www.istatistikciler.org
İstatistikçiler Dergisi 3 (2010) 63-68

İstatistikçiler Dergisi

# Distributions of order statistics arising from non-identical continuous variables

Mehmet Güngör
*Department of Econometrics*
*Inonu University*
*mgungor44@gmail.com*

Fahrettin Özbey
*Department of Statistics*
*Bitlis Eren University*
*fozbey2023@gmail.com*

**Abstract**

In this study, the probability density function (*pdf*) and the distribution function (*df*) of the *r*th order statistic arising from independent but not necessarily identically distributed (*innid*) continuous random variables are expressed. The results related to distributions of minimum and maximum order statistics of *innid* continuous random variables are given.

***Keywords:*** *Order statistics, Permanent, Continuous random variable, Probability density function, Distribution function.*

**Özet**

**Aynı dağılımlı olmayan sürekli değişkenlerin sıralı istatistiklerinin dağılımları**

*Bu çalışmada, bağımsız fakat aynı dağılımlı olmayan sürekli tesadüfi değişkenlerin r-inci sıralı istatistiğinin dağılım fonksiyonu ve olasılık yoğunluk fonksiyonu ifade edildi. Sonra, bağımsız fakat aynı dağılımlı olmayan sürekli tesadüfi değişkenlerin sıralı istatistiklerinin maksimum ve minimumunun dağılımlarıyla ilgili sonuçlar verildi.*

*Anahtar sözcükler: Sıralı istatistik, Permanent, Sürekli tesadüfi değişken, Olasılık yoğunluk fonksiyonu, Dağılım fonksiyonu.*

## 1. Introduction

Several identities and recurrence relations for the *pdf* and the *df* of order statistics of independent and identically distributed (*iid*) random variables were established by numerous authors including: Arnold et al.[1], Balasubramanian and Beg[3], David[13], and Reiss[18]. Furthermore; Arnold et al.[1], David[13], Gan and Bain[14], and Khatri[17] obtained the probability function and the *df* of order statistics of *iid* random variables from a discrete parent. Corley[11] defined a multivariate generalization of classical order statistics for random samples from a continuous multivariate distribution.

Expressions for generalized joint densities of the order statistics of the *iid* random variables in terms of Radon-Nikodym derivatives with respect to product measures based on *df* were derived by Goldie and Maller[15]. Guilbaud[16] expressed the probability of the functions of *innid* random vectors as a linear combination of probabilities of the functions of *iid* random vectors. This holds true for order statistics of random variables.

Recurrence relationships among the distribution functions of order statistics arising from *innid* random variables were obtained by Cao and West [9].

Vaughan and Venables[19] derived the joint *pdf* and marginal *pdf* of order statistics of *innid* random variables by means of permanents.

Balakrishnan[2], Bapat and Beg[7] obtained the joint *pdf* and *df* of order statistics of *innid* random variables by means of permanents. Using multinomial arguments, the *pdf* of $X_{r:n+1}$ $(1 \le r \le n+1)$ was obtained by Childs and Balakrishnan [10] by adding another independent random variable to the original *n* variables $X_1, X_2, ..., X_n$. Also, Balasubramanian et al.[6] established the identities which were satisfied by distributions of order statistics from non-independent non-identical variables through operator methods based on difference and differential operators.

In a paper published in 1991, Beg [8] obtained several recurrence relations and identities for product moments of order statistics of *innid* random variables using permanents. Recently, Cramer et al.[12] derived the expressions for the distribution and density functions by Ryser's method and the distributions of maxima and minima which was based on permanents.

In his first of two papers, Balasubramanian et al.[4] obtained the distribution of a single order statistic in terms of distribution functions of minimum and maximum order statistics using some subsets of $\{X_1, X_2, ..., X_n\}$ ( where $X_i$ 's are *innid* random variables). Later, Balasubramanian et al.[5] generalized their previous results[4] to the case of the joint distribution functions of several order statistics. Generally, the distribution theory for order statistics is complicated when the random variables are *innid*.

In this study, the *df* and *pdf* of order statistics of *innid* random variables are obtained.

Henceforth, 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.

If $\mathrm{a}_1, \mathrm{a}_2, ...$ are defined as column vectors, then the matrix obtained by taking $m_1$ copies of $\mathrm{a}_1$, $m_2$ copies of $\mathrm{a}_2, \ldots$ can be denoted by the following:

$$[\underset{m_1}{\mathrm{a}_1}\ \underset{m_2}{\mathrm{a}_2}\ \ldots]$$

and *per*A - which denotes the permanent of a square matrix A- is defined as similar to determinant except in a case where all terms in the expansion is accompanied by a positive sign.

