Volume 99, Number 4, July–August 1994
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 99, 543 (1994)]

On the Multivariate Extremal Index

Volume 99 Number 4 July–August 1994

S. Nandagopalan

Colorado State University,
Fort Collins, CO 80523

The exceedance point process approach of Hsing et al. is extended to multivariate stationary sequences and some weak convergence results are obtained. It is well known that under general mixing assumptions, high level exceedances typically have a limiting Compound Poisson structure where multiple events are caused by the clustering of exceedances. In this paper we explore (a) the precise effect of such clustering on the limit, and (b) the relationship between point process convergence and the limiting behavior of maxima. Following this, the notion of multivariate extremal index is introduced which is shown to have properties analogous to its univariate counterpart. Two examples of bivariate moving average sequences are presented for which the extremal index is calculated in some special cases.

Key words: dependence function; exceedance; extremal index; multivariate; point process; stationary.

Accepted: March 22, 1994

1. Introduction

Extreme value theory for multivariate iid sequences has been studied for quite some time now but attention to the dependent case has been relatively recent. For univariate sequences it is known that local dependence causes extreme values to occur in clusters, which in turn results in a stochastically smaller distribution for the maximum than if the observations were independent. We begin with a brief review of these results, which we shall later extend to the multivariate case.

Let $\{\xi_n\}$ be a univariate stationary sequence. Write $M_n=\max\{\xi_1,...,\xi_n\}$ and for $\tau>0$, let $\{u_n(\tau)\}$ denote a sequence satisfying $\lim_{n\to\infty} nP\{\xi_1>u_n(\tau)\}=\tau$. Under quite general mixing assumptions there exist constants $0\leq\theta'\leq\theta''\leq 1$ such that

$$\limsup_{n\to\infty} P\{M_n\leq u_n(\tau)\}=e^{-\theta'\tau} \text{ and}$$

$$\liminf_{n\to\infty} P\{M_n\leq u_n(\tau)\}=e^{-\theta''\tau}$$

for all τ. (See Ref. [1], although the idea actually dates back to Refs. [2–4].) Thus if $P\{M_n\leq u_n(\tau_0)\}$ converges for some τ_0, then $\theta'=\theta''(=\theta$, say) and hence $\lim_{n\to\infty} P\{M_n\leq u_n(\tau)\}=e^{-\theta\tau}$ for all $\tau>0$. The common value θ is then called the *extremal index* of $\{\xi_n\}$. We shall assume θ to be positive whenever it exists, since the case $\theta=0$ corresponds to a degenerate limiting distribution for M_n. Note that $\theta=1$ for iid sequences. Let $\{\hat{\xi}_n\}$ be an iid sequence with $\hat{\xi}_1 \stackrel{d}{=} \xi_1$, called the *associated iid sequence*, and write $\hat{M}_n=\max\{\hat{\xi}_1,...,\hat{\xi}_n\}$. If $\{\xi_n\}$ has extremal index θ and $\lim_{n\to\infty} P\{M_n\leq v_n(t)\}=H(t)$ for a suitable family of normalizing constants $\{v_n(t)\}$, then it follows (upon identifying $e^{-\theta\tau}$ with $H(t)$) that $\lim_{n\to\infty} P\{\hat{M}_n\leq v_n(t)\}=\hat{H}(t)$ where

$$H(t)=\hat{H}(t)^{\theta}. \tag{1}$$

The extremal index is thus a measure of the effect of dependence on the limiting distribution of M_n. The stochastically smaller limiting distribution of M_n is in fact a direct result of the clustering of extremes, as explained below. See Ref. [5] for details.

For fixed $\tau>0$ let the *exceedance point process* $N_n=N_n^{(\tau)}$ be defined by

$$N_n(B)=\sum_{i=1}^{n} I_{\{\xi_i>u_n(\tau)\}}I_{\{i/n\in B\}}, \quad B\subset[0,1],$$

where I_A denotes the indicator function of the event A. Then for a broad class of weakly dependent sequences, the limit in distribution of N_n, if it exists, is a Compound Poisson process with intensity $\theta\tau$ and multiplicity distribution π on $\{1,2,...\}$. The Poisson events may in fact be regarded as the positions of "exceedance clusters" while the multiplicities correspond to cluster sizes. More explicitly, one may divide the n observations into k_n blocks of roughly equal size and regard exceedances within each block as forming a single "cluster", so that the cluster sizes are given by $N_n(J_i)$, $i=1,...,k_n$, where $J_i(=J_{n,i})=(\frac{i-1}{k_n},\frac{i}{k_n}]$. For a suitable choice of k_n depending on the mixing rate of $\{\xi_n\}$, one then has

$$\lim_{n\to\infty} P\{N_n(J_1)=j \mid N_n(J_1)>0\}=\pi(j), \quad j\geq 1,$$

and

$$\lim_{n\to\infty} P^{k_n}\{N_n(J_1)=0\}=\lim_{n\to\infty} P\{N_n[0,1]=0\}$$
$$=\lim_{n\to\infty} P\{M_n\leq u_n(\tau)\}=e^{-\theta\tau},$$

so that in particular, $\lim_{n\to\infty} k_n P\{N_n(J_1)>0\}=\theta\tau$. Hence

$$\lim_{n\to\infty} E\,N_n[0,1]=\lim_{n\to\infty} k_n E\,N_n(J_1)$$
$$=\lim_{n\to\infty} k_n E(N_n(J_1)\mid N_n(J_1)>0)P\{N_n(J_1)>0\}$$
$$=\theta\tau \lim_{n\to\infty} E(N_n(J_1)\mid N_n(J_1)>0),$$

while on the other hand, $\lim_{n\to\infty} E\,N_n[0,1]=\lim_{n\to\infty} nP\{\xi_1>u_n(\tau)\}=\tau$. The cluster size distribution and the extremal index are therefore related by

$$\lim_{n\to\infty} E(N_n(J_1)\mid N_n(J_1)>0)=1/\theta. \tag{2}$$

Now let $\{\xi_n=(\xi_{n1},...,\xi_{nd}),\ n\in\mathbb{Z}\}$ be a multivariate stationary sequence where $d\geq 1$ is a fixed integer, and write $M_n=(M_{n1},...,M_{nd})$ where $M_{nj}=\max\{\xi_{1j},...,\xi_{nj}\}$, $j=1,...,d$. The study of multivariate extremes began in the early 1950s, focusing mainly on the limiting behavior of M_n under a linear normalization, when the observations are iid. The resulting class of limiting distributions was characterized in Ref. [6] and domains of attraction criteria were given in Ref. [7]. See also Ref. [8], Chapter 5, for an account of the literature surrounding this theory. For stationary sequences satisfying a general mixing assumption, it is known (see Refs. [9, 10], and Theorem 1.1 below) that the class of limiting distributions of M_n is the same as for iid sequences. In this paper we explore the precise effect of dependence on the limiting distribution by extending the univariate theory described above to the multivariate case. Essentially, this involves studying the inter-relationship between the two dependence structures present, one due to dependence over time and the other due to the dependence between the various components of the multivariate observations. The ideas become most transparent when presented in terms of so-called dependence functions [8]. Here we adopt the slightly modified definition found in Ref. [9]. A distribution function D on $[0,1]^d$ is called a *dependence function* if $D_j(D_j(u))=D_j(u)$, $u\in[0,1]$, $j=1,...,d$, where the subscript j signifies the jth marginal. The dependence function of a distribution F on $\mathbb{R}^d$ is defined by $D_F(\mathbf{u})=P\{F_1(X_1)\leq u_1,...,F_d(X_d)\leq u_d\}$, $\mathbf{u}=(u_1,...,u_d)\in[0,1]^d$, where $(X_1,...,X_d)$ is a random vector with distribution F. More generally, any dependence function satisfying $F(x)=D(F_1(x_1),...,F_d(x_d))$ could be defined to be a dependence function of F, although the present choice is a natural one.

