Journal of Computational and Applied Mathematics 235 (2011) 4629–4638

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Order statistics from mixed exchangeable random variables

Ismihan Bairamov [a,*], Safar Parsi [b]

[a] Department of Mathematics, Izmir University of Economics, Izmir, Turkey

[b] Department of Statistics, University of Mohaghegh Ardabili, Ardabil, Iran

ARTICLE INFO

Article history:
Received 23 October 2009
Received in revised form 9 April 2010

Keywords:
Order statistics
Exchangeable random variables
Distribution function

ABSTRACT

Two different exchangeable samples are considered and these two samples are assumed to be independent of each other. From these two samples a new sample is combined and treated as a single set of observations. The distribution of a single order statistic and the joint distribution of two order statistics for a new mixed sample are derived and expressed in terms of joint distribution functions. As a special case the distribution of a single order statistic and the joint distribution of two nonadjacent order statistics from exchangeable random variables are obtained. The results presented in this paper allows widespread applications in modelling of various lifetime data, biomedical sciences, reliability and survival analysis, actuarial sciences etc., where the assumption of independence of data cannot be accepted and the exchangeability is a more realistic assumption.

1. Introduction

The theory of order statistics has been extensively studied since the early part of the last century, and recent years have seen a particularly rapid growth of studies. For the basic theory of order statistics, description of their role in statistics and applications see [1–4]. Distributions of order statistics for independent and identically distributed (i.i.d.) random variables are well studied in both discrete and continuous cases. The distribution of rth order statistic for exchangeable and arbitrarily dependent random variables can be found, e.g. in [3] (formula (3.4.3) p. 46 and formula (5.3.1) p. 99). If $Y_1, Y_2, \ldots, Y_n$ are exchangeable random variables with $P\{Y_i \le y\} = F(y)$, then the distribution function of rth order statistic and its dual are given by

$$F_{r:n}(y) = \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} \binom{n}{j} F_{j:j}(y)$$

$$F_{n-r+1:n}(y) = \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} \binom{n}{j} F_{1:j}(y),$$

where $F_{j:j}(y) = P\{\max(Y_1, Y_2, \ldots, Y_j) \le y\}$ and $F_{1:j}(y) = P\{\min(Y_1, Y_2, \ldots, Y_j) \le y\}$.

Let $X_1, \ldots, X_{n_1}$ be exchangeable random variables, with joint cumulative distribution function (c.d.f.) $F(x_1, x_2, \ldots, x_{n_1})$, univariate marginal c.d.f. $F(x)$ and probability density function (p.d.f) $f(x)$ and $Y_1, \ldots, Y_{n_2}$ be exchangeable random variables with continuous joint c.d.f. $G(y_1, \ldots, y_{n_2})$ having univariate marginals $G(x)$ and p.d.f. $g(x)$. We assume that these two collections of random variables are independent of each other. Denote by $\{W_1, \ldots, W_n\}$ the $n = n_1 + n_2$ random variables combined from $n_1 X$s and $n_2 Y$s and treated as a single set of observations. It is clear that $W_1, \ldots, W_n$ are in general

* Corresponding author. Tel.: +90 02324888131; fax: +90 02324888339.
E-mail address: ismihan.bayramoglu@ieu.edu.tr (I. Bairamov).

0377-0427/$ – see front matter
doi:10.1016/j.cam.2010.04.030

not exchangeable. Let $W_{1:n} \leq \cdots \leq W_{n:n}$ be the order statistics of $W_1, \ldots, W_n$. If $n_1 = n$ and $n_2 = 0$ ($n_1 = 0$ and $n_2 = n$) then $W_1, \ldots, W_n$ are exchangeable random variables with joint c.d.f. $F(x_1, x_2, \ldots, x_n)(G(y_1, y_2, \ldots, y_{n_2}))$. Note that by definition of exchangeability $F(x_1, x_2, \ldots, x_n) = F(x_{i_1}, \ldots, x_{i_n})$ and $G(y_1, \ldots, y_n) = G(y_{i_1}, \ldots, y_{i_n})$, for every permutation $(i_1, i_2, \ldots, i_n)$ of the integers $(1, 2, \ldots, n)$.

In this paper we are interested in distributions of order statistics $W_{i:n}, i = 1, 2, \ldots, n$.

The need for ordering of mixed observation exists for example, in reliability analysis when n components of a technical system are randomly selected from two independent collections of elements. Each collection consists of dependent but identical components, this assumes that the life length of the components are exchangeable random variables. If components of the system are subject to the same set of stresses or shocks, then the failure of any component results in an increased load on the surviving components from the same collection and does not affect the life length of second type components. This means that the components from each collection work interactively between themselves and all of them are components of the unit system. The random variables $X_1, X_2, \ldots, X_{n_1}$ and $Y_1, Y_2, \ldots, Y_{n_2}$ with symmetric joint distribution functions $F(x_1, x_2, \ldots, x_{n_1})$ and $G(y_1, y_2, \ldots, y_{n_2})$ denote the life lengths of the components from the first and second collection, respectively. Then, for example, a $(n - s + 1)$-out-of-n system has life length $W_{s:n}$. The mean residual life function of such a system in the system level is $\Psi(t) = E(W_{s:n} - t \mid W_{r:n} > t), r < s$. The function $\Psi(t)$ expresses the mean residual life length of a $n - s + 1$ out-of-n system given that at least $n - r + 1$ components are alive at the moment t. (see e.g. [5–9]).

