

THE PARETO COPULA, AGGREGATION OF RISKS, AND THE EMPEROR'S SOCKS

CLAUDIA KLÜPPELBERG,* *Munich University of Technology*

SIDNEY I. RESNICK,** *Cornell University*

Abstract

The copula of a multivariate distribution is the distribution transformed so that one-dimensional marginal distributions are uniform. We review a different transformation of a multivariate distribution which yields standard Pareto for the marginal distributions, and we call the resulting distribution the *Pareto copula*. Use of the Pareto copula has a certain claim to naturalness when considering asymptotic limit distributions for sums, maxima, and empirical processes. We discuss implications for aggregation of risk and offer some examples.

Keywords: Regular variation; risk; maximal domain of attraction; copula; Pareto

2000 Mathematics Subject Classification: Primary 60E07; 60G51; 60G52; 60G70
Secondary 60F17; 62G30

1. Introduction

Religious copularians take as basic orthodoxy the desirability of transforming a multivariate distribution to have uniform marginals. Despite the shortcomings pointed out by the skeptic Mikosch (2005), (2006), this practice has become a fairly standard procedure. We argue that when one's objective is the study of limit distributions and asymptotic approximations, if one's religion requires transformation of marginal distributions, one would do better to transform marginals to the standard Pareto distribution. The resulting transformed distribution, which we call the Pareto copula, has natural interpretations for limit theory and heavy-tail analysis. This point of view will also show that several results attributed to properties of special copulas are, in fact, examples of more general properties of distributions.

Our transformation to Pareto marginals is not new and has been used in the study of multivariate domains of attraction to characterize these domains by means of multivariate regular variation. The method consists of transforming a domain of attraction condition to *standard* regular variation in which all components of the transformed vector are normalized by the same linear function. The technique dates at least to de Haan and Resnick (1977) and has been explained in de Haan and Ferreira (2006) and Resnick (1987), (2007).

Section 2 outlines the definition and basic properties of the Pareto copula in the context of a triangular array of random vectors $\{X_{n,j}; j \geq 1, n \geq 1\}$, where the rows consist of independent and identically distributed (i.i.d.) d -dimensional random vectors. We discuss the role of the Pareto copula in the study of asymptotic properties of empirical measures, extremes, and sums of entries in the n th row of the array as n tends to ∞ . Then, in Section 3, we specialize

Received 24 August 2007; revision received 11 December 2007.

* Postal address: Chair of Mathematical Statistics, Center for Mathematical Sciences, Munich University of Technology, Boltzmannstrasse 3, 85747 Garching bei München, Germany. Email address: cklu@ma.tum.de

** Postal address: School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA. Email address: sir1@cornell.edu

the triangular array setup to regular variation, where $X_{n,j} = X_j/b(n)$ for a suitable scaling function $b(\cdot)$ and i.i.d. random vectors $\{X_j\}$.

We also consider cases where the distribution of $\{X_j\}$ is in a maximal domain of attraction and study aggregation of risks: the asymptotic properties of the distribution of the sum of the components of $\{X_j\}$. We do this, when the vector's distribution is multivariate regularly varying, and also when the distribution of X_1 is in a maximal domain of attraction with equal one-dimensional marginals in a Gumbel domain and the distribution does not possess asymptotic independence. For this case, we obtain, without further assumptions, a reasonably explicit expression for the tail probabilities of the sum of the components.

1.1. Vector notation

Vectors are denoted by bold letters, random vectors are denoted by capital letters, and non-random vectors are denoted by lower-case letters. For example, $\mathbf{x} = (x^{(1)}, \dots, x^{(d)}) \in \mathbb{R}^d$. Operations between vectors should always be interpreted componentwise, so that, for two vectors \mathbf{x} and \mathbf{z} , $\mathbf{x} < \mathbf{z}$ means that $x^{(i)} < z^{(i)}$ for $i = 1, \dots, d$, with analogous notation for $\mathbf{x} \leq \mathbf{z}$ and $\mathbf{x} = \mathbf{z}$. If \mathbf{x}_j , for $j = 1, \dots, n$, are vectors, componentwise maxima are defined by $\bigvee_{j=1}^n \mathbf{x}_j = (\bigvee_{j=1}^n x_j^{(i)}, i = 1, \dots, d)$. Also, if $\boldsymbol{\alpha} = (\alpha^{(1)}, \dots, \alpha^{(d)}) \geq \mathbf{0}$, we write $\mathbf{x}^\alpha = ((x^{(1)})^{\alpha^{(1)}}, \dots, (x^{(d)})^{\alpha^{(d)}})$ for $\mathbf{x} \geq \mathbf{0}$. Furthermore, we define $\mathbf{0} = (0, \dots, 0)$, $\mathbf{1} = (1, \dots, 1)$, and $\infty = (\infty, \dots, \infty)$. For a real number c , we write as usual $c\mathbf{x} = (cx^{(1)}, \dots, cx^{(d)})$. We denote the rectangles (or the higher-dimensional intervals) by $[\mathbf{a}, \mathbf{b}] = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}\}$ with analogous notation for rectangles with one or both endpoints open.

To fix ideas, for now suppose that $\mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$. Complements are taken with respect to \mathbb{E} , so that, for $\mathbf{x} > \mathbf{0}$,

$$[\mathbf{0}, \mathbf{x}]^c = \mathbb{E} \setminus [\mathbf{0}, \mathbf{x}] = \left\{ \mathbf{y} \in \mathbb{E} : \bigvee_{i=1}^d \frac{y^{(i)}}{x^{(i)}} > 1 \right\}.$$

1.2. Symbol and concept list

Here is a glossary of miscellaneous symbols and nomenclature used throughout the paper.

- RV_ρ The class of regularly varying functions on $[0, \infty)$ with index $\rho \in \mathbb{R}$.
- f^\leftarrow The left-continuous inverse of a monotone function f , defined by $f^\leftarrow(x) = \inf\{y : f(y) \geq x\}$.
- $b(t)$ Usually the quantile function of a distribution function $F(x)$, defined by $b(t) = F^\leftarrow(1 - 1/t)$, but usage can vary somewhat with context.
- \xrightarrow{v} Vague convergence of measures.
- \xrightarrow{D} Convergence in distribution.
- $\stackrel{D}{=}$ Equality in distribution.
- ε_x The probability measure consisting of all mass at x .
- $M_+(\mathbb{E})$ The space of nonnegative Radon measures on \mathbb{E} .
- $M_p(\mathbb{E})$ The space of Radon point measures on \mathbb{E} .
- $\text{PRM}(\mu)$ Poisson random measure on \mathbb{E} with mean measure μ .

2. The Pareto copula

2.1. Basics

Consider a triangular array of random vectors $\{X_{n,j}; n \geq 1, j \geq 1\}$ in which the rows are i.i.d. The distribution of $X_{n,1}$ is F_n . We suppose that the random vectors are \mathbb{R}^d -valued and, for simplicity, we assume that the one-dimensional marginal distributions $F_n^{(i)}$ are continuous. If $X_{n,j} = (X_{n,j}^{(i)}; i = 1, \dots, d)$, we indicate the one-dimensional marginal distributions by

$$F_n^{(i)}(x) = P\{X_{n,1}^{(i)} \leq x\}.$$

Let K be a closed, compact cone contained in $[-\infty, \infty]$ and, for some $\mathbf{a} \in [-\infty, \infty)$, set

$$\mathbb{E} = K \setminus \{\mathbf{a}\},$$

so that \mathbb{E} is a one-point uncompactification of K (see Resnick (2007, p. 170)). Our interest is in the cases in which

- $\mathbb{E} = [0, \infty] \setminus \{0\}$,
- $\mathbb{E} = [-\infty, \infty] \setminus \{-\infty\}$,
- $\mathbb{E} = [-\infty, \infty] \setminus \{0\}$.