Write $T=(0,1)^d\setminus\{\mathbf{1}\}$ where $\mathbf{1}=(1,...,1)\in\mathbb{R}^d$, and for $\mathbf{t}=(t_1,...,t_d)\in T$, let $v_n(\mathbf{t})=(v_{n1}(t_1),...,v_{nd}(t_d))$ where $v_{nj}(t_j)$ satisfies $\lim_{n\to\infty} nP\{\xi_{1j}>v_{nj}(t_j)\}=-\log t_j$. Let H_n denote the distribution function of M_n (i.e., $H_n(x)=P\{M_n\leq x\}$), with marginals H_{nj}, $j=1,...,d$. Then (see Refs. [8, 11]),

$$H_n(v_n(\mathbf{t}))\to^w H(\mathbf{t}) \tag{3}$$

if and only if

$$H_{nj}(v_{nj}(t_j))\to^w H_j(t_j),\ j=1,...,d, \text{ and } D_{H_n}(\mathbf{t})\to^w D_H(\mathbf{t}).$$

The limiting behavior of M_n can therefore be separated into two parts, one pertaining to the convergence of the marginals (a univariate problem) and the other to the convergence of the dependence functions. Here we focus attention exclusively on the latter. It should be noted that the choice of normalising constants does not affect the dependence function of the limit distribution H, but only alters the marginals (see Ref. [9], Lemma 3.2). Since our main interest is in the dependence function, the present choice of normalising constants is appropriate in view of the fact that it results in Uniform$[0,1]$ marginals for the limit distribution when $\{\xi_n\}$ is iid, so that in particular $D_H=H$. According to Theorem 3.3 of Ref. [9], the class of all possible limits H in Eq. (3) (for iid $\{\xi_n\}$) is precisely the class of *extreme* dependence functions, that is those that satisfy

$$D^n(\mathbf{t})=D(t_1^n,\ldots,t_d^n) \tag{4}$$

for each $n>1$ and $\mathbf{t}=(t_1,\ldots,t_d)\in[0,1]^d$. Theorem 1.1 below shows that the same is true also if $\{\xi_n\}$ is a stationary sequence satisfying the following mixing condition.

For $\mathbf{t}\in T$, let

$$\mathscr{B}_k^l(v_n(\mathbf{t}))=\sigma\{(\xi_{ij}>v_{nj}(t_j)) : k\le i\le l, j=1,\ldots,d\}$$

and for $1\le l\le n-1$, define

$$\alpha_{n,l}=\sup\{|P(A\cap B)-P(A)P(B)| : A\in\mathscr{B}_1^k(v_n(\mathbf{t})),$$

$$B\in\mathscr{B}_{k+l}^n(v_n(\mathbf{t})), 1\le k<k+l\le n\}.$$

The mixing condition $\Delta(v_n(\mathbf{t}))$ is then said to hold if $\alpha_{n,l_n}\to 0$ for some sequence $\{l_n\}$ satisfying $l_n/n\to 0$. This is the multivariate version of the mixing condition used in Ref. [5] and is slightly stronger than the $D(u_n)$ condition in Ref. [9]. Henceforth $\{\xi_n\}$ will be assumed to satisfy $\Delta(v_n(\mathbf{t}))$, for some or all $\mathbf{t}$, as required.

THEOREM 1.1. *Let $\{\xi_n\}$ satisfy $\Delta(v_n(\mathbf{t}))$ for all $\mathbf{t}\in T$ and suppose that $P\{M_n\le v_n(\mathbf{t})\}\to^w H(\mathbf{t})$, non-degenerate. Then D_H is an extreme dependence function and hence, in particular, $H(\mathbf{t}^c)=H^c(\mathbf{t})$ for each $\mathbf{t}\in[0,1]^d$ and $c>0$ (where $\mathbf{t}^c=(t_1^c,\ldots,t_d^c)$).*

PROOF: The first part is an immediate consequence of Theorem 4.2 of Ref. [9] while the second part follows from the definition of extreme dependence functions upon noting that (by the univariate theory described above), the marginals of H are of the form $H_j(t_j)=t_j^{\theta_j}$ where θ_j is the extremal index of $\{\xi_{nj}\}$, the jth-component sequence of $\{\xi_n\}$.

In the next section we apply the exceedance point process approach to multivariate extremes and obtain some weak convergence results. The multivariate extremal index is then defined (in Sec. 3), based on the multivariate analogue of Eq. (1). It is seen to be a function of only $d-1$ variables and its properties naturally extend those of the univariate extremal index. Finally in Sec. 4 we consider two examples of bivariate moving average sequences for which the computation of the extremal index is demonstrated.

2. Exceedance Point Processes

Fix $t\in T$ and let $\delta_{ij}=I_{\{\xi_{ij}>v_{nj}(t_j)\}}$, $i=1,\ldots,n$, $j=1,\ldots,d$, and put $\delta_i=(\delta_{i1},\ldots,\delta_{id})$. The *multivariate exceedance point process* $N_n=N_n^{(\mathbf{t})}$ is then defined by

$$N_n(B)=\sum_{i=1}^n I_{\{i/n\in B\}}\delta_i,\quad B\in[0,1]. \tag{5}$$

Assume that $\{\xi_n\}$ satisfies $\Delta(v_n(\mathbf{t}))$. If also $N_n\to^d N_0$, then it may be shown (as in the univariate case) that the limit N_0 is a point process on $[0,1]$ which is of Compound Poisson type. More precisely, the Laplace Transform of N_0 is given by

$$-\log E\exp\{-\sum_{j=1}^d\int_{[0,1]} f_j\,dN_{0j}\}=\nu\int_{x\in[0,1]}\int_{\mathbf{y}\in\mathbb{Z}_+^d}$$

$$(1-\exp\{-\sum_{j=1}^d y_j f_j(x)\})d\pi(\mathbf{y})dx. \tag{6}$$

Here N_{0j} denotes the jth-component of N_0, ν is a positive constant, π is a probability distribution on $\mathbb{Z}_+^d=\{0,1,2,\ldots\}^d\setminus\{\mathbf{0}\}$ and f_j's are non-negative functions on $[0,1]$.

