Marquette University
e-Publications@Marquette

Mathematics, Statistics and Computer Science Faculty Research and Publications

Mathematics, Statistics and Computer Science, Department of

1-1-2015

Characterizations of Distributions via Conditional Expectation of Generalized Order Statistics

M. Ahsanullah
Rider University-Lawrenceville

Gholamhossein Hamedani
Marquette University, gholamhoss.hamedani@marquette.edu

Mehdi Maadooliat
Marquette University, mehdi.maadooliat@marquette.edu

Characterizations of Distributions via Conditional Expectation of Generalized Order Statistics

M. Ahsanullah[1], G.G. Hamedani[2] and M. Maadooliat[2]

[1] Department of Management Sciences, Rider University, Lawrenceville, NJ, USA

[2] Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI, USA

Correspondence: M. Maadooliat, Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI 53201-1881, USA. Tel: 1-414-288-6341. E-mail: mehdi@mscs.mu.edu

Abstract

Characterizations of probability distributions by different regression conditions on generalized order statistics has attracted the attention of many researchers. We present here, characterizations of certain continuous distributions based on the conditional expectation of generalized order statistics.

Mathematics Subject Classification: 62E10; 60E05; 60E15

Keywords: generalized order statistics, probability distribution

1. Introduction

Kamps (1995) introduced generalized order statistics (gos) in terms of their joint density function. The record values, k-record values, order statistics are special cases of the gos. The random variables $X(1,n,m,k)$, $X(2,n,m,k)$,..., $X(n,n,m,k)$, $k > 0$, $m \in \mathbb{R}$, are n gos from an absolutely continuous distribution F with density f if their joint density, $f_{1,2,...,n}(x_1, x_2, ..., x_n)$, is given by

$$f_{1,2,...,n}(x_1, x_2, ..., x_n) = k\left(\Pi_{j=1}^{n-1}\gamma_j\right)\left[\Pi_{j=1}^{n-1}\left(\overline{F}(x_j)\right)^m f(x_j)\right] \times$$

$$\left(\overline{F}(x_n)\right)^{k-1} f(x_n), F^{-1}(0+) < x_1 < x_2 < ... < x_n < F^{-1}(1-), \tag{1.1}$$

where '$x+$'('$x-$') stands for the right(left)-sided limit at x, $\overline{F}(x) = 1 - F(x)$ and $\gamma_j = k + (n-j)(m+1)$ for all j, $1 \le j \le n$, k is a positive integer and $m \ge -1$.

For $k = 1$ and $m = 0$, $X(r,n,m,k)$ will be the rth order statistic and for $k = 1$ and $m = -1$, it will be the upper record value of the $i.i.d.$ random variables with common distribution F and density f.

From (1.1) the density of $X(r,n,m,k)$, $f_{r,n,m,k}$, is

$$f_{r,n,m,k}(x) = \frac{c_r}{\Gamma(r)}\left(\overline{F}(x)\right)^{\gamma_r-1} f(x)\, h_m^{r-1}(F(x)), \tag{1.2}$$

where $c_r = \Pi_{j=1}^r \gamma_j$ and

$$h_m(x) = \frac{1}{m+1}\left[1 - (1-x)^{m+1}\right], \quad for\ all\ x \in (0,1)\ and\ all\ m\ with$$
$$h_{-1}(x) = \lim_{m\to-1} h_m(x) = -\ln(1-x).$$

The joint density of $X(r,n,m,k)$ and $X(r+2,n,m,k)$, $1 \le r \le n-2$, is (Kamps, 1995)

$$f_{r,r+2,n,m,k}(x,y) = \Pi_{i=1}^{r+2}\gamma_i\left(\sum_{i=r+1}^{r+2} a_i^{(r)}(r+2)\left(\frac{\overline{F}(y)}{\overline{F}(x)}\right)^{\gamma_i}\right)$$

$$\times\left(\sum_{i=1}^r a_i\left(\overline{F}(x)\right)^{\gamma_i-1} f(x)\right)\left(\frac{f(y)}{\overline{F}(y)}\right), \qquad x \le y,$$

where $a_i = a_i(r) = \Pi_{\substack{j=1 \\ j \neq i}}^{r} \frac{1}{\gamma_j - \gamma_i}$ and $a_i^{(r)}(s) = \Pi_{\substack{j=r+1 \\ j \neq i}}^{s} \frac{1}{\gamma_j - \gamma_i}$, $r + 1 \leq i \leq r + 2$ as borrowed from Bieniek and Szynal (2003).