Assume temporarily, for illustration, that $\mathbb{E} = [-\infty, \infty] \setminus \{-\infty\}$. Our basic assumption is that there exists a Radon measure ν on Borel subsets of \mathbb{E} such that

$$nF_n(\cdot) = nP\{X_{n,1} \in \cdot\} \xrightarrow{\nu} \nu(\cdot) \tag{2.1}$$

in $M_+(\mathbb{E})$. This entails

$$n\bar{F}_n^{(i)}(x) = nP\{X_{n,1}^{(i)} > x\} \xrightarrow{\nu} \nu^{(i)}(x, \infty], \quad i = 1, \dots, d, \tag{2.2}$$

in $M_+(-\infty, \infty]$, where

$$\nu^{(1)}(x, \infty] = \nu((x, \infty] \times [-\infty, \infty]^{d-1}).$$

Define the random vectors

$$\mathcal{P}_{n,j} = (\mathcal{P}_{n,j}^{(1)}, \dots, \mathcal{P}_{n,j}^{(d)}) = \left(\frac{1}{1 - F_n^{(i)}(X_{n,j}^{(i)})}, i = 1, \dots, d \right), \tag{2.3}$$

and note that $\mathcal{P}_{n,j}^{(i)}$ is standard Pareto distributed: for $i = 1, \dots, d$,

$$P\{\mathcal{P}_{n,j}^{(i)} > x\} = x^{-1}, \quad x \geq 1.$$

Definition 2.1. (*Pareto copula.*) Suppose that $X_{n,1}$ has distribution F_n with continuous marginals. Define $\mathcal{P}_{n,j}$ as in (2.3). Then we call the distribution ψ_n of $\mathcal{P}_{n,j}$ a *Pareto copula*.

A variant of (2.2) obtained by taking reciprocals is

$$\frac{1}{n(1 - F_n^{(i)}(x))} \rightarrow \frac{1}{\nu^{(i)}(x, \infty]}, \quad i = 1, \dots, d,$$

and inverting yields

$$\left(\frac{1}{1 - F_n^{(i)}}\right)^{\leftarrow}(ny) \rightarrow \left(\frac{1}{v^{(i)}(\cdot, \infty)}\right)^{\leftarrow}(y), \quad i = 1, \dots, d, y > 0. \tag{2.4}$$

To save writing, we define the nondecreasing functions

$$V^{(i)}(y) = \left(\frac{1}{v^{(i)}(\cdot, \infty)}\right)^{\leftarrow}(y), \quad i = 1, \dots, d, y > 0.$$

We summarize some properties of a Pareto copula; cf. de Haan and Resnick (1977), Resnick (1987, pp. 265, 277), or Resnick (2007, p. 204).

Proposition 2.1. *Let $X_{n,1}$ be a random vector with distribution F_n such that (2.1) holds. Let ψ_n be its Pareto copula. Then the following hold.*

(a) *There exists a Radon measure ψ_∞ on the Borel subsets of $[0, \infty] \setminus \{0\}$ such that*

$$n\psi_n(n\cdot) \xrightarrow{v} \psi_\infty(\cdot) \tag{2.5}$$

in $M_+([0, \infty] \setminus \{0\})$.

(b) *For $i = 1, \dots, d$,*

$$\psi_\infty^{(i)}(x, \infty] = \psi_\infty([0, \infty]^{i-1} \times (x, \infty] \times [0, \infty]^{d-i}) = x^{-1}, \quad x > 0.$$

(c) *The Radon measure ψ_∞ is a Lévy measure on \mathbb{R}_+^d .*

Proof. (a) From Lemma 6.1 of Resnick (2007, p. 174), it is enough to consider regions $[0, \mathbf{x}]^c$ for $\mathbf{x} \geq \mathbf{0}$. Then

$$\begin{aligned} n\psi_n([0, n\mathbf{x}]^c) &= n \mathbb{P}\{[\mathcal{P}_{n,1} \leq n\mathbf{x}]^c\} \\ &= n \mathbb{P}\left\{\left[X_{n,1}^{(i)} \leq \left(\frac{1}{1 - F_n^{(i)}}\right)^{\leftarrow}(nx^{(i)}); i = 1, \dots, d\right]^c\right\} \\ &= nF_n\left(\left[-\infty, \left(\frac{1}{1 - F_n^{(i)}}\right)^{\leftarrow}(nx^{(i)}); i = 1, \dots, d\right]^c\right), \end{aligned}$$

and from (2.1) and (2.4) this converges to

$$v([-\infty, (V^{(i)}(x^{(i)}); i = 1, \dots, d)]^c) =: \psi_\infty([0, \mathbf{x}]^c). \tag{2.6}$$

(b) This follows from $\mathcal{P}_{n,1}$ having Pareto marginal distributions.

(c) For simplicity suppose that $d = 2$. With $\|\mathbf{x}\| = |x^{(1)}| \vee |x^{(2)}|$ we have

$$\begin{aligned} \int_{\{\|\mathbf{x}\| \leq 1\}} \|\mathbf{x}\|^2 \psi_\infty(d\mathbf{x}) &= \iint_{0 \leq x^{(1)} < x^{(2)} \leq 1} (x^{(2)})^2 \psi_\infty(d\mathbf{x}) \\ &\quad + \iint_{0 \leq x^{(2)} \leq x^{(1)} \leq 1} (x^{(1)})^2 \psi_\infty(d\mathbf{x}) \\ &\leq 2 \int_0^1 s^2 s^{-2} ds \\ &= 2 \\ &< \infty. \end{aligned}$$

2.2. Point process interpretation

Continue to suppose for illustration that $\mathbb{E} = [-\infty, \infty] \setminus \{-\infty\}$. Condition (2.1) is equivalent to point process convergence (cf. Resnick (2007, pp. 179–180) or Resnick (1987)):

$$\sum_{j=1}^n \varepsilon_{X_{n,j}} \xrightarrow{D} \text{PRM}(\nu) =: \sum_k \varepsilon_{J_k} \tag{2.7}$$

in $M_p(\mathbb{E})$, or

$$\sum_{j=1}^{\infty} \varepsilon_{(j/n, X_{n,j})} \xrightarrow{D} \text{PRM}(\text{Leb} \times \nu) = \sum_k \varepsilon_{(t_k, J_k)} \tag{2.8}$$

in $M_p([0, \infty) \times \mathbb{E})$, where Leb stands for the Lebesgue measure, and recall that $\text{PRM}(\nu)$ denotes the Poisson random measure with mean measure ν . Similarly, (2.5) is equivalent to

$$\sum_{j=1}^n \varepsilon_{\mathcal{P}_{n,j}/n} \xrightarrow{D} \text{PRM}(\psi_{\infty}) = \sum_k \varepsilon_{j_k} \tag{2.9}$$

in $M_+([0, \infty] \setminus \{0\})$, or

$$\sum_{j=1}^{\infty} \varepsilon_{(j/n, \mathcal{P}_{n,j}/n)} \xrightarrow{D} \text{PRM}(\text{Leb} \times \psi_{\infty}) = \sum_k \varepsilon_{(t_k, j_k)} \tag{2.10}$$

in $M_+([0, \infty) \times [0, \infty] \setminus \{0\})$. From (2.3), (2.4), (2.7), and (2.9), we obtain the following result, which also explains the transformation of the points j_k to J_k .

Proposition 2.2. *When $\mathbb{E} = [-\infty, \infty] \setminus \{0\}$ and (2.1) holds,*

$$\begin{aligned} \sum_{j=1}^n \varepsilon_{X_{n,j}} &\stackrel{D}{=} \sum_{j=1}^n \varepsilon_{((1/(1-F_n^{(i)})) \leftarrow (n\mathcal{P}_{n,j}^{(i)}/n); i=1, \dots, d)} \\ &\xrightarrow{D} \sum_k \varepsilon_{(V^{(i)}(j_k^{(i)}); i=1, \dots, d)} \\ &\stackrel{D}{=} \sum_k \varepsilon_{J_k} \\ &= \text{PRM}(\nu). \end{aligned}$$

An analogous result holds when a time component is included.

