Frequency Analysis of Heavy Tail Phenomena

Yuwei Zhao[1,2] and Thomas Mikosch[1]

[1]Department of Mathematical Sciences, University of Copenhagen
[2]Contact Author: Yuwei Zhao, e-mail: zhao@math.ku.dk.

Abstract

The extremogram is an asymptotic correlogram for extreme events constructed from a regularly varying stationary sequence. Correspondingly, a frequency domain analog of the extremogram is introduced and its estimator is called *ex-periodogram*. Several asymptotic properties of the *ex-periodogram* are discussed in this paper. Furthermore, a functional central limit theorem for the integrated *ex-periodogram* is provided which can be used for constructing goodness-of-fit tests.
Key Words: Heavy tail phenomena, frequency analysis, extreme value theory, time series analysis.

1 Introduction

Heavy tail phenomena have been observed in various fields such as telecommunication networks (Internet), insurance, finance, seismology, to name a few. It is typical for these phenomena that they are rare and serially dependent. The tools of classical time series analysis (autocorrelation function, spectral density) are not suitable for describing extreme events and their dependence structure. However, in our approach we adapt the autocorrelation function from the time domain and the spectral density from the corresponding frequency domain of time series analysis: we apply them to indicator functions of the rare events of interest. Thus we deal with stationary sequences of indicator functions whose distributions change in dependence on a sufficiently high threshold. Therefore the classical results of time series analysis are not directly applicable: one deals with a triangular array of stationary sequences whose marginal distribution changes with the threshold. First, we introduce the extremogram as analog of the autocorrelation function of a stationary sequence and its sample version based on counts of rare events in a stationary sequence. Second, we define a corresponding periodogram, called *ex-periodogram*, as an estimator of the spectral density defined via the extremogram. The ex-periodogram shares various of the classical properties of the periodogram of a stationary weakly dependent sequence: the ex-periodogram ordinates at distinct (Fourier) frequencies are asymptotically independent and exponentially distributed and smoothed versions are consistent estimators of the spectral density. Having established some basic asymptotic theory for the ex-periodogram our next goal is to apply the theory for estimating the spectral distribution function, parameter estimation of suitable time series models and goodness-of-fit tests. In this context, we study different versions of the integrated ex-periodogram. It is our objective to derive results parallel to the classical theory for the integrated periodogram of a stationary sequence which can be interpreted as a spectral empirical distribution. We show that the integrated ex-periodogram satisfies a functional central limit theorem under mild conditions, which ensure that goodness-of-fit tests, such as the Grenander-Rosenblatt and the Cramér-von Mises tests, can be constructed for the ex-periodogram. This means that we can determine whether a model is suitable solely based on the behavior of extreme events.

2 Results

We assume that the sequence (X_t) is a d-dimensional strictly stationary sequence. Additionally, we also assume that the sequence (X_t) is regularly varying, which means that the finite-dimensional vector $(X_0, \ldots, X_h)$ is regularly varying for any finite number $h \geq 0$. The definition of multivariate regular variation can be founded in the literature on extreme value theory, such as Resnick (1987, 2007). Then it follows that the *extremogram* at lag $h \geq 0$

$$\rho(h) = \lim_{x \to \infty} P(x^{-1} X_h \in A \mid x^{-1} X_0 \in A)$$

exists where A is a smooth Borel set bounded away from zero.

Based on the extremogram, we define the corresponding spectral density

$$f(\lambda) = 1 + 2 \sum_{h=1}^{\infty} \rho(h) \cos(h\lambda)\,.$$

The *ex-periodogram* is an estimator of the spectral density. It is defined as

$$I_n(\lambda) = 1 + 2 \sum_{h=1}^{n-1} \widetilde{\rho}_n(h) \cos(h\lambda)\,.$$

Here $\widetilde{\rho}_n(h) = \sum_{h=1}^{n-h} I_t I_{t+h} / \sum_{h=1}^{n} I_t$ is the empirical extremogram where $I_t = I_{\{a_m^{-1} X_t \in A\}}$ and $P(|X_0| > a_m) \sim 1/m$ and $m = o(n) \to \infty$ as $n \to \infty$.

The mixing/dependence conditions which are used to prove the following theorems are listed below.

(M) The sequence (X_t) is strongly mixing with rate function (ξ_t). There exist $m = m_n \to \infty$ and $r_n \to \infty$ such that $m_n/n \to 0$ and $r_n/m_n \to 0$ and

$$\lim_{n\to\infty} m_n \sum_{h=r_n}^{\infty} \xi_h = 0\,, \tag{2.1}$$

and for all $\epsilon > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \sum_{h=k}^{r_n} P(|X_h| > \epsilon\, a_m \mid |X_0| > \epsilon\, a_m) = 0\,. \tag{2.2}$$

(M1) The sequences (m_n), (r_n), $k_n = [n/m_n]$ from (M) satisfy the growth conditions $k_n \xi_{r_n} \to 0$, and $m_n = o(n^{1/3})$.