Consequently, the conditional density of $X(r+2, n, m, k)$ given $X(r, n, m, k) = x$, for $m \geq -1$, is

$$f_{r+2|\,r,n,m,k}(y|x) = \frac{\gamma_{r+1}\gamma_{r+2}}{(m+1)\,[1 - F(x)]^{\gamma_{r+1}}} f(y) \left[\overline{F}(y)\right]^{\gamma_{r+2}-1} \times$$

$$\left\{ \left[\overline{F}(x)\right]^{m+1} - \left[\overline{F}(y)\right]^{m+1} \right\}, y > x. \tag{1.3}$$

The conditional density of $X(r, n, m, k)$ given $X(r+2, n, m, k) = y$, for $m \neq -1$, is

$$f_{r\,|\,r+2,n,m,k}(x|y) = r\,(r+1)\,(m+1)\left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{r-1}\left\{1 - \left[\overline{F}(y)\right]^{m+1}\right\}^{-r-1} \times$$

$$\left\{ \left[\overline{F}(x)\right]^{m+1} - \left[\overline{F}(y)\right]^{m+1} \right\} f(x)\left[\overline{F}(x)\right]^{m} , \quad y > x. \tag{1.4}$$

Different regression conditions on generalized order statistics have been employed to characterize distributions (e.g., see Bieniek & Szynal, 2003; Bieniek, 2009; Cramer, Kamps & Keseling, 2004), to name a few. In Bieniek and Szynal (2003), the authors consider all distributions F for which the following linearity of regression holds:

$$E\,[\,(X\,(r+l, n, m, k))\,|\,X(r, n, m, k)] = a\,X(r, n, m, k) + b.$$

They conclude that only exponential, Pareto and power function distributions satisfy this equation. Using this result they obtain characterizations of these distributions based on sequential order statistics, records and progressive type II censored order statistics. In the following section, we consider a certain class of *gos* and present two characterization results.

2. Characterization Results

Cramer et al. (2004), point out that characterizations of distributions based on linear regressions

$$E\,[s\,(X\,(r+l, n, m, k))\,|\,X(r, n, m, k) = \cdot] = g\,(\cdot) \tag{2.1}$$

have been studied extensively for order statistics and record values ($r \in \mathbb{N}$, $l = 1$) and set up a comprehensive solution related to characterization problems. They solved the case of *gos* ($l = 1$) and pointed out that for larger l the calculations become more difficult. They were able to obtain an explicit result when g is a linear function. They concluded that the linearity of the conditional expectation provides a characterization of generalized Pareto distribution. Ahsanullah and Hamedani (2013) presented characterizations of continuous distributions based on (2.1) for $l = 1$ but without the assumptions of monotonicity of $s\,(\cdot)$ and linearity of $g\,(\cdot)$. In what follows we present our characterizations for $l = 2$ with the assumption of monotonicity of $s\,(\cdot)$ but without assumption of linearity of $g\,(\cdot)$.

