème qui se pose est la longueur N de l'historique qui, si elle est trop faible, ne fournit pas un calcul précis de la VaR. Nous construisons pour cela une version aléatoire de même taille du vecteur $(\hat{R}_{t|t+h}, \dots, \hat{R}_{N|N+h})'$ dans laquelle les rentabilités du portefeuille peuvent apparaître pour certaines aucune fois, pour d'autres une fois, pour d'autres encore deux fois, etc.

Remark 56 *La méthode n'utilise pas d'hypothèse distributionnelle sur la variable aléatoire rendement. La procédure est en effet simple et fournit des résultats faciles à interpréter. Aucune relation linéaire des relations entre les prix et les facteurs de risque.*

2.10.2 Méthode paramétrique

La méthode paramétrique la plus reconnue provient de la banque JP Morgan qui développa ce modèle en 1994 et aussi sous le nom de méthode delta-normale par Jorion en 2001, elle s'est répandue à travers le monde sous l'impulsion de sa filiale Riskmetrics, elle est appelée également méthode analytique. Cette méthode détermine la valeur at risk d'un portefeuille d'actifs à partir d'une loi de distribution paramétrique explicite des gains/pertes, où la distribution est fixée dans ce modèle. la méthode analytique suppose que les variations des facteurs de risque suivent une loi normale et utilise les propriétés statistiques de ce type de loi pour estimer la VAR, la méthode présuppose la normalité de sorte que la moyenne μ et l'écart type δ des variations suffisent à leur description, l'objectif de la méthode analytique est d'estimer la variance du portefeuille à partir de variances et de corrélations standardisées

pour certaines échéances et certains produits, elle est basée sur les hypothèses suivantes:

1-La stationnarité des rendements des actifs, (Les distributions des rendements des actifs qui composent le portefeuille suivent une loi normale).

2-L'hypothèse de normalité des rendements.

3-Une décomposition linéaire des prix des actifs en facteurs de risques (La relation entre les variations de valeur du portefeuille et les variations des variables du marché sont linéaires), et Les produits dérivés (Futures, Swaps, ...) sont linéaires sauf les options.

Cette méthode se déroule en cinq étapes distinctes

1-On détermine un modèle d'évolution des gains/pertes du portefeuille en fonction des divers facteurs de risque (qui peuvent être les rendements des actifs financiers constitutifs du portefeuille),

2-On fait le choix de lois paramétriques pertinentes qui correspondent aux qualités distributionnelles des facteurs de risque,

3-On estime les paramètres de ces lois à partir des données passées suivant les méthodes statistiques classiques (méthode des moments ou du maximum de vraisemblance),

4-On détermine la loi de distribution des pertes et profits à partir du modèle choisi,

5-On calcule le quantile associé au seuil de confiance, la loi de probabilité la plus généralement utilisée est la loi normale.

Pour calculer la VAR par la méthode paramétrique elle s'exprime comme suit: Le choix du modèle paramétrique se porte sur la loi normale, supposons donc que les rentabilités du portefeuille R_{t+h}^p soient issues d'une variable aléatoire gaussienne, ainsi F est définie $N(\mu, \delta^2)$, nous obtenons

$$F\left(\frac{VAR_P - \mu}{\delta}\right) = 1 - \alpha \Leftrightarrow VAR_P = \mu + \delta F^{-1}(1 - \alpha) \quad (2.141)$$

les paramètres μ et δ sont données par leurs estimateurs du maximum de vraisemblance, et F^{-1} est la fonction quantile de la distribution normale centrée réduite au niveau de probabilité $1 - \alpha$.

Dans le calcul valeur de risque, l'hypothèse de normalité est à bannir car seules les valeurs extrêmes sont prises en considération et on sait que leurs distributions sont de type GEV ou GPD. Mesurer le risque revient à calculer la VaR donc un quantile extrême.

2.10.3 Estimation de quantiles extrêmes

La mesure et la gestion des risques exigent qu'on s'intéresse aux queues des distributions et notamment aux pertes que l'on peut subir avec des probabilités faibles mais non négligeables. Dans l'analyse des extrêmes, on s'intéresse souvent à l'estimation des quantiles extrêmes qui jouent un rôle important dans le contexte de la gestion des risques, où il est crucial d'évaluer de manière adéquate le risque d'une grande perte, de sorte que les probabilités de les dépasser soient très faibles c'est-à-dire proches de zéro. Nous nous sommes donc intéressés à l'estimation des quantiles extrêmes $Var(\alpha_n)$. Plus généralement si la taille de l'échantillon tend vers l'infini, nous avons

$$P(X_{n,n} < Var(\alpha_n)) = F^n(Var(\alpha_n)) = (1 - \alpha_n)^n$$

$$\exp(n \log(1 - \alpha_n)) = \exp(-n\alpha_n(1 + o(1))) \text{ quand } \alpha_n \rightarrow 0 \quad (2.142)$$

Plusieurs méthodes d'estimation d'un quantile extrême ont été proposées dans la littérature nous allons donner quelques méthodes d'estimation de quantiles extrêmes basées sur des estimateurs de l'indice de queue.

2.10.4 Approche basée sur l'estimateur du maximum de vraisemblance

L'idée générale de la théorie des valeurs extrêmes est de restreindre l'échantillon, sur lequel reposeront les estimations de la distribution de perte, à ses composantes assez grandes. On diminue ainsi tout ou une partie du centre de la distribution. La famille de lois qui ajuste le mieux la queue de distribution de perte est donnée par le théorème de Pickands-balkema-de Haan (voir par exemple Embrechts [15](1997), ou Beirlant [5] (2006)) qui donne, pour la

distribution conditionnelle, une loi limite appartenant à la classe de Distribution de Pareto Généralisée (GPD). C'est une méthode qui est basée sur l'approximation de la distribution des excès par la loi de Pareto généralisée. Soit (X_1, \dots, X_n) un échantillon, on choisit un certain seuil élevé u . Notons N_u le nombre d'observation qui dépassent le seuil u , on définit les v.a des excès $Y_i = X_i - u$, Le théorème de Belkema-De Haan-Pickands nous donne l'approximation de cette distribution $F_u(x) \approx G_{\xi, \beta}(x)$ encore $\bar{F}_u(x) \approx \bar{G}_{\xi, \beta}(x)$ lorsque u tend vers ∞ . On ajuste une loi GPD aux excès Y_1, \dots, Y_{N_u} pour obtenir les estimateurs du maximum de vraisemblance $\hat{\xi}$ et $\hat{\delta}$, on peut montre que La loi F peut se décomposer $\bar{F}(x) = \bar{F}(u)F(x - u)$ pour tout $x > u$ à partir de la définition des excès, nous pouvons estimer la queue $\bar{F}(x) = 1 - F(x)$ par

$$\bar{F}(x) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x - u}{\hat{\delta}}\right)^{-\frac{1}{\hat{\xi}}}, x \geq u \quad (2.143)$$

Finalement, on inverse la fonction définée et l'estimateur du quantile extrême est donnée par :

$$Var(\alpha_n) = u + \frac{\hat{\delta}}{\hat{\xi}} \left(\left(\frac{n}{N_u} \alpha_n \right)^{-\hat{\xi}u} - 1 \right) \quad (2.144)$$

Cette méthode présente un avantage, en ce sens qu'il est plus facile d'avoir un échantillon d'excès que des maximums. Dans la pratique, on remplace u par $(k_n + 1)$ qui représente la k plus grande observation de l'échantillon (X_1, \dots, X_n) , c'est-a-dire $X_{n-kn, n}$, qui correspond a la statistique d'ordre des observations situées au niveau de la queue de la distribution.

2.10.5 Approche basée sur l'estimateur des moments

la méthode des moments a été introduit par Dekkers, Einmahl et de Haan (1989). La normalité asymptotique de cet estimateur de quantile sous différentes conditions sur la queue de la distribution est prouvée par Dekkers, Einmahl et de Haan (1989), et par de Haan et Rootzén (1993) sous diverses hypothèses, portant d'une part sur la queue et d'autre

part sur α_n . L'estimateur des moments est donnée par

$$VaR(\alpha_n) = X_{n-k,n} + \hat{a} \left(\frac{k_n \left(\frac{n}{k_n} \alpha_n \right)^{-\hat{\xi}_{k_n}^M} - 1}{\hat{\xi}_{k_n}^M} \right) \quad (2.145)$$

$$\text{où } \hat{a}_{\frac{k_n}{n}} = X_{n-k,n} \hat{\xi}_{k_n}^M \max(1 - \hat{\xi}_{k_n}^M)$$

Conclusion

On a rappelé les fondamentaux sur la statistique des valeurs extrêmes, la modélisation par la méthode BM avec la distribution généralisée des extrêmes "GEV" et la modélisation des excès par la méthode POT avec la distribution généralisée de Pareto "GPD", on a discuté également sur la notion des domaines d'attraction d'une distribution des valeurs extrêmes qui sont étroitement liés au concept de variation régulière, ensuite on a présenté les estimateurs de l'indice des valeurs extrêmes et des quantiles extrêmes avec leurs propriétés asymptotiques enfin on a exposé quelques méthodes adaptatives pour déterminer le nombre de statistiques d'ordre extrêmes. on a présenté les estimateurs de l'indice ainsi que leurs propriétés asymptotiques par des méthodes paramétrique et semi-paramétrique parmi ces méthodes le maximum de vraisemblance, celle des moments pondérés, à celle de l'estimateur de Hill qui est l'estimateur le plus populaire pour l'indice de queue positif et l'estimateur de Pickands, on a donné également certaines méthodes de sélection du nombre de valeurs extrêmes utilisé lors de l'estimation du paramètre de l'indice de queue, dans ce chapitre on a traité aussi les mesures de risque financiers, et aussi présenté quelques notations et définitions pour la Value-at-Risk ainsi que les différentes méthodes d'estimation de cette dernière. Enfin on a étudié certaines méthodes d'estimation de quantiles extrêmes.