(M2) The sequences (m_n), (r_n) from (M) satisfy the growth conditions

$$m_n^2 n \sum_{h=r_n+1}^{n} \xi_h \to 0, \quad m_n r_n^3/n \to 0\,.$$

The conditions *(M)* and *(M1)* in Theorem 2.1 are introduced in Davis and Mikosch (2009). *(M2)* is used in Mikosch and Zhao (2012).

The ex-periodogram ordinates are asymptotically independent and exponentially distributed (see Mikosch and Zhao (2012)):

Theorem 2.1. *Consider a strictly stationary $\mathbb{R}^d$-valued sequence (X_t) which is regularly varying with index $\alpha > 0$. Let $A \subset \overline{\mathbb{R}}_0^d$ be bounded away from $\{0\}$. Assume that the conditions* (M), (M1) *and $\sum_{h\ge 1} \rho(h) < \infty$ hold. Let $(E_i)_{i=1,\dots,N}$ be a sequence of iid standard exponential random variables.*

1. *Consider any fixed frequencies $0 < \lambda_1 < \cdots < \lambda_N < \pi$ for some $N \ge 1$. Then,*

$$\big(I_n(\lambda_i)\big)_{i=1,\dots,N} \stackrel{d}{\to} \big(f(\lambda_i)E_i\big)_{i=1,\dots,N}\,, \quad n \to \infty\,.$$

2. *Consider any distinct Fourier frequencies $\omega_i(n) \to \lambda_i \in \Pi = (0,\pi)$ as $n \to \infty$, $i = 1,\dots,N$, where $\omega_i(n) = 2j\pi/n$ for some integer $1 \le j < n$. The limits λ_i do not have to be distinct. Then the following relations hold:*

$$\big(I_n(\omega_i(n))\big)_{i=1,\dots,N} \stackrel{d}{\to} \big(f(\lambda_i)E_i\big)_{i=1,\dots,N}\,, \quad n \to \infty\,,$$

To estimate the spectral density, we start by introducing the smoothed ex-periodogram. For a fixed frequency $\lambda \in (0,\pi)$ define

$$\lambda_0 = \min\{2\pi j/n : 2\pi j/n \geq \lambda\}\,, \quad \text{and} \quad \lambda_j = \lambda_0 + 2\pi j/n\,, \quad |j| \leq s_n\,.$$

Here we suppress the dependence of λ_j on n. In what follows, we will assume that $s_n \to \infty$ and $s_n/n \to 0$ as $n \to \infty$. For a given set $A \subset \overline{\mathbb{R}}_0^d$ bounded away from zero and any non-negative weight function $w = (w_n(j))_{|j|\leq s_n}$ satisfying the conditions

$$\textstyle\sum_{|j|\leq s_n} w_n(j) = 1 \quad \text{and} \quad \sum_{|j|\leq s_n} w_n^2(j) \to 0 \text{ as } n \to \infty\,, \tag{2.3}$$

we introduce the *smoothed periodogram*

$$\widetilde{f}_n(\lambda) = \sum_{|j|\leq s_n} w_n(j)\, I_n(\lambda_j)\,, \quad \lambda \in \Pi\,.$$

Theorem 2.2. *Assume the conditions of Theorem 2.1,* (2.3) *on the weight function* w *and* (M2)*. Then for every fixed frequency* $\lambda \in (0,\pi)$*, as* $n \to \infty$*,*

$$\widetilde{f}_n(\lambda) \stackrel{P}{\to} \rho(0)\, f(\lambda)\,.$$

Moreover, we define the empirical spectral distribution function with weight function g by

$$J_{nA}(x) = \int_{-\pi}^{x} I_{nA}(\lambda)\, g(\lambda)\, d\lambda\,, \quad x \in \Pi\,,$$

where g is a non-negative 2π-periodic continuous.