2.3. Partial sum convergence

As usual, we denote by $D([0, \infty), \mathbb{R}^d)$ the space of \mathbb{R}^d -valued càdlàg functions on $[0, \infty)$, i.e. those which are right continuous and have left limits. Since $\mathcal{P}_{n,1}^{(i)}$ has a standard Pareto distribution for $i = 1, \dots, d$, it follows that

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} n \mathbb{E} \left(\left(\frac{\mathcal{P}_{n,1}^{(i)}}{n} \right)^2 1_{\lfloor \mathcal{P}_{n,1}^{(i)} \rfloor \leq n\varepsilon} \right) = 0,$$

where $1_{[\cdot]}$ is the indicator function. Thus, by a standard result reviewed in Resnick (2007, p. 214), we obtain from (2.9) or (2.10) the following result.

Proposition 2.3. *Let $\{\mathcal{P}_{n,j}; n \geq 1, j \geq 1\}$ be a triangular array of random vectors with standard Pareto marginals, in which the rows are i.i.d. Then*

$$\sum_{j \leq nt} \left(\frac{\mathcal{P}_{n,j}}{n} - \mathbb{E} \left(\frac{\mathcal{P}_{n,1}}{n} 1_{\|\mathcal{P}_{n,1}\|/n \leq 1} \right) \right) \xrightarrow{D} X_1(t) \tag{2.11}$$

in $D([0, \infty), \mathbb{R}^d)$, where $X_1(\cdot)$ is a Lévy process with Lévy measure ψ_∞ .

Definition 2.2. (*Pareto Lévy copula.*) Let $\{X_1(t), t \geq 0\}$ be the limit process in (2.11). Then we call its Lévy measure ψ_∞ a *Pareto Lévy copula*.

Be aware that others have attached meaning to the phrase *Lévy copula* to indicate Lebesgue marginals. See Barndorff-Nielsen and Lindner (2006), Böcker and Klüppelberg (2006), Bregman and Klüppelberg (2005), Cont and Tankov (2004), and Kallsen and Tankov (2006). Our Pareto Lévy copula was also considered in Barndorff-Nielsen and Lindner (2006).

Remark 2.1. Marginally, for $i = 1, \dots, d$, $\{X_1^{(i)}(t), t \geq 0\}$ is a 1-stable process with only positive jumps. However, the multivariate process $\{X_1(t), t \geq 0\}$ is not stable unless ψ_∞ has the homogeneity property $\psi_\infty(t \cdot) = t^{-1} \psi_\infty(\cdot)$.

Now suppose that (2.1) holds with $\mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$. We restrict attention to the first quadrant for the convenience of having only one multivariate tail specifying probabilities near ∞ . The full case of partial sum convergence for vectors in \mathbb{R}^d and associated transformations to Pareto copulas can be considered in $[-\infty, \infty] \setminus \{\mathbf{0}\}$, but we would have to specify 2^d quadrants corresponding to the neighborhoods of the 2^d vertices of $[-\infty, \infty]$, which could be labeled $\{\mathbf{a} \cdot \infty: \mathbf{a} \in \{-1, 1\}^d\}$. (See the comments in Section 6.5.5 of Resnick (2007, p. 201).) The following is a consequence of Section 7.2.1 of Resnick (2007, p. 214).

Proposition 2.4. *With $X_{n,j} \geq \mathbf{0}$ and $\mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$, suppose that (2.1) holds in $M_+(\mathbb{E})$ and also that*

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} n \mathbb{E}(X_{n,1}^{(i)})^2 1_{\|X_{n,1}^{(i)}\| \leq \varepsilon} = 0.$$

Then

$$\sum_{j \leq nt} (X_{n,j} - \mathbb{E}(X_{n,1} 1_{\|X_{n,1}\| \leq 1})) \xrightarrow{D} X_2(t) \tag{2.12}$$

in $D([0, \infty), \mathbb{R}^d)$, where $X_2(\cdot)$ is a Lévy process with Lévy measure ν .

The following result links the processes $X_1(\cdot)$ and $X_2(\cdot)$. It is a consequence of Proposition 2.2.

Theorem 2.1. *If the Itô representation of $X_1(\cdot)$ in (2.11) is*

$$X_1(t) = \sum_{t_k \leq t} j_k 1_{\|j_k\| > 1} + \lim_{\varepsilon \downarrow 0} \left[\sum_{t_k \leq t} j_k 1_{\{\varepsilon < \|j_k\| \leq 1\}} - t \int_{\{\varepsilon < \|x\| \leq 1\}} x \psi_\infty(dx) \right],$$

which is consistent with the notation used in (2.10), then the Itô representation for $X_2(\cdot)$ is given by

$$X_2(t) = \sum_{t_k \leq t} J_k 1_{\|J_k\| > 1} + \lim_{\varepsilon \downarrow 0} \left[\sum_{t_k \leq t} J_k 1_{\{\varepsilon < \|J_k\| \leq 1\}} - t \int_{\{\varepsilon < \|x\| \leq 1\}} x \nu(dx) \right],$$

where

$$\mathbf{J}_k = (V^{(1)}(j_k^{(1)}), \dots, V^{(d)}(j_k^{(d)})).$$

2.4. Extremes

Again for simplicity assume that $\mathbb{E} = [-\infty, \infty] \setminus \{-\infty\}$. From (2.10), it is immediate that (Resnick (1987))

$$Y_n(t) := \bigvee_{j/n \leq t} \mathcal{P}_{n,j} \xrightarrow{D} \bigvee_{t_k \leq t} \mathbf{j}_k =: Y(t)$$

in $D([0, \infty), \mathbb{R}^d)$, where Y is the multivariate extremal process associated with the limit in (2.10).

Proposition 2.5. *Set*

$$\mathbf{Z}_n(t) = \bigvee_{j/n \leq t} \mathbf{X}_{n,j},$$

and assume that (2.1) holds in $M_+([-\infty, \infty] \setminus \{-\infty\})$. Then

$$\begin{aligned} \mathbf{Z}_n(t) &\stackrel{D}{=} \left(\left(\frac{1}{1 - F_n^{(i)}} \right)^{\leftarrow} \left(\bigvee_{j \leq nt} \mathcal{P}_{n,j}^{(i)} \right); i = 1, \dots, d \right) \\ &\xrightarrow{D} (V^{(i)}(Y^{(i)}(t)); i = 1, \dots, d) \\ &=: \mathbf{Z}(t) \end{aligned}$$

in $D([0, \infty), \mathbb{R}^d)$, where \mathbf{Z} is the multivariate extremal process associated with the limit in (2.8).

3. Regular variation

Suppose that $\mathbf{X} \geq \mathbf{0}$ is a random vector in \mathbb{R}_+^d with distribution F and one-dimensional marginal distributions $F^{(i)}$, $i = 1, \dots, d$. Define

$$\mathcal{P} = \left(\frac{1}{1 - F^{(i)}(X^{(i)})}; i = 1, \dots, d \right). \tag{3.1}$$

