People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

Mohamed Khider University of Biskra

Faculty of Exact Sciences and Siences of Nature and Life

Department of Mathematics

Thesis Submitted in Partial Execution of the Requirements of the Degree of

Master in "**Applied Mathematics**"

Option : Statistics

By Nacer Nour Imane

Title :

Estimation of extreme value distribution

Examination Committee Members :

Pr.	BRAHIMI Brahim	U. Biskra	President
Dr.	BOUREDJI Hind	U. Biskra	Supervisor
Dr.	CHINE Amel	U. Biskra	Examiner

02/06/2025

Dedication

I dedicate this work to the lights of my life:

To the one whose name I carry with pride, my dear father Saifi,

To my first supporter and the pillar of strength in my journey, my beloved mother Louiza,

For their continuous support and endless love, which have been the source of my strength and inspiration at every moment.

To my sisters Oumima and Manar, and my brothers Farouk, Amjad, and Mohamed, for their love and constant support.

To the soul of my grandmother Mabarka, who inspires me every day.

To my second family, Grandfather Hussein and Grandmother Jumaa, and my aunts Noura and Sadia, for their warmth and care.

To my dear aunt Safia and my dear uncle Wahid, for their continuous support.

To my cousins Asaad, Razan, and my little Intisar, who bring joy to my life.

To my friends Bouchra ,Noor Al-Huda, Imane, Ibtissem,, and Salsabil, for their unwavering friendship and encouragement.

I dedicate this work to all of you with love and gratitude.

Acknowledgment

I would like to express my deepest gratitude and appreciation to Dr. Bouredji Hind my esteemed supervisor, for her continuous support and valuable guidance throughout my academic journey. She has been a true source of inspiration, and her expertise and insightful advice have greatly broadened my knowledge and constantly motivated me to give my best at every stage of this work.

Thank you, Dr. Bouredji Hind , for your unwavering support.

I would like to thank all my teachers who have taught me throughout my academic journey. Additionally, I would also like to thank the jury members, Professor BRAHIMI Brahim and Dr. CHINE Amel, for accepting to examine and evaluate this work. I also wish to thank all our teachers in the Department of Mathematics at Mohamed Khider University, who contributed to our education during the Bachelor's and Master's years.

To all those who have contributed to the realization of this memory,

Thanks

Abbreviations and notations

The different abbreviations and notations used throughout this thesis are explained below:

$i.i.d$	: Independent and identically distributed.
CDF	: Cumulative distribution function note $F(.)$.
PDF	: Probability density function also density function note $f(.)$.
$\bar{F}(.)$	: Complementary cumulative distribution function.
$X_{k,n}$	: k^{th} statistical order.
$f_{k,n}(x)$	: pdf of the order statistic.
$F_{k,n}(x)$	: cdf of the order statistic.
$F_n(.)$	: Empirical (or sample) distribution function.
$\overleftarrow{F}(x)$	: Generalized inverse of a function F.
GEV	: Generalized Extreme Value.
GPD	: Generalized Pareto distribution.
$L(\theta \vert X)$	: Likelihood function.
MLE	: Maximum likelihood estimator.
LM	: L-moments
$\mathbf{I}_{\{X_i \leq x\}}$	: Indicator function
$\underset{n\to+\infty}{\overset{D}{\to}}$	: Convergence in probability
$\underset{n\to+\infty}{\overset{a.s}{\to}}$	: Almost sure convergence

Contents

List of Figures

List of Tables

Introduction

Extreme Value Theory (EVT), which emerged between 1920 and 1940 thanks to the work of scientists Fréchet, Fisher, Tippett, Gumbel, and Gnedenko, is considered one of the fundamental branches of mathematical statistics. It focuses on the analysis of rare events that have a low probability of occurrence. Despite their rarity, these events can have significant impacts, particularly in fields such as finance, insurance, and environmental science, where they may lead to severe losses or sudden changes.

Extreme events can result in significant human and material losses. While such disasters cannot always be prevented, societies can implement preventive measures to mitigate their impact. One valuable tool in this effort is the statistical theory of extreme values, which offers broadly applicable and insightful results for understanding and managing rare events.

In this context, EVT plays a crucial role in the statistical modeling of such phenomena by focusing on the behavior of distributions in the tails, i.e., at very large or very small values.

The importance of EVT lies in its ability to provide a reliable method for predicting rare events, even beyond the range of previously observed data (Embrechts et al., 1997). The most advanced statistical methods for extreme events will be studied both from the theoretical and application sides.

This master's thesis is divided into three chapters. In the first chapter, we will present an overview of the fundamental concepts related to Extreme Value Theory (EVT). This includes the definition of order statistics and record values (both upper and lower), their properties, and their distributions in the continuous case. We also review key theoretical foundations such as the Fisher-Tippett theorem (1928), domains of attraction, and the Generalized Pareto Distribution, which is considered a core tool for modeling tail behavior.

In the second chapter, we focus on the estimation of the tail index, a central quantity in Extreme Value Theory (EVT). In this context, we introduce and study several estimation methods, including semi-parametric estimators such as the Hill, Pickands, and Moment estimators, as well as the parametric approach of the Maximum Likelihood Estimator (MLE) and L-moment. Each method is analyzed separately with regard to its definition and theoretical properties.

Finally, the third chapter is devoted to the practical part of the creation of our application using the R software.

Chapter 1

Extreme value

This chapter introduces key concepts in extreme value theory. We begin with order statistics, followed by fundamental notations. Then, we present the Fisher-Tippett theorem, which characterizes the limiting distributions of extreme values. Next, we discuss the domain of attraction, which classifies distributions based on their asymptotic behavior. Finally, we examine the Pareto distribution, a crucial model for heavy-tailed data. These concepts form the theoretical foundation for the methods developed later.

1.1 Order statistics

The use of extreme value theory relies on properties of order statistics and extrapolation techniques. In other words, it is based on the law of large numbers for the maxima of properly renormalized random variables. Let $X_{1,\ldots,}X_n$ be a sequence of independent and identically distributed $(i.i.d)$ random variables following the distribution F_X and the associated density f_X.

1.1.1 Definitions and Notations

Definition 1.1.1 *The order statistic of a sample* $(X_1, ..., X_n)$ *refers to the sequence obtained by rearranging the values in increasing order, which is denoted by* $(X_{(1,n)}, ..., X_{(n,n)})$. *In other words, we have* $X_{(1,n)} \leq X_{(2,n)} \leq ... \leq X_{(n,n)}$. *For* $1 \leq i \leq n$, *we have:* $X_{i,n}$ *is the* i^{th} *order statistic. The minimum and maximum of an i.i.d sample of size* n *represent extreme values are,*

$$X_{(1,n)} = \min(X_{1,...,}X_n)$$

and,

$$X_{(n,n)} = \max(X_{1,...,}X_n)$$

1.1.2 Distributions of order statistics

Theorem 1.1.1 *The cumulative distribution function (CDF) of the order statistic* $X_{i,n}$ *is provided for all* $x \in \mathbb{R}$ *as described by David (*1970*) and Balakrishnan and Clifford Cohen (*1991*).*

$$F_{i,n}(x) = F_{X_{i,n}}(x) = P(X_{i,n} \leq x) = \sum_{r=i}^{n} C_n^r [F(x)]^r [1 - F(x)]^{n-r}, x \in \mathbb{R}.$$

the density of $X_{i,n}$ *is expressed as*

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

It can therefore be concluded that for the minimum statistic, the CDF and the PDF are, respectively

$$F_{1,n}(x) = F_{X_{1,n}}(x) = 1 - [1 - F(x)]^n$$

and,

$$f_{1,n}(x) = f_{X_{1,n}}(x) = n\left[1 - F(x)\right]^{n-1} f(x)$$

for the maximum statistics, we have:

$$F_{n,n}(x) = F_{X_{n,n}}(x) = \left[F(x)\right]^{n}$$

and,

$$f_{n,n}(x) = f_{X_{n,n}}(x) = n\left[F(x)\right]^{n-1} f(x)$$

Proof. Using the independence of the random variables $(X_{1,\ldots,}X_n)$ we can conclude that,

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

and,

$$F_{n,n}(x) = P\left(X_{n,n} \leq x\right) = \prod_{i=1}^{n} P\left(X_i \leq x\right) = \left[F(x)\right]^{n},$$

and we deduce the densities,

$$f_{1,n}(x) = f_{X_{1,n}}(x) = n\left[1 - F(x)\right]^{n-1} f(x)$$

$$f_{n,n}(x) = f_{X_{n,n}}(x) = n\left[F(x)\right]^{n-1} f(x)$$

■

1.2 Fundamental notions

1.2.1 Central Limit Theorem (C.L.T)

Definition 1.2.1 *If $X_1, ..., X_n$ are independent and identically distributed random variables (i.i.d.) with a mean μ and variance σ^2 ,then:*

$$\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \underset{n \to +\infty}{\overset{D}{\to}} N(0,1)$$

1.2.2 Empirical distribution function

Definition 1.2.2 *The empirical distribution function of the sample $(X_{1,...,}X_n)$ is estimated using order statistics as follows,*

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}_{\{X_i \leq x\}} = \begin{cases} 0 & if\ x < X_{1,n} \\ \frac{i-1}{n} & if\ X_{i-1,n} \leq x \leq X_{i,n},\ 2 \leq i \leq n \\ 1 & if\ x \geq X_{n,n} \end{cases}$$

1.2.3 Regular variations

The concept of a regularly varying function plays an important role in characterizing the domains of attraction in extreme value theory. Here, we present some key results and for further details, we refer to the work of Bingham et al. (1987).

