

## THE UPCROSSINGS INDEX AND THE EXTREMAL INDEX

H. FERREIRA,\* *University of Beira Interior*

### Abstract

For stationary sequences  $X = \{X_n\}_{n \geq 1}$  we relate  $\tau$ , the limiting mean number of exceedances of high levels  $u_n$  by  $X_1, \dots, X_n$ , and  $\nu$ , the limiting mean number of upcrossings of the same level, through the expression  $\theta = (\nu/\tau)\eta$ , where  $\theta$  is the extremal index of  $X$  and  $\eta$  is a new parameter here called the upcrossings index. The upcrossings index is a measure of the clustering of upcrossings of  $\mathbf{u}$  by variables in  $X$ , and the above relation extends the known relation  $\theta = \nu/\tau$ , which holds under the mild-oscillation local restriction  $D''(\mathbf{u})$  on  $X$ . 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  $\eta$  can both be computed from the  $k$ -variate distributions of  $X$ ; and (b) the cluster size distributions for the exceedances are asymptotically equivalent to those for the lengths of one run of exceedances or the lengths of several consecutive runs which are separated by at most  $k - 2$  nonexceedances and, except for the last one, each contain at most  $k - 2$  exceedances.

*Keywords:* Point process of exceedances; point process of upcrossings; extremal index; upcrossings index; mixing condition

2000 Mathematics Subject Classification: Primary 60F05

Secondary 60G55; 60G70

### 1. Introduction

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

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

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

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

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

Received 3 May 2006; revision received 8 September 2006.

\* Postal address: Department of Mathematics, University of Beira Interior, 6200 Covilhã, Portugal.

Email address: helena@mat.ubi.pt

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

$$L_N(f) = \exp\left(-\beta \int_0^1 \left(1 - \sum_{j=1}^{\infty} \pi(j)e^{-f(x)j}\right) dx\right) \tag{1.1}$$

for each nonnegative, measurable function  $f$  on  $[0, 1]$ , where

$$\beta = -\log \lim_{n \rightarrow \infty} P(N_n([0, 1]) = 0) \tag{1.2}$$

and

$$\pi(j) = \lim_{n \rightarrow \infty} P\left(\sum_{i=1}^{r_n} \mathbf{1}_{\{X_i > u_n\}} = j \mid \sum_{i=1}^{r_n} \mathbf{1}_{\{X_i > u_n\}} > 0\right), \quad j = 1, 2, \dots, \tag{1.3}$$

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

$$k_n \rightarrow \infty, \quad \frac{k_n l_n}{n} \rightarrow 0, \quad k_n \alpha_{n,l_n} \rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{1.4}$$

Moreover, under condition  $\Delta(\mathbf{u})$ , if (1.2) and (1.3) hold for some  $\beta > 0$ , a probability distribution  $\pi$ , and a sequence  $\{k_n\}$  satisfying (1.4), then  $\{N_n\}_{n \geq 1}$  converges in distribution to the above compound Poisson process.

For some  $j = 1, \dots, k_n$ , the exceedances of  $u_n$  by  $X_i$ ,  $i \in J_n, j = \{(j-1)r_n + 1, \dots, jr_n\}$ , are regarded as forming a cluster, and

$$\pi_n(j) = P\left(\sum_{i=1}^{r_n} \mathbf{1}_{\{X_i > u_n\}} = j \mid \sum_{i=1}^{r_n} \mathbf{1}_{\{X_i > u_n\}} > 0\right), \quad j = 1, 2, \dots,$$

is called the distribution of cluster sizes.

Let  $F$  denote the underlying distribution function of each of the  $X_n$ . In terms of the Poisson rate  $\beta$  and the limiting multiplicity distribution  $\pi$ , levels  $\mathbf{u}^{(\tau)} = \{u_n^{(\tau)}\}_{n \geq 1}$  satisfying  $n(1 - F(u_n^{(\tau)})) \rightarrow \tau > 0$  as  $n \rightarrow \infty$  exhibit additional interesting properties. We recall these 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  $\tau > 0$ , condition  $\Delta(\mathbf{u}^{(\tau)})$  holds for  $X$ . 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 Laplace transform given by (1.1)–(1.3), with Poisson rate  $\beta = \theta\tau$  and*

$$\theta = -\log \lim_{n \rightarrow \infty} P(N_n^{(1)}([0, 1]) = 0), \quad 0 \leq \theta \leq \sum_{j \geq 1} j\pi(j) \leq 1,$$

the probability measure  $\pi$  on  $\{1, 2, \dots\}$  and  $\theta$  both being independent of  $\tau$ .