Set

$$b_i(t) := \left(\frac{1}{1 - F^{(i)}} \right)^{\leftarrow} (t), \quad i = 1, \dots, d, \tag{3.2}$$

so

$$\mathbf{X} = (b_i(\mathcal{P}^{(i)}); i = 1, \dots, d).$$

Theorem 3.1. *Suppose that the distribution of \mathbf{X} is regularly varying (Resnick (2007, p. 204)); that is, for $i = 1, \dots, d$, as $t \rightarrow \infty$,*

$$t \mathbb{P} \left\{ \left(\frac{X^{(i)}}{b_i(t)}; i = 1, \dots, d \right) \in \cdot \right\} \xrightarrow{v} \nu(\cdot) \tag{3.3}$$

in $M_+([0, \infty] \setminus \{0\})$, where ν is a Radon measure. Assume that the marginal distributions $F^{(i)}$ have regularly varying tails $1 - F^{(i)} \in RV_{-\alpha^{(i)}}$ with $0 < \alpha^{(i)} < \infty$ for $i = 1, \dots, d$. Consider the Pareto copula ψ of \mathbf{X} , that is, the distribution of \mathcal{P} in (3.1). Then ψ is standard regularly varying,

$$t \psi(t \cdot) = t \mathbb{P} \left\{ \frac{\mathcal{P}}{t} \in \cdot \right\} \xrightarrow{v} \psi_\infty(\cdot) \tag{3.4}$$

in $M_+([\mathbf{0}, \infty] \setminus \{\mathbf{0}\})$, with

$$\psi_\infty(t \cdot) = t^{-1} \psi_\infty(\cdot),$$

and

$$v([\mathbf{0}, \mathbf{x}^{1/\alpha}]^c) = \psi_\infty([\mathbf{0}, \mathbf{x}]^c), \quad \mathbf{x} > \mathbf{0}.$$

Conversely, suppose that \mathbf{X} is a random vector in \mathbb{R}_+^d with Pareto copula ψ . If ψ is standard regularly varying, i.e. (3.4) holds and additionally we have marginal regular variation,

$$1 - F^{(i)} \in RV_{-\alpha^{(i)}}, \quad 0 < \alpha^{(i)} < \infty, \quad i = 1, \dots, d,$$

then \mathbf{X} is regularly varying and (3.3) holds.

Proof. The functions $b_i(\cdot) \in RV_{1/\alpha^{(i)}}$, where $\alpha^{(i)} > 0$, $i = 1, \dots, d$, and, for $\mathbf{x} > \mathbf{0}$,

$$\begin{aligned} t P\{\{\mathcal{P} \leq t\mathbf{x}\}^c\} &= t \psi([\mathbf{0}, t\mathbf{x}]^c) \\ &= t P\{\{X^{(i)} \leq b_i(tx^{(i)}); i = 1, \dots, d\}^c\} \\ &= t P\left\{\left\{\frac{X^{(i)}}{b_i(t)} \leq \frac{b_i(tx^{(i)})}{b_i(t)}; i = 1, \dots, d\right\}^c\right\} \\ &\rightarrow v([\mathbf{0}, ((x^{(i)})^{1/\alpha^{(i)}}); i = 1, \dots, d]^c) \\ &= v([\mathbf{0}, \mathbf{x}^{1/\alpha}]^c) \\ &= \psi_\infty([\mathbf{0}, \mathbf{x}]^c), \end{aligned} \tag{3.5}$$

by (2.6).

Suppose that $\{X, X_n, n \geq 1\}$ is i.i.d. with the regularly varying distribution F on \mathbb{R}_+^d . To link with the notation of Section 2, set

$$F_n(\mathbf{x}) = F(b_1(n)x^{(1)}, \dots, b_d(n)x^{(d)})$$

in $M_+([\mathbf{0}, \infty] \setminus \{\mathbf{0}\})$ and

$$F_n^{(i)}(x) = F^{(i)}(b_i(n)x), \quad i = 1, \dots, d.$$

In the notation of Section 2,

$$nF_n(\cdot) \xrightarrow{v} v(\cdot)$$

is equivalent to (3.3). Furthermore,

$$\mathcal{P}_{n,j}^{(i)} = \frac{1}{1 - F^{(i)}(b_i(n)X_j^{(i)}/b_i(n))} = \frac{1}{1 - F^{(i)}(X_j^{(i)})} = \mathcal{P}_j^{(i)}, \quad i = 1, \dots, d,$$

independent of n . The one-dimensional variables are standard Pareto distributed.

This allows us to rephrase Proposition 2.3 and Theorem 2.1 for the case of regular variation.

Corollary 3.1. *Suppose that $\{X_j, j \geq 1\}$ is i.i.d. on \mathbb{R}_+^d with equal continuous univariate marginal distributions $F^{(1)}$. Set*

$$\mathcal{P}_j^{(i)} = \left(\frac{1}{1 - F^{(1)}}\right)^\leftarrow(X_j^{(i)}), \quad b(t) = \left(\frac{1}{1 - F^{(1)}}\right)^\leftarrow(t).$$

The following statements are equivalent.

- (a) With $X_{n,j} = X_j/b(n)$, (2.12) holds, where $X_2(\cdot)$ is an α -stable Lévy motion ($0 < \alpha < 2$) with Lévy measure ν satisfying $\nu(t \cdot) = t^{-\alpha} \nu(\cdot)$.
- (b) X_1 has a multivariate regularly varying distribution on \mathbb{R}_+^d with index $\alpha \in (0, 2)$.
- (c) With $\bar{F}^{(1)} \in RV_{-\alpha}$, $0 < \alpha < 2$, and $\mathcal{P}_{n,j} = \mathcal{P}_j/n$, (2.11) holds, where $X_1(\cdot)$ is a 1-stable Lévy motion with Lévy measure ψ_∞ . The Pareto Lévy copula of $X_2(\cdot)$ in (a) is ψ_∞ .
- (d) $\bar{F}^{(1)} \in RV_{-\alpha}$, $0 < \alpha < 2$, and the Pareto copula of the random vector X_1 is standard regularly varying.

The equivalence of (a) and (b) was discussed in Resnick (2007, p. 214), and the equivalence of the rest follows from previous discussion.

3.1. Aggregation of risks

Assume that $\mathbb{E} = [0, \infty] \setminus \{0\}$. As before, suppose that $b_i(t)$, $i = 1, \dots, d$, is defined by (3.2). When the regular variation (3.3) holds, we obtain, as $t \rightarrow \infty$ (Resnick (2007, Section 7.3.1, p. 227)),

$$t \mathbb{P} \left\{ \sum_{i=1}^d \frac{X^{(i)}}{b_i(t)} > x \right\} \rightarrow \nu \left\{ \mathbf{x} \in \mathbb{E} : \sum_{i=1}^d x^{(i)} > x \right\}. \tag{3.6}$$

If in (3.3)

$$b_i(t) \sim b(t) \in RV_{1/\alpha}, \quad i = 1, \dots, d, \alpha > 0, t \rightarrow \infty,$$

then

$$\nu(t \cdot) = t^{-\alpha} \nu(\cdot), \tag{3.7}$$

and therefore from (3.6) we have

$$t \mathbb{P} \left\{ \sum_{i=1}^d X^{(i)} > b(t)y \right\} \rightarrow \nu \left\{ \mathbf{x} \in \mathbb{E} : \sum_{i=1}^d x^{(i)} > y \right\}.$$

By (3.7), this limit is

$$y^{-\alpha} \nu \left\{ \mathbf{x} \in \mathbb{E} : \sum_{i=1}^d x^{(i)} > 1 \right\}.$$

Thus,

$$t \mathbb{P} \left\{ \sum_{i=1}^d X^{(i)} > b(t) \right\} \rightarrow \nu \left\{ \mathbf{x} \in \mathbb{E} : \sum_{i=1}^d x^{(i)} > 1 \right\}$$

and

$$\frac{\mathbb{P} \{ \sum_{i=1}^d X^{(i)} > t \}}{\mathbb{P} \{ X^{(1)} > t \}} \rightarrow \nu \left\{ \mathbf{x} \in \mathbb{E} : \sum_{i=1}^d x^{(i)} > 1 \right\}, \tag{3.8}$$

since $\nu \{ \mathbf{x} \in \mathbb{E} : x^{(1)} > 1 \} = \lim_{t \rightarrow \infty} t \mathbb{P} \{ X^{(1)} > b(t) \} = 1$. The evaluation of the limit depends on the specific form of ν .

