Ann. Inst. Statist. Math.
Vol. 45, No. 3, 433–444 (1993)

STRONG CONVERGENCE OF MULTIVARIATE POINT PROCESSES OF EXCEEDANCES

E. KAUFMANN AND R.-D. REISS*

FB 6, Universität Gesamthochschule Siegen, Hölderlinstr. 3, D-57068 Siegen, Germany

(Received December 24, 1991; revised October 26, 1992)

Abstract. We study the asymptotic behavior of vectors of point processes of exceedances of random thresholds based on a triangular scheme of random vectors. Multivariate maxima w.r.t. marginal ordering may be regarded as a special case. It is proven that strong convergence—that is convergence of distributions w.r.t. the variational distance—of such multivariate point processes holds if, and only if, strong convergence of multivariate maxima is valid. The limiting process of multivariate point processes of exceedances is built by a certain Poisson process. Auxiliary results concerning upper bounds on the variational distance between vectors of point processes are of interest in its own right.

Key words and phrases: Poisson processes, exceedances, random threshold.

1. Introduction

Our attention will be restricted to $\boldsymbol{R}^2$-valued random vectors (X_n, Y_n), $n \in \boldsymbol{N}$, to keep the technical details as simple as possible. Let $(X_{n,i}, Y_{n,i})$, $i = 1, \dots, n$, be a sample of independent copies of (X_n, Y_n). Using the concept of point processes, this sample and the marginal samples may alternatively be represented by the empirical process

$$N_n = \sum_{i=1}^{n} \varepsilon_{(X_{n,i}, Y_{n,i})} \tag{1.1}$$

and the marginal empirical processes

$$N_{n,1} = \sum_{i=1}^{n} \varepsilon_{X_{n,i}} \quad \text{and} \quad N_{n,2} = \sum_{i=1}^{n} \varepsilon_{Y_{n,i}} \tag{1.2}$$

(with ε_x denoting the Dirac-measure at x).

* The author was supported by the Deutsche Forschungsgemeinschaft.

Given $\boldsymbol{t} = (t_1, t_2)$ denote by $N_{n,\boldsymbol{t}}$ and N_{n,j,t_j} the truncations of N_n and $N_{n,j}$ outside of $(\boldsymbol{t}, \boldsymbol{\infty}) := (t_1, \infty) \times (t_2, \infty)$ and (t_j, ∞), respectively. Hence, $N_{n,\boldsymbol{t}} = N_n(\cdot \cap (\boldsymbol{t}, \boldsymbol{\infty}))$ and $N_{n,j,t_j} = N_{n,j}(\cdot \cap (t_j, \infty))$. This notation will also be utilized for other point processes and for measures as well; e.g., if ν_n is the intensity measure of N_n then $\nu_{n,\boldsymbol{t}} := \nu_n(\cdot \cap (\boldsymbol{t}, \boldsymbol{\infty}))$ is the intensity measure of $N_{n,\boldsymbol{t}}$. We are going to study the asymptotic behavior of vectors of point processes of exceedances of random thresholds $t_j \equiv T_j(N_{n,j})$, namely,

$$(1.3) \qquad (N_{n,1,T_1(N_{n,1})}, N_{n,2,T_2(N_{n,2})}).$$

Within this framework we may also deal with the bivariate sample maximum w.r.t. marginal ordering given by

$$(1.4) \qquad (X_{n:n}, Y_{n:n})$$

where $X_{n:n} = \max(X_{n,1}, \ldots, X_{n,n})$ and $Y_{n:n} = \max(Y_{n,1}, \ldots, Y_{n,n})$.

Classical extreme value theory concerns (a) the weak convergence of normalized maxima $X_{n:n}$ of i.i.d. random variables and (b) the use of a sample of maxima within sub-periods (Gumbel or annual maxima method) to estimate functional parameters of the tail of a distribution. Statistical inference is carried out in parametric extreme value models that are motivated by (a). An adequate link between (a) and (b) is achieved when utilizing approximations in terms of the variational and Hellinger distances (for a comprehensive account see Reiss (1989)). Statistical inference based on extreme order statistics—instead of 'annual' maxima—was initiated by Pickands (1975). Likewise one may deal with point processes of exceedances of high thresholds introduced above (see Resnick (1987), Davison and Smith (1990), Falk and Reiss (1992) and the literature cited therein). There also exists a rich literature concerning the weak convergence of multivariate maxima taken w.r.t. marginal ordering (for results and further references see the monographs by Galambos (1987), Resnick (1987) and Reiss (1989)). As well one may jointly study extreme order statistics or point processes of exceedances in each component. The joint asymptotic behavior of point processes of exceedances $(N_{n,1,t_1}, N_{n,2,t_2})$ was studied in Reiss (1990) in the special case of asymptotic independence. The aim of the present paper is to discuss the probabilistic part of the questions indicated above in a general setting that includes the treatment of 'multivariate extreme order statistics'.

In Section 2, we study our main question, namely the convergence of vectors of point processes of exceedances of random thresholds. Moreover, upper bounds on the variational distance between distributions of truncated empirical processes and limiting vectors of Poisson processes and, in Section 3, between distributions of vectors of Poisson processes are established. In Section 3, the density of $(X_{n:n}, Y_{n:n})$ w.r.t. an appropriate dominating measure is established without imposing any regularity conditions on the underlying d.f. F.