The parameter  $\theta$  is called the extremal index of the sequence  $X$  and was introduced by Leadbetter (1983). Specifically,  $X$  has extremal index  $\theta$  if, for each  $\tau > 0$ , there exists a sequence  $\{u_n^{(\tau)}\}_{n \geq 1}$  such that

$$\lim_{n \rightarrow \infty} P(M_n \leq u_n^{(\tau)}) = \lim_{n \rightarrow \infty} P(N_n^{(\tau)}([0, 1]) = 0) = e^{-\theta\tau},$$

where  $M_n = \max\{X_1, \dots, X_n\}$ .

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

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

to hold, when the limit exists, where  $M_{i,j} = \max\{X_i, \dots, X_j\}$  for  $i \leq j$  and  $M_{i,j} = -\infty$  for  $i > j$ . The condition  $D^{(k)}(\mathbf{u})$  holds for  $X$  when, for some  $k_n$  as in (1.4),

$$n P(X_1 > u_n \geq M_{2,k_n}, M_{k_n+1,r_n} > u_n) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

In particular,  $D^{(1)}(\mathbf{u}^{(\tau)})$  and  $D'(\mathbf{u}^{(\tau)})$  are equivalent conditions which lead to  $\theta = 1$  (Leadbetter (1974)). Furthermore, condition  $D^{(2)}(\mathbf{u}^{(\tau)})$  is implied by condition  $D''(\mathbf{u}^{(\tau)})$  (Leadbetter and Nandagopalan (1989)) and leads to

$$\theta = \frac{\nu}{\tau}, \tag{1.5}$$

where

$$\nu = \lim_{n \rightarrow \infty} n P(X_1 \leq u_n < X_2). \tag{1.6}$$

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

$$\pi_n(j) - P(X_2 > u_n, \dots, X_{j+1} > u_n, X_{j+2} \leq u_n) \rightarrow 0 \quad \text{as } n \rightarrow \infty, \tag{1.7}$$

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

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

1. How is the limiting mean number of upcrossings of  $u_n$  related to the limiting mean number of exceedances of  $u_n$ ?
2. What is the structure of a cluster of exceedances?

The condition  $D^{(k)}(\mathbf{u})$  of Chernick *et al.* (1991) extends condition  $D''(\mathbf{u})$  in a direction which does not give sufficient insight into the relation between  $\theta$  and  $\nu$ , and those authors chose not to pursue the study of the structure of clusters of exceedances (see Chernick *et al.* (1991, p. 839)). If we replace exceedances with upcrossings in condition  $D^{(k)}(\mathbf{u})$  then we find a new direction in which to generalize condition  $D''(\mathbf{u})$ . Under this new family of local conditions, which are slightly stronger than the  $D^{(k)}(\mathbf{u})$ , we generalize (1.5) and (1.7). We shall prove that the limiting mean number of upcrossings of  $u_n$  is related to the limiting mean number of exceedances of  $u_n$  via

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

that is, through the extremal index and a new parameter,  $\eta$ , that indicates the presence of clustering of upcrossings, and show how the runs of exceedances are placed in a cluster. Under condition  $D''(\mathbf{u})$ ,  $\eta = 1$  and we find (1.5) and (1.7) 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 (1.5) does not hold in general and that more local

information than is given by condition  $D^{(k)}(\mathbf{u})$  is needed to obtain a complete description of the clusters. In Section 3 we also introduce the upcrossings index  $\eta$ , show that our local dependence condition  $\tilde{D}^{(k)}(\mathbf{u})$  is necessary and sufficient to calculate  $\lim_{n \rightarrow \infty} P(\tilde{N}_n([0, 1]) = 0)$  from a finite number  $k$  of consecutive variables of  $\mathbf{X}$ , and present a formula for the computation of  $\eta$ . In Section 4 we prove that, under condition  $\tilde{D}^{(k)}(\mathbf{u})$ , in a cluster with several runs of exceedances all the runs are separated by at most  $k - 2$  nonexceedances and, except for the last one, contain asymptotically at most  $k - 2$  exceedances.