Chapter 3

Distribution de Gumbel Exponentielle

Pouvons-nous utiliser d'autres distributions avec support comme générateur pour dériver différentes classes de distributions? Cette question est d'introduire une nouvelle technique pour dériver des familles de distributions en utilisant une fonction de densité de probabilité comme générateur. Une nouvelle méthode est proposée pour générer ces familles de distributions continues. On considère une variable aléatoire X appelée "le transformateur" qui est utilisée pour transformer une autre variable T "la transformée", ce résultat est la famille de distributions $T - X$.

3.1 Méthode du transformateur-transformée

Alzaatreh et al (2013)(3), ont introduit une méthode pour générer de nouvelles familles étiquetées $T - X$ qui remplit certaines conditions.

Definition 57 Soit X une variable aléatoire de densité f et F sa fonction de distribution cumulative. Soit T une variable aléatoire continue de densité de probabilité $r(x)$ définie sur $[a, b]$. $G(x)$ la fonction de répartition de la nouvelle famille de distribution définie comme suit :

$$G(x) = \int_a^{W(F(x))} r(t) dt \quad (3.1)$$

Où $W(F(x))$ vérifie les conditions suivantes

$$\left\{ \begin{array}{l} W(F(x)) \in [a, b] \\ W(F(x)) \text{ est différentiable et monotone non décroissante} \\ \text{si } x \rightarrow -\infty, W(F(x)) \rightarrow a \text{ et si } x \rightarrow \infty, W(F(x)) \rightarrow b \end{array} \right. \quad (3.2)$$

La fonction de densité de probabilité $g(x)$ est liée à la distribution de la fonction $G(x)$ avec la relation suivante

$$g(x) = \left\{ \frac{d}{dx} W(F(x)) \right\} r \{w(F(x))\} \quad (3.3)$$

Nous pouvons écrire

$$G(x) = \int_a^{W(F(x))} r(t) dt = R \{W(F(x))\} \quad (3.4)$$

Où $R(t)$ est la fonction de répartition de la variable aléatoire T .

La définition de $W(F(x))$ dépend du support de T .

Exemple de $W(\cdot)$

1) Lorsque $T \in [0, 1]$ (sans perte de généralisation), nous supposons que la distribution peut être considérée comme la loi, uniforme, beta, Kumaraswamy et autres types de distribution beta généralisée.

$W(F(x))$ peut être définie comme $F(x)$

2) Lorsque $T \in [a, b]$ (sans perte de généralisation) nous supposons que $a = 0$, $W(F(x))$ peut être défini comme

$$\left\{ \begin{array}{l} \left[\frac{F(x)}{1-F(x)} \right] \\ [\log(1 - F(x))] \\ [\log(1 - F^\alpha(x))] \\ \left[\frac{F^\alpha(x)}{1-F^\alpha(x)} \right] \end{array} \right. \quad (3.5)$$

3) Lorsque $T \in (-\infty, \infty)$, $W(F(x))$ peut être défini comme

$$\begin{cases} \log \left[\frac{F(x)}{1-F(x)} \right] \\ \log [-\log(1 - F(x))] \\ \log [-\log(1 - F^\alpha(x))] \\ \log \left[\frac{F^\alpha(x)}{1-F^\alpha(x)} \right] \end{cases} \quad (3.6)$$

3.2 Gumbel-Exponentielle Distribution "GE"

Dans cette section, on introduit une nouvelle famille de la distribution générée par une variable aléatoire T .

Theorem 58 *La fonction de densité de la probabilité $g(x)$ de la nouvelle distribution est donnée par*

$$g(x) = \frac{e^{-\frac{\mu}{\delta}}}{(x - \mu)^2} \left(\frac{\delta}{x - \mu} \right)^{\frac{1}{\delta}-1} \exp \left(-e^{-\frac{\mu}{\delta}} \left(\frac{\delta}{x - \mu} \right)^{\frac{1}{\delta}} \right), \mathbf{x} > \boldsymbol{\mu} \quad (3.7)$$

Proof. On a $T \in (-\infty, \infty)$ une variable aléatoire qui suit la distribution de Gumbel (68), de sorte que $r(t)$ soit la fonction de densité défini par :

$$r(t) = \frac{1}{\delta} \exp\left(\frac{t - \mu}{\delta}\right) \exp\left(-\exp\left(-\frac{t - \mu}{\delta}\right)\right), t \in \mathbb{R} \quad (3.8)$$

Avec

$$R(t) = \exp\left(-\exp\left(-\frac{t - \mu}{\delta}\right)\right), t \in \mathbb{R}$$

Et X une v.a qui suit la distribution exponentielle (68), sa fonction de répartition est:

$$F(x) = 1 - e^{-\frac{x-\mu}{\delta}}, x \geq 0 \quad (3.9)$$

Donc

$$w(F(x)) = \log[-\log(1 - F(x))] = \log\left(\frac{x - \mu}{\delta}\right) \quad (3.10)$$

On trouve

$$g(x) = \left\{ \frac{d}{dx} W(F(x)) \right\} r\{w(F(x))\} \quad (3.11)$$

et

$$G(x) = \int_a^{W(F(x))} r(t) dt = R\{W(F(x))\} \quad (3.12)$$

$$\begin{aligned} G(x) &= \int_{\mu}^{W(F(x))} r(t) dt = R\{W(F(x))\} \\ &= \int_{\mu}^{W(F(x))} \frac{1}{\delta} \exp\left(-\frac{t - \mu}{\delta}\right) \exp\left(-\exp\left(-\frac{t - \mu}{\delta}\right)\right) dt \\ &= \left[\exp\left(-\exp\left(-\frac{t - \mu}{\delta}\right)\right) \right]_{\mu}^{W(F(x))} \\ &= \exp\left(-\exp\left(-\frac{\log\left(\frac{x - \mu}{\delta}\right) - \mu}{\delta}\right)\right) - \exp\left(-\exp\left(-\frac{\mu - \mu}{\delta}\right)\right) \\ &= \exp\left(-\exp\left(-\frac{\log\left(\frac{x - \mu}{\delta}\right) - \mu}{\delta}\right)\right) - \exp(-1) \\ &= \exp\left(-\exp\left(\frac{-\log\left(\frac{x - \mu}{\delta}\right) - \mu}{\delta}\right)\right) - \frac{1}{e} \\ &= \exp\left(-\exp\left(\log\left(\frac{\delta}{x - \mu}\right)^{\frac{1}{\delta}} \exp\left(-\frac{\mu}{\delta}\right)\right)\right) - \frac{1}{e} \\ &= \exp\left(-\left(\frac{\delta}{x - \mu}\right)^{\frac{1}{\delta}} \exp\left(-\frac{\mu}{\delta}\right)\right) - \frac{1}{e} \end{aligned}$$

Alors la fonction de répartition $G(x)$ de la nouvelle distribution est donnée par

$$G(x) = \exp\left(-\left(\frac{\delta}{x - \mu}\right)^{\frac{1}{\delta}} \exp\left(-\frac{\mu}{\delta}\right)\right) - \frac{1}{e}, \quad \mathbf{x} > \boldsymbol{\mu} \quad (3.13)$$

on dérive $G(x)$ on trouve

$$\frac{d\left(\exp\left(-\left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}} \exp\left(-\frac{\mu}{\delta}\right)\right) - \frac{1}{e}\right)}{dx} \quad (3.14)$$

$$= \frac{e^{-\frac{\mu}{\delta}}}{(x-\mu)^2} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}-1} \exp\left(-e^{-\frac{\mu}{\delta}} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}}\right) \quad (3.15)$$

La fonction de densité de la probabilité $g(x)$ est

$$g(x) = \frac{e^{-\frac{\mu}{\delta}}}{(x-\mu)^2} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}-1} \exp\left(-e^{-\frac{\mu}{\delta}} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}}\right), \quad \mathbf{x} > \boldsymbol{\mu} \quad (3.16)$$

■

3.2.1 Fonction quantile

Theorem 59 La fonction quantile de X est $G_x^{-1}(\gamma)$, où $0 < \gamma < 1$. alors la fonction quantile de la distribution GE est :

$$G^{-1}(\gamma) = \frac{\delta}{-\left(\ln\left(\gamma + \frac{1}{e}\right) / \exp\left(-\frac{\mu}{\delta}\right)\right)^{\delta}} + \mu \quad (3.17)$$