Theorem 2.1. Let $X : \Omega \to [a, b)$ be a random variable with a twice differentiable distribution function $F(x)$ and with density function $f(x)$. Let $s(x)$ be a non-negative, non-decreasing and twice differentiable function on $[a, b)$ such that $s'(a) = f(a) > 0$, $s(a) = 0$ and $\lim_{x \to b} s(x) = \infty$. Then

$$E\,[s\,(X\,(r+2, n, m, k))\,|\,X(r, n, m, k) = x]$$
$$= \frac{1}{(\gamma_{r+1} + \gamma_{r+2} + 1 + \gamma_{r+1}\gamma_{r+2})}\,[\gamma_{r+1}\gamma_{r+2}\,s\,(x) - 1] , \tag{2.2}$$

implies

$$\overline{F}(x) = \left[(\gamma_{r+1} + \gamma_{r+2} + 1)\, s(x) + 1 \right]^{-1/(\gamma_{r+1} + \gamma_{r+2} + 1)}, \quad a \le x \le b. \tag{2.3}$$

Proof. Suppose (2.2) holds and let $A = \dfrac{\gamma_{r+1}\gamma_{r+2}}{(\gamma_{r+1} + \gamma_{r+2} + 1 + \gamma_{r+1}\gamma_{r+2})}$ and $B = \dfrac{-1}{(\gamma_{r+1} + \gamma_{r+2} + 1 + \gamma_{r+1}\gamma_{r+2})}$, then

$$\frac{\gamma_{r+1}\gamma_{r+2}}{m+1} \int_x^b s(y) f(y) \left[\overline{F}(y) \right]^{\gamma_{r+2} - 1} \left\{ \left[\overline{F}(x) \right]^{m+1} - \left[\overline{F}(y) \right]^{m+1} \right\} dy$$
$$= (As(x) + B) \left[\overline{F}(x) \right]^{\gamma_{r+1}}. \tag{2.4}$$

Differentiating both sides of (2.4) with respect to x, we have

$$-\gamma_{r+1}\gamma_{r+2}\, f(x) \left[\overline{F}(x) \right]^m \int_x^b s(y) f(y) \left[\overline{F}(y) \right]^{\gamma_{r+2} - 1} dy$$
$$= As'(x) \left[\overline{F}(x) \right]^{\gamma_{r+1}} - (As(x) + B)\, \gamma_{r+1} f(x) \left[\overline{F}(x) \right]^{\gamma_{r+1} - 1}.$$

Upon simplification, we get

$$\int_x^b s(y) f(y) \left[\overline{F}(y) \right]^{\gamma_{r+2} - 1} dy = -\frac{A}{\gamma_{r+1}\gamma_{r+2}} \left(\frac{s'(x)}{f(x)} \right) \left[\overline{F}(x) \right]^{\gamma_{r+2} + 1} +$$
$$\frac{1}{\gamma_{r+2}} (As(x) + B) \left[\overline{F}(x) \right]^{\gamma_{r+2}}. \tag{2.5}$$

Differentiating both sides of (2.5) with respect to x, we obtain

$$-s(x)\, f(x) \left[\overline{F}(x) \right]^{\gamma_{r+2} - 1} = -(As(x) + B)\, f(x) \left[\overline{F}(x) \right]^{\gamma_{r+2} - 1}$$
$$+ \frac{A(\gamma_{r+1} + \gamma_{r+2} + 1)}{\gamma_{r+1}\gamma_{r+2}} \left[\overline{F}(x) \right]^{\gamma_{r+2}} s'(x)$$
$$- \frac{A}{\gamma_{r+1}\gamma_{r+2}} \left\{ \frac{d}{dx} \left(\frac{s'(x)}{f(x)} \right) \right\} \left[\overline{F}(x) \right]^{\gamma_{r+2} + 1}. \tag{2.6}$$

Simplifying (2.6) results in

$$\frac{(1 - A)\, s(x) f(x)}{\left[\overline{F}(x) \right]^2} + \frac{A(\gamma_{r+1} + \gamma_{r+2} + 1)}{\gamma_{r+1}\gamma_{r+2}} \left[\frac{s'(x)}{\overline{F}(x)} \right]$$
$$- \frac{A}{\gamma_{r+1}\gamma_{r+2}} \left\{ \frac{d}{dx} \left(\frac{s'(x)}{f(x)} \right) \right\} - \frac{B f(x)}{\left[\overline{F}(x) \right]^2} = 0. \tag{2.7}$$

