ON THE COPULA FOR THE LIMITING DISTRIBUTION OF THE K LARGEST ORDER STATISTICS OF IID SAMPLES

MARCO AURÉLIO SANFINS AND GLAUCO VALLE

Abstract. We show that all multivariate Extreme Value distributions, which are the possible weak limits of the K largest order statistics of iid sequences, have the same copula, the so called K-extremal copula. This copula is described through exact expressions for its density and distribution functions. We also study measures of dependence, we obtain a weak convergence result and we propose a simulation algorithm for the K-extremal copula.

1. Introduction

In the study of extremes of iid sequences a question of interest is whether or not the dependence relation among the largest order statistics relies on the parent distribution function of the sequence. One way to evaluate nonlinear dependence between random variables is through the copula associated to them, this is already discussed in several books as the ones by Joe [7], Nelsen [10] and Drouet-Mari and Kotz [5]. In the present paper, we show that every multivariate extreme value distribution, which are the possible weak limits of the K largest order statistics of iid sequences, have the same copula called the K-extremal copula. This generalize the result in [9] obtained for $K = 2$. From the Extremal Types Theorem, see below, extremal distributions are obtained from linear transformations of one of three basic distributions, therefore the nonlinear dependence relation among the largest order statistics depends only on one of the three basic types. By our result, the non-linear dependence is uniquely caracterized by the K-extremal copula. This is not remarkable since the copula for any group of order statistics of an iid sample with continous parent distribution do not denpend on this distribution. However, a proper caracterization of the K-extremal copula is relevant as their consequences are.

The K-extremal copula is described by its distribution and density functions through exact expressions. We show that the copula of the K largest order statistics of iid sequences with continuos parent distribution converges in distribution to the K-extremal copula. We also study the assymptotic behavior of Spearman's rho and Kendall's tau for the first and the K largest order statistics. As a last result, we propose a simulation algorithm to sample from the K-extremal copula.

In section 2 we will present and discuss the results in this paper postponing the proofs to section 3.

2000 *Mathematics Subject Classification.* primary 60G70. secondary 60E99.

Key words and phrases. Copula, order statistics, independent random variables, extreme values.

2. Statements

Fix an interger $K \geq 2$. For every $n \geq K$, let $M_{1,n}, \ldots , M_{K,n}$ be the K largest order statistics of an iid sample of size n with the parent distribution of the sample not depending on n. The Extremal Types Theorem, see sections 2.2 and 2.3 in [8] and section 4.2 in [6], states that if for some sequences of real numbers $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ the random variables $a_n M_{1,n} + b_n$ converge in distribution then the random vectors

$$(a_n M_{1,n} + b_n, ..., a_n M_{K,n} + b_n) \tag{2.1}$$

converge in distribution and the limit has marginal distribution and density functions given respectively by

$$G_m(z) = \begin{cases} \exp\{-\Lambda(z)\} \sum_{j=0}^{m-1} \frac{\Lambda(z)^j}{j!} & \text{, if } \xi\left(\frac{z-\mu}{\sigma}\right) > -1 \text{ for } \xi \neq 0 \text{ or } z \in \mathbb{R} \text{ for } \xi = 0 \\ 0 & \text{, if } z < \mu - \frac{\sigma}{\xi} \text{ for } \xi > 0 \\ 1 & \text{, if } z > \mu - \frac{\sigma}{\xi} \text{ for } \xi < 0. \end{cases} \tag{2.2}$$

and

$$g_m(z) = \begin{cases} \exp\{-\Lambda(z)\} \frac{\Lambda'(z)\Lambda(z)^{m-1}}{(m-1)!} & \text{, if } \xi\left(\frac{z-\mu}{\sigma}\right) > -1 \text{ for } \xi \neq 0 \text{ or } z \in \mathbb{R} \text{ for } \xi = 0 \\ 0 & \text{, otherwise,} \end{cases} \tag{2.3}$$

where

$$\Lambda(z) = \Lambda_{\xi,\mu,\sigma}(z) = \begin{cases} \left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} & \text{, if } \xi \neq 0 \\ \exp\left(-\frac{z-\mu}{\sigma}\right) & \text{, if } \xi = 0, \end{cases}$$

for some $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$. A distribution with distribution function as above is called a Generalized Extreme Value (GEV) distribution and are classified in types I, II and III according respectively to $\xi = 0$, $\xi > 0$ and $\xi < 0$. Note that the function Λ is strictly decreasing positive function and satisfies

