ISSN:1306-3111
e-Journal of New World Sciences Academy
2008, Volume: 3, Number: 4
Article Number: A0105

NATURAL AND APPLIED SCIENCES
STATISTICS
❖ **STATISTICAL THEORY**
Received: May 2008
Accepted: September 2008

Yunus Bulut
Mehmet Güngör
University of Firat
ybulut79@hotmail.com
Elazig-Turkiye

ON THE DISTRIBUTIONS OF ORDER STATISTICS OF RANDOM VECTORS

ABSTRACT

In this study, the joint probability function and joint probability density function of order statistics of i.n.n.i.d random vectors which number of components are different are given respectively for discrete and discontinuous cases. Furthermore, continuous cases are derived from the discrete and discontinuous cases.

Keywords: Order Statistics, Discontinuous Random Vectors, Discrete Random Vectors, Continuous Random Vectors, Joint Probability Function, Joint Probability Density Function, Permanent

TESADÜFİ VEKTÖRLERİN SIRALI İSTATİSTİKLERİNİN DAĞILIMLARI ÜZERİNE

ÖZET

Bu çalışmada, bileşenlerinin sayısı farklı olan i.n.n.i.d. tesadüfi vektörlerin sıralı istatistiklerinin bileşik olasılık ve bileşik olasılık yoğunluk fonksiyonları kesikli ve süreksiz durumlar için sırasıyla verilmiştir.Ayrıca, kesikli ve süreksiz durumlardan sürekli durumlar türetilmiştir.

Anahtar Kelimeler: Sıralı İstatistikler, Süreksiz Tesadüfi Vektörler, Kesikli Tesadüfi Vektörler, Sürekli tesadüfi Vektörler, Bileşik Olasılık Fonksiyonu, Bileşik Olasılık Yoğunluk Fonksiyonu, Permanent

1. INTRODUCTION (GİRİŞ)

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

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

and per$\mathbf{A}$ denotes the permanent of a square matrix $\mathbf{A}$, which is defined as similar to determinants except that all terms in the expansion have a positive sign [1].

David [2] considered the fundamental distribution theory of order statistics.

Guilbaud [3] 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.

Khatri [4] examined marginal probability function (p.f.) of a single order statistic and the joint p.f. for any two order statistics of i.i.d. random variables for discrete case.

Reiss [5] 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..

Samuel and Thomas [6] and Vaughan and Venables [7] denoted the joint p.d.f. of order statistics of i.n.n.i.d. random variables by means of permanents.

In this study, the joint p.f., and joint p.d.f. of order statistics by marginal ordering of i.n.n.i.d. random vectors which are number of components under discontinuous distribution functions (d.f.'s) are given. Moreover, transitions from discontinuous and discrete cases to continuous cases are 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 $\mathbf{x}=(x^{(1)}, x^{(2)},\ldots, x^{(d)})$ and $\mathbf{y}=(y^{(1)}, y^{(2)},\ldots, y^{(d)})$, then it can be written as $\mathbf{x}\leq\mathbf{y}$ if $x^{(j)}\leq y^{(j)}$, $j=1,2,\ldots,d$.

Let $\boldsymbol{\xi}_i=(\xi_i^{(1)}, \xi_i^{(2)},\ldots, \xi_i^{(d)})$, $i=1,2,\ldots,n$ be n i.n.n.i.d. random vectors which components of $\boldsymbol{\xi}_i$ are independent. Moreover, $n(1)$ is the number of components of $\xi_i^{(1)}$, $n(2)$ is the number of components of $\xi_i^{(2)}$,….,$n(d)$ is the number of components of $\xi_i^{(d)}$ and $n(1)$, $n(2)$,…, $n(d)$ may be different. The expression

$$X^{(j)}_{r:n(j)} = Z_{r:n(j)}\left(\xi^{(j)}_1,\ \xi^{(j)}_2\ ,\ldots,\ \xi^{(j)}_{n(j)}\right) \qquad r=1,2,\ldots,n(j) \qquad (1)$$

is stated as the rth order statistic of the jth components of $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, …, $\boldsymbol{\xi}_n$. Using (1.1), the rth order statistic of $\boldsymbol{\xi}_i$'s in component wise can be expressed as