Let $X_1, X_2, ..., X_n$ be *innid* continuous random variables and $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ be the order statistics obtained by arranging the *n* ( $X_i$ ' s) in an increasing order of magnitude.

Let $F_i$ and $f_i$ be *df* and *pdf* of $X_i$ (*i=1, 2,…, n*), respectively.

This paper is organized as follows. In section 2, we give the theorems concerning *df* and *pdf* of order statistics of *innid* continuous random variables. In the final section, some results related to *df* and *pdf* will be provided.

## 2. Theorems for distribution and probability density functions

In this section, the theorems related to *df* and *pdf* of $X_{r:n}$ will be provided. Here, we will express the following theorem for *df* of the *r*th order statistic of *innid* continuous random variables.

**Theorem 2.1.**

$$F_{r:n}(x) = \sum_{m=r}^{n} \frac{1}{m!(n-m)!} \sum_{t=m}^{n} (-1)^{n-t} \binom{n-m}{t-m} \sum_{n_s=n-t+m} (t-m)!\, per[\underset{n-t+m}{\mathrm{F}(x)}][s/.) , \tag{2.1}$$

where $\mathrm{F}(x) = (F_1(x), F_2(x), ..., F_n(x))'$ is the column vector, $x \in R$. Here, $s$ is a non-empty subset of the integers $\{1, 2, ..., n\}$ with $n_s \geq 1$. $\mathrm{A}[s/.)$ is the matrix obtained from $\mathrm{A}$ by taking rows whose indices are in $s$.

**Proof.** It can be written

$$F_{r:n}(x) = P\{X_{r:n} \leq x\} . \tag{2.2}$$

(2.2) can be expressed as

$$F_{r:n}(x) = \sum_{m=r}^{n} \frac{1}{m!(n-m)!} per\mathrm{A} , \tag{2.3}$$

where $\mathrm{A} = [\underset{m}{\mathrm{F}(x)} \quad \underset{n-m}{1-\mathrm{F}(x)}]$ is the matrix and $1-\mathrm{F}(x) = (1-F_1(x), 1-F_2(x), ..., 1-F_n(x))'$.

Using the properties of permanent, we can write:

$$\begin{aligned} per\mathrm{A} &= per[\underset{m}{\mathrm{F}(x)} \quad \underset{n-m}{1-\mathrm{F}(x)}] \\ &= \sum_{t=0}^{n-m} (-1)^{n-m-t} \binom{n-m}{t} per[\underset{n-t}{\mathrm{F}(x)} \quad \underset{t}{1}\,] \\ &= \sum_{t=0}^{n-m} (-1)^{n-m-t} \binom{n-m}{t} \sum_{n_s=n-t} t!\, per[\underset{n-t}{\mathrm{F}(x)}][s/.) \\ &= \sum_{t=m}^{n} (-1)^{n-t} \binom{n-m}{t-m} \sum_{n_s=n-t+m} (t-m)!\, per[\underset{n-t+m}{\mathrm{F}(x)}][s/.) , \end{aligned} \tag{2.4}$$

where $1 = (1,1,...,1)'$. Using (2.4) in (2.3) and (2.1) is obtained.

**Theorem 2.2.**

$$F_{r:n}(x) = \sum_{m=r}^{n} \frac{1}{m!(n-m)!} \sum_{t=m}^{n} (-1)^{n-t} \binom{n-m}{t-m} \sum_{n_s=n-t+m} (t-m)!\, (n-t+m)! \prod_{l=1}^{n-t+m} F_{s^l}(x) , \tag{2.5}$$

where $s = \{s^1, s^2, ..., s^{n-t+m}\}$.

**Proof.** Omitted.

**Theorem 2.3.**

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!}\sum_{t=r}^{n}(-1)^{n-t}\binom{n-r}{t-r}\sum_{n_{\varsigma}=n+r-t}(t-r)!\sum_{n_{\varsigma}=n+r-1-t} per[\underset{n+r-1-t}{\mathrm{F}(x)}][\varsigma/.)\, per[\underset{1}{\mathrm{f}(x)}][\varsigma'/.)\,, \tag{2.6}$$

where $\mathrm{f}(x) = (f_1(x), f_2(x),...,f_n(x))'$, $\varsigma = \varsigma \cup \varsigma'$, $\varsigma \cap \varsigma' = \phi$ and $n_{\varsigma'} = 1$.