$$\begin{aligned} \lim_{z \to -\infty} \Lambda(z) = +\infty \quad &\text{and} \quad \lim_{z \to \infty} \Lambda(z) = 0, \text{ if } \xi = 0 \\ \lim_{z \downarrow (\mu - \frac{\sigma}{\xi})} \Lambda(z) = +\infty \quad &\text{and} \quad \lim_{z \to \infty} \Lambda(z) = 0, \text{ if } \xi > 0 \\ \lim_{z \to -\infty} \Lambda(z) = +\infty \quad &\text{and} \quad \lim_{z \uparrow (\mu - \frac{\sigma}{\xi})} \Lambda(z) = 0, \text{ if } \xi < 0. \end{aligned} \tag{2.4}$$

Futhermore, The joint density function $\tilde{g}_K$ of a limiting extreme value distribution for normalized sums of the K largest order statistics of an iid continuous random variables is given by

$$\tilde{g}_K(z_1, ..., z_K) = \begin{cases} (-1)^K \exp\{-\Lambda(z_K)\} \prod_{j=1}^{K} \Lambda'(z_j) & \text{, if } (z_1, ..., z_K) \in \Omega_\xi \\ 0 & \text{, otherwise.} \end{cases} \tag{2.5}$$

where

$$\Omega_\xi = \begin{cases} \mathbb{R}^K & \text{, if } \xi = 0 \\ \{(z_1, ..., z_K) \in \mathbb{R}^K : z_1 > ... > z_K > \mu - \frac{\sigma}{\xi}\} & \text{, if } \xi > 0 \\ \{(z_1, ..., z_K) \in \mathbb{R}^K : \mu - \frac{\sigma}{\xi} > z_1 > ... > z_K\} & \text{, if } \xi < 0. \end{cases}$$

A distribution with density as in (2.5) for parameters $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$ is called a Multivariate Generalized Extreme Value (MGEV) distribution.

Remark 2.1. *A broader class of stationary sequences of random variables have a MGEV distribution as the assimptotic distribution of the largest maxima.*

Our first result gives an explicity expression for the distribution function associated to the density $\tilde{g}_K$.

Proposition 2.1. *The distribution function $\tilde{G}_K$ of a limiting extreme value distribution for normalized sums of the K largest order statistics of iid continuous random variables has the following representation*

$$\tilde{G}_K(z_1,...,z_K) = H_K(z_1, \min(z_1,z_2), \min(z_1,z_2,z_3), ..., \min(z_1,...,z_K)),$$

for every $(z_1,...,z_K) \in \mathbb{R}^K$, where

$$H_K(z_1,...,z_K) = \exp\{-\Lambda(z_K)\}\, J_K(\Lambda(z_1),...,\Lambda(z_K))$$

for $\min(z_1,...,z_K) > \mu - \frac{\sigma}{\xi}$, if $\xi > 0$, or for $\min(z_1,...,z_K) < \mu - \frac{\sigma}{\xi}$, if $\xi < 0$, or $(z_1,...,z_K) \in \mathbb{R}^K$, if $\xi = 0$, otherwise $H_K(z_1,...,z_K) = 0$. The function $J_K : \mathbb{R}_+^K \to \mathbb{R}_+$ is a polynomial in K variables which is defined by induction by putting $J_1 \equiv 1$ and

$$J_m(x_1,...,x_m) = \sum_{j=0}^{m-1} \frac{x_m^j}{j!} - \sum_{j=1}^{m-1} \frac{x_j^j}{j!} J_{m-j}(x_{j+1},...,x_m), \qquad \text{for } m \geq 1.$$

We can now compute the density of the copula associated to the density $\tilde{g}_K$ of a MGEV distribution, which we call the K-extremal copula and tuns out to not depend on the distribution parameters ξ, μ and σ.