Definition 1.2.3 *A function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is said to be regularly varying at infinity if and only if there exists a real number α such that, for all $x > 0$,*

$$\lim_{t \to \infty} \frac{f(tx)}{f(t)} = x^{\alpha}$$

We write $f \in RV_\alpha$, where α is called the index (or exponent) of the regularly varying function f.

Remark 1.2.1 *When $\alpha = 0$, we obtain the case of the slowly varying function, as defined below.*

Proposition 1.2.1 *A function l is said to be slowly varying if $l(t) > 0$ for sufficiently large t and if, for all $x > 0$, we have:*

$$\lim_{t\to\infty} \frac{l(tx)}{l(t)} = 1$$

Proposition 1.2.2 *It is said that the function $g(x)$ is of regular variation of order α if and only if,*

$$\exists \alpha > 0 : g(x) = x^\alpha l(x)$$

1.2.4 Terminal point

Definition 1.2.4 *We denote by x_F the upper extreme point of the distribution F (i.e. the largest possible value for $X_{i,n}$, can take the value $+\infty$). Then, the terminal point of a function F is,*

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

Proposition 1.2.3 *We have when $n \to \infty$, $X_{n,n} \overset{a.s}{\to} x_F$.*

1.2.5 Survival Function

We assume that X is a continuous random variable with probability density function $f(x)$ and cumulative distribution function $F(x) = P(X < x)$. It is often useful to work with the complement of the cumulative distribution function, which is the survival function $\bar{F}(x)$, defined as follows,

$$\bar{F}(x) = 1 - F(x) = P\left(X > x\right) = \int_{t}^{\infty} f(t)dt$$

Remark 1.2.2 *Every survival function $\bar{F}(x)$ is monotonically decreasing.*

1.2.6 Generalized inverse

The generalized inverse remains defined even when F is not bijective either because this function is discontinuous or because it is constant on intervals with non-empty interior.

Definition 1.2.5 *The generalized inverse of a function F is the application defined by,*

$$\overleftarrow{F}(y) = \inf\{x \in \mathbb{R}, F(x) \geq y\}.$$

1.2.7 Quantile function and tail quantile

The quantile function of the distribution function F is the generalized inverse function of F defined by,

$$Q(y) = F^{\leftarrow}(y) = \inf\{x \in \mathbb{R}, F(x) \geq y\}$$

In extreme value theory, a function denoted by U and called the tail quantile function is defined by:

$$U(t) = Q(1 - \frac{1}{t}) = (\frac{1}{\bar{F}})^{\leftarrow}(t), \qquad 1 < t < \infty$$

1.3 Fisher-Tippet theorem (1928)

The extreme value theorem, according to Gnedenko (1943), specifies the form of the limiting distribution for the maxima or minima of samples when the observation period for the extremes becomes infinite. This theorem relies on four fundamental assumptions: independence, the same distribution for the variables, and the existence of a sequence of coefficients to normalize the random variable.

The limiting distribution can take three possible forms: the Gumbel distribution (Type I), the Fréchet distribution (Type II), and the Weibull distribution (Type III), each with specific characteristics depending on the distribution of the extremes.

Theorem 1.3.1 (Fisher's Theorem) *Under certain regularity conditions on the distribution function F, there exists a value $\gamma \in \mathbb{R}$ as well as two sequences of real normalizing constants $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ such that, for every $x \in \mathbb{R}$,*

$$\lim_{n \to +\infty} P\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) = G_\gamma(x)$$

where $G_\gamma(x)$ is a distribution function associated with a limiting distribution, with,

- If $\gamma > 0$ (Frechet law)

$$G_\gamma(x) = \phi_\gamma(x) = \begin{cases} 0 & \text{if } \; x \leq 0 \\ \exp[-(x)]^{-\frac{1}{\gamma}} & \text{if } \; x > 0 \end{cases}$$

- If $\gamma < 0$ (Weibull law)

$$G_\gamma(x) = \psi_\gamma(x) = \begin{cases} \exp[-(-x)]^{-\frac{1}{\gamma}} & \text{if } \; x < 0 \\ 1 & \text{if } \; x \geq 0 \end{cases}$$

- If $\gamma = 0$ (Gumbel law)

$$G_0(x) = \Lambda(x) = \exp[-\exp(-x)] \text{ for all } x \in \mathbb{R}$$

1.4 Generalized Extreme Value (GEV) Distribution

Jenkinson in 1995 combined the three families of extreme value distributions introduced by Fisher and Tippett in 1928 into a generalized version of the probability distributions which is called the generalized extreme value (GEV) distribution. The distribution function of GEV is written as

$$H(x) = \begin{cases} \exp\left(-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\left(\frac{1}{\gamma}\right)}\right) & \text{For } \gamma \neq 0,\ 1+\gamma\left(\frac{x-\mu}{\sigma}\right) > 0 \\ \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right) & \text{For } \gamma = 0 \end{cases} \tag{1.1}$$

where γ is called extreme values index. Then, we can easily show that the density function of the GEV distribution is

$$h(x) = \begin{cases} \frac{1}{\sigma}\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\left(\frac{1+\gamma}{\gamma}\right)} \exp\left(-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\left(\frac{1}{\gamma}\right)}\right) & \text{For } \gamma \neq 0 \\ \frac{1}{\sigma}\exp\left(-\left(\frac{x-\mu}{\sigma}\right) - \exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right) & \text{For } \gamma = 0 \end{cases}$$

1.5 Domains of attraction

Definition 1.5.1 *A distribution F is said to be in the domain of attraction of G_γ, written as $F \in D(G_\gamma)$, if $\exists$ ($a_n) > 0$ and $(b_n) \in \mathbb{R}$ such that:*

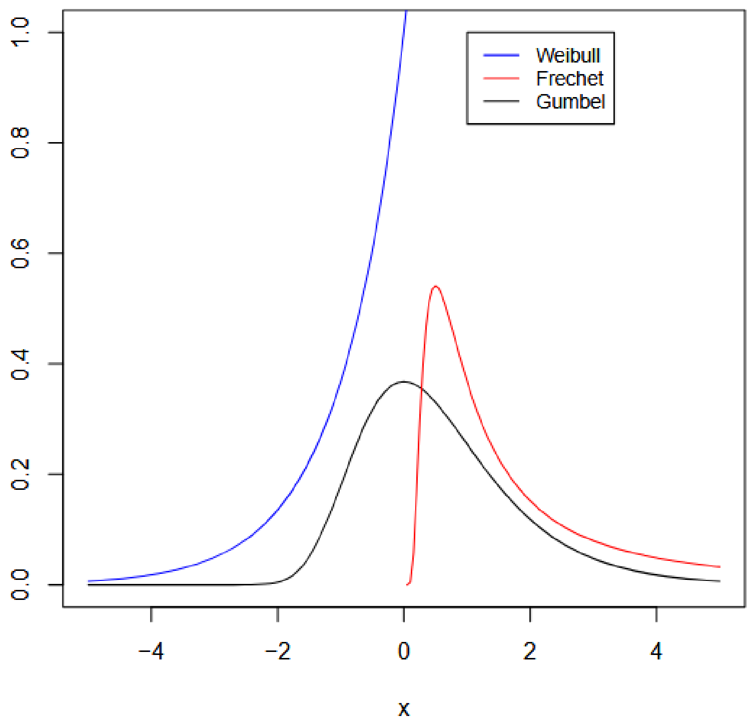

Figure 1.1: Densities of the standard extreme value distributions.

$$\lim_{n\to+\infty} P\left(\frac{X_{n,n}-b_n}{a_n}\leq x\right)=\lim_{n\to+\infty} F(a_n x+b_n)=G_\gamma(x)$$

1.5.1 Fréchet Attraction Domain

This concept is attributed to the French mathematician Maurice Fréchet $(1878-1973)$, who played a significant role in the development of the extreme value limit theory. Recall that the Fréchet domain of attraction contains heavy-tailed laws or Pareto-type laws. The laws in this domain have an infinite terminal point. Any function belonging to the Fréchet domain of attraction is a function with regular variations.

Definition 1.5.2 *A distribution function $F(x)$ belongs to the domain of attraction of the Fréchet distribution if and only if it can be written in the form,*

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

where $l(x)$ is a slowly varying function. In this case, the normalizing sequences(a_n) and (b_n) are given for all $n > 0$ by:

$$a_n = F^{-1}(1 - \frac{1}{n}),\ b_n = 0.$$

1.5.2 Weibull Attraction Domain

This concept emerged in the context of Extreme Value Theory and is primarily associated with the work of Swedish mathematician Ernst Hjalmar Waloddi Weibull. In 1939, Weibull presented his notable research on the Weibull distribution in the field of probability theory and statistics. A distribution function belongs to the domain of attraction of the Weibull distribution if and only if its terminal point is finite.

