Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v9n1p68

The Lawrence-Lewis Pareto process: an extremal approach

By Ferreira

Published: 26 April 2016

Electronic Journal of Applied Statistical Analysis
Vol. 09, Issue 01, April 2016, 68-82
DOI: 10.1285/i20705948v9n1p68

The Lawrence-Lewis Pareto process: an extremal approach

Marta Ferreira*

Center of Mathematics of University of Minho, Center for Computational and Stochastic Mathematics and Center of Statistics and Applications of University of Lisbon Portugal

Published: 26 April 2016

Pareto processes are more suitable for time series with heavy tailed marginals than the classical gaussian. Here we consider the Lawrence-Lewis Pareto process. In particular, we analyze long-range and local dependence and compute some extremal measures. This will provide us some more diagnostic tools and new estimators for the autoregressive parameter of the process. Based on a simulation study we will see that the new methods may be good alternatives in what concerns robustness.

keywords: extreme value theory, autoregressive processes, extremal index, tail dependence.

1 Introduction

Time series presenting large peaks are potentially well modeled by ARMA with heavy tailed innovations. When replacing the addition operator of the ARMA processes by geometric multiplication or minimization/maximization, we arrive at simpler models to work on under an extremal approach. In Pareto families of distributions we find closeness under geometric multiplication or minimization, as well as, Fréchet ones may be closed under maximization. These properties are the basis on the construction of, respectively, autoregressive Pareto and max-autoregressive MARMA processes. Applications and properties of MARMA processes are largely addressed in literature; Davis and Resnick (1989), Alpuim (1989), Alpuim and Athayde (1990), Lebedev (2008), Ferreira and Canto e Castro (2010), Ferreira (2012a), Ferreira (2012b), Ferreira and Ferreira

*Corresponding author: msferreira@math.uminho.pt.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

(2013) and references therein. The Pareto processes are not so well-known, in particular, in extreme values theory. A survey on these processes can be found in Arnold (2001); see also Yeh et al. (1988), Ferreira (2012b) and Carcea and Serfling (2015).

In this paper we consider the classical Lawrence-Lewis Pareto process given in Arnold (2001), whose marginals are distributed as classical Pareto, i.e, the distribution function (df) F is of the form:

$$F(x) = 1 - (x/\sigma)^{-\alpha},$$

with $\alpha > 0$ and $x > \sigma > 0$, and denoted Pareto(I)(σ, α). Reference (Arnold, 1985) provides a broad survey on the Pareto distribution and their generalizations. Our analysis is in the scope of extreme values. More precisely, we will see that it entails clustering of high levels, a typical feature in applications concerning risk analysis. The usual correlation misleads the dependence structure within a heavy tailed process, once linear dependence can not suitably accounts for the whole series, including the outlying observations. Moreover, it is not always defined since second moments may not exist. Therefore, a tail dependence measure which, roughly stated, describes the limiting proportion that one marginal exceeds a large value given that the other marginal has already exceeded it too, is more appropriate. We will derive the tail dependence coefficient as given in Joe (1997), applied to random pairs of the process which are lag$-m$ apart. These topics are developed in Section 2. Estimators for the autoregressive parameter of the process will be derived trough the association with the fluctuation probabilities derived in Arnold (2001) and the measures of clustering and tail dependence presented here. A simulation study is conducted to compare the methodologies. In order to evaluate the sensitivity of the estimators to the model assumptions, we also compare their performances for processes $\{X_n\}_{n\geq 1}$ perturbed by the presence of a white noise $\{Z_n\}_{n\geq 1}$, i.e., where each marginal of the "noisy" process is given by $X_n + \zeta Z_n$, $\zeta > 0$. This will be addressed in Section 3.

2 Main Results

The Lawrence-Lewis Pareto(I) process, in short LLP(I)(1), is given by

$$X_n = X_{n-1}^{1-U_n} \sigma^{U_n} \epsilon_n^p,\ n = 1, 2, \ldots, \tag{1}$$

where $\{U_n\}_{n\geq 1}$ is an independent and identically distributed (iid) sequence of random variables (rv's) Bernoulli(p), and the innovations sequence $\{\epsilon_n\}_{n\geq 1}$ is an iid sequence of rv's Pareto(I)$(1, \alpha)$. By considering X_0 ~Pareto(I)(σ, α), then X_n ~Pareto(I)(σ, α) for $n \geq 1$ and thus the process is stationary. Observe that larger values of p correspond to larger probability on the occurrence of the innovations an thus less dependence within the process. In the sequel we denote F_X the marginal df of the process and F_ϵ the marginal df of the innovations sequence. Sample paths simulated from the LLP(I)(1) process with $\sigma = \alpha = 1$ and $p = 0.25, 0.5, 0.75$ are displayed in Figure 1.

