International Journal of Computational and Theoretical Statistics
ISSN (2384-4795)
Int. J. Comp. Theo. Stat. 2, No. 1 (May -2015)

Some Characterization of Continuous Distributions Based on Order Statistics

Haseeb Athar[1] and Zuber Akhter[1]

[1] Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh-202002, India

Received Nov. 14, 2014, Revised Feb. 19, 2015, Accepted Mar.1, 2015, Published May 1, 2015

Abstract: A family of continuous probability distributions $F(x) = ah(x) + b,\ x \in (\alpha, \beta)$ have been characterized through conditional expectation of power of difference of two order statistics, conditioned on a pair of non-adjacent order statistics. The characterization results presented in this paper extend some of the existing results based on the order statistics. Further, some particular cases and examples are also discussed.

Keywords: Order statistics; Conditional expectation; Continuous probability distributions; Characterization.

1. Introduction

Characterization of distributions through the properties of conditional expectations of order statistics have been studied by several authors. Various approaches are available in literature. For detailed survey one may refer to Khan and Ali [4], Nagaraja [9], Khan and Abu-Salih [5], Balasubramanian and Beg [12], Franco and Ruiz ([13], [14]), Balasubramanian and Dey [11], López-Blázquez and Moreno-Rebello [6], Wesolowski and Ahsanullah [10], Dembińska and Wesolowski [1], Khan and Abouammoh [2], Athar *et al.* [7], Khan and Athar [3] and references therein.

In this paper, an attempt is made to characterize a general form of distribution $F(x) = ah(x) + b,\ x \in (\alpha, \beta)$ through conditional expectation of $p-th$ power of difference of functions of two order statistics, conditioned on a pair of non-adjacent order statistics.

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a continuous population having probability density function (pdf) $f(x)$ and distribution function (df) $F(x)$, over the support (α, β), where α, β may be finite or infinite and let $X_{1:n} \le X_{2:n} \le \ldots \le X_{n:n}$ be the corresponding order statistics. Then the conditional pdf of $X_{i:n}$ given $X_{r:n} = x$ and $X_{s:n} = y$, $1 \le r < i < s \le n$ is (David and Nagaraja [8])

$$f_{i|r,s}(t \mid x, y) = \frac{(s-r-1)!}{(i-r-1)!(s-i-1)!}\left[\frac{F(t)-F(x)}{F(y)-F(x)}\right]^{i-r-1}\left[\frac{F(y)-F(t)}{F(y)-F(x)}\right]^{s-i-1}\frac{f(t)}{[F(y)-F(x)]},\quad \alpha < x < t < y < \beta. \tag{1.1}$$

2. Characterization Results

Theorem 2.1: *Let X be an absolutely continuous random variable with the df $F(x)$ and pdf $f(x)$ over the support (α, β). Then for two consecutive values r and $r+1$, $1 \le r < i-1 \le s \le n$,*

$$E[\{h(X_{i:n}) - h(X_{l:n})\}^p \mid X_{l:n} = x, X_{s:n} = y] = g_{l,i,s,p}(x, y);\ l = r, r+1$$

$$= \frac{\Gamma(p+i-l)\Gamma(s-l)}{\Gamma(p+s-l)\Gamma(i-l)}\{h(y) - h(x)\}^p \tag{2.1}$$

if and only if

$$F(x) = ah(x) + b \tag{2.2}$$

E-mail: haseebathar@hotmail.com, akhterzuber022@gmail.com

where a, b *are so chosen that* $F(x)$ *is a* df *and* $h(x)$ *is a monotonic, continuous and differentiable function of* x *and* p *is a positive integer.*

Proof: To prove the necessary part, we have

$$E[\{h(X_{i:n})-h(X_{r:n})\}^p \mid X_{r:n}=x, X_{s:n}=y]=g_{r,i,s,p}(x,y)$$

$$=\frac{\Gamma(s-r)}{\Gamma(i-r)\Gamma(s-i)}\int_x^y (h(t)-h(x))^p\left[\frac{F(t)-F(x)}{F(y)-F(x)}\right]^{i-r-1}\left[1-\frac{F(t)-F(x)}{F(y)-F(x)}\right]^{s-i-1}\frac{f(t)}{F(y)-F(x)}dt.$$

