Selçuk J. Appl. Math.
Vol. 9. No. 2. pp. 19−27, 2008

Selçuk Journal of Applied Mathematics

Characterization of Distributions by Using The Conditional Expectations of Generalized Order Statistics

Tuğba Yıldız[1] and Ismihan Bairamov[2]

[1]Department of Statistics, Faculty of Art and Sciences, Dokuz Eylul University, 35160, Buca, Izmir, Turkey
e-mail: tugba.ozkal@deu.edu.tr
[2]Department of Mathematics, Faculty of Science and Literature, Izmir University of Economics, 35330, Balcova, Izmir, Turkey
e-mail: ismihan.bayramoglu@ieu.edu.tr

Abstract. In this study, some continuous distributions through the properties of conditional expectations of generalized order statistics are characterized. Let $X_{1:n:m:k}, ..., X_{n:n:m:k}$ be the generalized order statistics, where $n \in \mathbb{N}, k > 0, m_1, ..., m_{n-1} \in \mathbb{R}$, $M_r = \sum_{j=r}^{n-1} m_j, 1 \leq r \leq n-1, \gamma_r = k+n-r+M_r > 0$ for all $r \in \{1, ..., n-1\}$ and let $\mathbf{m} = \{m_1, ..., m_{n-1}\}$, if $n \geq 2$, $m \in \mathbb{R}$, arbitrary, if $n = 1$. Characterization theorems for a general class of distributions are presented in terms of the function $E\{g(X_{j:n:m:m+1}) \mid X_{j-p:n:m:m+1} = x, X_{j+q:n:m:m+1} = y\} = A(x, y)$, where $k = m+1$, p and q are positive integers such that $p+1 \leq j \leq n-q$ and $g(.)$, $A(., .)$ is a real valued function satisfying certain regularity conditions.

Keywords: Order statistics, progressively Type II censored order statistics, generalized order statistics, characterization of distributions.
2000 Mathematics Subject Classification 62G30, 46N30

1. Introduction

Generalized order statistics have been introduced by Kamps (1995a, 1995b) to unify several models of ordered random variables, e.g. ordinary order statistics, progressively Type II censored order statistics, records and sequential order statistics. The common approach makes it possible to deduce several distributional and moment properties at once. The structural similarities of these models are based on the similarity of their joint probability density functions. These models can be effectively applied in reliability theory and survival analysis.

Let $X_1, X_2, ..., X_n$ be independent random variables with a common absolutely continuous distribution function (d.f.) $F(x)$ and a probability density function (p.d.f.) $f(x)$. Let $X_{1:n}, X_{2:n}, ..., X_{n:n}$ be the corresponding order statistics. Characterizations of $F(x)$ based on the properties of conditional expectations with the constant sample size n have been discussed by many authors.

Zoroa and Ruiz (1985) give a characterization results through the right censored mean function $m_R(x) = E(X \mid X \leq x)$ obtaining the explicit expression of a general distribution $F(x)$ from $m_R(x)$.

If $\mathcal{F}$ denotes the set of real continuous distribution functions, for each $F \in \mathcal{F}$, considering the doubly truncated mean function $m(x, y)$ is given by (Ruiz and Navarro, 1996)

$$m(x,y) = E\left(X \mid x \leq X \leq y\right) = \frac{1}{F(y) - F(x)} \int_x^y t dF(t),$$

whose domain of definition is $D = \left\{(x, y) \in \mathbb{R}^2 \text{ such that } F(x) < F(y)\right\}$. In the doubly censored mean function, making the change $X = h(X)$ where $h(x)$ is a continuous and strictly monotonic function, then $m_h(x, y) = E\left(h\left(X\right) \mid x \leq X \leq y\right)$. They define the relationship between $m_h(x, y)$ and order statistics as

$$E\left(\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} h\left(X_{i:n}\right) \mid X_{r:n} = x, X_{s:n} = y\right) = E\left(h\left(X\right) \mid x \leq X \leq y\right),$$

for all $(x, y) \in D$, if $1 \leq r < s \leq n$, where $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ are the

order statistics from the sample of the random variable X. Balasubramanian and Beg (1992) presented similar results for a particular function $h(x)$. Gupta et al. (1993) consider the case when $r = 1$ and $s = n$.

In this work, we present characterization theorems for some continuous distributions by using the properties of conditional expectations of generalized order statistics.

a_X and b_X are the left and right extremities of $F(x)$, respectively, denoted by $a_X = \inf\{x : F(x) > 0\}$ and $b_X = \sup\{x : F(x) < 1\}$, where $F(x)$ is a d.f. of a random variable X. Throughout the paper we assume that $F(x)$ is absolutely continuous and strictly increasing d.f. Our characterization results are based on the following theorem.

