The upcrossings index and the extremal index

H. Ferreira*
Department of Mathematics
University of Beira Interior
Portugal

Abstract: For stationary sequences $\mathbf{X} = \{X_n\}_{n\geq 1}$ we relate the limiting mean number τ of exceedances of high levels u_n by $X_1, \ldots, X_n$ and the limiting mean number ν of upcrossings, of the same level, throughout $\theta = \frac{\nu}{\tau}\eta$, where θ is the extremal index of $\mathbf{X}$ and η is a new parameter here called the upcrossings index. The upcrossings index is a measure of clustering of upcrossings of $\mathbf{u}$ by variables in $\mathbf{X}$ and the above relation extends the known relation $\theta = \frac{\nu}{\tau}$ which holds under the mild oscillation local restriction $D''(\mathbf{u})$ for $\mathbf{X}$.

We present a new family of local mixing conditions $\tilde{D}^{(k)}(\mathbf{u})$ under which we prove that: a) the intensity of the limiting point process of upcrossings and η can be computed from the k-variate distributions of $\mathbf{X}$; b) the cluster size distributions for the exceedances are asymptotically equivalent to those for lenghs of one run of exceedances or lenghs of several consecutive runs with at most $k-2$ exceedances, except the last one, and separated by at most $k-2$ non-exceedances.

Keywords: Point process of exceedances, point process of upcrossings, extremal index, upcrossings index, mixing conditions.

*Helena Ferreira, Department of Mathematics, University of Beira Interior, 6200 Covilhã, Portugal. E-mail:helena@mat.ubi.pt

1 Introduction

Clustering of exceedances of high levels $\mathbf{u} = \{u_n\}_{n \geq 1}$ by the random variables of a stationary sequence $\mathbf{X} = \{X_n\}_{n \geq 1}$ may occur and, under wide dependence conditions for $\mathbf{X}$, any limiting point process for exceedances is necessarily a compound Poisson point process. Hsing *et al.* (1988) provides a detailed study of the limiting point processes of exceedances under the long range dependence condition $\Delta(\mathbf{u})$ which we now recall.

Definition 1.1 *The sequence* $\mathbf{X}$ *is said to satisfy the condition* $\Delta(\mathbf{u})$ *if* $\alpha_{n,l_n} \underset{n\to\infty}{\longrightarrow} 0$ *for some sequence* $l_n = o(n)$*, where*

$$\alpha_{n,l} = sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathbb{B}_1^k(u_n), B \in \mathbb{B}_{k+l}^n(u_n), 1 \leq k \leq n-l\},$$

and $\mathbb{B}_i^j(u_n)$ *denotes the* σ*-field generated by the events* $\{X_s \leq u_n\}$*,* $i \leq s \leq j$*.*

Let $\mathbb{I}_A$ denotes the indicator of the event A. For applications in the extreme value theory, the main result of Hsing *et al.* (1988) on the sequence of point processes of exceedances $N_n(B) = \sum_{i=1}^{n} \mathbb{I}_{\{X_i > u_n\}} \delta_{i/n}(B)$, $B \subset [0,1]$, can be stated as follows.

Proposition 1.1 *Suppose that* $\Delta(\mathbf{u})$ *holds for* $\mathbf{X}$ *and* $\{N_n\}_{n\geq 1}$ *converges in distribution to some point process* N*. Then* N *is necessarily a compound Poisson process with Laplace transform*

$$L_N(f) = exp\left(-\beta \int_0^1 (1 - \sum_{j=1}^{\infty} \pi(j) e^{-f(x)j}) dx\right),$$

for each non-negative measurable f *on* $[0,1]$*, where*

$$\beta = -log \lim_{n\to+\infty} P(N_n[0,1] = 0) \tag{1.1}$$

and

$$\pi(j) = \lim_{n\to+\infty} P\left(\sum_{i=1}^{r_n} \mathbb{I}_{\{X_i > u_n\}} = j | \sum_{i=1}^{r_n} \mathbb{I}_{\{X_i > u_n\}} > 0\right), j = 1, 2, \ldots, \tag{1.2}$$

for some sequence $r_n = [n/k_n]$ *with* $\{k_n\}$ *satisfying*

$$k_n \underset{n\to\infty}{\longrightarrow} +\infty, \ \frac{k_n l_n}{n} \underset{n\to\infty}{\longrightarrow} 0, \ k_n \alpha_{n,l_n} \underset{n\to\infty}{\longrightarrow} 0. \tag{1.3}$$

Moreover, if (1.1) and (1.2) hold for some sequence $\{k_n\}$ *satisfying (1.3) then* $\{N_n\}_{n\geq 1}$ *converges in distribution to the above compound Poisson process.*

The exceedances of u_n by X_i with $i \in J_{n,j} = \{(j-1)r_n + 1, \ldots, jr_n\}$, for some $j = 1, \ldots, k_n$, are regarded as forming a cluster and

$$\pi_n(j) = P\left(\sum_{i=1}^{r_n} \mathbb{I}_{\{X_i > u_n\}} = j | \sum_{i=1}^{r_n} \mathbb{I}_{\{X_i > u_n\}} > 0\right), \ j = 1, 2, \ldots,$$

is called the distribution of cluster sizes.