2. Rates and comparison of convergence

Let F_n denote the d.f. of (X_n, Y_n). Recall that F_n^n is the d.f. of $(X_{n:n}, Y_{n:n})$. If F_n^n weakly converges to a d.f. G then we know (cf. Resnick ((1987), Proposition 5.1)) that G is max-infinitely divisible (in short: max-i.d.); that is, for every positive integer n there exists a d.f. G_n such that $G_n^n = G$. Subsequently, let G denote a max-i.d. d.f. The starting point of our approach is a representation of max-i.d. d.f.'s due to Balkema and Resnick (1977) that was frequently applied since then. Define the 'lower endpoint' of a max-i.d. d.f. G by $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \equiv (\alpha(G_1), \alpha(G_2))$ where $\alpha(G_j)$ denotes the left endpoint of the support of the j-th marginal, say, G_j of G. Recall from Resnick ((1987), Proposition 5.8) that a max-i.d. d.f. G can be represented by

$$G(\boldsymbol{x}) = \begin{cases} \exp(-\nu([-\boldsymbol{\infty}, \boldsymbol{x}]^c)), & \boldsymbol{x} \geq \boldsymbol{\alpha}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.1}$$

where $\boldsymbol{\alpha} \in [-\boldsymbol{\infty}, \boldsymbol{\infty})$ and ν is the exponent measure (called max-Lévy measure in Giné *et al.* (1990)) pertaining to G. According to (2.1), $G(\boldsymbol{x}) > 0$ if $\boldsymbol{x} > \boldsymbol{\alpha}$ and hence $\boldsymbol{\alpha}$ is the lower endpoint described above. Keep in mind that an exponent measure ν has the following properties: (i) ν has its mass on $[\boldsymbol{\alpha}(G), \boldsymbol{\infty})\backslash\{\boldsymbol{\alpha}(G)\}$, (ii) $\nu([-\boldsymbol{\infty}, \boldsymbol{t}]^c) < \infty$ for every $\boldsymbol{t} > \boldsymbol{\alpha}(G)$, (iii) ν defines a d.f. in (2.1).

In the sequel, ν will always denote the exponent measure of the max-i.d. d.f. G. Note that an important sub-class of max-i.d. d.f.'s is built by max-stable d.f.'s G that possess the property

$$G^n(\boldsymbol{a}_n \boldsymbol{x} + \boldsymbol{b}_n) = G(\boldsymbol{x}), \qquad n \in \boldsymbol{N}, \tag{2.2}$$

for normalizing vectors $\boldsymbol{a}_n > 0$ and $\boldsymbol{b}_n$.

Denote by $\boldsymbol{M}(S)$ the set of point measures on $S = \boldsymbol{R}$ or $S = [-\boldsymbol{\infty}, \boldsymbol{\infty})$. The vector of processes $N_{n,1,t_1}$ and $N_{n,2,t_2}$ may be represented by means of N_n using a 'projection-truncation' operation. For that purpose consider the j-th projection π_j (given by $\pi_j(\boldsymbol{x}) = x_j$). Then the 'projection-truncation' map $\Pi_{\boldsymbol{t}}$ is defined by

$$\Pi_{\boldsymbol{t}}(\mu) = ((\pi_1\mu)(\cdot \cap (t_1, \infty)), (\pi_2\mu)(\cdot \cap (t_2, \infty))), \qquad \mu \in \boldsymbol{M}([-\boldsymbol{\infty}, \boldsymbol{\infty})), \tag{2.3}$$

where $\pi_j\mu$ is the measure induced by π_j and μ and the single components of $\Pi_{\boldsymbol{t}}(\mu)$ are regarded as measures on $(\boldsymbol{R}, \boldsymbol{B})$. Notice that

$$(N_{n,1,t_1}, N_{n,2,t_2}) = \Pi_{\boldsymbol{t}}(N_n). \tag{2.4}$$

The limiting distribution of that vector can be described by a Poisson process N^* with intensity measure ν. Let N_{1,t_1}^* and N_{2,t_2}^* be the Poisson processes with representation $(N_{1,t_1}^*, N_{2,t_2}^*) = \Pi_{\boldsymbol{t}}(N^*)$.

The intensity measures of N_{n,i,t_i} and N_{i,t_i}^* are denoted by ν_{n,i,t_i} and ν_{i,t_i}. Keep in mind that, in our notation, we do not distinguish between a d.f. and the pertaining probability measure. Moreover, the variational distance between distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ will be denoted by $\|\mathcal{L}(X) - \mathcal{L}(Y)\|$.

THEOREM 2.1. *Let G be a max-i.d. d.f. with exponent measure ν. We have*

$$\|\mathcal{L}(N_{n,1,t_1}, N_{n,2,t_2}) - \mathcal{L}(N^*_{1,t_1}, N^*_{2,t_2})\| \le C\left(\frac{-\log G(\boldsymbol{t})}{n} + A_{n,t}\right) \tag{2.5}$$

where

$$A_{n,t} = \|\nu_{n,1,t_1} - \nu_{1,t_1}\| + \|\nu_{n,2,t_2} - \nu_{2,t_2}\| + \|\nu_{n,t} - \nu_t\|.$$

