Computing the Moments of Order Statistics from Truncated Pareto Distributions Based on the Conditional Expectation

Gökhan Gökdere
Department of Mathematics
University of Bingol, Bingol, Turkey
g.g.gokdere@gmail.com

Abstract

In this paper, closed form expressions for the moments of the truncated Pareto order statistics are obtained by using conditional distribution. We also derive some results for the moments which will be useful for moment computations based on ordered data.

Keywords: Order statistics, Pareto distribution, Moments, Truncated distribution, Conditional distribution.

MSC 2010: 62G30, 62E15.

1. Introduction

Order statistics and their moments have assumed considerable interest in recent years. There is a vast literature on both theory and application of the moments of order statistics. Many recurrence relations for the single and product moments of the order statistics from different distributions have been revealed. These relations provide advantages in computation of the lower and higher moments of order statistics.

Joshi and Balakrishnan (1982) obtained several recurrence relations and identities for product moments of order statistics in a random sample of size n from an arbitrary continuous distribution. Balakrishnan and Joshi (1982) established independent and identically distributed results for the Pareto and doubly-truncated Pareto models, at the same time these results allow us to evaluate all the single and product moments of order statistics. Balakrishnan et al. (1988) and Malik, et al. (1988) reviewed several recurrence relations and identities for single and product moments of order statistics for specific distributions. Mohie El-Din et al. (1997) derived expressions for the moments and product moments of the order statistics from the doubly truncated linear-exponential distribution. Childs and Balakrishnan (1998) generalized the I.I.D. results for the Pareto and doubly-truncated Pareto models established by Balakrishnan and Joshi (1982). Ahmad (2001) derived some general recurrence relations satisfied by single and product moments of order statistics from doubly truncated continuous distributions. Afify (2006) derived some recurrence relations of single and product moments of order statistics from identical Pareto distribution and estimated the parameters of the first order statistics and the mean, variance and the coefficient of variation were also computed. Bekçi (2009) examined order statistics of a random sample of size n drawn from uniform distribution and derive some recurrence relations for the single and product moments of these order statistics. Nadarajah (2010) derived exact and explicit expressions for moments of order statistics from Pareto distributions.

A good deal of work has appeared in the literature on characterizing a distribution by means of the conditional expectation of $X_{k+1:n}$ (or its function) given $X_{k:n}$.

More specifically, Franco and Ruiz (1999) characterized a distribution by means of $E[\psi(X_{k+1:n}) | X_{k:n=z}] = g(z)$, under some mild conditions of $\psi(.)$ and $g(.)$. Gupta and Ahsanullah (2004) attempted to characterize a distribution by means of $E[\psi(X_{k+s:n}) | X_{k:n=z}] = g(z)$, under some mild conditions of $\psi(.)$ and $g(.)$. Khan and Faizan (2013) characterized two families of probability distributions through the conditional expectations of dual generalized order statistics. Various characterizations of the class of exponential distributions are presented by Hamedani (2013).

Suppose that the random variable X has the distribution function (*d.f.*) given by

$$F(x) = \begin{cases} 0, & x \le a \\ 1 - e^{-1/\lambda[h(x)-h(a)]}, & a < x \le b \\ 1, & x > b \end{cases} \tag{1}$$

where λ is a positive parameter, $h(x)$ is assumed to be non-decreasing, continuous and differentiable function on (a,b) such that $h(a) \ge 0$ and $h(b_{\_}) = \infty$. Different choices of $h(x)$ lead to distributions that is important in life testing as well as other areas of statistics (Ahmad, 2001). Put $h(x) = \ln x$, $x \in (1,\infty)$, $\lambda = \frac{1}{v}$, $h(a) = 0$, $v > 0$ then the *d.f.* (1) reduces to Pareto distribution in the form

$$F(x) = 1 - x^{-v}, \ x \in (1,\infty), v > 0 \tag{2}$$

and probability density function (*p.d.f.*)

$$f(x) = vx^{-(v+1)}, \ x \in (1,\infty), \ v > 0. \tag{3}$$

We note that $f(x)$ and $F(x)$ satisfy the relation

$$F(x) = 1 - \frac{x}{v} f(x).$$

Let $X_1, X_2, ..., X_n$ be independent random variables with *d.f.* $F_1, F_2, ..., F_n$ respectively and let $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ denote the corresponding order statistics. In the theory of order statistics it is usually assumed that $X_1, X_2, ..., X_n$ are identically distributed. However in many practical situations it is necessary to allow for non-identical $F_1, F_2, ..., F_n$.

