Federal State Autonomous Educational Institution for the Higher Education
NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS
Faculty of Economic Sciences

BACHELOR THESIS

Extreme Value Analysis for Mixture Models With Heavy-Tailed Impurity

Field of Study 38.03.01 Economics
Bachelor's Programme «Economics and Statistics»

Written by:
Student of the BST171
Ekaterina A. Morozova

Supervisor:
Associate Professor, Ph.D.
Vladimir A. Panov

Moscow, 2021

Abstract

This research is devoted to the extreme value analysis for mixture models with parameters depending on the number of available observations. Such models can be viewed as triangular arrays, for which there exist no general results concerning the asymptotic behaviour of maxima. The current study considers the case when one of the components has a heavy-tailed distribution and corresponds to the small mixing parameter, so that it can be associated with the «heavy-tailed impurity», which «pollutes» another component. We analyse two ways of modelling this «impurity», namely, by the non-truncated regularly varying law and its upper-truncated version with an increasing truncation level. The set of possible limit distributions for maxima turns out to be much more diverse than in the classical setting, especially for a mixture with the truncated component, where it includes four discontinuous laws not present in the Fisher-Tippett-Gnedenko theorem. For practical purposes we describe the procedure of application of the considered model to the analysis of price variations. Using the data on the returns of BMW shares in 2019, we demonstrate that the proposed model provides a reasonable fit for such quantities, becoming an efficient tool for the analysis of their maximal values.

Аннотация

Данное исследование посвящено анализу экстремальных значений в моделях смеси с параметрами, зависящими от количества наблюдений. Такие модели могут рассматриваться как схемы серий, для которых не существует общих результатов относительно асимптотического поведения максимумов. В настоящей работе изучается случай, когда одна из компонент имеет тяжелохвостное распределение и соответствует маленькому параметру смеси, «загрязняя» вторую компоненту. Мы рассматриваем два способа моделирования такого «загрязнения» — с помощью регулярно меняющегося или усечённого регулярно меняющегося закона с возрастающим уровнем урезания — и показываем, что множество возможных предельных распределений максимумов оказывается намного более разнообразным, чем в классической теории. В частности, в случае модели с усечённой компонентой оно включает в себя четыре разрывных закона, не присутствующих в теореме Фишера-Типпетта-Гнеденко. В практических целях мы описываем процедуру применения рассматриваемой модели к анализу ценовых колебаний. На примере доходностей BMW в 2019 году мы показываем, что предложенная модель хорошо описывает такие данные, становясь ценным инструментом для анализа максимальных значений доходностей финансовых активов.

Contents

Introduction

Extreme Value Analysis (EVA) is a special theory which concentrates on the extreme events, commonly referred to as anomalies. It has received a great deal of attention from both scientists and practitioners as a valuable and economic tool for studying such phenomena, and is widely employed in a variety of areas, ranging from engineering to economics and finance. To name a few, the applications of EVA include predicting the maximal possible claim size in insurance, determining the bond repayment values, and calculating the return periods and levels (Embrechts et al., 1997).

The current paper provides the extreme value analysis for mixture models of the form

$$F(x;\varepsilon,\vec{\theta}) = (1-\varepsilon)F^{(1)}(x;\vec{\theta}^{(1)}) + \varepsilon F^{(2)}(x,\vec{\theta}^{(2)}), \tag{1}$$

where $\varepsilon \in (0,1)$ is a mixture parameter, $F^{(1)}(x;\vec{\theta}^{(1)})$ and $F^{(2)}(x;\vec{\theta}^{(2)})$ are cumulative distribution functions (CDFs) of two distributions parametrised by vectors $\vec{\theta}^{(1)}, \vec{\theta}^{(2)}$, respectively, and $\vec{\theta} = (\vec{\theta}^{(1)}, \vec{\theta}^{(2)})$. Namely, we consider the case when $F^{(1)}(x;\vec{\theta}^{(1)})$ is a light-tailed Weibull law, while $F^{(2)}(x;\vec{\theta}^{(2)})$ is a heavy-tailed regularly varying distribution, which is either non-truncated or has a truncation level increasing in time. Such choice of distributions is mostly motivated by practical considerations and, particularly, the empirical facts regarding financial data, which will be discussed in the next chapter. The mixing parameter ε is assumed to be very small and moreover, tending to zero as the number of observations grows. In this setting, the second component models the occurrence of unusually large and rare events, and can be referred to as the heavy-tailed impurity (Grabchak, Molchanov, 2015).

Dependency of distribution parameters on the number of available observations allows to generalise the model (1) to the so-called row-wise independent triangular array, i.e., a collection of real random variables $\{X_{nj}, j = 1..k_n\}$, $k_n \to \infty$ as $n \to \infty$, such that $X_{n1}, ..., X_{nk_n}$ are independent for each n. While the classical limit theorems for such models are well-known (see, e.g., Petrov, 2012; Meerschaert and Scheffler, 2001), there exist no general statements concerning the asymptotic behaviour of their maxima, that is, the limits of

$$(\max_{j=1..k_n} X_{nj} - c_n)/s_n, \qquad n \to \infty, \tag{2}$$

with deterministic c_n, s_n. At the same time, such results appear to be valuable from both theoretical and practical perspectives, since softening the assumption of parameter invariability makes the model (1) much more flexible, and in fact, leads to a wider range of possible limits in (2) than in the classical setting of independent identically distributed (i.i.d.) sequences.

This determines the *novelty* and *professional significance* of the current research. Its *aim* is to provide the extreme value analysis for the model (1), and the main objectives are to

1. illustrate the difference between EVA for sequences of i.i.d. random variables and triangular arrays,

2. establish the limit distributions for maxima in the model (1) and formulate the conditions under which they are obtained,

3. analyse the effect of different types of truncation on the asymptotic behaviour of maxima,

4. apply the model (1) to the description of financial returns.

The *object* of the present study are mixture models with heavy-tailed impurity, the *subject* is the asymptotic behaviour of maxima in them. The scope of this research will be on deriving limit distributions for maxima in the aforementioned setting and demonstrating the applicability of the considered model to the analysis of price variations. For the latter purpose, we use the data on prices of BMW shares in 2019, retrieved from Finam[1]. All computations are performed in the `R` language.

This paper is organised as follows. In Chapter 1 we briefly recall the existing results for EVA for i.i.d. sequences and triangular arrays, as well as the critique against conventional methods of describing financial data. The same chapter contains discussion about different kinds of truncation and deciding

[1]Finam. URL: https://www.finam.ru. Retrieved on 05.11.20

between truncated and non-truncated regularly varying (RV) laws in practical applications. Chapter 2 is devoted to the derivation of limit distributions for maxima in the model (1). Our main results are formulated as Theorem 1 and Theorem 2. For both Weibull-RV and Weibull-truncated RV mixtures we get that the set of limits in (2) is much more diverse than in the classical setting, and in the latter case it includes 6 various laws, 4 of which are discontinuous and not present in the classical theory. The asymptotic behaviour of maxima in this model depends on the rate of growth of the truncation level, and as we show, the resulting conditions are related to the truncation regimes introduced by Chakrabarty and Samorodnisky (2012) and discussed in Section 1.4 of the current paper. Finally, Chapter 3 is devoted to the practical applications of the obtained results. In Section 3.1 our theoretical findings are exemplified by the simulation study. In Section 3.2 we develop a four-step procedure for the application of the model (1) to the asset price modelling and show that it can be successfully employed for the description of such data, becoming an efficient tool for the analysis of maximal values of financial returns. The final part of this paper summarises the results of our study and outlines the directions of further research.

Chapter 1. Background

1.1. Classical EVA

As was mentioned in the introduction, the classical EVA deals with the asymptotic behaviour of maxima of i.i.d. sequences. In this setting it has extensive theoretical results, the most fundamental of which was established by Fisher and Tippett (1928) and Gnedenko (1943), who have characterised the set of possible limits under linear transformation. More precisely, the Fisher-Tippett-Gnedenko theorem states that if for a sequence of i.i.d. random variables $X_1, \dots, X_n$, $n \geq 2$, there exists a deterministic normalising sequence $v_n(x) = s_n x + c_n$ with $s_n > 0$ and $c_n \in \mathbb{R}$ such that the limit distribution H in

$$\lim_{n\to\infty} \mathbb{P}\left\{ \max_{i=1,\dots,n} X_i \leq v_n(x) \right\} = H(x), \quad x \in \mathbb{R}, \tag{3}$$

is non-degenerate, then H is necessarily of one of the following forms

$$H(x) = \begin{cases} e^{-x^{-\alpha}} \cdot \mathbb{I}_{\{x>0\}} & \text{(Fréchet distribution)} \\ e^{-(-x)^{\alpha}} \cdot \mathbb{I}_{\{x\leq 0\}} + \mathbb{I}_{\{x>0\}} & \text{(Weibull distribution),} \\ e^{-e^{-x}} & \text{(Gumbel distribution)} \end{cases}$$

where $\alpha > 0$. The distribution of X is then said to belong to the maximum domain of attraction (MDA) of H. Due to the convergence to types theorem such limit, if exists, is unique, and the conditions for convergence to either of these distributions are well-known and can be found in any monograph on EVA: see, e.g., the book by Embrechts et al. (1997). Some extension to this result is due to Pantcheva (1985) and Mohan and Ravi (1993), who have obtained the results of the same kind for power normalisation, i.e., for $v_n(x) = a_n |x|^{b_n} \operatorname{sign}(x)$ with $a_n, b_n > 0$. All in all, the extreme value theory for i.i.d. sequences is well-established, and once the distribution is known, one can determine everything about the asymptotic behaviour of the corresponding maxima.

1.2. EVA for triangular arrays

It should be mentioned, however, that the distribution parameters may vary with time. For instance, in migration processes of population dynamics one may expect the proportion of «very active» species to decline as the overall number of species grows (Grabchak, Molchanov, 2015). In this case the model can be considered as a triangular array, for which the extreme value analysis is much more intricate. Unlike the case of i.i.d. sequences, there are no general characterisations of limit distributions for properly normalised maxima of triangular arrays, and the papers on this issue typically deal with some particular examples.

For instance, the paper by Anderson et al. (1997), which was one of the pioneering works in this field, focuses on the case when each random variable X_{nj} of a row-wise independent triangular array $\{X_{nj}, j = 1..n\}$, $n \to \infty$, can be represented as

$$X_{nj} \overset{\mathcal{L}}{=} \frac{\sum_{r=1}^{m_n} Y_r - m_n \mathbb{E}[Y_1]}{\sqrt{m_n \operatorname{Var} Y_1}}, \tag{4}$$

where $Y_1, \ldots, Y_{m_n}$ are i.i.d. random variables and m_n is some sequence of integers. The authors show that under certain additional conditions on the moment generating function of $Y_1, \ldots, Y_{m_n}$ and the moments themselves, as well as the asymptotic behaviour of their cdf and m_n, the properly normalised maxima converge to the Gumbel, Fréchet or nonextreme value distibution. It is worth mentioning that such a setting allows to obtain a non-degenerate limit for maxima of Poisson random variables, which is impossible in case of i.i.d. sequences due to the behaviour of cdf near the right endpoint. This work was generalised in the subsequent paper by Dkengne et al. (2016), who obtain the Gumbel limit for maxima of a row-wise stationary triangular array and do not require the sum-representation (4). The most general result is probably due to the paper by Freitas and Hüsler (2003), which deals with triangular arrays of row-wise independent random variables having twice differentiable cdfs F_n and provides conditions for the convergence of the form

$$\lim_{n\to\infty} (F_n^n(s_n x + c_n))^{(m)} = G^{(m)}(x), \quad s_n > 0, c_n \in \mathbb{R}, \quad m = 0, 1, 2,$$

where G can be any twice differentiable distribution function.

Nevertheless, it should be noted that the results presented above are quite restrictive. For instance, verification of the representation (4) required by Anderson et al. (1997), as well as that of the Leadbetter conditions assumed by Dkengne et al. (2016), is very difficult. Similarly, the results obtained by Freitas and Hüsler essentially employ the assumption of twice differentiability, which is violated in case of the model (1). Thus, the conditions for these results are either difficult to check or inconsistent with the proposed model, so the question of the asymptotic behaviour of maxima in this case remains open.

1.3. Applications to the asset price modelling

One of the areas that significantly benefit from the theory concentrated on extremes is finance, where there is an imminent need for a proper risk assessment. This is especially due to the heavy-tailedness of financial data, in particular, financial returns, first established by Mandelbrot (1963) and now considered as a well-known stylised fact (Cont, 2001). Most commonly, this property is accounted for by using pure heavy-tailed distributions from the Fréchet domain of attraction, such as stable or Pareto laws, for modelling financial quantities. However, such a decision is criticised by numerous papers, admitting that the real tails of returns are though heavier than normal, yet lighter than of a power law. For instance, Laherrère and Sornette (1998) demonstrate that daily price variations on the exchange market can be successfully described by the Weibull distribution with parameter smaller than one. Malevergne et al. (2005) intently analyse financial returns on different time scales, ranging from daily to 5- and 1-minute data, and come to the conclusion that the Pareto distribution fits the highest 5% of the data, while the remaining 95% are most efficiently described by the Weibull law. Therefore, it is reasonable to expect that the overall distribution of returns should be successfully described by the model which (in some sense) lies in between of these two distributions.

