People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

MOHAMED KHIDER UNIVERSITY, BISKRA

FACULTY of EXACT SCIENCES, SIENCE of NATURE and LIFE

DEPARTMENT of MATHEMATICS

A thesis presented for the degree of

Doctor in Mathematics

In The Field of Statistics

By

Nawel HAOUAS

Title :

Contribution to statistics of rare events of incomplete data

Examination Committee Members

Pr.	**Djamel MERAGHNI**	University of Biskra	Chairman
Pr.	**Abdelhakim NECIR**	University of Biskra	Supervisor
Pr.	**Brahim BRAHIMI**	University of Biskra	Examinator
Dr.	**Djabrane YAHIA**	University of Biskra	Examinator
Dr.	**Fateh BENATIA**	University of Biskra	Examinator
Dr.	**Halim ZEGHDOUDI**	University of Annaba	Examinator

December 12, 2017.


بسم الله الرحمن الرحيم

This work is dedicated to:

The memory of my dear brother: Azzedine "Azzo".

My dear parents.

My dear brothers and sisters.

To my dear husband.

and

To my dear daughter

Tasnime.

Acknowledgement

First, I would like to thank God for his blessing and guidance in the process of completing this research work.

I would like to express my deepest sense of gratitude to my advisor, Pr. Necir Abdelhakim. His patience and support helped me overcome many diffirent situations and finish this dissertation. I could not have imagined having a better advisor and mentor for my Ph.D. study. I would like also to express my gratitude to Pr. Meraghni Djamel and Pr. Brahimi Brahim for their time and effort to help me.

I would like to thank the examination committee members for having devoted some of their precious time to very careful readings of this thesis and for their comments, criticisms and suggestions.

I would like also to express my gratitude to all my teachers. Many thanks go in particular, to the head of Department of Mathematics and my advisor in the memory of Master: Dr. Djabrane Yahia and my advisor in the memory of License: The Professor Hafidhi Nadjibe.

I would like to thank my parents for their support and encouragement. Without them, I may never have gotten to where I am now. Sister's and Brather's, thanks for being supportive and caring siblings. My spesial thanks go to my dear husband Samir, has taken the load off my shoulder with all the love and released welcomed.

Achieved Works

Papers

- Estimating the second-order parameter of regular variation and bias reduction in tail index estimation under random truncation. Submitted. (with A. Necir and B. Brahimi, 2017).
- A Lynden-Bell integral estimator for the tail index of right-truncated data with a random threshold. Journal Afrika Statistika, 12(1); 1159–1170 (with A. Necir, D. Meraghni and B. Brahimi, 2017).
- Kernel estimate of the tail index based on log probability weighted moments (with A. Necir 2017).

Contents

Acknowledgement **4**

Achieved Works **5**

Table of Contents **6**

List of Figures **10**

List of tables **11**

General Introduction **1**

1 Extreme values **5**

1.1 Introduction 5

1.2 Concepts and definitions 5

1.3 Limite theorems 8

1.3.1 Laws of large numbers 8

1.3.2 Central Limit Theorem 9

1.4 Order statistics 10

1.5 Limit distributions of maxima 14

1.6 Generalized extreme value distribution (GEVD) 15

1.7 Regular variation function 18

1.7.1 First Order Regular Variation Assumption 20

1.7.2 Second Order Regular Variation Assumption 20
1.7.3 Third Order Regular Variation Assumption 21
1.8 Characterization of the Domain of Attraction 22
1.8.1 Domain of attraction of Fréchet 24
1.8.2 Domain of attraction of Weibull 25
1.8.3 Domain of attraction of Gumbel 26
1.8.4 General Characterizations 27

2 Tail index, extreme quantile and second-order parameter estimation 30

2.1 Introduction 30
2.2 Tail index estimation 31
2.2.1 Hill's estimator ($\gamma > 0$) 31
2.2.2 Pickands estimator 33
2.2.3 Moment estimator 35
2.2.4 Kernel type estimators 36
2.2.5 Probability-weighted moment estimator 38
2.3 Choice of optimal number of extremes k_n 40
2.3.1 Graphical method 41
2.3.2 Method bassed on the mean squared error 42
2.3.3 Numerical procedure 42
2.4 Quantile extreme estimation 44
2.4.1 Approach based on a positive index estimator 45
2.4.2 Approach based on a negative index estimator 46
2.4.3 Approach based on any index estimator 46
2.5 Second order parameter estimation 46

3 Incomplete data 48

3.1 Censored data 48

3.1.1 Types of censoring . . . 49
3.1.2 Right censoring . . . 50
3.1.3 Left censoring . . . 51
3.1.4 Interval censoring . . . 51
3.2 Truncated data . . . 51
3.2.1 Left truncation . . . 52
3.2.2 Right truncation . . . 52
3.2.3 Interval truncating . . . 53
3.3 Estimation in the case of incomplete data . . . 53
3.3.1 Estimation under random right censoring model . . . 53
3.3.2 Estimation under random right truncating model . . . 56

4 Estimating the second-order parameter of regular variation and bias reduction in tail index estimation under random truncation **59**

4.1 **Introduction** . . . 59
4.2 **Estimating the second-order parameter** ρ_1 . . . 62
4.3 **Reduced-bias tail index estimator** . . . 66
4.4 **Main results** . . . 67
4.5 **Simulation study** . . . 70
4.5.1 Automatic choice of the number of upper extremes . . . 71
4.5.2 Graphical diagnostics of the number k of upper extremes . . . 72
4.6 **High quantile estimation** . . . 81
4.7 **Real data example** . . . 83
4.8 **Proofs** . . . 84
4.8.1 **Proof of Theorem** 4.1 . . . 84
4.8.2 **Proof of Theorem** 4.2 . . . 90
4.8.3 **Proof of Theorem** 4.3 . . . 92

4.9 **Conclusion** . . . 92

4.10 **Appendix** . . . 92

5 A Lynden-Bell integral estimator for the tail index of right-truncated data with a random threshold **96**

5.1 **Introduction** . . . 96

5.2 **Main results** . . . 101

5.3 **Simulation study** . . . 102

5.4 **Proofs** . . . 106

5.4.1 Proof Theorem 5.1 . . . 106

5.4.2 Proof Theorem 5.2 . . . 109

5.4.3 Proof of Theorem 5.3 . . . 111

6 Kernel weighted moment estimator of extreme quantile of complete data **112**

6.1 Introduction . . . 112

6.2 **Main results** . . . 115

6.3 **Proofs** . . . 116

6.3.1 Proof of theorem 6.1 . . . 116

6.3.2 Proof of theorem 6.2 . . . 119

Conclusion **121**

Bibliography **123**

Appendix A **132**

Appendix B **134**

Abstract **i**

List of Figures

1.1 Densities of the standard extreme value distributions. We chose $\alpha = 1$ for the Frechet and the Weibull distribution. 17

2.1 Plot of Hill's estimator, for the EVI of a standard Pareto distrubution, as a function of the number of top statistics, based on 100 samples of size 3000. The horizontal line represents the true value of the tail index. 41

4.1 S1: $\gamma_1 = 0.6$, $p = 70\%$ 73

4.2 S2: $\gamma_1 = 0.6$, $p = 70\%$ 74

4.3 S4: $\gamma_1 = 0.6$, $p = 70\%$ 75

4.4 S3: $\gamma_1 = 0.6$, $p = 70\%$ 76

4.5 S1: $\gamma_1 = 0.6$, $p = 90\%$ 77

4.6 S2: $\gamma_1 = 0.6$, $p = 90\%$ 78

4.7 S3: $\gamma_1 = 0.6$, $p = 90\%$ 79

4.8 S4: $\gamma_1 = 0.6$, $p = 90\%$ 80

List of Tables

1.1 Some examples of laws from the three domains of attraction 27

2.1 Some example for the kernel function 37

4.1 Absolute biases and rmse's of the second-order parameter estimator based on 1000 right-truncated samples from Burr's models. 72

4.2 Absolute biases and rmse's of the tail index estimators based on 1000 right-truncated samples from Burr's models. 81

4.3 Absolute biases and rmse's of the second-order parameter estimator based on 1000 right-truncated samples from Fréchet's models. 82

4.4 Absolute biases and rmse's of the tail index estimators based on 1000 right-truncated samples from Fréchet's models. 83

4.5 Extreme quantiles for car brake pad lifetimes. 84

5.1 Estimation results of Lynden-Bell based (leftt pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.6$ of Burr's model through 1000 right-truncated samples with 45%-truncation rate. 103

5.2 Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.6$ of Burr's model through 1000 right-truncated samples with 30%-truncation rate. 103

5.3 Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.6$ of Burr's model through 1000 right-truncated samples with 10%-truncation rate. 104

5.4 Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.8$ of Burr's model through 1000 right-truncated samples with 45%-truncation rate. 104

5.5 Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.8$ of Burr's model through 1000 right-truncated samples with 30%-truncation rate. 105

5.6 Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.8$ of Burr's model through 1000 right-truncated samples with 10%-truncation rate. 105

General introduction

The classical statistical theory (and in particular the well-known theorem of the central limit) makes it possible to infer the central values of a sample but gives very little information on the distribution tail. The particularity of the extreme value theory (EVT) is that it focuses on the tail of distribution that generates the studied various extreme phenomenon. It is developed for the estimation of the probability of rare events and makes it possible to obtain reliable estimates of the extreme values, for which there are few observations. EVT or extreme value analysis (EVA) is a branch of statistics that aims to model and describe the occurrence and intensity of so-called rare events, ie those with very large amplitudes (a low probability of occurrence). When the behavior of these events is due to hasard, one can study their law. They are said to be extreme when they are much larger or smaller than those usually observed. The EVA seeks to assess, from a sample of ordered given a random variable, the probability of events which are more extreme than that any previously observed. It is widely used in many disciplines, such as structural engineering, finance, earth sciences, traffic prediction, and geological engineering. For example, EVA might be used in the field of hydrology to estimate the probability of an unusually wide flooding event, such as the 100-year flood. Similarly, for the design of a breakwater, a coastal engineer Would seek to estimate the 50-year wave and accordingly design the structure. There is a very good variety of textbooks which isdevoted to EVT and their applications for example [32], [20], [76] and [2]. This is what led Fisher and Tippet (1928) in [40] to develop the theory of extreme value.

In the other hand, when the observations of a phenomenon studied are not complete then we are in the case of incomplete data, where it can take various forms of censored or truncated data. Censoring is when an observation is incomplete due to some random case and the trancating is when an object can be detected only if its value is greater or less than some number and the value is completely known in the case of detection. As a result, the case of incomplete data is not that data to treat are not complete. Then, in this case the classical techniques does not apply.

Resently, in the letterature of statistics of extremes, the authors are more interested in estimating the extreme parameters (extreme value index and second order parameter) and the extreme quantile. In the case of complete data, several estimators have been proposed: for the extreme value index, recall the well know Hill estimator [60], the maximum likelihood estimator [88], the moment estimator, the Pickands estimator, probability-weighted moment estimator and a kernel type estimators of Csörgő and al. (1985) [16] and many more. We refer for the extreme value estimation in the case of censored data to Beirlant and al. (2007) [3], Einmahl and al. (2008) [31], Worms and Worms (2016) [91] and under random truncation to Gardes and Stupfler (2015) [42], Benchaira and al. (2015) [5]. In this thesis, we deal essentially with the case of right trancation with the estimation of extreme parameters and extreme quantile. The 4th chapter of this thesis contains the first published work on the estimate of the scond order parameter under random truncated.

Our thesis is organized in six chapters that allow us to present the different aspects of our work which is organized as follows:.

Chapter 1

The first chapter is an introduction to the extreme value: concepts and definitions of the distribution functions and order statistics. We review in this chapter the limit theorems: the lows of large numbers and central limit theorem. Then, we give generalized extreme value distribution, regular variation, the first and the second order of regular variation and domain of attraction.

Chapter 2

The second chapter, is a review of the estimating of the tail index, the second order parameter and the extreme quantile in the case of complete data by some classical methods. The informations that exists in this chapter enables understanding of the rest of the chapters.

Chapter 3

The third chapter contains definitions of incomplete data (censored and trancated data) and there forms, and the estimation of their distribution in every case.

Chapter 4

The 4th chapter contains our findings about the estimation of the second-order parameter of Pareto-type distributions under random right-truncation and its application by using the estimator of the second order parameter in the estimation of the tail index and this tail estimator is without bias. Our consedirations are based on results of Gomes and al. (2003) [46] (Semi-parametric estimation of the second order parameter in statistics of extremes) and on a useful Gaussian approximation of a tail product-limit process recently given by Benchaira et al. [6], we will prove their consistency and the asymptotic normality. Then, we will give the proves and the simulation of the estimators which can show the performance of our estimators in several size of the sample.

Chapter 5

In the 5th chapter, we will consider the random threshold case to derive a Hill-type estimator based on the recent results of Worms and Worms (2016) [91] (A Lynden-Bell integral estimator for extremes of randomly truncated data) introduced an asymptotically normal estimator of the tail index for Pareto-type (randomly right-truncated) data and we will establish the consistency and asymptotic normality of our estimator. A simulation study is carried out to evaluate the finite sample behavior of the proposed estimator and compare it with the existing ones.

Chapter 6

The last chapter contains an other method to estimate the quantile extreme which is based on both methods, the kernel type and the log probability weighted moment of estimation, where it is based on the results of Caeiro and Gomes (2015) [11] (A log probability weighted moment estimator of extreme quantiles) which consider the semi parametric estimation of extreme quantiles of a right heavy-tail model and propose a new probability weighted moment estimator of extreme quantiles. Then, we will prove the consistency and asymptotic normality of our estimator.

Finally, I would like to mention that the processing of data (numerical calculations and graphical representations) is carried out using the statistical analysis software R.

Chapter 1

Extreme values

1.1 Introduction

The EVT is a branch of statistics that aims to model and describe the occurrence and intensity of known events rare it is to say that present variations of great samplitudes (with a low probability of occurrence) .When the behavior of these events is due to chance, we can study their law. They expressed extreme values when it's about much larger or smaller than those usually observed. there is a very good variety of textbooks isdevoted to EVT and their applications for example [32], [20], [76] and [2].

1.2 Concepts and definitions

Let $(X_1, X_2, ..., X_n)$, be a sample of size $n \geq 1$ from a random variable (rv) $\mathbf{X}$ defined over some probability space $(\Omega, \mathcal{A}, \mathbf{P})$. The distribution of $\mathbf{X}$ may be caracterised by equivalents functions which are defined as follows.

Definition 1.1 (Distribution function) *The distribution function (df) of a rv X is the* *application F defined on $\mathbb{R}_+$ to $[0,1]$ by*

$$F(x) := P(X \leq x).$$

Definition 1.2 (Survival function) *If X is a rv defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ then, it's survival function is defined on $\mathbb{R}_+$ to $[0,1]$ by*

$$\overline{F}(x) := 1 - F(x) = P(X > x).$$

Definition 1.3 (Probability density function) *If F admits a derivative with respect to the Lebesgue measure on $\mathbb{R}_+$, the function of probability density (pdf) exists, defined for any $t \geq 0$, by*

$$f(t) := \frac{dF(t)}{dt} = \lim_{dx \to \infty} \frac{P(t < X < t + dx)}{dx}.$$

Definition 1.4 (Hasard function) *if X is a continuous positive rv representing a duration, the hasard function, noted by $h(t)$, is defined by*

$$h(t) := \frac{f(t)}{\overline{F}(t)} = \lim_{dx \to \infty} \frac{P(t < X < t + dx / X > t)}{dx}.$$

Remark 1.1 *Sometimes, it is useful to work with a cumulative (or integrated) which is given by*

$$\Lambda(t) := \int_0^t h(x)\,dx = \int_0^t \frac{f(x)}{\overline{F}(x)} dx, \tag{1.1}$$

it is easy to find the relationships between these different notions, for example (1.1) implies that

$$\Lambda(t) = -\log \overline{F}(x). \tag{1.2}$$

It is noted that, under (1.2), we can write

$$\overline{F}(x) = \exp\{-\Lambda(t)\} = \exp\left\{-\int_0^t \frac{f(x)}{\overline{F}(x)}\right\}. \tag{1.3}$$

This equality is the main exponential for survival analysis. It has a distribution characteristic and a survival function by intermediate of a hasard function.

Definition 1.5 (Quantile function) *The quantile function is defined for any* $0 < s < 1$ *by*

$$Q(s) = F^{\leftarrow}(s) := \inf\{t : F(t) \geq s\},$$

where $F^{\leftarrow}$ *is the generalised inverse function of df* F*, with the convention that* $\inf\{\phi\} = +\infty$.

Remark 1.2 *It is expressed in terms of the survival function by*

$$F^{\leftarrow}(s) = \inf\left\{t : \overline{F}(t) \leq 1 - s\right\} = \overline{F}^{\leftarrow}(1 - s),\ 0 < s < 1.$$

Definition 1.6 (Tail quantile function) *The tail quantile function is denoted by* U *and for any* $1 < t < \infty$

$$U(t) := Q(1 - 1/t) = \left(1/\overline{F}\right)^{\leftarrow}(t).$$

Proposition 1.1 (Quantile trasformation) *Let* U *be a* $(0,1)-$*uniformly distributed rv, then*

1. *For any df* F *of a rv* X*,* $F^{\leftarrow}(U) \overset{d}{=} X$.

2. *When* F *is continuous, we have* $F(X) \overset{d}{=} U$.

Definition 1.7 (Empirical df and survival function) *The empirical df and survival function of the sample* $(X_1, X_2, ..., X_n)$ *is defined respectly by*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{(X_i \leq x)},\ x \in \mathbb{R}$$

and

$$\overline{F}_n(x) := \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}_{(X_i>x)},\ \ x\in\mathbb{R},$$

where $\mathbb{I}_A$ *is the indicator function of* A.

Definition 1.8 (Empiriquale quantile and tail quantile function) *The empirical quantile function is defined by*

$$Q_n(s) := \inf\{t : F_n(t) \geq s\},\ \ 0<s<1.$$

The corresponding empirical tail quantile function is

$$U_n(t) := Q_n(1-1/t),\ \ 1<t<\infty.$$

Definition 1.9 (Sum and arithmetic mean) *Let* $(X_1, X_2, ..., X_n)$ *be a sample from a rv* X *defined over some probability space* $(\Omega, \mathcal{A}, \mathbf{P})$,. *For an integer* $n \geq 1$, *define the partial sum and the corresponding arithmetic mean by respectively*

$$S_n := \sum_{i=1}^{n} X_i \text{ and } \overline{X}_n := S_n/n.$$

$\overline{X}_n$ *is called sample mean or empirical mean.*

1.3 Limite theorems

In this section we reminded the Laws of large numbers and the Central limit theorem.

1.3.1 Laws of large numbers

In the classical theory, one is often interested in the behaviour of the mean or average. This average will then be described through the expected value EX of the distribution.

On the basis of the law of large numbers, the sample mean $\overline{X}_n$ is used as a consistent estimator of EX.

Theorem 1.1 (Laws of large numbers) *If* $(X_1, X_2, ..., X_n)$ *is a sample from a rv* X *such that* $E|X| < \infty$, *then*

$$\overline{X}_n \xrightarrow{P} \mu \text{ as } n \to \infty \text{ weak law,}$$

$$\overline{X}_n \xrightarrow{a.s.} \mu \text{ as } n \to \infty \text{ strong law,}$$

where $\mu := EX$.

Applying the strong law of large numbers on $F_n(x)$ yields the following result.

Corollary 1.1 *For every* $x \in \mathbb{R}$,

$$F_n(x) \xrightarrow{a.s.} F(x) \text{ as } n \to \infty.$$

Theorem 1.2 (Glivenko-Cantelli)

$$\underset{x \in \mathbb{R}}{Sup} |F_n(x) - F(x)| \xrightarrow{a.s.} 0 \text{ as } n \to \infty.$$

The proofs of Theorems 1.1 and 1.2 could be found in any standard textbook of probability theory such as [8].

1.3.2 Central Limit Theorem

The central limit theorem yields the asymptotic behaviour of the sample mean. This result can be used to provide a confidence interval for EX in case the sample size is sufficiently large, a condition necessary when invoking the central limit theorem.

Theorem 1.3 (Central Limit Theorem) *Let $X_1, X_2, ..., X_n$ be a sequence of iid rv's with mean μ and finite variance σ^2, then*

$$(S_n - n\mu)\,\sigma\sqrt{n} \xrightarrow{D} N(0,1) \ \ \text{as } n \to \infty.$$

The proof of the Central Limit Theorem (CLT) could be found in any standard book of statistics, see e.g., Saporta, G. (1990) , [79] page 66.

Note that a necessary condition for the CLT is that the variance be finite. That is, if the finite variance assumption is dropped, the limit distribution in Theorem 1.3 is no longer normal. In the case of infinite variance, there exists a result known as the generalized CLT which states that stable laws are the only possible limit distributions for properly normalized and centered sums of iid rv's.

1.4 Order statistics

Definition 1.10 (Order statistics) *The order statistics of a sample $(X_1, X_2, ..., X_n)$ are the X_i's arranged in non-decreasing order. They are denoted by $X_{1:n}, X_{2:n}, ..., X_{n:n}$ and for $k = 1, 2, ..., n$, the rv $X_{n-k+1:n}$ is called the kth upper order statistic. Order statistics satisfy $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$. Thus*

$$X_{1:n} := \min(X_1, X_2, ..., X_n) \ \ \text{and } X_{n:n} := \max(X_1, X_2, ..., X_n).$$

Remark 1.3 *Noted that it is easy to trouve the following relation*

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

On this thesis we shall concentrate on the study of the maximum.

Remark 1.4

- The empirical df of the sample $(X_1, X_2, ..., X_n)$ is evaluated using order statistics as follows:

$$F_n(x) = \begin{cases} 0 & \text{if } x < X_{1:n} \\ \frac{i-1}{n} & \text{if } X_{i-1:n} \leq x \leq X_{i:n} \\ 1 & \text{if } x > X_{n:n} \end{cases}, \text{ for } 1 < i \leq n.$$

- Q_n may be expressed as a simple function of the order statistics pertaining to the sample $(X_1, X_2, ..., X_n)$. Then, we have

$$Q_n(s) = X_{n-i+1:n} \text{ for } \frac{n-i}{n} < s \leq \frac{n-i+1}{n},\ 1 \leq i \leq n.$$

Proposition 1.2 (Distributions of rth order statistc) *The distributions of $X_{r:n}$ defined as follows :*

1. The df of $X_{r:n}$ is defined by

$$F_{X_{r:n}}(x) := \sum_{i=r}^{n} \binom{n}{i} F^r(x)\, \overline{F}^{n-r}(x).$$

2. The pdf of $X_{r:n}$ is defined if $F_{X_{r:n}}$ is continuous by

$$f_{X_{r:n}}(x) := \frac{n!}{(r-1)!\,(n-r)!} F^{r-1}(x)\, \overline{F}^{n-r}(x)\, f(x).$$

Remark 1.5 *The event $x \leq X_{r:n} \leq x + \delta x$ may be realized as follows :*

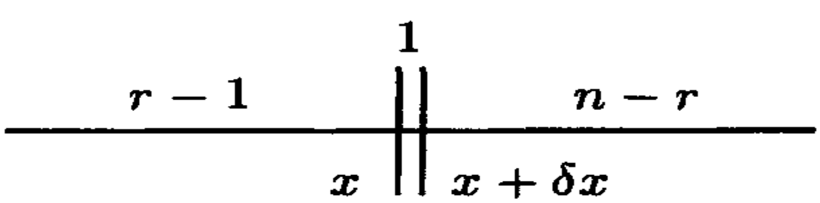

Illustration of the number of ways can be made for n observations compared to $X_{r:n}$.

$X_i < x$ for $r-1$ of the X_i, $x \leq X_i \leq x+\delta x$ for one X_i and $X_i > x+\delta x$ for the remaining $n-r$ the X_i. The number of ways in which the n observations can be so divided into three parcels is

$$\frac{n!}{(r-1)!1!\,(n-r)!}$$

and each such way has probability

$$F^{r-1}(x)\left[F(x+\delta x)-F(x)\right]\overline{F}^{n-r}(x+\delta x).$$

Proposition 1.3 (Joint distribution of two or more order statistics) *$X_{[np]+1:n}$ is the empiriquale quantile of order p for $0 < p < 1$.*

1. The joint pdf of $X_{1:n}, X_{2:n}, ..., X_{n:n}$ is defined by

$$f_{X_{1:n},X_{2:n},...,X_{n:n}}(x_1,x_2,...,x_n) := n!\prod_{i=1}^{n} f(x_i),\ x_1 \leq x_2 \leq ... \leq x_n.$$

2. The joint pdf of $X_{1:n}, X_{2:n}, ..., X_{k:n}$ is defined by

$$f_{X_{1:n},X_{2:n},...,X_{k:n}}(x_1,x_2,...,x_k) := \frac{n!}{(n-k)!}F^{n-k}(x_k)\prod_{i=1}^{k} f(x_i),\ x_1 \leq x_2 \leq ... \leq x_k.$$

3. The joint pdf of $X_{r:n}, X_{s:n}$ is defined by (for $1 \leq r < s \leq n$)

$$f_{X_{r:n},X_{s:n}}(x,y) := \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!}F^{r-1}(x)\,f(x)\left[F(y)-F(x)\right]^{s-r-1},\ x \leq y.$$

which may be realized as follows :

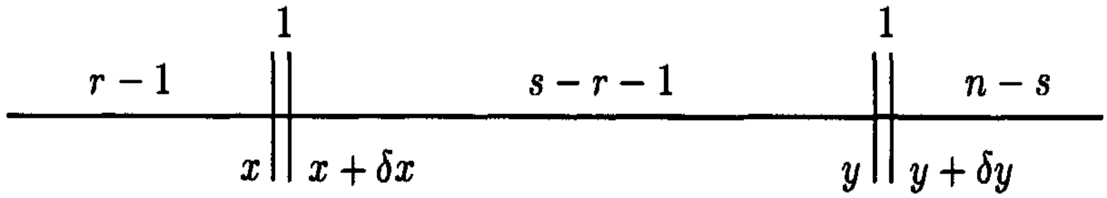

Illustration of the number of ways can be made for n observations compared to $X_{r:n}$ and $X_{s:n}$.

Proof. see e.g, David and Nagaraja (2003) [18] page 11. ∎

Distributional results for the smallest and largest order statistics ($X_{1:n}$ and $X_{n:n}$ successively) are immediate.

Proposition 1.4 (The Distributions functions of muximum and minimum)

1. The joint pdf of $X_{1:n}$ and $X_{n:n}$ is defined by

$$f_{X_{1:n},X_{n:n}}(x,y) := n(n-1)[F(y)-F(x)]^{n-2} f(x) f(y), \ x_1 < x_2.$$

2. The pdf of $X_{1:n}$ and $X_{n:n}$ are defined respectively by

$$f_{X_{1:n}}(x) := n\overline{F}^{n-1}(x) f(x) \text{ and } f_{X_{n:n}}(x) := nF^{n-1}(x) f(x).$$

3. The df of $X_{1:n}$ and $X_{n:n}$ are defined respectively by

$$F_{X_{1:n}}(x) := 1-\overline{F}^{n}(x) \text{ and } F_{X_{n:n}}(x) := F^{n}(x).$$

Proof. see e.g, Embrechts et al.(1997) [32] page 183. ∎

1.5 Limit distributions of maxima

The limit of the distribution of the maxima $X_{n:n}$, when n tends to infinity, is degenerate and, depending on F, the maximum will tend either to infinity or a finite number called upper (or right) endpoint of F, i.e :

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

This endpoint is may be finite or infinite (see [32], Exemple 3.3.2, page 139). Our interest is on the asymptotic distribution of the maximum

$$\lim_{n\to\infty} F_{X_{n:n}}(x) = \lim_{n\to\infty} [F(x)]^n = \begin{cases} 1 \text{ if } x \geq x_F, \\ 0 \text{ if } x < x. \end{cases}$$

Then, we have that

$$X_{n:n} \overset{a.s.}{\to} x_F, \tag{1.4}$$

as $n \to \infty$, the result 1.4 is immediate (see e.g., [32]). The central result on EVT which specifies the form of the limit distribution for centred and normalised maxima of independent and identically distributed random variables is the Fisher–Tippett theorem [40].

Definition 1.11 *Let F_1 and F_2 be two dfs. F_1 and F_2 are on the same type iff exist a real $a \in \mathbb{R}_+^*$ and $b \in \mathbb{R}$, sutch that for any $x \in \mathbb{R}$,*

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

Theorem 1.4 (Fisher and Tippett) *Let $(X_1, X_2, ..., X_n)$ be a sample from a rv X with continuous df F and $X_{n,n} = max(X_1, ..., X_n)$. If exist a non-degenerate df H and two real* *sequences* $\{a_n\}$ *and* $\{b_n\}$, $n \in \mathbb{N}$, *with* $a_n > 0$ *and* $b_n \in \mathbb{R}$, *such that for any* $x \in \mathbb{R}$,

$$\lim_{n\to\infty} P\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} F^n\left(a_n x + b_n\right) = H\left(x\right). \tag{1.5}$$

Then, H is on the same of the following three dfs :

$$\begin{aligned}
&\textit{Type I:} && \Lambda(x) && = \exp\left(-e^{-x}\right), && && \textit{(Gumbel df)},\\
&\textit{Type II:} && \Phi_a(x) && = \exp\left(-\left(x\right)^{-\alpha}\right) && \mathbb{I}_{(x\geq 0)},\ a > 0, && \textit{(Fréchet df)},\\
&\textit{Type III:} && \Psi_a(x) && = \exp\left(-\left(-x\right)^{\alpha}\right) && \mathbb{I}_{(x<0)} + \mathbb{I}_{(x\geq 0)},\ a > 0, && \textit{(Weibull df)},
\end{aligned}$$

where $\mathbb{I}_A$ *is the indicator function of the set A.*

For the proof of this Theorem, see e.g. [78] and [32].

Remark 1.6

- *H is the extreme value distribution (EVD).*
- $\{a_n\}$ *and* $\{b_n\}$ *are called norming sequences dependents with law of X.*
- *The theoritical norming sequences associated with the law of standard normal in [32] page* 145*, are*

$$a_n = \left(2\log n\right)^{-1/2} \text{ and } b_n = \left(2\log n\right)^{1/2} - \frac{\log\log n + \log 4\pi}{2\left(2\log n\right)^{1/2}}.$$

1.6 Generalized extreme value distribution (GEVD)

Definition 1.12 (Standard extreme value distributions) *The Three df's of theoreme of Fisher and Tippett are called standard extreme value distributions.* Λ *is known as Gumbel (or double exponential) type,* Φ_ξ *as Fréchet (or heavy-tailed) type and* Ψ_ξ *as (reverse) Weibull type.*

Remark 1.7 *Let X be a positive rv ($X > 0$), then*

$$(X \text{ has df } \Phi_\xi) \Leftrightarrow (\ln X^\xi \text{ has df } \Psi_\xi) \Leftrightarrow (-1/X \text{ has df } \Lambda).$$

Definition 1.13 (GEVD) *The GEVD is a df H_γ defined, for all $x \in \mathbb{R}$ such that $1+\gamma x > 0$, as follows :*

$$H_\gamma(x) = \begin{cases} \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\} & if \quad \gamma \neq 0, \\ \exp(-e^{-x}) & if \quad \gamma = 0. \end{cases} \tag{1.6}$$

Remark 1.8

- *The parameter γ is called Extreme Value Index (EVI), tail index or shape parameter.*
- *The corresponding pdf h_γ is defined for all $x \in \mathbb{R}$ by*

$$h_\gamma(x) = \begin{cases} H_\gamma(x)(1+\gamma x)^{-1/\gamma-1} & if \quad \gamma \neq 0, \\ \exp(-x-e^{-x}) & if \quad \gamma = 0, \end{cases}$$

  where $1+\gamma x > 0$.

- *We can writ $H_\gamma(x)$ in a more general form by replacing the argument x by $(x-\mu)/\sigma$ in the right hand side of 1.6, for $1+\gamma\frac{x-\mu}{\sigma} > 0$*

$$H_{\gamma,\mu,\sigma}(x) = \begin{cases} \exp\left\{-\left(1+\gamma\frac{x-\mu}{\sigma}\right)^{-1/\gamma}\right\} & if \quad \gamma \neq 0, \\ \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) & if \quad \gamma = 0. \end{cases}$$

  where $\mu \in \mathbb{R}$ and $\sigma > 0$ are respectively the location and scale parameters Weissman (1978) *[86].*

- *We can express the three extreme value distributions in terms of the GEVD H_γ as* *follows :*

$$\begin{aligned} \Lambda(x) &= H_0(x), & x \in \mathbb{R}, \\ \Phi_\xi(x) &= H_{1/\xi}[\xi(x-1)], & x > 0, \\ \Psi_\xi(x) &= H_{-1/\xi}[\xi(x+1)], & x < 0. \end{aligned}$$

- *The three extreme value distributions can be characterized by the sign of the tail index* γ *:*

$$H_\gamma = \begin{cases} \Psi_{-1/\gamma} & if \quad \gamma < 0, \\ \Lambda & if \quad \gamma = 0, \\ \Phi_{1/\gamma} & if \quad \gamma > 0. \end{cases}$$

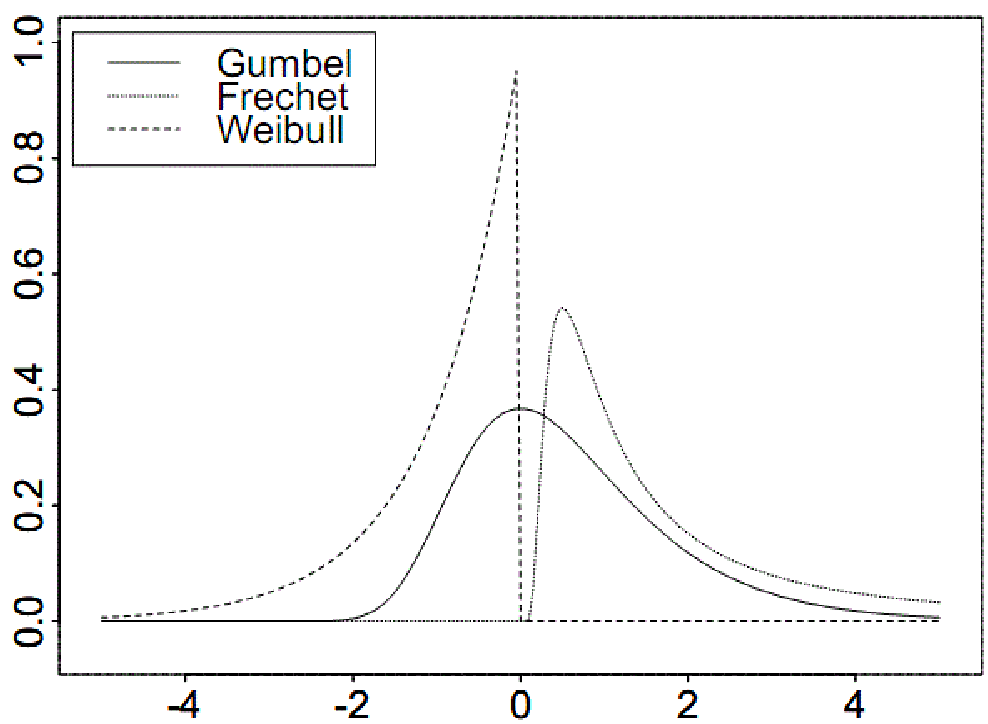

Figure 1.1: Densities of the standard extreme value distributions. We chose $\alpha = 1$ for the Frechet and the Weibull distribution.

1.7 Regular variation function

The concept of regular variation is frequently used in extreme value theory, for more details we refer to [9].

Definition 1.14

- *The measurable function* $f : \mathbb{R}^+ \to \mathbb{R}^+$ *is regular variation function at infinity with index* $\alpha \in \mathbb{R}$, *iff for any* $x \in \mathbb{R}$,

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

  notation $f \in RV_\alpha$, α *is named index of regular variation function* f.

- *The measurable function* $f : \mathbb{R}^+ \to \mathbb{R}^+$ *is regular variation function at* 0 *with index* $\alpha \in \mathbb{R}$, *iff for any* $x \in \mathbb{R}$,

  $$\lim_{t\to 0} \frac{f(tx)}{f(t)} = x^\alpha.$$

  Notation $f \in RV_\alpha^0$, *i.e:* $f(1/x)$ *is regular variation function with index* $-\alpha$ *at infinity.*

Remark 1.9 *If* $\alpha = 0$, *then function* f *is said to be slowly varying at infinity. i.e:*

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

Slowly varying functions are noted $\mathrm{l}(x)$.

