People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

Mohamed Khider University , Biskra,

Faculty of Exact Sciences

Department of Mathematics

Thesis presented with a view to obtaining the Diploma :

Master in "Applied Mathematics"

Option : statistics

By :MEDDAS Imene

Title :

Estimating the Parameters of the Frechet Distribution by Weighted Least-Squares Method

Examination Committee Members :

Pr	CHERFAOUI Mouloud	UMKB	President
Dr	KHEMISSI Zahia	UMKB	Supervisor
Dr	ROUBI Affef	UMKB	Examiner

02 June 2025

Dedication

$\mathcal{I}$ dedicate this work:

$\mathcal{T}$o my dear mother and my dear father MEDDAS Abd el hafid, who have never ceased to support me and encourage me during my years of study.

$\mathcal{M}$ay this work stand as a testament to my deep gratitude and the knowledge I have gained

$\mathcal{T}$o my be loved brother Youcef and sisters, for their constant moral support and unwavering presence.

$\mathcal{T}$o my respected loyal friends, who have always encouraged and supported me, I wish them continued success.

And above all, I thank Almighty God for granting me the strength, will, and perseverance to accomplish this achievement. All praise be to Him, first and foremost.

ACKNOWLEDGMENTS

*F*irst and foremost, I thank **Allah** for granting me strength, patience, and courage throughout my years of study.

I would like to express my sincere gratitude to my dear teacher and supervisor, Ms. **KHEMISSI Zahia**, for her continuous guidance and unwavering support during this project. I am deeply thankful for the time she devoted, the valuable insights she shared, and her understanding throughout the process.

*M*y sincere thanks also go to the members of the jury for kindly examining and evaluating this work. I extend my appreciation to the administration and all the teaching staff of the Mathematics Department for their support and dedication.

*M*ost importantly, I would like to express my deepest gratitude to those who worked the hardest to see me succeed, those who sacrificed their own dreams so I could achieve mine my beloved parents. They deserve all my love, respect, and heartfelt thanks.

M.IMENE....

Contents

List of Tables

List of Figures

Introduction

Extreme value distributions are gaining increasing importance in various applied fields such as engineering, finance, and environmental sciences, where they are used to model rare and extreme events. The Frechet distribution is among the most prominent of these distributions due to its ability to represent heavy-tailed data.

Estimating the parameters of such distributions is a crucial step to ensure the accuracy and effectiveness of statistical models. However, traditional methods such as maximum likelihood, moments, and ordinary least squares may perform poorly when dealing with small samples or highly extreme data.

In this context, the Weighted Least Squares (WLS) method emerges as a promising alternative potential to enhance estimation accuracy, especially when appropriate weights are applied to reflect the distributional characteristics of the data. Previous studies have shown that this method may outperform classical approaches in certain situations.

This thesis is structured into three chapters:

Chapter 1:

This chapter presents the theoretical framework of the study, including the fundamental concepts, extreme value theory, and tail index estimation. It concludes with an overview of tail index estimation methods, which are essential for analyzing the extreme behavior of distributions.

Chapter 2:

This chapter covers the mathematical background for applying regression-based estimation methods, with a particular focus on the weighted least squares (WLS) method. We begin by reviewing some basic concepts in regression theory, such as the regression model and the simple linear regression model, then discuss various weight expressions. The chapter concludes with the derivation and formulation of the estimators.

Chapter 3:

This chapter focuses on evaluating the performance of the estimators obtained using the WLS method through a comparative simulation study. The analysis is based on statistical indicators such as bias and mean squared error (MSE), with the results presented and discussed to highlight the efficiency of the proposed approach.

Chapter 1

Tail index estimation

Tail index estimation is a fundamental concept in extreme value theory and risk management, used to measure the "heaviness" or "extent" of the tail in a probability distribution. The tail index determines the behavior of extreme values, especially in heavy-tailed distributions such as the Pareto, Cauchy, and Frechet distributions.

This chapter focuses on the study of extreme values and their fundamental characteristics, with a particular emphasis on estimating the tail index and the various methods used for its estimation. ther are many refrence books on extreme value theory (**EVT**)for example: David et all., **1970**, Balakrishnan et all., **1991**, and De Haan et all., **2006**.

1.1 Fundamentals of EVT

Definition 1.1.1 ***(Distribution and survival functions)****:If X is a rv defined on aprobability space (Ω, F, P) then, its df $F(x)$ and survival function $\overline{F}(x)$(also called hazard function) are defined on $\mathbb{R}$ on follows:*

$$F(x) := P(X \leq x) \quad and \quad \overline{F}(x) := 1 - F(x)$$

Definition 1.1.2 (The empirical distribution function):*Let $X_1, ..., X_n$ be a sample of a positive r.v X of size $n \geq 1$. The empirical distribution function $F_n(x)$ is defined as:*

$$F_n := \frac{1}{n} \sum\nolimits_{i=1}^{n} I_{\{X_i \leq x\}}, \quad \forall x \geq 0$$

where $I_{\{X_i \leq x\}}$ is the indicator function of the set $\{X_i \leq x\}$, $F_n(x)$ is the proportion of the $X_i$variables which are less than or equal to x.

Definition 1.1.3 (The empirical survival function*): Let $X_1, ..., X_n$ be a sample of a positive r.v X, of size $n \geq 1$. The empirical survival function noted by S_n, is given by:*

$$S_n := 1 - F_n := \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i > x\}}, \quad \forall x \geq 0$$

S_n is the proportion of observations that exceeds x.

Definition 1.1.4 (Quantile function*):The quantile function of F is generalized inverse function of F defined by :*

$$Q(s) = F^{\longleftarrow}(s) := \inf \{x : F(x) \geq s\}$$

for all 0 < s < 1,with the convention that $\inf(\emptyset) = \infty$.

Definition 1.1.5 ***(Empirical quantile function)****:The empirical quantile function of the sample* $(X_1, ..., X_n)$ *is :*

$$Q_n(s) = \inf\{x : F_n(x) \geq s\} = \inf\left\{x : \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i > x\}} \geq s\right\}$$

for all 0 < s <1, where F_n *is the empirical distribution function.*

1.1.1 Order statistics

Definition 1.1.6 ***(Order statistics)****:Let* $(X_1, ..., X_n)$ *be* n *iid random variable with a common distribution* F *and density* f *. We call the order statistics (increasing order) the sequence of random variables* $(X_1, ..., X_n)$*which are ordered by ascending order, either:*

$$X_{1,n} \leq ...... \leq X_{n,n}$$

Remark 1.1.1 *:For* $1 \leq K \leq n$ *the variable*$X_{k,n}$ *is known under the name of the* K^{th} *order statistic or* K *order statistic.*

Definition 1.1.7 (Extreme order statistics):*Two order statistics are particular-interesting for the study of extreme events, are defined respectively by:*

- *the variable* $X_{1,n}$ *is the smallest statistic of order (or statistic of the minimum):*

$$X_{1,n} := \min(X_1, ..., X_n).$$

- *the variable $X_{n,n}$ is the greatest statistic of order (or maximum statistic):*

$$X_{n,n} := \max(X_1, ..., X_n).$$

Definition 1.1.8 (Extreme order statistics distributions): *The distributions $F_{X_{1,n}}$ and $F_{X_{n,n}}$ of the extreme order statistics $X_{1,n}$ and $X_{n,n}$ are respectively defined by:*

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

$$F_{X_{n,n}}(x) = [F_X(x)]^n$$

pdf of $X_{1,n}$ and $X_{n,n}$ is:

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

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

Definition 1.1.9 ***(Empirical df*** *) :The empirical df of the sample $(X_1, ..., X_n)$ is eval-uated using order statistics as follows:*

$$F_n(x) = \begin{cases} 0 & if \;\; x \leq X_{1,n} \\ \frac{i-1}{n} & if \;\; X_{i-1,n} \leq x \leq X_{i,n} \; for \; 2 \leq i \leq n \\ 1 & if \;\; x \geq X_{n,n} \end{cases}$$

Definition 1.1.10 ***(Distribution function of the K^{th} upper order statistic)****:for $k = 1, ..., n$ let $F_{X_{K,n}}$ denote the df of $X_{K,n}$, then*

$$F_{X_{k,n}} = \sum_{r=0}^{k-1} \binom{n}{r} \overline{F}^r(x) F^{n-r}(x)$$

if is continuous; then

$$F_{X_{K,n}} := \int_{-\infty}^{x} f_{K,n}(z)dF(z),$$

where,

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

i.e $f_{k,n}$ *is a density of* $F_{k,n}$ *with respect to* F

proof:see e.g Embrechts et al., 1997 page 183.

Definition 1.1.11 ***(Upper end point)**: We denote by* x_F *(resp* x_F^* *) the upper extreme point (resp. Lower) of the distribution* F *(i.e. the greatest possible value for* $X_{k,n}$ *which can take the value* $+\infty$ *(resp* $-\infty$*)) in the sense that:*

$$x_F := \sup\{x : F(x) < 1\} \leq \infty \qquad ,and$$
$$x_F^* := \inf\{x : F(x) > 0\},$$

1.1.2 Distribution of extreme values

This section deals the classical Extreme Value Theory (**EVT**), focusing on the fundamental result reached by Fisher-Tippett [11] 1928.The theory of extreme values shows that there are sequences $\{a_n\}$ and $\{b_n\}$; $n \in \mathbb{N}^*$, with $a_n > 0$ and $b_n \in \mathbb{R}$; as

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} F^n(a_n x + b_n) = \mathcal{H}(x) \qquad \forall x \in \mathbb{R}, \tag{1.1}$$

where$\mathcal{H}$ is a non-degenerate df. Since extreme value df's are continuous on $\mathbb{R}$, assumption1.1 is equivalent to the following weak convergence assumption

$$\frac{X_{n,n}-b_n}{a_n} \quad \overset{\mathcal{H}}{\rightarrow} \quad \text{as} \quad n\rightarrow\infty$$

Remark 1.1.2 *:The sequences $\{a_n\}$ and $\{b_n\}$, $n\geq 1$ are called sequences of normalization, the constants $a_n\in\mathbb{R}_+^*$ and $b_n\in\mathbb{R}$ are called constants of normalization and the random variable $\frac{1}{a_n}\left(X_{n,n}-b_n\right)$ is called the normalized maximum*

1.1.3 Limit distributions

The following theorem gives a necessary and suffcient condition for the existence of limit distribtion for the maximum.

