ProbStat Forum, Volume 09, October 2016, Pages 109–114
ISSN 0974-3235

ProbStat Forum is an e-journal.
For details please visit www.probstat.org.in

Characterisation of continuous probability distributions conditioned on a pair of record values

M J S Khan[a], Z Akhter[b], N Wahid[c], S Kumar[c]

[a] *Department of Statistics & Operations Research, Aligarh Muslim University, Aligarh, India.*
[b] *Department of Mathematics, Faculty of Physical Sciences, SGT University, Gurgaon, India.*
[c] *Babasaheb Bhimrao Ambedkar University, Lucknow, India.*

Abstract. In this article, a generalized family of continuous distributions have been characterized through conditional expectation of record values conditioned on a pair of non adjacent record values. Further some of its important deductions are also discussed.

1. Introduction

Record values were defined by Chandler (1952) as a model of successive extremes in a sequence of identically and independent random variables. Record values are often used as a model in extreme weather conditions, in reliability theory and in life insurance. For detailed survey on the record values, one may refer to Arnold et al. (1998) and Ahsanullah (2004).

Let $X_1, X_2, \ldots$ be a sequence of independently and identically distributed *(iid)* continuous random variables *(rv)* with the cumulative distribution function *(df)* and Lebesgue probability density function *(pdf)* over the support (α, β), where $\alpha = inf\{x : F(x) \geq 0\}$ and $\beta = sup\{x : F(x) \leq 1\}$. Define the upper record times by $U(1) = 1$ and $U(r) = \min\{k > U(r-1) : X_k > X_{U(r-1)}\}$,$r > 1$. The upper record value sequence is then defined by $X_{U(1)}, X_{U(2)}, \ldots$.

Let $X_{U(r)}$ be the r^{th} upper record from a continuous population with *df* $F(x)$ and the *pdf* $f(x)$ over the support (α, β), then the *pdf* of r^{th} upper record is given by (Ahsanullah; 2004)

$$f_r(x) = \frac{1}{(r-1)!}[-log\bar{F}(x)]^{r-1} f(x), \quad \alpha < x < \beta \tag{1}$$

Moreover, the joint *pdf* of $X_{U(r)}$ and $X_{U(s)}$, $1 \leq r < s \leq n$, is given by (Ahsanullah; 2004)

$$f_{r,s}(x, y) = \frac{1}{(r-1)!(s-r-1)!}[-log\bar{F}(x)]^{r-1}[B(x, y)]^{s-r-1}\frac{f(x)}{\bar{F}(x)}f(y), \quad \alpha < x < y < \beta, \tag{2}$$

2010 *Mathematics Subject Classification.* 62E10, 62G30
Keywords. Record values, Conditional expectation, Characterization, Probability distributions.
Received: 18 May 2015; Revised: 14 January 2016; Re-revised: 16 September 2016; Accepted: 02 October 2016
Email address: `author for correspondence: jahangirskhan@gmail.com` (M J S Khan)

where $\bar{F}(x) = 1 - F(x)$ and $B(x, y) = [-log\bar{F}(y) + log\bar{F}(x)]$.
The joint *pdf* of $X_{U(r)}, X_{U(j)}$ and $X_{U(s)},\ 1 \leq r < j < s \leq n$, can similarly be given as

$$f_{r,j,s}(x,t,y) = \frac{1}{(r-1)!(j-r-1)!(s-j-1)!}[-log\bar{F}(x)]^{r-1}[B(x,t)]^{j-r-1}[B(t,y)]^{s-j-1} \frac{f(x)}{\bar{F}(x)}\frac{f(t)}{\bar{F}(t)}f(y), \quad \alpha < x < t < y < \beta, \tag{3}$$

Hence, the conditional *pdf* of $X_{U(j)}$ given $X_{U(r)} = x$ and $X_{U(s)} = y,\ 1 \leq r < j < s \leq n$ is given by

$$f_{j|r,s}(t|x,y) = C_{r,j,s}\frac{[B(x,t)]^{j-r-1}[B(t,y)]^{s-j-1}}{[B(x,y)]^{s-r-1}}\frac{f(t)}{\bar{F}(t)},\ \alpha < x < t < y < \beta, \tag{4}$$

where $C_{r,j,s} = \frac{(s-r-1)!}{(j-r-1)!(s-j-1)!}$.

