Research Article

Helena Ferreira and Marta Ferreira*

The stopped clock model

https://doi.org/10.1515/demo-2022-0101
received January 7, 2022; accepted March 9, 2022

Abstract: The extreme value theory presents specific tools for modeling and predicting extreme phenomena. In particular, risk assessment is often analyzed through measures for tail dependence and high values clustering. Despite technological advances allowing an increasingly larger and more efficient data collection, there are sometimes failures in the records, which causes difficulties in statistical inference, especially in the tail where data are scarcer. In this article, we present a model with a simple and intuitive failures scheme, where each record failure is replaced by the last record available. We will study its extremal behavior with regard to local dependence and high values clustering, as well as the temporal dependence on the tail.

Keywords: extreme values, stationary sequences, failures model, extremal index, tail dependence coefficient

MSC 2020: 60G70

1 Introduction

Let \( \{X_n\}_{n \in \mathbb{Z}} \) and \( \{U_{n \in \mathbb{Z}}\} \) be stationary sequences of real random variables on the probability space \((\Omega, \mathcal{A}, P)\) and \(P(U_n \in \{0, 1\}) = 1\). We define, for \( n \geq 1 \),

\[
Y_n = \begin{cases} 
X_n, & U_n = 1 \\
Y_{n-1}, & U_n = 0. 
\end{cases}
\]  

(1)

Sequence \( \{Y_n\}_{n \geq 1} \) corresponds to a model of failures on records of \( \{X_n\}_{n \in \mathbb{Z}} \) replaced by the last available record, which occurs in some random past instant, if we interpret \( n \) as time. Thus, if, for example, it occurs \( \{U_1 = 1, U_2 = 0, U_3 = 1, U_4 = 0, U_5 = 0, U_6 = 0, U_7 = 1\} \), we will have \( \{Y_1 = X_1, Y_2 = X_1, Y_3 = X_2, Y_4 = X_2, Y_5 = X_2, Y_6 = X_3, Y_7 = X_3\} \). This constancy of some variables of \( \{X_{n \in \mathbb{Z}}\} \) for random periods of time motivates the designation of “stopped clock model” for sequence \( \{Y_n\}_{n \geq 1} \).

Failure models studied in the literature from the point of view of extremal behavior do not consider the stopped clock model (Hall and Hüsler [4]; Ferreira et al. [3] and references therein).

The model we will study can also be represented by \( \{X_{N_n}\}_{n \geq 1} \), where \( \{N_n\}_{n \geq 1} \) is a sequence of positive integer variables representable by

\[
N_n = nU_n + \sum_{n-1 \geq i \geq 1} \left( \prod_{j=0}^{n-1-i} (1-U_{n-j}) \right) U_{n-i}(n-i), \quad n \geq 1.
\]

We can also state a recursive formulation for \( \{Y_n\}_{n \geq 1} \) through

* Corresponding author: Marta Ferreira, Centro de Matemática e Departamento de Matemática, Universidade do Minho, CEMAT, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal; Statistics and Applications Center (CEAUL), University of Lisbon, Lisbon, Portugal, e-mail: msferreira@math.uminho.pt

Helena Ferreira: Universidade da Beira Interior, Centro de Matemática e Aplicações (CMA-UBI), Departamento de Matemática, Avenida Marquês d’Avila e Bolama, 6200-001 Covilhã, Portugal, e-mail: helena.ferreira@ubi.pt
\[
Y_n = X_n U_n + \sum_{n-1 \leq i \leq 1} \left( \prod_{j=0}^{i-1} (1 - U_{n-j}) \right) U_n \cdot X_{n-i} \cdot \prod_{i=0}^{j} (1 - U_{n-i}) Y_{n-k} \cdot n \geq 1, \ k \geq 1.
\]

Under any of the three possible representations (failures model, random index sequence or recursive sequence), we are not aware of an extremal behavior study of \( \{Y_n\}_{n=1}^{\infty} \) in the literature.

Our departure hypotheses about the base sequence \( \{X_n\}_{n \in \mathbb{Z}} \) and about sequence \( \{U_n\}_{n \in \mathbb{Z}} \) are as follows:

1. \( \{X_n\}_{n \in \mathbb{Z}} \) is a stationary sequence of random variables almost surely distinct and, without loss of generality, such that \( F_X(x) = F(x) = \exp(-1/x), \ x > 0 \), i.e., standard Fréchet distributed.

2. \( \{X_n\}_{n \in \mathbb{Z}} \) and \( \{U_n\}_{n \in \mathbb{Z}} \) are independent.

3. \( \{U_n\}_{n \in \mathbb{Z}} \) is stationary and \( p_{n, \ldots, n} (i_1, \ldots, i_s) = P(U_n = i_1, \ldots, U_n = i_s), i_j \in \{0, 1\}, j = 1, \ldots, s \), is such that
   \[
   p_n, n+1, \ldots, n+k-1 (0, \ldots, 0) = 0, \text{ for some } k \geq 1.
   \]

The trivial case \( \kappa = 1 \) corresponds to \( Y_n = X_n, n \geq 1 \). Hypothesis (3) means that we are assuming that it is almost impossible to lose \( \kappa \) or more consecutive values of \( \{X_n\}_{n \in \mathbb{Z}} \). We remark that, along the paper, the summations, products and intersections is considered to be nonexistent whenever the end of the counter is less than the beginning. We will also use notation \( a \lor b = \max(a, b) \).