Definition 1.5.3 *A distribution function $F(x)$ belonging to the Weibull attraction domain is written as follows,*

$$F(x) = 1 - (x_F - x)^{\frac{-1}{\gamma}} l((x_F - x)^{-1}), \quad l(.) \in RV_0 \textit{ for } x \leq x_F$$

The normalization sequences (a_n) and (b_n) are given by:

$$a_n = \ F^{-1}(1) - F^{-1}(1 - \frac{1}{n}), \quad b_n = \ F^{-1}(1)$$

and if the distribution function $F_*(.)$, defined by:

$$F_*(x) = \begin{cases} 0 & \textit{if } x < 0 \\ F(x_F - \frac{1}{x}) & \textit{if } x \geq 0 \end{cases}$$

belongs to the Fréchet attraction domain.

1.5.3 Gumbel Attraction Domain

This concept is attributed to the mathematician Emil Julius Gumbel , who was one of the pioneers in developing Extreme Value Theory. In 1958, Gumbel published his famous book "Statistics of Extremes",which became one of the key references in this field. The Gumbel domain of attraction contains the laws whose survival function is exponentially decreasing, i.e. the laws with light tails.

Definition 1.5.4 *A distribution function $F(x)$ belongs to the domain of attraction of Gumbel if and only if there exists $t < x_F$ such that $F(x)$ can be written in the following form,*

$$\bar{F}(x) = c(x) \exp\left(-\int_t^x \frac{1}{\alpha(u)} du\right),\ \{t < x \leq x_F\}$$

where $c(x) \to c > 0$ as $x \to x_F$, and $\alpha(x)$ is a positive and differentiable function, with $\alpha^{'}(x) \to 0$ as $x \to x_F$. In this case, a possible choice for the sequences (a_n) and (b_n) for all $n > 0$ is as follows,

$$a_n = F^{-1}(1 - \frac{1}{n \exp(1)}) - F^{-1}(1 - \frac{1}{n}), b_n = F^{-1}(1 - \frac{1}{n})$$

The table below give different examples of standard distributions in these three domains of attraction.

	Domain of attraction		
	Fréchet $\gamma > 0$	Gumbel $\gamma = 0$	weibull $\gamma < 0$
Law	Burr Fréchet Loggamma Loglogistic Pareto	Gamma Gumbel Logistic Lognormal Weibull	Uniform Reverse Burr

Table 1.1: Domains of attraction of common distributions

Example 1.5.1 *Let the distribution function of the exponential distribution with parameter*

$\lambda > 0$ *be:*

$$F(X) = 1 - exp(-\lambda x), x \geq 0$$

We define the normalization sequences .$a_n = \frac{1}{\lambda}$ *and* $b_n = \frac{1}{\lambda}\ln(n)$*. The normalized distribution function is given by:*

$$\begin{aligned} F^n(a_n x + b_n) &= [(1 - \exp(-\frac{1}{\lambda}\lambda x - \lambda\frac{1}{\lambda}\ln(n))]^n \\ &= \left(1 - \frac{\exp(-x)}{n}\right)^n \to \exp(-\exp(-x)) = \Lambda(x) \end{aligned}$$

1.6 Generalized Pareto Distribution

The generalized Pareto distribution (GPD) is widely used in engineering, environmental science, and finance to model low-probability events. Typically, the GPD is used to estimate extreme values, such as the 99^{th} percentile of a specific event.

The probability density function (PDF) of the Generalized Pareto distribution GPD is given by,

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

with $\sigma > 0$.

The cumulative distribution function is given,

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

Definition 1.6.1 *Let X be a random variable with a cumulative distribution function F and* *a terminal point x_F. For all $u < x_F$, the function $F_u(x)$ is defined as:*

$$F_u(x) = P[X - u \leq x | X > u], \text{ for } x \geq 0.$$

It is called the distribution function of the excess above the threshold u.

Definition 1.6.2 *The mean function of the excesses of the random variable X with respect to the threshold $u < x_F$, and we denote it by $e(u)$, de ned by:*

$$\forall u < x_F \quad e(u) = E\left(X - u | X > u\right) = \frac{1}{\overline{F}\left(u\right)} \int_u^{x_F} \overline{F}\left(t\right) dt \tag{1.2}$$

Remark 1.6.1 *By the definition of conditional probabilities, F_u can also be expressed as:*

$$F_u(x) = \begin{cases} \frac{F(u+x)-F(u)}{1-F(u)} & \text{if } x \geq 0, \\ 0 & \text{otherwise} \end{cases}.$$

Theorem 1.6.1 (Pickands, 1975) *A cumulative distribution function F belongs to the attraction domain D_γ if and only if there exists a positive function $\sigma(.)$ and a real number γ such that the excess distribution F_u can be uniformly approximated by a Generalized Pareto Distribution (GPD) denoted $G_{\sigma,\gamma}$.This can be formulated by the following relation:*

$$\underset{u \to x_F}{lim} \underset{x \in (0, x_F - u)}{sup} |F_u(x) - G_{\sigma(u),\gamma}(x)| = 0,$$

Example 1.6.1 *For the exponential distribution with parameter $\lambda = 1$, the cumulative distribution function is given by $F(x) = 1 - exp(-x)$ for $x \geq 0$ .By setting $\sigma = 1$, we have*

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

Simplifying, this gives:

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

Thus, the limiting distribution obtained is the Generalized Pareto Distribution (GPD) with parameter 1.

Chapter 2

Estimation of extreme value

The estimate of the tails index, plays an important role in limiting an extreme law, when it exists, is indexed by a parameter called extreme value index, there are two methods for estimating the extreme value index : parametric methods, meaning that the data follow an exact GEV distribution, and semi-parametric methods, where the parameter has both a finite-dimensional and an infinite-dimensional and are therefore based on partial properties of the underlying distribution, such as the Hill estimator, a widely used method for estimating the tail index of heavy-tailed distributions, next, the Pickands estimator, which provides an alternative approach based on order statistics and is particularly useful in extreme value analysis, finally, the Moment estimator, which leverages higher-order moments to estimate tail indices. Each of these estimators plays a crucial role in statistical modeling and risk assessment in various fields.

2.1 Parametric estimators

2.1.1 Maximum Likelihood Estimators

The Maximum Likelihood Estimation (MLE) method is one of the most widely used techniques for deriving estimators.

Principle of the Maximum Likelihood Method

Definition 2.1.1 *Suppose that $X_{1,\ldots,}X_n$ is an independent and identically distributed sample from a population with a probability density function or probability mass function given by $f(x|\theta_1, \ldots, \theta_k)$. The likelihood function is then defined as:*

$$L(\theta|X) = L(\theta_1, \ldots, \theta_k|x_{1,\ldots,}x_n) = \prod_{i=1}^{n} f(x_i|\theta_1, \ldots, \theta_k)$$

Definition 2.1.2 *Let $\hat{\theta}$ be the parameter value at which the likelihood function $L(\theta|x)$ attains its maximum when treated as a function of θ, the maximum likelihood estimator (MLE) of the parameter θ based on a sample X is denoted by $\hat{\theta}$.*

Proposition 2.1.1 *In most cases, the likelihood function is expressed as a product, so $\hat{\theta}$ will generally be calculated by maximizing the log-likelihood. When $\theta = (\theta_1, \ldots, \theta_k)$ and all the partial derivatives below exist, $\hat{\theta}$ is the solution to the system of equations:*

$$\frac{\partial}{\partial \theta_j} \log L(\theta_j|x_{1,\ldots,}x_n) = 0, \quad \forall j \in (1, 2, \ldots)$$

Application of the Maximum Likelihood Method to GEV

Proposition 2.1.2 *The log-likelihood function is given by:*

For $\gamma \neq 0$

$$\log(L(\gamma|X) = -n \log(\sigma) - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{n} \log\left(1 + \gamma \frac{X_i - \mu}{\sigma}\right) - \sum_{i=1}^{n} \left(1 + \gamma \frac{X_i - \mu}{\sigma}\right)^{-\frac{1}{\gamma}}$$

For $\gamma = 0$

$$\log(L(\gamma|X) = -n\log(\sigma) - \sum_{i=1}^{n}\exp\left(\frac{X_i - \mu}{\sigma}\right) - \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)$$

2.1.2 Method of L-moments

The L-moments, whose theory was unified by Hosking (1990), are linear combinations of order statistics and exhibit lower sample variances and higher robustness to outliers compared to the conventional moments.

Analogously to the classical method of moments, the L-moment estimation method derives parameter estimates by equating the first sample L-moments to their corresponding theoretical expressions. Theoretical L-moments are defined using the quantile function (i.e., the inverse of the cumulative distribution function).