Proposition 2.2. *The density of the copula of a MGEV distribution is given by*

$$c_K(u_1,...,u_K) = \left(\prod_{j=1}^{K-1} \frac{d\log\psi_j}{du_j}(u_j)\right) \frac{d\psi_K}{du_K}(u_K) \tag{2.6}$$

$$= \left(\prod_{j=1}^{K-1} (-1)^{j-1}\psi_j(u_j)\frac{(\log\psi_j(u_j))^{j-1}}{(j-1)!}\right)^{-1} \left(\frac{(-\log\psi_K(u_K))^{K-1}}{(K-1)!}\right)^{-1}, \tag{2.7}$$

for $(u_1,...,u_K) \in (0,1)^K$ such that $u_1 > \psi_2(u_2) > ... > \psi_K(u_K)$, where $\psi_m : (0,1) \to (0,1)$ is the increasing function that satisfies the following implicit equation

$$u = \psi_m(u) \sum_{j=0}^{m-1} (-1)^j \frac{(\log\psi_m(u))^j}{j!}. \tag{2.8}$$

Remark 2.2. *The function ψ_m which appears in the expression for the K-extremal copula can be explicitly computed from a MGEV distribution function as $\psi_m(u) = \exp\{-\Lambda(G_m^{-1}(u)\}$ for every $u \in (0,1)$ and $m \geq 1$.*

Also with the distribution function of the MGEV distribution it is straightforward to write the distribution function of the K-extremal copula which we present in the next result.

Proposition 2.3. *The copula of a MGEV is given by*

$$C_K(u_1,...,u_K) = \mathcal{H}_K(u_1, r_1(u_1,u_2), r_2(u_1,u_2,u_3), ..., r_{K-1}(u_1,...,u_K))\,.$$

for every $(u_1,...,u_K) \in [0,1]^K$, *where*

$$r_{m-1}(u_1,...,u_m) = \psi_m^{-1}(\psi_l(u_l)) = \psi_l(u_l) \sum_{j=0}^{m-1} (-1)^j \frac{(\log \psi_l(u_l))^j}{j!},$$

if $\psi_l(u_l) = \min(\psi_1(u_1),...,\psi_m(u_m))$ *and for every* $(u_1,...,u_K)$ *such that* $u_1 = \psi_1(u_1) \geq \psi_2(u_2) \geq ... \geq \psi_K(u_K)$

$$\begin{aligned} \mathcal{H}_K(u_1,...,u_K) &= \psi_K(u_K) J_K\left(-\log u_1, -\log \psi_2(u_2), ..., -\log \psi_K(u_K)\right), \\ &= u_K - \psi_K(u_K) \sum_{j=1}^{K-1} \frac{(-\log \psi_j(u_j))^j}{j!} J_{K-j}(-\log \psi_{j+1}(u_{j+1}), ..., -\log \psi_K(u_K)) \end{aligned}$$

with J_m *defined in the statement of Proposition 2.1.*

The next proposition is a convergence result for copulas that has the consequence that for continuous distributions the non-linear dependence structure of the K-largest order statistics of large iid samples is approximatedly captured by the K-extremal copula. By a simple generalization of Lemma 6 in [1], we have that the multivariate copula among the K largest order statistics of an iid sample do not depend on the continuous parent distribution of the sample. This copula will be denoted by $\tilde{C}_K^{(n)}$, where n denotes the size of the sample.

Proposition 2.4. *The copula* $\tilde{C}_K^{(n)}$ *converges in distribution to* C_K *as* $n \to \infty$.

From the K-extremal copula we can obtain the copula between the l largest and the m largest limiting order statistics for every choice of l and m, or between any two marginals of a MGEV distribution. Then we can use these bivariate copulas to obtain measures of dependence as the Spearman's rho and Kendall's tau. For a copula C, the Spearman's rho is defined by

$$12 \int_0^1 \int_0^1 C(u,v)dudv - 3 = 12 \int_0^1 \int_0^1 uvdC(u,v) - 3$$

and Kendall's tau by

$$4 \int_0^1 \int_0^1 C(u,v)dC(u,v) - 1\,.$$

We are going to study here the behavior of Spearman's rho and Kendall's tau for the first and the Kth marginals of the K-extremal copula in the limit as $K \to \infty$. We denote these measures respectively by ρ_K and τ_K, $K \geq 2$. Using the convergence result in proposition 2.4, this caracterizes the behavior of these measures for the first and the Kth largest order statistics of large samples with continuous parent distribution. We point out that $\rho_2 = 2/3$ and $\tau_2 = 1/2$, see [9]. For more on measures of dependence of order statistics see [1] and [11]. We have the following result.

Proposition 2.5. *Both sequences* (ρ_K) *and* (τ_K) *converges to zero as* $K \to \infty$.