$$\mathbf{x}_{r:n} = \left(X^{(1)}_{r:n(1)}, X^{(2)}_{r:n(2)},\ldots, X^{(d)}_{r:n(d)}\right)$$

where $n = \max\{n(j): 1\leq j\leq d\}, r\leq n, [F_i^{(j)}]^{-1}(1) = \sup\{x: F_i^{(j)}(x) < 1\}$ and

$$X_{r:n(j)} = \begin{cases} \left[F_i^{(j)}\right]^{-1}(1), & r > n(j) \\ X_{r:n(j)}, & r \leq n(j) \end{cases} .$$

From (1), the ordered values of the jth components of ξ_1, ξ_2, …, ξ_n are expressed as

$$X^{(j)}_{1:n(j)} \le X^{(j)}_{2:n(j)} \le \ldots \le X^{(j)}_{n(j):n(j)} . \tag{2}$$

2. RESEARCH SIGNIFICANCE (ÇALIŞMANIN ÖNEMİ)

In this study, the joint p.f. and joint p.d.f. of order statistics by marginal ordering of i.n.n.i.d. random vectors which are number of components under discontinuous distribution functions (d.f.'s) are given. Moreover, transitions from discontinuous and discrete cases to continuous cases are discussed.

3. ORDER STATISTICS OF I.N.N.I.D. RANDOM VECTORS FOR DISCRETE CASE (KESİKLİ I.N.N.I.D. TESADÜFİ VEKTÖRLERİN SIRALI İSTATİSTİKLERİ)

If the random variables are i.n.n.i.d. and $P(\xi = x) = f(x)$ is written for discrete case, the joint p.d.f. of $X^{(j)}_{r_1:n}, X^{(j)}_{r_2:n}, \ldots, X^{(j)}_{r_p:n}$ and $\mathbf{X}_{r_1:n}, \mathbf{X}_{r_2:n}, \ldots, \mathbf{X}_{r_p:n}$ will be given in Theorem 1 and Theorem 2, respectively.

Theorem 1. *Let F_i be p.f. of random vector ξ_i and $F^{(j)}_i$ be p.f. of $\xi^{(j)}_i$. Then, the joint p.f. of $X^{(j)}_{r_1:n}, X^{(j)}_{r_2:n}, \ldots, X^{(j)}_{r_p:n}$ is*

$$f^{(j)}_{r_1,r_2,\ldots,r_p:n(j)}\left(x^{(j)}_1, x^{(j)}_2, \ldots, x^{(j)}_p\right) = \sum \mathbf{C}.\,\mathrm{per}\mathbf{A} \tag{3}$$

and =0, otherwise, where $0 = r_0 < r_1 < \ldots < r_p < r_{p+1} = n+1$, **A=**

$$\left[\underset{r_1-k_1-1}{\mathbf{F}^{(j)}\left(x^{(j)}_1 - 1\right)} \quad \underset{k_1+m_1+1}{\mathbf{f}^{(j)}_1\left(x^{(j)}_1\right)} \quad \underset{r_2-r_1-m_1-k_2-1}{\mathbf{F}^{(j)}\left(x^{(j)}_2 - 1\right) - \mathbf{F}^{(j)}\left(x^{(j)}_1\right)} \ldots \quad \underset{\ldots}{\mathbf{f}^{(j)}_p\left(x^{(j)}_p\right)} \quad \underset{n(j)-r_p-m_p}{1 - \mathbf{F}^{(j)}\left(y^{(j)}_p, x^{(j)}_p\right)}\right]$$

which consists of column vectors with $n(j)$ components such as

$\mathbf{F}^{(j)}\left(x^{(j)}_l\right) = \left(F^{(j)}_1\left(x^{(j)}_l\right) F^{(j)}_2\left(x^{(j)}_l\right) \ldots F^{(j)}_{n(j)}\left(x^{(j)}_l\right)\right)'$, $l = 0,1,\ldots,p+1$ *and*

C=[$(r_1 - k_1 -)$!($k_1 + m_1 + 1$)!...($n(j) - r_p - m_p$)!]$^{-1}$ *and under the conditions such that* $m_l + k_{l+1} \le r_{l+1} - r_l - 1$, $m_0 = 0$, $k_{p+1} = 0$ *the sum is*

$$\sum = \sum_{k_1=0}^{r_1-1} \sum_{m_1=0}^{r_2-r_1-1} \sum_{k_2=0}^{r_2-r_1-1} \cdots \sum_{m_p=0}^{n(j)-r_p} .$$

Proof. Omitted.