Let $\{k_n\}$ be any sequence of positive integers satisfying

$$k_n\to\infty,\ k_nl_n/n\to 0,\text{ and } k_n\alpha_{n,l_n}\to 0,\quad\text{as } n\to\infty. \tag{7}$$

Set $r_n=[n/k_n]$ (the largest integer not exceeding n/k_n) and put $J_{n1}=[0,r_n/n]$. Define the probability distribution π_n on $\mathbb{Z}_+^d$ by

$$\pi_n(\mathbf{y})=P\{N_n(J_{n1})=\mathbf{y}\,|\,N_n(J_{n1})\ne\mathbf{0}\},\ \mathbf{y}\in\mathbb{Z}_+^d.$$

The following theorem which gives a useful characterization of the convergence of N_n is an immediate consequence of the results in Sec. 5 of Ref. [12].

THEOREM 2.1. *$N_n\to^d N_0$ if and only if $\pi_n\to^w\pi$ and $P\{M_n\le v_n(\mathbf{t})\}\to e^{-\nu}$, and in that case the Laplace Transform of N_0 is given by Eq. (6).*

Next we consider the iid case in some detail and obtain an interesting connection with Theorem 5.3.1 of Ref. [8].

PROPOSITION 2.2. *Let $\{\xi_n\}$ be iid and for fixed $\mathbf{t}\in T$ let N_n be defined by Eq. (5). If $N_n\to^d N_0$ then the multiplicity distribution π in Eq. (6) is supported on the set $S=\{0,1\}^d\setminus\{\mathbf{0}\}$.*

PROOF: Observe that $\Delta(v_n(\mathbf{t}))$ is trivially satisfied since $\alpha_{n,1}=0$ so that we may take $l_n\equiv 1$ and $k_n=n$. Then $\pi_n(\mathbf{y})=P\{\delta_1=\mathbf{y}\,|\,\xi_1\not\le v_n(\mathbf{t})\}$, $\mathbf{y}\in\mathbb{Z}_+^d$, which is clearly supported on S. The result is now immediate since S is a closed set and $\pi_n\to^w\pi$ by Theorem 2.1. □

Making the dependence on $\mathbf{t}$ explicit, we now write $N_n=N_n^{(\mathbf{t})}$, $N_0=N_0^{(\mathbf{t})}$, $\nu=\nu^{(\mathbf{t})}$ and $\pi=\pi^{(\mathbf{t})}$. In addition we shall require the following notation from Ref. [8]. For $1 \le k \le d$, let $\mathbf{j}(k)=(j_1,...,j_k)$ denote a vector with integer-valued components $1 \le j_1 < j_2 < \cdots < j_k \le d$, and for $\mathbf{x}=(x_1,...,x_d) \in \mathbb{R}^d$ write $\mathbf{x}_{\mathbf{j}(k)}=(x_{j_1},...,x_{j_k})$. Define the "survival function"

$$G(\mathbf{x})=P\{\xi_{11}>x_1,...,\xi_{1d}>x_d\}$$

and write $G_{\mathbf{j}(k)}(\mathbf{x})=P\{\xi_{1j_1}>x_{j_1},...,\xi_{1j_k}>x_{j_k}\}$. For each $\mathbf{j}(k)$, let $\mathbf{y}_{\mathbf{j}(k)}$ denote the element in $S=\{0,1\}^d\setminus\{0\}$ whose jth component equals 1 if and only if $j=j_i$ for some $i=1,...,k$. (This defines a natural 1-1 correspondence between S and the $\mathbf{j}(k)$'s.)

THEOREM 2.3. *Let $\{\xi_n\}$ be iid. Then $N_0^{(t)} \to^d N_0^{(t)}$ for some fixed $\mathbf{t} \in T$ if and only if*

$$\lim_{n\to\infty} nG_{\mathbf{j}(k)}(v_n(\mathbf{t}))=h_{\mathbf{j}(k)}(\mathbf{t})<\infty \tag{8}$$

for each $\mathbf{j}(k)$, $1 \le k \le d$. In that case $N_0^{(t)}$ has Laplace Transform given by Eq. (6) with

$$\nu^{(t)}=\sum_{k=1}^{d}(-1)^{k+1}\sum_{1\le j_1<\cdots<j_k\le d} h_{\mathbf{j}(k)}(\mathbf{t}) \tag{9}$$

and with $\pi^{(t)}$ determined by the relations

$$h_{\mathbf{j}(k)}(\mathbf{t})=\nu^{(t)}\sum_{\mathbf{y}\ge \mathbf{y}_{\mathbf{j}(k)}} \pi^{(t)}(\mathbf{y}). \tag{10}$$

PROOF: Write $S_k(\mathbf{t})=\sum_{1\le j_1<\cdots<j_k\le d} G_{\mathbf{j}(k)}(\mathbf{t})$, so that $P\{\xi_1 \not\le v_n(\mathbf{t})\}=\sum_{k=1}^{d}(-1)^{k+1}S_k(v_n(\mathbf{t}))$. If Eq. (8) holds for each k, then

$$\lim_{n\to\infty} nP\{\xi_1 \not\le v_n(\mathbf{t})\}=\sum_{k=1}^{d}(-1)^{k+1}\sum_{1\le j_1<\cdots<j_k\le d} h_{\mathbf{j}(k)}(\mathbf{t})=\nu^{(t)}, \tag{11}$$

and hence $\lim_{n\to\infty}P\{M_n \le v_n(\mathbf{t})\}=e^{-\nu^{(t)}}$.

Next observe that for each $\mathbf{j}(k)$, $1 \le k \le d$,

$$G_{\mathbf{j}(k)}(v_n(\mathbf{t}))=\sum_{\mathbf{y}\ge \mathbf{y}_{\mathbf{j}(k)}} P\{\delta_1=\mathbf{y}\}=P\{\xi_1 \not\le v_n(\mathbf{t})\}\sum_{\mathbf{y}\ge \mathbf{y}_{\mathbf{j}(k)}} \pi_n(\mathbf{y}), \tag{12}$$

where $\pi_n(\mathbf{y})=P\{\delta_1=\mathbf{y} \mid \xi_1 \not\le v_n(\mathbf{t})\}$. Moreover, this relationship is invertible in the sense that each of the probabilities $\pi_n(\mathbf{y})$, $\mathbf{y}\in S$, can be expressed as a linear combination of the $G_{\mathbf{j}(k)}(v_n(\mathbf{t}))$'s. Therefore by Eq. (8), $\lim_{n\to\infty}\pi_n(\mathbf{y})=\pi^{(t)}(\mathbf{y})$ (say) exists and satisfies Eq. (10). Hence by Theorem 2.1 $N_n^{(t)} \to^d N_0^{(t)}$ where $N_0^{(t)}$ has the specified parameters. Conversely if $N_n^{(t)} \to^d N_0^{(t)}$ then π_n and $P\{M_n \le v_n(\mathbf{t})\}$ converge (by Theorem 2.1 again), and hence Eq. (8) follows by virtue of Eqs. (11) and (12). □

COROLLARY 2.4. *Let $\{\xi_n\}$ be iid. Then $N_n^{(t)} \to^d N_0^{(t)}$ for each $\mathbf{t} \in T$ if and only if $P\{M_n \le v_n(\mathbf{t})\} \to^w H(\mathbf{t})$. Moreover H and $\{\nu^{(t)},\pi^{(t)}\}_{t\in T}$ determine each other.*

