Mohamed Khider University, Biskra

Faculty of Exact Sciences and the Natural Sciences and Life

Department of Mathematics

A Thesis Presented for the Degree of

DOCTOR in Mathematics

In The Field of Statistics

By

BENCHAIRA Souad

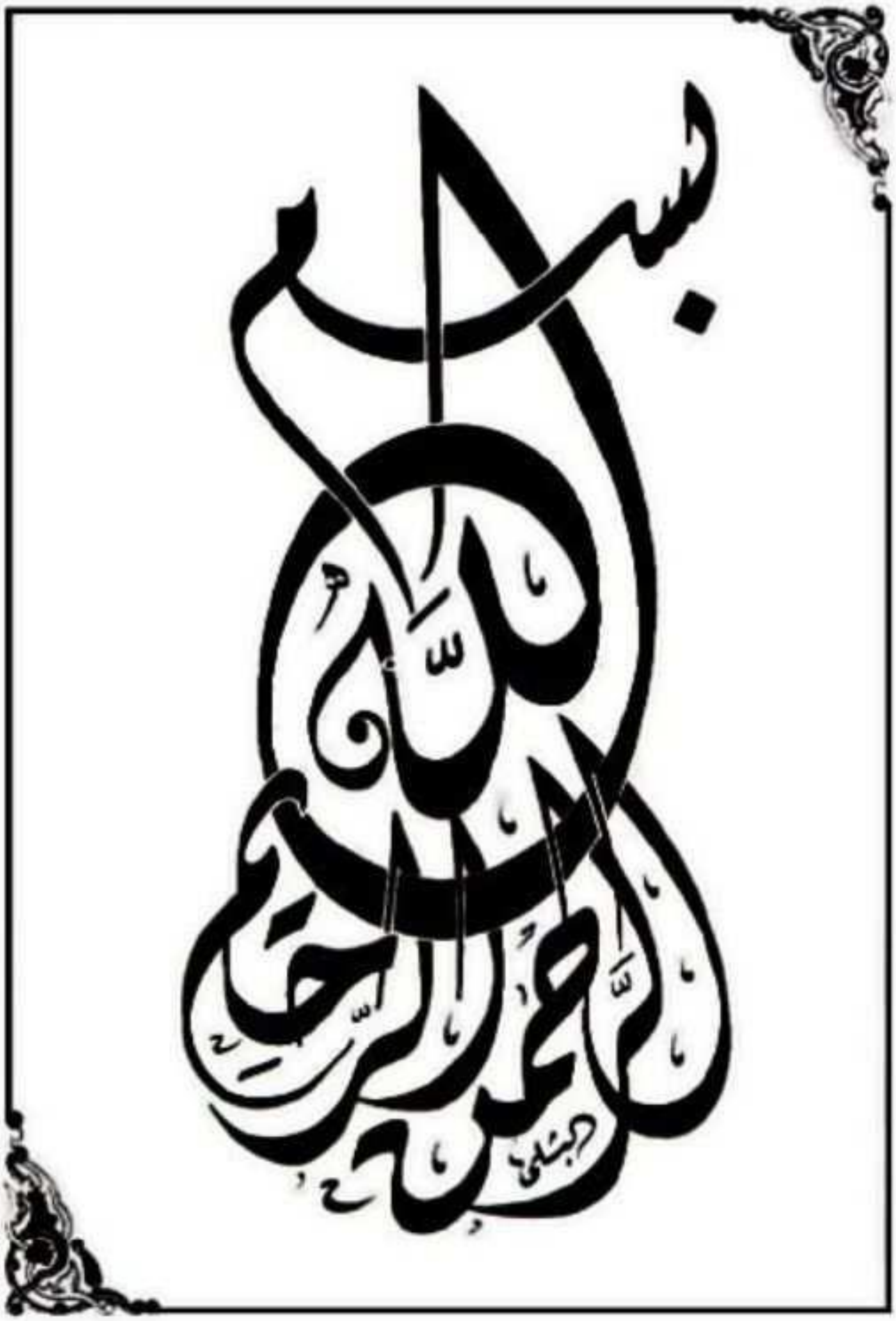
Title

Statistics of incomplete data

Examination Committee Members:

Pr.	NECIR	Abdelhakim	U. Biskra	Chairman
Pr.	MERAGHNI	Djamel	U. Biskra	Supervisor
Pr.	FELLAG	Hocine	U. Tizi Ouzou	Examinator
Dr.	BENATIA	Fateh	U. Biskra	Examinator
Dr.	BRAHIMI	Brahim	U. Biskra	Examinator
Dr.	YAHIA	Djabrane	U. Biskra	Examinator

February 02, 2017


بسم الله الرحمن الرحيم

Abstract

In this thesis, we are concerned with the estimation of the extreme value index and large quantiles for incompletely observed data, with a particular interest in the case of right-truncated data. We begin by exploiting the first work in this matter, which is due to [Gardes and Stupfler(2015)], to derive a simple tail index estimator based on a single sample fraction of extreme values. The asymptotic normality of the proposed estimator is established in the frameworks of tail dependence and second-order of regular variation. Second, starting from the first-order condition of regular variation, we construct a new estimator for the shape parameter of a right-truncated heavy-tailed distribution. We prove its asymptotic normality by making use of the tail Lynden-Bell process for which a weighted Gaussian approximation is provided. Also, a new approach of estimating high quantiles is proposed and applied to a real dataset consisting in lifetimes of automobile brake pads. Finally, a kernel-type asymptotically normal estimator is defined. Simulation experiments are carried out to evaluate the performances and illustrate the finite sample behaviors of the above estimators and make comparisons as well.

Keywords: Bivariate extremes; Empirical process; Extreme value index; Heavy-tails; High quantiles; Hill estimator; Kernel estimation; Lynden-Bell estimator; Regular variation; Random truncation; Tail dependence.

Achieved Works

Papers

- 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.

- Benchaira, S., Meraghni, D., Necir, A., 2016. Tail product-limit process for truncated data with application to extreme value index estimation. Extremes. 10.1007/s10687-016-0241-9.

- Benchaira, S., Meraghni, D., Necir, A.,2016. Kernel estimation of the tail index of a right-truncated Pareto-type distribution. Statist. Probab. Lett. 119, 186–193.

Dedication

This work is dedicated to my parents

My great parents, who never stop giving of themselves in countless ways

All I have and will accomplish are only possible due to their love and sacrifices

Aknowledgements

In the Name of Allah, the Most Gracious, the Most Merciful. All praise is due to Allah, Lord of the universe. Blessings and peace be upon the leader of the early and latter generations, our leader Muhammad and also upon his family and his companions.
First and foremost, I would like to express my deepest sense of gratitude to my supervisor Pr. Meraghni Djamel and Pr. Necir Abdelhakim, who offered their continuous advices and encouragement throughout the course of this thesis. Without their excellent guidance, great effort and persistent help their dissertation would not have been possible.
I am very grateful for their trust, their higher education, and their good natured disposition and I have also been very fortunate to have had the opportunity of taking courses from them.
I would be remiss if I did not thank them for engaging me in new ideas. Working with them is a great pleasure and I do not have enough words to express my deep and sincere appreciation. Thank you and the best of luck in your future endeavors.

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 the members, administrative personnel and lecturers of Department of Mathematics and the Laboratory of Applied Mathematics. Many thanks go in particular to Dr. Brahimi Brahim for his time and effort to help me, without forgetting head of Department Dr. Djabrane Yahia.

I want to thank my family for all their support through the years. Then, I would like to thank my parents for allowing me to realize my own potential. All the support they have provided me over the years was the greatest gift, who taught me the value of hard work and an education. Without them, I may never have gotten to where I am today.

Contents

List of Figures

List of Tables

Introduction

Incomplete data can take various forms, due to many different reasons; from censored or truncated data. Where, it is considered censored when the number of values in a set are known, but the values themselves are unknown. i.e, Censoring; Sources/events can be detected, but the values (measurements) are not known completely. Then, it is said to be truncated when there are values in a set that are excluded. i.e., Truncation; 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. So, it is not rare that data to treat are not complete. In this case a classical techniques don't adjust correctly.

In Statistics of Extremes we deal essentially with the estimation of parameters of extreme or even rare events. Where the formulation of the possible limiting distributions of the affinely transformed maximum of a sample, shows that the parameter γ, i.e, the extreme-value index is an important characteristic of the distribution. In the remainder of this thesis we will mainly be concerned with the estimation of that parameter under random truncation.

The estimation of the extreme-value index, for truncated data, has received a lot of attention in the extreme value literature. In this context, the treatment in this thesis is organized around two themes. The first is that the central analytic tool of extreme value theory, and the second is that the estimation under random right-truncation model. Accordingly we have presented an exposition of those aspects which are essential for a proper understanding of extreme value theory.

Extreme-value theory establishes the asymptotic behavior of the largest observations in a

sample. It provides methods for extending the empirical distribution function beyond the observed data. It is thus possible to estimate quantities related to the tail of a distribution such as small exceedance probabilities or extreme quantiles.
More specifically, let $X_1, ..., X_n$ be a sequence of random variables (rv), independent and identically distributed from a cumulative distribution function (cdf) F . Extreme-value theory establishes that the asymptotic distribution of the maximum $X_{n;n} = max\{X_1, ..., X_n\}$ properly rescaled is the extreme-value distribution with cdf

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

The parameter $\gamma \in \mathbb{R}$ is referred to as the extreme-value index. It plays an important parameter in univariate extreme-value theory since it controls the first order behavior of the distribution tail. The estimation of this paramater has first been considered in the case of complete data (no truncation). In the literature, numerous estimators of this parameter have been proposed, there exist a vast number of different approaches. For example the Hill estimator [Hill(1975)], the maximum likelihood estimator ([Hall(1982)]; [Smith(1985)];[Smith(1987)]; [Smith and Weissman(1985)]), the moment estimator [Dekkers et al.(1989)], the Pickands estimator ([Pickands(1975)]; [Drees(1996)]; [Segers(2005)]) and a kernel type estimator [Csörgö et al.(1985)], and many more. we will briefly discuss some of this estimators in this thesis. We will state their definitions and some of the (asymptotic) results obtained in the mentioned references.

In the recent years, several authors concentrated their efforts on obtaining good estimations of the EVI for incompletely observed data, i.e. randomly censored or truncated data (note here that, since the interest generally lies in the evaluation of the upper tail of the data, left censoring or left truncation is not a relevant framework, and therefore censoring or truncating is considered from the right). In those contexts, the usual estimators of the EVI need some modifications because otherwise they would lead to erroneous estimations when blindly applied to censored or truncated data. Some references for extreme value estimation in the context of randomly censored observations are [Beirlant et al.(2007)], [Einmahl et al.(2008)] and [Worms and Worms(2016)]. The first published work on extreme values estimation under random truncation was written by [Gardes and Stupfler(2015)], who dealt with heavy-tailed right truncated data. The framework of randomly right truncated data will be precisely defined in this thesis, which is organized as follows:

Chapter 1

Contains the essential definitions and results of incomplete data, with the main basic concepts on truncated data and some important and useful results existing in the literature for the random right truncation model. In this chapter we start by censored data, which it can be further classified into three categories: right censoring, left censoring and interval censoring. Afterwards, we will be interested in the truncated data. Which in turn has three type as follows: right truncation, left truncation and interval truncation, but in the present thesis, we are concerned with data that are right truncated.

Chapter 2

This chapter contains some mathematical preliminaries (the asymptotic properties of the sum of iid rv's including the CLT, order statistics and distributions of upper order statistics), also contains a derivation of the three families of classical Gnedenko limit distributions for extremes of iid variables and an account of regular variation and its extensions and domains of attraction. So, this chapter gives you an introduction to the mathematical and statistical theory underlying EVT.

Chapter 3

We will focus on the recent paper of [Gardes and Stupfler(2015)], Where recently they addressed the estimation of the extreme value index under random truncation. They proposed a consistent estimator based on two sample fractions k and k' of top observations from truncated and truncation data respectively. They also established its asymptotic normality in the case where k (resp. k') is asymptotically negligible with respect to k' (resp. k). Nevertheless, they did not cover the more interesting situation when $k = k'$. In this chapter, we consider this issue and derive a simple estimator based on a single sample fraction of extreme values. The asymptotic normality of the proposed estimator is proved by making use of weighted tail-copula processes and tail dependence frameworks and its finite sample behavior illustrated through some simulations.

Chapter 4

In addition,In chapter 4, we will introduce a tail Lynden-Bell process for heavy-tailed distributions of randomly truncated data and we give its weak approximation in terms of standard Wiener processes. In this chapter, a new estimator of the extreme value index for a heavy-tailed distribution is derived and its asymptotic normality is established. Extensive simulation study to investigate the performance of the proposed estimator is carried out. Our results will be of great interest to establish the limit distributions of many statistics in extreme value theory for randomly truncated data such as, the high quantiles, the actuarial risk measures and the goodness-of-fit functionals.

Chapter 5

Finally, in chapter 5 we will present the Kernel type estimator for the extreme value index, under random truncation, in the framework of Pareto-type distributions, and we establish the asymptotic normality of the proposed estimator by making use of the tail Lynden-Bell empirical process. Simulation experiments illustrate the finite sample behavior of some selected estimators.

We would not finish this introduction without mentioning that the statistical software R, is used in the treatment of the examples presented throughout this thesis.

Chapter 1

Incomplete data

Since 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. Truncation and censoring occur quite naturally in lifetime data, and one may refer to the books by [Cohen(1991)], [Balakrishnan and Cohen(1991)] and [Meeker and Escobar(1998)] for some detailed discussion in this regard.

There are three general types of censoring, right-censoring, left-censoring and interval-censoring. A second feature which may be present in some survival studies is that of truncation, there are three general types but in the present thesis, we are concerned with data that are right truncated. Since censoring and truncation are often confused, a brief discussion on censoring with examples is helpful to more fully understand right truncation.

1.1 Censoring

Definition 1.1.1 *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.*

Statistical techniques for analyzing censored data sets are quite well studied, especially in survival analysis, reliability and biostatistics. Also, the censoring mechanisms are very common and diversified. In this section, we will focus on discussing censored data. It gives partial information as events occurred to the right or left of a time boundary or within a time interval. It can be further classified into three categories: right censoring, left censoring and interval censoring, as follows:

1.1.1 Right censoring

The most common form of censoring is Right censoring, occurs when a time-to-event is only known to be greater than a censoring time due to end of study, loss to follow-up, or patient's withdrawal. It is convenient to use the following notation. For a specific individual under study, we assume that there is a lifetime X and a fixed censoring time, C_r (C_r for "right" censoring time). The X's are assumed to be independent and identically distributed. The exact lifetime X of an individual will be known if, and only if, X is less than or equal to C_r. If X is greater than C_r, the individual is a survivor, and his or her event time is censored at C_r. The data from this experiment can be conveniently represented by pairs of random variables (T, δ), where δ indicates whether the lifetime X corresponds to an event ($\delta = 1$) or is censored ($\delta = 0$), and T is equal to X, if the lifetime is observed, and to C_r if it is censored, i.e., $T = min(X, C_r)$.

1.1.2 Left censoring

Left censoring is much rare. A lifetime X associated with a specific individual in a study is considered to be left censored if it is less than a censoring time C_l (C_l for "left" censoring time), that is, the event of interest has already occurred for the individual before that person is observed in the study at time C_l. For such individuals, we know that they have experienced the event sometime before time C_l, but their exact event time is unknown. The exact lifetime X will be known if, and only if, X is greater than or equal to C_l. The data from a left-censored sampling scheme can be represented by pairs of random variables (T, δ), as in the previous kind, where T is equal to X if the lifetime is observed and δ

indicates whether the exact lifetime X is observed ($\delta = 1$) or not ($\delta = 0$). Note that, for left censoring as contrasted with right censoring, $T = \max(X, C_l)$.

Example 1.1.1 *In early childhood learning centers, interest often focuses upon testing children to determine when a child learns to accomplish certain specified tasks. The age at which a child learns the task would be considered the time-to-event. Often, some children can already perform the task when they start in the study. Such event times are considered left censored.*

1.1.3 Interval censoring

A more general type of censoring occurs when the lifetime is known to occur only within an interval. Such interval censoring occurs when patients in a clinical trial or longitudinal study have periodic follow-up and the patient's event time is only known to fall in an interval $(L_i, R_i]$ (L for left endpoint and R for right endpoint of the censoring interval). This type of censoring may also occur in industrial experiments where there is periodic inspection for proper functioning of equipment items. Animal tumorigenicity experiments may also have this characteristic.

Interval censoring is a generalization of left and right censoring because, when the left end point is 0 and the right end point is C_l we have left censoring and, when the left end point is C_r and the right end point is infinite, we have right censoring.

1.1.4 Estimation under random right-censoring model

This section deals with the nonparametric estimation of the df by means of the Kaplan–Meier estimator (also called the product–limit estimator) and with the estimator for the mean. We start with remarks about the statistics of extremes of randomly censored data. The topic was first mentioned in [Reiss and Thomas(1997)], Section 6.1, where an estimator of a positive extreme value index was introduced, but no (asymptotic) results were derived. In the last decade, several authors started to be interested in the estimation of the tail index along with large quantiles under random censoring as one can see in [Gomes and Oliveira(2003)], [Beirlant et al.(2007)], [Einmahl et al.(2008)] and [Worms and Worms(2014)]. [Gomes and Neves(2011)] also made a contribution to this eld by providing a detailed simulation study and applying the estimation procedures on some survival data sets.

Let $X_1, ..., X_n$ be $n \geq 1$ independent copies of a non-negative random variable X, defined over some probability space $(\Omega, \mathcal{A}, \mathbf{P})$, with continuous cumulative distribution function F. Rather then $X_1, ..., X_n$, the variables of interest, one observes

$$Z_i = \min(X_i, Y_i) \text{ and } \delta_i = 1_{X_i \leq Y_i}, \ 1 \leq i \leq n,$$

where $Y_1, ..., Y_n$ is another i.i.d. sequence from some (censoring) d.f. G being also independent of the $X's$. 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 context of this randomly right censored observations, the nonparametric maximum likelihood estimator of the survival distribution F is given by [Kaplan and Meier(1958)] as the product limit estimator defined by

$$F_n(x) := 1 - \prod_{Z_{i:n} \leq x} \left(1 - \frac{\delta_{[i:n]}}{n - i + 1}\right), \text{ for } x < Z_{n:n},$$

where $Z_{i:n}$ denote the order statistics associated to $Z_1, ..., Z_n$ and $\delta_{[i:n]}$ is the concomitant of the ith order statistics, that is, $\delta_{[i:n]} = \delta_j$ if $Z_{i:n} = Z_j$. This estimator may be expressed as follows

$$F_n(x) := \sum_{i=2}^{n} W_{i,n} 1_{\{Z_{i:n} \leq x\}}$$

where for $2 \leq i \leq n$,

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

(see, e.g., [Reiss and Thomas(2007)], page 162).

Now, we have the mean of X,

$$\mu := E[X] = \int_0^\infty \overline{F}(x)\, dx,$$

The sample mean for censored data is obtained by substituting, in the previous equation, the cdf F by its estimator F_n to have

$$\widetilde{\mu} := \sum_{i=2}^{n} \frac{\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{[j:n]}} Z_{i:n}.$$

The asymptotic normality of $\widetilde{\mu}$ is established by [Stute(1995)] under some assumptions.

1.2 Truncation

Definition 1.2.1 *Truncation is a variant of censoring which occurs when the incomplete nature of the observation is due to a systematic selection process inherent to the study design.*

Truncation appears when a time to the event is only observed in a study if the time-to-event variable is greater or smaller than the truncation variable. Some examples of truncated data from astronomy and economics can be found in [Woodroofe(1985)] and for applications in the analysis of AIDS data, see [Wang(1989)]. In reliability, a real dataset, consisting in lifetimes of automobile brake pads and already considered by [Lawless(2002)] in page 69, was recently analyzed in [Gardes and Stupfler(2015)] as an application of randomly truncated heavy-tailed models. One has three type of truncation, as follows:

1.2.1 Right truncation

Only individuals with event time less than some threshold 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 school are not included in the study.

Example 1.2.1 *Induction Times for AIDS data from [Lagakos, Barraj, and de Gruttola(1988)] are used to illustrate a situation in which one-sided (rather than two-sided) truncation appears. This data set is available from the book by [Klein and Moeshberger(2003)] (Table 1.10, pp. 20). The data include information on the infection and induction times for 258 adults and 37 children who were infected with HIV virus and developed AIDS by 1996-06-30 The data consist on the time in years, measured from 1978-04-01, when adults were infected by the virus from a contaminated blood transfusion, and the waiting time to development of AIDS, measured from the date of infection. In this sampling scheme, only individuals who had developed AIDS before the end of the study period were included and so the induction times suffer from right truncation. Let X be the induction time, that is, the time from HIV infection to the diagnosis of AIDS; and denote by T the time from HIV infection to the end of the study, which plays the role of right truncation time. Only those individuals (X, T) with $X \leq T$ are observed. In this example the sole information included is the infection and the induction times for the 258 adults. These variables X and T are reported in the second and the third column, respectively, of the matrix AIDSdata in Package DTDA.*

1.2.2 Left truncation

Due to structure of the study design, we can only observe those individuals whose event time is greater than some truncation threshold.

Example 1.2.2 *imagine you wish to study how long people who have been hospitalized for a heart attack survive taking some treatment at home. The start time is taken to be the time of the heart attack. Only those individuals who survive their stay in hospital are able to be included in the study.*

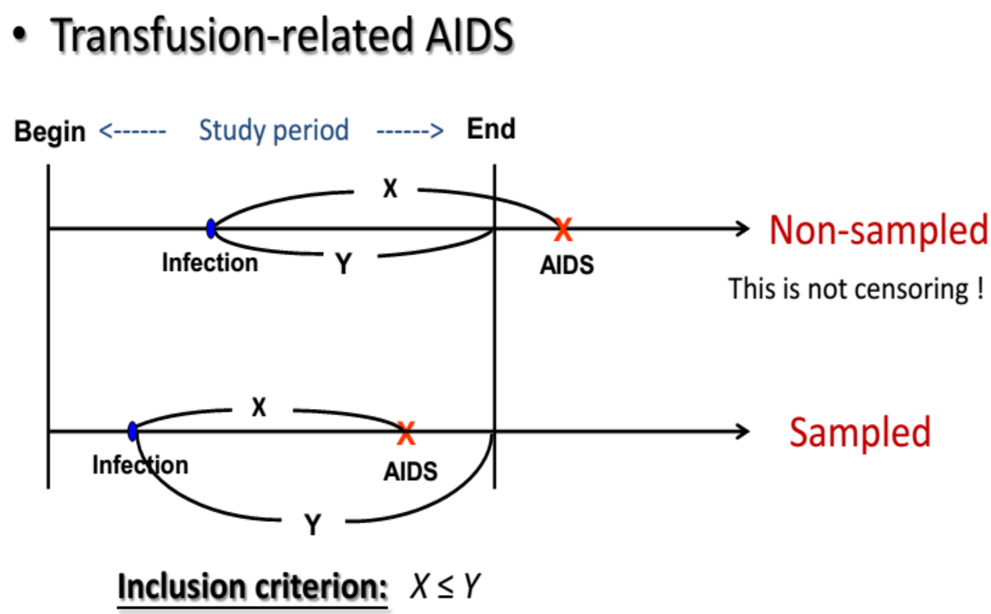


Figure 1.1: Example of right-truncated data

1.2.3 Interval truncation

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.

Example 1.2.3 *data on the luminosity of quasars in astronomy: One of the main aims of astronomers interested in quasars is to understand the evolution of the luminosity of quasars see [Efron and Petrosian(1999)]. The motivating example presented in this paper concerns a set of measurements on quasars in which there is double truncation, because the quasars are observed only if their luminosity occurs within a certain finite interval, that is bounded at both ends, with the interval varying for different observations.*

1.3 Estimation under random right-truncation model

In this section, we present some important and useful results existing in the literature for the random right truncation model:

1.3.1 Random right-truncation model

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 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. By the weak law of large numbers, we have

$$n/N \xrightarrow{\mathbf{P}} p := \mathbf{P}(\mathbf{X} \leq \mathbf{Y}), \text{ as } N \to \infty, \tag{1.1}$$

where the symbol $\xrightarrow{\mathbf{P}}$ stands for convergence in probability. We shall assume that $p > 0$, otherwise, nothing will be observed. The joint distribution of X_i and Y_i is

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

The marginal df's of the observed $X's$ and $Y's$, respectively denoted by F and G, are equal to $F(x) := p^{-1} \int_0^x \overline{\mathbf{G}}(z)\, d\mathbf{F}(z)$ and $G(y) := p^{-1} \int_0^y \mathbf{F}(z)\, d\mathbf{G}(z)$. It follows that the corresponding tails are

$$\overline{F}(x) = -p^{-1} \int_x^\infty \overline{\mathbf{G}}(z)\, d\overline{\mathbf{F}}(z) \text{ and } \overline{G}(y) = -p^{-1} \int_y^\infty \mathbf{F}(z)\, d\overline{\mathbf{G}}(z). \tag{1.2}$$

which are estimated by

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

Let $C(.)$ be a function defined by

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

with $\mathbf{Y}$ being the truncation rv. This quantity C is very crucial as it plays a prominent role is the statistical inference under random truncation. In other words, we have

$$C(z) := \mathbf{P}(X \leq z \leq Y) = F(z) - G(z).$$

with empirical estimator

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

1.3.2 Product-limit estimator

The focus of this section will be on the construction of the Lynden-Bell estimator of a distribution function in the random truncation model.

Since right endpoints of $\mathbf{F}$ and $\mathbf{G}$ are infinite and thus they are equal. Hence, from [Woodroofe(1985)], we may write

$$\int_x^{\infty}\frac{d\mathbf{F}(y)}{\mathbf{F}(y)} = \int_x^{\infty}\frac{dF(y)}{C(y)}, \tag{1.3}$$

Differentiating (1.3) leads to the following crucial equation

$$C(x)\,d\mathbf{F}(x) = \mathbf{F}(x)\,dF(x), \tag{1.4}$$

see, for instance, [Strzalkowska-Kominiak and Stute(2009)] whose solution is defined by

$\mathbf{F}(x) = \exp\left\{-\int_x^\infty dF(z)/C(z)\right\}$. Replacing F and C by their respective empirical counterparts yields the product-limit estimators of $\mathbf{F}$ and $\mathbf{G}$ given by

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

$$\mathbf{G}_n(x) := \prod_{i:Y_i \leq x} \exp\left\{-\frac{1}{nC_n(Y_i)}\right\},$$

The first mathematical investigation on this estimator may be attributed to [Woodroofe(1985)] and the central limit theorem under random truncation was established by [Stute and Wang(2008)]. Note that the approximation $\exp(-t) \sim 1 - t$, for small $t > 0$, results in the well-known estimator introduced by [Lynden-Bell(1971)].

Chapter 2

Extreme value theory

Extreme value theory EVT is an elegant and mathematically fascinating theory as well as a subject which pervades an enormous variety of applications. EVT is a classical topic, in probability theory and mathematical statistics, that was developed for the estimation of the occurrence probability of rare events. It permits to extrapolate the behavior of the distribution tails from the largest observed data. Classical EVT is well developed and a number of books are available in the area, see for example, [Gumbel(1958)], [Leadbetter et al.(1983)], [Resnick(1987)], [Embrechts et al.(1997)] and [de Haan and Ferreira(2006)] etc. In these next two section, we review some fundamental concepts of elementary probability and statistics. Then, we introduce various asymptotic models available in the classical EVT. Extreme value results are always phrased for maxima. One can convert results about maxima to apply to minima by using the rule $min(X_1, ..., X_n) = -max(-X_1, ..., -X_n)$.

2.1 Basic concepts

Definition 2.1.1 (Order statistics) *The order statistics pertaining to a sample* $(X_1, ..., X_n)$ *are the* X_i' *s arranged in non-decreasing order. They are denoted by* $X_{1,n}, ..., X_{n,n}$ *and for* $k = 1, ..., n,$ *the rv* $X_{n-k+1,n}$ *is called the kth upper order statistic. Order statistics satisfy*

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

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

Definition 2.1.2 (Distribution and survival functions) *If X is a rv defined on a probability space $(\Omega, \mathcal{F}, P)$ then, its df and survival function (also called hazard function) are respectively defined on $\mathbb{R}$ by*

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

Definition 2.1.3 (Quantile and tail quantile functions) *The quantile function of df F is the generalized inverse function of F defined by*

$$Q(s) := F^{\leftarrow}(s) = \inf\{x \in \mathbb{R} : F(x) \geq s\}, 0 < s < 1,$$

with the convention that the infimum of the empty set is ∞. In the theory of extremes, a function, denoted by U and called tail quantile function, is used quite often. It is defined by

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

Definition 2.1.4 (Empirical quantile and tail quantile functions) *The empirical (or sample) quantile function of the sample $(X_1, ..., X_n)$ is defined by*

$$\begin{aligned} Q_n(s) &:= \inf\{x \in \mathbb{R} : F_n(x) \geq s\}, 0 < s < 1, \\ &= X_{i,n} \text{ for } \tfrac{i-1}{n} < s \leq \tfrac{i}{n}, i = 1, .., n. \end{aligned}$$

Note that for $0 < p < 1$; $X_{[np]+1,n}$ is the sample quantile of order p, where $[np]$ denotes the integer part of np. If $s = 1/2$ then one also speaks of the sample median. The corresponding empirical tail quantile function is

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

Proposition 2.1.1 (Quantile transformation) *Let* $(U_1, ..., U_n)$ *be a sample from the standard uniform rv* U *and* $(U_{1,n}, ..., U_{n,n})$ *the corresponding ordered sample.*

- *For any df* F*, we have*

$$X_{i,n} \overset{d}{=} F^{\leftarrow}(U_{i,n}), i = 1, ..., n.$$

- *When* F *is continuous, we have*

$$F(X_{i,n}) \overset{d}{=} U_{i,n}, \ \ i = 1, ..., n.$$

In this case the rv's $F(X_1), ..., F(X_n)$ *are iid standard uniform.*

Definition 2.1.5 (Sum and arithmetic mean) *Let* $X_1, X_2, ..., X_n$ *be a sequence of iid rv's with common df* F*. For an integer* $n \geq 1$*, define the partial sum and the corresponding arithmetic mean by respectively*

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

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

2.1.1 Laws of large numbers

In what follows $(X_1, ..., X_n)$ will be considered as a sample from a rv X defined on a probability space $(\Omega, \mathcal{F}, P)$. If we want to get a rough idea about the fluctuations of the X_n we might ask for convergence of the sequence (X_n). Unfortunately, for almost all $\omega \in \Omega$ this sequence does not converge. However, we can obtain some information about how the X_n "behave in the mean". We have two kinds of laws describe the asymptotic behavior of the sample mean. The weak law is about the convergence in probability or the consistency of $\overline{X}_n$ while the strong law, concerns the strong convergence of $\overline{X}_n$, i.e. convergence with probability 1.

Theoreme 2.1.1 (Laws of large numbers) *If $(X_1, ..., X_n)$ is a sample from a rv X such that $E|X| < \infty$, then*

$$\text{weak law} \quad \overline{X}_n \overset{P}{\to} \mu \quad \text{as } n \to \infty,$$
$$\text{strong law} \quad \overline{X}_n \overset{a.s}{\to} \mu \quad \text{as } n \to \infty,$$

where $\mu := EX$.

Definition 2.1.6 (Empirical df) *The empirical df (or sample df) of the sample $(X_1, ..., X_n)$ is defined by*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}}, \; x \in \mathbb{R}.$$

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

$$F_n(x) = \begin{cases} 0 & x < X_{1,n} \\ \frac{i-1}{n} & X_{i-1,n} \le x < X_{i,n}, \text{for } i = 2, ..., n \\ 1 & x \ge X_{n,n} \end{cases}$$

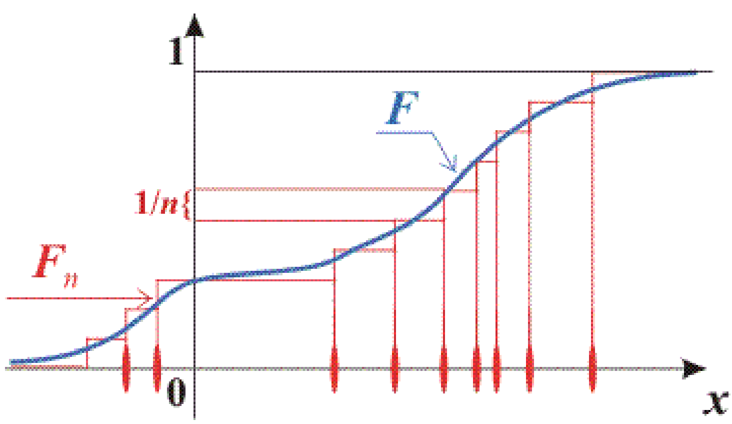


Figure 2.1: Empirical and theoretical distribution function

Example 2.1.1 (Glivenko–Cantelli theorem) *An application of the strong law of large* *numbers yields that*

$$F_n(x) \stackrel{a.s}{\to} EI_{\{X \leq x\}} = F(x)$$

for every $x \in \mathbb{R}$. The latter can be strengthened (and is indeed equivalent) to

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \stackrel{a.s}{\to} 0.$$

The latter is known as the Glivenko–Cantelli theorem. It is one of the fundamental results in non–parametric statistics.