Let

$$\frac{F(t)-F(x)}{F(y)-F(x)}=z,$$

implying

$$(h(t)-h(x))^p = z^p (h(y)-h(x))^p.$$

Then, we have

$$E[\{h(X_{i:n})-h(X_{r:n})\}^p \mid X_{r:n}=x, X_{s:n}=y]=g_{r,i,s,p}(x,y)$$

$$=\frac{\Gamma(s-r)}{\Gamma(i-r)\Gamma(s-i)}\int_0^1 (h(y)-h(x))^p z^{p+i-r-1}(1-z)^{s-i-1}dz,$$

and hence the necessary part.

To prove the sufficiency part, consider

$$E[\{h(X_{i:n})-h(X_{r:n})\}^p \mid X_{r:n}=x, X_{s:n}=y]=g_{r,i,s,p}(x,y)$$

or,

$$\frac{\Gamma(s-r)}{\Gamma(i-r)\Gamma(s-i)}\int_x^y (h(t)-h(x))^p[F(t)-F(x)]^{i-r-1}[F(y)-F(t)]^{s-i-1}f(t)dt = g_{r,i,s,p}(x,y)[F(y)-F(x)]^{s-r-1}. \quad (2.3)$$

Differentiating (2.3) *w.r.t.* x, we get

$$-ph'(x)\frac{\Gamma(s-r)}{\Gamma(i-r)\Gamma(s-i)}\int_x^y (h(t)-h(x))^{p-1}[F(t)-F(x)]^{i-r-1}[F(y)-F(t)]^{s-i-1}f(t)dt$$

$$-\frac{(i-r-1)\Gamma(s-r)}{\Gamma(i-r)\Gamma(s-i)}\int_x^y (h(t)-h(x))^p[F(t)-F(x)]^{i-r-2}[F(y)-F(t)]^{s-i-1}f(t)f(x)dt$$

$$=\frac{\partial}{\partial x}\{g_{r,i,s,p}(x,y)\}[F(y)-F(x)]^{s-r-1}-(s-r-1)g_{r,i,s,p}(x,y)[F(y)-F(x)]^{s-r-2}f(x) \quad (2.4)$$

or,

$$-ph'(x)g_{r,i,s,p-1}(x,y)-(s-r-1)g_{r+1,i,s,p}(x,y)\frac{f(x)}{[F(y)-F(x)]}$$

$$=\frac{\partial}{\partial x}g_{r,i,s,p}(x,y)-(s-r-1)g_{r,i,s,p}(x,y)\frac{f(x)}{[F(y)-F(x)]}. \quad (2.5)$$

Rearranging the terms of (2.5), we get

$$\frac{f(x)}{[F(y)-F(x)]}=\frac{1}{(s-r-1)}\frac{ph'(x)g_{r,i,s,p-1}(x,y)+\frac{\partial}{\partial x}g_{r,i,s,p}(x,y)}{g_{r,i,s,p}(x,y)-g_{r+1,i,s,p}(x,y)}. \quad (2.6)$$

Now consider,

$$ph'(x)g_{r,i,s,p-1}(x,y)+\frac{\partial}{\partial x}g_{r,i,s,p}(x,y)$$

$$= ph'(x)\{h(y)-h(x)\}^{p-1}\frac{\Gamma(p+i-r-1)\ \Gamma(s-r)}{\Gamma(p+s-r-1)\ \Gamma(i-r)} - ph'(x)\{h(y)-h(x)\}^{p-1}\frac{\Gamma(p+i-r)\ \Gamma(s-r)}{\Gamma(p+s-r)\ \Gamma(i-r)},$$

$$= p(s-i)h'(x)\{h(y)-h(x)\}^{p-1}\frac{\Gamma(p+i-r-1)\ \Gamma(s-r)}{\Gamma(p+s-r)\ \Gamma(i-r)} \tag{2.7}$$