We now describe a simulation algorithm to generate samples from the K-extremal copula. The method is based on a technique of conditional sampling to sample from multivariate copulas, see for instance Cherubini, U., Luciano, E., Vecchiato, W. (2004) [3]. Let x_i be an observation sampled from $U(0,1)$. We can resume the procedure with the following steps:

(i) Put $C_i(u_1, u_2, ..., u_m) = C(u_1, u_2, ..., u_m, 1, ..., 1)$ for $m = 2, ..., K$;
(ii) Sample u_1 from the uniform distribution in $(0,1)$;
(iii) Sample u_m from the conditional distribution $C_m(\cdot|u_1, ..., u_{m-1})$ for $m = 1, ..., K$;

We now are going to focus on how to sample u_k from the conditional distribution $C_k(\cdot|u_1, ..., u_{k-1})$. To sample u_m from $C_m(.|u_1, ..., u_{m-1})$, we sample q from $U(0,1)$ and we put $u_m = C_m^{-1}(q|u_1, ..., u_{m-1})$. Therefore we should know explicitly $C_m(\cdot|u_1, ..., u_{m-1})$. We compute it in the following lemma:

Lemma 2.6. *The condicional distribution function of $U_m|(U_1, U_2, ..., U_{m-1})$ when $(U_1, ..., U_K)$ has distribution given by the K-extremal copula is given by*

$$C_m(u_m|u_1, ..., u_{m-1}) = \frac{\psi_m(u_m)}{\psi_{m-1}(u_{m-1})}. \tag{2.9}$$

If we now put $q = C_m(u_m|u_1, ..., u_{m-1})$, we have that:

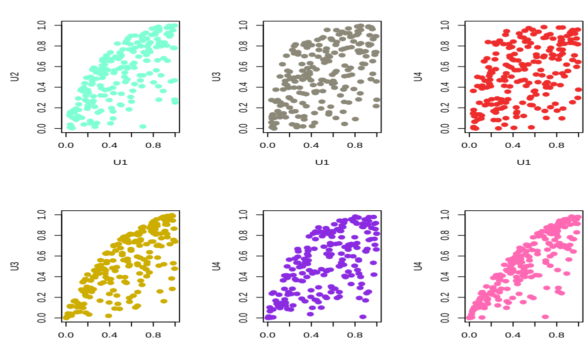
$$u_m = C_m^{-1}(q|u_1, ..., u_{m-1}) = \psi_m^{-1}(q.\psi_{m-1}(u_{m-1}))\,.$$

From definition 2.8 we get

$$u_m = \psi_m(q.\psi_{m-1}(u_{m-1})) \sum_{j=0}^{m-1} (-1)^j \frac{(\log \psi_m(q.\psi_{m-1}(u_{m-1})))^j}{j!}.$$

Therefore, we solve numerically $\psi_{m-1}(u_{m-1})$ and then $\psi_m(q.\psi_{m-1}(u_{m-1}))$ to obtain u_m.

We plot below a sample of size 200 from the 4-extremal copula.



3. Proofs

Proof of Proposition 2.1: We show that $\tilde{G}_K$ is a K-dimensional distribution function with density given by $\tilde{g}_K$. By the definition of $\tilde{g}_K$, the multiple integral

$$\int_{-\infty}^{z_1} ... \int_{-\infty}^{z_K} \tilde{g}_K(y_1, ...y_K)\ dy_1...dy_K$$

is equal to

$$\int_{-\infty}^{z_1} \int_{-\infty}^{\min(z_1,z_2)} ... \int_{-\infty}^{\min(z_1,...,z_K)} \tilde{g}_K(y_1, ...y_K)\ dy_1...dy_K.$$

Therefore $\tilde{G}_K(z_1, ..., z_K) = \tilde{G}_K(z_1, \min(z_1, z_2), ..., \min(z_1, ..., z_K))$ and we can suppose in the sequence that $z_1 > z_2 > ... > z_K$. Then

$$\tilde{G}_K(z_1, ..., z_K) = (-1)^K \int_{A_\xi}^{z_K} \int_{y_K}^{z_{K-1}} ... \int_{y_3}^{z_2} \int_{y_2}^{z_1} \exp\{-\Lambda(y_K)\} \prod_{j=1}^{K} \Lambda'(y_j)\ dy_1...dy_K\ .$$