Proof. Pour $\delta, \mu > 0$, nous devons résoudre l'équation $G_X(x) = \gamma$ par rapport à x pour $x > \mu$

$$\begin{aligned} G_X(x) &= \gamma & (3.18) \\ \exp\left(-\left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}} \exp\left(-\frac{\mu}{\delta}\right)\right) - \frac{1}{e} &= \gamma \\ \exp\left(-\left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}} \exp\left(-\frac{\mu}{\delta}\right)\right) &= \gamma + \frac{1}{e} \end{aligned}$$

En multipliant par log les deux côtés de l'équation (18), on obtient

$$\left(-\left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}} \exp\left(-\frac{\mu}{\delta}\right)\right) = \log\left(\gamma + \frac{1}{e}\right) \quad (3.19)$$

En résolvant la dernière équation par rapport à x , nous obtenons la fonction quantile de la distribution GE

$$G^{-1}(\gamma) = \frac{\delta}{-\left(\ln\left(\gamma + \frac{1}{e}\right) / \exp\left(\frac{-\mu}{\delta}\right)\right)^\delta} + \mu \quad (3.20)$$

■

Pour $\gamma = 0,25$, $\gamma = 0,50$ et $\gamma = 0,75$, les trois quartiles de la distribution GE sont

$$G^{-1}\left(\frac{1}{4}\right) = \frac{\delta}{-\left(\ln\left(\frac{1}{4} + \frac{1}{e}\right) / \exp\left(\frac{-\mu}{\delta}\right)\right)^\delta} + \mu \quad (3.21)$$

$$G^{-1}\left(\frac{1}{2}\right) = \frac{\delta}{-\left(\ln\left(\frac{1}{2} + \frac{1}{e}\right) / \exp\left(\frac{-\mu}{\delta}\right)\right)^\delta} + \mu \quad (3.22)$$

$$G^{-1}\left(\frac{3}{4}\right) = \frac{\delta}{-\left(\ln\left(\frac{3}{4} + \frac{1}{e}\right) / \exp\left(\frac{-\mu}{\delta}\right)\right)^\delta} + \mu \quad (3.23)$$

3.2.2 La fonction de risque

Lemma 60 *La fonction de risque de la distribution GE est*

$$h(x) = \frac{\frac{e^{-\frac{\mu}{\delta}}}{(x-\mu)^2} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}-1}}{\left(1 + \frac{1}{e}\right)z - 1} \quad (3.24)$$

$$\text{Où } z = \exp\left(e^{-\frac{\mu}{\delta}} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}}\right)$$

Proof. nous avons $h(x) = \frac{g(x)}{1-G(x)}$ Par simple calcul, on obtient

$$\begin{aligned} h(x) &= \frac{g(x)}{1-G(x)} \\ &= \frac{\frac{e^{-\frac{\mu}{\delta}}}{(x-\mu)^2} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}-1} \exp\left(-e^{-\frac{\mu}{\delta}} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}}\right)}{1 - \exp\left(-e^{-\frac{\mu}{\delta}} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}}\right) + \frac{1}{e}} \end{aligned}$$

on multiplie et on divise par $\exp\left(e^{-\frac{\mu}{\delta}}\left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}}\right)$ on trouve

$$h(x) = \frac{g(x)}{1 - G(x)} = \frac{\frac{e^{-\frac{\mu}{\delta}}}{(x-\mu)^2} \left(\frac{\delta}{x-\mu}\right)^{\frac{1}{\delta}-1}}{\left(1 + \frac{1}{e}\right)z - 1} \quad (3.25)$$

■

3.2.3 Moments et moments générateurs de fonctions

fonction génératrices des moments

La fonction de génératrice de moment (MGF) d'une variable aléatoire continue X est donnée par

$$M_X(t) = \int_{\mu}^{\infty} \exp(tx)g(x)dx = \int_{\mu}^{\infty} \left(\sum_{i=0}^{\infty} \frac{t^i x^i}{i!} g(x) \right) dx \quad (3.26)$$

$$M_X(t) = \sum_{i=0}^{\infty} \frac{t^i}{i!} E(X^i) = \frac{t^i}{i!} e^{-\frac{\mu}{\delta}} \sum_{k=0}^n C_k^n \delta^k \mu^{n-k} \frac{\Gamma(1 - k\delta)}{e^{-\frac{\mu}{\delta}(1-k\delta)}} \quad (3.27)$$

Moments

Le premier moment de X peut s'écrire

$$E(X) = \int_{\mu}^{\infty} xg(x)dx = -\delta e^{-\mu} \Gamma(1 - \delta) - \mu \quad (3.28)$$

Le moment d'ordre 2 de la distribution GE est

$$E(X^2) = -\delta^2 e^{-2\mu} \Gamma(1 - 2\delta) - \mu^2 - 2\mu\delta e^{-\mu} \Gamma(1 - \delta) \quad (3.29)$$

Le n ième moment de la distribution GE est

$$E(X^n) = \int_{\mu}^{\infty} x^n g(x)dx = \int_0^{\infty} (\delta y^{-\delta} + \mu)^n \exp\left(-e^{-\frac{\mu}{\delta}} y\right) dy \quad (3.30)$$

$$E(X^n) = \sum_{k=0}^n C_k^n \delta^k \mu^{n-k} \int_0^{\infty} y^{-k\delta} dy = e^{-\frac{\mu}{\delta}} \sum_{k=0}^n C_k^n \delta^k \mu^{n-k} \frac{\Gamma(1-k\delta)}{e^{-\frac{\mu}{\delta}(1-k\delta)}} \quad (3.31)$$

Pour $k = 1, \dots, n$

Les skewness et kurtosis de la distribution GE sont

$$skewness = \frac{E(x^3)}{V(X)^{\frac{3}{2}}} = \frac{e^{-\frac{\mu}{\delta}} \sum_{k=0}^3 C_k^3 \delta^k \mu^{3-k} \frac{\Gamma(1-k\delta)}{e^{-\frac{\mu}{\delta}(1-k\delta)}}}{\left[-\delta^2 e^{-2\mu} \Gamma(1-2\delta) + \Gamma(1-\delta) \left[\delta^2 e^{-2\mu} \Gamma(1-\delta) - \delta \mu e^{-\mu} \right] \right]^{\frac{3}{2}}} \quad (3.32)$$

$$skewness = \frac{e^{-\frac{\mu}{\delta}} \left[\mu^3 + 3\delta \mu^2 \frac{\Gamma(1-\delta)}{e^{-\frac{\mu}{\delta}(1-\delta)}} + 3\delta^2 \mu \frac{\Gamma(1-2\delta)}{e^{-\frac{\mu}{\delta}(1-2\delta)}} + \delta^3 \frac{\Gamma(1-3\delta)}{e^{-\frac{\mu}{\delta}(1-3\delta)}} \right]}{\left[-\delta^2 e^{-2\mu} \Gamma(1-2\delta) + \Gamma(1-\delta) \left[\delta^2 e^{-2\mu} \Gamma(1-\delta) - \delta \mu e^{-\mu} \right] \right]^{\frac{3}{2}}}$$

$$kurtosis = \frac{E(x^4)}{V(X)^2} = \frac{e^{-\frac{\mu}{\delta}} \sum_{k=0}^4 C_k^4 \delta^k \mu^{4-k} \frac{\Gamma(1-k\delta)}{e^{-\frac{\mu}{\delta}(1-k\delta)}}}{\left[-\delta^2 e^{-2\mu} \Gamma(1-2\delta) + \Gamma(1-\delta) \left[\delta^2 e^{-2\mu} \Gamma(1-\delta) - \delta \mu e^{-\mu} \right] \right]^2} \quad (3.33)$$

$$kurtosis = \frac{e^{-\frac{\mu}{\delta}} \left[\mu^4 + 4\delta \mu^3 \frac{\Gamma(1-\delta)}{e^{-\frac{\mu}{\delta}(1-\delta)}} + \delta^2 \mu^2 \frac{\Gamma(1-2\delta)}{e^{-\frac{\mu}{\delta}(1-2\delta)}} + 4\delta^3 \mu \frac{\Gamma(1-3\delta)}{e^{-\frac{\mu}{\delta}(1-3\delta)}} + \delta^4 \frac{\Gamma(1-4\delta)}{e^{-\frac{\mu}{\delta}(1-4\delta)}} \right]}{\left[-\delta^2 e^{-2\mu} \Gamma(1-2\delta) + \Gamma(1-\delta) \left[\delta^2 e^{-2\mu} \Gamma(1-\delta) - \delta \mu e^{-\mu} \right] \right]^2} \quad (3.34)$$

Tableau 01

μ	δ	Mean	skewnes	kurtosis
0.1	0.1	1,2766	0,40366	0,38322
	0.2	0,78267	0,31614	0,26913
	0.3	0,60279	0,29366	0,25471
	0.4	0,53335	0,28928	0,42978
0.5	0.1	0,7671	0,28669	0,28604
	0.2	0,58922	0,25535	0,25605
	0.3	0,54216	0,2513	0,25292
	0.4	0,54216	0,25032	0,25072

Le tableau affiche la moyenne, skewness et le kurtosis de la distribution GE pour différents choix du paramètre δ et μ .