3.2. An interesting special case

An interesting case of the regular variation result given in the previous subsection was discussed from the copula point of view by Alink *et al.* (2004) and reviewed by Albrecher *et al.* (2006). Suppose that $d = 2$ and that $\mathbf{X} = (X^{(1)}, X^{(2)})$, where $X^{(1)} \stackrel{D}{=} X^{(2)}$, so $F^{(1)} = F^{(2)}$. Write \mathcal{P} in the following way:

$$U^{(i)} = F^{(1)}(X^{(i)}), \quad i = 1, 2, \quad \text{and} \quad \mathcal{P}^{(i)} = \frac{1}{1 - U^{(i)}}, \quad i = 1, 2.$$

Then, for $\mathbf{x} > \mathbf{0}$,

$$\begin{aligned} \psi([\mathbf{0}, \mathbf{x}]) &= \mathbb{P}\{\mathcal{P} \leq \mathbf{x}\} \\ &= \mathbb{P}\left\{\frac{1}{1 - U^{(i)}} \leq x^{(i)}; i = 1, 2\right\} \\ &= \mathbb{P}\{1 - U^{(i)} \geq (x^{(i)})^{-1}; i = 1, 2\} \\ &= 1 - \mathbb{P}\{[1 - U^{(i)} \geq (x^{(i)})^{-1}; i = 1, 2]^c\} \\ &= 1 - \mathbb{P}\{[1 - U^{(1)} \leq (x^{(1)})^{-1}] \cup [1 - U^{(2)} \leq (x^{(2)})^{-1}]\} \\ &= 1 - ((x^{(1)})^{-1} + (x^{(2)})^{-1} - \hat{C}((x^{(1)})^{-1}, (x^{(1)})^{-1})), \end{aligned}$$

where \hat{C} , often called the survival copula, is the copula

$$\hat{C}(x^{(1)}, x^{(2)}) = \mathbb{P}\{1 - U^{(i)} \leq x^{(i)}; i = 1, 2\}.$$

Thus, to summarize,

$$\psi([\mathbf{0}, \mathbf{x}]^c) = (x^{(1)})^{-1} + (x^{(2)})^{-1} - \hat{C}((x^{(1)})^{-1}, (x^{(2)})^{-1}). \quad (3.9)$$

Now suppose that the copula \hat{C} is Archimedean, so that

$$\hat{C}(x^{(1)}, x^{(2)}) = \hat{\phi}^{-1}(\hat{\phi}(x^{(1)}) + \hat{\phi}(x^{(2)})),$$

where $\hat{\phi}$ is the proper generator of the copula so that $\hat{\phi}$ is continuous, convex, and strictly decreasing from $[0, 1] \rightarrow [0, \infty]$ such that $\hat{\phi}(1) = 0$. Here $\hat{\phi}^{-1}$ is the inverse function of $\hat{\phi}$. (See Albrecher *et al.* (2006) and Alink *et al.* (2004).) Additionally, suppose that $\hat{\phi}$ is regularly varying at 0 with index $-\xi$ for $\xi > 0$. Then

$$R(t) := \hat{\phi}\left(\frac{1}{t}\right) \in RV_{\xi}$$

at ∞ and

$$R^{\leftarrow}(x) = \inf\left\{s : \hat{\phi}\left(\frac{1}{s}\right) \geq x\right\} = \frac{1}{\hat{\phi}^{-1}(x)} \in RV_{1/\xi}$$

at ∞ . Therefore, with this assumption we obtain, from (3.9),

$$n\psi([\mathbf{0}, n\mathbf{x}]^c) = (x^{(1)})^{-1} + (x^{(2)})^{-1} - n\hat{\phi}^{-1}(\hat{\phi}((nx^{(1)})^{-1}) + \hat{\phi}((nx^{(2)})^{-1})).$$

The last term is

$$\begin{aligned} \frac{n}{R^{\leftarrow}(R(nx^{(1)}) + R(nx^{(2)}))} &= \left(\frac{R^{\leftarrow}(R(n)(R(nx^{(1)})/R(n) + R(nx^{(2)})/R(n)))}{R^{\leftarrow}(R(n))} \right)^{-1} \\ &\sim \left(\lim_{n \rightarrow \infty} \left(\frac{R(nx^{(1)})}{R(n)} + \frac{R(nx^{(2)})}{R(n)} \right) \right)^{-1/\xi} \\ &= ((x^{(1)})^\xi + (x^{(2)})^\xi)^{-1/\xi}, \end{aligned}$$

where ‘ \sim ’ denotes asymptotic equivalence. Thus,

$$\lim_{n \rightarrow \infty} n\psi([\mathbf{0}, nx]^c) = (x^{(1)})^{-1} + (x^{(2)})^{-1} - ((x^{(1)})^\xi + (x^{(2)})^\xi)^{-1/\xi} = \psi_\infty([\mathbf{0}, \mathbf{x}]^c). \tag{3.10}$$

Note that in this model \mathcal{P} does *not* possess asymptotic independence (Resnick (2007, p. 192)), since

$$\lambda := \lim_{t \rightarrow \infty} P\{\mathcal{P}^{(2)} > t \mid \mathcal{P}^{(1)} > t\} = \lim_{t \rightarrow \infty} t P\{\mathcal{P} > t(1, 1)\} = \psi_\infty([\mathbf{1}, \infty]),$$

where $\mathbf{1} = (1, 1)$. Observe that

$$\psi_\infty((\mathbf{x}, \infty]) = (x^{(1)})^{-1} + (x^{(2)})^{-1} - \psi_\infty([\mathbf{0}, \mathbf{x}]^c),$$

and so from (3.10) we obtain

$$\lambda := \psi_\infty([\mathbf{1}, \infty]) = 1 + 1 - (1 + 1 - (1^\xi + 1^\xi)^{-1/\xi}) = 2^{-1/\xi}.$$

The measure ψ_∞ has a density $\psi'_\infty(u, v)$, which after differentiating $\psi_\infty([\mathbf{0}, (u, v)]^c)$ is seen to be

$$\begin{aligned} \psi'_\infty(u, v) &= (1 + \xi)(u^\xi + v^\xi)^{-1/\xi-2}(uv)^{\xi-1} \\ &= (1 + \xi)u^{-2-\xi} \left(1 + \left(\frac{v}{u} \right)^\xi \right)^{-1/\xi-2} v^{\xi-1} \end{aligned} \tag{3.11}$$

for $u > 0$ and $v > 0$. Calculating the limit in (3.8), we obtain

$$\begin{aligned} &\psi_\infty\{(u, v) \in \mathbb{E}: u + v > 1\} \\ &= \int_{u=1}^\infty \int_{v=0}^\infty \psi'_\infty(u, v) \, du \, dv + \int_{u=0}^1 \int_{v=1-u}^\infty \psi'_\infty(u, v) \, du \, dv \\ &= \psi_\infty((1, \infty] \times [0, \infty]) \\ &\quad + \int_{u=0}^1 (1 + \xi)u^{-2-\xi} \left(\int_{v=1-u}^\infty \left(1 + \left(\frac{v}{u} \right)^\xi \right)^{-1/\xi-2} v^{\xi-1} \, dv \right) \, du, \end{aligned}$$

and after some changes of variables, this reduces to

$$1 + \int_0^\infty (1 + v^\xi)^{-1/\xi-1} \, dv = 1 + 1 = 2,$$

since the integrand in the second term is a probability density (Alink *et al.* (2004, Lemma 2.4)). This is an interesting limit because although this model does not possess asymptotic independence, the limit in (3.8) is the one predicted by asymptotic independence.