In Section 2 the distribution function $H_{(r)}(x)$ of rth order statistic $W_{r:n}$ has been derived and expressed in terms of $F(\underbrace{w, w, \ldots, w}_{i})$ and $G(\underbrace{w, w, \ldots, w}_{i})$. In Section 3 the joint distribution function of $W_{r:n}$ and $W_{s:n}$ has been derived and expressed in terms of $F(\underbrace{w, w, \ldots, w}_{i}, \underbrace{z, z, \ldots, z}_{t})$ and $G(\underbrace{w, w, \ldots, w}_{i}, \underbrace{z, z, \ldots, z}_{t})$. In special cases one obtains the distributions of order statistics from independent and exchangeable mixed collections and from exchangeable samples. The results presented in this paper allows wide spread applications in modelling of various lifetime data, biomedical sciences, reliability and survival analysis, actuarial sciences etc., where the assumption of independence of data cannot be accepted and the exchangeability is more realistic assumption.

2. Distribution of order statistic $W_{r:n}$

Denote the distribution function of order statistic $W_{r:n}$ by $H_{(r)}(w)$. Then according to the basic theory of order statistics one can write

$$\begin{aligned} H_{(r)}(w) &= P(W_{r:n} \leq w) = P(\text{at least } r \text{ of } W_1, \ldots, W_n \leq w) \\ &= \sum_{i=r}^{n} P(\text{exactly } i \text{ of } W_1, \ldots, W_n \leq w). \end{aligned} \tag{1}$$

Since $\mathbf{X} \equiv (X_1, X_2, \ldots, X_{n_1})$ and $\mathbf{Y} \equiv (Y_1, Y_2, \ldots, Y_{n_2})$ are vectors of exchangeable random variables and $\mathbf{X}$ and $\mathbf{Y}$ are independent, then applying the total probability formula one has

$$\begin{aligned} P\{\text{exactly } i \text{ of } W_1, \ldots, W_n \text{ are } \leq w\} &= \sum_{j=\max(0, n_1+i-n)}^{\min(i,n_1)} P\Big\{\text{exactly } j \text{ of } X's \text{ are } \leq w;\ (i-j) \text{ of } Y's \text{ are } \leq w; \\ &\quad n_1 - j \text{ of } X's \text{ are } > w \text{ and } n_2 - i + j \text{ of } Y's \text{ are } > w\Big\} \\ &= \sum_{j=\max(0, n_1+i-n)}^{\min(i,n_1)} P\Big\{\text{exactly } j \text{ of } X's \text{ are } \leq w \text{ and } n_1 - j \text{ of } X's \text{ are } > w\Big\} \\ &\quad \times P\Big\{\text{exactly } (i-j) \text{ of } Y's \text{ are } \leq w \text{ and } n_2 - i + j \text{ of } Y's \text{ are } > w\Big\} \\ &= \sum_{j=\max(0, n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j} P\Big\{X_1 \leq w, \ldots, X_j \leq w, X_{j+1} > w, \ldots, X_{n_1} > w\Big\} \\ &\quad \times \binom{n_2}{i-j} P\Big\{Y_1 \leq w, \ldots, Y_{i-j} \leq w, Y_{i-j+1} > w, \ldots, Y_{n_2} > w\Big\}. \end{aligned} \tag{2}$$

Formula (2) has been obtained by using the following consideration: using total probability formula for i places of $W \leq w$, one chooses exactly j Xs from total n_1 Xs, and for remaining free $i - j$ places of $W \leq w$ one chooses $i - j$ Ys from total $n_2 = n - n_1$ Ys. It is clear that j must satisfy $i - j \leq n - n_1$ and if $n_1 - n + i < 0$, then j starts from 0. Obviously, if $n_1 < i$, then one can choose at most i Xs, therefore $j \leq \min(i, n_1)$.

Throughout this paper we will denote $\mathbf{w}_t = (\underbrace{w, \ldots, w}_{t})$. Consider the function $F_{n_1-j,n_1}(\mathbf{w}_{n_1}) \equiv P(X_1 \leq w, \ldots, X_j \leq w, X_{j+1} > w, \ldots, X_{n_1} > w)$ and $G_{n_2-i+j,n_2}(\mathbf{w}_{n_2}) \equiv P(Y_1 \leq w, \ldots, Y_{i-j} \leq w, Y_{i-j+1} > w, \ldots, Y_{n_2} > w)$, appearing in the formula (2). In applications, for easy calculations it is important to express $H_{(r)}(w)$ in terms of marginal c.d.f.'s of $F(w_1, w_2, \ldots, w_{n_1})$ and $G(w_1, w_2, \ldots, w_{n_2})$. In the following lemma the expression for $F_{n_1-j,n_1}(\mathbf{w}_{n_1})$ in terms of $F(\underbrace{w, \ldots, w}_{t}) = P\{\max(X_1, \ldots, X_t) \le w\}$ has been obtained.

Lemma 1.

$$F_{n_1-j,n_1}(\mathbf{w}_{n_1}) = \sum_{t=j}^{n_1} (-1)^{t-j} \binom{n_1 - j}{t - j} F_{0,t}(\mathbf{w}_t). \tag{3}$$

where $F_{0,t}(\mathbf{w}_t) = F(\underbrace{w, \ldots, w}_{t})$.