This idea serves as a motivation for the application of the model (1) to modelling the financial returns. As can be seen, it exactly incorporates the idea of having a compromise between the Weibull and a RV law, where the latter one is responsible only for a small proportion of observations in the right tail of distribution. As we show in Section 3.2, this indeed leads to a better fit for the considered data, opening up a novel tool for the analysis of extreme values of price variations.

1.4. Truncated RV distributions

The common difficulty that arises with using regularly varying laws in practice is that the real distribution tails might be truncated. As pointed out by Beirlant et al. (2016), deviation from the pure power law can be caused by some natural bounds, though in many practical instances it is not apparent whether the underlying distribution should be unbounded or not. This motivates the study of the Weibull-truncated RV mixture in our paper.

In the context of truncated power laws, the asymptotic properties of random variables are known to strongly depend on the truncation level. This issue was analysed in detail by Chakrabarty and Samorodnitsky (2012), who have studied triangular arrays $\{X_{nj}, j = 1..n\}$, $n \to \infty$ of random variables X_{nj} with regularly varying distribution truncated at different levels $M_n \to \infty$ as $n \to \infty$. More precisely, they define X_{nj} as

$$X_{nj} = V_j \cdot \mathbb{I}\{\|V_j\| \le M_n\} + \frac{V_j}{\|V_j\|}(M_n + E_j) \cdot \mathbb{I}\{\|V_j\| > M_n\}, \quad j = 1..n,$$

where $V_1, \dots, V_n \in \mathbb{R}^d$ are i.i.d. random vectors having RV distribution with a tail index $\alpha \in (0, 2)$ and $E_1, \dots, E_n$ are non-negative i.i.d. random variables independent of $V_1, \dots, V_n$, and introduce the following truncation regimes in terms of the asymptotic behaviour of M_n:

1. soft, if $\lim\limits_{n\to\infty} n\mathbb{P}\{\|V_j\| > M_n\} = 0$,
2. hard, if $\lim\limits_{n\to\infty} n\mathbb{P}\{\|V_j\| > M_n\} = \infty$,
3. intermediate, if $\lim\limits_{n\to\infty} n\mathbb{P}\{\|V_j\| > M_n\} \in (0, \infty)$.

One of the key findings of that paper is that under soft truncation the random variables behave as in the unbounded case, in a sense that the properly normalised sums still converge to the same α-stable laws as before. To the contrary, in the hard truncation regime the normalised sums turn out to be asymptotically Gaussian, losing a significant trait of heavy-tailedness. For the intermediate case, the authors only remark that under the same normalisation the limit for partial sums is given by the α-stable law with jumps truncated in a certain manner.

In Section 2.2 we show that our findings are related to the truncation regimes introduced by Chakrabarty and Samorodnitsky (2012), and in some cases lead to similar conclusions. In addition, unlike that paper, we analyse the intermediate regime in detail.

Chapter 2. Theoretical part

2.1. Weibull-RV mixture

In this section, we focus on a particular case of the model (1), namely

$$F(x; \varepsilon, \vec{\theta}) = (1-\varepsilon)F_1(x; \lambda, \tau) + \varepsilon F_2(x; \alpha) \tag{5}$$

where $\vec{\theta} = (\lambda, \tau, \alpha)$, F_1 is the distribution function of the Weibull law,

$$F_1(x; \lambda, \tau) = F_1(x) = 1 - e^{-\lambda x^\tau}, \quad x \geq 0, \lambda > 0, \tau > 0, \tag{6}$$

and F_2 corresponds to the regularly varying distribution on $[m, \infty)$,

$$F_2(x; \alpha) = F_2(x) = 1 - x^{-\alpha}L(x), \quad \alpha > 0, \quad x \in [m, \infty), \tag{7}$$

with $m = \inf\{x > 0 : F_2(x) > 0\}$ and a continuous slowly varying function $L(\cdot)$. Let us recall that by definition,

$$\lim_{x\to\infty} L(tx)/L(x) = 1, \qquad \forall t > 0,$$

and the term «slow variation» comes from the property

$$x^{-\epsilon}L(x) \to 0 \quad \text{and} \quad x^{\epsilon}L(x) \to \infty \qquad \text{as } x \to \infty \tag{8}$$

for every $\epsilon > 0$. The extensive overview of the properties of slowly varying functions is given in the books by Bingham (1987) and Resnick (2013).

In the classical setting of i.i.d. sequences, when the parameters ε and $\vec{\theta}$ are fixed, it is not difficult to show that the mixture distribution function F is in the MDA of the Fréchet law. Indeed, let us analyse the asymptotic behaviour of maxima of a sequence of i.i.d. random variables $X_1, X_2, \ldots, X_n$, $n \geq 2$, with cumulative distribution function (5). That is, we consider

$$\lim_{n\to\infty} \mathbb{P}\left\{\max_{i=1,\dots,n} X_i \le v_n(x)\right\},$$

where $v_n(x) = s_n x + c_n$ is some non-decreasing normalising sequence unbounded in n and x. Since $X_i, i = 1..n$ are independent,

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\left\{\max_{i=1,\dots,n} X_i \le v_n(x)\right\} &= \lim_{n\to\infty} \left(\mathbb{P}\left\{X_1 \le v_n(x)\right\}\right)^n \\
&= \lim_{n\to\infty} \Big((1-\varepsilon)(1-e^{-\lambda v_n^\tau(x)}) \\
&\qquad\qquad +\varepsilon(1-v_n^{-\alpha}(x)L(v_n(x)))\Big)^n \\
&= \exp\Big\{-\lim_{n\to\infty} n\Big((1-\varepsilon)e^{-\lambda v_n^\tau(x)} \\
&\qquad\qquad +\varepsilon v_n^{-\alpha}(x)L(v_n(x))\Big)\Big\} \\
&= \exp\Big\{-\lim_{n\to\infty} n\varepsilon v_n^{-\alpha}(x)L(v_n(x)) \\
&\qquad \times\left(\frac{1-\varepsilon}{\varepsilon}\cdot\frac{v_n^\alpha(x)}{e^{\lambda(v_n(x))^\tau}L(v_n(x))}+1\right)\Big\}.
\end{aligned}$$

Since $v_n(x) \to \infty$ as $n \to \infty$, $\lim_{n\to\infty} v_n^\alpha(x)\left(e^{\lambda(v_n(x))^\tau}L(v_n(x))\right)^{-1} = 0$, and therefore

$$\lim_{n\to\infty} \mathbb{P}\left\{\max_{i=1,\dots,n} X_i \le v_n(x)\right\} = \exp\left\{-\lim_{n\to\infty} n\varepsilon v_n^{-\alpha}(x)L(v_n(x))\right\}.$$

Thus, the limit distribution for maxima is determined by the second component, leading to the Fréchet limit. In fact, choosing

$$c_n = 0, \qquad s_n = (1/\bar{F}_2)^{\leftarrow}(n),$$

we get

$$\lim_{n\to\infty} \mathbb{P}\left\{\max_{i=1,\dots,n} X_i \le v_n(x)\right\} = \exp\left\{-\varepsilon x^{-\alpha}\right\},$$

that is the Fréchet-type distribution.

It should be noted that such a result seems to be counterintuitive, since no matter the smallness of the mixture parameter ε, the first component does not bring any contribution to the asymptotic behaviour of maxima.

In what follows we consider the case when the mixing parameter $\varepsilon = \varepsilon_n$ decays to zero as n grows. It is natural to slightly generalise the model to the form of row-wise independent triangular array

$$X_{nj} \sim F(x; \varepsilon_n, \vec{\theta}), \qquad n \geq 1, \qquad j = 1..k_n, \tag{9}$$

where k_n is an unbounded increasing sequence, and for any $i = 1..n$, the r.v.'s $X_{nj}, j = 1..k_n$ are independent. The set-up allowing various numbers of elements in different rows is standard both in studying the classical limit laws (see Petrov, 2012) and in the extreme value theory (see Dkengne et al., 2016).

As we show in the next theorem, the asymptotic behaviour of the maximum in this model is determined by the rate of growth of k_n, the rate of decay of ε_n and the slowly varying function L. Note that the rates of $\log k_n$ and $k_n\varepsilon_n$ are compared in terms of the following three alternative conditions,

$$\exists \beta > {}^{\tau}/_{\alpha}: \qquad \lim_{n\to\infty} \frac{\log k_n}{(k_n\varepsilon_n)^{\beta}} = \infty, \tag{A1}$$

$$\exists \beta \in (0, {}^{\tau}/_{\alpha}): \qquad \lim_{n\to\infty} \frac{\log k_n}{(k_n\varepsilon_n)^{\beta}} = 0, \tag{A2}$$

$$\exists c > 0: \qquad k_n\varepsilon_n = c(\log k_n)^{\alpha/\tau}. \tag{A3}$$

The graphical representation of this result is presented in Figure 1.

Theorem 1. *Consider the row-wise independent triangular array (9). Assume that* $\lim_{x\to\infty} L(x) \in [0, \infty]$.[1] *Then for any sequences* $k_n \to \infty$, $\varepsilon_n \to 0$ *there exist deterministic sequences* c_n, s_n *such that*

$$\lim_{n\to\infty} \mathbb{P}\left\{\max_{j=1,\dots,k_n} X_{nj} \leq s_n x + c_n\right\} = H(x), \quad \forall x \in \mathbb{R} \tag{10}$$

with some non-degenerate limit law $H(x)$. More precisely, $H(x)$ belongs to the type of the following three distribution functions[2]*:*

(i) <u>Gumbel distribution,</u> $H(x) = e^{-e^{-x}}$, if and only if any of the following

[1] This assumption means that the slowly varying function $L(x)$ doesn't exhibit infinite oscillation. The counterexample to this condition is given in Mikosch (1999), Example 1.1.6.

[2] Due to the convergence to types theorem (Theorem A1.5 from Embrechts et al.,1997), if $H(x)$ is the distribution function of the limit law in (10), then any other non-degenerate law appearing in (10) under another normalisation is of the form $H(ax + b)$ with some constants a, b.

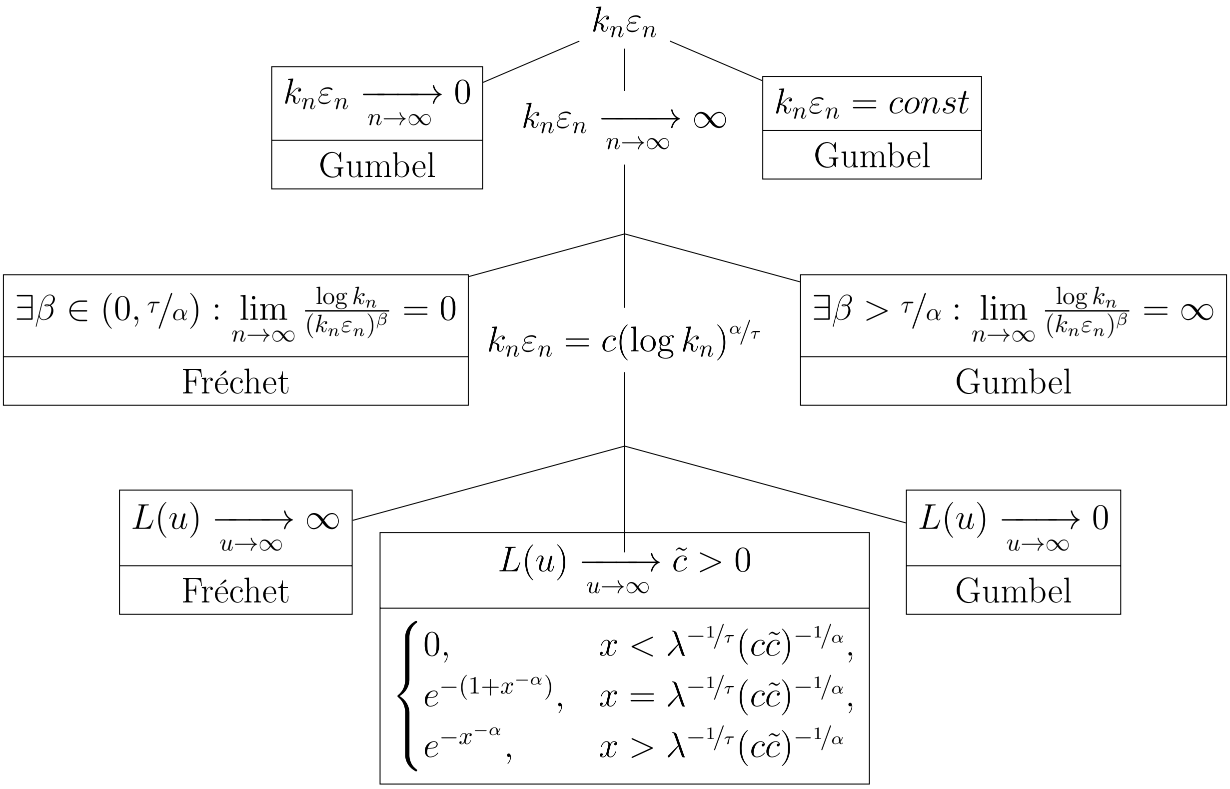


Figure 1: Possible limit distributions for maxima of the triangular array (5)

conditions is satisfied

(a) $k_n\varepsilon_n \to 0$ *as* $n \to \infty$ *or* $k_n\varepsilon_n = const > 0$*;*