Theorem 1.1.1 *(**Fisher & Tippett 1928**):Let $(X_n)_{n\geq1}$ be a sequence of i.i.d random variables with cdf F. If there exist two real sequences $\{a_n\}_{n\geq1}>0$ and $\{b_n\}_{n\geq1}\in\mathbb{R}$, and a non-degenerate distribution function $\mathcal{H}$ such that:*

$$\lim_{n\rightarrow\infty}\mathbf{P}\left[\frac{M_n-b_n}{a_n}\leq x\right]=\mathcal{H}_\gamma\left(x\right), \quad \text{where } M_n=\max(X_1,....,X_n), \tag{1.2}$$

then $\mathcal{H}_\gamma\left(x\right)$ must be one of the following three type of distributions:

$$\begin{array}{lll}
Frechet & \Phi_\alpha(x):=\left\{\begin{array}{cc} 0 & \text{if } x\leq 0 \\ \exp(-x^{-\frac{1}{\alpha}}) & \text{if } x>0 \end{array}\right| & \text{with } \alpha>0 \\
Weibull & \Psi_\alpha(x):=\left\{\begin{array}{cc} 1 & \text{if } x\geq 0 \\ \exp\left[-(-x)^{-\frac{1}{\alpha}}\right] & \text{if } x<0 \end{array}\right| & \text{with } \alpha<0 \\
Gumbel & \Lambda_\alpha(x):=\exp\left(-\exp\left(-x\right)\right), & \text{for all } x\in\mathbb{R}
\end{array}$$

We refer to Φ_α, Ψ_α and Λ_α as the extreme value distributions.

A detailed proof of this theorem is given in Resnick [24](1987).

Proposition 1.1.1 ***(Density function of extreme values)****:The density functions of the distribution of standard extreme values and the diferent types of extreme distribution, are as follows:*

$$\begin{aligned} \textit{Frechet} &: \quad \phi(x) = \alpha x^{-\alpha-1}\exp(-x^{-\alpha}) \quad && x > 0 \\ \textit{Weibull} &: \quad \Psi(x) = \gamma(-x)^{-\alpha-1}\exp(-(-x)^{-\alpha}) \quad && x < 0 \\ \textit{Gumbel} &: \quad \lambda(x) = exp\left[-\left\{x + \exp(-x)\right\}\right] \quad && x \in \mathbb{R} \end{aligned}$$

Figure1.1illustrates the Densities of the standard extreme value distributions. we chose $\alpha = 1$ for the Gumbel and the Frechet and the Weibull distribution.

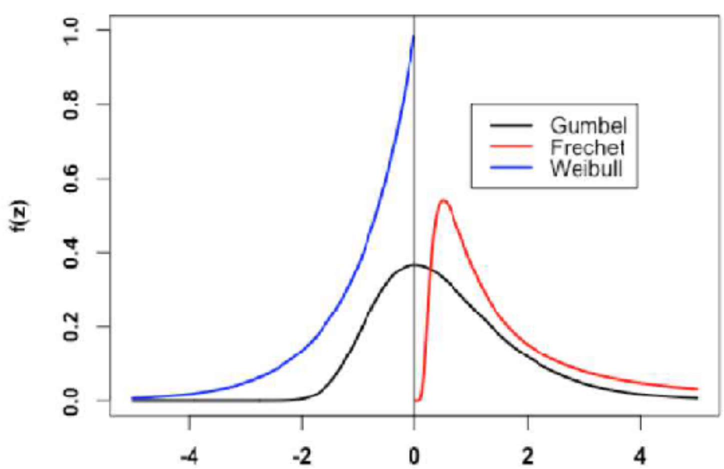

Figure 1.1: Densities of the standard extreme value distributions.

Definition 1.1.12 ***(Generalized extreme values distribution)(GEVD):***

The behaviour of $\Phi_{1/\gamma}$, $\Psi_{1/\gamma}$ and Λ is completely different but they can be combined into a single distribution dependent on a single parameter that controls the tail thickness of the distribution

Theorem 1.1.2 *If the limit (1.2) exists, then:*

$$\mathcal{G}_\gamma(x) := \begin{cases} \exp(-(1+\gamma x)^{-\frac{1}{\gamma}}) & \textit{if } \gamma \neq 0, \\ \exp(-\exp(-x)) & \textit{if } \gamma = 0. \end{cases}$$

where:

- $(1+\gamma x) > 0$.
- $\mathcal{G}_\gamma(x)$:is the cdf of the GEV Distribution.
- γ : is called the tail index or extreme value index,and it is a fundamental indicator of the shape of the tail.

Interpretation of the sign of the parameter γ:

1. The case $\gamma > 0$, corresponds to Frechet's distribution with parameter $1/\gamma > 0$.
2. The case $\gamma < 0$, corresponds to the Weibull's distribution with parameter $-1/\gamma < 0$.
3. The case $\gamma = 0$, corresponds to Gumbel's distribution.

Definition 1.1.13 ***(The Generalized Pareto Distribution):*** (GPD), *with parameters* $\gamma \in \mathbb{R}$, *and* $\sigma > 0$, *is defined by its distribution function, given by:*

$$\mathcal{G}_{\gamma,\sigma}(x) = \begin{cases} 1-(1+\frac{\gamma}{\sigma}x)^{-1/\gamma} & if\, \gamma \neq 0, \\ 1-\exp(-\frac{x}{\sigma}) & if\, \gamma = 0, \end{cases} \Bigg|$$

or, $x \geq 0$ *if* $\gamma \geq 0$ *and* $0 \leq x \leq \frac{\gamma}{\sigma}$ *if* $\gamma < 0$.

1.1.4 Domains of attraction

Before defining the domain of attraction, we begin by introducing functions that exhibit different types of variation.For more information, refer to Bingham et all., [3] 1987, where many results on regularly varying functions are discussed.

Definition 1.1.14 ***(Regularly varying and slowly varying functions):****.A measurable function* V *:*$\mathbb{R}_+ \longrightarrow \mathbb{R}_+$ *is regularly varying at* ∞ *with the index* ρ, *and we denote by* $V \in \mathcal{RV}_\rho$ *, if :*

$$\lim_{x\longrightarrow\infty} \frac{V(tx)}{V(t)} = x^\rho \qquad\qquad t > 0,$$

A measurable function $l:\ ]a,+\infty[\ \longrightarrow \mathbb{R}$ with $(t>0)$ is said slowly varying at infinity, if:

$$\lim_{x\longrightarrow\infty} \frac{l(tx)}{l(x)} = 1,$$

Theorem 1.1.3 *(**Kramata representation**):every slowly varying function (i.e* $l \in \mathcal{RV}_0$*),if and only if can be represented as:*

$$L(x) = c(x)\exp\left\{\int_1^x \frac{r(x)}{t}dt\right\},\ x > 0,$$

where $c(\cdot)$, $r(\cdot)$ *measurable functions, and*

$$\lim_{x\to\infty} c(x) = c_0 \in \,]0, +\infty[\;\; and \;\; \lim_{t\to\infty} r(t) = 0.$$

A function $V : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ *is regularly varying at* ∞ *with the index* ρ *if and only if* V *has the representation:*

$$V(x) = c(x) \exp\left\{ \int_1^x \frac{\rho(t)}{t} dt \right\}, \; x > 0,$$

proof:See Resnick[24], 1987, Corollary 2.1; page 29.

Definition 1.1.15 ***(Domains of Attraction)****:A distribution is said to belong to the* $\mathcal{DA}$ *of* G, *denoted* $F \in \mathcal{DA}(G)$, *if the distribution of the normalised maximum converges to* $\mathcal{G}$. *In other words, if there exist real constants* $a_n > 0$ *and* $b_n \in \mathbb{R}$ *such that*

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G_\gamma(x).$$

The table 1.1 presents the $\mathcal{DA}$ *for some models.*

Domaine of attraction	Frechet $(\gamma > 0)$	Gumbel $(\gamma = 0)$	Weibull $(\gamma < 0)$
Models	Burr student Log-gamma Chi-sequre Pareto	Gamma Normale Exponentielle	uniforme Beta

Table 1.1: Usual models and their domaine of attractions

1.1.5 Characterizations of the attraction domain

Diferent characterizations of three domain of attraction of Frechet,Weibull and Gumbel, according to the sign of γ, we can distinguish three domain of attraction:

Characterization of $\mathcal{D}\left(\Phi_{\gamma}\right)$:

if $\gamma > 0$, we say that F belongs to $\mathcal{D}\left(\Phi_{\gamma}\right)$, and F has an infinite right end point ($x_F = +\infty$). This indicates that falls within the domain of attraction of heavy-tailed distributions, which have a polynomially decaying survival function. This result was established by Gnedenko [13](1943), and a simplified proof can be found in Resnick's book, [Proposition 1.11].

Theorem 1.1.4 *:$F \in \mathcal{D}(\Phi_{\gamma})$ with parameter $\gamma > 0, x_F = +\infty$ if and only if :*

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

where l is a slowly varying function. In this case, a possible choice for the sequences a_n and b_n are $a_n = F^{-1}(1 - \frac{1}{n})$ and $b_n = 0$.

Characterization of $\mathcal{D}\left(\Psi_{\gamma}\right)$:

if $\gamma < 0$, we say that $F \in \mathcal{D}\left(\Psi_{\gamma}\right)$and that F has a finite right end point ($x_F < +\infty$). This implies that the domain of attraction of survival functions is restricted to distributions with an upper-bounded support.

The following result (see Gnedenko [13], 1943 and Resnick [24], 1987, [Proposition 1.3]) shows that we can transition from the domain of attraction of the Frechet distribution to that of the Weibull distribution by a simple change of variable in the distribution function.

Theorem 1.1.5 *:$F \in \mathcal{D}(\Psi_\gamma)$ with parameter $\gamma < 0$, if $x_F = +\infty$ and $1 - F^*$is a function with regular variations of index α*

$$1 - F(x) = x_F - x^{-1} = x^{-1/\gamma} l(x),$$

where the function l slowly varying of index $1/\gamma$ In this case, a possible choice for the sequences a_n and b_n is

$$a_n = x_{_F} - F^{-1}(1 - \frac{1}{n}) \quad \text{and} \quad b_n = x_F$$

this domain of attraction has been considered by Falk [10],1995, Gardes [12], 2010 to give an endpoint estimator of a distribution.