Proof. Denote $A = \{X_1 \le w, \ldots, X_j \le w\}$ and $B_l = \{X_l \le w\}$, for $l = j+1, \ldots, n_1$, then

$$\begin{aligned} P\left(A - \bigcup_{l=j+1}^{n_1} B_l\right) &= P(X_1 \le w, \ldots, X_j \le w, X_{j+1} > w, \ldots, X_{n_1} > w) \\ &= F_{n_1-j,n_1}(\mathbf{w}_{n_1}), \end{aligned} \tag{4}$$

where $P(A) = F_{0,j}(\mathbf{w}_j)$, $\mathbf{w}_j = (\underbrace{w, \ldots, w}_{j})$. Then one has

$$\begin{aligned} P\left(A - \bigcup_{l=j+1}^{n_1} B_l\right) &= P(A) - P\left[A \bigcap \bigcup_{l=j+1}^{n_1} B_l\right] \\ &= P(A) - P\left[\bigcup_{l=j+1}^{n_1} C_l\right], \end{aligned}$$

where $C_l = A \cap B_l = \{X_1 \le w, \ldots, X_j \le w, X_l \le w\}$, $l = j+1, \ldots, n_1$ and $C_j = A$. Using the inclusion–exclusion principle for the events $C_{j+1}, \ldots, C_{n_1}$, we have

$$P\left[\bigcup_{l=j+1}^{n_1} C_l\right] = \sum_{t=j+1}^{n_1} (-1)^{t-j-1} \sum_{\substack{I \subset \{j+1, \cdots, n_1\} \\ |I| = t}} P(C_I), \tag{5}$$

where the last sum runs over all subsets I of the indices $j+1, \ldots, n_1$ which contain exactly t elements, and $C_I \equiv \bigcap_{l \in I} C_l$ denotes the intersection of all those C_l with index in I.

Since $X_1, \ldots, X_{n_1}$ have symmetric c.d.f. $G(x_1, \ldots, x_{n_1})$, then in this case we have, $P(C_{I_1}) = P(C_{I_2})$ for any I_1 and I_2, where $I_1 \subset \{j+1, \ldots, n_2\}$ and $I_2 \subset \{j+1, \ldots, n_2\}$ and $|I_1| = |I_2| = t$ for $t = j+1, \ldots, n_1$. Also, the number of distinct $(t-j-1)$-subsets on a set of $n_1 - j$ elements is given by the $\binom{n_1-j}{t-j-1}$. Since $F_{0,j}(\mathbf{w}_j) = P(C_j) = P(A)$ then from (4) and (5) one has

$$\begin{aligned} F_{n_1-j,n_1}(\mathbf{w}_{n_1}) &= F_{0,j}(\mathbf{w}_j) - \sum_{t=j+1}^{n_1} (-1)^{t-j-1} \sum_{\substack{I \subset \{j+1, \cdots, n_1\} \\ |I| = t}} P(C_I) \\ &= F_{0,j}(\mathbf{w}_j) - \sum_{t=j+1}^{n_1} (-1)^{t-j-1} \binom{n_1 - j}{t - j - 1} F_{0,t}(\mathbf{w}_t) \\ &= \sum_{t=j}^{n_1} (-1)^{t-j} \binom{n_1 - j}{t - j - 1} F_{0,t}(\mathbf{w}_t). \end{aligned}$$

where $F_{0,t}(\mathbf{w}_t) = P\{X_1 \le w, \ldots, X_t \le w\}$.

Thus the lemma proved. □

From the Lemma 1 using n_2 instead of n_1, G instead of F, $i - j$ instead of j, one obtains for $G_{n_2-i+j,n_2}(\mathbf{w}_{n_2}) = P(Y_1 \le w, \ldots, Y_{i-j} \le w, Y_{i-j+1} > w, \ldots, Y_{n_2} > w)$

$$G_{n_2-i+j,n_2}(\mathbf{w}_{n_2}) = \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2 - i + j}{t - i + j} G_{0,t}(\mathbf{w}_t). \tag{6}$$

Using (3), (6) and formula (2) the distribution function of rth order statistic then is given in the following

Theorem 1.

$$H_{(r)}(w) = P\{W_{r:n} \le w\} = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j}\binom{n_2}{i-j} \left\{\sum_{t=j}^{n_1}(-1)^{t-j}\binom{n_1-j}{t-j} F_{0,t}(\mathbf{w}_t)\right\} \times \left\{\sum_{t=i-j}^{n_2}(-1)^{t-i+j}\binom{n_2-i+j}{t-i+j} G_{0,t}(\mathbf{w}_t)\right\}, \tag{7}$$

where

$$F_{0,t}(\mathbf{w}_t) = F(\underbrace{w,\dots,w}_{t}) = P\{X_1 \le w,\dots,X_t \le w\} = P\{X_{t:t} \le w\}.$$

$$G_{0,t}(\mathbf{w}_t) = G(\underbrace{w,\dots,w}_{t}) = P\{Y_1 \le w,\dots,Y_t \le w\} = P\{Y_{t:t} \le w\}.$$

Case 1. It follows from (7) that if $r = n = n_1 + n_2$, then the c.d.f. of $W_{n:n} = \max(W_1, W_2, \dots, W_n)$ is

$$H_{(n)}(w) = P\{W_{n:n} \le w\} = F(\mathbf{w}_{n_1})G(\mathbf{w}_{n_2}). \tag{8}$$

For this special case one obtains c.d.f. of $W_{n:n}$ without knowing (7) as follows:

$$\begin{aligned} P\{W_{n:n} \le w\} &= P\{W_1 \le w, W_2 \le w, \dots, W_n \le w\} \\ &= P\{X_1 \le w,\dots,X_{n_1} \le w, Y_1 \le w,\dots,Y_{n_2} \le w\} \\ &= F(\underbrace{w,w,\dots,w}_{n_1})G(\underbrace{w,w,\dots,w}_{n_2}). \end{aligned}$$

which agrees with (8).