(b) $k_n\varepsilon_n \to \infty$ *as* $n \to \infty$*, and* (A1) *holds;*

(c) (A3) *holds, and* $L(u) \to 0$ *as* $u \to \infty$*.*

In all cases, possible choice of the normalising sequences is

$$s_n = (\lambda\tau)^{-1}(\lambda^{-1}\log k_n)^{1/\tau - 1} \quad \text{and} \quad c_n = (\lambda^{-1}\log k_n)^{1/\tau}. \tag{11}$$

(ii) *<u>Fréchet distribution</u> with parameter* α*,* $H(x) = e^{-x^{-\alpha}}$*, if and only if any of the following conditions is satisfied*

(a) (A2) *holds;*

(b) (A3) *holds, and* $L(u) \to \infty$ *as* $u \to \infty$*.*

In all cases, one can take

$$s_n = F_2^{\leftarrow}(1 - (k_n\varepsilon_n)^{-1}), \qquad \text{and} \qquad c_n = 0, \tag{12}$$

where $F_2^{\leftarrow}(y) = \inf\{x \in \mathbb{R} : F_2(x) \geq y\}$ *for* $y \in [0,1]$.

(iii) *Discontinuous distribution with cdf*

$$H(x) = \begin{cases} 0, & x < \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ e^{-(1+x^{-\alpha})}, & x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ e^{-x^{-\alpha}}, & x > \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \end{cases}$$

if (A3) *holds, and* $L(u) \to \tilde{c}$ *for some* $\tilde{c} > 0$ *as* $u \to \infty$. *The normalising sequences can be fixed in the form* (12).

Proof. For given sequences s_n, c_n, the left-hand side of (10) can be represented as

$$\begin{aligned} &\lim_{n\to\infty} \left(\mathbb{P}\left\{X_1 \leq v_n(x)\right\}\right)^{k_n} \\ &\quad = \lim_{n\to\infty} \Big(\big(1-\varepsilon_n\big)\big(1-e^{-\lambda v_n^{\tau}(x)}\big) + \varepsilon_n\big(1 - v_n^{-\alpha}(x)L(v_n(x))\big)\Big)^{k_n} \\ &\quad\quad = \lim_{n\to\infty} (1 - e^{-\lambda v_n^{\tau}(x)} - \varepsilon_n v_n^{-\alpha}(x)L(v_n(x)))^{k_n} \\ &\quad\quad\quad = \exp\Big\{-\lim_{n\to\infty}\Big(k_n e^{-\lambda v_n^{\tau}(x)} + k_n\varepsilon_n v_n^{-\alpha}(x)L(v_n(x))\Big)\Big\}, \end{aligned} \tag{13}$$

where $v_n(x) = s_n x + c_n$. Our aim is to find the sequences s_n, c_n guarantying that this limit (denoted by $H(x)$) is non-degenerate. We divide the range of possible rates of convergence of k_n, ε_n into several essentially different cases.

(i) Let $k_n\varepsilon_n \to 0$ as $n \to \infty$ or $k_n\varepsilon_n = const > 0$. As $v_n^{-\alpha}(x)L(v_n(x)) \to 0$ as $n \to \infty$ by the slow variation of $L(\cdot)$ (see (8)), we get

$$H(x) = \exp\Big\{-\lim_{n\to\infty} k_n\varepsilon_n v_n^{-\alpha}(x)L(v_n(x))\Big\},$$

and therefore we deal with the extreme value analysis of the Weibull law. Since $\bar{F}_1(x) = e^{-\lambda x^{\tau}}, \lambda > 0, \tau \geq 1$ is a von Mises function, i.e., $\bar{F}_1(x)$ can be represented as

$$\bar{F}_1(x) = \breve{c} \cdot \exp\left\{-\int_0^x \frac{1}{a(u)}\, du\right\}, \quad 0 < x < \infty,$$

with $\breve{c} = 1$ and $a(u) = (\lambda\tau)^{-1}u^{1-\tau}, x > 0$, we get that the limit distribution is Gumbel under the choice $v_n(x) = s_n x + c_n$ with

$$c_n := F_1^{\leftarrow}\left(1 - \frac{1}{k_n}\right) = \left(\frac{\log k_n}{\lambda}\right)^{1/\tau}$$

$$s_n := a(c_n) = (\lambda\tau)^{-1}\left(\frac{\log k_n}{\lambda}\right)^{1/\tau - 1}.$$

(ii) Let $k_n\varepsilon_n \to \infty$ as $n \to \infty$. This case is divided into several subcases, depending on the relation between s_n and c_n.

1. First, let us consider $s_n \gtrsim |c_n|$[3]. Then

$$v_n(x) = s_n x + c_n = s_n x\left(1 + \frac{c_n}{s_n x}\right) = s_n x(1 + \bar{o}(1))$$

as $n \to \infty$. Therefore,

$$H(x) = \exp\Big\{-\lim_{n\to\infty}\Big(k_n e^{-\lambda s_n^\tau x^\tau} + k_n\varepsilon_n s_n^{-\alpha}x^{-\alpha}L(s_n)\Big)\Big\},$$

as $L(s_n x) \sim L(s_n)$ for all fixed $x \in \mathbb{R}$ as $n \to \infty$. Clearly, since c_n is not present in the above limit, one can take $c_n = 0$. As for s_n, the choice

$$s_n^{-\alpha}L(s_n) = \frac{1}{k_n\varepsilon_n}, \tag{14}$$

i.e., $s_n := F_2^{\leftarrow}\left(1 - (k_n\varepsilon_n)^{-1}\right)$. We have

$$\lim_{n\to\infty} k_n e^{-\lambda s_n^\tau x^\tau} = \lim_{n\to\infty}\exp\left\{-s_n^\tau\left(\lambda x^\tau - \frac{\log k_n}{s_n^\tau}\right)\right\},$$

and therefore the limit distribution in (13) is non-degenerate (and is actually the Fréchet distribution) if and only if

$$\lim_{n\to\infty}\frac{\log k_n}{s_n^\tau} = 0. \tag{15}$$

It would be a worth mentioning that s_n depends on the function L via the

[3] Here and below we mean by $f_n \gtrsim g_n$ that $\lim\limits_{n\to\infty}(f_n/g_n) = \infty$.

equality (14). Let us recall that $L(\cdot)$ is slowly varying and therefore

$$L(s_n) \gtrsim s_n^{-\epsilon} \quad \forall \epsilon > 0.$$

Thus, from (14) we get

$$\frac{1}{k_n \varepsilon_n} = s_n^{-\alpha} L(s_n) \gtrsim s_n^{-\alpha-\epsilon} \quad \forall \epsilon > 0,$$

and

$$s_n \gtrsim (k_n \varepsilon_n)^{1/(\alpha+\epsilon)} \quad \forall \epsilon > 0.$$

Now, since

$$\frac{\log k_n}{s_n^\tau} \lesssim \frac{\log k_n}{(k_n \varepsilon_n)^{\tau/(\alpha+\epsilon)}},$$

we get that for the condition (15) to be fulfilled, it is sufficient that the right-hand side tends to zero as $n \to \infty$, i.e.,

$$\lim_{n\to\infty} \frac{\log k_n}{(k_n \varepsilon_n)^{\tau/(\alpha+\epsilon)}} = 0,$$

or, equivalently,

$$\exists \beta \in (0, \tau/\alpha) : \quad \lim_{n\to\infty} \frac{\log k_n}{(k_n \varepsilon_n)^\beta} = 0. \tag{16}$$

We conclude that the condition (16) yields (15), and in this case the limit distribution is Fréchet.

2. Now, let $s_n > 0$ and $c_n \in \mathbb{R}$ be such that $s_n \lesssim |c_n|$. In this case

$$v_n(x) = s_n x + c_n = c_n \left(\frac{s_n}{c_n} x + 1 \right) = c_n(1 + \bar{o}(1))$$

as $n \to \infty$. Thus, the limit in (13) takes the form

$$H(x) = \exp\Big\{ - \lim_{n\to\infty} \Big(k_n e^{-\lambda(s_n x + c_n)^\tau} + k_n \varepsilon_n c_n^{-\alpha} L(c_n) \Big) \Big\}.$$

Let the norming constants c_n and s_n be chosen in the form (11). Then

$H(x)$ is the cdf of the Gumbel law if

$$\lim_{n\to\infty} \frac{k_n \varepsilon_n L\left((\log k_n)^{1/\tau}\right)}{(\log k_n)^{\alpha/\tau}} = 0. \tag{17}$$

As previously, we would like to replace (17) with another condition without $L(\cdot)$. Once more, we would like to recall that by slow variation of $L(\cdot)$

$$L(x) \lesssim x^{\epsilon} \quad \forall \epsilon > 0.$$

From this we conclude that

$$\frac{k_n \varepsilon_n L\left((\log k_n)^{1/\tau}\right)}{(\log k_n)^{\alpha/\tau}} \lesssim \frac{k_n \varepsilon_n}{(\log k_n)^{(\alpha-\epsilon)/\tau}}$$

and the fact that the right-hand side tends to zero as $n \to \infty$ will imply (17). In other words, we obtain the Gumbel limit if

$$\lim_{n\to\infty} \frac{k_n \varepsilon_n}{(\log k_n)^{(\alpha-\epsilon)/\tau}} = \lim_{n\to\infty} \left(\frac{(k_n \varepsilon_n)^{\tau/(\alpha-\epsilon)}}{\log k_n} \right)^{(\alpha-\epsilon)/\tau} = 0,$$

or, equivalently, if

$$\exists \beta > \tau/\alpha : \quad \lim_{n\to\infty} \frac{\log k_n}{(k_n \varepsilon_n)^{\beta}} = \infty. \tag{18}$$

3. The last possible situation is when neither (16), nor (18) is satisfied. Clearly, this is the case only when $k_n \varepsilon_n = c \cdot (\log k_n)^{\alpha/\tau}$ for some $c > 0$. Not surprisingly, it turns out that the final answer now depends on the asymptotic behaviour of $L(\cdot)$.

a) Let us first consider the case when $L(u) \to \infty$ as $u \to \infty$. Then one can take $c_n = 0$ and find s_n as the solution to the equation

$$s_n^{-\alpha} L(s_n) = \frac{1}{k_n \varepsilon_n} = \frac{1}{c \cdot (\log k_n)^{\alpha/\tau}}.$$

The limit for the second component in (13) coincides with the corresponding one in item $2(i)$ (and leads to the cdf of the Fréchet law), while for the first component we get

$$\lim_{n\to\infty} k_n e^{-\lambda s_n^\tau x^\tau} = \lim_{n\to\infty} k_n \exp\left\{-\lambda\left(\frac{s_n^{-\alpha}L(s_n)}{L(s_n)}\right)^{-\tau/\alpha} x^\tau\right\}$$
$$= \lim_{n\to\infty} k_n^{1-\lambda(cL(s_n))^{\tau/\alpha}x^\tau}. \tag{19}$$

The value of the latter limit is zero for all fixed $x > 0$ since $L(s_n) \to \infty$ as $n \to \infty$, and therefore the limit distribution is Fréchet.

b) Now, let $L(\cdot)$ be such that $L(u) \to \tilde{c} > 0$ as $u \to \infty$. Then the same choice of norming constants as when $L(u) \to \infty$ as $u \to \infty$ leads to the same limits as before. However, the value of (19) now depends on x, namely,

$$\lim_{n\to\infty} k_n^{1-\lambda(cL(s_n))^{\tau/\alpha}x^\tau} = \begin{cases}\infty, & x \in (0, \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}),\\ 1, & x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha},\\ 0, & x > \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}.\end{cases}$$

Thus, in this case the limit distribution is equal to

$$\lim_{n\to\infty} \mathbb{P}\left\{\max_{j=1,\dots,k_n} X_{nj} \le v_n(x)\right\} = \begin{cases}0, & x < \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha},\\ e^{-(1+x^{-\alpha})}, & x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha},\\ e^{-x^{-\alpha}}, & x > \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}.\end{cases}$$

An interesting point is that we get the limit distribution that is not from the extreme value family, having an atom at $x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}$.

c) Finally, let $L(\cdot)$ be such that $L(u) \to 0$ as $u \to \infty$. Then the normalising sequence can be chosen as in item 1, and for the first component we get

$$\lim_{n\to\infty} k_n e^{-\lambda v_n^\tau(x)} = e^{-x} \quad \forall x \in \mathbb{R},$$

while for the second one

$$\lim_{n\to\infty} \frac{k_n \varepsilon_n L(c_n)}{c_n^\alpha} = \lim_{n\to\infty} \frac{c \cdot (\log k_n)^{\alpha/\tau} L\left((\log k_n)^{1/\tau}\right)}{(\log k_n)^{\alpha/\tau}}$$
$$= \lim_{n\to\infty} c \cdot L\left((\log k_n)^{1/\tau}\right) = 0.$$

Therefore, in this case the limit distribution is again Gumbel.