PROOF: (Sketch) The first part follows from Theorem 2.3 above and Theorem 5.3.1 of Ref. [8] which states that $P\{M_n \le v_n(\mathbf{t})\} \to^w H(\mathbf{t})$ if and only if Eq. (8) holds for each $\mathbf{t} \in T$. Note that $H(\mathbf{t})=e^{-\nu^{(t)}}$ so that H and the $\nu^{(t)}$'s can be obtained from each other. Also the $\pi^{(t)}$'s can be obtained from the $\nu^{(t)}$'s by first inverting Eq. (9) to get the $h_{\mathbf{j}(k)}(\mathbf{t})$'s and then inverting Eq. (10). (The inversion of Eq. (9) is carried out inductively using the fact that the weak convergence of $H_n(v_n(\mathbf{t}))$ implies that of all lower dimensional marginals.) □

Analogous results for the dependent case take on a slightly different form. Let $\{\xi_n\}$ be a stationary sequence satisfying $\Delta(v_n(\mathbf{t}))$ for each $\mathbf{t} \in T$. As before let $r_n=[n/k_n]$ where $\{k_n\}$ is any sequence satisfying Eq. (7), and define

$$G_{r_n\mathbf{j}(k)}(v_n(\mathbf{t}))=P\{M_{r_nj_1}>v_{nj_1}(t_{j_1}),...,M_{r_nj_k}>v_{nj_k}(t_{j_k})\}.$$

THEOREM 2.5. *Let $\{\xi_n\}$ be a stationary sequence satisfying $\Delta(v_n(\mathbf{t}))$ for each $\mathbf{t} \in T$. Then $P\{M_n \le v_n(\mathbf{t})\} \to^w H(\mathbf{t})$ if and only if*

$$\lim_{n\to\infty} k_n G_{r_n\mathbf{j}(k)}(v_n(\mathbf{t}))=h_{\mathbf{j}(k)}(\mathbf{t})<\infty$$

for each $\mathbf{j}(k)$, $1 \le k \le d$ and $\mathbf{t} \in T$, and in that case

$$H(\mathbf{t})=\exp\{\sum_{k=1}^{d}(-1)^k\sum_{1\le j_1<\cdots<j_k\le d} h_{\mathbf{j}(k)}(\mathbf{t})\}.$$

PROOF: Observe that the mixing condition $\Delta(v_n(\mathbf{t}))$ implies that $\{\xi_{n\mathbf{j}(k)}\}$ satisfies $\Delta(v_{n\mathbf{j}(k)}(\mathbf{t}))$ for each $\mathbf{j}(k)$ (with obvious notation). Hence it may be shown as in the univariate case (see Lemma 2.1 of Ref. [1]) that

$$P\{M_{n\mathbf{j}(k)} \le v_{n\mathbf{j}(k)}(\mathbf{t})\}-P^{k_n}\{M_{r_n\mathbf{j}(k)} \le v_{n\mathbf{j}(k)}(\mathbf{t})\} \to 0, \tag{13}$$

for each $\mathbf{j}(k)$. The result may therefore be proved in exactly the same way as Theorem 5.3.1 of Ref. [8]. □

REMARK: Under the hypothesis of Theorem 2.5, if $N_n^{(t)} \to^d N_0^{(t)}$, $\mathbf{t} \in T$, with parameters $\nu^{(t)}$ and $\pi^{(t)}$ then $P\{M_n \le v_n(\mathbf{t})\} \to^w H(\mathbf{t})=e^{-\nu^{(t)}}$, as in the iid case. However it is not possible in general to recover the $\pi^{(t)}$'s from H since the clustering of exceedances may cause the support of $\pi^{(t)}$ to extend beyond S. References [9, 10] give sufficient conditions (analogous to the $D'(u_n)$ condition of Ref. [13]) under which clustering does not occur, so that Corollary 2.4 can be extended to stationary sequences satisfying this condition.

A distribution function F on $\mathbb{R}^d$ is said to be *independent* if $F(\mathbf{x})=\prod_{j=1}^d F_j(x_j)$, $\mathbf{x}\in\mathbb{R}^d$. If $\{\xi_n\}$ is iid and $P\{M_n\le v_n(\mathbf{t})\}\to^w H(\mathbf{t})$, then it follows from Corollary 5.3.1 of Ref. [8] that H is independent if and only if the marginals of H are pairwise independent. The analogous result for the dependent case is stated below. The proof (which is omitted) is essentially the same as for the iid case, but uses Theorem 2.5 instead of Theorem 5.3.1 of Ref. [8].

COROLLARY 2.6. *Let $\{\xi_n\}$ be a stationary sequence satisfying $\Delta(v_n(\mathbf{t}))$ for each $\mathbf{t}\in T$ and suppose that $P\{M_n\le v_n(\mathbf{t})\}\to^w H(\mathbf{t})$. Then H is independent if and only if $k_nP\{M_{r_n j}>v_{n,j}(t_j),\ M_{r_n l}>v_{n,l}(t_l)\}\to 0$ for each $1\le j<l\le d$, $\mathbf{t}\in T$, i.e., if and only if $k_nG_{r_n\mathbf{j}(2)}(v_n(\mathbf{t}))\to 0$ for each $\mathbf{j}(2)$ and each $\mathbf{t}\in T$.*

It is shown in Ref. [14] that H is independent if $H(\mathbf{t})=\prod_{j=1}^d H_j(t_j)$ for some $\mathbf{t}\in(0,1)^d$. Although the result in [14] only stated for iid sequences under a linear normalization, the proof essentially rests on the defining property of extreme dependence functions, namely Eq. (1). Consequently the result extends to the present more general situation allowing dependence and non-linear normalizations. Corollary 2.6 can therefore be improved as follows.

COROLLARY 2.7. *Let $\{\xi_n\}$ be as in Corollary 2.6 and suppose that $P\{M_n\le v_n(\mathbf{t})\}\to^w H(\mathbf{t})$. Then the following are equivalent:*

(i) H is independent,
(ii) $H(\mathbf{t})=\prod_{j=1}^d H_j(t_j)$ for some $\mathbf{t}\in(0,1)^d$,
(iii) $k_nG_{r_n\mathbf{j}(2)}(v_n(\mathbf{t}))\to 0$ for each $\mathbf{j}(2)$, for some $\mathbf{t}\in(0,1)^d$.

It should be noted that Refs. [9, 10] give some interesting sufficient conditions for H to be independent when $\{\xi_n\}$ is a stationary sequence. A natural question to ask in the present context is whether H is independent whenever $\hat{H}$ is. Proposition 3.4 gives a necessary and sufficient condition for this in terms of the extremal index, but the answer in general is negative and a counter-example can be found in [10]. It seems more plausible that the converse may be true, i.e., that $\hat{H}$ is independent whenever H is. In fact however, this too is not the case, as shown by an interesting counter-example in [15].

We conclude this section by stating a result which extends Theorem 5.1 of [5] and is proved similarly.