**Proof.** Consider

$$P\{x < X_{r:n} \le x + \delta x\}\,. \tag{2.7}$$

Dividing (2.7) by $\delta x$ and then letting $\delta x$ tend to zero, we obtain

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!}\, per\mathrm{B}\,, \tag{2.8}$$

where $\mathrm{B} = [\underset{r-1}{\mathrm{F}(x)} \quad \underset{1}{\mathrm{f}(x)} \quad \underset{n-r}{1-\mathrm{F}(x)}]$ is the matrix. Using properties of permanent, we can write:

$$\begin{aligned} per\mathrm{B} &= per[\underset{r-1}{\mathrm{F}(x)} \quad \underset{1}{\mathrm{f}(x)} \quad \underset{n-r}{1-\mathrm{F}(x)}] \\ &= \sum_{t=0}^{n-r}(-1)^{n-r-t}\binom{n-r}{t} per[\underset{n-1-t}{\mathrm{F}(x)} \quad \underset{1}{\mathrm{f(x)}} \quad \underset{t}{1}] \\ &= \sum_{t=0}^{n-r}(-1)^{n-r-t}\binom{n-r}{t}\sum_{n_{\varsigma}=n-t} t!\,\mathrm{per}[\underset{\mathrm{n-1-t}}{\mathrm{F}(x)} \quad \underset{1}{\mathrm{f}(x)}][\varsigma/.) \\ &= \sum_{t=r}^{n}(-1)^{n-t}\binom{n-r}{t-r}\sum_{n_{\varsigma}=n+r-t}(t-r)!\,\mathrm{per}[\underset{\mathrm{n+r-1-t}}{\mathrm{F}(x)} \quad \underset{1}{\mathrm{f}(x)}][\varsigma/.) \\ &= \sum_{t=r}^{n}(-1)^{n-t}\binom{n-r}{t-r}\sum_{n_{\varsigma}=n+r-t}(t-r)!\sum_{n_{\varsigma}=n+r-1-t} per[\underset{n+r-1-t}{\mathrm{F}(x)}][\varsigma/.)\, per[\underset{1}{\mathrm{f}(x)}][\varsigma'/.)\,. \end{aligned} \tag{2.9}$$

Using (2.9) in (2.8), (2.6) is obtained.

**Theorem 2.4.**

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!}\sum_{t=r}^{n}(-1)^{n-t}\binom{n-r}{t-r}\sum_{n_{\mathrm{s}}=n+r-t}(t-r)!\sum_{n_{\varsigma}=n+r-1-t}(n+r-1-t)!\left(\prod_{l=1}^{n+r-1-t}F_{\varsigma^l}(x)\right)f_{\varsigma'^1}(x)\,, \tag{2.10}$$

where $\varsigma = \{\varsigma^1, \varsigma^2, ..., \varsigma^{n+r-1-t}\}$ and $\varsigma' = \{\varsigma'^1\}$.

**Proof.** Omitted.

**3. Results for distribution and probability density functions**

In this section, the results related to *df* and *pdf* of $X_{r:n}$ will be provided. We will now express the following result for *df* of minimum order statistic:

**Result 3.1.**

$$F_{1:n}(x) = 1 - \frac{1}{n!}\sum_{t=0}^{n}(-1)^{n-t}\binom{n}{t}\sum_{n_s=n-t} t!\, per\underset{n-t}{[\mathrm{F}(x)]}[s/.)$$

$$= 1 - \frac{1}{n!}\sum_{t=0}^{n}(-1)^{n-t}\binom{n}{t}\sum_{n_s=n-t} t!\,(n-t)!\prod_{l=1}^{n-t}F_{s^l}(x)\,. \tag{3.1}$$

**Proof.** In section (2.1) and (2.5), if $r = 1$, (3.1) is obtained.

We will express the following result for *df* of maximum order statistic.

**Result 3.2.**

$$F_{n:n}(x) = \frac{1}{n!}\, per\underset{n}{[\mathrm{F}(x)]}$$

$$= \prod_{l=1}^{n}F_l(x)\,. \tag{3.2}$$

**Proof.** In (2.1) and (2.5), if $r = n$, (3.2) is obtained.

In the following result, we will express *pdf* of minimum order statistic.