**Example 1.1.** Consider an independent and identically distributed sequence \( \{W_n\}_{n \in \mathbb{Z}} \) of real random variables on \( (\Omega, \mathcal{A}, P) \) and a Borelian set \( A \). Let \( p = P(A_n), \) where \( A_n = \{W_n \in A\}, n \in \mathbb{Z} \). The sequence of Bernoulli random variables
\[
U_n = 1_{0 \leq i \leq \kappa-1} X_{n-i} + \left( 1 - 1_{0 \leq i \leq \kappa-1} X_{n-i} \right) 1_{|A_n|}, \ n \in \mathbb{Z},
\]
where \( 1_{|A_n|} \) denotes the indicator function, defined for some fixed \( \kappa \geq 2 \), is such that \( p_{n, n+1, \ldots, n+k-1}(0, \ldots, 0) = 0 \), i.e., it is almost sure that after \( \kappa - 1 \) consecutive variables equal to zero, the next variable takes value one. We also have
\[
p_n(0) = P(\bigcup_{i=0}^{\kappa-1} A_{n-i} \cap A_n),
\]
where
\[
p_n(0) = P(\bigcup_{i=0}^{\kappa-1} A_{n-i} = 0, 1_{|A_n|} = 0) = P \left( \bigcup_{i=1}^{\kappa-1} A_{n-i} \cap A_n \right) = P(\overline{A}_n) - P \left( \bigcap_{i=0}^{\kappa-1} \overline{A}_{n-i} \right) = 1 - p - (1 - p)^\kappa.
\]

**Figure 1:** Sample path of 100 observations simulated from \( \{Y_n\} \) defined in (1) based on independent standard Fréchet \( \{X_n\} \) and on \( \{U_n\} \) given in (2) where we take random variables \( \{W_n\} \) standard exponential distributed, \( A = [0, 1/2) \), and thus, \( p = 0.3935 \) and considering \( \kappa = 3 \).
since the independence of random variables $W_n$ implies the independence of events $A_n$, and, for $\kappa > 2$,

$$p_{n-1,n}(0, 0) = P\left( \bigcup_{i=1}^{\kappa-1} A_{n-i} \cap \overline{A}_n \right) \cap \left( \bigcup_{i=1}^{\kappa-1} A_{n-1-i} \cap \overline{A}_{n-1} \right)$$

$$= P(\overline{A}_n \cap \overline{A}_{n-1}) - P\left( \bigcap_{i=1}^{\kappa-1} A_{n-i} \cap \bigcup_{i=1}^{\kappa-1} A_{n-1-i} \right)$$

$$= P(\overline{A}_n)P(\overline{A}_{n-1}) - P\left( \bigcap_{i=0}^{\kappa-1} A_{n-i} \right) = (1 - p)^2 - (1 - p)^\kappa,$$

$p_{n-1,n}(1, 0) = p_n(0) - p_{n-1,n}(0, 0) = p(1 - p)$.

In Figure 1, we illustrate with a particular example based on independent standard Fréchet $\{X_n\}_{n \in \mathbb{Z}}$, $\{W_n\}_{n \in \mathbb{Z}}$ with standard exponential marginals, $A = [0, 1/2]$ and thus, $p = 0.3935$ and considering $\kappa = 3$.

Therefore, $p_{n,n-1,n+2}(0, 0, 0) = 0$, $p_{n,n+1,n+2}(1, 0, 0) = p_{n,n+1}(0, 0) = p(1 - p)^2$.

In the next section, we propose an estimator for probabilities $p_{n,n-1,s}(1, 0, \ldots, 0)$, $0 \leq s < \kappa - 1$. In Section 3, we analyze the existence of the extremal index for $\{Y_n\}_{n \geq 1}$, an important measure to evaluate the tendency to occur clusters of its high values (see, e.g., Kulik and Solier [6], and references therein). A characterization of the tail dependence will be presented in Section 4. The results are illustrated with an ARMAX sequence.

For the sake of simplicity, we will omit the variation of $n$ in sequence notation whenever there is no doubt, taking into account that we will keep the designation $\{Y_n\}$ for the stopped clock model and $\{X_n\}$ and $\{U_n\}$ for the sequences that generate it.

## 2 Inference on $\{U_n\}$

Assuming that $\{U_n\}$ is not observable, as well as the values of $\{X_n\}$ that are lost, it is of interest to retrieve information about these sequences from the available sequence $\{Y_n\}$.