Theorem 1.1 (Bairamov and Özkal, 2007). Let $h(x)$ be a differentiable real valued function on $[0, 1]$ and the condition

$$h^{'}(y) \neq \frac{h(y)-h(x)}{y-x},$$

is valid for all $0 < x < y < 1$. Furthermore, let $G(x)$ be an absolutely continuous and strictly increasing d.f. with support $[a_X, b_X]$. Then $F(x) = G(x)$ if and only if the representation

$$E\left\{h^{'}(G(X)) \mid x \leq G(X) \leq y\right\} = \frac{h(y)-h(x)}{y-x},$$

is valid for all $0 < x < y < 1$.

Denote by $X_{1:n}, ..., X_{n:n}$ the ordinary order statistics based on the sample $X_1, X_2, ..., X_n$. The following fact shows that the conditional distribution of $X_{j:n}$ given $X_{j-p:n} = x$ and $X_{j+q:n} = y$ does not depend on the sample size n . It depends on x, y and $F(x)$.

Theorem 1.2 (Bairamov and Özkal, 2007). It is true that if $p+1 \leq j \leq n-q$, then

$$(X_{j:n} \mid X_{j-p:n} = x, X_{j+q:n} = y) \stackrel{d}{=} Z_{p:p+q-1},$$

where

$$Z \stackrel{d}{=} X \mid x < X < y.$$

We use Theorem 1.1 and Theorem 1.2 to characterize several distributions by properties of conditional expectations of order statistic $X_{j:n}$ given $X_{j-p:n} = x$ and $X_{j+q:n} = y$.

2. Characterizations Through The Conditional Expectations of Generalized Order Statistics

Assume that the d.f. $F(x)$ of the random variable X is absolutely continuous and strictly increasing on (a_F, b_F), where $a_F = \inf\{x : F(x) > 0\}$ and $b_F = \sup\{x : F(x) < 1\}$.

Theorem 2.1 (Bairamov and Özkal, 2007). Let $G(x)$ be an absolutely continuous and strictly increasing d.f. and left and right extremities of $G(x)$ be

$a_G = a_F$ and $b_G = b_F$, respectively. Then, $F(x) = G(x)$ if and only if the representation

$$\begin{aligned} & \frac{1}{s}\sum_{p=1}^{s} E\left[h^{'}(G(X_{j:n})) \ \mid \ X_{j-p:n} = x, X_{j+s+1-p:n} = y\right] \\ = \ & \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}, \end{aligned}$$

holds for all $a_F < x < y < b_F$, where $h(x)$ satisfies conditions of Theorem 1.1. and j, n, s are fixed integers such that $s + 1 \leq j \leq n - s$.

Proof Taking into account the results of Theorem 1.2 we have

$$\begin{aligned} & \frac{1}{k}\sum_{p=1}^{k} E\left[h^{'}(G(X_{j:n})) \ \mid \ X_{j-p:n} = x, X_{j+k+1-p:n} = y\right] \\ = \ & E\left[\frac{1}{k}\sum_{p=1}^{k} h'\left(G\left(Z_{p:k}\right)\right)\right] = E\left[\frac{1}{k}\sum_{i=1}^{k} h'\left(G\left(Z_{i}\right)\right)\right] \\ = \ & \frac{1}{k} kE\left[\ h'\left(G\left(Z\right)\right)\right] = E\left[h'\left(G\left(Z\right)\right)\right] \\ = \ & E\ \left[h'\left(G\left(X\right)\right) \mid x \leq X \leq y\right]. \end{aligned}$$

Then the proof is completed by using Theorem 1.1.◀

Let $X_1, X_2, ..., X_n$ be continuous independent and identically distributed (i.i.d.) random variables with d.f. $F(x)$ and p.d.f. $f(x)$. Let $X_{1:n}, ..., X_{n:n}$ be the corresponding order statistics. $X_{1:n:m:k}$, $X_{2:n:m:k}$, ..., $X_{n:n:m:k}$ is denoted by generalized order statistics, where $n \in \mathbb{N}, k > 0, m_1, ..., m_{n-1} \in \mathbb{R}$, $M_r = \sum_{j=r}^{n-1} m_j, 1 \leq r \leq n-1$, $\gamma_r = k + n - r + M_r > 0$ for all $r \in \{1, ..., n-1\}$ and let $\mathbf{m} = \{m_1, ..., m_{n-1}\}$,if $n \geq 2$, $m \in \mathbb{R}$, arbitrary, if $n = 1$. In the case when $m_i = m$ and $k = m + 1$, the joint distribution of generalized order statistics can be reduced to the joint distribution of usual order statistics of a sample size n from a continuous random variable.