The logarithm of this process leads to the NEAR(1) processes of Lawrance and Lewis (1981) where the marginals are exponential. The LLP(I)(1) process is little known in

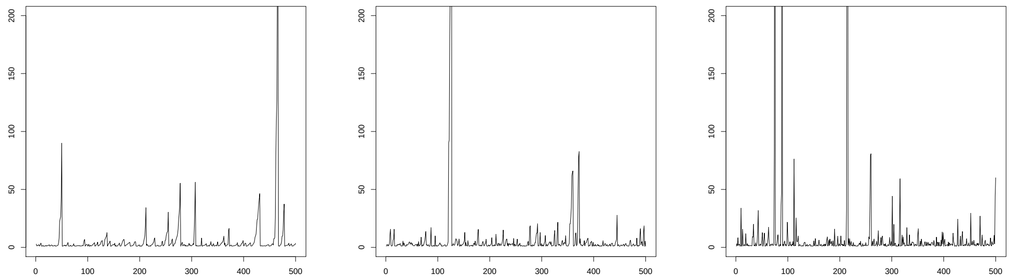

Figure 1: Simulated sample paths of LLP(I)(1) processes with marginals Pareto(I)$(1,1)$ for $p=0.25$ (left), $p=0.5$ (middle) and $p=0.75$ (right).

literature. It was introduced in the monograph of Arnold (2001), where some properties were presented, namely the fluctuation probability and the autocorrelation function (only defined for $\alpha>2$), as will be seen below.

The fluctuation probability, f_1, is given by

$$f_1 := P(X_{n-1} < X_n) = \frac{1}{1+p} \tag{2}$$

and may be used in a modeling framework.

In order to deal with probabilities of events involving multiple marginals, we derive the transition probability function (tpf):

$$Q(x,]\sigma,y]) := P(X_2 \leq y | X_1 = x)$$

$$= \begin{cases} F_\epsilon\left(\left(\dfrac{y}{\sigma}\right)^{1/p}\right)p + F_\epsilon\left(\left(\dfrac{y}{x}\right)^{1/p}\right)(1-p) & , y > x \\ F_\epsilon\left(\left(\dfrac{y}{\sigma}\right)^{1/p}\right)p & , y \leq x. \end{cases}$$

Moreover, for each positive integer m, the m-steps tpf is given by

$$Q^m(x,]\sigma,y]) := P(X_{1+m} \leq y | X_1 = x)$$

$$= \begin{cases} \displaystyle\sum_{k=1}^m F_{\prod_{i=1}^k \epsilon_i}\left(\left(\frac{y}{\sigma}\right)^{1/p}\right)p(1-p)^{k-1} + F_{\prod_{i=1}^m \epsilon_i}\left(\left(\frac{y}{x}\right)^{1/p}\right)(1-p)^m & , y > x \\ \displaystyle\sum_{k=1}^m F_{\prod_{i=1}^k \epsilon_i}\left(\left(\frac{y}{\sigma}\right)^{1/p}\right)p(1-p)^{k-1} & , y \leq x, \end{cases}$$

where

$$F_{\prod_{i=1}^k \epsilon_i}(x) = 1 - x^{-\alpha}\sum_{j=1}^k \frac{(\log x^\alpha)^{k-j}}{(k-j)!}. \tag{3}$$

A Pareto-type distribution F is of the form

$$F(x) = 1 - x^{-\alpha} L(x),$$

with $\alpha > 0$ and where $L(x)$ is a slowly varying function (i.e. $L(tx)/L(x) \to \infty$ as $x \to \infty$, for every $t > 0$). Observe that $F_{\prod_{i=1}^{k} \epsilon_i}(x)$ in (3) is a Pareto-type distribution.

In order to study the extremal behavior of the LLP(I)(1) process we will analyze some long-range and local dependence conditions.