3.2.4 L'entropie

L'entropie (Rényi, 1961), d'une variable aléatoire X est une mesure de variation de l'incertitude, celle de la distribution GE est donnée par

$$\begin{aligned} I_R(s) &= \frac{1}{1-s} \left\{ \ln \left(\int_1^{\infty} g^s(x) dx \right) \right\}, \quad s > 0, s \neq 1 \\ &= \frac{e^{-\frac{s\mu}{\delta}}}{e^{-\frac{s\mu}{\delta}(\delta s + s - \delta)}(1-s)} \ln \{ \delta^{-2s+2} \Gamma(\delta s + s - \delta) \}, \quad s > 0, s \neq 1 \end{aligned}$$

De plus, l'entropie de Shannon (1948) définie par $E \{-\ln(g(x))\}$, est un cas particulier de $\lim_{s \rightarrow 1} I_R(s)$.

3.2.5 La fiabilité

La mesure de la fiabilité des composants industriels a de nombreuses applications, en particulier dans le domaine des tests de durée de vie et de l'ingénierie, le composant échoue au moment où la contrainte aléatoire X_2 qui lui est appliquée dépasse la résistance aléatoire X_1 , et le composant fonctionnera de manière satisfaisante chaque fois que $X_1 > X_2$. Par conséquent, $R = P[X_2 < X_1]$ est une mesure de la fiabilité des composants. Nous dérivons la fiabilité R lorsque X_1 et X_2 ont des $GE(\delta_1; \mu)$ et $GE(\delta_2; \mu)$ des distributions avec le même paramètre de forme μ et le paramètre fixe δ . La fiabilité est définie comme suit

$$R = \int_0^{\infty} g_1(x) G_2(x) dx = -k_1 \Gamma(2 - \delta_1) G_2(x), \quad k = e^{-\frac{\mu}{\delta_1} + \exp(-\frac{\mu}{\delta_1})} \quad (3.35)$$

3.2.6 Estimations du maximum de vraisemblance (MV)

Soit X_i suit $GE(\delta, \mu)$, $i = 1, \dots, n$, variables aléatoires, la fonction de vraisemblance maximale $l(x, \mu, \delta)$ est

$$L(\theta) = \ln l(x, \mu, \delta) = \sum_{i=1}^n \left(-\frac{\mu}{\delta}\right) - 2n \log(x - \mu) + n\left(\frac{1}{\delta} - 1\right) \log \delta - \left(\frac{1}{\delta} - 1\right) \sum_{i=1}^n \log(x_i - \mu) + \sum_{i=1}^n \left(\left(\frac{1}{x_i - \mu}\right) * \exp\left(\frac{\mu}{\delta}\right) \right) \quad (3.36)$$

Les dérivées de $L(\theta)$ par rapport à δ et μ sont

$$\frac{\partial \log \Pi g(x)}{\partial \delta} = \sum_{i=1}^n \left(\frac{\mu}{\delta^2}\right) - \frac{n}{\delta^2} \log \delta + n\frac{1}{\delta} \left(\frac{1}{\delta} - 1\right) - \frac{\mu}{\delta^2} \sum_{i=1}^n \left(\frac{\mu}{\delta^2} \left(\frac{\delta}{x_i - \mu}\right) * \exp\left(\frac{\mu}{\delta}\right) \right) \quad (3.37)$$

$$\begin{aligned} \frac{\partial \log \Pi g(x)}{\partial \mu} &= \sum_{i=1}^n \left(-\frac{1}{\delta}\right) + \frac{2n}{(x - \mu)} + \left(\frac{1}{\delta} - 1\right) \sum_{i=1}^n \frac{1}{(x_i - \mu)} - \sum_{i=1}^n \frac{1}{(x_i - \mu)^2} \\ &\times \exp\left(\frac{\mu}{\delta}\right) - \frac{1}{\delta^2} \sum_{i=1}^n \left(\frac{1}{x_i - \mu}\right) * \exp\left(\frac{\mu}{\delta}\right) \end{aligned} \quad (3.38)$$

Les deux équations ne peuvent pas être résolues directement, nous devons utiliser une méthode numérique par exemple la méthode de Fisher. Nous avons

$$\begin{bmatrix} \frac{\partial L(\theta)}{\partial \delta^2} & \frac{\partial^2 L(\theta)}{\partial \delta \partial \mu} \\ \frac{\partial^2 L(\theta)}{\partial \mu \partial \delta} & \frac{\partial L(\theta)}{\partial \mu^2} \end{bmatrix}_{\substack{\hat{\delta} = \delta_0 \\ \hat{\mu} = \mu_0}} \begin{bmatrix} \hat{\delta} = \delta_0 \\ \hat{\mu} = \mu_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial L(\theta)}{\partial \delta} \\ \frac{\partial L(\theta)}{\partial \mu} \end{bmatrix}_{\substack{\hat{\delta} = \delta_0 \\ \hat{\mu} = \mu_0}} \quad (3.39)$$

L'équation (42) peut être résolue de manière itérative où δ_0, μ_0 sont les valeurs initiales de δ et μ .

3.2.7 Existence et unicité des paramètres MV

Lemma 61 *il existe un sous-ensemble compact $K \equiv K(\eta) \subset (0, \infty) \times (0, \infty)$, tel que*

$$\{(\delta, \mu) : L(\theta) \geq -\eta\} \subset K \quad (3.40)$$

Theorem 62 *Supposons que $X_i \sim GE(\delta, \mu)$, $i = 1, \dots, n$, alors les estimateurs des paramètres δ et μ de la distribution exponentielle de Gumbel existent et sont uniques.*

Proof. Il suffit de montrer que les estimateurs des paramètres MV δ et μ existent uniquement. Selon les résultats de Mékeläinen et al. (1981), dans l'ordre pour montrer l'existence et l'unicité les estimateurs des paramètres MV δ et μ , il suffit pour vérifier les deux conditions suivantes

- i) Pour tout $\eta > 0$ donné, l'équation est donnée
- ii) La matrice de H est définie négative en tout point $(\delta; \mu) \in (0, \infty) \times (0, \infty)$.

$$H = \begin{bmatrix} \frac{\partial L(\theta)}{\partial \delta^2} & \frac{\partial^2 L(\theta)}{\partial \delta \partial \mu} \\ \frac{\partial^2 L(\theta)}{\partial \mu \partial \delta} & \frac{\partial L(\theta)}{\partial \mu^2} \end{bmatrix} \quad (3.41)$$

■

La condition i est certainement satisfaite par le lemme 1. Par conséquent, pour prouver le théorème, il suffit de montrer ii, on a $x^t H x < 0$ où $x^t = (x_1 \ x_2)$ alors H est définie négative.

3.2.8 Simulation

Dans cette section, nous étudions le comportement des estimateurs MV pour une taille d'échantillon finie (n). Une étude de simulation suit les étapes suivantes pour chacun triplet (δ, μ, n) .

- Choisir les valeurs initiales des paramètres pour les éléments correspondants au vecteur

de paramètres (δ, μ) pour spécifier la distribution $GE(\delta, \mu)$ et la taille de l'échantillon, nous avons cinq tailles d'échantillons ($n = 30, 50, 100, 500, 1000$);

-Générer $N = 5000$ échantillons indépendants de taille n à partir de la distribution $GE(\delta, \mu)$;

-Calculer l'estimation MV de $\theta\{n\}$ pour chacun des N échantillons;

-Calculer la moyenne des estimateurs obtenus sur tous les N échantillons

$$\text{biais moyen}(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_0) \quad (3.42)$$

Et l'erreur quadratique moyenne

$$MSE(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_0)^2 \quad (3.43)$$

Tableaux03

N	$\text{biais}(\delta) \delta = 0.15$	$\text{biais}(\mu)\delta = 0.15$	$\text{biais}(\delta) \delta = 0.3$	$\text{biais}(\mu)\delta = 0.3$
$n = 30$	0,0668	-0,2034	-0,971	0,0210
$n = 50$	0,0394	0,0653	0,1445	0,0589
$n = 100$	0,0186	0,0488	0,0521	0,0117
$n = 500$	0,0112	0,0087	0,0376	0,0022
$n = 1000$	0,00894	0,0058	0,0091	0,00194

Tableaux 04

N	$\text{biais}(\delta) \delta = 0.08$	$\text{biais}(\mu)\delta = 0.08$	$\text{biais}(\delta) \delta = 0.2$	$\text{biais}(\mu)\delta = 0.2$
$n = 30$	0,0355	0,0097	0,5102	0,3361
$n = 50$	0,0258	0,0066	0,5010	-0,0616
$n = 100$	0,0128	0,0066	0,1447	-0,2307
$n = 500$	0,0095	0,0244	0,0499	-0,444
$n = 1000$	0,0073	-0,244	0,0069	0,0790

Tableaux 05

N	$biais(\delta) \delta = 0.25$	$biais(\mu)\delta = 0.25$	$biais(\delta) \delta = 0.05$	$biais(\mu)\delta = 0.05$
$n = 30$	-0.9183	-0.7425	-0.744221	-0.494221
$n = 50$	-0.96743	-0.7485	-0.7453921	-0.4953921
$n = 100$	0.9112	-0.74925	-0.7457845	-0.4957845
$n = 500$	0.099	0.794221	0.2457845	0.195784
$n = 1000$	0.0998	-0.7953921	0.02457845	0.1953921

Les tableaux 03,04,05 représentent le biais moyen des estimations simulées avec $\mu = 0.05, \mu = 0.1, 0.75$.