□

2.2. Weibull-truncated RV mixture

Now we consider one more complicated model, such that the distribution of the second component in (5) also changes as n grows. Consider the mixture distribution

$$F(x; \varepsilon, M, \vec{\theta}) = (1-\varepsilon) F_1(x; \lambda, \tau) + \varepsilon \widetilde{F}_2(x; \alpha, M), \tag{20}$$

where as before, $\vec{\theta} = (\lambda, \tau, \alpha)$, F_1 is the distribution function of the Weibull law (see (6)), while $\widetilde{F}_2$ is the upper-truncated regularly varying distribution,

$$\widetilde{F}_2(x; \alpha, M) = \begin{cases} \frac{F_2(x;\alpha)}{F_2(M;\alpha)}, & \text{if } x \in [m, M], \\ 1, & \text{if } x > M \end{cases} \tag{21}$$

with $F_2(x; \alpha)$ corresponding to a regularly varying distribution (7).

It would be an interesting mentioning that the components in this model correspond to different maximum domains of attraction: the maximum for the first component under proper normalisation converges to the Gumbel law, while the second — to the Weibull law. The latter claim is shown in the following lemma.

Lemma 1. *Let $\widetilde{F}_2(x) = F_2(x; m, M, L)$ be the upper-truncated regularly varying distribution defined as (21). Then $\widetilde{F}_2$ is in the maximum domain of attraction of the Weibull law Ψ_τ having the distribution function*

$$\Psi_\tau = e^{-(-x)^\tau} \cdot \mathbb{I}\{x \le 0\} + \mathbb{I}\{x > 0\}$$

with $\tau = 1$.

Proof. As it is known, $F \in MDA(\Psi_\tau)$ for some $\tau > 0$ if and only if $x^* = \sup\{x \in \mathbb{R} : F(x) < 1\} < \infty$ and $\bar{F}\left(x^* - \frac{1}{x}\right) \in RV_{-\tau}$, see, e.g, Embrechts et al. (1997). Thus, $\widetilde{F}_2 \in MDA(\Psi_\tau)$ for some $\tau > 0$ if and only if

$$\bar{F}\left(M - \frac{1}{x}\right) = x^{-\tau}\tilde{L}(x), \quad \tau > 0,$$

for some slowly varying function $\tilde{L}(\cdot)$, or, equivalently, iff

$$G(x) := x^{\tau}\bar{F}\left(M - \frac{1}{x}\right) \in RV_0, \quad \tau > 0.$$

Therefore, to prove this statement of this lemma, we need to show that $\lim_{x\to\infty}(G(tx)/G(x)) = 1$ for any $t > 0$, that is,

$$\begin{aligned} \lim_{x\to\infty} \frac{G(tx)}{G(x)} &= \lim_{x\to\infty} t^{\tau} \frac{\bar{F}\left(M - \frac{1}{tx}\right)}{\bar{F}\left(M - \frac{1}{x}\right)} \\ &= \lim_{x\to\infty} t^{\tau} \frac{\left(M - \frac{1}{tx}\right)^{-\alpha} L\left(M - \frac{1}{tx}\right) - M^{-\alpha}L(M)}{\left(M - \frac{1}{x}\right)^{-\alpha} L\left(M - \frac{1}{x}\right) - M^{-\alpha}L(M)} = 1 \quad \forall t > 0. \end{aligned} \tag{22}$$

First,

$$\lim_{x\to\infty}\left(M - \frac{1}{x}\right)^{-\alpha} = \lim_{x\to\infty} M^{-\alpha}\left(1 - \frac{1}{Mx}\right)^{-\alpha} = M^{-\alpha} + \alpha M^{-\alpha-1}\cdot\frac{1}{x}(1 + \bar{o}(1)).$$

Then, assuming that $L(\cdot)$ is continuous and differentiable,

$$\lim_{x\to\infty} L\left(M - \frac{1}{x}\right) = L(M) - L'(M)\cdot\frac{1}{x}(1 + \bar{o}(1)).$$

Therefore,

$$\lim_{x\to\infty} t^{\tau} \frac{\left(M - \frac{1}{tx}\right)^{-\alpha} L\left(M - \frac{1}{tx}\right) - M^{-\alpha}L(M)}{\left(M - \frac{1}{x}\right)^{-\alpha} L\left(M - \frac{1}{x}\right) - M^{-\alpha}L(M)}$$

$$= \lim_{x\to\infty} t^{\tau} \frac{\left(M^{-\alpha} + \frac{\alpha M^{-\alpha-1}}{tx}(1+\bar{o}(1))\right)\left(L(M) - \frac{L'(M)}{tx}(1+\bar{o}(1))\right) - M^{-\alpha}L(M)}{\left(M^{-\alpha} + \frac{\alpha M^{-\alpha-1}}{x}(1+\bar{o}(1))\right)\left(L(M) - \frac{L'(M)}{x}(1+\bar{o}(1))\right) - M^{-\alpha}L(M)}$$

$$= \lim_{x\to\infty} t^{\tau} \frac{\left(\frac{M^{-\alpha}L'(M)}{tx} + \frac{\alpha M^{-\alpha-1}}{tx}L(M)\right)(1+\bar{o}(1))}{\left(\frac{M^{-\alpha}L'(M)}{x} + \frac{\alpha M^{-\alpha-1}}{x}L(M)\right)(1+\bar{o}(1))}.$$

Clearly, the latter limit is equal to one if $\tau = 1$, meaning that $\widetilde{F}_2 \in MDA(\Psi_1)$. □

By analogue with (9), we consider the triangular array

$$X_{nj} \sim F(x; \varepsilon_n, M_n, \vec{\theta}), \qquad n \geq 1, \qquad j = 1..k_n, \tag{23}$$

where k_n, M_n are unbounded increasing sequences, and for any $i = 1..n$, the r.v.'s $X_{nj}, j = 1..k_n$ are independent. Note that the classical limit laws for this model (law of large numbers and limit theorems for the sums) are essentially established by Panov (2017).

The next theorem reveals the asymptotic behaviour of the maximal value depending on the rates of ε_n, M_n, k_n, and the properties of the slowly varying function L. An important difference from the model considered in Section 2.1 is that in some cases the limit distribution is degenerate for any (also non-linear) normalising sequence.

It turns out that if $k_n\varepsilon_n$ tends to any finite constant, then the limit distribution is Gumbel. Our findings in the remaining case $k_n\varepsilon_n \to \infty$ are presented in Table 1. The asymptotic behaviour of the maximum is determined by the asymptotic properties of the sequences k_n, ε_n in terms of (A1)-(A3), and the rate of growth of M_n in terms of the following three alternating conditions:

$$\exists \gamma > 1/\alpha : \qquad \lim_{n\to\infty} \frac{M_n}{(k_n\varepsilon_n)^{\gamma}} = \infty, \tag{M1}$$

$$\exists \gamma \in (0, {}^1\!/\!{}_\alpha) : \qquad \lim_{n\to\infty} \frac{M_n}{(k_n\varepsilon_n)^\gamma} = 0; \tag{M2}$$

$$\exists \breve{c} > 0 : \qquad M_n = \breve{c}(k_n\varepsilon_n)^{1/\alpha}. \tag{M3}$$

The conditions (M1)-(M3) are related to the notion of hard- and soft truncation. Following Chakrabarty and Samorodnitsky (2012), we say that that a variable η is truncated softly, if

$$\lim_{n\to\infty} k_n \mathbb{P}\big\{|\eta| \geq M_n\big\} = 0. \tag{24}$$

For a regularly varying distribution of η, the condition (24) holds if there exists $\gamma > 1/\alpha$ such that $M_n/k_n^\gamma \to \infty$. This fact follows from

$$k_n \mathbb{P}\big\{|\eta| \geq M_n\big\} = k_n M_n^{-\alpha} L(M_n) \lesssim k_n M_n^{-\alpha+\epsilon} = (M_n/k_n^{1/(\alpha-\epsilon)})^{(-\alpha+\epsilon)}$$

for any $\epsilon > 0$. Analogously, η is truncated hard, that is,

$$\lim_{n\to\infty} k_n \mathbb{P}\big\{|\eta| \geq M_n\big\} = \infty,$$

if there exists $\gamma \in (0, 1/\alpha)$ such that $M_n/k_n^\gamma \to 0$.

Our results for the case (M1) (see first raw in Table 1) coincide with the findings from Chakrabarty and Samorodnitsky (2012): *«in the soft truncation regime, truncated power tails behave, in important respects, as if no truncation took place»* (p.109). In fact, in our setup, the results are completely the same as for the non-truncated distribution considered in Theorem 1.

Our outcomes for (M2) (second raw in Table 1) are quite close to another finding from Chakrabarty and Samorodnitsky (2012), namely, *«in the hard truncation regime much of "heavy tailedness" is lost»* (p.109). Actually, we get that the behaviour is determined by the first component except the case (A2) with $\lim_{n\to\infty} k_n e^{-\lambda M_n^\tau} \neq \infty$.

Finally, the intermediate case (M3) (third raw in Table 1) is divided into various subcases. The comparison with Chakrabarty and Samorodnitsky (2012) is not possible because the authors decide to *«largely leave* [this question] *aside in this article, in order to keep its size manageable»* (p.111). In our research, we provide the complete study of this case.

The exact result is formulated below.

Table 1

Possible limit distributions for maxima of the triangular array (23)

<table>
<tr><th>$k_n\varepsilon_n \underset{n\to\infty}{\longrightarrow} \infty$</th><th>(A1)</th><th>(A2)</th><th>(A3)</th></tr>
<tr><td rowspan="3">(M1)</td><td rowspan="3">Gumbel</td><td rowspan="3">Fréchet</td><td>Gumbel, if $L(u) \underset{u\to\infty}{\longrightarrow} 0$</td></tr>
<tr><td>Fréchet, if $L(u) \underset{u\to\infty}{\longrightarrow} \infty$</td></tr>
<tr><td>Distribution I, if $L(u) \underset{u\to\infty}{\longrightarrow} \tilde{c} \in (0,\infty)$</td></tr>
<tr><td rowspan="2">(M2)</td><td rowspan="2">Gumbel</td><td>Gumbel, if $k_n \gtrsim e^{\lambda M_n^\tau}$</td><td rowspan="2">Gumbel</td></tr>
<tr><td>no limit, if $k_n \gtrsim e^{\lambda M_n^\tau}$ is not fulfilled</td></tr>
<tr><td rowspan="5">(M3)</td><td rowspan="5">Gumbel</td><td rowspan="2">Fréchet, if $L(u) \underset{u\to\infty}{\longrightarrow} 0$</td><td>Gumbel, if $L(u) \underset{u\to\infty}{\longrightarrow} 0$</td></tr>
<tr><td>Gumbel, if $L(u) \underset{u\to\infty}{\longrightarrow} \tilde{c} \in (0,\infty]$, and $\lambda^{1/\tau}\breve{c}c^{1/\alpha} \in (0,1)$</td></tr>
<tr><td rowspan="2">Distribution II, if $L(u) \underset{u\to\infty}{\longrightarrow} \tilde{c} \in (0,\infty)$</td><td>Distribution III, if $L(u) \underset{u\to\infty}{\longrightarrow} \tilde{c} \in (0,\infty)$, and $\lambda^{1/\tau}\breve{c}c^{1/\alpha} > 1$</td></tr>
<tr><td>Distribution IV, if $L(u) \underset{u\to\infty}{\longrightarrow} \tilde{c} \in (0,\infty)$, and $\lambda^{1/\tau}\breve{c}c^{1/\alpha} = 1$</td></tr>
<tr><td>no limit, if $L(u) \underset{u\to\infty}{\longrightarrow} \infty$</td><td>no limit, if $L(u) \underset{u\to\infty}{\longrightarrow} \infty$, and $\lambda^{1/\tau}\breve{c}c^{1/\alpha} \geq 1$</td></tr>
</table>

Theorem 2. *Consider the row-wise independent triangular array* (23) *under the assumption that* $\varepsilon_n \to 0, M_n \to \infty, k_n \to \infty$ *as* $n \to \infty$. *Assume also that* $\lim_{x\to\infty} L(x) \in [0,\infty]$.