Theoreme 2.1.2 (Central Limit Theorem) *If $X_1, X_2, ..., X_n$ is a sequence of iid rv's with mean μ and finite variance σ^2, then*

$$(S_n - n\mu) / \sigma\sqrt{n} \stackrel{d}{\to} \mathcal{N}(0,1) \quad \text{as } n \to \infty.$$

The proof of the CLT could be found in any standard book of statistics, see e.g., [Saporta(1990)].

2.1.2 Order statistics

Order statistics are very instrumental in EVT because they (more precisely the upper ones) provide information on the distribution (right) tail. In this section, we will summarize some of their properties and results. After having investigated the behaviour of the maximum, i.e. the largest value of a sample, we now consider the joint behaviour of several upper order statistics.

The relationship between the order statistics and the empirical df of a sample is immediate, we have

$$X_{k,n} \leq x \text{ if and only if } \sum_{i=1}^{n} I_{\{X_i > x\}} < k,$$

which implies that

$$P(X_{k,n} \leq x) = P\left(F_n(x) > 1 - \frac{k}{n}\right).$$

Upper order statistics estimate tails and quantiles, and also excess probabilities over certain thresholds. we have

$$F^{\leftarrow}(t) = X_{k,n} \text{ for } 1 - \frac{k}{n} < t \leq 1 - \frac{k-1}{n},$$

for $k = 1, .., n$. Next we calculate the df $F_{k,n}$ of the kth upper order statistic explicitly.

Proposition 2.1.2 (Distribution function of the k th upper order statistic) *For $k = 1, ..., n$ let $F_{k,n}$ denote the df of $X_{k,n}$. Then*

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

(b) If F is continuous, then

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

where

$$f_{k,n}(x) = \frac{n!}{(k-1)!\,(n-k)!} F^{n-k}(x)\, \overline{F}^{k-1}(x);$$

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

Proof. see e.g, [Embrechts et al.(1997)] p183. ■

Similar arguments lead to the joint distribution of a finite number of different order statistics. If for instance F is absolutely continuous with density f, then the joint density of $(X_1, ..., X_n)$ is

$$f_{X_1,...,X_n}(x_1, ..., x_n) = \prod_{i=1}^{n} f(x_i), \quad (x_1, ..., x_n) \in \mathbb{R}^n.$$

Since the n values of $(X_1, ..., X_n)$ can be rearranged in $n!$ ways (by absolute continuity there are a.s. no ties), every specific ordered collection $(X_{k,n})_{k=1,...,n}$ could have come from $n!$ different samples.

Proposition 2.1.3 (Distributions of order statistics)

- *Joint pdf of* $X_{1,n},...,X_{n,n}$

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

- *Joint pdf of* $X_{1,n},...,X_{k,n}$

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

Note that the df of the ith order statistic is a tail distribution of a binomial distribution with parameters n and $F(x)$. Distributional results for the smallest and largest order statistics are immediate.

Corollary 2.1.1 (Distributions of the minimum and maximum)

- Joint pdf of $X_{1,n}$ and $X_{n,n}$

$$f_{X_{1,n},X_{n,n}}(x,y) = n(n-1)\{F(y) - F(x)\}^{n-2} f(x) f(y), -\infty < x < y < \infty.$$

- pdf of min $(X_1,...,X_n)$

$$f_{X_{1,n}}(x) = n\left\{\overline{F}(x)\right\}^{n-1} f(x), -\infty < x < \infty.$$

- pdf of max $(X_1,...,X_n)$

$$f_{X_{n,n}}(x) = n\{F(x)\}^{n-1} f(x), -\infty < x < \infty.$$

- df of min $(X_1,...,X_n)$

$$F_{X_{1,n}}(x) = 1 - \left\{\overline{F}(x)\right\}^{n}, -\infty < x < \infty.$$

- df of max $(X_1, ..., X_n)$

$$F_{X_{n,n}}(x) = \{F(x)\}^n, -\infty < x < \infty.$$

2.2 Fluctuations of maxima

This section is concerned with classical EVT. The central result is the Fisher–Tippett theorem which specifies the form of the limit distribution for centred and normalised maxima. The three families of possible limit laws are known as extreme value distributions. We remind that throughout this section, $(X_1, .., X_n)$ is a sample from a rv X with continuous df F and $X_{n,n} = \max(X_1, .., X_n)$.

2.2.1 Limit distributions

Our interest is in finding possible limit distributions for sample maxima of independent and identically distributed random variables. Let F be the underlying distribution function and x_F its right endpoint, i.e.,

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

which may be infinite.Then $X_{n,n} \xrightarrow{P} x_F$, $n \to \infty$, since $P(X_{n,n} \leq x) = F^n(x)$, which converges to zero for $x < x_F$ and to 1 for $x \geq x_F$. Hence,in order to obtain a nondegenerate limit distribution, a normalization is necessary.

Suppose there exists a sequence of constants $a_n > 0$,and b_n real $(n = 1, 2, ...)$, such that $\frac{X_{n,n}-b_n}{a_n}$ has a nondegenerate limit distribution as $n \to \infty$, i.e.,

$$\lim_{n\to\infty} P\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} F^n(a_n x + b_n) = H(x), \ x \in \mathbb{R}, \tag{2.1}$$

where H is a non-degenerate df. Since extreme value df's are continuous on $\mathbb{R}$, assumption

2.1 is equivalent to the following weak convergence assumption

$$\frac{X_{n,n} - b_n}{a_n} \xrightarrow{d} H \text{ as } n \to \infty. \tag{2.2}$$

a_n and b_n are called norming constants.

We shall find all distribution functions H that can occur as this limit. These distributions are called extreme value distributions. The class of distributions F satisfying 2.1 is called the maximum domain of attraction or simply domain of attraction of H. We are going to identify all extreme value distributions and their domains of attraction.

Theoreme 2.2.1 (Fisher-Tippett) *Let (X_n) be a sequence of iid rvs. If there exist norming constants $a_n > 0$, $b_n \in \mathbb{R}$ and some non–degenerate df H satisfies assumption 2.2, then H belongs to the type of one of the following three dfs:*

$$\begin{aligned}
\mathit{TypeI}: \quad & \Lambda(x) &&= \exp\left(-e^{-x}\right), \qquad x \in \mathbb{R}. \\
\mathit{TypeII}: \quad & \Phi_\xi(x) &&= \begin{cases} 0, & x \le 0 \\ \exp\left(-x^{-\xi}\right), & x > 0 \end{cases} \quad \xi > 0. \\
\mathit{TypeIII}: \quad & \Psi_\xi(x) &&= \begin{cases} \exp\left(-(-x)^{\xi}\right), & x \le 0 \\ 1, & x > 0 \end{cases} \quad \xi > 0.
\end{aligned}$$

For the Sketch of the proof see [Embrechts et al.(1997)] p122.

Definition 2.2.1 (Standard extreme value distributions) *The Three df 's of this theorem are called standard extreme value distributions. Λ is known as Gumbel (or double exponential) type, Φ_ξ as Frechet (or heavy-tailed) type and Ψ_ξ as (reverse) Weibull type.*

The types of Λ, Φ_ξ and Ψ_ξ are very different, from a mathematical point of view they are closely linked. Indeed, one immediately verifies the following properties. Suppose $X > 0$, then

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

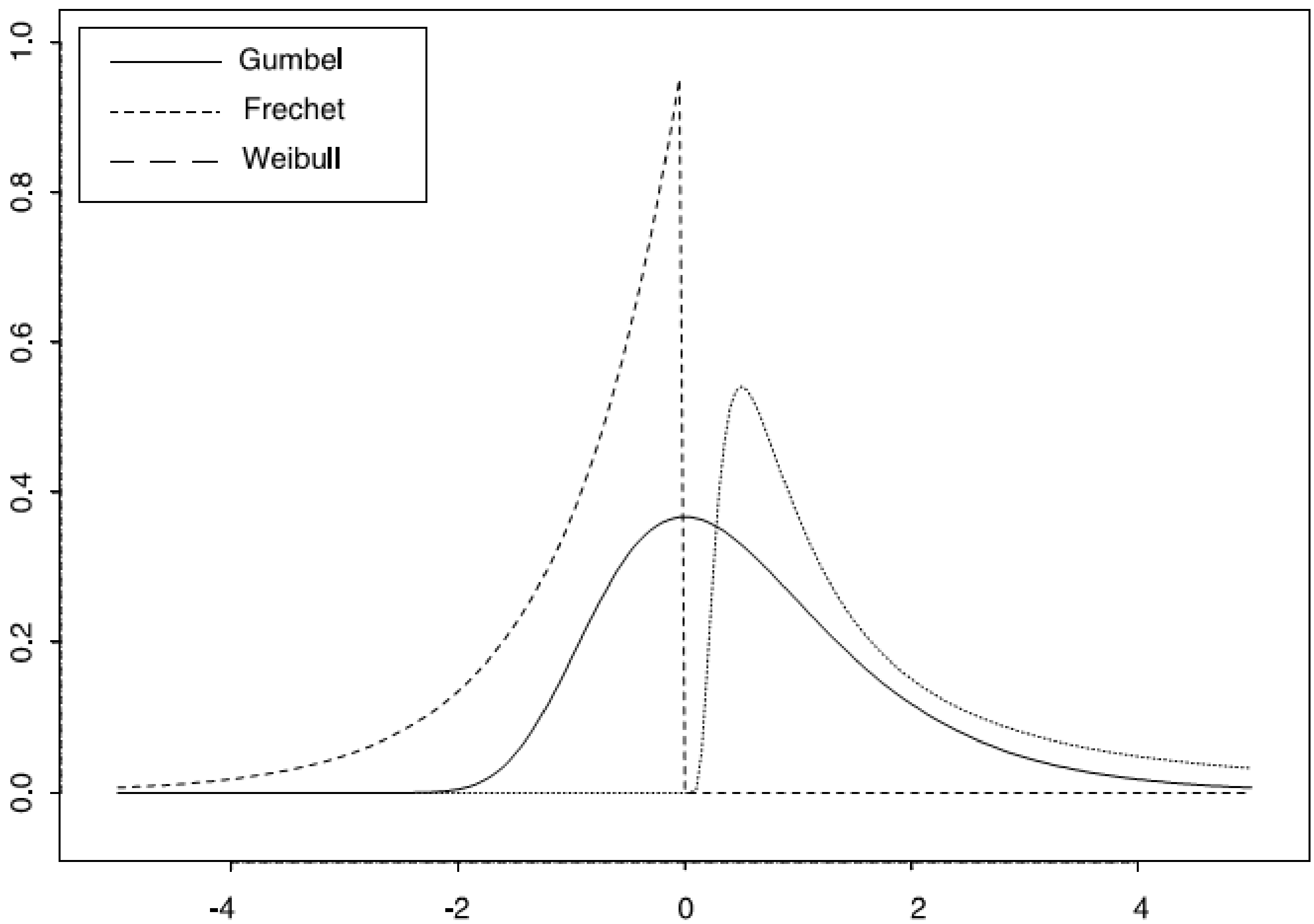


Figure 2.2: Densities of the standard EV distributions. We chose $\alpha = 1$ for the Fréchet and the Weibull distribution

There exists some freedom in the choice of the norming constants in 2.1 because the uniqueness of the limit H is only up to affine transformations. The three limit types of theorem 2.2.1 may be combined into a single form known as the Generalized Extreme Value Distribution (GEVD) and widely accepted as the standard representation of the extreme value distributions.

Definition 2.2.2 (GEVD) *The GEVD is a df H_γ defined , for all $x \in \mathbb{R}$ such that $1 + \gamma x > 0$, as follows:*

$$H_\gamma(x) = \begin{cases} \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\} & \text{if} \quad \gamma \neq 0, \\ \exp\left(-e^{-x}\right) & \text{if} \quad \gamma = 0. \end{cases} \tag{2.3}$$

The parameter γ is called Extreme Value Index (EVI), tail index or shape parameter.
The GEVD H_γ can be written in a more general form by replacing the argument x by $(x-\mu)/\sigma$ in the right hand side of equation 2.3, where $\mu \in \mathbb{R}$ and $\sigma > 0$ are respectively the location and scale parameters.
The parametrization in 2.3 is due to [von Mises(1936)] and [Jenkinson(1955)] and is known as the GEVD or the von Mises-Jenkinson family, which unifies all possible non-degenerate weak limits of the maximum. We express the three extreme value distributions in terms of the GEVD H_γ as follows:

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

In other words,

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

Hence the three extreme value distributions can be characterized by the sign of the tail index γ : Gumbel type corresponds to $\gamma = 0$, Fréchet type to $\gamma > 0$ and Weibull type to $\gamma < 0$.

Recall that if relation 2.3 holds with $H = H_\gamma$ for some $\gamma \in \mathbb{R}$, we say that the distribution function F is in the domain of attraction of H_γ. Notation $F \in \mathcal{D}(H_\gamma)$.

Remark 2.2.1 *Let us consider the subclasses separately:*

- *For* $\gamma > 0$*clearly* $H_\gamma(x) < 1$ *for all* x, *i.e., the right endpoint of the distribution is infinity. Moreover, as* $x \to \infty$, $1 - H_\gamma(x) \sim \gamma^{-1/\gamma} x^{-1/\gamma}$, *i.e., the distribution has a rather heavy right tail; for example, moments of order greater than or equal to* $1/\gamma$ *do not exist.*

- *For* $\gamma = 0$ *the right endpoint of the distribution equals infinity. The distribution, however,is rather light-tailed:* $1 - H_0(x) \sim e^{-x}$ *as* $x \to \infty$, *and all moments exist.*

- *For* $\gamma < 0$ *the right end point of the distribution is* $-1/\gamma$ *so it has a short tail, verifying* $1 - H_\gamma(-\gamma^{-1} - x) \sim (-\gamma x)^{-1/\gamma}$, *as* $x \downarrow 0$.

2.2.2 Domains of attraction

In this part of the section, we shall derive sufficient conditions on the distribution function F that ensure that there are sequences of constants $a_n > 0$ and b_n such that

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

for some given real γ and all x. These conditions, basically due to [von Mises(1936)], require the existence of one or two derivatives of F.

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

Theoreme 2.2.2 *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 left neighborhood of x_F.If

$$\lim_{t\uparrow x_F}\left(\frac{1-F}{F'}\right)'(t)=\gamma \tag{2.4}$$

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.

Theoreme 2.2.3

1. ($\gamma>0$) Suppose $x_F=\infty$ and F' exists.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. ($\gamma<0$) Suppose $x_F<\infty$ and F' exists for $x<x_F$. If

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

   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 consult [de Haan and Ferreira(2006)] p15.

Now in the following theorem, we shall establish necessary and sufficient conditions for a distribution function F to belong to the domain of attraction of H_γ.

Theoreme 2.2.4 *The distribution function F is in the domain of attraction of the extreme value distribution H_γ if and only if*

1. for $\gamma>0$: x_F is infinite and

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

for all $x > 0$. This means that the function $1 - F$ is regularly varying at infinity with index $-1/\gamma$.

2. for $\gamma < 0 : x_F$ is finite and

$$\lim_{t \downarrow 0} \frac{1 - F(x_F - tx)}{1 - F(x_F - t)} = x^{-1/\gamma}$$

for all $x > 0$.

3. for $\gamma = 0 : x_F$ can be finite or infinite and

$$\lim_{t \uparrow x_F} \frac{1 - F(t + xf(t))}{1 - F(t)} = e^{-x} \tag{2.5}$$

for all real x, where f is a suitable positive function. If 2.5 holds for some f, then $\int_t^{x_F} (1 - F(s))\, ds < \infty$ for $t < x_F$ and 2.5 holds with

$$f(t) := \frac{\int_t^{x_F} (1 - F(s))\, ds}{1 - F(t)}.$$

In addition to this formulations, for the domain of attraction assumption, there exist other alternative assertions stated in the following proposition. The First one illustrates the restriction on the upper distribution tail, the second form is in terms of function Q and the third assertion is in terms of function U.

Proposition 2.2.1 (Characterizations of $\mathcal{D}(H_\gamma)$) *For $\gamma \in \mathbb{R}$, the following assertions are equivalent.*

(a) $F \in \mathcal{D}(H_\gamma)$.

(b) *For some positive function b*

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

for all $x > 0$ *with* $(1 + \gamma x) > 0$.

(*c*) *For some positive function* $\widetilde{a}$

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

for $x > 0$.

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

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

for $x > 0$.

The latter two assertions are respectively equivalent to

$$\lim_{s \to 0} \frac{Q(1 - sx) - Q(1 - s)}{Q(1 - sy) - Q(1 - s)} = \begin{cases} \frac{x^{-\gamma} - 1}{y^{-\gamma} - 1} & \text{if } \gamma \neq 0, \\ \frac{\log x}{\log y} & \text{if } \gamma = 0, \end{cases}$$

and

$$\lim_{t \to \infty} \frac{U(tx) - U(t)}{U(ty) - U(t)} = \begin{cases} \frac{x^{\gamma} - 1}{y^{\gamma} - 1} & \text{if } \gamma \neq 0, \\ \frac{\log x}{\log y} & \text{if } \gamma = 0, \end{cases} \tag{2.6}$$

for $x, y > 0, y \neq 1$.

2.3 Regular variation

In the previous subsection we stated necessary and sufficient conditions that characterised the domains of attraction of the three extreme value distributions. These conditions are closely related to the concept of regularly varying functions. Moreover, in the analysis of the behaviour of estimators in the field of EVT, properties of regularly varying functions and so called Π-varying functions are frequently used.

We define the second order assumption that strengthens the regular variation condition of the distribution tail $\overline{F}$, with a reminder on that first condition in the case of heavy tailed distributions.

Proposition 2.3.1 (First Order Regular Variation Condition) *The following assertions are equivalent :*

(a) *F heavy tailed*

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

(b) *$\overline{F}$ regularly varying at ∞ with index $-1/\gamma$*

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

(c) *$Q(1-s)$ regularly varying at 0 with index $-\gamma$*

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

(d) *U regularly varying at ∞ with index γ*

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

Definition 2.3.1 (Second Order Regular Variation Assumption) *We say that $F \in D\left(\Phi_{1/\gamma}\right)$, $\gamma > 0$, is second order regularly varying at infinity if it satisfies one of the following (equivalent) conditions:*

(a) *There exist some parameter $\rho \leq 0$ and a function A^*, tending to 0 and not changing sign near infinity, such that for all $x > 0$*

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

(*b*) *There exist some parameter* $\rho \le 0$ *and a function* A^{**}, *tending to* 0 *and not changing sign near* 0, *such that for all* $x > 0$

$$\lim_{s \to 0} \frac{\left(Q\left(1 - sx\right)/Q\left(1 - s\right) - x^{-\gamma}\right)}{A^{**}\left(s\right)} = x^{-\gamma} \frac{x^{\rho} - 1}{\rho}.$$

(*c*) *There exist some parameter* $\rho \le 0$ *and a function* A, *tending to* 0 *and not changing sign near* 0, *such that for all* $x > 0$

$$\lim_{t \to \infty} \frac{\left(U\left(xt\right)/U\left(t\right) - x^{\gamma}\right)}{A\left(t\right)} = x^{\gamma} \frac{x^{\rho} - 1}{\rho}. \tag{2.7}$$

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

A, A^{*} and A^{**} are regularly varying functions with $A^{*}\left(t\right) = A\left(1/\overline{F}\left(t\right)\right)$ and $A^{**}\left(s\right) = A\left(1/s\right)$. Their role is to control the speed of convergence in First Order Regular Variation Condition $(a), (b)$ and (c) respectively.

The equivalent assumptions above specify the rates (necessary to derive the asymptotic normality of tail index estimators) of convergence in Proposition 2.3.1.

Hall's class of df 's

As an example of heavy tailed distributions satisfying the second order assumption, we have the so called and frequently used Hall's model (introduced in [Hall(1982)]) which is a class of df's

$$F\left(x\right) = 1 - cx^{-1/\gamma}\left(1 + dx^{\rho/\gamma} + o\left(x^{\rho/\gamma}\right)\right) \text{ as } x \to \infty, \tag{2.8}$$

where $\gamma > 0$, $\rho \le 0$, $c > 0$, and $d \in \mathbb{R} \backslash \{0\}$. This sub-class of heavy-tailed distributions contains the Pareto, Burr, Fréchet and t-Student df's usually used, in insurance mathematics, as models for dangerous risks. Relation 2.8 may be reformulated in terms of functions Q and U as follows:

$$Q\left(1 - s\right) = c^{\gamma} s^{-\gamma}\left(1 + \gamma d c^{\rho} s^{-\rho} + o\left(s^{-\rho}\right)\right) \text{ as } s \to 0,$$

and

$$U(t) = c^{\gamma} t^{\gamma} \left(1 + \gamma d c^{\rho} t^{\rho} + o\left(t^{\rho}\right)\right) \text{ as } t \to \infty.$$

Straightforward computations show that, in the case of Hall model, functions $A(t)$ and $A^*(t)$ are respectively equivalent to $d\rho\gamma c^{\rho} t^{\rho}$ and $d\rho\gamma t^{\rho/\gamma}$ as $t \to \infty$, whereas function $A^{**}(s)$ is equivalent to $d\rho\gamma c^{\rho} s^{-\rho}$ as $s \to 0$.

2.4 Tail index estimation

In this section we study different estimators of the shape parameter γ for $F \in \mathcal{D}(H_{\gamma})$. We also give some of their statistical properties. That is, the data $(X_1, ..., X_n)$ are assumed to be drawn from a population X with df F. This semi-parametric statistical procedures don't assume the knowledge of the whole distribution but only focus on the distribution tails. The case $\gamma > 0$ has got more interest because data sets in most real-life applications, exhibit heavy tails. Classical estimators of γ may be based on k upper order statistics $X_{k,n}, ..., X_{n,n}$ where

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

we will present some algorithms, in the next section, on how to determine this crucial number "k".

2.4.1 Pickands estimator

The simplest and oldest estimator for γ is the Pickands estimator, was introduced in 1975 for any $\gamma \in \mathbb{R}$, but, as it is rather unworkable in practise for small or moderate samples. The basic idea behind this estimator consists of finding a condition equivalent to $F \in D(H_{\gamma})$ which involves the parameter γ in an easy way, namely assertion 2.6, 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

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

The basic idea now consists of constructing an empirical estimator using this formule. To that effect, let the ordered $(Y_{1,n}, ..., Y_{n,n})$ from a standard Pareto rv Y with df $F_Y(y) = 1 - 1/y$, $y \geq 1$ (namely, $(k/n) Y_{n-k+1,n} \overset{P}{\to} 1$ and $Y_{n-k+1,n}/Y_{n-2k+1,n} \overset{P}{\to} 2$ as $n \to \infty$), yields

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

Finally, we use the distributional identity

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

and we now define the Pickands estimator

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

This was already observed by [Pickands(1975)]. A full analysis on $\widehat{\gamma}^{(p)}$ is to be found in [Dekkers and de Haan(1989)] from which the following result is taken.

Theoreme 2.4.1 (Asymptotic Properties of $\widehat{\gamma}^{(p)}$) *Assume that* $F \in D(H_{\gamma})$, $\gamma \in \mathbb{R}$, $k \to \infty$ *and* $k/n \to 0$ *as* $n \to \infty$.

(a) *Weak Consistency:*

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

(b) *Strong consistency: If* $k/\log\log n \to \infty$ *as* $n \to \infty$, *then*

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

(c) *Asymptotic normality: Under further conditions on* k *and* F *(see [Dekkers and de Haan(1989)], p. 1799),*

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

where

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

2.4.2 Hill's estimator

This estimator is only applicable in case the EVI γ is known to be positive, which corresponds to distributions belonging to the Fréchet type domain of attraction. In order to introduce the Hill estimator, a simple and widely used estimator, let us start from: $F\in D\left(\Phi_{1/\gamma}\right)$ for $\gamma>0$ if and only if

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

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

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

which, by an integration by parts, becomes

$$\lim_{t\to\infty}\frac{1}{\overline{F}\left(t\right)}\int_t^{\infty}\log\frac{x}{t}dF(x)=\gamma.$$

Replacing F by F_n and letting $t=X_{n-k:n}$ yields the Hill's(1975) estimator $\widehat{\gamma}^{(H)}$, defined by

$$\widehat{\gamma}^{(H)}:=\frac{1}{\overline{F}_n\left(X_{n-k:n}\right)}\int_{X_{n-k:n}}^{\infty}\log\frac{x}{X_{n-k:n}}dF_n\left(x\right),$$

or

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

Hill's estimator is usual and easy-to-explain. It can be derived through several other approaches (see [Embrechts et al.(1997)] p. 330). In his original paper [Hill(1975)], Hill did not investigate the asymptotic behavior of the estimator. It was Mason who proved the weak consistency in [Mason(1982)], The strong consistency was proved by [Deheuvels, Häusler and Mason(1988)] who gave an optimal rate of convergence for an appropriately chosen sequence k_n. The asymptotic normality was established, under some extra condition on F, in several papers such as, e.g [Csörgo and Mason(1985)] and [Davis and Resnick(1984)].

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

Theoreme 2.4.2 (Asymptotic Properties of $\widehat{\gamma}^{(H)}$) *Assume that* $F \in D\left(\Phi_{1/\gamma}\right)$, $\gamma > 0$, $k \to \infty$ *and* $k/n \to 0$ *as* $n \to \infty$.

(a) *Weak Consistency:*

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

(b) *Strong consistency: If* $k/\log\log n \to \infty$ *as* $n \to \infty$, *then*

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

(c) *Asymptotic normality: Assume that* F *satisfies* (2.7). *If* $\sqrt{k}A(n/k) \to \lambda$ *as* $n \to \infty$, *then*

$$\sqrt{k}\left(\widehat{\gamma}^{(H)} - \gamma\right) \xrightarrow{d} N\left(\frac{\lambda}{1-\rho}, \gamma^2\right) \text{ as } n \to \infty.$$

2.4.3 Moment estimator

The moment estimator has been introduced by [Dekkers et al.(1989)] as a direct generalization of the Hill estimator, is similar to the Hill estimator but one that can be used for general $\gamma \in \mathbb{R}$, not only for $\gamma > 0$.

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

where

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

Because $M_n^{(1)}$ and $M_n^{(2)}$ can be interpreted as empirical moments, $\widehat{\gamma}^{(M)}$ is also referred to as amoment estimator of γ.

Theoreme 2.4.3 (Asymptotic Properties of $\widehat{\gamma}^{(M)}$) *Assume that $F \in D(H_\gamma)$, $\gamma \in \mathbb{R}$, $k \to \infty$ and $k/n \to 0$ as $n \to \infty$.*

(a) *Weak Consistency:*

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

(b) *Strong consistency: If $k/(\log n)^\delta \to \infty$ as $n \to \infty$, for some $\delta > 0$, then*

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

(c) *Asymptotic normality: (see Theorem 3.1 and Corollary 3.2 of [Dekkers et al.(1989)]),*

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

where

$$\eta^2 := \begin{cases} 1 + \gamma^2, & \gamma \geq 0 \\ \frac{\left(1-\gamma^2\right)(1-2\gamma)\left(1-\gamma+6\gamma^2\right)}{(4\gamma-1)(3\gamma-1)}, & \gamma < 0. \end{cases}$$

2.4.4 Kernel type estimators

In 1985, using a kernel function K, Csörgő, Deheuvels and Mason proposed a smoother version of Hill's estimator and proved its consistency and asymptotic normality. To define their estimator we need a kernel function K that satisfies the following condition: non-negative, non-increasing and right continuous function on $(0,\infty)$ such that $\int_0^\infty K(u)\,du = 1$ and $\int_0^\infty u^{-1/2}K(u)\,du < \infty$. Then the estimator is defined as

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

where $h > 0$ is called bandwidth. Notice that, using the uniform kernel $K = I_{(0,1)}$ and $h = k/n$ yields Hill's estimator $\widehat{\gamma}^{(H)}$ as a special case. This estimator depends in a continuous way on the bandwidth h representing the proportion of top order statistics used. Under von Mises's condition, the kernel type estimators have been generalized by [Groeneboom *et al.*(2003)] for all real tail indices.

To be able to state the asymptotic normality of the kernel type estimator, we will need some additional conditions on the kernel K, for a discussion of these conditions we refer to [Csörgö et al.(1985)].

Remark 2.4.1 *Recall that under appropriate conditions, the Hill estimator is consistent only for positive values of γ, the Pickands and moment estimators are defined and consistent for all real values of γ. For instance, we shall see that for an important range of values of γ,the Pickands estimator has larger asymptotic variance than the others. Some simulation results for some common distributions are given in [de Haan and Ferreira(2006)] p116. On Figure 2.3, Pickands estimators appears to be the least stable.*

2.5 Optimal sample fraction selection

Most widely used semi-parametric estimators of the extreme value parameter depend on the number of upper extremes which locate where the tail of a distribution begins. In the presence of a random sample with finite size, the problem concerning the choice of the number of upper extremes is not easy to handle. This number k is not only

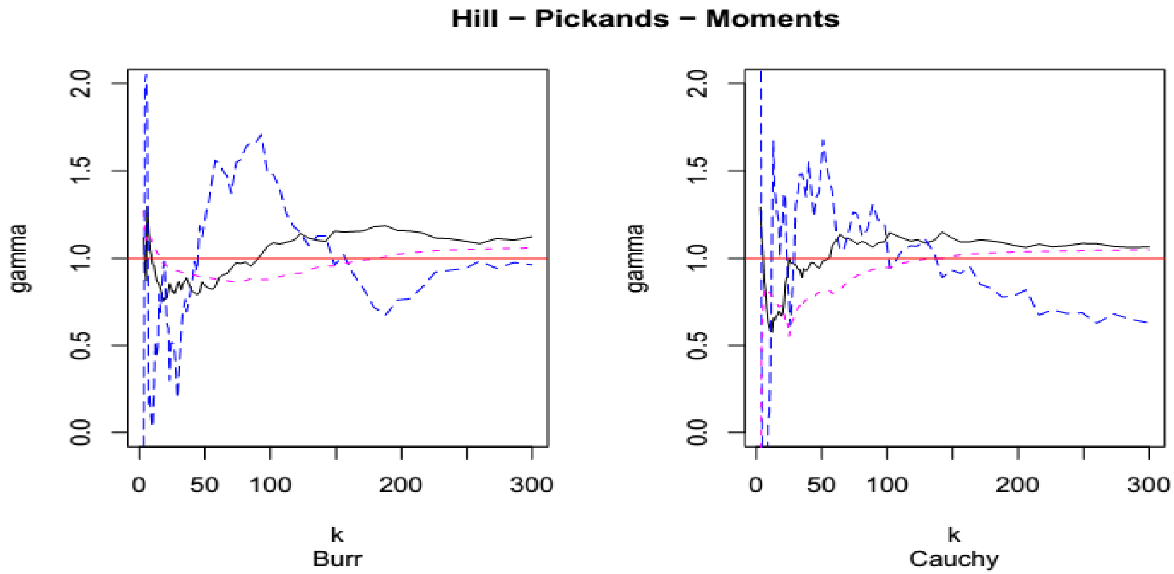


Figure 2.3: Hill (solid line), Pickands (dashed line) and the moment (dotted line) estimators for the EVI of the Burr(1,1,1) (left) and standard Cauchy (right) distributions, based on 100 samples of 3000 observations.

governed by the sample size n, but also ruled by parameters characterizing F. When the underlying distribution function is known, the optimum value k can be attained through the minimization of the asymptotic mean squared error of the considered estimator. In this section, we present some of the proposed methods in order to get an optimal number k. In this thesis we are interesting only with algorithm of Reiss and Thomas for esteablished this fraction.

2.5.1 Graphical method

We start by presenting a universal graphical method which should be applied prior to any numerical investigation. The method consists of using the plot $\{(k, \widehat{\gamma}), k = 1, ..., n\}$, in order to make an optimal choice of k. where it is clear that one should choose k in the (first) region where the plot is roughly horizontal. Some other graphical procedures for selecting an optimal k-value are extensively discussed and compared in [Sousa(2002)]. For an illustration see Figure 2.4, where it seems that any k between 80 and 100 would be a good choice.