Remark 1. If $X_1, X_2, \dots, X_{n_1}$ and $Y_1, Y_2, \dots, Y_{n_2}$ are two independent samples with c.d.f. $F(x)$ and $G(x)$, respectively, then

$$H_{(r)}(w) = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j}\binom{n_2}{i-j} F(w)^j(1-F(w))^{n_1-j}G^{i-j}(w)(1-G(w))^{n_2-i+j}.$$

Remark 2. If $n_1 = 0$ and $n_2 = n$, then the c.d.f. of rth order statistics of exchangeable sample $Y_1, Y_2, \dots, Y_n$ is

$$H_{(r)}(w) = \sum_{i=r}^{n} C_n^i \sum_{t=i}^{n}(-1)^{t-i}C_{n-i}^{t-i}G_{0,t}(\mathbf{w}_t).$$

Corollary 1. *If $X_1, X_2, \dots, X_{n_1}$ are i.i.d. with c.d.f. $F(x)$ and $Y_1, Y_2, \dots, Y_{n_2}$ are exchangeable with joint c.d.f. $G(y_1, y_2, \dots, y_{n_2})$ then*

$$H_{(r)}(w) = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j}\binom{n_2}{i-j} (F(w))^j(1-F(w))^{n_1-j} \sum_{t=i-j}^{n_2}(-1)^{t-i+j}C_{n_2-i+j}^{t-i+j}G_{0,t}(\mathbf{w}_t).$$

2.1. Expression of c.d.f. of $W_{r:n}$ in terms of joint survival functions

In Theorem 1 the c.d.f. of $W_{r:n}$ is expressed in terms of joint distribution functions F and G given at the diagonal point $(w, w, \dots, w)$. In many practical applications, especially in reliability theory the joint distribution of random variables is given in terms of the survival function. The simple modifications of Lemma 1 and Theorem 1 allows us to express the c.d.f of $W_{r:n}$ in terms of $\bar{F}(w, w, \dots, w) = P\{X_1 > w, X_2 > w, \dots, X_i > w\}$, $2 \le i \le n_1$ and $\bar{G}(w, w, \dots, w) = P\{Y_1 > w, Y_2 > w, \dots, Y_j > w\}$, $2 \le j \le n_2$.

Lemma 1A.

$$F_{n_1-j,n_1}(\mathbf{w}_{n_1}) = \sum_{t=0}^{j}(-1)^t\binom{j}{t}\bar{F}(\underbrace{w,w,\dots,w}_{n_1-t}). \tag{9}$$

Proof. Taking $A = \{X_{j+1} > w, \dots, X_{n_1} > w\}$ and $B_i = \{X_i > w\}$ and repeating proof of Lemma 1 the proof easily can be completed. □

Theorem 1A.

$$H_{(r)}(w) = P\{W_{r:n} \le w\} = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j} \binom{n_2}{i-j} \left\{ \sum_{t=0}^{j} (-1)^t \binom{j}{t} \bar{F}(\underbrace{w, w, \ldots, w}_{n_1-t}) \right\}$$
$$\times \left\{ \sum_{t=0}^{j} (-1)^t \binom{j}{t} \bar{G}(\underbrace{w, w, \ldots, w}_{n_2-t}) \right\}. \quad (10)$$

2.2. The p.d.f. of $W_{r:n}$

The probability density function $h_{(r)}(w)$ of $W_{r:n}$ is

$$h_{(r)}(w) = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j} \binom{n_2}{i-j} \left\{ \sum_{t=j}^{n_1} (-1)^{t-j} \binom{n_1-j}{t-j} \frac{\mathrm{d}}{\mathrm{d}w} F_{0,t}(\mathbf{w}_t) \right\}$$
$$\times \left\{ \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \frac{\mathrm{d}}{\mathrm{d}w} G_{0,t}(\mathbf{w}_t) \right\} \quad (11)$$

or

$$h_{(r)}(w) = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j} \binom{n_2}{i-j} \left\{ \sum_{t=0}^{j} (-1)^t \binom{j}{t} \frac{\mathrm{d}}{\mathrm{d}w} \bar{F}(\underbrace{w, w, \ldots, w}_{n_1-t}) \right\}$$
$$\times \left\{ \sum_{t=0}^{j} (-1)^t \binom{j}{t} \frac{\mathrm{d}}{\mathrm{d}w} \bar{G}(\underbrace{w, w, \ldots, w}_{n_2-t}) \right\}. \quad (12)$$

Example 1. Suppose that $X_1, \ldots, X_{n_1}$, are i.i.d. random variables, with *Uniform*$(0, 1)$ distribution. Let $Y_1, \ldots, Y_{n_2}$ be exchangeable random variables with joint c.d.f.

$$G(y_1, \ldots, y_{n_2}) = \prod_{i=1}^{n_2} y_i \left\{ 1 + \alpha_{n_2} \sum_{1 \le k < j \le n_2} (1-y_k)(1-y_j) \right\}. \quad (13)$$

This distribution is simple Farlie–Gumbel–Morgenstern multivariate copula. The admissible range for an association parameter α_m allowing (13) to be a multivariate distribution function has been investigated in [10] and it is

$$-\frac{1}{\binom{n_2}{2}} \le \alpha_{n_2} \le \frac{1}{\left[\frac{n_2}{2}\right]},$$

where $[x]$ denotes the integer part of the number x.