Next, set $\alpha = (\alpha, \alpha)$, $\alpha > 0$, and following (3.5) we suppose that

$$v([\mathbf{0}, \mathbf{x}]^c) = \psi_\infty([\mathbf{0}, \mathbf{x}^\alpha]^c),$$

so that with $\mathbf{x} = (u, v)$ we have

$$v([\mathbf{0}, (u, v)]^c) = u^{-\alpha} + v^{-\alpha} - (u^{\alpha\xi} + v^{\alpha\xi})^{-1/\xi}.$$

Observe that

$$v((1, \infty] \times [0, \infty]) = 1.$$

Furthermore, v has a density $v'(u, v)$ given by

$$\begin{aligned} v'(u, v) &= \alpha^2(1 + \xi)(u^{\alpha\xi} + v^{\alpha\xi})^{-1/\xi-2}(uv)^{\alpha\xi-1} \\ &= \alpha^2(1 + \xi)u^{-\alpha(1+\xi)-1} \left(1 + \left(\frac{v}{u} \right)^{\alpha\xi} \right)^{-1/\xi-2} v^{\alpha\xi-1} \end{aligned}$$

for $u \geq 0$ and $v \geq 0$.

We may now compute the limit in (3.8) for this model. We have the limit

$$\begin{aligned} v\{\mathbf{x} \in \mathbb{E}: x^{(1)} + x^{(2)} > 1\} &= \int_{u=1}^\infty \int_{v=0}^\infty v'(u, v) \, du \, dv + \int_{u=0}^1 \int_{v=1-u}^\infty v'(u, v) \, dv \, du \\ &= 1 + \int_{u=0}^1 \alpha^2(1 + \xi)u^{-\alpha(1+\xi)-1} \left(\int_{v=1-u}^\infty \left(1 + \left(\frac{v}{u} \right)^{\alpha\xi} \right)^{-1/\xi-2} v^{\alpha\xi-1} \, dv \right) du, \end{aligned}$$

and after changes of variables, this reduces to

$$1 + \int_0^\infty \alpha(1 + v^{\alpha\xi})^{-1/\xi-1}(1 + v)^{\alpha-1} \, dv = 1 + \int_0^\infty (1 + s^\xi)^{-1/\xi-1}(1 + s^{-1/\alpha})^{\alpha-1} \, ds.$$

If Y_ξ has the probability density $(1 + s^{-1/\xi})^{\alpha-1}$, $s > 0$, this can be expressed as (Alink *et al.* (2004))

$$1 + E(1 + Y_\xi^{-1/\alpha})^{\alpha-1}.$$

Thus, for $d = 2$ with equal marginals, whenever $F \in RV_{-\alpha}$ for $\alpha > 0$, and ψ_∞ is given by (3.10), we have

$$\lim_{t \rightarrow \infty} \frac{P\{X^{(1)} + X^{(2)} > t\}}{P\{X^{(1)} > t\}} = 1 + E(1 + Y_\xi^{-1/\alpha})^{\alpha-1}.$$

4. The Pareto copula and distributions in the multivariate maximal domain of attraction

Suppose that $\{X, X_n, n \geq 1\}$ are i.i.d. random vectors with common distribution F . Then X or F is in a multivariate maximal domain of attraction if there exist

$$\mathbf{b}(t) = (b^{(1)}(t), \dots, b^{(d)}(t)) \in \mathbb{R}^d \quad \text{and} \quad \mathbf{a}(t) = (a^{(1)}(t), \dots, a^{(d)}(t)) \in \mathbb{R}_+^d$$

such that

$$\begin{aligned} P^n \left\{ \frac{X - \mathbf{b}(n)}{\mathbf{a}(n)} \leq \mathbf{x} \right\} &= F^n(\mathbf{a}(n)\mathbf{x} + \mathbf{b}(n)) \\ &= \left(P \left\{ \frac{X^{(i)} - b^{(i)}(n)}{a^{(i)}(n)} \leq x^{(i)}; i = 1, \dots, d \right\} \right)^n \\ &\rightarrow G(\mathbf{x}), \end{aligned} \tag{4.1}$$

where G is a nondegenerate distribution called a max-stable or extreme value distribution. The marginal distributions $G_{\gamma^{(i)}}^{(i)}$, $i = 1, \dots, d$, of G are one-dimensional extreme value distributions of the type

$$G_{\gamma^{(i)}}^{(i)}(x^{(i)}) = \exp(-(1 + \gamma^{(i)}x^{(i)})^{-1/\gamma^{(i)}}), \quad 1 + \gamma^{(i)}x^{(i)} > 0,$$

and $G_{\gamma^{(i)}}^{(i)}(x^{(i)})$ concentrates on $\{u \in \mathbb{R}: 1 + \gamma^{(i)}u > 0\}$. See, for example, de Haan and Ferreira (2006), Embrechts *et al.* (1997), and Resnick (1987).

In the notation of Section 2 we may write

$$F_n(\cdot) = P \left\{ \frac{X - \mathbf{b}(n)}{\mathbf{a}(n)} \in \cdot \right\} = F(\mathbf{a}(n)(\cdot) + \mathbf{b}(n)),$$

and then, after the customary logarithmic transformation, it is seen that (4.1) is equivalent to (2.1). Furthermore, matching the notation with that used in Section 2, we set

$$X_{n,j} = \frac{X_j - \mathbf{b}(n)}{\mathbf{a}(n)}.$$

The transformation given in (2.3) becomes

$$\begin{aligned} \mathcal{P}_{n,j}^{(i)} &= \frac{1}{1 - F_n^{(i)}(X_{n,j}^{(i)})} \\ &= \frac{1}{1 - F^{(i)}(\mathbf{a}^{(i)}(n)X_{n,j}^{(i)} + b^{(i)}(n))} \\ &= \frac{1}{1 - F^{(i)}(\mathbf{a}^{(i)}(n)((X_j^{(i)} - b^{(i)}(n))/\mathbf{a}^{(i)}(n)) + b^{(i)}(n))} \\ &= \frac{1}{1 - F^{(i)}(X_j^{(i)})}, \end{aligned}$$

independent of n .

As in Section 3, write, for $\mathbf{x} > \mathbf{0}$,

$$\psi([\mathbf{0}, \mathbf{x}]) = P\{\mathcal{P}_{n,1} \leq \mathbf{x}\} = P \left\{ \frac{1}{1 - F^{(i)}(X_1^{(i)})} \leq x^{(i)}; i = 1, \dots, d \right\}.$$

Then (4.1) is equivalent to ψ being standard regularly varying,

$$n\psi(n\cdot) \xrightarrow{v} \psi_\infty(\cdot)$$

in $M_+([\mathbf{0}, \infty] \setminus \{\mathbf{0}\})$ as n tends to ∞ with $\psi_\infty(t\cdot) = t^{-1}\psi_\infty(\cdot)$ for $t > 0$, and, for every $i = 1, \dots, d$, the random variable $X_1^{(i)}$ is in a one-dimensional maximal domain of attraction of a univariate extreme value distribution $G_{\gamma^{(i)}}$. See de Haan and Resnick (1977), Resnick (1987, Chapter 5), and de Haan and Ferreira (2006, Chapter 6).

4.1. Aggregation of risks when marginals are in the maximal domain of attraction of the Gumbel

We now discuss aggregation of risks when (4.1) holds with $\gamma^{(i)} = 0$ for $i = 1, \dots, d$, so that each marginal is in the domain of attraction of the Gumbel distribution. This is equivalent to supposing that, for $i = 1, \dots, d$, there exists a self-neglecting function $e^{(i)}(t)$ with derivative converging to 0 such that

$$\frac{\bar{F}^{(i)}(t + xe^{(i)}(t))}{\bar{F}^{(i)}(t)} \rightarrow e^{-x}, \quad x \in \mathbb{R}, \tag{4.2}$$

as t converges to the right endpoint of $F^{(i)}$ (see de Haan (1970), de Haan and Ferreira (2006), Embrechts *et al.* (1997), and Resnick (1987)). An acceptable choice of $e^{(i)}$ is the mean excess function (see Bingham *et al.* (1987), de Haan (1970), and Geluk and de Haan (1987)). Then we may take

$$b^{(i)}(t) = \left(\frac{1}{1 - F^{(i)}(\cdot)} \right)^{\leftarrow}(t) \quad \text{and} \quad a^{(i)}(t) = e^{(i)}(b^{(i)}(t)), \quad i = 1, \dots, d. \tag{4.3}$$

To obtain attractive formulae, it is necessary to assume that all marginals of F are the same, so we proceed under the assumption that

$$F^{(i)}(\cdot) = F^{(1)}(\cdot), \quad i = 1, \dots, d. \tag{4.4}$$

Formulae for aggregation of risks may be readily obtained when F does *not* possess asymptotic independence.