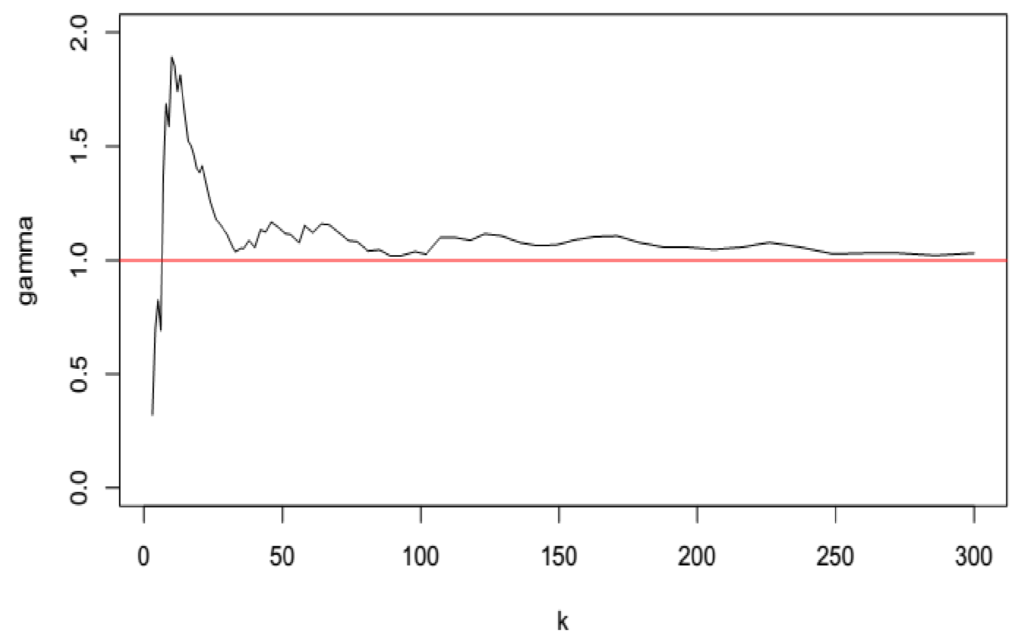


Figure 2.4: Plot of Hill's estimator, for the EVI of a standard Pareto distrubution, as a function of the number of top statistics, based on 100 samples of size 3000. The horizontal line represents the true value of the tail index.

2.5.2 Adaptive procedures

A large variety of algorithms and data-adaptive procedures of computing consistent estimate $\widehat{k}_{opt}$ for k_{opt} in the sense that $\frac{\widehat{k}_{opt}}{k_{opt}} \overset{P}{\to} 1$ as $n \to \infty$. In the remainder of this subsection, we will outline some of the most known data-driven methods of choosing the number of largest statistics suitable for an accurate estimation, for example: Hall and Welsh approach, Bootstrap approach, Sequential approach, Coverage accuracy approach,Cheng and Peng approach, Reiss and Thomas approach, for more details see thesis of Pr. [Meraghni(2008)] p63. In this thesis, to determine the optimal number of upper order statistics used in the computation of $\widehat{\gamma}$, we apply the algorithm of page 137 in [Reiss and Thomas(2007)].

Reiss and Thomas approach

They propose an automatic manner to choose k by minimizing

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

or the following suggested modification

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

In our simulation study, we apply this procedure and we choose $\theta = 0.3$, for a discussion on the choice of θ, one refers to the paper of [Neves and Fraga Alves(2004)].

Chapter 3

On the asymptotic normality of the EVI for right-truncated data

We introduce in this chapter, a consistent estimator of the extreme value index under random truncation based on a single sample fraction of top observations from truncated and truncation data. We establish the asymptotic normality of the proposed estimator by making use of the weighted tail-copula process framework.

3.1 Tail index estimation

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 indices $-1/\gamma_1$ and $-1/\gamma_2$. That is, for any $s > 0$

$$\lim_{x\to\infty} \frac{\overline{\mathbf{F}}(sx)}{\overline{\mathbf{F}}(x)} = s^{-1/\gamma_1} \text{ and } \lim_{y\to\infty} \frac{\overline{\mathbf{G}}(sy)}{\overline{\mathbf{G}}(y)} = s^{-1/\gamma_2}. \tag{3.1}$$

Being characterized by their heavy tails, these distributions play a prominent role in extreme value theory. They include distributions such as Pareto, Burr, Fréchet, stable and log-gamma, known to be appropriate models for fitting large insurance claims, log-returns, large fluctuations, etc. see, e.g., [Resnick(2006)].

Making use of Proposition B.1.10 in [de Haan and Ferreira(2006)], for the regularly varying functions $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$, we may readily show that both $\overline{G}$ and $\overline{F}$ are regularly varying at infinity as well, with respective indices γ_2 and $\gamma := \gamma_1\gamma_2/(\gamma_1+\gamma_2)$. That is, we have, for any $s>0$,

$$\lim_{x\to\infty}\frac{\overline{F}(sx)}{\overline{F}(x)}=s^{-1/\gamma}\text{ and }\lim_{y\to\infty}\frac{\overline{G}(sy)}{\overline{G}(y)}=s^{-1/\gamma_2}. \tag{3.2}$$

Recently [Gardes and Stupfler(2015)] addressed the estimation of the extreme value index γ_1 under random truncation. They used the definition of γ to derive the following consistent estimator:

$$\widehat{\gamma}_1(k,k') := \frac{\widehat{\gamma}(k)\,\widehat{\gamma}_2(k')}{\widehat{\gamma}_2(k')-\widehat{\gamma}(k)},$$

where

$$\widehat{\gamma}(k) := \frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{n-i+1:n}}{X_{n-k:n}}\text{ and }\widehat{\gamma}_2(k') := \frac{1}{k'}\sum_{i=1}^{k'}\log\frac{Y_{n-i+1:n}}{Y_{n-k':n}}, \tag{3.3}$$

are the well-known Hill estimators of γ and γ_2, with $X_{1:n}\leq\ldots\leq X_{n:n}$ and $Y_{1:n}\leq\ldots\leq Y_{n:n}$ being the order statistics pertaining to the samples $(X_1,\ldots,X_n)$ and $(Y_1,\ldots,Y_n)$ respectively. The two sequences $k=k_n$ and $k'=k'_n$ of integer rv's, which satisfy

$$1<k,k'<n,\ \min(k,k')\to\infty\text{ and }\max(k/n,k'/n)\to 0\text{ as }n\to\infty,$$

respectively represent the numbers of top observations from truncated and truncation data. By considering the two situations $k/k'\to 0$ and $k'/k\to 0$ as $n\to\infty$, the authors established the asymptotic normality of $\widehat{\gamma}_1(k,k')$, but when $k/k'\to 1$, they only showed, in Theorem 3, that $\sqrt{\min(k,k')}\,(\widehat{\gamma}_1(k,k')-\gamma_1)=O_{\mathbf{p}}(1)$, as $n\to\infty$. It is obvious that an accurate computation of the estimate $\widehat{\gamma}_1(k,k')$ requires good choices of both k and k'. However from a practical point of view, it is rather unusual in extreme value analysis to handle two distinct sample fractions simultaneously, which is mentioned by [Gardes and Stupfler(2015)] in their conclusion as well. In the present work, we consider the situation when $k = k'$ (rather than $k/k' \to 1$), to obtain an estimator

$$\widehat{\gamma}_1 := \widehat{\gamma}_1(k) = k^{-1} \frac{\sum_{i=1}^{k} \log \frac{X_{n-i+1:n}}{X_{n-k:n}} \sum_{i=1}^{k} \log \frac{Y_{n-i+1:n}}{Y_{n-k:n}}}{\sum_{i=1}^{k} \log \frac{X_{n-k:n} Y_{n-i+1:n}}{Y_{n-k:n} X_{n-i+1:n}}}, \tag{3.4}$$

of simpler form, expressed in terms of a single sample fraction k of truncated and truncation observations. Thereby, the number of extreme values used to compute the optimal estimate value $\widehat{\gamma}_1$ may be obtained by applying one of the various heuristic methods available in the literature such that, for instance, the algorithm of page 137 in [Reiss and Thomas(2007)]. This estimator is used by [Gardes and Stupfler(2015)] in their simulation study (to evaluate the performance high quantile estimators) where they took $k = k'$ as it is mentioned in their conclusion. The task of establishing the asymptotic normality of $\widehat{\gamma}_1$ is a bit delicate as one has to take into account the dependence structure of X and Y. The authors of [Gardes and Stupfler(2015)] handled this issue by putting conditions on the sample fractions k and k'. In our case we require that the joint df H have a stable tail dependence function ℓ (see [Huang(1992)] and [Drees and Huang(1998)]), in the sense that the following limit exists:

$$\lim_{t \downarrow 0} t^{-1} \mathbf{P}\left(\overline{F}(X) \leq tx \text{ or } \overline{G}(Y) \leq ty\right) =: \ell(x, y), \tag{3.5}$$

for all $x, y \geq 0$ such that $\max(x, y) > 0$. Note that the corresponding tail copula function is defined by

$$\lim_{t \downarrow 0} t^{-1} \mathbf{P}\left(\overline{F}(X) \leq tx,\ \overline{G}(Y) \leq ty\right) =: R(x, y), \tag{3.6}$$

which equals $x + y - \ell(x, y)$. In on other words, we assume that H belongs to the domain of attraction of a bivariate extreme value distribution. This may be split into two sets of conditions, namely conditions for the convergence of the marginal distributions (3.2) and others for the convergence of the dependence structure (3.5). For details on this topic, see for instance Section 6.1.2 in [de Haan and Ferreira(2006)] and the papers of [Huang(1992)] and [Einmahl et al.(2006)], [de Haan *et al.*(2008)].

3.2 Main results

Weak approximations of extreme value theory based statistics are achieved in the second-order framework, see [de Haan and Stadtmüller(1996)]. Thus, it seems quite natural to suppose that both df's F and G satisfy the well-known second-order condition of regular variation. That is, we assume that for any $x > 0$

$$\begin{aligned}\lim_{z\to\infty}\frac{1}{A(z)}\left(\frac{\mathbb{U}(zx)}{U(z)}-x^{\gamma}\right)&=x^{\gamma}\frac{x^{\tau}-1}{\tau},\\ \lim_{z\to\infty}\frac{1}{A_2(z)}\left(\frac{\mathbb{U}_2(zx)}{\mathbb{U}_2(z)}-x^{\gamma_2}\right)&=x^{\gamma_2}\frac{x^{\tau_2}-1}{\tau_2},\end{aligned}\tag{3.7}$$

where $\mathbb{U} := \left(1/\overline{F}\right)^{\leftarrow}$, $\mathbb{U}_2 := \left(1/\overline{G}\right)^{\leftarrow}$ (with $E^{\leftarrow}(u) := \inf\{v : E(v) \geq u\}$, for $0 < u < 1$, denoting the quantile function pertaining to a function E), $|A|$ and $|A_2|$ are some regularly varying functions with negative indices (second-order parameters) τ and τ_2 respectively.

Theoreme 3.2.1 *Assume that the second-order regular variation condition* (3.7) *and* (3.5) *hold. Let $k := k_n$ be a sequence of integers such that $k \to \infty$, $k/n \to 0$. Then, there exist two standard Wiener processes $\{W_i(t),\ t \geq 0\}$, $i = 1, 2$, defined on the probability space $(\Omega, \mathcal{A}, \mathbf{P})$ with covariance function $R(\cdot,\cdot)$, and two functions $A^* \sim A$ and $A_2^* \sim A_2$ with $\sqrt{k}A^*(n/k) = O(1) = \sqrt{k}A_2^*(n/k)$, such that*

$$\begin{aligned}&\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)-\mu(k)\\ &=\int_0^1 t^{-1}\left(cW_1(t)-c_2W_2(t)\right)dt-cW_1(1)+c_2W_2(1)+o_{\mathbf{p}}(1),\end{aligned}$$

where $c := \gamma_1^2/\gamma$, $c_2 := \gamma_1^2/\gamma_2$ and

$$\mu(k) := \frac{c\sqrt{k}A^*(n/k)}{\gamma(1-\tau)} + \frac{c_2\sqrt{k}A_2^*(n/k)}{\gamma_2(1-\tau_2)}.$$

If in addition we have $\sqrt{k}A^{*}(n/k) \to \lambda$ *and* $\sqrt{k}A_{2}^{*}(n/k) \to \lambda_{2}$, *then*

$$\sqrt{k}\left(\widehat{\gamma}_{1}-\gamma_{1}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mu, \sigma^{2}\right), \text{ as } n \to \infty,$$

where

$$\mu := \frac{c\lambda}{\gamma(1-\tau)} + \frac{c_{2}\lambda_{2}}{\gamma_{2}(1-\tau_{2})} \text{ and } \sigma^{2} := 2c^{2} + 2c_{2}^{2} - 2cc_{2}\delta,$$

with

$$\delta = \delta(R) := \int_{0}^{1}\int_{0}^{1}\frac{R(s,t)}{st}dsdt - \int_{0}^{1}\left(R(s,1) - R(1,s)\right)ds + R(1,1).$$

Remark 3.2.1 *Note that* σ^{2} *is finite. Indeed, the fact that* $\ell(x,y)$ *is a tail copula function, implies that* $\max(x,y) \le \ell(x,y) \le x+y$ *see, e.g., [Gudendorf and Segers(2010)] and since* $R(x,y) = x + y - \ell(x,y)$, *then* $0 \le R(x,y) \le \min(x,y)$. *It follows that*

$$\int_{0}^{1}\int_{0}^{1}\frac{R(s,t)}{st}dsdt \le \int_{0}^{1}\int_{0}^{1}\frac{\min(s,t)}{st}dsdt = 2,$$

$$\int_{0}^{1} R(s,1)\,ds \le \frac{1}{2}, \quad \int_{0}^{1} R(1,s)\,ds \le \frac{1}{2} \text{ and } R(1,1) \le 1.$$

Therefore $|\delta| \le 4$, *which yields that* $\sigma^{2} < \infty$.

The following corollary directly leads to a practical construction of confidence intervals for the tail index γ_{1}.

Corollary 3.2.1 *Under the assumptions of Theorem 3.2.1, we have*

$$\frac{\sqrt{k}\left(\widehat{\gamma}_{1}-\gamma_{1}\right) - \widehat{\mu}}{\widehat{\sigma}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \text{ as } n \to \infty,$$

where $\widehat{\mu} = \dfrac{\widehat{c}\widehat{\lambda}}{\widehat{\gamma}(1-\widehat{\tau})} + \dfrac{\widehat{c}_{2}\widehat{\lambda}_{2}}{\widehat{\gamma}_{2}(1-\widehat{\tau}_{2})}$ *and* $\widehat{\sigma}^{2} := 2\widehat{c}^{2} + 2\widehat{c}_{2}^{2} - 2\widehat{c}\widehat{c}_{2}\widehat{\delta}$, *with*

$$\widehat{c} := \widehat{\gamma}_{1}^{2}/\widehat{\gamma}, \ \widehat{c}_{2} := \widehat{\gamma}_{1}^{2}/\widehat{\gamma}_{2}, \ \widehat{\delta} := \delta\left(\widehat{R}\right),$$

$$\widehat{\lambda} := \sqrt{k}\widehat{\tau}\frac{X_{n-2k:n} - 2^{-\widehat{\gamma}}X_{n-k:n}}{2^{-\widehat{\gamma}}\left(2^{-\widehat{\tau}}-1\right)X_{n-k:n}} \text{ and } \widehat{\lambda}_{2} := \sqrt{k}\widehat{\tau}_{2}\frac{Y_{n-2k:n} - 2^{-\widehat{\gamma}_{2}}Y_{n-k:n}}{2^{-\widehat{\gamma}_{2}}\left(2^{-\widehat{\tau}_{2}}-1\right)Y_{n-k:n}}.$$

Here $\widehat{\gamma}$ and $\widehat{\gamma}_2$ are the respective Hill estimators of γ and γ_2 defined in (3.3) *with $k' = k$, $\widehat{\tau}$ (resp. $\widehat{\tau}_2$) is one of the estimators of τ (resp. τ_2) see, e.g., [Gomes and Pestana(2007)] and $\widehat{R}$ is a nonparametric estimator of R given in [Peng(2010)] by*

$$\widehat{R}(s,t) := k^{-1} \sum_{i=1}^{n} \mathbf{1}\left(X_i \geq X_{n-[ks]:n},\ Y_i \geq Y_{n-[kt]:n}\right),$$

with $[x]$ standing for the integer part of the real number x and $\mathbf{1}(\cdot)$ for the indicator function.

3.3 Proofs

Proof of Theorem 3.2.1

We begin by a brief introduction on the weak approximation of a weighed tail copula process given in Proposition 1 of [Einmahl et al.(2006)]. Set $U_i := \overline{F}(X_i)$ and $V_i := \overline{G}(Y_i)$, for $i = 1, ..., n$, and let $C(x, y)$ be the joint df of (U_i, V_i). The copula function C and its corresponding tail R, defined in (3.6), are linked by $t^{-1}C(tx, ty) - R(x, y) = O(t^{\epsilon})$, as $t \downarrow 0$, for some $\epsilon > 0$, uniformly for $x, y \geq 0$ and $\max(x, y) \leq 1$ [Huang(1992)]. Let us define

$$\upsilon_n(x, y) := \sqrt{k}\left(\mathbf{T}_n(x, y) - R_n(x, y)\right),\ x, y > 0,$$

where

$$\mathbf{T}_n(x, y) := \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}\left(U_i < \frac{k}{n}x,\ V_i < \frac{k}{n}y\right) \text{ and } R_n(x, y) := \frac{n}{k}C\left(\frac{kx}{n}, \frac{ky}{n}\right).$$

In the sequel, we will need the following two empirical processes:

$$\alpha_n(x) := \upsilon_n(x, \infty) = \sqrt{k}\left(\mathbf{U}_n(x) - x\right) \text{ and } \beta_n(y) := \upsilon_n(\infty, y) = \sqrt{k}\left(\mathbf{V}_n(y) - y\right),$$

where

$$\mathbf{U}_n(x) := \mathbf{T}_n(x, \infty) = \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}\left(U_i < \frac{k}{n}x\right),$$

and

$$\mathbf{V}_n(y) := \mathbf{T}_n(\infty, y) = \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}\left(V_i < \frac{k}{n}y\right).$$

From assertions (3.8) and (3.9) in [Einmahl et al.(2006)], there exists a Gaussian process $W_R(x, y)$, defined on the probability space $(\Omega, \mathcal{A}, \mathbf{P})$, with mean zero and covariance

$$\mathbf{E}\left[W_R(x_1, y_1) W_R(x_2, y_2)\right] = R(\min(x_1, x_2), \min(y_1, y_2)), \tag{3.8}$$

such that, for any $M > 0$,

$$\sup_{0<x,y\leq M} \frac{|\upsilon_n(x, y) - W_R(x, y)|}{\{\max(x, y)\}^{\eta}} = o_{\mathbf{p}}(1),$$

and

$$\sup_{0<x\leq M} \frac{|\alpha_n(x) - W_1(x)|}{x^{\eta}} = o_{\mathbf{p}}(1) = \sup_{0<y\leq M} \frac{|\beta_n(y) - W_2(y)|}{y^{\eta}}, \tag{3.9}$$

as $n \to \infty$, for any $0 \leq \eta < 1/2$, where

$$W_1(x) := W_R(x, \infty) \text{ and } W_2(y) := W_R(\infty, y),$$

are two standard Wiener processes such that $\mathbf{E}\left[W_1(x) W_2(y)\right] = R(x, y)$. To prove our result, we will write the tail index estimator $\widehat{\gamma}_1$ in terms of the processes $\alpha_n(\cdot)$ and $\beta_n(\cdot)$. We start by splitting $\widehat{\gamma}_1 - \gamma_1$ into the sum of

$$T_{n1} := \frac{\widehat{\gamma}_2(\gamma_2 - \gamma) + \gamma_2\gamma}{(\widehat{\gamma}_2 - \widehat{\gamma})(\gamma_2 - \gamma)}(\widehat{\gamma} - \gamma) \text{ and } T_{n2} := -\frac{\gamma^2}{(\widehat{\gamma}_2 - \widehat{\gamma})(\gamma_2 - \gamma)}(\widehat{\gamma}_2 - \gamma_2).$$

The consistency of both Hill estimators $\widehat{\gamma}$ and $\widehat{\gamma}_2$ yields that

$$T_{n1} = \frac{c}{\gamma}(\widehat{\gamma} - \gamma)(1 + o_{\mathbf{p}}(1)) \text{ and } T_{n2} = -\frac{c_2}{\gamma_2}(\widehat{\gamma}_2 - \gamma_2)(1 + o_{\mathbf{p}}(1)),$$

where c and c_2 are those defined in Theorem 3.2.1. On the other hand, we have, from Theorem 3.2.5 in [de Haan and Ferreira(2006)], $\widehat{\gamma}-\gamma=O_{\mathbf{p}}\left(k^{-1/2}\right)=\widehat{\gamma}_2-\gamma_2$, as $n\rightarrow\infty$. Since $k\rightarrow\infty$, it follows that

$$\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)=\frac{c}{\gamma}\sqrt{k}\left(\widehat{\gamma}-\gamma\right)-\frac{c_2}{\gamma_2}\sqrt{k}\left(\widehat{\gamma}_2-\gamma_2\right)+o_{\mathbf{p}}\left(1\right). \tag{3.10}$$

Next, we represent $\sqrt{k}\left(\widehat{\gamma}-\gamma\right)$ and $\sqrt{k}\left(\widehat{\gamma}_2-\gamma_2\right)$ in terms of $\alpha_n\left(\cdot\right)$ and $\beta_n\left(\cdot\right)$ respectively. By using the first-order condition of regular variation of $\overline{F}$ (3.2) and applying Theorem 1.2.2 in [de Haan and Ferreira(2006)] we get $\lim_{n\rightarrow\infty}\frac{n}{k}\int_{a_k}^{\infty}t^{-1}\overline{F}\left(t\right)dt=\gamma$, with $a_k:=\mathbb{U}\left(n/k\right).$ This allows us to write $\widehat{\gamma}=\frac{n}{k}\int_{X_{n-k:n}}^{\infty}t^{-1}\overline{F}_n\left(t\right)dt$. Now, we consider the following decomposition $\widehat{\gamma}-\gamma=S_{n1}+S_{n2}+S_{n3}$, where

$$S_{n1}:=\frac{n}{k}\int_{X_{n-k:n}}^{\infty}t^{-1}\left(\overline{F}_n\left(t\right)-\overline{F}\left(t\right)\right)dt,\ S_{n2}:=-\frac{n}{k}\int_{a_k}^{X_{n-k:n}}t^{-1}\overline{F}\left(t\right)dt$$

and $S_{n3}:=\frac{n}{k}\int_{a_k}^{\infty}t^{-1}\overline{F}\left(t\right)dt-\gamma$. It is easy to verify that, almost surely, we have $\frac{n}{k}\overline{F}_n\left(t\right)=\mathbf{U}_n\left(\frac{n}{k}\overline{F}\left(t\right)\right)$, it follows, after a change of variables, that

$$S_{n1}=\int_1^{\infty}t^{-1}\left(\mathbf{U}_n\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)-\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)dt.$$

In other words $\sqrt{k}S_{n1}=\int_1^{\infty}t^{-1}\alpha_n\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)dt$, which may be decomposed into

$$\begin{aligned}&\int_1^{\infty}t^{-1}W_1\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)dt\\&+\int_1^{\infty}t^{-1}\left\{\alpha_n\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)-W_1\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)\right\}dt.\end{aligned}$$

Let $0<\eta<1/2$ and apply approximation (3.9), to the second term above, to get

$$\sqrt{k}S_{n1}=\int_1^{\infty}t^{-1}W_1\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)dt+o_{\mathbf{p}}\left(1\right)\int_1^{\infty}t^{-1}\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)^{\eta}dt.$$

Note that $k/n=\overline{F}(a_k)$ and write

$$\int_1^\infty t^{-1}\left(\frac{n}{k}\overline{F}(tX_{n-k:n})\right)^\eta dt=\left(\frac{\overline{F}(X_{n-k:n})}{\overline{F}(a_k)}\right)^\eta\int_1^\infty t^{-1}\left(\frac{\overline{F}(tX_{n-k:n})}{\overline{F}(X_{n-k:n})}\right)^\eta dt.$$

Let $\epsilon>0$ be sufficiently small. From Potter's inequalities to $\overline{F}$, see, e.g., Proposition B.1.9, assertion 5 in [de Haan and Ferreira(2006)], we have, for all large n and for any $t>1$,

$$(1-\epsilon)\,t^{-1/\gamma-\epsilon}\leq\frac{\overline{F}(tX_{n-k:n})}{\overline{F}(X_{n-k:n})}\leq(1+\epsilon)\,t^{-1/\gamma+\epsilon}.$$

This implies that

$$\int_1^\infty t^{-1}\left(\frac{\overline{F}(X_{n-k:n}t)}{\overline{F}(X_{n-k:n})}\right)^\eta dt=O_{\mathbf{p}}(1)\int_1^\infty t^{\eta\epsilon-\eta/\gamma-1}dt.$$

On the other hand, by combining Corollary 2.2.2 with Proposition B.1.10 in [de Haan and Ferreira(2006)] (applied to $\mathbb{U}$), we show that $X_{n-k:n}/a_k\overset{\mathbf{P}}{\to}1$ as $n\to\infty$. By, once again, using Potter's inequalities above, we infer that $\overline{F}(X_{n-k:n})/\overline{F}(a_k)\overset{\mathbf{P}}{\to}1$ as well. Since $\int_1^\infty t^{\eta\epsilon-\eta/\gamma-1}dt=1/(\eta/\gamma-\eta\epsilon)<\infty$, then

$$\sqrt{k}S_{n1}=(1+o_{\mathbf{p}}(1))\int_1^\infty t^{-1}W_1\left(\frac{n}{k}\overline{F}(tX_{n-k:n})\right)dt+o_{\mathbf{p}}(1).$$

Let us decompose the previous integral into

$$\int_1^\infty t^{-1}W_1\left(t^{-1/\gamma}\right)dt+\int_1^\infty t^{-1}\left\{W_1\left(\frac{n}{k}\overline{F}(tX_{n-k:n})\right)-W_1\left(t^{-1/\gamma}\right)\right\}dt.$$

Next, we show that the second term is negligible in (probability). To this end, we write

$$\begin{aligned}&\left|\int_1^\infty t^{-1}\left\{W_1\left(\frac{n}{k}\overline{F}(tX_{n-k:n})\right)-W_1\left(t^{-1/\gamma}\right)\right\}dt\right|\\&\leq\int_1^\infty t^{-1}\left|W_1\left(\frac{n}{k}\overline{F}(tX_{n-k:n})\right)-W_1\left(t^{-1/\gamma}\right)\right|dt.\end{aligned}$$

From Levy's modulus of continuity of the Wiener process, see, e.g., Theorem 1.1.1 in [Csörgő and Révész, 1981], we have almost surely for all large n

$$\left|W_1\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)-W_1\left(t^{-1/\gamma}\right)\right|\leq\left(1+\epsilon\right)\sqrt{2h_n\left(t\right)\log\left(1/h_n\left(t\right)\right)}, \tag{3.11}$$

uniformly on $t>1$, where $h_n\left(t\right):=\left|\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)-t^{-1/\gamma}\right|$ that we show to be equal to $o_{\mathbf{p}}\left(1\right)$. Indeed, we have

$$\begin{aligned} h_n\left(t\right) &= \left|\frac{\overline{F}\left(tX_{n-k:n}\right)}{\overline{F}\left(a_k\right)}-t^{-1/\gamma}\right| \\ &\leq \left|\frac{\overline{F}\left(tX_{n-k:n}\right)}{\overline{F}\left(a_k\right)}-\left(t\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma}\right|+t^{-1/\gamma}\left|\left(\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma}-1\right|, \end{aligned}$$

and in view of Proposition B.1.10 in [de Haan and Ferreira(2006)], we may write

$$\left|\frac{\overline{F}\left(tX_{n-k:n}\right)}{\overline{F}\left(a_k\right)}-\left(t\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma}\right|<\epsilon\left(t\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma\pm\epsilon},$$

for all large n and $t>1$. In other words, we have

$$\ell_n\left(t\right)<\left\{\epsilon\left(\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma\pm\epsilon}+\left|\left(\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma}-1\right|\right\}t^{-1/\gamma}.$$

Since $X_{n-k:n}/a_k=1+o_{\mathbf{p}}\left(1\right),$ then $h_n\left(t\right)=\left(\epsilon+o_{\mathbf{p}}\left(1\right)\right)t^{-1/\gamma}$, uniformly on $t>1$, this means that $\sup_{t>1}h_n\left(t\right)\rightarrow0$ as $n\rightarrow\infty$. Going back to $\left(3.11\right)$, we use the fact that, $\log u<\epsilon u^{-\epsilon}$ as $u\downarrow0$, we end up with

$$\sqrt{h_n\left(t\right)\log\left(1/h_n\left(t\right)\right)}=O_{\mathbf{p}}\left(1\right)\epsilon^{1/2}\left(\epsilon+o_p\left(1\right)\right)^{\left(1-\epsilon\right)/2}t^{-\left(1/\gamma+\epsilon\right)/2}.$$

It follows that

$$\int_1^\infty t^{-1}\left\{W_1\left(\frac{n}{k}\overline{F}\left(tX_{n-k:n}\right)\right)-W_1\left(t^{-1/\gamma}\right)\right\}dt$$
$$=O_{\mathbf{p}}\left(1\right)\epsilon^{1/2}\left(\epsilon+o_{\mathbf{p}}\left(1\right)\right)^{(1-\epsilon)/2}\int_1^\infty t^{-(1/\gamma+\epsilon)/2-1}dt.$$

We have $\int_1^\infty t^{-(1/\gamma+\epsilon)/2-1}dt$ is finite and $\epsilon^{1/2}\left(\epsilon+o_{\mathbf{p}}\left(1\right)\right)^{(1-\epsilon)/2}\xrightarrow{\mathbf{P}}0$ as $\varepsilon\downarrow 0$, therefore

$$\sqrt{k}S_{n1}=\left(1+o_{\mathbf{p}}\left(1\right)\right)\int_1^\infty t^{-1}W_1\left(t^{-1/\gamma}\right)dt+o_{\mathbf{p}}\left(1\right).$$

It is easy to verify that $\mathbf{E}\left|\int_1^\infty t^{-1}W_1\left(t^{-1/\gamma}\right)dt\right|\leq 2\gamma$, then after a change of variables we get

$$\sqrt{k}S_{n1}=\gamma\int_0^1 s^{-1}W_1\left(s\right)ds+o_{\mathbf{p}}\left(1\right). \tag{3.12}$$

As for the second term S_{n2}, we use the mean value theorem to get

$$S_{n2}=-\frac{n}{k}\left(X_{n-k:n}-a_k\right)z_n^{-1}\overline{F}\left(z_n\right),$$

where z_n is a sequence of rv's lying between $X_{n-k:n}$ and a_k. Observe that

$$S_{n2}=-\frac{\overline{F}\left(z_n\right)}{\overline{F}\left(a_k\right)}\frac{a_k}{z_n}\left(\frac{X_{n-k:n}}{a_k}-1\right).$$

Since $X_{n-k:n}/a_k\xrightarrow{\mathbf{P}}1$, then $z_n/a_k\xrightarrow{\mathbf{P}}1$ and $\frac{n}{k}\overline{F}\left(z_n\right)\xrightarrow{\mathbf{P}}1$. It follows that

$$S_{n2}=-\left(1+o_{\mathbf{p}}\left(1\right)\right)\left(\frac{X_{n-k:n}}{a_k}-1\right).$$

Recall that $U_i=\overline{F}\left(X_i\right)$ and note that $U_{i:n}=\overline{F}\left(X_{n-i+1:n}\right)$, therefore

$$S_{n2}=-\left(1+o_{\mathbf{p}}\left(1\right)\right)\left(\frac{\mathbb{U}\left(1/U_{k+1:n}\right)}{\mathbb{U}\left(n/k\right)}-1\right).$$

By using Proposition B.1.10 in [de Haan and Ferreira(2006)] (applied to $\mathbb{U}$) together with the mean value theorem, we write $S_{n2} = (1 + o_{\mathbf{p}}(1))\,\gamma\left(\frac{n}{k}U_{k+1:n} - 1\right).$ Since $\mathbf{U}_n\left(\frac{n}{k}U_{k+1:n}\right) = 1$, then

$$\sqrt{k}S_{n2} = -\left(1 + o_{\mathbf{p}}(1)\right)\gamma\alpha_n\left(\frac{n}{k}U_{k+1:n}\right).$$

By applying approximation (3.9) and the stochastic boundednees of W_1, we get $\sqrt{k}S_{n2} = -\gamma W_1\left(\frac{n}{k}U_{k+1:n}\right) + o_{\mathbf{p}}(1)$. Recall that, from Corollary 2.2.2 in [de Haan and Ferreira(2006)], we have $\frac{n}{k}U_{k+1:n} \overset{\mathbf{P}}{\to} 1$, then by using similar arguments based on Levy's modulus of continuity of the Wiener process, we show that

$$\sqrt{k}S_{n2} = -\gamma W_1(1) + o_{\mathbf{p}}(1). \tag{3.13}$$

By summing up (3.12) and (3.13), we get

$$\sqrt{k}\left(S_{n1} + S_{n2}\right) = \gamma\int_0^1 s^{-1}W_1(s)\,ds - \gamma W_1(1) + o_{\mathbf{p}}(1).$$

For the third term S_{n3}, it suffices to use the inequality (2.3.23) of Theorem 2.3.9 in [de Haan and Ferreira(2006)] to get

$$\sqrt{k}S_{n3} = (1 + o(1))\,\frac{\sqrt{k}A^*(n/k)}{1-\tau} \text{ as } n \to \infty,$$

for a suitable function $A^* \sim A$. In summary, by using the fact that $\sqrt{k}A^*(n/k) = O(1)$, we obtain

$$\sqrt{k}\left(\widehat{\gamma} - \gamma\right) = \gamma\int_0^1 s^{-1}W_1(s)\,ds - \gamma W_1(1) + \frac{\sqrt{k}A^*(n/k)}{1-\tau} + o_{\mathbf{p}}(1). \tag{3.14}$$

Likewise, we write $\widehat{\gamma}_2 = \frac{n}{k}\int_{Y_{n-k:n}}^{\infty} t^{-1}\overline{G}_n(t)\,dt$, where $G_n(x) := n^{-1}\sum_{i=1}^{n}\mathbf{1}(Y_i \leq x)$ is the usual empirical df based on the fully observed sample $(Y_1, ..., Y_n)$. Then, by using similar arguments, we express $\widehat{\gamma}_2$ in terms of the process $\beta_n(\cdot)$, whose approximation (3.9) yields

$$\sqrt{k}\left(\widehat{\gamma}_2-\gamma_2\right)=\gamma_2\int_0^1 s^{-1}W_2(s)\,ds-\gamma_2 W_2(1)+\frac{\sqrt{k}A_2^*(n/k)}{1-\tau_2}+o_{\mathbf{p}}(1), \tag{3.15}$$

where $A_2^*\sim A_2$. Finally, substituting results (3.14) and (3.15) in equation (3.10) achieves the proof of the first part of the theorem. Finally, some elementary calculations, using the covariance formula (3.8) and the fact that $E\left[\int_0^1 s^{-1}W_i(s)\,ds\right]^2=2$, $i=1,2$,straightforwardly lead to the asymptotic normality result. □

Proof of Corollary 3.2.1

It suffices to plug the estimate of each parameter in the result of Corollary 3.2.1. To estimate the limits λ and λ_2, we exploit the second-order conditions of regular variation (3.7). We have, as $z\rightarrow\infty$,

$$A(z)\sim\tau\frac{U(zx)/U(z)-x^{\gamma}}{x^{\gamma}\left(x^{\tau}-1\right)},\ \text{for any } x>0.$$

In particular, for $x=1/2$, and $z=n/k$, we have

$$A(n/k)\sim\tau\frac{U\left(\dfrac{n}{2k}\right)/U\left(\dfrac{n}{k}\right)-2^{-\gamma}}{2^{-\gamma}\left(2^{-\tau}-1\right)}.$$

Hence, we take

$$\widehat{A}(n/k)=\widehat{\tau}\frac{X_{n-2k:n}/X_{n-k:n}-2^{-\widehat{\gamma}}}{2^{-\widehat{\gamma}}\left(2^{-\widehat{\tau}}-1\right)}=\widehat{\tau}\frac{X_{n-2k:n}-2^{-\widehat{\gamma}}X_{n-k:n}}{2^{-\widehat{\gamma}}\left(2^{-\widehat{\tau}}-1\right)X_{n-k:n}},$$

an estimate of $A(n/k)$. Thus, the expression of $\widehat{\lambda}$ readily follows. The same idea applies to λ_2 as well. □

Chapter 4

Tail product-limit process for truncated data with application to EVI estimation

In this chapter a weighted Gaussian approximation to tail product-limit process for Pareto-like distributions of randomly right-truncated data is provided and a new consistent and asymptotically normal estimator of the extreme value index is introduced. A simulation study is carried out to evaluate the finite sample behavior of the proposed estimator and compare it to that recently proposed by [Gardes and Stupfler(2015)]. Also, a new approach of estimating extreme quantiles, under random right truncation, is derived and applied to a real dataset of lifetimes of automobile brake pads.