Characterization of $\mathcal{D}(\Lambda)$:

If $\gamma = 0$, we say that F belongs to $\mathcal{D}(\Lambda)$. In this case, the upper end x_F point can be either finite or infinite. This domain includes distributions with light tails, meaning those that have an exponentially decaying survival function. This result was notably proven in Resnick[24] , 1987 [Proposition 1.4].

Theorem 1.1.6 *:A distribution function F belongs to the Gumbel domain of attraction if and only if there exists $z < x_F < \infty$ such that*

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

where$c(x) \to c > 0$ when $x \to x_F$ and $a(.)$ is a positive and differentiable function with derivative $\acute{a}(.)$ such that $\lim_{x \to x_F} \acute{a}(.) \to 0$.

The tables 1.2, 1.3 and 1.4 give different examples of standard distributions in these three domains of attraction.

Distributions	$\bar{F}(x)$	γ
Uniforme$[0,1]$	$1-x$	-1
Inverse Burr$(\beta,\tau,\lambda,x_\tau)$ $\beta,\lambda,\tau>0$	$\left(\frac{\beta}{\beta+(x_\tau+x)^{-\tau}}\right)^\lambda$	$-\frac{1}{\lambda}$

Table 1.2: Some distributions associated with a negative index.

Distributions	$\bar{F}(x)$ or density f	γ
Burr(β,τ,λ) $\beta>0,\tau>0,\lambda>0$	$\left(\frac{\beta}{\beta-x^\tau}\right)^\lambda$	$\frac{1}{\lambda^\tau}$
Frechet$\frac{1}{\alpha},\alpha>0$	$1-\exp(-x^{-\alpha})$	$\frac{1}{\alpha}$
Log-gamma $\lambda>0,m>0$	$\frac{\lambda^m}{\Gamma(m)}\int\limits_x^\infty(\log(u))^{m-1}u^{-(\lambda+1)}du$	$\frac{1}{\lambda}$
log-logistic $\beta>0,\alpha>0$	$\frac{1}{1+\beta^\alpha}$	$\frac{1}{\alpha}$
Pareto $\alpha>0$	$x^{-\alpha},x>0$	$\frac{1}{\alpha}$

Table 1.3: Some distributions associated with a positive index

Distributions	$\bar{F}(x)$ or density f	γ
Gamma (m,λ) $m\in\mathbb{N},\lambda>0$	$f(x)=\frac{\lambda^m}{\Gamma(m)}\int\limits_x^\infty u^{m-1}\exp(-\lambda u)du$	0
Gumbel (μ,β) $\mu\in\mathbb{R},\beta>0$	$f(x)=\exp(-\exp(-\frac{x-u}{\beta}))$	0
Logistic	$\bar{F}(x)=\frac{2}{1+\exp(x)}$	0
Log nomale $\beta>0,\alpha>0$	$f(x)=\frac{1}{2\pi}\int\limits_x^\infty\frac{1}{u}\exp(-\frac{1}{2\sigma^2}(\log u-u)^2)du$	0
Weibull $\alpha>0$	$\bar{F}(x)=\exp(-\lambda u^\tau)$	0

Table 1.4: Some distributions associated with a null index

1.2 Methods of semi-parametric estimation

These are methods that combine flexibility with assumptions; they do not assume a specific form for the entire distribution, but only for the tail (e.g., the Generalized Pareto Distribution – GPD).

Among the most commonly used semi-parametric estimators for estimating the tail index in heavy-tailed distributions are the Hill and Pickands estimators. These are classical methods that rely on extreme values.

1.2.1 Hill's estimator

Hill's estimator is one of the most widely used estimators for the tail index of heavy-tailed distributions. Research in this field has primarily focused on the case where the EVI is positive ($\gamma = \frac{1}{\alpha} > 0$) since most real-world data follows distributions that belong to the domain of attraction of the Frechet distribution $F \in \mathcal{D}(\Phi_\alpha)$.

Hill[14] ,1975 identified that the distribution tail follows a Pareto shape,

$$\hat{\gamma}^H = \hat{\gamma}_K^{(H)} := \frac{1}{K} \sum_{j=1}^{K} \log X_{n-j+1,n} - \log X_{n-K,n}$$

The estimator's construction was detailed by De Haan et all.,[5], 2006 and Beirlant et all.,[1] 2016. Alternative estimators include Beirlant et all.,[1], 2016 exponential regression model and Csörgő et all., [4], 1985 kernel-based approach.

The asymptotic properties of Hill's estimator are summarized in the following theorem.

Theorem 1.2.1 *(Asymptotic Properties of $\hat{\gamma}^H$):Assume that $F \in \mathcal{D}(\Phi_{\frac{1}{\gamma}}), \gamma > 0$, $k \to \infty$ and $K/n \to 0$ when $n \to \infty$.*

1. *Mason [21],1982 has proven weak consistency :*

$$\hat{\gamma}^H \quad \xrightarrow{p} \quad \gamma \quad \textit{when} \quad n \to \infty,$$

2. *Strong consistency was established by Deheuvels et all. [7],1985 under the*

condition that $:K/\log n \to \infty$*,then*

$$\hat{\gamma}^H \overset{a.s}{\longrightarrow} \gamma \text{ when } n \to \infty,$$

and more recently by Necir [22], 2006.

3. *Asymptotic normality was established under a suitable extra assumption, known as the second-order regular variation condition (see De haan and Stadtmüller* **[6],** *1996 and De haan and Ferreira* **[5],** *2006), with mean* γ *and variance* γ^2/k.

$$\sqrt{K}\left(\frac{\hat{\gamma}^H - \gamma}{\gamma}\right) \overset{d}{\longrightarrow} \mathcal{N}(0,1)$$

Figure 1.2, show that the Hill estimator against k performs well with both the Frechet distribution, the sample size is $n = 1000$, with parameter $\gamma = 0.6$.

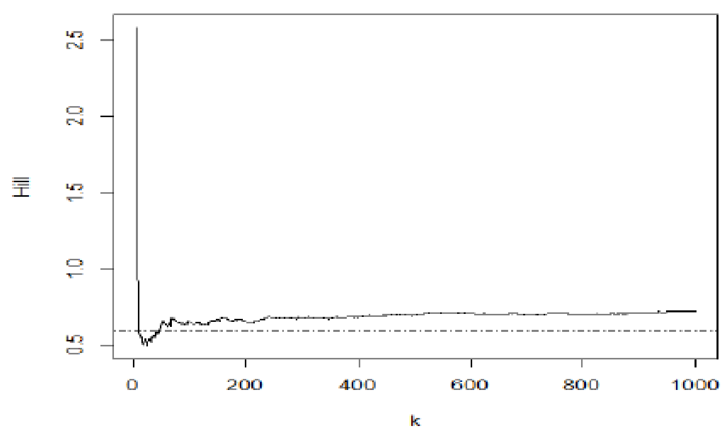

Figure 1.2: Hill estimator for samples of a Frechet distribution ($\gamma = 0.6$)

1.2.2 Pickand's estimator

Pickands estimator is a classical alternative to Hill's estimator and is primarily used in analyzing heavy-tailed distributions.

James Pickands proposed his estimator in 1975,[23] to estimate the tail index in heavy-tailed distributions, based on the order statistics of extreme values.the Pickand estimator is defined by:

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

This estimator was later improved by Dekkers and de Haan (1989)[8] and Drees (1995)[9].

Theorem 1.2.2 ***(Asymptotic Properties of*** $\hat{\gamma}^{(p)}$***):*** *Assume that* $F \in \mathcal{D}(\mathcal{H})$, $\gamma \in \mathbb{R}, k \to \infty$ *and* $\frac{k}{n}$ *when* $n \to \infty$.

1. Weak Consistency :

$$\hat{\gamma}^{(p)} \xrightarrow{p} \gamma \qquad \text{when} \;\; n \to \infty.$$

2. Strong consistency:

$$\hat{\gamma}^{(p)} \xrightarrow{a.s} \gamma \qquad \text{when} \;\; n \to \infty.$$

3. Asymptotic normality: under further conditions on k and F,

$$\sqrt{K}(\hat{\gamma}_{K}^{(p)} - \gamma) \xrightarrow{d} \mathcal{N}(0, \eta^2) \qquad \text{when} \;\; n \to \infty.$$

where

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

1.2.3 Moment estimator

Another estimator which can be considered as an adaptation of Hill's estimator, to obtain the consistency for all $\gamma \in \mathbb{R}$, has been proposed by Dekkers et all., [8],1989. This is the moment estimator, given by:

$$\hat{\gamma}^M = \hat{\gamma}_K^{(M)} := M_1 + 1 - \frac{1}{2}\left(1 - \frac{\left(M_{(K)}^{(1)}\right)^2}{M_{(K)}^{(2)}}\right)^{-1},$$

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

Theorem 1.2.3 *(Asymptotic Properties of $\hat{\gamma}^M$):Assume that $F \in \mathcal{D}(\mathcal{H}), \gamma \in \mathbb{R}$, $k \to \infty$ and $K/n \to 0$ when $n \to \infty$:*

1. *weak consistency :*

$$\hat{\gamma}^M \quad \xrightarrow{p} \quad \gamma \quad when \quad n \to \infty,$$

2. *Strong consistency:if $K/(\log n)^\delta \to \infty$, when $n \to \infty$ for certain $\delta > 0$, so*

$$\hat{\gamma}^M \quad \xrightarrow{a.s} \quad \gamma \quad when \quad n \to \infty,$$

3. *Asymptotic normality:(see Theorem 3.1 and Corollary 3.2 of [8])*

$$\sqrt{k}(\hat{\gamma}^M - \gamma) \quad \xrightarrow{d} \quad N(0, \eta^2) \quad when \quad n \to \infty,$$

or

$$\eta^2 := \begin{cases} 1+\gamma^2 & \gamma \geq 0, \\ 1-(1-2\gamma)\left(4-8\frac{1-2\gamma}{1-3\gamma}+\frac{(5-11\gamma)(1-2\gamma)}{(1-3\gamma)(1-4\gamma)}\right) & \gamma < 0, \end{cases}$$

The normality of this estimator was established by Dekkers et all..,[8] under suitable regularity conditions.