Tableaux 06

N	$MSE(\delta) \delta = 0.15$	$MSE(\mu)\delta = 0.15$	$MSE(\delta) \delta = 0.03$	$MSE(\mu)\delta = 0.03$
$n = 30$	4,4654 10^{-3}	3,6786 10^{-4}	8,0237 10^{-3}	2,4364 10^{-3}
$n = 50$	1,5552 10^{-3}	7,4996 10^{-5}	4,6262 10^{-3}	7,2359 10^{-4}
$n = 100$	3,4546 10^{-4}	3,3405 10^{-5}	1,4167 10^{-3}	2,7785 10^{-4}
$n = 500$	2,1128 10^{-5}	2,3903 10^{-5}	8,3529 10^{-5}	1,0651 10^{-4}
$n = 1000$	1,8561 10^{-5}	5,0401 10^{-6}	8,3529 10^{-5}	4,3718 10^{-5}

Tableaux 07

N	$MSE(\delta) \delta = 0.08$	$MSE(\mu)\delta = 0.08$	$MSE(\delta) \delta = 0.2$	$MSE(\mu)\delta = 0.2$
$n = 30$	4,4654 10^{-3}	3,6786 10^{-4}	8,0237 10^{-3}	2,4364 10^{-3}
$n = 50$	1,5552 10^{-3}	7,4996 10^{-5}	4,6262 10^{-3}	7,2359 10^{-4}
$n = 100$	3,4546 10^{-4}	3,3405 10^{-5}	1,4167 10^{-3}	2,7785 10^{-4}
$n = 500$	2,1128 10^{-5}	2,3903 10^{-5}	8,3529 10^{-5}	1,0651 10^{-4}
$n = 1000$	1,8561 10^{-5}	5,0401 10^{-6}	3,3000 10^{-5}	4,3718 10^{-5}

Tableaux08

N	$MSE(\delta) \delta = 0.25$	$MSE(\mu)\delta = 0.25$	$MSE(\delta) \delta = 0.05$	$MSE(\mu)\delta = 0.05$
$n = 30$	$2,2451 \cdot 10^{-4}$	$1,0935 \cdot 10^{-2}$	$9,3359 \cdot 10^{-5}$	$4,5248 \cdot 10^{-3}$
$n = 50$	$2,9504 \cdot 10^{-4}$	$6,8303 \cdot 10^{-3}$	$4,3801 \cdot 10^{-5}$	$2,4916 \cdot 10^{-3}$
$n = 100$	$2,7258 \cdot 10^{-5}$	$9,0554 \cdot 10^{-5}$	$4,0024 \cdot 10^{-5}$	$6,3244 \cdot 10^{-4}$
$n = 500$	$1,3610 \cdot 10^{-5}$	$2,05867 \cdot 10^{-5}$	$2,8911 \cdot 10^{-6}$	$1,8568 \cdot 10^{-4}$
$n = 1000$	$1,3610 \cdot 10^{-5}$	$2,5698 \cdot 10^{-5}$	$2,0155 \cdot 10^{-7}$	$4,7630 \cdot 10^{-5}$

Les tableaux 06,07,08 représentent MSE des estimations simulées avec $\mu = 0.05, \mu = 0.1, \mu = 0.75$.

Discusion

Les tableaux précédents montrent comment le biais et les erreurs quadratiques moyennes varient en fonction de n . Les erreurs quadratiques moyennes et biais moyens pour chaque paramètre diminuent vers zéro lorsque $n \rightarrow \infty$. Ces résultats numériques coïncident avec les résultats théoriques établis.

Chapter 4

Application

La caractérisation et la quantification du comportement de la queue d'événements rares est un problème de gestion du risque financier, il est d'un grand intérêt pour les chercheurs, les régulateurs et les gestionnaires de fonds, en raison des développements négatifs dramatiques que la survenue d'événements extrêmes comme une crise financière mondiale provoque sur les marchés financiers. La quantification du risque des marchés financiers consiste fondamentalement d'écrire l'incertitude prix/rendement résultant des mouvements du marché. La prédiction d'événements financiers extrêmes est liés aux queues de la distribution du processus de génération de données, ce qui est crucial pour la gestion des risques et la diversification du portefeuille. Par conséquent, pour quantifier adéquatement le risque (catastrophique) associé aux queues des distributions empiriques via la VaR, il est important de supposer des modèles distributionnels flexibles qui peuvent correctement rendre compte de la valeur extrême, en finance et en économie, les applications de la théorie des valeurs extrêmes ont été très intensives, la première application de la théorie des valeurs extrêmes en finance, a été proposé par Parkinson en (1980). Par la suite, plusieurs autres chercheurs ont fourni de nouvelles applications aux données financières. Par exemple, Longina a utilisé la TVE pour décrire les mouvements de prix élevés des actions de la Bourse de New York pour la période 1985 – 1990, il a identifié la distribution de Fréchet comme le meilleur modèle pour les données utilisées, Tolikas et Brown ont utilisé trois distributions dans le cadre de travail pour étudier la distribution asymptotique des queues inférieures des rendements quotidiens de la Bourse d'Athènes sur la période 1986 – 2001. Gettinby et

al, ont caractérisé les extrêmes des rendements des actions aux états-Unis, au Royaume-Uni et au Japon pour la période 1963 – 2000. Ils ont conclu que la distribution logistique généralisée surpassait les autres distributions. Plus récemment Karmakar et Shukla ont utilisé l’approche en deux étapes de McNeil et Frey également appelé TVE conditionnelle pour étudier la performance relative de la VaR des indices quotidiens des cours des actions de six pays différents. Izgi et Duran ont régulé le co-mouvement et la polarisation des taux d’intérêt et des rendements quotidiens de l’indice BIST entre 2010 et 2013 en utilisant la TVE.

4.0.9 Les données réelles

Les données utilisées dans cette thèse sont les indices boursiers de 2 pays la Russie et la France. Le CAC 40 est un indice boursier qui réunit les quarante plus importantes sociétés françaises cotées en Bourse. Ces quarante dernières sociétés sont choisies parmi les cent sociétés françaises au sein desquelles les volumes d’échanges de titres sont les plus importants.

Le sigle CAC 40 correspond à la cotation assistée en continu, ce qui signifie que sa valeur augmente ou diminue en permanence tous les jours.

L’indice RTS (en anglais RTS Index ou RTSI) est un indice boursier basé sur les 50 entreprises russes cotées de la bourse des valeurs RTS de Moscou. L’indice 100 correspond à leur capitalisation totale au 1er septembre 1995. (Code ISIN : RU000A0JPEB3 - Mnemo : RTSI\$). L’indice du système commercial russe comprend 50 des plus grands titres capturant 85% de la capitalisation boursière totale de la Russie. On note également que l’horizon temporel utilisé couvre une période de stress financier. Le choix de cette action et l’indice peuvent s’expliquer par le niveau élevé de liquidité.

4.0.10 Analyse descriptive des indices financiers CAC 40 et RTS50

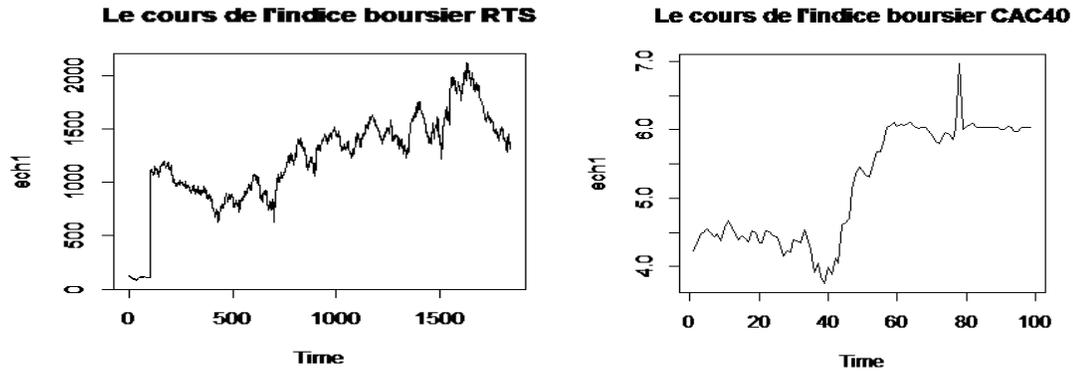


Figure 4 .02 :des cours d'indice boursier
"RTS50"

Figure 4.01: des cours d'indice boursier
"CAC40"

Ces figures présentent les variations des cours des indices boursier "CAC40 et "RTS50 de la période allant de 20/12/2015 jusqu 14/05 2020 et 01/01/1995 jusqu 01/ 01/2018 respectivement. On remarque qu'il existe quelques chutes brutales à cause de l'influence des indices par les crises financières.