Principle of the L-moments Method

Definition 2.1.3 *Let* $X_1, X_2, .., X_n$ *be a sample of size* n *from a continuous distribution* $F_X(x)$ *with the quantile function* $Q(u) = F_X^{-1}(u)$*, and let* $X_{(1,n)}, X_{(2,n)}, ..., X_{(n,n)}$ *be the order statistics associated with this sample, for* $r \geq 1$ *the* r^{th} *L-moments* l_r *is given by:*

$$l_r = \frac{1}{r}\sum_{i=0}^{r-1}(-1)^i\binom{r-1}{i}E\left(X_{(r-i,r)}\right), \qquad r = 1, 2, ... \tag{2.1}$$

With

$$E(X_{(j,r)}) = \frac{r!}{(j-1)!(r-j)!}\int x(F(x))^{(j-1)}(1 - F(x))^{(r-j)}dF(x) \tag{2.2}$$

The L in "L-moments" emphasizes that l_r *is a linear function of the expected order statistics. The first few L-moments are*

The first L-moment is used to calculate the mean (position measurement) and is defined by:

$$l_1 = E(X_{(1,1)})$$

The second L-moment is used to calculate (measure of dispersion) is given by:

$$l_2 = \frac{1}{2}E(X_{(2,2)} - X_{(1,2)})$$

The third L-moment for studying symmetry (skewnes measure) is given by:

$$l_3 = \frac{1}{3}E(X_{(3,3)} - 2X_{(2,3)} + X_{(1,3)})$$

The fourth L-moment to study kurtosis (kurtosis measure) is defined by:

$$l_4 = \frac{1}{4}E(X_{(4,4)} - 3X_{(3,4)} + 3X_{(2,4)} - X_{(1,4)})$$

L-skewness, τ_3, can then be found by taking the ratio of l_3 to l_2; i.e.

$$\tau_3 = \frac{l_3}{l_2}$$

The approximation for L-moments and L-skewness which are known as sample L-moments and sample L-skewness respectively, can be found from a finite sample of size n, arranged in ascending order. The sample L-moments, λ, are written as

$$\lambda_{j+1} = \sum_{k=0}^{J} (-1)^{j-k} \binom{j}{k} \binom{j+1}{k} b_k, \quad j = 0, 1, ..., n-1 \tag{2.3}$$

with,

$$b_k = \frac{1}{n} \binom{n-1}{k}^{-1} \sum_{i=k+1}^{n} \binom{i-1}{k} X_{i,n}$$

Thus, the sample L-skewness, t_3, can then be found by taking the ratio of λ_3 to λ_2 such that

$$t_3 = \frac{\lambda_3}{\lambda_2}$$

Application of the L-moment method to GEV

By substituting equation 1.1 in equation 2.2 and equating the population L-moments in equation 2.1 to their corresponding sample L-moments in equation 2.3, the parameter estimates for the GEV distribution can be found as follows,

$$\hat{\gamma} = 7.8590c + 2.9554c^2$$

with $c = \dfrac{2}{(3+t_3)} - \dfrac{\log(2)}{\log(3)}$

$$\hat{\sigma} = \frac{l_2 \hat{\gamma}}{\left(1 - 2^{-\hat{\gamma}}\right)\left(\Gamma\left(1-\hat{\gamma}\right)\right)}$$

$$\hat{\mu} = l_1 - \frac{\hat{\sigma}}{\hat{\gamma}}\left[1 - \Gamma\left(1+\hat{\gamma}\right)\right]$$

with Γ typically refers to the Gamma function.

2.2 Semi-parametric estimators

2.2.1 Hill Estimator

The Hill estimator was introduced by Hill (1975) as a nonparametric method for estimating the tail parameter of distributions belonging to the Fréchet domain of attraction. To construct this estimator, Hill applied the maximum likelihood method to the k largest observations in a sample. A significant number of theoretical studies have been dedicated to investigating its properties. Mason (1982) demonstrated its weak consistency, while Deheuvels, Haeusler, and Mason established its strong consistency in Deheuvels et al. (1988). The asymptotic normality of the estimator was studied by Davis and Resnick (1984), Csörgö and Mason (1985), Haeusler and Teugels (1985), and Smith (1987), among others.

Theorem 2.2.1 *Let $F \in D(G_\gamma)$ for $\gamma > 0$ if and only if,*

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

Theorem 2.2.2 *Gives an equivalent form of this condition:*

$$\lim_{t\to\infty} \frac{\int_t^\infty (1-F(x))\frac{dx}{x}}{1-F(t)} = \gamma$$

Now partial integration yields

$$\int_t^\infty (1-F(x))\frac{dx}{x} = \int_t^\infty (\log u - \log t)dF(u).$$

Hance we have

$$\lim_{t\to\infty} \frac{\int_t^\infty (\log u - \log t)dF(u)}{1-F(t)} = \gamma$$

To develop an estimator based on this asymptotic result, we replace in the last equation the parameter t with the intermediate order statistic $X_{n-k,n}$, and the distribution function F with the empirical distribution function F_n. Thus, we obtain the estimator proposed by Hill in 1975, denoted by $\hat{\gamma}_H$,defined by

$$\hat{\gamma}_H = \frac{\int_{X_{n-k,n}}^\infty (\log u - \log X_{n-k,n})dF(u)}{1-F(X_{n-k,n})}$$

or

$$\hat{\gamma}_H = \frac{1}{k}\sum_{i=1}^{k} [\log X_{n-i+1,n} - \log X_{n-k,n}]$$

Proof. This can be interpreted as follows:

$$P(\frac{X}{t} > x|X > t) \to x^{-\frac{1}{\gamma}}$$

as $t \to \infty$ and $x > 1$.

Moreover, if we define $Y_j(t)$ as the relative excesses beyond t, that is:

$$Y_j(t) = \frac{X_i}{t}, \text{with} \quad X_i > t$$

where i corresponds to the index of the j-th excess in the original sample, and $j = 1, ..., N_t$, then N_t represents the total number of excesses above t.By constructing the log-likelihood based on the excesses $Y_1(t), ..., Y_{N_t}(t)$ conditionally on N_t we obtain:

$$\log L(Y_1(t), ..., Y_{N_t}(t)) = -N_t \log \gamma - (1 + \frac{1}{\gamma})\sum_{j=1}^{N_t} \log Y_j(t)$$

Thus, the derivative of the log-likelihood with respect to γ is given by:

$$\frac{d \log L}{d\gamma} = -\frac{N_t}{\gamma} + \frac{1}{\gamma^2}\sum_{j=1}^{N_t} \log Y_j(t) = 0$$

Solving this equation, we obtain the estimator of γ in the following form:

$$\gamma = \frac{1}{N_t}\sum_{j=1}^{N_t} \log Y_j(t)$$

For an excess threshold t defined as the order statistic $X_{n-k_n,n}$ where $1 \leq k \leq n$ and $k \to \infty$, and replacing N_t with k , we obtain Hill's estimator (1975):

$$\hat{\gamma}_H = \frac{1}{k}\sum_{i=1}^{k} [\log X_{n-i,n} - \log X_{n-k,n}]$$

■

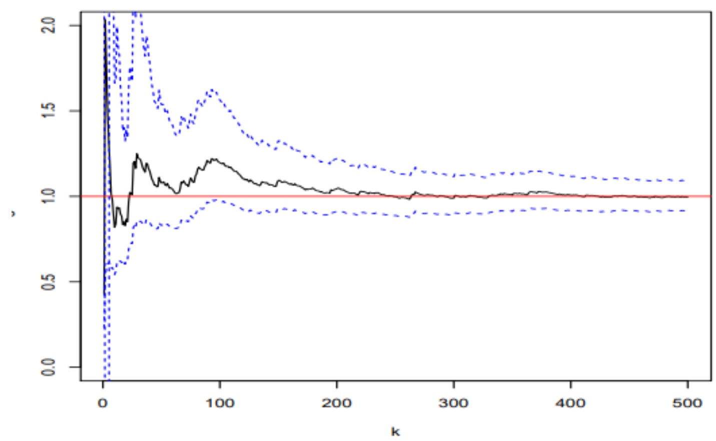

Figure 2.1: Hill estimator

Proposition 2.2.1 (Properties of the Hill Estimator) *Let $(k)_{n\geq 1}$ be a sequence of integers such that $1 < k \leq n$, $k \to \infty$ and $\frac{k}{n} \to 0$ as $n \to \infty$.*

- *Weak consistency : Then, $\hat{\gamma}_H$ converges in probability to γ.*
- *Strong consistency : Moreover, if $\frac{k}{\log\log n} \to \infty$ an $n \to \infty$, then $\hat{\gamma}_H$ converges almost surely to γ.*

Proof. See Mason (1982), Deheuvels and all (1988) ■

Proposition 2.2.2 (Asymptotic Normality of the Hill Estimator) *Let F be a distribution function that belongs to the domain of attraction of the extreme value distribution G_γ with $\gamma > 0$. Then, for any sequence $k \to \infty$ satisfying ....,*

$$\sqrt{k}(\hat{\gamma}_H - \gamma) \xrightarrow{D} N(0, \gamma^2)$$

as $n \to \infty$, where $N(0, \gamma^2)$ is the normal distribution with mean 0 and variance γ^2.

Proof. See Davis and Resnick (1984). ■

2.2.2 Pickands Estimator

This estimator is based on the calculation of quantiles. It was first introduced by Pickands (1975) and later revisited by Drees (1995) and Drees and Kaufmann (1998). Additionally, Dekkers and de Haan (1989) studied its properties, establishing its weak consistency and asymptotic normality. Constructed using three order statistics, this estimator has the advantage of being valid regardless of the domain of attraction of the distribution.