The Poisson rate β and the limiting multiplicity distribution π present additional interesting properties for levels $\mathbf{u} \equiv \mathbf{u}^{(\tau)} = \{u_n^{(\tau)}\}_{n\geq 1}$ satisfying $n(1 - F(u_n^{(\tau)})) \underset{n\to\infty}{\longrightarrow} \tau > 0$, which we recall in the following proposition of Hsing *et al.* (1988). Let $N_n^{(\tau)}$ denote the point process of exceedances of $u_n^{(\tau)}$.

Proposition 1.2 *Suppose that, for each* $\tau > 0$, $\Delta(\mathbf{u}^{(\tau)})$ *holds for* $\mathbf{X}$. *If, for some* $\tau_0 > 0$, $\{N_n^{(\tau_0)}\}_{n\geq 1}$ *converges in distribution to some point process* $N^{(\tau_0)}$, *then for all* $\tau > 0$, $\{N_n^{(\tau)}\}_{n\geq 1}$ *converges in distribution to a compound Poisson process with Poisson rate* $\beta = \theta\tau$, $\theta = -log \lim_{n\to+\infty} P(N_n^{(1)}[0,1] = 0)$, $0 \leq \theta \leq \sum_{j\geq 1} j\pi(j) \leq 1$, θ *and* π *being independent of* τ.

The parameter θ is called the extremal index of the sequence $\mathbf{X}$ and was introduced by Leadbetter (1983). Specifically, $\mathbf{X}$ has extremal index θ if, for each $\tau > 0$, there exists $\{u_n^{(\tau)}\}_{n\geq 1}$ and $\lim_{n\to+\infty} P(M_n \leq u_n^{(\tau)}) \equiv \lim_{n\to+\infty} P(N_n^{(\tau)}[0,1] = 0) = e^{-\theta\tau}$, where $M_n = max\{X_1, \ldots, X_n\}$.

Several local dependence conditions provide formulas for the computation of θ from the distribution of a finite number of consecutive variables of $\mathbf{X}$. The family of conditions $D^{(k)}(\mathbf{u}^{(\tau)})$, for $k \geq 1$, considered in Chernick *et al.* (1991) are sufficient to

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

when the limit exists, where $M_{i,j} = max\{X_i, \ldots, X_j\}$ for $i \leq j$ and $M_{i,j} = -\infty$ for $i > j$. The condition $D^{(k)}(\mathbf{u})$ holds for $\mathbf{X}$ when for some k_n as in (1.3),

$$nP(X_1 > u_n \geq M_{2,k}, M_{k+1,r_n} > u_n) \underset{n\to\infty}{\longrightarrow} 0.$$

In particular: $D^{(1)}(\mathbf{u}^{(\tau)}) \equiv D'(\mathbf{u}^{(\tau)})$ leads to $\theta = 1$ (Leadbetter (1974)); $D^{(2)}(\mathbf{u}^{(\tau)})$ is implied by $D''(\mathbf{u}^{(\tau)})$ (Leadbetter and Nandagopalan (1988)) and leads to

$$\theta = \frac{\nu}{\tau}, \tag{1.4}$$

where

$$\nu = \lim_{n\to+\infty} nP(X_1 \leq u_n < X_2), \tag{1.5}$$

that is, the limiting mean number of upcrossings of u_n by the first n variables is equal to θ times the limiting mean number of excedances of u_n by the first n variables of $\mathbf{X}$. Under the condition $D''(\mathbf{u}^{(\tau)})$ we have also, for each $j = 1, 2, \ldots$

$$\pi_n(j) - P(X_2 > u_n, \ldots, X_{j+1} > u_n, X_{j+2} \leq u_n) \underset{n\to\infty}{\longrightarrow} 0, \tag{1.6}$$

that is, a cluster of exceedances is asymptotically a run of exceedances.

Despite the important contributions in the above cited papers, two questions remain without answer when $D''(\mathbf{u}^{(\tau)})$ doesn't hold:

1) How is the limiting mean number of upcrossings of u_n related with the limiting mean number of excedances of u_n?

2) How is the structure of a cluster of exceedances?

The condition $D^{(k)}(\mathbf{u})$ in Chernick *et al.* (1991) extends the $D''(\mathbf{u})$ in a direction which doesn't give sufficient insight into the relation between θ and ν and on what concerns the structure of clusters of exceedances those authors choose not to pursue it (pag. 839 in Chernick *et al.* (1991)).

If we replace exceedances with upcrossings in the condition $D^{(k)}(\mathbf{u})$ we find a new direction to generalize the condition $D''(\mathbf{u})$. Under this new family of local conditions slightly stronger than $D^{(k)}(\mathbf{u})$ we generalize (1.4) and (1.6). We shall prove that the limiting mean number of upcrossings of u_n is related with the limiting mean number of excedances of u_n throughout the extremal index and a new parameter η that indicates the presence of clustering of upcrossings,

$$\theta = \frac{\nu}{\tau}\eta,$$

and show how the runs of exceedances are placed in a cluster. Under the $D''(\mathbf{u})$ condition $\eta = 1$ and we find (1.4) and (1.6) as particular results.