Lemma 1.1 *Inverse of regular variation function*

- *If* f *is regular variation at infinity with index* $\alpha > 0$, *then* $f^{\leftarrow}$ *is regular variation at infinity with index* $1/\alpha > 0$.

- *If* f *is regular variation at infinity with index* $\alpha < 0$, *then* $f^{\leftarrow}$ *is regular variation at infinity with index* $-1/\alpha > 0$.

The proof of Lemma could be found in [9]. Then, if l is slowly varying function and $\alpha \in \mathbb{R}$, then the function $f(x) := x^{\alpha} l(x) \in RV_{\alpha}$, for all $x > 0$.

Proposition 1.5 *Let be $\alpha \in \mathbb{R}$ and $f \in RV_{\alpha}$. Then, there is a slowly varying function at infinity* l *where for all* $\forall x > 0$,

$$f(x) = x^{\alpha} \mathrm{l}(x).$$

Example 1.1 *The slowly varying functions at infinity for example:*

1. *Functions have a strict positive limite at infinity,*
2. *Functions of forms $f : x \to |\log x|^{\beta}$, $\beta \in \mathbb{R}$.*
3. *Functions f where*

$$\exists M > 0,\ \forall x \geq M,\ g(x) = c + dx^{-\beta}(1 + o(1)),$$

   where $c, \beta > 0$ and $d \in \mathbb{R}$. The set of this functions is named Hall's class.

Theorem 1.5 (Karamata representation (Resnick, 1987)) *Evry slowly varying function* l *at infinity is defined as*

$$\mathrm{l}(x) = c(x) \exp\left(\int_1^x r(t) t^{-1} dt\right),$$

where $c(.) > 0$ and $r(.)$ are two measurable functions, sutch that

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

if the function $c(.)$ is a constant, then we said l *is normalised.*

For the proof see Resnick [78], corollary 2.1; page 29.

Proposition 1.6 *For evry slowly varying function* l *at infinity we have*

$$\lim_{x\to\infty}\frac{\log\left(\mathrm{l}\left(x\right)\right)}{\log\left(x\right)}=0.$$

For more details on this issue, see de Haan [20], [22] and [9].

1.7.1 First Order Regular Variation Assumption

For a df function F and U the tail quantile function, the following assertions (assumptions) are equivalent :

- $\overline{F}$ is regularly varying at infuinity with index $-1/\gamma$

$$\lim_{z\to\infty}\frac{\overline{F}\left(xz\right)}{\overline{F}\left(z\right)}=x^{-1/\gamma},\ x>0.$$

- $Q\left(1-s\right)$ is regularly varying at 0 with index $-\gamma$

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

- U is regularly varying at ∞ with index γ

$$\lim_{z\to\infty}\frac{U\left(xz\right)}{U\left(z\right)}=x^{\gamma},\ x>0.$$

- F is heavy tailed.

1.7.2 Second Order Regular Variation Assumption

We say that F is second order regularly varying at infinity if it satisfies one of the following (equivalent) conditions :

- There exist some parameter $\rho \leq 0$ and a function A^*, such that for all $x > 0$

$$\lim_{t\to\infty} \frac{\overline{F}(tx)/\overline{F}(t) - x^{-1/\gamma}}{A^*(t)} = x^{-1/\gamma}\frac{x^\rho - 1}{\rho}. \tag{1.7}$$

- There exist some parameter $\rho \leq 0$ and a function A^{**}, such that for all $x > 0$

$$\lim_{s\to\infty} \frac{Q(1-sx)/Q(1-s) - x^{-1/\gamma}}{A^{**}(s)} = x^{-\gamma}\frac{x^\rho - 1}{\rho}.$$

- There exist some parameter $\rho \leq 0$ and a function A, such that for all $x > 0$

$$\lim_{t\to\infty} \frac{U(tx)/U(t) - x^{\gamma}}{A(t)} = x^{\gamma}\frac{x^\rho - 1}{\rho}. \tag{1.8}$$

Where A^*, A^{**} and A are regularly varying functions with

$$A^*(t) = A\left(1/\overline{F}(t)\right) \text{ and } A^{**}(t) = A(1/t),$$

their role is to control the speed of convergence in First Order Regular Variation Condition.

If $\rho = 0$, interpret $(x^\rho - 1)/\rho$ as $\log x$.

For the proofs see de Haan ana Ferrira (2006) [20].

1.7.3 Third Order Regular Variation Assumption

There exists a positive real parameter γ, negative real parameters ρ and β, functions b and $\widetilde{b}$ with $b(t) \to 0$ and $\widetilde{b}(t) \to 0$ for $t \to \infty$, both of constant sign for large values of t, such that

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{b(t)} - \frac{x^\rho - 1}{\rho}}{\widetilde{b}(t)} = \frac{1}{\beta}\left(\frac{x^{\rho+\beta} - 1}{\rho + \beta} - \frac{x^\rho - 1}{\rho}\right), \text{ for } x > 0$$

where $\left|\widetilde{b}\right|$ is regularly varying of index β. We refer to [20] for further details.

1.8 Characterization of the Domain of Attraction

An important problem is to define the conditions (necessary and sufficient) of membership of a distribution to a domain of attraction. Different characterizations of the three domains of attraction of Fréchet, Gumbel and Weibull have been proposed in Resnick and al. (1987) [78], Embrechts and al. (1997) [32] and de Haan ana Ferreira (2006) [20]. These characterizations involve classes of functions with regular variation. In the following, a df F is said to be in the domain of attraction of a non-degenerate df H_γ, denoted by $F \in D(H_\gamma), \gamma \in \mathbb{R}$.

The following theorem states a sufficient condition for belonging to a domain of attraction. The condition is called von Mises condition.

Theorem 1.6 *Let F be a distribution function and x_F its right endpoint. Suppose $F''(x)$ exists and $F'(x)$ is positive for all x in some lefte neighborhood of x_F. If*

$$\lim_{t \uparrow x_F} \left(\frac{1-F}{F'} \right)' (t) = \gamma, \tag{1.9}$$

or equivalently

$$\lim_{t \uparrow x_F} \frac{(1-F(t)) F''(t)}{(F'(t))^2} = -\gamma - 1,$$

then F is in the domain of attraction of H_γ.

Remark 1.10 *Under 1.9 we have 1.5 with $b_n = U(n)$ and $a_n = nU'(n) = 1/(nF'(b_n))$.*

Theorem 1.7

1. *For $\gamma > 0$, suppose $x_F = \infty$ and F' exists. Then, if*

$$\lim_{t \to \infty} \frac{tF'(t)}{1-F(t)} = \frac{1}{\gamma},$$

for some positive γ, then F is in the domain of attraction of H_γ.

2. *For $\gamma < 0$, suppose $x_F = \infty$ and F' exists for $x < x_F$. Then, if*

$$\lim_{t \uparrow x_F} \frac{(x_F - 1) F'(t)}{1 - F(t)} = \frac{-1}{\gamma},$$

for some negative γ. then F is in the domain of attraction of H_γ.

For the proofs and more details on this issue, one may consulte de Haan ana Ferreira (2006) [20] (see page 15).

Theorem 1.8 *The distribution function F is in the domain of attraction of the extreme value distribution $D(H_\gamma)$ if and only if*

1. *for $\gamma > 0 : F(x) < 0$ for all x, $\int_1^\infty [(1 - F(x))/x]\, dx < \infty$, and*

$$\lim_{t \to \infty} \frac{\int_t^\infty (1 - F(x)) \frac{dx}{x}}{1 - F(t)} = \gamma. \tag{1.10}$$

2. *for $\gamma < 0 :$ there is $x_F < \infty$ such that, $\int_{x_F - t}^{x_F} [(1 - F(x))/(x_F - x)]\, dx < \infty$, and*

$$\lim_{t \downarrow 0} \frac{\int_{x_F - t}^{x_F} (1 - F(x)) \frac{dx}{x_F - x}}{1 - F(x_F - t)} = -\gamma. \tag{1.11}$$

3. *for $\gamma = 0 : \int_x^{x_F} \int_t^{x_F} (1 - F(s))\, dsdt < \infty$ (here the right endpoint x_F may be finite or infinite) and*

$$\lim_{t \uparrow x_F} \frac{(1 - F(x)) \int_x^{x_F} \int_t^{x_F} (1 - F(s))\, dsdt}{\left(\int_x^{x_F} (1 - F(s))\, ds\right)^2} = 1. \tag{1.12}$$

Remark 1.11 *Limit 1.10 is equivalent to*

$$\lim_{t \to \infty} E(\log X - \log t \mid X > t) = \gamma.$$

In fact,

$$\lim_{t\to\infty}\frac{\int_t^{\infty}(1-F(x))\frac{dx}{x}}{1-F(t)}=\lim_{t\to\infty}E\left(\log X-\log t\mid X>t\right),$$

since

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

Relation 1.10 will be the basis for the construction of the Hill estimator of γ. Similarly, 1.11 can be interpreted as

$$\lim_{t\downarrow 0}E\left(\log(x_F-X)-\log t\mid X>x_F-t\right)=\gamma,$$

which will be the basis for the construction of the negative Hill estimator and 1.12 is equivalent to

$$\lim_{t\uparrow x_F}\frac{E\left((X-t)^2\mid X>t\right)}{E^2\left(\log(X-t)\mid X>t\right)}=2,$$

and this relation leads to the moment estimator of γ. In chapter 2 of estimation of parameters we defined the estimators of γ.

For the proofs see [20]. Now in following, we shall establish necessary and sufficient condition for a distribution function F to belong to one of the three domains of attraction of Fréchet, Gumbel and Weibull.

1.8.1 Domain of attraction of Fréchet

The result below stated in Gnedenko (1943) and we find a demonstration in Resnick (1987).

Theorem 1.9 *For the endpoit x_F, a df F is belonging to domain of attraction of Fréchet $D(H_\gamma)$, $\gamma>0$ iff $x_F=\infty$ and $\overline{F}$ is regularing variation with index $-1/\gamma$ at infinity i.e :*

$$\lim_{z\to\infty}\frac{\overline{\mathbf{F}}(xz)}{\overline{\mathbf{F}}(z)}=x^{-1/\gamma}. \tag{1.13}$$

The normalised constants are

$$a_n = U(n) = F^{\leftarrow}(1 - 1/n) \text{ and } b_n = 0, \; \forall n > 0$$

Remark 1.12

1. *From the proposition (existe slowly function $g(x) = xl(x)$) we have that*

$$F \in D(H_\gamma), \; \gamma > 0 \iff \overline{F}(x) = x^{-1/\gamma} l(x),$$

   where l is a slowly regular variation function at infinity.

2. *For any $s \in (0,1)$, $Q(1-s) = F^{\leftarrow}(1/s)$. The equation 1.13 is iquivalent to $Q(1-.) \in RV^0_{-\gamma}$ ie : $Q(1-s)$ is a rugular variation function at 0 with index $-\gamma$ and*

$$Q(1-s) = s^{-\gamma} l(1/s),$$

   sutch that l is a slowly regular variation function at infinity $(l \in RV_0)$, and the tail quantile function U is regularly varying with index γ at infinity $(U \in RV_\gamma)$.

3. *From 1.13, $\overline{F} \in RV_{-1/\gamma}$, $\gamma > 0$, then the representation of karamata we have that*

$$\overline{F}(x) = c(x) x^{-1/\gamma} l(x) \left(\int_1^x r(t) t^{-1} dt \right), \; x < x_F,$$

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

1.8.2 Domain of attraction of Weibull

The result in following shows that we can pass from the domain of attraction of Fréchet to that of Weibull by a simple change in the distribution function (see Gnedenko (1943) [52] or Resnick (1987) [78]).

Theorem 1.10 *For the endpoit x_F, a df F is belonging to domain of attraction of Weibull $D(H_\gamma)$, $\gamma < 0$ iff $x_F < \infty$ and the df F^* defined as*

$$F^*(x) = \begin{cases} 0 & if \quad x \leq 0, \\ F(x_F - 1/x) & if \quad x > 0. \end{cases}$$

is belonging to domain of attraction of Fréchet with index $-\gamma > 0$ at infinity i.e : $\overline{F}^*$ is a regular variation function with index $1/\gamma$ at infinity ($\overline{F}^* \in RV_{1/\gamma}$). In this case the normalised constants are

$$a_n = x_F - F^{\leftarrow}(1 - 1/n) \text{ and } b_n = x_F, \ \forall n > 0.$$

For the proof of this theorem we refer to Resnick and al. (1987) [78], proposition1.13 or to Embrechts and al. (1997) [32], Théorème3.3.12.

1.8.3 Domain of attraction of Gumbel

The following result is prouved in Resnick and al. (1987) [78].

Theorem 1.11 *For the endpoit x_F, a df F is belonging to domain of attraction of Gumbel $D(H_\gamma)$, $\gamma = 0$ iff exists a reel $z < x_F \leq \infty$ sutch that*

$$\overline{F}(x) = c(x) \exp\left(-\int_z^x \frac{g(t)}{a(t)} dt\right), \ z < x < x_F, \tag{1.14}$$

where c and g are positive measurable functions and such that

$$\lim_{x \to x_F} c(x) = c > 0 \text{ and } \lim_{x \to x_F} g(x) = 1,$$

and a a positive and absolutely continuous function (with respect to Lebesgue measure) *with density* a' *with* $\lim_{x \to x_F} a'(x) = 0$. *In this case we can choose*

$$a_n = U(n) = F^{\leftarrow}(1 - 1/n) \text{ and } b_n = a(a_n), \ \forall n > 0,$$

as norming constants. A possible choice for the function a *is*

$$a(x) = \int_x^{x_F} \frac{\overline{F}(t)}{\overline{F}(x)} dt, \ x < x_F. \tag{1.15}$$

Remark 1.13 *The function* a *defined in 1.15 is named auxiliary function and the function* F *defined in 1.14 is named von Mises function with auxiliary function* a.

In the following table 1.1 some examples of laws from the three domains of attraction

	Frechet	Gumbel	Weibull
Tail index	$\gamma > 0$	$\gamma = 0$	$\gamma < 0$
Laws	Cauchy Pareto Student Burr Loggamma	Normale Exponentielle Lognormale Gamma Weibull	Uniforme Beta

Table 1.1: Some examples of laws from the three domains of attraction

1.8.4 General Characterizations

The following results gives general characterizations of three domains of attraction on this issue, one may consulte [20].

Theorem 1.12 *For* $\gamma \in \mathbb{R}$, *the following assertions are equivalent.*

- $F \in D(H_\gamma)$.

- *For some positive function a, for $x > 0$*

$$\lim_{t\to\infty}\frac{U(tx)-U(t)}{a(t)}=\begin{cases}\frac{x^{\gamma}-1}{\gamma} & if \quad \gamma\neq 0,\\ \log x & if \quad \gamma=0.\end{cases} \tag{1.16}$$

- *For some positive function b, for $x > 0$ with $(1+\gamma x) > 0$*

$$\lim_{t\to x_F}\frac{\overline{F}(t+xb(t))}{\overline{F}(t)}=\begin{cases}(1+\gamma x)^{-1/\gamma} & if \quad \gamma\neq 0,\\ \exp(-x) & if \quad \gamma=0,\end{cases}$$

sutch that $b(t)=a\left(1/\overline{F}(t)\right)$.

- *For some positive function* $\widetilde{a}(t)=a(1/t)$,

$$\lim_{t\to 0}\frac{Q(1-sx)-Q(1-s)}{\widetilde{a}(t)}=\begin{cases}\frac{x^{-\gamma}-1}{\gamma} & if \quad \gamma\neq 0,\\ \log x & if \quad \gamma=0.\end{cases}$$

Suppose $U(\infty) > 0$, the condition 1.16 yields

$$\lim_{t\to\infty}\frac{U(tx)-U(t)}{a(t)/U(t)}=\begin{cases}\frac{x^{\gamma}-1}{\gamma} & if \quad \gamma< 0,\\ \log x & if \quad \gamma\geq 0.\end{cases} \tag{1.17}$$

In Dekkers and al. (1987), the result 1.17 is used to proposed the moments estimator of parameter γ.

Proposition 1.7 *For $\gamma\in\mathbb{R}$, $F\in D(H_{\gamma})$ iff for any $x>0$, $y>0$ and $y\neq 1$,*

$$\lim_{t\to\infty}\frac{U(tx)-U(t)}{U(ty)-U(t)}=\begin{cases}\frac{x^{\gamma}-1}{y^{\gamma}-1} & if \quad \gamma\neq 0,\\ \frac{\log x}{\log y} & if \quad \gamma=0.\end{cases} \tag{1.18}$$

This Proposition is used for construction the Pickands estimator of γ. For the proofs of this results, see Embrechts et al. (1997) [32].

Chapter 2

Tail index, extreme quantile and second-order parameter estimation

2.1 Introduction

In this chapter we interested to the tail index parameter, extreme quantile and second-order parameter estimation. In the following we defined some estimators (in semi parametric estimation) constructed under maximum domain of attraction conditions. That is, we set $X_1, X_2, ..., X_n$ be iid random variables with distribution function $F \in D(H_\gamma)$ i.e.

$$\lim_{t\to\infty} \frac{\overline{F}(xt)}{\overline{F}(t)} = x^{-1/\gamma},\ x > 0. \tag{2.1}$$

and let $X_{1:n}, X_{2:n}, ..., X_{n:n}$ be the nth order statistics. The sample fraction $k = k_n$ being a (random) sequence of integers such that,

$$k_n \to \infty \text{ and } k_n/n \to 0 \text{ as } n \to \infty.$$

2.2 Tail index estimation

There is many estimators for the tail index γ, a simple one is Hill's (1975) estimator for $\gamma > 0$, Pickands (1975) and a Moment estimator in general case where $\gamma \in \mathbb{R}$, the Maximum Likelihood Estimator for $\gamma > -1/2$, and other estimators like the Probability-Weighted Moment Estimator $(\gamma < 1)$, the Negative Hill Estimator $(\gamma < -1/2)$. We interested only in this section with Hill estimator, Pickand's, Moment estimator and Probability-Weighted Moment Estimator. For more details on this issu see de Haan ana Ferreira (2006) [20].

2.2.1 Hill's estimator $(\gamma > 0)$

In this case the parameter $\alpha := 1/\gamma > 0$ is called the tail index of F, an equivalent form of the condition 2.1 is :

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

Now partial integration yields

$$\int_t^\infty x^{-1}\overline{F}(x)\,dx = \int_t^\infty (\log s - \log t)\,dF(s).$$

Hence we have

$$\lim_{t\to\infty} \frac{\int_t^\infty (\log s - \log t)\,dF(s)}{\overline{F}(t)} = \gamma.$$

Then by remplaced the parameter t by the kth order statistic $X_{n-k:n}$ and F by the empirical distribution function F_n we obtain the Hill's (1975) estimator $\widehat{\gamma}_n^{(H)}$ defined [60].

Definition 2.1 (Hill's estimator $(\gamma > 0)$)

$$\widehat{\gamma}_n^{(H)} = \widehat{\gamma}_n^{(H)}(k) := \frac{\int_{X_{n-k:n}}^\infty (\log s - \log X_{n-k:n})\,dF_n(s)}{\overline{F}(X_{n-k:n})},$$

or

$$\widehat{\gamma}_n^{(H)} := \frac{1}{k}\sum_{i=1}^{k} \log X_{n-i+1:n} - \log X_{n-k:n}.$$

This estimator is only applicable in case the EVI γ is known to be positive, which corresponds to distributions belonging to the Frechet type domain of attraction.

Hill's estimator is usual and easy to explain. It can be derived through several other approaches (see Embrechts et al.(1997) [32] p. 330). In [60], Hill did not investigate the asymptotic behavior of the estimator.

Mason who proved the weak consistency in (1982) [67], The strong consistency was proved by Deheuvels, Hausler and Mason (1988) in [24] who gave an optimal rate of convergence for an appropriately chosen sequence k_n. The a symptotic normality was established, under some extra condition on F, by Csörgő and Mason [17] and Hausler and Teugels [61] in (1985). Recently, Beirlant, Bouquiaux and Werker in (2006) [2] derived a local asymptotic normality result showing that the asymptotic variance of Hill's estimator attains a lower bound.

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

Theorem 2.1 (Asymptotic properties of $\widehat{\gamma}_n^{(H)}$) *Assume that $F \in D\left(\Phi_{1/\gamma}\right)$ (is belonging to the domain of attraction of Frechet) i.e. F satisfied the condition 2.1, then for $\gamma > 0$, $k := k_n \to \infty$ and $k/n \to 0$ as $n \to \infty$.*

1. *For the weak consistency,*
$$\widehat{\gamma}_n^{(H)} \xrightarrow{P} \gamma \text{ as } n \to \infty.$$

2. *For the strong consistency, if $k/\log\log n \to \infty$ as $n \to \infty$, then*
$$\widehat{\gamma}_n^{(H)} \xrightarrow{a.s.} \gamma \text{ as } n \to \infty.$$

3. *For the symptotic normality, assume that F satisfies the second order regular vari-*

ation assumption that defined in subsection (1.7.2), i.e., for $x > 0$,

$$\lim_{t\to\infty} \frac{U(tx)/U(t) - x^{\gamma}}{A(t)} = x^{\gamma}\frac{x^{\rho}-1}{\rho},$$

or equivalently,

$$\lim_{t\to\infty} \frac{\overline{F}(tx)/\overline{F}(t) - x^{-1/\gamma}}{A\left(1/\overline{F}(t)\right)} = x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho},$$

where $\gamma > 0$, $\rho \leq 0$, *and* A *is a function with* $\lim_{t\to\infty} A(t) = 0$. *Then, if* $\sqrt{k}A(n/k) \to \lambda < \infty$ *as* $n \to \infty$,

$$\sqrt{k}\left(\widehat{\gamma}_n^{(H)} - \gamma\right) \to \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right).$$

2.2.2 Pickands estimator

This estimator is simplest and oldest estimator for γ, was introduced in 1975 by J. Pickands in [75] for any $\gamma \in \mathbb{R}$, and thus it can be used to estimate the shape parameter of any one of the three types of extreme value distributions. But, as it is rather unworkable in practise for small or moderate samples, several refinements were introduced mainly by Drees (1996) in [30]. The derivation of the estimator is based on an equivalent condition to $F \in D(H_{\gamma})$, namely assertion 1.18 in proposition 1.5 which for $x = 2$ and $y = 1/2$ yields

$$\lim_{t\to\infty} \frac{U(2t) - U(t)}{U(t) - U(t/2)} = 2^{\gamma}.$$

Furthermore, for any positive function c such that $\lim_{t\to\infty} c(t) = 2$, we have that

$$\lim_{t\to\infty} \frac{U(c(t)\,t) - U(t)}{U(t) - U(t/c(t))} = 2^{\gamma}.$$

The basic idea now consists of constructing an empirical estimator using this formule. To that effect, let the ordered $(Y_{1:n}, Y_{2:n}, ..., Y_{n:n})$ from a standard Pareto rv Y i.e. the df is defined by

$$F(y) = 1 - \frac{1}{y}, \text{ for any } y \geq 1.$$

In the fact that

$$\frac{k}{n}Y_{n-k+1:n}\overset{P}{\to}1 \text{ and } \frac{Y_{n-k+1:n}}{Y_{n-2k+1:n}}\overset{P}{\to}2, \text{ as } n\to\infty,$$

see e.g., [35], yields

$$\frac{U\left(Y_{n-k+1:n}\right)-U\left(Y_{n-2k+1:n}\right)}{U\left(Y_{n-2k+1:n}\right)-U\left(Y_{n-4k+1:n}\right)}=2^{\gamma}.$$

Finally, we use the distributional identity

$$X_{n-i+1:n}\overset{d}{=}U\left(Y_{n-i+1:n}\right),\ i=1,2,...,n$$

we get the following definition of the Pickands estimator $\widehat{\gamma}_n^{(P)}$.

Definition 2.2 (Pickand's estimator $(\gamma\in\mathbb{R})$**)**

$$\widehat{\gamma}_n^{(P)}=\widehat{\gamma}_n^{(P)}(k):=(\log 2)^{-1}\log\frac{X_{n-k:n}-X_{n-2k:n}}{X_{n-2k:n}-X_{n-4k:n}}.$$

In the following theorem the consistensy and asymptotic normality of $\widehat{\gamma}_n^{(P)}$.

Theorem 2.2 (Asymptotic properties of $\widehat{\gamma}_n^{(P)}$) *Assume that $F\in D(H_\gamma)$, $\gamma\in\mathbb{R}$, $k:=k_n\to\infty$ and $k/n\to 0$ as $n\to\infty$.*

1. *For the weak consistency,*

$$\widehat{\gamma}_n^{(P)}\overset{P}{\to}\gamma \text{ as } n\to\infty.$$

2. *For the strong consistency, if $k/\log\log n\to\infty$ as $n\to\infty$, then,*

$$\widehat{\gamma}_n^{(P)}\overset{a.s.}{\to}\gamma \text{ as } n\to\infty.$$

3. *For the asymptotic normality, suppose that U has a positive derivative U' and that $\pm t^{1-\gamma}U'(t)$ (with either choice of sign) is satisfies the second order regular variation assumption at infinity with auxiliary function a. If $k=o\left(n/g^{\leftarrow}(n)\right)$, where $g(t):=$*

$t^{3-2\gamma}\left(U'(t)/a(t)\right)^2$, *then*

$$\sqrt{k}\left(\widehat{\gamma}_n^{(P)}-\gamma\right)\rightarrow\mathcal{N}\left(0,\eta^2\right)\ \ as\ n\rightarrow\infty,$$

where

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

2.2.3 Moment estimator

This estimator has been introduced as a direct extension or generalization of the Hill estimator by Dekkers and al. (1989) in [26], that is similar to the Hill estimator but one that can be used for general case of γ $(\gamma\in\mathbb{R})$. Before stating the definition of the Moment estimator, we reminde on the following statistics

$$M_n^{(\alpha)}=M_n^{(\alpha)}(k):=\frac{1}{k}\sum_{i=1}^{k}\left[\log\left(X_{n-i+1:n}\right)-\log\left(X_{n-k:n}\right)\right]^\alpha, \tag{2.2}$$

then, for $\alpha=1,2$ we obtains $M_n^{(1)}$ and $M_n^{(2)}$ as empirical moments (of $(\log X)^\alpha$ calculated at the threshold $t=\log\left(X_{n-k:n}\right)$ respectively of the 1st order moment and the 2th order moment) and with them we can proposed the Moment estimator $\widehat{\gamma}_n^{(M)}$ of γ as in the following definition.

Definition 2.3 (Moment estimator $(\gamma\in\mathbb{R})$**)**

$$\widehat{\gamma}_n^{(M)}=\widehat{\gamma}_n^{(M)}(k):=M_n^{(1)}+1-\frac{1}{2}\left(1-\frac{\left(M_n^{(1)}\right)^2}{M_n^{(2)}}\right)^{-1}.$$

Remark 2.1 *Notice that, the case of $\alpha=1$ in 2.2 yields Hill's estimator $\widehat{\gamma}_n^{(H)}$.*

After that we go to the asymptotic properties of the estimator $\widehat{\gamma}_n^{(M)}$.

Theorem 2.3 (Asymptotic properties of $\widehat{\gamma}_n^{(M)}$) *Assume that* $F \in D(H_\gamma)$, $\gamma \in \mathbb{R}$, $k := k_n \to \infty$ *and* $k/n \to 0$ *as* $n \to \infty$.

1. *For the weak consistency,*

$$\widehat{\gamma}_n^{(M)} \xrightarrow{P} \gamma \text{ as } n \to \infty.$$

2. *For the strong consistency, if* $k/(\log n)^\delta \to \infty$ *as* $n \to \infty$, *for some* $\delta > 0$, *then*

$$\widehat{\gamma}_n^{(M)} \xrightarrow{a.s.} \gamma \text{ as } n \to \infty.$$

3. *For the symptotic normality,*

$$\sqrt{k}\left(\widehat{\gamma}_n^{(M)} - \gamma\right) \to \mathcal{N}\left(0, \eta^2\right) \text{ as } n \to \infty,$$

where

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

For the the symptotic normality of $\widehat{\gamma}_n^{(M)}$ see Theorem 3.1and Corollary 3.2 of [26]. For more details on the moment estimator, we reffer to [23].

2.2.4 Kernel type estimators

In the kernel estimation method, K, Csörgő, Deheuvels and Mason (1985) [16] proposed a smoother version of Hill's estimator and proved its consistency and asymptotic normality. For define the Kernel type estimators denoted by $\widehat{\gamma}_n^{(CDM)}$ or by $\widehat{\gamma}_n^{(K)}$, suppose that there is a function K (named kernel function) satisfies :

1. $K(x) \geq 0$ for $0 < x < \infty$ (non-negative),

2. $K(.)$ is non-increasing and right continuous function on $(0, \infty)$,

3. $\int_0^\infty K(x)\,dx = 1$,

4. $\int_0^\infty x^{-1/2} K(x)\,dx < \infty$.

Definition 2.4 (Kernel type estimators)

$$\widehat{\gamma}_n^{(K)} = \widehat{\gamma}_n^{(K)}(h) := \frac{\sum_{i=1}^{n-1} \left(\frac{i}{nh}\right) K\left(\frac{i}{nh}\right) (\log X_{n-i+1:n} - \log X_{n-i:n})}{\int_0^{1/h} K(x)\,dx},$$

or (by remplacing $\int_0^{1/h} K(x)\,dx$ *whith* $(nh)^{-1}\sum_{i=1}^{n} K\left(\frac{i}{nh}\right)$ *because the both are the same),*

$$\widehat{\gamma}_n^{(K)} = \frac{\sum_{i=1}^{n-1} \left(\frac{i}{nh}\right) K\left(\frac{i}{nh}\right) (\log X_{n-i+1:n} - \log X_{n-i:n})}{(nh)^{-1}\sum_{i=1}^{n} K\left(\frac{i}{nh}\right)} \tag{2.3}$$

where $h > 0$ *is called bandwidth or parameter of smoothing.*

Remark 2.2 *Notice that, using the uniform kernel* $K = I(0,1)$ *and* $h = k/n$ *in 2.3 yields Hill's estimator* $\widehat{\gamma}_n^{(H)}$ *as a special case.*

Remark 2.3 *There are many examples for the kernel function* K*, we reminded of some kernel functions in the table 2.1.*

The name of kernel function	The formula
Biweight	$K(s) := \frac{15}{8}\left(1-s^2\right)^2 \mathbb{I}_{\{0\le s<1\}}$
Triweight	$K(s) := \frac{35}{16}\left(1-s^2\right)^3 \mathbb{I}_{\{0\le s<1\}}$
Gaussian	$K(s) := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2}$

Table 2.1: Some example for the kernel function

The Kernel type estimators $\widehat{\gamma}_n^{(K)}$ depends in a continuous way on the bandwidth h representing the proportion of top order statistics used. Under some additional conditions on the kernel K Csörgő and al. (1985) proved its consistency and asymptotic normality of their Kernel type estimators and for a discussion of these conditions we refer to [16].

Under von Mises's condition, the kernel tail index estimators of Csörgő and al. (1985) have been generalized by Groeneboon and al. (2003) in [46] for all real tail indices. Weak consistency and asymptotic normality of their kernel estimators have been established. In [70] (A functional law of the iterated logarithm for kernel-type estimators of the tail index), Necir present a characterization of the almost sure behavior of these estimators and they show their strong consistency.

2.2.5 Probability-weighted moment estimator

This estimator is only applicable on the case that $\gamma < 1$, First let us consider the probability-weighted moment estimator of Hosking and Wallis (1987) in [56]. The starting point is the observation that if Y is a random variable with a generalized Pareto distribution, i.e., with distribution function

$$H_{\gamma,\alpha}(x) := 1 - \left(1 + \frac{\gamma x}{\alpha}\right)^{-1/\gamma},\ 0 < x < \frac{\alpha}{0 \vee (-\gamma)},$$

where $\alpha > 0$ and γ are real parameters, then for $\gamma < 1$,

$$EY = \int_0^{\alpha/0\vee(-\gamma)} \overline{H}_{\gamma,\alpha}(x)\,dx = \frac{\alpha}{1-\gamma}, \tag{2.4}$$

then, define which can be called a probability-weighted moment as follows :

$$E\left[Y\left(\overline{H}_{\gamma,\alpha}(Y)\right)\right] = \int_0^{\alpha/0\vee(-\gamma)} \overline{H}^2_{\gamma,\alpha}(x)\,dx = \frac{\alpha}{2(2-\gamma)}. \tag{2.5}$$

By solving relations 2.4 and 2.5 obtain

$$\gamma = \frac{EY - 4E\left[Y\left(\overline{H}_{\gamma,\alpha}(Y)\right)\right]}{EY - 2E\left[Y\left(\overline{H}_{\gamma,\alpha}(Y)\right)\right]} \tag{2.6}$$

and

$$\alpha = \frac{2\left(EY\right)E\left[Y\left(\overline{H}_{\gamma,\alpha}\left(Y\right)\right)\right]}{EY - 2E\left[Y\left(\overline{H}_{\gamma,\alpha}\left(Y\right)\right)\right]}. \tag{2.7}$$

Then, for independent and identically distributed random variables $X_1, X_2, ..., X_n$ with distribution function F and suppose that F is in the domain of attraction of an extreme value distribution $D\left(H_\gamma\right)$. $X_{1:n}, X_{2:n}, ..., X_{n:n}$ are the corresponding order statistics, let the intermidiate $X_{n-k:n}$ where $k := k_n \to \infty$ and $k/n \to 0$ as $n \to \infty$. Finally, for define the Probability-weighted moment estimator we need to define the following empirical statistics

$$P_n := \frac{1}{k}\sum_{i=1}^{k} \log X_{n-i:n} - \log X_{n-k:n},$$

$$Q_n := \frac{1}{k}\sum_{i=1}^{k} \frac{i}{k} \log X_{n-i:n} - \log X_{n-k:n}.$$

For the construction of P_n and Q_n read de Haan ana Ferreira (2006) [22].

Definition 2.5 (Probability-weighted moment estimator $(\gamma < 1)$)

$$\widehat{\gamma}_n^{(PWM)} = \widehat{\gamma}_n^{(PWM)}\left(k\right) := \frac{P_n - 4Q_n}{P_n - 2Q_n} = 1 - \left(\frac{P_n}{2Q_n} - 1\right)^{-1}.$$

Theorem 2.4 ***Proposition 2.1 (Asymptotic properties of $\widehat{\gamma}_n^{(PWM)}$)*** *Assume that $F \in D\left(H_\gamma\right)$, $\gamma < 1$, $k := k_n \to \infty$ and $k/n \to 0$ as $n \to \infty$.*

1. *For the weak consistency,*

$$\widehat{\gamma}_n^{(PWM)} \xrightarrow{P} \gamma \text{ as } n \to \infty.$$

2. *For the symptotic normality, If F satisfies the second-order condition with $\gamma < 1/2$, and $\lim_{n\to\infty} \sqrt{k}A\left(n/k\right) \to \lambda < \infty$ as $n \to \infty$, then*

$$\sqrt{k}\left(\widehat{\gamma}_n^{(PWM)} - \gamma\right) \to \mathcal{N}\left(E, M\right) \text{ as } n \to \infty,$$

where E is the mean vector

$$\begin{cases} \frac{\gamma}{(1-\gamma-\rho)(2-\gamma-\rho)}\left((1-\gamma)(2-\gamma),-\rho\right), & \rho<0, \\ \lambda(1,0), & \gamma\neq\rho=0, \\ \lambda\left(1,-\frac{1}{2}\right), & \gamma=\rho=0, \end{cases}$$

and M is the covariance matrix

$$\begin{pmatrix} \frac{(1-\gamma)(2-\gamma)^2\left(1-\gamma+2\gamma^2\right)}{(1-2\gamma)(3-2\gamma)} & \frac{(2-\gamma)\left(-2+6\gamma-7\gamma^2+2\gamma^3\right)}{(1-2\gamma)(3-2\gamma)} \\ \frac{(2-\gamma)\left(-2+6\gamma-7\gamma^2+2\gamma^3\right)}{(1-2\gamma)(3-2\gamma)} & \frac{31-94\gamma+102\gamma^2-126\gamma^3+144\gamma^4-80\gamma^5+16\gamma^6}{(1-2\gamma)(3-2\gamma)} \end{pmatrix}.$$

For $1/2<\gamma<1$ the convergence of $\widehat{\gamma}_n^{(PWM)}$ to γ is slower than that for $\gamma<1/2$, and for the proove or more details see de Haan ana Ferreira (2006) [20]. Hosking and Wallis (1987) in [56] derive a simple method of moments to estimate γ and σ, but this only works if $\gamma<1/2$. They also apply a variant with probability weighted moments (PWM) and find that the corresponding EVI estimator is a good alternative to the Maximum Likelihood estimator for $\gamma<1$.