Definition 2.2.1 *We assume that $(X_i, i = 1, ..., n)$ is a sequence of independent random variables following a distribution F that belongs to one of the domains of attraction. Let $(k)_{n\geq 1}$ be a sequence of integers with $1 \leq k \leq n$. The Pickands estimator is defined by:*

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

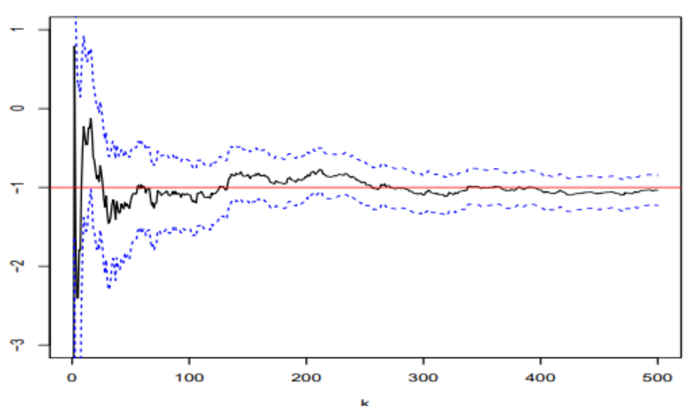

Figure 2.2: Pickands' Estimator

Proposition 2.2.3 (Properties of Pickands' Estimator) *Let $(\ k)_{n\geq 1}$ be a sequence of integers such that $1 \leq k \leq n, k \to \infty$, and $\frac{k}{n} \to 0$ as $n \to \infty$.*

- *Weak consistency : Then,$\hat{\gamma}_p$converges in probability to γ.*
- *Strong consistency : If $\frac{k}{\log \log n} \to \infty$ an $n \to \infty$, then $\hat{\gamma}_p$ converges almost surely to γ.*

- *Asymptotic normality: The strong convergence as well as the asymptotic normality were established by Dekkers and de Haan (*1989*) as*

$$\sqrt{k}(\hat{\gamma}_p - \gamma) \xrightarrow{D} N(0, \frac{\gamma^2(2^{2\gamma+1}+1)}{4(\log 2)^2(2^\gamma - 1)^2}).$$

  as $n \to \infty$ and $k \to \infty$.

Proof. See Dekkers and de Haan (1989). ■

2.2.3 Moment Estimator

Another estimator, which can be considered as an adaptation of the Hill estimator to ensure consistency regardless of the sign of the index γ , was proposed by Dekkers et al. (1989). This is known as the moment estimator.

Definition 2.2.2 *For $\gamma \in \mathbb{R}$, the moment estimator is defined by*

$$\hat{\gamma}_M = \hat{\gamma}_H + 1 - \frac{1}{2}\left(1 - \frac{(\hat{\gamma}_H)^2}{M_n^{(2)}}\right)^{-1}$$

Where:

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

The moment estimator is also known as the Dekkers-Einmahl-de Haan estimator.

Proposition 2.2.4 (Properties of the Moment Estimator) *Suppose that $F \in D(G_\gamma), \gamma \in \mathbb{R}$, $k \to \infty$ and $\frac{k}{n} \to 0$ as $n \to \infty$*

- *Weak consistency : Then,$\hat{\gamma}_M$ converges in probability to γ.*

- *Strong consistency : If, moreover, $\frac{k}{(\log n)^\delta} \to \infty$ as $n \to \infty$ for some $\delta > 0$, then $\hat{\gamma}_M$ converges almost surely to γ*

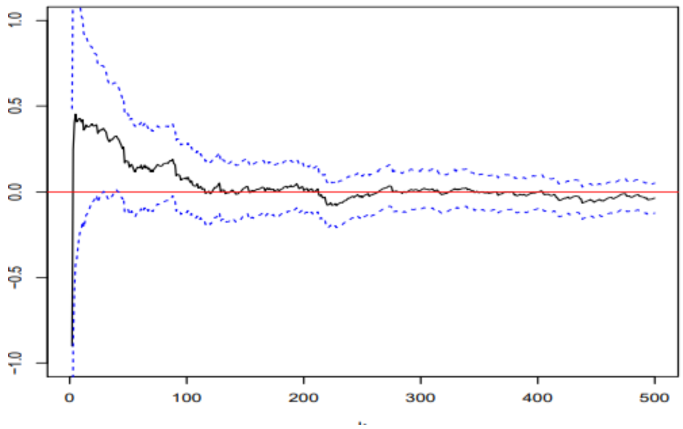

Figure 2.3: Moment estimator

- *Asymptotic normality: Under certain conditions on the distribution F (see Dekkers et al. (1987)),*

$$\sqrt{k}(\hat{\gamma}_M - \gamma) \xrightarrow{D} N(0, \sigma_M^2) \text{ as } n \to \infty$$

where

$$\sigma_M^2 = \begin{cases} 1 + \gamma^2 \text{ si } \gamma \geq 0 \\ (1-\gamma^2)(1-2\gamma)\left(4 - 8\frac{1-2\gamma}{1-3\gamma} + \frac{(5-11\gamma)(1-2\gamma)}{(1-3\gamma)(1-4\gamma)}\right) \text{ si } \gamma \leq 0 \end{cases}$$

Proof. See Dekkers et al. (1987). ■

2.3 The choice of the number k

Choosing the number k in extreme value estimation is a key challenge as it depends on the shape of the tail of the distribution and must be estimated based on the available data. The problem lies in balancing bias and variance: increasing k leads to higher bias, while reducing the data increases the variance. The optimal value of k can be determined using the numerical and graphical methods.

2.3.1 Numerical Method

To achieve an accurate estimation of the tail index using the Hill estimator, it is necessary to compute the Mean Squared Error (MSE), which depends on the number of extreme observations k. The MSE is expressed as:

$$\begin{aligned} MSE(\hat{\gamma}_k) &= MSE(\hat{\gamma}_k - \gamma)^2 \\ &= biais^2(\hat{\gamma}_k) + var(\hat{\gamma}_k) \end{aligned}$$

Thus, the objective is to select the optimal value of k that minimizes the MSE, achieving the best balance between bias and variance.

ln this context, de Haan and Peng (1998) proposed determining the optimal number of observations k_{opt} for estimating the tail index as follows :

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

2.3.2 Graphical Method

It is the simplest method for determining k. It consists of plotting the graph $(k, \hat{\gamma}_H)$ and selecting the value at which $(k, \hat{\gamma}_H)$ becomes horizontal. This estimator is only valid in the domain of attraction of the Fréchet distribution, that is, when $\gamma > 0$.

2.4 Threshold selection

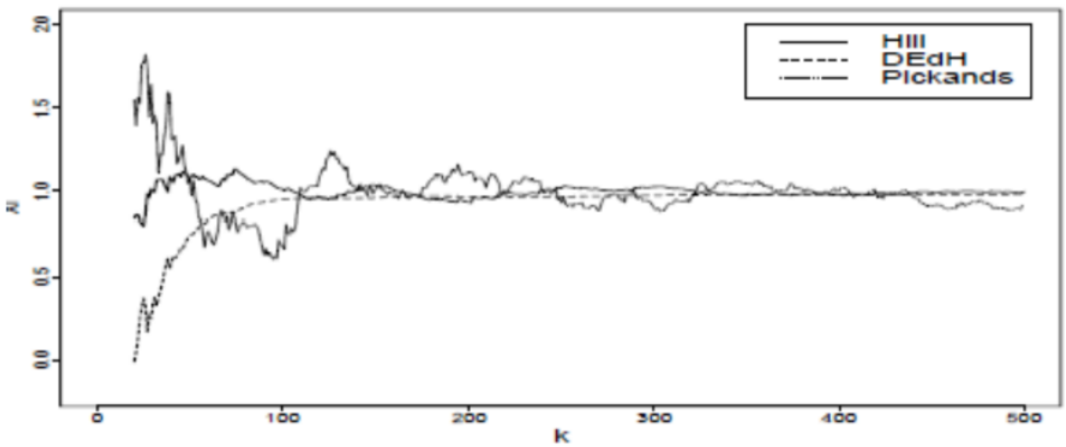

Figure 2.4: The graph $(k, \hat{\gamma}_H)$

2.4.1 Rule of thumb

One way to approach setting a threshold is by using a rule of thumb to choose the k largest observations and modeling. Commonly used is the 90th percentile, but others have also been proposed, such as $k = \sqrt{n}$ and $k = n^{(2/3)}/log(log(n))$ all of which are practical but to some level theoretically improper[. It is the fastest method of setting a threshold but due to the difference in behavior between different data it is not necessarily a reliable way of setting a good threshold due to the inevitable difference between most data. From the view of an insurance company it is possible that the information of interest may be the distribution of claims above some certain value, or the size of some upper quantile of claims. It is therefore sometimes of interest to use a certain threshold to get information even if it depreciates the theoretical analysis of the data.

2.4.2 Graphical approach

In a statistical framework, the choice of the threshold u is very important because it induces a large variability in the estimation of the extreme quantiles and the parameters of the law of excesses. There are different approaches for the choice of the threshold u of the POT method.