We organize the presentation as follows: section 2 presents results on the convergence of the point process $\tilde{N}_n$ of upcrossings $\{X_i \leq u_n < X_{i+1}\}$ analogous to those above for N_n; section 3 begins with an example showing that (1.4) does not hold in general and more local information than $D^{(k)}(\mathbf{u})$ is needed in order to get a complete description of the clusters and in this section we introduce the upcrossings index η; section 3 shows also that our local dependence condition $\tilde{D}^{(k)}(\mathbf{u})$ is a necessary and sufficient condition to obtain $\lim_{n\to+\infty} P(\tilde{N}_n[0,1]=0)$ from a finite number k of consecutive variables of $\mathbf{X}$ and presents a formula for the computation of η; in section 4 we prove that, under $\tilde{D}^{(k)}(\mathbf{u})$, in a cluster with several runs of exceedances all the runs except the last one have asymptotically at most $k-2$ exceedances and the runs in the same cluster are separated by at most $k-2$ non-exceedances.

2 Point processes of upcrossings

Let the sequence of point processes of upcrossings of u_n by $X_1, \ldots, X_n$ be defined by $\tilde{N}_n(B) = \sum_{i=1}^{n} \mathbb{I}_{\{X_i \leq u_n < X_{i+1}\}} \delta_{i/n}(B)$, $B \subset [0,1]$. We first state a lemma on the asymptotic independence of upcrossings over disjoint blocks $J_{n,i} = \{(i-1)r_n + 1, \ldots, ir_n\}$, $i = 1, \ldots, k_n$, for each k_n satisfying (1.3).

Lemma 2.1 *Suppose that* $\mathbf{X}$ *satisfies the condition* $\Delta(\mathbf{u})$ *and let* k_n *satisfy (1.3). Then*

$$P(\tilde{N}_n[0,1]=0) - P^{k_n}(\tilde{N}_n[0,r_n]=0) \underset{n\to\infty}{\longrightarrow} 0. \tag{2.7}$$

Proof: We can apply the Lemma 2.2 in Hsing *et al.* (1988) with $\chi_{n,i}$ being the indicator of the event $\{X_i \leq u_n < X_{i+1}\}$, $f = 1$, $a_n \underset{n\to\infty}{\longrightarrow} +\infty$ and such that $k_n e^{-a_n} \underset{n\to\infty}{\longrightarrow} 0$, and get

$$E(e^{-a_n\tilde{N}_n[0,1]}) - \prod_{i=1}^{k_n} E(e^{-a_n\tilde{N}_n(J_{n,i})}) \underset{n\to\infty}{\longrightarrow} 0.$$

Then (2.7) follows since $E(e^{-a_n\tilde{N}_n[0,1]}) - P(\tilde{N}_n[0,1]=0) = \sum_{s=1}^{+\infty} e^{-a_n s}P(\tilde{N}_n[0,1]=s) \leq \frac{e^{-a_n}}{1-e^{-a_n}} = o(1)$ and $|\prod_{i=1}^{k_n} E(e^{-a_n\tilde{N}_n(J_{n,i})}) - \prod_{i=1}^{k_n} P(\tilde{N}_n(J_{n,i})=0)| \leq k_n\frac{e^{-a_n}}{1-e^{-a_n}} = o(1)$.

□

The results of sections 3 and 4 in Hsing *et al.* (1988) can be applied to the point processes of upcrossings and we have the following analogous of the above proposition 1.1 for $\tilde{N}_n$.

Proposition 2.1 *Suppose that* $\Delta(\mathbf{u})$ *holds for* $\mathbf{X}$ *and* $\{\tilde{N}_n\}_{n\geq 1}$ *converges in distribution to some point process* $\tilde{N}$. *Then* $\tilde{N}$ *is necessarily a compound Poisson process with Laplace transform*

$$L_{\tilde{N}}(f) = exp\left(-\tilde{\beta}(0,+\infty)\int_0^1 (1 - \sum_{j=1}^{\infty}\tilde{\pi}(j)e^{-f(x)j})dx\right),$$

for each non-negative measurable f *on* $[0,1]$, *where* $\tilde{\beta}(.)$ *is a finite measure concentraded on the positive integers* $\mathbb{N}$;

$$\tilde{\beta} \equiv \tilde{\beta}(0,+\infty) = -log \lim_{n\to+\infty} P(\tilde{N}_n[0,1]=0); \tag{2.8}$$

$$\tilde{\pi}(.) = \tilde{\beta}(.)/\tilde{\beta}(\mathbb{N}) = \lim_{n\to+\infty} \tilde{\pi}_n(.) \tag{2.9}$$

where

$$\tilde{\pi}_n(j) = P\left(\sum_{i=1}^{r_n} \mathbb{I}_{\{X_i\leq u_n<X_{i+1}\}} = j | \sum_{i=1}^{r_n} \mathbb{I}_{\{X_i\leq u_n<X_{i+1}\}} > 0\right),\ j=1,2,\ldots,$$

for some sequence $r_n = [n/k_n]$ *with* $\{k_n\}$ *satisfying (1.3). Moreover, if the convergences in (2.8) and (2.9) hold for some sequence* $\{k_n\}$ *satisfying (1.3) then* $\{\tilde{N}_n\}_{n\geq 1}$ *converges in distribution to the above compound Poisson process.*