Suppose $X_1, X_2, ..., X_n$ have independent and identical *d.f.* $F(x)$ and *p.d.f.* $f(x)$. The *p.d.f.* of $X_{r:n}$ $(1 \le r \le n)$ is denoted by (David, 1981)

$$f_{X_{r:n}}(x) = \frac{n!}{(r-1)!(n-r)!}[F(x)]^{r-1} f(x)[1-F(x)]^{n-r}, \tag{4}$$

where $-\infty < x < \infty$.

The joint *p.d.f.* of $X_{r:n}$ and $X_{s:n}$ $(1 \le r < s \le n)$ is denoted by (David, 1981)

$$f_{X_{r:n},X_{s:n}}(x,y) = \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!}[F(x)]^{r-1} f(x)[F(y)-F(x)]^{s-r-1} f(y)[1-F(y)]^{n-s}, \quad (5)$$

where $x < y$.

The joint *p.d.f.* of $X_{r_1:n}, X_{r_2:n}, \ldots, X_{r_d:n}$ $(1 \le d \le n$ and $0 = r_0 < r_1 < r_2 < \ldots < r_d < r_{d+1} = n+1)$ is denoted by (Reiss, 1989)

$$f_{X_{r_1:n},X_{r_2:n},\ldots,X_{r_d:n}}(x_1, x_2, \ldots, x_d) = n!\prod_{i=1}^{d} f(x_i)\prod_{i=1}^{d+1}\frac{[F(x_i)-F(x_{i-1})]^{r_i - r_{i-1}-1}}{(r_i - r_{i-1}-1)!}, \quad (6)$$

where $x_1 < x_2 < \ldots < x_d$, $F(x_0) = 0$ and $F(x_{d+1}) = 1$.

The aim of this paper is to obtain closed form expressions for the moments of the truncated Pareto order statistics. The contents of this paper are organized as follows. At first we give materials and methods. Then, we compute the moments of order statistics from left truncated Pareto variables, right truncated Pareto variables and doubly truncated Pareto variables. Later, we derive some results for the moments of the truncated Pareto order statistics. We consider Pareto distribution since it has a wide use in economic and finance.

2. Materials and Methods

The conditional distribution of the *s*th order statistic, $X_{s:n}$ given that the *r*th order statistic $X_{r:n}$ is equal to x_*, is expressed as

$$f_{X_{s:n}|X_{r:n}=x_*}(y_*) = \frac{f_{X_{r:n},X_{s:n}}(x_*, y_*)}{f_{X_{r:n}}(x_*)}, \quad (7)$$

where $1 \le r < s \le n$ and $x_* < y_*$. After substituting (4) and (5) into (7) and some manipulation, we can be obtained that

$$f_{X_{s:n}|X_{r:n}=x_*}(y_*) = (s-r)\binom{n-r}{n-s}\left[\frac{F(y_*)-F(x_*)}{1-F(x_*)}\right]^{s-r-1}\left[\frac{1-F(y_*)}{1-F(x_*)}\right]^{n-s}\frac{f(y_*)}{1-F(x_*)}, \quad (8)$$

where $f_{X_{s:n}|X_{r:n}=x_*}(y_*)$ is the left truncated *p.d.f.* of *s*th order statistic. We note that $\frac{F(y_*)-F(x_*)}{1-F(x_*)}$ and $\frac{f(y_*)}{1-F(x_*)}$ are the *d.f.* and *p.d.f.* of the population whose distributions is obtained by truncating the distribution $F(x)$ on the left at x_*.

