

On Multivariate Order Statistics by Marginal Ordering of i.n.n.i.d. Random Vectors Under Discontinuous d.f.'s

M. Güngör, A. Gökhan and Y. Bulut

Department of Mathematics, Firat University, 23119, Elazig, Turkey

Abstract: In this study, the joint Probability Density Function (p.d.f.) of multivariate order statistics by marginal ordering of the Independent and Not Necessarily Identically Distributed (i.n.n.i.d.) random vectors under discontinuous d.f.'s is examined. Furthermore, the continuous cases are derived from the discontinuous cases.

Key words: Multivariate Order Statistics, Discontinuous d.f., Joint p.d.f., Permanent

INTRODUCTION

Reiss [1] considered the joint Probability Density Function (p.d.f.) of any k order statistics of Independent and Identically Distributed (i.i.d.) random variables under a continuous Distribution Function (d.f.) and discontinuous d.f.. He also considered p.d.f. of bivariate order statistics by marginal ordering of bivariate i.i.d. random vectors with a continuous d.f. by means of multinomial probabilities of appropriate "cell frequency vectors", defining multivariate order statistics by marginal ordering of i.i.d. random vectors with a continuous d.f.

Guilbaud [2] expressed probability of the functions of Independent and Not Necessarily Identically Distributed (i.n.n.i.d.) random vectors as a linear combination of probabilities of the functions of i.i.d. random vectors and thus also for order statistics of random variables.

Vaughan and Venables [3], Samuel and Thomas [4] denoted the joint p.d.f. of order statistics of i.n.n.i.d. random variables by means of permanents.

David [5] considered the fundamental distribution theory of order statistics.

In this study, the joint p.d.f. of multivariate order statistics by marginal ordering of i.n.n.i.d. random vectors under discontinuous Distribution Functions (d.f.'s) is given. Moreover, transition from discontinuous cases to continuous cases is discussed.

From now on, the subscripts and superscripts are defined in the first place in which they are used and these definitions will be valid unless they are redefined. Consider $\mathbf{x}=(x^{(1)}, x^{(2)}, \dots, x^{(d)})$ and $\mathbf{y}=(y^{(1)}, y^{(2)}, \dots, y^{(d)})$, then it can be written as $x_j \leq y_j$ if $x_j^{(j)} \leq y_j^{(j)}$, $j=1,2,\dots,d$.

Let $\mathbf{x}_i=(x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(d)})$, $i=1,2,\dots,n$ be n i.n.n.i.d. random vectors under discontinuous d.f.'s and suppose that the components of \mathbf{x}_i are independent. The expression

$$\mathbf{X}_{r:n}^{(j)} = Z_{r:n}^{(j)} (\xi_1^{(j)}, \xi_2^{(j)}, \dots, \xi_n^{(j)}) \quad r=1,2,\dots,n \quad (1)$$

is stated as the r th order statistic of the j th components of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$. From (1), the ordered values of the j th components of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are expressed as:

$$\mathbf{X}_{r:n}^{(j)} \leq \mathbf{X}_{2:n}^{(j)} \leq \dots \leq \mathbf{X}_{n:n}^{(j)} \quad (2)$$

From (2), $\mathbf{X}_{r:n} = (\mathbf{X}_{r:n}^{(1)}, \mathbf{X}_{r:n}^{(2)}, \dots, \mathbf{X}_{r:n}^{(d)})$ can be written [1].

If $\mathbf{a}_1, \mathbf{a}_2, \dots$ are defined as column vectors, then the matrix obtained by taking i_1 copies of \mathbf{a}_1 , i_2 copies of \mathbf{a}_2, \dots can be denoted as:

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots \\ i_1 & i_2 & \dots \end{bmatrix}$$

and per A denotes the permanent of a square matrix A , which is defined as similar to determinants except that all terms in the expansion have a positive sign [4].

Joint p.d.f. of Multivariate Order Statistics by Marginal Ordering of i.n.n.i.d. Random Vectors Under Discontinuous d.f.'s:

Let $F_i^{(j)}$ be d.f. of $x_i^{(j)}$ and be discontinuous at any $x_1^{(j)}, 1=1,2,\dots,p, p=1, 2, \dots, n$ which is a realization of $x_1^{(j)}$. Thus, the random variable $F_i^{(j)}(x_i^{(j)+}) + \eta_i^{(j)}(F_i^{(j)}(x_i^{(j)}) - F_i^{(j)}(x_i^{(j)-}))$ which is uniformly distributed on the interval $(F_i^{(j)}(x_1^{(j)-}), F_i^{(j)}(x_1^{(j)+}))$ can be taken instead of the realization $F_i^{(j)}(x_1^{(j)})$ where, $F_i^{(j)}(x_1^{(j)-})$ denotes the left-hand limit of $F_i^{(j)}$ and $x_1^{(j)-}$ is uniformly distributed on $(0,1)$. If $y_1^{(j)}$ is a realization of $x_1^{(j)}$, then $y_1^{(j)}$ is taken instead of $x_1^{(j)}$. Moreover, $x_1^{(j)}$ and $x_2^{(j)}$ are assumed to be independent. If $H^{(j)}(y^{(j)}, x^{(j)}) = F^{(j)}(x_i^{(j)-}) + y_1^{(j)}(F^{(j)}(x_i^{(j)}) - F^{(j)}(x_i^{(j)-}))$ then $H_i^{(j)}(x_1^{(j)}, x_2^{(j)})$ is uniformly distributed on $(0,1)$ [1]. Now, the joint p.d.f. of $\mathbf{X}_{r1:n}^{(j)}, \mathbf{X}_{r2:n}^{(j)}, \mathbf{X}_{r3:n}^{(j)}, \dots, \mathbf{X}_{rpn}^{(j)}$ and $\mathbf{X}_{r1:n}, \mathbf{X}_{r2:n}, \mathbf{X}_{r3:n}, \dots, \mathbf{X}_{rpn}$ will be given in Theorem 1 and 2, respectively.

Theorem 1: Let F_i be d.f. of random vector \mathbf{x}_i and $F_i^{(j)}$ be d.f. of $x_i^{(j)}$ and discontinuous at any $x_1^{(j)}$, and $B_p^{(j)}$ denote the set of all $(x_1^{(j)}, y_1^{(j)}, x_2^{(j)}, y_2^{(j)}, \dots, x_p^{(j)}, y_p^{(j)})$

satisfying $0 < y_1^{(i)} < 1$ and $x_{1-1}^{(i)} < x_1^{(i)}$ or $x_{1-1}^{(i)} = x_1^{(i)}$ and $y_{1-1}^{(i)} < y_1^{(i)}$. Then, the joint p.d.f. of $X_{r_1:n}^{(i)}, X_{r_2:n}^{(i)}, \dots, X_{r_p:n}^{(i)}$ is:

$$f_{r_1, r_2, \dots, r_p, n}^{(i)}(x_1^{(i)}, x_2^{(i)}, \dots, x_p^{(i)}) = \int C^{-1} \text{per } A \, dy_1^{(i)} dy_2^{(i)} \dots dy_p^{(i)} \tag{3}$$

for $B_p^{(i)}$ and $=0$, otherwise, where, $0 = r_0 < r_1 < \dots < r_p < r_{p+1} = n + 1$, $A =$

$$\begin{bmatrix} H^{(i)}(y_1^{(i)}, x_1^{(i)}) & 1_1^{(i)} & H^{(i)}(y_2^{(i)}, x_2^{(i)}) - H^{(i)}(y_1^{(i)}, x_1^{(i)}) & \dots & 1_p^{(i)} & 1 - H^{(i)}(y_p^{(i)}, x_p^{(i)}) \\ r_1 - 1 & 1 & r_2 - r_1 - 1 & & 1 & n - r_p \end{bmatrix}$$

$$H^{(i)}(y_l^{(i)}, x_l^{(i)}) = \left(H_1^{(i)}(y_l^{(i)}, x_l^{(i)}) H_2^{(i)}(y_l^{(i)}, x_l^{(i)}) \dots H_n^{(i)}(y_l^{(i)}, x_l^{(i)}) \right)^{l-1}, l = 0, 1, \dots, p + 1$$

and $1 = (1 \ 1 \ \dots \ 1)$, $1_1^{(i)} = (1 \ 1 \ \dots \ 1)$, $l = 1, 2, \dots, p$ are column vectors with n components and $C = \prod_{l=0}^p (r_{l+1} - r_l - 1)!$

$$\left(H_1^{(i)}(y_0^{(i)}, x_0^{(i)}) = 0, H_1^{(i)}(y_{p+1}^{(i)}, x_{p+1}^{(i)}) = 1 \right).$$