First, we will show that the β-mixing condition holds. This is a slightly stronger condition than strong-mixing which basically states that the realization of two rv's tends to be independent as they are getting increasingly separated in time. A stationary sequence $\{X_i\}_{i\geq 1}$ is said to be β-mixing if

$$\beta(l) := \sup_{p\in\mathbb{N}} E\big(\sup_{B\in\mathcal{F}(X_{p+l+1},\ldots)} |P(B|\mathcal{F}(X_1, ..., X_p)) - P(B)|\big) \underset{l\to\infty}{\longrightarrow} 0,$$

with $\mathcal{F}(.)$ denoting the $\sigma-$field generated by the indicated random variables.

Proposition 2.1. *The LLP(I)(1) process is β-mixing.*

Proof. We state the β-mixing condition by proving that the process is regenerative and aperiodic (Asmussen, 2008).

In what concerns regeneration we will see that it has a regeneration set R, i.e., a recurrent set R such that, for some $m \in \mathbb{N}$, a distribution ψ and $\delta \in (0, 1)$, we have

$$Q^m(x, B) \geq \delta\, \psi(B),\ x \in \mathbb{R},$$

for all real borelian B. The process is aperiodic if, for any regeneration set R and any real borelian B, we have

$$Q^{m+1}(x, B) \geq \delta_1\, \psi(B) \ \text{ and } \ Q^m(x, B) \geq \delta_2\, \psi(B),\ \forall x \in \mathbb{R}, \tag{4}$$

for some $m \in \mathbb{N}$ and $\delta_1, \delta_2 \in (0, 1)$.

Let $R =]r, \infty[\subset]\sigma, \infty[$ and B a real borelian set. We have that R is recurrent since it is in the support of the process. Consider $x \in R$, $S =]0, r]$ and $W \sim$Pareto$(\alpha/p, \sigma)$. For all $x \in R$,

$$Q(x, B) \geq \int_{B\cap S} dQ(x, z) = P(W \in B \cap S)p$$

and thus regeneration holds by considering $m = 1$, $\psi(B) = P(W \in B \cap S)$ and $\delta = p$. For aperiodicity, see that

$$Q^2(x, B) \geq \int_S pP(W \in B)Q(x, dz) \geq pP(W \in B\cap S)Q(x, S) = pP(W \in B\cap S)pP(W \in S)$$

and thus (4) holds with $\delta_1 = pP(W \in S)$ and $\delta_2 = \delta$. □

If we consider a mixing condition but only required to hold for events of the form $\{X_i \leq u_n\}$ or their intersections, we arrive at the weaker long-range dependence condition $D(u_n)$ of Leadbetter (1974), for any real sequence $\{u_n\}_{n\geq 1}$. In stationary sequences satisfying the dependence condition $D(u_n)$, it is possible to watch some clustering phenomena of high values and a dependence parameter known as extremal index arrives in this context. Formerly, a stationary sequence $\{X_n\}_{n\geq 1}$ has extremal index $\theta \in [0,1]$ if, for each $\tau > 0$, there is a sequence of normalized levels $\{u_n \equiv u_n^{(\tau)}\}_{n\geq 1}$, i.e.,

$$n(1 - F(u_n)) \to \tau, \tag{5}$$

as $n \to \infty$, such that

$$P(M_n \leq u_n) \to e^{-\theta\tau}$$

(Leadbetter et al., 2012). Parameter θ is a measure associated with the degree of clustering. More precisely, under quite general conditions it corresponds to the arithmetic inverse of the cluster size. In particular, a unit value of θ means a behavior that mimics an iid sequence where we have an isolated occurrence of high values and thus no clustering phenomena.

Local dependence conditions $D^{(k)}(u_n)$ considered in Chernick et al. (1991), allow us to compute the extremal index by

$$\theta = \lim_{n\to\infty} P(M_{2,k} \leq u_n | X_1 > u_n), \tag{6}$$

where $M_{i,j} = \max(X_i, \ldots, X_j)$ and u_n are normalized levels, i.e., satisfy (5).

Condition $\mathrm{D}^{(k)}(u_n)$ is said to hold for $\{X_n\}_{n\geq 1}$ if, under condition $D(u_n)$, we have

$$nP\left(X_1 > u_n, M_{1,k} \leq u_n < M_{k,r_n}\right) \underset{n\to\infty}{\longrightarrow} 0,$$

with $\{r_n = [n/k_n]\}_{n\geq 1}$, where $[x]$ denotes the integer part of x, for some sequence $\{k_n\}_{n\geq 1}$ satisfying

$$k_n \to \infty,\ k_n\alpha_{n,l_n} \to 0,\ k_n l_n/n \to 0,$$

as $n \to \infty$. This is implied by condition

$$n \sum_{j=k+1}^{r_n} P\left(X_1 > u_n, M_{1,k} \leq u_n < X_j\right) \underset{n\to\infty}{\longrightarrow} 0\,, \tag{7}$$

which corresponds to condition $D'(u_n)$ of Leadbetter et al. (2012) if $k = 1$ and condition $D''(u_n)$ of Leadbetter and Nandagopalan (1989) if $k = 2$. The first enables local clustering of exceedances of large values, corresponding to $\theta = 1$, and the second allows local clustering of exceedances but restricts upcrossings.