Substituting for A and B in (2.7) and regrouping terms we arrive at

$$(\gamma_{r+1} + \gamma_{r+2} + 1) \left\{ \frac{d}{dx} \left(\frac{s(x)}{\overline{F}(x)} \right) \right\} - \left\{ \frac{d}{dx} \left(\frac{s'(x)}{f(x)} \right) \right\}$$
$$+ \frac{f(x)}{\left[\overline{F}(x) \right]^2} = 0.$$

Integrating both sides of the above equation from a to x and using the assumptions $s'(a) = f(a) > 0$ and $s(a) = 0$, we obtain

$$(\gamma_{r+1} + \gamma_{r+2} + 1) \left(\frac{s(x)}{\overline{F}(x)} \right) - \left(\frac{s'(x)}{f(x)} \right) + \frac{1}{\left[\overline{F}(x) \right]} = 0,$$

or

$$\frac{f(x)}{\overline{F}(x)} = \frac{s'(x)}{(\gamma_{r+1} + \gamma_{r+2} + 1) s(x) + 1}. \qquad (2.8)$$

Integrating both sides of (2.8) from a to x and in view of the assumption $\lim_{x \to b} s(x) = \infty$, we obtain the distribution function

$$\overline{F}(x) = \left[(\gamma_{r+1} + \gamma_{r+2} + 1) s(x) + 1 \right]^{-1/(\gamma_{r+1} + \gamma_{r+2} + 1)}.$$

Remark 2.1. For different functions s satisfying the conditions of Theorem 2.1, one can obtain characterizations of various well-known distributions. We will mention some of these distributions below.

(1) For $s(x) = \frac{e^{\lambda(\gamma_{r+1} + \gamma_{r+2} + 1)x} - 1}{(\gamma_{r+1} + \gamma_{r+2} + 1)}$, $\lambda > 0$, $x \geq 0$, (2.3) presents $\overline{F}(x) = e^{-\lambda x}$, i.e., X has an exponential distribution.

(2) For $s(x) = \frac{\left(\frac{x}{\beta} \right)^{\alpha(\gamma_{r+1} + \gamma_{r+2} + 1)} - 1}{(\gamma_{r+1} + \gamma_{r+2} + 1)}$, $\alpha, \beta > 0$ and $x \geq \beta$, (2.3) presents $\overline{F}(x) = (x/\beta)^{-\alpha}$, $x \geq \beta$, i.e., X has a Pareto distribution.

(3) For $s(x) = \frac{\exp\left\{ \left(\frac{x}{\lambda} \right)^p (\gamma_{r+1} + \gamma_{r+2} + 1) \right\} - 1}{(\gamma_{r+1} + \gamma_{r+2} + 1)}$, p, $\lambda > 0$, $x \geq 0$, (2.3) presents $\overline{F}(x) = e^{-\left(\frac{x}{\lambda} \right)^p}$, $x \geq 0$, i.e., X has Weibull (also Rayleigh) distribution.

(4) For $s(x) = \frac{\left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{\frac{(\gamma_{r+1} + \gamma_{r+2} + 1)}{\xi}} - 1}{(\gamma_{r+1} + \gamma_{r+2} + 1)}$, ξ, $\sigma > 0$, $x \geq \mu$ (2.3) presents $\overline{F}(x) = \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{-1/\xi}$, $x \geq \mu$, i.e., X has a generalized Pareto distribution.

(5) For $s(x) = \frac{\left(1 - \frac{x}{\beta} \right)^{-\alpha(\gamma_{r+1} + \gamma_{r+2} + 1)} - 1}{(\gamma_{r+1} + \gamma_{r+2} + 1)}$, $\alpha > 0$, $0 \leq x \leq \beta$, (2.3) presents $\overline{F}(x) = \left(1 - \frac{x}{\beta} \right)^{\alpha}$, $0 \leq x \leq \beta$, , i.e., X has a power function distribution.