Theorem 2.2 Let $X_{1:n:m:k}, X_{2:n:m:k}, ..., X_{n:n:m:k}$ be the generalized order statistics, with underlying continuous d.f. $F(x)$ and p.d.f. $f(x)$. Then in the special case when $m_i = m$ and $k = m + 1$ we have

$$(X_{1:n:m:m+1},...,X_{n:n:m:m+1}) \overset{d}{=} (Y_{1:n},...,Y_{n:n}),$$

where $Y_1, Y_2, ..., Y_n$ sample with d.f. $P(x) = 1-(1-F(x))^{m+1}$.

Proof Let $X(1,n,m,k),...,X(n,n,m,k)$ possess the joint probability density function of the form

$$f^{X(1,n,m,k),...,X(n,n,m,k)}(x_1,...,x_n) = k\left(\prod_{j=1}^{n-1}\gamma_j\right)\left(\prod_{i=1}^{n-1}(1-F(x_i))^{m_i}f(x_i)\right)(1-F(x_n))^{k-1}f(x_n).$$

In the case $m_i = m$, $M_1 = (m_1 + ... + m_{n-1}) = (n-1)m$, $M_2 = (m_2 + ... + m_{n-1}) = (n-2)m, ..., M_{n-1} = m$ and $\gamma_1 = k+n-1+(n-1)m$, $\gamma_2 = k+n-2+(n-2)m$, $\gamma_{n-1} = k+1+m$, $\gamma_n = k$, then we have

$= k(k+(n-(n-1))(m+1))(k+(n-(n-2))(m+1))...(k+(n-1)(m+1))\left(\prod_{i=1}^{n-1}(1-F(x_i))^m f(x_i)\right)(1-F(x_n))^{k-1}f(x_n)$

$= k(k+1(m+1))(k+2(m+1))...(k+(n-1)(m+1))\left(\prod_{i=1}^{n-1}(1-F(x_i))^{m_i}f(x_i)\right)(1-F(x_n))^{k-1}f(x_n)$

$= k^n(1+\frac{m+1}{k})(1+2\frac{m+1}{k})...(1+(n-1)\frac{m+1}{k})$
$\times\left(\prod_{i=1}^{n-1}(1-F(x_i))^{m_i}f(x_i)\right)(1-F(x_n))^{k-1}f(x_n)$

if $k = m+1$,

$f^{X(1,n,m,m+1),...,X(n,n,m,m+1)}(x_1,...,x_n) = n!(m+1)^n\left(\prod_{i=1}^{n-1}(1-F(x_i))^m f(x_i)\right)(1-F(x_n))^m f(x_n),$

then we obtain the joint p.d.f. of ordinary order statistics with sample size n,

$$f_{1,2,...,n}(x_1,...,x_n) = n!(m+1)^n\prod_{i=1}^{n}(1-F(x_i))^m f(x_i)$$

$= n!(m+1)^n(1-F(x_1))^m(1-F(x_2))^m...(1-F(x_n))^m f(x_1)...f(x_n)$

$= n!g(x_1)...g(x_n)$,

where $g(x) = (m+1) f(x) (1 - F(x))^m$, $G(x) = 1 - (1 - F(x))^{m+1}$.

Due to Theorem 2.2, any property of ordinary order statistics can be easily transformed to the corresponding property of generalized order statistics.

Theorem 2.3 Assume that the random variable X has absolutely continuous and strictly increasing d.f. $F(x)$ with left and right extremities a_F and b_F, respectively. Let $G(x)$ be also an absolutely continuous and strictly increasing d.f. with left and right extremities $a_G = a_F$ and $b_G = b_F$. Then $F(x) = 1 - (1 - G(x))^{1/(m+1)}$ if and only if the representation

$$\begin{aligned} & E\left\{ \frac{1}{s} \sum_{p=1}^{s} h^{'}(G(X_{j:n:m:m+1})) \mid X_{j-p:n:m:m+1} = x, X_{j+s+1-p:n:m:m+1} = y \right\} \\ = \quad & \frac{h(G(y)) - h(G(x))}{G(y) - G(x)}, \end{aligned}$$

holds for all $a_X < x < y < b_X$. The number j, n, s are fixed and satisfies the condition $s + 1 \leq j \leq n - s$. Proof of Theorem 2.3 follows easily from Theorem 2.1 and Theorem 2.2.