THEOREM 2.8. *Let $\{\xi_n\}$ be a stationary sequence satisfying $\Delta(v_n(\mathbf{t}))$ for each $\mathbf{t}\in T$ and suppose that $N_n^{(\mathbf{t})}\to^d N_0^{(\mathbf{t})}$ for some $\mathbf{t}\in T$. Then $N_n^{(\mathbf{t}^c)}\to^d N_0^{(\mathbf{t}^c)}$ for each $c>0$ and furthermore, $\nu^{(\mathbf{t}^c)}=c\nu^{(\mathbf{t})}$ and $\pi^{(\mathbf{t}^c)}=\pi^{(\mathbf{t})}$ (where $\mathbf{t}^c=(t_1^c,\dots,t_d^c)$).*

3. The Multivariate Extremal Index

Let $\{\xi_n\}$ be a stationary sequence and $\{\hat{\xi}_n\}$ the associated iid sequence. Suppose that $P\{M_n\le v_n(\mathbf{t})\}\to^w H(\mathbf{t})$ and $P\{\hat{M}_n\le v_n(\mathbf{t})\}\to^w \hat{H}(\mathbf{t})$. The *multivariate extremal index* of $\{\xi_n\}$ is then defined by the relation $H(\mathbf{t})=\hat{H}^{\theta(\mathbf{t})}(\mathbf{t})$ (see Eq. (1)), or more explicitly

$$\theta(\mathbf{t})=\log H(\mathbf{t})/\log \hat{H}(\mathbf{t}),\quad \mathbf{t}\in T.$$

Observe that $\theta(\mathbf{t})$ is well defined since $\hat{H}$ has Uniform[0,1] marginals and hence, $0<\hat{H}(\mathbf{t})<1$ on T. The following results describe some basic properties of the multivariate extremal index.

PROPOSITION 3.1. *Assume that $\{\xi_n\}$ satisfies $\Delta(v_n(\mathbf{t}))$ for each $\mathbf{t}\in T$ and has extremal index $\theta(\mathbf{t})$. Then*

(i) $\theta(\mathbf{t})=\theta(\mathbf{t}^c)$ for each $\mathbf{t}\in T$ and $c>0$, and
(ii) for each $j=1,\dots,d$, $\{\xi_{nj}\}$ has extremal index $\theta_j=\theta(\mathbf{t})$ where $\mathbf{t}\in T$ has all coordinates equal to 1 except the jth.

(Note that by (*i*), $\theta(\mathbf{t})$ is constant along the contours $L_t=\{\mathbf{t}^c:c>0\}$, $\mathbf{t}\in T$, and hence θ_j in (*ii*) is well-defined.)

PROOF: Recall that (by Theorem 1.1) $H(\mathbf{t}^c)=H^c(\mathbf{t})$ and $\hat{H}(\mathbf{t}^c)=\hat{H}^c(\mathbf{t})$ so that (*i*) follows from the definition of the extremal index. Next, for $\mathbf{t}\in T$ with all coordinates but the jth equal to 1, $P\{M_n\le v_n(\mathbf{t})\}=P\{M_{nj}\le v_{nj}(t_j)\}$ and hence

$$\lim_{n\to\infty} P\{M_{nj}\le v_{nj}(t_j)\}=\lim_{n\to\infty} P\{M_n\le v_n(\mathbf{t})\}=H(\mathbf{t})=H_j(t_j).$$

Therefore by Theorem 2.2 of Ref. [1], $\{\xi_{nj}\}$ has extremal index θ_j (say) so that $H_j(t_j)=t_j^{\theta_j}$. Now $H(\mathbf{t})=\hat{H}^{\theta(\mathbf{t})}(\mathbf{t})$ by definition of the extremal index, and for the present choice of $\mathbf{t}$ this is the same as $H_j(t_j)=t_j^{\theta(\mathbf{t})}$, whence it follows that $\theta(\mathbf{t})=\theta_j$ for all such $\mathbf{t}$. □

For $\mathbf{t}\in T$, let $\tilde{N}_n^{(\mathbf{t})}$ denote the one-dimensional point process obtained from $N_n^{(\mathbf{t})}$ via the map $\mathbf{y}\to I_{\{\mathbf{y}\ne 0\}}$ from $\{0,1\}^d$ to $\{0,1\}$, i.e., $\tilde{N}_n^{(\mathbf{t})}(B)=\sum_{i=1}^n I_{\{i/n\in B\}}I_{\{\delta_i\ne 0\}}$, $B\in\mathscr{B}$. Thus $\tilde{N}_n^{(\mathbf{t})}$ has unit mass at i/n if and only if $\xi_i\nleq v_n(\mathbf{t})$. Assume that $\{\xi_n\}$ satisfies $\Delta(v_n(\mathbf{t}))$ and with J_{n1} as in Sec. 2, let

$$\tilde{\pi}_n^{(\mathbf{t})}(y)=P\{\tilde{N}_n^{(\mathbf{t})}(J_{n1})=y\mid \tilde{N}_n^{(\mathbf{t})}(J_{n1})>0\},\quad y\ge 1.$$

PROPOSITION 3.2. *Assume that $\{\xi_n\}$ satisfies $\Delta(v_n(\mathbf{t}))$ for each $\mathbf{t}\in T$ and has extremal index $\theta(\mathbf{t})$. Then $\theta(\mathbf{t})=(\lim_{n\to\infty}\sum_{y\ge 1} y\tilde{\pi}_n^{(\mathbf{t})}(y))^{-1}$.*

PROOF: Observe that

$$\sum_{y\geq 1} y\tilde{\pi}_n^{(\mathbf{t})}(y)=E(\tilde{N}_n^{(\mathbf{t})}(J_{n1})\mid \tilde{N}_n^{(\mathbf{t})}(J_{n1})>0)=\frac{r_nP\{\xi_1\not\leq v_n(\mathbf{t})\}}{P\{\tilde{N}_n^{(\mathbf{t})}(J_{n1})>0\}}=$$

$$\frac{k_nr_n}{n}\frac{nP\{\xi_1\not\leq v_n(\mathbf{t})\}}{k_nP\{M_{r_n}\not\leq v_n(\mathbf{t})\}}.$$

Now $\lim_{n\to\infty} P\{\hat{M}_n\leq v_n(\mathbf{t})\}=\hat{H}(\mathbf{t})$ and $\lim_{n\to\infty} P\{M_n\leq v_n(\mathbf{t})\}=H(\mathbf{t})$ (by assumption), so that $\lim_{n\to\infty} nP\{\xi_1\not\leq v_n(\mathbf{t})\}=-\log\hat{H}(\mathbf{t})$ and (by Eq. (13)) $\lim_{n\to\infty} k_nP\{M_{r_n}\not\leq v_n(\mathbf{t})\}=-\log H(\mathbf{t})$. Therefore $\lim_{n\to\infty}\sum_{y\geq 1} y\tilde{\pi}_n^{(\mathbf{t})}(y)=\log\hat{H}(\mathbf{t})/\log H(\mathbf{t})=1/\theta(\mathbf{t})$, as required. □

REMARK: Proposition 3.2 is simply the multivariate version of Eq. (1) and shows how the extremal index is related to the clustering of "exceedances." Indeed, according to the present viewpoint, an exceedance occurs at time i if $\xi_i\not\leq v_n(\mathbf{t})$, i.e., if $\xi_{ij}>v_{nj}(t_j)$ for at least one j. Thus Propositions 3.1 and 3.2 show that while the degree of clustering may depend on $\mathbf{t}$, it is constant on each $L_{\mathbf{t}}$. Note also the connection to Theorem 2.8.