Then the non-degenerate limit law $H(x)$ *for the properly normalised row-wise maximum* $\max\limits_{j=1,\ldots,k_n} X_{nj}$ *(see* (10)*) belongs to the type of the following dis-*

tributions.

1. *Gumbel distribution,* $H(x) = e^{-e^{-x}}$, *if and only if any of the following conditions is satisfied*

   1.1 $k_n\varepsilon_n \to 0$ *as* $n \to \infty$ *or* $k_n\varepsilon_n = const > 0$;

   1.2 $k_n\varepsilon_n \to \infty$ *as* $n \to \infty$, *and moreover*

   - (M1) *and* (A1) *hold;*
   - (M1) *and* (A3) *hold, and* $L(u) \to 0$ *as* $u \to \infty$;
   - (M2) *and* (A1) *hold;*
   - (M2) *and* (A2) *hold, and* $k_n \gtrsim e^{\lambda M_n^\tau}$;
   - (M2) *and* (A3) *hold;*
   - (M3) *and* (A1) *hold;*
   - (M3) *and* (A3) *hold, and* $L(u) \to 0$ *as* $u \to \infty$;
   - (M3) *and* (A3) *hold,* $L(u) \to \tilde{c} \in (0,\infty]$ *as* $u \to \infty$, *and* $\lambda^{1/\tau}\breve{c}c^{1/\alpha} \in (0,1)$.

   In all cases, possible choice of the normalising sequences is given by (11).

2. *Fréchet distribution with parameter* α, $H(x) = e^{-x^{-\alpha}}$, *if and only if any of the following conditions is satisfied*

   - (M1) *and* (A2) *hold;*
   - (M1) *and* (A3) *hold, and* $L(u) \to \infty$ *as* $u \to \infty$;
   - (M3) *and* (A2) *hold, and* $L(u) \to 0$ *as* $u \to \infty$.

   In all cases, one can take s_n, c_n *in the form* (12).

3. *Special cases:*

   (I)

$$H(x) = \begin{cases} 0, & x < \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ e^{-(1+x^{-\alpha})}, & x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ e^{-x^{-\alpha}}, & x > \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \end{cases}$$

   provided (M1) *and* (A3) *hold, and* $L(u) \to \tilde{c} \in (0,\infty)$;

(II)

$$H(x) = \begin{cases} e^{\tilde{c}\breve{c}^{-\alpha} - x^{-\alpha}}, & x \in (0, \breve{c}\tilde{c}^{-1/\alpha}], \\ 1, & x > \breve{c}\tilde{c}^{-1/\alpha}, \end{cases}$$

provided (M3) *and* (A2) *hold,* $L(u) \to \tilde{c} \in (0, \infty)$*;*

(III)

$$H(x) = \begin{cases} 0, & x < \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ \exp\left\{-1 + \tilde{c}\breve{c}^{-\alpha} - \lambda^{\alpha/\tau} c\tilde{c}\right\}, & x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ \exp\left\{\tilde{c}\breve{c}^{-\alpha} - x^{-\alpha}\right\}, & x \in (\lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \breve{c}\tilde{c}^{-1/\alpha}], \\ 1, & x > \breve{c}\tilde{c}^{-1/\alpha}, \end{cases}$$

provided (M3) *and* (A3) *hold,* $L(u) \to \tilde{c} \in (0, \infty)$ *as* $u \to \infty$*, and* $\lambda^{1/\tau}\breve{c}c^{1/\alpha} > 1$*;*

(IV)

$$H(x) = \begin{cases} 0, & x < \breve{c}\tilde{c}^{-1/\alpha}, \\ e^{-1}, & x = \breve{c}\tilde{c}^{-1/\alpha}, \\ 1, & x > \breve{c}\tilde{c}^{-1/\alpha}, \end{cases}$$

provided (M3) *and* (A3) *hold,* $L(u) \to \tilde{c} \in (0, \infty)$ *as* $u \to \infty$*, and* $\lambda^{1/\tau}\breve{c}c^{1/\alpha} = 1$*.*

In all cases the normalising sequences can be chosen as in (12)*.*

The limit distribution is degenerate for any sequences s_n *and* c_n *in the following three cases:*

- (M2) *and* (A2) *hold, and the condition* $k_n \gtrsim e^{\lambda M_n^\tau}$ *is not fulfilled;*
- (M3) *and* (A2) *hold, and* $L(u) \to \infty$ *as* $u \to \infty$*;*
- (M3) *and* (A3) *hold,* $L(u) \to \infty$ *as* $u \to \infty$*, and* $\lambda^{1/\tau}\breve{c}c^{1/\alpha} \geq 1$*.*

Moreover, in these cases the distribution of

$$\lim_{n\to\infty} \mathbb{P}\left\{\max_{j=1,\dots,k_n} X_{nj} \leq v_n(x)\right\}$$

is degenerate for any increasing sequence $v_n(x)$, which is unbounded in n and x.

Proof. **Step 1. Several simple cases.** As in the proof of Theorem 1, we use the notation $v_n(x) = s_n x + c_n$. We have

$$\begin{aligned} H(x) &= \lim_{n\to\infty}\Bigg((1-\varepsilon_n)(1-e^{-\lambda v_n^\tau(x)}) \\ &\qquad +\varepsilon_n\left(\frac{1-v_n^{-\alpha}(x)L(v_n(x))}{1-M_n^{-\alpha}L(M_n)}\mathbb{I}_{\{v_n(x)\in[m,M_n]\}}\right. \\ &\qquad\qquad \left.+\mathbb{I}_{\{v_n(x)>M_n\}}\right)\Bigg)^{k_n} \\ &= \lim_{n\to\infty}\Bigg(1-(1-\varepsilon_n)e^{-\lambda v_n^\tau(x)} \\ &\qquad +\varepsilon_n\left(\frac{M_n^{-\alpha}L(M_n)-v_n^{-\alpha}(x)L(v_n(x))}{1-M_n^{-\alpha}L(M_n)}\mathbb{I}_{\{v_n(x)\in[m,M_n]\}}\right. \\ &\qquad\qquad \left.-\mathbb{I}_{\{v_n(x)<m\}}\right)\Bigg)^{k_n}, \end{aligned}$$

where we use that $\mathbb{I}_{\{v_n(x)>M_n\}} = 1-\mathbb{I}_{\{v_n(x)\in[m,M_n]\}}-\mathbb{I}_{\{v_n(x)<m\}}$. Since $\varepsilon_n\to 0$, $M_n\to\infty$ and $M_n^{-\alpha}L(M_n)\to 0$ as $n\to\infty$ by slow variation of $L(\cdot)$, we get

$$\begin{aligned} H(x) &= \lim_{n\to\infty}\Big(1-e^{-\lambda v_n^\tau(x)} \\ &\qquad +\varepsilon_n((M_n^{-\alpha}L(M_n)-v_n^{-\alpha}(x)L(v_n(x)))\mathbb{I}_{\{v_n(x)\in[m,M_n]\}} \\ &\qquad\qquad -\mathbb{I}_{\{v_n(x)<m\}})\Big)^{k_n} \\ &= \exp\Big\{-\lim_{n\to\infty}\Big(k_n e^{-\lambda v_n^\tau(x)} \\ &\qquad -k_n\varepsilon_n(M_n^{-\alpha}L(M_n)-v_n^{-\alpha}(x)L(v_n(x)))\mathbb{I}_{\{v_n(x)\in[m,M_n]\}} \\ &\qquad\qquad +k_n\varepsilon_n\mathbb{I}_{\{v_n(x)<m\}}\Big)\Big\}. \end{aligned}$$

(*i*) Let $k_n\varepsilon_n\to 0$ as $n\to\infty$. By slow variation of $L(\cdot)$, $v_n^{-\alpha}(x)L(v_n(x))\to 0$ as $n\to\infty$, therefore, the whole second component disappears. We deal with maxima of a Weibull random variable and obtain the Gumbel limit under the normalisation $v_n(x) = s_n x + c_n$ with s_n, c_n in the form (15).

(*ii*) Let $k_n\varepsilon_n = const$. By a similar argument as in the previous item,

$$H(x) = \lim_{n\to\infty}\left(1 - \frac{1}{k_n}[k_n e^{-\lambda v_n^\tau(x)} + k_n\varepsilon_n \mathbb{I}_{\{v_n(x)<m\}}]\right)^{k_n}.$$

Under the same choice of normalising sequences, we get

$$v_n(x) = \frac{x}{\lambda\tau}\left(\frac{\log k_n}{\lambda}\right)^{1/\tau - 1} + \left(\frac{\log k_n}{\lambda}\right)^{1/\tau}$$
$$= \left(\frac{\log k_n}{\lambda}\right)^{1/\tau}\left(\frac{x}{\tau\log k_n} + 1\right) \to \infty$$

for all fixed $x \in \mathbb{R}$ as $n \to \infty$. Therefore, the limit distribution is again Gumbel.

(*iii*) Let $k_n\varepsilon_n \to \infty$ as $n \to \infty$. The further analysis depends on the asymptotic properties of M_n. If (M1) holds, the proof is based on the observation that

$$M_n^{-\alpha}L(M_n) \lesssim M_n^{-\alpha+\epsilon} \quad \forall\epsilon > 0,$$

and therefore,

$$k_n\varepsilon_n M_n^{-\alpha}L(M_n) \lesssim k_n\varepsilon_n M_n^{-\alpha+\epsilon} \quad \forall\epsilon > 0.$$

From (M1) it follow that for any $\epsilon \in (0, \alpha - 1/\gamma]$ the right-hand side tends to zero as $n \to \infty$, and therefore $k_n\varepsilon_n M_n^{-\alpha}L(M_n) \to 0$ as $n \to \infty$. The rest of the proof in this situation follows the same lines as the proof of Theorem 1. Other cases are more complicated, and we divide the further proof into several steps.

Step 2. Case (M2). Recalling again that $L(\cdot) \in RV_0$, we get that

$$k_n\varepsilon_n M_n^{-\alpha}L(M_n) \gtrsim k_n\varepsilon_n M_n^{-\alpha-\epsilon} \quad \forall\epsilon > 0.$$

Therefore, for any $\epsilon \in (0, 1/\gamma - \alpha]$, $k_n\varepsilon_n M_n^{-\alpha}L(M_n) \to \infty$ as $n \to \infty$. Now, assume that $v_n(x)$ is such that $v_n(x) \le M_n$ for all $x \in \mathbb{R}$ and n large enough.

Then

$$\begin{aligned}
\lim_{x\to\infty} & H(x) \\
&= \exp\Big\{-\lim_{n\to\infty}\lim_{x\to\infty}\Big(k_n e^{-\lambda v_n^\tau(x)} \\
&\qquad -k_n\varepsilon_n(M_n^{-\alpha}L(M_n)-v_n^{-\alpha}(x)L(v_n(x)))\mathbb{I}_{\{v_n(x)\in[m,M_n]\}} \\
&\qquad\qquad + k_n\varepsilon_n\mathbb{I}_{\{v_n(x)<m\}}\Big)\Big\} = e^{\lim_{n\to\infty} k_n\varepsilon_n M_n^{-\alpha}L(M_n)} = \infty,
\end{aligned}$$

and therefore the limit distribution doesn't exist. We conclude that there is a non-degenerate limit distribution only if $v_n(x) > M_n$ for all $x \in \mathbb{R}$ and n large enough. In this case,

$$k_n e^{-\lambda v_n^\tau(x)} = -\log(H(x))(1+\bar{o}(1)).$$

The condition $v_n(x) > M_n$ leads to the inequality

$$\frac{k_n}{e^{\lambda M_n^\tau}} > -\log(H(x))(1+\bar{o}(1)) \tag{25}$$

for n sufficiently large. Finally, as $-\log(H(x))$ takes only non-negative values, and the sequences k_n, M_n tend to infinity as $n \to \infty$, we conclude that the necessary condition for the existence of non-degenerate limit distribution is

$$M_n^\tau \lesssim \log k_n. \tag{26}$$

1. Now let us consider the case when (A1) holds. We have

$$\frac{M_n^\tau}{\log k_n} \lesssim \frac{M_n^\tau}{(k_n\varepsilon_n)^\beta} \lesssim \left(\frac{M_n}{(k_n\varepsilon_n)^\gamma}\right)^\tau,$$

since $\beta > \tau/\alpha > \tau\gamma$. The right-hand side tends to zero as $n \to \infty$ and therefore (26) holds. Choosing s_n and c_n as in (11), we get that $v_n(x) > M_n$ for all $x \in \mathbb{R}$, and therefore the limit distribution is Gumbel.