Considering the following change of variables in the last integral, $x_j = \Lambda(y_j)$, for $1 \leq j \leq K$, we get the following integral

$$I_K(w_1, ..., w_K) := (-1)^K \int_{w_K}^{+\infty} \int_{w_{K-1}}^{x_K} ... \int_{w_2}^{x_3} \int_{w_1}^{x_2} e^{-x_K} dx_1...dx_K\ ,$$

where $w_j = \Lambda(z_j)$. We shall prove by induction that

$$I_K(w_1, ..., w_K) = e^{-w_K} J_K(w_1, ..., w_K)\ .$$

For $K = 1$, a simple verification shows that the result holds. Now suppose that it holds for $1 \leq K \leq L - 1$ then for $K = L$ we have that $I_K(w_1, ..., w_K)$ is equal to

$$(-1)^K \int_{w_K}^{+\infty} \int_{w_{K-1}}^{x_K} ... \int_{w_2}^{x_3} x_2 e^{-x_K} dx_2...dx_K - w_1 I_{K-1}(w_2, ..., w_K)$$

which is equal to

$$(-1)^K \int_{w_K}^{+\infty} \int_{w_{K-1}}^{x_K} ... \int_{w_3}^{x_4} \frac{x_3}{2} e^{-x_K} dx_3...dx_K - \frac{w_2}{2} I_(K-2)(w_3, ..., w_K) - w_1 I_{K-1}(w_1, ..., w_K)\ .$$

Following recursively this procedure we get

$$I_K(w_1, ..., w_K) = e^{-w_K} \sum_{j=0}^{m-1} \frac{w_K^j}{j!} - \sum_{j=1}^{m-1} \frac{w_j^j}{j!} I_(K-j)(w_{j+1}, ..., w_K).$$

By the definition of J_K and the induction hypotheses we complete the proof. □

Proof of Proposition 2.2: Let us fix a limiting extreme value distribution function $\tilde{G}_K$. We have that

$$c_K(u_1, ..., u_K) = \frac{\tilde{g}_K(G_1^{-1}(u_1), ..., G_K^{-1}(u_K))}{\prod_{j=1}^{K} g_j(G_j^{-1}(u_j))}\ .$$

Therefore we just apply formulas (2.3) and (2.5) to obtain that

$$c_K(u_1, ..., u_K) = \left(\prod_{j=1}^{K-1} \exp\{-\Lambda(G_j^{-1}(u_j))\} \frac{\Lambda(G_j^{-1}(u_j))^{j-1}}{(j-1)!} \right)^{-1} \left(\frac{\Lambda(G_K^{-1}(u_K))^{K-1}}{(K-1)!} \right)^{-1}.$$

From this formula, if we put $\psi_m(u) = \exp\{-\Lambda(G_m^{-1}(u)\}$ we get (2.7) in the statement. Now 2.8 is a direct consequence of the explicit formulas for the distribution functions G_m given in (2.2).

It remains to verify (2.6). If we derive both sides of (2.8), we get that

$$1 = \left(\sum_{j=0}^{m-1} (-1)^j \frac{(\log \psi_m)^j}{(j)!} - \sum_{j=0}^{m-2} (-1)^j \frac{(\log \psi_m)^j}{(j)!} \right) \frac{d\psi_m}{du} = (-1)^{m-1} \frac{(\log \psi_m)^{m-1}}{(m-1)!} \frac{d\psi_m}{du},$$

which implies that

$$\frac{d\psi_m}{du} = (-1)^{m-1} \left(\frac{(\log \psi_m)^{m-1}}{(m-1)!} \right)^{-1} \tag{3.1}$$

and

$$\frac{d \log \psi_m}{du} = (-1)^{m-1} \left(\psi_m \frac{(\log \psi_m)^{m-1}}{(m-1)!} \right)^{-1}. \tag{3.2}$$

From (3.1), (3.2) and (2.7) we arrive at (2.6). □

Proof of Proposition 2.3: Let us fix a limiting extreme value distribution function $\tilde{G}_K$. Then the distribution function of the K-extremal copula is given by

$$C_K(u_1, ..., u_K) = \tilde{G}_K(G_1^{-1}(u_1), ..., G_K^{-1}(u_K))$$

for every $(u_1, ..., u_K) \in [0,1]^K$ which by Proposition 2.1 is equal to

$$H_K(G_1^{-1}(u_1), \min(G_1^{-1}(u_1), G_2^{-1}(u_2)), ..., \min(G_1^{-1}(u_1), ..., G_K^{-1}(u_K))).$$