4.1 Tail product-limit process

In the present section, we introduce a tail product-limit process for which we provide a weighted Gaussian approximation as well. This tool will be very helpful when dealing with the estimation of any tail related quantity. In particular, it will lead to the asymptotic normality of the extreme value index estimator that we define, under random right-truncation, as a function of a single sample fraction of upper order statistics. But, prior to describing our estimation methodology, let us note that, as mentioned by [Gardes and Stupfler(2015)], in order to ensure that it remains enough extreme data for the inference to be accurate, we need to impose the condition $\gamma_1 < \gamma_2$. In other words, we consider the situation where the tail of the rv of interest $\mathbf{X}$ is not too contaminated by the truncation rv $\mathbf{Y}$.

Now, let $k = k_n$ be a (random) sequence of integers such that, given $n = m = m_N$,

$$1 < k_m < m,\ k_m \to \infty \text{ and } k_m/m \to 0 \text{ as } N \to \infty, \tag{4.1}$$

and introduce a tail product-limit process corresponding to $\mathbf{F}_n$ as follows:

$$\mathbf{D}_n(x) := \sqrt{k}\left(\frac{\overline{\mathbf{F}}_n(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} - x^{-1/\gamma_1}\right),\ x > 0, \tag{4.2}$$

where, given $n = m$, $X_{1:m} \leq ... \leq X_{m:m}$ denote the order statistics pertaining to $X_1, ..., X_m$. It is worth mentioning, since $n = n_N$ is a (random) sequence of integers tending in probability to ∞ as $N \to \infty$, then for any sequence of real numbers b_N, such that $b_N \to b$ (finite or not) as $N \to \infty$, we have $b_n \xrightarrow{\mathbf{P}} b$ as $N \to \infty$. Hence assumptions (4.1), also imply that, $1 < k < n$, $k \xrightarrow{\mathbf{P}} \infty$ and $k/n \xrightarrow{\mathbf{P}} 0$ as $N \to \infty$. Observe now that, in the case of complete data, we have $n \equiv N$ and $\mathbf{F}_n \equiv F_n$ with $\overline{F}_n(X_{n-k:n}) = k/n$ and thus the process defined in (4.2) becomes

$$D_n(x) := \sqrt{k}\left(\frac{n}{k}\overline{F}_n(xX_{n-k:n}) - x^{-1/\gamma_1}\right).$$

By jointly applying Theorems 2.4.8 and 5.1.4 (pages 52 and 161) in [de Haan and Ferreira(2006)] we have that, for $x_0 > 0$ and $0 < \epsilon < 1/2$,

$$\sup_{x \geq x_0} x^{(1/2-\epsilon)/\gamma_1}\left|D_n(x) - \Gamma(x;W) - x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1} - 1}{\tau_1\gamma_1}\sqrt{k}A_0(n/k)\right| \xrightarrow{\mathbf{P}} 0, \tag{4.3}$$

provided that F fulfills the second-order regular variation condition with auxiliary function A_0 tending to zero, not changing sign near infinity, having a regularly varying absolute value with index $\tau_1 < 0$ and satisfying $\sqrt{k}A_0(n/k) = O(1)$, where

$$\Gamma(x; W) := W\left(x^{-1/\gamma_1}\right) - x^{-1/\gamma_1}W(1),$$

with $\{W(s);\ s \geq 0\}$ being a standard Wiener process. Many authors used this approximation to establish the limit distributions of several statistics of heavy-tailed distributions, such as tail index estimators (see, e.g., [de Haan and Ferreira(2006)], p76) and goodness-of-fit statistics [Koning and Peng, 2008]. The main goal of this chapter is to provide an analogous result to (4.3) in the random truncation setting through the tail product-limit process (4.2), which, to the best of our knowledge, was not addressed yet in the extreme value theory literature.

Main results

We present our main result which consists in a Gaussian approximation to the tail product-limit process $\mathbf{D}_n(x)$.

Weak approximations of extreme value theory based statistics are achieved in the second-order framework, see, e.g., [de Haan and Stadtmüller(1996)]. Thus, it seems quite natural to suppose that df's $\mathbf{F}$ and $\mathbf{G}$ satisfy the well-known second-order condition of regular variation that we express in terms of the tail quantile functions. That is, we assume that for $x > 0$, we have

$$\lim_{t\to\infty} \frac{\mathbb{U}_{\mathbf{F}}(tx)/\mathbb{U}_{\mathbf{F}}(t) - x^{\gamma_1}}{\mathbf{A}_{\mathbf{F}}(t)} = x^{\gamma_1}\frac{x^{\tau_1} - 1}{\tau_1}, \tag{4.4}$$

and

$$\lim_{t\to\infty} \frac{\mathbb{U}_{\mathbf{G}}(tx)/\mathbb{U}_{\mathbf{G}}(t) - x^{\gamma_2}}{\mathbf{A}_{\mathbf{G}}(t)} = x^{\gamma_2}\frac{x^{\tau_2} - 1}{\tau_2}, \tag{4.5}$$

where $\tau_1, \tau_2 < 0$ 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 τ_1, τ_2 respectively. For any df K, the functions $K^{\leftarrow}(s) := \inf\{x : K(x) \geq s\}$, $0 < s < 1$, and $\mathbb{U}_K(t) := K^{\leftarrow}(1 - 1/t)$, $t > 1$, respectively stand for the quantile and tail quantile functions. For convenience, we set $\mathbf{A}^*_{\mathbf{F}}(t) := \mathbf{A}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}\left(\mathbb{U}_F(t)\right)\right), t > 1.$

Theoreme 4.1.1 *Assume that both second-order conditions* (4.4) *and* (4.5) *hold with* $\gamma_1 < \gamma_2$. *Let* k *be a sequence satisfying* (4.1), *then there exist a function* $\mathbf{A}_0(t) \sim \mathbf{A}^*_{\mathbf{F}}(t), t \to \infty$, *and a standard Wiener process* $\{\mathbf{W}(s);\ s \geq 0\}$, *defined on the probability space* $(\Omega, \mathcal{A}, \mathbf{P})$, *such that, for* $0 < \xi < 1/2 - \gamma/\gamma_2$ *and* $x_0 > 0$, *we have*

$$\sup_{x \geq x_0} x^{(1/2-\xi)/\gamma - 1/\gamma_2} \left| \mathbf{D}_n(x) - \mathbf{\Gamma}(x;\mathbf{W}) - x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \mathbf{A}_0(n/k) \right| \xrightarrow{\mathbf{P}} 0,$$

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*

$$\begin{aligned} \mathbf{\Gamma}(x;\mathbf{W}) \ :=& \ \frac{\gamma}{\gamma_1} x^{-1/\gamma_1} \left\{ x^{1/\gamma} \mathbf{W}\left(x^{-1/\gamma}\right) - \mathbf{W}(1) \right\} \\ &+ \frac{\gamma}{\gamma_1 + \gamma_2} x^{-1/\gamma_1} \int_0^1 s^{-\gamma/\gamma_2 - 1} \left\{ x^{1/\gamma} \mathbf{W}\left(x^{-1/\gamma} s\right) - \mathbf{W}(s) \right\} ds. \end{aligned} \tag{4.6}$$

Remark 4.1.1 *A very large value of* γ_2 *yields a* γ*-value that is very close to* γ_1, *meaning that the really observed sample is almost the whole dataset. In other words, the complete data case corresponds to the situation when* $1/\gamma_2 \equiv 0$, *in which case we have* $\gamma \equiv \gamma_1$. *It follows that in that case,* $\gamma(\gamma_1 + \gamma_2)^{-1} \int_0^1 s^{-\gamma/\gamma_2 - 1} \left\{ x^{1/\gamma} \mathbf{W}\left(x^{-1/\gamma} s\right) - \mathbf{W}(s) \right\} ds \equiv 0$ *and therefore* $\Gamma(x;\mathbf{W}) = \mathbf{W}\left(x^{-1/\gamma_1}\right) - x^{-1/\gamma_1}\mathbf{W}(1)$, *which agrees with the weak approximation* (4.3).

4.2 Tail index estimation

As an application, we introduce a new Hill-type estimator [Hill(1975)] for the tail index γ_1 and establish its consistency and asymptotic normality. The proposed estimator is compared with that of [Gardes and Stupfler(2015)] and its finite sample behavior is checked by simulation in next Section.

We start the construction of our estimator by noting that from Theorem 1.2.2 in [de Haan and Ferreira(20 the first-order condition (3.1) (for $\overline{\mathbf{F}}$) implies that

$$\lim_{t\to\infty}\frac{1}{\overline{\mathbf{F}}(t)}\int_t^\infty x^{-1}\overline{\mathbf{F}}(x)\,dx=\gamma_1,$$

which, by an integration by parts, becomes

$$\lim_{t\to\infty}\frac{1}{\overline{\mathbf{F}}(t)}\int_t^\infty \log\frac{x}{t}d\mathbf{F}(x)=\gamma_1. \tag{4.7}$$

Replacing $\mathbf{F}$ by $\mathbf{F}_n$ and letting $t=X_{n-k:n}$ yields

$$\widehat{\gamma}_1:=\frac{1}{\overline{\mathbf{F}}_n(X_{n-k:n})}\int_{X_{n-k:n}}^\infty \log\frac{x}{X_{n-k:n}}d\mathbf{F}_n(x),$$

as a new estimator to γ_1. By setting $\varphi_n^{(1)}(x):=\mathbf{1}\{x\geq X_{n-k:n}\}\log(x/X_{n-k:n})$ and $\varphi_n^{(2)}(x):=\mathbf{1}\{x\geq X_{n-k:n}\}$, this may be rewritten into $\widehat{\gamma}_1=\int_0^\infty\varphi_n^{(1)}(x)\,d\mathbf{F}_n(x)\,/\int_0^\infty\varphi_n^{(2)}(x)\,d\mathbf{F}_n(x)$. From the empirical counterpart of equation (1.4) we get

$$\int_0^\infty\varphi_n^{(1)}(x)\,d\mathbf{F}_n(x)=\frac{1}{n}\sum_{i=n-k}^{n}\frac{\mathbf{F}_n(X_{i:n})}{C_n(X_{i:n})}\log(X_{i:n}/X_{n-k:n}),$$

and

$$\int_0^\infty\varphi_n^{(2)}(x)\,d\mathbf{F}_n(x)=\frac{1}{n}\sum_{i=n-k}^{n}\frac{\mathbf{F}_n(X_{i:n})}{C_n(X_{i:n})}.$$

Finally, changing i to $n-i+1$ yields

$$\widehat{\gamma}_1=\left(\sum_{i=1}^{k}\frac{\mathbf{F}_n(X_{n-i+1:n})}{C_n(X_{n-i+1:n})}\right)^{-1}\sum_{i=1}^{k}\frac{\mathbf{F}_n(X_{n-i+1:n})}{C_n(X_{n-i+1:n})}\log\frac{X_{n-i+1:n}}{X_{n-k:n}}. \tag{4.8}$$

Note that a similar estimator (with deterministic threshold) is proposed in the independent parallel working paper [Worms and Worms(2016)]. Its asymptotic normality is established by means of the classical Lindeberg-Feller central limit theorem.

Remark 4.2.1 *For complete data, we have $n\equiv N$ and $\mathbf{F}_n\equiv F_n\equiv C_n$. Consequently $\widehat{\gamma}_1$ reduces to the classical Hill estimator [Hill(1975)].*

Main results

Theoreme 4.2.1 *Assume that* (3.1) *holds with* $\gamma_1 < \gamma_2$ *and let* k *be an integer sequence satisfying* (4.1). *Then* $\widehat{\gamma}_1 \to \gamma_1$ *in probability, as* $N \to \infty$. *Assume further that both second-order conditions* (4.4) *and* (4.5) *hold. Then*

$$\sqrt{k}\left(\widehat{\gamma}_1 - \gamma_1\right) = \int_1^\infty x^{-1}\mathbf{D}_n(x)\,dx = \frac{\sqrt{k}\mathbf{A}_0(n/k)}{1-\tau_1} - \gamma\mathbf{W}(1)$$
$$+ \frac{\gamma}{\gamma_1+\gamma_2}\int_0^1 \left(\gamma_2 - \gamma_1 - \gamma\log s\right) s^{-\gamma/\gamma_2 - 1}\mathbf{W}(s)\,ds + o_{\mathbf{P}}(1),$$

provided that, given $n = m$, $\sqrt{k_m}\mathbf{A}_0(m/k_m) = O(1)$, *as* $N \to \infty$.

Corollary 4.2.1 *If, in addition to the assumptions of Theorem 4.2.1, we suppose that, given* $n = m$, $\sqrt{k_m}\mathbf{A}^*_{\mathbf{F}}(m/k_m) \to \lambda$, *as* $N \to \infty$, *then*

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

where

$$\sigma^2 := \gamma^2\left(1 + \gamma_1/\gamma_2\right)\left(1 + \left(\gamma_1/\gamma_2\right)^2\right) / \left(1 - \gamma_1/\gamma_2\right)^3.$$

Remark 4.2.2 *In the complete data case, we have* $n \equiv N$ *and* $\sigma^2 \equiv \gamma_1^2$ *(from Remark* 4.1.1). *It follows that* $\sqrt{k}\left(\widehat{\gamma}_1 - \gamma_1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\lambda/\left(1-\tau_1\right), \gamma_1^2\right)$, *as* $N \to \infty$, *which meets the asymptotic normality of the classical Hill estimator [Hill(1975)], see for instance, Theorem 3.2.5 in [de Haan and Ferreira(2006)].*

4.3 Simulation study

This study is intended for illustrating the performance of our estimator and comparing it to that introduced by [Gardes and Stupfler(2015)] (respectively denoted by $\widehat{\gamma}_1$ and $\widehat{\gamma}_1^{GS}$ in the tables below), with respect to bias and root of the mean squared error (rmse). It is realized through two sets of truncated and truncation data, both drawn from Burr's model:

$$\overline{\mathbf{F}}(x) = \left(1 + x^{1/\delta}\right)^{-\delta/\gamma_1}, \ \overline{\mathbf{G}}(x) = \left(1 + x^{1/\delta}\right)^{-\delta/\gamma_2}, \ x \geq 0, \tag{4.9}$$

where $\delta, \gamma_1, \gamma_2 > 0$. The corresponding percentage of observed data equals $p = \gamma_2/(\gamma_1 + \gamma_2)$. We fix $\delta = 1/4$ and choose the values 0.6, 0.7, 0.8 and 0.9 for γ_1 and 70%, 80% 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, summarized in Tables 4.1, 4.2, 4.3 and 4.4 for the respective γ_1-values above, are taken as the empirical means of the results obtained through all repetitions. For the selection of the optimal numbers of upper order statistics used in the computation of estimators $\widehat{\gamma}_1$ and $\widehat{\gamma}_1^{GS}$, we apply the algorithm of [Reiss and Thomas(2007)], page 137. First, we note that, as expected, the estimation accuracy of both estimators decreases when the truncation percentage increases. Second, we see in all four tables that our estimator performs better as far as small samples are concerned (those of sizes less than 300 which indeed might be considered as small in the context of extreme values). This is especially advantageous in case studies where datasets are not so large as we will see in next section. It is also notable that, with regard to samples of average sizes (300 to 500 observations), both estimators roughly produce similar results, whereas for large data series (e.g. of size 1000) the estimator of [Gardes and Stupfler(2015)] is more suitable. Finally, we observe that, when the values of the tail index get larger, both estimation procedures are less precise.

4.4 High quantile estimation

We devote this Section to the estimation of high quantiles, under random right truncation, which we apply to a real dataset composed of lifetimes of car brake pads.

In many real-life fields, such as insurance, finance, hydrology and reliability, a typical requirement is to find values, large enough, so that the chances of exceeding them are very

N	n	$\widehat{\gamma}_1$	bias	rmse	$\widehat{\gamma}_1^{GS}$	bias	rmse
				$p = 0.7$			
100	69	0.403	−0.197	0.447	0.178	−0.422	7.310
150	105	0.446	−0.154	0.399	0.375	−0.225	1.892
200	139	0.452	−0.148	0.363	0.373	−0.227	0.993
300	209	0.510	−0.090	0.335	0.506	−0.094	1.003
500	350	0.537	−0.063	0.290	0.553	−0.047	0.453
1000	699	0.551	−0.049	0.205	0.576	−0.024	0.168
				$p = 0.8$			
100	79	0.462	−0.138	0.471	0.408	−0.192	2.546
150	120	0.501	−0.099	0.378	0.447	−0.153	0.907
200	159	0.517	−0.083	0.333	0.508	−0.092	0.424
300	240	0.528	−0.071	0.297	0.531	−0.069	0.367
500	400	0.547	−0.053	0.244	0.557	−0.043	0.194
1000	800	0.577	−0.023	0.169	0.579	−0.021	0.136
				$p = 0.9$			
100	90	0.550	−0.050	0.556	0.478	−0.122	4.751
150	134	0.539	−0.061	0.392	0.528	−0.072	0.537
200	180	0.532	−0.068	0.309	0.516	−0.084	0.651
300	269	0.554	−0.046	0.249	0.559	−0.041	0.245
500	449	0.557	−0.043	0.173	0.563	−0.037	0.167
1000	900	0.579	−0.021	0.126	0.582	−0.018	0.114

Table 4.1: Biases and RMSE's of the new estimator (left panel) and that of Gardes and Stupfler (right panel) of the tail index $\gamma_1 = 0.6$ based on 1000 samples of Burr models

N	n	$\widehat{\gamma}_1$	bias	rmse	$\widehat{\gamma}_1^{GS}$	bias	rmse
				$p = 0.7$			
100	69	0.468	−0.232	0.516	0.400	−0.300	5.129
150	104	0.545	−0.155	0.494	0.545	−0.155	2.562
200	140	0.545	−0.155	0.447	0.491	−0.209	1.285
300	210	0.584	−0.116	0.379	0.597	−0.103	0.744
500	350	0.637	−0.063	0.334	0.639	−0.061	0.566
1000	699	0.651	−0.049	0.255	0.668	−0.032	0.189
				$p = 0.8$			
100	79	0.565	−0.135	0.601	0.48	−0.212	5.068
150	119	0.604	−0.096	0.500	0.576	−0.124	1.039
200	159	0.609	−0.091	0.388	0.602	−0.098	0.620
300	240	0.626	−0.074	0.340	0.644	−0.056	0.477
500	399	0.638	−0.062	0.280	0.652	−0.048	0.238
1000	800	0.675	−0.025	0.201	0.680	−0.020	0.168
				$p = 0.9$			
100	90	0.629	−0.071	0.630	0.624	−0.076	2.856
150	135	0.601	−0.099	0.438	0.593	−0.107	0.832
200	180	0.635	−0.065	0.393	0.634	−0.066	0.436
300	269	0.630	−0.070	0.289	0.636	−0.064	0.264
500	450	0.659	−0.041	0.222	0.659	−0.041	0.203
1000	900	0.669	−0.031	0.149	0.672	−0.028	0.136

Table 4.2: Biases and RMSE's of the new estimator (left panel) and that of Gardes and Stupfler (right panel) of the tail index $\gamma_1 = 0.7$ based on 1000 samples of Burr models

N	n	$\widehat{\gamma}_1$	bias	rmse	$\widehat{\gamma}_1^{GS}$	bias	rmse
$p = 0.7$							
100	70	0.553	−0.247	0.617	0.485	−0.315	9.594
150	104	0.610	−0.190	0.515	0.492	−0.308	2.803
200	139	0.600	−0.200	0.513	0.544	−0.256	1.192
300	209	0.655	−0.145	0.417	0.662	−0.138	0.722
500	350	0.694	−0.106	0.350	0.719	−0.081	0.521
1000	700	0.753	−0.047	0.318	0.772	−0.028	0.212
$p = 0.8$							
100	80	0.652	−0.148	0.673	0.594	−0.206	5.210
150	120	0.660	−0.140	0.534	0.657	−0.143	0.883
200	159	0.688	−0.112	0.499	0.673	−0.127	0.737
300	240	0.686	−0.114	0.386	0.721	−0.079	0.534
500	400	0.728	−0.072	0.319	0.739	−0.061	0.286
1000	799	0.745	−0.055	0.226	0.756	−0.044	0.182
$p = 0.9$							
100	89	0.710	−0.090	0.713	0.707	−0.093	5.440
150	134	0.663	−0.137	0.467	0.662	−0.138	0.786
200	179	0.698	−0.102	0.407	0.690	−0.110	0.487
300	270	0.727	−0.073	0.327	0.730	−0.070	0.290
500	450	0.742	−0.058	0.240	0.745	−0.055	0.216
1000	899	0.759	−0.041	0.165	0.765	−0.035	0.156

Table 4.3: Biases and RMSE's of the new estimator (left panel) and that of Gardes and Stupfler (right panel) of the tail index $\gamma_1 = 0.8$ based on 1000 samples of Burr models

N	n	$\widehat{\gamma}_1$	bias	rmse	$\widehat{\gamma}_1^{GS}$	bias	rmse
$p = 0.7$							
100	70	0.613	−0.287	0.678	0.229	−0.671	7.099
150	104	0.666	−0.234	0.608	0.595	−0.305	1.823
200	140	0.657	−0.243	0.545	0.599	−0.301	1.430
300	209	0.755	−0.145	0.501	0.739	−0.161	1.413
500	349	0.814	−0.086	0.435	0.852	−0.048	0.673
1000	700	0.853	−0.047	0.425	0.869	−0.031	0.259
$p = 0.8$							
100	79	0.719	−0.181	0.747	0.534	−0.366	6.285
150	119	0.753	−0.147	0.585	0.691	−0.209	1.292
200	159	0.796	−0.104	0.511	0.769	−0.131	0.928
300	240	0.790	−0.110	0.460	0.787	−0.113	0.751
500	399	0.844	−0.056	0.347	0.863	−0.037	0.316
1000	799	0.860	−0.040	0.261	0.862	−0.038	0.244
$p = 0.9$							
100	90	0.786	−0.114	0.774	0.724	−0.176	8.239
150	134	0.776	−0.124	0.551	0.742	−0.158	0.793
200	179	0.788	−0.112	0.537	0.765	−0.135	0.499
300	269	0.838	−0.062	0.393	0.842	−0.058	0.365
500	450	0.837	−0.063	0.265	0.844	−0.056	0.252
1000	899	0.864	−0.036	0.197	0.866	−0.034	0.178

Table 4.4: Biases and RMSE's of the new estimator (left panel) and that of Gardes and Stupfler (right panel) of the tail index $\gamma_1 = 0.9$ based on 1000 samples of Burr models

small. That is, the interest is in estimating extreme quantiles of df $\mathbf{F}$, that we denote by $q_\nu := \mathbf{F}^{\leftarrow}(1-\nu) = \mathbb{U}_{\mathbf{F}}(1/\nu)$, when ν is close to 0. As we use asymptotic theory, ν must depend on the observed sample size n, i.e., $\nu = \nu_n$. The position of a high quantile with respect to the data depends on how small ν is. The most relevant case for purposes of real-world applications is when ν is much smaller than $1/n$, in which case q_ν is outside the available observations. Consequently, we are led to infer beyond the limits of the sample by extrapolating from intermediate quantiles. Obviously, this cannot be done without some kind of information on the distribution tail and so an appropriate modelling of the latter is needed. In other words, for a heavy-tailed distribution, an accurate estimation of the extreme value index is essential to the process of high quantile estimation. Needless to say that estimating such quantiles is a central issue in the context of risk management, where it is crucial to adequately evaluate the risk of a big loss that occurs very rarely. In the presence of complete data, where the most celebrated large quantile estimator is the due to [Weissman(1978)], this subject was extensively studied in the literature as was the case for the tail index estimation, see, for instance [de Haan and Ferreira(2006)]. In this section, we propose

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

as a Weissman-type estimator for q_ν under random right truncation. A similar estimator is proposed by [Gardes and Stupfler(2015)], but instead of $\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)$ they consider an arbitrary sequence of deterministic order asymptotically negligible with respect to ν. Before we state our result on the asymptotic normality of $\widehat{q}_\nu$, we set $d_n := \overline{\mathbf{F}}\left(\mathbb{U}_F\left(n/k\right)\right)/\nu_n$ with $\mathbb{U}_F$ regularly varying at infinity with index γ, and assume that, given $n = m$,

$$d_m \to 0 \text{ and } \sqrt{k_m}/\log d_m \to \infty \text{ as } N \to \infty. \tag{4.10}$$

4.4.1 Main results

Theoreme 4.4.1 *Assume that both second-order conditions* (4.4) *and* (4.5) *hold with* $\gamma_1 < \gamma_2$ *and let* k *be an integer sequence satisfying* (4.1) *and* (4.10). *Then*

$$\frac{\sqrt{k}}{\log d_n}\left(\frac{\widehat{q}_\nu}{q_\nu}-1\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(\frac{\lambda}{1-\tau_1},\sigma^2\right),\ \ \textit{as } N\to\infty,$$

provided that, given $n=m$,

$$\sqrt{k_m}\mathbf{A}_{\mathbf{F}}^{*}\left(m/k_m\right)\to\lambda \textit{ and } \sqrt{k_m}\mathbf{A}_{\mathbf{G}}^{*}\left(m/k_m\right)=O\left(1\right),\ \ \textit{as } N\to\infty, \tag{4.11}$$

where $\mathbf{A}_{\mathbf{G}}^{*}\left(t\right):=\mathbf{A}_{\mathbf{G}}\left(1/\overline{\mathbf{G}}\left(\mathbb{U}_F\left(t\right)\right)\right),\ t>1.$

Remark 4.4.1 *In the case of untruncated data, we have* $\mathbf{F}\equiv F$ *and* $d_N = k_N/(\nu_N N)$ *which coincides with* d_N *in Theorem 4.3.8 in [de Haan and Ferreira(2006)], page 138. Thereby, the conditions* (4.10), (4.11) *and the limiting distribution above agree with those in the same theorem.*

4.4.2 Real data example

Case study: automobile brake pad lifetime

As a real data example, we analyze the lifetimes of car brake pads already considered by [Lawless(2002)], page 69. We follow the same steps as those of [Gardes and Stupfler(2015)] who transformed this sample, which originally is left-truncated, into a right-truncated one, then checked the heavy-tail nature of its distribution. Since we are dealing with a dataset of small size $(n=98)$, then our estimator should be preferred to that of [Gardes and Stupfler(2015)]. After selecting the optimal number of top statistics, via the numerical procedure of [Reiss and Thomas(2007)], page 137, used in the previous section, we find 0.47 as an estimate value for the tail index γ_1. The estimations of the tail indices γ_2 and γ (using the same notation as before) are represented in Figure 4.2 and $\widehat{\gamma}_1$ are represented in Figure 4.1. The corresponding extreme quantiles that we obtain for three different levels are

quantile level	transformed data	original data
0.990	0.094	17604
0.995	0.130	14641
0.999	0.277	10559

Table 4.5: Extreme quantiles for car brake pad lifetimes

represented in Figure 4.3 and summarized in Table 4.5. For instance, we conclude that the brake pad lifetime is estimated to be less than 17, 600 km for 1% of the cars while only one out of a thousand brake pads lasts less than 10, 600 km.

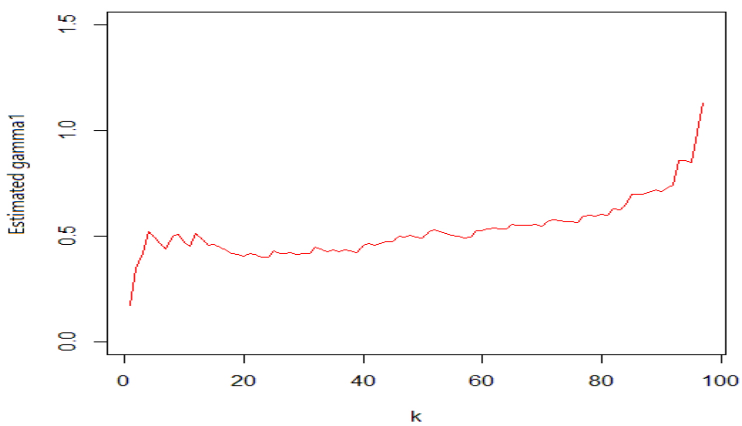


Figure 4.1: Plot of $\widehat{\gamma}_1$ as functions of k_n

4.5 Proofs

The proofs are postponed to this Section and some results that are instrumental to our needs are gathered in five lemmas in the Appendix.