Let $\tilde{\mathbf{u}}^{(\nu)} = \{\tilde{u}_n^{(\nu)}\}_{n\geq 1}$ denote a sequence satisfying (1.5) and $\tilde{N}_n(\tilde{u}_n^{(\nu)})$ the corresponding point process of upcrossings of $\tilde{u}_n^{(\nu)}$. In general, for two levels $u_{n,1} \equiv \tilde{u}_{n,1}^{(\nu)}$ and $u_{n,2} \equiv \tilde{u}_{n,2}^{(\nu)}$ we can not guarantee that $P(\tilde{N}_n(\tilde{u}_{n,1}^{(\nu)}) \neq \tilde{N}_n(\tilde{u}_{n,2}^{(\nu)})) \underset{n\to\infty}{\longrightarrow} 0$. Such convergence holds if

$$nP(min\{u_{n,1}, u_{n,2}\} < X_1 \leq max\{u_{n,1}, u_{n,2}\}) \underset{n\to\infty}{\longrightarrow} 0, \tag{2.10}$$

which occurs for instance when the two levels are normalized for the same limiting number of exceedances, that is, $u_{n,1} = u_{n,1}^{(\tau)}$ and $u_{n,2} = u_{n,2}^{(\tau)}$ for some $\tau > 0$. However for $\tilde{u}_n^{(1)}$ and $\tilde{v}_n^{(\nu)} = \tilde{u}_{[n/\nu]}^{(1)}$ we can apply the same arguments used in the proof of proposition 2.1 and get the following result.

Proposition 2.2 *Suppose that, for each* $\nu > 0$, $\Delta(\tilde{\mathbf{u}}^{(\nu)})$ *holds for* $\mathbf{X}$. *If* $\{\tilde{N}_n(\tilde{u}_n^{(1)})\}_{n\geq 1}$ *converges in distribution to some point process* $\tilde{N}^{(1)}$, *then for all* $\nu > 0$, $\{\tilde{N}_n(\tilde{u}_{[n/\nu]}^{(1)})\}_{n\geq 1}$ *converges in distribution to a compound Poisson process with Poisson rate* $\tilde{\beta} = \eta\nu$, *with* $\eta = -log \lim_{n\to+\infty} P(\tilde{N}_n(\tilde{u}_n^{(1)})[0,1] = 0)$,

$$0 \leq \eta \leq \sum_{j\geq 1} j\tilde{\pi}(j) \leq 1, \tag{2.11}$$

η *and* $\tilde{\pi}$ *being independent of* ν. *For each* $\tilde{v}_n^{(\nu)}$ *satisfying (2.10) with respect* $\tilde{u}_{[n/\nu]}^{(1)}$ *the same compound Poisson process arises in the limit of* $\{\tilde{N}_n(\tilde{v}_n^{(\nu)})\}_{n\geq 1}$.

The parameter η, when it exists for each $\tilde{\mathbf{u}}^{(\nu)}$, will be referred as the upcrossing index of $\mathbf{X}$. We formalize this definition in the next section.

3 The condition $\tilde{D}^{(k)}(\mathbf{u})$ and the upcrossings index

Identifying clusters of high level exceedances is a key issue for estimators of θ based on proposition 1.2 and the equality $\theta = \sum_{j\geq 1} j\pi(j)$ (see Ancona-Navarrete and Tawn (2000) and references therein). Should two runs of exceedances separated by one single non-exceedance be considered in the same cluster? Suppose that the sequence $\mathbf{X}$ satisfies the condition $D^{(k)}(\mathbf{u})$, for some $k \geq 3$. We then say that runs in the same cluster must be separated by at most $k-2$ non-exceedances. However, if the sequence satisfies additional local restrictions on the distance between upcrossings, such maximum for the distance between runs can be misleading.

We first consider an example of one max-autoregressive sequence for which $D^{(3)}(\mathbf{u})$ holds but some runs separated by one single exceedance must be considered in different clusters, (1.4) does not hold and, beyond the extremal index, a measure of clustering of upcrossings can be computed.

Example Let $\mathbf{Y} = \{Y_n\}_{n \geq -2}$ be a sequence of independent and uniformly distributed on $[0, 1]$ variables. Let $u_n = 1 - \frac{\tau'}{n}$, $\tau' > 0$, and F denote the common distribution function in $\mathbf{Y}$.

We shall consider a max-autoregressive sequence of the form $max\{Y_{n-t_1}, Y_{n-t_2}, \ldots, Y_{n-t_k}\}$ with non-consecutive fixed integers $t_1, \ldots, t_k$.