PROOF. Arguing like in Reiss ((1990), proof of Theorem 2), we obtain

$$\|\mathcal{L}(N_{n,1,t_1}, N_{n,2,t_2}) - \mathcal{L}(N^*_{n,1,t_1}, N^*_{n,2,t_2})\| \le \nu_n[-\infty, \boldsymbol{t}]^c/n$$

where $(N^*_{n,1,t_1}, N^*_{n,2,t_2}) = \Pi_t(N^*_n)$ and N^*_n is a Poisson process having the same intensity measure ν_n as N_n. Moreover, applying the triangle inequality we deduce from Lemma 3.1 that

$$\begin{aligned}\|\mathcal{L}(N_{n,1,t_1}, N_{n,2,t_2}) - \mathcal{L}(N^*_{1,t_1}, N^*_{2,t_2})\| &\le C\left(\nu_n[-\infty, \boldsymbol{t}]^c/n + A_{n,t}\right)\\ &\le 2C\left(\frac{-\log G(\boldsymbol{t})}{n} + A_{n,t}\right).\end{aligned}$$

The proof is complete. □

One possible direction of further research work is to find conditions under which the upper bound in (2.5) can be replaced by a more accessible one. Omey and Rachev (1991) obtained rates of convergence for distributions of multivariate maxima w.r.t. weighted Kolmogorov distance and Lévy-Prohorov distance.

Another ingredient of our theory is the notion of an admissible threshold. A measurable map $T : \boldsymbol{M}(\boldsymbol{R}) \to [-\infty, \infty)$ is an admissible threshold (for a univariate d.f. G) if

$$\max(x, T(\mu)) = \max(x, T(\mu_x)), \qquad \mu \in \boldsymbol{M}(\boldsymbol{R}), \quad x \in \boldsymbol{R}, \tag{2.6}$$

with μ_x denoting the truncation of μ left of x, and

$$P\{T(N^*) \le x\} \to 0 \quad \text{as} \quad x \downarrow \alpha(G) \tag{2.7}$$

where N^* is a Poisson process satisfying $E\,N(x, \infty) = -\log G(x)$.

In the bivariate case, a 'threshold' $T = (T_1, T_2)$ is admissible (for G) if T_j is admissible for the j-th marginals G_j of G for $j = 1, 2$.

Example 2.1. (i) The constant, univariate threshold $T(\mu) = t$ is admissible if $t > \alpha(G)$.

(ii) If G is continuous at $\alpha(G)$ then

$$T(\mu) = \begin{cases} k\text{-th largest point of } \mu & \text{if } \mu(\boldsymbol{R}) \ge k \\ -\infty & \text{if } \mu(\boldsymbol{R}) < k \end{cases}$$

is admissible. To see this notice that

$$P\{T(N^*) \le x\} = P\{N^*(x,\infty) \le k-1\} = G(x)\sum_{i=0}^{k-1} \frac{(-\log G(x))^i}{i!}$$

and the last expression converges to zero as $x \downarrow \alpha(G)$ due to the continuity of G at $\alpha(G)$. If $k=1$ we will write $T(\mu) = \max(\mu)$.

Our main result unifies and extends several results known in the literature. It deals with the joint distribution of the k largest order statistics as well as with point processes of exceedances of non-random thresholds.

THEOREM 2.2. *Let G be a max-i.d. d.f. with exponent measure ν. If the marginal d.f.'s G_j are continuous at $\alpha(G_j)$ then the following assertions are equivalent:*

(i) *For every $\mathbf{t} > \boldsymbol{\alpha}(G)$,*

$$\|\mathcal{L}(N_{n,1,t_1}, N_{n,2,t_2}) - \mathcal{L}(N^*_{1,t_1}, N^*_{2,t_2})\| \to 0, \qquad n \to \infty. \tag{2.8}$$

(ii) *For every admissible threshold $T = (T_1, T_2)$, as $n \to \infty$,*

$$\|\mathcal{L}(N_{n,1,T_1(N_{n,1})}, N_{n,2,T_2(N_{n,2})}) - \mathcal{L}(N^*_{1,T_1(N^*_1)}, N^*_{2,T_2(N^*_2)})\| \to 0. \tag{2.9}$$

(iii)

$$\|\mathcal{L}(X_{n:n}, Y_{n:n}) - G\| \to 0, \qquad n \to \infty. \tag{2.10}$$

(iv) *For every $\mathbf{t} > \boldsymbol{\alpha}(G)$, as $n \to \infty$,*

$$A_{n,t} = \|\nu_{n,1,t_1} - \nu_{1,t_1}\| + \|\nu_{n,2,t_2} - \nu_{2,t_2}\| + \|\nu_{n,t} - \nu_t\| \to 0. \tag{2.11}$$

The continuity condition is merely required to verify in the proof of (iii) that the maximum satisfies condition (2.7). This condition cannot be omitted without compensation. This is simply due to the fact that assertion (2.8) merely determines the asymptotic behavior of d.f.'s F_n^n of maxima on the set $(-\boldsymbol{\infty}, \boldsymbol{\alpha}(G)]^c$. In the univariate case (explicitly formulated in Remark 2.1), assertion (i) holds for $F_n = G_1^{1/n}$, where G_1 is continuous at $\alpha(G_1)$ and $G = \tilde{G}_1 = G_1 1_{[u,\infty)}$ with $u > \alpha(G_1)$. Yet $F_n^n = G_1$ and, hence, the convergence in (2.10) is not valid.

Remark 2.1. Theorem 2.2 entails the following univariate version (for convenience written down in the same notation): Assume that G_1 is continuous at $\alpha(G_1)$. Then the following four assertions are equivalent:

(i) For every $t_1 > \alpha(G_1)$, $\|\mathcal{L}(N_{n,1,t_1}) - \mathcal{L}(N^*_{1,t_1})\| \to 0$, $n \to \infty$.

(ii) For every admissible threshold T_1, $\|\mathcal{L}(N_{n,1,T_1(N_{n,1})}) - \mathcal{L}(N^*_{1,T_1(N^*_1)})\| \to 0$, $n \to \infty$.