2.3 Choice of optimal number of extremes k_n

The results of estimation of the extreme values are asymptotic: they are obtain as $k_n\to\infty$ and $k_n/n\to 0$. Then, using too many data, in the estimation procedure, results in a substantial bias whereas using too few observations leads to a considerable variance. For this, it is necessary to make a trade-off between bias and variance by the choice of optimal number of extremes. We refer to Meraghni-thesis [69] for more details in this section.

2.3.1 Graphical method

This method is the more simple to make an optimal choice of $k := k_n$, consists of using the plot

$$\{(k, \widehat{\gamma}_n(k)) : 1 \leq k < n\}.$$

Some other graphical procedures for selecting an optimal k_n value are extensively discussed and compared in Sousa (2002) [83] (a good reference for the interested reader) where the author introducesa new method which he calls the Sum plot. Then, the graphical should choose the optimal k_n (noted k_{opt}) in the first region where the plot of the estimator $\widehat{\gamma}_n(k)$ is roughly horizontal. For an illustration see 2.1, where it seems that any k_n between 80 and 100 would be a good choice.

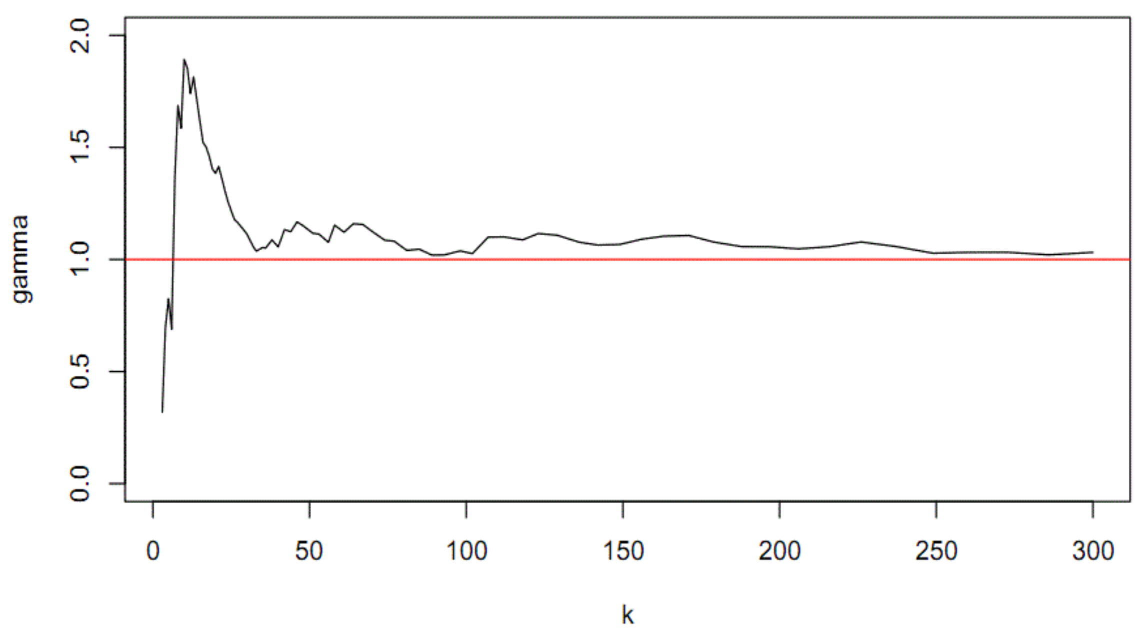

Figure 2.1: Plot of Hill's estimator, for the EVI of a standard Pareto distrubution, as a function of the number of top statistics, based on 100 samples of size 3000. The horizontal line represents the true value of the tail index.

2.3.2 Method bassed on the mean squared error

This is an outher method in the choice of optimal number of extremes k_{opt}, it is based on the minimisation of the asymptotic mean of squared error of the estimator $\widehat{\gamma}_n(k)$ ($AMSE$), that is defined by

$$AMSE\left(\widehat{\gamma}_n(k)\right) := E_\infty\left[\left(\widehat{\gamma}_n(k) - \gamma\right)^2\right],$$

where E_∞ is the mathematical expectation corresponding to the asymptotic distribution. It is therefore easy to assume that the mean squared error of $\widehat{\gamma}_n(k)$ that is a function of k is equal to bias squared plus variance. for any classical estimator, for a precise estimate and an accurate estimation of the tail index γ, it is necessary to make a trade-off between bias and variance. It seems reasonable that a minimization of the $AMSE$ makes it possible to find an intermediate value between the components of the bias and the variance for this compromise. That is, the optimal choice of k, denoted by k_{opt}, corresponds to the smallest $AMSE$, this optimal number of extremes k_{opt} defined by

$$k_{opt} := \arg\min_k AMSE\left(\widehat{\gamma}_n(k)\right).$$

2.3.3 Numerical procedure

This method is proposed because the process of choosing k_{opt} is made dificult by the fact that the latter does not exclusively depend on the sample size and the index but it also depends upon unknown features (the second order parameter ρ among others) of the underlying df F. En effet, for example we used the theorem 2.1 the asymptotic mean of squared error of the Hill estimator $\widehat{\gamma}_n^{(H)}(k)$ is $AMSE\left(\widehat{\gamma}_n^{(H)}(k)\right) := AMSE\left(\widehat{\gamma}_n^{(H)}(k), A(n/k), \rho\right)$ defined by

$$AMSE\left(\widehat{\gamma}_n^{(H)}(k)\right) = \frac{\gamma^2}{k} + \left(\frac{A\left(\frac{n}{k}\right)}{1-\rho}\right)^2.$$

To overcome this hurdle, a large variety of algorithms and data-adaptive procedures of computing consistent estimate $\widehat{k}_{opt}$ for k_{opt} in the sense that $\widehat{k}_{opt}$ satisfes as $n \rightarrow \infty$

$$\frac{\widehat{k}_{opt}}{k_{opt}} \xrightarrow{P} 1.$$

The estimator of the tail index correspond $\widehat{\gamma}_n^{(H)}\left(\widehat{k}_{opt}\right)$ is asymptotically as efficient as $\widehat{\gamma}_n^{(H)}\left(k_{opt}\right)$. There are many papers dealing with this issue of selecting the optimal fraction of top statistics to be used when estimating an EVI and we refer to Hall and Welsh (1985) [53], Reiss and Thomas (1997) [76] and Cheng Peng (2001) [14]. Since we are used in this thesis the algorithm of Reiss and Thomas (1997) (see [76] page 137) in the choice of optimal number of extremes k_{opt}, then, we just remind of them.

Reiss and Thomas approach

This approach based of choosing the adequate number of highest observations on minimizing a mean distance summing up a penalty term in [76]. More precisely, they propose an automatic manner to choose k by minimizing

$$\frac{1}{k}\sum_{i\leq k} i^{\theta}\left|\widehat{\gamma}_n\left(i\right)-med\left(\widehat{\gamma}_n\left(1\right),...,\widehat{\gamma}_n\left(k\right)\right)\right|,\ 0\leq\theta\leq 1/2,$$

or the following suggested modification

$$\frac{1}{k-1}\sum_{i<k} i^{\theta}\left(\widehat{\gamma}_n\left(i\right)-\widehat{\gamma}_n\left(k\right)\right)^2,\ 0\leq\theta\leq 1/2.$$

The performance of this methodology is discussed and evaluated by Neves and Fraga Alves in [72] by substantially reducing the domain of variation of the weight θ of the penalty term i^{θ}, considering Hill's and the moment estimators. Depending on the prior information one might have about the value of the EVI, the authors provide, for each estimator, suitable values of θ and indicate which expression (out of the two above) to minimize. On the light of a thorough simulation study they come up with the overall conclusion that the most proper choice for θ is 0. A fully detailed list of all possible combinations, with particular θ-values, is to be found in [72]. The latter reference also contains a brief discussion with summarized results of the methods of Dekkers and de Haan [27] and de Haan and Peng [21]. For more details in numerical procedure see thesis of Pr. Meraghni (2008) in [69] page 63.

2.4 Quantile extreme estimation

In this section, we are interested in estimating the quantile extreme of order $p \in (0,1)$ corresponds to p is close to 0. the quantile of order $1-p$ (or $(1-p)$-quantile) of df F is defined by

$$x_p := F^{\leftarrow}(1-p) = Q(1-p) = U(1/p).$$

For a sample $(X_1, X_2, ..., X_n)$ withdrawn from a df F, p must depend on the sample size n i.e., $p =: p_n$. There are two cases possible for x_p :

- The first situation where $p_n \to 0$ with $np_n \to c \in [1, \infty]$ as $n \to \infty$ the quantile of order $1-p$ is within the sample.
- The other situation is $p_n \to 0$ with $np_n \to c \in [0, 1)$ as $n \to \infty$ the quantile of order $1-p$ is outside the sample.

In other words, the with in-sample estimation is possible up to the quantile of order $1/n$ where as for $p < 1/n$, quantile estimates are beyond the range of the data. The latter case is the most relevant for purposes of real-life applications.

In the first case the natural estimator of $(1-p_n)$-quantile is the order statistic $X_{n-i+1:n}$ with $p_n = (i-1)/n$. En effet, for $s = 1 - p_n$

$$Q_n\left(1 - \frac{i-1}{n}\right) = X_{n-i+1:n}, \text{ for } i =, 2, ..., n.$$

For the other case there are many estimation based on the estimating of the tail index.

2.4.1 Approach based on a positive index estimator

Weissman in 1978 provided the classic semi-parametric extreme quantile estimator where defined by

$$\widehat{x}_{k_n}^{W}(p_n) := X_{n-k_n:n}\left(\frac{k_n}{np_n}\right)^{\widehat{\gamma}_{k_n}}$$

where $\widehat{\gamma}_{k_n}$ is some consistent positive index estimator of γ, if we remplace $\widehat{\gamma}_{k_n}$ by $\widehat{\gamma}_{k_n}^{H}$ the Hill estimator, we get (the Weissman-Hill estimator) in [86] the well know estimator of the extreme quantile $x(p_n)$, which is defined by

$$\widehat{x}_{k_n}^{W-H}(p_n) := X_{n-k_n:n}\left(\frac{k_n}{np_n}\right)^{\widehat{\gamma}_{k_n}^{H}}.$$

The asymptotic properties of this estimator are discussed and confidence intervals constructed under some conditions on F, k_n and p_n in, e.g. [32], [68], [38] and [66].
Frederico Caeiro and Dora Prata Gomes (2015) in [11] proposed Pareto Log Probability Weighted Moment (PLPWM) estimator in the semi- parametric estimation of extreme quantiles of a right heavy-tail model, which is defined as the following form:

$$\widehat{x}_{k_n}^{PLPWM}(p_n) := \left(\frac{k_n}{np_n}\right)^{\widehat{\gamma}_{k_n}^{PLPWM}} \exp\left\{\frac{1}{k}\sum_{i=1}^{k}\left(4\frac{i-1}{k-1}-1\right)\ln X_{n-i+1:n}\right\},\ k_n = 2, ..., n$$

where

$$\widehat{\gamma}_{k_n}^{PLPWM} := \frac{1}{k}\sum_{i=1}^{k}\left(2-4\frac{i-1}{k-1}\right)\ln X_{n-i+1:n}$$

$\widehat{\gamma}_{k_n}^{PLPWM}$ is the PLPWM estimator of γ. For the asymptotic results of the estimators see [11].

2.4.2 Approach based on a negative index estimator

For $\gamma < 0$. If we use the Pickands estimator $\widehat{\gamma}_{k_n}^P$, we get the estimator $\widehat{x}_{k_n}^P(p_n)$ which is derived in [25] by Dekkers and de Haan (1989), it is defined as following

$$\widehat{x}_{k_n}^P(p_n) := X_{n-k_n+1:n} + (X_{n-k_n+1:n} - X_{n-2k_n+1:n}) \frac{(k_n/np_n)^{\widehat{\gamma}_{k_n}^P} - 1}{1 - 2^{-\widehat{\gamma}_{k_n}^P}}.$$

For more details, several asymptotic results and examples see [26].

2.4.3 Approach based on any index estimator

By using the estimator of moments, Dekkers and al (1989) proposed an estimate of $x(p_n) = U(1/p_n)$ pour $p_n \to 0$ as following:

$$\widehat{x}_{k_n}^M(p_n) := X_{n-k_n:n} + X_{n-k_n:n}\widehat{\gamma}_{k_n}^H \max\left(1 - \widehat{\gamma}_{k_n}^M, 1\right)\left(\frac{k_n}{np_n}\right)^{\widehat{\gamma}_{k_n}^M}.$$

Asymptotic results are established in [26] provided conditions on F, k_n and p_n.

2.5 Second order parameter estimation

The theory of extreme values establishes that the asymptotic law of the normalised maximum $X_{n:n} = \max(X_1, X_2, ..., X_n)$ is given by

$$G_\gamma(x) = \exp\left(-(1+\gamma x)_+\right)^{-1/\gamma}$$

where $y_+ = \max(y, 0)$. The unknown parameter $\gamma \in \mathbb{R}$ is called the extreme value index or a tail index. Note that γ is the parameter of importance in the theory of extreme values, it controls the behavior of the queue to the first order. The larger the γ, the larger the tail, it controls the heaviness of the tail of this distribution. Several estimators have been proposed in the statistical literature and their asymptotic distributions have been established from a second-order condition given by 1.7 or 1.8, that they are respectively equivalent to

$$\lim_{t\to\infty}\frac{\log\left(\overline{F}\left(tx\right)\right)-\log\left(\overline{F}\left(t\right)\right)-\gamma^{-1}\log x}{A^{*}\left(t\right)}=\frac{x^{\rho}-1}{\rho}.$$

and

$$\lim_{t\to\infty}\frac{\log\left(U\left(tx\right)\right)-\log\left(U\left(t\right)\right)-\gamma\log x}{A\left(t\right)}=\frac{x^{\rho}-1}{\rho}.$$

This parameter is of practical relevance in extreme value analysis due its crucial importance in selecting the optimal number of upper order statistics k in tail index estimation (see, e.g., [20]) and the estimator of ρ is also used for the reduction of the bias of the estimators of the index of extreme values (see for example [47] and [13]), or for Weibull's tail coefficient, in [29], even if the bias reduction is obtained with the canonical choice $\rho=-1$, as is suggested in [39]. In the case of complete data, this problem has received a lot of attention from many authors like, for instance, [73], [36], [46], [74], [50], [87], [89], [28], [53] and [14].

In the case of censored data this issue is addressed recently in [4] by considering the adapted tail index estimator introduced by [31].

Inspired by the paper of [46], we propose an estimator for ρ adapted to the random right-truncation case in the chapter four.

Chapter 3

Incomplete data

In lifetime data, truncation and censoring occur quite naturally and the estimation in this cases is very important. For this, our work deals with incomplete data, and in order to make the thesis easier to read, we give some definitions and examples of the incomplete data, i.e. truncated or censored data. There are three general types of censoring and of truncating: right, left and interval. In this thesis our motivation is based on the right truncation.

3.1 Censored data

In this section, we will focus on discussing censored data. The phenomenon of censorship is linked to the disturbing events which may occur in the time required for the collection of data. Subsequently, censoring is when an observation is incomplete due to some random case. The cause of the censoring must be independent of the event of interest if we are to use standard methods of analysis. So, When a data set contains observations within a restricted range of values, but otherwise not measured, it is called a censored data set.

Definition 3.1 *The censoring variable Y is defined by the non-observation of the studied event. if instead of observing X, we observe Y, and we know that $X > Y$ (respectively*

$X < Y$, $Y1 < X < Y2$), we say that there is right censoring (respectively left censoring, interval censoring).

3.1.1 Types of censoring

In the literature we find the following types:

Type 1: fixed censoring

The experimenter fixes a value (a date for example non-random end of experience). For example, in epidemiology, the maximum duration of participation is fixed and, for each observation, the difference between the end of experience and the date of the patient's entry into the study. The number of events observed is, on the other hand, random.
Let Y be a fixed value. For example, in the right censoring, instead of observing the variables $X_1, ..., X_n$ which are of interest to us, we observe Y_j when it is less than or equal to a fixed duration Y. We observe a variable Z_j where $Z_j := \min(X_j, Y_j)$, $j = 1, 2, ..., n$. This mechanism of censorship is frequently encountered in industrial applications.

Type 2: censorship waiting

The experimenter fixes a priori the number of events to be observed. The end date of the experiment then becomes random, the number of events being non-random. This model is often used in reliability and epidemiology studies. For example, in epidemiology, it is decided to observe the survival durations of the n patients until r $(1 \leq r \leq n)$ of them deceaded and stop the study at that time. Let $X_{j:n}$ and $Z_{j:n}$ be the order statistics of variables X_j and Z_j. The date of censorship is therefore $X_{r:n}$ and one observes

$$\begin{cases} Z_{j:n} = X_{j:n}, \text{ if } j \leq r, \\ Z_{j:n} = X_{j:n}, \text{ if } j \geq r. \end{cases}$$

Type 3: random censoring

It is typically this model that is used for therapeutic trials. In this type of experiment, the date of inclusion of the patient in the study is fixed, but the end of observation date is unknown (this corresponds, for example, to the patient's hospital stay). Let $X_1, ..., X_n$ be a sample of a positive rv X, we say that there is random censoring of this sample if there exists another positive rv Y of sample $Y_1, ..., Y_n$ in this case instead of observing the X_j, observe a couple of rv $(Z_j, \delta j)$ with

$$Z_j := \min(X_j, Y_j), \text{ and } \delta_j := 1(X_j \leq Y_j) \text{ for } j = 1, 2, ..., n, \quad (3.1)$$

where δ_j is named the indicator of censore, which determines whether X has been censored or not:

If $\delta_j = 1$ the duration of interest is observed $(Z_j = X_j)$

If $\delta_j = 0$ it is censored $(Z_j = Y_j)$. There are incomplete durations.

Censored data can be further classified into three categories: right censoring, left censoring and interval censoring, as follows:

3.1.2 Right censoring

The variable of interest is said to be censored to the right if the individual concerned has no information about his last observation. Thus, in the presence of the right censoring, the variables of interest are not observed all.

A typical example is where the event considered is the death of a sick patient and the duration of observation is a total duration of hospitalization. This phenomenon is also found in reliability studies when the failure of an electronic device or component does not allow observation to continue for another device or component.

These kinds of phenomena can also be found in hydrology, rainfall, etc. The experimenter can set an end-of-experience date and observations for individuals for whom the event of interest has not been observed before this date will be censored on the right.

3.1.3 Left censoring

There is left censoring when the individual has already suffered the event before it is observed. It is known only that the variable of interest is less than or equal to a known variable. For example, if a certain electronic component is to be reliably studied, which is connected in parallel with one or more other components: the system can continue to function, albeit aberrantly, until this fault is detected (for example during a check or when the system is shut down). Thus, the duration observed for this component is censored on the left. In life there are several phenomena which present both censored data on the right and on the left.

Remark 3.1 *We say that we have double or mixed censorship if we have censored data on the right and censored data on the left in the same sample, see Lawless* (2002) *page 66 in [65].*

3.1.4 Interval censoring

In this case, as its name indicates, we observe both a lower bound and an upper bound of the variable of interest. This model is generally found in medical follow-up studies where patients are monitored periodically if a patient does not show up for one or more controls and then presents himself after the event of interest has occurred. We also have this kind of data that is censored on the right or, more rarely, on the left. An advantage of this type is that it makes it possible to present the censored data to the right or to the left by intervals of the type $[a, +\infty]$ or $[0, a]$ respectively.

3.2 Truncated data

Censored data are not the only type of incomplete data. The other classic case of incomplete data is that of the so-called truncated data.The truncated data are frequently used on the life time study. Some examples of truncated data from astronomy and economics can be found in Woodroofe (1985) [88] and for applications in the analysis of AIDS data, see Wang (1989) [85]. In reliability, are al data set, consisting in life times of automobile brake pads and already considered by Lawless (2002) in [65] page 69, was recently analyzed in Gardes and Stupfer (2015) [42] as an application of randomly truncated heavy-tailed models.

The phenomenon of truncation is very different from censorship. Truncation is not uncommon for the variable of interest X not to be observable when it is less than a random threshold Y.

Definition 3.2 *Truncation eliminates from the study a part of the $X_j's$ of the variable of interest X and suppose a random threshold Y, the analysis in this case can only relate to the conditional law of X where $X > Y$.*

There are three types of truncation: left $(X > Y)$, right $(X < Y)$ and interval $(Y_1 < X < Y_2)$ truncation defined as follows:

3.2.1 Left truncation

due to structure of the study design, we can only observe those individuals whose event time is greater than some truncation threshold. As example, imagine you wish to study how long people who have been hospitalized for a heart at tack survive taking some treatment at home. The start time is taken to be the time of the heart at tack. Only those individuals who survive their stayin hospital are able to be included in the study.

3.2.2 Right truncation

only individuals with event time less than some thre shold are included in the study. As example, if you ask a group of smoking school pupils at what age they started smoking, you necessarily have truncated data, as individuals who start smoking after leaving schoolare not included in the study.

3.2.3 Interval truncating

Or doubly truncated failure-time arises if an individual is potentially observed and only if its failure-time falls within a certain interval,unique to that individual. Doubly truncated data play an important role in the statistical analysis of astronomical observations as well as in survival analysis.

3.3 Estimation in the case of incomplete data

When the empirical data is incomplete (truncated or censored), empirical estimators will not produce good results. In this section we show the importante estimators of the distribution function wish is used to estimate any statistics for example to estimate the mean, tail index and quantiles.

3.3.1 Estimation under random right censoring model

There are two techniques available to determine the distribution function based on the data. The Kaplan-Meier product limit estimator can be used to generate a survival distribution function. The Nelson-Aalen estimator can be used to generate a cumulative hazard rate function.

Let $X_1, ..., X_n$ be n $(n \geq 1)$ independent copies of a non-negative random variable (rv) X, defined over some probability space (Ω, A, P), with continuous cumulative distribution function (cdf) F. These rv's are censored to the right by a sequence of independent copies $Y_1, ..., Y_n$ of a non-negative rv Y from some (censoring) d.f. G, being also independent of X. Let $\{(Z_j, \delta j)\,,\ 1 \leq j \leq n\}$ is a couple of rv observed and defined in 3.1. Let be Z

from cdf H. This model is very useful in a variety of areas where random censoring is very likely to occur such as in biostatistics, medical research, reliability analysis, actuarial science,.... In the following we represent two types of df estimators.

Kaplan-Meier estimator

In the context of this randomly right censored observations, the nonparametric maximum likelihood estimator of the survival distribution F is given in [63] by Kaplan and Meier in (1958) as the product limit estimator defined by

$$\widehat{F}_n(x) := \begin{cases} 1 - \prod_{Z_{i:n} \leq x}^{n} \left[1 - \frac{\delta_{[i:n]}}{n-i+1}\right] & \text{si} \quad x < Z_{n:n}, \\ 1 & \text{si} \quad x \geq Z_{n:n}. \end{cases} . \tag{3.2}$$

where $Z_{1:n} \leq ... \leq Z_{n:n}$ denote the order statistics associated to $Z_1, ..., Z_n$ and $\delta_{[1:n]}, ..., \delta_{[n:n]}$ is the associated concomitants, that is, $\delta_{[i:n]} = j$ if $Z_{i:n} = Z_j$. This estimator may be expressed in the form of sum. This writing can be found in the book of Reiss and Thomas [76] (page 162) as follows

$$\widehat{F}_n(x) := \sum_{i=2}^{n} W_{i:n} 1_{\{x < Z_{n:n}\}}$$

where for $i = 2, ..., n$,

$$W_{i:n} := \frac{\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n]}}.$$

In the literature of survival analysis, many authors have been devoted to the study of the asymptotic properties of the Kaplan-Meier estimator. For example, the uniform consistency has been studied by Shorack et Wellner (1986) [80], Stute et Wang (1993) [81] and Gill (1994) [45]. Asymptotic normality was studied by Breslow et Crowley (1974) [10], Gill (1980) [43] et Gill (1983) [44].

In the last decade, several authors started to be interested in the estimation of tail index along with large quantiles under random censoring as one can see in Gomes and Oliveira (2003) [49], Beirlant et al.(2007) [3], Einmahl et al.(2008) [31] and Worms and Worms

(2014) [90]. Gomes and Neves (2011) [48] also made a contribution to this eld by providing a detailed simulation study and applying the estimation procedures

Nelson-Aalen estimator

Here we consider the outher nonparametric estimator of cdf F, based on Nelson-Aalen estimator (Nelson 1972 [71], Aalen 1976 [1]) of the cumulative hazard function

$$\Lambda(z) := \int_0^z \frac{dF(v)}{\overline{F}(v)} = \int_0^z \frac{dH^{(1)}(v)}{\overline{H}(v)}$$

where $H^{(1)}(z) := P(Z_1 \leq z,\ \delta_1 = 1) = \int_0^z \overline{G}(y)\, dF(y),\ z \geq 0.$

A natural nonparametric estimator Λ_n of Λ is obtained by replacing H and $H^{(1)}$ by their respective empirical counterparts

$$H_n(v) := n^{-1} \sum_{i=1}^{n} 1(Z_i \leq v) \text{ and } H_n^{(1)} := n^{-1} \sum_{i=1}^{n} \delta_i 1(Z_i \leq v)$$

pertaining to the observed Z-sample. Then, the Nelson-Alen estimator is defined by

$$\Lambda_n(z) = \int_0^z \frac{dH_n^{(1)}(v)}{\overline{H}_n(v)} := \begin{cases} \sum_{Z_{j:n} \leq z}^{n} \frac{\delta_{[j:n]}}{n-j+1}, & z < Z_{j:n}, \\ 1, & z \geq Z_{j:n}. \end{cases}$$

Under 1.3, $F(z) = 1 - \exp\{-\Lambda(z)\}$ which by substituting Λ_n for Λ, yields Nelson-Aalen estimator of cdf F, given by

$$F_n^{NA}(z) = \begin{cases} 1 - \prod_{i:Z_{i:n} \leq x}^{n} \exp\left\{-\frac{\delta_{[i:n]}}{n-i+1}\right\}, & \text{for } z \leq Z_{n:n}, \\ 1, & \text{for } z > Z_{n:n}. \end{cases}$$

By considering samples of various sizes, Fleming and Harrington (1984) [34] numerically compared F_n with Kaplan-Meier (non parametric maximum likelihood) estimator of F (Kaplan and Meier, 1958), given in 3.2 and pointed out that they are asymptotically equivalent and usually quite close to each other. A nice discussion on the tight relationship between the two estimators may be found in Huang and Strawderman (2006) [58].

3.3.2 Estimation under random right truncating model

In the truncated data, there are also two estimators of the distribution function one is Woodroofe estimator and the other is Lynden-Bell. The following of the thesis is interested to them.

Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \leq i \leq N$ be a sample of size $N \geq 1$ from a couple $(\mathbf{X}, \mathbf{Y})$ of independent random variables (rv's) defined over some probability space $(\Omega, \mathcal{A}, \mathbf{P})$, with continuous marginal distribution functions (df's) $\mathbf{F}$ and $\mathbf{G}$ respectively. Suppose that $\mathbf{X}$ is truncated to the right by $\mathbf{Y}$, in the sense that $\mathbf{X}_i$ is only observed when $\mathbf{X}_i \leq \mathbf{Y}_i$. This model of randomly truncated data commonly finds its applications in such areas like astronomy, economics, medicine and insurance.

Let us now denote (X_i, Y_i), $i = 1, ..., n$ to be the observed data, as copies of a couple of rv's (X, Y), corresponding to the truncated sample $(\mathbf{X}_i, \mathbf{Y}_i)$, $i = 1, ..., N$, where $n = n_N$ is a sequence of discrete rv's which, in virtue of the weak law of large numbers, satisfies

$$n_N/N \overset{\mathbf{P}}{\to} p := \mathbf{P}(\mathbf{X} \leq \mathbf{Y}), as N \to \infty.$$

We shall assume that $p > 0$, otherwise, nothing will be observed. We denote the joint df of X and Y by

$$H(x, y) := \mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(\mathbf{X} \leq \min(x, \mathbf{Y}), \mathbf{Y} \leq y \mid \mathbf{X} \leq \mathbf{Y}),$$

which is equal to $p^{-1} \int_0^y \mathbf{F}(\min(x, z))\, d\mathbf{G}(z)$. The marginal distributions of the rv's X and Y, respectively denoted by F and G, are given by

$$F(x) = p^{-1} \int_0^x \overline{\mathbf{G}}(z)\, d\mathbf{F}(z) \text{ and } G(y) = p^{-1} \int_0^y \mathbf{F}(z)\, d\mathbf{G}(z).$$

Since right endpoints of F and G are infinite and thus they are equal. Hence, from [88], we may write

$$\int_x^\infty d\mathbf{F}(y)/\mathbf{F}(y) = \int_x^\infty dF(y)/C(y),$$

where

$$C(z) := P(X \leq z \leq Y \mid X \leq Y) = p^{-1}F(x)G(x).$$

Differentiating the previous equation leads to the following crucial equation

$$C(x)\,d\mathbf{F}(x) = \mathbf{F}(x)\,dF(x). \tag{3.3}$$

Woodroofe estimator

The solution of 3.3 is defined by

$$\mathbf{F}(x) = \exp\left\{-\int_x^\infty dF(z)/C(z)\right\},$$

see, for instance, Strzalkowska-Kominiak and Stute (2009) [84].This leads to nonparametric estimator [88] of df $\mathbf{F}$, Woodroofe estimator given by

$$\mathbf{F}_n^{(\mathbf{W})}(x) := \prod_{i:X_i>x} \exp\left\{-\frac{1}{nC_n(X_i)}\right\},$$

which is derived only by replacing df's F and C by their respective empirical counterparts

$$F_n(x) := n^{-1}\sum_{i=1}^n \mathbf{1}(X_i \leq x)$$

and

$$C_n(x) := n^{-1}\sum_{i=1}^n \mathbf{1}(X_i \leq x \leq Y_i). \tag{3.4}$$

The central limit theorem under random truncation was established by Stute and Wang

(2008) in [82].

Lynden-Bell estimator

There exists a more popular estimator for $\mathbf{F}$. Note that the approximation $\exp(t) \sim 1-t$, for small $t > 0$, results in the well-known estimator introduced by Lynden-Bell (1971), known as Lynden-Bell nonparametric maximum likelihood estimator [64], defined by

$$\mathbf{F}_n^{(\mathbf{LB})}(x) := \prod_{i:X_i>x} \left(1 - \frac{1}{nC_n(X_i)}\right),$$

where C_n is defined in 3.4.

There is several applications to randomly right-truncated insurance claims, one refers to [33], and in reliability, one consults for instance [42] and [6] in which they considered lifetimes of automobile brake pads, given by [65] in page 69, as example of randomly truncated heavy-tailed models.

In the last decade, several authors started to be interested in the estimation of tail index along with large quantiles under random truncation as one can see in [42], [5] and [91],

Chapter 4

Estimating the second-order parameter of regular variation and bias reduction in tail index estimation under random truncation

In this chapter, we propose a consistent estimator of the second-order parameter of Pareto-type distributions under random right-truncation and establish its asymptotic normality. Moreover, we derive an asymptotically unbiased estimator for the tail index and study its asymptotic behaviour.

4.1 Introduction

We assume that both survival functions $\overline{\mathbf{F}} := 1 - \mathbf{F}$ and $\overline{\mathbf{G}} := 1 - \mathbf{G}$ are regularly varying at infinity with negative tail indices $-1/\gamma_1$ and $-1/\gamma_2$ respectively, that is, for any $x > 0$

$$\lim_{z\to\infty} \frac{\overline{\mathbf{F}}(xz)}{\overline{\mathbf{F}}(z)} = x^{-1/\gamma_1} \text{ and } \lim_{z\to\infty} \frac{\overline{\mathbf{G}}(xz)}{\overline{\mathbf{G}}(z)} = x^{-1/\gamma_2}. \tag{4.1}$$

To specifying the rates of convergence in (4.1), we consider the second-order conditions in terms of the quantile functions [22], that is, for any $x > 0$

$$\lim_{t\to\infty} \frac{\mathbb{U}_{\mathbf{F}}(tx)/\mathbb{U}_{\mathbf{F}}(t) - x^{\gamma_1}}{\mathbf{A}_{\mathbf{F}}(t)} = x^{\gamma_1}\frac{x^{\rho_1}-1}{\rho_1}, \tag{4.2}$$

and

$$\lim_{t\to\infty} \frac{\mathbb{U}_{\mathbf{G}}(tx)/\mathbb{U}_{\mathbf{G}}(t) - x^{\gamma_2}}{\mathbf{A}_{\mathbf{G}}(t)} = x^{\gamma_2}\frac{x^{\rho_2}-1}{\rho_2}, \tag{4.3}$$

where $\rho_1, \rho_2 < 0$ are the second-order parameters and $\mathbf{A}_{\mathbf{F}}$, $\mathbf{A}_{\mathbf{G}}$ are functions tending to zero and not changing signs near infinity with regularly varying absolute values at infinity with indices ρ_1, ρ_2 respectively. For any df K, we write $\mathbb{U}_K(t) := K^{\leftarrow}(1-1/t)$, $t > 1$, where $K^{\leftarrow}(s) := \inf\{x : K(x) \geq s\}$, $0 < s < 1$ stands for the quantile function.