Since, for $n \geq 1$ and $s \geq 1$, we have

$$p_n(1) = E(1_{\{Y_n \neq Y_{n-1}\}}), \quad p_n(0) = E(1_{\{Y_n = Y_{n-1}\}}) \quad \text{and} \quad p_{n-n-1,\ldots,1,0,\ldots,0} = E(1_{\{Y_{n-1} \neq Y_{n-2} \neq \cdots \neq Y_0\}}),$$

we propose to estimate these probabilities from the respective empirical counterparts of a random sample $(\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_m)$ from $\{Y_n\}$, i.e.,

$$\hat{p}_n(1) = \frac{1}{m} \sum_{i=2}^{m} 1_{\{\hat{Y}_i \neq \hat{Y}_{i-1}\}}, \quad \hat{p}_n(0) = \frac{1}{m} \sum_{i=2}^{m} 1_{\{\hat{Y}_i = \hat{Y}_{i-1}\}} \quad \text{and}$$

$$\hat{p}_{n-s,n-1,\ldots,1,0,\ldots,0} = \frac{1}{m} \sum_{i=s+2}^{m} 1_{\{\hat{Y}_{i-s} \neq \hat{Y}_{i-s-1} \neq \cdots \neq \hat{Y}_1\}}$$

which are consistent by the weak law of large numbers. The value of $\kappa$ can be inferred from

$$\hat{\kappa} = \bigvee_{i=s+2}^{m} \bigvee_{\leq 1} s 1_{\{\hat{Y}_{i-s} \neq \hat{Y}_{i-s-1} \neq \cdots \neq \hat{Y}_1\}}$$

In order to evaluate the finite sample behavior of the aforementioned estimators, we have simulated 1,000 independent replicas with size $m = 100, 1,000, 5,000$ of the model in Example 1.1. The absolute bias (abias) and root mean squared error (rmse) are presented in Table 1. The results reveal a good performance of the estimators, even in the case of smaller sample sizes. Parameter $\kappa$ was always estimated with no error.
Table 1: The absolute bias (abias) and rmse obtained from 1,000 simulated samples with size $m = 100, 1,000, 5,000$ of the model in Example 1.1

<table>
<thead>
<tr>
<th>$\hat{\rho}_n(0)$</th>
<th>abias</th>
<th>rmse</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 100$</td>
<td>0.0272</td>
<td>0.0335</td>
</tr>
<tr>
<td>$m = 1,000$</td>
<td>0.0087</td>
<td>0.0108</td>
</tr>
<tr>
<td>$m = 5,000$</td>
<td>0.0039</td>
<td>0.0048</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\hat{\rho}_{n-1,0}(1, 0)$</th>
<th>abias</th>
<th>rmse</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 100$</td>
<td>0.0199</td>
<td>0.0253</td>
</tr>
<tr>
<td>$m = 1,000$</td>
<td>0.0065</td>
<td>0.0080</td>
</tr>
<tr>
<td>$m = 5,000$</td>
<td>0.0030</td>
<td>0.0037</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\hat{\rho}_{n-2,n-1,0}(1, 0, 0)$</th>
<th>abias</th>
<th>rmse</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 100$</td>
<td>0.0160</td>
<td>0.0200</td>
</tr>
<tr>
<td>$m = 1,000$</td>
<td>0.0051</td>
<td>0.0064</td>
</tr>
<tr>
<td>$m = 5,000$</td>
<td>0.0022</td>
<td>0.0028</td>
</tr>
</tbody>
</table>

3 The extremal index of $\{Y_n\}$

The sequence $\{Y_n\}$ is stationary because the sequences $\{X_n\}$ and $\{U_n\}$ are stationary and independent from each other. In addition, the common distribution for $Y_n$, $n \geq 1$, is also standard Fréchet, as is the common distribution for $X_n$, since

$$F_n(x) = \sum_{i=1}^{k-1} P(X_{n-i} \leq x, U_{n-i} = 1, U_{n-i+1} = 0 = \ldots = U_n) + P(X_n \leq x)P(U_n = 1) = F(x)\left(p_d(1) + \sum_{i=1}^{k-1} p_{n-i, \ldots, n}(1, 0, \ldots, 0)\right) = F(x).$$

For any $r > 0$, if we define $u_n \equiv u_d(\tau) = n / r$, $n \geq 1$, it turns out that $E(\sum_{i=1}^{n} 1_{\{Y_i > u_n\}}) = nP(Y_i > u_n) \longrightarrow \tau$ and $nP(X_i > u_n) \longrightarrow \tau$, so we refer to these levels $u_n$ by normalized levels for $\{Y_n\}$ and $\{X_n\}$.

In this section, in addition to the general assumptions about the model presented in Section 1, we start by assuming that $\{X_n\}$ and $\{U_n\}$ present dependency structures such that variables sufficiently apart can be considered approximately independent. Concretely, we assume that $\{U_n\}$ satisfies the strong-mixing condition (Rosenblatt [9]) and $\{X_n\}$ satisfies condition $D(u_n)$ (Leadbetter [7]) for normalized levels $u_n$.

**Proposition 3.1.** If $\{U_n\}$ is strong-mixing and $\{X_n\}$ satisfies condition $D(u_n)$, then $\{Y_n\}$ also satisfies condition $D(u_n)$.