Indeed, the threshold must be large enough to satisfy the asymptotic character of the model, but not too high either, in order to keep a sufficient number of excesses to properly estimate the parameters of the model. Generally, u is determined graphically by the graph of the mean excess function (Mean Excess Plot) 1.2.

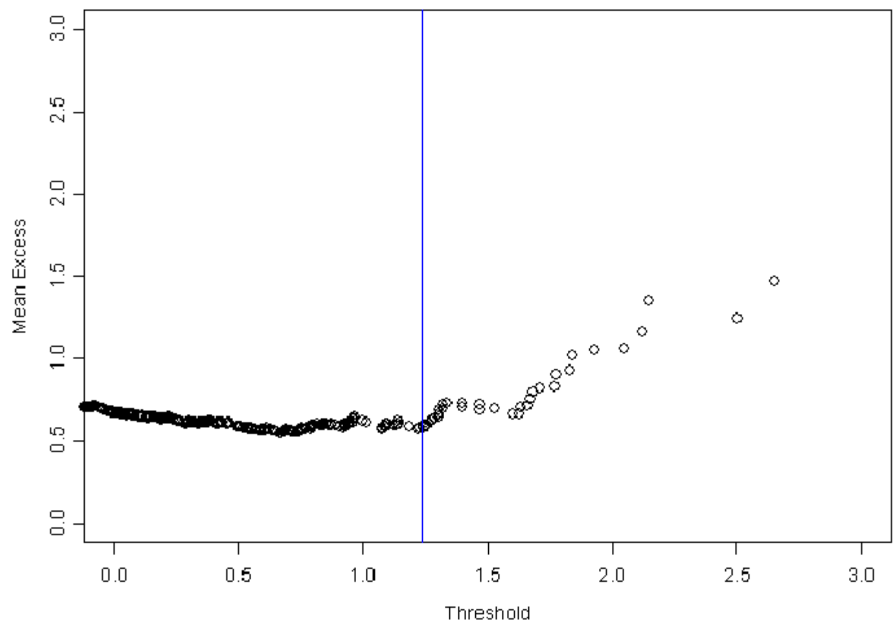

Figure 2.5: The average distribution of excesses.

Chapter 3

Application

3.1 Simulation

In this simulation study, we explore the behavior of extreme values using the Generalized Extreme Value (GEV) distribution. We generate synthetic data to model such events and analyze their properties within the GEV framework. Our objectives include generating synthetic data, estimating the parameters of the GEV distribution, analyzing the characteristics of extreme values, and validating the model by comparing it with real-world data. This study aims to deepen our understanding of extreme value theory and demonstrate the application of the GEV distribution in contexts where forecasting and managing extreme risks are essential.

	$n = 100$		$n = 200$	
	ML	LM	ML	LM
$\hat{\gamma}$	0.3249	0.2805	0.2838	0.2817
$\hat{\sigma}$	0.9339	0.9705	0.9034	0.8995
$\hat{\mu}$	-0.0064	0.0079	-0.0941	-0.0953

Table 3.1: The results of the parameters estimated by the two parametric methods.

3.1.1 Simulate a data sample for a GEV and estimated the parameters

R code

```
n =100, 200 # Sample size
shape =0.3
scale =1
location = 0
rgev_alternative = function(n, loc, scale, shape) {
u = runif(n)
x= loc + (scale / shape) * ((-log(u))^(-shape) - 1)
return(x)
}
simulated_data = rgev_alternative(n, loc = location, scale = scale, shape= shape)
hist(simulated_data, breaks = 30, freq = FALSE,
main = "Simulated Data rom GEV Distribution", xlab = "Value")
LM= fevd(simulated_data, method = "Lmoments")
ML=fevd(simulated_data,method="MLE")
```

3.1.2 Discussion and conclusion

Based on the results of this code, in particular the results of the form parameter estimated by the two methods, which were positive and included in the 0.3 range. This leads us to conclude

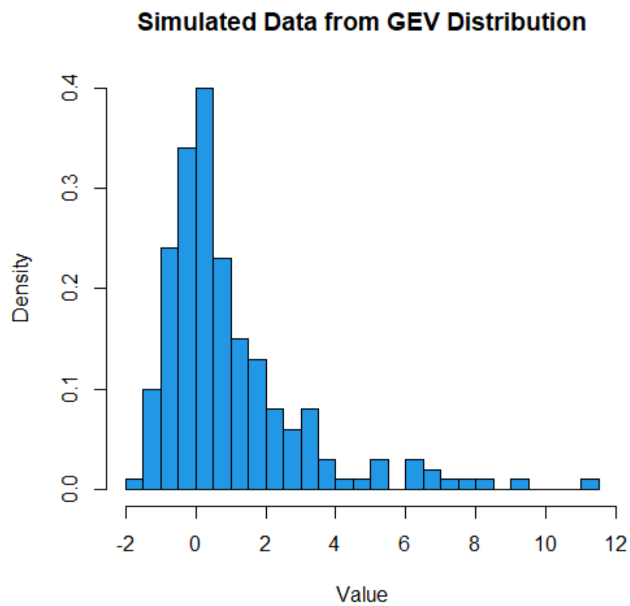

Figure 3.1: Histogram of the simulated Data from GEV Distribution

that the exact distribution of the data is the distribution of Fréchet, which is characterized by a heavy tail.

3.2 Real data

This part focuses on estimating the extreme value index using real-world data. The dataset consists of confirmed cases of coronavirus infection recorded between January $22, 2020$, and December $23, 2022$. The data were collected from Constantine Hospital.

The objectives of this empirical study are:

- To conduct a descriptive analysis of the dataset.
- To perform threshold selection for extreme value modeling.
- To estimate the index using parametric methods.
- To estimate the index using semi-parametric methods.

3.2.1 Descriptive analysis of the data

First, we do a simple statistical analysis on our data set, the results are as follows:

Min	1st Qu	Median	Mean	3rd Qu	Max	skewness	kurtosis
0.0	14.25	153.50	254.41	328.75	2521.00	2.68	12.30

Table 3.2: Statistical analysis of data.

These statistics indicate a high degree of dispersion in the number of daily cases, with values ranging from 0 to over 2500. The presence of extreme values is evident, making this dataset suitable for extreme value theory (EVT) analysis. Kortosis is 12.30 greater than 3, hence the heavy-tailed distribution.