For this distribution, $G_{0,t}(\mathbf{w}_t) = w^t[1 + \alpha_t \binom{t}{2} (1-w)^2]$. Then using Corollary 1 the c.d.f. of rth order statistic $W_{r:n}$ of the set of mixed observations $W_1, W_2, \ldots, W_n$ combined from $X_1, \ldots, X_{n_1}$ and $Y_1, Y_2, \ldots, Y_{n_2}$ is

$$H_{(r)}(w) = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \binom{n_1}{j} \binom{n_2}{i-j} w^j (1-w)^{n_1-j}$$
$$\times \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} w^t \left[1 + \alpha_t \binom{t}{2} (1-w)^2 \right]$$
$$= \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} \binom{n_2-i+j}{t-i+j} w^{j+t} (1-w)^{n_1-j}$$
$$+ \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} \alpha_t \binom{t}{2} w^j (1-w)^{n_1-j+2}$$

which is a mixture of the Beta distribution. The p.d.f. is

$$h_{(r)}(w) = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} (j+t) w^{j+t-1} (1-w)^{n_1-j}$$
$$+ \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} (j-n_1) w^{j+t} (1-w)^{n_1-j-1}$$
$$+ \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} \alpha_t j \binom{t}{2} w^{j-1} (1-w)^{n_1-j+2}$$
$$+ \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} \alpha_t (j-n_1-1) \binom{t}{2} w^{j} (1-w)^{n_1-j+1}.$$

The moments also can be calculated as

$$E(W_{r:n}^{\kappa}) = \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} (j+t) B(\kappa+j+t, n_1-j-1)$$
$$+ \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} (j-n_1) B(\kappa+j+t-1, n_1-j)$$
$$+ \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} \alpha_t j \binom{t}{2} B(\kappa+j, n_1-j+1)$$
$$+ \sum_{i=r}^{n} \sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)} \sum_{t=i-j}^{n_2} (-1)^{t-i+j} \binom{n_2-i+j}{t-i+j} \binom{n_1}{j} \binom{n_2}{i-j} \alpha_t (j-n_1-1) \binom{t}{2} B(\kappa+j-1, n_1-j),$$

where $B(a,b)$ is the beta function.

The distribution function of the extreme order statistic $W_{n:n}$, (i.e. for $r = n$ and $n_1 \geq 1$ and $n_2 \geq 2$) has simple expression

$$H_{(n)}(w) = w^{n_1+n_2} \left[1 + \alpha_{n_2} \binom{n_2}{2} (1-w)^2 \right].$$

3. The joint distribution of $W_{r:n}$ and $W_{s:n}$

Let $1 \leq r < s \leq n$. The joint distribution function of two order statistics $W_{r:n}$ and $W_{s:n}$ can be found by using similar considerations for order statistics. We have

$$H_{(r)(s)}(w,z) = P(W_{r:n} \leq w, W_{s:n} \leq z) = P(\text{at least } r\ W_i \leq w, \text{at least } s\, W_i \leq z)$$
$$= \sum_{j=s}^{n} \sum_{i=r}^{j} P(\text{exactly } i\, W_i \leq w, \text{exactly } j\, W_i \leq z). \tag{14}$$

By using total probability formula one can write

$$P\{\text{exactly } i\, W_i \leq w, \text{exactly } j\, W_i \leq z\}$$
$$= \sum_{p=\max(0,i-n_2)}^{\min(i,n_1)} \sum_{q=\max(0,j-i-n_2)}^{\min(j-i,n_1-p)} P\{\text{ exactly } p \text{ of } X'\text{s are} \leq w,\ q \text{ of } X\text{s} \in (w,z],\ n_1-p-q \text{ of } X'\text{s are } > z;$$
$$i-p \text{ of } Y'\text{s are } \leq w,\ (j-i-q) \text{ of } Y'\text{s } \in (w,z],\ n_1-p-q \text{ are } > z\}$$
$$= \sum_{p=\max(0,i-n_2)}^{\min(i,n_1)} \sum_{q=\max(0,j-i-n_2)}^{\min(j-i,n_1-p)} P\{\text{exactly } p \text{ of } X'\text{s are} \leq w,\ q \text{ of } X\text{s } \in (w,z],\ n_1-p-q \text{ of } X'\text{s are } > z\}$$
$$P\{\text{exactly } (i-p) \text{ of } Y'\text{s are } \leq w,\ (j-i-q) \text{ of } Y'\text{s } \in (w,z],\ (n_1-p-q) \text{ of } Y'\text{s are } > z\}$$
$$= \sum_{p=\max(0,i-n_2)}^{\min(i,n_1)} \sum_{q=\max(0,j-i-n_2)}^{\min(j-i,n_1-p)} \binom{n_1}{p} \binom{n_1-p}{q}$$
$$\times P\{X_1 \leq w, \ldots, X_p \leq w, X_{p+1} \in (w,z], \ldots, X_{p+q} \in (w,z], X_{p+q+1} > z, \ldots, X_{n_1} > z\}$$

$$\times \binom{n_2}{i-p}\binom{n_2-i+p}{j-i-q} P\{Y_1 \le w, \ldots, Y_{i-p} \le w, w < Y_{i-p+1} \le z, \ldots, w < Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \ldots, Y_{n_2} > z\}. \tag{15}$$

Formula (15) obtained by using the following consideration: by total probability formula, from the mixed sample $W_1, W_2, \ldots, W_n$ for i places of "$W \le w$" one chooses pXs from total n_1 Xs and q Xs from $n_1 - p$ Xs for $j - i$ places of "$w < W \le z$". Remaining $n_1 - p - q$ Xs will be used for free $n - (j - i)$ places of "$W > z$". For remaining $i - p$ free places of "$W \le w$" we select $i - p$ Ys from total $n_2 = n - n_1$ Ys; for remaining $j - i - q$ free places of "$w < W \le z$" one selects $j - i - q$ Ys from remaining $n_2 - (i - p)$ Ys; remaining $n_2 - (i - p) - (j - i - q) = n_2 - j + p + q$ Ys will be replaced to $n_2 - j + p + q$ free places of "$W > z$". It is clear that $p \le \min(i, n_1)$ and $i - p \le n - n_1$, i.e. $p \ge i - n_2$. If $i - n_2 < 0$, then p starts from 0. Analogously, $q \le \min(j - i, n_1 - p)$ and $j - i - q \le n - n_1$, i.e $q \ge j - i - n_2$. If $j - i - n_2 < 0$ then q starts from 0.