(6) For $s(x) = \frac{(1 + x^p)^{-\alpha(\gamma_{r+1} + \gamma_{r+2} + 1)} \exp\{\alpha(\gamma_{r+1} + \gamma_{r+2} + 1) x^p\} - 1}{(\gamma_{r+1} + \gamma_{r+2} + 1)}$, α, $p > 0$, $x \geq 0$, (2.3) presents $\overline{F}(x) = (1 + x^p)^{\alpha} e^{-\alpha x^p}$, $x \geq 0$. For $\alpha = p = 1$, X has a generalized gamma distribution and for $\alpha = 1$, $p = 2$, X has a generalized normal distribution.

Theorem 2.2. Let $X : \Omega \to (a, b)$ be a random variable with a twice differentiable distribution function $F(x)$ and with density function $f(x)$. Let $s(x)$ be a non-positive, non-decreasing and twice differentiable function on (a, b) such that for some x_0 $(a < x_0 < b)$, $f(x_0) > 0$, $\lim_{x \to b} s(x) = 0$ and $\lim_{x \to a} s(x) = -\infty$. Then

$$E\left[s(X(r, n, m, k)) \mid X(r+2, n, m, k) = x \right] = s(x), \qquad (2.9)$$

implies

$$\overline{F}(x) = \left\{ 1 - \left[1 - \frac{1}{C} (2r+1)(m+1) s(x) \right]^{-1/(2r+1)} \right\}^{1/(m+1)}, \quad a \leq x \leq b, \qquad (2.10)$$

where C, which depends on x_0, is a positive normalizing constant.

Proof. Suppose (2.9) holds, then we have

$$r(r+1)(m+1) \int_a^x s(y) f(y) \left[\overline{F}(y)\right]^m \left\{1 - \left[\overline{F}(y)\right]^{m+1}\right\}^{r-1} \times$$

$$\left[\left[\overline{F}(y)\right]^{m+1} - \left[\overline{F}(x)\right]^{m+1}\right] dy = s(x) \left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{r+1} \tag{2.11}$$

Differentiating both sides of (2.11) with respect to x, we get

$$r(r+1)(m+1) \, f(x) \left[\overline{F}(x)\right]^m \int_a^x (m+1) s(y) f(y) \left[\overline{F}(y)\right]^m \left\{1 - \left[\overline{F}(y)\right]^{m+1}\right\}^{r-1} dy$$
$$= s'(x) \left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{r+1} + (r+1)(m+1) s(x) f(x) \left[\overline{F}(x)\right]^m \left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{r}.$$

Upon simplification, we have

$$\int_a^x s(y) f(y) \left[\overline{F}(y)\right]^m \left\{1 - \left[\overline{F}(y)\right]^{m+1}\right\}^{r-1} dy =$$
$$\frac{1}{r(r+1)(m+1)^2} \left(\frac{s'(x)}{f(x)}\right) \left[\overline{F}(x)\right]^{-m} \left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{r+1}$$
$$+ \frac{1}{r(m+1)} s(x) \left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{r}. \tag{2.12}$$

Differentiating both sides of (2.12) with respect to x and then simplifying, we obtain

$$\frac{d}{dx}\left(\frac{s'(x)}{f(x)}\right) + \frac{2(r+1)(m+1) s'(x) \left[\overline{F}(x)\right]^m}{\left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}} + \frac{m s'(x)}{\overline{F}(x)} = 0,$$

or

$$\frac{\frac{d}{dx}\left(\frac{s'(x)}{f(x)}\right)}{\left(\frac{s'(x)}{f(x)}\right)} + 2(r+1)(m+1) \frac{f(x) \left[\overline{F}(x)\right]^m}{\left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}} + \frac{m f(x)}{\overline{F}(x)} = 0.$$

Integrating the above equation from x_0 to x, we arrive at

$$s'(x) = C f(x) \left[\overline{F}(x)\right]^m \left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{-2r-2}. \tag{2.13}$$

Integrating both sides of (2.13) from x to b and using the assumption $\lim_{x \to b} s(x) = 0$, we obtain

$$s(x) = \frac{C}{(2r+1)(m+1)} \left[1 - \left\{1 - \left[\overline{F}(x)\right]^{m+1}\right\}^{-2r-1}\right],$$

and

$$\overline{F}(x) = \left\{1 - \left[1 - \frac{1}{C}(2r+1)(m+1) s(x)\right]^{-1/(2r+1)}\right\}^{1/(m+1)}.$$

Remark 2.2. Similar to Remark 2.1, for different functions s satisfying the conditions of Theorem 2.2, one can obtain characterizations of various well-known distributions of which, we mention some of them below.

