Journal of Multivariate Analysis 166 (2018) 50–61

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Extreme-value copulas associated with the expected scaled maximum of independent random variables

Jan-Frederik Mai

XAIA Investment, Sonnenstr. 19, 80331 München, Germany

ARTICLE INFO

Article history:
Received 4 December 2017

AMS subject classifications:
62G32
65C10
60E07

Keywords:
Extreme-value copula
De Finetti's theorem
Lévy measure
Simulation
Stable tail dependence function

ABSTRACT

It is well-known that the expected scaled maximum of non-negative random variables with unit mean defines a stable tail dependence function associated with some extreme-value copula. In the special case when these random variables are independent and identically distributed, min-stable multivariate exponential random vectors with the associated survival extreme-value copulas are shown to arise as finite-dimensional margins of an infinite exchangeable sequence in the sense of De Finetti's Theorem. The associated latent factor is a stochastic process which is strongly infinitely divisible with respect to time, which induces a bijection from the set of distribution functions F of non-negative random variables with finite mean to the set of Lévy measures ν on $(0, \infty]$. Since the Gumbel and the Galambos copula are the most popular examples of this construction, the investigation of this bijection contributes to a further understanding of their well-known analytical similarities. Furthermore, a simulation algorithm based on the latent factor representation is developed, if the support of F is bounded. Especially in large dimensions, this algorithm is efficient because it makes use of the De Finetti structure.

1. Introduction

A d-dimensional copula C is a distribution function on $[0, 1]^d$ with all one-dimensional margins being uniformly distributed on $[0, 1]$. The importance of copulas in multivariate statistics stems from Sklar's Theorem, see [36], which states that for arbitrary one-dimensional distribution functions $G_1, \ldots, G_d$ the function $C\{G_1(t_1), \ldots, G_d(t_d)\}$ (resp. $C\{1 - G_1(t_1), \ldots, 1 - G_d(t_d)\}$) defines a multivariate distribution function (resp. survival function) with the pre-defined one-dimensional margins $G_1, \ldots, G_d$. A copula C is of *extreme-value kind* if it satisfies

$$\forall_{t \in (0,\infty)} \, \forall_{u_1,\ldots,u_d \in [0,1]} \quad \{C(u_1, \ldots, u_d)\}^t = C(u_1^t, \ldots, u_d^t). \tag{1}$$

This analytical property is usually interpreted in one of the following two ways.

On one hand, a random vector $\mathbf{Y} = (Y_1, \ldots, Y_d)$ with survival function defined, for all $t_1, \ldots, t_d \in [0, \infty)$, by

$$\Pr(Y_1 > t_1, \ldots, Y_d > t_d) = C(e^{-\lambda_1 t_1}, \ldots, e^{-\lambda_d t_d})$$

for $\lambda_1, \ldots, \lambda_d \in (0, \infty)$ has a *min-stable multivariate exponential distribution*, which means that the scaled minimum $\min(t_1 X_1, \ldots, t_d X_d)$ is exponentially distributed for all $t_1, \ldots, t_d \in (0, \infty)$; see [12]. If one wishes to focus on the dependence structure, it is convenient to normalize the margins to $\lambda_1 = \cdots = \lambda_d = 1$, which we do henceforth.

E-mail address: jan-frederik.mai@xaia.com.

https://doi.org/10.1016/j.jmva.2018.02.005
0047-259X/

On the other hand, a random vector $\mathbf{Z} = (Z_1, \ldots, Z_d)$ with distribution function

$$\Pr(Z_1 \le t_1, \ldots, Z_d \le t_d) = C\{G_1(t_1), \ldots, G_d(t_d)\}, \tag{2}$$

for univariate extreme-value distribution functions $G_1, \ldots, G_d$, has a *multivariate extreme-value distribution*, meaning that it arises as the limit of appropriately normalized componentwise maxima of independent and identically distributed random vectors. If one wishes to focus on the dependence structure, it is convenient to normalize the margins to $G_1(t) = \cdots = G_d(t) = e^{-1/t}$ for all $t \in [0, \infty)$, which we do henceforth. In particular, the distributional relation between **Y** and **Z** after their respective margin normalizations becomes

$$\mathbf{Y} \stackrel{d}{=} 1/\mathbf{Z},$$

with "$\stackrel{d}{=}$" denoting equality in distribution.