**Proof.** For any choice of $p + q$ integers, $1 \leq i_1 < \ldots < i_p < j_1 < \ldots < j_q \leq n$ such that $j_k \geq i_p + l$, we have that

$$P\left(\bigcap_{s=1}^{p} X_{i_s} \leq u_n, \bigcap_{s=1}^{q} X_{j_s} \leq u_n\right) - P\left(\bigcap_{s=1}^{p} X_{i_s} \leq u_n\right)P\left(\bigcap_{s=1}^{q} X_{j_s} \leq u_n\right) \leq a_{n,l},$$

with $a_{n,l} \to 0$, as $n \to \infty$, for some sequence $l_n = o(n)$, and

$$|P(A \cap B) - P(A)P(B)| \leq g(l),$$

with $g(l) \to 0$, as $l \to \infty$, where $A$ belongs to the $\sigma$-algebra generated by $\{U_i, i = 1, \ldots, i_p\}$ and $B$ belongs to the $\sigma$-algebra generated by $\{U_i, i = j_1, j_1 + 1, \ldots\}$. Thus, for any choice of $p + q$ integers, $1 \leq i_s < \ldots < i_p < j_1 < \ldots < j_q \leq n$ such that $j_k \geq i_p + l + \kappa$, we will have

$$\left|P\left(\bigcap_{s=1}^{p} Y_{i_s} \leq u_n, \bigcap_{s=1}^{q} Y_{j_s} \leq u_n\right) - P\left(\bigcap_{s=1}^{p} Y_{i_s} \leq u_n\right)P\left(\bigcap_{s=1}^{q} Y_{j_s} \leq u_n\right)\right| \leq \sum_{i_p - \kappa \leq l_i \leq i_p} \left|P\left(\bigcap_{s=1}^{p} X^{*}_{i_s} \leq u_n, \bigcap_{s=1}^{q} X^{*}_{j_s} \leq u_n\right)P(A' \cap B') - P\left(\bigcap_{s=1}^{p} X^{*}_{i_s} \leq u_n\right)P\left(\bigcap_{s=1}^{q} X^{*}_{j_s} \leq u_n\right)P(A')P(B')\right|,$$
where $A^* = \bigcap_{\substack{i=1}}^p \{U_i = 0 = \ldots = U_{i+1}^*, U_{i+1}^* = 1\}$ and $B^* = \bigcap_{\substack{j=1}}^q \{U_k = 0 = \ldots = U_{k+1}^*, U_{k+1}^* = 1\}$ and $j^* > j_1 - \kappa > i_1^* + 1$.

Therefore, the aforementioned summation is upper limited by

$$
\sum_{i_2^{-\kappa} < i \leq i_2^{-\kappa} + l_1} \left( P\left( \bigcap_{s=1}^q X_{i_s}^* \leq u_n, \bigcap_{s=1}^q X_{i_s}^* \leq u_n \right) - P\left( \bigcap_{s=1}^q X_{i_s}^* \leq u_n \right) P\left( \bigcap_{s=1}^q X_{i_s}^* \leq u_n \right) \right) + \left| P(A^* \cap B^*) - P(A^*)P(B^*) \right|
$$

which allows to conclude that $D(u_n)$ holds for $\{Y_n\}$ with $l_{n(Y)} = l_n + \kappa$.

The tendency for clustering of values of $\{Y_n\}$ above $u_n$ depends on the same tendency within $\{X_n\}$ and the propensity of $\{U_n\}$ for consecutive null values. The clustering tendency can be assessed through the extremal index (Leadbetter, [7]). More precisely, $\{X_n\}$ is said to have extremal index $\theta_X \in (0, 1]$ if

$$
\lim_{n \to \infty} P\left( \bigvee_{i=1}^{n} X_i \leq n/\tau \right) = e^{-\theta_X \tau}.
$$

If $D(u_n)$ holds for $\{X_n\}$, we have

$$
\lim_{n \to \infty} P\left( \bigvee_{i=1}^{n} X_i \leq u_n \right) = \lim_{n \to \infty} P^k\left( \bigvee_{i=1}^{[n/k]} X_i \leq u_n \right)
$$

for any integers sequence $\{k_n\}$, such that,

$$
k_n \to \infty, \quad k_n/\alpha_n \to 0 \quad \text{and} \quad k_n\alpha_{n,\alpha} \to 0, \quad \text{as} \quad n \to \infty.
$$

We can therefore say that

$$
\theta_X = \lim_{n \to \infty} k_n P\left( \bigvee_{i=1}^{[n/k_n]} X_i > u_n \right).
$$

Now we compare the local behavior of sequences $\{X_n\}$ and $\{Y_n\}$, i.e., of $X_i$ and $Y_i$ for $i \in \left\{ (j - 1)\left[ \frac{n}{k} \right] + 1, \ldots, j\left[ \frac{n}{k} \right] \right\}$, $j = 1, \ldots, k_n$, with regard to the oscillations of their values in relation to $u_n$. To this end, we will use local dependency conditions $D^{(s)}(u_n)$, $s \geq 2$, whenever

$$
\lim_{n \to \infty} n \sum_{j=s}^{[n/k_n]} P(X_j > u_n, X_j \leq u_n < X_{j+1}) = 0,
$$

for some integers sequence $\{k_n\}$ satisfying (4). Condition $D^{(1)}(u_n)$ translates into