### 2. Point processes of upcrossings

Let the sequence of point processes of upcrossings of  $u_n$  by  $X_1, \dots, X_n$  be defined by  $\tilde{N}_n(B) = \sum_{i=1}^n \mathbf{1}_{\{X_i \leq u_n < X_{i+1}\}} \delta_{i/n}(B)$ ,  $B \subset [0, 1]$ . We first state a lemma on the asymptotic independence of upcrossings over disjoint blocks  $J_{n,i} = \{(i - 1)r_n + 1, \dots, ir_n\}$ ,  $i = 1, \dots, k_n$ , for each  $k_n$  satisfying (1.4).

**Lemma 2.1.** *Suppose that  $\mathbf{X}$  satisfies condition  $\Delta(\mathbf{u})$  and let the sequence  $\{k_n\}$  satisfy (1.4). Then*

$$P(\tilde{N}_n([0, 1]) = 0) - P^{k_n}(\tilde{N}_n([0, r_n/n]) = 0) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{2.1}$$

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

$$E(e^{-a_n \tilde{N}_n([0,1])}) - \prod_{i=1}^{k_n} E(e^{-a_n \tilde{N}_n(J_{n,i}/n)}) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Equation (2.1) then 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

$$\left| \prod_{i=1}^{k_n} E(e^{-a_n \tilde{N}_n(J_{n,i}/n)}) - \prod_{i=1}^{k_n} P\left(\tilde{N}_n\left(\frac{J_{n,i}}{n}\right) = 0\right) \right| \leq k_n \frac{e^{-a_n}}{1 - e^{-a_n}} = o(1).$$

The results of Sections 3 and 4 of Hsing *et al.* (1988) can be applied to the point processes of upcrossings and yield the following analogue of Proposition 1.1 for  $\tilde{N}_n$ .

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

$$L_{\tilde{N}}(f) = \exp\left(-\tilde{\beta}((0, \infty)) \int_0^1 \left(1 - \sum_{j=1}^{\infty} \tilde{\pi}(j) e^{-f(x)j}\right) dx\right) \tag{2.2}$$

for each nonnegative, measurable function  $f$  on  $[0, 1]$ , where

$$\tilde{\beta} \equiv \tilde{\beta}((0, \infty)) = -\log \lim_{n \rightarrow \infty} P(\tilde{N}_n([0, 1]) = 0) \tag{2.3}$$

is a finite measure concentrated on the positive integers  $\mathbb{N}$  and

$$\tilde{\pi}(\cdot) = \tilde{\beta}(\cdot) / \tilde{\beta}(\mathbb{N}) = \lim_{n \rightarrow \infty} \tilde{\pi}_n(\cdot), \tag{2.4}$$

where

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

for some sequence  $\{r_n = \lfloor n/k_n \rfloor\}$  with  $\{k_n\}$  satisfying (1.4). Moreover, under condition  $\Delta(\mathbf{u})$ , if the limits in (2.3) and (2.4) hold for some  $\tilde{\beta} > 0$ , a probability distribution  $\tilde{\pi}$ , and a sequence  $\{k_n\}$  satisfying (1.4), then  $\{\tilde{N}_n\}_{n \geq 1}$  converges in distribution to the above compound Poisson process.