and

$$g_{r,i,s,p}(x,y) - g_{r+1,i,s,p}(x,y)$$

$$= \{h(y)-h(x)\}^{p}\frac{\Gamma(p+i-r)\ \Gamma(s-r)}{\Gamma(p+s-r)\ \Gamma(i-r)} - \{h(y)-h(x)\}^{p}\frac{\Gamma(p+i-r-1)\ \Gamma(s-r-1)}{\Gamma(p+s-r-1)\ \Gamma(i-r-1)}$$

$$= p(s-i)\{h(y)-h(x)\}^{p}\frac{\Gamma(p+i-r-1)\ \Gamma(s-r-1)}{\Gamma(p+s-r)\ \Gamma(i-r)}. \tag{2.8}$$

Therefore in view of (2.6), we have

$$\frac{f(x)}{[F(y)-F(x)]} = \frac{h'(x)}{[h(y)-h(x)]}.$$

Hence the theorem.

Remark 2.1: At $p=1$, (2.1) reduces to

$$E[h(X_{i:n}) \mid X_{r:n}=x, X_{s:n}=y] = \frac{(s-i)h(x)+(i-r)h(y)}{(s-r)}, \quad i=r+1, r+2, \ldots, s-1 \tag{2.9}$$

or equivalently

$$E[h(X_{r+i:n}) \mid X_{r:n}=x, X_{s:n}=y] = \frac{(m-i+1)h(x)+ih(y)}{(m+1)}, \quad m=s-r-1,\ 1\le r<s\le n,\ i=1,2,\ldots,m. \tag{2.10}$$

as obtained by Khan and Athar [3].

Further, at $s=n+1$, $X_{n+1:n}=y=\beta$ and $m=n-r$, then we have

$$E[h(X_{i:n}) \mid X_{r:n}=x] = \frac{(n-i+1)h(x)+(i-r)h(\beta)}{(n-r+1)}, \quad i=r+1, r+2, \ldots, n \tag{2.11}$$

or,

$$E[h(X_{r+i:n}) \mid X_{r:n}=x] = \frac{(n-r-i+1)h(x)+ih(\beta)}{(n-r+1)}, \quad 1\le r<s\le n,\ i=1,2,\ldots,n-r. \tag{2.12}$$

and at $r=0$, $X_{0:n}=x=\alpha$ and $m=s-1$, we get

$$E[h(X_{i:n}) \mid X_{s:n}=y] = \frac{(s-i)h(\alpha)+ih(y)}{s},$$

as given by Khan and Abu-Salih [5], Franco and Ruiz [14], Khan and Abouammoh [2] and Khan and Athar [3].

Remark 2.2: At $p=1$, $i=r+1$ and $s=r+2$ in Theorem 2.1, we have

$$E[h(X_{r+1:n}) \mid X_{r:n}=x, X_{r+2:n}=y] = E[h(X) \mid x\le X\le y] = \frac{h(x)+h(y)}{2} \tag{2.13}$$

which may also be expressed as,

$$\frac{1}{(s-r-1)}\sum_{i=r+1}^{s-1} E[h(X_{i:n}) \mid X_{r:n}=x, X_{s:n}=y] = \frac{h(x)+h(y)}{2} \tag{2.14}$$

as obtained by Balasubramanian and Beg [12].

3. EXAMPLES

(i) **Power function distribution**

$F(x) = \theta^{-\nu} x^{\nu}, \quad 0 \le x \le \theta; \; \theta, \nu > 0.$

Then $F(x)$ is given by (2.2) with $a = \theta^{-\nu}$, $b = 0$ and $h(x) = x^{\nu}$.

(ii) **Pareto distribution**

$F(x) = 1 - \theta^{\nu} x^{-\nu}, \quad \theta \le x < \infty; \; \theta, \nu > 0.$

Then $F(x)$ is given by (2.2) with $a = -\theta^{\nu}$, $b = 1$ and $h(x) = x^{-\nu}$.