By the definition of H_K, monotonicity and the expression for ψ_m in remark 2.2, see also the proof of Proposition 2.2, the previous expression is given by

$$\min_{1\le l\le K}(\psi_l(u_l)) \; J_K \left(-\log u_1, -\log \min_{l=1,2}(\psi_l(u_l)), ..., -\log \min_{1\le l\le K}(\psi_l(u_l)) \right).$$

Using the definition of r_m in the statement, write the above expression as

$$\psi_K(r_K(u_1, ..., u_m)) \; J_K \left(-\log u_1, -\log \psi_2(r_2(u_1, u_2)), ..., -\log \psi_K(r_K(u_1, ..., u_m)) \right),$$

which completes the proof. □

Proof of Proposition 2.4: Let $M_{1,n}$, ... , $M_{K,n}$ be the K-largest order statistics of a sample of size n with a given continuous parent distribution function F which belongs to the domain of atraction of a GEV distribution. This means that there exists $(a_n)_{n=1}^{+\infty}$ and $(b_n)_{n=1}^{+\infty}$ sequences of real numbers such that the random vector

$$(a_n M_{1,n} + b_n, ..., a_n M_{K,n} + b_n)$$

converges in distribution to some $\tilde{G}_K$ which is MGEV distribution. By invariance concerning composition with affine transformations the copula associated to $(M_{1,n}, ..., M_{K,n})$ and $(a_n M_{1,n} + b_n, ..., a_n M_{K,n} + b_n)$ is $\tilde{C}_K^{(n)}$ independently of F.

Let $F_{j,n}$ be the distribution function of $a_n M_{j,n} + b_n$. Therefore, if we define the function $V_n(x_1, ..., x_K) = (F_{1,n}(x_1), ..., F_{K,n}(x_K))$, $(x_1, ..., x_K) \in \mathbb{R}^n$ then

$$V_n(a_n M_{1,n} + b_n, ..., a_n M_{K,n} + b_n) \tag{3.3}$$

has the distribution of the copula $\tilde{C}_K^{(n)}$.

The K-extremal copula has the distribution of $V(Y_1, ..., Y_K)$, where $V(x_1, ..., x_K) = (G_1(x_1), ..., G_K(x_K))$, $(x_1, ..., x_K) \in \mathbb{R}^n$. By Theorem 5.1 in [2], (3.3) converges in distribution to the K-extremal copula if V_n converges uniformly to V on compact intervals, but this is a consequence of Plyas's Theorem which implies that $F_{j,n}$ converges uniformly to G_j since the last is absolutely continuous. □

Proof of Proposition 2.5: We shall prove through estimates on exact expressions that $\rho_K \to 0$. The analogous result can be applied to τ_K since $\rho_K \geq \tau_K \geq 0$ which can be verified through Theorem 5.1 of Fredricks and Nelsen in [4], since for two order statistics the largest is always left-tail decreasing and right-tail increasing in the smallest.

Applying directly the definition we can write $(\rho_K + 3)/12$ as

$$\int_0^1 \int_{\psi_{K-1}^{-1}(\psi_K(u_K))}^1 \dots \int_{\psi_2^{-1}(\psi_3(u_3))}^1 \int_{\psi_2(u_2)}^1 u_1\, u_K\, c_K(u_1,...,u_K)\, du_1...du_K. \tag{3.4}$$

which we are going to show that converges to $1/4$ as $K \to \infty$ resulting in $\rho_K \to 0$. By (2.6) the previous iterated integral can be rewritten as

$$\int_0^1 \int_{\psi_{K-1}^{-1}(\psi_K(u_K))}^1 \dots \int_{\psi_2^{-1}(\psi_3(u_3))}^1 \int_{\psi_2(u_2)}^1 u_1\, u_K \left(\prod_{j=1}^{K-1} \frac{d\log\psi_j}{du_j}(u_j)\right) \frac{d\psi_K}{du_K}(u_K)\, du_1...du_K.$$

By induction in $1 \leq m \leq K-1$, we show that

$$\int_{\psi_m^{-1}(\psi_{m+1}(u_{m+1}))}^1 \dots \int_{\psi_2^{-1}(\psi_3(u_3))}^1 \int_{\psi_2(u_2)}^1 u_1 \prod_{j=1}^{m} \frac{d\log\psi_j}{du_j}(u_j)\, du_1...du_m.$$

is equal to

$$(-1)^m \left[\psi_{m+1}(u_{m+1}) - \sum_{j=0}^{m-1} \frac{(\log\psi_{m+1}(u_{m+1}))^j}{j!}\right]. \tag{3.5}$$