Characterization of probability distributions by regression condition through various models of ordered random variables were investigated by several authors. In this paper, we have characterized the continuous distribution function using the conditional expectation of record values conditioned on a pair of non-adjacent records. Probability distribution can be characterized in many ways and the method under study here is one of them. That is, if the regression equation is known, we can find the family of distributions P_F from which the observations have been drawn.

Characterization of distributions through conditional expectation of record values was first considered by Nagaraja (1988). Nagaraja (1988) obtained the characterization result based on the linear regression of adjacent record values. He characterized the distributions by considering the regression equation $g_{r+1|r}(x) = E[X_{U(r+1)}|X_{U(r)} = x] = ax + b$. Later Franco and Ruiz (1996, 1997), Lopez-Bláquez and Rebollo (1997), Ahsanullah and Wesolowski (1998), Dembińska and Wesolowski (2000) and Athar et al. (2003) extended the result of Nagaraja (1988) and characterized the continuous distributions conditioned on non-adjacent records. Characterization of continuous distributions conditioned on a pair of adjacent records was investigated recently by Bairamov et al. (2005). Bairamov et al. (2005) characterized the exponential distribution and continuous distributions by taking the monotone transformations. Further Yanev et al. (2008), Yanev and Ahsanullah (2009) and Khan and Khan (2009) characterized the continuous distributions conditioned on a pair of non-adjacent records. Recently Noor and Athar (2014) characterized the continuous distributions by taking the conditional expectation $g^p_{r,s}(x) = E[\{\Psi(X_{U(s)}) - \Psi(X_{U(r)})\}^p|X_{U(r)} = x]$. In this paper, we have extended the result of Noor and Athar (2014) and investigated the conditional expectation $g^p_{r,s}(x,y) = E[\{\Psi(X_{U(j)}) - \Psi(X_{U(r)})\}^p|X_{U(r)} = x, X_{U(s)} = y]$, conditioned on a pair of non-adjacent records.

The paper is divided into two sections. In section 2, we have deduced the characterization result for the continuous distributions conditioned on a pair of non-adjacent records. Finally in Table 3.1, the results obtained in section 2 are used to characterize some specific continuous distributions using monotone transformations.

2. Characterization of continuous distributions.

Theorem 2.1. *Let $X_{U(i)}, i = 1, 2, \dots$, be the i^{th} record from a continuous population with* df $F(x)$ *and* pdf *$f(x)$ over the support (α, β). Let $\Psi : \mathbb{R} \to \mathbb{R}$ is a monotonic and differentiable function and $p \in \mathbb{N}$. Then for $1 \leq l < j < s \leq n,\ l = r, r+1$*

$$g^p_{l,s}(x,y) = E[\{\Psi(X_{U(j)}) - \Psi(X_{U(l)})\}^p|X_{U(l)} = x, X_{U(s)} = y] = [\Psi(y) - \Psi(x)]^p\frac{\Gamma(s-l)\Gamma(p+j-l)}{\Gamma(j-l)\Gamma(p+s-l)} \tag{5}$$

if and only if

$$F(x) = 1 - e^{-[a\Psi(x)+b]}, \qquad \alpha \leq x \leq \beta, \tag{6}$$

where $g^p_{r,s}(x,y)$ *is a finite and differentiable function of* x *and* $\Gamma(.)$ *is a gamma function.*

Proof. Here first we shall prove that (2.2) implies (2.1), then

$$g^p_{r,s}(x,y) = \frac{C_{r,j,s}}{[B(x,y)]}\int_x^y [\Psi(t)-\Psi(x)]^p \left[\frac{B(x,t)}{B(x,y)}\right]^{j-r-1}\left[1-\frac{B(x,t)}{B(x,y)}\right]^{s-j-1}\frac{dF(t)}{\bar{F}(t)}$$

$$= \frac{C_{r,j,s}}{[\Psi(y)-\Psi(x)]}\int_x^y [\Psi(t)-\Psi(x)]^p \left[\frac{\Psi(t)-\Psi(x)}{\Psi(y)-\Psi(x)}\right]^{j-r-1}\left[1-\frac{\Psi(t)-\Psi(x)}{\Psi(y)-\Psi(x)}\right]^{s-j-1} d\Psi(t)$$