(iii) **Exponential distribution**

$F(x) = 1 - e^{-\theta x}, \quad 0 < x < \infty; \; \theta > 0.$

Then $F(x)$ is given by (2.2) with $a = -1$, $b = 1$ and $h(x) = e^{-\theta x}$.

(iv) **Rayleigh distribution**

$F(x) = 1 - e^{-\theta x^2}, \quad 0 < x < \infty; \; \theta > 0.$

Then $F(x)$ is given by (2.2) with $a = -1$, $b = 1$ and $h(x) = e^{-\theta x^2}$.

(v) **Weibull distribution**

$F(x) = 1 - e^{-\theta x^{\nu}}, \quad 0 < x < \infty; \; \theta, \nu > 0.$

Then $F(x)$ is given by (2.2) with $a = -1$, $b = 1$ and $h(x) = e^{-\theta x^{\nu}}$.

(vi) **Inverse Weibull distribution**

$F(x) = e^{-\theta x^{-\nu}}, \quad 0 < x < \infty; \; \theta, \nu > 0.$

Then $F(x)$ is given by (2.2) with $a = 1$, $b = 0$ and $h(x) = e^{-\theta x^{-\nu}}$.

(vii) **Extreme value distribution**

$F(x) = 1 - e^{-e^{x}}, \quad -\infty < x < \infty.$

Then $F(x)$ is given by (2.2) with $a = -1$, $b = 1$ and $h(x) = e^{-e^{x}}$.

(viii) **Cauchy distribution**

$F(x) = \frac{1}{\pi} \tan^{-1} x + \frac{1}{2}, \quad -\infty < x < \infty.$

Then $F(x)$ is given by (2.2) with $a = \frac{1}{\pi}$, $b = \frac{1}{2}$ and $h(x) = \tan^{-1} x.$

(ix) **Gumbel distribution**

$F(x) = e^{-e^{-x}}, \quad -\infty < x < \infty.$

Then $F(x)$ is given by (2.2) with $a = 1$, $b = 0$ and $h(x) = e^{-e^{-x}}$.

(x) **Burr type XII distribution**

$F(x) = 1 - (1 + \theta x^{k})^{-m}, \quad 0 < x < \infty; \; \theta, k, m > 0.$

Then $F(x)$ is given by (2.2) with $a=-1$, $b=1$ and $h(x)=(1+\theta x^k)^{-m}$.

Similarly, characterization results for other distributions may be obtained with proper choice of a, b and $h(x)$. One may refer to Khan and Athar [3].

4. Numerical Illustrations

Let us consider here the example of Mann and Fertig [15], who gave the failure times of airplane components for a life test in which 13 components were placed on test, with the test terminating after the $10-th$ failure. Failure times in hours are recorded as below

0.22, 0.50, 0.88, 1.00, 1.32, 1.33, 1.54, 1.76, 2.50, 3.00.

Now, if we calculate the expected values of the failure times of airplane components $\hat{X}_{2:10}, \hat{X}_{4:10}, \hat{X}_{6:10}$ and $\hat{X}_{8:10}$ on using (2.9), we get

					Distributions				
					Power ($\nu=1$)	**Pareto** ($\nu=1$)	**Exponential** ($\theta=1$)	**Rayleigh** ($\theta=1$)	**Weibull** ($\nu=0.5, \theta=0.5$)
n	r	i	s	$X_{i:n}$	$\hat{X}_{i:n}$	$\hat{X}_{i:n}$	$\hat{X}_{i:n}$	$\hat{X}_{i:n}$	$\hat{X}_{i:n}$
10	1	2	3	0.50	0.55	0.35	0.50	0.59	0.49
	1	2	4	0.50	0.48	0.30	0.42	0.52	0.40
	1	2	6	0.50	0.44	0.26	0.36	0.48	0.35
	1	2	8	0.50	0.44	0.25	0.34	0.44	0.33
	1	2	10	0.50	0.53	0.25	0.33	0.41	0.36
10	1	4	5	1.00	1.05	0.58	0.91	0.99	0.91
	2	4	6	1.00	0.92	0.73	0.83	0.86	0.84
	3	4	5	1.00	1.10	1.06	1.07	1.07	1.00
	3	4	9	1.00	1.15	0.99	1.02	0.98	1.06
	3	4	10	1.00	1.18	0.98	1.01	0.96	1.07
10	1	6	7	1.33	1.32	0.77	1.16	1.20	1.18
	2	6	8	1.33	1.34	0.96	1.15	1.11	1.21
	3	6	8	1.33	1.47	1.26	1.31	1.25	1.33
	5	6	7	1.33	1.43	1.42	1.42	1.42	1.43
	5	6	10	1.33	1.67	1.49	1.49	1.40	1.57
10	1	8	9	1.76	2.22	1.09	1.76	1.45	1.97
	2	8	9	1.76	2.21	1.59	1.85	1.48	2.04
	6	8	9	1.76	2.11	1.93	1.95	1.69	2.04
	7	8	9	1.76	2.02	1.90	1.91	1.74	1.97
	7	8	10	1.76	2.03	1.83	1.84	1.67	1.94
$MSE=\frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i-Y_i)^2$					0.0371	0.0615	0.0060	0.0070	0.0169