Define $\mathbf{X} = \{X_n\}_{n \geq 1}$ by $X_n = max\{Y_n, Y_{n-2}, Y_{n-3}\}$. The sequence $\mathbf{X}$ satisfies the condition $\Delta(\mathbf{u})$ since it is 4-dependent and $D^{(3)}(\mathbf{u})$ since, for all k_n as in (1.3),

$$nP(X_1 > u_n \geq M_{2,3}, M_{4,r_n} > u_n) \leq Anr_n P(Y_1 > u_n, Y_4 > u_n) = Anr_n \overline{F}^2(u_n) \underset{n\to\infty}{\longrightarrow} 0,$$

where A is a constant. It doens't satisfy $D''(\mathbf{u})$ since, for all k_n as in (1.3), we have

$$nP(X_1 > u_n \geq M_{2,2}, M_{3,r_n} > u_n) \geq n\overline{F}(u_n)F^5(u_n) \underset{n\to\infty}{\longrightarrow} \tau > 0.$$

It holds $u_n \equiv u_n^{(\tau)}$ with $\tau = 3\tau'$ and $u_n \equiv \tilde{u}_n^{(\nu)}$ with $\nu = 2\tau'$. This sequence has extremal index

$$\theta = \lim_{n\to+\infty} nP(X_1 > u_n \geq M_{2,3}) / \lim_{n\to+\infty} nP(X_1 > u_n) = \tau'/\tau = 1/3$$

and

$$\nu \neq \theta\tau.$$

Moreover, some runs separated by one non-exceedance must be considered in different clusters since, if we denote $\tilde{N}_n[\frac{i}{n}, \frac{j}{n}]$ simply by $\tilde{N}_{i,j}$, we have

$$nP(X_1 \leq u_n < X_2, \tilde{N}_{3,3} = 0, \tilde{N}_{4,r_n} > 0) =$$

$$nP(X_1 \leq u_n < X_2, X_3 > u_n, \tilde{N}_{4,r_n} > 0) + nP(X_1 \leq u_n < X_2, X_3 \leq u_n, \tilde{N}_{4,r_n} > 0) \leq$$

$$Bnr_n\overline{F}^2(u_n) = o(1).$$

where B is a constant. Therefore, asymptoticaly, the probability of one run with more than one exceedance be followed by another run in the same cluster is negligible, even if they are separeted by one single non-exceedance. In this example we find, for each $\nu > 0$,

$$\lim_{n\to+\infty} P(\tilde{N}_n(\tilde{u}_n^{(\nu)})[0,1] = 0) = e^{-\eta\nu},$$

with $\eta = 1/2$.

We now introduce a local dependence condition which is necessary and sufficient for $\lim_{n\to+\infty} P(\tilde{N}_n[0,1] = 0)$ can be computed by knowledge of the joint distribution of k consecutive terms of $\mathbf{X}$. For this purpose we replace exceedances with upcrossings in the condition $D^{(k)}(\mathbf{u})$. We shall assume that $\tilde{N}_{i,j} = 0$ if $j < i$.

Definition 3.1 *Let* $\mathbf{X}$ *be a sequence satisfying the condition* $\Delta(\mathbf{u})$. *For any* $k \geq 2$, $\mathbf{X}$ *satisfies* $\tilde{D}^{(k)}(\mathbf{u})$ *if*

$$\lim_{n\to+\infty} nP(X_1 \leq u_n < X_2, \tilde{N}_{3,k} = 0, \tilde{N}_{k+1,r_n} > 0) = 0, \tag{3.12}$$

for some sequence $r_n = [n/k_n]$ *with* $\{k_n\}$ *satisfying (1.3).*

If $R_i^{p,q} = \{X_{i+1} > u_n, \ldots, X_{i+p} > u_n, X_{i+p+1} \leq u_n, \ldots, X_{i+p+q} \leq u_n\}$ and $R_i^{p,0} = \{X_{i+1} > u_n, \ldots, X_{i+p} > u_n\}$, for $p \geq 1$, $q \geq 1$, then (3.12) is equivalent to

$$\lim_{n\to+\infty} n \sum_{\substack{p+q=k-1 \\ p\geq 1, q\geq 0}} P(X_1 \leq u_n, R_1^{p,q}, \tilde{N}_{k+1,r_n} > 0) = 0.$$

When $k = 2$ we find the slightly weakened condition $D''(\mathbf{u})$ (pg. 72,73 in Leadbetter and Nandagopalan (1988)).

Proposition 3.1 *Suppose that* $\Delta(\mathbf{u})$ *holds for* $\mathbf{X}$ *and* $\liminf_{n\to+\infty} P(\tilde{N}_n[0,1] = 0) > 0$. *Then, for each positive integer* k,

$$P(\tilde{N}_n[0,1] = 0) - exp(-nP(X_1 \leq u_n < X_2, \tilde{N}_{3,k} = 0)) \underset{n\to\infty}{\longrightarrow} 0 \tag{3.13}$$

if and only if $\mathbf{X}$ *satisfies the condition* $\tilde{D}^{(k)}(\mathbf{u})$.