Similarly, the conditional distribution of the *r*th order statistic, $X_{r:n}$ given that the *s*th order statistic $X_{s:n}$ is equal to y_*, is expressed as

$$f_{X_{r:n}|X_{s:n}=y_*}(x_*) = (s-r)\binom{s-1}{r-1}\left[\frac{F(x_*)}{F(y_*)}\right]^{r-1}\left[\frac{F(y_*)-F(x_*)}{F(y_*)}\right]^{s-r-1}\frac{f(x_*)}{F(y_*)}, \quad (9)$$

where $1 \le r < s \le n$, $x_* < y_*$ and $f_{X_{r:n}|X_{s:n}=y_*}(x_*)$ is the right truncated *p.d.f.* of *r*th order statistic.

We note that $\frac{F(x_*)}{F(y_*)}$ and $\frac{f(x_*)}{F(y_*)}$ are the *d.f.* and *p.d.f.* of the population whose distributions is obtained by truncating the distribution $F(x)$ on the left at y_*.

The conditional distribution of the *u*th order statistic, $X_{u:n}$ given that the *r*th order statistic $X_{r:n}$ is equal to x_* and *s*th order statistic $X_{s:n}$ is equal to y_*, is expressed as,

$$f_{X_{u:n}|X_{r:n}=x_*,X_{s:n}=y_*}(z_*) = \frac{f_{X_{r:n},X_{u:n},X_{s:n}}(x_*, z_*, y_*)}{f_{X_{r:n},X_{s:n}}(x_*, y_*)}, \tag{10}$$

where $1 \le r < u < s \le n$ and $x_* < z_* < y_*$. After substituting (5) and (6) into (10) and some manipulation, we can be obtained that

$$f_{X_{u:n}|X_{r:n}=x_*,X_{s:n}=y_*}(z_*) = (u-r)\binom{s-r-1}{s-u-1}\left[\frac{F(z_*)-F(x_*)}{F(y_*)-F(x_*)}\right]^{u-r-1}\left[\frac{F(y_*)-F(z_*)}{F(y_*)-F(x_*)}\right]^{s-u-1}\frac{f(z_*)}{F(y_*)-F(x_*)} \tag{11}$$

where $f_{X_{u:n}|X_{r:n}=x_*,X_{s:n}=y_*}(z_*)$ is the doubly truncated *p.d.f.* of *r*th and *s*th order statistic. We note that $\frac{F(z_*)-F(x_*)}{F(y_*)-F(x_*)}$ and $\frac{f(z_*)}{F(y_*)-F(x_*)}$ are the *d.f.* and *p.d.f.* of the population whose distributions is obtained by truncating the distribution $F(x)$ on the left at x_* and right at y_*.

3. Moments of Truncated Pareto Order Statistics

In this section, we compute the moments of order statistics from truncated Pareto variables by means of the conditional expectation. First, we will establish the moment of the left truncated order statistics from the Pareto distribution. Let us denote the moments $E(X_{s:n}|X_{r:n=x_*})$ by $\mu_{s:n|r:n}$, is given by

$$\mu_{s:n|r:n} = \int_{x_*}^{\infty} y_*\, f_{X_{s:n}|X_{r:n}=x_*}(y_*)dy_*, \tag{12}$$

where $1 \le r < s \le n$ and $1 < x_* < y_* < \infty$.

Theorem 3.1.

The moment of the left truncated order statistics from Pareto distribution by means of the conditional expectation is given by

$$\mu_{s:n|r:n} = (s-r)\binom{n-r}{n-s} v x_* \sum_{k=1}^{s-r}\binom{s-r-1}{k-1}(-1)^k \frac{1}{v(s-n-k)+1} \tag{13}$$

Proof. Let $F(x_*)$ and $f(x_*)$ be as in (2) and (3). By using properties of binom and some manipulation, (8) can be expressed as

$$f_{X_{s:n}|X_{r:n}=x_*}(y_*) = (s-r)\binom{n-r}{n-s} v \sum_{k=1}^{s-r} \binom{s-r-1}{k-1} (-1)^{k-1} (x_*)^{v(n+k-s)} (y_*)^{v(s-n-k)-1} \tag{14}$$

After substituting (14) into (12) and then integrating them we can be obtained (13).