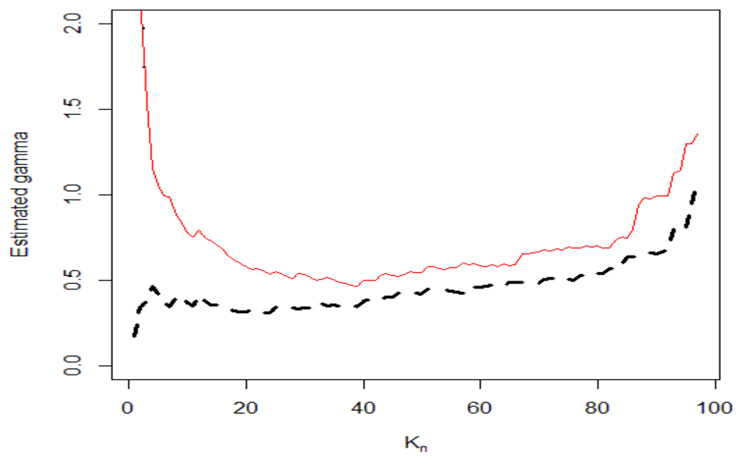


Figure 4.2: Hill estimators of γ_2 (full line) and γ (dashed line) as functions of k_n

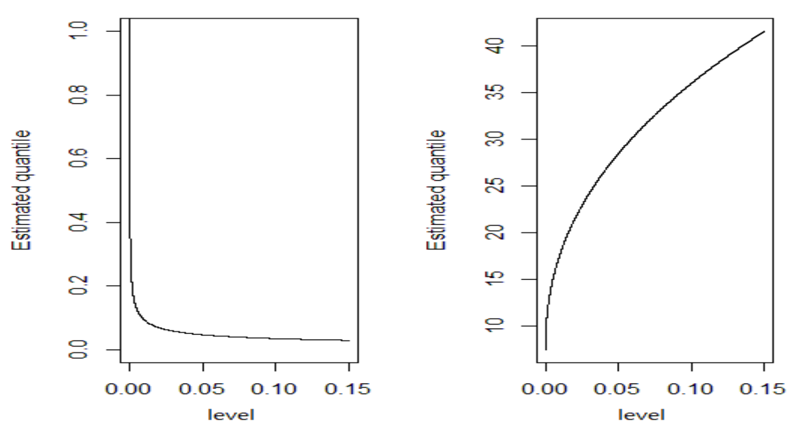


Figure 4.3: Estimated quantiles for the transformed data (left) and the original data (right)

Proof of Theorem 4.1.1

We fix $x_0 > 0$, then decompose $k^{-1/2}\mathbf{D}_n(x)$, for $x \geq x_0$, as the sum of the following four terms:

$$\mathbf{M}_{n1}(x) := x^{-1/\gamma_1} \frac{\overline{\mathbf{F}}_n(xX_{n-k:n}) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})},$$

$$\mathbf{M}_{n2}(x) := -\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} \frac{\overline{\mathbf{F}}_n(X_{n-k:n}) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})},$$

$$\mathbf{M}_{n3}(x) := \left(\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n})} - x^{-1/\gamma_1}\right) \frac{\overline{\mathbf{F}}_n(xX_{n-k:n}) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})},$$

and

$$\mathbf{M}_{n4}(x) := \frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})} - x^{-1/\gamma_1}.$$

Our goal is to provide an approximation to the tail product-limit process $\mathbf{D}_n(x)$ by means of a Wiener process. To this end, we introduce the sequence of iid rv's $U_i := \overline{F}(X_i)$, $i = 1, ..., n$. Since df's $\mathbf{F}$ and $\mathbf{G}$ are assumed to be continuous, then df F (of the $X_i's$) is continuous as well. On the other hand, we have, conditionally on $n = m$, for $0 \leq s \leq 1$, $\mathbf{P}\left(\overline{F}(X_i) \leq s\right) = s$, $i = 1, ..., m$, which means that $\{U_i\}_{i=1,m}$ are uniformly distributed on $(0, 1)$. Let us now define the corresponding uniform tail empirical process by $\alpha_n(s) := \sqrt{k}(\mathbf{U}_n(s) - s)$, for $0 \leq s \leq 1$, where $\mathbf{U}_n(s) := k^{-1}\sum_{i=1}^n \mathbf{1}(U_i < ks/n)$. In view of Proposition 3.1 of [Einmahl et al.(2006)], we show in Lemma 4.6.3, that there exists a Wiener process $\mathbf{W}$, such that for every $0 \leq \eta < 1/2$,

$$\sup_{0<s\leq 1} s^{-\eta} |\alpha_n(s) - \mathbf{W}(s)| \xrightarrow{\mathbf{P}} 0, \text{ as } N \to \infty. \tag{4.12}$$

In order to establish the result of the theorem, we will successively show that, under the first-order of regular variation conditions, we have uniformly on $x \geq x_0$, for $\gamma/\gamma_2 < \eta < 1/2$

and $\epsilon > 0$ sufficiently small

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n1}(x)$$
$$= x^{1/\gamma}\left\{\frac{\gamma}{\gamma_1}\mathbf{W}\left(x^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1+\gamma_2}\int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}\left(x^{-1/\gamma}t\right)dt\right\} + o_{\mathbf{P}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right),$$

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n2}(x) = -\left\{\frac{\gamma}{\gamma_1}\mathbf{W}(1) + \frac{\gamma}{\gamma_1+\gamma_2}\int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}(t)\,dt\right\} + o_{\mathbf{P}}\left(x^{\pm\epsilon}\right),$$

and $x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n3}(x) = o_{\mathbf{P}}\left(x^{-1/\gamma_1+(1-\eta)/\gamma\pm\epsilon}\right)$, where $x^{\pm\epsilon} := \max\left(x^{-\epsilon}, x^{\epsilon}\right)$. Moreover, if we assume the second-order condition we will show that

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n4}(x) = (1+o_{\mathbf{P}}(1))\frac{x^{\tau_1/\gamma_1}-1}{\gamma_1\tau_1}\sqrt{k}\mathbf{A}_0(n/k).$$

Let $a_k := \mathbb{U}_F(n/k)$ be a random sequence tending to ∞ in probability, as $N\to\infty$. From Lemma 4.6.4 we have $X_{n-k:n}/a_k \xrightarrow{\mathbf{P}} 1$ as $N\to\infty$, which, in virtue of the regular variation of $\overline{\mathbf{F}}$, implies that $\overline{\mathbf{F}}(xa_k)/\overline{\mathbf{F}}(xX_{n-k:n}) = 1+o_{\mathbf{P}}(1)$. Therefore

$$\mathbf{M}_{n1}(x) = (1+o_{\mathbf{P}}(1))\,\mathbf{M}_{n1}^{*}(x), \tag{4.13}$$

where

$$\mathbf{M}_{n1}^{*}(x) := x^{-1/\gamma_1}\frac{\overline{\mathbf{F}}_n(xX_{n-k:n}) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xa_k)}.$$

Now, observe that, in view of equation (1.4), we may write

$$\mathbf{F}(x) = \exp\{-\Lambda(x)\} \text{ and } \mathbf{F}_n(x) = \exp\{-\Lambda_n(x)\},$$

where $\Lambda(x)$ and its empirical counterpart $\Lambda_n(x)$ are defined, respectively, by

$$\int_x^\infty dF(z)/C(z) \text{ and } \int_x^\infty dF_n(z)/C_n(z).$$

Note that $\overline{\mathbf{F}}_n(xX_{n-k:n})$, $\overline{\mathbf{F}}(xX_{n-k:n})$ and $\overline{\mathbf{F}}(xa_k)$ tend to zero in probability, uniformly on $x \geq x_0$, it follows that $\Lambda_n(xX_{n-k:n})$, $\Lambda(xX_{n-k:n})$ and $\Lambda(xa_k)$ go to zero in probability as well. Using the approximation $1-\exp(-t) \sim t$, as $t \downarrow 0$, we may write

$$x^{1/\gamma_1}\mathbf{M}_{n1}^*(x) = (1+o_{\mathbf{p}}(1))\frac{\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})}{\Lambda(xa_k)}.$$

Next, we provide a Gaussian approximation to the expression

$$\sqrt{k}\frac{\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})}{\Lambda(xa_k)},$$

then we deduce one to $\sqrt{k}x^{1/\gamma_1}\mathbf{M}_{n1}^*(x)$. For this, we decompose $\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})$ into the sum of

$$S_{n1}(x) := -\int_{xa_k}^{\infty}\frac{d\left(\overline{F}_n(z) - \overline{F}(z)\right)}{C(z)},$$

$$S_{n2}(x) := -\int_{xX_{n-k:n}}^{\infty}\left\{\frac{1}{C_n(z)} - \frac{1}{C(z)}\right\}d\overline{F}_n(z),$$

and

$$S_{n3}(x) := -\int_{xX_{n-k:n}}^{xa_k}\frac{d\left(\overline{F}_n(z) - \overline{F}(z)\right)}{C(z)}.$$

For the first term, we use the fact that $\overline{F}_n(z) = 0$ for $z \geq X_{n:n}$, to write, after an integration by parts and a change of variables, $S_{n1}(x) = S_{n1}^{(1)}(x) - S_{n1}^{(2)}(x)$, with

$$S_{n1}^{(1)}(x) := \frac{\overline{F}_n(a_kx) - \overline{F}(a_kx)}{C(a_kx)} \text{ and } S_{n1}^{(2)}(x) := \int_x^{\infty}\frac{\overline{F}_n(a_kz) - \overline{F}(a_kz)}{C^2(a_kz)}dC(a_kz).$$

It is easy to verify that $\overline{F}_n(xa_k) - \overline{F}(xa_k) = \dfrac{\sqrt{k}}{n}\alpha_n\left(\dfrac{n}{k}\overline{F}(xa_k)\right)$, it follows that

$$\frac{\sqrt{k}S_{n1}^{(1)}(x)}{\Lambda(a_kx)} = h_n(x)\,\alpha_n\left(\frac{n}{k}\overline{F}(a_kx)\right),$$

where $h_n(x) := \dfrac{k/n}{\Lambda(xa_k)\, C(a_k x)}$. From Lemma 4.6.2 (iii), we have

$$h_n(x) = (\gamma/\gamma_1)\, x^{1/\gamma} + o_{\mathbf{p}}\left(x^{1/\gamma \pm \epsilon}\right), \text{ as } N \to \infty, \tag{4.14}$$

uniformly on $x \geq x_0$, it follows that

$$\frac{\sqrt{k} S_{n1}^{(1)}(x)}{\Lambda(a_k x)} = \left\{(\gamma/\gamma_1)\, x^{1/\gamma} + o_{\mathbf{p}}\left(x^{1/\gamma \pm \epsilon}\right)\right\} \alpha_n\left(\frac{n}{k}\overline{F}(a_k x)\right).$$

On the other hand, for $0 < \eta < 1/2$, the sequence of rv's $\sup_{0<s\leq 1} |\alpha_n(s)| / s^{\eta}$ is stochastically bounded. This comes from the inequality

$$\sup_{0<s\leq 1} s^{-\eta} |\alpha_n(s)| \leq \sup_{0<s\leq 1} s^{-\eta} |\alpha_n(s) - \mathbf{W}(s)| + \sup_{0<s\leq 1} s^{-\eta} |\mathbf{W}(s)|,$$

with approximation (4.12) and the fact $\sup_{0<s\leq 1} s^{-\eta} |\mathbf{W}(s)| = O_{\mathbf{p}}(1)$, see, e.g., Lemma 3.2 in [Einmahl et al.(2006)]. Now, for $\epsilon > 0$ be sufficiently small, we write, by applying Potter's inequalities to $\overline{F}$, see, e.g., Proposition B.1.9, assertion 5 in [de Haan and Ferreira(2006)] together with (1.1), $\dfrac{n}{k}\overline{F}(a_k x) \leq (1+\epsilon)\, x^{-1/\gamma \pm \epsilon}$, it follows that

$$\alpha_n\left(\frac{n}{k}\overline{F}(a_k x)\right) = O_{\mathbf{p}}\left(x^{-\eta/\gamma \pm \eta\epsilon}\right),$$

uniformly on $x \geq x_0$. For notational simplicity and without loss of generality, we attribute ϵ to any constant times ϵ and $v^{\pm\epsilon}$ to any linear combinations of $v^{\pm c_1 \epsilon}$ and $v^{\pm c_2 \epsilon}$, for every $c_1, c_2 > 0$. Therefore

$$\frac{\sqrt{k} S_{n1}^{(1)}(x)}{\Lambda(a_k x)} = \frac{\gamma}{\gamma_1} x^{1/\gamma} \alpha_n\left(\frac{n}{k}\overline{F}(a_k x)\right) + o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma \pm \epsilon}\right).$$

For $S_{n1}^{(2)}(x)$, let us write

$$\frac{\sqrt{k} S_{n1}^{(2)}(x)}{\Lambda(a_k x)} = h_n(x) \frac{C(a_k x)}{C(a_k)} \int_x^{\infty} \frac{C^2(a_k)}{C^2(a_k z)} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{C(a_k z)}{C(a_k)}.$$

From Lemma 4.6.2 (i), the function C is regularly varying at infinity with index $(-1/\gamma_2)$, as $\overline{G}$ is, this implies that $C(xa_k)/C(a_k) = x^{-1/\gamma_2} + o_{\mathbf{p}}\left(x^{-1/\gamma_2 \pm \epsilon}\right)$. Then by using (4.14), we get

$$h_n(x)\frac{C(a_k x)}{C(a_k)} = (\gamma/\gamma_1)\, x^{1/\gamma_1} + o_{\mathbf{p}}\left(x^{1/\gamma_1 \pm \epsilon}\right), \text{ as } N \to \infty. \tag{4.15}$$

For convenience, we set $\sqrt{k} S_{n1}^{(2)}(x)/\Lambda(a_k x) = (1 + o_{\mathbf{p}}(x^{\pm\epsilon}))\, \mathcal{T}_n(x)$, where

$$\mathcal{T}_n(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty \frac{C^2(a_k)}{C^2(a_k z)} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{C(a_k z)}{C(a_k)},$$

which we decompose in the sum of

$$I_n(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty \frac{C^2(a_k)}{C^2(a_k z)} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{F(a_k z)}{C(a_k)},$$

$$J_n(x) := -\frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty \left\{\frac{C^2(a_k)}{C^2(a_k z)} - z^{2/\gamma_2}\right\} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{G(a_k z)}{C(a_k)},$$

and

$$K_n(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty z^{2/\gamma_2} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{\overline{G}(a_k z)}{C(a_k)}.$$

Recall that $a_k \xrightarrow{\mathbf{P}} \infty$, which implies that $C(a_k)/\overline{G}(a_k) \xrightarrow{\mathbf{P}} 1$ and $\overline{F}(a_k)/\overline{G}(a_k) \xrightarrow{\mathbf{P}} 0$ as $N \to \infty$ and note that all the inequalities below, corresponding to C and $\overline{F}$, occur with probabilities close to 1 as $N \to \infty$. By using, once again, Potter's inequalities to C, (regularly varying at infinity with index $-1/\gamma_2$), we write, for $z \geq x$

$$(1-\epsilon)\, z^{-1/\gamma_2} \min\left(z^\epsilon, z^{-\epsilon}\right) \leq \frac{C(a_k z)}{C(a_k)} \leq (1+\epsilon)\, z^{-1/\gamma_2} \max\left(z^\epsilon, z^{-\epsilon}\right). \tag{4.16}$$

It is clear this implies that $C^2(a_k)/C^2(a_k z) \leq (1-\epsilon)^{-2} z^{2/\gamma_2 \pm 2\epsilon}$. On the other hand, we have $\sup_{0<s\leq 1} |\alpha_n(s)|/s^\eta = O_{\mathbf{p}}(1)$ and $\frac{n}{k}\overline{F}(a_k z) \leq (1+\epsilon)\, z^{-1/\gamma \pm \epsilon}$, then

$$I_n(x) = o_{\mathbf{p}}(1)\, x^{1/\gamma_1} \int_x^\infty z^{2/\gamma_2 \mp 2\epsilon} \left(z^{-1/\gamma \pm \epsilon}\right)^\eta d\frac{\overline{F}(a_k z)}{\overline{F}(a_k)}.$$

Integrating by parts, we readily get $I_n(x) = o_{\mathbf{p}}\left(x^{1/\gamma_2-\eta/\gamma\pm\epsilon}\right) = o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right)$. Let us now consider $J_n(x)$. From Proposition B.1.10 in [de Haan and Ferreira(2006)], we have $\left|C(a_k z)/C(a_k) - z^{-1/\gamma_2}\right| \leq \epsilon z^{-1/\gamma_2\pm\epsilon}$. Applying the mean value theorem, then combining this inequality with (4.16), yield

$$\left|\frac{C^2(a_k)}{C^2(a_k z)} - z^{2/\gamma_2}\right| \leq \epsilon \frac{2\left(z^{\pm\epsilon}+1\right)}{(1-\epsilon)^3} z^{2/\gamma_2\pm\epsilon}.$$

Similar arguments as the above lead to $J_n(x) = o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right)$. Now, we focus on $K_n(x)$. Since $C(a_k)/\overline{G}(a_k) \xrightarrow{\mathbf{p}} 1$, then

$$K_n(x) = (1+o_{\mathbf{p}}(1)) \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_x^\infty z^{2/\gamma_2} \alpha_n\left(\frac{n}{k}\overline{F}(a_k z)\right) d\frac{\overline{G}(a_k z)}{\overline{G}(a_k)}.$$

By using the change of variables $z = G^{\leftarrow}\left(1 - s\overline{G}(a_k)\right)/a_k$ we get

$$K_n(x) = -(1+o_{\mathbf{p}}(1)) \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{\frac{\overline{G}(a_k x)}{\overline{G}(a_k)}} \left(\frac{G^{\leftarrow}\left(1-s\overline{G}(a_k)\right)}{a_k}\right)^{2/\gamma_2} \alpha_n(\ell_n(s))\, ds,$$

where $\ell_n(s) := \frac{n}{k}\overline{F}\left(G^{\leftarrow}\left(1-s\overline{G}(a_k)\right)\right)$. It is easy to check that

$$K_n(x) = -(1+o_{\mathbf{p}}(1)) \sum_{i=1}^3 K_{ni}(x),$$

where

$$K_{n1}(x) := \frac{\gamma}{\gamma_1} x^{\frac{1}{\gamma_1}} \int_0^{\frac{\overline{G}(a_k x)}{\overline{G}(a_k)}} \left\{\left(\frac{G^{\leftarrow}\left(1-s\overline{G}(a_k)\right)}{a_k}\right)^{2/\gamma_2} - s^{-2}\right\} \alpha_n(\ell_n(s))\, ds,$$

$$K_{n2}(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_{x^{-1/\gamma_2}}^{\frac{\overline{G}(a_k x)}{\overline{G}(a_k)}} s^{-2} \alpha_n(\ell_n(s))\, ds,$$

and

$$K_{n3}(x) := \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{x^{-1/\gamma_2}} s^{-2} \alpha_n(\ell_n(s))\, ds.$$

By similar arguments based on stochastic boundedness of $\sup_{0<s\leq 1}|\alpha_n(s)|/s^\eta$ and the aforementioned Proposition B.1.10 in [de Haan and Ferreira(2006)] applied to the regularly varying functions $\overline{G}$ and $G^{\leftarrow}(1-\cdot)$, we show that $K_{ni}(x)=o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\mp\epsilon}\right)$, $i=1,2$ and $K_{n3}(x) = O_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\mp\epsilon}\right)$, therefore we omit the details. Up to this stage, we have shown that $\mathcal{T}_n(x)=O_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\mp\epsilon}\right)$. It follows that $\sqrt{k}S_{n1}^{(2)}(x)/\Lambda(a_kx)=\mathcal{T}_n(x)+o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\mp\epsilon}\right)$, which, after gathering the components of $\mathcal{T}_n(x)$, is equal to

$$\frac{\gamma}{\gamma_1}x^{1/\gamma_1}\int_0^{x^{-1/\gamma_2}}s^{-2}\alpha_n(\ell_n(s))\,ds+o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\mp\epsilon}\right).$$

Therefore

$$\begin{aligned}&\frac{\sqrt{k}S_{n1}(x)}{\Lambda(a_kx)}\\&=\frac{\gamma}{\gamma_1}x^{1/\gamma}\alpha_n\left(\frac{n}{k}\overline{F}(a_kx)\right)+\frac{\gamma}{\gamma_1}x^{1/\gamma_1}\int_0^{x^{-1/\gamma_2}}s^{-2}\alpha_n(\ell_n(s))\,ds+o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\mp\epsilon}\right).\end{aligned}$$

Recall that $\gamma_1<\gamma_2$ and $\gamma/\gamma_2=\gamma_1/(\gamma_1+\gamma_2)$, then we may choose the constant η in such a way that $\gamma/\gamma_2<\eta<1/2$. Making use of weak approximation (4.12), we obtain

$$\begin{aligned}&\frac{\sqrt{k}S_{n1}(x)}{\Lambda(a_kx)}\\&=\frac{\gamma}{\gamma_1}x^{1/\gamma}\mathbf{W}\left(\frac{n}{k}\overline{F}(a_kx)\right)+\frac{\gamma}{\gamma_1}x^{1/\gamma_1}\int_0^{x^{-1/\gamma_2}}s^{-2}\mathbf{W}(\ell_n(s))\,ds+o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\mp\epsilon}\right).\end{aligned}$$

Note that $k/n=\overline{F}\left(G^{\leftarrow}\left(1-\overline{G}(a_k)\right)\right)$, hence

$$\ell_n(s)=\frac{\overline{F}\left(G^{\leftarrow}\left(1-s\overline{G}(a_k)\right)\right)}{\overline{F}\left(G^{\leftarrow}\left(1-\overline{G}(a_k)\right)\right)}.$$

Since $s\to\overline{F}\circ G^{\leftarrow}(1-s)$ is regularly varying at infinity with index γ_2/γ, then, from Proposition B.1.10 in [de Haan and Ferreira(2006)], we have

$$\omega_n(s):=\left|\ell_n(s)-s^{\gamma_2/\gamma}\right|\leq\epsilon s^{\gamma_2/\gamma\pm\epsilon},\tag{4.17}$$

with high probability for large N. Recall that $x_0 > 0$ is fixed, then

$$\sup_{x \geq x_0} \sup_{0 < s \leq x^{-1/\gamma_2}} \omega_n(s) \xrightarrow{\mathbf{P}} 0, \text{ as } N \to \infty.$$

On the other hand, by using Levy's modulus of continuity of the Wiener process, see, e.g., Theorem 1.1.1 in [Csörgő and Révész, 1981], we have with probability one

$$\left| \mathbf{W}(\ell_n(s)) - \mathbf{W}\left(s^{\gamma_2/\gamma}\right) \right| \leq 2\sqrt{\omega_n(s) \log(1/\omega_n(s))},$$

uniformly on $s \geq x^{-1/\gamma_2}$. By using the fact that, $\log u < \epsilon u^{-\epsilon}$ as $u \downarrow 0$, together with inequality (4.17), we get $\left| \mathbf{W}(\ell_n(s)) - \mathbf{W}\left(s^{\gamma_2/\gamma}\right) \right| \leq 2\epsilon s^{(\gamma_2/\gamma)(1-\epsilon)/2}$, almost surely. Following our convention, we may write that $(\gamma_2/\gamma \pm \epsilon)(1 - \epsilon/2) \equiv \gamma_2/\gamma \pm \epsilon$. Since $\gamma_1 < \gamma_2$ then $\gamma_2/(2\gamma) > 1$ and after elementary calculation, we show that uniformly on $x \geq x_0$

$$\frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{x^{-1/\gamma_2}} s^{-2} \mathbf{W}(\ell_n(s))\, ds = \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{x^{-1/\gamma_2}} s^{-2} \mathbf{W}\left(s^{\gamma_2/\gamma}\right) ds + o_{\mathbf{p}}\left(x^{1/(2\gamma) \pm \epsilon}\right).$$

By similar arguments, we get

$$\frac{\gamma}{\gamma_1} x^{1/\gamma} \mathbf{W}\left(\frac{n}{k} \overline{F}(a_k x)\right) = \frac{\gamma}{\gamma_1} x^{1/\gamma} \mathbf{W}\left(x^{-1/\gamma}\right) + o_{\mathbf{p}}\left(x^{1/(2\gamma) \pm \epsilon}\right).$$

It is obvious that $o_{\mathbf{p}}\left(x^{1/(2\gamma) \pm \epsilon}\right) + o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma \pm \epsilon}\right) = o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma \pm \epsilon}\right)$, it follows that

$$\frac{\sqrt{k} S_{n1}(x)}{\Lambda(a_k x)} = \frac{\gamma}{\gamma_1} x^{1/\gamma} \mathbf{W}\left(x^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1} x^{1/\gamma_1} \int_0^{x^{-1/\gamma_2}} s^{-2} \mathbf{W}\left(s^{\gamma_2/\gamma}\right) ds + o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma \pm \epsilon}\right).$$

After a change of variables, this may be rewritten into

$$\begin{aligned} &\frac{\sqrt{k} S_{n1}(x)}{\Lambda(a_k x)} \\ &= \frac{\gamma}{\gamma_1} x^{1/\gamma} \mathbf{W}\left(x^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1 + \gamma_2} x^{1/\gamma} \int_0^1 t^{-\gamma/\gamma_2 - 1} \mathbf{W}\left(x^{-1/\gamma} t\right) dt + o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma \pm \epsilon}\right). \end{aligned} \tag{4.18}$$

Now, we consider the second term $S_{n2}(x)$. We have $\overline{F}_n(z)=0$, for $z\geq X_{n:n}$, thus

$$S_{n2}(x)=\int_{xX_{n-k:n}}^{X_{n:n}}\frac{C_n(z)-C(z)}{C_n(z)\,C(z)}d\overline{F}_n(z).$$

Therefore $|S_{n2}(x)|\leq\theta_n\int_{xX_{n-k:n}}^{\infty}\dfrac{|C_n(z)-C(z)|}{C^2(z)}dF_n(z)$, where $\theta_n:=\sup_{X_{1:n}\leq z\leq X_{n:n}}\left\{\dfrac{C(z)}{C_n(z)}\right\}$. Note, that from, see, e.g., [Stute and Wang(2008)] , given $n=m$, the sequence θ_m is stochastically bounded, then by using the total probability formula, we show easily that θ_n is also stochastically bounded. By recalling that we have $C=\overline{G}-\overline{F}$ and $C_n=\overline{G}_n-\overline{F}_n$, with G_n denoting the empirical df of G, we write $|S_{n2}(x)|\leq\theta_n(T_{n1}(x)+T_{n2}(x))$, where

$$T_{n1}(x):=\int_{xX_{n-k:n}}^{\infty}\frac{\left|\overline{F}_n(z)-\overline{F}(z)\right|}{C^2(z)}dF_n(z)$$

and

$$T_{n2}(x):=\int_{xX_{n-k:n}}^{\infty}\frac{\left|\overline{G}_n(z)-\overline{G}(z)\right|}{C^2(z)}dF_n(z).$$

It is easy to verify that, by a change of variables, we have

$$\begin{aligned}\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_kx)}&=h_n(x)\frac{k/n}{C(a_k)}\frac{C(a_kx)}{C(a_k)}\\&\quad\times\frac{C^2(a_k)}{C^2(xX_{n-k:n})}\int_1^{\infty}\frac{\left|\alpha_n\left(\frac{n}{k}\overline{F}(xX_{n-k:n}z)\right)\right|}{C^2(xX_{n-k:n}z)/C^2(xX_{n-k:n})}d\frac{F_n(xX_{n-k:n}z)}{\overline{F}(a_k)}.\end{aligned}$$

Recall that, uniformly on $x\geq x_0$, we have $C(a_k)/C(xX_{n-k:n})=O_{\mathbf{p}}\left(x^{1/\gamma_2\pm\epsilon}\right)$. Moreover, we use (4.16) and (4.15) to write

$$\begin{aligned}\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_kx)}&=O_{\mathbf{p}}\left(\frac{k/n}{C(a_k)}\right)x^{1/\gamma\pm\epsilon}\\&\quad\times\int_1^{\infty}z^{2/\gamma_2}\left|\alpha_n\left(\frac{n}{k}\overline{F}(xX_{n-k:n}z)\right)\right|d\frac{F_n(xX_{n-k:n}z)}{\overline{F}(a_k)}.\end{aligned}$$

On the other hand, by using the stochastic boundedness of $\sup_{0<s\leq 1}|\alpha_n(s)|/s^{\eta}$ we get

$$\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_k x)} = \frac{k/n}{C(a_k)} O_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right) \int_1^{\infty} z^{2/\gamma_2-\eta/\gamma\pm\epsilon} d\frac{\overline{F}_n(xX_{n-k:n}z)}{\overline{F}(a_k)},$$

where the integral may be split as follows

$$\int_1^{\infty} z^{2/\gamma_2-\eta/\gamma\pm\epsilon} d\frac{\overline{F}_n(xX_{n-k:n}z)}{\overline{F}(a_k)} = P_n(x) + Q_n(x),$$

where

$$P_n(x) := \int_1^{\infty} z^{2/\gamma_2-\eta/\gamma\pm\epsilon} d\left\{\frac{\overline{F}_n(xX_{n-k:n}z) - \overline{F}(xX_{n-k:n}z)}{\overline{F}(a_k)}\right\},$$

and

$$Q_n(x) := \int_1^{\infty} z^{2/\gamma_2-\eta/\gamma\pm\epsilon} d\frac{\overline{F}(xX_{n-k:n}z)}{\overline{F}(a_k)}.$$

It is clear that

$$P_n(x) = k^{-1/2} \int_1^{\infty} z^{2/\gamma_2-\eta/\gamma\pm\epsilon} d\alpha_n\left(\frac{n}{k}\overline{F}(xX_{n-k:n}z)\right).$$

By similar arguments as those used above, we show that, uniformly on $x \geq x_0$, we have $P_n(x) = o_{\mathbf{p}}\left(x^{-\eta/\gamma\pm\epsilon}\right)$ and $Q_n(x) = O_{\mathbf{p}}\left(x^{-1/\gamma\pm\epsilon}\right)$, therefore

$$\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_k x)} = \frac{k/n}{C(a_k)} O_{\mathbf{p}}\left(x^{-\eta/\gamma\pm\epsilon}\right).$$

Next, let $V_i := \overline{G}(Y_i)$, $i = 1, ..., n$, and define the corresponding tail empirical process $\beta_n(s) := \sqrt{k}(\mathbf{V}_n(s) - s)$, for $0 \leq s \leq 1$, where $\mathbf{V}_n(s) := k^{-1}\sum_{i=1}^{n} \mathbf{1}(V_i < ks/n)$. Like for $\alpha_n(s)$, we also have $\sup_{0<s\leq 1}|\beta_n(s)|/s^{\eta} = O_{\mathbf{p}}(1)$, therefore by similar arguments as those used for $T_{n1}(x)$, with the facts that $\overline{G}(t) \sim C(t)$ as $t \to \infty$ and $\gamma_2 > \gamma$, we show that

$$\frac{\sqrt{k}T_{n2}(x)}{\Lambda(a_k x)} = \frac{k/n}{C^{1-\eta}(a_k)} O_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right).$$

From Lemma 4.6.2 (ii), we have that both $\frac{n}{k}C(a_k)$ and $\frac{n}{k}C^{1-\eta}(a_k)$ tend to infinity in probability, it follows that

$$\frac{\sqrt{k}T_{n1}(x)}{\Lambda(a_k x)}=o_{\mathbf{p}}\left(x^{-\eta/\gamma\pm\epsilon}\right) \text{ and } \frac{\sqrt{k}T_{n2}(x)}{\Lambda(a_k x)}=o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right).$$

Since $o_{\mathbf{p}}\left(x^{-\eta/\gamma\pm\epsilon}\right)+o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right)=o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right)$, then

$$\frac{\sqrt{k}S_{n2}(x)}{\Lambda(a_k x)}=o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right). \tag{4.19}$$

Let us now focus on the third term S_{n3}, which, by integration by parts, equals the sum of

$$S_{n3}^{(1)}(x):=-\int_{xX_{n-k:n}}^{xa_k}\frac{\overline{F}_n(z)-\overline{F}(z)}{C^2(z)}dC(z),$$

and

$$S_{n3}^{(2)}(x)=-\frac{\overline{F}_n(a_k x)-\overline{F}(a_k x)}{C(a_k x)}+\frac{\overline{F}_n(xX_{n-k:n})-\overline{F}(xX_{n-k:n})}{C(xX_{n-k:n})}.$$

By using the change of variables $z=txa_k$ we get

$$\frac{\sqrt{k}S_{n3}^{(1)}(x)}{\Lambda(a_k x)}=-h_n(x)\int_{X_{n-k:n}/a_k}^{1}\frac{\alpha_n\left(\frac{n}{k}\overline{F}(a_k xz)\right)}{\left(C(a_k xz)/C(a_k x)\right)^2}d\frac{C(a_k xz)}{C(a_k x)},$$

and

$$\frac{\sqrt{k}S_{n3}^{(2)}(x)}{\Lambda(a_k x)}=-h_n(x)\left\{\alpha_n\left(\frac{n}{k}\overline{F}(a_k x)\right)-\frac{C(a_k x)}{C(xX_{n-k:n})}\alpha_n\left(\frac{n}{k}\overline{F}(xX_{n-k:n})\right)\right\}.$$

Routine manipulations, including Proposition B.1.10 in [de Haan and Ferreira(2006)] and the stochastic boundedness of $\sup_{0<s\leq 1}|\alpha_n(s)|/s^{\eta}$, yield

$$\frac{\sqrt{k}S_{n3}^{(1)}(x)}{\Lambda(a_k x)}=o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right) \text{ and } \frac{\sqrt{k}S_{n3}^{(2)}(x)}{\Lambda(a_k x)}=o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right).$$

It follows that

$$\sqrt{k}S_{n3}(x)/\Lambda(a_k x) = o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right). \tag{4.20}$$

By gathering results (4.18), (4.19) and (4.20), we obtain

$$\begin{aligned}&\sqrt{k}\frac{\Lambda_n(xX_{n-k:n}) - \Lambda(xX_{n-k:n})}{\Lambda(a_k x)} \qquad (4.21)\\ &= \frac{\gamma}{\gamma_1}x^{1/\gamma}\mathbf{W}\left(x^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1+\gamma_2}x^{1/\gamma}\int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}\left(x^{-1/\gamma}t\right)dt + o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right),\end{aligned}$$

which yields that

$$\begin{aligned}&x^{1/\gamma_1}\sqrt{k}\mathbf{M}^*_{n1}(x)\\ &= x^{1/\gamma}\left\{\frac{\gamma}{\gamma_1}\mathbf{W}\left(x^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1+\gamma_2}\int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}\left(x^{-1/\gamma}t\right)dt\right\} + o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right).\end{aligned}$$