Let us now denote the observed observations of the truncated sample $(\mathbf{X}_i, \mathbf{Y}_i)$, $i = 1, ..., N$, by (X_i, Y_i), $i = 1, ..., n$, as a sample of a couple of rv's (X, Y), where $n = n_N$ is a sequence of discrete rv's. Note that by the weak law of large numbers we have $n_N/N \overset{\mathbf{P}}{\to} p := \mathbf{P}(\mathbf{X} \leq \mathbf{Y})$, as $N \to \infty$, which implies that $n_N \overset{\mathbf{P}}{\to} \infty$ as $N \to \infty$, where $\overset{\mathbf{P}}{\to}$ stands for convergence in probability. The joint distribution of X_i and Y_i is

$$\begin{aligned} H(x,y) &:= \mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(\mathbf{X} \leq \min(x, \mathbf{Y}), \mathbf{Y} \leq y \mid \mathbf{X} \leq \mathbf{Y}) \\ &= p^{-1}\int_0^y \mathbf{F}(\min(x,z))\, d\mathbf{G}(z). \end{aligned}$$

Thereby, the marginal distributions of the rv's X and Y are equal to

$$F^*(x) := p^{-1}\int_0^x \overline{\mathbf{G}}(z)\, d\mathbf{F}(z) \text{ and } G^*(y) := p^{-1}\int_0^y \mathbf{F}(z)\, d\mathbf{G}(z),$$

respectively. The survival function $\overline{F}^*$ simultaneously depends on $\overline{\mathbf{G}}$ and $\overline{\mathbf{F}}$ while $\overline{G}^*$ only relies on $\overline{\mathbf{G}}$. Since $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$ are regularly varying functions, then by making use of Proposition B.1.10 in [20], we may readily show that both $\overline{G}^*$ and $\overline{F}^*$ are also regularly varying at infinity, with respective indices $-1/\gamma_2$ and $-1/\gamma := -(\gamma_1+\gamma_2)/(\gamma_1\gamma_2)$. In this context, Gardes and Stupfler 2015, [42] recently proposed a consistent estimator to the extreme value index γ_1 by using the definition of γ as a quotient depending only on two Hill estimators [60] of tail indices γ and γ_2 which are based on the upper order statistics $X_{n-k:n} \leq ... \leq X_{n:n}$ and $Y_{n-k:n} \leq ... \leq Y_{n:n}$ pertaining to the samples $(X_1, ..., X_n)$ and $(Y_1, ..., Y_n)$ respectively. The sample fraction $k = k_n$ being a (random) sequence of integers such that, given $n = m = m_N$, $k_m \to \infty$ and $k_m/m \to 0$ as $N \to \infty$. The asymptotic normality of this estimator is established in Benchaira et al. 2015, [5] by assuming both the tail dependence and the second-order regular variation conditions. Recently, Wormes and Wormes 2016, [91] proposed an asymptotically normal estimator for γ_1 by considering a Lynden-Bell integration. Independently, Benchaira et al. 2016a, [6] provided a Hill-type estimator for randomly right-truncated data, defined by

$$\widehat{\gamma}_1^{(BMN)} = \sum_{i=1}^{k} a_n^{(i)} \log \frac{X_{n-i+1:n}}{X_{n-k:n}}, \tag{4.4}$$

where

$$a_n^{(i)} := \frac{\mathbf{F}_n\left(X_{n-i+1:n}\right)/C_n\left(X_{n-i+1:n}\right)}{\sum_{i=1}^{k} \mathbf{F}_n\left(X_{n-i+1:n}\right)/C_n\left(X_{n-i+1:n}\right)}, \tag{4.5}$$

with $\mathbf{F}_n(x)$ is the well-known product-limit Woodroofe's estimator defined by 3.4 of the underlying df $\mathbf{F}$. The authors show by simulation that, for small datasets, their estimator behaves better in terms of bias and root of the mean squared error (rmse), than the Gardes-Supfler estimator. Moreover, they establish the asymptotic normality by considering the second-order regular variation conditions (4.2) and (4.3) with the assumption $\gamma_1 < \gamma_2$. More precisely, they show that, for a sufficiently large N,

$$\widehat{\gamma}_1^{(BMN)} = \gamma_1 + k^{-1/2}\Lambda(\mathbf{W}) + \frac{\mathbf{A}_0(n/k)}{1-\rho_1}\left(1 + o_{\mathbf{P}}(1)\right), \tag{4.6}$$

where $\mathbf{A}_0(t) := \mathbf{A}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}(\mathbb{U}_{F^*}(t))\right), t > 1$, and $\Lambda(\mathbf{W})$ is a centred Gaussian rv defined by

$$\Lambda(\mathbf{W}) := \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 (\gamma_2 - \gamma_1 - \gamma \log s) s^{-\gamma/\gamma_2 - 1} \mathbf{W}(s) ds - \gamma \mathbf{W}(1),$$

with $\{\mathbf{W}(s); s \geq 0\}$ being a standard Wiener process defined on the probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Thereby, for a given $n = m$, such that $\sqrt{k_m}\mathbf{A}_0(m/k_m) \to \lambda < \infty$, they conclude that

$$\sqrt{k}\left(\widehat{\gamma}_1^{(BMN)} - \gamma_1\right) \overset{\mathcal{D}}{\to} \mathcal{N}\left(\lambda/(1-\rho_1), \sigma^2\right), \text{ as } N \to \infty,$$

where $\sigma^2 := \gamma^2(1+\gamma_1/\gamma_2)\left(1+(\gamma_1/\gamma_2)^2\right)/(1-\gamma_1/\gamma_2)^3$. Recently, Benchaira et al. 2016b, [7] adopted the same approach to introduce a kernel estimator to the tail index γ_1 which improves the bias of $\widehat{\gamma}_1^{(BMN)}$. It is worth mentioning that the assumption $\gamma_1 < \gamma_2$ is required in order to ensure that it remains enough extreme data for the inference to be accurate. In other words, they consider the situation where the tail of the rv of interest is not too contaminated by the truncation rv.

4.2 Estimating the second-order parameter ρ_1

The aim of this paper is the estimation of the second order-parameter ρ_1 given in condition (4.2) which, to our knowledge, is not addressed yet in the extreme value literature. Inspired by the paper of Gomes et al. 2003, [46], we propose an estimator for ρ_1 adapted to the random right-truncation case. To this end, for $\alpha > 0$ and $t > 0$, we introduce the following tail functionals

$$M^{(\alpha)}(t; \mathbf{F}) := \frac{1}{\overline{\mathbf{F}}(\mathbb{U}_{F^*}(t))} \int_{\mathbb{U}_{F^*}(t)}^{\infty} \log^{\alpha}(x/\mathbb{U}_{F^*}(t)) d\mathbf{F}(x), \tag{4.7}$$

$$Q^{(\alpha)}(t; \mathbf{F}) := \frac{M^{(\alpha)}(t; \mathbf{F}) - \Gamma(\alpha+1)\left(M^{(1)}(t; \mathbf{F})\right)^{\alpha}}{M^{(2)}(t; \mathbf{F}) - 2\left(M^{(1)}(t; \mathbf{F})\right)^2}, \tag{4.8}$$

and

$$S^{(\alpha)}(t;\mathbf{F}) := \delta(\alpha) \frac{Q^{(2\alpha)}(t;\mathbf{F})}{\left(Q^{(\alpha+1)}(t;\mathbf{F})\right)^2}, \tag{4.9}$$

where $\log^\alpha x := (\log x)^\alpha$ and $\delta(\alpha) := \alpha(\alpha+1)^2 \Gamma^2(\alpha) / (4\Gamma(2\alpha))$, with $\Gamma(\cdot)$ standing for the usual Gamma function. The following Lemma is instrumental for our needs.

Lemma 4.1 *Assume that the second-order regular variation condition* (4.2) *holds, then for any* $\alpha > 0$

$$(i) \quad \frac{M^{(\alpha)}(t;\mathbf{F}) - \mu_\alpha^{(1)} \left(M^{(1)}(t;\mathbf{F})\right)^\alpha}{\left(M^{(1)}(t;\mathbf{F})\right)^{\alpha-1} \mathbf{A}_0(t)} \to \alpha \left(\mu_\alpha^{(2)}(\rho_1) - \mu_\alpha^{(1)} \mu_1^{(2)}(\rho_1)\right),$$

$$(ii) \quad Q^{(\alpha)}(t;\mathbf{F}) \to q_\alpha(\rho_1) \ \text{and} \ (ii) \quad S^{(\alpha)}(t;\mathbf{F}) \to s_\alpha(\rho_1), \ \text{as } t \to \infty,$$

where

$$q_\alpha(\rho_1) := \frac{\gamma_1^{\alpha-2} \mu_\alpha^{(1)} \left(1 - (1-\rho_1)^\alpha - \alpha\rho_1 (1-\rho_1)^{\alpha-1}\right)}{2\rho_1^2 (1-\rho_1)^{\alpha-2}} \tag{4.10}$$

and

$$s_\alpha(\rho_1) := \frac{\rho_1^2 \left(1 - (1-\rho_1)^{2\alpha} - 2\alpha\rho_1 (1-\rho_1)^{2\alpha-1}\right)}{\left(1 - (1-\rho_1)^{\alpha+1} - (\alpha+1)\rho_1 (1-\rho_1)^\alpha\right)^2}, \tag{4.11}$$

with

$$\mu_\alpha^{(1)} := \Gamma(\alpha+1), \ \mu_\alpha^{(2)}(\rho_1) := \frac{\Gamma(\alpha)\left(1 - (1-\rho_1)^\alpha\right)}{\rho_1 (1-\rho_1)^\alpha}. \tag{4.12}$$

Proof. See the Appendix. ∎

To summarize, from Lemma 4.1, for any $\alpha > 0$,

$$M^{(\alpha)}(t;\mathbf{F}) \to \gamma_1^\alpha \Gamma(\alpha+1), \ Q^{(\alpha)}(t;\mathbf{F}) \to q_\alpha(\rho_1) \ \text{and} \ S^{(\alpha)}(t;\mathbf{F}) \to s_\alpha(\rho_1), \tag{4.13}$$

as $t \to \infty$. The three results (4.13) allow us to construct an estimator for the second-order parameter ρ_1. Indeed, by recalling that $n = n_N$ is a random sequence of integers, let $\upsilon = \upsilon_n$ be a subsequence of n, different than k, such that given $n = m$, $\upsilon_m \to \infty$, $\upsilon_m / m \to 0$ as $N \to \infty$. The sequence υ has to be chosen so that $\sqrt{\upsilon_m} \left|\mathbf{A}_0(m/\upsilon_m)\right| \to \infty$, which is a necessary condition to ensure the consistency of ρ_1 estimator. On the other hand, as already pointed out, the asymptotic normality of $\widehat{\gamma}_1^{(BMN)}$ requires that, for a given $n = m$, $\sqrt{k_m}\mathbf{A}_0 (m/k_m) \rightarrow \lambda < \infty$. This means that both sample fractions k and υ have to be distinctly chosen. Since $\overline{F}^*$ is regularly varying at infinity with index $-1/\gamma$, then from Lemma 3.2.1 in de Haan [20] page 69, we infer that, given $n = m$, we have $X_{m-\upsilon_m:m} \rightarrow \infty$ as $N \rightarrow \infty$ almost surely. Then by using the total probability formula, we show that $X_{n-\upsilon:n} \rightarrow \infty$, almost surely too. By letting, in (4.7) , $t = n/\upsilon$ then by replacing $\mathbb{U}_{F^*} (n/\upsilon)$ by $X_{n-\upsilon:n}$ and $\mathbf{F}$ by the product-limit estimator $\mathbf{F}_n$, we get an estimator $M_n^{(\alpha)} (\upsilon) = M^{(\alpha)} (t; \mathbf{F}_n)$ for $M^{(\alpha)} (t; \mathbf{F})$ as follows:

$$M_n^{(\alpha)} (\upsilon) = \frac{1}{\overline{\mathbf{F}}_n (X_{n-\upsilon:n})} \int_{X_{n-\upsilon:n}}^{\infty} \log^{\alpha} (x/X_{n-\upsilon:n})\, d\mathbf{F}_n (x) . \tag{4.14}$$

Next, we give an explicit formula for $M_n^{(\alpha)} (\upsilon)$ in terms of observed sample $X_1, ..., X_n$. Since $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$ are regularly varying with negative indices $-1/\gamma_1$ and $-1/\gamma_2$ respectively, then their right endpoints are infinite and thus they are equal. Hence, from Woodroofe 1985, [88], we may write $\int_x^{\infty} d\mathbf{F} (y) / \mathbf{F} (y) = \int_x^{\infty} dF^* (y) / C (y)$, where $C (z) := \mathbf{P} (X \leq z \leq Y)$ is the theoretical counterpart of $C_n (z)$ given in (3.4). Differentiating the previous two integrals leads to the crucial equation $C (x)\, d\mathbf{F} (x) = \mathbf{F} (x)\, dF^* (x)$, which implies that $C_n (x)\, d\mathbf{F}_n (x) = \mathbf{F}_n (x)\, dF_n^* (x)$, where $F_n^* (x) := n^{-1} \sum_{i=1}^{n} 1 (X_i \leq x)$ is the usual empirical df based on the observed sample $X_1, ..., X_n$. It follows that

$$M_n^{(\alpha)} (\upsilon) = \frac{1}{\overline{\mathbf{F}}_n (X_{n-\upsilon:n})} \int_{X_{n-\upsilon:n}}^{\infty} \frac{\mathbf{F}_n (x)}{C_n (x)} \log^{\alpha} (x/X_{n-\upsilon:n})\, dF_n^* (x) ,$$

which equals

$$M_n^{(\alpha)} (\upsilon) = \frac{1}{n\overline{\mathbf{F}}_n (X_{n-\upsilon:n})} \sum_{i=1}^{\upsilon} \frac{\mathbf{F}_n (X_{n-i+1:n})}{C_n (X_{n-i+1:n})} \log^{\alpha} \frac{X_{n-i+1:n}}{X_{n-\upsilon:n}}.$$

Similarly, we show that $\overline{\mathbf{F}}_n(X_{n-v:n}) = n^{-1}\sum_{i=1}^{n} \mathbf{F}_n(X_{n-i+1:n})/C_n(X_{n-i+1:n})$. This leads to following form of $M^{(\alpha)}(t;\mathbf{F})$ estimator:

$$M_n^{(\alpha)}(v) := \sum_{i=1}^{v} a_n^{(i)} \log^{\alpha} \frac{X_{n-i+1:n}}{X_{n-v:n}},$$

where $a_n^{(i)}$ is as in (4.5). It is readily observable that $M_n^{(1)}(k) = \widehat{\gamma}_1^{(BMN)}$. Making use of (4.9) with the expression above, we get an estimator of $S^{(\alpha)}(t;\mathbf{F})$, that we denote $S_n^{(\alpha)} = S_n^{(\alpha)}(v)$. This, in virtue of the third limit in (4.13), leads to estimating $s_\alpha(\rho_1)$. It is noteworthy that the function $\rho_1 \to s_\alpha(\rho_1)$, defined and continuous on the set of negative real numbers, is increasing for $0 < \alpha < 1/2$ and decreasing for $\alpha > 1/2$, $\alpha \neq 1$. Then, for suitable values of α, we may invert s_α to get an estimator $\widehat{\rho}_1^{(\alpha)}$ for ρ_1 as follows:

$$\widehat{\rho}_1^{(\alpha)} := s_\alpha^{\leftarrow}\left(S_n^{(\alpha)}\right), \text{ provided that } S_n^{(\alpha)} \in \mathcal{A}_\alpha, \tag{4.15}$$

where $\mathcal{A}_\alpha$ is one of the following two regions:

$$\left\{s : (2\alpha-1)/\alpha^2 < s \leq 4(2\alpha-1)/(\alpha(\alpha+1))^2, \text{ for } \alpha \in (0,1/2)\right\},$$

or

$$\left\{s : 4(2\alpha-1)/(\alpha(\alpha+1))^2 \leq s < (2\alpha-1)/\alpha^2, \text{ for } \alpha \in (1/2,\infty)\setminus\{1\}\right\}.$$

For more details, regarding the construction of these two sets, one refers to Remark 2.1 and Lemma 3.1 in [46]. It is worth mentioning that, for $\alpha = 2$, we have

$$s_2(\rho_1) = \left(3\rho_1^2 - 8\rho_1 + 6\right)/(3-2\rho_1)^2$$

and $s_2^{\leftarrow}(s) = \left(6s - 4 + \sqrt{3s-2}\right)/(4s-3)$, for $2/3 < s < 3/4$. Thereby, we obtain an explicit formula to the estimator of ρ_1 as follows

$$\widehat{\rho}_1^{(2)} = \frac{6S_n^{(2)} - 4 + \sqrt{3S_n^{(2)} - 2}}{4S_n^{(2)} - 3}, \text{ provided that } 2/3 < S_n^{(2)} < 3/4. \tag{4.16}$$

4.3 Reduced-bias tail index estimator

Next, we derive an asymptotically unbiased estimator for γ_1, that improves $\widehat{\gamma}_1^{(BMN)}$ by estimating the asymptotic bias $\mathbf{A}_0(n/k)/(1-\rho_1)$, given in weak approximation (4.6). Indeed, let v be equal to $u_n := [n^{1-\epsilon}]$, for a fixed $\epsilon > 0$ close to zero (say $\epsilon = 0.01$) so that, given $n = m$, $u_m \to \infty$, $u_m/m \to \infty$ and $\sqrt{u_m}\,|\mathbf{A}_0(m/u_m)| \to \infty$. The validity of such a sequence is discussed in [46] (Subsection 6.1, conclusions 2 and 5). The estimator of ρ_1 pertaining to this choice of v will be denoted by $\widehat{\rho}_1^{(*)}$. Let us now define an estimator for $\mathbf{A}_0(n/k)$. From assertion (i) in Lemma 4.1, taking $\alpha = 2$, we have

$$\mathbf{A}_0(t) \sim (1-\rho_1)^2 \left(M^{(2)}(t;\mathbf{F}) - 2\left(M^{(1)}(t;\mathbf{F})\right)^2\right)/\left(2\rho_1 M^{(1)}(t;\mathbf{F})\right), \text{ as } t \to \infty.$$

Then, by letting $t = n/k$ and by replacing, in the previous quantity, $\mathbb{U}_{F^*}(n/k)$ by $X_{n-k:n}$, $\mathbf{F}$ by $\mathbf{F}_n$ and ρ_1 by $\widehat{\rho}_1^{(*)}$, we end up with

$$\widehat{\mathbf{A}}_0(n/k) := \left(1 - \widehat{\rho}_1^{(*)}\right)^2 \left(M_n^{(2)}(k) - 2\left(M_n^{(1)}(k)\right)^2\right)/\left(2\widehat{\rho}_1^{(*)} M_n^{(1)}(k)\right),$$

as an estimator for $\mathbf{A}_0(n/k)$. Thus, we obtain an asymptotically unbiased estimator

$$\widehat{\gamma}_1 := M_n^{(1)}(k) + \frac{M_n^{(2)}(k) - 2\left(M_n^{(1)}(k)\right)^2}{2M_n^{(1)}(k)} \left(1 - \frac{1}{\widehat{\rho}_1^{(*)}}\right),$$

for the tail index γ_1, as an adaptation of Peng's estimator [73] to the random right-truncation case. The rest of the paper is organized as follows. In Section 4.2, we present our main results which consist in the consistency and the asymptotic normality of the estimators $\widehat{\rho}_1^{(\alpha)}$ and $\widehat{\gamma}_1$ whose finite sample behaviors are checked by simulation in Section 4.3. All proofs are gathered in Section 4.4. The proofs of two instrumental lemmas are postponed to the Appendix.

4.4 Main results

It is well known that, weak approximations of the second-order parameter estimators are achieved in the third-order condition of regular variation framework, [37]. Thus, it seems quite natural to suppose that df $\mathbf{F}$ satisfies

$$\lim_{t\to\infty}\left\{\frac{\mathbb{U}_{\mathbf{F}}(tx)/\mathbb{U}_{\mathbf{F}}(t)-x^{\gamma_1}}{\mathbf{A}_{\mathbf{F}}(t)}-x^{\gamma_1}\frac{x^{\rho_1}-1}{\rho_1}\right\}/\mathbf{B}_{\mathbf{F}}(t)=\frac{x^{\gamma_1}}{\rho_1}\left(\frac{x^{\rho_1+\beta_1}-1}{\rho_1+\beta_1}-\frac{x^{\rho_1}-1}{\rho_1}\right), \tag{4.17}$$

where $\beta_1<0$ is the third-order parameter and $\mathbf{B}_{\mathbf{F}}$ is a function tending to zero and not changing sign near infinity with regularly varying absolute value at infinity with index β_1. For convenience, we set $\mathbf{B}_0(t):=\mathbf{B}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}\left(\mathbb{U}_{F^*}(t)\right)\right)$ and by keeping similar notations to those used in [46], we write

$$\mu_\alpha^{(3)}(\rho_1):=\begin{cases}\dfrac{1}{\rho_1^2}\log\dfrac{(1-\rho_1)^2}{1-2\rho_1}, & \text{if } \alpha=1,\\ \dfrac{\Gamma(\alpha)}{\rho_1^2(\alpha-1)}\left\{\dfrac{1}{(1-2\rho_1)^{\alpha-1}}-\dfrac{2}{(1-\rho_1)^{\alpha-1}}+1\right\}, & \text{if } \alpha\neq 1,\end{cases}$$

$$\mu_\alpha^{(4)}(\rho_1,\beta_1):=\beta_1^{-1}\left(\mu_\alpha^{(2)}(\rho_1+\beta_1)-\mu_\alpha^{(2)}(\rho_1)\right),$$

$$m_\alpha:=\mu_\alpha^{(2)}(\rho_1)-\mu_\alpha^{(1)}\mu_1^{(2)}(\rho_1),\ c_\alpha:=\mu_\alpha^{(3)}(\rho_1)-\mu_\alpha^{(1)}\left(\mu_1^{(2)}(\rho_1)\right)^2$$

and $d_\alpha:=\mu_\alpha^{(4)}(\rho_1,\beta_1)-\mu_\alpha^{(1)}\mu_1^{(4)}(\rho_1,\beta_1)$. For further use, we set $r_\alpha:=2q_\alpha\gamma_1^{2-\alpha}/\Gamma(\alpha+1)$, where q_α defined (4.10), and let

$$\eta_1:=\frac{1}{2\gamma_1 m_2 r_{\alpha+1}^2}\left\{\frac{(2\alpha-1)c_{2\alpha}}{\Gamma(2\alpha)}+c_2 r_{2\alpha}-\frac{2c_{\alpha+1}r_{2\alpha}}{r_{\alpha+1}\Gamma(\alpha)}\right\},$$

$$\eta_2 := \frac{1}{\gamma_1 m_2 r_{\alpha+1}^2} \left\{ \frac{d_{2\alpha}}{\Gamma(2\alpha)} + d_2 r_{2\alpha} - \frac{2d_{\alpha+1} r_{2\alpha}}{r_{\alpha+1}\Gamma(\alpha+1)} \right\},$$

$$\xi := \gamma \left(\frac{1-2\alpha-3r_{2\alpha}}{r_{\alpha+1}^2 m_2} + \frac{2\alpha r_{2\alpha}}{r_{\alpha+1}^3 m_2} \right),$$

$$\tau_1 := \frac{1}{\gamma_1^{2\alpha-1} r_{\alpha+1}^2 \Gamma(2\alpha+1) m_2}, \ \tau_2 := -\frac{2r_{2\alpha}}{\gamma_1^{\alpha} r_{\alpha+1}^3 \Gamma(\alpha+2) m_2},$$

$$\tau_3 := \frac{r_{2\alpha}}{\gamma_1 r_{\alpha+1}^2 2m_2}, \ \tau_4 := \frac{-2\alpha r_{\alpha+1} + 2(\alpha+1) r_{2\alpha} - 4r_{\alpha+1} r_{2\alpha}}{r_{\alpha+1}^3 m_2},$$

$$\tau_5 := \frac{\rho_1 - 1}{2\gamma_1\rho_1}, \ \tau_6 := 1 + 2\frac{1-\rho_1}{\gamma_1\rho_1} \text{ and } \mu := \gamma\left(2 + 2\frac{1-\rho_1}{\gamma_1\rho_1} - \frac{1}{\rho_1}\right).$$

Theorem 4.1 *Assume that both df's* $\mathbf{F}$ *and* $\mathbf{G}$ *satisfy the second-order conditions* (4.2) *and* (4.3) *respectively with* $\gamma_1 < \gamma_2$. *Let* $\alpha \in \mathcal{A}_\alpha$ *be fixed and let* υ *be a random sequence of integers such that, given* $n = m$, $\upsilon = \upsilon_m \to \infty$, $\upsilon/m \to 0$ *and* $\sqrt{\upsilon}\left|\mathbf{A}_0(m/\upsilon)\right| \to \infty$, *then*

$$\widehat{\rho}_1^{(\alpha)} \xrightarrow{\mathbf{P}} \rho_1, \ \text{as } N \to \infty.$$

If in addition, we assume that the third-order condition (4.17) *holds, then whenever, given* $n = m$, $\sqrt{\upsilon}\mathbf{A}_0^2(m/\upsilon)$ *and* $\sqrt{\upsilon}\mathbf{A}_0(m/\upsilon)\mathbf{B}_0(m/\upsilon)$ *are asymptotically bounded, then there exists a standard Wiener process* $\{\mathbf{W}(s);\ s \geq 0\}$, *defined on the probability space* $(\Omega, \mathcal{A}, \mathbf{P})$, *such that*

$$s'_\alpha(\rho_1)\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\left(\widehat{\rho}_1^{(\alpha)} - \rho_1\right) = \int_0^1 s^{-\gamma/\gamma_2-1}\Delta_\alpha(s)\mathbf{W}(s)\,ds - \xi\mathbf{W}(1)$$
$$+ \eta_1\sqrt{\upsilon}\mathbf{A}_0^2(n/\upsilon) + \eta_2\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\mathbf{B}_0(n/\upsilon) + o_{\mathbf{P}}(1),$$

where s'_α *is the Lebesgue derivative of* s_α *given in* (4.11) *and*

$$\begin{aligned}\Delta_\alpha(s) \;:=\;& \frac{\tau_1\gamma\log^{2\alpha}s^{-\gamma}}{\gamma_1+\gamma_2}+\frac{2\alpha\tau_1\gamma^2\log^{2\alpha-1}s^{-\gamma}}{\gamma_1}+\frac{\tau_2\gamma\log^{\alpha+1}s^{-\gamma}}{\gamma_1+\gamma_2}\\ &+\frac{\tau_2\left(\alpha+1\right)\gamma^2\log^{\alpha}s^{-\gamma}}{\gamma_1}+\frac{\tau_3\gamma\log^2 s^{-\gamma}}{\gamma_1+\gamma_2}\\ &+\left(\frac{2\tau_3\gamma^2}{\gamma_1}+\frac{\tau_4\gamma}{\gamma_1+\gamma_2}\right)\log s^{-\gamma}+\frac{\tau_4\gamma^2}{\gamma_1}-\frac{\gamma_1\xi}{\gamma_1+\gamma_2}.\end{aligned}$$

If, in addition, we suppose that given $n=m$,

$$\sqrt{\upsilon}\mathbf{A}_0^2\left(m/\upsilon\right)\rightarrow\lambda_1<\infty \text{ and } \sqrt{\upsilon}\mathbf{A}_0\left(m/\upsilon\right)\mathbf{B}_0\left(m/\upsilon\right)\rightarrow\lambda_2<\infty,$$

then $\sqrt{\upsilon}\mathbf{A}_0\left(n/\upsilon\right)\left(\widehat{\rho}_1^{(\alpha)}-\rho_1\right)\overset{\mathcal{D}}{\rightarrow}\mathcal{N}\left(\eta_1\lambda_1+\eta_2\lambda_2,\sigma_\alpha^2\right),$ *as* $N\rightarrow\infty$, *where*

$$\sigma_\alpha^2:=\int_0^1\int_0^1 s^{-\gamma/\gamma_2-1}t^{-\gamma/\gamma_2-1}\min\left(s,t\right)\Delta_\alpha(s)\Delta_\alpha(t)dsdt-2\xi\int_0^1 s^{-\gamma/\gamma_2}\Delta_\alpha(s)ds+\xi^2.$$

Theorem 4.2 *Let* k *be a random sequence of integers, different from* υ, *such that, given* $n=m$, $k=k_m\rightarrow\infty$, $k/m\rightarrow0$ *and* $\sqrt{k}\mathbf{A}_0\left(m/k\right)$ *is asymptotically bounded, then with the same Wiener process* $\left\{\mathbf{W}\left(s\right);\ s\geq0\right\}$ *as in Theorem 4.1, we have*

$$\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)=\int_0^1 s^{-\gamma/\gamma_2-1}\mathbf{D}(s)\mathbf{W}\left(s\right)ds-\mu\mathbf{W}\left(1\right)+o_{\mathbf{P}}\left(1\right),$$

for any $\epsilon>0$, *where*

$$\mathbf{D}(s):=\frac{\gamma^3\tau_5}{\gamma_1+\gamma_2}\log^2 s-\left(\frac{2\tau_5\gamma^3}{\gamma_1}+\frac{\gamma^2\tau_6}{\gamma_1+\gamma_2}\right)\log s+\frac{\tau_6\gamma^2}{\gamma_1}-\frac{\gamma_1\mu}{\gamma_1+\gamma_2}.$$

If, in addition, we suppose that, given $n=m$, $\sqrt{k}\mathbf{A}_0\left(m/k\right)\rightarrow\lambda<\infty$, *then*

$$\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)\overset{\mathcal{D}}{\rightarrow}\mathcal{N}\left(0,\sigma_*^2\right),\ \text{as } N\rightarrow\infty,$$

where $\sigma_*^2 := \int_0^1 \int_0^1 s^{-\gamma/\gamma_2 - 1} t^{-\gamma/\gamma_2 - 1} \min(s,t) \mathbf{D}(s) \mathbf{D}(t) ds dt - 2\mu \int_0^1 s^{-\gamma/\gamma_2} \mathbf{D}(s) ds + \mu^2.$

4.5 Simulation study

We begin by choosing $\alpha = 2$ to study the performance of $\widehat{\rho}_1^{(2)}$ and compare the newly introduced bias-reduced estimator $\widehat{\gamma}_1$ with $\widehat{\gamma}_1^{(BMN)}$. To this end, let us consider sets of truncated and truncation data drawn from Burr's and Fréchet's models:

- Burr (a,b) distribution with survival function $\overline{\mathbf{F}}(x) = \left(1 + x^b\right)^{-1/(ab)}$, $x > 0$, $a, b > 0$.
- Fréchet (c) distribution with survival function $\overline{\mathbf{F}}(x) = \exp\left(-x^{-1/c}\right)$, $x \geq 0$, $c > 0$.

Both models satisfies the third-order condition (4.17), with:

- $\rho = \beta = -ab$, $\mathbf{A_F}(x) = abx^\rho / (1 - x^\rho)$ and $\mathbf{B_F}(x) = \rho x^\rho / (1 - x^\rho)$.
- $\rho = \beta = -1$, $\mathbf{A_F}(x) = \gamma^{-1} x^{-1}/2$ and $\mathbf{B_F}(x) = 5x^{-1}/6$.

Let us consider the following four scenarios:

- $[S_1]$ Burr (a_1, b_1) truncated by Burr (a_2, b_2)
- $[S_2]$ Fréchet (c_1) truncated by Fréchet (c_2)
- $[S_3]$ Fréchet (c) truncated by Burr (a, b)
- $[S_4]$ Burr (a, b) truncated by Fréchet (c)

Both tail indices corresponding to the scenarios above are, respectively,

$$(\gamma_1, \gamma_2) = (a_1, a_2), (a_1, c), (c, a_2) \text{ and } (c_1, c_2). \tag{4.18}$$

For all cases, we fix $b = 1/4$ and choose the values 0.6 and 0.8 for γ_1 and 70% and 90% for the percentage of observed data $p = \gamma_2/(\gamma_1 + \gamma_2)$. For each couple (γ_1, p), we solve the latter equation to get the pertaining γ_2-value. We vary the common size N of both samples $(\mathbf{X}_1, ..., \mathbf{X}_N)$ and $(\mathbf{Y}_1, ..., \mathbf{Y}_N)$, then for each size, we generate 1000 independent replicates. Next, we study the performance of our estimators by means of two adaptive methods to select the sample fractions, namely the automatic choice and the graphical diagnostics. For the first method we will consider both scenarios $[S_1]$ and $[S_2]$ will for the second one take all the four scenarios $[S_1] - [S_4]$.

4.5.1 Automatic choice of the number of upper extremes

There exists in the literature several heuristic methods for choosing the optimal number of upper extremes used in the computation of the tail index estimate. An exhaustive bibliography to this topic is gathered in the nice survey given by Caeiro and Gomes 2015, [12]. Our choice fell on the method of Reiss and Thomas given in their text book [76], page 137, and also done in software program incorporated in the "Xtremes" package. We will apply this method to select the optimal numbers υ^* and k^* of upper order statistics used in the computation of $\widehat{\rho}_1^{(2)}$, $\widehat{\gamma}_1$ and $\widehat{\gamma}_1^{(BMN)}$. For each estimator, we compute the average of the resulting 1000 observations of bias, as well as its corresponding rmse. The performance of the estimators and their comparison are made with respect to the absolute biases (abias) and rmse's, which are summarized in Tables 4.1, 4.2, 4.3 and 4.4. On the light the results, we see that the estimation quality of three estimators decreases when the truncation percentage increases. On the other hand, the results of Table 4.1, show that $\widehat{\rho}_1^{(2)}$ behaves well in terms of bias and rmse and those of Tables 4.2, 4.3 and 4.4, clarify that the newly proposed estimator $\widehat{\gamma}_1$ performs better than $\widehat{\gamma}_1^{(BMN)}$ both in bias and rmse. It is worth mentioning that, the kernel estimator for the tail index γ_1, introduced by Benchaira et al. 2016b, [7], also performs better than $\widehat{\gamma}_1^{(BMN)}$ from the bias viewpoint but with higher rmse. Furthermore, we observe that the kernel estimator works only when the sample sizes are greater than, approximately, $N = 100$. In other words, this is not recommended when the number n of observed data is less that 100.

	$p = 0.7$				$p = 0.9$			
N	n	υ^*	abias	rmse	n	υ^*	abias	rmse
				$\gamma_1 = 0.6$				
100	73	31	0.011	0.048	88	37	0.007	0.047
200	149	68	0.009	0.045	178	69	0.005	0.046
500	352	211	0.007	0.041	448	244	0.003	0.038
1000	702	672	0.002	0.029	895	643	0.001	0.031
				$\gamma_1 = 0.8$				
100	70	31	0.012	0.055	89	42	0.017	0.049
200	137	65	0.010	0.047	178	75	0.011	0.048
500	351	196	0.006	0.040	450	230	0.004	0.044
1000	729	300	0.002	0.035	901	376	0.002	0.027

Table 4.1: Absolute biases and rmse's of the second-order parameter estimator based on 1000 right-truncated samples from Burr's models.

4.5.2 Graphical diagnostics of the number k of upper extremes

Being $b = 1/4$ and (γ_1, γ_2) are fixed, then in view of (4.18) , we may suppose that the second-order parameter ρ is also fixed and equals $\gamma_1 b$ for both first and second scenarios $[S_1] - [S_2]$ and equals -1 for both third and fourth ones $[S_3] - [S_4]$. Note also that for all cases we take $N = 300$ and since p is fixed then the number of observed sample $n = [pN]$ is known. Thereby, for each scenario and for each $k = 1, ..., n$, we compute the average of 1000 values of biases and rmse's with respect to both tail index estimators $\widehat{\gamma}_1$ and $\widehat{\gamma}_1^{(BMN)}$. The results are reported in Figures 4.1-4.8, where the solid and dashed lines correspond respectively to the reduced-bias estimator and the original one. Likewise we clearly see that, for the all scenarios, $\widehat{\gamma}_1$ performed better $\widehat{\gamma}_1^{(BMN)}$ both in bias and rmse.

