Journal of Multivariate Analysis 160 (2017) 1–9

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Extremal properties of order statistic distributions for dependent samples with partially known multidimensional marginals

Andrzej Okolewski

Institute of Mathematics, Lodz University of Technology, Wolczanska 215, 90–924 Lodz, Poland

A R T I C L E   I N F O

Article history:
Received 2 February 2017
Available online 1 June 2017

AMS 2000 subject classifications:
62G30
52A40

Keywords:
Order statistics
Dependent observations
Copula
Diagonal section
Empirical distribution function
Coherent systems

A B S T R A C T

Let $\mathbb{X} = (X_1, \ldots, X_n)$ be an n-tuple of random variables where each X_j has the same known distribution function F and where there is a number $k \le n$ such that for each $i \in \{1, \ldots, k\}$, all i-tuples have copulas with the same known diagonal δ_i. A reliability system with such nonnegative component lifetimes $X_1, \ldots, X_n$ is a system with the property that for each $i \le k$, all of its structurally identical sub-systems of i components have the same known reliability function. We provide a characterization for empirical distributions from the X_j's, and apply it to derive two-sided bounds (depending on F and δ_i's) for arbitrary linear combinations of distribution functions of the associated order statistics as well as to establish necessary and sufficient conditions for uniform sharpness of these bounds. Moreover, for $k = 2$ and some classes of δ_2's, we determine stochastically extremal distributions of single order statistics.

1. Introduction

Let $X_1, \ldots, X_n$ be random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$ and let $X_{1:n} \le \cdots \le X_{n:n}$ denote the associated order statistics. The extremal properties of order statistic distributions under various assumptions about the marginals of the underlying sample have been the object of considerable attention. Mallows [9] and Lai and Robbins [8] investigated the behavior of sample maxima when the X_j's are identically distributed and no restrictions are imposed on the structure of their interdependence; these authors determined the stochastically extremal distributions of $X_{n:n}$ when the marginals are uniform and arbitrary, respectively. Rychlik [16,19] and Caraux and Gascuel [2] extended the result of Lai and Robbins to all order statistics. Uniformly sharp bounds for linear combinations of distribution functions of arbitrarily dependent identically distributed observations were provided by Rychlik [17]. Furthermore, Rychlik [20] stated necessary and sufficient conditions on possibly nonidentical marginal distributions for the existence of a random vector such that the distribution function of a given order statistic attains a sharp bound uniformly.

Under the assumption that the X_j's are maximally (resp. minimally) stable of order $i \in \{1, \ldots, n\}$, i.e., if the distribution $F_{(i)}$ of $\max(X_{k_1}, \ldots, X_{k_i})$ (resp. $G_{(i)}$ of $\min(X_{k_1}, \ldots, X_{k_i})$) is the same for any subset $\{k_1, \ldots, k_i\}$ of size i of $\{1, \ldots, n\}$, Papadatos [15] established one-sided uniformly sharp bounds for distribution functions of single order statistics in terms of $F_{(i)}$ (resp. $G_{(i)}$). Two-sided extensions of Papadatos' bounds to the case of linear combinations of distribution functions of order statistics were given by Okolewski [13].

Mallows [10] considered the special case of three element 2-independent samples with uniform univariate marginals and presented sharp bounds for expected values of the associated order statistics. The solution to this problem turned out

E-mail address: oko@p.lodz.pl.

http://dx.doi.org/10.1016/j.jmva.2017.05.007

to be very complicated and the author stated in his conclusion that, for this reason, he had not considered extensions of the problem to three-dimensional marginals. However, some point-wise attainable distribution bounds for order statistics from k-independent identically distributed n-element samples (see [7]) are known as well as some extensions to dependence structures other than k-independence (see [14]).

In this paper, we derive uniformly sharp bounds for linear combinations of distribution functions of order statistics from identically distributed random variables $X_1, \ldots, X_n$ having the same known distribution function F and such that for each $i \le k$, all i-tuples have copulas with the same known diagonal section δ_i. To do that, we provide necessary and sufficient conditions for a stochastic process to be an empirical distribution function of such X_j's, and use this characterization to determine bounds for order statistics in terms of F and the δ_i's, as well as to establish necessary and sufficient conditions for the uniform sharpness of these bounds. Moreover, for $k = 2$ and a wide class of diagonally 2-dependent samples, we derive stochastically extremal distributions of order statistics. Finally, for $k = 2$ and some other class of diagonally 2-dependent samples, we present uniformly sharp bounds for arbitrary linear combinations of distribution functions of order statistics. These results can be applied to study the properties of order statistics in the case when the interdependence among observations is partially known, e.g., to determine sharp bounds for expected values of order statistics and L-estimates.