Theorem 2. *Let F_i be p.f. of random vector ξ_i and $F^{(j)}_i$ be p.f. of $\xi^{(j)}_i$. Moreover, assume that the same conditions in Theorem 1 hold. Then, the joint p.f. of $\mathbf{X}_{r_1:n}$, $\mathbf{X}_{r_2:n}$, …, $\mathbf{X}_{r_p:n}$ is*

$$f_{r_1,r_2,\ldots,r_p:n}\left(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\right) = \prod_{j=1}^{d} f^{(j)}_{r_1,r_2,\ldots,r_p:n}\left(x^{(j)}_1, x^{(j)}_2, \ldots, x^{(j)}_p\right) \tag{4}$$

and =0, otherwise, where $\mathbf{x}_l = \left(x^{(1)}_l, x^{(2)}_l, \ldots., x^{(d)}_l\right)$.

Proof. Omitted.

4. ORDER STATISTICS OF I.N.N.I.D. RANDOM VECTORS FOR DISCONTINUOUS CASE (SÜREKSİZ I.N.N.I.D. TESADÜFİ VEKTÖRLERİN SIRALI İSTATİSTİKLERİ)

Let $F_i^{(j)}$ be d.f. of $\xi_i^{(j)}$ and be discontinuous at any $x_l^{(j)}, l=1,2,\ldots,p,\ p=1,2,\ldots,n(j)$ which is a realization of $\xi_l^{(j)}$. Thus, the random variable $F_i^{(j)}\left(x_l^{(j)-}\right)+\eta_l^{(j)}\left(F_i^{(j)}\left(x_l^{(j)}\right)-F_i^{(j)}\left(x_l^{(j)-}\right)\right)$ which is uniformly distributed on the interval $\left(F_i^{(j)}\left(x_l^{(j)-}\right), F_i^{(j)}\left(x_l^{(j)}\right)\right)$ can be taken instead of the realization $F_i^{(j)}\left(x_l^{(j)}\right)$ where $F_i^{(j)}\left(x_l^{(j)-}\right)$ denotes the left-hand limit of $F_i^{(j)}$, and $\eta_l^{(j)}$ is uniformly distributed on (0,1). If $y_l^{(j)}$ is a realization of $\eta_l^{(j)}$, then $y_l^{(j)}$ is taken instead of $\eta_l^{(j)}$. Moreover, $\xi_l^{(j)}$ and $\eta_l^{(j)}$ are assumed to be independent. If $H_i^{(j)}\left(y_l^{(j)},x_l^{(j)}\right) = F_i^{(j)}\left(x_l^{(j)-}\right)+y_l^{(j)}\left(F_i^{(j)}\left(x_l^{(j)}\right)-F_i^{(j)}\left(x_l^{(j)-}\right)\right)$, then $H_i^{(j)}\left(\eta_l^{(j)},\xi_l^{(j)}\right)$ is uniformly distributed on (0,1) [5]. Furthermore, it can be written

$$X_{r:n(j)}=\begin{cases}\left[H_i^{(j)}\right]^{-1}(1,1), & r>n(j)\\ X_{r:n(j)'} & r\le n(j)\end{cases}$$

where

$$\left[H_i^{(j)}\right]^{-1}(1,1)=\sup\left\{(x,y): H_i^{(j)}(x,y)<1\right\}.$$

Now, the joint p.d.f. of $X_{r_1:n}^{(j)},X_{r_2:n}^{(j)},\ldots,X_{r_p:n}^{(j)}$ and $\mathbf{X}_{r_1:n},\mathbf{X}_{r_2:n},\ldots,\mathbf{X}_{r_p:n}$ will be given in Theorem 3 and Theorem 4, respectively.