Indeed, ψ_1 is the identity in $(0,1)$ and therefore

$$\int_{\psi_2(u_2)}^1 u_1 \frac{d\log\psi_1}{du_1}(u_1)du_1 = (-1)[\psi_2(u_2) - 1].$$

Now suppose that (3.5) holds for some $1 \leq m = l \leq K-2$ then

$$(-1)^l \left[\psi_{l+1}(u_{l+1}) - \sum_{j=0}^{l-1} \frac{(\log\psi_{l+1}(u_{l+1}))^j}{j!}\right] \frac{\log\psi_{l+1}}{du_{l+1}}(u_{l+1}).$$

is equal to

$$(-1)^l \frac{d}{du_{l+1}} \left(\psi_{l+1}(u_{l+1}) - \sum_{j=1}^{l} \frac{(\log\psi_{l+1}(u_{l+1}))^j}{j!}\right)$$

and, since $\psi_{l+1}(1) = 1$, integrating on u_{l+1} over the interval $(\psi_{l+1}^{-1}(\psi_{l+2}(u_{l+2})), 1)$ we obtain that (3.4) holds for $m = l+1$.

Therefore the integral in (3.4) is equal to

$$\int_0^1 u\, \frac{d\psi_K}{du}(u)(-1)^{K-1} \left[\psi_K(u) - \sum_{j=0}^{K-2} \frac{(\log\psi_K(u))^j}{j!}\right] du.$$

Put $v = \psi_K(u)$, $u \in (0,1)$ and uses the power series expansion

$$v = \sum_{j=0}^{\infty} \frac{\log(v)^j}{j!}$$

to write the previous integral as

$$(-1)^{K-1} \int_0^1 \psi_K^{-1}(v) \left(\sum_{j=K-1}^{\infty} \frac{\log(v)^j}{j!} \right) dv\,.$$

Another change of variables and (2.8) allows us to write the integral in (3.4) as

$$(-1)^{K-1} \sum_{l=0}^{K-1} \sum_{j=K-1}^{\infty} \frac{(-1)^j}{j!l!} \int_0^{+\infty} y^{l+j} e^{-2y} dy = (-1)^{K-1} \sum_{l=0}^{K-1} \sum_{j=K-1}^{\infty} (-1)^j \binom{l+j}{l} \frac{1}{2^{l+j+1}}$$

since

$$\int_0^{+\infty} y^{l+j} e^{-2y} dy = \frac{(l+j)!}{2^{l+j+1}}$$

We finish the proof showing that

$$(-1)^{K-1} \sum_{l=0}^{K-1} \sum_{j=K-1}^{\infty} (-1)^j \binom{l+j}{l} \frac{1}{2^{l+j}} \to \frac{1}{2}\,.$$

The left hand side term in the previous convergence statement is equal to

$$\sum_{l=0}^{K-1} \sum_{j=\lfloor \frac{K-1}{2} \rfloor}^{\infty} \binom{l+2j}{l} \frac{1}{2^{l+2j}} \left(1 - \left(\frac{j+2l+1}{2l+1} \right) \frac{1}{2} \right)\,.$$

which for K large can be replaced by

$$\frac{1}{2} \sum_{l=0}^{K-1} \sum_{j=\lfloor \frac{K-1}{2} \rfloor}^{\infty} \binom{l+2j}{l} \frac{1}{2^{l+2j}}$$

Some combinatorial estimates allow us to show that

$$\sum_{l=0}^{K-1} \sum_{j=\lfloor \frac{K-1}{2} \rfloor}^{\infty} \binom{l+2j}{l} \frac{1}{2^{l+2j}} \to 1$$

as $k \to \infty$.