We will show that $D''(u_n)$ condition is satisfied by LLP(I)(1) processes.

Proposition 2.2. *Condition $D''(u_n)$ holds for process LLP(I)(1), for levels u_n satisfying (5).*

Proof. We have,

$$
\begin{aligned}
& P\left(X_1 > u_n, X_j \leq u_n < X_{j+1}\right) \\
= \; & P\left(X_1 > u_n, X_j \leq u_n\right)\left(1 - F_\epsilon\left(\left(\frac{u_n}{\sigma}\right)^{1/p}\right)\right) p \\
& + \int_1^\infty P\left(X_1 > u_n, X_j \leq u_n, X_j > \frac{u_n}{x^p}\right) F_\epsilon(dx)\,(1-p) \\
\leq \; & \underbrace{\left(1 - F_X(u_n)\right)\left(1 - F_\epsilon\left(\left(\frac{u_n}{\sigma}\right)^{1/p}\right)\right) p}_{I_1} \\
& + \underbrace{\int_1^\infty \int_{u_n}^\infty Q^{j-1}\left(y, \left]\frac{u_n}{x^p}, u_n\right]\right) F_X(dy) F_\epsilon(dx)\,(1-p)}_{I_2}.
\end{aligned}
$$

Observe that

$$
I_1 = \left(1 - F_X(u_n)\right)\left(1 - F_\epsilon\left(\left(\frac{u_n}{\sigma}\right)^{1/p}\right)\right) p = O\left(\frac{1}{n^{1/p+1}}\right).
$$

In what concerns the second term, by the dominated convergence theorem, we have successively

$$
\begin{aligned}
I_2 \leq \; & \int_1^\infty \int_{u_n}^\infty \left(1 - Q^{j-1}\left(y, \left]\sigma, \frac{u_n}{x^p}\right]\right)\right) F_X(dy) F_\epsilon(dx)\,(1-p) \\
\leq \; & \int_1^\infty \int_{u_n}^\infty \left[1 - F_{\prod_{i=1}^k \epsilon_i}\left(\left(\frac{u_n}{x^p \sigma}\right)^{1/p}\right)\right] F_X(dy) F_\epsilon(dx)\,(1-p) \\
\lesssim \; & \left(1 - F_X(u_n)\right) \int_1^\infty \left[1 - F_\epsilon\left(\frac{u_n}{x^p \sigma}\right)\right] F_\epsilon(dx)\,(1-p) \\
= \; & O\left(\tfrac{\tau}{n^2}\right).
\end{aligned}
$$

Therefore,

$$
n \sum_{j=2}^{r_n - 1} P\left(X_1 > u_n, X_j \leq u_n < X_{j+1}\right) \leq O\left(\frac{1}{k_n n}\right).
$$

□

In the next result we will see that the degree of clustering in the LLP(I) process depends on the parameter p.

Proposition 2.3. *The LLP(I)(1) process has extremal index $\theta = p$.*

Proof. By the Chernick et al. (1991) result in (6), we have

$$\theta = \lim_{n\to\infty} P(X_2 \leq u_n | X_1 > u_n),$$

with $u_n \sim (n/\tau)^{1/\alpha}\sigma$, $n \geq 1$, since the quantile function is

$$F_X^{-1}(t) = \sigma(1-t)^{-1/\alpha}.$$

Observe that

$$P(X_2 \leq u_n, X_1 > u_n) = \int_{u_n}^{\infty} Q(x,]\sigma, u_n])F_X(dx) = F_\epsilon\left(\left(\frac{y}{\sigma}\right)^{1/p}\right)p(1 - F_X(u_n))$$

and thus

$$\theta = \lim_{n\to\infty}\left(1 - \left(\frac{\tau}{n}\right)^{1/p}\right)p = p.$$

□

Dependence measures like autocorrelation, based on the central part of the series and usually considered in the modeling from linear ARMA, may not be defined and poorly describe the dependence for large levels (see Embrechts et al. (2002)). In the LLP(I)(1) process, the autocorrelation is derived as (Arnold, 2001)

$$\rho(X_{n-1}, X_n) = -\frac{(1-p)\alpha}{\alpha - p} \tag{8}$$

and the negative sign may be a useful tool in diagnosing the suitability of a LLP(I)(1) process to describe a given time series. Nevertheless, it is only defined for $\alpha > 2$.