We show that the expectation of the absolute value of the first term in the right-hand side of the previous equation equals $O_{\mathbf{p}}\left(x^{1/(2\gamma)}\right)$. Since we already have $1/(2\gamma) < (1-\eta)/\gamma$, then $x^{1/\gamma_1}\sqrt{k}\mathbf{M}^*_{n1}(x) = O_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right)$, which leads to

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n1}(x) = x^{1/\gamma_1}\sqrt{k}\mathbf{M}^*_{n1}(x) + o_{\mathbf{p}}\left(x^{(1-\eta)/\gamma\pm\epsilon}\right).$$

Recall that $\epsilon > 0$ is chosen sufficiently small, then for any $0 < \eta < 1/2$, we have

$$\begin{aligned}&x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n1}(x)\\ &= x^{1/\gamma}\left\{\frac{\gamma}{\gamma_1}\mathbf{W}\left(x^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1+\gamma_2}\int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}\left(x^{-1/\gamma}t\right)dt\right\} + o_{\mathbf{p}}(1)\,x^{(1-\eta)/\gamma\pm\epsilon}.\end{aligned}$$

Before we treat the term $\mathbf{M}_{n2}(x)$, it is worth mentioning that by letting $x = 1$ in the previous approximation, we infer that

$$\frac{\overline{\mathbf{F}}_n(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})} - 1 = O_{\mathbf{p}}\left(k^{-1/2}\right) = o_{\mathbf{p}}(1). \tag{4.22}$$

This, with the regular variation of $\overline{\mathbf{F}}$, implies that

$$\frac{\overline{\mathbf{F}}\left(xX_{n-k:n}\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)} = \left(1 + O_{\mathbf{p}}\left(x^{\pm\epsilon}\right)\right) x^{-1/\gamma_1}. \tag{4.23}$$

To represent $\sqrt{k}\mathbf{M}_{n2}\left(x\right)$, we apply results (4.21) (for $x = 1$) and (4.23) to get

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n2}\left(x\right) = -\left\{\frac{\gamma}{\gamma_1}\mathbf{W}\left(1\right) + \frac{\gamma}{\gamma_1+\gamma_2}\int_0^1 t^{-\gamma/\gamma_2-1}\mathbf{W}\left(t\right)dt\right\} + o_{\mathbf{p}}\left(x^{\pm\epsilon}\right).$$

For the third term $\mathbf{M}_{n3}\left(x\right)$, we write

$$x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n3}\left(x\right) = \left(\frac{\overline{\mathbf{F}}\left(xX_{n-k:n}\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)} - x^{-1/\gamma_1}\right) x^{1/\gamma_1}\sqrt{k}\mathbf{M}_{n1}\left(x\right),$$

which, by equation (4.23), is equal to $o_{\mathbf{p}}\left(1\right) x^{-1/\gamma_1+(1-\eta)/\gamma\pm\epsilon}$. Let η_0 be a real number such that $\gamma/\gamma_2 < \eta_0 < \eta < 1/2$, then $\eta_0 - \eta < 0$ and for $\epsilon > 0$ sufficiently small, we have $\left(\eta_0 - \eta\right)/\gamma + \epsilon < 0$. Since $x \geq x_0 > 0$, then $o_{\mathbf{p}}\left(1\right) x^{(\eta_0-\eta)/\gamma\pm\epsilon} = o_{\mathbf{p}}\left(1\right)$ and thus

$$x^{1/\gamma_1-(1-\eta_0)/\gamma}\left\{\sqrt{k}\left(\mathbf{M}_{n1}\left(x\right) + \mathbf{M}_{n2}\left(x\right) + \mathbf{M}_{n3}\left(x\right)\right) - \mathbf{\Gamma}\left(x;\mathbf{W}\right)\right\} = o_{\mathbf{p}}\left(1\right), \tag{4.24}$$

where $\mathbf{\Gamma}\left(x;\mathbf{W}\right)$ is the Gaussian process given in Theorem 4.1.1. For the fourth term $\mathbf{M}_{n4}\left(x\right)$, it suffices to use the uniform inequality, corresponding to the second-order condition (4.4), given in assertion (2.3.23) of Theorem 2.3.9 in [de Haan and Ferreira(2006)], to get

$$\sqrt{k}\mathbf{M}_{n4}\left(x\right) = \left(1 + o_{\mathbf{p}}\left(1\right)\right) x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1\tau_1}\sqrt{k}\mathbf{A}_0\left(1/\overline{F}\left(X_{n-k:n}\right)\right).$$

Then Proposition B.1.10 in [de Haan and Ferreira(2006)] and the facts that $\mathbf{A}_0\left(t\right)$ is regularly varying at infinity with index τ_1/γ_1 and $X_{n-k:n}/a_k \overset{\mathbf{P}}{\to} 1$, imply that

$$\frac{\mathbf{A}_0\left(1/\overline{F}\left(X_{n-k:n}\right)\right)}{\mathbf{A}_0\left(1/\overline{F}\left(a_k\right)\right)} \overset{\mathbf{P}}{\to} 1, \text{ as } N \to \infty.$$

By assumption, given $n = m$, we have $\sqrt{k_m}\mathbf{A}_0\left(m/k_m\right) = O\left(1\right)$ which implies that $\sqrt{k}\mathbf{A}_0\left(n/k\right) = O_{\mathbf{P}}\left(1\right)$ and thus $\sqrt{k}\mathbf{A}_0\left(1/\overline{F}\left(a_k\right)\right) = O_{\mathbf{P}}\left(1\right).$ On the other hand, we have

$$o_{\mathbf{p}}\left(1\right)x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1}-1}{\gamma_1\tau_1} = o_{\mathbf{p}}\left(x^{-1/\gamma_1+(1-\eta)/\gamma\pm\epsilon}\right).$$

It follows that

$$\sqrt{k}\mathbf{M}_{n4}\left(x\right) = x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1}-1}{\gamma_1\tau_1}\sqrt{k}\mathbf{A}_0\left(n/k\right) + o_{\mathbf{p}}\left(x^{-1/\gamma_1+(1-\eta)/\gamma\pm\epsilon}\right).$$

Finally, by letting $\epsilon \downarrow 0$ in (4.24), we end up with

$$\sup_{x\geq x_0} x^{1/\gamma_1-(1-\eta_0)/\gamma}\left|\mathbf{D}_n\left(x\right) - \mathbf{\Gamma}\left(x;\mathbf{W}\right) - x^{-1/\gamma_1}\frac{x^{\tau_1/\gamma_1}-1}{\gamma_1\tau_1}\sqrt{k}\mathbf{A}_0\left(n/k\right)\right| \xrightarrow{\mathbf{P}} 0,$$

as $N \to \infty$, for every $x_0 > 0$ and $\gamma/\gamma_2 < \eta_0 < \eta < 1/2$. Letting $\eta_0 := 1/2 - \xi$ and recalling that $1/\gamma_1 = 1/\gamma - 1/\gamma_2$ yields that $0 < \xi < 1/2 - \gamma/\gamma_2$ and achieves the proof. □

Proof of Theorem 4.2.1

We start by proving the consistency of $\widehat{\gamma}_1$ that we write as

$$\widehat{\gamma}_1 = \int_1^\infty x^{-1}\frac{\overline{\mathbf{F}}_n\left(xX_{n-k:n}\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)}dx.$$

It is readily checked that this may be decomposed into the sum of

$$I_{1n} := \int_1^\infty x^{-1}\frac{\overline{\mathbf{F}}\left(xX_{n-k:n}\right)}{\overline{\mathbf{F}}\left(X_{n-k:n}\right)}dx \text{ and } I_{2n} := \int_1^\infty x^{-1}\sum_{i=1}^{3}\mathbf{M}_{ni}\left(x\right)dx.$$

First, we show that $I_{1n} \xrightarrow{\mathbf{P}} \gamma_1$ as $N \to \infty$. Indeed, let N be sufficiently large, hence n is also sufficiently large (in probability), then given $n = m \to \infty$, we have $z := X_{m-k_m:m} \xrightarrow{\mathbf{P}} \infty$ (because $\mathbf{F}$ is a Pareto-type distribution). Then by using the first inequality in (4.29), we write that: for all large m and any $0 < \epsilon < 1$, there exists $m_0 = m_0\left(\epsilon\right)$, such that for all

$m > m_0$ and $x \geq 1$,

$$\left| \frac{\overline{\mathbf{F}}\left(xX_{m-k_m:m}\right)}{\overline{\mathbf{F}}\left(X_{m-k_m:m}\right)} - x^{-1/\gamma_1} \right| < \epsilon x^{-1/\gamma_1 \pm \epsilon},$$

with probability greater than $1-\epsilon$. Multiplying by x^{-1} then integrating, on $(1,\infty)$, the two sides of the previous inequality yield that $|I_{1m} - \gamma_1| < \dfrac{\epsilon}{1/\gamma_1 \pm \epsilon}$, with probability greater than $1-\epsilon$, where $I_{1m} := \int_1^\infty x^{-1} \dfrac{\overline{\mathbf{F}}\left(xX_{m-k_m:m}\right)}{\overline{\mathbf{F}}\left(X_{m-k_m:m}\right)} dx$. By using the total probabilities formula and similar arguments as those used in Lemma 4.6.3 and Lemma 4.6.4, we get $\mathbf{P}\left\{ |I_{1n} - \gamma_1| > \dfrac{\epsilon}{1/\gamma_1 \pm \epsilon} \right\} < \epsilon$, meaning that $I_{1n} \xrightarrow{\mathbf{P}} \gamma_1$ as $N \to \infty$. Now, it remains to show that $I_{2n} \xrightarrow{\mathbf{P}} 0$. Observe that, from (4.24), we have

$$I_{2n} = \frac{1}{\sqrt{k}} \int_1^\infty x^{-1} \mathbf{\Gamma}\left(x; \mathbf{W}\right) dx + \frac{1}{\sqrt{k}} \int_1^\infty x^{-1} o_{\mathbf{P}}\left(x^{(1-\eta)/\gamma - 1/\gamma_1}\right) dx.$$

On the one hand, since $\gamma/\gamma_2 < \eta$, the second integral above is finite and therefore the second term of I_{2n} is negligible in probability. On the other hand, we have

$$\begin{aligned} \int_1^\infty x^{-1} \mathbf{\Gamma}\left(x; \mathbf{W}\right) dx &= \frac{\gamma}{\gamma_1} \int_1^\infty x^{1/\gamma_2 - 1} \left\{ \mathbf{W}\left(x^{-1/\gamma}\right) - x^{-1/\gamma} \mathbf{W}\left(1\right) \right\} dx + \frac{\gamma}{\gamma_1 + \gamma_2} \\ &\times \int_1^\infty x^{1/\gamma_2 - 1} \left\{ \int_0^1 s^{-\gamma/\gamma_2 - 1} \left\{ \mathbf{W}\left(x^{-1/\gamma} s\right) - x^{-1/\gamma} \mathbf{W}\left(s\right) \right\} ds \right\} dx, \end{aligned}$$

which, after some elementary but tedious manipulations of integral calculus (change of variables and integration by parts), becomes

$$\begin{aligned} \int_1^\infty x^{-1} \mathbf{\Gamma}\left(x; \mathbf{W}\right) dx &= -\gamma \mathbf{W}\left(1\right) \\ &+ \frac{\gamma}{\gamma_1 + \gamma_2} \int_0^1 \left(\gamma_2 - \gamma_1 - \gamma \log s\right) s^{-\gamma/\gamma_2 - 1} \mathbf{W}\left(s\right) ds. \end{aligned} \tag{4.25}$$

Since $\mathbf{E}\left|\mathbf{W}\left(s\right)\right| \leq s^{1/2}$ and $\gamma_1 < \gamma_2$, then $\int_1^\infty x^{-1} \mathbf{\Gamma}\left(x; \mathbf{W}\right) dx$ is stochastically bounded and therefore the first term of I_{2n} is negligible in probability as well. Consequently, we have $I_{2n} = o_{\mathbf{P}}\left(1\right)$ when $N \to \infty$, as sought. As for the Gaussian representation result, it is easy to verify that $\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)=\int_1^{\infty}x^{-1}\mathbf{D}_n\left(x\right)dx$. Then, applying Theorem 4.1.1 yields that

$$\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)=\frac{\sqrt{k}\mathbf{A}_0\left(n/k\right)}{1-\tau_1}+\int_1^{\infty}x^{-1}\mathbf{\Gamma}\left(x;\mathbf{W}\right)dx+o_{\mathbf{P}}\left(1\right),$$

and finally, using result (4.25) completes the proof. □

Proof of Corollary 4.2.1

We set

$$\sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)=\gamma\Delta+\frac{\sqrt{k}\mathbf{A}_0\left(n/k\right)}{1-\tau_1}+o_{\mathbf{P}}\left(1\right),$$

where $\Delta:=a\Delta_1+b\Delta_2-\Delta_3$, with $a:=\left(\gamma_2-\gamma_1\right)/\left(\gamma_1+\gamma_2\right)$, $b:=-\gamma/\left(\gamma_1+\gamma_2\right)$ and

$$\Delta_1:=\int_0^1 s^{\rho-2}\mathbf{W}\left(s\right)ds,\ \ \Delta_2:=\int_0^1 s^{\rho-2}\mathbf{W}\left(s\right)\log s ds,\ \ \Delta_3:=\mathbf{W}\left(1\right),$$

with $\rho:=1-\gamma/\gamma_2>0$. It is clear that the asymptotic mean is equal

$$\frac{\sqrt{k}\mathbf{A}_0\left(n/k\right)}{1-\tau_1}\overset{\mathbf{P}}{\rightarrow}\frac{\lambda}{1-\tau_1},\ \text{ as } N\rightarrow\infty.$$

For the asymptotic variance we find, after elementary but tedious computations, the following covariances:

$$\mathbf{E}\left[\Delta_1^2\right]=\frac{2}{\rho\left(2\rho-1\right)},\ \mathbf{E}\left[\Delta_2^2\right]=\frac{2\left(4\rho-1\right)}{\rho^2\left(2\rho-1\right)^3},\ \mathbf{E}\left[\Delta_3^2\right]=1,$$
$$\mathbf{E}\left[\Delta_1\Delta_2\right]=\frac{1-4\rho}{\rho^2\left(2\rho-1\right)^2},\ \mathbf{E}\left[\Delta_1\Delta_3\right]=\frac{1}{\rho},\ \mathbf{E}\left[\Delta_2\Delta_3\right]=-\frac{1}{\rho^2}.$$

It follows that $\mathbf{E}\left[\Delta^2\right]=\dfrac{2a^2}{\rho\left(2\rho-1\right)}+\dfrac{2b^2\left(4\rho-1\right)}{\rho^2\left(2\rho-1\right)^3}+\dfrac{2ab\left(1-4\rho\right)}{\rho^2\left(2\rho-1\right)^2}+\dfrac{2b}{\rho^2}-\dfrac{2a}{\rho}+1$. Replacing a, b and ρ by their values achieves the proof. □

Proof of Theorem 4.4.1

Recall that the high quantile of level $1-\nu$ and its Weissman-type estimator are respectively defined by

$$q_\nu := \mathbf{F}^{\leftarrow}(1-\nu) \text{ and } \widehat{q}_\nu := X_{n-k:n}\left(\frac{\nu}{\overline{\mathbf{F}}_n(X_{n-k:n})}\right)^{-\widehat{\gamma}_1}.$$

It is readily checked that

$$\frac{\widehat{q}_\nu}{q_\nu} = \frac{X_{n-k:n}}{a_k}\left(\frac{\overline{\mathbf{F}}(a_k)}{\overline{\mathbf{F}}_n(X_{n-k:n})}\right)^{-\widehat{\gamma}_1}\left\{\frac{\mathbf{F}^{\leftarrow}\left(1-\overline{\mathbf{F}}(a_k)\right)}{\mathbf{F}^{\leftarrow}(1-\nu)}\left(\frac{\nu}{\overline{\mathbf{F}}(a_k)}\right)^{-\widehat{\gamma}_1}\right\},$$

and its logarithm is equal to the sum of

$$T_{n1} := \log\frac{X_{n-k:n}}{a_k},\ T_{n2} := \log\left(\frac{\overline{\mathbf{F}}(a_k)}{\overline{\mathbf{F}}_n(X_{n-k:n})}\right)^{-\widehat{\gamma}_1}$$

and

$$T_{n3} := \log\left\{\frac{\mathbf{F}^{\leftarrow}\left(1-\overline{\mathbf{F}}(a_k)\right)}{\mathbf{F}^{\leftarrow}(1-\nu)}\left(\frac{\nu}{\overline{\mathbf{F}}(a_k)}\right)^{-\widehat{\gamma}_1}\right\}.$$

We will show that $\sqrt{k}T_{ni}/\log d_n \overset{\mathbf{P}}{\to} 0$, $i=1,2$, and $\sqrt{k}T_{n3}/\log d_n$ is asymptotically Gaussian with mean $\lambda/(1-\tau_1)$ and variance σ^2, were $(\lambda,\tau_1,\sigma^2)$ are those given in Corollary 4.2.1. From Lemma 4.6.4, we have $X_{n-k:n}/a_k \overset{\mathbf{P}}{\to} 1$, then by using the approximation $\log(1+x)\sim x$, as $x\downarrow 0$, we write

$$T_{n1} = \log\frac{X_{n-k:n}}{a_k} = (1+o_{\mathbf{P}}(1))\left(\frac{X_{n-k:n}}{a_k}-1\right),\ \text{as } N\to\infty.$$

Next we show that $\sqrt{k}\,(X_{n-k:n}/a_k-1)$ is asymptotically stochastically bounded. Indeed, let us write

$$\frac{X_{n-k:n}}{a_k}-1 = \left\{\frac{X_{n-k:n}}{a_k}-\left(\frac{1/U_{k+1:n}}{n/k}\right)^{\gamma}\right\}+\left\{\left(\frac{1/U_{k+1:n}}{n/k}\right)^{\gamma}-1\right\} =: H_{n1}+H_{n2},$$

where $U_{k+1:n}=\overline{F}\left(X_{n-k:n}\right)$ is the $(k+1)$-th order statistic pertaining to the $(0,1)$-uniform sample $U_1,...,U_n$. Recall that $\overline{F}\left(a_k\right)=k/n$ and rewrite H_{n1} into

$$H_{n1}=\left(\left(\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma}\right)^{-\gamma}-\left(\frac{\overline{F}\left(X_{n-k:n}\right)}{\overline{F}\left(a_k\right)}\right)^{-\gamma}.$$

By using the mean value theorem, together with the fact that $X_{n-k:n}/a_k\overset{\mathbf{P}}{\to}1$ and the regular variation of $\overline{F}$, we get

$$H_{n1}=\gamma\left(1+o_{\mathbf{P}}\left(1\right)\right)\left(\frac{\overline{F}\left(X_{n-k:n}\right)}{\overline{F}\left(a_k\right)}-\left(\frac{X_{n-k:n}}{a_k}\right)^{-1/\gamma}\right),\text{ as } N\to\infty.$$

By applying Lemma 4.6.5 with similar arguments as those used in the proof of Lemma 4.6.4, we get $H_{n1}=O_{\mathbf{P}}\left(\left|\mathbf{A}_{\mathbf{F}}^*\left(n/k\right)\right|+\left|\mathbf{A}_{\mathbf{G}}^*\left(n/k\right)\right|\right)$, as $N\to\infty$. On the other hand, by assumptions, we infer that $\sqrt{k}\mathbf{A}_{\mathbf{F}}^*\left(n/k\right)\overset{\mathbf{P}}{\to}\lambda$ and $\sqrt{k}\mathbf{A}_{\mathbf{G}}^*\left(n/k\right)=O_{\mathbf{P}}\left(1\right)$ as $N\to\infty$. It follows, by analogous manipulations as those of the proof of Lemma 4.6.3, that $\sqrt{k}H_{n1}=O_{\mathbf{P}}\left(1\right)$, as $N\to\infty$. Let us now focus on the term H_{n2}. Note that in the proof of Lemma 4.6.4, we showed that $\frac{n}{k}U_{k+1:n}\overset{\mathbf{P}}{\to}1$, as $N\to\infty$, then by using the mean value theorem, we get

$$H_{n2}=\left(1+o_{\mathbf{P}}\left(1\right)\right)\gamma\left(1-\frac{n}{k}U_{k+1:n}\right),\text{ as } N\to\infty.$$

Observe that $\sqrt{k}\left(1-\frac{n}{k}U_{k+1:n}\right)=\alpha_n\left(\frac{n}{k}U_{k+1:n}\right)$, then by using approximation (4.12), we have that

$$\sqrt{k}\left(1-\frac{n}{k}U_{k+1:n}\right)=\mathbf{W}\left(\frac{n}{k}U_{k+1:n}\right)+o_{\mathbf{P}}\left(1\right).$$

Once again, by making use of Levy's modulus of continuity of the Wiener process with the fact that $\frac{n}{k}U_{k+1:n}\overset{\mathbf{P}}{\to}1$, as $N\to\infty$, we readily show, by similar arguments as those used for inequality (4.17), that

$$\sqrt{k}\left(1-\frac{n}{k}U_{k+1:n}\right)=\mathbf{W}\left(1\right)+o_{\mathbf{P}}\left(1\right).$$

It follows that $\sqrt{k}\left(1-\dfrac{n}{k}U_{k+1:n}\right)$ is an asymptotically centred Gaussian rv with variance 1, hence $\sqrt{k}H_{n2}$ is asymptotically stochastically bounded. By assumptions, we have $d_n \overset{\mathbf{P}}{\to} 0$, that is $\log d_n \overset{\mathbf{P}}{\to} -\infty$, as $N\to\infty$, therefore $\sqrt{k}T_{n1}/\log d_n \overset{\mathbf{P}}{\to} 0$, as $N\to\infty$. For the term T_{n2}, we write

$$T_{n2} = -\widehat{\gamma}_1 \log\left(\frac{\overline{\mathbf{F}}\left(a_k\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)}\right).$$

In view of Lemma 4.6.4 and the regular variation of $\overline{\mathbf{F}}$, we infer that $\overline{\mathbf{F}}\left(a_k\right)/\overline{\mathbf{F}}\left(X_{n-k:n}\right)$ tends to 1 in probability, as $N\to\infty$. Then by using assertion (4.22), we get $\overline{\mathbf{F}}\left(a_k\right)/\overline{\mathbf{F}}_n\left(X_{n-k:n}\right) \overset{\mathbf{P}}{\to}$ 1, as $N\to\infty$ as well. Now, we (once again) use the approximation $\log(1+x)\sim x$, as $x\to 0$, to write

$$\sqrt{k}T_{n2} = -\widehat{\gamma}_1\left(1+o_{\mathbf{P}}(1)\right)\sqrt{k}\left(\frac{\overline{\mathbf{F}}\left(a_k\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)}-1\right),$$

which, by assertion (4.22), is asymptotically stochastically bounded. Consequently, we have $\sqrt{k}T_{n2}/\log d_n \overset{\mathbf{P}}{\to} 0$, as $N\to\infty$. Finally, the third term T_{n3} may be rewritten into

$$T_{n3} = \log\left\{\frac{\mathbf{F}^{\leftarrow}\left(1-\overline{\mathbf{F}}\left(a_k\right)\right)}{\mathbf{F}^{\leftarrow}\left(1-\nu\right)}\left(\frac{\nu}{\overline{\mathbf{F}}\left(a_k\right)}\right)^{-\gamma_1}\left(\frac{\nu}{\overline{\mathbf{F}}\left(a_k\right)}\right)^{-\widehat{\gamma}_1+\gamma_1}\right\},$$

which equals

$$\left(-\widehat{\gamma}_1+\gamma_1\right)\log\left(\nu/\overline{\mathbf{F}}\left(a_k\right)\right)+\log\left\{\frac{\mathbf{F}^{\leftarrow}\left(1-\overline{\mathbf{F}}\left(a_k\right)\right)}{\mathbf{F}^{\leftarrow}\left(1-\nu\right)}\left(\frac{\nu}{\overline{\mathbf{F}}\left(a_k\right)}\right)^{-\gamma_1}\right\} =: K_{n1}+K_{n2}.$$

By substituting $1/d_n$ for $\nu/\overline{\mathbf{F}}\left(a_k\right)$, we get $\sqrt{k}K_{n1}/\log d_n = \sqrt{k}\left(\widehat{\gamma}_1-\gamma_1\right)$ which, by Corollary 4.2.1, is asymptotically Gaussian with mean $\lambda/\left(1-\tau_1\right)$ and variance σ^2. It remains to show that $\sqrt{k}K_{n2}/\log d_n \overset{\mathbf{P}}{\to} 0$, as $N\to\infty$. Indeed, it is easy to check that K_{n2} may be rewritten into

$$K_{n2} = -\left\{\log\frac{\mathbb{U}_{\mathbf{F}}\left(1/\nu\right)}{\mathbb{U}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}\left(a_k\right)\right)}-\gamma_1\log\left(\frac{1/\nu}{1/\overline{\mathbf{F}}\left(a_k\right)}\right)\right\},$$

Note that the second-order regular variation condition (4.4) is equivalent to

$$\lim_{t\to\infty}\frac{\log\dfrac{\mathbb{U}_{\mathbf{F}}(tx)}{\mathbb{U}_{\mathbf{F}}(t)}-\gamma_1\log x}{\mathbf{A}_{\mathbf{F}}(t)}=\frac{x^{\tau_1}-1}{\tau_1},\text{ for } x>0.$$

From the inequality given in Theorem B.2.18 in[de Haan and Ferreira(2006)] page 383, the previous limit implies that: for a possibly different function $\widetilde{\mathbf{A}}_{\mathbf{F}}$, with $\widetilde{\mathbf{A}}_{\mathbf{F}}(t)\sim\mathbf{A}_{\mathbf{F}}(t)$, as $t\to\infty$, and for each $\epsilon>0$, there exists a t_0 such that for $t>t_0$ and $x>1$ we have

$$\left|\frac{\log\dfrac{\mathbb{U}_{\mathbf{F}}(tx)}{\mathbb{U}_{\mathbf{F}}(t)}-\gamma_1\log x}{\widetilde{\mathbf{A}}_{\mathbf{F}}(t)}-\frac{x^{\tau_1}-1}{\tau_1}\right|\leq\epsilon x^{\tau_1+\epsilon}.$$

We apply the inequality above for $x=x_n=d_n$ and $t=t_n=1/\overline{\mathbf{F}}(a_k)$ and we use the fact that $\tau_1<0$, to readily show that $K_{n2}=O_{\mathbf{p}}\left(\widetilde{\mathbf{A}}_{\mathbf{F}}(n/k)\right)$. Since, by assumption, we have $\sqrt{k}\mathbf{A}_{\mathbf{F}}(n/k)\overset{\mathbf{P}}{\to}\lambda<\infty$, then $\sqrt{k}K_{n2}=O_{\mathbf{p}}(1)$ and therefore $\sqrt{k}K_{n2}/\log d_n\overset{\mathbf{P}}{\to}0$, as $N\to\infty$. Finally, we end up with $\left(\sqrt{k}/\log d_n\right)\log\left(\widehat{q}_\nu/q_\nu\right)\overset{\mathcal{D}}{\to}\mathcal{N}\left(\frac{\lambda}{1-\tau_1},\sigma^2\right)$, as $N\to\infty$, and we achieve the proof by noting that $\log\left(\widehat{q}_\nu/q_\nu\right)=\left(1+o_{\mathbf{p}}(1)\right)\left(\widehat{q}_\nu/q_\nu-1\right)$.

4.6 Appendix

The following lemmas are instrumental for our needs.

Lemma 4.6.1 *Assume that both second-order conditions* (4.4) *and* (4.5) *hold. Then, for all large x, there exist constants $s_1,s_2>0$, such that*

$$\overline{F}(x)=(1+o(1))\,s_1x^{-1/\gamma}\text{ and }\overline{G}(x)=(1+o(1))\,s_2x^{-1/\gamma_2}.$$

Proof. We only show the first statement since the second one follows by similar arguments. To this end, we rewrite the first equation of (1.2) into

$$\overline{F}(x) = -p^{-1}\overline{\mathbf{G}}(x)\overline{\mathbf{F}}(x)\int_1^\infty \frac{\overline{\mathbf{G}}(xz)}{\overline{\mathbf{G}}(x)} d\frac{\overline{\mathbf{F}}(xz)}{\overline{\mathbf{F}}(x)}.$$

By applying Proposition B.1.10 in [de Haan and Ferreira(2006)] to both $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$, it is easy to check that

$$\int_1^\infty \frac{\overline{\mathbf{G}}(xz)}{\overline{\mathbf{G}}(x)} d\frac{\overline{\mathbf{F}}(xz)}{\overline{\mathbf{F}}(x)} = -(1+o(1))\,\gamma/\gamma_1.$$

On the other hand, since $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$ satisfy the aforementioned second-order conditions, then in view of Lemma 3 in [Hua and Joe(2011)], there exist two constants $r_1, r_2 > 0$, such that $\overline{\mathbf{F}}(x) = (1+o(1))\,r_1 x^{-1/\gamma_1}$ and $\overline{\mathbf{G}}(x) = (1+o(1))\,r_2 x^{-1/\gamma_2}$, as $x \to \infty$. Therefore $\overline{F}(x) = (1+o(1))\,s_1 x^{-1/\gamma}$ with $s_1 = p^{-1} r_1 r_2 \gamma/\gamma_1$. ■

Lemma 4.6.2 *Under the assumptions of Lemma* 4.6.1, *we have*

(i) $\lim_{t\to\infty} C(t)/\overline{G}(t) = 1.$

(ii) $\lim_{t\to\infty} t^{1/\nu} C(\mathbb{U}_F(t)) = \infty$, *for each* $0 < \nu \le 1.$

(iii) $\lim_{t\to\infty} \sup_{x\ge x_0} x^{-1/\gamma\pm\epsilon} \left|(t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)))^{-1} - (\gamma/\gamma_1)\,x^{1/\gamma}\right| = 0,$

for $x_0 > 0$ *and any sufficiently small* $\epsilon > 0.$

Proof. For assertion (i), write $C(t)/\overline{G}(t) = 1 - \overline{F}(t)/\overline{G}(t)$ and observe that from Lemma 4.6.1 we have $\overline{F}(t)/\overline{G}(t) = (1+o(1))\,(d_1/d_2)\,t^{1/\gamma_2 - 1/\gamma}$. Since $1/\gamma_2 - 1/\gamma < 0$, then $\overline{F}(t)/\overline{G}(t) = o(1)$, that is $C(t)/\overline{G}(t) = 1 + o(1)$ as sought. For result (ii), Lemma 4.6.1 implies that $\mathbb{U}_F(t) = (1+o(1))\,(d_1 t)^{\gamma}$ (as $t \to \infty$), it follows that $C(\mathbb{U}_F(t)) = (1+o(1))\,d_2\,(d_1 t)^{-\gamma/\gamma_2}$. Since $0 < \gamma/\gamma_2 < 1$, then for every $0 < \nu \le 1$, $t^{1/\nu} C(\mathbb{U}_F(t)) \to \infty$ as $t \to \infty$. To prove (iii), we first show that

$$t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)) - (\gamma_1/\gamma)\,x^{-1/\gamma} = o\left(x^{-1/\gamma\pm\epsilon}\right). \tag{4.26}$$

Recalling that $\Lambda(x) = \int_x^\infty dF(z)/C(z)$ and $\overline{F}(\mathbb{U}_F(t)) = 1/t$, we write

$$t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)) = -\frac{C(x\mathbb{U}_F(t))}{C(\mathbb{U}_F(t))}\int_x^\infty \frac{C(\mathbb{U}_F(t))}{C(z\mathbb{U}_F(t))}\frac{d\overline{F}(z\mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))}. \tag{4.27}$$

Observe now that $t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)) - \dfrac{\gamma_1}{\gamma}x^{-1/\gamma}$ may be decomposed into the sum of

$$L_1(s;t) := -\left(\frac{C(x\mathbb{U}_F(t))}{C(\mathbb{U}_F(t))} - x^{-1/\gamma_2}\right)\int_x^\infty \frac{C(\mathbb{U}_F(t))}{C(z\mathbb{U}_F(t))}\frac{d\overline{F}(z\mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))},$$

$$L_2(s;t) := -x^{-1/\gamma_2}\int_x^\infty \left(\frac{C(\mathbb{U}_F(t))}{C(z\mathbb{U}_F(t))} - z^{1/\gamma_2}\right)\frac{d\overline{F}(z\mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))}$$

and

$$L_3(s;t) := -x^{-1/\gamma_2}\int_x^\infty z^{1/\gamma_2}d\left(\frac{\overline{F}(z\mathbb{U}_F(t))}{\overline{F}(\mathbb{U}_F(t))} - z^{-1/\gamma}\right).$$

By applying Proposition B.1.10 in [de Haan and Ferreira(2006)] to both C and $\overline{F}$ with integrations by parts, it is easy to verify that

$$\left|t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)) - (\gamma_1/\gamma)\,x^{-1/\gamma}\right| \le \epsilon x^{-1/\gamma\pm\epsilon}.$$

Observe now that $t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)) - (\gamma/\gamma_1)\,x^{1/\gamma}$ is equal to

$$\left((t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)))^{-1}\right)^{-1} - \left((\gamma_1/\gamma)\,x^{-1/\gamma}\right)^{-1}.$$