To obtain most appropriate formula for $H_{(r)(s)}(w, z)$ which is suitable for calculations in practical applications, the probabilities

$$P\{X_1 \le w, \ldots, X_p \le w, X_{p+1} \in (w, z], \ldots, X_{p+q} \in (w, z], X_{p+q+1} > z, \ldots, X_{n_1} > z\}$$

and

$$P\{Y_1 \le w, \ldots, Y_{i-p} \le w, w < Y_{i-p+1} \le z, \ldots, w < Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \ldots, Y_{n_2} > z\}$$

appearing in formula (15) must be expressed in terms of marginal distributions of $F(x_1, x_2, \ldots, x_{n_1})$ and $G(y_1, y_2, \ldots, y_{n_2})$. The following two lemmas serve for this purpose.

Lemma 2.

$$\begin{aligned} &P\{X_1 \le w, \ldots, X_p \le w, w \le X_{p+1} \le z, \ldots, w \le X_{p+q} \le z, X_{p+q+1} > z, \ldots, X_{n_1} > z\} \\ &\quad = P\{X_1 \le w, \ldots, X_p \le w, X_{p+1} \le z, \ldots, X_{p+q} \le z, X_{p+q+1} > z, \ldots, X_{n_1} > z\} \\ &\qquad + \sum_{l=1}^{q} (-1)^l \binom{q}{l} P\{X_1 \le w, \ldots, X_p \le w, X_{p+1} \le w, \ldots, X_{p+l} \le w, \\ &\qquad X_{p+l+1} \le z, \ldots, X_{p+q} \le z, X_{p+q+1} > z, \ldots, X_{n_1} > z\}. \end{aligned} \tag{16}$$

Proof. The probability that the random point $(\xi_1, \xi_2, \ldots, \xi_n)$ falls into parallelepiped $a_i \le \xi_i \le b_i$ $(i = 1, 2, \ldots, n)$, where a_i and b_i are arbitrary constants is

$$\begin{aligned} &P\{a_1 \le \xi_1 \le b_1, \ldots, a_n \le \xi_n \le b_n\} \\ &\qquad = P\{\xi_1 \le b_1, \ldots, \xi_n \le b_n\} - \sum_{i=1}^{n} p_i + \sum_{i<j} p_{ij} \pm \cdots + (-1)^n P\{\xi_1 \le a_1, \ldots, \xi_n \le a_n\}, \end{aligned} \tag{17}$$

where $p_{ij\cdots k}$ denotes the probability $P\{\xi_1 \le c_1, \xi_2 \le c_2, \ldots, \xi_n \le c_n\}$ for $c_i = a_i, c_j = a_j, \ldots, c_k = a_k$ and for the other indices c_s equal to b_s. (see, [11], page 135). A modification of (17) for $P\{M, K, a_1 \le \xi_1 \le b_1, \ldots, a_n \le \xi_n \le b_n\}$, where M and K are events is

$$\begin{aligned} &P\{M, a_1 \le \xi_1 \le b_1, \ldots, a_n \le \xi_n \le b_n, K\} \\ &\quad = P\{M, \xi_1 \le b_1, \ldots, \xi_n \le b_n, K\} - \sum_{i=1}^{n} p'_i + \sum_{i<j} p'_{ij} \pm \cdots + (-1)^n P\{M, K, \xi_1 \le a_1, \ldots, \xi_n \le a_n\}, \end{aligned} \tag{18}$$

where $p'_{ij\cdots k}$ denotes the probability $P\{M, \xi_1 \le c_1, \xi_2 \le c_2, \ldots, \xi_n \le c_n, K\}$ for $c_i = a_i, c_j = a_j, \ldots, c_k = a_k$ and for the other indices c_s equal to b_s.

In formula (18) taking $n = q$, by using $(X_{p+1}, \ldots, X_{p+q})$ instead of $(\xi_1, \ldots, \xi_n)$ and $a_i = w$ and $b_i = z, M = \{X_1 \le w, \ldots, X_p \le w\}, K = \{X_{p+q+1} > z, \ldots, X_{n_1} > z\}$ and recalling that X_i's are exchangeable, one obtains (16).

The lemma thus proved. □

From the Lemma 1 one easily obtains

$$\begin{aligned} &P\{Y_1 \le w, \ldots, Y_{i-p} \le w, w \le Y_{i-p+1} \le z, \ldots, w \le Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \ldots, Y_{n_2} > z\} \\ &\quad = P\{Y_1 \le w, \ldots, Y_{i-p} \le w, Y_{i-p+1} \le z, \ldots, Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \ldots, Y_{n_2} > z\} \\ &\qquad + \sum_{l=1}^{j-i-q} (-1)^l \binom{j-i-q}{l} P\{Y_1 \le w, \ldots, Y_{i-p} \le w, Y_{i-p+1} \le w, \ldots, Y_{i-p+l} \le w, \\ &\qquad \times Y_{i-p+l+1} \le z, \ldots, Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \ldots, Y_{n_2} > z\}. \end{aligned} \tag{19}$$

Now, the probability

$$P\{X_1 \le w, \dots, X_p \le w, X_{p+1} \le w, \dots, X_{p+l} \le w, X_{p+l+1} \le z, \dots, X_{p+q} \le z, X_{p+q+1} > z, \dots, X_{n_1} > z\}$$

appearing in the formula (16) can be expressed in terms of $F(\underbrace{w, w, \dots, w}_{p}, \underbrace{z, z, \dots, z}_{t})$ as given in the following lemma.