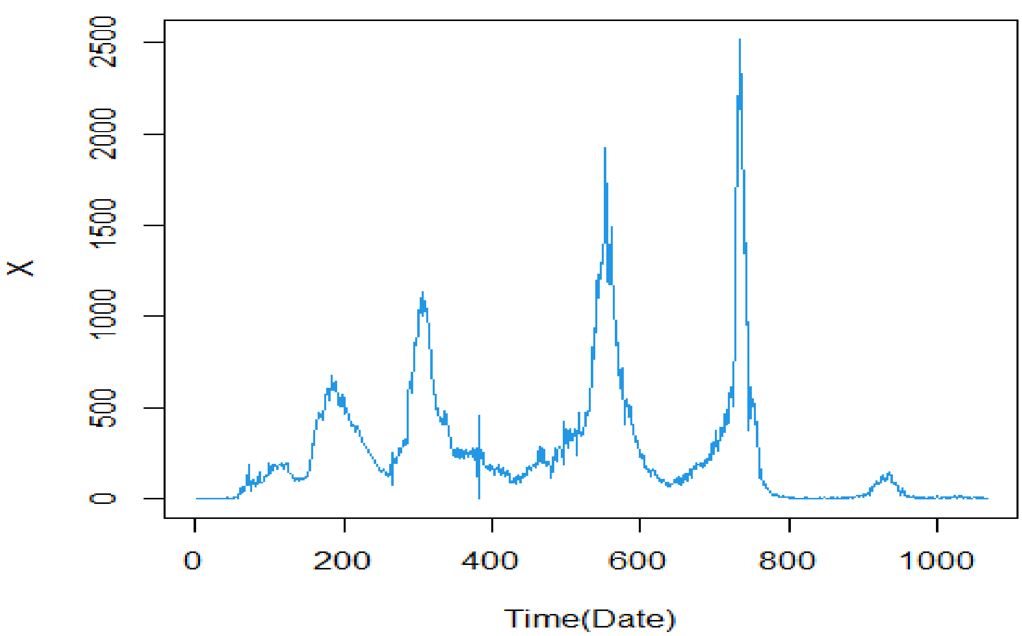

Figure 3.2: Confirmed cases of coronavirus infection in Algeria

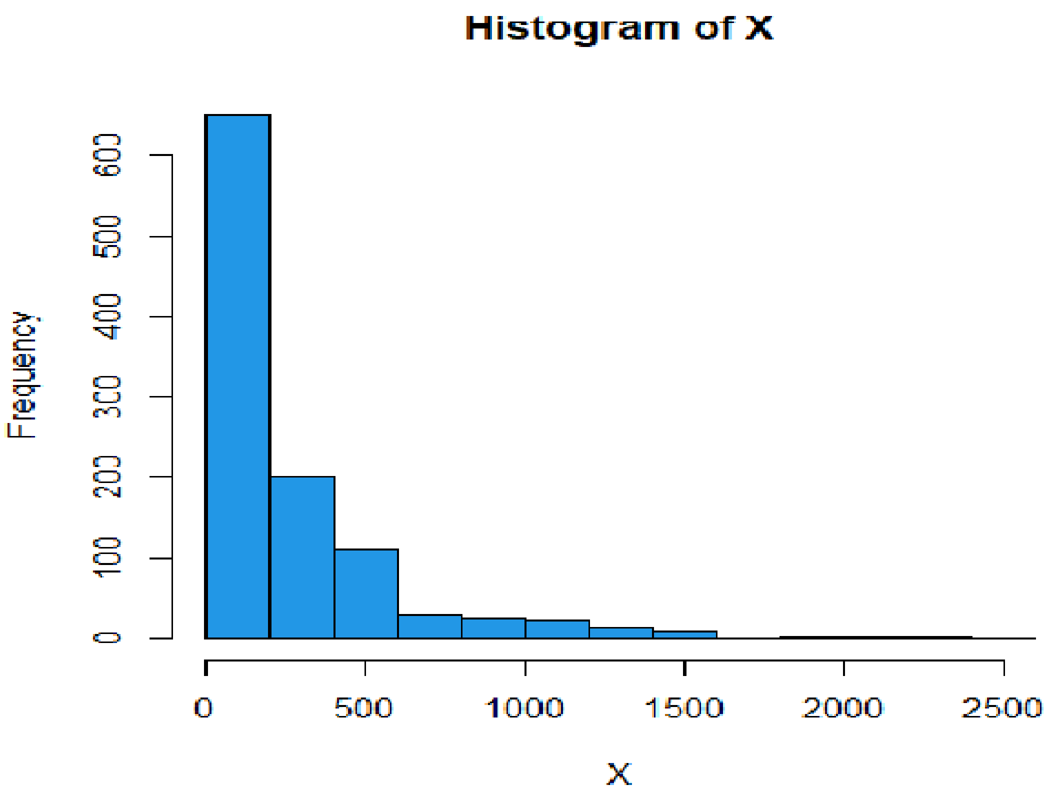

Figure 3.3: Histogram of the confirmed cases of coronavirus infection

R code

```
library(moments)
summary(X)
kurtosis(X)
skewness(X)
```

3.2.2 Adjustment by normal law

For the normal distribution adjustment using the Shapiro-Wilk test, the following hypotheses are tested:

$$\begin{cases} H_0: & \text{The sample follows a normal distribution.} \\ H_1: & \text{The sample does not follow a normal distribution.} \end{cases}$$

We set a significance threshold of $\alpha = 0.05$. The result: p-value $= 2.2 \times 10^{-16}$.

Test decision:

Fitting the data of our sample by the normal distribution gives a p-value equal to 2.2×10^{-16}, this value is less than the significance threshold $\alpha = 0.05$, therefore the sample does not follow a normal distribution, further confirming the appropriateness of EVT

R code

```
shapiro.test(X)
```

3.2.3 Threshold selection

We use the POT method to choice the threshold u,

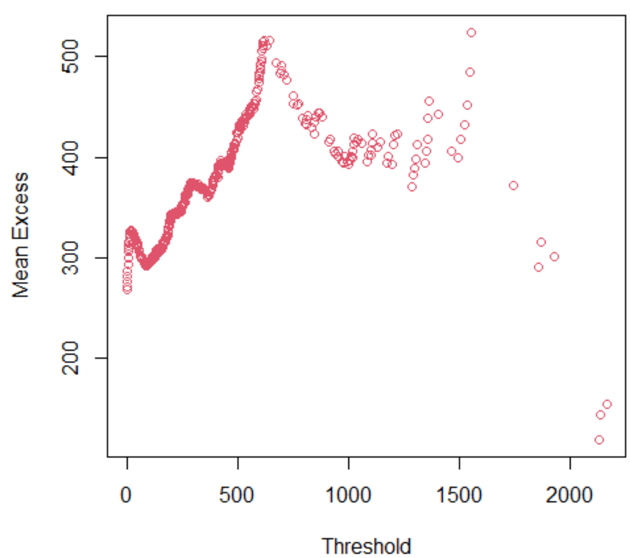

Figure 3.4: The average distribution of excesses for the real data.

the graph appears linear from threshold 1500 approximately, this allowed us to consider the value $u = 1500$ as the optimal threshold.

R code

```
library(evir)
meplot(X)
```

3.2.4 Estimation by parametric methods

In this section, we will estimate the parameters of the GEV using the following methods:

- Maximum likelihood (ML) method.
- L-moments (LM) method.

	ML	LM
$\hat{\gamma}$	1.5277	0.3680

Table 3.3: Parametric estimation.

Interpretation of Results

The estimate obtained by the maximum likelihood method is $\hat{\gamma} = 1.5277$, which is a very high value, indicating that the data follow a heavy-tailed distribution. This suggests a high probability of observing extremely large values, typical of a Fréchet-type behavior. Such a large value may also reflect the presence of strong outliers or extreme variability in the sample. In contrast, the estimate obtained using the L-moment method is $\hat{\gamma} = 0.3680$. This value is still positive, confirming that the distribution retains a heavy-tailed.

R code

```
library(extRemes)
library(evd)
LM= fevd(X, method = "Lmoments")
ML=fevd(X,threshold=1500,method="MLE")
```

3.2.5 Estimation by semi-parametric methods

Using some semi-parametric methods to estimate the extreme value index γ (Hill, Moments, Pickands). The results are as follows:

$\hat{\gamma}_H$	$\hat{\gamma}_P$	$\hat{\gamma}_M$
0.2169	0.1536	0.1497

Table 3.4: Semi-parametric estimation.

R code

```
N = length(X)
k = length(X[which(X > 1500)])
Y = sort(X,decreasing = TRUE)
t=numeric(k)
for(i in 1:k){
t[i]=Y[i]
}
# Hill estimator
Z = log(t)
H = sum(Z)/k-log(Y[k])
# Pickands estimator
Pickands=(1/log(2))*log(((Y[k])-(Y[(2*k)]))
/(((Y[(2*(k))])-(Y[(4 *(k))]))))
# Moments estimator
M2=(1/k)*sum((log(t)-log(Y[k]))^2)
M=H+1-(1/2)*((1-((H^2)/M2))^(-1))
```

Interpretation of Results

A positive extreme value index γ (as in your Hill and Moment estimates) suggests the data belong to the Fréchet domain, i.e., the distribution has a heavy tail. This is consistent with rare but very large spikes in infection counts.

This code in R confirms the reliability of our conclusions:

R code

```
library(VGAM)
x=rfrechet(1000,location=0,scale=1,shape=0.2169)
hist(X,prob=TRUE,col=4,main="HistogramwithFittedfrechetDistribution")
curve(dfrechet(x,location=0,scale=1,shape=0.2324),col=2,add=TRUE,lwd=2)
```

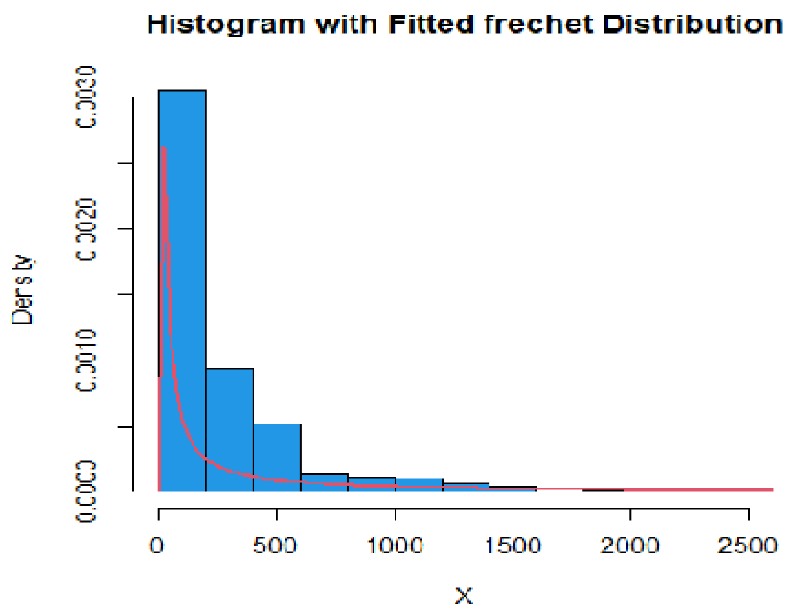

Figure 3.5: Histogram with fitted Frechet distribution

Conclusion

This master's thesis addresses the estimation of extreme values through a systematic comparison between semi-parametric estimators, such as the Pickands, Hill, and Moment estimators, and parametric estimators like L-moments and Maximum Likelihood Estimators (MLE). These methods are pivotal tools for analyzing extreme data, particularly when dealing with rare events that have significant impacts.

The study demonstrated that semi-parametric estimators are characterized by their flexibility, as they do not assume a specific distributional form, making them suitable for cases where the statistical model is difficult to define precisely. In contrast, parametric estimators rely on clear assumptions about the distribution and provide greater accuracy and efficiency when these assumptions hold true.

Practically, the analysis showed that the performance of each estimator varies based on sample characteristics and data distribution. For example, the Hill estimator excelled in heavy-tailed distributions, while the L-moments estimator provided more stable results for small samples. On the other hand, Maximum Likelihood Estimators offered high accuracy under ideal conditions, but they are highly sensitive to model specification.