Then, we will establish the moment of the right truncated order statistics from the Pareto distribution. Let us denote the moments $E(X_{r:n}|X_{s:n=y_*})$ by $\mu_{r:n|s:n}$, is given by

$$\mu_{r:n|s:n} = \int_1^{y_*} x_* \; f_{X_{r:n}|X_{s:n}=y_*}(x_*) dx_* \,, \tag{15}$$

where $1 \le r < s \le n$ and $1 < x_* < y_* < \infty$.

Theorem 3.2.

The moment of the right truncated order statistics from Pareto distribution by means of the conditional expectation is given by

$$\mu_{r:n|s:n} = (s-r)\binom{s-1}{r-1} \frac{v}{[1-y_*^{-v}]^{s-1}} \sum_{k_1=1}^{r} \sum_{k_2=1}^{s-r} \binom{r-1}{k_1-1}\binom{s-r-1}{k_2-1} (-1)^{k_1+k_2-2} \frac{(y_*)^{v(r-k_1-s+1)+1} - (y_*)^{v(-k_2+1)}}{v(r+k_2-k_1-s)+1} \tag{16}$$

Proof. Let $F(x_*)$ and $f(x_*)$ be as in (2) and (3). By using properties of binom and some manipulation, (9) can be expressed as

$$f_{X_{r:n}|X_{s:n}=y_*}(x_*) = (s-r)\binom{s-1}{r-1} \frac{v}{[1-y_*^{-v}]^{s-1}} \sum_{k_1=1}^{r} \sum_{k_2=1}^{s-r} \binom{r-1}{k_1-1}\binom{s-r-1}{k_2-1} (-1)^{k_1+k_2-2} (x_*)^{v(r+k_2-k_1-s)-1} (y_*)^{v(-k_2+1)} \tag{17}$$

After substituting (17) into (15) and then integrating them we can be obtained (16).

Finally, we will establish the moment of the doubly truncated order statistics from the Pareto distribution. Let us denote the moments $E(X_{u:n}|X_{r:n=x_*}, X_{s:n=y_*})$ by $\mu_{u:n|r:n,s:n}$, is given by

$$\mu_{u:n|r:n,s:n} = \int_{x_*}^{y_*} z_* \; f_{X_{u:n}|X_{r:n}=x_*, X_{s:n}=y_*}(z_*) dz_* \,, \tag{18}$$

where $1 \le r < u < s \le n$ and $1 < x_* < z_* < y_* < \infty$.

Theorem 3.3

The moment of the doubly truncated order statistics from Pareto distribution by means of the conditional expectation is given by

$$\mu_{u:n|r:n,s:n} = (u-r)\binom{s-r-1}{s-u-1} \frac{v}{[x_*^{-v} - y_*^{-v}]^{s-r-1}} \sum_{k_3=1}^{u-r} \sum_{k_4=1}^{s-u} \binom{u-r-1}{k_3-1}\binom{s-u-1}{k_4-1} (-1)^{k_3+k_4-2}$$

$$\frac{(x_*)^{-v(u-r-k_3)} (y_*)^{v(u+1-s-k_3)+1} - (x_*)^{v(r+k_4-s)+1} (y_*)^{v(-k_4+1)}}{v(u+k_4-k_3-s)+1} \tag{19}$$

Proof. Let $F(x_*)$ and $f(x_*)$ be as in (2) and (3). By using properties of binom and some manipulation, (11) can be expressed as

$$f_{X_{u:n}|X_{r:n}=x_*,X_{s:n}=y_*}(z_*) = (u-r)\binom{s-r-1}{s-u-1}\frac{v}{[x_*^{-v} - y_*^{-v}]^{s-r-1}}\sum_{k_3=1}^{u-r}\sum_{k_4=1}^{s-u}\binom{u-r-1}{k_3-1}\binom{s-u-1}{k_4-1}(-1)^{k_3+k_4-2}$$

$$.\,(x_*)^{-v(u-r-k_3)}(y_*)^{-v(k_4-1)}(z_*)^{-v(s+k_3-u-k_4)-1}. \qquad (20)$$

After substituting (20) into (18) and then integrating them we can be obtained (19).

4. Results for Moments of Truncated Pareto Order Statistics

In this section, we derive some results for the moments of the truncated Pareto order statistics.