By using the mean value theorem, the latter equals

$$\frac{(\gamma_1/\gamma)\,x^{-1/\gamma} - t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t))}{(\psi(x;t))^2},$$

where $\psi(x;t)$ is between $(\gamma_1/\gamma)\,x^{-1/\gamma}$ and $t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t))$. In view of the representation (4.27) and Potter's inequalities, applied to C and $\overline{F}$, with an integration by parts, we get $t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)) \ge (1-\epsilon)\,x^{-1/\gamma\pm\epsilon}$. It follows that $(\psi(x;t))^2 \ge (1-\epsilon)^2\,x^{-2/\gamma\pm2\epsilon}$ and therefore $\left|t\Lambda(x\mathbb{U}_F(t))\,C(x\mathbb{U}_F(t)) - (\gamma/\gamma_1)\,x^{1/\gamma}\right| \le (1-\epsilon)^{-2}\,\epsilon x^{1/\gamma\pm\epsilon}$, as sought. ■

Lemma 4.6.3 *On the probability space* $(\Omega, \mathcal{A}, \mathbf{P})$, *there exists a standard Wiener process* $\{\mathbf{W}(s)\,;\ s \geq 0\}$, *such that for any* $0 < \eta < 1/2$,

$$\sup_{0<s\leq 1} s^{-\eta} \left|\alpha_n(s) - \mathbf{W}(s)\right| \xrightarrow{\mathbf{P}} 0, \ \text{as } N \to \infty.$$

Proof. Let ϵ, $\delta > 0$ be arbitrary and N an integer sufficiently large. We will show that

$$\mathbf{P}\left\{\sup_{0<s\leq 1} s^{-\eta} \left|\alpha_n(s) - \mathbf{W}(s)\right| > \delta\right\} < \epsilon.$$

Indeed, from (1.1), we infer that $n = n_N \xrightarrow{\mathbf{P}} \infty$ as $N \to \infty$, this means that for any fixed integer $M > 0$,

$$\mathbf{P}\{n \leq M\} < \epsilon/(2M), \text{ for all large } N. \tag{4.28}$$

On the other hand, from Proposition 3.1 of [Einmahl et al.(2006)], there exists a standard Wiener process $\{\mathbf{W}(s)\,;\ s \geq 0\}$ defined on $(\Omega, \mathcal{A}, \mathbf{P})$, such that for given $n = m \to \infty$, we have $Z_m := \sup_{0<s\leq 1} s^{-\eta} \left|\alpha_m(s) - \mathbf{W}(s)\right| \xrightarrow{\mathbf{P}} 0$, for any $0 < \eta < 1/2$. In other words, there exists N_* such that we have $\mathbf{P}\{Z_m > \delta\} < \epsilon/2$, for any $m > N_*$. By the total probabilities formula we have $\mathbf{P}\{Z_n > \delta\} = \sum_{m=1}^{N} \mathbf{P}\{Z_n > \delta, n = m\}$. From Lemma 1 in [Gardes and Stupfler(2015)], we infer that for arbitrary Borel subsets $(A_i)_{i\geq 1}$ of $[x_0, \infty[$, we have

$$\mathbf{P}\{X_1 \in A_1, ..., X_n \in A_n, n = m\} = \mathbf{P}\{n = m\} \prod_{i=1}^{m} \mathbf{P}\{X_i \in A_i\}.$$

Since Z_n is a statistic based on the sample $(X_1, ..., X_n)$, this yields that

$$\mathbf{P}\{Z_n > \delta, n = m\} = \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\}$$

and therefore $\mathbf{P}\{Z_n > \delta\}$ may be written into $\mathbf{P}\{Z_n > \delta\} = \sum_{m=1}^{N} \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\}$. Let us write $\mathbf{P}\{Z_n > \delta\} = \sum_{m=1}^{N_*} \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\} + \sum_{m=N_*+1}^{N} \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\}$. Observe that by taking $M = N_*$ in (4.28), we have $\mathbf{P}\{n = m\} < \epsilon/(2N_*)$, for any $m \leq N_*$, it follows that $\sum_{m=1}^{N_*} \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\} < \epsilon/2$, because $\mathbf{P}\{Z_m > \delta\}$ is less to 1. On the other hand, for any $m > N_*$, we have $\mathbf{P}\{Z_m > \delta\} < \epsilon/2$, therefore

$$\sum_{m=N_*+1}^{N} \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\} \leq \sum_{m=N_*+1}^{\infty} \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\}$$
$$\leq (\epsilon/2) \sum_{m=N_*+1}^{\infty} \mathbf{P}\{n = m\}.$$

Since $\sum_{m=1}^{\infty} \mathbf{P}\{n = m\} = 1$, then $\sum_{m=N_*+1}^{\infty} \mathbf{P}\{Z_m > \delta\} \mathbf{P}\{n = m\} < \epsilon/2$. Hence we showed that, for any $N > N_*$, we have $\mathbf{P}\{Z_n > \delta\} < \epsilon$, as sought. ■

Lemma 4.6.4 *Let k be an integer sequence satisfying* (4.1), *then*

$$X_{n-k:n}/\mathbb{U}_F(n/k) \xrightarrow{\mathbf{P}} 1, \text{ as } N \to \infty.$$

Proof. Given $n = m$, the rv $U_{k_m+1:m}$ is the $(k_m + 1)$-th order statistic pertaining to the sequence of iid $(0,1)$-uniform rv's $U_i = \overline{F}(X_i)$, $i = 1, ..., m$. It is well-known that $U_{j:m} \overset{\mathcal{D}}{=} S_j/S_{m+1}$, $j = 1, ..., m$, where S_j is the j-partial sum of iid standard exponential rv's, see, e.g., Proposition 1 in page 335 of [Shorak and Wellner(1986)]. Hence $U_{k_m+1:m} \overset{\mathcal{D}}{=} S_{k_m+1}/S_{m+1}$ and by using the law of total probabilities, we easily show that $U_{k+1:n} \overset{\mathcal{D}}{=} S_{k+1}/S_{n+1}$ as well. Observe that

$$\frac{n}{k} U_{k+1:n} \overset{\mathcal{D}}{=} \frac{n}{n+1} \frac{k+1}{k} \left\{ \frac{S_{k+1}}{k+1} \right\} \left\{ \frac{S_{n+1}}{n+1} \right\}^{-1}.$$

Then by using the law of large numbers, we get $\frac{n}{k} U_{k+1:n} \xrightarrow{\mathbf{P}} 1$, as $N \to \infty$. To achieve the proof, it suffices to write $X_{n-k:n}/\mathbb{U}_F(n/k) = \mathbb{U}_F(1/U_{k+1:n})/\mathbb{U}_F(n/k)$ and apply Proposition B.1.10 in [de Haan and Ferreira(2006)] to the regularly varying function $\mathbb{U}_F$, therefore the details are omitted. ■

Lemma 4.6.5 *Under the assumptions of Lemma* 4.6.1, *we have for any $x_0 > 0$*

$$\sup_{x \geq x_0} \left| \frac{\overline{F}(ux)}{\overline{F}(u)} - x^{-1/\gamma} \right| = O\left\{ \left| \mathbf{A_F}\left(1/\overline{\mathbf{F}}(u)\right) \right| + \left| \mathbf{A_G}\left(1/\overline{\mathbf{G}}(u)\right) \right| \right\}, \text{ as } u \to \infty.$$

Proof. Recall (1.2) and observe that, for $x \geq x_0 > 0$, we have

$$\overline{F}(x) = -p^{-1}\overline{\mathbf{F}}(x)\overline{\mathbf{G}}(x)\int_1^\infty \frac{\overline{\mathbf{G}}(xt)}{\overline{\mathbf{G}}(x)} d\frac{\overline{\mathbf{F}}(xt)}{\overline{\mathbf{F}}(x)},$$

which implies that

$$\frac{\overline{F}(ux)}{\overline{F}(u)} = \frac{\overline{\mathbf{F}}(ux)\overline{\mathbf{G}}(ux)}{\overline{\mathbf{F}}(u)\overline{\mathbf{G}}(u)} \frac{\int_1^\infty \frac{\overline{\mathbf{G}}(uxt)}{\overline{\mathbf{G}}(ux)} d\frac{\overline{\mathbf{F}}(uxt)}{\overline{\mathbf{F}}(ux)}}{\int_1^\infty \frac{\overline{\mathbf{G}}(ut)}{\overline{\mathbf{G}}(u)} d\frac{\overline{\mathbf{F}}(ut)}{\overline{\mathbf{F}}(u)}}.$$

Let us write $\overline{F}(ux)/\overline{F}(u) - x^{-1/\gamma}$ into the sum of

$$L_1(x;u) := \left\{ \frac{\int_1^\infty \frac{\overline{\mathbf{G}}(uxt)}{\overline{\mathbf{G}}(ux)} d\frac{\overline{\mathbf{F}}(uxt)}{\overline{\mathbf{F}}(ux)}}{\int_1^\infty \frac{\overline{\mathbf{G}}(ut)}{\overline{\mathbf{G}}(u)} d\frac{\overline{\mathbf{F}}(ut)}{\overline{\mathbf{F}}(u)}} \right\} \left\{ \frac{\overline{\mathbf{F}}(ux)\overline{\mathbf{G}}(ux)}{\overline{\mathbf{F}}(u)\overline{\mathbf{G}}(u)} - x^{-1/\gamma} \right\}$$

and

$$L_2(x;u) := x^{-1/\gamma} \left\{ \frac{\int_1^\infty \frac{\overline{\mathbf{G}}(uxt)}{\overline{\mathbf{G}}(ux)} d\frac{\overline{\mathbf{F}}(uxt)}{\overline{\mathbf{F}}(ux)}}{\int_1^\infty \frac{\overline{\mathbf{G}}(ut)}{\overline{\mathbf{G}}(u)} d\frac{\overline{\mathbf{F}}(ut)}{\overline{\mathbf{F}}(u)}} - 1 \right\}.$$

Note that, the first factor in $L_1(x;u)$ tends to 1 as $u \to \infty$, uniformly on $x \geq x_0$. To prove this, we use Proposition B.1.10 in [de Haan and Ferreira(2006)], page 369, to the regularly varying functions $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$: for any $0 < \epsilon < 1$, there exists $y_0 = y_0(\epsilon)$, such that for all $yz \geq y_0$ we have

$$\left|\overline{\mathbf{F}}(yz)/\overline{\mathbf{F}}(y) - z^{-1/\gamma_1}\right| < \epsilon z^{-1/\gamma_1 \pm \epsilon} \text{ and } \left|\overline{\mathbf{G}}(yz)/\overline{\mathbf{G}}(y) - z^{-1/\gamma_2}\right| < \epsilon z^{-1/\gamma_2 \pm \epsilon}. \quad (4.29)$$

To achieve the proof, it suffices to apply successively the previous inequalities and the uniform inequalities to the second-order regularly varying functions $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$, which say that for possibly different functions $\widetilde{\mathbf{A}}_{\mathbf{F}}$ and $\widetilde{\mathbf{A}}_{\mathbf{G}}$, with $\widetilde{\mathbf{A}}_{\mathbf{F}}(t) \sim \mathbf{A}_{\mathbf{F}}(t)$ and $\widetilde{\mathbf{A}}_{\mathbf{G}}(t) \sim \mathbf{A}_{\mathbf{G}}(t)$, as $t \to \infty$, and for each $0 < \epsilon < 1$, there exists $y_0 = y_0(\epsilon)$, such that for all

$yz \geq y_0$ we have

$$\left| \frac{\overline{\mathbf{F}}(yz)/\overline{\mathbf{F}}(y) - z^{-1/\gamma_1}}{\widetilde{\mathbf{A}}_{\mathbf{F}}\left(1/\overline{\mathbf{F}}(y)\right)} - z^{-1/\gamma_1} \frac{z^{\tau_1/\gamma_1} - 1}{\tau_1/\gamma_1} \right| \leq \epsilon z^{-1/\gamma_1 \pm \epsilon},$$

$$\left| \frac{\overline{\mathbf{G}}(yz)/\overline{\mathbf{G}}(y) - z^{-1/\gamma_2}}{\widetilde{\mathbf{A}}_{\mathbf{G}}\left(1/\overline{\mathbf{G}}(y)\right)} - z^{-1/\gamma_2} \frac{z^{\tau_2/\gamma_2} - 1}{\tau_2/\gamma_2} \right| \leq \epsilon z^{-1/\gamma_2 \pm \epsilon}.$$

see, e.g., Proposition 4 and Remark 1 in [Hua and Joe(2011)]. The rest of the proof consists in elementary calculations, therefore we omit the details. ■

Chapter 5

Kernel estimation of the tail index for right-truncated data

In this chapter, we define a kernel estimator for the tail index of a Pareto-type distribution under random right-truncation and establish its asymptotic normality. A simulation study shows that, compared to the estimators recently proposed by [Gardes and Stupfler(2015)] and [Benchaira *et al.*(2016)], this newly introduced estimator behaves better, in terms of bias and mean squared error, for small samples.

5.1 Tail index estimation

In this section, we derive a kernel version of $\widehat{\gamma}_1$ in the spirit of what is called kernel estimator of [Csörgö et al.(1985)]. Thereby, for a suitable choice of the kernel function, we obtain an improved estimator of γ_1 in terms of bias and mean squared error. To this end, let $\mathbb{K} : \mathbb{R} \to \mathbb{R}_+$ be a fixed function, that will be called kernel, satisfying:

[$\mathbb{C}$1] $\mathbb{K}$ is non increasing and right-continuous on $\mathbb{R}$;

[$\mathbb{C}$2] $\mathbb{K}(s) = 0$ for $s \notin [0,1)$ and $\mathbb{K}(s) \geq 0$ for $s \in [0,1)$;

[$\mathbb{C}$3] $\int_{\mathbb{R}} \mathbb{K}(s)ds = 1$;

[$\mathbb{C}$4] $\mathbb{K}$ and its first and second Lebesgue derivatives $\mathbb{K}'$ and $\mathbb{K}''$ are bounded on $\mathbb{R}$.

As examples of such functions see, e.g., [Groeneboom *et al.*(2003)], we have the indicator kernel $\mathbb{K} = \mathbf{1}_{[0,1)}$ and the biweight and triweight kernels respectively defined by

$$\mathbb{K}_2(s) := \frac{15}{8}\left(1-s^2\right)^2 \mathbf{1}_{\{0\leq s<1\}},\ \mathbb{K}_3(s) := \frac{35}{16}\left(1-s^2\right)^3 \mathbf{1}_{\{0\leq s<1\}}. \tag{5.1}$$

For an overview of kernel estimation of the extreme value index with complete data, one refers to, for instance, [Hüsler *et al.*(2006)] and [?]. By using Potter's inequalities, see e.g. Proposition B.1.10 in [de Haan and Ferreira(2006)], to the regularly varying function $\overline{\mathbf{F}}$ together with assumptions [$\mathbb{C}$1]-[$\mathbb{C}$3] , we may readily show that

$$\lim_{u\rightarrow\infty}\int_u^\infty x^{-1}\frac{\overline{\mathbf{F}}(x)}{\overline{\mathbf{F}}(u)}\mathbb{K}\left(\frac{\overline{\mathbf{F}}(x)}{\overline{\mathbf{F}}(u)}\right)dx = \gamma_1\int_0^\infty \mathbb{K}(s)ds = \gamma_1. \tag{5.2}$$

An integration by parts yields

$$\lim_{u\rightarrow\infty}\frac{1}{\overline{\mathbf{F}}(u)}\int_u^\infty g_{\mathbb{K}}\left(\frac{\overline{\mathbf{F}}(x)}{\overline{\mathbf{F}}(u)}\right)\log\frac{x}{u}d\mathbf{F}(x) = \gamma_1, \tag{5.3}$$

where $g_{\mathbb{K}}$ denotes the Lebesgue derivative of the function $s \rightarrow \Psi_{\mathbb{K}}(s) := s\mathbb{K}(s)$. Note that, for $\mathbb{K} = \mathbf{1}_{[0,1)}$, we have $g_{\mathbb{K}} = \mathbf{1}_{[0,1)}$, then the previous two limits meet assertion (1.2.6) given in Theorem 1.2.2 by [de Haan and Ferreira(2006)]. For kernels $\mathbb{K}_2$ and $\mathbb{K}_3$, we have

$$g_{\mathbb{K}_2}(s) := \frac{15}{8}\left(1-s^2\right)\left(1-5s^2\right)\mathbf{1}_{\{0\leq s<1\}},\ g_{\mathbb{K}_3}(s) := \frac{35}{16}\left(1-s^2\right)^2\left(1-7s^2\right)\mathbf{1}_{\{0\leq s<1\}}.$$

Since $\overline{F}$ is regularly varying at infinity with tail index $\gamma > 0$, then $X_{n-k:n}$ tends to ∞ almost surely. By replacing, in (5.3), u by $X_{n-k:n}$ and $\mathbf{F}$ by its empirical counterpart $\mathbf{F}_n$, we get

$$\widehat{\gamma}_{1,\mathbb{K}} = \frac{1}{\overline{\mathbf{F}}_n(X_{n-k:n})}\int_{X_{n-k:n}}^\infty g_{\mathbb{K}}\left(\frac{\overline{\mathbf{F}}_n(x)}{\overline{\mathbf{F}}_n(X_{n-k:n})}\right)\log\frac{x}{X_{n-k:n}}d\mathbf{F}_n(x),$$

as a kernel estimator for γ_1. Next, we give an explicit formula for $\widehat{\gamma}_{1,\mathbb{K}}$. Since $\overline{\mathbf{F}}$ and $\overline{\mathbf{G}}$ are regularly varying at infinity with tail indices $\gamma_1 > 0$ and $\gamma_2 > 0$ respectively, then their right endpoints are infinite and so they are equal. Hence, we have the empirical counterpart of equation (1.4). This allow us to rewrite $\widehat{\gamma}_{1,\mathbb{K}}$ into

$$\widehat{\gamma}_{1,\mathbb{K}}=\frac{1}{\overline{\mathbf{F}}_{n}\left(X_{n-k:n}\right)}\int_{X_{n-k:n}}^{\infty}\frac{\mathbf{F}_{n}\left(x\right)}{C_{n}\left(x\right)}g_{\mathbb{K}}\left(\frac{\overline{\mathbf{F}}_{n}\left(x\right)}{\overline{\mathbf{F}}_{n}\left(X_{n-k:n}\right)}\right)\log\frac{x}{X_{n-k:n}}dF_{n}\left(x\right),$$

which is equal to

$$\frac{1}{n\overline{\mathbf{F}}_{n}\left(X_{n-k:n}\right)}\sum_{i=1}^{k}\frac{\mathbf{F}_{n}\left(X_{n-i+1:n}\right)}{C_{n}\left(X_{n-i+1:n}\right)}g_{\mathbb{K}}\left(\frac{\overline{\mathbf{F}}_{n}\left(X_{n-i+1:n}\right)}{\overline{\mathbf{F}}_{n}\left(X_{n-k:n}\right)}\right)\log\frac{X_{n-i+1:n}}{X_{n-k:n}}.$$

In view of this equation

$$C_{n}\left(x\right)d\mathbf{F}_{n}\left(x\right)=\mathbf{F}_{n}\left(x\right)dF_{n}\left(x\right),$$

[Benchaira *et al.*(2016)] showed that

$$\overline{\mathbf{F}}_{n}\left(X_{n-k:n}\right)=\frac{1}{n}\sum_{i=1}^{k}\frac{\mathbf{F}_{n}\left(X_{n-i+1:n}\right)}{C_{n}\left(X_{n-i+1:n}\right)}.$$

Thereby, by setting $a_{n}^{(i)}:=\mathbf{F}_{n}\left(X_{n-i+1:n}\right)/C_{n}\left(X_{n-i+1:n}\right)$, we end up with the final formula of our new kernel estimator

$$\widehat{\gamma}_{1,\mathbb{K}}:=\frac{\sum\limits_{i=1}^{k}a_{n}^{(i)}g_{\mathbb{K}}\left(\frac{\overline{\mathbf{F}}_{n}\left(X_{n-i+1:n}\right)}{\overline{\mathbf{F}}_{n}\left(X_{n-k:n}\right)}\right)\log\frac{X_{n-i+1:n}}{X_{n-k:n}}}{\sum\limits_{i=1}^{k}a_{n}^{(i)}}. \tag{5.4}$$

Note that in the complete data situation, $\mathbf{F}_{n}$ is equal to C_{n} and both reduce to the classical empirical df. As a result, we have in that case $a_{n}^{(i)}=1$ and $\overline{\mathbf{F}}_{n}\left(X_{n-i,n}\right)/\overline{\mathbf{F}}_{n}\left(X_{n-k,n}\right)=i/k$ meaning that $\widehat{\gamma}_{1,\mathbb{K}}=k^{-1}\sum_{i=1}^{k}g_{\mathbb{K}}\left(\frac{i-1}{k}\right)\log\left(X_{n-i+1:n}/X_{n-k:n}\right)$. By applying the mean value theorem to function $\Psi_{\mathbb{K}}$, we get

$$\frac{i}{k}\mathbb{K}\left(\frac{i}{k}\right)-\frac{i-1}{k}\mathbb{K}\left(\frac{i-1}{k}\right)=\frac{1}{k}g_{\mathbb{K}}\left(\frac{i-1}{k}\right)+O\left(\frac{1}{k^{2}}\right),\text{ as }N\rightarrow\infty.$$

It follows that

$$\widehat{\gamma}_{1,\mathbb{K}}=\sum_{i=1}^{k}\left\{\frac{i}{k}\mathbb{K}\left(\frac{i}{k}\right)-\frac{i-1}{k}\mathbb{K}\left(\frac{i-1}{k}\right)\right\}\log\frac{X_{n-i+1:n}}{X_{n-k:n}}+O\left(\frac{1}{k}\right)\widehat{\gamma}_1^{Hill},$$

where $\widehat{\gamma}_1^{Hill}:=k^{-1}\sum\limits_{i=1}^{k}\log\left(X_{n-i+1:n}/X_{n-k:n}\right)$ is Hill's estimator of the tail index γ_1. In view of the consistency of $\widehat{\gamma}_1^{Hill}$ [Mason(1982)], we obtain

$$\widehat{\gamma}_{1,\mathbb{K}}=\sum_{i=1}^{k}\frac{i}{k}\mathbb{K}\left(\frac{i}{k}\right)\log\frac{X_{n-i+1:n}}{X_{n-i:n}}+O_{\mathbf{P}}\left(\frac{1}{k}\right),\text{ as } N\rightarrow\infty,$$

which is an approximation of the aforementioned CDM's kernel estimator of the tail index γ_1 with untrucated data.

5.2 Main results

Theoreme 5.2.1 *Assume that the second-order conditions of regular variation* (4.4) *and* (4.5) *hold with $\gamma_1<\gamma_2$, and let $\mathbb{K}$ be a kernel function satisfying assumptions* [$\mathbb{C}$1]-[$\mathbb{C}$4] *and $k=k_n$ a random sequence of integers such that given $n=m$, $k_m\rightarrow\infty$ and $k_m/m\rightarrow 0$, as $N\rightarrow\infty$. Then, there exist a function $\mathbf{A}_0(t)\sim\mathbf{A}_{\mathbf{F}}^*(t)$, as $t\rightarrow\infty$, and a standard Wiener process $\{\mathbf{W}(s);\ s\geq 0\}$, defined on the probability space $(\Omega,\mathcal{A},\mathbf{P})$ such that*

$$\begin{aligned}&\sqrt{k}\left(\widehat{\gamma}_{1,\mathbb{K}}-\gamma_1\right)\\&=\left(\gamma^2/\gamma_1\right)\int_0^1 s^{-1}\mathbf{W}(s)\,d\{s\varphi_{\mathbb{K}}(s)\}+\sqrt{k}\mathbf{A}_0(n/k)\int_0^1 s^{-\tau_1}\mathbb{K}(s)\,ds+o_{\mathbf{P}}(1),\end{aligned}$$

provided that, given $n=m$, $\sqrt{k_m}\mathbf{A_0}(m/k_m)=O(1)$, as $N\rightarrow\infty$, where

$$\varphi_{\mathbb{K}}(s):=s^{-1}\int_0^s t^{-\gamma/\gamma_2}\left\{\mathbb{K}\left(t^{\gamma/\gamma_1}\right)-\frac{\gamma_1}{\gamma_2}t^{-\gamma_2/\gamma_1}\mathbb{K}\left(t^{\gamma/\gamma_1}\right)+t^{\gamma/\gamma_1}\mathbb{K}'\left(t^{\gamma/\gamma_1}\right)\right\}dt.$$

If in addition, we suppose that, given $n = m$, $\sqrt{k_m}\mathbf{A_0}(m/k_m) \to \lambda$, *then* $\sqrt{k}(\widehat{\gamma}_{1,\mathbb{K}} - \gamma_1) \overset{\mathcal{D}}{\to} \mathcal{N}(\mu_{\mathbb{K}}, \sigma_{\mathbb{K}}^2)$, *as* $N \to \infty$, *where*

$$\mu_{\mathbb{K}} := \lambda \int_0^1 s^{-\tau_1}\mathbb{K}(s)\,ds \text{ and } \sigma_{\mathbb{K}}^2 := \left(\gamma^2/\gamma_1\right)^2 \int_0^1 \varphi_{\mathbb{K}}^2(s)\,ds.$$

Remark 5.2.1 *A very large value of* γ_2 *yields a* γ*-value that is very close to* γ_1*, meaning that the really observed sample is almost the whole dataset. In other words, the complete data case corresponds to the situation when* $1/\gamma_2 \equiv 0$*, in which case we have* $\gamma \equiv \gamma_1$*. It follows that in that case*

$$\varphi_{\mathbb{K}}(s) = \gamma_1 s^{-1} \int_0^s \{\mathbb{K}(t) + t\mathbb{K}'(t)\}\,dt = \gamma_1 s^{-1} \int_0^s d\{t\mathbb{K}(t)\} = \gamma_1 \mathbb{K}(s),$$

and therefore $\sigma_{\mathbb{K}}^2 = \gamma_1^2 \int_0^1 \mathbb{K}^2(s)\,ds$*, which agrees with the asymptotic variance given in Theorem 1 of [Csörgö et al.(1985)].*

5.3 Simulation study

In this section, we check the finite sample behavior of $\widehat{\gamma}_{1,\mathbb{K}}$ and, at the same time, we compare it with $\widehat{\gamma}_1$ and $\widehat{\gamma}_1^{(\mathbf{GS})}$ respectively proposed by [Benchaira *et al.*(2016)] and [Gardes and Stupfler(2015 and defined in (4.8) and (3.4). To this end, we consider two sets of truncated and truncation data, both drawn from Burr's model defined in 4.9. 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 70%, 80% 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. For the construction of our estimator $\widehat{\gamma}_{1,\mathbb{K}}$, we select the biweight and the triweight kernel functions defined in (5.1). 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 each one of the three estimators, we use the algorithm of [Reiss and Thomas(2007)], 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 and 5.2 for $\gamma_1 = 0.6$ and in Tables 5.3 and 5.4 for $\gamma_1 = 0.8$. In light of all four tables, we first note that, as expected, the estimation accuracy of all estimators decreases when the truncation percentage increases. Second, with regard to the bias, the comparison definitely is in favour of the newly proposed tail index estimator $\widehat{\gamma}_{1,\mathbb{K}}$, whereas it is not as clear-cut when the rmse is considered. Indeed, the kernel estimator preforms better than the other pair as far as small samples are concerned while for large datasets, it is $\widehat{\gamma}_1^{(\mathbf{GS})}$ that seems to have the least rmse but with greater bias. As an overall conclusion, one may say that, for case studies where not so many data are at one's disposal, the kernel estimator $\widehat{\gamma}_{1,\mathbb{K}}$ is the most suitable among the three estimators.

5.4 Proofs

The proof is based on a very useful weak approximation to the tail product-limit process recently provided by [Benchaira *et al.*(2016)]. From (5.2), the estimator $\widehat{\gamma}_{1,\mathbb{K}}$ may be rewritten into

$$\widehat{\gamma}_{1,\mathbb{K}} = \int_1^\infty x^{-1}\Psi_{\mathbb{K}}\left(\frac{\overline{\mathbf{F}}_n\left(xX_{n-k:n}\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)}\right)dx.$$

Recall that $\Psi_{\mathbb{K}}(s) = s\mathbb{K}(s)$, then it is easy to verify that $\int_1^\infty x^{-1}\Psi_{\mathbb{K}}\left(x^{-1/\gamma_1}\right)dx = \gamma_1$. Hence

$$\widehat{\gamma}_{1,\mathbb{K}} - \gamma_1 = \int_1^\infty x^{-1}\left\{\Psi_{\mathbb{K}}\left(\frac{\overline{\mathbf{F}}_n\left(xX_{n-k:n}\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n}\right)}\right) - \Psi_{\mathbb{K}}\left(x^{-1/\gamma_1}\right)\right\}dx.$$

Let $\mathbf{D}_n(x)$ defined in 4.2 be the tail product-limit process, then Taylor's expansion of $\Psi_{\mathbb{K}}$ yields that

$$\sqrt{k}\left(\widehat{\gamma}_{1,\mathbb{K}} - \gamma_1\right) = \int_1^\infty x^{-1}\mathbf{D}_n(x)\, g_{\mathbb{K}}\left(x^{-1/\gamma_1}\right)dx + R_{n1},$$

		$\widehat{\gamma}_{1,\mathbb{K}}$		$\widehat{\gamma}_1$		$\widehat{\gamma}_1^{GS}$	
N	n	abs bias	rmse	abs bias	rmse	abs bias	rmse
$p = 0.7$							
150	104	0.073	0.665	0.133	0.408	0.136	3.341
200	140	0.008	0.614	0.152	0.392	0.258	1.647
300	210	0.003	0.467	0.095	0.321	0.102	0.962
500	349	0.007	0.439	0.063	0.296	0.022	0.409
1000	699	0.020	0.284	0.042	0.210	0.023	0.211
1500	1049	0.009	0.255	0.024	0.189	0.013	0.142
2000	1399	0.011	0.245	0.018	0.177	0.013	0.116
$p = 0.8$							
150	120	0.054	0.608	0.093	0.398	0.100	0.989
200	160	0.030	0.520	0.085	0.353	0.109	0.488
300	239	0.022	0.467	0.067	0.322	0.069	0.353
500	399	0.002	0.340	0.049	0.240	0.040	0.196
1000	799	0.013	0.217	0.033	0.168	0.029	0.135
1500	1199	0.003	0.190	0.017	0.140	0.019	0.109
2000	1599	0.005	0.149	0.011	0.113	0.005	0.095
$p = 0.9$							
150	134	0.031	0.492	0.082	0.387	0.149	2.740
200	180	0.019	0.404	0.069	0.313	0.072	0.334
300	270	0.016	0.299	0.051	0.238	0.043	0.231
500	449	0.002	0.236	0.045	0.176	0.037	0.160
1000	899	0.006	0.163	0.024	0.131	0.020	0.123
1500	1350	0.010	0.131	0.021	0.103	0.018	0.093
2000	1799	0.002	0.116	0.010	0.088	0.009	0.078

Table 5.1: Biweight-kernel estimation results for the shape parameter $\gamma_1 = 0.6$ of Burr's model based on 1000 right-truncated samples

		$\widehat{\gamma}_{1,\mathbb{K}}$		$\widehat{\gamma}_1$		$\widehat{\gamma}_1^{GS}$	
N	n	abs bias	rmse	abs bias	rmse	abs bias	rmse
$p = 0.7$							
150	104	0.134	0.808	0.142	0.408	0.245	1.242
200	139	0.097	0.705	0.129	0.373	0.184	0.857
300	209	0.045	0.566	0.090	0.313	0.091	0.582
500	349	0.002	0.430	0.074	0.268	0.064	0.550
1000	699	0.003	0.399	0.031	0.237	0.023	0.161
1500	1050	0.010	0.362	0.013	0.217	0.010	0.130
2000	1401	0.010	0.244	0.018	0.164	0.009	0.117
$p = 0.8$							
150	119	0.096	0.730	0.109	0.397	0.117	0.729
200	159	0.060	0.580	0.091	0.340	0.108	0.874
300	239	0.037	0.496	0.067	0.315	0.080	0.490
500	399	0.009	0.303	0.057	0.231	0.047	0.280
1000	799	0.001	0.265	0.027	0.177	0.021	0.139
1500	1199	0.008	0.194	0.018	0.139	0.015	0.109
2000	1600	0.001	0.183	0.013	0.124	0.012	0.095
$p = 0.9$							
150	134	0.066	0.660	0.080	0.392	0.081	0.450
200	179	0.047	0.454	0.061	0.314	0.061	0.359
300	270	0.003	0.299	0.064	0.243	0.062	0.230
500	449	0.001	0.226	0.043	0.174	0.037	0.164
1000	899	0.009	0.175	0.016	0.124	0.014	0.113
1500	1350	0.002	0.146	0.017	0.108	0.017	0.098
2000	1799	0.003	0.134	0.010	0.093	0.008	0.081

Table 5.2: Triweight-kernel estimation results for the shape parameter $\gamma_1 = 0.6$ of Burr's model based on 1000 right-truncated samples

$p = 0.7$							
		$\widehat{\gamma}_{1,\mathbb{K}}$		$\widehat{\gamma}_1$		$\widehat{\gamma}_1^{GS}$	
N	n	abs bias	rmse	abs bias	rmse	abs bias	rmse
150	105	0.090	0.893	0.187	0.548	0.294	2.126
200	139	0.014	0.863	0.199	0.542	0.316	1.351
300	210	0.022	0.573	0.140	0.412	0.173	0.812
500	349	0.031	0.519	0.103	0.372	0.053	0.593
1000	699	0.004	0.462	0.042	0.324	0.020	0.253
1500	1049	0.017	0.356	0.031	0.255	0.020	0.174
2000	1399	0.008	0.424	0.017	0.267	0.017	0.150
$p = 0.8$							
150	120	0.088	0.862	0.122	0.553	0.248	1.947
200	159	0.040	0.684	0.121	0.472	0.178	1.143
300	239	0.006	0.516	0.084	0.406	0.099	0.494
500	399	0.022	0.372	0.078	0.285	0.058	0.247
1000	800	0.003	0.297	0.029	0.221	0.021	0.189
1500	1199	0.004	0.239	0.020	0.180	0.012	0.157
2000	1599	0.001	0.209	0.013	0.156	0.014	0.121
$p = 0.9$							
150	134	0.034	0.585	0.113	0.479	0.118	0.543
200	180	0.002	0.512	0.120	0.402	0.127	0.459
300	270	0.003	0.389	0.082	0.320	0.073	0.310
500	450	0.002	0.305	0.052	0.246	0.045	0.228
1000	900	0.004	0.223	0.024	0.169	0.020	0.153
1500	1349	0.005	0.176	0.020	0.141	0.021	0.124
2000	1800	0.006	0.166	0.013	0.126	0.013	0.110

Table 5.3: Biweight-kernel estimation results for the shape parameter $\gamma_1 = 0.8$ of Burr's model based on 1000 right-truncated samples

$p = 0.7$							
		$\widehat{\gamma}_{1,\mathbb{K}}$		$\widehat{\gamma}_1$		$\widehat{\gamma}_1^{GS}$	
N	n	abs bias	rmse	abs bias	rmse	abs bias	rmse
150	104	0.159	0.976	0.202	0.511	0.386	3.264
200	139	0.064	0.905	0.205	0.493	0.247	1.355
300	209	0.090	0.831	0.101	0.469	0.141	1.082
500	349	0.014	0.589	0.090	0.371	0.063	0.586
1000	700	0.013	0.458	0.049	0.296	0.023	0.264
1500	1050	0.008	0.561	0.023	0.315	0.020	0.189
2000	1400	0.012	0.381	0.027	0.241	0.013	0.164
$p = 0.8$							
150	120	0.103	0.886	0.151	0.511	0.180	1.906
200	160	0.058	0.775	0.131	0.466	0.153	1.311
300	239	0.023	0.629	0.106	0.398	0.078	0.502
500	399	0.005	0.515	0.069	0.339	0.060	0.256
1000	800	0.005	0.330	0.036	0.226	0.030	0.186
1500	1200	0.017	0.242	0.035	0.176	0.029	0.145
2000	1600	0.001	0.225	0.017	0.160	0.012	0.133
$p = 0.9$							
150	135	0.039	0.611	0.117	0.465	0.133	1.103
200	180	0.047	0.603	0.102	0.435	0.127	0.845
300	270	0.020	0.414	0.078	0.308	0.071	0.301
500	449	0.008	0.321	0.049	0.256	0.050	0.223
1000	900	0.011	0.230	0.024	0.173	0.020	0.153
1500	1350	0.008	0.197	0.016	0.137	0.015	0.120
2000	1800	0.001	0.162	0.014	0.115	0.011	0.105

Table 5.4: Triweight-kernel estimation results for the shape parameter $\gamma_1 = 0.8$ of Burr's model based on 1000 right-truncated samples

with

$$R_{n1} := \frac{1}{2\sqrt{k}} \int_1^\infty x^{-1} \mathbf{D}_n^2 (x) \, g'_{\mathbb{K}} (\xi_n (x)) \, dx.$$

where $\xi_n (x)$ is a stochastic intermediate value lying between $\overline{\mathbf{F}}_n (x X_{n-k:n}) / \overline{\mathbf{F}}_n (X_{n-k:n})$ and x^{-1/γ_1}. According to [Benchaira *et al.*(2016)], we have, for $0 < \epsilon < 1/2 - \gamma/\gamma_2$

$$\sup_{x \geq 1} x^{(1/2-\epsilon)/\gamma - 1/\gamma_2} \left| \mathbf{D}_n (x) - \mathbf{\Gamma} (x; \mathbf{W}) - x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \mathbf{A}_0 (n/k) \right| \overset{\mathbf{P}}{\to} 0, \text{ as } N \to \infty, \tag{5.5}$$

where $\{\Gamma (x; \mathbf{W}) ; \ x > 0\}$ is a Gaussian process defined in 4.6.