Despite the challenges associated with each estimation approach, the comparison highlights the potential benefits of combining both methods in future research. Integrating the flexibility of semi-parametric estimators with the efficiency of parametric ones could lead to more adaptable models that better capture the complexities of extreme data.

Thus, this thesis underscores the importance of continued research in this area, focusing on developing hybrid and adaptive approaches that account for the nature of extreme phenomena and contribute to improved estimation accuracy and a deeper understanding of extreme data behavior.

R Software

R is a system, commonly known as language and software, which allows statistical analyzes to be carried out. More particularly, it comprises means which make possible the manipulation of the data, the calculations and the graphical representations. R also has the ability to run programs stored in text files and includes a large number of statistical procedures called packets. The latter make it possible to deal fairly quickly with subjects as varied as linear models (simple and generalized), regression (linear and non-linear), time series, classic parametric and non-parametric tests, the various methods of data analysis , ... Several packages, such ade4, MASS, multivariate, scatterplot3d among others are intended for the analysis of multidimensional statistical data.

It was originally created in 1996 by Robert Gentleman and Ross Ihaka of the Department of Statistics at the University of Auckland in New Zealand. It is designed to be used with Unix, Linux, Windows and MacOS operating systems. A key element in R's development mission is the Comprehensive R Archive Network (CRAN) which is a collection of sites that provides everything needed for the distribution of R, its extensions, documentation, source files and files. binaries. The master site of CRAN is located in Austria in Vienna, it can be accessed by the URL: "http://cran.r-project.org/". The other CRAN sites, called mirror sites, are spread all over the world.

Bibliography

[1] Abdelli, J. (2018). On the theory of extreme values, risk measures and applications [Doctoral thesis, Mohamed Khider University - Biskra]..

[2] Abu Bakar, M. A., Mohd Ariff, N., & Mohd Nadzir, M. S. (2023). Comparative analysis between L-moments and the maximum product spacing method for extreme PM10 concentration. In ICMSS 2022 (pp. 214–227). Atlantis Press. https://doi.org/10.2991/978-94-6463-014-5_21

[3] Ahsanullah, M., Nevzorov, V. B., & Shakil, M. (2013). An introduction to order statistics. Atlantis Studies in Probability and Statistics (Vol. 3). https://doi.org/10.2991/978-94-91216-83-1-1

[4] Balakrishnan, N., & Cohen, A. C. (1991). Order statistics and inference: Estimation methods. Academic Press.

[5] Berkane, H. (2022). Empirical processes and order statistics [Master's course]. Department of Mathematics, Mohamed Khider University - Biskra.

[6] Bingham, N. H., Goldie, C. M., & Teugels, J. L. (1987). Regular variation. Cambridge University Press.

[7] Borchani, A. (2010). Extreme value statistics in the case of discrete distributions. Higher School of Statistics and Information Analysis, Tunis.

[8] Csörgő, S., & Mason, D. M. (1985). Limit distributions of the Hill estimator. The Annals of Statistics, 13(2), 847–859.

[9] David, H. A. (1970). Order statistics. Wiley..

[10] Davis, R. A., & Resnick, S. I. (1984). The asymptotic distribution of Hill's estimator. The Annals of Probability, 12(4), 1216–1237.

[11] Dekkers, A. L. M., & de Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation. The Annals of Statistics, 17, 1795–1832.

[12] Dekkers, A. L. M., & de Haan, L. (1989). The generalized Pareto distribution and the estimation of extremes. Statistica Neerlandica, 43(1), 1–15.

[13] Dekkers, A. L. M., de Haan, L., & Einmahl, J. H. J. (1989). A moment estimator for the index of the extreme value distribution. The Annals of Statistics, 17(4), 1047–1063.

[14] Deheuvels, P., Haeusler, E., & Mason, D. M. (1988). On the strong consistency of the Hill estimator. The Annals of Probability, 16(3), 1185–1202.

[15] Deme, E. H. (2013). Contributions to univariate extreme value theory and actuarial risk measures for heavy-tailed losses [Doctoral thesis, Université Gaston Berger de Saint-Louis].

[16] Diamoutene, A. (2018). Contribution of extreme value theory to the management and health of systems [Doctoral thesis, National Polytechnic Institute of Toulouse, in collaboration with Joseph Ki-Zerbo University of Ouagadougou].

[17] Drees, H. (1995). The asymptotic normality of the Pickands estimator. The Annals of Statistics, 23(2), 212–233.

[18] Drees, H., & Kaufmann, R. (1998). Estimating the extreme tail of a distribution using order statistics. Statistica Sinica, 8(3), 829–847.

[19] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. Annals of Mathematics, 44(3), 423–453.

[20] Gumbel, E. J. (1958). Statistics of extremes. Columbia University Press.

[21] Haeusler, E., & Teugels, J. L. (1985). Asymptotic results for the Hill estimator. Statistics & Probability Letters, 3(3), 155–160.

[22] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. The Annals of Statistics, 3(5), 1163–1174.

[23] Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society: Series B, 52(1), 105–124.

[24] Kotz, S., & Nadarajah, S. (2000). Extreme value distributions: Theory and applications. Imperial College Press.

[25] Laurens,D,H & Ferreira, A. (2006). Extreme value theory: An introduction. Springer Series in Operations Research and Financial Engineering.

[26] Lekina, A. (2010). Non-parametric estimation of conditional extreme quantiles [Doctoral thesis, University of Grenoble].

[27] Longin, F. (1995). Extreme value theory: Presentation and initial applications in finance. Journal of the Statistical Society of Paris, 136(1), 77–97.

[28] Mason, D. M. (1982). On the weak consistency of the Hill estimator. Statistics & Probability Letters, 1(3), 173–176

[29] Pickands, J. (1975). Statistical inference using extreme order statistics. The Annals of Statistics, 3(1), 119–131.

[30] Real Statistics Using Excel. (n.d.). Generalized Pareto distribution. Retrieved. https://real-statistics.com/other-key-distributions/pareto-distribution/generalized-pareto-distribution

[31] Rodríguez, G. (n.d.). Survival models (Chapter 7). Retrieved from https://grodri.github.io/glms/notes/c7.pdf

[32] Yoann Novaro. RStudio Pubs. (n.d.). Extreme value theory. Retrieved April 17, 2025, from

[33] Sharpe, J., & Juárez, M. A. (2023). Estimation of the Pareto and related distributions — A reference-intrinsic approach. Communications in Statistics - Theory and Methods, 52(3), 523–542.

[34] Smith, R. L. (1987). Estimating the extreme tail of distributions. The Annals of Statistics, 15(4), 1172–1187.

Abstract

The tail index, which is one of the fundamental concepts in extreme value theory, is widely used in the analysis of rare events and large deviations. It has significant applications in various fields such as economics, insurance, and environmental studies.

This thesis focuses on estimating extreme values using parametric and semi-parametric methods. The parametric methods include Linear Moment Estimators (L-moments) and Maximum Likelihood Estimator (MLE), while the semi-parametric methods rely on order statistics such as Hill, Pickands, and Moment estimators. The theoretical foundations of these methods are presented, along with numerical applications using the R software.

Keywords: Extreme values, The tail index,Parametric estimation, Semi-parametric estimation.

Résumé

L'indicateur de queue, qui est l'un des concepts fondamentaux dans la théorie des valeurs extrêmes, est largement utilisé dans l'analyse des phénomènes rares et des grandes déviations. Il trouve des applications importantes dans divers domaines tels que l'économie, l'assurance et l'environnement. Cette thèse se concentre sur l'estimation des valeurs extrêmes à l'aide de méthodes paramétriques et semi-paramétriques. Les méthodes paramétriques incluent les estimateurs des moments linéaires (L-moments) et l'estimateur du maximum de vraisemblance (MLE), tandis que les méthodes semi-paramétriques reposent sur des statistiques ordonnées telles que les estimateurs de Hill, Pickands et Moments. Les bases théoriques de ces méthodes sont présentées, ainsi que des applications numériques utilisant le programme R.

Mots-clés : Valeurs extrêmes, L'indice de queue, Estimation paramétrique, Estimation semi-paramétrique.

الملخص

مؤشر الذيل، الذي يُعدّ من المفاهيم الأساسية في نظرية القيم المتطرفة، يُستخدم على نطاق واسع في تحليل الظواهر النادرة والانحرافات الكبيرة، وله تطبيقات هامة في مجالات متعددة مثل الاقتصاد، التأمين، والبيئة .

تركز هذه الأطروحة على تقدير القيم المتطرفة باستخدام طرق معلمية وشبه معلمية. تشمل الطرق المعلمية مقدّرات العزوم الخطية (L-moments)ومقدّر الاحتمالية العظمى(MLE) ، بينما تعتمد الطرق شبه المعلمية على الإحصاءات الترتيبية مثل مقدّرات هيل، وبيكاندس، والمومنت. تُعرض الأسس النظرية لهذه الطرق، إلى جانب تطبيقات عددية باستخدام برنامج R.

كلمات مفتاحية :القيم المتطرفة، مؤشر الذيل، التقدير المعلمي، التقدير شبه المعلمي.