Some Characterization of Continuous Distributions Based on Order Statistics

Haseeb Athar¹ and Zuber Akhter¹

¹Department 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 \( F(x) = ab(x) + b, \ x \in (\alpha, \beta) \) 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-Salah [5], Balasubramanian and Beg [12], Franco and Ruiz ([13], [14]), Balasubramanian and Dey [11], López-Blázquez and Moreno-Rebello [6], Wesolowski and Ahsanullah [10], Dębińska 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 \( F(x) = ab(x) + b, \ x \in (\alpha, \beta) \) 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 \((\alpha, \beta)\), where \( \alpha, \beta \) may be finite or infinite and let \( X_{1n} \leq X_{2n} \leq \ldots \leq X_{nn} \) be the corresponding order statistics. Then the conditional pdf of \( X_{ln} \) given \( X_{rn} = x \) and \( X_{sn} = y \), \( 1 \leq r < s \leq n \) is (David and Nagaraja [8])

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

(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 \((\alpha, \beta)\). Then for two consecutive values \( r \) and \( r+1, \ 1 \leq r < i-1 \leq s \leq n \),

\[
E[h(X_{ln}) - h(X_{rn})^p | X_{ln} = x, X_{rn} = y] = \frac{s!}{l!x_p(x, y)} \cdot \frac{\Gamma(p + i - l) \Gamma(s - l)}{\Gamma(p + s - l) \Gamma(l - i)} [h(y) - h(x)]^p
\]

(2.1)

if and only if

\[
F(x) = ab(x) + b
\]

(2.2)

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

http://journals.uob.edu.bh
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

\[
E[\bar{h}(X_{1n}) - h(X_{2n})]^{p} | X_{1n} = x, X_{2n} = y] = g_{r,i,s,p}(x,y)
\]

\[
= \frac{1}{(s-r)} \int_{0}^{1} (h(t) - h(x))^{p} \left[ F(t) - F(x) \right]^{s-r-1} \left[ F(y) - F(t) \right]^{s-i-1} \frac{f(t)}{F(y) - F(x)} \, dt.
\]

Let

\[
F(t) - F(x) = z,
\]

implying

\[
(h(t) - h(x))^{p} = z^{p} (h(y) - h(x))^{p}.
\]

Then, we have

\[
E[\bar{h}(X_{1n}) - h(X_{2n})]^{p} | X_{1n} = x, X_{2n} = y] = g_{r,i,s,p}(x,y)
\]

\[
= \frac{1}{(s-r)} \int_{0}^{1} (h(y) - h(x))^{p} z^{p+i-r-1} (1-z)^{s-i-1} \, dz,
\]

and hence the necessary part.

To prove the sufficiency part, consider

\[
E[\bar{h}(X_{1n}) - h(X_{2n})]^{p} | X_{1n} = x, X_{2n} = y] = g_{r,i,s,p}(x,y)
\]

or,

\[
\frac{1}{(s-r)} \int_{0}^{1} (h(t) - h(x))^{p}[F(t) - F(x)]^{i-r-1}[F(y) - F(t)]^{s-i-1} f(t) \, dt = g_{r,i,s,p}(x,y)[F(y) - F(x)]^{i-r-1}.
\]

(2.3)

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

\[
- ph'(x) \frac{1}{(s-r)} \int_{0}^{1} (h(t) - h(x))^{p-1}[F(t) - F(x)]^{i-r-1}[F(y) - F(t)]^{s-i-1} f(t) \, dt
\]

\[
= \frac{(i-r-1)}{(i-r) \Gamma(s-i)} \int_{0}^{1} (h(t) - h(x))^{p}[F(t) - F(x)]^{i-r-2}[F(y) - F(t)]^{s-i-1} f(t) \, dt
\]

\[
= \frac{\partial}{\partial x} \left[ g_{r,i,s,p}(x,y)[F(y) - F(x)]^{s-i-1} - (s-r-1) g_{r,i,s,p}(x,y)[F(y) - F(x)]^{s-r-1} f(x) \right]
\]

(2.4)

or,

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

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

(2.5)

Rearranging the terms of (2.5), we get

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

(2.6)

Now consider,

\[
ph'(x) g_{r,i,s,p-1}(x,y) + \frac{\partial}{\partial x} g_{r,i,s,p}(x,y)
\]
\[
= p(s-i)h'(x)[h(y) - h(x)]^p - 1 \frac{\Gamma(p+i-r-1) \Gamma(s-r)}{\Gamma(p+s-r-1) \Gamma(i-r)} - p(s-i)h'(x)[h(y) - h(x)]^p - 1 \frac{\Gamma(p+i-r) \Gamma(s-r)}{\Gamma(p+s-r) \Gamma(i-r)}
\]
and
\[
g_{r,i,s}(x,y) - g_{r+1,i,s}(x,y)
\]
\[
= (h(y) - h(x))^p \frac{\Gamma(p+i-r) \Gamma(s-r)}{\Gamma(p+s-r) \Gamma(i-r)} - (h(y) - h(x))^p \frac{\Gamma(p+i-r-1) \Gamma(s-r-1)}{\Gamma(p+s-r-1) \Gamma(i-r-1)}
\]
\[
= p(s-i)(h(y) - h(x))^p \frac{\Gamma(p+i-r-1) \Gamma(s-r-1)}{\Gamma(p+s-r) \Gamma(i-r)}.
\]
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
\[
E[h(X_{r+i}) | X_{r+i} = x, X_{s+i} = y] = \frac{(s-i)h(x) + (i-r)h(y)}{(s-r)}, \quad i = r+1, r+2, \ldots, s-1
\]
or equivalently
\[
E[h(X_{r+i}) | X_{r+i} = x, X_{s+i} = y] = \frac{(m-i+1)h(x) + ih(y)}{(m+1)}, \quad m = s-r-1, \quad 1 \leq r < s \leq n, \quad i = 1, 2, \ldots, m.
\]
as obtained by Khan and Athar [3].