Proof: In the arguments of O'Brien (1987) to obtain Equation (2.6), if we replace exceedances with upcrossings then we get the analogous convergence result for the probability of no upcrossings:

$$P(\tilde{N}_n[0,1]=0) - exp(-nP(X_1 \leq u_n < X_2, \tilde{N}_{3,r_n}=0)) \underset{n\to\infty}{\longrightarrow} 0.$$

Since

$$nP(X_1 \leq u_n < X_2, \tilde{N}_{3,r_n}=0) = nP(X_1 \leq u_n < X_2, \tilde{N}_{3,k}=0) -$$
$$nP(X_1 \leq u_n < X_2, \tilde{N}_{3,k}=0, \tilde{N}_{k+1,r_n}>0),$$

the convergence in (3.13) holds if and only if $\tilde{D}^{(k)}(\mathbf{u})$ holds. □

We now define the upcrossings index η, which by (2.11) can be viewed as a measure of clustering of upcrossings of $\mathbf{u}$ by variables in $\mathbf{X}$.

Definition 3.2 *If for each $\nu > 0$ there exists $\{\tilde{u}_n^{(\nu)}\}_{n\geq 1}$ and $\lim_{n\to+\infty} P(\tilde{N}_n(\tilde{u}_n^{(\nu)})=0) = e^{-\eta\nu}$, for some constant $0 \leq \eta \leq 1$, then we say that the sequence $\mathbf{X}$ has upcrossings index η.*

If for each $\nu > 0$ there exists $\tilde{u}_n^{(\nu)}$ and $\tilde{u}_n^{(\nu)} = u_n^{(\tau)}$ for some $\tau > 0$, then $P(\tilde{N}_n = 0) - P(N_n = 0) \underset{n\to\infty}{\longrightarrow} 0$ and the upcrossings index η exists if and only if there exists the extremal index θ. In that case,

$$\theta = \frac{\nu}{\tau}\eta.$$

When $\eta = 1$, which holds in particular under $D''(\mathbf{u})$, we find the formula for θ of Leadbetter and Nandagopalan (1988). We generalize such result, under the condition $\tilde{D}^{(k)}(\tilde{\mathbf{u}}^{(\nu)})$, by computing the upcrossings index and the extremal index as follows.

Corollary 3.1 *If, for some $k \geq 2$, $\mathbf{X}$ satisfies $\tilde{D}^{(k)}(\tilde{\mathbf{u}}^{(\nu)})$ for each $\nu > 0$, then the upcrossings index of $\mathbf{X}$ exists and is equal to η if and only if*

$$P(\tilde{N}_{3,k}(\tilde{u}_n^{(\nu)})=0 \mid X_1 \leq \tilde{u}_n^{(\nu)} < X_2) \underset{n\to\infty}{\longrightarrow} \eta,$$

for each $\nu > 0$.

Proof: If η exists then, by (3.13), $P(\tilde{N}_{3,k}(\tilde{u}_n^{(\nu)}) = 0 | X_1 \leq \tilde{u}_n^{(\nu)} < X_2) \underset{n\to\infty}{\longrightarrow} \eta$.

If the last convergence holds then $\liminf_{n\to+\infty} P(\tilde{N}_n(\tilde{u}_n^{(\nu)}) = 0) > 0$. Otherwise we would have $P(\tilde{N}_n(\tilde{u}_n^{(\nu)}) = 0) \to 0$ along some subsequence of $\mathbb{N}$ and then $nP(X_1 \leq \tilde{u}_n^{(\nu)} < X_2, \tilde{N}_{3,k}(\tilde{u}_n^{(\nu)}) = 0) \to +\infty \neq \eta\nu$ along that subsequence. In fact, under $\Delta(\tilde{\mathbf{u}}^{(\nu)})$, $P(X_1 \leq \tilde{u}_n^{(\nu)} < X_2) \underset{n\to\infty}{\longrightarrow} 0$ and $P(\tilde{N}_n(\tilde{u}_n^{(\nu)}) = 0) \underset{n\to\infty}{\longrightarrow} 0$ leads to

$$nP(X_1 \leq \tilde{u}_n^{(\nu)} < X_2, \tilde{N}_{3,k}(\tilde{u}_n^{(\nu)}) = 0) \underset{n\to\infty}{\longrightarrow} +\infty, \tag{3.14}$$

for each $k \geq 2$. To conclude it, we can choose $\{k_n\}$ satisfying (1.3) and $k_n P(X_1 \leq \tilde{u}_n^{(\nu)} < X_2) \underset{n\to\infty}{\longrightarrow} 0$, and, by applying the Lemma 2.1, we have

$$k_n P(\tilde{N}_{r_n}(\tilde{u}_n^{(\nu)}) > 0) \underset{n\to\infty}{\longrightarrow} +\infty. \tag{3.15}$$

Since, for each $k \geq 2$,

$$k_n P(\tilde{N}_{r_n}(\tilde{u}_n^{(\nu)}) > 0) \leq$$
$$nP(X_1 \leq \tilde{u}_n^{(\nu)} < X_2, \tilde{N}_{3,k}(\tilde{u}_n^{(\nu)}) = 0) + k_n(k-2)P(X_1 \leq \tilde{u}_n^{(\nu)} < X_2),$$

then from (3.15) we would have (3.14). Therefore, under the conditions of the corollary and $P(\tilde{N}_{3,k}(\tilde{u}_n^{(\nu)}) = 0 | X_1 \leq \tilde{u}_n^{(\nu)} < X_2) \underset{n\to\infty}{\longrightarrow} \eta$, for each $\nu > 0$, we can apply Proposition 3.1 and conclude that $\lim_{n\to+\infty} P(\tilde{N}_n(\tilde{u}_n^{(\nu)}) = 0) = e^{-\eta\nu}$. □