Now, we write

$$\sqrt{k} \left(\widehat{\gamma}_{1,K} - \gamma_1 \right) = \int_1^\infty x^{-1} \mathbf{\Gamma} (x; \mathbf{W}) \, g_{\mathbb{K}} \left(x^{-1/\gamma_1} \right) dx + \sum_{i=1}^{3} R_{ni},$$

where R_{n1} is as defined above and

$$R_{n2} := \int_1^\infty x^{-1} \left\{ \mathbf{D}_n (x) - \mathbf{\Gamma} (x; \mathbf{W}) - x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \mathbf{A}_0 (n/k) \right\} g_{\mathbb{K}} \left(x^{-1/\gamma_1} \right) dx,$$

and

$$R_{n3} := \int_1^\infty x^{-1} \left\{ x^{-1/\gamma_1} \frac{x^{\tau_1/\gamma_1} - 1}{\gamma_1 \tau_1} \sqrt{k} \mathbf{A}_0 (n/k) \right\} g_{\mathbb{K}} \left(x^{-1/\gamma_1} \right) dx.$$

Elementary calculation yields that

$$\int_1^\infty x^{-1} \mathbf{\Gamma} (x; \mathbf{W}) \, g_{\mathbb{K}} \left(x^{-1/\gamma_1} \right) dx = \left(\gamma^2 / \gamma_1 \right) \int_0^1 s^{-1} \mathbf{W} (s) \, d \{ s \varphi_{\mathbb{K}} (s) \} =: Z,$$

where $\varphi_{\mathbb{K}} (s)$ is that defined in the theorem. Next, we evaluate the remainder terms R_{ni}, for $i = 1, 2, 3$. First, we show that R_{n1} tends to zero in probability, as $N \to \infty$. Recall that $\gamma_1 < \gamma_2$ and $0 < \epsilon < 1/2 - \gamma/\gamma_2$, then $(1/2 - \epsilon) / \gamma - 1/\gamma_2 > 0$. It follows that $\int_1^\infty x^{2(1/\gamma_2 - (1/2 - \xi)/\gamma) - 1} dx$ is finite and, from Lemma 5.4.1, we get $\sup_{x \geq 1} \left| \mathbf{D}_n^2 (x) \right| = O_{\mathbf{p}} (1)$. On the other hand, from assumption $[\mathbb{C}4]$, we infer that $g'_{\mathbb{K}}$ is bounded on $(0, 1)$. Consequently, we have $R_{n1} = o_{\mathbf{p}} (1)$. Second, for the term R_{n2}, we use approximation

(5.5), to get

$$R_{n2}=o_{\mathbf{p}}\left(1\right)\int_{1}^{\infty}x^{1/\gamma_{2}-(1/2-\epsilon)/\gamma-1}\left|g_{\mathbb{K}}\left(x^{-1/\gamma_{1}}\right)\right|dx.$$

Since $g_{\mathbb{K}}$ is bounded on $(0,1)$, then $R_{n2}=o_{\mathbf{p}}(1)$. Finally, we show that the third term R_{n3} is equal to $\sqrt{k}\mathbf{A}_0(n/k)\int_0^1 s^{-\tau_1}\mathbb{K}(s)\,ds$. Observe that

$$R_{n3}=\sqrt{k}\mathbf{A}_{0}\left(n/k\right)\int_{1}^{\infty}x^{-1/\gamma_{1}-1}\frac{x^{\tau_{1}/\gamma_{1}}-1}{\gamma_{1}\tau_{1}}g_{\mathbb{K}}\left(x^{-1/\gamma_{1}}\right)dx.$$

Keeping in mind that $g_{\mathbb{K}}(s)=(s\mathbb{K}(s))'$, we end up, after a change of variables and an integration by parts, with

$$\int_{1}^{\infty}x^{-1/\gamma_{1}-1}\frac{x^{\tau_{1}/\gamma_{1}}-1}{\gamma_{1}\tau_{1}}g_{\mathbb{K}}\left(x^{-1/\gamma_{1}}\right)dx=\int_{0}^{1}s^{-\tau_{1}}\mathbb{K}\left(s\right)ds.$$

Gathering all the results above leads to the first part of the theorem. For the second part, it suffices to use Lemma 8 in [Csörgö et al.(1985)], to show that the variance of the centred Gaussian rv Z equals $\sigma^2_{\mathbb{K}}$. Finally, whenever (given $n=m$) $\sqrt{k_m}\mathbf{A_0}(m/k_m)\to\lambda$, we have

$$R_{n3}\overset{\mathbf{P}}{\to}\lambda\int_{0}^{1}s^{-\tau_{1}}\mathbb{K}\left(s\right)ds,\ \text{as } N\to\infty,$$

which corresponds to the asymptotic bias $\mu_{\mathbb{K}}$, as sought.

Lemma 5.4.1 *Under the assumptions of Theorem 5.2.1, we have, for any* $0<\epsilon<1/2-\gamma/\gamma_2$

$$\sup_{x\geq1}x^{(1/2-\epsilon)/\gamma-1/\gamma_{2}}\left|\mathbf{D}_{n}\left(x\right)\right|=O_{\mathbf{p}}\left(1\right),\ \text{as } N\to\infty.$$

Proof. This result is straightforward from the weak approximation (5.5). Indeed, it is clear that $\sup_{x\geq1}x^{(1/2-\epsilon)/\gamma-1/\gamma_2}\left|\mathbf{D}_n(x)\right|\leq T_{1,n}+T_{2,n}+T_3$, where

$$T_{1,n}:=\sup_{x\geq1}x^{(1/2-\epsilon)/\gamma-1/\gamma_{2}}\left|\mathbf{D}_{n}\left(x\right)-\mathbf{\Gamma}\left(x;\mathbf{W}\right)-x^{-1/\gamma_{1}}\frac{x^{\tau_{1}/\gamma_{1}}-1}{\gamma_{1}\tau_{1}}\sqrt{k}\mathbf{A}_{0}\left(n/k\right)\right|,$$

$$T_{2,n}:=\frac{\sqrt{k}\left|\mathbf{A}_{0}\left(n/k\right)\right|}{\gamma_{1}\left|\tau_{1}\right|}\sup_{x\geq1}\left\{x^{-(1/2+\epsilon)/\gamma}\left|1-x^{\tau_{1}/\gamma_{1}}\right|\right\}\text{ and }T_{3}:=\sup_{x\geq1}x^{(1/2-\epsilon)/\gamma-1/\gamma_{2}}\left|\mathbf{\Gamma}\left(x;\mathbf{W}\right)\right|.$$

First, it is readily checked, from (5.5), that $T_{1,n} = o_{\mathbf{p}}(1)$. Second, observe that, in addition to the assumption $\sqrt{k}\mathbf{A}_0(n/k) = O_{\mathbf{p}}(1)$, we have $0 \leq x^{-(1/2+\epsilon)/\gamma}\left(1 - x^{\tau_1/\gamma_1}\right) \leq 1$, for $x \geq 1$, it follows that $T_{2,n} = O_{\mathbf{p}}(1)$. Finally, note that $x^{(1/2-\epsilon)/\gamma - 1/\gamma_2}\mathbf{\Gamma}(x; \mathbf{W})$ is equal to

$$x^{-(1/2+\epsilon)/\gamma}\left\{\frac{\gamma}{\gamma_1}\left(x^{1/\gamma}\mathbf{W}\left(x^{-1/\gamma}\right) - \mathbf{W}(1)\right)\right.$$
$$\left. + \frac{\gamma}{\gamma_1+\gamma_2}\int_0^1 s^{-\gamma/\gamma_2-1}\left(x^{1/\gamma}\mathbf{W}\left(x^{-1/\gamma}s\right) - \mathbf{W}(s)\right)ds\right\},$$

where the quantity between brackets is a Gaussian rv and $x^{-(1/2+\epsilon)/\gamma} \leq 1$, for $x \geq 1$. Therefore, $T_3 = O_{\mathbf{p}}(1)$ and the proof is completed. ■

.

Concluding notes

As we have seen throughout this thesis, we proposed an estimator of the tail index for randomly truncated heavy-tailed data based on the same number of extreme observations from both truncated and truncation variables. Thus, the determination of the optimal sample fraction becomes standard, in the sense of applying any convenient algorithm available in the literature. The asymptotic normality of the estimator is established by taking into account the dependence structure of the observations and a practical way to construct confidence bounds for the extreme value index is given. The obtained Gaussian approximations are of great usefulness as they allow to determine the limiting distributions of several statistics related to the extreme value index such that high quantiles and risk measures estimators (see, for instance, [Necir and Meraghni(2009)]). As an application, we can provided an estimator for the excess-of-loss reinsurance premium in the case of large randomly truncated claims.

In chapter 4, We introduced a product-limit process for the tail of a Pareto-like distribution under random right-truncation. The weak approximation of this process proved to be a very useful tool in establishing the asymptotic normality of the estimators of tail indices and related statistics such as high quantiles. Moreover, we proposed a natural Hill-type estimator for the extreme value index, that behaves well in the case of small datasets. An interesting point, which is beyond the scope of the present thesis and deserves to be considered in a future work, is to reduce estimation biases under random truncation. Similar anterior works were done with complete datasets by, for instance, [Peng and Qi(2004)], [Li et al.(2010)] and [Brahimi *et al.*(2013)].

We finish this work by making a comment on relation (4.7), which actually is a special case of a more general functional of the distribution tail defined by

$$\Gamma_t(g,\alpha) := \frac{\dfrac{1}{\overline{F}(t)}\int_t^\infty g\left(\dfrac{\overline{F}(x)}{\overline{F}(t-)}\right)\left(\log\dfrac{x}{t}\right)^\alpha dF(x)}{\int_0^1 g(x)(-\log x)^\alpha dx},\ t \geq 0,$$

where g is some weight function and α some positive real number. As a consequence of the fact that $\lim_{t\to\infty}\Gamma_t(g,\alpha) = \gamma^\alpha$, this functional can be considered as the starting point to constructing a whole class of estimators for distribution tail parameters. Indeed, in the complete data case, we replace F by its empirical counterpart F_n and t by $X_{n-k:n}$ to get the following statistic which generalizes several extreme value theory based procedures of estimation already existing in the literature:

$$\Gamma_{n,k}(g,\alpha) := \frac{\dfrac{1}{k}\sum_{i=1}^k g\left(\dfrac{i}{k+1}\right)\left(\log\dfrac{X_{n-i+1:n}}{X_{n-k:n}}\right)^\alpha}{\int_0^1 g(x)(-\log x)^\alpha dx}.$$

When $g = \alpha = 1$, we recover the famous Hill estimator [Hill(1975)]. For a detailed list of extreme value index estimators drawn from the statistic above, we refer to the paper of [Ciuperca and Mercadier(2010)], where the authors propose an estimation approach of the second-order parameter by considering differences and quotients of several forms of $\Gamma_{n,k}(g,\alpha)$. By analogy, when we deal with randomly truncated observations, we substitute the product-limit estimator $\mathbf{F}_n$ for F in the formula of $\Gamma_t(g,\alpha)$ in order to obtain the following family of parameter estimators under random truncation:

$$\Gamma_{n,k}(g,\alpha) := \frac{\sum_{i=1}^k a_n^{(i)} g\left(\dfrac{\overline{\mathbf{F}}_n(X_{n-i+1:n})}{\overline{\mathbf{F}}_n(X_{n-k-1:n})}\right)\left(\log\dfrac{X_{n-i+1:n}}{X_{n-k:n}}\right)^\alpha}{\sum_{i=1}^k a_n^{(i)}\int_0^1 g(x)(-\log x)^\alpha dx},$$

where $a_n^{(i)} := \mathbf{F}_n(X_{n-i+1:n})/C_n(X_{n-i+1:n})$. This would have fruitful consequences on the statistical analysis of extremes under random truncation.

Annexe A: Abbreviations and Notations

Abbreviations and Notations that is largely confined to sections or chapters is mostly excluded from the list below:

Abbreviation	**Signification**
a.s.	almost sure
CLT	central limit theorem
df	distribution function
e.g.	for example
EVI	extreme value index
EVT	extreme value theory
df	distribution function
cdf	cumulative distribution function
GEVD	generalized extreme value distribution
i.e.	in other words
iff	if and only if
iid	independent and identically distributed
MSE	mean squared error
RMSE	Root mean squared error
rv	random variable

Notations	**Signification**
$Cov(X,Y)$	covariance between X and Y
$\xrightarrow{\mathcal{D}}$	convergence in distribution
$\xrightarrow{p}$	convergence in probability
$a(t) \sim b(t)$	$\lim_t a(t)/b(t) = 1$
$\mathcal{D}(.)$	domain of attraction
$\stackrel{d}{=}$	equality in distribution
γ	extreme value index
F	distribution function
F_n	empirical distribution function
$1_{\{p\}}$	indicator function: equals 1 if p is true and 0 otherwise
$\inf A$	infimum of set A
l	dependence function
a_+	$max(a,0)$
a_-	$min(a,0)$
$a \vee b$	$max(a,b)$
$a \wedge b$	$min(a,b)$
n	integer number greater than 1
$\mathbb{N}$	set of non-negative integers
$\mathcal{N}(\mu,\sigma^2)$	normal or Gaussian distribution with mean μ and variance σ^2
$o(.)$	$f(x) = o(g(x))$ as $x \to x_0 : f(x)/g(x) \to 0$, 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 $O_p(.)$	stochastic order symbols
$(\Omega, \mathcal{F}, P)$	probability space
$\mathbb{R}$	set of real numbers
$\sup A$	supremum of set A
X	rv defined on $(\Omega, \mathcal{F}, P)$, population
$(X_1, ..., X_n)$	sample of size n from X
$(X_{1,n}, ..., X_{n,n})$	order statistics pertaining to $(X_1, ..., X_n)$

Bibliography

- Balakrishnan, N., Cohen, A. C., 1991. Order Statistics and Inference. Estimation Methods, Academic Press, San Diego.

- Beirlant, J., Dierckx, G., Fils-Villetard, A., Guillou, A., 2007. Estimation of the extreme value index and extreme quantiles under random censoring. Extremes 10, 151–174.

- 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.

- Benchaira, S., Meraghni, D., Necir, A., 2016. Tail product-limit process for truncated data with application to extreme value index estimation. Extremes. 10.1007/s10687-016-0241-9.

- Brahimi, B., Meraghni, D., Necir, A., Yahia, D., 2013. A bias-reduced estimator for the mean of a heavy-tailed distribution with an infinite second moment. J. Statist. Plann. Inference 143, no. 6: 1064-1081.

- Ciuperca, G., Mercadier, C., 2010. Semi-parametric estimation for heavy tailed distributions. Extremes 13: 55-87.

- Cohen, A. C., 1991. Truncated and Censored Samples. Theory and Applications, Marcel Dekker, New York.

- Csörgö, S., Deheuvels, P., Mason, D. ,1985. Kernel estimates of the tail index of a distribution. Ann. Statist 13, 1050–1077.

- Csörgo, S., Mason, D.M., 1985. Central Limit Theorems for Sums of Extreme Values. Mathematical Proceedings of the Cambridge Philosophical Society 98, 547-558.
- Csörgő, M., Révész, P., 1981. Strong approximations in probability and statistics. Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London.
- Davis, R., Resnick, S.I., 1984. Tail Estimates Motivated by Extreme Value Theory. Annals of Statistics 12, 1467-1487.
- Deheuvels, P., Häusler, E., Mason, D.M., 1988. Almost Sure Convergence of the Hill Estimator. Mathematical Proceedings of the Cambridge Philosophical Society 104, 371-381.
- Dekkers, A.L.M., de Haan, L., 1989. On the estimation of the extreme value index and large quantile estimation.Ann. Statist.17, 1795–1832.
- Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L., 1989. A Moment Estimator for the Index of an Extreme Value Index. Annals of Statistics 17, 1833-1855.
- Drees, H., 1996. Refined Pickands Estimator of the Extreme Value Index. Annals of Statistics 23, 2059-2080.
- Drees, H., de Haan, L., Li, D., 2006. Approximations to the tail empirical distribution function with application to testing extreme value conditions. J. Statist. Plann. Inference 136: 3498-3538.
- Drees, H., Huang, X., 1998. Best attainable rates of convergence for estimators of the stable tail dependence function. J. Multivariate Anal. 64, 25-47.
- Efron, B., Petrosian, V., 1999. Noparametric methods for doubly truncated data.Journal of the American Statistical Association, 94, 824–834.
- Einmahl, J.H.J., 1992. Limit theorems for tail processes with application to intermediate quantile estimation. J. Statist. Plann. Inference 32: 137-145.

- Einmahl, J.H.J., de Haan, L., Li, D., 2006. Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. Ann. Statist. 34: 1987-2014.

- Einmahl, J., Fils-Villetard, A., Guillou, A., 2008. Statistics of extremes under random censoring. Bernoulli 14, 207–227.

- Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

- Escudero, F., Ortega, E., 2008. Actuarial comparisons for aggregate claims with randomly right-truncated claims. Insurance Math. Econom. 43: 255-262.

- Gardes, L., Stupfler, G., 2015. Estimating extreme quantiles under random truncation. TEST 24: 207-227.

- Gomes, M.I., Neves, M.M., 2011. Estimation of the extreme value index for randomly censored data. Biometrical Lett. 48 (1), 1-22.

- Gomes M.I., Oliveira O., 2003. Censoring estimators of a positive tail index. Statist. Probab. Lett. 65 (3), 147-159.

- Gomes, M.I., Pestana, D., 2007. A simple second-order reduced bias' tail index estimator. J. Stat. Comput. Simul. 5-6: 487-504.

- Groeneboom, P., Lopuhaä, H. P., de Wolf, P. P., 2003. Kernel-type estimators for the extreme value index. *Ann. Statist.* 31, 1956-1995.

- Gudendorf, G., Segers, J., 2010. Extreme-value copulas. Copula theory and its applications, 127–145, Lect. Notes Stat. Proc., 198, Springer, Heidelberg.

- Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press, New York.

- de Haan, L., Stadtmüller, U., 1996. Generalized regular variation of second order. J. Australian Math. Soc. (Series A) 61: 381-395.

- de Haan, L., Neves, C., Peng, L., 2008. Parametric tail copula estimation and model testing. J. Multivariate Anal. 99, 1260-1275.
- de Haan, L., Ferreira, A., 2006. Extreme Value Theory: An Introduction. Springer.
- Hall, P.,1982. On Some Simple estimates of an Exponent of Regular Variation. Journal of the Royal Statistical Society 44, 37-42.
- Hall, P., 1982. On estimating the endpoint of a distribution. Ann. Statist. 10, 556–568.
- Herbst, T., 1999. An application of randomly truncated data models in reserving IBNR claims. Insurance Math. Econom. 25: 123-131.
- Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. Ann. Statist. 3: 1163-1174.
- Hua, L., Harry, J., 2011. Second order regular variation and conditional tail expectation of multiple risks. Insurance Math. Econom. 49: 537-546.
- Huang, X.,1992. Statistics of Bivariate Extreme Values. Tinbergen Institute Research Series 22.
- Hüsler, J., Li D. and Müller, S., 2006. Weighted least squares estimation of the extreme value index. *Statist. Probab. Lett.* 76, 920–930
- Jenkinson, A. F., 1955. The frequency distribution of the annual maximum (or minimum) values of metereological elements. Q. J. Roy. Meteor. Soc. 81, 158-171.
- Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-481.
- Klein, J.P. and Moeshberger, M.L., 2003. Survival Analysis. Techniques for Censored and Truncated Data. Springer-Verlag, New York.
- Koning, A.J., Peng, L., 2008. Goodness-of-fit tests for a heavy tailed distribution. J. Statist. Plann. Inference 138, no. 12: 3960-3981.

- Lagakos, S.W., Barraj, L.M., de Gruttola, V., 1988. Nonparametric Analysis of Truncated Survival Data, with Applications to AIDS. Biometrika, 75, 515-523.
- Lawless, J.F., 2002. Statistical Models and Methods for Lifetime Data, Second Edition. Wiley Series in Probability and Statistics.
- Leadbetter, M.R., Lindgren, G., Rootzèn, H., 1983. Extremes and Related Properties of Random Sequences and Processes. Springer, Berlin.
- Li, D., Peng, L., Yang, J., 2010. Bias reduction for high quantiles. J. Statist. Plann. Inference 140, no. 9: 2433-2441.
- 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.
- Mason, D.M., 1982. Laws of large numbers for sums of extreme values. Ann. Probab. 10: 756-764.
- Meeker, W. Q., Escobar, L. A., 1998. Statistical Methods for Reliability Data, New York. Wiley.
- Meraghni, D., 2008. Modeling Distribution Tails. Doctorat thesis, Med Khider University, Biskra-Algeria.
- Necir, A., Meraghni, D., 2009. Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. Insurance Math. Econom. 45: 49-58.
- Neves, C., Fraga Alves, M.I., 2004. Reiss and Thomas' Automatic Selection of the Number of Extremes. Computational Statistics and Data Analysis. 47, 689-704.
- Peng, L., 2010. A practical way for estimating tail dependence functions. Statist. Sinica 20: 365-378.
- Peng, L., Qi, Y., 2004. Estimating the first- and second-order parameters of a heavy-tailed distribution. Aust. N. Z. J. Stat. 46, no. 2: 305-312.

- Pickands, J., 1975. Statistical Inference Using Extreme Order Statistics. Annals of Statistics 3, 119-131.
- Reiss, R.D., Thomas, M., 1997. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhauser.
- Reiss, R.D., Thomas, M., 2007. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, 3rd ed. Birkhäuser Verlag, Basel, Boston, Berlin.
- Rényi, A., 1953. On the theory of order statistics. Acta Mathematica Scient. Hungar. Tomus IV, 191-227.
- Resnick, S.I., 1987. Extreme Values, Regular Variation and Point Processes. Springer, New York.
- Resnick, S., 2006. Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer.
- Saporta, G., 1990. Probabilités, Analyse des Données et Statistique. Editions Technip, Paris.
- Segers, J., 2005. Generalized Pickands estimators for the extreme value index. J. Statist. Plann. Inference 128, 381–396.
- Shorack, G.R., Wellner, J. A., 1986. Empirical processes with applications to statistics. New York: John Wiley & Sons.
- Smith, R.L., 1985. Maximum likelihood estimation in a class of nonregular cases. Biometrika 72, 67–90.
- Smith, R.L., 1987. Estimating tails of probability distributions. Ann. Statist. 15, 1174–1207.

- Smith, R.L., Weissman, I., 1985. Maximum likelihood estimation of the lower tail of a probability distribution. J. Roy. Statist. Soc. Ser. B47, 285–298.
- Sousa, B.C., 2002. A Contribution to the Estimation of the Tail Index of Heavy-Tailed Distributions. PhD. Thesis, University of Michigan.
- Strzalkowska-Kominiak, E., Stute, W., 2009. Martingale representations of the Lynden-Bell estimator with applications. Statist. Probab. Lett. 79: 814-820.
- Stute, W., 1995. The central limit theorem under random censorship. Ann. Statist. 23, 422-439.
- Stute, W., Wang, J., 2008. The central limit theorem under random truncation. Bernoulli 14: 604-622.
- von Mises, R., 1936. La distribution de la plus grande de n valeurs. Reprinted in Selected Papers Volumen II, Amer. Math. Soc., Providence, R.I., 271-294.
- Wang M.C., Jewell N.P., Tsai, W.Y., 1986. Asymptotic properties of the product-limit estimate under random truncation. Ann Statist 14: 1597-1605.
- Wang, M.C., 1989. A semiparametric model for randomly truncated data. Statist. Assoc., 84: 742-748
- Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. J. Am. Statist. Assoc. 73: 812-815.
- Woodroofe, M., 1985. Estimating a distribution function with truncated data. Ann. Statist. 13: 163-177.
- Worms, J., Worms, R., 2014. New estimators of the extreme value index under random right censoring, for heavy-tailed distributions. Extremes 17, 337-358.
- Worms, J., Worms, R., 2016. A Lynden-Bell integral estimator for extremes of randomly truncated data. Statist. Probab. Lett. 109, 106-117.

Abstract

In this thesis, we are concerned with the estimation of the extreme value index and large quantiles for incompletely observed data, with a particular interest in the case of right-truncated data. We begin by exploiting the first work in this matter, which is due to [Gardes and Stupfler(2015)], to derive a simple tail index estimator based on a single sample fraction of extreme values. The asymptotic normality of the proposed estimator is established in the frameworks of tail dependence and second-order of regular variation. Second, starting from the first-order condition of regular variation, we construct a new estimator for the shape parameter of a right-truncated heavy-tailed distribution. We prove its asymptotic normality by making use of the tail Lynden-Bell process for which a weighted Gaussian approximation is provided. Also, a new approach of estimating high quantiles is proposed and applied to a real dataset consisting in lifetimes of automobile brake pads. Finally, a kernel-type asymptotically normal estimator is defined. Simulation experiments are carried out to evaluate the performances and illustrate the finite sample behaviors of the above estimators and make comparisons as well.

Keywords: Bivariate extremes; Empirical process; Extreme value index; Heavy-tails; High quantiles; Hill estimator; Kernel estimation; Lynden-Bell estimator; Regular variation; Random truncation; Tail dependence.

Résumé

Dans cette thèse, nous nous intéressons à l'estimation de l'indice des valeurs extrêmes et des quantiles extrêmes pour des données incomplètement observées, avec un intérêt particulier au cas des données tronquées à droite. Nous commençons par l'exploitation du premier travail sur ce sujet, qui est dù à Gardes et Stupfler (2015), pour obtenir un estimateur simple d'indice de queue basé sur une seule fraction d'échantillon de valeurs extrêmes. La normalité asymptotique de l'estimateur proposé est établie dans le cadre de la dépendance de queue et de la condition du second ordre de variation régulière. Deuxièmement, à partir de la condition du premier ordre de variation régulière, nous construisons un nouvel estimateur pour le paramètre de forme d'une distribution à queue lourde tronquée à droite. Nous prouvons sa normalité asymptotique en utilisant le processus de queue de Lynden-Bell pour lequel une approximation gaussienne pondérée est fournie. En outre, une nouvelle approche de l'estimation des quantiles extrêmes est proposée et appliquée sur des données réelles consistant en les durées de vie des plaquettes de frein automobile. Enfin, un estimateur de type noyau asymptotiquement normal est défini. Des expériences de simulation sont effectuées pour évaluer les performances et illustrer les comportements des estimateurs ci-dessus sur des échantillons finis et aussi pour faire des comparaisons.

Mots clés: Extrêmes bivariées; Processus empirique; Indice des valeurs extrêmes; Queues lourdes; Quantiles extrêmes; Estimateur de Hill; Estimation à noyau; Estimateur de Lynden-Bell; Variation régulière; Troncature aléatoire.

ملخص

في هذه الأطروحة، نحن نهتم بتقدير مؤشر القيم القصوى الخاص ببيانات غير كاملة، مع اهتمام خاص بحالة بيانات مقطوعة من اليمين. نبدأ من خلال استغلال أول عمل في هذا الشأن لغاردس و ستوفلر(2015)، والذي يرجع إلى اشتقاق مقدر مؤشر ذيل بسيط يستند إلى عينة واحدة من القيم المتطرفة. الحالة السوية المقاربة للمقدر المقترح تأسست في أطر العمل بذيل الترابط و الدرجة الثانية من الاختلاف المنتظم. ثانيا، ننطلق من أول شرط للاختلاف المنتظم و نستخرج مقدر جديد خاص بمؤشر القيم القصوى لتوزيعات ذات أذيال ثقيلة. نثبت الحالة السوية له من خلال استعمال نهج ذيل ليندن-بيل الذي تم توفير تقريب غوص له. أيضا، نقترح نهجا جديدا لتقدير قيم عالية تقسم البيانات الى مجالات تحتوي على نفس العدد من البيانات, ونطبقها على مجموعة بيانات حقيقية تتمثل في أعمار فرامل السيارات. وأخيرا، نعرف مقدر مؤشر ذيل بسيط من نوع نواة. وتجرى تجارب المحاكاة لتقييم الأداء وتوضيح السلوكيات لعينة محدودة من المقدرات المذكورة أعلاه، وكذلك إجراء مقارنات بينها.

الكلمات المفتاحية: قصوى ذات متغيرين ; عملية تجريبية ; مؤشر قيمة قصوى ; ذيول ثقيلة ; أجزاء عالية ; تقدير هيل ; تقدير النواة ;تقدير ليندن-بيل ; اختلاف منتظم ; اقتطاع عشوائي.