2.1. Characterization for Uniform Distribution

X is absolutely continuous random variable with strictly increasing d.f. having support $[a, b] = [0, 1]$ has Uniform distribution over if and only if the representation

$$\begin{aligned} & \frac{1}{s} \sum_{p=1}^{s} E\left[X_{j:n:m:m+1} \mid X_{j-p:n:m:m+1} = x, X_{j+s+1-p:n:m:m+1} = y\right] \\ = \quad & 1 - \frac{m+1}{m+2} \frac{a(x, y; m+2)}{a(x, y; m+1)}, \; s + 1 \leq j \leq n - s, \end{aligned}$$

holds for all $0 < x < y < 1$, where $a(x, y; m+1) = (1-x)^{m+1} - (1-y)^{m+1}$. The result follows from the Theorem 2.3 by a choice of $h(x) = x + \frac{m+1}{m+2}(1 - x)^{(m+2)/(m+1)} - \frac{m+1}{m+2}$, $h^{'}(x) = 1 - (1-x)^{1/(m+1)}$, $G(x) = 1 - (1-x)^{m+1}$.

2.2. Characterization for Weibull Distribution

The absolutely continuous random variable X strictly increasing d.f. having support $[0,\infty)$ has Weibull distribution $F(x)=1-\exp(-\alpha x^{\beta})$, $x\geq 0$, $\alpha>0$, $\beta>0$ if and only if the representation

$$
\begin{aligned}
&\frac{1}{s}\sum_{p=1}^{s} E\left[X_{j:n:m:m+1} \mid X_{j-p:n:m:m+1}=x, X_{j+s+1-p:n:m:m+1}=y\right] \\
=\quad & \frac{1}{m+1}+\frac{\alpha\left[x^{\beta}(\exp(-\alpha x^{\beta}))^{m+1}-y^{\beta}(\exp(-\alpha y^{\beta}))^{m+1}\right]}{(\exp(-\alpha x^{\beta}))^{m+1}-(\exp(-\alpha y^{\beta}))^{m+1}} \\
& .
\end{aligned}
$$

holds for all $0\leq x<y<\infty$, we take $h(x)=-\ln(1-x)^{1/(m+1)}x+\frac{x}{m+1}+\frac{\ln(1-x)}{m+1}-\frac{1}{m+1}$, $h^{'}(x)=-\ln(1-x)^{1/(m+1)}$and $G(x)=1-(\exp(-\alpha x^{\beta}))^{m+1}$.

In special case when $\beta=1$, X has exponential distribution. Then the following characterization for the exponential distribution can be given X has Exponential distribution $F(x)=1-\exp(-\alpha x)$, $x\geq 0$ if and only if the representation

$$
\begin{aligned}
&\frac{1}{s}\sum_{p=1}^{s} E\left[X_{j:n:m:m+1} \mid X_{j-p:n:m:m+1}=x, X_{j+s+1-p:n:m:m+1}=y\right] \\
=\quad & \frac{1}{m+1}+\frac{\alpha\left[x(\exp(-\alpha x))^{m+1}-y(\exp(-\alpha y))^{m+1}\right]}{(\exp(-\alpha x))^{m+1}-(\exp(-\alpha y))^{m+1}}
\end{aligned}
$$

holds for all $0\leq x<y<\infty$ and $G(x)=1-(\exp(-\alpha x))^{m+1}$.

2.3. Characterization for Generalized Beta Distribution

The absolutely continuous random variable X strictly increasing d.f. having support $[a,b)$ has Generalized Beta distribution $F(x)=1-\frac{(b-x)^{\theta}}{(b-a)^{\theta}}$, $\theta>0$, $-\infty<a<b<\infty$, if and only if the representation

$$
\begin{aligned}
&\frac{1}{s}\sum_{p=1}^{s} E\left[X_{j:n:m:m+1} \mid X_{j-p:n:m:m+1}=x, X_{j+s+1-p:n:m:m+1}=y\right] \\
=\quad & b+\frac{\theta(m+1)(a-b)}{\theta(m+1)+1}\frac{b_1(x,y;\theta+1)}{b_1(x,y;\theta)},
\end{aligned}
$$

holds for all $a \leq x < y < b$, where $b_1(x,y;\theta) = \left(\frac{b-x}{b-a}\right)^{\theta(m+1)} - \left(\frac{b-y}{b-a}\right)^{\theta(m+1)}$. The result is obtained by using target function $h(x) = bx - \frac{(a-b)(1-x)^{1+1/(\theta(m+1))}}{1/(\theta(m+1))+1}$ and $h^{'}(x) = b - (b-a)(1-x)^{1/(\theta(m+1))}$ and $G(x) = 1 - \frac{(b-x)^{\theta(m+1)}}{(b-a)^{\theta(m+1)}}$.