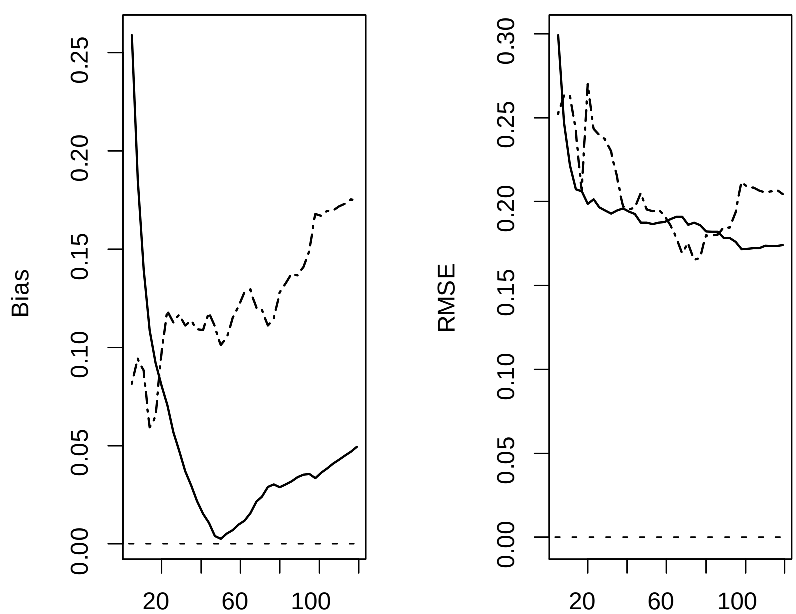

Figure 4.1: S1: $\gamma_1 = 0.6$, $p = 70\%$

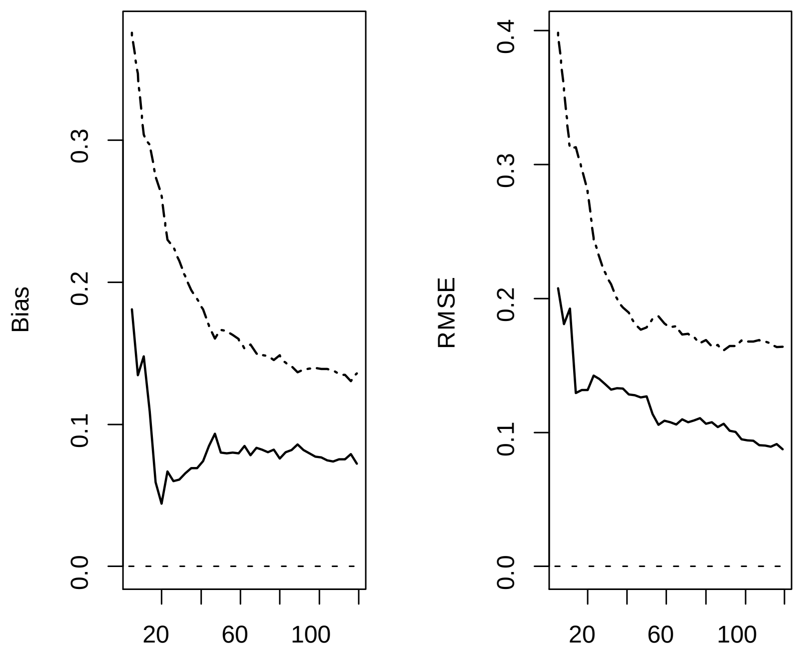

Figure 4.2: S2: $\gamma_1 = 0.6$, $p = 70\%$

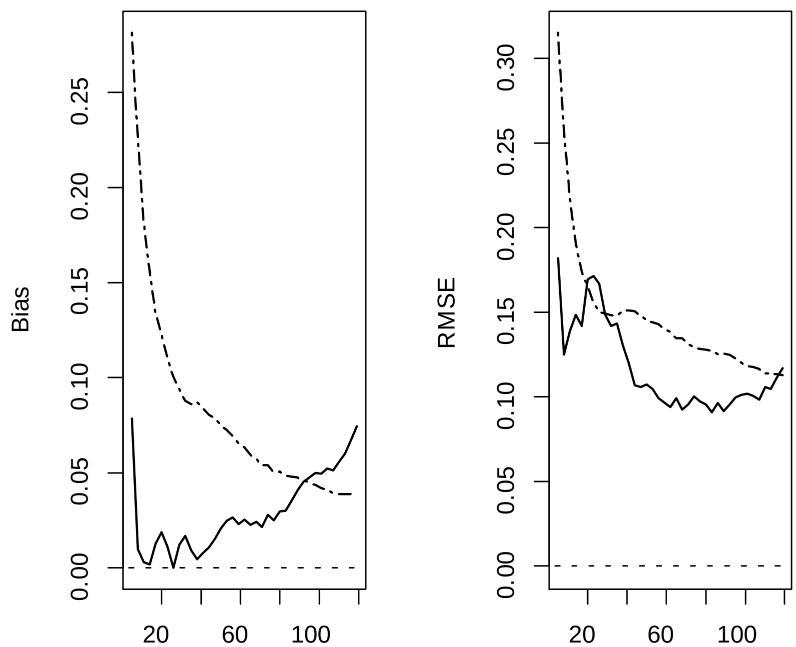

Figure 4.3: S4: $\gamma_1 = 0.6$, $p = 70\%$

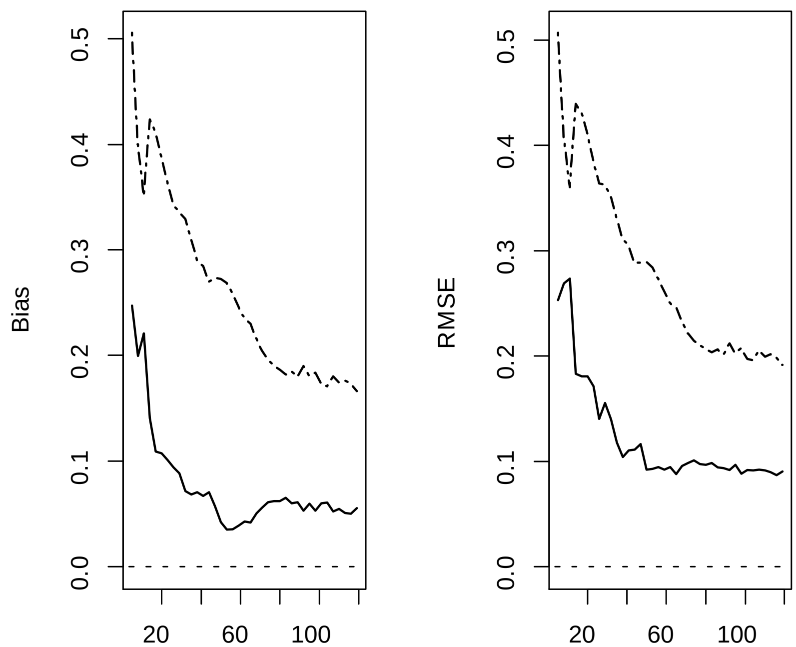

Figure 4.4: S3: $\gamma_1 = 0.6$, $p = 70\%$

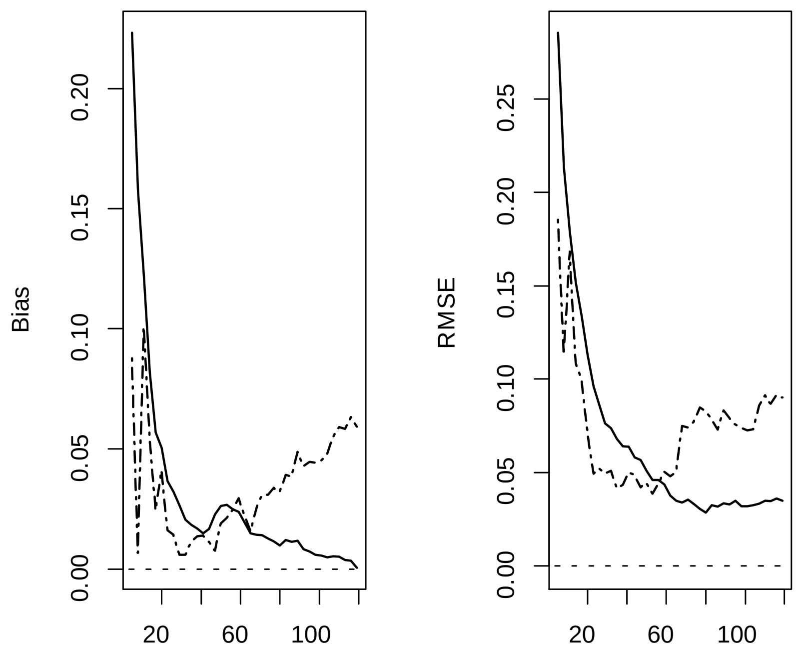

Figure 4.5: S1: $\gamma_1 = 0.6$, $p = 90\%$

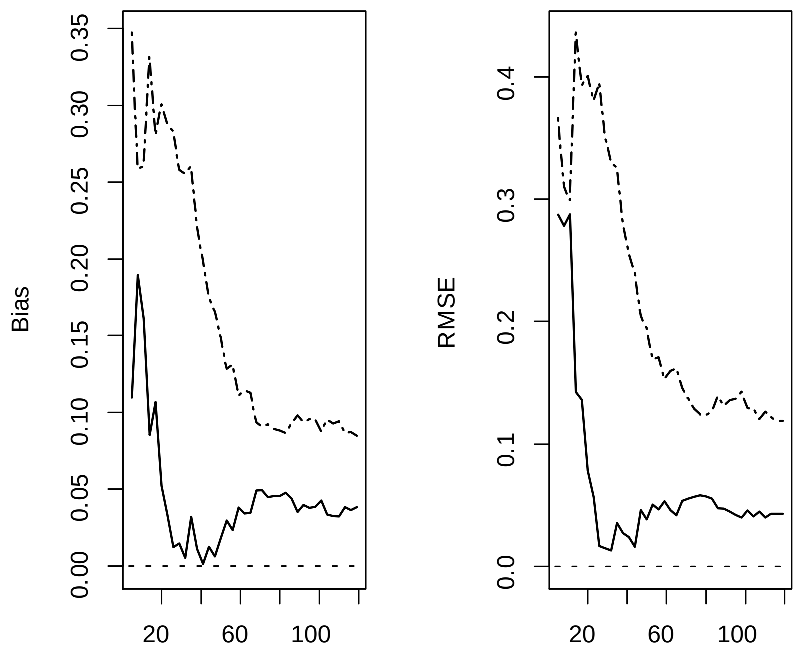

Figure 4.6: S2: $\gamma_1 = 0.6$, $p = 90\%$

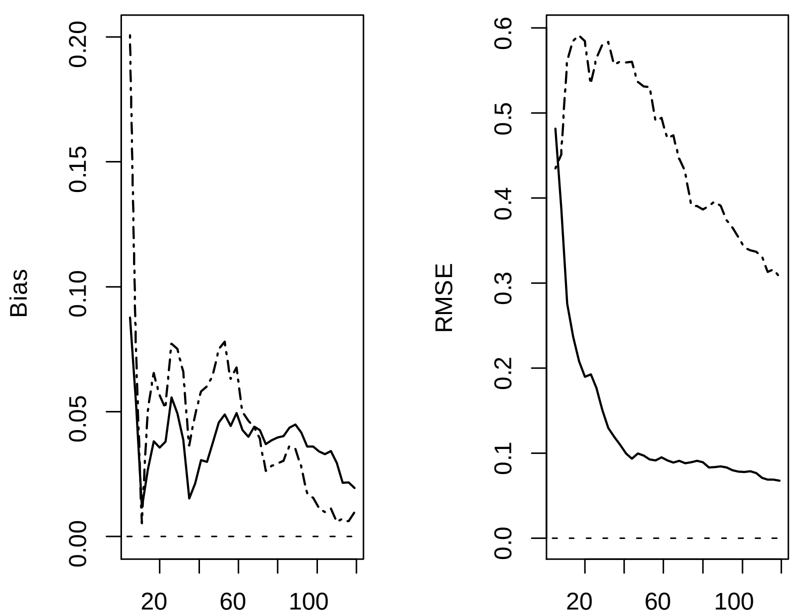

Figure 4.7: S3: $\gamma_1 = 0.6$, $p = 90\%$

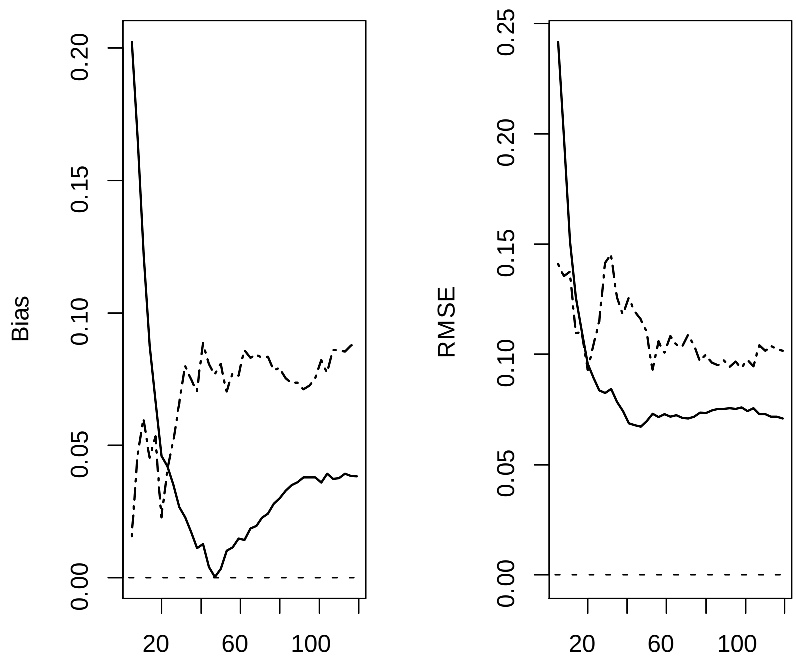

Figure 4.8: S4: $\gamma_1 = 0.6$, $p = 90\%$

	$p = 0.7$							$p = 0.9$						
	$\widehat{\gamma}_1$				$\widehat{\gamma}_1^{(BMN)}$			$\widehat{\gamma}_1$				$\widehat{\gamma}_1^{(BMN)}$		
N	n	k^*	abias	rmse	k^*	abias	rmse	n	k^*	abias	rmse	k^*	abias	rmse
$\gamma_1 = 0.6$														
100	71	11	0.067	0.258	11	0.125	0.259	89	16	0.015	0.155	15	0.119	0.220
200	139	25	0.047	0.202	24	0.092	0.223	180	34	0.010	0.118	31	0.088	0.169
500	352	67	0.026	0.125	58	0.084	0.173	449	83	0.005	0.069	78	0.049	0.132
1000	704	113	0.008	0.089	112	0.015	0.121	898	176	0.003	0.035	174	0.018	0.052
$\gamma_1 = 0.8$														
100	70	12	0.064	0.311	11	0.221	0.218	87	15	0.068	0.221	14	0.195	0.321
200	141	26	0.015	0.219	25	0.164	0.279	172	32	0.034	0.151	30	0.130	0.242
500	348	61	0.012	0.152	60	0.032	0.223	443	88	0.020	0.098	81	0.089	0.157
1000	702	142	0.007	0.054	124	0.020	0.131	894	180	0.011	0.056	168	0.015	0.086

Table 4.2: Absolute biases and rmse's of the tail index estimators based on 1000 right-truncated samples from Burr's models.

4.6 High quantile estimation

Theoretically, the extreme quantile for df $\mathbf{F}$ is the value, in $v \downarrow 0$, of the generalized inverse $q_\nu := \mathbb{U}_{\mathbf{F}}(1/\nu)$; in other words the value, sufficiently large, so that the probability of exceeding q_ν is so small. As being a risk measure, this quantity, known as the value-at-risk (VAR), is used in several fields, such as in finance, insurance, hydrology and reliability. For asymptotic needs, we suppose that v is a function of the observed sample size n, denoted by $v = v_n$, and assumed to be much smaller than $1/n$. The estimation of high quantiles for heavy-tailed distributions, in the case of complete data, is extensively studied in the literature [20]. Benchaira et al. 2016a [6] adapted the well-known Weissman estimator [86] to the random truncation case and proposed the following estimator

$$\widehat{q}_\nu := X_{n-k:n}\left(\frac{\nu}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)}\right)^{-\widehat{\gamma}_1^{(BMN)}}.$$

Otherwise, Gardes and Stupfler 2015, [42] also suggested a similar estimator, but instead of $\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)$ they considered a sequence of deterministic order asymptotically negligible

	$p = 0.7$				$p = 0.9$			
N	n	v^*	abias	rmse	n	v^*	abias	rmse
$\gamma_1 = 0.6$								
100	72	32	0.009	0.039	90	32	0.008	0.038
200	139	66	0.007	0.042	168	67	0.006	0.036
500	349	202	0.005	0.037	439	238	0.005	0.033
1000	698	651	0.003	0.027	867	641	0.002	0.029
$\gamma_1 = 0.8$								
100	70	31	0.012	0.055	89	42	0.017	0.049
200	137	65	0.010	0.047	178	75	0.011	0.048
500	351	196	0.006	0.040	450	230	0.004	0.044
1000	729	300	0.002	0.035	901	376	0.002	0.027

Table 4.3: Absolute biases and rmse's of the second-order parameter estimator based on 1000 right-truncated samples from Fréchet's models.

with respect to ν. As stated in Section 4.5, our new estimator $\widehat{\gamma}_1$ performed better than $\widehat{\gamma}_1^{(BMN)}$ in terms of bias, therefore its corresponding high quantile estimator

$$\widetilde{q}_\nu := X_{n-k:n}\left(\frac{\nu}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)}\right)^{-\widehat{\gamma}_1}$$

will be systematically better than $\widehat{q}_\nu$ as well. It is worth mentioning, from Tables 4.2 and 4.4, that the sample fraction k in the computation of $\widetilde{q}_\nu$ is not necessarily the same for $\widehat{q}_\nu$. The asymptotic normality of $\widetilde{q}_\nu$ is established in the following theorem.

Theorem 4.3 *Assume that both second-order conditions (4.2) and (4.3) hold with $\gamma_1 < \gamma_2$. Then*

$$\frac{\sqrt{k}}{\log d_n}\left(\frac{\widetilde{q}_\nu}{q_\nu}-1\right)\overset{\mathcal{D}}{\to}\mathcal{N}\left(0,\sigma_*^2\right),\ \text{as } N\to\infty,$$

provided that, given $n = m$, $k = k_m \to \infty$, $k/m \to 0$, $d_m \to 0$, $\sqrt{k_m}/\log d_m \to \infty$, $\sqrt{k_m}\mathbf{A}_{\mathbf{F}}^*\left(m/k_m\right) \to \lambda$ and $\sqrt{k_m}\mathbf{A}_{\mathbf{G}}^*\left(m/k_m\right) = O\left(1\right)$, as $N \to \infty$, where $d_n := \overline{\mathbf{F}}\left(\mathbb{U}_F\left(n/k\right)\right)/\nu_n$ and $\mathbf{A}_{\mathbf{G}}^*\left(t\right) := \mathbf{A}_{\mathbf{G}}\left(1/\overline{\mathbf{G}}\left(\mathbb{U}_F\left(t\right)\right)\right)$, $t > 1$.

	$p=0.7$							$p=0.9$						
	$\widehat{\gamma}_1$				$\widehat{\gamma}_1^{(BMN)}$			$\widehat{\gamma}_1$				$\widehat{\gamma}_1^{(BMN)}$		
N	n	k^*	abias	rmse	k^*	abias	rmse	n	k^*	abias	rmse	k^*	abias	rmse
$\gamma_1 = 0.6$														
100	69	13	0.078	0.347	13	0.065	0.278	90	14	0.055	0.275	17	0.049	0.320
200	138	29	0.054	0.325	28	0.040	0.214	178	32	0.039	0.219	31	0.032	0.269
500	362	59	0.035	0.245	58	0.012	0.193	459	79	0.018	0.127	76	0.009	0.129
1000	709	105	0.011	0.149	119	0.010	0.117	887	169	0.012	0.107	172	0.019	0.078
$\gamma_1 = 0.8$														
100	68	15	0.055	0.281	11	0.211	0.219	87	17	0.077	0.231	13	0.184	0.317
200	135	24	0.025	0.247	25	0.158	0.274	172	29	0.046	0.174	31	0.131	0.219
500	347	59	0.022	0.142	60	0.047	0.231	443	87	0.031	0.085	79	0.097	0.149
1000	717	139	0.018	0.042	124	0.030	0.148	894	175	0.018	0.062	172	0.043	0.095

Table 4.4: Absolute biases and rmse's of the tail index estimators based on 1000 right-truncated samples from Fréchet's models.

4.7 Real data example

We consider, as it is done in both Gardes and Stupfler 2015, [42] and Benchaira et al. 2016a, [6], the (left-truncated) lifetimes of car brake pads given in Lawless [65], page 69. We first make a transformation on the dataset to be right-truncated and then verify the Pareto-like nature of its distribution, see Section 5 in [42]. Since the sample size, $n = 98 < 100$, is relatively small then the simulation study results in Section 4.5 suggest that the new estimator $\widehat{\gamma}_1$ is a better candidate to estimate the tail index γ_1. Making use of the algorithm of Reiss and Thomas [76], we select the optimal sample fraction k^* and then compute the corresponding value of $\widehat{\gamma}_1$. The result gives $\widehat{\gamma}_1^* = 0.492$, however Benchaira et al. 2016a, [6] obtained the value 0.470. Thereby, we compute, for three different high levels $\overline{\nu} = 1 - \nu = 0.990$, 0.995 and 0.999 their corresponding extreme quantiles. Finally, via the aforementioned transformation, we obtain the pertaining extreme quantiles of the original dataset. The results are summarized in Table 4.5. Then, we may conclude that the estimated value of the brake pad lifetime is less than $17,063$ km for 1% of the cars. However, only one out of a thousand brake pads lasts less than 10.200 km.

$\overline{\nu}$	Transformed data	Original data
0.990	0.094	17063
0.995	0.130	14138
0.999	0.227	10203

Table 4.5: Extreme quantiles for car brake pad lifetimes.

4.8 Proofs

4.8.1 Proof of Theorem 4.1

We begin by proving the consistency of $\widehat{\rho}_1^{(\alpha)}$ defined in (4.15). We let

$$\mathbf{L}_n(x;\upsilon) := \frac{\overline{\mathbf{F}}_n(X_{n-\upsilon:n}x)}{\overline{\mathbf{F}}_n(X_{n-\upsilon:n})} - \frac{\overline{\mathbf{F}}(X_{n-\upsilon:n}x)}{\overline{\mathbf{F}}(X_{n-\upsilon:n})},$$

and we show that for any $\alpha > 0$

$$M_n^{(\alpha)}(\upsilon) = \gamma_1^\alpha \mu_\alpha^{(1)} + \int_1^\infty \mathbf{L}_n(x;\upsilon)\, d\log^\alpha x + (1 + o_{\mathbf{P}}(1))\, \alpha \gamma_1^{\alpha-1} \mu_\alpha^{(2)}(\rho_1)\, \mathbf{A}_0(n/\upsilon), \quad (4.19)$$

where $\mu_\alpha^{(1)}$ and $\mu_\alpha^{(2)}(\rho_1)$ are as in (4.12). It is clear that from formula (4.14), $M_n^{(\alpha)}(\upsilon)$ may be rewritten into $-\int_1^\infty \log^\alpha x d\overline{\mathbf{F}}_n(X_{n-\upsilon:n}x)/\overline{\mathbf{F}}_n(X_{n-\upsilon:n})$, which by an integration by parts equals $\int_1^\infty \overline{\mathbf{F}}_n(X_{n-\upsilon:n}x)/\overline{\mathbf{F}}_n(X_{n-\upsilon:n})\, d\log^\alpha x$. The latter may be decomposed into

$$\int_1^\infty \mathbf{L}_n(x;\upsilon)\, d\log^\alpha x + \int_1^\infty \left(\frac{\overline{\mathbf{F}}_n(X_{n-\upsilon:n}x)}{\overline{\mathbf{F}}_n(X_{n-\upsilon:n})} - x^{-1/\gamma_1}\right) d\log^\alpha x + \int_1^\infty x^{-1/\gamma_1} d\log^\alpha x.$$

It is easy to verify that $\int_1^\infty x^{-1/\gamma_1} d\log^\alpha x$ equals $\gamma_1^\alpha \mu_\alpha^{(1)}$. Since, $X_{n-\upsilon:n} \to \infty$, almost surely, then by making use of the uniform inequality of the second-order regularly varying functions, to $\overline{\mathbf{F}}$, given in Proposition 4 of Hua [57], we write: with probability one, for any

$0<\epsilon<1$ and large N

$$\left|\frac{\overline{\mathbf{F}}\left(X_{n-\upsilon:n}x\right)/\overline{\mathbf{F}}\left(X_{n-\upsilon:n}\right)-x^{-1/\gamma_1}}{\gamma_1^{-2}\widetilde{\mathbf{A}}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}\left(X_{n-\upsilon:n}\right)\right)}-x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\gamma_1/\rho_1}\right|\leq\epsilon x^{-1/\gamma_1+\epsilon},\ \text{for any } x\geq1, \tag{4.20}$$

where $\widetilde{\mathbf{A}}_{\mathbf{F}}(t)\sim\mathbf{A}_{\mathbf{F}}(t)$, as $t\to\infty$. This implies, almost surely, that

$$\int_1^\infty\left(\frac{\overline{\mathbf{F}}_n\left(X_{n-\upsilon:n}x\right)}{\overline{\mathbf{F}}_n\left(X_{n-\upsilon:n}\right)}-x^{-1/\gamma_1}\right)d\log^\alpha x$$
$$=\widetilde{\mathbf{A}}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}\left(X_{n-\upsilon:n}\right)\right)\left\{\int_1^\infty x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\gamma_1\rho_1}d\log^\alpha x+o_{\mathbf{P}}\left(\int_1^\infty x^{-1/\gamma_1+\epsilon}d\log^\alpha x\right)\right\}.$$

We check that $\int_1^\infty x^{-1/\gamma_1}\dfrac{x^{\rho_1/\gamma_1}-1}{\gamma_1\rho_1}d\log^\alpha x=\alpha\gamma_1^{\alpha-1}\mu_\alpha^{(2)}(\rho_1)$ and $\int_1^\infty x^{-1/\gamma_1+\epsilon}d\log^\alpha x$ is finite. From Lemma 7.4 in [6], $X_{n-\upsilon:n}/\mathbb{U}_{F^*}(n/\upsilon)\xrightarrow{\mathbf{P}}1$, as $N\to\infty$, then by using the regular variation property of $\left|\mathbf{A}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}(\cdot)\right)\right|$ and the corresponding Potter's inequalities (see, for instance, Proposition B.1.10 in [20]), we get

$$\widetilde{\mathbf{A}}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}\left(X_{n-\upsilon:n}\right)\right)=\left(1+o_{\mathbf{P}}(1)\right)\mathbf{A}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}\left(\mathbb{U}_{F^*}(n/\upsilon)\right)\right)=\left(1+o_{\mathbf{P}}(1)\right)\mathbf{A}_0(n/\upsilon),$$

therefore $M_n^{(\alpha)}(\upsilon)=\gamma_1^\alpha\mu_\alpha^{(1)}+\int_1^\infty\mathbf{L}_n(x;\upsilon)\,d\log^\alpha x+\alpha\gamma_1^{\alpha-1}\mu_\alpha^{(2)}(\rho_1)\mathbf{A}_0(n/\upsilon)\left(1+o_{\mathbf{P}}(1)\right)$. In the second step, we use the Gaussian approximation of $\mathbf{L}_n(x)$ recently given by [6] (assertion (6.26)), saying that: for any $0<\epsilon<1/2-\gamma/\gamma_2$, there exists a standard Wiener process $\{\mathbf{W}(s);\ s\geq0\}$, defined on the probability space $(\Omega,\mathcal{A},\mathbf{P})$ such that

$$\sup_{x\geq1}x^{(1/2-\epsilon)/\gamma-1/\gamma_2}\left|\sqrt{\upsilon}\mathbf{L}_n(x;\upsilon)-\mathcal{L}(x;\mathbf{W})\right|\xrightarrow{\mathbf{P}}0,\ \text{as } N\to\infty, \tag{4.21}$$

where $\{\mathcal{L}(x;\mathbf{W});\ x>0\}$ is a Gaussian process defined by

$$\frac{\gamma}{\gamma_1}x^{-1/\gamma_1}\left\{x^{1/\gamma}\mathbf{W}\left(x^{-1/\gamma}\right)-\mathbf{W}(1)\right\}$$
$$+\frac{\gamma}{\gamma_1+\gamma_2}x^{-1/\gamma_1}\int_0^1 s^{-\gamma/\gamma_2-1}\left\{x^{1/\gamma}\mathbf{W}\left(x^{-1/\gamma}s\right)-\mathbf{W}(s)\right\}ds.$$

Let us decompose $\sqrt{\upsilon}\int_1^\infty \mathbf{L}_n(x;\upsilon)\,d\log^\alpha x$ into

$$\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\log^\alpha x + \int_1^\infty \left\{\sqrt{\upsilon}\mathbf{L}_n(x;\upsilon) - \mathcal{L}(x;\mathbf{W})\right\}d\log^\alpha x.$$

By using approximation (4.21), we obtain $\int_1^\infty \{\sqrt{\upsilon}\mathbf{L}_n(x;\upsilon) - \mathcal{L}(x;\mathbf{W})\}\,d\log^\alpha x = o_{\mathbf{P}}(1)$. We showed in Lemma 4.2 that $\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\log^\alpha x = O_{\mathbf{P}}(1)$, therefore $\int_1^\infty \mathbf{L}_n(x;\upsilon)\,d\log^\alpha x = O_{\mathbf{P}}\left(\upsilon^{-1/2}\right)$, it follows that

$$\begin{aligned} M_n^{(\alpha)}(\upsilon) = \gamma_1^\alpha \mu_\alpha^{(1)} + \upsilon^{-1/2}\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\log^\alpha x \\ + \alpha\gamma_1^{\alpha-1}\mu_\alpha^{(2)}(\rho_1)\,\mathbf{A}_0(n/\upsilon)(1+o_{\mathbf{P}}(1)) + o_{\mathbf{P}}\left(\upsilon^{-1/2}\right). \end{aligned} \tag{4.22}$$

Once again, by using the fact that $\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\log^\alpha x = O_{\mathbf{P}}(1)$, we get

$$M_n^{(\alpha)}(\upsilon) = \gamma_1^\alpha \mu_\alpha^{(1)} + \alpha\gamma_1^{\alpha-1}\mu_\alpha^{(2)}(\rho_1)\,\mathbf{A}_0(n/\upsilon)(1+o_{\mathbf{P}}(1)) + o_{\mathbf{P}}\left(\upsilon^{-1/2}\right).$$

It particular, for $\alpha = 1$, we have $\mu_1^{(1)} = 1$, this means that

$$M_n^{(1)}(\upsilon) = \gamma_1 + \mu_1^{(2)}(\rho_1)\,\mathbf{A}_0(n/\upsilon)(1+o_{\mathbf{P}}(1)) + o_{\mathbf{P}}\left(\upsilon^{-1/2}\right),$$

which implies that

$$\left(M_n^{(1)}(\upsilon)\right)^2 = \gamma_1^2 + 2\gamma_1\mu_1^{(2)}(\rho_1)\,\mathbf{A}_0(n/\upsilon)(1+o_{\mathbf{P}}(1)) + o_{\mathbf{P}}\left(\upsilon^{-1/2}\right). \tag{4.23}$$

Likewise, for $\alpha = 2$, we have $\mu_2^{(1)} = 2$, then

$$M_n^{(2)}(\upsilon) = 2\gamma_1^2 + 2\gamma_1\mu_2^{(2)}(\rho_1)\,\mathbf{A}_0(n/\upsilon)(1+o_{\mathbf{P}}(1)) + o_{\mathbf{P}}\left(\upsilon^{-1/2}\right). \tag{4.24}$$

Similar to the definition of $M_n^{(\alpha)}(\upsilon)$, let $Q_n^{(\alpha)}(\upsilon)$ be $Q^{\alpha}(t;\mathbf{F})$ with $\mathbb{U}_{F^*}(t)$ and $\mathbf{F}$ respectively replaced by $X_{n-\upsilon:n}$ and $\mathbf{F}_n$. From (4.8), we may write

$$Q_n^{(\alpha)}(\upsilon)=\frac{M_n^{(\alpha)}(\upsilon)-\Gamma(\alpha+1)\left(M_n^{(1)}(\upsilon)\right)^{\alpha}}{M_n^{(2)}(\upsilon)-2\left(M_n^{(1)}(\upsilon)\right)^{2}}.$$

Then, by using the approximations above, we end up with

$$Q_n^{(\alpha)}(\upsilon)=(1+o_{\mathbf{P}}(1))\frac{\alpha\gamma_1^{\alpha-1}\left(\mu_\alpha^{(2)}(\rho_1)-\mu_\alpha^{(1)}\mu_1^{(2)}(\rho_1)\right)}{2\gamma_1\left(\mu_2^{(2)}(\rho_1)-\mu_2^{(1)}\mu_1^{(2)}(\rho_1)\right)}.$$

By replacing $\mu_\alpha^{(1)}$, $\mu_1^{(1)}$, $\mu_\alpha^{(2)}(\rho_1)$ and $\mu_1^{(2)}(\rho_1)$ by their corresponding expressions, given in (4.12), with the fact that $\alpha\Gamma(\alpha)=\Gamma(\alpha+1)$, we show that the previous quotient equals $q_\alpha(\rho_1)$ given in (4.10). This implies that $Q_n^{(\alpha)}(\upsilon)\overset{\mathbf{P}}{\to}q_\alpha(\rho_1)$ and therefore $S_n^{(\alpha)}(\upsilon)\overset{\mathbf{P}}{\to}s_\alpha(\rho_1)$, as $N\to\infty$, as well. By using the mean value theorem, we infer that $\widehat{\rho}_1^{(\alpha)}=s_\alpha^{\leftarrow}\left(S_n^{(\alpha)}(\upsilon)\right)\overset{\mathbf{P}}{\to}\rho_1$, as sought. Let us now focus on the asymptotic representation of $\widehat{\rho}_1^{(\alpha)}$. We begin by denoting $\widetilde{M}_n^{(\alpha)}(\upsilon)$, $\widetilde{S}_n^{(\alpha)}(\upsilon)$ and $\widetilde{Q}_n^{(\alpha)}(\upsilon)$ the respective values of $M^{(\alpha)}(t;\mathbf{F})$, $S^{(\alpha)}(t;\mathbf{F})$ and $Q^{(\alpha)}(t;\mathbf{F})$ when replacing $\mathbb{U}_{F^*}(t)$ by $X_{n-\upsilon:n}$. It is clear that the quantity $S_n^{(\alpha)}(\upsilon)-s_\alpha(\rho_1)$ may be decomposed into the sum of

$$T_{n1}:=-\delta(\alpha)\frac{\left(Q_n^{(\alpha+1)}(\upsilon)\right)^2-\left(\widetilde{Q}_n^{(\alpha+1)}(\upsilon)\right)^2}{\left(Q_n^{(\alpha+1)}(\upsilon)\,\widetilde{Q}_n^{(\alpha+1)}(\upsilon)\right)^2}Q_n^{(2\alpha)}(\upsilon;\mathbf{F}_n),$$

$$T_{n2}:=\delta(\alpha)\frac{Q_n^{(2\alpha)}(\upsilon)-\widetilde{Q}_n^{(2\alpha)}(\upsilon)}{\left(\widetilde{Q}_n^{(\alpha+1)}(\upsilon)\right)^2}\text{ and }T_{n3}:=\widetilde{S}_n^{(\alpha)}(\upsilon)-s_\alpha(\rho_1).$$

Since $Q_n^{(\alpha)}(\upsilon)\overset{\mathbf{P}}{\to}q_\alpha(\rho_1)$, then by using the mean value theorem, we get

$$T_{n1}=-(1+o_{\mathbf{p}}(1))\,2\delta(\alpha)\,q_{2\alpha}q_{\alpha+1}^{-3}\left(Q_n^{(\alpha+1)}(\upsilon)-\widetilde{Q}_n^{(\alpha+1)}(\upsilon)\right).$$

Making use of the third-order condition (4.17), by analogy of the weak approximation given in [46] (page 411), we write

$$M_n^{(\alpha)}(\upsilon) = \gamma_1^{\alpha}\mu_{\alpha}^{(1)} + \upsilon^{-1/2}\int_1^{\infty}\mathcal{L}(x;\mathbf{W})\,d\log^{\alpha}x + \alpha\gamma_1^{\alpha-1}\mu_{\alpha}^{(2)}(\rho_1)\,\mathbf{A}_0(n/\upsilon) \\ + \alpha\gamma_1^{\alpha-1}\mu_{\alpha}^{(4)}(\rho_1,\beta_1)\,\mathbf{A}_0(n/\upsilon)\,\mathbf{B}_0(n/\upsilon)\,(1+o_{\mathbf{P}}(1)) + o_{\mathbf{P}}\left(\upsilon^{-1/2}\right). \tag{4.25}$$

Since $\int_1^{\infty}\mathcal{L}(x;\mathbf{W})\,d\log^{\alpha}x = O_{\mathbf{P}}(1)$, then

$$M_n^{(\alpha)}(\upsilon) = \gamma_1^{\alpha}\mu_{\alpha}^{(1)} + \alpha\gamma_1^{\alpha-1}\mu_{\alpha}^{(2)}(\rho_1)\,\mathbf{A}_0(n/\upsilon) \\ + \alpha\gamma_1^{\alpha-1}\mu_{\alpha}^{(4)}(\rho_1,\beta_1)\,\mathbf{A}_0(n/\upsilon)\,\mathbf{B}_0(n/\upsilon)\,(1+o_{\mathbf{P}}(1)) + o_{\mathbf{P}}\left(\upsilon^{-1/2}\right). \tag{4.26}$$

Let us write

$$Q_n^{(\alpha)}(\upsilon) - \widetilde{Q}_n^{(\alpha)}(\upsilon) \\ = \frac{M_n^{(\alpha)}(\upsilon) - \Gamma(\alpha+1)\left(M_n^{(1)}(\upsilon)\right)^2}{M_n^{(2)}(\upsilon) - 2\left(M_n^{(1)}(\upsilon)\right)^2} - \frac{\widetilde{M}_n^{(\alpha)}(\upsilon) - \Gamma(\alpha+1)\left(\widetilde{M}_n^{(1)}(\upsilon)\right)^2}{\widetilde{M}_n^{(2)}(\upsilon) - 2\left(\widetilde{M}_n^{(1)}(\upsilon)\right)^2}.$$

By reducing to the common denominator and by using the weak approximations (4.25) and (4.26) with the fact that $\mathbf{A}_0(n/\upsilon) \overset{\mathbf{P}}{\to} 0$, $\sqrt{\upsilon}\mathbf{A}_0^2(n/\upsilon)$ and $\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\,\mathbf{B}_0(n/\upsilon)$ are stochastically bounded, we get

$$\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\left(Q_n^{(\alpha)}(\upsilon) - \widetilde{Q}_n^{(\alpha)}(\upsilon)\right) \\ = \int_1^{\infty}\mathcal{L}(x;\mathbf{W})\,dg_1(x;\alpha) + \theta_1(\alpha)\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\,\mathbf{B}_0(n/\upsilon) + o_{\mathbf{P}}(1),$$

where

$$g_1(x;\alpha) := \frac{\gamma_1^{\alpha-1}}{2m_2}\left\{\gamma_1^{-\alpha}\log^{\alpha}x - \frac{\alpha\Gamma(\alpha)}{2}r_{\alpha}\gamma_1^{-2}\log^2 x - \left(\alpha\mu_{\alpha}^{(1)} - 2\alpha\Gamma(\alpha)\,r_{\alpha}\right)\gamma_1^{-1}\log x\right\},$$

and $\theta_1(\alpha) := \alpha\gamma_1^{\alpha-2}\{d_\alpha - \Gamma(\alpha) r_\alpha d_2\}/(2m_2)$ with d_α, r_α and m_2 being those defined in the beginning of Section 4.2. It follows that

$$\begin{aligned}&\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon) T_{n1}\\&= -2\delta(\alpha) q_{2\alpha} q_{\alpha+1}^{-3}\left\{\int_1^\infty \mathcal{L}(x;\mathbf{W})\, dg_1(x;\alpha+1) + \theta_1(\alpha+1)\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\mathbf{B}_0(n/\upsilon) + o_\mathbf{P}(1)\right\}.\end{aligned}$$

Likewise, by similar arguments, we also get

$$\begin{aligned}&\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon) T_{n2}\\&= \delta(\alpha) q_{\alpha+1}^{-2}\left\{\int_1^\infty \mathcal{L}(x;\mathbf{W})\, dg_1(x;2\alpha) + \theta_1(2\alpha)\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\mathbf{B}_0(n/\upsilon) + o_\mathbf{P}(1)\right\}.\end{aligned}$$

Therefore

$$\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)(T_{n1}+T_{n2}) = \int_1^\infty \mathcal{L}(x;\mathbf{W})\, dg(x;\alpha) + K(\alpha)\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\mathbf{B}_0(n/\upsilon) + o_\mathbf{P}(1),$$

where $K(\alpha) := \delta(\alpha)\left(q_{\alpha+1}^{-2}\theta_1(2\alpha) - 2q_{2\alpha}q_{\alpha+1}^{-3}\theta_1(\alpha+1)\right)$ and

$$g(x;\alpha) := \delta(\alpha)\left(q_{\alpha+1}^{-2} g_1(x;2\alpha) - 2q_{2\alpha}q_{\alpha+1}^{-3} g_1(x;\alpha+1)\right).$$

Once again by using the third-order condition (4.17) with the fact that $\mathbf{A}_0(n/\upsilon) \xrightarrow{\mathbf{P}} 0$ and $\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\mathbf{B}_0(n/\upsilon) = O_\mathbf{P}(1)$, we show that $\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon) T_{n3} = \eta_1\sqrt{\upsilon}\mathbf{A}_0^2(n/\upsilon) + o_\mathbf{P}(1)$. It is easy to check that $K(\alpha) \equiv \eta_2$, hence we have

$$\begin{aligned}&\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\left(S_n^{(\alpha)}(\upsilon) - s_\alpha(\rho_1)\right)\\&= \int_1^\infty \mathcal{L}(x;\mathbf{W})\, dg(x;\alpha) + \eta_1\sqrt{\upsilon}\mathbf{A}_0^2(n/\upsilon) + \eta_2\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\mathbf{B}_0(n/\upsilon) + o_\mathbf{P}(1),\end{aligned}$$

where η_1 and η_2 are those defined in the beginning of Section 4.2. Recall that $S_n^{(\alpha)}(\upsilon) = s_\alpha\left(\widehat{\rho}_1^{(\alpha)}\right)$, then in view of the mean value theorem and the consistency of $\widehat{\rho}_1^{(\alpha)}$, we end up with

$$s'_{\alpha}(\rho_1)\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\left(\widehat{\rho}_1^{(\alpha)}-\rho_1\right)$$
$$=\int_1^{\infty}\mathcal{L}(x;\mathbf{W})\,dg(x;\alpha)+\eta_1\sqrt{\upsilon}\mathbf{A}_0^2(n/\upsilon)+\eta_2\sqrt{\upsilon}\mathbf{A}_0(n/\upsilon)\mathbf{B}_0(n/\upsilon)+o_{\mathbf{P}}(1).$$

Finally, integrating by parts with elementary calculations completes the proof of the second part of the theorem, namely the Gaussian approximation of $\widehat{\rho}_1^{(\alpha)}$. For the third assertion, it suffices to calculate $\mathbf{E}\left[\int_0^1 s^{-\gamma/\gamma_2-1}\Delta_{\alpha}(s)\mathbf{W}(s)\,ds-\xi\mathbf{W}(1)\right]^2$ to get the asymptotic variance σ_{α}^2, therefore we omit details.