(iii) $\|\mathcal{L}(X_{n:n}) - G_1\| \to 0$, $n \to \infty$.

(iv) For every $t_1 > \alpha(G_1)$, $\|\nu_{n,1,t_1} - \nu_{1,t_1}\| \to 0$, $n \to \infty$.

From Example 2.1(ii) we know that the k-th largest point defines an admissible threshold. Hence the equivalence of (ii) and (iii) in Remark 2.1 yields that the joint distribution of the k largest order statistics strongly converges if, and only if, such a result holds for the maximum. This is well-known in case of weak convergence (cf. Galambos ((1987), Theorem 2.8.2)) and was proven in the case of strong convergence in Reiss (1989) under the condition that the underlying d.f. possesses a Lebesgue density. Moreover, in the univariate case it was proven by Falk and Reiss (1992) that the standardized maxima $X_{n:n}$ strongly converge to a max-stable r.v. if, and only if, strong convergence of the pertaining empirical point processes $N_{n,1,t_1}$ to a certain Poisson process holds. Our present conditions are slightly weaker. For a related result concerning the weak convergence of univariate empirical point processes see Resnick ((1987), Corollary 4.19).

PROOF OF THEOREM 2.2. The implication (iv) $\Rightarrow$ (i) is immediate from Theorem 2.1. It remains to prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv).

(i) $\Rightarrow$ (ii): Let $\boldsymbol{t} > \boldsymbol{\alpha}$. Denote by $N_{t,s}$ subsequent truncation left of t and s. We have

$$\begin{aligned}(2.12)\quad &\|\mathcal{L}(N_{n,1,T_1(N_{n,1})}, N_{n,2,T_2(N_{n,2})}) - \mathcal{L}(N_{n,1,t_1,T_1(N_{n,1})}, N_{n,2,t_2,T_2(N_{n,2})})\| \\ &\quad\le P\{(N_{n,1,T_1(N_{n,1})}, N_{n,2,T_2(N_{n,2})}) \\ &\qquad\qquad \neq (N_{n,1,t_1,T_1(N_{n,1})}, N_{n,2,t_2,T_2(N_{n,2})})\} \\ &\quad\le P\{T_1(N_{n,1}) \le t_1\} + P\{T_2(N_{n,2}) \le t_2\} \\ &\quad= P\{T_1(N_{n,1,t_1}) \le t_1\} + P\{T_2(N_{n,2,t_2}) \le t_2\}\end{aligned}$$

where the final step is immediate from condition (2.6). According to (i), N_{n,j,t_j} can asymptotically be replaced by N^*_{j,t_j} as $n \to \infty$ and applying (2.6) again we see that N^*_{j,t_j} can be replaced by N^*_j. Moreover, (2.7) implies that $P\{T_j(N^*_j) \le t_j\}$ can be made arbitrarily small for t_j sufficiently close to α_j. Notice that (2.6) implies $\mu_{x,T(\mu)} = \mu_{x,T(\mu_x)}$ and, hence, $N_{n,j,t_j,T_j(N_{n,j})} = N_{n,j,t_j,T_j(N_{n,j,t_j})}$. Because the last term only depends on N_{n,j,t_j} (i) implies

$$\begin{aligned}(2.13)\quad &\|\mathcal{L}(N_{n,1,t_1,T_1(N_{n,1})}, N_{n,2,t_2,T_2(N_{n,2})}) - \mathcal{L}(N^*_{1,t_1,T_1(N^*_{1,t_1})}, N^*_{2,t_2,T_2(N^*_{2,t_2})})\| \\ &\quad\to 0, \qquad n \to \infty.\end{aligned}$$

Again, $N^*_{j,t_j,T_j(N^*_{j,t_j})} = N^*_{j,t_j,T_j(N^*_j)}$ and repeating the arguments in (2.12) we see that

$$(2.14)\qquad \|\mathcal{L}(N^*_{1,t_1,T_1(N^*_1)}, N^*_{2,t_2,T_2(N^*_2)}) - \mathcal{L}(N^*_{1,T_1(N^*_1)}, N^*_{2,T_2(N^*_2)})\|$$

can be made arbitrarily small for t_j sufficiently close to α_j.

Combining (2.12)–(2.14) the proof can easily be completed.

(ii) $\Rightarrow$ (iii): Apply (ii) to thresholds $T_j(\mu) = \max(\mu)$ as defined in Example 2.1(ii). Identify $(N_{n,1,\max(N_{n,1})}, N_{n,2,\max(N_{n,2})})$ with $(X_{n:n}, Y_{n:n})$ and identify $N^*_{j,\max(N^*_j)}$ with $\max(N^*_j)$. Moreover, notice that

$$P\{(\max(N^*_1), \max(N^*_2)) \le \boldsymbol{x}\} = P\{N^*([-\infty, \boldsymbol{x}]^c) = 0\} = G(\boldsymbol{x}).$$

(iii) $\Rightarrow$ (iv): Let ρ be a σ-finite measure with marginals ρ_1 and ρ_2 such that $\mathcal{L}(X_{n,i}, Y_{n,i})$ has the ρ-density f_n, G has the ρ-density g, and $\rho_1 \times \rho_2$ has the ρ-density $h_{1,2}$; e.g., let $\rho = \tilde{\rho} + \tilde{\rho}_1 \times \tilde{\rho}_2$ where $\tilde{\rho} = G + \sum_{n=1}^{\infty} 2^{-n}\mathcal{L}(X_{n,1}, Y_{n,1})$ (see Section 3 for details). Denote by $F_{n,j}$ the j-th marginal of F_n and by $f_{n,j}$ the ρ_j-density of $F_{n,j}$.