Let  $\tilde{\mathbf{u}}^{(v)} = \{\tilde{u}_n^{(v)}\}_{n \geq 1}$  denote a sequence satisfying (1.6) and let  $\tilde{N}_n(\tilde{\mathbf{u}}^{(v)})$  denote the corresponding point process of upcrossings of  $\tilde{u}_n^{(v)}$ . In general, for two normalized levels with the same limiting number of upcrossings,  $u_{n,1} \equiv \tilde{u}_{n,1}^{(v)}$  and  $u_{n,2} \equiv \tilde{u}_{n,2}^{(v)}$ , we cannot guarantee that  $\mathbb{P}(\tilde{N}_n(\tilde{u}_{n,1}^{(v)}) \neq \tilde{N}_n(\tilde{u}_{n,2}^{(v)})) \rightarrow 0$  as  $n \rightarrow \infty$ . This convergence holds if

$$n \mathbb{P}(\min\{u_{n,1}, u_{n,2}\} < X_1 \leq \max\{u_{n,1}, u_{n,2}\}) \rightarrow 0 \quad \text{as } n \rightarrow \infty, \tag{2.5}$$

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

**Proposition 2.2.** *Suppose that, for each  $v > 0$ , condition  $\Delta(\tilde{\mathbf{u}}^{(v)})$  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  $v > 0$ ,  $\{\tilde{N}_n(\tilde{u}_{\lfloor n/v \rfloor}^{(1)})\}_{n \geq 1}$  converges in distribution to a compound Poisson process with Laplace transform given by (2.2)–(2.4), with Poisson rate  $\tilde{\beta} = \eta v$  and*

$$\eta = -\log \lim_{n \rightarrow \infty} \mathbb{P}(\tilde{N}_n(\tilde{u}_n^{(1)})([0, 1]) = 0), \quad 0 \leq \eta \leq \sum_{j \geq 1} j \tilde{\pi}(j) \leq 1, \tag{2.6}$$

the probability distribution  $\tilde{\pi}$  on  $\{1, 2, \dots\}$  and  $\eta$  both being independent of  $v$ . If  $\tilde{v}_n^{(v)}$  and  $\tilde{u}_{\lfloor n/v \rfloor}^{(1)}$  satisfy (2.5) then the same compound Poisson process arises in the limit of  $\{\tilde{N}_n(\tilde{v}_n^{(v)})\}_{n \geq 1}$ .

The parameter  $\eta$ , when it exists for each  $\tilde{\mathbf{u}}^{(v)}$ , will be referred to as the upcrossings 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  $\theta$  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 nonexceedance be considered parts of 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  $k - 2$  nonexceedances. However, if the sequence satisfies additional local restrictions on the distances between upcrossings, such maxima for the distances between runs can be misleading.

We first consider an example of a max-autoregressive sequence in which condition  $D^{(3)}(\mathbf{u})$  holds but some runs separated by a single exceedance must be considered in different clusters, (1.5) does not hold, and, beyond the extremal index, a measure of clustering of upcrossings can be computed.

**Example 3.1.** Let  $Y = \{Y_n\}_{n \geq -2}$  be a sequence of independent variables uniformly distributed on  $[0, 1]$ . Let  $u_n = 1 - \tau'/n$  for  $\tau' > 0$ , and let  $F$  denote the common distribution function of the  $Y_n$ .

We shall consider a max-autoregressive sequence of the form  $\max\{Y_{n-t_1}, Y_{n-t_2}, \dots, Y_{n-t_k}\}$ , for nonconsecutive fixed integers  $t_1, \dots, t_k$ . Define  $X = \{X_n\}_{n \geq 1}$  by

$$X_n = \max\{Y_n, Y_{n-2}, Y_{n-3}\}.$$

The sequence  $X$  satisfies condition  $\Delta(\mathbf{u})$  since it is four-dependent, and satisfies condition  $D^{(3)}(\mathbf{u})$  since, for all  $k_n$  as in (1.4),

$$n \mathbb{P}(X_1 > u_n \geq M_{2,3}, M_{4,r_n} > u_n) \leq nr_n \mathbb{P}(Y_1 > u_n, Y_4 > u_n) = nr_n \bar{F}^2(u_n) \rightarrow 0$$

as  $n \rightarrow \infty$ ,

where  $\bar{F} = 1 - F$ . However, it does not satisfy condition  $D''(\mathbf{u})$  since, for all  $k_n$  as in (1.4),

$$n \mathbb{P}(X_1 > u_n \geq M_{2,2}, M_{3,r_n} > u_n) \geq n \bar{F}(u_n) F^3(u_n) \rightarrow \tau > 0 \quad \text{as } n \rightarrow \infty.$$