1.3 Methods of parametric estimation

These are methods that rely on the assumption that the data follow a specific, known distribution, such as the Pareto or Frechet distribution.

1.3.1 Maximum likelihood estimator

Maximum Likelihood Estimation (MLE) is a statistical method used to estimate the extreme value index based on maximum values in a sample. MLE aims to find the optimal parameter values that maximize the likelihood of the observed data under the extreme value distribution.

Let's assume we have a sample of maximum values where n is the number of observation.

1. case when $\gamma \neq 0$: the likelihood function takes the forms :

$$L = ((0, a_n, b_n); X) = -n \log a_n + (\frac{1}{\gamma}+1)\sum_{i=1}^{n} \log(1+\gamma\frac{X_i-b_n}{a_n}) - \sum_{i=1}^{n}(1+\gamma\frac{X_i-b_n}{a_n})^{\frac{-1}{\gamma}}$$

2. case when $\gamma = 0$:the likelihood function takes the forms :

$$L = ((0, a_n, b_n); X) = -n \log a_n - \sum_{i=1}^{n} \frac{X_i-b_n}{a_n} - \sum_{i=1}^{n} \exp(-\frac{X_i-b_n}{a_n})$$

According to Smith[25] (1985), the maximum likelihood estimator is consistent, meaning it converges to the true parameter values as $n \to \infty$ That is:

$$\sqrt{n}((\hat{\gamma}, \hat{a}_n, \hat{b}_n) - (\gamma, a_n, b_n)) \to N(0, J^{-1})$$

where J is the Fisher information matrix estimated by its empirical version

$$J(\theta) = -E\left[\frac{\partial^2 L(\theta; X)}{\partial \theta^2}\right]$$

$L(\theta; X)$ is the log-likelihood function associated with the law of the random variable X,θ parameterized by a set of parameters θ.

1.3.2 Weighted moment estimator

The Weighted Moment Estimator (WME) is a statistical method introduced by Hosking [15],(1985) is based on the following quantity, called the weighted moment of order r :

$$w_r := E(X\ \mathcal{H}^r_{\gamma,\mu,\sigma}(x)), r \in \mathbb{N}.$$

This quantity exists for $\gamma < 1$ and given by:

$$w_r := \frac{1}{1+r}\left\{\mu - \frac{\sigma}{\gamma}(1 - \Gamma(1-\gamma)(r+1)^{\gamma})\right\}.$$

where Γ is Euler's gamma function. In this case, three weighted moments are enough to calculate μ, σ and γ.

$$\begin{cases} \hat{w}_0(\theta) = \mu - \frac{\sigma}{\gamma}(1 - \Gamma(1-\gamma)), \\ 2\hat{w}_1(\theta) - \hat{w}_0(\theta) = \frac{\sigma}{\gamma}\Gamma(1-\gamma)(2^\gamma - 1), \\ \frac{3\hat{w}_2(\theta) - \hat{w}_0(\theta)}{2\hat{w}_1(\theta) - \hat{w}_0(\theta)} = \frac{3^\gamma - 1}{2^\gamma - 1}. \end{cases}$$

Thus by replacing respectively w_r, $r \in \{0, 1, 2\}$ by its empirical estimator

$$\hat{w}_{r,n} = \frac{1}{n}\sum_{i=1}^{n} X_{i,n}(\frac{i-1}{n})^r$$

The weighted moment estimator (WME) is obtained by solving the system of three equations

$$w_r = \hat{w}_{r,n} \qquad r = 0, 1, 2.$$

The solution to this equation is the WM estimator $\hat{\gamma}$ of γThe other parameters σ and μ are estimated respectively by:

$$\hat{\sigma} = \frac{(2\hat{w}_1 - \hat{w}_0)\hat{\gamma}}{\Gamma(1-\gamma)(2^\gamma - 1)}$$

and

$$\hat{\mu} = \hat{w}_0 + \frac{\hat{\sigma}}{\hat{\gamma}}(1 - \Gamma(1-\hat{\gamma}))$$

1.3.3 L-moment estimator

The L-moments method is a development of the traditional method of moments, and are linear combinations of order statistics. This concept was first introduced by Hosking J. R. M.in 1990[16].

Definition 1.3.1 *: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 $X_{1:r} \leq X_{2:r} \leq ...... \leq X_{r,r}$ be the order statistics associated with this sample. For $r \geq 1$, the r^{th} L-moments λ_r are given by:*

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \begin{pmatrix} r-1 \\ k \end{pmatrix} E(X_{r-k:r}), \forall r \geq 1.$$

such that $E(X_{r-k:r})$ presents the expectation of the order statistic.

In particular, for $r = 1, 2, 3, 4$, we obtain the first L-moments which are given by: The first L-moment λ_1 is used to calculate the mean (measure of position) and is defined by:

$$\lambda_1 = E(X_{1:1})$$

The second L-moment λ_2 is used to calculate (measure of dispersion or scale) and is given by:

$$\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2})$$

The third L-moment λ_3 to study symmetry (skewness measure) is given by:

$$\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})$$

The fourth L-moment, λ_4, to study kurtosis (kurtosis measure) is defined by:

$$\lambda_4 = \frac{1}{4}E(X_{4:4} - X_{3:4} + 3X_{2:4} - X_{1:4})$$

Representation of L-moment in terms of orthogonal polynomials

The L-moments can be written in terms of displaced Legendre polynomials P_r^* which are defined by:

$$P_r^*(u) = \sum_{k=0}^{r} P_{r,k}^* \; u^k = \sum_{k=0}^{r} (-1)^{r+k} \begin{pmatrix} r \\ k \end{pmatrix} \begin{pmatrix} r+1 \\ k \end{pmatrix} u^k$$

such that r is a positive integer, for $r = 0, 1, 2$ and 4 we have:

$$\begin{aligned} P_0^*(u) &= 1 \\ P_1^*(u) &= (2u - 1) \\ P_2^*(u) &= (6u^2 - 6u + 1) \\ P_3^*(u) &= (20u^3 - 30u^2 + 12u - 1) \end{aligned}$$

Then $\lambda_r$is written as:

$$\lambda_r = \int_0^1 Q(u) P_{r-1}^*(u) du \qquad r = 1, 2, ...$$

and the first L-moment are given as :

$$\begin{aligned} \lambda_1 &= \int_0^1 Q(u) du \\ \lambda_2 &= \int_0^1 Q(u)(2u - 1) du \\ \lambda_3 &= \int_0^1 Q(u)(6u^2 - 6u + 1) du \\ \lambda_4 &= \int_0^1 Q(u)(20u^3 - 30u^2 + 12u - 1) du \end{aligned}$$

Chapter 2

A weighted least-squares estimation method for Frechet distribution parameters

Estimating statistical distribution parameters is a fundamental topic in applied statistics, as these parameters play a crucial role in data analysis and modeling. However, when regression models are used for this purpose, issues such as heteroscedasticity may arise, leading to inefficient estimates. To address this issue, the Weighted Least Squares (WLS) method is often applied, where different weights are assigned to observations to enhance estimation accuracy.In this context, the distribution function is transformed into a linear regression model, enabling the use of WLS for parameter estimation. This approach has gained wide adoption., such as those by Hossain and Howlader [17] (1996), Zhang et all., [28](2007) Lu and Tao [21](2007), Zyl and Schall [29] (2012), Kantar and Arik [18](2014), Kantar and Yildirim [19](2018), and Khemissi (2022)[20].

In this chapter, we introduce fundamental concepts of linear regression, with a specific focus on applying WLS to estimate the parameters of the Frechet distribution. We will particularly consider the weighting scheme proposed by Zyl and Schall [29] to improve the accuracy and efficiency of the estimators.

2.1 Fundamentals of regression

Regression is a statistical method used to model the relationship between a dependent variable and one or more independent (explanatory) variables. When this relationship is assumed to be linear, it is called linear regression. If there is only one independent variable, it is known as simple linear regression, whereas if multiple independent variables are involved, it is called multiple linear regression.

2.1.1 Simple regression model

Simple linear regression examines the relationship between a dependent variable and a single independent variable.

Definition 2.1.1 ***(Simple linear regression model)****: A simple linear regression model is defined by an equation of the form ,*

$$Y_i = a_0 + a_1 x_i + \varepsilon_i \qquad \forall i = \overline{1, n}$$

where:

- y_i: the i^{th} observation of the random variable to be explained Y.
- x_i:the i^{th} observation of the explanatory variable X.
- a_0 and a_1:unknown contante called the model parameters.

- ε_i :the random error of the model.

- n :the sample size.

In addition, the simple linear regression model defined by 2.1.1 can be written in matrix form:

$$Y = Xa + \varepsilon \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ : \\ : \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ : & : \\ : & : \\ 1 & x_n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ : \\ : \\ \varepsilon_n \end{pmatrix}.$$

The assumptions related to this model are as follows:

The errors ε_i are centered, have the same variance, and are uncorrelated:

$$E(\varepsilon_i) = 0 \quad , \quad E(\varepsilon_i^2) = \sigma_\varepsilon^2 < \infty \quad , \quad i = 1, ..., n.$$

The error is independent of X:

$$cov(\varepsilon, X) = 0.$$

2.1.2 Least-squares (LS) method

The least squares method is used to determine the best-fitting straight line for a given dataset. The main idea is to minimize the sum of the squared differences between the actual values and the predicted values.

Definition 2.1.2 *:To estimate parameters a_0 and a_1, by minimizing the sum of the squares of the diferences between observations and model 2.1.1,Least squares are given by the following formulas :*

$$\hat{a}_1 = \frac{S_{xy}}{S_x^2} \qquad \& \qquad \hat{a}_0 = \bar{y} - \hat{a}_1\bar{x},$$

where:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad , \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

$$S_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}),$$

$$S_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \quad , \quad S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Statistical properties of LS estimates

- These estimators are unbiased estimators :

$$E(\hat{a}_1) = \hat{a}_1 \quad , \quad E(\hat{a}_0) = \hat{a}_0$$

- $var(\hat{a}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{nS_x^2}.$
- $var(\hat{a}_0) = \frac{\sigma^2 \sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{n}(1 + \frac{\bar{x}^2}{S_x^2})$
- $cov(\hat{a}_0, \hat{a}_1) = cov(\hat{a}_1, \hat{a}_0) = -\frac{\sigma^2\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = -\frac{\bar{x}^2}{nS_x^2}$

Remark 2.1.1 *:An unbiased estimator of $\sigma_\varepsilon^2$is given by :*

$$S_\varepsilon^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y})^2 = \frac{1}{n-2}\sum_{i=1}^{n}\varepsilon_i^2$$

2.1.3 Weighted least-squares (WLS) method

The Weighted Least Squares method is an extension of OLS method, used when the variance of the errors is not constant (heteroscedasticity). In such cases, assigning different weights to observations can improve the estimation accuracy.