In special case $a \to 0$, $b \to 1$, we have

$$\begin{aligned} &\frac{1}{s}\sum_{p=1}^{s} E\left[X_{j:n:m:m+1} \mid X_{j-p:n:m:m+1} = x, X_{j+s+1-p:n:m:m+1} = y\right] \\ = \quad & 1 - \frac{\theta(m+1)}{\theta(m+1)+1}\frac{b_2(x,y;\theta+1)}{b_2(x,y;\theta)}, \end{aligned}$$

where $b_2(x,y;\theta) = (1-x)^{\theta(m+1)} - (1-y)^{\theta(m+1)}$. Here the target function $h(x) = x + \frac{(1-x)^{1+1/(\theta(m+1))}}{1+1/(\theta(m+1))}$, $h^{'}(x) = 1 - (1-x)^{1/(\theta(m+1))}$ and $G(x) = 1 - (1-x)^{\theta(m+1)}$.

2.4. Characterization for Pareto Distribution

X is absolutely continuous random variable with strictly increasing d.f. having support $[\gamma, \infty)$ has Pareto distribution $F(x) = 1 - \frac{(\gamma+\delta)^{\theta}}{(x+\delta)^{\theta}}$, $x \geq \gamma$, $\theta > 0$,$\gamma + \delta > 0$, if and only if the representation

$$\begin{aligned} &\frac{1}{s}\sum_{p=1}^{s} E\left[X_{j:n:m:m+1} \mid X_{j-p:n:m:m+1} = x, X_{j+s+1-p:n:m:m+1} = y\right] \\ = \quad & -\frac{\theta(m+1)}{\theta(m+1)-1}\frac{c(x,y;\theta-1)}{c(x,y;\theta)} - \delta, \end{aligned}$$

holds for all $\gamma \leq x < y < \infty$, where $c(x,y;\theta) = \left(\frac{1}{x+\delta}\right)^{\theta(m+1)} - \left(\frac{1}{y+\delta}\right)^{\theta(m+1)}$. The target function $h(x) = \frac{(\gamma+\delta)}{(-1/(\theta(m+1)))+1}(1-x)^{(-1/(\theta(m+1)))+1} - \delta x$, $h^{'}(x) = (\gamma+\delta)(1-x)^{-1/(\theta(m+1))} - \delta$ and $G(x) = 1 - \frac{(\gamma+\delta)^{\theta(m+1)}}{(x+\delta)^{\theta(m+1)}}$.

References

1. Bairamov, I. and Özkal, T., On characterization of distributions through the properties of conditional expectations of order statistics, Communications in Statistics: Theory and Methods, 36, 2007, p:1319-1326.
2. Balakrishnan, N. and Aggarwala, R., Progressive Censoring: Theory, Methods and Application, Boston: Birkhauser, 2000.

3. Balasubramanian, K. and Beg, M. I., Distributions determined by conditioning on a pair of Order Statistics, Metrika, 39, 1992, p:107-112.
4. Gupta, S. D., Goswami, A. and Rao, B. V., On a characterization of Uniform distribution, Journal of Multivariate Analysis, 44,1993, p:102-114.
5. Kamps, U., A Concept of Generalized Order Statistics, Teubner, Stuttgart, 1995a.
6. Kamps, U., A concept of generalized order statistics, Journal of Statistical Planning and Inference, 48,1995b,p:1-23.
7. Kotlarski, I.I., On a characterization of some probability distributions by conditional expectations, Sankhyã, Ser. A., 34, 1972, p:461-466.
8. Navarro, J., Ruiz, J. M. and Zoroa, N., A unified approach characterization problems using conditional expectations, Journal of Statistical Planning and Inference, 69, 1998, p:193-207.
9. Ruiz, J.M., Navarro, J., Characterizations based on conditional expectations of the doubled truncated distribution, Ann. Inst. Statist. Math., 48,1996,p:563-572.
10. Shanbhag, D.N., Bhaskara Rao, M., A note on characterizations of probability distributions based on conditional expected values, Sankhyã, Ser. A., 37, 1975, p:297-300.
11. Zoroa, P., Ruiz, J.M., El producto integral de Riemann-Stieltjes y su aplicacion alas funciones de medias de las distribuciones truncadas, Revista de la Real Academia de las Ciencias Exactas, Fisicas y Naturales, 79, 1985, p:39-51.