We have  $u_n \equiv u_n^{(\tau)}$  for  $\tau = 3\tau'$  and  $u_n \equiv \tilde{u}_n^{(\nu)}$  for  $\nu = 2\tau'$ . The sequence  $X$  has extremal index

$$\theta = \frac{\lim_{n \rightarrow \infty} n \mathbb{P}(X_1 > u_n \geq M_{2,3})}{\lim_{n \rightarrow \infty} n \mathbb{P}(X_1 > u_n)} = \frac{\tau'}{\tau} = \frac{1}{3},$$

and  $\nu \neq \theta\tau$ . Moreover, some runs separated by one nonexceedance must be considered in different clusters since, denoting  $\tilde{N}_n([i/n, j/n])$  simply by  $\tilde{N}_{i,j}$ , we have

$$\begin{aligned} n \mathbb{P}(X_1 \leq u_n < X_2, \tilde{N}_{3,3} = 0, \tilde{N}_{4,r_n} > 0) &= n \mathbb{P}(X_1 \leq u_n < X_2, X_3 > u_n, \tilde{N}_{4,r_n} > 0) \\ &\quad + n \mathbb{P}(X_1 \leq u_n < X_2, X_4 \leq u_n, \tilde{N}_{4,r_n} > 0) \\ &\leq Bnr_n \bar{F}^2(u_n) \\ &= o(1), \end{aligned}$$

where  $B > 0$  is a constant. Therefore, asymptotically, the probability of one run with more than one exceedance being followed by another run in the same cluster is negligible, even if they are separated by a single nonexceedance. This means that the example satisfies condition  $\tilde{D}^{(3)}(\mathbf{u})$ , which is formalized in Definition 3.1.

In this example, for each  $\nu > 0$  we find that

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

with  $\eta = \frac{1}{2}$ .

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

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

$$\lim_{n \rightarrow \infty} n \mathbb{P}(X_1 \leq u_n < X_2, \tilde{N}_{3,k} = 0, \tilde{N}_{k+1,r_n} > 0) = 0 \tag{3.1}$$

for some sequence  $\{r_n = [n/k_n]\}$  with  $\{k_n\}$  satisfying (1.4).

Let

$$R_i^{p,q} = \{X_{i+1} > u_n, \dots, X_{i+p} > u_n, X_{i+p+1} \leq u_n, \dots, X_{i+p+q} \leq u_n\}$$

and

$$R_i^{p,0} = \{X_{i+1} > u_n, \dots, X_{i+p} > u_n\}$$

for  $p \geq 1$  and  $q \geq 1$ . Then (3.1) is equivalent to

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

For  $k = 2$  we find that  $X$  satisfies the slightly weakened condition  $D''(\mathbf{u})$  (see Leadbetter and Nandagopalan (1989, pp. 72–73)).

**Proposition 3.1.** *Suppose that condition  $\Delta(\mathbf{u})$  holds for  $X$  and that*

$$\liminf_{n \rightarrow \infty} P(\tilde{N}_n([0, 1]) = 0) > 0.$$

Then, for each positive integer  $k$ ,

$$P(\tilde{N}_n([0, 1]) = 0) - \exp(-n P(X_1 \leq u_n < X_2, \tilde{N}_{3,k} = 0)) \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad (3.2)$$

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

*Proof.* If, in the arguments used by O’Brien (1987) to obtain his Equation (2.6), we replace exceedances with upcrossings, then we obtain the analogous convergence result for the probability of no upcrossings:

$$P(\tilde{N}_n([0, 1]) = 0) - \exp(-n P(X_1 \leq u_n < X_2, \tilde{N}_{3,r_n} = 0)) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Since

$$\begin{aligned} n P(X_1 \leq u_n < X_2, \tilde{N}_{3,r_n} = 0) &= n P(X_1 \leq u_n < X_2, \tilde{N}_{3,k} = 0) \\ &\quad - n P(X_1 \leq u_n < X_2, \tilde{N}_{3,k} = 0, \tilde{N}_{k+1,r_n} > 0), \end{aligned}$$

the convergence in (3.2) holds if and only if condition  $\tilde{D}^{(k)}(\mathbf{u})$  holds.