The paper is organized as follows. In Section 2, we introduce a concept of diagonal dependency and give a characterization of empirical distributions from diagonally dependent identically distributed observations (Theorem 2.1). In Section 3, we apply Theorem 2.1 to determine uniformly sharp distribution bounds for order statistics (Theorem 3.1), and provide some explicit results (Propositions 3.3 and 3.7 and Corollaries 3.5 and 3.9).

2. A concept of diagonal dependency

Let Q^* be an arbitrary fixed joint distribution of n exchangeable random variables having uniform distribution on $(0, 1]$. In the sequel, k and n are fixed integers, $1 \le k \le n$, $n \ge 2$, while F is a given distribution function. Let $I_0(u) = [0, u]$ and $I_1(u) = (u, 1]$, $u \in [0, 1]$, and let $J_0(x) = (-\infty, x]$ and $J_1(x) = (x, \infty)$ for all $x \in \mathbb{R}$. We write C^* for the copula of Q^* and C^*_ℓ for the copula of ℓ-dimensional marginal distributions Q^*_ℓ of Q^*. By convention, $C^*_1(u) = u$ for $u \in [0, 1]$. We will assume that $\mathbb{X} = (X_1, \ldots, X_n)$ is a random vector such that each X_j has F as its distribution function and

$$\Pr\{X_{j_1} \in J_{s_1}(x), \ldots, X_{j_i} \in J_{s_i}(x)\} = Q^*_i\left[I_{s_1}\{F(x)\} \times \cdots \times I_{s_i}\{F(x)\}\right] \tag{1}$$

for any $1 \le i \le k$, $1 \le j_1 < \cdots < j_i \le n$, $s_1, \ldots, s_i \in \{0, 1\}$ and $x \in \mathbb{R}$. From now on, we will say that $\mathbb{X}$ is a vector of *diagonally C^*_k-dependent* random variables with the same distribution function F, or equivalently, that $\mathbb{X}$ is diagonally C^*_k-dependent and has the same (one-dimensional) marginal distribution function F. For $k = 1$, the condition (1) is equivalent to the fact that the X_j's are arbitrarily dependent and identically distributed with common distribution F. The X_j's satisfying the condition (1) with Q^* being the uniform distribution on the cube $(0, 1]^n$ will be called diagonally k-independent. Note that if the X_j's are diagonally k-dependent, then they are maximally and minimally stable of all orders no greater than k; cf. Papadatos [15].

The concept of diagonal dependency has a natural interpretation in terms of reliability. In fact, a reliability system with nonnegative diagonally C^*_k-dependent component lifetimes $X_1, \ldots, X_n$ is a system with the property that for each $i \le k$, all of its structurally identical sub-systems of i components (e.g., all of its i component series sub-systems) have the same known reliability. To see this, observe that for any fixed $i \le k$, structure function $\phi_i : \{0, 1\}^i \to \{0, 1\}$ (cf. Barlow and Proschan [1]) and $x > 0$, the probability $\Pr[\phi_i\{\mathbf{1}(X_{j_1} > x), \ldots, \mathbf{1}(X_{j_i} > x)\} = 1]$ is the same for all $1 \le j_1 < \cdots < j_i \le n$. Here and in what follows, $\mathbf{1}(s) = 1$ if s is true and $\mathbf{1}(s) = 0$ otherwise.

We shall need a characterization of empirical distributions from diagonally C^*_k-dependent samples. Let $X^*_1, \ldots, X^*_n$ stand for some random variables with the copula C^* and $\Pr(X^*_j \le x) = F(x)$ for all $j \in \{1, \ldots, n\}$. Define, for all $x \in \mathbb{R}$,

$$Z(x) = \sum_{j=1}^{n} \mathbf{1}(X^*_j \le x). \tag{2}$$

We will use the symbol $\stackrel{d}{=}$ to denote the equality in distribution and for any integers a, b write $[a; b] = \{a, a + 1, \ldots, b\}$ if $a \le b$ and $[a; b] = \emptyset$ otherwise.

Theorem 2.1. *Let $\mathbb{Y} = \{Y(x) : x \in \mathbb{R}\}$ be any stochastic process taking its values in $[0; n]$. Then the following statements are equivalent.*

(i) *There exists a random vector $\mathbb{X} = (X_1, \ldots, X_n)$ of exchangeable and diagonally C^*_k-dependent random variables with distribution function F such that, for every $x \in \mathbb{R}$,*

$$Y(x) \stackrel{d}{=} \sum_{j=1}^{n} \mathbf{1}(X_j \le x). \tag{3}$$

(ii) *$\mathbb{R} \ni x \mapsto \Pr\{Y(x) \ge j\}$ is a distribution function for every $j \in [1; n]$ and for $i \le k$ the following equality holds:*

$$\forall_{x\in\mathbb{R}} \quad \mathrm{E}\{Y^i(x)\} = \mathrm{E}\{Z^i(x)\}. \tag{4}$$