**Result 3.3.**

$$f_{1:n}(x) = \frac{1}{(n-1)!}\sum_{t=1}^{n}(-1)^{n-t}\binom{n-1}{t-1}\sum_{n_s=n+1-t}(t-1)!\sum_{n_\varsigma=n-t} per\underset{n-t}{[\mathrm{F}(x)]}[\varsigma/.)\; per\underset{1}{[\mathrm{f}(x)]}[\varsigma'/.)$$

$$= \frac{1}{(n-1)!}\sum_{t=1}^{n}(-1)^{n-t}\binom{n-1}{t-1}\sum_{n_s=n+1-t}(t-1)!\sum_{n_\varsigma=n-t}(n-t)!\left(\prod_{l=1}^{n-t}F_{\varsigma^l}(x)\right)f_{\varsigma'^1}(x)\,. \tag{3.3}$$

**Proof.** In (2.6) and (2.10), if $r = 1$, (3.3) is obtained.

We will express the following result for *pdf* of maximum order statistic.

**Result 3.4.**

$$f_{n:n}(x) = \frac{1}{(n-1)!}\sum_{n_\varsigma=n-1} per\underset{n-1}{[\mathrm{F}(x)]}[\varsigma/.)\; per\underset{1}{[\mathrm{f}(x)]}[\varsigma'/.)$$

$$= \sum_{n_\varsigma = n-1} \left( \prod_{l=1}^{n-1} F_{\varsigma^l}(x) \right) f_{\varsigma'^1}(x) \,. \tag{3.4}$$

**Proof.** In (2.6) and (2.10), if $r = n$, (3.4) is obtained.

## References


[1] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, 1992, *A first course in order statistics*, John Wiley and Sons Inc., New York.

[2] N. Balakrishnan, 2007, Permanents, order statistics, outliers and robustness, *Rev. Mat. Complut*. 20, no.1, 7-107.

[3] K. Balasubramanian and M. I. Beg, 2003, On special linear identities for order statistics, *Statistics* 37, no.4, 335-339.

[4] K. Balasubramanian, M. I. Beg and R. B. Bapat, 1991, On families of distributions closed under extrema, *Sankhyā* Ser. A 53, no.3, 375-388.

[5] K. Balasubramanian, M. I. Beg and R. B. Bapat, 1996, An identity for the joint distribution of order statistics and its applications, *J. Statist. Plann. Inference* 55, no.1, 13-21.

[6] K. Balasubramanian, N. Balakrishnan and H. J. Malik, 1994, Identities for order statistics from non-independent non- identical variables, *Sankhyā* Ser. B 56, no.1, 67-75.

[7] R. B. Bapat and M. I. Beg, 1989, Order statistics for nonidentically distributed variables and permanents, *Sankhyā* Ser. A 51, no.1, 79-93.

[8] M. I. Beg, 1991, Recurrence relations and identities for product moments of order statistics corresponding to nonidentically distributed variables, *Sankhyā* Ser. A 53, no.3, 365-374.

[9] G. Cao and M. West, 1997, Computing distributions of order statistics, *Communications in Statistics Theory and Methods* 26, no.3, 755-764.

[10] A. Childs and N. Balakrishnan, 2006, Relations for order statistics from non-identical logistic random variables and assessment of the effect of multiple outliers on bias of linear estimators, *Journal of Statistical Planning and Inference* 136, no.7, 2227-2253.

[11] H. W. Corley, 1984, Multivariate order statistics, *Commun. Statist.- Theor. Meth.* 13, no.10, 1299-1304.

[12] E. Cramer, K. Herle and N. Balakrishnan, 2009, Permanent Expansions and Distributions of Order Statistics in the INID Case, *Communications in Statistics - Theory and Methods* 38, no.12, 2078-2088.

[13] H. A. David, 1981, *Order statistics*, John Wiley and Sons Inc., New York.

[14] G. Gan and L. J. Bain, 1995, Distribution of order statistics for discrete parents with applications to censored sampling, *J. Statist. Plann. Inference* 44, no.1, 37-46.

[15] C. M. Goldie and R. A. Maller, 1999, Generalized densities of order statistics, *Statistica Neerlandica* 53, no.2, 222-246.

[16] O. Guilbaud, 1982, Functions of non-i.i.d. random vectors expressed as functions of i.i.d. random vectors, *Scand. J. Statist*. 9, no.4, 229-233.

[17] C. G. Khatri, 1962, Distributions of order statistics for discrete case, *Annals of the Ins. of Stat. Math.* 14, no.1, 167-171.

[18] R. -D. Reiss, 1989, *Approximate distributions of order statistics*, Springer-Verlag, New York.

[19] R. J. Vaughan and W. N. Venables, 1972, Permanent expressions for order statistics densities, *Journal of the Royal Statistical Society* Ser. B 34, no.2, 308-310.