$$
\lim_{n \to \infty} n \sum_{j=2}^{[n/k_n]} P(X_j > u_n, X_j > u_n) = 0.
$$

Observe that if $D^{(s)}(u_n)$ holds for some $s \geq 1$, then $D^{(m)}(u_n)$ also holds for $m > s$. Condition $D^{(1)}(u_n)$ is known as $D'(u_n)$ (Leadbetter et al. [8]) and relates to a unit extremal index, i.e., absence of extreme values clustering. In particular, this is the case of independent variables. Although $\{X_n\}$ satisfies $D'(u_n)$, this condition is not generally valid for $\{Y_n\}$. Observe that

$$
n \sum_{j=2}^{[n/k_n]} P(Y_j > u_n, Y_j > u_n)
$$

$$
= \sum_{i=2}^{n} \sum_{j=2}^{[n/k_n]} \sum_{j'=i-1}^{j} P(X_i > u_n, X_{j'} > u_n) \cdot P_{i,j',\ldots,1,j'+1,\ldots,j}(0, \ldots, 0, 1, 0, \ldots, 0).
$$

For $i = 1$ and $j = \kappa$, we have $j^* = 1$ and the corresponding term becomes $nP(X_i > u_n) \to \tau > 0$, as $n \to \infty$, and this is the reason why, in general, $\{Y_n\}$ does not satisfy $D'(u_n)$ even if $\{X_n\}$ satisfies it.
Proposition 3.2. The following statements hold:

(i) If \( \{Y_n\} \) satisfies \( D^{(s)}(u_n) \), \( s \geq 2 \), then \( \{X_n\} \) satisfies \( D^{(s)}(u_n) \).

(ii) If \( \{X_n\} \) satisfies \( D^{(s)}(u_n) \), \( s \geq 2 \), then \( \{Y_n\} \) satisfies \( D^{(s+1)}(u_n) \).

(iii) If \( \{X_n\} \) satisfies \( D'(u_n) \), then \( \{Y_n\} \) satisfies \( D^{(2)}(u_n) \).

Proof. Consider \( n = [n / k_n] \). We have that

\[
n \sum_{j=s}^{n} \mathbb{P}(Y_i > u_n, Y_j \leq u_n < Y_{j+1}) \]

\[
= \frac{1}{n} \sum_{i=2-k}^{n} \sum_{j=s}^{n} \mathbb{P}(X_i > u_n, X_j \leq u_n < X_{j+1}, U_i = 1, U_{i+1} = 0 = ... = U_j, U_{j+1} = 1) \tag{5}
\]

\[
= \frac{1}{n} \sum_{i=2-k}^{n} \sum_{j=s}^{n} \sum_{j'=i+1}^{n} \sum_{j''=j'-1}^{j'-1} \mathbb{P}(X_i > u_n, X_j' \leq u_n < X_{j+1} \leq u_n, X_{j+1} > u_n) \tag{6}
\]

Since \( \{Y_n\} \) satisfies \( D^{(s)}(u_n) \), with \( s \geq 2 \), and thus, the first summation in (5) converges to zero, as \( n \to \infty \), then all the terms in the last summations also converge to zero. In particular, when \( i = 1 \) and \( j'' = j'' = 1 \), we have

\[
n \sum_{j=s}^{n} \mathbb{P}(X_i > u_n, X_j' \leq u_n < X_{j+1} \leq u_n, X_{j+1} > u_n) \to 0, \text{ as } n \to \infty,
\]

which proves (i).

Conversely, writing the first summation in (5) with \( j \) starting at \( s + k - 1 \), we have

\[
n \sum_{j=s+k-1}^{n} \mathbb{P}(Y_i > u_n, Y_j \leq u_n < Y_{j+1}) \]

\[
= \frac{1}{n} \sum_{i=2-k}^{n} \sum_{j=s+k-1}^{n} \sum_{j'=i+1}^{n} \sum_{j''=j'-1}^{j'-1} \mathbb{P}(X_i > u_n, X_j'' \leq u_n, X_{j+1} \leq u_n, X_{j+1} > u_n) \tag{6}
\]

As for (iii), observe that

\[
n \sum_{j=2}^{n} \mathbb{P}(Y_i > u_n, Y_j \leq u_n < Y_{j+1}) \]

\[
= \frac{1}{n} \sum_{i=2-k}^{n} \sum_{j=2}^{n} \mathbb{P}(X_i > u_n, Y_j \leq u_n < X_{j+1}, U_i = 1, U_{i+1} = 0 = ... = U_j, U_{j+1} = 1) \tag{7}
\]

\[
\leq \frac{1}{n} \sum_{i=2-k}^{n} \sum_{j=2}^{n} \mathbb{P}(X_i > u_n, X_{j+1} > u_n) = \frac{1}{n} \sum_{i=2-k}^{n} \sum_{j=2}^{n} \mathbb{P}(X_i > u_n, X_{j-2} > u_n) \leq \frac{1}{n} \sum_{j=2}^{n} \mathbb{P}(X_i > u_n, X_j > u_n).
\]