Set $u = \frac{\Psi(t)-\Psi(x)}{\Psi(y)-\Psi(x)}$, we get

$$g^p_{r,s}(x,y) = [\Psi(y)-\Psi(x)]^p \frac{\Gamma(s-r)\Gamma(p+j-r)}{\Gamma(j-r)\Gamma(p+s-r)}$$

To prove (2.1) implies (2.2), we have

$$g^p_{r,s}(x,y)[B(x,y)]^{s-r-1} = C_{r,j,s}\int_x^y [\Psi(t)-\Psi(x)]^p [B(x,t)]^{j-r-1}[B(t,y)]^{s-j-1}\frac{dF(t)}{\bar{F}(t)}$$

Differentiating both the sides *w.r.t.* x, we get

$$\frac{\partial}{\partial x}g^p_{r,s}(x,y)[B(x,y)]^{s-r-1} - g^p_{r,s}(x,y)(s-r-1)[B(x,y)]^{s-r-2}\frac{f(x)}{\bar{F}(x)}$$

$$= C_{r,j,s}\int_x^y \left[-p\Psi^{'}(x)[\Psi(t)-\Psi(x)]^{p-1}[B(x,t)]^{j-r-1}[B(t,y)]^{s-j-1}\right]\frac{dF(t)}{\bar{F}(t)}$$

$$-C_{r,j,s}\int_x^y \left[[\Psi(t)-\Psi(x)]^p(j-r-1)[B(x,t)]^{j-r-2}[B(t,y)]^{s-j-1}\frac{f(x)}{\bar{F}(x)}\right]\frac{dF(t)}{\bar{F}(t)}$$

This implies that

$$\frac{f(x)}{\bar{F}(x)B(x,y)} = \frac{p\Psi^{'}(x)g^{p-1}_{r,s}(x,y) + \frac{\partial}{\partial x}g^p_{r,s}(x,y)}{(s-r-1)[g^p_{r,s}(x,y) - g^p_{r+1,s}(x,y)]} \tag{7}$$

Now consider,

$$p\Psi^{'}(x)g^{p-1}_{r,s}(x,y) + \frac{\partial}{\partial x}g^p_{r,s}(x,y)$$

$$= p\Psi^{'}(x)[\Psi(y)-\Psi(x)]^{p-1}\frac{\Gamma(s-r)\Gamma(p+j-r-1)}{\Gamma(j-r)\Gamma(p+s-r-1)} - p\Psi^{'}(x)[\Psi(y)-\Psi(x)]^{p-1}\frac{\Gamma(s-r)\Gamma(p+j-r)}{\Gamma(j-r)\Gamma(p+s-r)}$$

$$= p(s-j)\Psi^{'}(x)[\Psi(y)-\Psi(x)]^{p-1}\frac{\Gamma(s-r)\Gamma(p+j-r-1)}{\Gamma(j-r)\Gamma(p+s-r)} \tag{8}$$

and

$$g^p_{r,s}(x,y) - g^p_{r+1,s}(x,y) = [\Psi(y)-\Psi(x)]^p\frac{\Gamma(s-r)\Gamma(p+j-r)}{\Gamma(j-r)\Gamma(p+s-r)} - [\Psi(y)-\Psi(x)]^p\frac{\Gamma(s-r-1)\Gamma(p+j-r-1)}{\Gamma(j-r-1)\Gamma(p+s-r-1)}$$

$$= p(s-j)[\Psi(y)-\Psi(x)]^p\frac{\Gamma(s-r-1)\Gamma(p+j-r-1)}{\Gamma(j-r)\Gamma(p+s-r)} \tag{9}$$