4.1.1. *Asymptotic independence is absent.* Special cases of this result have been given in Maulik *et al.* (2002, Proposition 3.1), Albrecher *et al.* (2006), and Alink *et al.* (2004). We assume that condition (4.4) of equal marginal distributions holds, and write $\mathbf{b}(t) = b(t)\mathbf{1}$ and $a^{(1)}(t) = a(t)$.

Set $\mathbb{E} = [-\infty, \infty] \setminus \{-\infty\}$. When the marginal distributions of F are in the maximal domain of attraction of a Gumbel distribution, (4.1) is equivalent to (see, for example, Resnick (2007, p. 138))

$$\sum_{j=1}^n \varepsilon_{a(n)^{-1}(X_j - b(n)\mathbf{1})} \xrightarrow{D} \varepsilon_{j_k} \tag{4.5}$$

in $M_p(\mathbb{E})$. Choose a large M . The restriction map $\mathbb{E} \rightarrow \mathbb{E}^M := (-M\mathbf{1}, \infty]$ is almost surely continuous, so from (4.5) we obtain the same convergence restricted to $M_p(\mathbb{E}^M)$. Define the addition map $T: \mathbb{E}^M := (-\infty, \infty] \rightarrow (-\infty, \infty]$ by

$$T\mathbf{x} = \sum_{i=1}^d x^{(i)}.$$

The map T is almost surely continuous from $\mathbb{E}^M \rightarrow (-\infty, \infty]$ and applying it to the restricted version of (4.5) we obtain

$$\begin{aligned} N_n^M &:= \sum_{j=1}^n 1_{[(X_j - b(n)\mathbf{1})/a(n)] \geq -M\mathbf{1}} \varepsilon_{(\sum_{i=1}^d X_j^{(i)} - db(n))/a(n)} \xrightarrow{D} N_\infty^M \\ &:= \sum_k 1_{[j_k > -M\mathbf{1}]} \varepsilon_{\sum_{i=1}^d j_k^{(i)}} \end{aligned} \tag{4.6}$$

Note that asymptotic independence would require all points of the limit Poisson process to be on the lines through $-\infty$, which would render the limit in (4.6) identically 0 and, hence, useless. However, asymptotic independence has been excluded.

We now proceed with a converging-together argument (cf. Resnick (2007, Theorem 3.5, p. 56) or Billingsley (1999)). Define

$$N_n := \sum_{j=1}^n \varepsilon_{(\sum_{i=1}^d X_j^{(i)} - db(n))/a(n)}.$$

We make two claims. First, we have, as $M \rightarrow \infty$,

$$N_\infty^M \xrightarrow{D} N_\infty := \sum_k 1_{[j_k > -\infty]} \varepsilon_{\sum_{i=1}^d j_k^{(i)}}, \tag{4.7}$$

considered as convergence in $M_p(-\infty, \infty]$. Second, we claim that if $d(\cdot, \cdot)$ is the vague metric on $M_p(-\infty, \infty]$ then, for any $\eta > 0$,

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} P\{d(N_n^M, N_n) > \eta\} = 0. \tag{4.8}$$

We are now in a position to state the following proposition.

Proposition 4.1. *Suppose that (4.1) holds, where all marginals of $F(\mathbf{x})$ are equal and all marginals of $G(\mathbf{x})$ are Gumbel, and that (4.2) and (4.4) hold. Suppose that F does not possess asymptotic independence, and define $v(\cdot)$ by*

$$v([-\infty, \mathbf{x}]^c) = -\log G(\mathbf{x}), \quad \mathbf{x} \neq -\infty.$$

Then

$$N_n := \sum_{j=1}^n \varepsilon_{(\sum_{i=1}^d X_j^{(i)} - db(n))/a(n)} \xrightarrow{D} \sum_k 1_{[j_k > -\infty]} \varepsilon_{\sum_{i=1}^d j_k^{(i)}}$$

in $M_p(-\infty, \infty]$, where the limit N_∞ is a Poisson random measure with mean measure

$$v\left\{\mathbf{x} \in (-\infty, \infty]: \sum_{i=1}^d x^{(i)} \in \cdot\right\}.$$

Therefore (Resnick (2007, p. 138)), as $n \rightarrow \infty$,

$$n P\left\{\frac{\sum_{i=1}^d X^{(i)} - db(n)}{a(n)} > y\right\} \rightarrow v\left\{\mathbf{x} \in (-\infty, \infty]: \sum_{i=1}^d x^{(i)} > y\right\}. \tag{4.9}$$

Corollary 4.1. *Under the conditions of Proposition 4.1, from (4.9) we have*

$$\lim_{t \rightarrow \infty} \frac{P\{\sum_{i=1}^d X^{(i)} > dt\}}{P\{X^{(1)} > t\}} = v\left\{\mathbf{x} \in (-\infty, \infty]: \sum_{i=1}^d x^{(i)} > 0\right\}. \tag{4.10}$$

To verify (4.10), set $y = 0$ in (4.9) and from (4.3) note that $P\{X^{(1)} > b(t)\} \sim t^{-1}$ as $t \rightarrow \infty$.

Proof of Proposition 4.1. The convergence in (4.7) is clear as it occurs almost surely. To prove (4.8), it suffices to take an arbitrary test function $g(\cdot)$ which is continuous with compact support in $(-\infty, \infty]$ and show that, for any $\eta > 0$,

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\{|N_n(g) - N_n^M(g)| > \eta\} = 0,$$

which reduces to showing that

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left\{\sum_{j=1}^n 1_{[a^{-1}(n)(\bigwedge_{i=1}^d X_j^{(i)} - b(n)) \leq -M]} \mathbb{E} a^{-1}(n)(\sum_{i=1}^d X_j^{(i)} - db(n))(g) > \eta\right\} = 0. \tag{4.11}$$

Suppose that the compact support of g is contained in $[-K, \infty]$ for some fixed K . Then the probability on the left-hand side of (4.11) is bounded by

$$n \mathbb{P}\left\{a^{-1}(n)\left(\bigwedge_{i=1}^d X_1^{(i)} - b(n)\right) \leq -M, a^{-1}(n)\left(\sum_{l=1}^d X_1^{(l)} - db(n)\right) > -K\right\}.$$

We drop the subscript ‘1’ for typographical simplicity. For the minimum to be less than $-M$, at least one of the terms must be less than $-M$, so the previous probability yields the upper bound

$$\sum_{i=1}^d n \mathbb{P}\left\{a^{-1}(n)(X^{(i)} - b(n)) \leq -M, a^{-1}(n)\left(\sum_{l=1}^d X^{(l)} - db(n)\right) > -K\right\}.$$

For the sum in the line above to be big when $a^{-1}(n)(X^{(i)} - b(n))$ is small, we require the sum of the $d - 1$ other terms with $l \neq i$ to be big, which yields the next upper bound,

$$\sum_{i=1}^d n \mathbb{P}\left\{a^{-1}(n)(X^{(i)} - b(n)) \leq -M, a^{-1}(n)\left(\sum_{l \neq i} X^{(l)} - db(n)\right) > -K + M\right\},$$

and, for the sum of the $d - 1$ terms to be bigger than $-K + M$, at least one summand must be bigger than $(-K + M)/(d - 1)$. This leads to the upper bound

$$\sum_{i=1}^d \sum_{l \neq i} n \mathbb{P}\left\{a^{-1}(n)(X^{(i)} - b(n)) \leq -M, a^{-1}(n)(X^{(l)} - db(n)) > \frac{-K + M}{d - 1}\right\},$$

and as n tends to ∞ , this converges to

$$\begin{aligned} & \sum_{i=1}^d \sum_{l \neq i} \nu\left\{x \in [-\infty, \infty] \setminus \{-\infty\}: x^{(i)} \leq -M, x^{(l)} > \frac{-K + M}{d - 1}\right\} \\ & \leq \sum_{i=1}^d \sum_{l \neq i} \nu\left\{x \in [-\infty, \infty] \setminus \{-\infty\}: x^{(i)} \leq -1, x^{(l)} > \frac{-K + M}{d - 1}\right\}. \end{aligned}$$

As M tends to ∞ , this converges to 0 since all bivariate marginals of the limit distribution $G(x)$ in (4.1) being proper precludes the limit from being positive.