If \( \{X_n\} \) satisfies \( D'(u_n) \), then (7) converges to zero, as \( n \to \infty \), and \( D^{(2)}(u_n) \) holds for \( \{Y_n\} \). □

Under conditions \( D'(u_n) \) and \( D^{(s)}(u_n) \) with \( s \geq 2 \), we can also compute the extremal index \( \theta_X \) defined in (3) by (Chernick et al. [1]; Corollary 1.3)

\[
\theta_X = \lim_{n \to \infty} \mathbb{P}(X_i \leq u_n, \ldots, X_s \leq u_n | X_i > u_n).
\]

\[
\theta_X = \frac{1}{n} \sum_{i=2-k}^{n} \sum_{j=2}^{n} \mathbb{P}(X_i > u_n, X_j > u_n).
\]

\[
\theta_X = \frac{1}{n} \sum_{j=2}^{n} \mathbb{P}(X_i > u_n, X_j > u_n).
\]

\[
\theta_X = \frac{1}{n} \sum_{j=2}^{n} \mathbb{P}(X_i > u_n, X_j > u_n).
\]
If \( \{X_n\} \) and \( \{Y_n\} \) have extremal indexes \( \theta_X \) and \( \theta_Y \), respectively, then \( \theta_Y \leq \theta_X \), since \( P(\sqrt[n]{1}X_i \leq n/\tau) \leq P(\sqrt[n]{1}Y_i \leq n/\tau) \). This corresponds to the intuitively expected, if we remember that the possible repetition of variables \( X_n \) leads to larger clusters of values above \( u_n \). In the following result, we establish a relationship between \( \theta_X \) and \( \theta_Y \).

**Proposition 3.3.** Suppose that \( \{U_n\} \) is strong-mixing and \( \{X_n\} \) satisfies conditions \( D(u_n) \) and \( D(s')(u_n) \), \( s \geq 2 \), for normalized levels \( u_n = u_0(\tau) \). If \( \{X_n\} \) has extremal index \( \theta_X \), then \( \{Y_n\} \) has extremal index \( \theta_Y \) given by

\[
\theta_Y = \theta_X \sum_{j=0}^{k-1} p_{1,2,\ldots,j+1,1,2}(1,0,\ldots,0,1) \beta_j,
\]

where

\[
\beta_j = \lim_{n \to \infty} P(X_{s+j} > u_0 | X_i \leq u_0, \ldots, X_{s-1} \leq u_0 < u_0).
\]

**Proof.** By Proposition 3.1, \( \{Y_n\} \) also satisfies condition \( D(u_n) \). Thus, we have

\[
\lim_{n \to \infty} P\left( \bigvee_{i=1}^{n} Y_i \leq u_0 \right) = \exp\left\{ - \lim_{n \to \infty} k_n P\left( \bigvee_{i=1}^{[n/\beta]} Y_i > u_0 \right) \right\}
\]

and

\[
\lim_{n \to \infty} k_n P\left( \bigvee_{i=1}^{[n/\beta]} Y_i > u_0 \right) = \lim_{n \to \infty} k_n P\left( Y_i \leq u_0, \bigvee_{i=1}^{[n/\beta]} Y_i > u_0 \right)
\]

\[
= \lim_{n \to \infty} k_n P\left( \bigcup_{i=1}^{[n/\beta]} \{ Y_i \leq u_0 < Y_{i+1} \} \right)
\]

\[
= \lim_{n \to \infty} k_n P\left( \bigcup_{i=1}^{[n/\beta]} \{ Y_i \leq u_0 < X_{i+1}, U_{i+1} = 1 \} \right)
\]

\[
= \lim_{n \to \infty} k_n \sum_{i=0}^{[n/\beta]} \sum_{j=0}^{k-1} P(X_{i+j} \leq u_0, X_{i+j} < X_{i+j+1}, X_{i+j+1} > u_0) \ p_{1,1,\ldots,i+j,1,2}(1,0,\ldots,0,1)
\]

\[
= \lim_{n \to \infty} k_n \sum_{i=0}^{[n/\beta]} \sum_{j=0}^{k-1} P(X_{i+j} \leq u_0, X_{i+j} < X_{i+j+1}, X_{i+j+1} > u_0) \ p_{1,2,\ldots,i+j,1,2}(1,0,\ldots,0,1)
\]

since \( \{X_n\} \) satisfies condition \( D(s')(u_n) \) for some \( s \geq 2 \). The stationarity of \( \{X_n\} \) leads to

\[
\lim_{n \to \infty} k_n \sum_{i=0}^{[n/\beta]} \sum_{j=0}^{k-1} P(X_{i+j} \leq u_0, X_{i+j} < X_{i+j+1}, X_{i+j+1} > u_0) \ p_{1,2,\ldots,i+j,1,2}(1,0,\ldots,0,1)
\]

\[
= \lim_{n \to \infty} k_n \sum_{i=0}^{[n/\beta]} \sum_{j=0}^{k-1} P(X_{i+j} \leq u_0, X_{i+j} < X_{i+j+1}, X_{i+j+1} > u_0) \ p_{1,2,\ldots,i+j,1,2}(1,0,\ldots,0,1)
\]

\[
= \lim_{n \to \infty} k_n \sum_{i=0}^{[n/\beta]} \sum_{j=0}^{k-1} P(X_{i+j} \leq u_0, X_{i+j} < X_{i+j+1}, X_{i+j+1} > u_0) \ p_{1,2,\ldots,i+j,1,2}(1,0,\ldots,0,1)
\]
where the last step follows from (8).