Alternatively, we can use tail dependence measures as the lag-m tail dependence coefficient,λ_m, defined by

$$\lambda_m = \lim_{t\downarrow 0} P(X_{1+m} > F_X^{-1}(1-t)|X_1 > F_X^{-1}(1-t)),$$

and thus analyze the probability of X_{1+m} being extreme given that X_1 is extreme too. We have X_1 and X_{1+m} asymptotically independent if $\lambda_m = 0$, and asymptotically dependent if $0 < \lambda_m \leq 1$, with the boundary cases $\lambda_m = 1$ and $\lambda_m \sim P(X_{1+m} > F_X^{-1}(1-t))$ corresponding to complete dependence and independence, respectively. Under slightly weaker conditions than (7), we can express the extremal index θ in terms of coefficients λ_m. In particular, under $D''(u_n)$, we have $\theta = 1 - \lambda_1$ (Ferreira and Ferreira (2012), Proposition 4).

Proposition 2.4. *The LLP(I)(1) process has lag-m tail dependence coefficient* $\lambda_m = (1-p)^m$.

Proof. Let $a_t = F_X^{-1}(1-t)$. We shall consider $m = 3$, but the proof runs along the same lines for any value of m. We have

$$P(X_1 > a_t, X_4 > a_t) = \int_{a_t}^{\infty}(1 - Q^3(x,]\sigma, a_t]))F_X(dx)$$

$$= \left[1 - F_{\epsilon}(t^{-1/(\alpha p)})p - F_{\epsilon_1\epsilon_2}(t^{-1/(\alpha p)})p(1-p) - F_{\epsilon_1\epsilon_2\epsilon_3}(t^{-1/(\alpha p)})p(1-p)^2\right](1 - F_X(a_t)).$$

Therefore,

$$\lambda_3 = \lim_{t\downarrow 0}\left[1 - (1-t^{1/p})p - (1-t^{1/p})p(1-p) - (1-t^{1/p})p(1-p)^2\right] = (1-p)^3.$$

□

We state the following result relating the several measures and thus providing additional diagnostic tools for modeling purposes.

Corollary 2.5. *In the LLP(I)(1) process, we have*

a) $\theta = \frac{1}{f_1} - 1$;

b) $\lambda = 2 - \frac{1}{f_1}$.

.

3 Estimation

Measures like the fluctuation probabilities, the extremal index and the TDC are related with the parameter p, and thus provide direct estimation procedures.

Based on (2), we have

$$p = \frac{1}{f_1} - 1,\ f_1 \geq 1/2.$$

If we replace f_1 by the respective empirical counterpart

$$\widehat{f_1} = \frac{1}{n-1}\sum_{j=2}^{n} \mathbb{1}_{\{X_{j-1}<X_j\}},$$

where $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function, we have

$$\widehat{p}^{FP} = \frac{1}{\widehat{f_1}} - 1 = \frac{n-1}{\sum_{j=2}^{n}\mathbb{1}_{\{X_{j-1}<X_j\}}} - 1,$$

provided that $\widehat{f_1} \geq 1/2$. This restriction may be an indicator of an unappropriate choice of the model.

Unrestricted estimators will be derived from the extremal measures, whose inference concerns only the large values and ignores the rest of the data.

For the extremal index θ, we consider the estimator of Nandagopalan (1990) which works under condition $D''(u_n)$:

$$\widehat{\theta} = \frac{C_n(u)}{N_n(u)},$$

where $C_n(u_n)$ and $N_n(u_n)$ are, respectively, the number of downcrossings and the number of exceedances of a high threshold u. Therefore,

$$\widehat{p}^{EI} \equiv \widehat{p}_u^{EI} = \frac{C_n(u)}{N_n(u)}\,. \tag{9}$$

The estimator of the TDC that we apply corresponds to the empirical counterpart of λ_1, considered in Schmidt and Stadtmüller (2006):

$$\widehat{\lambda} \equiv \widehat{\lambda}_u = \frac{N_n^*(u)}{N_n(u)}, \tag{10}$$

where $N_n^*(u)$ denotes the number of exceedances of a high threshold u among $\min(X_j, X_{j+1})$, $j = 1, \ldots, n-1$. We consider $\widehat{p}^{TDC} \equiv \widehat{p}_u^{TDC} = 1 - \widehat{\lambda}_u$.