Lemma 3.

$$P\{X_1 \le w, \dots, X_p \le w, X_{p+1} \le z, \dots, X_{p+q} \le z, X_{p+q+1} > z, \dots, X_{n_1} > z\}$$
$$= \sum_{t=q}^{n_1-p} (-1)^{t-q} \binom{n_1-p-q}{t-q} F(\mathbf{w}_p, \mathbf{z}_t) \tag{20}$$

and

$$P\{X_1 \le w, \dots, X_p \le w, X_{p+1} \le w, \dots, X_{p+l} \le w, X_{p+l+1} \le z, \dots, X_{p+q} \le z, X_{p+q+1} > z, \dots, X_{n_1} > z\}$$
$$= \sum_{t=q-l}^{n_1-p-l} (-1)^{t-q-l} \binom{n_1-p-q}{t-q+l} F(\mathbf{w}_{p+l}, \mathbf{z}_t), \tag{21}$$

where $F(\mathbf{w}_p, \mathbf{z}_t) = F(\underbrace{w, w, \dots, w}_{p}, \underbrace{z, z, \dots, z}_{t})$.

Proof. Proof of (20) is similar to the proof of Lemma 1. By considering $A = \{X_1 \le w, \dots, X_p \le w, X_{p+1} \le z, \dots, X_{p+q} \le z\}$ and $B_l = \{Y_l \le z\}$, for $l = p+q+1, \dots, n_1$ repeating similar considerations as in Lemma 1 the proof is completed. Analogously, (21) can be proved.

It follows from the Lemma 3 that

$$P\{Y_1 \le w, \dots, Y_{i-p} \le w, Y_{i-p+1} \le z, \dots, Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \dots, Y_{n_2} > z\}$$
$$= \sum_{t=j-i-q}^{n_2-i+p} (-1)^{t-j+i+q} \binom{n_2-j+p+q}{t-j+i+q} G(\mathbf{w}_{i-p}, \mathbf{z}_t) \tag{22}$$

and

$$P\{Y_1 \le w, \dots, Y_{i-p} \le w, Y_{i-p+1} \le w, \dots, Y_{i-p+l} \le wY_{i-p+l+1} \le z, \dots,$$
$$Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \dots, Y_{n_2} > z, Y_{i-p+\ell} \le w, Y_{i-p+\ell+1} > z, \dots, Y_{n_2-j+i+q-p+\ell} > z)$$
$$= \sum_{t=j-i-l-q}^{n_2-i+p-l} (-1)^{t-j+i+l+q} \binom{n_2-j+p+q}{t-j+i+q+l} G(\mathbf{w}_{i-p+l}, \mathbf{z}_t) \tag{23}$$

where $G(\mathbf{w}_{i-p}, \mathbf{z}_t) = G(\underbrace{w, w, \dots, w}_{i-p}, \underbrace{z, z, \dots, z}_{t})$. □

Remark 3. Accuracy of (22) and (23) can be verified by using independent variables. Assuming Ys being independent with c.d.f. F, in the left hand side of (22) we have

$$= P\{Y_1 \le w, \dots, Y_{i-p} \le w, Y_{i-p+1} \le z, \dots, Y_{j-p-q} \le z, Y_{j-p-q+1} > z, \dots, Y_{n_2} > z\}$$
$$= F^{i-p}(w) F^{j-i-q}(z)(1-F(z))^{n_2-j+p+q}.$$

In the right hand side of (18) we have

$$\sum_{t=j-i-q}^{n_2-i+p} (-1)^{t-j+i+q} \binom{n_2-j+p+q}{t-j+i+q} G(\mathbf{w}_{i-p}, \mathbf{z}_t)$$
$$= \sum_{t=j-i-q}^{n_2-i+p} (-1)^{t-j+i+q} \binom{n_2-j+p+q}{t-j+i+q} P\{Y_1 \le w, \dots, Y_{i-p} \le w, Y_{i-p+1} \le z, \dots, Y_{i-p+t} \le z\}$$
$$= F^{i-p}(w) \sum_{t=j-i-q}^{n_2-i+p} (-1)^{t-j+i+q} \binom{n_2-j+p+q}{t-j+i+q} F^k(z)$$
$$= F^{i-p}(w) F^{j-i-q}(z) \sum_{k=0}^{n_2-j+p+q} (-1)^k \binom{n_2-j+p+q}{k} F^k(z)$$
$$= F^{i-p}(w) F^{j-i-q}(z)(1-F(z))^{n_2-j+p+q}.$$

Therefore using Lemmas 2 and 3, taking into account (16), (19)–(23) in (15) the joint distribution of rth and sth order statistics $W_{r:n}$ and $W_{s:n}$ can be written as in the following

Theorem 2.

$$\begin{aligned} H_{(r)(s)}(w,z) &= P(W_{r:n} \le w, W_{s:n} \le z) \\ &= \sum_{j=s}^{n}\sum_{i=r}^{j}\sum_{p=\max(0,i-n_2)}^{\min(i,n_1)}\sum_{q=\max(0,j-i-n_2)}^{\min(j-i,n_1-p)} \binom{n_1}{p}\binom{n_1-p}{q}\binom{n_2}{i-p}\binom{n_2-i+p}{j-i-q} \\ &\quad \times \left\{\sum_{l=0}^{q}(-1)^l\binom{q}{l}\sum_{t=q-l}^{n_1-p-l}(-1)^{t-q-l}\binom{n_1-p-q}{t-q+l}F(\mathbf{w}_{p+l},\mathbf{z}_t)\right\} \\ &\quad \times \left\{\sum_{l=0}^{j-i-q}(-1)^l\binom{j-i-q}{l}\sum_{t=j-i-l-q}^{n_2-i+p-l}(-1)^{t-j+i+l+q}\binom{n_2-j+p+q}{t-j+i+q+l}G(\mathbf{w}_{i-p+l},\mathbf{z}_t)\right\}. \end{aligned}$$