We now define the upcrossings index  $\eta$ , which by (2.6) can be viewed as a measure of the clustering of upcrossings of  $\mathbf{u}$  by variables of  $X$ .

**Definition 3.2.** If, for each  $\nu > 0$ , there exists a sequence  $\{\tilde{u}_n^{(\nu)}\}_{n \geq 1}$  such that

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

for some constant  $\eta$ ,  $0 \leq \eta \leq 1$ , then we say that the sequence  $X$  has upcrossings index  $\eta$ .

If, for each  $\nu > 0$ , there exists a  $\tilde{u}_n^{(\nu)}$  such that  $\tilde{u}_n^{(\nu)} = u_n^{(\tau)}$  for some  $\tau > 0$ , then

$$P(\tilde{N}_n(\tilde{u}_n^{(\nu)}) = 0) - P(N_n(u_n^{(\tau)}) = 0) \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

and the upcrossings index  $\eta$  exists if and only if the extremal index  $\theta$  exists. In that case,

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

For  $\eta = 1$ , which holds in particular under condition  $D''(\mathbf{u})$ , we find the formula for  $\theta$  of Leadbetter and Nandagopalan (1989). We generalize this result under condition  $\tilde{D}^{(k)}(\tilde{\mathbf{u}}^{(v)})$  by computing the upcrossings index and the extremal index as follows.

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

$$P(\tilde{N}_{3,k}(\tilde{u}_n^{(v)}) = 0 \mid X_1 \leq \tilde{u}_n^{(v)} < X_2) \rightarrow \eta \text{ as } n \rightarrow \infty$$

for each  $v > 0$ .

*Proof.* If  $\eta$  exists then, by (3.2),

$$P(\tilde{N}_{3,k}(\tilde{u}_n^{(v)}) = 0 \mid X_1 \leq \tilde{u}_n^{(v)} < X_2) \rightarrow \eta \text{ as } n \rightarrow \infty.$$

If this convergence holds then  $\liminf_{n \rightarrow \infty} P(\tilde{N}_n(\tilde{u}_n^{(v)}) = 0) > 0$ , since otherwise we would have  $P(\tilde{N}_n(\tilde{u}_n^{(v)}) = 0) \rightarrow 0$  along some subsequence of  $\mathbb{N}$  and, thus,

$$n P(X_1 \leq \tilde{u}_n^{(v)} < X_2, \tilde{N}_{3,k}(\tilde{u}_n^{(v)}) = 0) \rightarrow \infty \neq \eta\nu$$

along that subsequence. In fact, under condition  $\Delta(\tilde{\mathbf{u}}^{(v)})$ , the facts that  $P(X_1 \leq \tilde{u}_n^{(v)} < X_2) \rightarrow 0$  and  $P(\tilde{N}_n(\tilde{u}_n^{(v)}) = 0) \rightarrow 0$  as  $n \rightarrow \infty$  imply that

$$n P(X_1 \leq \tilde{u}_n^{(v)} < X_2, \tilde{N}_{3,k}(\tilde{u}_n^{(v)}) = 0) \rightarrow \infty \text{ as } n \rightarrow \infty \tag{3.3}$$

for each  $k \geq 2$ . To prove this, we can choose a sequence  $\{k_n\}$  satisfying (1.4) such that  $k_n P(X_1 \leq \tilde{u}_n^{(v)} < X_2) \rightarrow 0$  as  $n \rightarrow \infty$ , and, from Lemma 2.1, find that

$$k_n P(\tilde{N}_{r_n}(\tilde{u}_n^{(v)}) > 0) \rightarrow \infty \text{ as } n \rightarrow \infty. \tag{3.4}$$

Since, for each  $k \geq 2$ ,

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

we obtain (3.3) from (3.4). Therefore, under the conditions of the corollary, and because  $P(\tilde{N}_{3,k}(\tilde{u}_n^{(v)}) = 0 \mid X_1 \leq \tilde{u}_n^{(v)} < X_2) \rightarrow \eta$  as  $n \rightarrow \infty$ , for each  $v > 0$ , we can apply Proposition 3.1 to conclude that  $\lim_{n \rightarrow \infty} P(\tilde{N}_n(\tilde{u}_n^{(v)}) = 0) = e^{-\eta\nu}$ .