In reality, it may happens that the data do not exactly satisfies the functional equation (1) but some arbitrarily close formula. Here we consider "noisy" processes of the form $X_n^{(\zeta)} = X_n + \zeta Z_n$, $n \geq 1$, where $\{Z_n\}_{n\geq 1}$ is an iid sequence of standard Gaussian rv's and $\zeta > 0$, in order to analyze the sensitivity of the estimators to these perturbed versions.

Our simulation study consists in 1000 samples of size $n = 1000$ drawn from the LLP(I)(1) model by considering marginals Pareto(I)(σ, α), with $\sigma = 1$ and $\alpha = 1$, and taking $p = 0.25, 0.5, 0.75$. The results are compared in terms of root mean squared error and absolute bias and are placed in Tables 1 and 2, including the case of the "noisy" processes $\{X_n^{(\zeta)}\}_{n\geq 1}$ for $\zeta = 1, 0.1$. Estimator $\widehat{p}^{FP}$ is the best for simple LLP(I)(1) processes but it is not robust for noisy versions, particularly for lower values of p which corresponds to larger dependence within the process, besides restricted to $\widehat{f_1} \geq 1/2$. On the other hand, the estimators based on the tail, $\widehat{p}_u^{TDC}$ and $\widehat{p}_u^{EI}$, are robust, with the best results attained for quantile $u = 0.8$, except in the case $p = 0.75$.

Table 1: Root mean squared errors obtained for the LLP(I)(1) process ($\zeta = 0$) and for noisy versions ($\zeta = 1, 0.1$).

$p = 0.25$	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 1$
$\widehat{p}^{FP}$	-	-	-
$\widehat{p}^{TDC}_{q_{0.95}}$	0.063	0.063	0.071
$\widehat{p}^{TDC}_{q_{0.8}}$	0.032	0.032	0.063
$\widehat{p}^{EI}_{q_{0.95}}$	0.063	0.063	0.071
$\widehat{p}^{EI}_{q_{0.8}}$	0.032	0.032	0.063
$p = 0.5$	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 1$
$\widehat{p}^{FP}$	0.032	0.063	0.263
$\widehat{p}^{TDC}_{q_{0.95}}$	0.071	0.071	0.071
$\widehat{p}^{TDC}_{q_{0.8}}$	0.032	0.032	0.032
$\widehat{p}^{EI}_{q_{0.95}}$	0.071	0.071	0.071
$\widehat{p}^{EI}_{q_{0.8}}$	0.032	0.032	0.032
$p = 0.75$	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 1$
$\widehat{p}^{FP}$	0.032	0.045	0.126
$\widehat{p}^{TDC}_{q_{0.95}}$	0.055	0.055	0.055
$\widehat{p}^{TDC}_{q_{0.8}}$	0.089	0.089	0.084
$\widehat{p}^{EI}_{q_{0.95}}$	0.063	0.063	0.063
$\widehat{p}^{EI}_{q_{0.8}}$	0.095	0.095	0.089

Table 2: Absolute biases obtained for the LLP(I)(1) process ($\zeta = 0$) and for noisy versions ($\zeta = 1, 0.1$).

$p = 0.25$	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 1$
$\widehat{p}^{FP}$	-	-	-
$\widehat{p}^{TDC}_{q_{0.95}}$	0.026	0.027	0.032
$\widehat{p}^{TDC}_{q_{0.8}}$	0.004	0.005	0.047
$\widehat{p}^{EI}_{q_{0.95}}$	0.012	0.013	0.018
$\widehat{p}^{EI}_{q_{0.8}}$	0.002	0.003	0.045

$p = 0.5$	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 1$
$\widehat{p}^{FP}$	0.001	0.056	0.261
$\widehat{p}^{TDC}_{q_{0.95}}$	0.02	0.019	0.021
$\widehat{p}^{TDC}_{q_{0.8}}$	0.016	0.015	0.001
$\widehat{p}^{EI}_{q_{0.95}}$	0.011	0.011	0.012
$\widehat{p}^{EI}_{q_{0.8}}$	0.018	0.017	0.001