4.8.2 Proof of Theorem 4.2

Let us write

$$\sqrt{k}(\widehat{\gamma}_1-\gamma_1)=\sqrt{k}\left(M_n^{(1)}(k)-\gamma_1\right)+\frac{\widehat{\rho}_1^{(*)}-1}{2\widehat{\rho}_1^{(*)}M_n^{(1)}(k)}\sqrt{k}\left(M_n^{(2)}(k)-2\left(M_n^{(1)}(k)\right)^2\right).$$

From, Theorem 3.1 in [6] and Theorem 4.1 above both $M_n^{(1)}(k)=\widehat{\gamma}_1^{(BMN)}$ and $\widehat{\rho}_1^{(*)}$ are consistent for γ_1 and ρ_1 respectively. It follows that

$$\sqrt{k}(\widehat{\gamma}_1-\gamma_1)$$
$$=\sqrt{k}\left(M_n^{(1)}(k)-\gamma_1\right)+\frac{\rho_1-1}{2\gamma_1\rho_1}\sqrt{k}\left(M_n^{(2)}(k)-2\left(M_n^{(1)}(k)\right)^2\right)(1+o_{\mathbf{P}}(1)).$$

By applying the weak approximation (4.22), for $\alpha=1$, we get

$$\sqrt{k}\left(M_n^{(1)}(k)-\gamma_1\right)=\int_1^{\infty}\mathcal{L}(x;\mathbf{W})\,d\log x+\frac{\sqrt{k}\mathbf{A}_0(n/k)}{1-\rho_1}+o_{\mathbf{P}}(1). \quad (4.27)$$

Using the mean value theorem and the consistency of $M_n^{(1)}(k)$ yields

$$\begin{aligned}&\sqrt{k}\left(\left(M_n^{(1)}(k)\right)^2-\gamma_1^2\right)\\&=2\gamma_1\left\{\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\log x+\frac{\sqrt{k}\mathbf{A}_0(n/k)}{1-\rho_1}+o_{\mathbf{P}}(1)\right\}(1+o_{\mathbf{P}}(1)).\end{aligned}$$

In the sequel we need to the following Lemma.

Lemma 4.2 *For any* $\alpha>0$, *we have* $\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\log^\alpha x=O_{\mathbf{P}}(1)$.

Proof. See the Appendix. ■

From Lemma 4.2 and the assumption $\sqrt{k}\mathbf{A}_0(n/k)=O_{\mathbf{P}}(1)$ as $N\to\infty$ we have

$$\sqrt{k}\left(\left(M_n^{(1)}(k)\right)^2-\gamma_1^2\right)=\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d(2\gamma_1\log x)+\frac{2\gamma_1}{1-\rho_1}\sqrt{k}\mathbf{A}_0(n/k)+o_{\mathbf{P}}(1).$$

Once again, by applying the weak approximation (4.22 , for $\alpha=2$, we write

$$\sqrt{k}\left(M_n^{(2)}(k)-2\gamma_1^2\right)=\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\log^2 x+2\gamma_1\mu_2^{(2)}(\rho_1)\sqrt{k}\mathbf{A}_0(n/k)+o_{\mathbf{P}}(1),$$

where $\mu_2^{(2)}(\rho_1)=\left(1-(1-\rho_1)^2\right)/\left(\rho_1(1-\rho_1)^2\right)$. It follows that

$$\begin{aligned}&\sqrt{k}\left(M_n^{(2)}(k)-2\left(M_n^{(1)}(k)\right)^2\right)\\&=\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\left(\log^2 x-4\gamma_1\log x\right)+\frac{2\gamma_1\rho_1}{(1-\rho_1)^2}\sqrt{k}\mathbf{A}_0(n/k)+o_{\mathbf{P}}(1).\end{aligned}\tag{4.28}$$

By combining approximations (4.27) and (4.28), we obtain

$$\sqrt{k}(\widehat{\gamma}_1-\gamma_1)=\int_1^\infty \mathcal{L}(x;\mathbf{W})\,d\Psi(x)+o_{\mathbf{P}}\left(\sqrt{k}\mathbf{A}_0(n/k)\right),\ \text{as } N\to\infty,$$

where $\Psi(x):=\tau_6\log x+\tau_5\log^2 x$. Finally, making an integration by parts wit a change of variables and elementary calculations, achieves the proof of the first assertion of the theorem. The second part is straightforward.

4.8.3 Proof of Theorem 4.3

It is straightforward. Indeed, by using similar arguments as those used in the proof of Theorem 5.1 in [6], we infer that

$$\frac{\sqrt{k}}{\log d_n}\left(\frac{\widetilde{q}_\nu}{q_\nu}-1\right)=\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)+o_{\mathbf{P}}\left(1\right),\ \text{as } N\rightarrow\infty,$$

then, by making use of Theorem 4.2, the result of Theorem 4.3 comes.

4.9 Conclusion

We proposed an estimation method to the second-order parameter of Pareto-type distributions for randomly right-truncated data which conduced us to a new bias-reduced estimator of the tail index. The useful weak approximation of the tail empirical process, given in [6], allowed us to establish the asymptotic normality of the proposed estimators. We emphasize, that our approach may also be employed to derive several asymptotically normal estimators to both parameters. Indeed, it suffices to define the pertaining theoretical functionals, in terms of the underlying df $\mathbf{F}$ to be replaced by the corresponding product-empirical estimator $\mathbf{F}_n$.

4.10 Appendix

Poof of Lemma. 4.1Let us consider assertion (i) . We begin by letting

$$U^{(\alpha)}:=-\int_1^\infty \log^\alpha s ds^{-1/\gamma_1}\text{ and }\ell\left(t\right):=\frac{M^{(\alpha)}\left(t;\mathbf{F}\right)-\mu_\alpha^{(1)}\left(M^{(1)}\left(t;\mathbf{F}\right)\right)^\alpha}{\mathbf{A}_0\left(t\right)},$$

to show that, for any $\alpha>0$

$$\lim_{t\rightarrow\infty}\ell\left(t\right)=\alpha\gamma_1^{\alpha-1}\left(\mu_\alpha^{(2)}\left(\rho_1\right)-\mu_\alpha^{(1)}\mu_1^{(2)}\left(\rho_1\right)\right). \tag{4.29}$$

Indeed, let us first decompose $\ell(t)$ into

$$\frac{M^{(\alpha)}(t;\mathbf{F})-U^{(\alpha)}}{\mathbf{A}_0(t)}-\mu_\alpha^{(1)}\frac{\left(M^{(1)}(t;\mathbf{F})\right)^\alpha-\left(U^{(1)}\right)^\alpha}{\mathbf{A}_0(t)}+\frac{U^{(\alpha)}(t)-\mu_\alpha^{(1)}\left(U^{(1)}\right)^\alpha}{\mathbf{A}_0(t)},$$

and note that $\mu_\alpha^{(1)}=\Gamma(\alpha+1)$ where $\Gamma(a)=\int_0^\infty e^{-x}x^{a-1}dx=\int_1^\infty t^{-2}\log^{a-1}t dt$, $a>0$. It is easy to verify that $U^{(\alpha)}-\mu_\alpha^{(1)}\left(U^{(1)}\right)^\alpha=0$, therefore

$$\ell(t)=\frac{M^{(\alpha)}(t;\mathbf{F})-U^{(\alpha)}}{\mathbf{A}_0(t)}-\mu_\alpha^{(1)}\frac{\left(M^{(1)}(t;\mathbf{F})\right)^\alpha-\left(U^{(1)}\right)^\alpha}{\mathbf{A}_0(t)}. \tag{4.30}$$

Recall that $M^{(\alpha)}(t;\mathbf{F})=\int_u^\infty\log^\alpha(x/u)\,d\mathbf{F}(x)/\overline{\mathbf{F}}(u)$, where $u:=\mathbb{U}_{F^*}(t)$, which by a change of variables and an integration by parts, may be rewritten into

$$\int_1^\infty\frac{\overline{\mathbf{F}}(ux)}{\overline{\mathbf{F}}(u)}d\log^\alpha x=:M_u^{(\alpha)}(\mathbf{F}).$$

Making use, once again, of Proposition 4 of Hua [57], we write: for possibly different function $\widetilde{\mathbf{A}}_\mathbf{F}$, with $\widetilde{\mathbf{A}}_\mathbf{F}(y)\sim\mathbf{A}_\mathbf{F}(y)$, as $y\to\infty$, for any $0<\epsilon<1$ and $x\geq1$, we have

$$\left|\frac{\overline{\mathbf{F}}(ux)/\overline{\mathbf{F}}(u)-x^{-1/\gamma_1}}{\gamma_1^{-2}\widetilde{\mathbf{A}}_\mathbf{F}\left(1/\overline{\mathbf{F}}(u)\right)}-x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\gamma_1/\rho_1}\right|\leq\epsilon x^{-1/\gamma_1+\epsilon},\text{ as }u\to\infty.$$

By using elementary analysis, we easily show that this inequality implies that

$$\frac{M_u^{(\alpha)}(\mathbf{F})-U^{(\alpha)}}{\widetilde{\mathbf{A}}_\mathbf{F}\left(1/\overline{\mathbf{F}}(u)\right)}\to\alpha\gamma_1^{\alpha-1}\mu_\alpha^{(2)}(\rho_1),\text{ as }u\to\infty.$$

Hence, since $1/\overline{\mathbf{F}}(u)\to\infty$ as $u\to\infty$, then $\widetilde{\mathbf{A}}_\mathbf{F}\left(1/\overline{\mathbf{F}}(u)\right)\sim\mathbf{A}_\mathbf{F}\left(1/\overline{\mathbf{F}}(u)\right)=\mathbf{A}_0(t)$. This means that

$$\frac{M^{(\alpha)}(t,\mathbf{F})-U^{(\alpha)}}{\mathbf{A}_0(t)}\to\alpha\gamma_1^{\alpha-1}\mu_\alpha^{(2)}(\rho_1),\text{ as }t\to\infty. \tag{4.31}$$

Note that for $\alpha=1$, we have $U^{(1)}=\gamma_1$ and therefore $\left(M^{(1)}(t,\mathbf{F})-\gamma_1\right)/\mathbf{A}_0(t)\to\mu_1^{(2)}(\rho_1)$, which implies that $M^{(1)}(t,\mathbf{F})\to\gamma_1$. By using the mean value theorem and the previous two results we get

$$\frac{\left(M^{(1)}\left(t,\mathbf{F}\right)\right)^{\alpha}-\gamma_{1}^{\alpha}}{\mathbf{A}_{0}\left(t\right)}\rightarrow\alpha\gamma_{1}^{\alpha-1}\mu_{1}^{(2)}\left(\rho_{1}\right),\text{ as }t\rightarrow\infty. \tag{4.32}$$

Combining (4.30) , (4.31) and (4.32) leads to (4.29) . Finally, we use the fact that $M^{(1)}\left(t,\mathbf{F}\right)\rightarrow\gamma_{1}$ to achieve the proof of assertion (i) . To show assertion (ii) , we apply assertion (i) twice, for $\alpha>0$ and for $\alpha=2$, then we divide the respective results to get

$$\begin{aligned}Q^{(\alpha)}\left(t;\mathbf{F}\right)&=\frac{M^{(\alpha)}\left(t;\mathbf{F}\right)-\mu_{\alpha}^{(1)}\left[M^{(1)}\left(t;\mathbf{F}\right)\right]^{\alpha}}{M^{(2)}\left(t;\mathbf{F}\right)-2\left[M^{(1)}\left(t;\mathbf{F}\right)\right]^{2}}\\&\sim\frac{\left(M^{(1)}\left(t;\mathbf{F}\right)\right)^{\alpha-1}}{M^{(1)}\left(t;\mathbf{F}\right)}\frac{\alpha\left(\mu_{\alpha}^{(2)}\left(\rho_{1}\right)-\mu_{\alpha}^{(1)}\mu_{1}^{(2)}\left(\rho_{1}\right)\right)}{2\left(\mu_{2}^{(2)}\left(\rho_{1}\right)-\mu_{2}^{(1)}\mu_{1}^{(2)}\left(\rho_{1}\right)\right)}.\end{aligned}$$

By replacing $\mu_{\alpha}^{(1)}$ and $\mu_{\alpha}^{(2)}\left(\rho_{1}\right)$ by their expressions, given in (4.12), we get

$$\begin{aligned}&\alpha\left(\mu_{\alpha}^{(2)}\left(\rho_{1}\right)-\mu_{\alpha}^{(1)}\mu_{1}^{(2)}\left(\rho_{1}\right)\right)\\&=\alpha\left\{\frac{\Gamma\left(\alpha\right)\left(1-\left(1-\rho_{1}\right)^{\alpha}\right)}{\rho_{1}\left(1-\rho_{1}\right)^{\alpha}}-\Gamma\left(\alpha+1\right)\frac{\Gamma\left(1\right)\left(1-\left(1-\rho_{1}\right)\right)}{\rho_{1}\left(1-\rho_{1}\right)}\right\}\\&=\alpha\left\{\frac{\Gamma\left(\alpha\right)\left(1-\left(1-\rho_{1}\right)^{\alpha}\right)}{\rho_{1}\left(1-\rho_{1}\right)^{\alpha}}-\frac{\Gamma\left(\alpha+1\right)}{1-\rho_{1}}\right\}.\end{aligned}$$

Since $M^{(1)}\left(t,\mathbf{F}\right)\rightarrow\gamma_{1}$, then

$$Q^{(\alpha)}\left(t;\mathbf{F}\right)\rightarrow\frac{\gamma_{1}^{\alpha-2}\Gamma\left(\alpha+1\right)\left(1-\left(1-\rho_{1}\right)^{\alpha}-\alpha\rho_{1}\left(1-\rho_{1}\right)^{\alpha-1}\right)}{2\rho_{1}^{2}\left(1-\rho_{1}\right)^{\alpha-2}},\text{ as }t\rightarrow\infty,$$

which is $q_{\alpha}\left(\rho_{1}\right)$ given in (4.10) . For assertion (iii) , it is clear that

$$\delta\left(\alpha\right)\frac{Q_{t}^{(2\alpha)}}{\left(Q_{t}^{(\alpha+1)}\right)^{2}}\rightarrow\frac{\rho_{1}^{2}\left(1-\left(1-\rho_{1}\right)^{2\alpha}-2\alpha\rho_{1}\left(1-\rho_{1}\right)^{2\alpha-1}\right)}{\left(1-\left(1-\rho_{1}\right)^{\alpha+1}-\left(\alpha+1\right)\rho_{1}\left(1-\rho_{1}\right)^{\alpha}\right)^{2}},$$

which meets the expression of $s_{\alpha}\left(\rho_{1}\right)$ given in (4.11) . ■

Proof of Lemma. 4.2. Observe that $\int_{1}^{\infty}\mathcal{L}\left(x;\mathbf{W}\right)d\log^{\alpha}x$ may be decomposed into the sum of

$$I_1 := \frac{\gamma}{\gamma_1}\int_1^\infty x^{1/\gamma_2}\mathbf{W}\left(x^{-1/\gamma}\right)d\log^\alpha x,\ \ I_2 := -\frac{\gamma}{\gamma_1}\mathbf{W}\left(1\right)\int_1^\infty x^{-1/\gamma_1}d\log^\alpha x,$$

$$I_3 := \frac{\gamma}{\gamma_1+\gamma_2}\int_1^\infty x^{1/\gamma_2}\left\{\int_0^1 s^{-\gamma/\gamma_2-1}\mathbf{W}\left(x^{-1/\gamma}s\right)ds\right\}d\log^\alpha x,$$

and

$$I_4 := -\frac{\gamma}{\gamma_1+\gamma_2}\left\{\int_0^1 s^{-\gamma/\gamma_2-1}\mathbf{W}\left(s\right)ds\right\}\int_1^\infty x^{-1/\gamma_1}d\log^\alpha x.$$

Next we show that $I_i = O_{\mathbf{P}}\left(1\right),\ i=1,...,4$. To this end, we will show that $\mathbf{E}\left|I_i\right|$ is finite for $i=1,...,4$. Indeed, we have $\mathbf{E}\left|I_1\right| \leq \left(\gamma/\gamma_1\right)\int_1^\infty x^{-1/\gamma_1}x^{1/\gamma}\mathbf{E}\left|\mathbf{W}\left(x^{-1/\gamma}\right)\right|d\log^\alpha x$. Since $\mathbf{E}\left|\mathbf{W}\left(y\right)\right| \leq \sqrt{y}$, for any $0 \leq y \leq 1$, then $\mathbf{E}\left|I_1\right| \leq \left(\alpha\gamma/\gamma_1\right)\int_1^\infty x^{1/\gamma_2-1/(2\gamma)-1}\log^{\alpha-1}xdx$. By successively making two changes of variables $\log x = t$, then $\left(-1/\gamma_2+1/\left(2\gamma\right)+1\right)t = s$, we end up with $\mathbf{E}\left|I_1\right| \leq \gamma\gamma_1^{\alpha-1}\left(2\gamma/\left(2\gamma-\gamma_1\right)\right)^\alpha\Gamma\left(\alpha+1\right)$ which is finite for any $\alpha > 0$. By similar arguments we also show that $\mathbf{E}\left|I_2\right| \leq \gamma\gamma_1^{\alpha-1}\Gamma\left(\alpha+1\right)$ which is finite as well. For the third term I_3, we have

$$\mathbf{E}\left|I_3\right| \leq \gamma/\left(\gamma_1+\gamma_2\right)\int_1^\infty x^{1/\gamma_2}\left\{\int_0^1 s^{-\gamma/\gamma_2-1}\mathbf{E}\left|\mathbf{W}\left(x^{-1/\gamma}s\right)\right|ds\right\}d\log^\alpha x.$$

By elementary calculations, we get

$$\mathbf{E}\left|I_3\right| \leq \frac{\gamma_2\left(2\gamma\right)^{\alpha+1}}{\left(\gamma_2-2\gamma\right)\left(\gamma_1+\gamma_2\right)}\left(\frac{\gamma_1}{2\gamma-\gamma_1}\right)^\alpha\Gamma\left(\alpha+1\right),$$

which is also finite. By using similar arguments, we get

$$\mathbf{E}\left|I_4\right| \leq \frac{2\gamma_2\gamma\gamma_1^\alpha\Gamma\left(\alpha+1\right)}{\left(\gamma_1+\gamma_2\right)\left(\gamma_2-2\gamma\right)} < \infty,$$

as sought. ■

Chapter 5

A Lynden-Bell integral estimator for the tail index of right-truncated data with a random threshold

By means of a Lynden-Bell integral with deterministic threshold, recently Worms and Worms [A Lynden-Bell integral estimator for extremes of randomly truncated data. Statist. Probab. Lett. 2016; 109: 106-117] introduced an asymptotically normal estimator of the tail index for Pareto-type (randomly right-truncated) data. In this context, we consider the random threshold case to derive a Hill-type estimator and establish its consistency and asymptotic normality. A simulation study is carried out to evaluate the finite sample behavior of the proposed estimator and compare it to the existing ones.

5.1 Introduction

Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \leq i \leq N$ be a sample of size $N \geq 1$ from a couple $(\mathbf{X}, \mathbf{Y})$ of independent random variables (rv's) defined over some probability space $(\Omega, \mathcal{A}, \mathbf{P})$, with continuous marginal distribution functions (df's) $\mathbf{F}$ and $\mathbf{G}$ respectively. Suppose that $\mathbf{X}$ is truncated to the right by $\mathbf{Y}$, in the sense that $\mathbf{X}_i$ is only observed when $\mathbf{X}_i \leq \mathbf{Y}_i$. We assume that both survival functions $\overline{\mathbf{F}} := 1 - \mathbf{F}$ and $\overline{\mathbf{G}} := 1 - \mathbf{G}$ are regularly varying at infinity with respective negative indices $-1/\gamma_1$ and $-1/\gamma_2$. (4.1) holds.

It is well known that, in extreme value analysis, weak approximations are achieved in the second-order framework (see, e.g., de Haan 2006 [20] page 48). Thus, it seems quite natural to suppose that $\mathbf{F}$ and $\mathbf{G}$ satisfy the second-order condition of regular variation 4.2 and 4.3 respectively with $\tau_1, \tau_2 < 0$ are the second-order parameters and $\mathbf{A_F}$, $\mathbf{A_G}$ are functions tending to zero and not changing signs near infinity with regularly varying absolute values at infinity with indices τ_1, τ_2 respectively, which we express in terms of the tail quantile functions pertaining to both df's.

For any df K, the function $\mathbb{U}_K(t) := K^{\leftarrow}(1 - 1/t)$, $t > 1$, stands for the tail quantile function, with $K^{\leftarrow}(u) := \inf\{v : K(v) \geq u\}$, $0 < u < 1$, denoting the generalized inverse of K. From Lemma 3 in [62] , the second-order conditions (4.2) and (4.3) imply that there exist constants $d_1, d_2 > 0$, such that

$$\overline{\mathbf{F}}(x) = d_1 x^{-1/\gamma_1} \ell_1(x) \text{ and } \overline{\mathbf{G}}(x) = d_2 x^{-1/\gamma_2} \ell_2(x), \ x > 0, \tag{5.1}$$

where $\lim_{x\to\infty} \ell_i(x) = 1$ and $|1 - \ell_i|$ is regularly varying at infinity with tail index $\tau_i \gamma$, $i = 1, 2$. This ccondition is fullfilled by many commonly used models such as Burr, Fréchet, Generalized Pareto, absolute Student, log-gamma distributions, to name but a few. Also known as heavy-tailed, Pareto-type or Pareto-like distributions, these models take a prominent role in extreme value theory and have important practical applications as they are used rather systematically in certain branches of non-life insurance, as well as in finance, telecommunications, hydrology, etc... [77].

Let us now denote (X_i, Y_i), $i = 1, ..., n$ to be the observed data, as copies of a couple of rv's (X, Y), corresponding to the truncated sample $(\mathbf{X}_i, \mathbf{Y}_i)$, $i = 1, ..., N$, where $n = n_N$ is a sequence of discrete rv's which, in virtue of the weak law of large numbers, satisfies $n_N/N \overset{\mathbf{P}}{\rightarrow} p := \mathbf{P}(\mathbf{X} \leq \mathbf{Y})$, as $N \rightarrow \infty$. We denote the joint df of X and Y by $H(x,y) := \mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(\mathbf{X} \leq \min(x, \mathbf{Y}), \mathbf{Y} \leq y \mid \mathbf{X} \leq \mathbf{Y})$, which is equal to $p^{-1}\int_0^y \mathbf{F}(\min(x,z))\, d\mathbf{G}(z)$. The marginal distributions of the rv's X and Y, respectively denoted by F and G, are given by $F(x) = p^{-1}\int_0^x \overline{\mathbf{G}}(z)\, d\mathbf{F}(z)$ and $G(y) = p^{-1}\int_0^y \mathbf{F}(z)\, d\mathbf{G}(z)$. Since $\mathbf{F}$ and $\mathbf{G}$ are heavy-tailed, then their right endpoints are infinite and thus they are equal. Hence, from Woodrofee 1985, [88], we may write $\int_x^\infty d\mathbf{F}(y)/\mathbf{F}(y) = \int_x^\infty dF(y)/C(y)$, where $C(z) := \mathbf{P}(X \leq z \leq Y)$. Differentiating the previous equation leads to the following crucial equation $C(x)\, d\mathbf{F}(x) = \mathbf{F}(x)\, dF(x)$, whose solution is defined by $\mathbf{F}(x) = \exp\left\{-\int_x^\infty dF(z)/C(z)\right\}$. This leads to Woodroofe's nonparametric estimator [88] of df $\mathbf{F}$, given by

$$\mathbf{F}_n^{(\mathbf{W})}(x) := \prod_{i:X_i>x} \exp\left\{-\frac{1}{nC_n(X_i)}\right\},$$

which is derived only by replacing df's F and C by their respective empirical counterparts $F_n(x) := n^{-1}\sum_{i=1}^n \mathbf{1}(X_i \leq x)$ and $C_n(x) := n^{-1}\sum_{i=1}^n \mathbf{1}(X_i \leq x \leq Y_i)$. There exists a more popular estimator for $\mathbf{F}$, known as Lynden-Bell nonparametric maximum likelihood estimator [64], defined by

$$\mathbf{F}_n^{(\mathbf{LB})}(x) := \prod_{i:X_i>x} \left(1 - \frac{1}{nC_n(X_i)}\right),$$

which will be considered in this paper to derive a new estimator for the tail index of df $\mathbf{F}$. Note that the tail of df F simultaneously depends on $\overline{\mathbf{G}}$ and $\overline{\mathbf{F}}$ while that of $\overline{G}$ only relies on $\overline{\mathbf{G}}$. By using Proposition B.1.10 in [20], to the regularly varying functions $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$, we show that both $\overline{F}$ and $\overline{G}$ are regularly varying at infinity as well, with respective indices $-1/\gamma := -(\gamma_1+\gamma_2)/(\gamma_1\gamma_2)$ and $-1/\gamma_2$. In view of the definition of γ, Gardes and Stupfler 2015, [42] derived a consistent estimator, for the extreme value index γ_1, whose asymptotic normality is established in [5], under the tail dependence and the second-order conditions of regular variation. Recently, by considering a Lynden-Bell integration with a deterministic threshold $t_n > 0$, Wormes and Wormes 2016, [91] proposed another asymptotically normal estimator for γ_1 as follows:

$$\widehat{\gamma}_1^{(\mathbf{LB})}(t_n) := \frac{1}{n\overline{\mathbf{F}}_n^{(\mathbf{LB})}(t_n)} \sum_{i=1}^{n} \mathbf{1}(X_i > t_n) \frac{\mathbf{F}_n^{(\mathbf{LB})}(X_i)}{C_n(X_i)} \log \frac{X_i}{t_n}.$$

Likewise, Benchaira et al. 2016a, [6] considered a Woodroofe integration (with a random threshold) to propose a new estimator for the tail index γ_1 given by

$$\widehat{\gamma}_1^{(\mathbf{W})} := \frac{1}{n\overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n})} \sum_{i=1}^{k} \frac{\mathbf{F}_n^{(\mathbf{W})}(X_{n-i+1:n})}{C_n(X_{n-i+1:n})} \log \frac{X_{n-i+1:n}}{X_{n-k:n}},$$

where, given $n = m = m_N$, $Z_{1:m} \leq ... \leq Z_{m:m}$ denote the order statistics pertaining to a sample $Z_1, ..., Z_m$, and $k = k_n$ is a (random) sequence of integers such that, given $n = m$, $1 < k_m < m$, $k_m \to \infty$ and $k_m/m \to 0$ as $N \to \infty$. The consistency and asymptotic normality of $\widehat{\gamma}_1^{(\mathbf{W})}$ are established in [6] through a weak approximation to Woodroofe's tail process

$$\mathbf{D}_n^{(\mathbf{W})}(x) := \sqrt{k}\left(\frac{\overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n}x)}{\overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n})} - x^{-1/\gamma_1}\right),\ x > 0.$$

More precisely, the authors showed that, under (4.2) and (4.3) with $\gamma_1 < \gamma_2$, there exist a function $\mathbf{A}_0(t) \sim \mathbf{A}_{\mathbf{F}}^*(t) := \mathbf{A}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}(\mathbb{U}_F(t))\right)$, $t \to \infty$, and a standard Wiener process $\{\mathbf{W}(s);\ s \geq 0\}$, defined on the probability space $(\Omega, \mathcal{A}, \mathbf{P})$, such that, for $0 < \epsilon < 1/2 - \gamma/\gamma_2$ and $x_0 > 0$,

$$\sup_{x \geq x_0} x^{(1/2-\epsilon)/\gamma - 1/\gamma_2}\left|\mathbf{D}_n^{(\mathbf{W})}(x) - \mathbf{\Gamma}(x;\mathbf{W}) - x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1}-1}{\gamma_1\tau_1}\sqrt{k}\mathbf{A}_0(n/k)\right| = o_{\mathbf{P}}(1), \quad (5.2)$$

as $N \to \infty$, provided that given $n = m$, $\sqrt{k_m}\mathbf{A}_0(m/k_m) = O(1)$, where $\{\Gamma(x;\mathbf{W});\ x > 0\}$

is a Gaussian process defined by

$$\begin{aligned}\mathbf{\Gamma}\left(x;\mathbf{W}\right) := & \frac{\gamma}{\gamma_1}x^{-1/\gamma_1}\left\{x^{1/\gamma}\mathbf{W}\left(x^{-1/\gamma}\right)-\mathbf{W}\left(1\right)\right\} \\ & +\frac{\gamma}{\gamma_1+\gamma_2}x^{-1/\gamma_1}\int_0^1 s^{-\gamma/\gamma_2-1}\left\{x^{1/\gamma}\mathbf{W}\left(x^{-1/\gamma}s\right)-\mathbf{W}\left(s\right)\right\}ds.\end{aligned}$$

In view of the previous weak approximation, the authors also proved that if, given $n=m$, $\sqrt{k_m}\mathbf{A}_{\mathbf{F}}^{*}\left(m/k_m\right)\to\lambda$, then $\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)\overset{\mathcal{D}}{\to}\mathcal{N}\left(\dfrac{\lambda}{1-\tau_1},\sigma^2\right)$, as $N\to\infty$, where $\sigma^2 := \gamma^2\left(1+\gamma_1/\gamma_2\right)\left(1+\left(\gamma_1/\gamma_2\right)^2\right)/\left(1-\gamma_1/\gamma_2\right)^3$. Recently, Benchaira et al. 2016b, [7] followed this approach to introduce a kernel estimator to γ_1 which improves the bias of $\widehat{\gamma}_1^{(\mathbf{W})}$. In this paper, we are interested in Worm's estimator $\widehat{\gamma}_1^{(\mathbf{LB})}\left(t_n\right)$, but with a threshold t_n that is assumed to be random and equal to $X_{n-k:n}$. This makes the estimator more convenient for numerical implementation than the one with a deterministic threshold. In other words, we will deal with the following tail index estimator:

$$\widehat{\gamma}_1^{(\mathbf{LB})} := \frac{1}{n\overline{\mathbf{F}}_n^{(\mathbf{LB})}\left(X_{n-k:n}\right)}\sum_{i=1}^{k}\frac{\mathbf{F}_n^{(\mathbf{LB})}\left(X_{n-i+1:n}\right)}{C_n\left(X_{n-i+1:n}\right)}\log\frac{X_{n-i+1:n}}{X_{n-k:n}}.$$

Note that $\mathbf{F}_n^{(\mathbf{LB})}\left(\infty\right)=1$ and write $\overline{\mathbf{F}}_n^{(\mathbf{LB})}\left(X_{n-k:n}\right)=\int_{X_{n-k:n}}^{\infty}d\mathbf{F}_n^{(\mathbf{LB})}\left(y\right)$. On the other hand, we have $C_n\left(x\right)d\mathbf{F}_n^{(\mathbf{LB})}\left(x\right)=\mathbf{F}_n^{(\mathbf{LB})}\left(x\right)dF_n\left(x\right)$ (see, e.g., [84]), then

$$\overline{\mathbf{F}}_n^{(\mathbf{LB})}\left(X_{n-k:n}\right)=\int_{X_{n-k:n}}^{\infty}\frac{\mathbf{F}_n^{(\mathbf{LB})}\left(x\right)}{C_n\left(x\right)}dF_n\left(x\right)=\frac{1}{n}\sum_{i=1}^{k}\frac{\mathbf{F}_n^{(\mathbf{LB})}\left(X_{n-i+1:n}\right)}{C_n\left(X_{n-i+1:n}\right)}.$$

This allows us to rewrite the new estimator into

$$\widehat{\gamma}_1^{(\mathbf{LB})} := \sum_{i=1}^{k}a_n^{(i)}\log\frac{X_{n-i+1:n}}{X_{n-k:n}},$$

where

$$a_n^{(i)} := \frac{\mathbf{F}_n^{(\mathbf{LB})}\left(X_{n-i+1:n}\right)}{C_n\left(X_{n-i+1:n}\right)}\Big/\sum_{i=1}^{k}\frac{\mathbf{F}_n^{(\mathbf{LB})}\left(X_{n-i+1:n}\right)}{C_n\left(X_{n-i+1:n}\right)}.$$

It is worth mentioning that for complete data, we have $n \equiv N$ and $\mathbf{F}_n \equiv F_n \equiv C_n$, it follows that $a_n^{(i)} \equiv k^{-1}$, $i = 1, ..., k$ and consequently both $\widehat{\gamma}_1^{(\mathbf{LB})}$ and $\widehat{\gamma}_1^{(\mathbf{W})}$ reduce to the classical Hill estimator [60]. The consistency and asymptotic normality of $\widehat{\gamma}_1^{(\mathbf{LB})}$ will be achieved through a weak approximation of the corresponding tail Lynden-Bell process that we define by

$$\mathbf{D}_n^{(\mathbf{LB})}(x) := \sqrt{k}\left(\frac{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}x)}{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n})} - x^{-1/\gamma_1}\right), \ x > 0.$$

The rest of the paper is organized as follows. In Section 5.2, we provide our main results whose proofs are postponed to Section 5.4. The finite sample behavior of the proposed estimator $\widehat{\gamma}_1^{(\mathbf{LB})}$ is checked by simulation in Section 5.3, where a comparison with the one recently introduced by Benchaira et al. 2016a, [6] is made as well.