(7) For $s(x) = \frac{C}{(2r+1)(m+1)} \left\{ 1 - \left[1 - e^{-\lambda(m+1)x} \right]^{-(2r+1)} \right\}$, $x > 0$, (2.10) presents $\overline{F}(x) = e^{-\lambda x}$, i.e., X has an exponential distribution.

(8) For $s(x) = \frac{C}{(2r+1)(m+1)} \left\{ 1 - \left[1 - (x/\beta)^{-\alpha(m+1)} \right]^{-(2r+1)} \right\}$, α, $\beta > 0$ and $x > \beta$, (2.10) presents $\overline{F}(x) = (x/\beta)^{-\alpha}$, i.e., X has a Pareto distribution.

(9) For $s(x) = \frac{C}{(2r+1)(m+1)} \left\{ 1 - \left[1 - e^{-\left(\frac{x}{\lambda}\right)^p (m+1)} \right]^{-(2r+1)} \right\}$, $p, \lambda > 0$, and $x > 0$, (2.10) presents $\overline{F}(x) = e^{-\left(\frac{x}{\lambda}\right)^p}$, i.e., X has Weibull (also Rayleigh) distribution.

(10) For $s(x) = \frac{C}{(2r+1)(m+1)} \left\{ 1 - \left\{ 1 - \left[1 + \xi\left(\frac{x-\mu}{\sigma}\right) \right]^{-\frac{(m+1)}{\xi}} \right\}^{-(2r+1)} \right\}$, $\xi, \sigma > 0$, and $x > \mu$, (2.10) presents $\overline{F}(x) = \left[1 + \xi\left(\frac{x-\mu}{\sigma}\right) \right]^{-\frac{1}{\xi}}$, i.e., X has generalized Pareto distribution.

(11) For $s(x) = \frac{C}{(2r+1)(m+1)} \left\{ 1 - \left[1 - \left(1 - \frac{x}{\beta} \right)^{\alpha(m+1)} \right]^{-(2r+1)} \right\}$, $\alpha > 0$, and $0 < x \leq \beta$, (2.10) presents $\overline{F}(x) = \left(1 - \frac{x}{\beta} \right)^{\alpha}$, i.e., X has power function distribution.

Acknowledgments

References

Ahsanullah, M., & Hamedani, G. G. (2013). Characterizations of continuous distributions based on conditional expectation of generalized order statistics. *Communications in Statistics - Theory and Methods, 42(19)*, 3608-3613. http://dx.doi.org/10.1080/03610926.2011.635255

Bieniek, M., & Szynal, D. (2003). Characterizations of distributions via linearity of regression of generalized order statistics. *Metrika, 58(3)*, 259-271. http://dx.doi.org/10.1007/s001840300263

Bieniek, M. (2009). A note on characterizations of distributions by regressions of non-adjacent generalized order statistics. *Australian & New Zealand Journal of Statistics, 51(1)*, 89-99. http://dx.doi.org/10.1111/j.1467-842X.2008.00530.x

Cramer, E., Kamps, U., & Keseling, C. (2004). Characterizations via linear regression of ordered random variables: A unifying approach. *Communications in Statistics - Theory and Methods, 33(12)*, 2885 - 2911. http://dx.doi.org/10.1081/STA-200038832

Kamps, U. (1995). *A concept of generalized order statistics*. Teubner Skripten zur Mathematischen Stochastik [Teubner Texts on Mathematical Stochastics]. B. G. Teubner, Stuttgart.

Copyrights