Corollary 3.2 *Suppose that, for some $k \geq 2$, $\mathbf{X}$ satisfies $\tilde{D}^{(k)}(\tilde{\mathbf{u}}^{(\nu)})$ for each $\nu > 0$, and, for each $\nu > 0$, $\tilde{u}_n^{(\nu)} = u_n^{(\tau)}$ for some $\tau > 0$. Then the extremal index of $\mathbf{X}$ exists and is equal to $\theta = \frac{\nu}{\tau}\eta$ if and only if*

$$P(\tilde{N}_{3,k}(\tilde{u}_n^{(\nu)}) = 0 | X_1 \leq \tilde{u}_n^{(\nu)} < X_2) \underset{n\to\infty}{\longrightarrow} \eta,$$

for each $\nu > 0$.

Define now the point process of cluster positions of upcrossings of $\tilde{u}_n^{(\nu)}$ as $\tilde{N}_n^*(\tilde{u}_n^{(\nu)})(B) = \sum_{i=1}^{k_n} \mathbb{I}_{\{\tilde{N}_n(\tilde{u}_n^{(\nu)})(J_{n,i})>0\}} \delta_{i/n}(B)$, $B \subset [0,1]$, for some k_n as in (1.3). If $\mathbf{X}$ satisfies the condition $\Delta(\tilde{\mathbf{u}}^{(\nu)})$, for each ν, and has upcrossings index η then $\{\tilde{N}_n^*(\tilde{u}_n^{(\nu)})\}_{n\geq 1}$ converges in distribution to a Poisson Process $\tilde{N}^{*(\nu)}$ with intensity parameter $\eta\nu$. In fact, for each $0 \leq a < b \leq 1$,

$$\lim_{n\to+\infty} E(\tilde{N}_n^*(\tilde{u}_n^{(\nu)})(a,b]) =$$

$$\lim_{n\to+\infty} k_n(b-a)P(\tilde{N}_n^*(\tilde{u}_n^{(\nu)})(J_{n,1}) > 0) = (b-a)\eta\nu = E(\tilde{N}^{*(\nu)}(a,b]) \tag{3.16}$$

and, by the condition $\Delta(\tilde{\mathbf{u}}^{(\nu)})$, for each $0 \leq a_1 < b_1 \leq \ldots \leq a_k < b_k \leq 1$,

$$\lim_{n\to+\infty} P(\tilde{N}_n^*(\tilde{u}_n^{(\nu)})(\cup_{i=1}^k (a_i, b_i]) = 0) =$$

$$\lim_{n\to+\infty} \prod_{i=1}^k P(\tilde{N}_n^*(\tilde{u}_n^{(\nu)})(a_i, b_i] = 0) = \prod_{i=1}^k e^{-\eta\nu(b_i - a_i)} =$$

$$\prod_{i=1}^k P(\tilde{N}^{*(\nu)})(a_i, b_i] = 0) = P(\tilde{N}^{*(\nu)}(\cup_{i=1}^k (a_i, b_i]) = 0). \tag{3.17}$$

The convergences in (3.16) and (3.17) are sufficient to conclude that $\tilde{N}_n^*(\tilde{u}_n^{(\nu)})$ converges to the simple point process $\tilde{N}^{*(\nu)}$ (Kallenberg (1976)).

4 Cluster size distributions for exceedances

The cluster size distributions are now investigated for the class of sequences that satisfy the condition $\tilde{D}^{(k)}(\mathbf{u})$.

Under the condition $\tilde{D}^{(2)}(\mathbf{u})$, (1.6) holds as proved in Proposition 3.5 of Leadbetter and Nandagopalan (1988). For $k > 2$, we will use the notation $R_i^{p,q}$ introduced in the previous section to describe how the runs of exceedances are placed in a cluster.

The cluster size distributions for the exceedances are asymptotically equivalent to those for lenghs of one run of exceedances or lenghs of several consecutive runs with at most $k-2$ exceedances, except the last one, and separated by at most $k-2$ non-exceedances.

Proposition 4.1 *Suppose that* $\mathbf{X}$ *satisfies* $\tilde{D}^{(k)}(\mathbf{u})$*, for some* $k > 2$*, and* $\mathbf{u} \equiv \tilde{\mathbf{u}}^{(\nu)}$*, for some* $\nu > 0$*. Then* $\pi_n(j) - \pi_n^*(j) \underset{n\to\infty}{\longrightarrow} 0$ *for each* $j = 1, 2, \ldots$*, where*

$$\pi_n^*(j) = \sum_{s=1}^{j} \sum_{\{p_1,q_1,\ldots,p_s,q_s\}\in S(j)} P(R_1^{p_1,q_1,p_2,q_2,\ldots,p_s,1\vee(k-p_s)} \mid X_1 \leq u_n < X_2), \tag{4.18}$$

$$R_1^{p_1,q_1,p_2,q_2,\ldots,p_s,1\vee(k-p_s)} = R_1^{p_1,q_1} \cap R_{1+p_1+q_1}^{p_2,q_2} \cap \ldots \cap R_{1+\sum_{i\leq s-1} p_i+q_i}^{p_s,1\vee(k-p_s)}$$

and

$$S(j) = \{\{p_1, q_1, \ldots, p_s, q_s\} : \sum_{i=1}^{s} p_i = j, p_i > 0, q_i > 0, max_{i\leq s-1}(p_i + q_i) < k\}.$$