Definition 2.1.3 *(**WLS in Simple Regression**):Consider the following model:*

$$Y_i = a_0 + a_1 x_i + \varepsilon_i$$

where $\varepsilon_i \sim N(0; \sigma^2/w_i)$*for known constants* $w_1, ..., w_n$. *The weighted least squares estimates of*a_0 *and* a_1 *minimize the quantity*

$$Q_w(a_0, a_1) = \sum_{i=1}^{n} w_i(y_i - a_0 - a_1 x_i)^2$$

Definition 2.1.4 *:To estimate parameters* a_0*and* a_1*,the WLS estimates are then given as :*

$$\hat{a}_1 = \frac{\sum_{i=1}^{n} w_i(y_i - \bar{y}_w)(x_i - \bar{x}_w)}{\sum_{i=1}^{n} w_i(x_i - \bar{x}_w)^2} \quad \& \quad \hat{a}_0 = \bar{y}_w - \hat{a}_1 \bar{x}_w.$$

where $\bar{x}_w$ *and* $\bar{y}_w$ *are the weighted means with ;*

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \quad \& \quad \bar{y}_w = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}.$$

Statistical properties of WLS estimates

- These estimators are unbiased estimators:

$$E\left(\hat{a}_1\right)=\hat{a}_1,\ E\left(\hat{a}_0\right)=\hat{a}_0.$$

- $var(\hat{a}_1)=\frac{\sigma^2}{\sum\limits_{i=1}^{n}w_i(x_i-\bar{x}_w)^2}.$
- $var(\hat{a}_0)=[\frac{1}{\sum\limits_{i=1}^{n}w_i}+\frac{\bar{x}_w^2}{\sum\limits_{i=1}^{n}w_i(x_i-\bar{x}_w)^2}].$
- The weighted error mean square $Q_w(\hat{a}_0,\hat{a}_1)/(n-2)$ also gives us an unbiased estimator of σ^2.

Definition 2.1.5 ***(General WLS Solution)****:If W is a diagonal matrix with diagonal elements $w_1,w_2,.....,w_n$, the weighted residual sum of squares is given by:*

$$\begin{aligned}Q_w(\beta)&=\sum_{i=1}^{n}w_i(Y_i-x_i^t\beta)^2,\\&=(Y-X\beta)^tW(Y-X\beta)\end{aligned}$$

The general solution to this is:

$$\hat{\beta}=(X^tWX)^{-1}X^tWY$$

Definition 2.1.6 ***(WLS as a Transformation)****:In general suppose we have the linear model*

$$Y = X\beta + \varepsilon$$

wher $var(\varepsilon) = W^{-1}\sigma^2$.*Let* $W^{1/2}$ *be a diagonal matrix with diagonal entries equal to* $\sqrt{w_i}$.*Then we have* $var(W^{1/2}/\varepsilon) = \sigma^2 I_n$. *Hence we consider the transformation*

$$\acute{Y} = W^{1/2}Y \ , \ \acute{X} = W^{1/2}X \ \textit{and} \ \acute{\varepsilon} = W^{1/2}\varepsilon$$

This gives rise to the usual least squares model

$$\acute{Y} = \acute{X}\beta + \acute{\varepsilon}$$

using the results from regular least squares we then get the solution

$$\begin{aligned} \hat{\beta} &= ((\acute{X})^t X)^{-1}(\acute{X})^t \acute{Y} \\ &= (X^t W X)^{-1} X^t W Y \end{aligned}$$

hence this is the weighted least squares solution.

Example 2.1.1 *:The data taken from Tomassone et all., [27], (1998).*

We consider data comprising 10 observations with the explanatory variable X. The variable Y is generated using the following model :

$$y_i = 3 + 2x_i + \varepsilon_i,$$

where the ε_i are normally distributed $E(\varepsilon_i) = 0, var(\varepsilon_i) = (0.2x_i)^2$,we present the data thus generated in the following table :

x_i	1	2	3	4	5	6	7	8	9	10
Y_i	4.90	6.55	8.67	12.59	17.38	13.81	14.60	32.46	18.73	20.27

Table 2.1: Values xi and Yi generated by the model studied.

A simple regression study always begins with a plot of the observations $(x_i, y_i), i = \overline{1, 10}$.

This first representation makes it possible to know if the linear model is relevant.

Graphic Representation : in figure 2.1, we plot Y_i and individuals x_i:

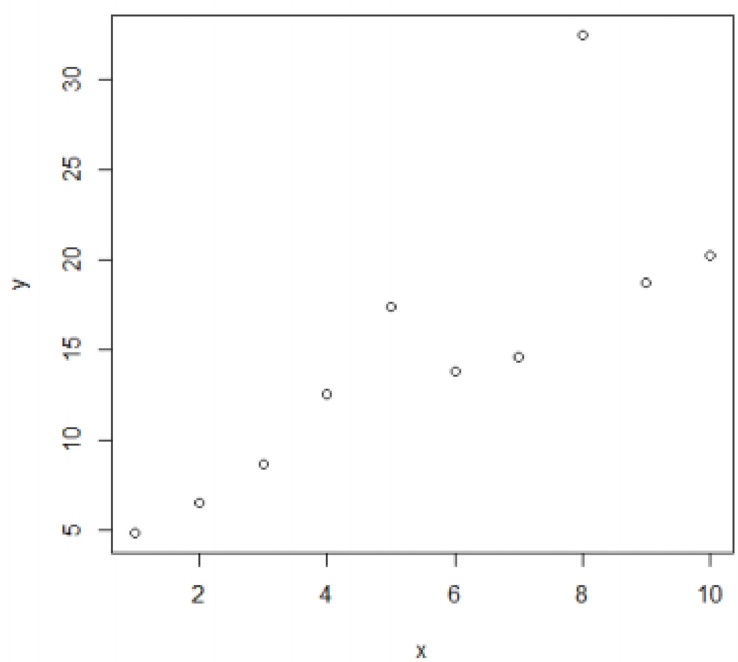

Figure 2.1: plot $Y_i$and individuals x_i

- **The least squares method :**

The least squares method provides the following estimated coefficients on the example, for all$i = \overline{1, 10}$. The regression equation is ,

$$\hat{Y}_i = 3.49 + 2.09x_i$$

The estimated slope of the line :$\hat{a}_1 = 2.09$.

The estimated y-intercept :$\hat{a}_0 = 3.49$.

Least squares regression line

We are looking for the line for which the sum of the squares of the vertical deviations of the points from the line is minimum. On the graph, we have drawn any line through the data and we represent the errors for some points, figure2.2 below illustrates the regression line by least squares.

$$R^2 = \frac{SSR}{SST} = 0.6294$$

the regression model explains$62, 94\%$ of the total variation

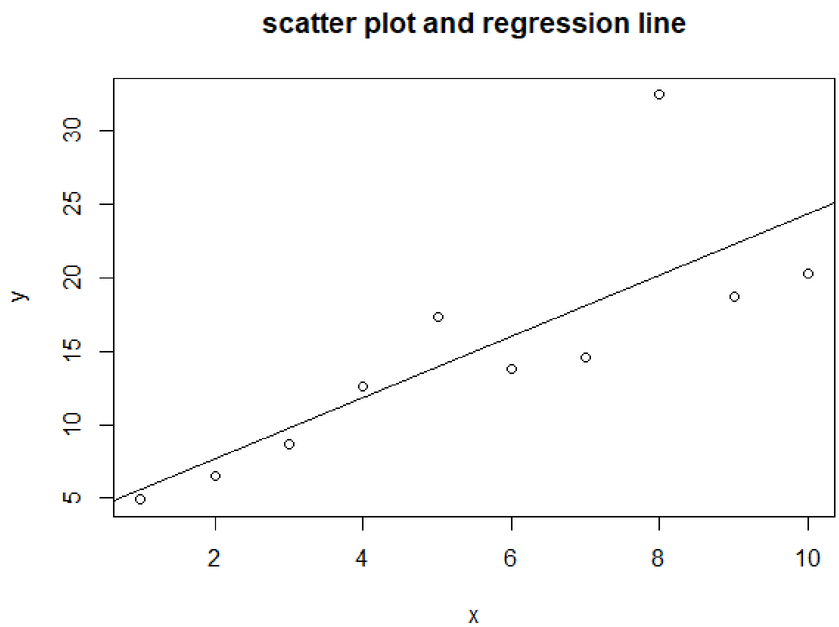

Figure 2.2: Linear regression line and scatter plot.

- **The weighted least squares method :**

A weighted regression study, using the values $(1/x^2)$ as weights. These weights are known since they must be proportional to the true variances, the occurrence equal to$(0.2x_i)^2$:

The weighted least squares method provides the following estimated coefficients on the example :

$$\hat{Y}_i = 2.53 + 2.28x_i$$

$$R^2 = \frac{SSR}{SST} = 0.8611$$

the regression model explains 86.11% of the total variation.

Remark 2.1.2 *:On this basis, the fol lowing comments can be made:*

1. All quantities related to the sum of the squares of the dependent variable are assigned by weights and are not comparable to those obtained by the least squares regression.

2. The estimated coefficients are relatively close to those of the least squares regression

2.2 Estimation of Frechet distribution parameters using regression model

When analyzing phenomena characterized by extreme or rare values, it becomes essential to employ statistical distributions capable of accurately representing such behavior. Among these, extreme value distributions have emerged as effective tools for modeling data with heavy tails and capturing the behavior of extreme events. The Frechet distribution (Extreme Value Type II) is one such distribution commonly used to model extreme events. It is particularly prevalent in engineering statistics for representing phenomena involving large maximum observations. This distribution was first introduced by the French mathematician Maurice Frechet in 1927.