5.2 Main results

We basically have three main results. The first one, that we give in Theorem 5.1, consists in an asymptotic relation between the above mentioned estimators of the distribution tail, namely $\overline{\mathbf{F}}_n^{(\mathbf{W})}$ and $\overline{\mathbf{F}}_n^{(\mathbf{LB})}$. This in turn is instrumental to the Gaussian approximation of the tail Lynden-Bell process $\mathbf{D}_n^{(\mathbf{LB})}(x)$ stated in Theorem 5.2. Finally, in Theorem 5.3, we deduce the asymptotic behavior of the tail index estimator $\widehat{\gamma}_1^{(\mathbf{LB})}$.

Theorem 5.1 *Assume that both* $\mathbf{F}$ *and* $\mathbf{G}$ *satisfy the second-order conditions (4.2) and (4.3) respectively with* $\gamma_1 < \gamma_2$. *Let* $k = k_n$ *be a random sequence of integers such that, given* $n = m$, $k_m \to \infty$ *and* $k_m/m \to 0$, *as* $N \to \infty$, *then, for any* $x_0 > 0$, *we have*

$$\sup_{x \geq x_0} x^{1/\gamma_1} \frac{\left|\overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n}x) - \overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}x)\right|}{\overline{\mathbf{F}}(X_{n-k:n})} = O_{\mathbf{P}}\left((k/n)^{\gamma_1/\gamma}\right).$$

Theorem 5.2 *Assume that the assumptions of Theorem 5.1 hold and given* $n = m$,

$$k_m^{1+\gamma_1/(2\gamma)}/m \to 0, \tag{5.3}$$

and $\sqrt{k_m}\mathbf{A}_0(m/k_m) = O(1)$, *as* $N \to \infty$. *Then, for any* $x_0 > 0$ *and* $0 < \epsilon < 1/2 - \gamma/\gamma_2$, *we have*

$$\sup_{x \geq x_0} x^{(1/2-\epsilon)/\gamma - 1/\gamma_2} \left| \mathbf{D}_n^{(\mathbf{LB})}(x) - \mathbf{\Gamma}(x; \mathbf{W}) - x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \mathbf{A}_0(n/k) \right| = o_{\mathbf{P}}(1).$$

Theorem 5.3 *Assume that* (4.1) *holds with* $\gamma_1 < \gamma_2$ *and let* $k = k_n$ *be a random sequence of integers such that given* $n = m$, $k_m \to \infty$ *and* $k_m/m \to 0$, *as* $N \to \infty$, *then* $\widehat{\gamma}_1^{(\mathbf{LB})} \overset{\mathbf{P}}{\to} \gamma_1$. *Assume further that the assumptions of Theorem* 5.2 *hold, then*

$$\sqrt{k}\left(\widehat{\gamma}_1^{(\mathbf{LB})} - \gamma_1\right) = \frac{\sqrt{k}\mathbf{A}_0(n/k)}{1 - \tau_1} - \gamma \mathbf{W}(1) + \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 (\gamma_2 - \gamma_1 - \gamma \log s) s^{-\gamma/\gamma_2 - 1} \mathbf{W}(s) ds + o_{\mathbf{P}}(1).$$

If, in addition, we suppose that, given $n = m$, $\sqrt{k_m}\mathbf{A}_{\mathbf{F}}^*(m/k_m) \to \lambda < \infty$, *then*

$$\sqrt{k}\left(\widehat{\gamma}_1^{(\mathbf{LB})} - \gamma_1\right) \overset{\mathcal{D}}{\to} \mathcal{N}\left(\frac{\lambda}{1 - \tau_1}, \sigma^2\right), \text{ as } N \to \infty.$$

5.3 Simulation study

In this section, we illustrate the finite sample behavior of $\widehat{\gamma}_1^{(\mathbf{LB})}$ and, at the same time, we compare it with $\widehat{\gamma}_1^{(\mathbf{W})}$. To this end, we consider two sets of truncated and truncation data, both drawn from Burr's model: $\overline{\mathbf{F}}(x) = \left(1 + x^{1/\delta}\right)^{-\delta/\gamma_1}$, $\overline{\mathbf{G}}(x) = \left(1 + x^{1/\delta}\right)^{-\delta/\gamma_2}$, $x \geq 0$, where $\delta, \gamma_1, \gamma_2 > 0$. The corresponding percentage of observed data is equal to $p = \gamma_2/(\gamma_1 + \gamma_2)$. We fix $\delta = 1/4$ and choose the values 0.6 and 0.8 for γ_1 and 55%, 70% and 90% for p. For each couple (γ_1, p), we solve the equation $p = \gamma_2/(\gamma_1 + \gamma_2)$ to get the pertaining γ_2-value. We vary the common size N of both samples $(\mathbf{X}_1, ..., \mathbf{X}_N)$ and $(\mathbf{Y}_1, ..., \mathbf{Y}_N)$, then for each size, we generate 1000 independent replicates. Our overall results are taken as the empirical means of the results obtained through all repetitions. To determine the optimal number of top statistics used in the computation of the tail index

$\gamma_1 = 0.6;\ p = 0.55$

		$\widehat{\gamma}_1^{(\mathbf{LB})}$			$\widehat{\gamma}_1^{(\mathbf{W})}$		
N	n	abs bias	rmse	k^*	abs bias	rmse	k^*
100	54	0.0407	0.2381	26	0.0443	0.2328	26
200	109	0.0378	0.2610	36	0.0358	0.2532	37
300	165	0.0352	0.2359	36	0.0323	0.2315	37
500	274	0.0199	0.2290	61	0.0185	0.2238	61
1000	549	0.0074	0.1763	112	0.0068	0.1748	112
3000	1649	0.0036	0.0982	350	0.0037	0.0981	352
5000	2747	0.0007	0.1066	432	0.0007	0.1065	432

Table 5.1: Estimation results of Lynden-Bell based (leftt pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.6$ of Burr's model through 1000 right-truncated samples with 45%-truncation rate.

$\gamma_1 = 0.6;\ p = 0.7$

		$\widehat{\gamma}_1^{(\mathbf{LB})}$			$\widehat{\gamma}_1^{(\mathbf{W})}$		
N	n	abs bias	rmse	k^*	abs bias	rmse	k^*
100	69	0.0158	0.2451	25	0.0144	0.2428	25
200	140	0.0095	0.1871	39	0.0089	0.1866	39
300	210	0.0085	0.1590	61	0.0082	0.1587	61
500	348	0.0074	0.1294	76	0.0072	0.1293	76
1000	699	0.0063	0.1014	124	0.0062	0.1014	124
3000	2096	0.0053	0.0962	246	0.0053	0.0962	246
5000	3498	0.0036	0.0984	400	0.0036	0.0984	400

Table 5.2: Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.6$ of Burr's model through 1000 right-truncated samples with 30%-truncation rate.

estimate values, we use the algorithm of Reiss and Thomas [76], page 137. Our illustration and comparison are made with respect to the estimators absolute biases (abs bias) and the roots of their mean squared errors (rmse). We summarize the simulation results in Tables 5.1, 5.2 and 5.3 for $\gamma_1 = 0.6$ and in Tables 5.4, 5.5 and 5.6 for $\gamma_1 = 0.8$. After the inspection of all the tables, two conclusions can be drawn regardless of the situation. First, the estimation accuracy of both estimators decreases when the truncation percentage increases and this was quite expected. Second, we notice that the newly proposed estimator $\widehat{\gamma}_1^{(\mathbf{LB})}$ and $\widehat{\gamma}_1^{(\mathbf{W})}$ behave equally well.

$\gamma_1 = 0.6;\ p = 0.9$

		$\widehat{\gamma}_1^{(\mathbf{LB})}$			$\widehat{\gamma}_1^{(\mathbf{W})}$		
N	n	abs bias	rmse	k^*	abs bias	rmse	k^*
100	90	0.0073	0.1779	21	0.0070	0.1778	21
200	180	0.0066	0.1208	54	0.0064	0.1208	54
300	270	0.0055	0.1133	88	0.0056	0.1133	88
500	450	0.0050	0.0864	125	0.0050	0.0863	125
1000	898	0.0030	0.0614	189	0.0029	0.0614	189
3000	2702	0.0016	0.0494	398	0.0016	0.0494	398
5000	4496	0.0010	0.0112	467	0.0010	0.0112	467

Table 5.3: Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.6$ of Burr's model through 1000 right-truncated samples with 10%-truncation rate.

$\gamma_1 = 0.8;\ p = 0.55$

		$\widehat{\gamma}_1^{(\mathbf{LB})}$			$\widehat{\gamma}_1^{(\mathbf{W})}$		
N	n	abs bias	rmse	k^*	abs bias	rmse	k^*
100	55	0.0570	0.3330	30	0.0636	0.3167	31
200	110	0.0401	0.3604	33	0.0347	0.3453	35
300	164	0.0252	0.2563	69	0.0272	0.2530	71
500	276	0.0227	0.1807	112	0.0216	0.1794	113
1000	551	0.0148	0.1795	196	0.0142	0.1788	197
3000	1647	0.0124	0.1794	525	0.0121	0.1783	525
5000	2751	0.0075	0.1260	688	0.0074	0.1259	688

Table 5.4: Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.8$ of Burr's model through 1000 right-truncated samples with 45%-truncation rate.

$\gamma_1 = 0.8$; $p = 0.7$

		$\widehat{\gamma}_1^{(\mathbf{LB})}$			$\widehat{\gamma}_1^{(\mathbf{W})}$		
N	n	abs bias	rmse	k^*	abs bias	rmse	k^*
100	69	0.0217	0.3827	28	0.0195	0.3787	28
200	139	0.0203	0.2918	59	0.0194	0.2905	59
300	210	0.0189	0.1857	66	0.0184	0.1852	66
500	348	0.0143	0.1593	113	0.0140	0.1591	113
1000	700	0.0049	0.1205	230	0.0049	0.1204	230
3000	2100	0.0037	0.0886	449	0.0038	0.0886	449
5000	3500	0.0031	0.0857	500	0.0031	0.0857	500

Table 5.5: Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.8$ of Burr's model through 1000 right-truncated samples with 30%-truncation rate.

$\gamma_1 = 0.8$; $p = 0.9$

		$\widehat{\gamma}_1^{(\mathbf{LB})}$			$\widehat{\gamma}_1^{(\mathbf{W})}$		
N	n	abs bias	rmse	k^*	abs bias	rmse	k^*
100	89	0.0380	0.1833	38	0.0369	0.1827	38
200	179	0.0345	0.1383	80	0.0342	0.1383	80
300	269	0.0173	0.1014	99	0.0175	0.1013	99
500	450	0.0108	0.0927	143	0.0106	0.0926	143
1000	899	0.0021	0.0729	260	0.0021	0.0729	260
3000	2697	0.0013	0.0591	443	0.0013	0.0591	443
5000	4500	0.0001	0.0309	997	0.0001	0.0309	997

Table 5.6: Estimation results of Lynden-Bell based (left pannel) and Woodroofe based (right pannel) estimators of the shape parameter $\gamma_1 = 0.8$ of Burr's model through 1000 right-truncated samples with 10%-truncation rate.

5.4 Proofs

5.4.1 Proof Theorem 5.1

For $x \geq x_0$ we have

$$\mathbf{F}_n^{(\mathbf{W})}\left(X_{n-k:n}x\right) = \exp\left\{-\int_{X_{n-k:n}x}^{\infty} \frac{dF_n\left(y\right)}{C_n\left(y\right)}\right\}.$$

We show that the latter exponent is negligible in probability uniformly over $x \geq x_0$. Indeed, note that both $F_n\left(y\right)/F\left(y\right)$ and $C\left(y\right)/C_n\left(y\right)$ are stochastically bounded from above on $y < X_{n:n}$ (see, e.g., [80] page 415 and [84], respectively), it follows that

$$-\int_{X_{n-k:n}x}^{\infty} \frac{dF_n\left(y\right)}{C_n\left(y\right)} = O_{\mathbf{P}}\left(1\right)\int_{X_{n-k:n}x}^{\infty} \frac{dF\left(y\right)}{C\left(y\right)}. \tag{5.4}$$

By a change of variables we have

$$\int_{X_{n-k:n}x}^{\infty} \frac{d\overline{F}\left(y\right)}{C\left(y\right)} = \frac{\overline{F}\left(X_{n-k:n}\right)}{C\left(X_{n-k:n}\right)}\left(\int_{x}^{\infty} \frac{C\left(X_{n-k:n}\right)}{C\left(X_{n-k:n}t\right)} d\frac{\overline{F}\left(X_{n-k:n}t\right)}{\overline{F}\left(X_{n-k:n}\right)}\right). \tag{5.5}$$

Recall that $X_{n-k:n} \overset{\mathbf{P}}{\to} \infty$ and that $\overline{F}$ is regularly varying at infinity with index $-1/\gamma$. On the other hand, from Assertion (i) of Lemma A2 [6] we deduce that $1/C$ is also regularly varying at infinity with index $1/\gamma_2$. Thus, we may apply Potters inequalities, see e.g. Proposition B.1.10 in [20], to both $\overline{F}$ and $1/C$ to write: for all large N, any $t \geq x_0$ and any sufficiently small $\delta, \nu > 0$, with large probability,

$$\left|\frac{\overline{F}\left(X_{n-k:n}t\right)}{\overline{F}\left(X_{n-k:n}\right)} - t^{-1/\gamma}\right| < \delta t^{-1/\gamma \pm \nu} \text{ and } \left|\frac{C\left(X_{n-k:n}\right)}{C\left(X_{n-k:n}t\right)} - t^{1/\gamma_2}\right| < \delta t^{1/\gamma_2 \pm \nu}, \tag{5.6}$$

where $t^{\pm a} := \max\left(t^a, t^{-a}\right)$. These two inequalities may be rewritten, into

$$\frac{\overline{F}\left(X_{n-k:n}t\right)}{\overline{F}\left(X_{n-k:n}\right)} = t^{-1/\gamma}\left(1 + o_{\mathbf{P}}\left(t^{\pm\nu}\right)\right) \text{ and } \frac{C\left(X_{n-k:n}\right)}{C\left(X_{n-k:n}t\right)} = t^{1/\gamma_2}\left(1 + o_{\mathbf{P}}\left(t^{\pm\nu}\right)\right),$$

uniformly on $t \geq x_0$. This leads to

$$\int_x^\infty \frac{C(X_{n-k:n})}{C(X_{n-k:n}t)} d\frac{\overline{F}(X_{n-k:n}t)}{\overline{F}(X_{n-k:n})} = -\frac{\gamma_1}{\gamma} x^{-1/\gamma_1} \left(1 + o_{\mathbf{P}}\left(x^{\pm\nu}\right)\right). \tag{5.7}$$

In view of (5.1), Benchaira et al. 2016a, [6] showed, in Lemma A1, that $\overline{F}(y) = (1 + o(1)) c_1 y^{-1/\gamma}$ and $\overline{G}(y) = (1 + o(1)) c_2 y^{-1/\gamma_2}$ as $y \to \infty$, for some constants $c_1, c_2 > 0$. In other words, $\mathbb{U}_F(s) = (1 + o(1)) (c_1 s)^\gamma$ as $s \to \infty$, and $C(y) = (1 + o(1)) c_2 y^{-1/\gamma_2}$ as $y \to \infty$. On the other hand, from Lemma A4 in [6], we have $X_{n-k:n} = (1 + o_{\mathbf{P}}(1)) \mathbb{U}_F(n/k)$, it follows that $X_{n-k:n} = (1 + o_{\mathbf{P}}(1)) c_1^\gamma (k/n)^{-\gamma}$. Note that $1 - \gamma/\gamma_2 = \gamma/\gamma_1$, hence

$$\frac{\overline{F}(X_{n-k:n})}{C(X_{n-k:n})} = (1 + o_{\mathbf{P}}(1)) c_1^{\gamma/\gamma_2} c_2^{-1} (k/n)^{\gamma/\gamma_1}. \tag{5.8}$$

Plugging results (5.7) and (5.8) in equation (5.5) yields

$$\int_{X_{n-k:n}x}^\infty \frac{d\overline{F}(y)}{C(y)} = (k/n)^{\gamma/\gamma_1} c_1^{\gamma/\gamma_2} c_2^{-1} \gamma_1 x^{-1/\gamma_1} \left(1 + o_{\mathbf{P}}\left(x^{\pm\nu}\right)\right). \tag{5.9}$$

By combining equations (5.4) and (5.9), we obtain

$$\int_{X_{n-k:n}x}^\infty \frac{dF_n(y)}{C_n(y)} = O_{\mathbf{P}}(1) (k/n)^{\gamma/\gamma_1} x^{-1/\gamma_1} \left(1 + o_{\mathbf{P}}\left(x^{\pm\nu}\right)\right), \tag{5.10}$$

which obviously tends to zero in probability (uniformly on $x \geq x_0$). We may now apply Taylor's expansion $e^t = 1 + t + O(t^2)$, as $t \to 0$, to get

$$\exp\left\{-\int_{X_{n-k:n}x}^\infty \frac{dF_n(y)}{C_n(y)}\right\} = 1 - \int_{X_{n-k:n}x}^\infty \frac{dF_n(y)}{C_n(y)} + O_{\mathbf{P}}\left(\int_{X_{n-k:n}x}^\infty \frac{dF_n(y)}{C_n(y)}\right)^2, \ N \to \infty.$$

In other words, we have

$$\overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n}x) = \int_{X_{n-k:n}x}^\infty \frac{dF_n(y)}{C_n(y)} + R_{n1}(x), \ N \to \infty, \tag{5.11}$$

where $R_{n1}(x) := O_{\mathbf{P}}\left((k/n)^{2\gamma/\gamma_1}\right) x^{-2/\gamma_1}\left(1 + o_{\mathbf{P}}\left(x^{\pm\nu}\right)\right)$. Next, we show that

$$\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}x) = \int_{X_{n-k:n}x}^{\infty} \frac{dF_n(y)}{C_n(y)} + R_{n2}(x), \ N \to \infty. \tag{5.12}$$

Observe that, by taking the logarithme then its exponential in the definition of $\mathbf{F}_n^{(\mathbf{LB})}(x)$, we have

$$\mathbf{F}_n^{(\mathbf{LB})}(X_{n-k:n}x) = \exp\left\{\sum_{i=1}^{n} \mathbf{1}(X_{i:n} > X_{n-k:n}x)\log\left(1 - \frac{1}{nC_n(X_{i:n})}\right)\right\},$$

which may be rewritten into $\exp\left\{n\int_x^\infty \log\left(1 - \dfrac{1}{nC_n(X_{n-k:n}y)}\right) dF_n(X_{n-k:n}y)\right\}$. To get approximation (5.12) it suffices to apply successively, in the previous quantity, Taylor's expansions $e^t = 1 + t + O(t^2)$ and $\log(1-t) = -t + O(t^2)$ (as $t \to 0$) with similar arguments as above (we omit further details). Combining (5.11) and (5.12) and setting $R_n(x) := R_{n1}(x) - R_{n2}(x)$ yield

$$\overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n}x) - \overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}x) = R_n(x), \ N \to \infty. \tag{5.13}$$

On the other hand, by once again using Taylor's expansion, we write

$$\overline{\mathbf{F}}(X_{n-k:n}) = \int_{X_{n-k:n}}^{\infty} \frac{dF(y)}{C(y)} + \widetilde{R}_n(x), \ N \to \infty.$$

From equation (5.9), we infer that $\overline{\mathbf{F}}(X_{n-k:n}) = c_2^{-1} c_1^{1-\gamma/\gamma_1}(k/n)^{\gamma/\gamma_1}(1 + o_{\mathbf{P}}(1))$, which implies, in view of (5.13), that

$$x^{1/\gamma_1}\frac{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}x) - \overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n}x)}{\overline{\mathbf{F}}(X_{n-k:n})} = O_{\mathbf{P}}\left((k/n)^{\gamma/\gamma_1}\right) x^{-1/\gamma_1 \pm \nu}.$$

Observe now that, for a sufficiently small $\nu > 0$, we have $x^{-1/\gamma_1 \pm \nu} = O_{\mathbf{P}}(1)$, uniformly on $x \geq x_0 > 0$, as sought.

5.4.2 Proof Theorem 5.2

In a similar way to what is done with $\mathbf{D}_n^{(\mathbf{W})}(x)$, in the proof of Theorem 2.1 in [6], we decompose $k^{-1/2}\mathbf{D}_n^{(\mathbf{LB})}(x)$ into the sum of

$$\mathbf{N}_{n1}(x) := x^{-1/\gamma_1} \frac{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}x) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})},$$

$$\mathbf{N}_{n2}(x) := -\frac{\overline{\mathbf{F}}(X_{n-k:n}x)}{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n})} \frac{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})},$$

$$\mathbf{N}_{n3}(x) := \left(\frac{\overline{\mathbf{F}}(X_{n-k:n}x)}{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n})} - x^{-1/\gamma_1} \right) \frac{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n}x) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n}x)},$$

and $\mathbf{N}_{n4}(x) := \overline{\mathbf{F}}(X_{n-k:n}x) / \overline{\mathbf{F}}(X_{n-k:n}) - x^{-1/\gamma_1}$. If we let

$$\mathbf{M}_{n1}(x) := x^{-1/\gamma_1} \frac{\overline{\mathbf{F}}_n^{(\mathbf{W})}(X_{n-k:n}x) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})},$$

then, by applying Theorem 5.1, we have $x^{1/\gamma_1}\mathbf{N}_{n1}(x) = x^{1/\gamma_1}\mathbf{M}_{n1}(x) + x^{-1/\gamma_1} O_{\mathbf{P}}\left((k/n)^{\gamma/\gamma_1}\right)$, uniformly on $x \geq x_0$. By assumption we have $k^{1+\gamma_1/(2\gamma)}/n \overset{\mathbf{P}}{\to} 0$, which is equivalent to $\sqrt{k}(k/n)^{\gamma/\gamma_1} \overset{\mathbf{P}}{\to} 0$ as $N \to \infty$, therefore

$$x^{1/\gamma_1}\sqrt{k}\mathbf{N}_{n1}(x) = x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n1}(x) + o_{\mathbf{P}}\left(x^{-1/\gamma_1}\right). \tag{5.14}$$

In view of this representation we show that, both $\mathbf{D}_n^{(\mathbf{W})}(x)$ and $\mathbf{D}_n^{(\mathbf{LB})}(x)$ are (weakly) approximated, in the probability space $(\Omega, \mathcal{A}, \mathbf{P})$, by the same Gaussian process $\mathbf{\Gamma}(x; \mathbf{W})$ given in (5.2). Indeed, for a sufficiently small $\epsilon > 0$, and $0 < \eta < 1/2$, [6] (see the beginning of the proof of Theorem 2.1 therein), showed that

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n1}(x) = \Phi(x) + o_{\mathbf{P}}\left(x^{(1-\eta)/\gamma \pm \epsilon}\right),$$

where $\Phi(x) := x^{1/\gamma}\left\{\dfrac{\gamma}{\gamma_1}\mathbf{W}\left(x^{-1/\gamma}\right)+\dfrac{\gamma}{\gamma_1+\gamma_2}\int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}\left(x^{-1/\gamma}t\right)dt\right\}$. Then by using representation (5.14), we get $x^{1/\gamma_1}\sqrt{k}\mathbf{N}_{n1}(x)=\Phi(x)+o_{\mathbf{P}}\left(x^{-1/\gamma_1}\right)+o_{\mathbf{P}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right)$. In particular for $x=1$, we have

$$\sqrt{k}\left(\frac{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})}-1\right)=\sqrt{k}\mathbf{N}_{n1}(1)=\Phi(1)+o_{\mathbf{P}}(1), \tag{5.15}$$

leading to $\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n})/\overline{\mathbf{F}}(X_{n-k:n})\overset{\mathbf{P}}{\to}1$, as $N\to\infty$. By applying Potters inequalities to $\overline{\mathbf{F}}$ (as it was done for $\overline{F}$ in (5.8)) together with the previous limit, we obtain

$$\frac{\overline{\mathbf{F}}(X_{n-k:n}x)}{\overline{\mathbf{F}}_n^{(\mathbf{LB})}(X_{n-k:n})}=\left(1+O_{\mathbf{P}}\left(x^{\pm\epsilon}\right)\right)x^{-1/\gamma_1}. \tag{5.16}$$

By combining (5.15) and(5.16), we get $x^{1/\gamma_1}\sqrt{k}\mathbf{N}_{n2}(x)=-\Phi(1)+o_{\mathbf{P}}\left(x^{\pm\epsilon}\right)$. For the third term $\mathbf{N}_{n3}(x)$, we use similar arguments to show that

$$x^{1/\gamma_1}\sqrt{k}\mathbf{N}_{n3}(x)=o_{\mathbf{P}}\left(x^{-1/\gamma_1\pm\epsilon}\right)+o_{\mathbf{P}}\left(x^{-1/\gamma_1+(1-\eta)/\gamma\pm\epsilon}\right).$$

Observe that $x^{1/\gamma_1-(1-\eta_0)/\gamma}o_{\mathbf{P}}\left(x^{-1/\gamma_1\pm\epsilon}\right)$ and $x^{1/\gamma_1-(1-\eta_0)/\gamma}o_{\mathbf{P}}\left(x^{-1/\gamma_1+(1-\eta)/\gamma\pm\epsilon}\right)$ respectively equal $o_{\mathbf{P}}\left(x^{-(1-\eta_0)/\gamma\pm\epsilon}\right)$ and $o_{\mathbf{P}}\left(x^{(\eta-\eta_0)/\gamma\pm\epsilon}\right)$, for $\gamma/\gamma_2<\eta_0<\eta<1/2$, and that both the last two quantities are equal to $o_{\mathbf{P}}(1)$ for any small $\epsilon>0$ and $x\geq x_0>0$. Finally, by following the same steps at the end of the proof of Theorem 2.1 in [6], we get

$$\sqrt{k}\mathbf{N}_{n4}(x)=x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1}-1}{\gamma_1\tau_1}\sqrt{k}\mathbf{A}_0(n/k)+o_{\mathbf{p}}\left(x^{-1/\gamma_1+(1-\eta)/\gamma\pm\epsilon}\right).$$

Consequently, we have

$$x^{1/\gamma_1-(1-\eta_0)/\gamma}\left\{\mathbf{D}_n^{(\mathbf{LB})}(x)-\mathbf{\Gamma}(x;\mathbf{W})-x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1}-1}{\gamma_1\tau_1}\sqrt{k}\mathbf{A}_0(n/k)\right\}=o_{\mathbf{P}}(1),$$

uniformly over $x\geq x_0$. Recall that $1/\gamma_1=1/\gamma-1/\gamma_2$, then letting $\eta_0:=1/2-\xi$ yields $0<\xi<1/2-\gamma/\gamma_2$ and achieves the proof.

5.4.3 Proof of Theorem 5.3

The proof is similar, mutatis mutandis, as that of Corollary 3.1 in [6]. Therefore we omit the details.

Concluding notes.

1. Therougout the simulation, we may conclude that are equivalent
2. From theoritical poin of view k/n more strong assumption $k^{1+\gamma_1/(2\gamma)}/n$.

Chapter 6

Kernel weighted moment estimator of extreme quantile of complete data

In this thapter we introduce a new estimator for the extreme quantile but in the case of complete data, we use in this estimation the kernel weighted moment. Our consederations are based on the results of Caeiro and Gomes 2015 in [11].

6.1 Introduction

Let $X_1, ..., X_n$ be n $(n \geq 1)$ independent copies of a non-negative random variable (rv) X, defined over some probability space (Ω, A, P), with continuous cumulative distribution function (cdf) F. Note $X_{1:n}, ..., X_{n:n}$ the order statistics of X. We assume that $\overline{F} := 1 - F$ hase a right Pareto-type i,e

$$\overline{F}(x) \sim (x/C)^{-1/\gamma}, \ x \to \infty$$

with $\gamma > 0$ and $C > 0$ denoting the shape and scale parameters, respectively. Then the quantile function $U(t) := F^{\leftarrow}(1 - 1/t) = \inf\{x : F(x) \geq 1 - 1/t\}$, $t > 1$ is a regularly varying function with a positive index of regular variation equal to γ, i.e.,

$$\lim_{t\to\infty} U(tx)/U(t) = x^{\gamma}. \tag{6.1}$$

According Caeiro and Gomes 2015 [11], the probability weighted moments (PWM) method is a generalization of the method of moments. The PWM of a rv X are defined by $M_{p,r,s} := E\left[X^p F^r(X)\overline{F}^s(X)\right]$, with $p, r, s \in \mathbb{R}$, and the usualy work with one of the moments $a_r := M_{1,0,r}$ or $b_r := M_{1,r,0}$. Caeiro and Gomes 2015 introduce a new semi-parametric estimators Pareto log PWM (PLPWM) based on the log moments $l_r := E\left[(\ln X)\left(1 - F\left(X\right)\right)^r\right]$ for non-negative integer r. For the strict Pareto model with d.f. $F(x) = 1 - (x/C)^{-1/\gamma}$, $x > C > 0$, $\gamma > 0$ the PLPWM are

$$l_r = \ln(C)/(1+r) + \gamma/(1+r)^2. \tag{6.2}$$

Next, we introduce a kernel PLPWM estimator, we write

$$l_r = -\int \ln x \left(\overline{F}\left(x\right)\right)^r d\left[\overline{F}\left(x\right) K\left(\overline{F}\left(x\right)\right)\right]$$

where K is a function will be called kernel, satisfiying:

(1)K is nonincreasing and right-continuous;

(2)$K(s) = 0$ for $s \notin [0,1)$ and $K\left(s\right) \geq 0$, for $s \in [0,1)$;

(3)$\int_0^{\infty} K\left(s\right) ds = 1$;

(4)K and their Lebesgue derivatives K' and K" are bounded on $\mathbb{R}$. In the extreme quantile we write

$$\begin{aligned} l_r &= -\int_{t_-}^{\infty} \ln\frac{x}{t}\left(\frac{\overline{F}\left(x_-\right)}{\overline{F}\left(t_-\right)}\right)^r d\left[\frac{\overline{F}\left(x_-\right)}{\overline{F}\left(t_-\right)} K\left(\frac{\overline{F}\left(x_-\right)}{\overline{F}\left(t_-\right)}\right)\right] \\ &= \int_{t}^{\infty} \ln\frac{x}{t}\left(\frac{\overline{F}\left(x_-\right)}{\overline{F}\left(t_-\right)}\right)^r g\left(\frac{\overline{F}\left(x_-\right)}{\overline{F}\left(t_-\right)}\right)\frac{dF\left(x_-\right)}{\overline{F}\left(t_-\right)} \end{aligned}$$

where $g\left(x\right) := \left[xK\left(x\right)\right]'$. Then, $X_{n-k:n}$ tends almost surely to ∞ where $k := k_n$ is an integer sequence satisfying

$$1 < k < n,\ k \to \infty \text{ and } k/n \to 0 \text{ as } n \to \infty. \tag{6.3}$$

By replacing t by $X_{n-k:n}$ and F by F_n (the well know empirical estimator of F) in the last formula of l_r we get

$$\widehat{l}_r = \frac{1}{n\overline{F}_n\left(X_{n-k:n}\right)} \sum_{i=n-k+1}^{n} \ln \frac{X_{i:n}}{X_{n-k:n}} \left(\frac{\overline{F}_n\left(X_{i:n}\right)}{\overline{F}_n\left(X_{n-k:n}\right)}\right)^r g\left(\frac{\overline{F}_n\left(X_{i:n}\right)}{\overline{F}_n\left(X_{n-k:n}\right)}\right),$$

by replacing $n-i+1$ par i and $\overline{F}_n\left(X_{n-i:n}\right) = i/n$, for $i = 1, 2, ..., k$, then

$$\widehat{l}_r = \frac{1}{k} \sum_{i=1}^{k} \ln \frac{X_{n-i+1:n}}{X_{n-k:n}} \left(\frac{i}{k+1}\right)^r g\left(\frac{i}{k+1}\right).$$

According (6.2) we can write

$$\gamma = 2\left(l_0 - 2l_1\right),\ C = \left(\frac{k}{n}\right)^{\gamma} \exp\left(4l_1 - l_0\right) \text{ and } q_p = \left(\frac{k}{np}\right)^{\gamma} \exp\left(4l_1 - l_0\right). \tag{6.4}$$

Our estimators are

$$\widehat{\gamma} := 2\frac{1}{k} \sum_{i=1}^{k} \ln \frac{X_{n-i+1:n}}{X_{n-k:n}} g\left(\frac{i}{k+1}\right)\left(1 - 2\frac{i}{k+1}\right),$$

$$\widehat{C}_{k,n} := \left(\frac{k}{n}\right)^{\widehat{\gamma}} \exp\left\{\frac{1}{k} \sum_{i=1}^{k} \ln \frac{X_{n-i+1:n}}{X_{n-k:n}} g\left(\frac{i}{k+1}\right)\left(4\frac{i}{k+1} - 1\right)\right\}$$

and

$$\widehat{Q}_{k,n}\left(p\right) := \left(\frac{k}{np}\right)^{\widehat{\gamma}} \exp\left\{\frac{1}{k} \sum_{i=1}^{k} \ln \frac{X_{n-i+1:n}}{X_{n-k:n}} g\left(\frac{i}{k+1}\right)\left(4\frac{i}{k+1} - 1\right)\right\}.$$

6.2 Main results

We need the second order regular variation condition with a parameter $\rho \leq 0$, that measures the rate of convergence in (6.1) and is given by

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho} \Leftrightarrow \lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho} \tag{6.5}$$

for all $x > 0$, with $|A|$ a regular varying function with index ρ and $\frac{x^\rho - 1}{\rho} = \ln x$ if $\rho = 0$.