For background on extreme-value copulas, the interested reader is referred to [18], and to [27] for general background on copulas. Due to the defining property (1) of an extreme-value copula, its so-called *stable tail dependence function,* defined, for all $t_1, \ldots, t_d \in [0, \infty)$, by

$$\ell(t_1, \ldots, t_d) = -\ln\{C(e^{-t_1}, \ldots, e^{-t_d})\} \tag{3}$$

is homogeneous of order 1, i.e., $t \times \ell(t_1, \ldots, t_d) = \ell(t \times t_1, \ldots, t \times t_d)$ for all $t \in [0, \infty)$. This property gives rise to a canonical integral representation for the stable tail dependence function, see [9,31], given by

$$\ell(t_1, \ldots, t_d) = d\,\mathrm{E}\{\max(t_1 Q_1, \ldots, t_d Q_d)\}, \tag{4}$$

where the random vector $\mathbf{Q} = (Q_1, \ldots, Q_d)$ takes values on the unit simplex $S_d \equiv \{\mathbf{q} = (q_1, \ldots, q_d) \in [0, 1]^d : q_1 + \cdots + q_d = 1\}$, and each component has mean $1/d$. The finite measure $d\Pr(\mathbf{Q} \in \mathrm{d}\mathbf{q})$ on S_d is called the *Pickands dependence measure* associated with C, a nomenclature which dates back to [28].

While the Pickands dependence measure stands in unique correspondence with an extreme-value copula, this does not mean that the stable tail dependence function cannot have an alternative stochastic representation. In particular, if $X_1, \ldots, X_d$ are arbitrary non-negative random variables with unit mean, Segers [35] showed that setting, for all $t_1, \ldots, t_d \in [0, \infty)$,

$$\ell(t_1, \ldots, t_d) \equiv \mathrm{E}\{\max(t_1 X_1, \ldots, t_d X_d)\},$$

defines a proper stable tail dependence function of some extreme-value copula, which yields a useful construction device for parametric models. In the present article, we study the associated extreme-value copulas in the special case when $X_1, \ldots, X_d$ are independent. Denoting their distribution functions by $\mathbf{F} = (F_1, \ldots, F_d)$, we denote, for all $t_1, \ldots, t_d \in [0, \infty)$,

$$\ell_{\mathbf{F}}(t_1, \ldots, t_d) \equiv \mathrm{E}\{\max(t_1 X_1, \ldots, t_d X_d)\}, \tag{5}$$

and the extreme-value copula associated with $\ell_{\mathbf{F}}$ via (3) is denoted by $C_{\mathbf{F}}$.

The main contribution of the present article is a detailed study of the De Finetti structure of $C_{\mathbf{F}}$ in the special case when $F_1 = \cdots = F_d = F$. The computations in [10] point out that the two most prominent representatives in this family of extreme-value copulas are the Gumbel copula (F is a certain Fréchet distribution) and the Galambos copula (F is a certain Weibull distribution). The Gumbel copula is named after Emil Gumbel [19,20], whereas the Galambos copula is named after János Galambos [14]. Moreover, the recent articles [2,15] point out some further striking similarities between the Gumbel and the Galambos extreme-value copulas.

The remainder of the article is organized as follows. Section 2 considers the case when $F_1 = \cdots = F_d = F$, in which case we also write $\ell_{\mathbf{F}} = \ell_F$ and $C_{\mathbf{F}} = C_F$. An infinite exchangeable sequence $(Y_k)_{k\in\mathbb{N}}$ of random variables is constructed such that for each integer $d \in \mathbb{N}$, the random vector $(Y_1, \ldots, Y_d)$ has a min-stable multivariate exponential distribution with associated stable tail dependence function ℓ_F. It follows that the conditional cumulative hazard process $H_t \equiv -\ln\{\Pr(Y_1 > t \,|\, \mathcal{H})\}$ is strongly infinitely divisible with respect to time in the sense of [24], where $\mathcal{H}$ denotes the tail-σ-field of $(Y_k)_{k\in\mathbb{N}}$ in the sense of De Finetti's Theorem; see [1,7,8]. The relation between the associated Lévy measure ν_F on $(0, \infty]$ and the distribution function F is explored.

Section 3 enhances the stochastic model to allow for the non-exchangeable case of arbitrary $F_1, \ldots, F_d$. In particular, the De Finetti construction of the preceding section is slightly enhanced to derive a similar stochastic model for a min-stable multivariate exponential random vector $(Y_1, \ldots, Y_d)$ with stable tail dependence function $\ell_{\mathbf{F}}$. It is based on d latent frailty processes $(H_t^{(1)})_{t\ge 0}, \ldots, (H_t^{(d)})_{t\ge 0}$ which are dependent. Simulation algorithms for the new family are discussed. If the supports of $F_1, \ldots, F_d$ are all bounded, the aforementioned frailty model can be used for exact simulation. The latent frailty processes on which this simulation algorithm is based, resemble shot-noise processes in this case. In the general case of possibly unbounded supports of $F_1, \ldots, F_d$, an exact simulation strategy of [10], based on the Pickands dependence measure, can be applied. In particular, the simulation of **Q** in Eq. (4) is straightforward for the family of extreme-value copulas $C_{\mathbf{F}}$. Section 4 concludes.