Remark 4.1. Note that the notion of copula played no role here. What was crucial to this argument was the absence of asymptotic independence.

4.2. Back to our interesting special case

Consider again the example in Subsection 3.2, where the standard $\psi_\infty(\cdot)$ is given in (3.10). Since G has Gumbel marginals, we have $v^{(i)}(x, \infty] = e^{-x}$, which makes $V^{(i)}(x) = \log x$ for $x > 0$. From the analogue of (2.6) with $-\infty$ replacing $\mathbf{0}$, we have

$$v([-\infty, (V^{(i)}(x^{(i)}; i = 1, 2]^c)) = \psi_\infty([\mathbf{0}, \mathbf{x}]^c),$$

and thus

$$v([-\infty, \mathbf{x}]^c) = \psi_\infty([\mathbf{0}, \mathbf{e}^{\mathbf{x}}]^c),$$

for $\mathbf{x} \neq -\infty$, where $\mathbf{e}^{\mathbf{x}} = (\exp(x^{(1)}), \exp(x^{(2)}))$. So v has a density, which we call $v'(u, v)$, and from (3.11),

$$v'(u, v) = \psi'_\infty(e^u, e^v)e^ue^v = (1 + \xi)(e^{\xi u} + e^{\xi v})^{-1/\xi-2}e^{\xi u}e^{\xi v} \quad \text{for } (u, v) \in \mathbb{R}^2.$$

For this example, the limit in (4.10) is

$$v\{(u, v) \in \mathbb{R}^2 : u + v > 0\} = \iint_{\{(u,v) \in \mathbb{R}^2 : u+v>0\}} v'(u, v) du dv,$$

which we may evaluate as follows. Write $s = e^{\xi u}$ and $t = e^{\xi v}$, and the integral becomes

$$\iint_{\{(s,t) \in \mathbb{R}_+^2, st>1\}} \frac{1 + \xi}{\xi^2} (s + t)^{-1/\xi-2} ds dt.$$

Writing the double integral as $\int_{s=0}^\infty \int_{t>1/s}$ and evaluating the inner integral, we obtain

$$\int_0^\infty \frac{1}{\xi} \left(s + \frac{1}{s}\right)^{-1/\xi-1} ds.$$

Factor out $1/s$ from $s + 1/s$ to obtain

$$\int_0^\infty \frac{1}{\xi} s^{1/\xi+1} (1 + s^2)^{-1/\xi-1} ds.$$

With the intent to convert this to a beta integral, we now substitute $y = 1/(1 + s^2) \in (0, 1)$ to obtain

$$\int_0^1 \frac{1}{2\xi} y^{1/2\xi-1} (1 - y)^{1/2\xi+1-1} dy = \frac{1}{2\xi} B\left(\frac{1}{2\xi}, \frac{1}{2\xi} + 1\right),$$

where $B(x, y) = \int_0^1 t^{x-1} (1 - t)^{y-1} dt$. Expressing this in terms of the gamma function yields

$$\frac{(1/2\xi)\Gamma(1/2\xi)\Gamma(1/2\xi + 1)}{\Gamma(1/\xi + 1)} = \frac{\Gamma(1/2\xi + 1)^2}{\Gamma(1/\xi + 1)}.$$

Let us summarize this example. Suppose that $d = 2$ and that F is in a maximal domain of attraction, as in (4.1), with the limit G having Gumbel marginals. Furthermore, suppose that ψ_∞ has the form given in (3.10). Then Corollary 4.1 gives

$$\lim_{t \rightarrow \infty} \frac{P\{X^{(1)} + X^{(2)} > 2t\}}{P\{X^{(1)} > t\}} = \frac{\Gamma(1/2\xi + 1)^2}{\Gamma(1/\xi + 1)}.$$

5. Concluding remarks

Religious copularians have unshakable faith in the value of transforming a multivariate distribution to its copula. For the skeptics who believe the Emperor wears no clothes (Mikosch (2006)), perhaps use of the Pareto copula convinces some of them that the Emperor at least wears socks.

Constructing Lévy measures by transforming to the case of Lebesgue marginals seems, to the authors, uncritical transference of the copula philosophy to the domain of Lévy processes, and we believe that our transformation of random vectors to those having Pareto marginals has a much stronger probabilistic interpretation.

Adding dependent random variables in the domain of attraction of the Gumbel distribution, as discussed in Proposition 4.1, produces a specific tail behavior when asymptotic independence is absent. When the random variables are independent, the result requires a generalization of the concept of subexponentiality. We are actively thinking about the case of asymptotic independence specifically ruled out by Proposition 4.1 and Corollary 4.1.

Acknowledgements

Sidney Resnick's research was partially supported by ARO contract W911NF-07-1-0078 and NSA grant H98230-06-1-0069 at Cornell University. Much of the work was accomplished during May–June 2007 when Sidney Resnick was visiting the Chair of Mathematical Statistics at Munich University of Technology, whose hospitality and support is gratefully acknowledged.

References

- ALBRECHER, H., ASMUSSEN, S. AND KORTSCHAK, D. (2006). Tail asymptotics for the sum of two heavy-tailed dependent risks. *Extremes* **9**, 107–130.
- ALINK, S., LÖWE, M. AND WÜTHRICH, M. V. (2004). Diversification of aggregate dependent risks. *Insurance Math. Econom.* **35**, 77–95.
- BARNDORFF-NIELSEN, O. AND LINDNER, A. (2006). Lévy copulas: dynamics and transforms of epsilon type. *Scand. J. Statist.* **34**, 298–316.
- BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
- BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press.
- BÖCKER, K. AND KLÜPPELBERG, C. (2006). Multivariate models for operational risk. Submitted.
- BREGMAN, Y. AND KLÜPPELBERG, C. (2005). Ruin estimation in multivariate models with Clayton dependence structure. *Scand. Actuarial J.* **2005**, 462–480.
- CONT, R. AND TANKOV, P. (2004). *Financial Modelling With Jump Processes*. Chapman & Hall/CRC, Boca Raton, FL.
- DE HAAN, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Mathematisch Centrum, Amsterdam.
- DE HAAN, L. AND FERREIRA, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York.
- DE HAAN, L. AND RESNICK, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrscheinlichkeitsth.* **40**, 317–337.
- EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extreme Events for Insurance and Finance*. Springer, Berlin.
- GELUK, J. L. AND DE HAAN, L. (1987). *Regular Variation, Extensions and Tauberian Theorems (CWI Tract. 40)*. CWI, Amsterdam.
- KALLSEN, J. AND TANKOV, P. (2006). Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *J. Multivariate Anal.* **97**, 1551–1572.
- MAULIK, K., RESNICK, S. I. AND ROOTZÉN, H. (2002). Asymptotic independence and a network traffic model. *J. Appl. Prob.* **39**, 671–699.
- MIKOSCH, T. (2005). How to model multivariate extremes if one must? *Statist. Neerlandica* **59**, 324–338.
- MIKOSCH, T. (2006). Copulas: Tales and facts. *Extremes* **9**, 3–20.
- RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- RESNICK, S. I. (2007). *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.