We will show that the ρ_t-densities of $\nu_{n,t}$, namely, nf_n converge to the ρ_t-density of ν_t in $L_1(\rho_t)$ for every $\boldsymbol{t} > \boldsymbol{\alpha}$.

Lemma 3.2 implies that $(X_{n:n}, Y_{n:n})$ has the ρ-density

$$\begin{aligned} f_{(n)}(x, y) = {} & nf_n(x, y)S_{n,1}(x, y) \\ & + (h_{n,1}(x, y)S_{n,2}(x, y)h_{n,2}(x-, y) \\ & \quad + h_{n,1}(x, y-)S_{n,3}(x, y)h_{n,2}(x, y))h_{1,2}(x, y), \end{aligned}$$

where $S_{n,i}$ and $h_{n,j}$ are defined as S_i and h_j, respectively, in (3.5) with F and f replaced by F_n and f_n, respectively. Let

$$F_{n,1}(\cdot \mid y) := P(X_n \le \cdot \mid Y_n = y) \quad \text{and} \quad F_{n,2}(\cdot \mid x) := P(Y_n \le \cdot \mid X_n = x).$$

The rest of the proof is organized as follows. We verify that

(a) $S_{n,1}(\boldsymbol{x})$, $(2/n(n-1))S_{n,2}(\boldsymbol{x})$, $(2/n(n-1))S_{n,3}(\boldsymbol{x})$ are bounded away from 0 and ∞ uniformly for $\boldsymbol{x} \ge \boldsymbol{t}$ and large n and converge on $(\boldsymbol{t}, \boldsymbol{\infty})$ for every $\boldsymbol{t} > \boldsymbol{\alpha}$.

(b) $(S_{n,2}/S_{n,1})h_{n,1}h_{n,2}(\cdot-, \cdot) + (S_{n,3}/S_{n,1})h_{n,1}(\cdot, \cdot-)h_{n,2}$ converges in $L_1((\rho_1 \times \rho_2)_t)$.

(c) If $nf_n \to \tilde{g}$, $n \to \infty$, in $L_1(\rho_t)$ for every $\boldsymbol{t} > \boldsymbol{\alpha}$, then $\tilde{g}$ is a $\rho_{\boldsymbol{\alpha}}$-density of $\nu_{\boldsymbol{\alpha}}$.

It is immediate from (a) and (b) that

$$nf_n = \frac{f_{(n)}}{S_{n,1}} - \left(\frac{S_{n,2}}{S_{n,1}} h_{n,1}h_{n,2}(\cdot-, \cdot) + \frac{S_{n,3}}{S_{n,1}} h_{n,1}(\cdot, \cdot-)h_{n,2}\right) h_{1,2}$$

has a limit, say $\tilde{g}$, in $L_1(\rho_t)$ for every $\boldsymbol{t} > \boldsymbol{\alpha}$. According to (c), $\tilde{g}$ is a $\rho_{\boldsymbol{\alpha}}$-density of $\nu_{\boldsymbol{\alpha}}$. Hence, $\|\nu_{n,t} - \nu_t\| \to 0$, $n \to \infty$, for every $\boldsymbol{t} > \boldsymbol{\alpha}$. Analogously, one may prove $\|\nu_{n,j,t_j} - \nu_{j,t_j}\| \to 0$, $n \to \infty$ for $j = 1, 2$.

To prove (a) notice that $S_{n,1}(x, y) \in [F_n^{n-1}(x-, y-), F_n^{n-1}(x, y)]$. This implies $S_{n,1}(x, y) \to G(x, y)$ if (x, y) is a continuity point of G. If $G(x-, y-) < G(x, y)$ then, as $n \to \infty$,

$$\begin{aligned} S_{n,1}(x, y) &= n^{-1}F_n^{n-1}(x, y)\sum_{i=1}^{n}(F_n(x-, y-)/F_n(x, y))^{n-i} \\ &= n^{-1}\frac{F_n^n(x, y) - F_n^n(x-, y-)}{F_n(x, y) - F_n(x-, y-)} \to \frac{G(x, y) - G(x-, y-)}{\log G(x, y) - \log G(x-, y-)}. \end{aligned}$$

Moreover, (iii) implies that there exist S_1, S_2, S_3 such that

$$\left(S_{n,1},\ \frac{2}{n(n-1)}S_{n,2},\ \frac{2}{n(n-1)}S_{n,3}\right) \to (S_1, S_2, S_3) \tag{2.15}$$

as $n \to \infty$. Each component on the left-hand side is bounded from below and from above by $F_n^{n-1}(\cdot -, \cdot -)$ and F_n^{n-2}, respectively. Moreover, $F_n^{n-1}(\cdot -, \cdot -)$ and F_n^{n-2} converge pointwise to $G(\cdot -, \cdot -)$ and G. This implies (a).

To prove (b) define

$$T_{n,1}(x,y) = \sum_{i=1}^{n} F_n(x,y)^{i-1} F_n(x-,y)^{n-i} \qquad \text{and}$$
$$T_{n,2}(x,y) = \sum_{i=1}^{n} F_n(x,y)^{i-1} F_n(x,y-)^{n-i}$$

and denote by $g_1(x,y)$ and $g_2(x,y)$ the ρ_1-density of $G(\cdot \times (-\infty, y])$ and the ρ_2-density of $G((-\infty, x] \times \cdot)$.