The next result gives the relation between the dependence functions of H and $\hat{H}$, which is seen to involve the extremal index in an intricate manner.

PROPOSITION 3.3. *If $\{\xi_n\}$ has extremal index $\theta(\mathbf{t})$, $\mathbf{t}\in T$, then*

$$D_H(t_1^{\theta_1},\ldots,t_d^{\theta_d})=D_{\hat{H}}^{\theta(\mathbf{t})}(\mathbf{t}),\quad \mathbf{t}\in T, \tag{14}$$

where θ_j is the extremal index of $\{\xi_{nj}\}$, $j=1,\ldots,d$.

PROOF: By definition of the dependence function, $D_H(\mathbf{t})=P\{H_1(X_1)\leq t_1,\ldots,H_d(X_d)\leq t_d\}$ where $(X_1,\ldots,X_d)$ is a random vector with distribution H. Therefore, since $H_j(t_j)=t_j^{\theta_j}$,

$$D_H(t_1^{\theta_1},\ldots,t_d^{\theta_d})=P\{X_1\leq t_1,\ldots,X_d\leq t_d\}=H(\mathbf{t})=\hat{H}^{\theta(\mathbf{t})}(\mathbf{t})$$

$$=D_{\hat{H}}^{\theta(\mathbf{t})}(\mathbf{t}),$$

or required. □

REMARKS

1.) Note that $\mathbf{s}=\mathbf{t}^c$ (for some $c>0$) if and only if $\log s_j/\log s_d=\log t_j/\log t_d=a_j$ (say), $j=1,\ldots,d-1$. Therefore we may write $L_{\mathbf{t}}=L_{\mathbf{a}}$ where $\mathbf{a}=(a_1,\ldots,a_{d-1})$, and hence by the remark following Proposition 3.2, $\theta(\mathbf{t})=\theta(\mathbf{a})$, i.e., the extremal index is a function of $d-1$ variables only.
2.) By Proposition 3.3, $D_H(t_1^{\theta_1},\ldots,t_d^{\theta_d})=D_{\hat{H}}^{\theta(\mathbf{t})}(\mathbf{t})=D_{\hat{H}}(\mathbf{t}^{\theta(\mathbf{t})})$. Also, if $\mathbf{t}\in L_{\mathbf{a}}$ then $(t_1^{\theta_1},\ldots,t_d^{\theta_d})\in L_{\mathbf{a}^*}$ where $\mathbf{a}^*=(a_1\theta_1/\theta_d,\ldots,a_{d-1}\theta_{d-1}/\theta_d)$. Thus D_H is obtained by translating the values of $D_{\hat{H}}(=\hat{H})$ on $L_{\mathbf{a}}$ onto $L_{\mathbf{a}^*}$.
3.) While the above results illustrate some of the basic properties of the multivariate extremal index, they are far from complete. For instance, it would be useful to identify the set of all "admissible" $\theta(\cdot)$ for a given $\hat{H}$, that is the set of all $\theta(\cdot)$ such that $D_H(\cdot)$ defined by Eq. (14) is a probability distribution on $[0,1]^d$. It would also be of interest to study the properties of $\theta(\cdot)$ when one or both of $\hat{H}$ and H are independent. In this context we have the following simple result.

PROPOSITION 3.4. *If $\hat{H}$ is independent, then H is independent if and only if*

$$\theta(\mathbf{t})=\sum_{j=1}^{d}\theta_j\log t_j/\sum_{j=1}^{d}\log t_j,\quad \textit{for some } \mathbf{t}\in(0,1)^d.$$

In particular, if both $\hat{H}$ and H are independent then $\theta(\mathbf{t})$ is a convex combination of the θ_j's.

PROOF: If $\hat{H}$ is independent, then $H(\mathbf{t})=\hat{H}(\mathbf{t})^{\theta(\mathbf{t})}=(\prod_{j=1}^{d}t_j)^{\theta(\mathbf{t})}$. The conclusion follows immediately from Corollary 2.7 (iii) upon taking logarithms and noting that if H is independent, then $H(\mathbf{t})=\prod_{j=1}^{d}t_j^{\theta_j}$.

The extremal index can be given the following more general formulation. Let $\hat{\mu}$ and μ be the probability measures on $(0,1)^d$ corresponding to $\hat{H}$ and H, respectively. Thus for instance,

$$\mu(A)=\lim_{n\to\infty}P\{M_n\in v_n(A)\}$$

where $v_n(A)=\{v_n(\mathbf{s}) : \mathbf{s}\in A\}$, $A\subset(0,1)^d$. We now define $\tilde{\theta}(A)$ via the relationship $\mu(A)=\hat{\mu}^{\tilde{\theta}(A)}(A)$, or more directly $\tilde{\theta}(A)=\log\mu(A)/\log\hat{\mu}(A)$, for subsets $A\subset(0,1)^d$ such that $\hat{\mu}(A)>0$ and $\mu(A)>0$.

Note that $\theta(\mathbf{t})=\tilde{\theta}((0,t_1)\times\cdots\times(0,t_d))$ for $\mathbf{t}\in T$. Thus if $\{\theta(\mathbf{t}) : \mathbf{t}\in T\}$ is known along with either of H or $\hat{H}$, then it is possible at least in theory to obtain $\{\tilde{\theta}(A) : A\subset(0,1)^d\}$. In practice, however, it may not be possible to obtain $\tilde{\theta}(A)$ in a tractable form, but frequently one is only interested in certain special sets, typically rectangles of the form $\prod_{j=1}^{d}(a_j,b_j)$, and for such sets the computation is easy.

The definition of M_n as the vector of componentwise maxima actually corresponds to regarding ξ_i as an extreme observation if $\xi_{ij}>v_{nj}(t_j)$ for some j. More generally, one may define ξ_i to be an extreme value if $\xi_i\in v_n(A)$ for some $A\subset(0,1)^d$, in which case $\tilde{\theta}(A)$ has an interpretation as a measure of the clustering of such extremes. Note that the original definition of extremes corresponds to letting $A=((0,t_1)\times\cdots\times(0,t_d))^c$. Alternately one may consider taking $A=(t_1,1)\times\cdots\times(t_d,1)$ which corresponds to defining ξ_i as an extreme observation if $\xi_{ij}>v_{nj}(t_j)$ for all j. Yet another choice is $A=\{\mathbf{t}:\Sigma t_j^2>c\}$.

4. Examples

We conclude with two examples, both involving bivariate stationary sequences.