The Mean Sum of Square (MSE) of the expected values of the failure times of airplane components using exponential distribution is minimum. Thus, we can conclude that the exponential distribution gives the best fit.

Acknowledgments

The authors acknowledge with thanks to learned referees and Editor, *IJCTS* for their fruitful suggestions and comments which led the overall improvement in the manuscript.

References

[1] A. Dembińska and J. Wesolowski, "Linearity of regression for non-adjacent order statistics," Metrika, 1998, vol. 48, pp. 215-222.

[2] A.H. Khan and A.M. Abouammoh, "Characterizations of distributions by conditional expectation of order statistics," J. Appl. Statist. Sci., 2000, vol. 9, pp. 159-167.

[3] A.H. Khan and H. Athar, "Characterization of distributions through order statistics," J. Appl. Statist. Sci., 2004, vol. 13(2), pp. 147-154.

[4] A.H. Khan and M.M. Ali, "Characterization of probability distributions through higher order gap," Commun. Statist. Theory Meth., 1987, vol. 16, pp. 1281-1287.

[5] A.H. Khan and M.S. Abu-Salih, "Characterization of probability distributions by conditional expectation of order statistics," Metron, 1989, vol. 48, pp. 157-168.

[6] F. López-Blázquez and J.L. Moreno-Rebello, "A characterization of distributions based on linear regression of order statistics and record values," Sankhya Ser. A, 1997, vol. 59, pp. 311-323.

[7] H. Athar, M. Yaqub and H.M. Islam, "On characterization of distributions through linear regression of record values and order statistics, " Aligarh J. Statist., 2003, vol. 23, pp. 97-105.

[8] H.A. David and H.N. Nagaraja, Order Statistics. John Wiley, New York, 2003.

[9] H.N. Nagaraja, "Some characterizations of continuous distributions based on regressions of adjacent order statistics and record values," Sankhya Ser. A, 1988, vol. 50, pp. 70-73.

[10] J. Wesolowski and M. Ahsanullah "On characterizing distributions via linearity of regression for order statistics," Aust. J. Statist., 1997, vol. 39, pp. 69-78.

[11] K. Balasubramanian and A. Dey, "Distribution characterized through conditional expectations," Metrika, 1997, vol. 45, pp. 189-196.

[12] K. Balasubramanian and M.I. Beg, "Distributions determined by conditioning on a pair of order statistic," Metrika, 1992, vol. 39, pp. 107-112.

[13] M. Franco and J.M. Ruiz, "On characterization of continuous distributions with adjacent order statistics," Statistics, 1995, vol. 26, pp. 375-385.

[14] M. Franco and J.M. Ruiz, "On characterization of distributions by expected values of order statistics and record values with gap," Metrika, 1997, vol. 45, pp. 107-119.

[15] N.R. Mann and E.W. Fertig, "Tables for obtaining Weibull confidence bounds and tolerance bounds based on best linear invariant estimates of parameters of the extreme-value distribution," Technometrics, 1973, vol. 15, pp. 87-101.