Theorem 3. *Let F_i be d.f. of random vector $\boldsymbol{\xi}_i$ and $F_i^{(j)}$ be d.f. of $\xi_i^{(j)}$ and discontinuous at any $x_l^{(j)}$, and $B_p^{(j)}$ denote the set of all $\left(x_1^{(j)},y_1^{(j)},x_2^{(j)},y_2^{(j)},\ldots,x_p^{(j)},y_p^{(j)}\right)$ satisfying $0<y_l^{(j)}<1$, and $x_{l-1}^{(j)}<x_l^{(j)}$ or $x_{l-1}^{(j)}=x_l^{(j)}$ and $y_{l-1}^{(j)}<y_l^{(j)}$. Then, the joint p.d.f. of $X_{r_1:n}^{(j)},X_{r_2:n}^{(j)},\ldots,X_{r_p:n}^{(j)}$ is*

$$f_{r_1,r_2,\ldots,r_p:n}^{(j)}\left(x_1^{(j)},x_2^{(j)},\ldots,x_p^{(j)}\right)=\int C\ \mathrm{per}\mathbf{A}\,dy_1^{(j)}dy_2^{(j)}\ldots dy_p^{(j)} \tag{5}$$

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

$$\Big[\ \underbrace{\mathbf{H}^{(j)}\left(y_1^{(j)},x_1^{(j)}\right)}_{r_1-1}\ \ \underbrace{\mathbf{1}_1^{(j)}}_{1}\ \ \underbrace{\mathbf{H}^{(j)}\left(y_2^{(j)},x_2^{(j)}\right)-\mathbf{H}^{(j)}\left(y_1^{(j)},x_1^{(j)}\right)}_{r_2-r_1-1}\ \ldots\underbrace{\mathbf{1}_p^{(j)}}_{1}\ \ \underbrace{\mathbf{1}-\mathbf{H}^{(j)}\left(y_p^{(j)},x_p^{(j)}\right)}_{n-r_p}\ \Big]$$

$\mathbf{H}^{(j)}\left(y_l^{(j)},x_l^{(j)}\right)=\left(H_1^{(j)}\left(y_l^{(j)},x_l^{(j)}\right) H_2^{(j)}\left(y_l^{(j)},x_l^{(j)}\right)\ \ldots\ H_n^{(j)}\left(y_l^{(j)},x_l^{(j)}\right)\right)', l=0,1,\ldots,p+1$,

and $\mathbf{1}=\left(1\ 1\ \ldots\ 1\right)'$, $\mathbf{1}_l^{(j)}=\left(1\ 1\ \ldots\ 1\right)'$, *$l=1,2,\ldots,p$ are column vectors with n components and* $C=\prod_{i=0}^{p}[(r_{i+1}-r_i-1)!]^{-1}$

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

Proof. Omitted.

Theorem 4. *Let F_i be d.f. of random vector $\boldsymbol{\xi}_i$ and $F_i^{(j)}$ be d.f. of $\xi_i^{(j)}$, and $F_i^{(j)}$ be discontinuous at $x_l^{(j)}$. Moreover, assume that the same conditions in Theorem 2.1 hold. Then, the joint p.d.f. of $\mathbf{X}_{r_1:n}$, $\mathbf{X}_{r_2:n}$,…, $\mathbf{X}_{r_p:n}$ is*

$$f_{r_1,r_2,\ldots,r_p:n}(\mathbf{x}_1,\ \mathbf{x}_2,\ldots,\ \mathbf{x}_p)=\prod_{j=1}^{d} f^{(j)}_{r_1,r_2,\ldots,r_p:n}\left(x_1^{(j)},x_2^{(j)},\ldots,x_p^{(j)}\right) \tag{6}$$

for $\bigcup_{j=1}^{d} B_p^{(j)}$ *and =0, otherwise, where* $\mathbf{x}_l=\left(x_l^{(1)},x_l^{(2)},\ldots,x_l^{(d)}\right)$ *and* $\mathbf{y}_l=\left(y_l^{(1)},y_l^{(2)},\ldots,y_l^{(d)}\right)$.