Therefore in view of (2.3), we have

$$\frac{f(x)}{\bar{F}(x)B(x,y)} = \frac{\Psi^{'}(x)}{[\Psi(y)-\Psi(x)]}$$

Integrating both the sides w .r .t. x, over (α, x), we get

$$-[logB(x,y) - logB(\alpha,y)] = -log[\Psi(y)-\Psi(t)]|_{\alpha}^{x}$$

or,

$$log\left[1 - \frac{log\bar{F}(x)}{log\bar{F}(y)}\right] = log\left[\frac{\Psi(y) - \Psi(x)}{\Psi(y) - \Psi(\alpha)}\right]$$

or,

$$\frac{-log\bar{F}(x)}{-log\bar{F}(y)} = \left[\frac{\Psi(x) - \Psi(\alpha)}{\Psi(y) - \Psi(\alpha)}\right]$$

This implies that, $-log\bar{F}(x) = a[\Psi(x) - \Psi(\alpha)] = a\Psi(x) + b$, where, $b = -a\Psi(\alpha)$. This proves the theorem. □

Theorem 2.2. *Let $X_{U(i)}, i = 1, 2, \dots,$ be the i^{th} record from a continuous population with* df $F(x)$ *and* pdf $f(x)$ *over the support (α, β). Let $\Psi : \mathbb{R} \to \mathbb{R}$ is a monotonic and differentiable function and $p \in \mathbb{N}$. Then for $1 \le r < j < l \le n,\ l = s - 1, s$*

$$\xi^p_{r,s}(x, y) = E[\{\Psi(X_{U(l)}) - \Psi(X_{U(j)})\}^p | X_{U(r)} = x, X_{U(l)} = y] = [\Psi(y) - \Psi(x)]^p \frac{\Gamma(l - r)\Gamma(p + l - j)}{\Gamma(l - j)\Gamma(p + l - r)} \quad (10)$$

if and only if

$$F(y) = 1 - e^{-[a\Psi(y)+b]}, \qquad \alpha \le y \le \beta, \quad (11)$$

provided that $\xi^p_{r,s}(x, y)$ is a finite and differentiable function of y and there exists a $q \in (\alpha, \beta)$ such that

$$q = inf\left[x : x \ge F^{-1}\left(\frac{e - 1}{e}\right)\right] \quad (12)$$

Proof. First we prove that (2.7) implies (2.6),
$\xi^p_{r,s}(x, y) = \frac{C_{r,j,s}}{[B(x,y)]} \int_x^y [\Psi(y) - \Psi(t)]^p \left[\frac{B(t,y)}{B(x,y)}\right]^{s-j-1} \left[1 - \frac{B(t,y)}{B(x,y)}\right]^{j-r-1} \frac{dF(t)}{\bar{F}(t)}$

$$= \frac{C_{r,j,s}}{[\Psi(y) - \Psi(x)]} \int_x^y [\Psi(y) - \Psi(t)]^p \left[\frac{\Psi(y) - \Psi(t)}{\Psi(y) - \Psi(x)}\right]^{s-j-1} \left[1 - \frac{\Psi(y) - \Psi(t)}{\Psi(y) - \Psi(x)}\right]^{j-r-1} d\Psi(t)$$

Set $v = \frac{\Psi(y)-\Psi(t)}{\Psi(y)-\Psi(x)}$, to get

$$\xi^p_{r,s}(x, y) = C_{r,j,s}[\Psi(y) - \Psi(x)]^p \int_0^1 v^{p+s-j-1}(1 - v)^{j-r-1} dv$$

$$\xi^p_{r,s}(x, y) = [\Psi(y) - \Psi(x)]^p \frac{\Gamma(s - r)\Gamma(p + s - j)}{\Gamma(s - j)\Gamma(p + s - r)}$$

To prove that (2.6) implies (2.7), note that

$$\xi^p_{r,s}(x, y)[B(x, y)]^{s-r-1} = C_{r,j,s} \int_x^y [\Psi(y) - \Psi(t)]^p [B(x, t)]^{j-r-1} [B(t, y)]^{s-j-1} \frac{dF(t)}{\bar{F}(t)}$$