Proof: Since $1, 2, \dots, n$ are multivariate i.n.n.i.d. random vectors under discontinuous d.f.'s, the joint p.d.f. of $X_{r_1:n}^{(i)}, X_{r_2:n}^{(i)}, \dots, X_{r_p:n}^{(i)}$ at $x_1^{(i)}, x_2^{(i)}, \dots, x_p^{(i)}$ is obtained as follows. For this purpose, the permanent of matrix A_1 of order n will be used, which has $r_{l+1} - r_l - 1$ columns of $H^{(i)}(y_{l+1}^{(i)}, x_{l+1}^{(i)}) - H^{(i)}(y_l^{(i)}, x_l^{(i)})$ between column r_l and column r_{l+1} ($l = 0, 1, \dots, p$) and a column $dy_l^{(i)} = (dy_1^{(i)} dy_2^{(i)} \dots dy_l^{(i)})$ at r_l th column ($l = 1, 2, \dots, p$). Moreover, when $X_{r_1:n}^{(i)}, X_{r_2:n}^{(i)}, \dots, X_{r_p:n}^{(i)}$ is considered, the permanent should be multiplied by C^{-1} . If the integration of $C^{-1} \text{per } A = C^{-1} \text{per } A \, dy_1^{(i)} dy_2^{(i)} \dots dy_p^{(i)}$ is taken over $B_p^{(i)}$, then the desired result is obtained.

Theorem 2: Let F_1 be d.f. of random vector y_1 and $F_1^{(i)}$ be d.f. of $y_1^{(i)}$ and $F_1^{(i)}$ be discontinuous at $x_1^{(i)}$. Moreover, assume that the same conditions in Theorem 1 hold. Then, the joint p.d.f. of $X_{r_1:n}, X_{r_2:n}, \dots, X_{r_p:n}$ is:

$$f_{r_1, r_2, \dots, r_p, n}(x_1, x_2, \dots, x_p) = \prod_{j=1}^d f_{r_1, r_2, \dots, r_p, n}^{(j)}(x_1^{(j)}, x_2^{(j)}, \dots, x_p^{(j)}) \tag{4}$$

for $\bigcup_{j=1}^d B_p^{(j)}$ and $=0$, otherwise, where, $x_i = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(d)})$ and $y_i = (y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(d)})$.

Proof: Omitted.

RESULTS AND DISCUSSION

In (3) (or in (4) for $d=1$), if $H_i^{(i)}(y_1^{(i)}, x_1^{(i)}) = H^{(i)}(y_1^{(i)}, x_1^{(i)})$, then the joint p.d.f. of order statistics of i.i.d. random variables under discontinuous d.f. is obtained. This result gives Corollary 1.5.7. in [1].

In (3) (or in (4) for $d=1$), if $H_i^{(i)}(y_1^{(i)}, x_1^{(i)}) = F_1^{(i)}(x_1^{(i)})$, then the joint p.d.f. of order statistics of i.n.n.i.d. random variables with continuous d.f.'s is obtained. In this case, the obtained joint p.d.f. is equal to $\text{per } A'$ if $\text{per } A'$ is determined by taking the $(f_1^{(i)}(x_1^{(i)}), f_2^{(i)}(x_1^{(i)}), \dots, f_n^{(i)}(x_1^{(i)}))$ instead of $1_1^{(i)}$ in $\text{per } A$. This result gives the opened form of the expression which is constructed by

taking order statistics instead of T in (4) in [2] and the general case [3].

In (3) (or in (4) for $d=1$), if $H_i^{(i)}(y_1^{(i)}, x_1^{(i)}) = F^{(i)}(x_1^{(i)})$, then the joint p.d.f. of order statistics of i.i.d. random variables with continuous d.f. is obtained. This result gives (2.2.3.) in Theorem 1.4.5. in [1].

In (4), if $H_i(y_1, x_1) = H(y_1, x_1)$, then the joint p.d.f. of multivariate order statistics by marginal ordering of i.i.d. random vectors under discontinuous d.f. is obtained.

In (4), if $H_i(y_1, x_1) = F_1(x_1)$, then the joint p.d.f. of multivariate order statistics by marginal ordering of i.n.n.i.d. random vectors with continuous d.f.'s is obtained.

In (4), if $H_i(y_1, x_1) = F(x_1)$, then the joint p.d.f. of multivariate order statistics by marginal ordering of i.i.d. random vectors with continuous d.f. is obtained. Whenever the components of vectors are independent in Lemma 2.2.2 in [1], this result gives the Lemma for $d=2$ and $p=1$.

Note that the discrete case occurs, if the number of discontinuous points are infinite.

REFERENCES

1. Reis, R.D., 1989. Approximate Distributions of Order Statistics. Springer-Verlag New York Inc.
2. Guilbaud, O., 1982. Functions of non-i.i.d. random vectors expressed as functions of i.i.d. random vectors. Scand. J. Statist., 9: 229-233.
3. Vaughan, R.J. and W.N. Venables, 1972. Permanent expressions for order statistic densities. J. Roy. Statist. Soc. Ser. B, 34: 308-310.
4. Samuel, P. and P.Y. Thomas, 1998. Modified expression for a recurrence relation on the product moments of order statistics. Statistics and Probability Lett., 37: 89-95.
5. David, H.A., 1981. Order Statistics. 2nd Ed. Wiley, New York.