For $s = n$ in (13) the moment of the conditional *p.d.f.* of largest order statistics given $X_{r:n=x_*}$ is given by

$$\mu_{n:n|r:n} = (n-r)vx_*\sum_{k=1}^{n-r}\binom{n-r-1}{k-1}(-1)^k \frac{1}{-vk+1}$$

For $r = 1$ in (16) the moment of the conditional *p.d.f.* of smallest order statistics given $X_{s:n=y_*}$ is given by

$$\mu_{1:n|s:n} = (s-1)\frac{v}{[1-y^{-v}]^{s-1}}\sum_{k_2=1}^{s-1}\binom{s-2}{k_2-1}(-1)^{k_2-1}\frac{(y_*)^{v(s+1)+1} - (y_*)^{v(-k_2+1)}}{v(k_2-s)+1}.$$

For $r = k$, $u = k+1$ and $r = k+2$ in (19) the moment of the conditional *p.d.f.* of order statistics given $X_{k:n=x_*}$ and $X_{k+2:n=y_*}$ is given by

$$\mu_{k+1:n|k:n,k+2:n} = \left(\frac{v}{v-1}\right)\left(\frac{x_*^{-v+1} - y_*^{-v+1}}{x_*^{-v} - y_*^{-v}}\right).$$

References

1. Afify, E. d-E. (2006). Order statistics from Pareto distributions. *Journal of Applied Science* 6, 2151-2157.
2. Ahmad, Abd el-baset A. (2001). Moments of order statistics from doubly truncated continuous distributions and characterizations. *Statistics: A journal of Theoretical and Applied Statistics* 35, 479-494.
3. Balakrishnan, N. and Joshi, P. C. (1982). Moments of order statistics from doubly truncated Pareto distribution. *Journal of the Indian Statistical Association* 20, 109-117.
4. Balakrishnan, N., Malik, H.J. and Ahmed, S.E. (1988). Recurrence relations and identities for moments of order statistics, II: Specific continuous distributions. *Commun.Statist.-Theor. Meth.* 17, 2657-2694.

5. Bekçi, M. (2009). Recurrence relations for the moment of order statistics from uniform distribution. *Scientific Research and Essay* 4, 1302-1305.

6. Childs, A. and Balakrishnan, N. (1998). Generalized recurrence relations for moment of order statistics from non-identical Pareto and truncated Pareto random variables with applications to robustness. *Handbook of Statistics,* 16, 403-438 (North-Holland Amsterdam).

7. David, H. A. (1981). *Order statistics*. John Wiley and Sons Inc., New York.

8. Franco, M. and Ruiz, J.M. (1999). Characterization based on conditional expectation of adjacent order statistics: A unified approach. *Proceedings of the American Mathematical Society,* 123(3), 861-874.

9. Gupta, R.C. and Ahsanullah M. (2004). Some characterization results based on the conditional expectation of a function of non-adjacent order statistic (record value). *Ann. Inst. Statist. Math.*, 56(4), 721-732.

10. Hamedani, G. H. (2013). Characterizations of Exponential Distributions. *Pak.j.stat.oper.res.,* 4(1), 17-24.

11. Joshi, P. C. and Balakrishnan, N. (1982).Recurrence relations and identities for the product moment of order statistics. *Sankhyâ B* 44, 39-49.

12. Khan, M.I. and Faizan, M. (2013). Some Characterizations Results Based on Conditional Expectation of Function of Dual Generalized Order Statistics. *Pak.j.stat.oper.res.,* 8(4), 789-799.

13. Malik, H.J., Balakrishnan, N. and Ahmed, S.E. (1988). Recurrence relations and identities for moments of order statistics, I: Arbitrary continuous distributions. *Commun.Statist.-Theor. Meth.,* 17, 2623-2655.

14. Mohie, El-Din M.M., Mahmoud, M.A.W., Abu-Youssef, S.E. Sultan. K.S. (1997). Order statistics from doubly truncated linear-exponential distribution and its characterizations.*Commun.Statist.-Simul. and Comput.* 26(1), 281-290.

15. Nadarajah, S. (2010). Explicit expressions for moment of Pareto order statistics. *Quantitative Finance* 10, 585-589.

16. Reiss, R. D. (1989). *Approximate distributions of order statistics*. Springer, Verlag. New York. Inc., USA.