Further, at \( s=n+1 \), \( X_{n+i} = y = \beta \) and \( m = n-r \), then we have
\[
E[h(X_{r+i}) | X_{r+i} = x, X_{r+i} = y] = \frac{(n-i+1)h(x) + (i-r)h(\beta)}{(n-r+1)}, \quad i = r+1, r+2, \ldots, n
\]
or,
\[
E[h(X_{r+i}) | X_{r+i} = x] = \frac{(n-r-i+1)h(x) + ih(\beta)}{(n-r+1)}, \quad 1 \leq r < s \leq n, \quad i = 1, 2, \ldots, n-r.
\]
and at \( r=0 \), \( X_{i} = x = \alpha \) and \( m = s-1 \), we get
\[
E[h(X_{r+i}) | X_{r+i} = y] = \frac{(s-i)h(\alpha) + ih(y)}{s},
\]
as given by Khan and Abu-Salah [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
\[
E[h(X_{r+1+i}) | X_{r+i} = x, X_{r+i+2} = y] = E[h(X) | x \leq X \leq y] = \frac{h(x) + h(y)}{2}
\]
which may also be expressed as,
\[
\frac{1}{(s-r-1)} \sum_{i=r+1}^{s-1} E[h(X_{i+r}) | X_{r+i} = x, X_{s+i} = y] = \frac{h(x) + h(y)}{2}
\]
as obtained by Balasubramanian and Beg [12].

3. EXAMPLES

(i) **Power function distribution**

\[ F(x) = \theta^{-\nu} x^\nu, \quad 0 \leq x \leq \theta; \, \theta, \nu > 0. \]

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

(ii) **Pareto distribution**

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

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**

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

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\textsuperscript{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 \( \hat{X}_{210}, \hat{X}_{430}, \hat{X}_{640}, \) and \( \hat{X}_{830} \) on using (2.9), we get

<table>
<thead>
<tr>
<th>( n )</th>
<th>( r )</th>
<th>( i )</th>
<th>( s )</th>
<th>( X_{in} )</th>
<th>( \hat{X}_{in} )</th>
<th>( \hat{X}_{in} )</th>
<th>( \hat{X}_{in} )</th>
<th>( \hat{X}_{in} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0.50</td>
<td>0.55</td>
<td>0.35</td>
<td>0.50</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0.50</td>
<td>0.48</td>
<td>0.30</td>
<td>0.42</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>0.50</td>
<td>0.44</td>
<td>0.26</td>
<td>0.36</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>0.50</td>
<td>0.44</td>
<td>0.25</td>
<td>0.34</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>10</td>
<td>0.50</td>
<td>0.53</td>
<td>0.25</td>
<td>0.33</td>
<td>0.41</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>1.00</td>
<td>1.05</td>
<td>0.58</td>
<td>0.91</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>1.00</td>
<td>0.92</td>
<td>0.73</td>
<td>0.83</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1.00</td>
<td>1.10</td>
<td>1.06</td>
<td>1.07</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>1.00</td>
<td>1.15</td>
<td>0.99</td>
<td>1.02</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>1.00</td>
<td>1.18</td>
<td>0.98</td>
<td>1.01</td>
<td>0.96</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>1.33</td>
<td>1.32</td>
<td>0.77</td>
<td>1.16</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>1.33</td>
<td>1.34</td>
<td>0.96</td>
<td>1.15</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>1.33</td>
<td>1.47</td>
<td>1.26</td>
<td>1.31</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>1.33</td>
<td>1.43</td>
<td>1.42</td>
<td>1.42</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>1.33</td>
<td>1.67</td>
<td>1.49</td>
<td>1.49</td>
<td>1.40</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>8</td>
<td>9</td>
<td>1.76</td>
<td>2.22</td>
<td>1.09</td>
<td>1.76</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8</td>
<td>9</td>
<td>1.76</td>
<td>2.21</td>
<td>1.59</td>
<td>1.85</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>1.76</td>
<td>2.11</td>
<td>1.93</td>
<td>1.95</td>
<td>1.69</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>1.76</td>
<td>2.02</td>
<td>1.90</td>
<td>1.91</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>8</td>
<td>10</td>
<td>1.76</td>
<td>2.03</td>
<td>1.83</td>
<td>1.84</td>
<td>1.67</td>
</tr>
</tbody>
</table>

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2
\]

\[
\begin{array}{cccccc}
\text{Power} & \text{Pareto} & \text{Exponential} & \text{Rayleigh} & \text{Weibull} \\
(\nu = 1) & (\nu = 1) & (\theta = 1) & (\theta = 1) & (\nu = 0.5, \ \theta = 0.5) \\
0.0371 & 0.0615 & 0.0060 & 0.0070 & 0.0169
\end{array}
\]

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