On differentiating both the sides *w.r.t.* y, we get
$\frac{\partial}{\partial y}\xi^p_{r,s}(x, y)[B(x, y)]^{s-r-1} + \xi^p_{r,s}(x, y)(s - r - 1)[B(x, y)]^{s-r-2} \frac{f(y)}{\bar{F}(y)}$

$$= C_{r,j,s} \int_x^y \left[p\Psi^{'}(y)[\Psi(y) - \Psi(t)]^{p-1} [B(x, t)]^{j-r-1} [B(t, y)]^{s-j-1}\right] \frac{dF(t)}{\bar{F}(t)}$$

$$+C_{r,j,s} \int_x^y \left[[\Psi(y) - \Psi(t)]^p (s - j - 1)[B(x, t)]^{j-r-1} [B(t, y)]^{s-j-2} \frac{f(y)}{\bar{F}(y)}\right] \frac{dF(t)}{\bar{F}(t)}$$

implying that

$$\frac{f(y)}{\bar{F}(y)B(x,y)} = \frac{p\Psi^{'}(y)\xi^{p-1}_{r,s}(x,y) - \frac{\partial}{\partial y}\xi^{p}_{r,s}(x,y)}{(s-r-1)[\xi^{p}_{r,s}(x,y) - \xi^{p}_{r,s-1}(x,y)]} \tag{13}$$

Now consider,
$p\Psi^{'}(y)\xi^{p-1}_{r,s}(x,y) - \frac{\partial}{\partial y}\xi^{p}_{r,s}(x,y)$

$$= p\Psi^{'}(y)[\Psi(y)-\Psi(x)]^{p-1}\frac{\Gamma(s-r)\Gamma(p+s-j-1)}{\Gamma(s-j)\Gamma(p+s-r-1)} - p\Psi^{'}(y)[\Psi(y)-\Psi(x)]^{p-1}\frac{\Gamma(s-r)\Gamma(p+s-j)}{\Gamma(s-j)\Gamma(p+s-r)}$$

$$= p(j-r)\Psi^{'}(y)[\Psi(y)-\Psi(x)]^{p-1}\frac{\Gamma(s-r)\Gamma(p+s-j-1)}{\Gamma(s-j)\Gamma(p+s-r)} \tag{14}$$

and
$\xi^{p}_{r,s}(x,y) - \xi^{p}_{r,s-1}(x,y) = [\Psi(y)-\Psi(x)]^{p}\frac{\Gamma(s-r)\Gamma(p+s-j)}{\Gamma(s-j)\Gamma(p+s-r)} - [\Psi(y)-\Psi(x)]^{p}\frac{\Gamma(s-r-1)\Gamma(p+s-j-1)}{\Gamma(s-j-1)\Gamma(p+s-r-1)}$

$$= p(j-r)[\Psi(y)-\Psi(x)]^{p}\frac{\Gamma(s-r-1)\Gamma(p+s-j-1)}{\Gamma(s-j)\Gamma(p+s-r)} \tag{15}$$

Therefore in view of (2.9), we have

$$\frac{f(y)}{\bar{F}(y)B(x,y)} = \frac{\Psi^{'}(y)}{[\Psi(y)-\Psi(x)]}$$

Integrating both the sides $w.r.t.\ y$, over (y,q), we get

$$\frac{1+log\bar{F}(y)}{1+log\bar{F}(x)} = 1 - e^{\left[-\int_y^q \frac{[\Psi(t)}{\Psi(t)-\Psi(x)}dt\right]} = \frac{\Psi(q)-\Psi(y)}{\Psi(q)-\Psi(x)},$$

that is, $1 + log\bar{F}(y) = a_1[\Psi(q) - \Psi(y)] = a\Psi(y) + b$, where, $a = -a_1$ and $b = a_1\Psi(q)$. This proves the theorem. □

Table 2.1 shows that for particular choices of a, b and $\Psi(x)$ the following distribution can be characterized using Theorem 2.1 and Theorem 2.2.