Observe that \( \sum_{j=0}^{k-1} p_{1,2,\ldots,j+1,j+2}(1,0,\ldots,0,1) = p_n(1) = P(U_n = 1) \) and thus, \( \theta_Y \leq \theta_X p_n(1) \leq \theta_X \), as expected.

**Proposition 3.4.** Suppose that \( \{U_n\} \) is strong-mixing and \( \{X_n\} \) satisfies conditions \( D(u_n) \) and \( D'(u_n) \), for normalized levels \( u_n = u_n(\tau) \). Then, \( \{Y_n\} \) has extremal index \( \theta_Y \) given by \( \theta_Y = p_{1,2}(1,1) \).

**Proof.** By condition \( D'(u_n) \), the only term to consider in (9) corresponds to \( j = 0 \), and we obtain

\[
\lim_{n \to \infty} nP(Y_n \leq u_n) = \lim_{n \to \infty} \sum_{i=1}^{\lceil n/k \rceil} P(Y_n \leq u_n, U_1 = \ldots = U_i = 1, U_{i+1} = 0) = \lim_{n \to \infty} nP(X_n > u_n)p_{1,2}(1,1) = \tau p_{1,2}(1,1). \]

Observe that we can obtain the aforementioned result by applying Proposition 3.2 (iii) and calculating directly \( \tau \theta_Y = \lim_{n \to \infty} nP(Y_n \leq u_n < Y_{n+1}). \) More precisely, we have that \( \{Y_n\} \) satisfies \( D(2)(u_n) \), and by applying (8), we obtain

\[
\tau \theta_Y = \lim_{n \to \infty} nP(Y_n \leq u_n < Y_{n+1}) = \lim_{n \to \infty} nP(Y_n \leq u_n < X_n, U_1 = 1) = \lim_{n \to \infty} nP(Y_n \leq u_n < X_n, U_1 = 1, U_2 = 0).
\]

The same result can also be seen as a particular case of Proposition 3.3, where, if we take \( s = 1 \), we have \( \beta_j = 0 \) for \( j \neq 0 \), and we obtain \( \theta_Y = \theta_Y p_{1,2}(1,1) = p_{1,2}(1,1) \), since \( \beta_0 = 1 \) and under \( D'(u_n) \) it comes \( \theta_X = 1 \).

**Example 3.1.** Consider \( \{Y_n\} \) such that \( \{X_n\} \) is an ARMAX sequence, i.e., \( X_n = \phi X_{n-1} + (1 - \phi) Z_n, n \geq 1 \), where \( \{Z_n\} \) is an independent sequence of random variables with standard Fréchet marginal distribution and \( \{X_n\} \) and \( \{Z_n\} \) are independent. We show that \( \{X_n\} \) has also standard Fréchet marginal distribution, satisfies condition \( D(2)(u_n) \) and has extremal index \( \theta_X = 1 - \phi \) (see, e.g., Ferreira and Ferreira [2] and references therein).

Observe that, for normalized levels \( u_n = u_n(\tau) \), \( \tau > 0 \), we have

\[
\beta_1 = \lim_{n \to \infty} P(X_n > u_n | X_1 \leq u_n < X_2) = \lim_{n \to \infty} \frac{P(X_1 \leq u_n, X_2 \leq u_n) - P(X_1 \leq u_n, X_2 \leq u_n)}{P(X_1 \leq u_n, X_2 \leq u_n)} = \lim_{n \to \infty} \frac{1 - \frac{\tau}{\pi} - \left(1 - \frac{\tau}{\pi}(2 - \phi)\right) - \left(1 - \frac{\tau}{\pi}(2 - \phi^2)\right) + 1 - \frac{\tau}{\pi}(3 - 2\phi)}{1 - \frac{\tau}{\pi} - \left(1 - \frac{\tau}{\pi}(2 - \phi)\right)} = \phi.
\]
Analogous calculations lead to $\beta_2 = \phi^2$. Considering $\kappa = 3$, we have $\theta_T = (1 - \phi)(p_{1,2}(1, 1) + \phi p_{1,2,3}(1, 0, 1) + \phi^2 p_{1,2,3,4}(1, 0, 0, 1))$.

The observed sequence is $\{Y_n\}$, and therefore, results that allow retrieving information about the extreme behavior of the initial sequence $\{X_n\}$, subject to the failures determined by $\{U_n\}$, may be of interest.