Elementary calculations as in (2.15) show that $S_{n,2}/(S_{n,1}T_{n,1}T_{n,2}(\cdot -, \cdot))$ converges to a limiting function on $(\boldsymbol{\alpha}, \boldsymbol{\infty})$ as $n \to \infty$, which is bounded on $(\boldsymbol{t}, \boldsymbol{\infty})$ for every $\boldsymbol{t} > \boldsymbol{\alpha}$.

Next, we show

$$g_{n,1}g_{n,2}(\cdot -, \cdot) \to g_1 g_2(\cdot -, \cdot), \; n \to \infty, \qquad \text{in} \qquad L_1((\rho_1 \times \rho_2)_t)$$

where $g_{n,1}(x,y) = T_{n,1}(x,y)h_{n,1}(x,y)$ is the density of $P\{X_{n:n} \in \cdot, Y_{n:n} \le y\}$ w.r.t. ρ_1, and $g_{n,2}(x-,y) = T_{n,2}(x-,y)h_{n,2}(x-,y)$ is the density of $P\{X_{n:n} < x, Y_{n:n} \in \cdot\}$ w.r.t. ρ_2. It follows from (iii) that

$$\begin{aligned}
&\int |g_{n,1}(\cdot, y) - g_1(\cdot, y)| d\rho_1 \to 0 \text{ uniformly over } y \in \boldsymbol{R} \cup \{\infty\}, \\
&\int |g_{n,2}(x-, \cdot) - g_2(x-, \cdot)| d\rho_2 \to 0 \text{ uniformly over } x \in \boldsymbol{R} \cup \{\infty\}
\end{aligned} \tag{2.16}$$

as $n \to \infty$. Hence,

$$\begin{aligned}
&\int |g_{n,1}(x,y)g_{n,2}(x-,y) - g_1(x,y)g_2(x-,y)| d(\rho_1 \times \rho_2)(x,y) \\
&\quad \le \int g_{n,1}(x,y)|g_{n,2}(x-,y) - g_2(x-,y)| d(\rho_1 \times \rho_2)(x,y) \\
&\qquad + \int g_2(x-,y)|g_{n,1}(x,y) - g_1(x,y)| d(\rho_1 \times \rho_2)(x,y) \\
&\quad \le \int g_{n,1}(x,\infty) \int |g_{n,2}(x-,y) - g_2(x-,y)| d\rho_2(y) d\rho_1(x) \\
&\qquad + \int g_2(\infty-,y) \int |g_{n,1}(x,y) - g_1(x,y)| d\rho_1(x) d\rho_2(y) \\
&\quad \to 0, \qquad n \to \infty.
\end{aligned}$$

Using the preceding arguments once more, one may show that

$$S_{n,3}/(S_{n,1}h_{n,1}(\cdot, \cdot -)h_{n,2})h_{1,2}$$

has a limit in $L_1(\rho_t)$ for $\boldsymbol{t} > \boldsymbol{\alpha}$.

Moreover, (c) is valid if $\tilde{g}$ is a $\rho_{\boldsymbol{\alpha}}$-density of $\nu_{\boldsymbol{\alpha}}$. Since

$$n(1 - F_n(\boldsymbol{x})) \to -\log G(\boldsymbol{x}), \qquad n \to \infty, \quad \text{for } \boldsymbol{x} > \boldsymbol{\alpha}$$

we obtain

$$\begin{aligned}\nu(\boldsymbol{x}, \boldsymbol{\infty}) &= \int_{(\boldsymbol{x},\boldsymbol{\infty})} d\log(G) = \lim_n \int_{(\boldsymbol{x},\boldsymbol{\infty})} d(nF_n) \\ &= \lim_n \int_{(\boldsymbol{x},\boldsymbol{\infty})} nf_n d\rho = \int_{(\boldsymbol{x},\boldsymbol{\infty})} \tilde{g} d\rho,\end{aligned}$$

for every $\boldsymbol{x} > \boldsymbol{\alpha}$. Hence, $\tilde{g}$ is a $\rho_{\boldsymbol{\alpha}}$-density of $\nu_{\boldsymbol{\alpha}}$. The proof is complete. □

3. Auxiliary results

The following lemma provides an upper bound on the variational distance between certain vectors of Poisson processes by means of the sum of the variational distances between intensity measures.

LEMMA 3.1. *Let N^* and N^{**} be Poisson processes on $[-\boldsymbol{\infty}, \boldsymbol{\infty})$ with finite intensity measures ν^* and ν^{**}. Put $(N^*_{1,t_1}, N^*_{2,t_2}) = \Pi_{\boldsymbol{t}} N^*$ and $(N^{**}_{1,t_1}, N^{**}_{2,t_2}) = \Pi_{\boldsymbol{t}} N^{**}$. Denote by ν^*_j and ν^{**}_j the intensity measures of N^*_j and N^{**}_j, respectively, for $j = 1, 2$. We have*

$$\begin{aligned}&\|\mathcal{L}(N^*_{1,t_1}, N^*_{2,t_2}) - \mathcal{L}(N^{**}_{1,t_1}, N^{**}_{2,t_2})\| \\ &\quad \le C(\|\nu^*_{1,t_1} - \nu^{**}_{1,t_1}\| + \|\nu^*_{2,t_2} - \nu^{**}_{2,t_2}\| + \|\nu^*_{\boldsymbol{t}} - \nu^{**}_{\boldsymbol{t}}\|)\end{aligned}$$

for some universal constant $C > 0$.