4.0.11 Analyse descriptive des log des rendements de CAC 40 et RTS50

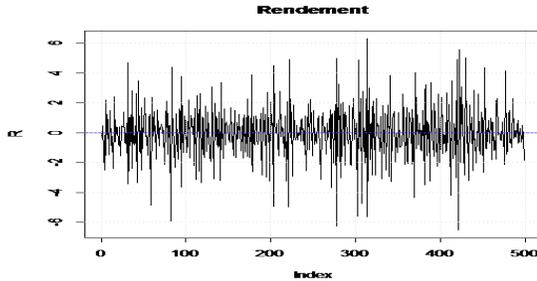
Definition 63 Rendement

Soit p_t le prix d'un actif à la date t et r_t le logarithme du rendement de l'indice i pour la même t , on va s'intéresser à une transformation de ce prix par rapport une durée

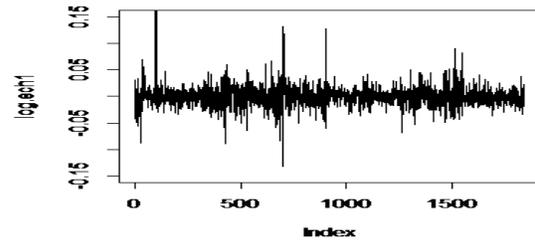
$$r_{i,t} = \log\left(\frac{p_{i,t}}{p_{i,t-1}}\right) \quad (4.1)$$

où $p_{i,t}$ est le cours de clôture de l'indice à la date t et $p_{i,t-1}$ est le prix de l'indice à la date $t - 1$.

Grphe des logarithmes des rendements



rendements d'indice boursier "RTS50"



rendements d'indice boursier "CAC40"

Ces figures présentent l'évolution de la série des rendements des indices boursiers "CAC40" et "RTS50", il existe de grandes volatilités dans les périodes des crises financières. On constate que le rendement se fluctue autour de la moyenne, il a une tendance normale. On montrera cette remarque avec les tests qui suivent.

4.0.12 Tests de normalité

Les tests de normalité permettent de savoir si l'échantillon suit bien une loi normale ou non,

en supposant que l'hypothèse $H_0 \rightarrow$ le rendement ne suit pas une distribution normale.

si la p-value $< \alpha$ on rejette l'hypothèse nulle H_0

si la p-value $\geq \alpha$ on accepte l'hypothèse nulle H_0

Nous mentionnons les tests suivants, le test de Shapiro-Wilk, le test de Shapiro Francia et le test de Lilliefors, voir le tableau suivant pour leur valeur de p (p-value).

	CAC40	RTS50
Shapiro-Wilk	0.125×10^{-49}	0.624×10^{-46}
Shapiro-Francia	0.870×10^{-48}	0.3×10^{-44}
Lilliefors	0.270×10^{-46}	0.7×10^{111}

Ces tests de normalité pour les données de "RTS50" et "CAC40" dans tous les cas, on

trouve que la p-value est strictement inférieure $2.2e - 16$ c-à-d que la p-value est strictement inférieure α , alors on rejette l'hypothèse. Donc le rendement suit bien une loi normale.

4.0.13 Modélisation de la distribution "GE" avec l'indice "CAC40"

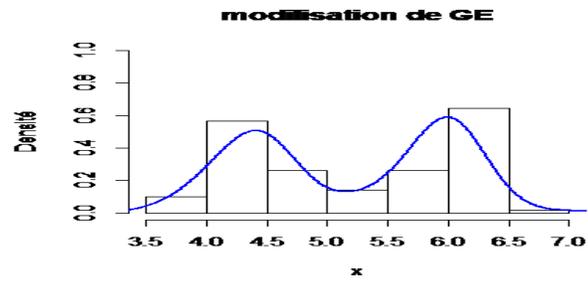
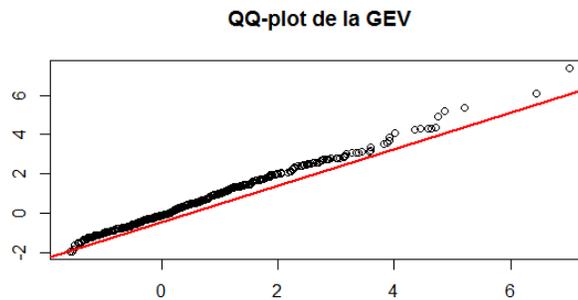


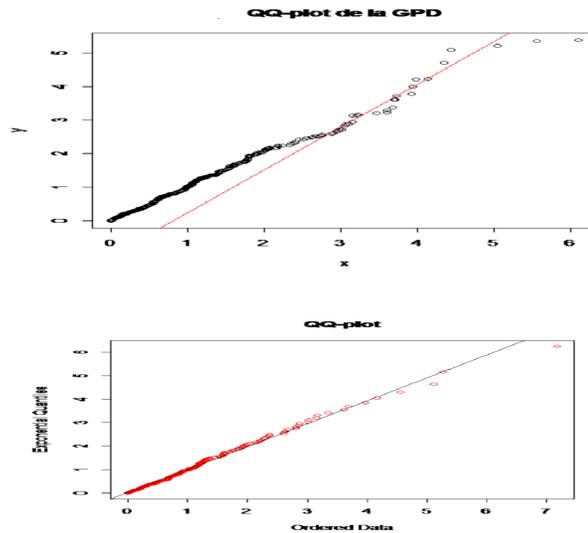
Figure 4.04 : la modélisation de distribution GE avec CAC 40

La figure présente la modélisation de la distribution GE avec "CAC 40". La densité de la distribution des extrêmes du cours de ce indice boursiers est une distribution de GE avec les paramètres estimés par la méthode de MV.

4.0.14 Outils graphiques pour l'identification de distribution

Quantile-Quantile plot

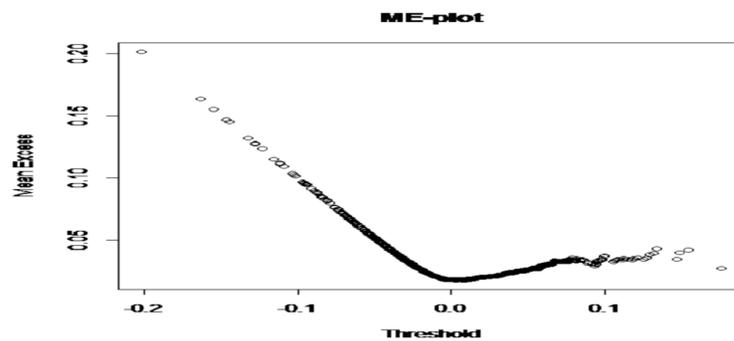




Ces figures présentent le QQ-plot, on observe une courbe linéaire qui signifie que la distribution de quantile empirique est semblable la distribution quantile théorique de la GEV,GPD et GE respectivement.

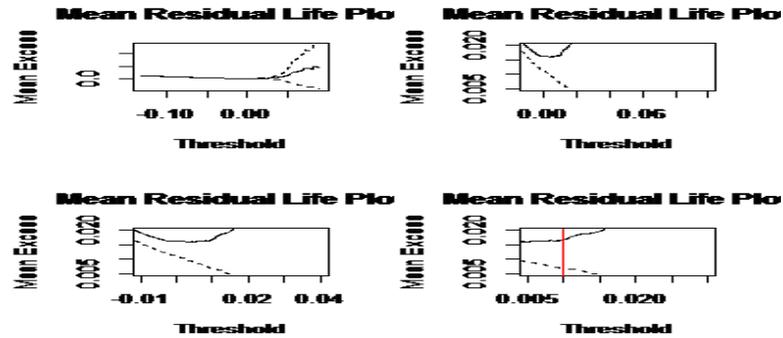
Le choix du seuil

Le choix du seuil pour des estimateurs précis doit être grand, pour cela on va le calculer par différentes méthodes. Le graphe de la fonction moyenne des excès



La figure présente le graphe de ME-plot, on remarque qu'il devient horizontal avec l'axe des abscisses depuis 0.0 jusqu'à 0.04, on peut donc choisir notre seuil dans cet intervalle.

le graphe de MRL plot



La figure présente le graphe MRL plot, on zoome la partie horizontale avec l'axe des abscisses pour bien préciser notre seuil, puis on choisit un seuil, qu'on a présenté avec la droite rouge.

4.0.15 Illustration: Application sur les données des indices "CAC 40" et "RTS50"

On montre à présent que la distribution GE peut être un meilleur modèle pour certaines distributions, par exemple Nadarajah (2006) ont présenté la distribution exponentielle de Gumbel, Persson et Ryden (2010) ont examiné cette distribution dans l'application océanographique, de plus, Pinheiro et Ferrari (2016) comparent plusieurs généralisations de la distribution de Gumbel. Nous avons fourni une applicabilité de la distribution GE en considérant les données réelles utilisées par les différents chercheurs. La distribution de GE est ajustée aux données dans chaque cas, les paramètres sont estimés par le maximum de vraisemblance en utilisant la commanderMaxlik (<https://CRAN.R-project.org/package=maxLik>) à l'aide du logiciel R-software. Nous adaptons les distributions (GE), (GPD), (GEV).