If we assume that $\{Y_n\}$ satisfies $D^{(s)}(u_n)$, then $\{X_n\}$ also satisfies $D^{(s)}(u_n)$ by Proposition 3.2 (i), thus coming

$$\tau = \lim_{n \to \infty} nP(X_i < u_n, \ldots, X_{i+1} < u_n)$$

$$= \lim_{n \to \infty} nP(Y_i < u_n, \ldots, Y_{i+1} < u_n)$$

$$= \lim_{n \to \infty} nP(Y_i < u_n, \ldots, Y_{i+1} < Y_i Y_0 \neq Y_i \ldots Y_j).$$

Thereby, we can write

$$\theta = \lim_{n \to \infty} \frac{P(Y_i < u_n, \ldots, Y_{i+1} < u_n | Y_i Y_0 \neq Y_i \ldots Y_j)}{P(Y_i > u_n)}.$$

4 Tail dependence

Now we will analyze the effect of this failure mechanism on the dependency between two variables, $Y_n$ and $Y_{n+m}$, $m \geq 1$. More precisely, we are going to evaluate the lag-$m$ tail dependence coefficient

$$\lambda(Y_{n+m}|Y_n) = \lim_{x \to \infty} P(Y_{n+m} > x | Y_n > x),$$

which incorporates the tail dependence between $X_n$ and $X_{n+j}$, with $j$ regulated by the maximum number of failures $\kappa - 1$ and by the relation between $m$ and $\kappa$. In particular, independent variables present null tail dependence coefficients. If $m = 1$, we obtain the tail dependence coefficient in Joe [5]. For results related to lag-$m$ tail dependence in the literature, see, e.g., Zhang [10,11].

**Proposition 4.1.** Sequence $\{Y_n\}$ has lag-$m$ tail dependence coefficient, with $m \geq 1$,

$$\lambda(Y_{n+m}|Y_n) = p_{1,\ldots,m}(0,\ldots,0)\mathbf{1}_{[m \leq \kappa - 1]} + \sum_{i=1}^{m} \sum_{i'=0}^{\kappa-1} \lambda(X_{n+i+i'}|X_n).$$

provided all coefficients $\lambda(X_{n+i+i'}|X_n)$ exist.

**Proof.** Observe that

$$P(Y_n > x, Y_{n+m} > x)$$

$$= P(Y_n > x, U_{n+1} = 0 = \ldots = U_{n+m}) + \sum_{i=1}^{m} P(Y_n > x, X_{n+i} > x, U_{n+i} = 1, U_{n+i+1} = 0 = \ldots = U_{n+m})$$

$$= \sum_{i=0}^{\kappa-1-m} P(X_{n-i} > x, U_{n-i+1} = 0 = \ldots = U_{n+m})$$

$$+ \sum_{i=1}^{m} \sum_{i'=0}^{\kappa-1} P(X_{n-i} > x, X_{n+i} > x) p_{n-i,n-i+1,\ldots,n+m}(1,0,\ldots,0)\mathbf{1}_{[m \leq \kappa - 1]}$$

and $\sum_{i=0}^{\kappa-1-m} p_{1,\ldots,m+i+1}(1,0,\ldots,0) = p_{1,\ldots,m}(0,\ldots,0).$
Taking \( m = 1 \) in (10), we obtain the tail dependence coefficient

\[
\lambda(Y_{n+1}|Y_n) = p_m(0) + \sum_{i=0}^{k-1} \lambda(X_{n+i+1}|X_n)p_{1+i,2+i+2}(1,0,\ldots,0,1),
\]

provided all coefficients \( \lambda(X_{n+i+1}|X_n) \) exist.

If \( \{X_n\} \) is lag-\( m^* \) tail independent for all integer \( m^* \geq 1 \), we have \( \lambda(X_{n+i+1}|X_n) = 0 \) in the second term of (10), and thus, \( \lambda(Y_{n+m}|Y_n) = p_{1,\ldots,m}(0,\ldots,0)1_{[m<\kappa-1]} \) and \( \{Y_n\} \) is lag-\( m \) tail independent for all integer \( m \geq \kappa \).

**Example 4.1.** Consider again \( \{Y_n\} \) based on ARMAX sequence \( \{X_n\} \) as in Example 3.1. We have that \( \{X_n\} \) has lag-\( m \) tail dependence coefficient \( \lambda(X_{n+m}|X_n) = \phi^m \) (Ferreira and Ferreira [2]), and thus,

\[
\lambda(Y_{n+m}|Y_n) = p_{1,\ldots,m}(0,\ldots,0)1_{[m<\kappa-1]} + \sum_{i=1}^{m} \sum_{i' = \frac{(m-\kappa+1)}{i}}^{k-1} \phi^{i+i'}p_{1,2,\ldots,i',1,i'+1,i'+2,i',\ldots,i'+m}(1,0,\ldots,0,1,0,\ldots,0).
\]

**Acknowledgements:** The authors thank the reviewers for the comments that contributed to the improvement of this work. Helena Ferreira was partially supported by the research unit Centre of Mathematics and Applications of University of Beira Interior UIDB/00212/2020 and UIDP/00212/2020 of Centre of Mathematics of the University of Minho, UIDB/00006/2020 of Centre of Statistics and its Applications of University of Lisbon, UIDB/04621/2020 and UIDP/04621/2020 of Center for Computational and Stochastic Mathematics and PTDC/MAT-STA/28243/2017.

**Conflict of interest:** The authors state no conflict of interest.

**References**