PROOF. Put $D(1) = (t_1, \infty) \times [-\infty, t_2]$ and $D(2) = [-\infty, t_1] \times (t_2, \infty)$. For $N \in \{N^*, N^{**}\}$ we get $N_{j,t_j} = \pi_j(N|_{D(j)}) + \pi_j(N_{\boldsymbol{t}})$, where $N|_D$ denotes the truncation of N outside of D. Recall that $N_{\boldsymbol{t}} = N|_{(\boldsymbol{t},\boldsymbol{\infty})}$. Applying the monotonicity theorem (see, e.g., Liese and Vajda (1987)) we obtain

$$\begin{aligned}&\|\mathcal{L}(N^*_{1,t_1}, N^*_{2,t_2}) - \mathcal{L}(N^{**}_{1,t_1}, N^{**}_{2,t_2})\| \\ &\quad \le \|\mathcal{L}(\pi_1(N^*|_{D(1)}), \pi_2(N^*|_{D(2)}), N^*_{\boldsymbol{t}}) \\ &\qquad - \mathcal{L}(\pi_1(N^{**}|_{D(1)}), \pi_2(N^{**}|_{D(2)}), N^{**}_{\boldsymbol{t}})\| \\ &\quad \le \|\mathcal{L}(\pi_1(N^*|_{D(1)})) - \mathcal{L}(\pi_1(N^{**}|_{D(1)}))\| \\ &\qquad + \|\mathcal{L}(\pi_2(N^*|_{D(2)})) - \mathcal{L}(\pi_2(N^{**}|_{D(2)}))\| + \|\mathcal{L}(N^*_{\boldsymbol{t}}) - \mathcal{L}(N^{**}_{\boldsymbol{t}})\|\end{aligned}$$

where the second inequality holds because of the independence of the involved processes. From Matthes *et al.* ((1978), Proposition 1.12.1), it follows that $\|\mathcal{L}(N^*_{\boldsymbol{t}}) - \mathcal{L}(N^{**}_{\boldsymbol{t}})\| \le C\|\nu^*_{\boldsymbol{t}} - \nu^{**}_{\boldsymbol{t}}\|$.

Moreover, check that for $j = 1, 2$,

$$\begin{aligned}\|\mathcal{L}(\pi_j(N^*|_{D(j)})) - \mathcal{L}(\pi_j(N^{**}|_{D(j)}))\| &\le C\|\pi_j(\nu^*|_{D(j)}) - \pi_j(\nu^{**}|_{D(j)})\| \\ &\le C(\|\nu^*_{j,t_j} - \nu^{**}_{j,t_j}\| + \|\nu^*_{\boldsymbol{t}} - \nu^{**}_{\boldsymbol{t}}\|)\end{aligned}$$

and, hence, the asserted inequality holds. □

Hereafter, let $(X_{n:n}, Y_{n:n})$ be the sample maximum of n copies (X_i, Y_i) of a random vector (X, Y) with d.f. F. Denote by F_j the j-th marginal of F. Let

$$F_1(\cdot \mid y) := P(X \le \cdot \mid Y = y) \quad \text{and} \quad F_2(\cdot \mid x) := P(Y \le \cdot \mid X = x).$$

If F possesses a Lebesgue density f then one knows (see Reiss ((1989), (2.2.7))) that $(X_{n:n}, Y_{n:n})$ has the Lebesgue density

$$f_{(n)} = nF^{n-1} f + n(n-1)F^{n-2} h_1 h_2 \tag{3.1}$$

where $h_1(x,y) = f_1(x)F_2(y \mid x)$ and $h_2(x,y) = f_2(y)F_1(x \mid y)$ with f_1, f_2 denoting the marginals of f. In the sequel, we will prove an extension of this result that will also confirm a conjecture stated in Reiss ((1989), Problem 2.8).

LEMMA 3.2. *Let ρ be a σ-finite measure with marginals ρ_1 and ρ_2 such that*

$$(X, Y) \text{ has the } \rho\text{-density } f \tag{3.2}$$

and

$$\rho_1 \times \rho_2 \text{ has the } \rho\text{-density } h_{1,2}. \tag{3.3}$$

Then, $(X_{n:n}, Y_{n:n})$ has the ρ-density $f_{(n)}$ given by

$$\begin{aligned} f_{(n)}(x,y) = {} & nS_1(x,y)f(x,y) \\ & + (h_1(x,y)S_2(x,y)h_2(x-,y) + h_1(x,y-)S_3(x,y)h_2(x,y)) \\ & \cdot h_{1,2}(x,y) \end{aligned} \tag{3.4}$$

where $h_2(x-,y) = \lim_{z \uparrow x} h_2(z,y)$ etc. and

$$\begin{aligned} S_1(x,y) &= n^{-1} \sum_{i=1}^{n} F(x,y)^{i-1} F(x-,y-)^{n-i}, \\ S_2(x,y) &= \sum_{1 \le i < j \le n} F(x,y)^{i-1} F(x-,y)^{j-i-1} F(x-,y-)^{n-j}, \\ S_3(x,y) &= \sum_{1 \le j < i \le n} F(x,y)^{j-1} F(x,y-)^{i-j-1} F(x-,y-)^{n-i}, \\ h_1(x,y) &= f_1(x)F_2(y \mid x) \quad \text{and} \quad h_2(x,y) = f_2(y)F_1(x \mid y) \end{aligned} \tag{3.5}$$

with f_j denoting the ρ_j-density of F_j for $j = 1, 2$.