2. Now assume that (A2) holds. In this case, (26) can be violated. In fact,

$$\frac{M_n^\tau}{\log k_n} = \left(\frac{M_n}{(\log k_n)^{1/\tau}}\right)^\tau \lesssim \left(\frac{(k_n\varepsilon_n)^\gamma}{(\log k_n)^{1/\tau}}\right)^\tau = \frac{(k_n\varepsilon_n)^{\tau\gamma}}{\log k_n}$$

with some $\gamma \in (0, 1/\alpha)$. From (M2), it follows that the right-hand side is infinite if $\tau\gamma \geq \beta$, and has an unknown asymptotic behaviour otherwise. The lower bound is given by

$$\frac{M_n^\tau}{\log k_n} \gtrsim \frac{M_n^\tau}{(k_n\varepsilon_n)^\beta} = \left(\frac{M_n}{(k_n\varepsilon_n)^{\beta/\tau}}\right)^\tau,$$

where for $\beta \geq \tau\gamma$, the right-hand side tends to zero as $n \to \infty$, while otherwise the asymptotic behaviour is again unknown. In this case, we conclude that if M_n is such that (26) holds, the non-degenerate limit distribution exists and is in fact the Gumbel distribution.

3. Finally, in the case (A3),

$$\frac{M_n^\tau}{\log k_n} = \left(\frac{c^{1/\alpha} M_n}{(k_n\varepsilon_n)^{1/\alpha}}\right)^\tau \lesssim \left(\frac{c^{1/\alpha} M_n}{(k_n\varepsilon_n)^{\gamma}}\right)^\tau,$$

because $\gamma \in (0, 1/\alpha)$. Since by (26) there exists $\gamma \in (0, 1/\alpha)$ such that the right-hand side tends to zero as $n \to \infty$, we get under proper normalisation the Gumbel limit distribution.

Step 3. Case (M3).

1. If (A1) is satisfied, it is possible to obtain the Gumbel limit under the same choice of normalising sequence (11). Indeed, in this case

$$\frac{M_n^\tau}{\log k_n} = \frac{\breve{c}^\tau (k_n\varepsilon_n)^{\tau/\alpha}}{\log k_n} \lesssim \frac{(k_n\varepsilon_n)^\beta}{\log k_n},$$

because $\beta > \tau/\alpha$. Therefore, (26) follows from (A1), and we obtain the Gumbel distribution as a limit.

2. If (A2) holds, then the result turns out to depend on the asymptotic behaviour of $L(\cdot)$. Let us recall that $k_n\varepsilon_n M_n^{-\alpha}$ is equal to a constant.

   a) If $L(u) \to \infty$ as $u \to \infty$, we have that $k_n\varepsilon_n M_n^{-\alpha} L(M_n) \to \infty$ as $n \to \infty$. Thus, as was argued above, the non-degerated limit $H(x)$ exists if and only if (26) holds for all x and n large enough. However, for any $\beta \in (0, \tau/\alpha)$,

$$\frac{M_n^\tau}{\log k_n} = \frac{\breve{c}^\tau (k_n \varepsilon_n)^{\tau/\alpha}}{\log k_n} \gtrsim \frac{(k_n \varepsilon_n)^\beta}{\log k_n},$$

and therefore the assumption (26) is violated due to (A2). Thus, in this case there exists no non-degenerate limit distribution.

b) Now consider the case $L(u) \to \tilde{c}$ for some $\tilde{c} > 0$ as $u \to \infty$. Let us fix s_n, c_n in the form (12). The inequality $v_n(x) = s_n x > M_n$ is equivalent to $x > \breve{c}\tilde{c}^{-1/\alpha}$. Under this normalisation, we have

$$\lim_{n\to\infty} k_n e^{-\lambda v_n^\tau(x)} = 0,$$

and therefore

$$H(x) = \begin{cases} e^{\tilde{c}\breve{c}^{-\alpha} - x^{-\alpha}}, & x \in (0, \breve{c}\tilde{c}^{-1/\alpha}], \\ 1, & x > \breve{c}\tilde{c}^{-1/\alpha}. \end{cases}$$

c) If $L(u) \to 0$ as $u \to \infty$, one can take the norming constants as in the previous item and obtain the Fréchet limit distribution since $k_n \varepsilon_n M_n^{-\alpha} L(M_n) \to 0$ and $k_n \varepsilon_n s_n^{-\alpha} L(s_n) x^{-\alpha} \to x^{-\alpha}$ as $n \to \infty$. The last thing which is crucial here is to check that $v_n(x) \leq M_n$ for all $x \in \mathbb{R}$. This inequality follows from

$$s_n x = (k_n \varepsilon_n)^{1/\alpha} L^{1/\alpha}(s_n) x = \breve{c}^{-1} M_n L^{1/\alpha}(s_n) x \lesssim M_n.$$

3. Finally, let us consider the case (A3). As in the previous situations, the limit distribution depends on the asymptotic behaviour of $L(\cdot)$.

a) Let $L(u) \to \infty$ as $u \to \infty$. Since then $k_n \varepsilon_n M_n^{-\alpha} L(M_n) \to \infty$ as $n \to \infty$, we conclude that the non-degenerate limit exists only if (25) holds. In the considered case, (25) is equivalent to

$$\lambda^{1/\tau} \breve{c} c^{1/\alpha} < 1$$

The normalising sequence (11) again leads to the Gumbel limit distribution.

b) If $L(u) \to \tilde{c}$ for some $\tilde{c} > 0$ as $u \to \infty$, we have that $k_n \varepsilon_n M_n^{-\alpha} L(M_n) = \tilde{c}\breve{c}^{-\alpha}$.

- If $\lambda^{1/\tau}\breve{c}c^{1/\alpha} < 1$, a linear normalising sequence as in (11) leads to the Gumbel limit, since $v_n(x) > M_n$ for all $x \in \mathbb{R}$ and n large enough; see the previous item.

- If $\lambda^{1/\tau}\breve{c}c^{1/\alpha} > 1$, the choice (12) of normalising constants yields $s_n x > M_n$ for all $x > \breve{c}\tilde{c}^{-1/\alpha}$, while

$$\lim_{n\to\infty} k_n e^{-\lambda s_n^\tau x^\tau} = k_n^{1-(c\tilde{c})^{\tau/\alpha} x^\tau} = \begin{cases} \infty, & x \in (0, \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}), \\ 1, & x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ 0, & x > \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha} \end{cases}$$

  and

$$\lim_{n\to\infty} k_n \varepsilon_n s_n^{-\alpha} L(s_n) x^{-\alpha} = x^{-\alpha}.$$

  Therefore, we get

$$H(x) = \begin{cases} 0, & x \in (0, \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}), \\ \exp\left\{-1 + \tilde{c}\breve{c}^{-\alpha} - \lambda^{\alpha/\tau} c\tilde{c}\right\}, & x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \\ \exp\left\{\tilde{c}\breve{c}^{-\alpha} - x^{-\alpha}\right\}, & x \in (\lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}, \breve{c}\tilde{c}^{-1/\alpha}], \\ 1, & x > \breve{c}\tilde{c}^{-1/\alpha}. \end{cases}$$

  As we see, the limit distribution does not belong to the extreme value family, and has an atom at $x = \lambda^{-1/\tau}(c\tilde{c})^{-1/\alpha}$.

- If $\lambda^{1/\tau}\breve{c}c^{1/\alpha} = 1$, the choice (12) leads to the discrete limit distribution having a unique atom at $x = \breve{c}\tilde{c}^{-1/\alpha}$ with probability mass $1/e$.

c) Lastly, let $L(\cdot)$ be such that $L(u) \to 0$ as $u \to \infty$. Then the Gumbel limit can be obtained under the normalisation (11) since $k_n \varepsilon_n M_n^{-\alpha} L(M_n) \to 0$ as $n \to \infty$, see item (i,c) in Theorem 1. This observation completes the proof.

□

Chapter 3. Empirical part

3.1. Simulation study

The aim of the current section is to illustrate the dependence of limit distribution for maxima in the model (23) on the rates of the mixing parameter ε_n and of the truncation level M_n. For this purpose we consider four triangular arrays (23) with $k_n = n$ having all permanent parameters the same, namely, $\vec{\theta} = (1, 1, 1.5, 1)$ and $m = 0.1$. The sequences ε_n and M_n are chosen to satisfy the following pairs of conditions: (A1)-(M1), (A2)-(M1), (A1)-(M2) and (A2)-(M2). The exact form of mixing and truncation parameters are presented in Table 2.

As previously, the primary separation is made due to the rates of $\log k_n$ and $k_n\varepsilon_n$: we fix $\varepsilon_n = n^{-1}(\log n)$ and $\varepsilon_n = n^{-1}(\log n)^2$, which imply conditions (A1) and (A2), respectively. Next, the models are divided according to the rate of growth of M_n in the form $M_n = (\log(n+1))^a$ with $a = 1/2, 1, 2$. Recall that from Theorem 2, it follows that the limit distribution is Gumbel for the pairs (A1)-(M1), (A1)-(M2) and (A2)-(M2) (note that for the last two cases $k_n \gtrsim e^{\lambda M_n^\tau}$ under our choice), and Fréchet for the pair (A2)-(M1).

For each case we simulate 1000 samples of length 1000, and find the maximal value of each sample. The goodness-of-fit of the limit distributions of the maximal values suggested by Theorem 2 is tested by the Kolmogorov-Smirnov criterion. Figure 2 depicts the kernel density estimates of the densities of normalised maxima in each case superimposed with the limit distributions implied by Theorem 2. It can be seen that for all groups the density estimates are quite close to the theoretical densities, and the Kolmogorov-Smirnov test does not reject the null of the corresponding theoretical distributions (corresponding p-values are given on the same figure).

Table 2

The values of ε_n and M_n chosen for the numerical study

	(A1)	(A2)
(M1)	$\varepsilon_n = n^{-1}\log n$, $M_n = \log(n+1)$	$\varepsilon_n = n^{-1}(\log n)^2$, $M_n = (\log(n+1))^2$
(M2)	$\varepsilon_n = n^{-1}\log n$, $M_n = \sqrt{\log(n+1)}$	$\varepsilon_n = n^{-1}(\log n)^2$, $M_n = \sqrt{\log(n+1)}$

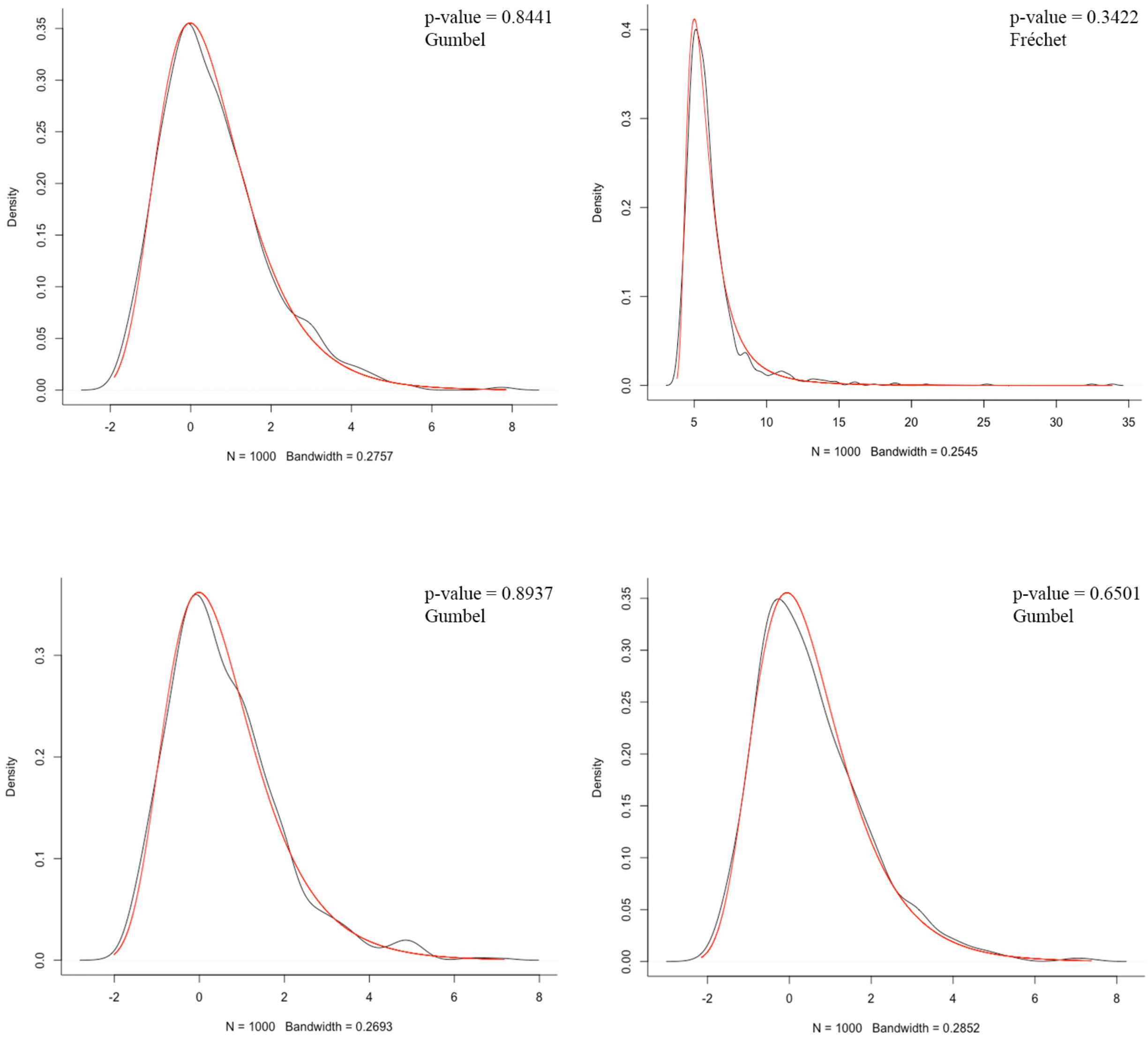


Figure 2: Densities of sample maxima for groups (A1)-(M1) (top left), (A2)-(M1) (top right), (A1)-(M2) (bottom left) and (A2)-(M2) (bottom right) superimposed with the theoretical limit distributions suggested by Theorem 2
Source: computations performed by author

3.2. Modelling the log-returns of BMW shares

In this section we turn to the application of the model (1) to the analysis of price variations. For this purpose, we consider hourly logarithmic returns of BMW shares in 2019. Following Malevergne et al. (2005), we analyse positive and negative returns separately. The sample sizes are equal to 1130 and 1062, respectively. The plots for positive and negative log returns are presented in the first plot in Figure 3. In what follows, we assume that the log-returns are jointly independent. This assumption was checked by the chi-squared test resulting in p-values 0.234 and 0.223 for positive and negative returns, respectively.