Proof of Lemma 2.6: Let $(U_1, U_2, ..., U_K)$ be a randon vector whose distribution function is C, then the conditional distribution of U_m given $U_1, U_2, ..., U_{m-1}$ has distribution function

$$\begin{aligned} C_m(u_m|u_1,...,u_{m-1}) &= Pr(U_m \le u_m | U_1 = u_1, ..., U_{m-1} = u_{m-1}) \\ &= \left(\frac{[\partial^{m-1} C_m(u_1,...,u_m)]/[\partial u_1,...,\partial u_{m-1}]}{\partial^{m-1} C_{m-1}(u_1,...,u_{m-1})]/[\partial u_1,...,\partial u_{m-1}]} \right) \end{aligned} \tag{3.6}$$

for every $m = 2, ..., k$.

We first deal with the numerator in (3.6) which by the formula in Proposition 2.3 can be written as

$$\frac{\partial^{m-1}\left[-\psi_m(u_m)\sum_{j=1}^{m-1}\frac{-log(\psi_j(u_j))^j}{j!}J_{m-j}(-\log\psi_{j+1}(u_{j+1}),...,-\log\psi_m(u_m))\right]}{\partial u_1...\partial u_{m-1}}$$

If we remove the terms that do not depend on the variables $u_1,...,u_{m-1}$ we obtain that the last partial derivative is equal to

$$\frac{\partial^{m-1}\left[-\psi_m(u_m)\prod_{j=1}^{m-1}(-log(\psi(u_j)))\right]}{\partial u_1...\partial u_{m-1}}. \tag{3.7}$$

Using that

$$\frac{d\log\psi_m}{du}=(-1)^{m-1}\left(\psi_m\frac{(\log\psi_m)^{m-1}}{(m-1)!}\right)^{-1},$$

we obtain that (3.7) is

$$=(-1)^m\psi_m(u_m)(-1)^{m-1}\prod_{j=1}^{m-1}(-1)^{j-1}\left(\psi_j(u_j)\frac{log(\psi_j(u_j))^{j-1}}{(j-1)!}\right)^{-1}. \tag{3.8}$$

Now we consider the denominator in (3.6) which is equal to the density function of the (m-1)-extremal copula. Hence it is equal to

$$\left(\prod_{j=1}^{m-2}(-1)^{j-1}\psi_j(u_j)\frac{(\log\psi_j(u_j))^{j-1})}{(j-1)!}\right)^{-1}\left(-\frac{(\log\psi_{m-1}(u_{m-1}))^{m-2}}{(m-2)!}\right)^{-1}. \tag{3.9}$$

Finally replace the expressions in (3.8) and (3.9) respectively in the numerator and denominator in (3.6) to obtain that

$$C_m(u_m|u_1,...,u_{m-1})=\frac{\psi_m(u_m)}{\psi_{m-1}(u_{m-1})}. \quad \square$$

References

[1] J. Averous, C. Genest, S. Kochar: On dependence structure of order statistics, Journal of Mult. Anal. 94, 159-171 (2005)

[2] P. Billingsley: *Convergence of Probability Measures*, John Wiley & Sons (1968)

[3] U. Cherubini, E. Luciano, W. Vecchiato: *Copula Methods in Finance.* Wiley Finance: Chichester. (2004)

[4] G. Fredricks, R. Nelsen: On the relation between Spearman's rho and Kendall's tau for pairs of continuous random variables, Journal of Stat. Plan. and Inf. 137, 2143-2150 (2007)

[5] D. Drouet-Mari, S. Kotz: *Correlation and dependence*, Imperial College Press, London (2001).

[6] P. Embrechts, C. Kluppelberg, T. Mikosch: *Modelling Extremal Events for Insurance and Finance*, Applications of Mathematics 33, Springer

[7] H. Joe: *Multivariate models and dependence concepts*, Chapman & Hall (1997)

[8] M. R. Leadbetter; G. Lindgreen; H. Rootzen: *Extremes and related properties of random sequences and processes*, Springer-Verlag New York (1983)

[9] B. Mendes, M. Sanfins: The limiting copula of the two largest order statistics of iid samples, *Brazilian Jounal of prob. and Statisc.* (2007)

[10] R. Nelsen: *An introduction to copulas*, Springer series in statistics, 2^o edition (2007).

[11] Y. Chen: A note on the relantionship between Spearman's ρ and Kendall's τ for extreme order statistics, *Journal of Stat. Plan. and Inf.* 137, 2165-2171 (2007)

UFRJ - Departamento de métodos estatísticos do Instituto de Matemática.
Caixa Postal 68530, 21945-970, Rio de Janeiro, Brasil
e-mail: glauco.valle@im.ufrj.br
e-mail: marcosanfins@dme.ufrj.com