If condition (3.2) holds and (3.3) is not valid then ρ may be replaced by $\rho + \rho_1 \times \rho_2$. If F is continuous then (3.4) reduces to

$$f_{(n)} = nF^{n-1} f + n(n-1)F^{n-2} h_1 h_2 h_{1,2}. \tag{3.6}$$

In addition, $h_{1,2} \equiv 1$ if $\rho = \rho_1 \times \rho_2$. Thus, (3.1) is immediate from (3.6) with ρ being the Lebesgue measure on $\boldsymbol{R}^2$. Moreover, one may easily check that

$$(x, y) \to f_1(x) f_2(y) h_{1,2}(x, y)$$

is a ρ-density of $F_1 \times F_2$. Hence,

$$\begin{aligned} &P\{(X_{n:n}, Y_{n:n}) \in B\} \\ &\quad = \int_B nF^{n-1} dF + \int_B n(n-1) F^{n-2} F_1(x \mid y) F_2(y \mid x) d(F_1 \times F_2)(x, y) \end{aligned}$$

which is the formula given in Reiss ((1989), Problem 2.8).

Proof of Lemma 3.2. Let $(x, y) \in \boldsymbol{R}^2$. Then,

$$\begin{aligned} &P\{(X_{n:n}, Y_{n:n}) \le (x, y)\} \\ &\quad = \sum_{1 \le i,j \le n} P\{(X_i, Y_j) \le (x, y), X_{i_0} \le X_i, X_{i_1} < X_i, i_0 < i < i_1, \\ &\qquad\qquad Y_{j_0} \le Y_j, Y_{j_1} < Y_j, j_0 < j < j_1\} \\ &\quad = \sum_{1 \le i = j \le n} \int_{(-\infty, (x,y)]} F(u, v)^{i-1} F(u-, v-)^{n-i} d\mathcal{L}(X_i, Y_j)(u, v) \\ &\qquad + \sum_{1 \le i \ne j \le n} \int_{(-\infty, (x,y)]} P\{X_{i_0} \le X_i, X_{i_1} < X_i, i_0 < i < i_1, \\ &\qquad Y_{j_0} \le Y_j, Y_{j_1} < Y_j, j_0 < j < j_1 \mid X_i = u, Y_j = v\} d\mathcal{L}(X_i, Y_j)(u, v). \end{aligned}$$

The first term is equal to $\int_{(-\infty,(x,y)]} (nS_1 f) d\rho_1$. Moreover, we get the following equalities for the integrands in the second term for $\mathcal{L}(X_i, Y_j)$ almost all (u, v) if $1 \le i < j \le n$:

$$\begin{aligned} &P\{X_{i_0} \le X_i, X_{i_1} < X_i, i_0 < i < i_1, \\ &\quad Y_{j_0} \le Y_j, Y_{j_1} < Y_j, j_0 < j < j_1 \mid X_i = u, Y_j = v\} \\ &\quad = \int_{(-\infty, v]} P\{X_{i_0} \le u, X_{i_1} < u, i_0 < i < i_1, Y_{j_0} \le v, Y_{j_1} < v, j_0 < j < j_1 \mid \\ &\qquad Y_i = s, X_i = u, Y_j = v\} P(Y_i \in ds \mid X_i = u, Y_j = v) \\ &\quad = \int_{(-\infty, v]} F(u, v)^{i-1} F(u-, v)^{j-i-1} F(u-, v-)^{n-j} \\ &\qquad \times P(X_j < u \mid Y_j = v) P(Y_i \in ds \mid X_i = u) \\ &\quad = F(u, v)^{i-1} F(u-, v)^{j-i-1} F(u-, v-)^{n-j} F_1(u- \mid v) F_2(v \mid u). \end{aligned}$$

Similar calculations may be carried out for $1 \le j < i \le n$. □

Acknowledgement

The work of the second author was partly done while visiting Cornell University. Thanks are due to S. I. Resnick for his hospitality and advice.

References

Balkema, A. A. and Resnick, S. (1977). Max-infinite divisibility, *J. Appl. Probab.*, **14**, 309–319.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds, *J. Roy. Statist. Soc. Ser. B*, **52**, 393–442.

Falk, M. and Reiss, R.-D. (1992). Poisson approximation of empirical processes, *Statist. Probab. Lett.*, **14**, 39–48.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, 2nd ed., Krieger, Malabar, Florida (1st ed. Wiley (1978)).

Giné, E., Hahn, M. G. and Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes, *Probab. Theory Related Fields*, **87**, 139–165.

Liese, F. and Vajda, I. (1987). *Convex Statistical Distances*, Teubner-Texte zur Mathematik, Bd. 95, Teubner, Leipzig.

Matthes, K., Kerstan, J. and Mecke, J. (1978). *Infinitely Divisible Point Processes*, Wiley, Chichester.

Omey, E. and Rachev, S. T. (1991). Rates of convergence in multivariate extreme value theory, *J. Multivariate Anal.*, **38**, 36–50.

Pickands, J. (1975). Statistical inference using extreme order statistics, *Ann. Statist.*, **3**, 119–131.

Reiss, R.-D. (1989). Approximate distributions of order statistics: with applications to nonparametric statistics, *Springer Ser. Statist.*, Springer, New York.

Reiss, R.-D. (1990). Asymptotic independence of marginal point processes of exceedances, *Statist. Decisions*, **8**, 153–165.

Resnick, S. I. (1987). Extreme values, regular variation, and point processes, *Appl. Probab. Ser. Appl. Probab. Trust*, **4**, Springer, New York.