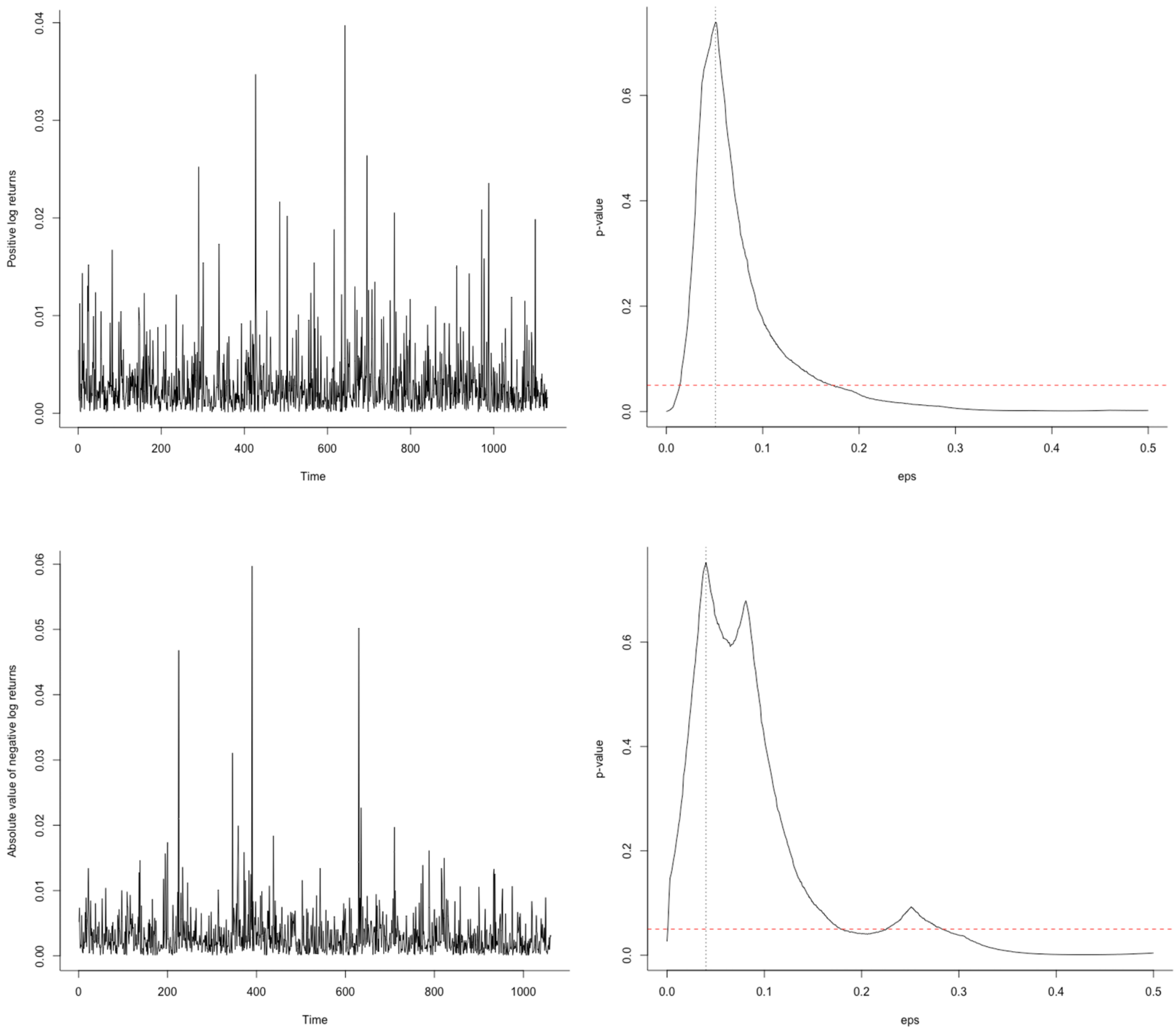


Figure 3: First (second) row: the plot of data; p-values for Weibull fit for positive (absolute negative) log returns
Source: computations performed by author

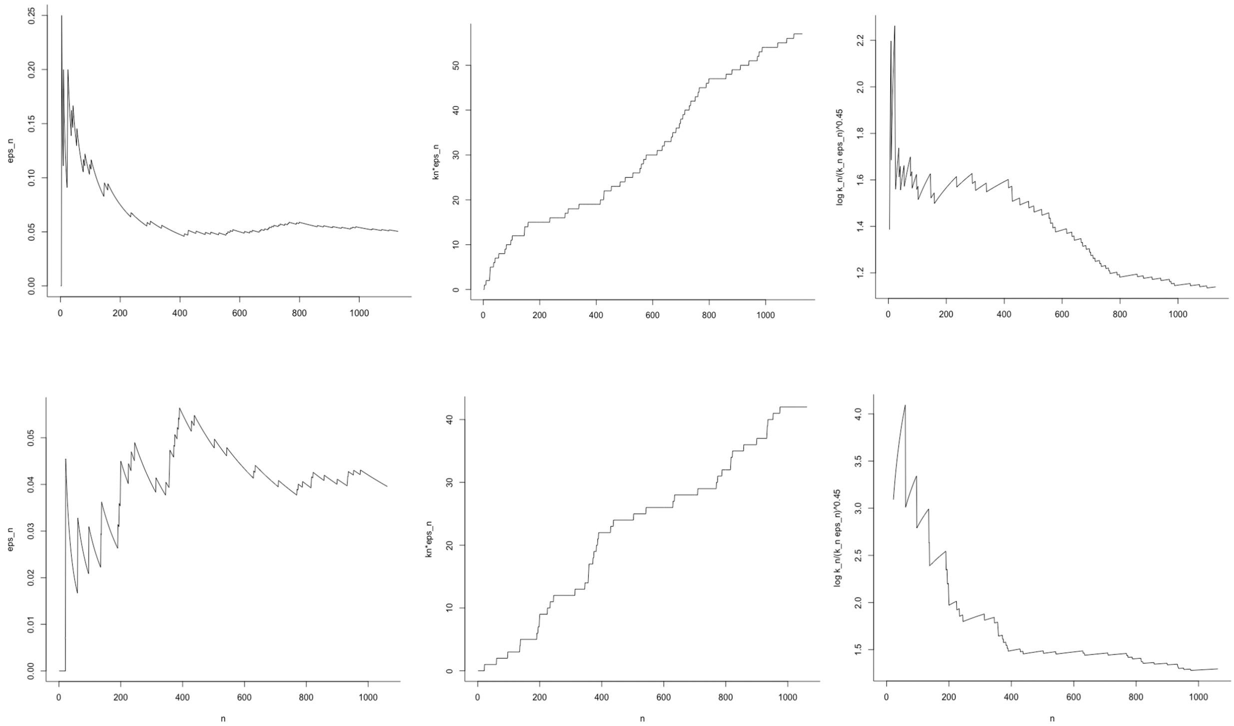
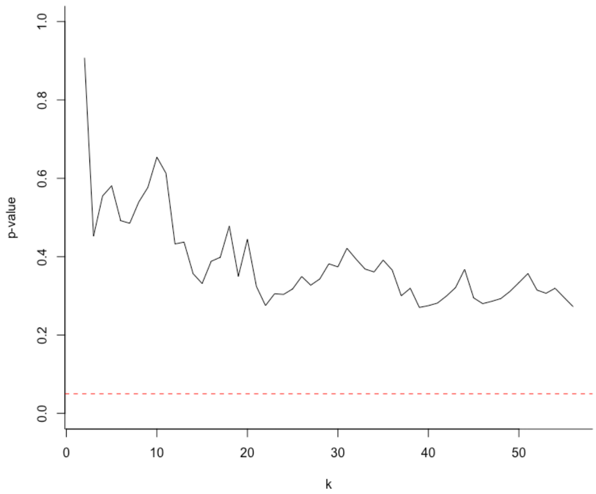


Figure 4: $\hat{\varepsilon}_n$ against n (left), $k_n\hat{\varepsilon}_n$ against n (middle), $\log k_n(k_n\varepsilon_n)^{-0.45}$ against n (right) for positive log returns of BMW (first row) and absolute values of negative log returns of BMW (second row)
Source: computations performed by author

The analysis consists of 4 steps. Below we denote the positive log-returns by $X_1, ..., X_{1130}$ and the negative log-returns by $Y_1, ..., Y_{1062}$.

1. **Separation of components.** For each $n = 1...1130$, the sample $X_1, ..., X_n$ is divided into 2 parts corresponding to the first and the second components in (20), where the slowly varying function L is equal to a constant. For such partition we assign all observations except the $\lfloor 1130 \cdot \varepsilon_{1130} \rfloor$ greatest order statistics of the whole sample to the first component and test the goodness-of-fit for Weibull distribution by the Kolmogorov-Smirnov criterion. The values of ε_{1130} are taken on a grid from 0 to 0.5 with a step of 0.001. From the second plot in Figure 3 it can be seen that there is an evident peak in p-values. The corresponding value $\hat{\varepsilon}_{1130}$ is considered to be an estimate of the mixing parameter for $n = 1130$, and the $\lfloor 1130 \cdot \hat{\varepsilon}_{1130} \rfloor$ upper order statistics are assumed to come from the heavy-tailed part. For all $n = 1...1129$ the parameter ε_n is then estimated as the proportion of elements corresponding to the second component. The same procedure is



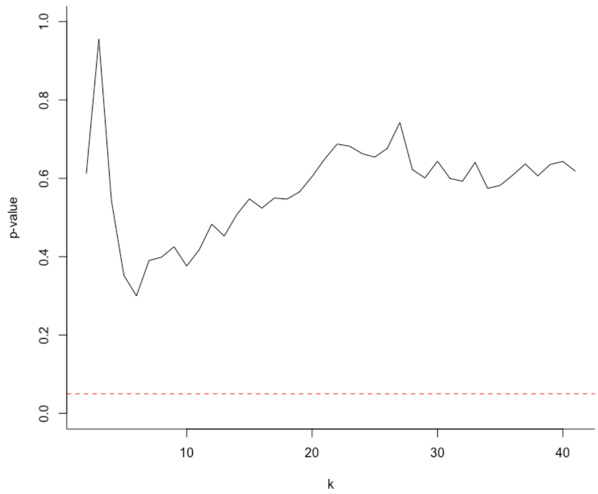


Figure 5: The p-values of the Aban's test for positive (left) and absolute values of negative (right) log returns
Source: computations performed by author

applied for each $n = 1..1062$ to the sample $Y_1, ..., Y_n$.

The results are illustrated by Figure 4. The first plot in two rows indicates that in both cases ε_n declines with n, though in case of negative log returns the decrease is not so evident. From the second plot one can see that $k_n\varepsilon_n$ with $k_n = n$ appear to tend to infinity. Therefore, as suggested by Theorems 1 and 2, we examine the asymptotic behaviour of the ratio $\log k_n/(k_n\varepsilon_n)^\beta$ for different values of $\beta > 0$. For both positive and absolute values of negative log returns we get that for $\beta = 0.45$ this ratio decreases rapidly.