Proof. Omitted.

5. RESULTS (SONUÇLAR)

Result 1. In (3) (or in (4) for $d=1$), $f_i^{(j)}\left(x_l^{(j)}-1\right)=f^{(j)}\left(x_l^{(j)}\right)$, then the joint p.f. of order statistics of i.i.d. random variables from discrete parent is obtained. This result gives an alternative representation of (3) in [3], and the special cases of this result for $p=1$ and $p=2$ give (2) and (6) in [4], respectively.

Result 2. In (3) (or in (4) for $d=1$), if discrete case approaches to continuous case, i.e., $k_l=m_l=0$ and $F_i^{(j)}\left(x_l^{(j)}-1\right)=F_i^{(j)}\left(x_l^{(j)}\right)$ for $1\leq l\leq p$, then the joint p.d.f. of order statistics of i.n.n.i.d. random variables with continuous d.f.'s is obtained. This result gives the opened form of the expression which constructed by taking order statistics instead of *T* in (4) in [3], and the general case in [7].

Result 3. In (3) (or in (4) for $d=1$), if discrete case approaches to continuous case, i.e., $k_l=m_l=0$ and $F_i^{(j)}\left(x_l^{(j)}-1\right)=F_i^{(j)}\left(x_l^{(j)}\right)=F^{(j)}\left(x_l^{(j)}\right)$ for $1\leq l\leq p$, then the joint p.d.f. of order statistics of i.i.d. random variables with continuous d.f.'s is obtained. This result gives (2.2.3) in [2] and Theorem 1.4.5 in [5].

Result 4. In (5) (or in (6) for $d=1$), $H_i^{(j)}\left(y_l^{(j)},x_l^{(j)}\right)=H^{(j)}\left(y_l^{(j)},x_l^{(j)}\right)$, then the joint p.f. of order statistics of i.i.d. random variables under discontinuous d.f.'s is obtained. This result gives Corollary 1.5.7 in [5].

Result 5. In (5) (or in (6) for $d=1$), if $H_i^{(j)}\left(y_l^{(j)},x_l^{(j)}\right)=F_i^{(j)}\left(x_l^{(j)}\right)$ for $1\leq l\leq p$, 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 joint p.d.f. is obtained by taking $(f_1^{(j)}(x_l^{(j)})\ f_2^{(j)}(x_l^{(j)})\ldots f_n^{(j)}(x_l^{(j)}))'$ instead of $\mathbf{1}_l^{(j)}$ in per**A**. This result gives the opened form of the expression which constructed by taking order statistics instead of *T* in (2.2) in [3], and the general case in [5].

Result 6. In (3) (or in (4) for $d=1$), if $H_i^{(j)}\left(y_l^{(j)},x_l^{(j)}\right)=F^{(j)}\left(x_l^{(j)}\right)$, then the joint p.d.f. of order statistics of i.i.d. random variables with continuous d.f.'s is obtained. This result gives (2.2.3) in [2] and Theorem 1.4.5 in [5].

REFERENCES (KAYNAKLAR)

1. Barakat, H.M. and Abdelkader, Y.H., (2004). Computing the moments of order statistics from nonidentical random variables, Statistical Methods and Applications 13, pp:13-24.
2. David, H.A., (1981). Order Statistics, 2nd ed. Wiley, New York.
3. Guilbaud, O., (1982). Functions of non-i.i.d. random vectors expressed as functions of i.i.d. random vectors, Scand. J. Statist. 9, pp:229-233.
4. Khatri, C.G., (1962). Distributions of order statistics for discrete case, Annals of Institute of Statistical Mathematics 14, pp:167-171.
5. Reiss, R.D., (1989). Approximate Distributions of Order Statistics, Springer-Verlag, New York Inc..
6. Samuel, P. and Thomas, P.Y., (1998). Modified expression for a recurrence relation on the product moments of order statistics, Statistics and Probability Letters 37, pp:89-95.
7. Vaughan, R.J. and Venables, W.N., (1972). Permanent expressions for order statistic densities, J. Roy. Statist. Soc. Ser. B 34 pp:308-310.