2.2.1 Model transformation to linear form

The distribution functions are transformed into a linear regression model to estimate the parameter of the considered distribution.

We consider two-parameter Frechet distribution with shape parameter α, scale parameter β.The probability density function Frechet distribution is,

$$f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} \exp\left(-\left(\frac{\beta}{x}\right)^{\alpha}\right) \quad , x > 0,\ \alpha, \beta > 0 \tag{2.1}$$

and The cumulative distribution function is

$$F(x, \beta, \alpha) = \exp\left(-\left(\frac{\beta}{x}\right)^{\alpha}\right) \quad , x > 0,\ \alpha,\ \beta > 0 \tag{2.2}$$

The CDF of Frechet distribution 2.2 will be transformed to a linear function :

$$\ln(-\ln(F(x, \beta, \alpha))) = \alpha \ln \beta - \alpha \ln x \tag{2.3}$$

For a sample of size n and $x_{(1)} \leq x_{(2)} \leq ..... \leq x_{(n)}$, equation 2.3 the regression model can be rewritten as follows :

$$\ln(-\ln(F(x_{(i)}))) = \alpha \ln \beta - \alpha \ln x_{(i)} \tag{2.4}$$

where i the order number .

For estimates of $F(x_{(i)})$, Bernard and Bosi-Levenbach [2],1953 using the following methods of estimation summary in , 2.2 where $\hat{F}_i$ is some non-parametric estimate of $F(x_{(i)})$:

Method	$\hat{F}_i$
Mean Rank	$\frac{i}{n+1}$
Median Rank	$\frac{i-0.3}{n+0.4}$
Symmetric CDF	$\frac{i}{n}$

Table 2.2: Methods of estimation

comparing the equation 2.4 with

$$Y_i = \beta_1 + \beta_2 X_i$$

we get,

$$Y_i = \ln(-\ln(\hat{F}(x_{(i)}))), \quad \beta_1 = \alpha \ln \beta, \qquad \beta_2 = -\alpha, \qquad X_i = \ln x_{(i)}.$$

the regression model with error term occurs as :

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i \tag{2.5}$$

2.3 Estimators and main results

In this section, we present the methods used to estimate the shape and scall parameters of the Frechet distribution.

2.3.1 Least-squares estimator

Least squares, or least sum of squares, requires that a straight line be fitted to a set of data points, such that the sum of the squares of the distance of the points to the fitted line is minimized.

Suppose that random variables $X_1, X_2, ..., X_n$ are independent and identically distributed from the Frechet distribution. After algebraic manipulation, Equation 2.2 can be linearized as follows :

$$\ln(-\ln F(x, \beta, \alpha)) = \alpha \ln \beta - \alpha \ln x$$

The estimator of $F(x_{(i)})$ can be considered to follow the mean rank estimator :

$$\hat{F}(x_{(i)}) = \frac{i}{n+1}$$

where i is the rank of the data point in the sample in ascending order and $\hat{F}_i$ is non-parametric estimate of $F(x_{(i)};\alpha)$. See Bernard[2].

In estimation, the sum of the squares of the errors, which is defined below, should be minimized

$$\min \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_1 - \beta_2 \ln x_i)^2 \tag{2.6}$$

Therefore, the estimate the parameter α and β is given by differentiating equation 2.6 partially and equaling to zero we get:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \ln(-\ln\left(F(x_{(i)})\right)) + \hat{\alpha} \sum_{i=1}^{n} \ln x_{(i)}}{n}$$

$$\hat{\beta}_2 = \frac{n \sum_{i=1}^{n} \ln(-\ln\left(F(x_{(i)})\right)) \ln x_{(i)} - \sum_{i=1}^{n} \ln(-\ln\left(F(x_{(i)})\right)) \sum_{i=1}^{n} \ln x_{(i)}}{\left(n \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2 - \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2\right)}$$

Finally, estimates and of the parameter $\hat{\beta}_{LSE}$ and $\hat{\alpha}_{LSE}$ are given as:

$$\hat{\alpha}_{LSE} = -\left(\frac{n \sum_{i=1}^{n} \ln(-\ln\left(F(x_{(i)})\right)) \ln x_{(i)} - \sum_{i=1}^{n} \ln(-\ln\left(F(x_{(i)})\right)) \sum_{i=1}^{n} \ln x_{(i)}}{\left(n \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2 - \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2\right)}\right)$$

$$\hat{\beta}_{LSE} = \exp\left(\frac{\sum_{i=1}^{n} \ln(-\ln\left(F(x_{(i)})\right)) + \hat{\alpha}_{LSE} \sum_{i=1}^{n} \ln x_{(i)}}{n\hat{\alpha}_{LSE}}\right)$$

2.3.2 Weighted least-squares estimator

One of the main advantages of using regression approach for estimating the parameters of the Frechet distribution is its simplicity. Given a sample $x_{(1)} \leq x_{(2)} \leq .... \leq x_{(n)}$ from a Frechet distribution , we can use the following regression model:

$$Y_i = \beta_1 + \beta_2 X_i \tag{2.7}$$

where Y_i is derived from the empirical distribution function, and $X_i = \ln x_i$. To estimate the parameters $\beta_1$and β_2, we minimize the weighted sum of squared errors:

$$\min \sum_{i=1}^{n} \varepsilon_i^2 = \min \sum_{i=1}^{n} w_i \left[y_i - \beta_1 - \beta_2 \ln x_i\right]^2 \tag{2.8}$$

Where $w_i$is the weighted factor , $i = 1, 2, ..., n$.

Next to,derived from the inverse of the asymptotic variances of the order statistics. These weights stabilize the variance, ensuring efficient estimation as proposed by Zyl (2012) [29], specified in the following formula:

$$var(\Lambda(x_{(i)})) \approx \frac{m_i(1 - m_i)}{(n + 2)(f(x_i))^2} \left[\frac{d\Lambda(x_{(i)})}{dx_{(i)}}\right]^2_{x_{(i)=x_i}} \qquad i = \overline{1, n} \tag{2.9}$$

Now, Applying this expression for estimation of parameters of frechet distributions, yield the following results:

Let $\Lambda(x_{(i)}) = \ln[-\ln F(x_{(i)}]$ and $\mu_i = E[\Lambda(x_{(i)})]$

$$\begin{aligned}
\Lambda(x_{(i)}) &= \alpha \ln \beta - \alpha \ln x_{(i)} \\
\Lambda(x_{(i)}) + (\mu_i - \mu_i) &= \alpha \ln \beta - \alpha \ln x_{(i)} \\
\mu_i &= \alpha \ln \beta - \alpha \ln x_{(i)} + (\mu_i - \Lambda(x_{(i)})) \\
\mu_i &= \alpha \ln \beta - \alpha \ln x_{(i)} + u_i
\end{aligned}$$

Where $u_i = \mu_i - (\alpha \ln \beta - \alpha \ln x_{(i)})\ \ i = \overline{1, n}$, are the residuals for the regression and the weights are the inverses of the variances of the residuals.

The approximate variance of $\ln(-\ln F(x_{(i)}))$ by 2.9,is:

$$\begin{aligned}
var(\ln(-\ln F(x_{(i)})) &\approx \frac{m_i(1-m_i)}{(n+2)(f(x_i))^2}\left[\frac{d\ln(-\ln(F(x_{(i)}))}{dx_i}\right]^2_{x_{(i)}=x_i} \\
&\approx \frac{m_i}{(n+2)(m_i \ln(m_i))^2(1-m_i)} \qquad , m_i = \frac{i}{n+1} \\
&\approx \frac{(n+1-i)}{(n+2)i(\ln\left(\frac{i}{n+1}\right))^2}
\end{aligned}$$

For this reason, the WLS regression equation is solved by letting :$\hat{\beta}_{WLS} = (X^tWX)^{-1}X^tWY$,

$$X = \begin{pmatrix} 1 & \dots & \dots & 1 \\ \ln(x_{(1)}) & \dots & \dots & \ln(x_{(i)}) \end{pmatrix}^t$$

$$Y^t = \left(\ln(-\ln(\hat{F}_{(1)})), \dots\dots\dots, \ln(-\ln(\hat{F}_{(i)}))\right) \quad \text{and}$$

$$w = diag(w_1, w_2, \dots\dots., w_n)$$

$$w_i = \frac{(n+2)i(\ln\left(\frac{i}{n+1}\right))^2}{(n+1-i)} \qquad , \quad i = 1, \dots., n$$

Therefore, the estimate the parameter α and β is given by differentiating equation 2.8 partially and equaling to zero we get:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} w_i \ln(-\ln\left(F(x_{(i)})\right)) + \hat{\alpha}\sum_{i=1}^{n} w_i \ln x_{(i)}}{\sum_{i=1}^{n} w_i}$$

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i \ln(-\ln\left(F(x_{(i)})\right)) \ln x_{(i)} - \sum_{i=1}^{n} w_i \ln(-\ln\left(F(x_{(i)})\right)) \sum_{i=1}^{n} \ln x_{(i)}}{\left(\sum_{i=1}^{n} w_i \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2 - \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2\right)}$$

Finally, estimates and of the parameter $\hat{\beta}_{WLS}$ and $\hat{\alpha}_{WLS}$ are given as:

$$\hat{\alpha}_{WLS} = -\left(\frac{\sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i \ln(-\ln\left(F(x_{(i)})\right)) \ln x_{(i)} - \sum_{i=1}^{n} w_i \ln(-\ln\left(F(x_{(i)})\right)) \sum_{i=1}^{n} w_i \ln x_{(i)}}{\left(\sum_{i=1}^{n} w_i \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2 - \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2\right)}\right)$$

$$\hat{\beta}_{WLS} = \exp\left(\frac{\sum_{i=1}^{n} w_i \ln(-\ln\left(\hat{F}(x_{(i)})\right)) + \hat{\alpha}_{WLS}\sum_{i=1}^{n} w_i \ln x_{(i)}}{\hat{\alpha}_{WLS}\sum_{i=1}^{n} w_i}\right)$$

$$w_i \approx 1/var\left(\ln[-\ln F(x_{(i)}]\right)$$

Chapter 3

Simulation study

In this chapter examine the performance of three methods for estimating the parameters of the Frechet distribution: the least lquares (LS) method, the weighted least squares (WLS) method, and the maximum likelihood estimation (MLE) method, using Monte Carlo simulation.