2. **Model selection.** Based on the partition obtained on the previous step, the decision between truncated and non-truncated distributions for the second component is made based on the test proposed by Aban (2006). From Figure 5 one observes that the null hypothesis of non-truncated law is not rejected for both positive and negative log returns since the p-values are significantly larger than 0.05. Thus, it can be concluded that the model (5) is more appropriate for the considered data. The Kolmogorov-Smirnov test does not reject the null of Pareto distribution for the observations assigned to the second component with p-values 0.971 and 0.925 for positive and

Table 3

Estimated values of the parameters of mixture distribution for positive and absolute values of negative log returns of BMW, 2019

Estimates	$\hat{\lambda}$	$\hat{\tau}$	$\hat{\alpha}$	$\hat{m}$	$\hat{\varepsilon}$
Positive log returns	0.003	1.278	2.649	0.009	0.051
Abs. negative log returns	0.003	1.246	2.573	0.01	0.04

Source: computations performed by author

Table 4

Empirical quantiles of positive log returns of BMW, 2019, and the estimated confidence intervals

Quantiles$\cdot 10^3$	10%	20%	30%	40%	50%	60%	70%	80%	90%
Lower CI	0.416	0.783	1.154	1.548	1.982	2.509	3.164	4.013	5.679
Estimate	0.452	0.825	1.232	1.64	2.085	2.611	3.289	4.183	6.383
Upper CI	0.537	0.946	1.33	1.747	2.223	2.79	3.498	4.493	6.657

Source: computations performed by author

negative log returns, respectively. It should be noted that in both cases the Pareto distribution does not fit the whole sample, since the p-values are smaller than 10^{-16}.

3. **Estimation of parameters.** The parameters of the first and second components are estimated by the maximum-likelihood approach. The estimated values are presented in Table 3. Since $\hat{\tau}/\alpha$ is equal to 0.483 for positive log returns and 0.484 for absolute values of negative, we conclude that the assumption (A2) is fulfilled with $\beta = 0.45$, and therefore the limit distribution for maxima is the Fréchet distribution, see item (ii) in Theorem 1.

   It is worth mentioning that for both positive and absolute values of negative log returns we get $\hat{\alpha} > 2$, which is completely coherent with general empirical results for financial returns and addresses the common critique against models with infinite variance, see Cont (2001).

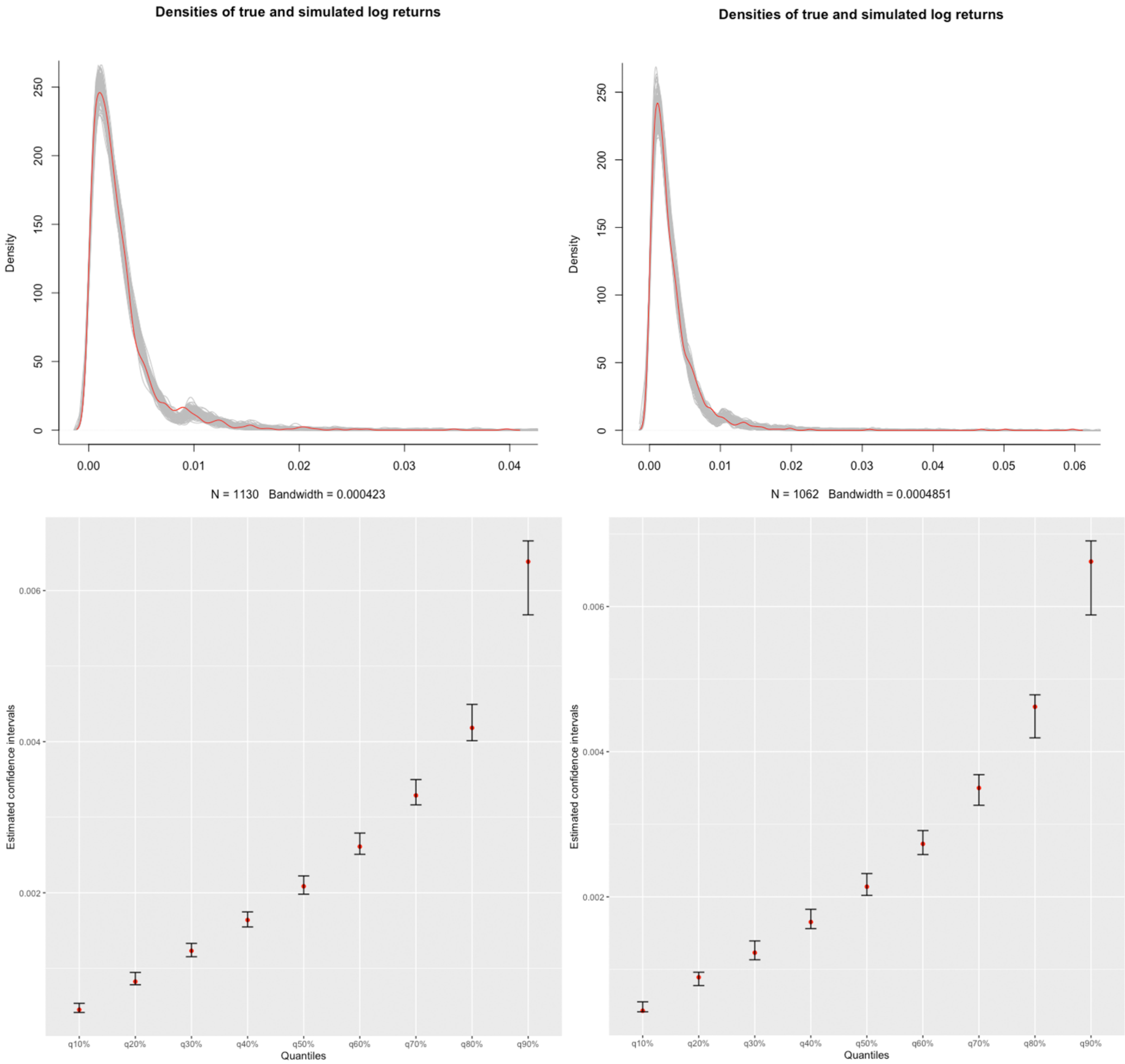


Figure 6: Top left (right): real (red) and simulated (grey) density of positive (negative) log returns; bottom left (right): empirical quantiles of positive (negative) log returns and the corresponding confidence intervals
Source: computations performed by author

4. **Validation of the model.** Figure 6 depicts the true density of positive (top left) and absolute values of negative (top right) log returns superimposed with densities of 100 simulations from the mixture (5) with the corresponding parameter estimates. The constructed model is also verified by the empirical confidence intervals for the sample quantiles based on 100 simulations. The results are given in Tables 4 and 5 and illustrated by Figure 6. These intervals are reasonably small and contain all true values

Table 5

Empirical quantiles of absolute values of negative log returns of BMW, 2019, and the estimated confidence intervals

Quantiles $\cdot 10^3$	10%	20%	30%	40%	50%	60%	70%	80%	90%
Lower CI	0.417	0.778	1.134	1.562	2.021	2.583	3.262	4.191	5.884
Estimate	0.431	0.892	1.231	1.654	2.14	2.73	3.501	4.619	6.62
Upper CI	0.554	0.962	1.393	1.829	2.321	2.914	3.684	4.784	6.905

Source: computations performed by author

of quantiles. Next, from the Kolmogorov-Smirnov test we conclude that the null of distribution (5) is not rejected with p-values 0.769 and 0.775 for positive and negative log returns, respectively. Finally, we arrive at the outcome that the model (5) is appropriate both for positive and absolute values of negative log returns of BMW at the considered time scale.

Conclusion

To sum up, we have obtained the following results. First, we develop a model for describing the case of heavy-tailed impurity, when the majority of observations comes from a light-tailed distribution, while some small part is generated by a heavy-tailed law. For this purpose, we use mixture models with varying parameters, which can be viewed as triangular arrays. We analyse the asymptotic behavior of maxima in the two models of this kind, and show that the set of possible limit distributions is much more diverse than in the classical setting. In particular, in one of the considered models it includes 6 various distributions, 4 of which are discontinuous and not present in the classical theory. Next, we analyse the effect of different types of truncation on the asymptotic behavior of maxima. Our findings turn out to be related, though not completely the same, to the ones from Chakrabarty and Samorodnisky (2012). In addition, unlike that paper, we study in detail the case of intermediate truncation regime. Finally, we develop a procedure for applying the proposed model to the asset price modelling, and demonstrate its efficiency for the analysis of maximal values of financial returns.

The conducted study also points to the directions of further research. First of all, there is an evident need for more general results for the extreme value analysis for triangular arrays. As illustrated by our examples, even when the asymptotic behaviour of maxima can be analysed by direct calculations, they might be very extensive. In addition, softening the assumption of independence would bring significant contribution to the existing theory. While certain results for the case of dependence already exist, they typically rely on the assumption of «approximate independence», when the maxima behave similarly to those of the corresponding independent sequence. In particular, Leadbetter et al. (1974) have introduced the mixing and anti-clustering conditions assuring such a behaviour and proved the analogue of the Fisher-Tippett-Gnedenko theorem for sequences of stationary identically distributed random variables. Later, it was shown by Hüsler (1986) that the same holds also in case of non-stationarity. In the context of triangular arrays, the Leadbetter conditions are often imposed to ensure «approximate independence» in rows; see, e.g., Dkengne et al. (2016). However, as was mentioned earlier, these conditions are very difficult to verify, and there exist few examples when this can be made directly. At the same time, the results concerning limit distributions for extremes under dependence are highly important, since it is commonly observed in practical instances. Therefore, more general characterisation of the asymptotic behaviour of row-wise maxima of triangular arrays, as well as extending the existing theory to the case of dependence, would open up even more applications for the extreme value theory, making it an essential tool for the analysis of real-world problems.

Bibliography

[1] Aban, I.B. and Meerschaert, M.M. and Panorska, A.K. (2006). Parameter estimation for the truncated Pareto distribution. *Journal of the American Statistical Association*, 101(473):270–277.

[2] Anderson, C., Coles, S., and Hüsler, J. (1997). Maxima of Poisson-like variables and related triangular arrays. *The Annals of Applied Probability*, pages 953–971.

[3] Beirlant, J., Alves, I. and Gomes, I. (2016). Tail fitting for truncated and non-truncated Pareto-type distributions. *Extremes*, 19(3):429–462.

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

[5] Chakrabarty, A. and Samorodnitsky, G. (2012). Understanding heavy tails in a bounded world or, is a truncated heavy tail heavy or not? *Stochastic models*, 28(1):109–143.

[6] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.

[7] Dkengne, P.S. and Eckert, N. and Naveau, P. (2016). A limiting distribution for maxima of discrete stationary triangular arrays with an application to risk due to avalanches. *Extremes*, 19(1):25–40.

[8] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling extremal events for insurance and finance.* Springer.

[9] Fisher, R.A. and Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press.

[10] Freitas, A.V. and Hüsler, J. (2003). Condition for the convergence of maxima of random triangular arrays. *Extremes*, 6(4):381–394.

[11] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, pages 423–453.

[12] Grabchak, M. and Molchanov, S. (2015). Limit theorems and phase transitions for two models of summation of independent identically distributed random variables with a parameter. *Theory Prob. Appl.*, 59(2):222–243.

[13] Hüsler, J. (1986). Extreme values of non-stationary random sequences. *Journal of applied probability*, pages 937–950.

[14] Laherrere, J. and Sornette, D. (1998). Stretched exponential distributions in nature and economy:"fat tails" with characteristic scales. *The European Physical Journal B-Condensed Matter and Complex Systems*, 2(4):525–539.

[15] Leadbetter, M.R. (1974). On extreme values in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 28(4):289–303.

[16] Malevergne, Y., Pisarenko, V. and Sornette, D. (2005). Empirical distributions of stock returns: between the stretched exponential and the power law? *Quantitative Finance*, 5(4):379–401.

[17] Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, (36).

[18] Meerschaert, M. and Scheffler, H.-P. (2001). *Limit distributions for sums of independent random vectors: Heavy tails in theory and practice*, volume 321. John Wiley & Sons.

[19] Mikosch, T. (1999). Regular variation, subexponentiality and their applications in probability theory. Report Eurandom; Vol. 99013.

[20] Mohan, N.R. and Ravi, S. (1993). Max domains of attraction of univariate and multivariate p-max stable laws. *Theory of Probability & Its Applications*, 37(4):632–643.

[21] Panov, V. (2017). Limit theorems for sums of random variables with mixture distribution. *Statistics and Probability Letters*, 129:379 – 386.

[22] Pantcheva, E. (1985). Limit theorems for extreme order statistics under nonlinear normalization. In *Stability problems for stochastic models*, pages 284–309. Springer.

[23] Petrov, V. (2012). *Sums of independent random variables*, volume 82. Springer Science & Business Media.

[24] Resnick, S. (2013). *Extreme values, regular variation and point processes*. Springer.

[25] Finam (2019). German shares, BMW AG. URL: https://www.finam.ru. Retrieved on 05.11.20.