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

$$P(\tilde{N}_{3,k}(\tilde{u}_n^{(v)}) = 0 \mid X_1 \leq \tilde{u}_n^{(v)} < X_2) \rightarrow \eta \text{ as } n \rightarrow \infty$$

for each  $v > 0$ .

Now define the point process of cluster positions of upcrossings of  $\tilde{u}_n^{(v)}$  as

$$\tilde{N}_n^*(\tilde{u}_n^{(v)})(B) = \sum_{i=1}^{k_n} \mathbf{1}_{\{\tilde{N}_n(\tilde{u}_n^{(v)})(J_{n,i}) > 0\}} \delta_{i/n}(B), \quad B \subset [0, 1],$$

for some  $k_n$  as in (1.4). If  $X$  satisfies condition  $\Delta(\tilde{u}^{(v)})$  for each  $v$  and has upcrossings index  $\eta$ , then  $\{\tilde{N}_n^*(\tilde{u}_n^{(v)})\}_{n \geq 1}$  converges in distribution to a Poisson process  $\tilde{N}^{*(v)}$  with intensity parameter  $\eta v$ . In fact, for each  $a$  and  $b$ ,  $0 \leq a < b \leq 1$ , we have

$$\begin{aligned} \lim_{n \rightarrow \infty} E(\tilde{N}_n^*(\tilde{u}_n^{(v)})([a, b])) &= \lim_{n \rightarrow \infty} k_n(b - a) P(\tilde{N}_n^*(\tilde{u}_n^{(v)})(J_{n,1}) > 0) \\ &= (b - a)\eta v \\ &= E(\tilde{N}^{*(v)}([a, b])) \end{aligned} \tag{3.5}$$

and, by condition  $\Delta(\tilde{u}^{(v)})$ , for each  $a_i$  and  $b_i$ ,  $0 \leq a_1 < b_1 \leq \dots \leq a_k < b_k \leq 1$ , we have

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(\tilde{N}_n^*(\tilde{u}_n^{(v)})\left(\bigcup_{i=1}^k (a_i, b_i]\right) = 0\right) &= \lim_{n \rightarrow \infty} \prod_{i=1}^k P(\tilde{N}_n^*(\tilde{u}_n^{(v)})([a_i, b_i]) = 0) \\ &= \prod_{i=1}^k e^{-\eta v(b_i - a_i)} \\ &= \prod_{i=1}^k P(\tilde{N}^{*(v)}([a_i, b_i]) = 0) \\ &= P\left(\tilde{N}^{*(v)}\left(\bigcup_{i=1}^k (a_i, b_i]\right) = 0\right). \end{aligned} \tag{3.6}$$

The limits in (3.5) and (3.6) are sufficient to conclude that  $\tilde{N}_n^*(\tilde{u}_n^{(v)})$  converges to the simple point process  $\tilde{N}^{*(v)}$  (Kallenberg (1976)).

#### 4. Cluster size distributions for exceedances

We now investigate the cluster size distributions for the class of sequences that satisfy condition  $\tilde{D}^{(k)}(\mathbf{u})$ . Under condition  $\tilde{D}^{(2)}(\mathbf{u})$ , (1.7) holds, as proved in Proposition 3.5 of Leadbetter and Nandagopalan (1989). 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 the lengths of one run of exceedances or the lengths of several consecutive runs which are separated by at most  $k - 2$  nonexceedances and, except for the last one, each contain  $k - 2$  exceedances.