$p = 0.75$	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 1$
$\widehat{p}^{FP}$	0	0.025	0.123
$\widehat{p}^{TDC}_{q_{0.95}}$	0.007	0.007	0.007
$\widehat{p}^{TDC}_{q_{0.8}}$	0.086	0.086	0.081
$\widehat{p}^{EI}_{q_{0.95}}$	0.012	0.012	0.011
$\widehat{p}^{EI}_{q_{0.8}}$	0.087	0.087	0.082

4 An application to financial series

Financial data often present heavy tailed marginals. Volatility within stock market indexes is usually characterized by sudden large atypical observations and thus better modeled by heavy tailed processes like Pareto ones. We consider the daily closing values of the FTSE100 index from January 1980 to March 2004. Volatility can be measured through the absolute values of the log-returns (the difference between the logarithms of successive daily prices), or equivalently, through the squared log-returns. In Figure 2 is plotted the time series of both log-returns and squared log-returns (volatility), respectively. Observe that volatility presents bursts of high peaks similar to LLP(I)(1) sample paths in Figure 1. In order to assure that our data is likely to be modeled with Pareto marginals (for instance, the observations are not in the support $[1, \infty[$), we conducted a robust regression and obtained scale and location estimates, approximately, 11560 and 1 (see the resulting Pareto quantile-quantile plot in Figure 3). From now on we address the transformed data, denoted X_i, $i = 1, \ldots, n = 3596$. First we estimate the weight of the tail, i.e., the shape parameter α. We consider the Hill estimator, $\widehat{\alpha}^{-1} = k^{-1} \sum_{i=1}^{k} \log(X_{n-i+1:n}/X_{n-k:n})$, $k = 1, \ldots, n-1$, based on the $k+1$ top order statistics (Hill et al. (1975)), whose sample path is in Figure 3 (right). The estimate is inferred from a flat region sought after the high variability in the beginning of the path due to the small amount of observations that are being used, but not to far from the tail where the bias starts to dominate. We can find some stability around 0.5 and also around 0.6 leading to an estimate of α, respectively, about 2 and 1.67. The autocorrelation function in (8) is only stated for $\alpha > 2$. Since we cannot guarantee this condition, we do not proceed in this pathway. Now we estimate parameter p. We used estimator $\widehat{p}^{EI}$ in (9) and obtained $\widehat{p}^{EI}_{q_{0.80}} = .50$ and $\widehat{p}^{EI}_{q_{0.95}} = 0.60$. The estimator $\widehat{p}^{FP}$ was not considered since it is very sensitive to noisy processes, a more realistic assumption in real data. Computing estimates of $\widehat{\lambda}$ in (10), lead us to $1 - \widehat{\lambda}_{q_{0.80}} = 0.54$ and $1 - \widehat{\lambda}_{q_{0.95}} = 0.65$, which approximately comprises the relation $\lambda = 1 - p$, a characterizing feature of this process.

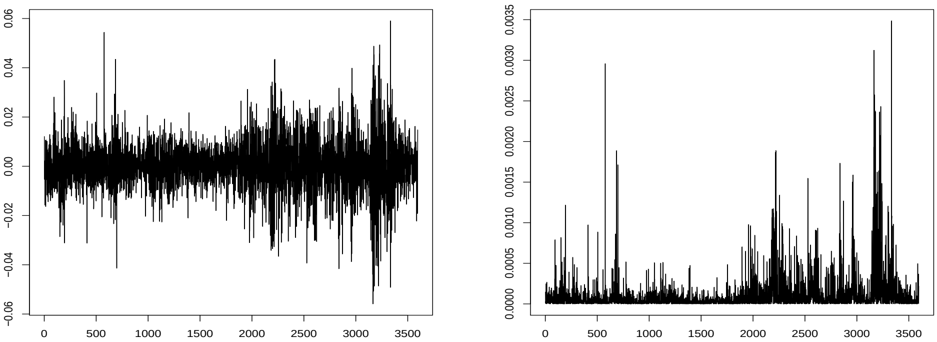

Figure 2: From left to right: Log-returns and squared log-returns (volatility) of the FTSE100 index from January 1980 to March 2004, amounting a sample of size 3596.