EXAMPLE 4.1 Let $\{\eta_n\}$ be an iid sequence, and put $\xi_{n1}=\eta_n$, and $\xi_{n2}=\max\{\eta_{n-1},\eta_n\}$. Let F denote the distribution of $\xi_n=(\xi_{n1},\xi_{n2})$ with marginals F_1 and F_2. Then $F_2(x)=P\{\xi_{n2}\le x\}=P\{\eta_{n-1}\le x,\ \eta_n\le x\}=F_1^2(x)$ and

$$F(x_1,x_2)=P\{\xi_{n1}\le x_1,\xi_{n2}\le x_2\}=\begin{cases}F_1^2(x_2), & \text{if } x_1\ge x_2\\ F_1(x_1)F_1(x_2), & \text{if } x_1<x_2.\end{cases}$$

If $v_{nj}(t_j)$ satisfies $F_j^n(v_{nj}(t_j))\to t_j$, $j=1,2$, then $\lim_{n\to\infty}F_1^n(v_{n2}(t_2))=t_2^{1/2}$ so that $v_{n2}(t_2)=v_{n1}(t_2^{1/2})$. Moreover $v_{n1}(t_1)\ge v_{n2}(t_2)$ if and only if $t_1\ge t_2^{1/2}$, and so

$$H_n(v_n(t))=P\{M_n\le v_n(t)\}\to^w H(t)=\begin{cases}t_2^{1/2}, & \text{if } t_1\ge t_2^{1/2}\\ t_1, & \text{if } t_1<t_2^{1/2}.\end{cases}$$

The marginals of H are therefore $H_1(t_1)=t_1$ and $H_2(t_2)=t_2^{1/2}$ so that $\theta_1=1$ and $\theta_2=1/2$, and the dependence function of H is $D_H(t)=H(t_1,t_2^2)=t_1\wedge t_2$. For the associated iid sequence $\{\hat{\xi}_n\}$ on the other hand, it is easily verified that

$$\hat{H}_n(v_n(t))=P\{\hat{M}_n\le v_n(t)\}\to\hat{H}(t)=\begin{cases}t_2, & \text{if } t_1\ge t_2^{1/2}\\ t_1t_2^{1/2}, & \text{if } t_1<t_2^{1/2}.\end{cases}$$

from which it follows that

$$\theta(t)=\begin{cases}\frac{1}{2}, & \text{if } t_1\ge t_2^{1/2}\\ \frac{\log t_1}{\log t_1t_2^{1/2}}, & \text{if } t_1<t_2^{1/2}.\end{cases}$$

We next consider a moving average sequence studied in Ref. [16].

EXAMPLE 4.2. Let $\{\mathbf{Z}_k=(Z_{k1},Z_{k2})'\}$, $-\infty<k<\infty$, be a sequence of iid random vectors in IR^2. We assume the existence of a sequence of positive constants $a_n\to\infty$, and a measure ν on IR^2 which is finite on sets of the form $\{\mathbf{x}:\|\mathbf{x}\|>r\}$, $r>0$ (where $\|\cdot\|$ denotes the Euclidean norm in IR^2), such that $nP\{a_n^{-1}\mathbf{Z}_0\in\cdot\}\to^v \nu(\cdot)$. (Here '$\to^v$' denotes vague convergence of measures on IR^2 with respect to the metric $d(\mathbf{x}_1,\mathbf{x}_2)=|r_1^{-1}-r_2^{-1}|\vee|\theta_1-\theta_2|$, where for $i=1,2$, r_i and θ_i denote the polar coordinates of $\mathbf{x}_i$, and $a\vee b=\max\{a,b\}$.) The measure ν is necessarily of the form $\nu(\{\mathbf{x}:\|\mathbf{x}\|>r,\theta(\mathbf{x})\in A\})=r^{-\alpha}S(A)$ for $r>0$ and $A\subset[0,2\pi)$, where $S(\cdot)$ is a probability measure on $[0,2\pi)$ and $\alpha>0$. Hence in particular [17],

$$\nu(cA)=c^{-\alpha}\nu(A) \tag{15}$$

for all $c>0$ and all sets A with $\nu(A)<\infty$.

Define the bivariate moving average process $\mathbf{X}_n=\Sigma_{j=0}^{\infty}C_j\mathbf{Z}_{n-j}$, where $\{C_j=[c_{j,kl}]_{k,l=1}^2\}_{j\ge0}$ is a sequence of real 2×2 matrices satisfying $\Sigma_{j=0}^{\infty}|c_{j,kl}|^{\delta}<\infty$, $k,l=1,2$, for some $\delta\in(0,\alpha)$, $\delta\le1$. For $\mathbf{x}=(x_1,x_2)'\in\mathrm{IR}^2$, write $A_{\mathbf{x},j}=\{\mathbf{z}:C_j\mathbf{z}\in((-\infty,x_1)\times(-\infty,x_2))^c\}$, where A^c denotes the complement of a set $A\subset\mathrm{IR}^2$. Then [16],

$$\lim_{n\to\infty}P\{a_n^{-1}\hat{M}_n\le\mathbf{x}\}=\exp\{-\hat{\gamma}(\mathbf{x})\},\text{ and}$$

$$\lim_{n\to\infty}P\{a_n^{-1}M_n\le\mathbf{x}\}=\exp\{-\gamma(\mathbf{x})\},\ \mathbf{x}\in\mathrm{IR}^2,$$

where $\hat{\gamma}(\mathbf{x})=\Sigma_{j=0}^{\infty}\nu(A_{\mathbf{x},j})$ and $\gamma(\mathbf{x})=\nu(\cup_{j=0}^{\infty}A_{\mathbf{x},j})$. The extremal index is therefore $\theta(\mathbf{x})=\gamma(\mathbf{x})/\hat{\gamma}(\mathbf{x})$, $\mathbf{x}\in\mathrm{IR}^2$. It follows from the definition of $A_{\mathbf{x},j}$ and Eq. (15) that this is in fact a function of x_1/x_2. Note that the extremal index defined above differs from that in Sec. 3 in that it is defined on IR^2 rather than $[0,1]^2$. However the two definitions are equivalent as may be seen by means of a suitable transformation from IR^2 to $[0,1]^2$.

The actual calculation of $\theta(\mathbf{x})$ may be quite difficult in general, but possible to carry out under appropriate simplifying assumptions.

Case (i). If $C_j=c_jC$ where $C=[c_{kl}]_{k,l=1}^2$ and the c_j's are non-negative constants, then $A_{\mathbf{x},j}=c_j^{-1}B(\mathbf{x})$ and $\cup_{j=0}^{\infty}A_{\mathbf{x},j}=cB(\mathbf{x})$, where $B(\mathbf{x})=\{\mathbf{z}:C\mathbf{z}\in((-\infty,x_1)\times(-\infty,x_2))^c\}$ and $c=\max\{c_j:j\ge0\}$. Therefore by Eq. (15), $\nu(A_{\mathbf{x},j})=c_j^{\alpha}\nu(B(\mathbf{x}))$ and $\nu(\cup_{j=0}^{\infty}A_{\mathbf{x},j})=c^{\alpha}\nu(B(\mathbf{x}))$ so that $\theta(\mathbf{x})\equiv c^{\alpha}/\Sigma_{j=0}^{\infty}c_j^{\alpha}$.