3.1 Performance of the estimators

We conducted a Monte Carlo study using 10000 randomly generated samples for each scenario, with various sample sizes: $10, 20, 30, 50, 100, 200, 500, 1000$, and 2000 for Frechet distribution, and the results are presented in tables:3.1 , 3.2:shape parameters ($\alpha = 0.5$; $1/0.6$) with two scale parameters ($\beta = 1$; 2). and tables: 3.3, 3.4: scale parameters ($\beta = 0.5$; 1.5) with two shape parameters ($\alpha = 1$; 2).

The efficiency of the estimation methods was assessed using two statistical measures:the bias, defined as the difference between the expected value of the estimator and the true parameter value : $Bias(\hat{\alpha}) = E(\hat{\alpha}) - \alpha$,and the mean squared error , which combines the variance of the estimator and the root square of its bias:$RMSE(\hat{\alpha}) = \sqrt{Var(\hat{\alpha}) + (Bias(\hat{\alpha}))^2}$.

3.2 Results and discussion

According to the Bias criterion:

In the estimation of α , the WLS estimator shows the lowest bias among the three methods, indicating its superior performance in small samples. The LS estimator shows a higher bias, reflecting its sensitivity due to the lack of weighting. The MLE estimator displays a moderate bias, which may be due to the small sample size and numerical instability during the optimization process.

As for the estimation of β, a similar pattern is observed: the WLS estimator provides the lowest bias, confirming its efficiency and stability under small and large-sample conditions. In contrast, the LS estimator shows a higher bias, while the MLE also exhibits noticeable bias due to the fixed value of during optimization and the limited sample size.

In addition, bias decreases with increasing sample size and shape and scall parameters cases.

According to the Root Mean Squared Error (RMSE) criterion:

In the estimations of α and β, the WLS estimator achieves the lowest $RMSE$ than the least square estimation method for each sample size, in addition the $RMSE$ of the MLE is the smallest. As sample size increases the root mean square error decreases for each methods and and shape parameter cases, and thus conclude that there are accurate parameter increments.

n	Method	$\beta = 1$		$\beta = 2$	
		Biais	*RMSE*	*Biais*	*RMSE*
10	WLS	0.37819	0.47505	0.38012	0.97727
	LS	0.98365	1.65191	0.73631	1.16336
	MLE	0.72451	0.99468	0.81430	1.01413
30	WLS	0.21712	0.27207	0.22013	0.27489
	LS	0.39307	0.81559	0.39566	0.61546
	MLE	0.63879	0.53201	0.39870	0.53424
50	WLS	0.16877	0.21180	0.16876	0.21104
	LS	0.29332	0.46678	0.29858	0.72975
	MLE	0.48145	0.39070	0.68321	0.39624
100	WLS	0.11998	0.04442	0.11919	0.14910
	LS	0.19987	0.26056	0.20283	0.85137
	MLE	0.35045	0.15015	0.55058	0.26655
200	WLS	0.08313	0.03938	0.08367	0.79765
	LS	0.13699	0.17676	0.13363	0.17412
	MLE	0.38116	0.10428	0.53341	0.10489
1000	WLS	0.03755	0.03684	0.03730	0.04914
	LS	0.05673	0.07184	0.05739	0.07285
	MLE	0.29610	0.04687	0.05591	0.04660
2000	WLS	0.02649	0.03368	0.02658	0.03162
	LS	0.03899	0.04919	0.03923	0.04964
	MLE	0.03189	0.03321	0.03453	0.03331

Table 3.1: Simulated bias and RMSE when shape parameter(α=0.5)

n	Method	$\beta = 1$		$\beta = 2$	
		Bais	*RMSE*	*Biais*	*RMSE*
10	WLS	0.31879	0.39864	0.94736	0.60411
	LS	0.76495	0.955391	0.61050	0.83877
	MLE	0.60984	0.83416	0.31883	0.40127
30	WLS	0.18462	0.23021	0.29693	0.30160
	LS	0.32575	0.43519	0.32812	0.44023
	MLE	0.21886	0.23137	0.18208	0.22926
50	WLS	0.14000	0.17579	0.48361	0.17714
	LS	0.35822	0.37431	0.25019	0.50616
	MLE	0.24726	0.32793	0.14060	0.33095
100	WLS	0.07879	0.12341	0.94271	0.15959
	LS	0.16654	0.21692	0.16823	0.22036
	MLE	0.09854	0.17788	0.09941	0.13507
200	WLS	0.01396	0.08772	0.18617	0.08766
	LS	0.11402	0.14638	0.11388	0.10992
	MLE	0.07013	0.14079	0.06978	0.14682
1000	WLS	0.04160	0.06006	0.03086	0.05916
	LS	0.04735	0.06895	0.04737	0.06027
	MLE	0.03102	0.03891	0.03119	0.03919
2000	WLS	0.02239	0.02779	0.02412	0.02778
	LS	0.03284	0.04144	0.03318	0.04168
	MLE	0.02213	0.02963	0.02223	0.03449

Table 3.2: Simulated bias and RMSE when shape parameter(α=1/0.6)

n	Method	$\alpha = 1$		$\alpha = 2$	
		$Bais$	$RMSE$	$Bais$	$RMSE$
10	WLS	0.13939	0.18893	0.30157	0.42791
	LS	0.42105	0.20475	0.45459	0.49921
	MLE	0.14341	0.19149	0.34094	0.58587
30	WLS	0.07508	0.09064	0.15572	0.20511
	LS	0.07908	0.10031	0.15820	0.21126
	MLE	0.07637	0.09970	0.15932	0.22128
50	WLS	0.04996	0.06910	0.19986	0.14998
	LS	0.06256	0.07892	0.12307	0.15970
	MLE	0.05837	0.07542	0.12040	0.16144
100	WLS	0.03999	0.04999	0.08299	0.10149
	LS	0.04404	0.05523	0.08705	0.11086
	MLE	0.04064	0.05165	0.08352	0.10848
200	WLS	0.02536	0.03514	0.55391	0.05134
	LS	0.03159	0.03942	0.06147	0.07778
	MLE	0.02887	0.03640	0.05725	0.07368
1000	WLS	0.01135	0.01504	0.02251	0.02995
	LS	0.01367	0.01712	0.02778	0.03472
	MLE	0.01258	0.01588	0.02562	0.03215
2000	WLS	0.00685	0.01055	0.01597	0.02051
	LS	0.00977	0.01221	0.01932	0.02434
	MLE	0.00898	0.01126	0.01802	0.02265

Table 3.3: Simulated bias and RMSE when scale parameter β=0.5)

n	Method	$\alpha = 1$		$\alpha = 2$	
		Bais	*RMSE*	*Bais*	*RMSE*
10	WLS	0.41989	0.32721	0.89506	0.87052
	LS	0.42300	0.57731	0.88965	0.92597
	MLE	0.43762	0.62281	0.90508	0.78606
30	WLS	0.21726	0.28476	0.47100	0.64048
	LS	0.23748	0.30029	0.48173	0.64701
	MLE	0.22691	0.29728	0.48262	0.67121
50	WLS	0.16491	0.22049	0.36063	0.48502
	LS	0.18311	0.23067	0.37021	0.48397
	MLE	0.17179	0.22127	0.49952	0.49988
100	WLS	0.12037	0.49998	0.23645	0.31685
	LS	0.12900	0.16190	0.25870	0.32940
	MLE	0.14999	0.15239	0.24688	0.32038
200	WLS	0.08115	0.10115	0.16581	0.23007
	LS	0.09291	0.11668	0.18343	0.23036
	MLE	0.08573	0.10815	0.17198	0.21836
1000	WLS	0.03115	0.04515	0.05364	0.09952
	LS	0.04093	0.05123	0.08193	0.10250
	MLE	0.03772	0.04730	0.07590	0.09526
2000	WLS	0.02001	0.02115	0.04215	0.06587
	LS	0.02896	0.03627	0.05817	0.07255
	MLE	0.02649	0.03325	0.05388	0.06741

Table 3.4: Simulated bias and RMSE when scale parameter (β=1.5)

Conclusion

This thesis addresses the estimation of the parameters of the Frechet distribution using the Weighted Least Squares (WLS) method, which was developed to handle heteroscedasticity that affects the efficiency of Least Squares (LS) estimators. Previous studies supporting the use of WLS were reviewed, emphasizing the importance of selecting an appropriate weighting function to ensure variance stability and improve estimation accuracy.

The proposed methodology integrates previous concepts for determining suitable weights, particularly the approach introduced by Zyl & Schall (2012), which enhanced estimation performance. Simulation results demonstrated that the proposed method provides higher accuracy and efficiency compared to traditional methods.

Accordingly, the proposed approach can be considered an alternative for estimating the parameters of statistical distributions, with potential for future development by integrating it with other estimation techniques , it also can be applied to real and censured data.

Bibliography

[1] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. 2004. Statistics of extremes :theory and applications. Wiley series in probability and statistics. Chichester.

[2] Bernard, A., Bosi-Levenbach, E.C. 1953. The plotting of observations on probability paper. Stat. Neederlandica, No :7, page :163 − 173.

[3] Bingham, N., Goldie, C., Teugels, J. 1987. Regular variation. Cambridge university press.

[4] Csörgö, S., Deheuvels, P., Mason, D. 1985. Kernel estimates of the tail index of a distribution.The analysis of statistics, page:1050 − 1077

[5] De Haan, L., Ferreira, A.2006. Extreme value theory : An introduction. Springer., Vol: 21.

[6] De Haan, L., Stadtmüller, U. 1996. Generalized regular variation of second order. Jornal of australian math. Soc. (Series A) No:61, page :381 − 395.

[7] Deheuvels, P., Haeusler, E., Mason, D.M. 1988. Almost sure convergence of the hill estimator. Math. Proc. Cambridge philos. Soc., Vol:104,No:2, page:371 − 381.

[8] Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L. 1989. A moment estimator for the index of an extreme-value distribution. Analysis of statistics. No:17, page:1833 − 1855.

[9] Drees, H. 1995. Refined pickands estimators of the extreme value index. The analysis of statistics, page: 2059 − 2080.

[10] Falk, M. 1995. Some best parameter estimates for distributions with finite endpoint. Statistics : A journal of theoretical and applied statistics. Vol:27, No:2, page:115 − 125.