Theorem 2.3. *Assume that the sequence* (X_t) *satisfies the assumptions in Theorem 2.1. If the conditions* (M), (M1) *and* $\sum_{l=1}^{\infty} \rho(l) < \infty$ *hold then in* $\mathbb{C}(\Pi)$*, the space of continuous functions on* Π *equipped with the uniform topology,*

$$(n/m)^{0.5}(J_{nA} - EJ_{nA}) \quad \stackrel{d}{\to} \quad G\,, \tag{2.4}$$

where the limit process is given by the infinite series

$$G = \psi_0 Z_0 + 2\sum_{h=1}^{\infty} \psi_h\, Z_h\,,$$

where $\psi_h(x) = \int_{-\pi}^{x} g(\lambda)\cos(h\lambda)\, d\lambda$ *for* $h \geq 0$ *and* (Z_h) *is a mean zero Gaussian sequence whose structure is determined by the asymptotic covariance matrix* $\Sigma_h = (\sigma_{ij})_{i,j=0,\dots,h}$ *with entries*

$$\sigma_{ij} \quad = \quad \gamma_A(i,j) + \sum_{l=1}^{\infty} \big[\gamma_A(i,l,l+j) + \gamma_A(j,l,l+i)\big]\,, \quad i,j = 0,\dots,h\,,$$

and for $u, s, t \geq 0$,

$$\gamma_A(u,s,t) \;=\; \lim_{m\to\infty} m\, P(a_m^{-1}X_0 \in A\,, a_m^{-1}X_u \in A, a_m^{-1}X_s \in A, a_m^{-1}X_t \in A)\,,$$

with the convention that $\gamma_A(u,t) = \gamma_A(u,u,t)$.

Based on Theorem 2.3, we can construct goodness-of-fit test statistics, for example,

1. Grenander-Rosenblatt test:

$$(n/m)^{0.5} \sup_{x\in\Pi} |J_{nA}(x) - EJ_{nA}(x)| \;\stackrel{d}{\to}\; \sup_{x\in\Pi} |G(x)|\,.$$

2. ω^2-statistic or Cramér-von Mises test:

$$(n/m) \int_{x\in\Pi} (J_{nA}(x) - EJ_{nA}(x))^2\, dx \stackrel{d}{\to} \int_{x\in\Pi} G^2(x)\, dx\,.$$

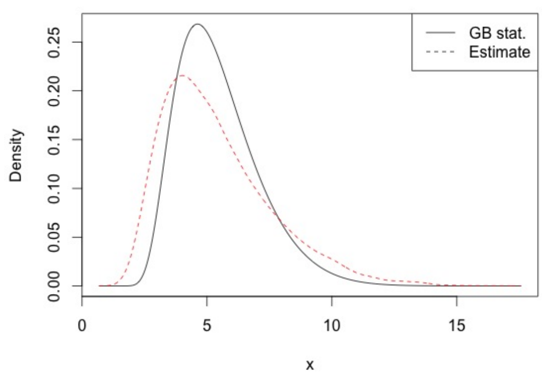

Figure 1: The densities of the Grenander-Rosenblatt statistic and the corresponding estimates from simulated iid t-distributed sequences with coefficient $\alpha = 3$. The threshold is 95%. The size of the sequence is $10,000$ and the number of the sequences is $10,000$.

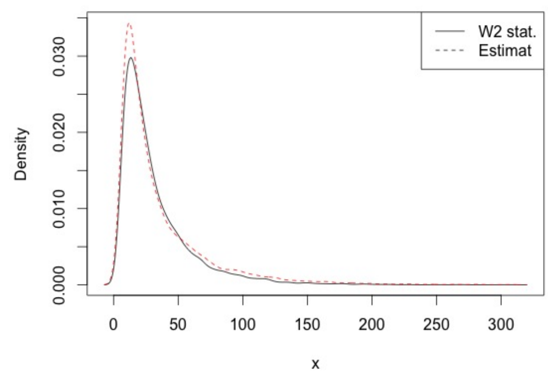

Figure 2: The densities of the ω^2-statistic and the corresponding estimates from simulated iid t-distributed sequences with coefficient $\alpha = 3$. The threshold is 95%. The size of the sequence is $10,000$ and the number of the sequences is $10,000$.

References

[1] Davis, R.A. and Mikosch, T. (2009) The extremogram: a correlogram for extreme events. *Bernoulli* **15**, 977–1009.

[2] Davis, R.A., Mikosch, T. and Cribben, I. (2012) Towards estimating extremal serial dependence via the boostrapped extremogram. *J. Econometrics* **170**, 142–152.

[3] Davis, R.A., Mikosch, T. and Zhao, Y. (2013) Measures of serial extremal dependence and their estimation. *Stoch. Proc. Appl.*, to appear.

[4] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin. Springer, New York.

[5] Mikosch, T. and Zhao, Y. (2012) A Fourier analysis of extremal events. *Bernoulli*, to appear.

[6] Mikosch, T. and Zhao, Y. (2013) Goodness-of-fit tests for extremograms. Working paper.

[7] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[8] Resnick, S.I. (2006) *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.*