Case (ii). If the C_j's are diagonal, i.e., $C_j=\mathrm{diag}[c_{j1},c_{j2}]$ with $c_{ji}\ge0$, $i=1,2$, then $A_{\mathbf{x},j}=\{\mathbf{z}:c_{j1}z_1>x_1\text{ or }c_{j2}z_2>x_2\}$ and $\cup_{j=0}^{\infty}A_{\mathbf{x},j}=((-\infty,x_1/c_1)\times(-\infty,x_2/c_2))^c$ where $c_i=\max\{c_{ji}:j\ge0\}$, $i=1,2$. In particular, taking $x_2=\infty$ and using Eq. (15) as in Case (i), we have $\nu(A_{\mathbf{x},j})=c_{j1}^{\alpha}\nu(\{\mathbf{z}:z_1>x_1\})$ and $\nu(\cup_{j=0}^{\infty}A_{\mathbf{x},j})=c_1^{\alpha}\nu(\{\mathbf{z}:z_1>x_1\})$, so that the extremal index of $\{X_{n1}\}$ is $\theta_1=c_1^{\alpha}/\Sigma_{j=0}^{\infty}c_{j,1}^{\alpha}$.

Case (iii). Let D denote the support of ν. If $D\subset\{\mathbf{z}:z_1=0\text{ or }z_2=0\}$ (which is the case if the coordinates of $\mathbf{Z}_0$ are independent), then we may write

$$\nu((-\infty,x_1)\times(-\infty,x_2))^c=a_1x_1^{-\alpha}+a_2x_2^{-\alpha},\quad x_1,x_2\ge0, \tag{16}$$

for suitable constants $a_1\ge0$ and $a_2\ge0$. Once again, assuming the $c_{j,kl}$'s to be non-negative and writing $c_{kl}=\max\{c_{j,kl}:j\ge0\}$ for $k,l=1,2$, we have (writing $a\wedge b=\min\{a,b\}$)

$$A_{\mathbf{x},j}\cap D=((-\infty,x_1/c_{j,11}\wedge x_2/c_{j,21})\times(-\infty,x_1/c_{j,12}\wedge x_2/c_{j,22}))^c\cap D$$

and

$$\cup_{j=0}^{\infty} A_{x,j} \cap D = ((-\infty, x_1/c_{11} \wedge x_2/c_{21})$$

$$\times(-\infty, x_1/c_{12} \wedge x_2/c_{22}))^c \cap D,$$

so that using Eq. (16)

$$\theta(\mathbf{x}) = \frac{a_1(x_1/c_{11} \wedge x_2/c_{21})^{-\alpha} + a_2(x_1/c_{12} \wedge x_2/c_{22})^{-\alpha}}{a_1 \sum_{j=0}^{\infty}(x_1/c_{j,11} \wedge x_2/c_{j,21})^{-\alpha} + a_2 \sum_{j=0}^{\infty}(x_1/c_{j,12} \wedge x_2/c_{j,22})^{-\alpha}}$$

Thus putting $x_2=\infty$, we have $\theta_1=(a_1 c_{11}^{\alpha}+a_2 c_{12}^{\alpha})/(a_1 \sum_{j=0}^{\infty} c_{j,11}^{\alpha} + a_2 \sum_{j=0}^{\infty} c_{j,12}^{\alpha})$ and similarly, $\theta_2=(a_1 c_{21}^{\alpha}+a_2 c_{22}^{\alpha})/(a_1 \sum_{j=0}^{\infty} c_{j,21}^{\alpha} + a_2 \sum_{j=0}^{\infty} c_{j,22}^{\alpha})$.

If also $c_{j,12}=c_{j,21}=0$ for each j, (that is if the C_j's are diagonal), then

$$\theta(\mathbf{x}) = \frac{a_1 x_1^{-\alpha} c_{11}^{\alpha} + a_2 x_2^{-\alpha} c_{22}^{\alpha}}{a_1 x_1^{-\alpha} \sum_{j=0}^{\infty} c_{j,11}^{\alpha} + a_2 x_2^{-\alpha} \sum_{j=0}^{\infty} c_{j,22}^{\alpha}},$$

and in particular, $\theta_1=c_{11}^{\alpha}/\sum_{j=0}^{\infty} c_{j,11}^{\alpha}$ and $\theta_2=c_{22}^{\alpha}/\sum_{j=0}^{\infty} c_{j,22}^{\alpha}$. Note that in this case the limiting distributions of M_n and $\hat{M}_n$ are both independent, and hence (in accordance with Proposition 3.4) $\theta(\mathbf{x})$ is a convex combination of θ_1 and θ_2.

The non-negativeness of the C_j's assumed above is not crucial and may be relaxed, although at the cost of more involved calculations.

Acknowledgments

It is a pleasure to thank Professor M. R. Leadbetter for his encouragement and guidance during the course of this research. Part of this work was carried out during a brief visit to the University of Bern (supported by the Swiss National Science Foundation), and I am very grateful to Professor J. Hüsler for several useful discussions. Research supported by the Air Force Office of Scientific Research Contract No. F49620 85C 0144.

5. References

[1] M. R. Leadbetter, Extremes and local dependence in a stationary sequence, Z. Wahr-sch. verw. Gebiete. **65**, 291-306 (1983).
[2] G. F. Newell, Asymptotic extremes for m-dependent random variables, Ann. Math. Statist. **35**, 1322-1325 (1964).
[3] R. M. Loynes, Extreme values in uniformly mixing stochastic processes, Ann. Math. Statist. **36**, 993-999 (1965).
[4] G. L. O'Brien, The maximum term of uniformly mixing stationary processes, Z. Wahr-sch. verw. Gebiete. **30**, 57-63 (1974).
[5] T. Hsing, J. Hüsler, and M. R. Leadbetter, On the exceedance point process for a stationary sequence, Probab. Theory Rel Fields. **78**, 97–112 (1988).
[6] L. de Haan and S. Resnick, Limit theory for multivariate sample extremes, Z. Wahrsch. verw. Gebiete. **40**, 317-337 (1977).
[7] A. W. Marshall and I. Olkin, Domains of attraction of multivariate extreme value distributions, Ann. Probab. **11**, 168–177 (1983).
[8] J. Galambos, The asymptotic theory of extreme order statistics, Wiley, New York (1987).
[9] T. Hsing, Extreme value theory for multivariate stationary sequences, J. Multivariate Anal. **29**, 274–291 (1989).
[10] J. Hüsler, Multivariate extreme values in stationary random sequences, Stochastic Proc. Appl. **35**, 99–108 (1990).
[11] P. Deheuvels, Caractèrisation compléte des lois extrêmes mulitvariées et de la convergence des types extrêmes, Publ. Instit. Statist. Univ. Paris. **23**, 1–36 (1978).
[12] S. Nandagopalan, M. R. Leadbetter, and J. Hüsler, Limit theorems for multi-dimensional random measures, 92-14, Department of Statistics, Colorado State University, November 1992.
[13] M. R. Leadbetter, (1974). On extreme values in stationary sequences, Z. Wahrsch. verw. Gebiete. **28**, 298–303 (1974).
[14] R. Takahashi, Some properties of multivariate extreme value distributions and multivariate tail equivalence, Ann. Inst. Statist. Math. **39**, 637–647 (1987).
[15] N. Catkan, University of Bern, Private communication, (1993).
[16] R. A. Davis, J. Marengo, and S. Resnick, Extremal properties of a class of multivariate moving averages, L1-4, Proceedings of the 45th Session of the I.S.I., Amsterdam, (1985) pp. 1–14.
[17] S. Resnick, Extreme values, regular variation, and point processes, Springer-Verlag, New York (1987).

About the author: *S. Nandagopalan is an Assistant Professor of Statistics at Colorado State University.*