Theorem 6.1 *Under the second order regular variation condition (6.5), and for intermediate $k = k_n$ satisfying (6.3),*

$$\begin{aligned}\sqrt{k}\left(\widehat{\gamma} - \gamma\right) &= \gamma \int_0^1 s^{-1} W_n(s)\,\varphi(s)\,ds - 4\gamma W_n(1) \int_0^1 sK(s)\,ds \\ &\quad + \sqrt{k}\widetilde{A}(n/k) \int_0^1 \left\{2s^{-\tau} + \frac{4(-\tau+1)}{\tau} s^{-\tau+1} - \frac{4}{\tau} s\right\} K(s)\,ds + o_p(1),\end{aligned}$$

provided that $\sqrt{k}\widetilde{A}(n/k) = O(1)$, where $\varphi(s) := 2\psi_0'(s) - 4\psi_1'(s)$ and $\psi_r(s) := \int_0^s g(t)\,t^r dt$. If in addition we suppose that $\sqrt{k}\widetilde{A}(n/k) = \lambda$, then $\sqrt{k}\left(\widehat{\gamma} - \gamma\right) \xrightarrow{D} N(\mu, \sigma^2)$, as $n \to \infty$, where

$$\mu := \lambda \int_0^1 \left\{2s^{-\tau} + \frac{4(-\tau+1)}{\tau} s^{-\tau+1} - \frac{4}{\tau} s\right\} K(s)\,ds$$

and

$$\sigma^2 := \gamma^2 \left\{16\left(\int_0^1 sK(s)\,ds\right)^2 + 2\int_0^1 s^{-1} \int_0^s \varphi(t)\,dt\varphi(s)\,ds - 32\left(\int_0^1 sK(s)\,ds\right)^3\right\}.$$

Remark 6.1 *If $K = 1$ we get $\mu = 2/(-\tau+1)(-\tau+2)$ and $\sigma^2 = (2\gamma^2)/3$ are the same as in caeiro 2015.*

Theorem 6.2 *Under the conditions of Theoreme 6.1, if $p = p_n$ is a sequence of probabil-*

ities such that $c_n := k/(np) \to \infty$, *as* $n \to \infty$, *then*

$$\frac{\sqrt{k}}{\ln c_n}\left(\frac{\widehat{Q}_{k,n}(p)}{q_p}-1\right) \overset{D}{=} \sqrt{k}\left(\widehat{\gamma}-\gamma\right)\left(1+o_p(1)\right)+o_p(1).$$

6.3 Proofs

6.3.1 Proof of theorem 6.1

We have

$$\begin{aligned}
\widehat{l}_r &= -\int_1^\infty \ln x g\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right)\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right)^r d\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right)\\
&= -\int_1^\infty \ln x \psi_r'\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right) d\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right)\\
&= -\int_1^\infty \ln x d\psi_r\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right)\\
&= \int_1^\infty x^{-1}\psi_r\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right)dx.
\end{aligned}$$

Then,

$$l_r = \int_1^\infty x^{-1}\psi\left(x^{-1/\gamma}\right)dx.$$

Hence

$$\sqrt{k}\left(\widehat{l}_r - l_r\right) = \sqrt{k}\int_1^\infty x^{-1}\left[\psi_r\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}\right)-\psi_r\left(x^{-1/\gamma}\right)\right]dx.$$

Let $\mathbf{D}_n(x) := \sqrt{k}\left(\frac{\overline{F}_n(xX_{n-k:n})}{\overline{F}_n(X_{n-k:n})}-x^{-1/\gamma}\right)$, be the tail product-limit process, then Taylor's expansion of ψ_r yields that

$$\sqrt{k}\left(\widehat{l}_r - l_r\right) = \int_1^\infty x^{-1}\mathbf{D}_n(x)\,\psi_r'\left(x^{-1/\gamma}\right)dx + R_{n1},$$

where $R_{n1} = 2^{-1}k^{-1/2}\int_1^\infty x^{-1}\mathbf{D}_n^2(x)\,\psi_r''(\zeta)\,dx$ and ζ is between the minimum and the maximum of $\overline{F}_n(xX_{n-k:n})/\overline{F}_n(X_{n-k:n})$ and $x^{-1/\gamma}$,

$$|R_{n1}| \leq \sup_{x\geq 1}\left|\mathbf{D}_n^2(x)\right| 2^{-1}k^{-1/2}\int_1^{\infty} x^{-1}\left|\psi_r''(\zeta)\right| dx,$$

In the fact that $\sup_{x\geq 1}\left|\mathbf{D}_n^2(x)\right| = O_p(1)$ and from assumption (4), we infer that ψ_r'' is bounded on $(0,1)$. Consequently, we have $R_{n1} = o_p(1)$.

Then by combining the theorem 2.4.8 and 5.1.4 in de Haan and Ferreira 2006, we infer that under the second-order regular variation condition (6.5), there exist a constant $\tau \leq 0$, a function $\widetilde{A}(t) \sim A(t)$ at infinity with $\sqrt{k}\widetilde{A}(n/k) = O(1)$ and a sequence of Brownian motions $\{W_n(s); 0 \leq s \leq 1\}$ such that for all $x_0 > 0$

$$\sup_{x\geq x_0} x^{(1/2-\epsilon)/\gamma}\left|\mathbf{D}_n(x) - \left\{W_n\left(x^{-1/\gamma}\right) - x^{-1/\gamma}W_n(1)\right\}\right. \tag{6.6}$$

$$\left.-\sqrt{k}\widetilde{A}(n/k)\, x^{-1/\gamma}\frac{x^{\tau/\gamma}-1}{\tau\gamma}\right| \overset{P}{\to} 0, \text{ as } n \text{ tends to } \infty, \tag{6.7}$$

for any $\varepsilon > 0$. Then

$$\begin{aligned}\sqrt{k}\left(\widehat{l}_r - l_r\right) &= \int_1^{\infty} x^{-1}\left\{W_n\left(x^{-1/\gamma}\right) - x^{-1/\gamma}W_n(1)\right\}\psi_r'\left(x^{-1/\gamma}\right)dx \\ &\quad + \sqrt{k}\widetilde{A}(n/k)\int_1^{\infty} x^{-1-1/\gamma}\frac{x^{\tau/\gamma}-1}{\gamma\tau}\psi_r'\left(x^{-1/\gamma}\right)dx \\ &\quad + R_{n1} + R_{n2},\end{aligned}$$

where $R_{n2} := \int_1^{\infty} x^{-1}\left(\mathbf{D}_n(x) - \left\{W_n\left(x^{-1/\gamma}\right) - x^{-1/\gamma}W_n(1)\right\} - \sqrt{k}\widetilde{A}(n/k)\, x^{-1/\gamma}\frac{x^{\tau/\gamma}-1}{\tau\gamma}\right)\psi_r'\left(x^{-1/\gamma}\right)dx$, we have that

$$|R_{n2}| \leq \sup_{x\geq x_0} x^{(1/2-\epsilon)/\gamma}\left|\mathbf{D}_n(x) - \left\{W_n\left(x^{-1/\gamma}\right) - x^{-1/\gamma}W_n(1)\right\} - \sqrt{k}\widetilde{A}(n/k)\, x^{-1/\gamma}\frac{x^{\tau/\gamma}-1}{\tau\gamma}\right|$$
$$\int_1^{\infty} x^{-1-(1/2-\epsilon)/\gamma}\left|\psi_r'\left(x^{-1/\gamma}\right)\right| dx,$$

by using (6.6) since ψ_r' is bounded on $(0,1)$ and $\int_1^{\infty} x^{-1-(1/2-\epsilon)/\gamma}dx$ is finite, then $R_{n2} =$

$o_p(1)$

For the fact that

$$\sqrt{k}\left(\widehat{\gamma}-\gamma\right)=2\sqrt{k}\left(\widehat{l_0}-l_0\right)-4\sqrt{k}\left(\widehat{l_1}-l_1\right). \tag{6.8}$$

Hence

$$\begin{aligned}\sqrt{k}\left(\widehat{\gamma}-\gamma\right)=&\int_1^\infty x^{-1}\left\{W_n\left(x^{-1/\gamma}\right)-x^{-1/\gamma}W_n\left(1\right)\right\}\varphi\left(x^{-1/\gamma}\right)dx\\&+\sqrt{k}\widetilde{A}\left(n/k\right)\int_1^\infty x^{-1-1/\gamma}\frac{x^{\tau/\gamma}-1}{\gamma\tau}\varphi\left(x^{-1/\gamma}\right)dx+0_p\left(1\right).\end{aligned}$$

Then by using the change of variable $s=x^{-1/\gamma}$, we get

$$\begin{aligned}\sqrt{k}\left(\widehat{\gamma}-\gamma\right)=&\,\gamma\int_0^1 s^{-1}\left\{W_n\left(s\right)-sW_n\left(1\right)\right\}\varphi\left(s\right)ds\\&+\sqrt{k}\widetilde{A}\left(n/k\right)\int_0^1\frac{s^{-\tau}-1}{\tau}\varphi\left(s\right)ds+0_p\left(1\right).\end{aligned}$$

By calculing we get $\int_0^1\varphi\left(s\right)ds=4\int_0^1 sK\left(s\right)ds$ and

$$\int_0^1\frac{s^{-\tau}-1}{\tau}\varphi\left(s\right)ds=\left\{2\int_0^1 s^{-\tau}K\left(s\right)ds+\frac{4\left(-\tau+1\right)}{\tau}\int_0^1 s^{-\tau+1}K\left(s\right)ds-\frac{4}{\tau}\int_0^1 sK\left(s\right)ds\right\}.$$

Then

$$\begin{aligned}\sqrt{k}\left(\widehat{\gamma}-\gamma\right)=&\,\gamma\left\{\int_0^1 s^{-1}W_n\left(s\right)\varphi\left(s\right)ds-4W_n\left(1\right)\int_0^1 sK\left(s\right)ds\right\}\\&+\sqrt{k}\widetilde{A}\left(n/k\right)\left\{\int_0^1\left\{2s^{-\tau}+\frac{4\left(-\tau+1\right)}{\tau}s^{-\tau+1}-\frac{4}{\tau}s\right\}K\left(s\right)ds\right\}+0_p\left(1\right).\end{aligned}$$

Finally, if $\sqrt{k}\widetilde{A}\left(n/k\right)\to\lambda$ as $n\to\infty$, $\sqrt{k}\left(\widehat{\gamma}-\gamma\right)\overset{D}{\to}N\left(\mu,\sigma^2\right)$, where

$$\mu=\lambda\left\{\int_0^1\left\{2s^{-\tau}+\frac{4\left(-\tau+1\right)}{\tau}s^{-\tau+1}-\frac{4}{\tau}s\right\}K\left(s\right)ds\right\}$$

and

$$\sigma^2 = \gamma^2 E\left[\int_0^1 s^{-1} W_n(s)\varphi(s)\,ds - 4W_n(1)\int_0^1 sK(s)\,ds\right]^2,$$

by calculing σ^2 we get the asymptotic bias as sought.

6.3.2 Proof of theorem 6.2

We have that

$$\begin{aligned}\frac{\widehat{Q}_{k,n}(p)}{q_p} &= c_n^{(\widehat{\gamma}-\gamma)} \exp\left\{4\left(\widehat{l}_1 - l_1\right) - \left(\widehat{l}_0 - l_0\right)\right\}\\ &= \exp\left\{\ln c_n^{(\widehat{\gamma}-\gamma)} + 4\left(\widehat{l}_1 - l_1\right) - \left(\widehat{l}_0 - l_0\right)\right\}\\ &= \exp\left\{(\widehat{\gamma}-\gamma)\ln c_n + 4\left(\widehat{l}_1 - l_1\right) - \left(\widehat{l}_0 - l_0\right)\right\}.\end{aligned}$$

Then

$$\frac{\widehat{Q}_{k,n}(p)}{q_p} - 1 = \exp\left\{(1-\ln c_n)\,4\left(\widehat{l}_1 - l_1\right) + (2\ln c_n - 1)\left(\widehat{l}_0 - l_0\right)\right\} - \exp(0)$$

then Taylor's expansion of exp yields that

$$\frac{\sqrt{k}}{\ln c_n}\left(\frac{\widehat{Q}_{k,n}(p)}{q_p} - 1\right) = \frac{1-\ln c_n}{\ln c_n}4\sqrt{k}\left(\widehat{l}_1 - l_1\right) + \frac{2\ln c_n - 1}{\ln c_n}\sqrt{k}\left(\widehat{l}_0 - l_0\right) + R,$$

where

$$R := 2^{-1}k^{-1/2}\left\{\frac{1-\ln c_n}{\ln c_n}4\sqrt{k}\left(\widehat{l}_1 - l_1\right) + \frac{2\ln c_n - 1}{\ln c_n}\sqrt{k}\left(\widehat{l}_0 - l_0\right)\right\}^2 \exp(\zeta)$$

and ζ is between the minimum and the maximum of $\left\{(1-\ln c_n)\,4\left(\widehat{l}_1 - l_1\right) + (2\ln c_n - 1)\left(\widehat{l}_0 - l_0\right)\right\}$ and 0, in the fact that the bouth of $\sqrt{k}\left(\widehat{l}_1 - l_1\right)$ and $\sqrt{k}\left(\widehat{l}_0 - l_0\right)$ are Gaussian process,

$R = O_p\left(2^{-1}k^{-1/2}\right)$, then $R = o_p(1)$. Since, $\lim_{c_n\to\infty} \frac{1-\ln c_n}{\ln c_n} = -1$ and $\lim_{c_n\to\infty} \frac{2\ln c_n - 1}{\ln c_n} = 2$, so

$$\begin{aligned}\frac{\sqrt{k}}{\ln c_n}\left(\frac{\widehat{Q}_{k,n}(p)}{q_p} - 1\right) &= (-1 + o_p(1))\, 4\sqrt{k}\left(\widehat{l}_1 - l_1\right) + (2 + o_p(1))\sqrt{k}\left(\widehat{l}_0 - l_0\right) + o_p(1) \\ &= \left\{2\sqrt{k}\left(\widehat{l}_0 - l_0\right) - 4\sqrt{k}\left(\widehat{l}_1 - l_1\right)\right\}(1 + o_p(1)) + o_p(1).\end{aligned}$$

Finally, according (6.8) we get the result as sought.

Conclusion

This thesis is divided into four distinct parts, to which is added one introduction. In the first part, we recalled the areas where incomplete data (censored and truncated) and to facilitate the reading of the document, it was recalled First chapter some basic notions of survival analysis then the theory of extreme value and the estimation of the tail index, extreme quantiles and second order parameter.

Then, we propose a consistent estimator of the second-order parameter of Pareto-type distributions under random right-truncation and establish its asymptotic normality. Our considerations are based on a useful Gaussian approximation of a tail product-limit process given recently by Benchaira *et al.* [Tail product-limit process for truncated data with application to extreme value index estimation. *Extremes,* 2016; 19: 219-251] and on the results of Gomes *et al.* [Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes,* 2003; **5**: 387-414]. We show, by simulation, that the proposed estimators behave well, in terms of bias and mean square error.

After that, by means of a Lynden-Bell integral with deterministic threshold, recently Worms and Worms [A Lynden-Bell integral estimator for extremes of randomly truncated data. Statist. Probab. Lett. 2016; 109: 106-117] introduced an asymptotically normal estimator of the tail index for Pareto-type (randomly right-truncated) data. In this context, we consider the random threshold case to derive a Hill-type estimator and establish its consistency and asymptotic normality. A simulation study is carried out to evaluate the finite sample behavior of the proposed estimator and compare it to the existing ones.

Finally, we estimate the quantile extreme that is based on the both of methods the kernel type and the log probability weighted moment of estimation, where it is based on the results of Caeiro and Gomes (2015) [11] (A log probability weighted moment estimator of extreme quantiles) which consider the semi parametric estimation of extreme quantiles of a right heavy-tail model and propose a new probability weighted moment estimator of extreme quantiles. Then, we will prove the consistency and asymptotic normality of our estimator.

Bibliography

[1] Aalen, O., 1976. Nonparametric inference in connection with multiple decrement models. Scand. J. Statist. 3, 15-27.

[2] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2006). Statistics of extremes : theory and applications. John Wiley.

[3] Beirlant, J., Guillou, A., Dierckx, G., and Fils-Villetard, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring. Extremes. 10(3), 151-174.

[4] Beirlant, J., Bardoutsos, A., de Wet, T., Gijbels, I., 2016. Bias reduced tail estimation for censored Pareto type distributions. *Statist. Probab. Lett.* **109,** 78-88.

[5] Benchaira, S., Meraghni, D., Necir, A., 2015. On the asymptotic normality of the extreme value index for right-truncated data. *Statist. Probab. Lett.* **107,** 378-384.

[6] Benchaira, S., Meraghni, D., Necir, A., 2016a. Tail product-limit process for truncated data with application to extreme value index estimation. *Extremes* **19**, 219-251.

[7] Benchaira, S., Meraghni, D., Necir, A., 2016b. Kernel estimation of the tail index of a right-truncated Pareto-type distribution. *Statist. Probab. Lett.* **119,** 186-193.

[8] Billingsley, P. (1995). Probability and Measure, 3rd edition. Wiley, New York.

[9] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). Regular Variation. Cambridge University Press, Cambridge.

[10] Breslow, N., and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. Ann. Statist., 2(3), 437 453.

[11] F. Caeiro and D., P., Gomes (2015). A Log Probability Weighted Moment Estimator of Extreme Quantiles. Springer Proceedings in Mathematics & Statistics.

[12] Caeiro, F., Gomes, M.I, 2015. Threshold Selection in Extreme Value Analysis. Chapter in: Dipak Dey and Jun Yan, *Extreme Value Modeling and Risk Analysis: Methods and Applications*, Chapman-Hall/CRC, ISBN 9781498701297, pp. 69-87.

[13] Caeiro, F., M. Gomes, and D. Pestana (2005). Direct reduction of bias of the classical hill estimator. Statistical Journal (3(2)), 113–136.

[14] Cheng, S. and L. Peng (2001). Confidence intervals for the tail index. Bernoulli7, 751–760.

[15] Ciuperca, G., Mercadier, C., 2010. Semi-parametric estimation for heavy tailed distributions. *Extremes* **13**, 55-87.

[16] Csörgő, S., Deheuvels, P., Mason, D., 1985. Kernel estimates of the tail index of a distribution. Ann. Statist 13,1050-1077.

[17] Csörgő, S. and Mason, D.M. (1985). Central Limit Theorems for Sums of Extreme Values. Mathematical Proceedings of the Cambridge Philosophical Society 98, 547–558.

[18] David H., A., and Nagaraja H., N. (2003). Order Statistics. John Wiley & Sons, Inc., Hoboken, New Jersey.

[19] Davis, R. and Resnick, S.I. (1984). Tail Estimates Motivated by Extreme Value Theory. Annals of Statistics 12, 1467–1487.

[20] de Haan, L., Ferreira, A., 2006. Extreme Value Theory: An Introduction. *Springer.*

[21] de Haan, L. and Peng, L. (1998). Comparison of Tail Index Estimators. Statistica Neerlandica 52, 60-70.

[22] de Haan, L., Stadtmüller, U., 1996. Generalized regular variation of second order. *J. Australian Math. Soc.* (Series A) **61**, 381-395.

[23] de Wolf, P.P. (1999). Estimating the Extreme Value Index. Ph.D. Thesis, Delft University of Technology.

[24] Deheuvels, P., Häeusler, E., and Mason, D.M. (1988). Almost sure convergence of the Hill estimator. Math. Proc. Cambridge Philos. Soc., 104(02). 371-381.

[25] Dekkers, A. L. M. and de Haan, L. (1989). On the Estimation of the Extreme-Value Index and Large Quantile Estimation. Annals of Statistics 17, 1795-1832.

[26] Dekkers, A. L. M., Einmahl, J. H. J. and de Haan, L. (1989). A Moment Estimator for the Index of an Extreme Value Index. Annals of Statistics. 17.1833-1855.

[27] Dekkers, A. L. M. and de Haan, L. (1993). Optimal Choice of Sample Fraction in Extreme-Value Estimation. Journal of Multivariate Analysis 47, 173-195.

[28] Deme, E., Gardes, L., Girard, S., 2013. On the estimation of the second order parameter for heavy-tailed distributions. *REVSTAT* **11**, 277-299.

[29] Diebolt, J., L. Gardes, S. Girard, and A. Guillou (2008b). Bias-reduced extreme quantiles estimators of weibull distributions. Journal of Statistical Planning and Inference138, 1389–1401.

[30] Drees, H.(1996). Refined Pickands Estimator of the Extreme Value Index. Annals of Statistics 23, 2059–2080.

[31] Einmahl, J. H., Fils-Villetard, A., and Guillou, A. (2008). Statistics of extremes under random censoring. Bernoulli. 14(1), 207-227.

[32] Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

[33] Escudero, F., Ortega, E., 2008. Actuarial comparisons for aggregate claims with randomly right-truncated claims. *Insurance Math. Econom.* 43: 255-262.

[34] Fleming, T. R., and Harrington, D. P. (1984). Nonparametric estimation of the survival distribution in censored data. Comm. Statist. Theory Methods, 13(20), 2469 2486.

[35] Fraga Alves, M. I. (1995). Estimation of the Tail Parameter in the Domain of Attraction of an Extremal Distribution. Journal of Statistical Planning and inference 45,143-173.

[36] Fraga Alves, M. I., de Haan, L., Lin, T., 2003. Estimation of the parameter controlling the speed of convergence in extreme value theory. *Math. Methods Statist.* **12**, 155-176.

[37] Fraga Alves, M. I., de Haan, L., Lin, T., 2006. Third order extented regular variation. *Pub. de l'institut Mathématique,* tome 80(94), 109-120.

[38] Ferreira, A., de Haan, L. and Peng, L. (2003). On Optimizing the Estimation of High Quantiles of a Probability Distribution. Statistics 37, 401-434.

[39] Feuerverger, A. and P. Hall (1999). Estimating a tail exponent by modeling departure from a pareto distribution. Annals of Statistics 27, 760–781.

[40] Fisher, R. A., and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Math. Proc. Cambridge Philos. Soc., 24(02), 180 190.

[41] Fraga Alves, M., M. Gomes, and L. de Haan (2003). A new class of semi-parametric estimators of the second order parameter. Portugaliae Mathematica 60, 193–214.

[42] Gardes, L., Stupfler, G., 2015. Estimating extreme quantiles under random truncation. *TEST* **24**, 207-227.

[43] Gill, R. D. (1980). Censoring and Stochastic Integrals," Mathematical Centre Tracts 124. Mathematisch Centrum, Amsterdam

[44] Gill, R. (1983). Large sample behaviour of the product-limit estimator on the whole line. Ann. Statist., 49 58.

[45] Gill, R. D. (1994). Glivenko-Cantelli for Kaplan-Meier. Math. Methods Statist., 3(1), 76.

[46] Gomes, M. I., de Haan, L., Peng, L., 2003. Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes* **5**, 387-414.

[47] Gomes, M. I., M. Martins, and M. Neves (2007). Improving second order reduced bias extreme value index estimator. REVSTAT-Statistical Journal (5(2)), 177–207.

[48] Gomes, M.I., Neves, M.M., 2011. Estimation of the extreme value index for randomly censored data. Biometrical Lett. 48 (1), 1-22.

[49] Gomes M.I., Oliveira O., 2003. Censoring estimators of a positive tail index. Statist. Probab. Lett. 65 (3), 147-159.

[50] Goegebeur, Y., Beirlant, J., de Wet, T., 2010. Kernel estimators for the second order parameter in extreme value statistics. *J. Statist. Plann. Inference* **140**, 2632-2652.

[51] Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press, New York.

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

[53] Hall, P. and A. Welsh (1985). Adaptive estimates of parameters of regular variation Annals of Statistics 13,331–341.

[54] Haouas, N., Necir, A. and Brahimi, B. 2016. Estimating the second-order parameter of regular variation and bias reduction in tail index estimation under random truncation. Submitted.

[55] Haouas, N., Necir, A. Djamel Meraghni and Brahimi, B. 2017. A Lynden-Bell integral estimator for the tail index of right-truncated data with a random threshold. Journal Afrika Statistika, 12(1); 1159–1170

[56] Hosking, J. R. M., Wallis, J. R. and Wood, E. F. (1985). Estimation of the Generalized Extreme Value Distribution by the Method of Probability-Weighted Moments. Technometrics 27.251-261.

[57] Hua, L., Joe, H., 2011. Second order regular variation and conditional tail expectation of multiple risks. *Insurance Math. Econom.* **49,** 537-546.

[58] Huang, X., and Strawderman, R. L. (2006). A note on the Breslow survival estimator. J. Nonparametr. Stat., 18(1), 45 56.

[59] Herbst, T., 1999. An application of randomly truncated data models in reserving IBNR claims. *Insurance Math. Econom.* **25**, 123-131.

[60] Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163-1174.

[61] Häusler, E. and Teugels, J.L. (1985). On Asymptotic Normality of Hill's estimator for the Exponent of Regular Variation. Annals of Statistics 13, 743-756.

[62] Hua, L., Joe, H., 2011. Second order regular variation and conditional tail expectation of multiple risks. Insurance Math. Econom. 49: 537-546.

[63] Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-481.

[64] Lynden-Bell, D., 1971. A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices Roy. Astronom. Soc.* **155**, 95-118.

[65] Lawless, J.F., 2002. Statistical Models and Methods for Lifetime Data, Second Edition. Wiley Series in Probability and Statistics.

[66] Markovich, N. M. (2005). High Quantile Estimation for Heavy-Tailed Distributions. Performance Evaluation 62, 178-192.

[67] Mason, D.M., 1982. Laws of large numbers for sums of extreme values. *Ann. Probab.* **10**, 756-764.

[68] Matthys, G. and Beirlant, J. (2003). Estimating the Extreme Value Index and High Quantiles with Exponential Regression Models. Statistica Sinica 13,853-880.

[69] Meraghni, D., 2008. Modeling Distribution Tails. Doctorat thesis, Med Khider University, Biskra-Algeria.

[70] Necir, A., 2006. A functional law of the iterated logarithm for kernel-type estimators of the tail index. Journal of Statistical Planning and Inference. 136 (2006) 780–802

[71] Nelson,W., 1972. Theory and applications of hazard plotting for censored failure data. Techno-metrics 14, 945-966.

[72] Neves, C. and Fraga Alves, M.I. (2004). Reiss and Thomas' Automatic Selection of the Number of Extremes. Computational Statistics and Data Analysis 47, 689-704.

[73] Peng, L., 1998. Asymptotically unbiased estimators for the extreme-value index. *Statist. Probab. Lett.* **38**, 107-115.

[74] Peng, L., Qi, Y., 2004. Estimating the first- and second-order parameters of a heavy-tailed distribution. *Aust. N. Z. J. Stat.* **46**, 305-312.

[75] Pickands, J. (1975) .Statistical Inference Using Extreme Order Statistics. Annals of Statistics 3. 119-131.

[76] Reiss, R.D., Thomas, M., 2007. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, 3rd ed. *Birkhäuser.*

[77] Resnick, S., 2006. Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer.

[78] Resnick, S.I. 1987. Extreme values, regular variation and point processes. Springer,NewYork.

[79] Saporta,G.(1990). Probabilités, Analyse des Donnée set Statistique. Editions Technip, Paris.

[80] Shorack, G.R., Wellner, J. A., 1986. Empirical processes with applications to statistics. *New York: John Wiley & Sons.*

[81] Stute, W., and Wang, J. L. (1993). A strong law under random censorship. Ann. Statist., 1591-1607.

[82] Stute, W., Wang, J., 2008. The central limit theorem under random truncation. Bernoulli 14: 604-622.

[83] Sousa, B. (2002). A contribution to the estimation of the tail index of heavy-tailed distributions. phd. thesis, university of michigan.

[84] Strzalkowska-Kominiak, E., Stute, W., 2009. Martingale representations of the Lynden-Bell estimator with applications. *Statist. Probab. Lett.* **79**, 814-820.

[85] Wang, M.C., 1989. A semiparametric model for randomly truncated data. *Statist. Assoc.* **84**, 742-748.

[86] Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. J. Am. Statist. Assoc. 73: 812-815.

[87] de Wet, T., Goegebeur, Y., Guillou, A., 2012. Weighted moment estimators for the second order scale parameter. *Methodol. Comput. Appl. Probab.* **14**, 753-783.

[88] Woodroofe, M., 1985. Estimating a distribution function with truncated data. *Ann. Statist.* **13**, 163-177.

[89] Worms, J., Worms, R., 2012. Estimation of second order parameters using probability weighted moments. *ESAIM Probab. Stat.* **16**, 97-113.

[90] Worms, J., Worms, R., 2014. New estimators of the extreme value index under random right censoring, for heavy-tailed distributions. Extremes 17, 337-358.

[91] Worms, J., Worms, R., 2016. A Lynden-Bell integral estimator for extremes of randomly truncated data. *Statist. Probab. Lett.* **109**, 106-117

Appendix A

Software R

R is a system, commonly known as language and software, which allows statistical analysis. More particularly, it includes means which make it possible to manipulate data, calculations and graphical representations. R also has the ability to run programs stored in text files and has a large number of statistical procedures called packets. The latter allow to deal fairly quickly with subjects as varied as linear (simple and generalized) models, regression (linear and nonlinear), time series, classical parametric and nonparametric tests, different methods of data analysis , ... Several packages, such as ade4, FactoMineR, MASS, multivariate, scatterplot3d and rgl, 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 of the University of Auckland in New Zealand. Since 1997, a "R Core Team" has been formed that is developing R. It is designed to be used with Unix, Linux, Windows and MacOS operating systems.

A key element in the development mission of R is the Comprehensive R Archive Network (CRAN) which is a set of sites that provides everything needed for R distribution, extensions, documentation, source files and files binaries. The master site of the CRAN is located in Austria in Vienna, it can be accessed by URL: "http://cran.r-project.org/". Other CRAN sites, known as mirror sites, are widespread around the world.

R is free software distributed under the terms of the "GNU Public License". It is an integral part of the GNU project and has an official website at http://www.R-project.org. It is often presented as an S clone which is a high-level language developed by AT & T Bell Laboratories and more particularly by Rick Becker, John Chambers and Allan Wilks. S is usable through the S-Plus software which is marketed by the company Insightful (http://www.splus.com/).

Appendix B

Abbreviations and notations

The different symbols and abbreviations used in this thesis.

$(\Omega, \mathcal{A}, \mathbf{P})$	:	probability space
rv	:	random variable
X	:	rv defined on $(\Omega, \mathcal{A}, \mathbf{P})$
$(X_1, X_2, ..., X_n)$	:	sample of size n from X
$(X_{1:n}, X_{2:n}, ..., X_{n:n})$	:	order statistics pertaining to $(X_1, X_2, ..., X_n)$
$X_{k:n}$	:	kth order statistic $(i = 1, ..., n)$
$X_{1:n}$	:	minimum of $(X_1, X_2, ..., X_n)$
$X_{n:n}$	:	maximum of $(X_1, X_2, ..., X_n)$
$\mathbb{R}$	:	set of real numbers
$\mathbb{R}^+$	:	set of positive real numbers
$[a, b]$	:	closed interval
(a, b)	:	open interval
df	:	distribution function
pdf	:	probability density function
F	:	df of a rv X
f	:	pdf of rv X
$F^{\leftarrow}$	:	generalized inverse of F, quantile function

F_n	:	empirical df
Q	:	quantile function, generalized inverse of F
Q_n	:	empirical quantile function
$\mathbb{I}_A(.)$	:	indicator function of set A
S_n	:	the partial sum X
$\inf A$	:	infimum of set A
$\sup A$	:	supremum of set A
$\overline{X}$	:	arithmetic mean of X
$U(a,b)$	:	uniform distribution on (a,b)
$U(0,1)$	:	standard uniform distribution
EX	:	expectation or mean of X
$a.s.$	:	almost sure
$\overset{a.s.}{\rightarrow}$	:	$a.s.$ convergence
$\overset{P}{\rightarrow}$	:	convergence in probability
$\overset{D}{\rightarrow}$	:	convergence in distribution
$\mathcal{N}(\mu,\sigma^2)$	:	normal or Gaussian distribution with mean μ and variance σ^2
$\mathcal{N}(0,1)$	:	standard normal or standard Gaussian distribution

$\exists$	:	exists
$\forall$	:	$\forall x$ i.e. for any x
RV_α	:	regular variation at ∞ with index α
RV_α^0	:	regular variation at 0 with index α
$D(H_\gamma)$	:	domain of attraction of the distribution H with tail index γ
$e.g.$	:	for example
EVI	:	extreme value index
EVT	:	extreme value theory
$i.e.$	:	in other words
iff	:	if and only if
iid	:	independent identically distributed
n	:	integer number greater than 1
$\mathbb{N}$	:	set of non-negative integers
$o(.)$	:	$f(x) = o(g(x))$ as $x \to x_0 : f(x)/g(x) = 0$ as $x \to x_0$
$O(.)$	:	$f(x) = O(g(x))$ as $x \to x_0 : \exists M > 0 : \lvert f(x)/g(x)\rvert \leq M,$ as $x \to x_0$
$o_p(.)$ and $O_p(.)$	:	stochasticordersymbols
	:	

$\max(A)$	:	maximum of the set A
$\min(A)$	:	minimum of the set A
x_F	:	the upper endpoint of df F
$\in$	:	belongs
$\vee$	:	$(a \vee b) = \max(a, b)$
$\wedge$	:	$(a \wedge b) = \min(a, b)$
$AMSE$	:	asymptotic mean squared error
log	:	logarithme
e	:	exponentiel (exp)

Abstract

In this thesis, we interested to statistics of rare events for incompletely observed data, with a particular interest in the estimate of extreme value (tail index, quantile extreme and second-order parameter) of distributions under random right-truncation. In this context, we proposed a consistent estimator of the second-order parameter of Pareto-type distributions under random right-truncation and establish its asymptotic normality. Moreover, we derived an asymptotically unbiased estimator for the tail index and study its asymptotic behaviour. We show, by simulation, that the proposed estimators behave well, in terms of bias and mean square error. After that, by means of a Lynden-Bell integral with deterministic threshold, we consider the random threshold case to derive a Hill-type estimator and establish its consistency and asymptotic normality. A simulation study is carried out to evaluate the finite sample behavior of the proposed estimator and compare it to the existing ones. And finally we estimate the quantile extreme in the case of complete data that is based on the both of methods the kernel type and the log probability weighted moment of estimation. Then, we prove the consistency and asymptotic normality of our estimator.

Keywords: Bias-reduction; Extreme value index; Heavy-tails; Second-order parameter; Random truncation; Product-limit estimator; Lynden-Bell estimator.

Résumé

Dans cette thèse, nous intéressons à statistique des événements rares pour des données incomplètement observées, en particulier au l'estimation des valeurs extrêmes aux distributions dans le cas ou les données sont tronquées à droite. Dans ce contexte, nous proposons un estimateur consistent du paramètre de second ordre des distributions de type Pareto et établissons sa normalité asymptotique. Nous obtenons un estimateur asymptotiquement sans biais pour l'indice de la queue et étudions son comportement asymptotique. Par simulation, on montre que les estimateurs proposés se comportent bien, en termes de biais et d'erreur quadratique moyenne. A travers l'intégrale de Lynden-Bell avec un seuil déterministe, nous considérons le cas du seuil aléatoire pour obtenir un estimateur de type Hill et établit sa consistent et sa normalité asymptotique. Une étude de simulation est réalisée afin d'évaluer le comportement des échantillons finis de l'estimateur proposé et de le comparer à ceux qui existent. Ensuite, nous estimons l'extrême quantile dans le cas des données complètes, qui est basé sur les deux méthodes, type noyau et le moment pondéré de probabilité logarithmique. Finalement, nous prouverons la consistent et la normalité asymptotique de notre estimateur.

Mots-clés:

Réduction des biais; Indice de valeur extrême; Heavy-tails; Paramètre de second ordre; Troncature aléatoire; Estimateur de produit -limite; Estimateur de Lynden-Bell.

ملخص

في هذه الأطروحة، اهتمينا بإحصاء القيم النادرة لبيانات غير كاملة و نهتم بصفة خاصة بتقدير القيم القصوى لتوزيعات في حالة بيانات مقطوعة من اليمين. في هذا السياق اقترحنامقدر متسق للمؤشر من الدرجة الثانية لتوزيعات باريتو و أثبتنا تقاربها الطبيعي. من جهة أخرى، تحصلنا على مقدر لمؤشر الذيل بدون تحيز مع دراسة تقاربه الطبيعي. وتبين لنا عن طريق المحاكاة أن للمقدرات المقترحة أداء جيد من حيث التحيز ومعدل مربع الخطأ. باستخدام مكاملة ليندن-بيل مع عتبة عشوائية قمنا بإثبات مقدر هيل واثبات اتساقه و تقاربه الطبيعي. *دراسة المحاكاة طبقت لمعاينة سلوك العينات بمقدرنا و مقارنته مع المقدرات الموجودة.* ثم قدمنا مقدر للربيعيات القصوى في حالة البيانات الكاملة المعتمدة على الطريقتين التقدير من نوع النواة و اللحظات المتزنة للاحتمالات اللوغاريتمية. ثم أثبتنا الاتساق و الحالة السوية للمقدر.

كلمات مفتاحية:

الحد من التحيز، مؤشر القيم القصوى، الذيول الثقيلة، المؤشر من الدرجة الثانية، اقتطاع عشوائي، مقدر نهاية المنتج، مقدر لندن-بيل.