**Proposition 4.1.** *Suppose that  $X$  satisfies condition  $\Delta(\mathbf{u})$  and condition  $\tilde{D}^{(k)}(\mathbf{u})$  for some  $k > 2$ , and that  $\mathbf{u} \equiv \tilde{u}^{(v)}$  for some  $v > 0$ . Then, for each  $j = 1, 2, \dots$ ,  $\pi_n(j) - \pi_n^*(j) \rightarrow 0$  as  $n \rightarrow \infty$ , where*

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

with

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

and

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

*Proof.* For each  $j = 1, 2, \dots$ ,

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

The first term in the sum can be handled as in the proof of Proposition 3.5 of Leadbetter and Nandagopalan (1989), and we obtain

$$\begin{aligned} & \frac{k_n}{v} \mathbb{P}(X_1 \leq u_n, N_{r_n} = j, \tilde{N}_{r_n} = 1) \\ &= \frac{k_n}{n} (r_n - j + 1) \mathbb{P}(X_2 > u_n, \dots, X_{j+1} > u_n, X_{j+2} \leq u_n, \dots, \\ & \quad X_{j+1+(1 \vee (k-j))} \leq u_n \mid X_1 \leq u_n < X_2)(1 + o(1)) \\ &= \mathbb{P}(R_1^{j, 1 \vee (k-j)} \mid X_1 \leq u_n < X_2)(1 + o(1)), \end{aligned} \tag{4.1}$$

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

For the second term in the initial sum, we have

$$\begin{aligned} & \frac{k_n}{v} \mathbb{P}(X_1 \leq u_n, N_{r_n} = j, \tilde{N}_{r_n} > 1) \\ &= \frac{k_n}{v} \sum_{i=1}^{r_n-j+1} \mathbb{P}(X_i \leq u_n < X_{i+1}, N_{i+1, r_n} = j, \tilde{N}_{i, r_n} > 1) \\ &= \frac{k_n}{v} \sum_{i=1}^{r_n-j+1} \sum_{s=2}^j \mathbb{P}(X_i \leq u_n < X_{i+1}, N_{i+1, r_n} = j, \tilde{N}_{i, r_n} = s) \\ &= a_n \\ & \quad + \frac{k_n}{v} \sum_{i=1}^{r_n-j+1} \sum_{s=2}^j \sum_{p_1, q_1, \dots, p_s, q_s \in S(j)} \mathbb{P}(X_i \leq u_n < X_{i+1}, R_i^{p_1, q_1}, R_{i+p_1+q_1}^{p_2, q_2}, \dots, \\ & \quad R_{i+\sum_{j \leq s-1} p_j + q_j}^{p_s, 1 \vee (k-p_s)}). \end{aligned} \tag{4.2}$$

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

$$a_n \leq \frac{k_n}{v} (r_n - j + 1)(j - 1) \text{card } S(j)^c \mathbb{P}(X_1 \leq u_n < X_2, \tilde{N}_{3, k} = 0, \tilde{N}_{k+1, r_n} > 0) = o\left(\frac{1}{n}\right),$$

where  $S(j)^c$  denotes the complement of  $S(j)$ . Using stationarity and assuming that  $\sum_{i \leq 0} p_i + q_i = 0 = \max_{i \leq 0} (p_i + q_i)$ , we can now rewrite (4.2) and (4.1) in a single expression and obtain the result.

We conclude by remarking that, beyond the examples of max-autoregressive sequences like those considered at the beginning of this section, it would be interesting to know if the results valid under 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.

### Acknowledgement

I am grateful to the referee for his corrections and rigorous report.

### References

- ANCONA-NAVARRETE, M. A. AND TAWN, J. A. (2000). A comparison of methods for estimating the extremal index. *Extremes* **3**, 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., HÜSLER, J. AND LEADBETTER, M. R. (1988). On the exceedance point process for a stationary sequence. *Prob. Theory Relat. Fields* **78**, 97–112.
- KALLENBERG, O. (1976). *Random Measures*. Academic Press, London.
- LEADBETTER, M. R. (1974). On extreme values in stationary sequences. *Z. Wahrscheinlichkeitsth.* **28**, 289–303.
- LEADBETTER, M. R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrscheinlichkeitsth.* **65**, 291–306.
- LEADBETTER, M. R. AND NANDAGOPALAN, S. (1989). On exceedance point processes for stationary sequences under mild oscillation restrictions. In *Extreme Value Theory* (Oberwolfach 1987; Lecture Notes Statist. **51**), eds J. Hüslér and D. Reiss, Springer, Berlin, pp. 69–80.
- O'BRIEN, G. L. (1987). Extreme values for stationary and Markov sequences. *Ann. Prob.* **15**, 281–291.