Proof: For each $j = 1, 2, \ldots,$

$$\pi_n(j) = \frac{k_n}{\nu}\left(P(X_1 \leq u_n, N_{r_n} = j, \tilde{N}_{r_n} = 1) + P(X_1 \leq u_n, N_{r_n} = j, \tilde{N}_{r_n} > 1)\right)(1 + o(1)).$$

The first term in the sum can be handed as in the proof of proposition 3.5 in Leadbetter and Nandagopalan (1988) and we get

$$\frac{k_n}{\nu} P(X_1 \leq u_n, N_{r_n} = j, \tilde{N}_{r_n} = 1) =$$

$$\frac{k_n}{n}(r_n - j + 1) P(X_2 > u_n, \ldots, X_{j+1} > u_n, X_{j+2} \leq u_n, \ldots, X_{j+1+(1\vee(k-j))} \leq u_n | X_1 \leq u_n < X_2)(1 + o(1)) =$$

$$P(R_1^{j,1\vee(k-j)} | X_1 \leq u_n < X_2)(1 + o(1)), \qquad (4.19)$$

since $P(X_1 \leq u_n < X_2, R_1^{j,1\vee(k-j)}, N_{j+2+(1\vee(k-j)),r_n} > 0) = o(1/n)$ by the condition $\tilde{D}^{(k)}(\mathbf{u})$.

On what concerns the second term in the initial sum, it holds

$$\frac{k_n}{\nu} P(X_1 \leq u_n, N_{r_n} = j, \tilde{N}_{r_n} > 1) =$$

$$\frac{k_n}{\nu} \sum_{i=1}^{r_n - j + 1} P(X_i \leq u_n < X_{i+1}, N_{i+1,r_n} = j, \tilde{N}_{i,r_n} > 1) =$$

$$\frac{k_n}{\nu} \sum_{i=1}^{r_n - j + 1} \sum_{s=2}^{j} P(X_i \leq u_n < X_{i+1}, N_{i+1,r_n} = j, \tilde{N}_{i,r_n} = s) = a_n +$$

$$\frac{k_n}{\nu} \sum_{i=1}^{r_n - j + 1} \sum_{s=2}^{j} \sum_{\{p_1,q_1,\ldots,p_s,q_s\}\in S(j)} P(X_i \leq u_n < X_{i+1}, R_i^{p_1,q_1}, R_{i+p_1+q_1}^{p_2,q_2}, \ldots, R_{i+\sum_{j\leq s-1} p_j+q_j}^{p_s,1\vee(k-p_s)})$$

$$ \tag{4.20} $$

where a_n is the sum of the terms for which $max_{i=1,\dots,s-1}(p_i+q_i) \geq k$ and therefore

$$a_n \leq \frac{k_n}{\nu}(r_n - j + 1)(j-1)\sharp S(j) P(X_1 \leq u_n < X_2, \tilde{N}_{3,k} = 0, \tilde{N}_{k+1,r_n} > 0) = o(1/n).$$

Now, by using the stationarity and assuming $\sum_{i \leq 0} p_i + q_i = 0 = max_{i \leq 0}(p_i + q_i)$, we can rewrite (4.20) and (4.19) in a single expression and find the result. □

We conclude by remarking that, beyond the examples of max-autoregressive sequences as considered in the beginning of this section, it would be interesting to know if the results under the condition $\tilde{D}^{(k)}(\mathbf{u})$ can be applied to generalized moving averages $\sum_{s=-\infty}^{+\infty} c_{t_s} Y_{j-t_s}$, where $\{t_s\}$ is a sequence of integers.

References

Ancona-Navarrete, M.A. and Tawn, J.A. (2000) A comparison of methods for estimating the extremal index. *Extremes* **3:1**, 5-38.

Chernick, M.R., Hsing, T. and McCormick, W.P. (1991) Calculating the extremal index for a class of stationary sequences. *Adv. Appl. Prob.* **23**, 835-850.

Hsing, T., Husler, J. and Leadbetter, M.R. (1988) On the exceedance point process for a stationary sequence. *Probab. Th. Rel. Fields* **78**, 97-112.

Kallenberg, O. (1976) *Random measures.* Berlin: Akademie-Verlag, London-New York: Academic Press.

Leadbetter, M.R. (1974) On extreme values in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **28**, 289-303.

Leadbetter, M.R. and Nandagopalan, S. (1988) On exceedance point process for stationary sequences under mild oscillation restrictions. In *Extreme Value Theory: Proceedings, Oberwolfach 1987*, ed. J. Husler and D. Reiss, Lecture Notes in Statistics **51**, 69-80. Springer-Verlag, Berlim.

O'Brien, G.L (1987) Extreme values for stationary and Markov sequences. *Ann. Prob.* **15**, 281-291.