Tableau4.01

Indice	Distribution	Paramètres
CAC40	GED	$\hat{\delta} = 0.001193883, \hat{\mu} = 0.01323464, \hat{\xi} = 0$
	GEV	$\hat{\delta} = 0.001205942, \hat{\mu} = 0.01336832, \hat{\xi} = 0.05$
	GPD	$\hat{\delta} = 0.001205942, \hat{\mu} = 0.01336832; \hat{\xi} = 0.05$
RTS50	GED	$\hat{\delta} = 3.5 \times 10^{-5}, \hat{\mu} = 2.7 \times 10^{-7}, \hat{\xi} = 0$
	GEV	$\hat{\delta} = 1.7 \times 10^{-3}, \hat{\mu} = 4.3 \times 10^{-6}, \hat{\xi} = 0.5$
	GPD	$\hat{\delta} = 9.6 \times 10^{-4}, \hat{\mu} = 8.0 \times 10^{-6}; \hat{\xi} = 0.5$

Le tableau présente les estimation des paramètres pour les données des indices.

On a utilisé des outils statistiques pour comparer les modèles et choisir le meilleur possible parmi les autres distribution à partir des ensembles de données, ces outils sont AIC(4) , BIC(71) AICc(41), HQC(42) et CAIC(9) décrits comme suit :

$$AIC = -2LL + 2p \quad (4.2)$$

$$BIC = -2LL + p \ln(n) \quad (4.3)$$

$$AICc = AIC + \frac{2p(p+1)}{n-p-1}$$

$$HQC = -2LL + 2p(\ln(\ln(n)))$$

$$CACI = -2LL + p(\ln(n) + 1) \quad (4.4)$$

où p est le nombre de paramètres à estimer du modèle et L est le maximum de la fonction de vraisemblance du modèle.avec n le nombre d'observations dans l'échantillon.

Tableau 4.03

Indice	Distribution	-LL	AIC	BIC	AICc	HQC	CAIC
CAC40	GEV	3007.501	6009.002	5994.202	6008.979	6003.384	5991.202
	GPD	2998.798	5991.596	5976.796	5991.573	5985.978	5973.796
	GED	2980.791	5955.582	5940.782	5955.559	5949.964	5937.782
RTS50	GEV	2632.706	5259.412	5244.612	5259.389	5253.794	5241.612
	GPD	2583.680	5161.360	5146.560	5161.337	5155.742	5143.560
	GED	2483.680	4961.360	4946.560	4961.337	4955.742	4943.560

Le tableau affiche les valeurs des outils statistiques AIC, BIC, AICc, HQC et CAIC pour les données des indices "CAC40" et "RTS50".

Discusion A partir du tableau 4.03 et après une lecture visuelle des résultats numériques des paramètres estimés, on a observé que la distribution de GE donne les plus petites valeurs de log-vraisemblance négative, AIC, BIC, AICc, HQC et CAIC, on conclue que le modèle GE est plus puissant pour la manipulation des données comparativement aux autres modèles populaires utilisés.

En plus de ces critères, on a également utilisé les tests suivants pour comparer aussi et montrer les performances de notre distribution.

Le test de Cramér-von Mises

La statistique de test donnée par(55) :

$$CM = \frac{1}{12n} + \sum_i^n \left[\frac{2i-1}{2n} - \hat{F}(x_i) \right] \quad (4.5)$$

Le test de Anderson-Darling

La statistique de test donnée par(5) :

$$AD = -n - \sum_i^n \left\{ \ln \hat{F}(x_i) + \ln \left[1 - \hat{F}(x_{n+1-i}) \right] \right\} \quad (4.6)$$

Le test de Kolmogorov-Smirnov

La statistique de test donnée par(38) :

$$KS = \sup \left| \frac{1}{n} \sum_i^n I \{x_i \leq x\} - \hat{F}(x) \right| \quad (4.7)$$

Ces tests résument essentiellement la différence entre les fonctions de distribution hypothétiques et empiriques.

En supposant que l'hypothese H_0 est : la distribution GE décrit correctement les données réelles.

si la p-value $< \alpha$ on accepte l'hypothèse nulle H_0

si la p-value $\geq \alpha$ on rejette l'hypothèse nulle H_0

Tous les tests donnent des valeurs de p qui sont donnés dans le tableau suivant :

	RTS50	CAC40
CM	0.235	0.273
AD	0.126	0.174
KS	0.146	0.198

Dans tous les cas on trouve que la p-value inferieur α , alors on accepte l'hypothèse H_0 , encore une fois les valeurs de p confirment que notre distribution GE décrit correctement les données réelles avec $\alpha = 5\%$.

Conclusion générale

Dans cette thèse, on s'est intéressé à la théorie des valeurs extrêmes, on a modélisé des outils et des techniques pour identifier les distributions de la TVE. L'étude de ces valeurs extrêmes repose sur l'analyse des plus grandes observations d'un échantillon, cette analyse se fait selon deux approches, la première, qui est le comportement asymptotique du maximum appelée la méthode (BM) qui modélise la distribution des extrêmes par les lois GEV et l'autre méthode (POT) modélise la distribution des excès au-dessus d'un seuil élevé par les lois GPD. Ainsi, on a réussi à introduire une nouvelle distribution "GE" avec une discussion sur les caractéristiques de ces propriétés, on a également illustré cette distribution par une application sur des données réelles des indices boursiers "CAC40" et "RTS50" utilisées par certains chercheurs. On a comparé notre distribution GE avec d'autres modèles de valeurs extrêmes à l'aide des outils statistiques, on conclut que la GE est un bon modèle de distribution pour la prédiction et peut être très flexible pour générer ces données, qui ajoutera une certaine valeur à la littérature existante sur la modélisation des données de survie. Par la suite, il serait intéressant d'utiliser des données réelles dans d'autres domaines d'applications. Pour nos futures travaux de recherches, on envisage de proposer et d'étudier d'autre distribution dans la théorie des valeurs extrêmes.

References

- Anwar Hassan , Ishfaq S. Ahmad and Peer Bilal Ahmad :A New Mixed Generalized Negative Binomial Distribution, arXiv:1905.12852v (2019).
- Alzaatreh,F.Famoye,C.Lee :A new method for generating families of continuous distribution.Metron 71:63-79.
- A. Alzaatreh,F.Famoye,C, Lee :Exponentiated T-X family of distribution with some application,Int J Prob Stat 2,31 OCo49, (2013b).
- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–723.
- Anderson TW, Darling DA (1954) A test of goodness of fit. J Am Stat Assoc 49:765–769.
- Balkema, A. A., De Haan, L., 1974. Residual life time at great age. The Annals of probability, 792–804.
- Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2006a. Statistics of extremes : theory and applications. John Wiley & Sons.
- Beirlant, J., Vynckier, P., Teugels, J. L., et al., 1996b. Excess functions and estimation of the extreme-value index. Bernoulli 2 (4), 293–318.
- Bozdogan H (1987) Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. Psychometrika 52:345–370.
- BETTEKA Samah. les varuers extrêmes bivariees, Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie, département de Mathématiques, Biskra 2017.

- Beirlant, J., Goegebeur, Y., Segers, J. & Teugels, J., 2004, *Statistics of Extremes - Theory and Applications*, Wiley, England.
- Beirlant, J., Matthys, G. (2003). Estimating the extreme value index and high quantiles.
- Brahimi, B. (2010). *Statistics of Bivariate Extreme Values*. Thèse de doctorat, Université de Biskra, Algérie.
- de Wolf, P.P. (1999). *Estimating the Extreme Value Index*. Thèse de doctorat, Delft University of Technology.
- Beirlant, J., Goegebeur, Y., Segers, J. & Teugels, J., 2004, *Statistics of Extremes - Theory and Applications*, Wiley, England
- Chicheportiche, R. Bouchaud, J. P. 2012. Weighted Kolmogorov-Smirnov test : Accounting for the tails. *Phys. Rev. E*. Vol 86. Issue 4.
- Coles, S., Bawa, J., Trenner, L., Dorazio, P., 2001. *An introduction to statistical modeling of extreme values*. Vol. 208. Springer.
- Csrrgoo, S., Mason, D. M., 1985. Central limit theorems for sums of extreme values. In *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 98. Cambridge University Press, pp. 547–558.
- Cordeiro, GM, Nadarajah, S, Ortega, EMM :The Kumaraswamy Gumbel distribution. *Statistical Methods and Applications*, 2012 21(2), 139–168 (2011).
- de Haan, L. and Ferreira, A., 2006. *Extreme Values Theory : An introduction*. Springer.
- Danielsson, J., Jansen, D. W., De vries, C. G., 1996. The method of moments ratio estimator for the tail shape parameter. *Communications in Statistics-Theory and Methods* 25 (4), 711–720.
- de Haan, L., Ferreira, A., 2007. *Extreme value theory : an introduction*. Springer Science & Business Media.