Table 2.1: $F(x) = 1 - e^{-[a\Psi(x)+b]}$

Distribution	$F(x)$	a	b	$\Psi(x)$
Power function	$a^{-p}x^{p}$	1	$p\ loga$	$-log(a^p - x^p)$
Pareto	$1 - a^{p}x^{-p}$	p	$-p\ loga$	$log\ x$
Beta of I kind	$1 - (1-x)^p$	p	0	$-log(1-x)$
Exponential	$1 - e^{-\theta x}$	θ	0	x
Rayleigh	$1 - e^{-\theta x^2}$	θ	0	x^2
Weibull	$1 - e^{-\theta x^p}$	θ	0	x^p
Extreme value II	$1 - e^{e^{\theta x}}$	1	0	$e^{\theta x}$
Burr Type XII	$1 - (1+\theta x^p)^{-m}$	m	0	$log(1+\theta x^p)$

3. Applications

A probability distribution can be characterized in many ways and the method under study here is one of them. We have used here the conditional expectation of record statistics conditioned on a pair of non-adjacent records to characterize the probability distribution. That is, we have characterized the probability distribution if the regression equation truncated from both sides is given, i. e. the data are truncated from

left side at x and truncated from right side at y. In real practice, several times we get the data of which observations are missing either in beginning or in the end. In such type of data we can use the result of this paper, i.e. when the data are in form of records and when the data are missing at both.
Keeping this in view, we have characterized probability distributions through conditional expectation conditioned on a pair of non-adjacent records.

Acknowledgements The authors would like to thanks the referee's for carefully reading the paper and for helpful suggestions which greatly improved the paper. First author is also thankful to University Grant Commission for awarding UGC-BSR start up grant ((No. F.30-90/2015(BSR)).

References

[1] Ahsanullah, M. (2004), *Record Values - Theory and Applications.* University Press of America, Lanham, Maryland.
[2] Ahsanullah, M. and Wesolowski, J. (1998), Linearity of best predictors for non-adjacent record values. *Sankhyā Ser. B*, **60**, 221-227.
[3] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1998), *Records.* John Wiley & Sons, New York.
[4] Athar, H., Yaqub, M. and Islam, H.M. (2003), On characterization of distributions through linear regression of record values and order statistics. *Aligarh J. Statist.*, **23**, 97-105.
[5] Bairamov, I., Ahsanullah, M. and Pakes, A.G. (2005), A characterization of continuous distributions via regression on pairs of record values. *Aust. N. Z. J. Stat.*, **47**, 543-547.
[6] Chandler, K.N. (1952), The distribution and frequency of record values. *J. Roy. Statist. Soc. Ser. B*, **14**, 220-228.
[7] Dembińska, A. and Wesolowski, J. (2000), Linearity of regression for non adjacent record values. *J. Statist. Plann. Inference*, **90**, 195-205.
[8] Franco, M. and Ruiz, J.M. (1996), On characterization of continuous distributions by conditional expectation of record values. *Sankhyā Ser.* A, **58**, 135-141.
[9] Franco, M. and Ruiz, J.M. (1997), On characterization of distributions by expected values of order statistics and record values with gap. *Metrika*,**45**, 107-119.
[10] Khan, A.H. and Khan, M.J.S. (2009), On characterization of continuous distributions conditioned on a pair of non-adjacent records. *J. Appl. Probab. Statist.*, **4(1)**, 65-75.
[11] Lopez-Blázquez,F. and Moreno-Rebollo, J. L. (1997), A characterization of distributions based on linear regression of order statistics and record values. *Sankhyā Ser. A*, **59**, 311-323.
[12] Nagaraja, H. N. (1988), Some characterizations of continuous distributions based on regressions of adjacent order statistics and record values. *Sankhyā Ser. A*, **50**, 70-73.
[13] Noor, Z. and Athar, H. (2014), Characterization of probability distributions by conditional expectation of function of record statistics. *J. Egypt. Math. Soc.*, **22**, 275-279.
[14] Yanev, G. P. and Ahsanullah, M. (2009), On Characterizations based on regression of linear combinations of record values. *Sankhyā Ser. A* **71(1)**, 109-121.
[15] Yanev, G.P., Ahsanullah, M. and Beg, M.I. (2008), Characterizations of probability distributions via bivariate regression of record values. *Metrika*, **68**, 51-64.