Remark 4. If $n_1 = n,\ n_2 = 0$ and $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with c.d.f. F then from Theorem 2 one has

$$\begin{aligned} H_{(r)(s)}(w,z) &= \sum_{j=s}^{n}\sum_{i=r}^{j}\sum_{q=j}^{j}\sum_{p=i}^{i}\binom{n_1}{p}\binom{n_1-p}{q}\binom{n_2}{i-p}\binom{n_2-i+p}{j-i-q}(F(w))^p(F(z)-F(w))^{q-p}(1-F(z))^{n-q} \\ &= \sum_{j=s}^{n}\sum_{i=r}^{j}\frac{n!}{i!(j-i)!(n-j)!}(F(w))^i(F(z)-F(w))^{j-i}(1-F(z))^{n-j}. \end{aligned} \quad (24)$$

Remark 5. If $n_1 = 0$ and $n_2 = n$, then from Theorem 2 one obtains the joint distribution of order statistics $W_{r:n}$ and $W_{s:n}$ from exchangeable random variables $W_1, W_2, \ldots, W_n$ having joint distribution function $G(w_1, w_2, \ldots, w_n)$

$$\begin{aligned} H_{(r)(s)}(w,z) &= \sum_{j=s}^{n}\sum_{i=r}^{j}\frac{n!}{i!(j-i)!(n-j)!} \\ &\quad \times \left\{\sum_{l=0}^{j-i}(-1)^l C_{j-i}^l \sum_{t=j-i-l}^{n-i-l}(-1)^{t-j+i+l}\binom{n-j}{t-j+i}G(\underbrace{w,w,\ldots,w}_{i+l},\underbrace{z,z,\ldots,z}_{t})\right\}. \end{aligned} \quad (25)$$

It is easy to verify that if all variables $W_1, W_2, \ldots, W_n$ are i.i.d. with c.d.f. F, then (25) reduces to (24).

Corollary 2. *If $X_1, X_2, \ldots, X_{n_1}$ and $Y_1, Y_2, \ldots, Y_{n_2}$ are two independent samples with c.d.f. $F(x)$ and $G(x)$, respectively, then*

$$\begin{aligned} H_{(r)}(w) &= P\{W_{r:n} \le w\} \\ &= \sum_{i=r}^{n}\sum_{j=\max(0,n_1+i-n)}^{\min(i,n_1)}\binom{n_1}{j}\binom{n_2}{i-j}(F(w))^j(1-F(w))^{n_1-j}G^{j-i}(w)(1-G(w))^{n_2-i+j} \end{aligned}$$

and

$$\begin{aligned} H_{(r)(s)}(w,z) &= \sum_{j=s}^{n}\sum_{i=r}^{j}\sum_{p=\max(0,i-n_2)}^{\min(i,n_1)}\sum_{q=\max(0,j-i-n_2)}^{\min(j-i,n_1-p)}\binom{n_1}{p}\binom{n_1-p}{q}\binom{n_2}{i-p}\binom{n_2-i+p}{j-i-q}(F(w))^p \\ &\quad \times (F(z)-F(w))^q(1-F(z))^{n_1-p-q}G^{i-p}(w)(G(z)-G(w))^{j-i-q}(1-G(z))^{n_2-j+p+q}. \end{aligned}$$

Acknowledgements

The authors are grateful to the referees and the editor for the careful reading and constructive suggestions, which were helpful in improving the presentation.

References

[1] H.A. David, Order Statistics, second edition, Wiley, New York, 1981.
[2] B. Arnold, N. Balakrishnan, H.N. Nagaraja, A First Course in Order Statistics, Wiley, New York, 1992.
[3] H.A. David, H.N. Nagaraja, Order Statistics, third edition, Wiley, New York, 2003.
[4] N. Balakrishnan, Permanents, order statistics, outliers and robustness, Revista Matematica Complutense 20 (1) (2007) 7–107.

[5] I. Bairamov, M. Ahsanullah, I. Akhundov, A residual life function of a system having parallel or series structures, Journal of Statistical Theory and Applications 1 (2) (2002) 19–132.
[6] M. Asadi, I. Bairamov, A note on the mean residual life function of the parallel systems, Communications in Statistics — Theory and Methods 34 (2005) 1–12.
[7] M. Asadi, I. Bairamov, On the mean residual life function of the k-out-of- n systems at system level, IEEE Transactions on Reliability 55 (2006) 314–318.
[8] M. Asadi, S. Goliforushani, On the mean residual life function of coherent systems, IEEE Transactions on Reliability 57 (4) (2008) 574–580.
[9] S. Gurler, I. Bairamov, Parallel and k-out-of-n: G systems with nonidentical components and their mean residual life functions, Applied Matematical Modelling 33 (2) (2009) 1116–1125.
[10] I. Bairamov, S. Eryılmaz, Characterization of symmetry and exceedance models in multivariate FGM distributions, Journal of Applied Statistical Science 13 (2) (2004) 87–99.
[11] B.V. Gnedenko, The Theory of Probability, Mir Publishers, Moskow, 1978.