- Dekkers, A. L., De Haan, L., et al., 1989a. On the estimation of the extreme-value index and large quantile estimation. *The Annals of Statistics* 17 (4), 1795–1832.
- Dekkers, A. L., Einmahl, J. H., De Haan, L., 1989b. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, 1833–1855.
- Diebolt, J., Guillou, A., Rached, I., 2007. Approximation of the distribution of excesses through a generalized probability-weighted moments method. *Journal of Statistical Planning and Inference* 137 (3), 841–857.
- Diebolt, J., Guillou, A., Rached, I., 2007. Approximation of the distribution of excesses through a generalized probability-weighted moments method. *Journal of Statistical Planning and Inference* 137 (3), 84.
- Drees, H., Kaufmann, E., 1998. Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and their Applications* 75 (2), 149–172.
- Dekkers, A.L.M. et de Haan, L. (1989). On the Estimation of the Extreme-Value Index and Large Quantile Estimation. *Annals of Statistics* 17, 1795-183.
- Drees, H., Ferreira, A. and de Haan, L. (2002). On maximum likelihood estimation of the extreme value index. To appear in *Ann. Appl. Probab.*
- Drees, H., Ferreira, A. and de Haan, L. (2002). On maximum likelihood estimation of the extreme value index. To appear in *Ann. Appl. Probab.*
- Dehaimi,S, Hadji,m,l, 2020. Gumbel-Exponential distribution. *Italian Journal of Pure and Applied Mathematics. Laps Laboratoir, Badji Mokhtar University, Box 12, Annaba, 23000, Algeria (soumet).*
- El Adlouni, S.-E., Bobée, B., Ouarda, T. B., 2007. Caractérisation des distributions à queue lourde pour l’analyse des crues. No. R929. INRS, Centre Eau, Terre et Environnement.
- Embrechts, P., 2000. Extreme Value Theory : Potential and Limitations as an Integrated Risk Management Tool Derivatives Use, Trading & Regulation 6, 449-456.

- Embrechts, P., Klüppelberg, C., Mikosch, T., 2013. Modelling extremal events : fo insurance and finance. Vol. 33. Springer Science & Business Medi.
- Feller, W. (1971). An Introduction to Probability Theory and its Applications, 2nd edition. Wiley, New York..
- Fisher, R.A. et Tippett, L.H.C. (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.
- Jeena Joseph, K.K. Jose : Gumbel - Frechet Distribution and its Applications. International Journal of Mathematics and Computer Research, ISSN: 2320-7167 (2019).
- Juan M. Astorga, Yuri A. Iriarte, Héctor W. Gomez ,Heleno Bolfarine : Modified slashed generalized exponential distribution, journal of Communications in Statistic- Theory and Methods,ISSN: 0361-0926 (Print) 1532-415(2019)
- Kolmogorov A (1933) Sulla determinazione empirica di una legge di distribuzione. Giornale dellstituto Italiano degli Attuari 4:83–9.
- Hamedani, Haitham M. Yousof, Mahdi Rasekhi, Morad Alizadeh, Seyed Morteza Najibi : Type I General Exponential Class of Distributions, Pak.j.stat.oper.res, pp39-55 (2018).
- Hosking, J. et Wallis, J. (1987). Parameter and Quantile Estimation for the Gene ralized Pareto Distribution. Technometrics 29, 339-349.
- Hurvich CM, Tsai C-L (1989) Regression and time series model selection in small samples. Biometrika 76:297–307.
- Hannan EJ, Quinn BG (1979) The determination of the order of an autoregression. J R Stat Soc B 41:190–195.
- Hubert, P., Bendjoudi, H., 1996. Introduction à l'étude des longues séries.
- http://fr.wikipedia.org/wiki/Indice_RTS.

- Gilli, M., et al., 2006. An application of extreme value theory for measuring financial risk. *Computational Economics* 27 (2-3), 207–228.
- Gnedenko, B., 1943. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, 423–453.
- Gumbel, E. J. Les valeurs extrêmes des distributions statistiques. *Ann. Inst. Henri Poincaré* 5 (2), 115–15.
- Gardes, L., and Girard, S. (2013). Estimation de quantiles extrêmes pour les lois à queue de type Weibull : une synthèse bibliographique. *J. de la Société Française de Statistique*, 154(2), 98-118.
- K. Fathi1, S.F. Bagheri, M. Alizadeh, M. Alizadeh : A study of methods for estimating in the exponentiated Gumbel distribution, *Journal of Statistical Theory and Applications*, Vol. 16, No. 1 (March 2017) 81–95.
- M.Bourguignon,R.B.Silva,G.M.Corddeiro :The Weibull-G family of probability distribution,*J Data Sci* 12,53-68, (2014).
- Matthys, G., Beirlant, J., 2003. Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica*, 853–880.
- Matthys, G. et Beirlant, J. (2003). Estimating the Extreme Value Index and High Quantiles with Exponential Regression Models. *Statistica Sinica* 13, 853-880.
- M. R. Mahmoud, R. M. Mandouh and R. E. Abdelatty : Lomax-Gumbel (Frechet) A New Distribution, *JAMCS*.47331(2019).
- Morad Alizadeh, Thiago G. Ramires, S. M. T. K. MirMostafae, Maryam Samizadeh and Edwin M. M. Ortega : A new useful four-parameter extension of the Gumbel distribution: Properties, regression model and applications using the GAMLSS framework, *journal of communications in statistics simulation and computation*, ISSN: 0361-0918 (Print) 1532-4141 (2018).

86. Mises Von, Richard E (1928) *Wahrscheinlichkeit, Statistik und Wahrheit*. Julius Springer, Heidelberg.
- Neves, C. et Fraga Alves, M.I., (2004). Reiss and Thomas'Automatic Selection of the Number of Extremes. *Computational Statistics and Data Analysis*. 47, 689–704.
- Neves, C. and Fraga Alves, M.I., 2004. Reiss and Thomas'automatic selection of the number of extremes. *Comput. Statist. Data Anal.* 47(4), 689-704.
- Neves, C. et Fraga Alves, M.I., (2004). Reiss and Thomas'Automatic Selection of the Number of Extremes. *Computational Statistics and Data Analysis*. 47, 689–704.
- Neves, C. and Fraga Alves, M.I., 2004. Reiss and Thomas'automatic selection of the number of extremes. *Comput. Statist. Data Anal.* 47(4), 689-704.
- Nagaraja, H. N. et David, H. A. . *Order Statistics*. Third Edition. Wiley. Nadarajah, S. and Kotz, S :The beta Gumbel distribution . *Math. Probability. Eng.*, 10, 323-332.
- Persson, K. and RydÈn, J. (2010). Exponentiated Gumbel distribution for estimation of return levels of significant wave height. *Journal of Environmental Statistics*, 1(3): 1-12.
- Peng, L., 1998. Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters* 38 (2), 107–115.
- Pickands III, J., et al., 1975. Statistical inference using extreme order statistics. *the Annals of Statistics* 3 (1), 119–131.
- Pinheiro, E.C. and Ferrari, S.L.P. (2016). A comparative review of generalizations of the Gumbel extreme value distribution with an application to wind speed data. *Journal of Statistical Computing and Simulation*, 86(11): 2241-2261.
- Poon, S.H., Rockinger, M., Tawn, J., 2004. Extreme-value dependence in . . . nancial markets : diagnostics, models, and . . . nancial implications. *Rev. Financial Studies* 17, 581–610.

- Reiss, R.D. et Thomas, M. (1997). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhuser, Basel.
- Rubén Gomez González, M. Isabel Parra, Francisco Javier Acero, Jacinto Martín : An improved method for the estimation of the Gumbel distribution parameters, *Journal of Computational Statistics and Data Analysis* (2019).
- Rootzen, H. et Tajvidi, N. (1996). *Extreme Value Statistics and Wind Storm Losses : a Case Study*, *Scandinavian Actuarial Journal*, 70-94.
- Rinne, H., 2008. *The Weibull distribution : a handbook*. Chapman and Hall/CRC.
- Toulemonde, G., 2008. *Estimation et tests en théorie des valeurs extrêmes*. Ph.D. thesis, Université Pierre et Marie Curie-Paris VI.
- Smith, R., 1985, Maximum likelihood estimation in a class of non regular cases, *Biometrika*, 72, 6790.
- Schwarz GE (1978) Estimating the dimension of a model. *Ann stat* 6:461–464.
- Thiago Alexandro Nascimento , Frank Gomes-Silva and Luz Milena Zea : Mathematical properties, application and simulation for the exponentiated generalized standardized Gumbel distribution, *Acta Scientiarum. Technology*, v. 41, e39867, (2019).
- Toulemonde, G., 2008. *Estimation et tests en théorie des valeurs extrêmes*. Ph.D. thesis, Université Pierre et Marie Curie-Paris VI.
- www.investing.com.

وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR -ANNABA
UNIVERSITY
UNIVERSITE BADJI MOKHTAR
ANNABA



جامعة باجي مختار
- عنابة -

Faculté des Sciences

Année : 2021

Département de Mathématiques



THÈSE

Présentée en vue de l'obtention du diplôme de
Doctorat

Etude d'un Modèle dans la Théorie des Valeurs Extrêmes

Option :

Actuariat

Par

Dehaimi Samira

DIRECTEUR DE THÈSE : Dr Hadji Mohammed lakhdar U.B.M. ANNABA

Devant le jury

PRESIDENT :	Taallah Frekh	MCA	U.B.M. ANNABA
EXAMINATRICE :	Seddik Ameer Nacira	Prof	U.B.M. ANNABA
EXAMINATEUR :	Aissaoui Mohamed Zine	Prof	U. GUELMA
EXAMINATEUR :	Ezzebsa Abdelali	M.C.A	U. GUELMA