[11] Fisher, R., Tippet, L. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings of the cambridge, philosophical society,No:24, page: 180 − 190.

[12] Gardes, L., Girard, S., Lekina, A. 2010. Functional nonparametric estimation of conditional extreme quantiles. Journal of multivariate Anal. Vol:101, No:2, page:419 − 433.

[13] Gnedenko, B. 1943. Sur la distribution limite du terme maximum d'une série aléatoire. The annals of mathematics, Vol:44, No:3; page: 423 − 453.

[14] Hill, B.M. 1975. A simple general approach to inference about the tail of a distribution.Ann. Statist.,No:3, page:1163 − 1174.

[15] Hosking, J. 1985. Algorithm as 215 : maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. Journal of the royal statistical society. Series C (Applied statistics), Vol:3 ,No:34; page: 301 310

[16] 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 (Methodological), 52(1), 105–124.

[17] Hossain, A. Howlader, H.A. 1996. Unweighted least squares estimation of weibull parameters. Journal of statistical computation and simulation, No:54, page:$265 - 271$

[18] Kantar, Y.M., Arik, I. 2014. M-estimation of log-logistic distribution parameters with outliers. International journal of agricultural and statistical sciences.

[19] Kantar, Y.M., Yildirim, V. 2015. Robust estimation for parameters of the extended burr type III distribution. Communications in statistics ñtheory and methods, simulation and computation,Vol:44, No:7, page: $1901 - 1930$.

[20] Khemissi, Z., Brahimi, B., & Benatia, F. (2022). Heavy Tail Index Estimator through Weighted Least-squares Rank Regression. Journal of Siberian Federal University. Mathematics & Physics, 15(6), 797–805. https://doi.org/10.17516/1997-1397-2022-15-6-797-805

[21] Lu, H.L., Tao, S.H. 2007. The estimation of pareto distribution by a weighted least square method. Quality & Quantity, No:41, page: $913 - 926$. Springer

[22] Mason, D.M. 1982. Laws of large numbers for sums of extreme values. Ann.Probab,Vol:10, No:3,page:$754 - 764$.

[23] Necir, A. 2006. A functional law of the Iterated logarithm for kernel type estimators of the tail index. Journal of statistical planning and inference No: 136, page:$780 - 802$.

[24] Pickands, J. 1975. Statistical inference using extreme order statistics. The annals of statistics, Vol :3, No: 1,page: 119 − 131.

[25] Resnick, S. 1987. Extreme values, regular variation, and point processes. Springer verlag.

[26] Smith, R. 1987. Estimating tails of probability distributions. The annals of statistics, Vol: 15, No: 3, page :1174 − 1207.

[27] Tomassone R., Charles-Bajard S., Bellanger L. 1998. Introduction à la planifcation expérimentale, DEA" Analyse et modélisation des systémes biologiques" .Fiche pratique splus 2 /05/01/99.

[28] Zhang, L.F., Xie, M., Tang, L.C. 2007. A study of two estimation approaches for parameters of weibull distribution based on WPP, Reliab. Eng. Syst. Safe.,No: 92, page:360 − 368..

[29] Zyl, J.M.V., Schall, R. 2012. Parameter estimation through weighted least-squares rank regression with specific reference to the weibull and gumbel distributions. Communications in statistics-simul. No: 41, page: 1654 − 1666.

Annex A: R Software

3.3 What is the R language?

R is a system and programming language specifically designed for conducting statistical analyses. It is characterized by its ability to process data, perform calculations, and create various graphical representations. R also offers the capability to run stored programs (packages) to execute advanced statistical procedures such as linear and non-linear models, time series, parametric and non-parametric tests, and methods for analyzing multidimensional data.

Code R:

chapter 2 : example 2.1.1

 # Data from the table

```
x<-c(1,2,3,4,5,6,7,8,9,10)

y<-c(4.90,6.55,8.67,12.59,17.38,13.81,14.60,32.46,18.73,20.27)

plot(x,y)
```

 # LS estimation

```
lm1<-lm(y~x)

lm1
```

```
abline(lm1)
title("scatter plot and regression line")
summary(lm1)
```

WLS estimation

```
lm2<-lm(y~x,weights=(1/x^2))
lm2
summary(lm2)
```

chapter 3 :

#Simulation of shape parameter estimation α

```
library(evd)
Fchapeau <- function(x) {rank(x) / (length(x) + 1)}
```

WLS estimator

```
wls <- function(x) {
y <- log(x)
n <- length(x)
t <- 1:n
w <- ((n + 1 - t) / (n + 2)) * t * (log(t / (n + 1)))^2
f <- log(-log(Fchapeau(x)))
alpha <- -(sum(w) * sum(w * f * y) - sum(w * f) * sum(w * y)) / (sum(w) *
sum((w * y)^2) - (sum(w * y))^2)
return(1 / alpha)
}
```

 # LS estimator

```
ls <- function(x) {
y <- log(x)
n <- length(x)
f <- log(-log(Fchapeau(x)))
alpha <- -(n * sum(f * y) - sum(f) * sum(y)) / (n * sum(y^2) - (sum(y))^2)
return(1 / alpha)
}
```

 # simulation Function

```
my_sample <- function(n, alpha, beta) {
T <- numeric(3)
x <- rfrechet(n, loc = 0, scale = beta, shape = 1 / alpha)
# MLE via optimization
logL <- function(alpha_est) {
-sum(log(dfrechet(x, loc = 0, scale = beta, shape = 1 / alpha_est)))
}
opt <- optimize(logL, c(0.01, 5))
T[1] <- wls(x)
T[2] <- ls(x)
T[3] <- opt$minimum
return(T)
}
```

 # Matrix to store the estimators

```
n <- 2000 # sample size

M <- 10000 # numbre of iterations

alpha <- 1/0.6

beta <- 1

B <- matrix(0, nrow = M, ncol = 3)

for (j in 1:M) {

print(j)

A <- my_sample(n, alpha, beta)

B[j, 1] <- A[1] # WLS

B[j, 2] <- A[2] # LS

B[j, 3] <- A[3] # MLE

}

# Bias and RMSE calculation

bias <- means(abs(B - alpha))

rmse <- sqrt(means((B - alpha)^2))

# Show results

cat("Bias WLS =", bias[1], "\n")

cat("RMSE WLS =", rmse[1], "\n\n")

cat("Bias LS =", bias[2], "\n")

cat("RMSE LS =", rmse[2], "\n\n")

cat("Bias MLE =", bias[3], "\n")

cat("RMSE MLE =", rmse[3], "\n")
```

Annex B:Abreviations and notations

The following is an explanation of the various abbreviations and notations which are in use throughout this report:

(Ω, A, P)	probability space
rv	random variable
$i.i.d$	Independent andidentically distributed
X	rv dened on (Ω, A, P) , population
$E[X]$	expectation of (or mean of X)
$Var(X)$	variance of X
pdf	probabilty density function
df	distribution function
F_n	empirical df
$F^{\leftarrow}$	generalized inverse of F,quantile function
Q	quantile function ,generalized inverse of X
Q_n	empirical quantile function.

$X_{1,n} \leq .... \leq X_{n,n}$	order statistics pertaining to the sample$(X_1, ..., X_n)$
k	numbers of top statistics (upper observations)
x_F	upper endpoint
EVI	extreme value index
EVT	extreme value theory
$GEVD$	generalized extreme value distribution
GPD	generalized Pareto distribution
$\mathcal{D}(.)$	domain of attraction
$\mathcal{RV}_\rho$	regular variation at ∞ with index ρ
$\mathcal{RV}_0$	regular variation at 0 with index ρ
$\overset{a.s}{\longrightarrow}$	almost sure convergence
$\overset{p}{\rightarrow}$	convergence in probability
$\overset{d}{\rightarrow}$	convergence in distribution
$i.e$	in other words
α	tail index
γ	extreme value index
MLE	maximum likelihood estimator
$\exp$ or e	exponential
$\log$	logarithm
LSE	least squares estimator
WLS	weighted least squares
$RMSE$	root mean square ereur

Abstract

This thesis is devoted to studing the estimation of the Frechet distribution parameters using the weighted least squares method, as an alternative, effective, and more efficient approach than traditional method, especially in the presence of the problem of heteroscedasticity.

The main objective of this thesis is to improve the accuracy and efficiency of estimating distribution parameters by proposing appropriate weights based on previous concepts (Zyl and Schall (2012)).

An **R** simulation study was conducted to test the performance of the proposed method, and the results showed its superiority in terms of efficiency and accuracy.

Keywords : Frechet distribution, tail index estimation, weighted least squares, shape parameter, scale parameter .

Résumé

Cette thèse est consacrée à l'étude de l'estimation des paramètres de la distribution de Fréchet par la méthode des moindres carrés pondérés, une approche alternative, efficace et plus performante que les méthodes traditionnelles, notamment face au problème d'hétéroscédasticité.

L'objectif principal de cette thèse est d'améliorer la précision et l'efficacité de l'estimation des paramètres de distribution en proposante des pondérations appropriées sur des concepts antérieurs (Zyl et Schall (2012)).

Une étude de simulation **R** a été réalisée pour tester les performances de la méthode proposée, et les résultats ont montré sa supériorité en termes d'efficacité et de précision.

Mots clés : distribution de Fréchet, estimation de l'indice de queue, paramètre de forme, paramètre d'echelle.

ملخـــــــص

الأطروحة مخصصة لدراسة تقدير معلمات توزيع فريشيه باستخدام طريقة المربعات الصغرى الموزونة، كنهج بديل وفعال وأكثر كفاءة من الطرق التقليدية. خاصة في ظل وجود مشكلة عدم تجانس التباين.

الهدف الرئيسي من هذه الرسالة هو تحسين دقة وكفاءة تقدير معلمات التوزيع من خلال اقتراح أوزان مناسبة مستندة الى مفاهيم سابقة (زيل وشال 2012).

تم اجراء دراسة محاكاة باستخدام **R** لاختبار أداء الطريقة المقترحة، واظهرت النتائج تفوقها من حيث الكفاءة والدقة.

الكلمات المفتاحية: توزيع فريشيه، تقدير مؤشر الذيل، المربعات الصغرى الموزونة، معلمات المقياس، معلمات الشكل.