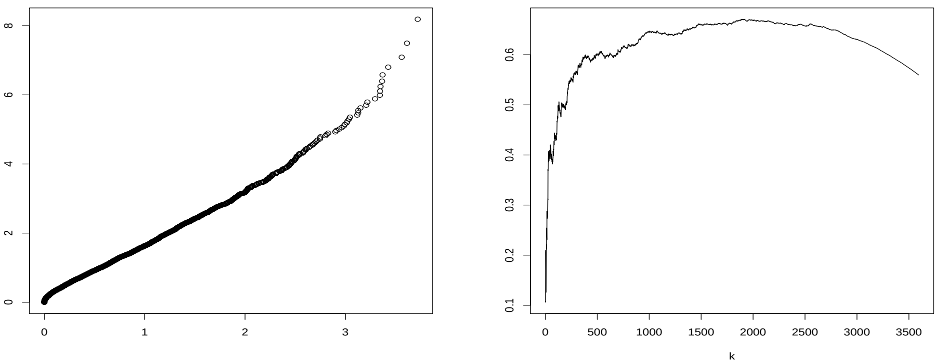

Figure 3: From left to right: Pareto quantile-quantile plot and sample path of Hill estimator of α^{-1} of the transformed data.

5 Discussion

The extremal approach presented in this paper provides new methods concerning inference within the LLP(I)(1) processes, namely, estimators for the autoregressive parameter p valid in all of the domain $[0, 1]$. As observed in Arnold (2001), these processes are potentially suitable for variables of economic nature. We intend to pursue on an extreme values scope to develop modeling technics similar to Ferreira and Canto e Castro (2010) and apply to real data.

Acknowledgement

The author research was supported by the Research Center CEMAT through the Project UID/Multi/04621/2013.

References

Alpuim, M. T. (1989). An extremal markovian sequence. *Journal of Applied Probability*, 26(2):219–232.

Alpuim, M. T. and Athayde, E. (1990). On the stationary distribution of some extremal markovian sequences. *Journal of Applied Probability*, 27(2):291–302.

Arnold, B. C. (1985). *Pareto distribution*. Wiley Online Library.

Arnold, B. C. (2001). Pareto processes. In *Stochastic Processes: Theory and Methods*, volume 19 of *Handbook of Statistics*, pages 1–33. Elsevier.

Asmussen, S. (2008). *Applied probability and queues*, volume 51. Springer Science & Business Media.

Carcea, M. and Serfling, R. (2015). A gini autocovariance function for time series modelling. *Journal of Time Series Analysis*, 36(6):817–838.

Chernick, M. R., Hsing, T., and McCormick, W. P. (1991). Calculating the extremal index for a class of stationary sequences. *Advances in Applied Probability*, 23(4):835–850.

Davis, R. A. and Resnick, S. I. (1989). Basic properties and prediction of max-arma processes. *Advances in Applied Probability*, 21(4):781–803.

Embrechts, P., McNeil, A., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond*, pages 176–223.

Ferreira, M. (2012a). On the extremal behavior of a pareto process: an alternative for armax modeling. *Kybernetika*, 48(1):31–49.

Ferreira, M. (2012b). Parameter estimation and dependence characterization of the mar(1) process. *ProbStat Forum*, 5(12):107–111.

Ferreira, M. and Canto e Castro, L. (2010). Modeling rare events through a prarmax process. *Journal of Statistical Planning and Inference*, 140(11):3552–3566.

Ferreira, M. and Ferreira, H. (2012). On extremal dependence: some contributions. *Test*, 21(3):566–583.

Ferreira, M. and Ferreira, H. (2013). Extremes of multivariate armax processes. *Test*, 22(4):606–627.

Hill, B. M. et al. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.

Lawrance, A. and Lewis, P. (1981). A new autoregressive time series model in exponential variables (near (1)). *Advances in Applied Probability*, 13(4):826–845.

Leadbetter, M. R. (1974). On extreme values in stationary sequences. *Probability Theory and Related Fields*, 28(4):289–303.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (2012). *Extremes and related properties of random sequences and processes*. Springer Science & Business Media.

Leadbetter, M. R. and Nandagopalan, S. (1989). *On exceedance point processes for stationary sequences under mild oscillation restrictions*. Springer.

Lebedev, A. V. (2008). Statistical analysis of first-order marma processes. *Mathematical Notes*, 83(3-4):506–511.

Nandagopalan, S. (1990). *Multivariate extremes and estimation of the extremal index*. PhD thesis, University of North Carolina, Chapel Hill.

Schmidt, R. and Stadtmüller, U. (2006). Non-parametric estimation of tail dependence. *Scandinavian Journal of Statistics*, 33(2):307–335.

Yeh, H.-C., Arnold, B. C., and Robertson, C. A. (1988). Pareto processes. *Journal of Applied Probability*, 25(2):291–301.