International Journal of Mathematical Archive-3(8), 2012, 2997-3003
Available online through www.ijma.info ISSN 2229 – 5046

Identifying A Certain Class Of Distributions Using Some Recurrence Relations

Ali A. A-Rahman*
Associate Professor, Institute of Statistical Studies and Research, Cairo University, Egypt

(Received on: 18-07-12; Revised & Accepted on: 05-08-12)

ABSTRACT

In this paper, three recurrence relations for a certain class of probability distributions are presented. The first one is a recurrence relation between conditional moments of h(X) given X < y. The second is the relationship between the moments $E(h^m(Y_{k+1}))$, $E(h^m(Y_k))$ and $E(h^{m-1}(Y_k))$, where Y_k is the k^{th} order statistic from a sample of size n. The last one is the relationship between the conditional moments $E(h^m(Y_k)|Y_k<t)$ and $E(h^{m-1}(Y_k)|Y_k<t)$. Some results concerning exponentiated Weibull, modified Weibull, exponentiated Pareto, inverse Weibull, inverse Rayleigh, linear failure rate distribution, Burr, power and uniform distributions are obtained as special cases.

Keywords: *Characterization, right truncated moments, order statistics, recurrence relations, exponentiated Weibull, expnentiated Pareto, modified Weibull, inverse Weibull, inverse Rayleigh, linear failure rate, Burr, Power, beta, uniform distributions.*

1. INTRODUCTION

Characterization theorems are located on the borderline between probability theory and mathematical statistics and utilize numerous classical tools of mathematical analysis such as complex variable and differential equations. Some excellent references are Azlarove and Volodin [7], Galambos and Kotz [10], Kagan, Linnik and Rao [13]and Mchlachlan and Peel[17], among others. Several tools have been used to characterize the probability distributions. Gupta [12] ,Ouyang [19], Talwalker [21], have used the concept of right truncated moments to identify different distributions like Weibull, exponential, pareto, and power distributions. In fact characterizations by right truncated moments are very important in practice since, for example, in reliability studies some measuring devices may be unable to record values greater than time t. On the other hand characterizations of some particular distributions based on conditional moments of order statistics have been considered by several authors such as Pakes et al [20], Wu and Ouyang [22], Ahsanullah and Nevzorov [4], Asadi et al [6] and Govindarajulu [11], among others.

Let X be a continuous random variable with distribution function F(x) defined by:

$$F(x) = (d - h(x))^c, \; x \in (a, b) \quad (1.1)$$

Such that:
(1) d and c are constants such that $c \notin \{-1, 0\}$.

(2) h(X) is a real valued differentiable function defined on (a, b) with
(a) $\lim_{x\to a^+} h(x) = d$ and
(b) $\lim_{x\to b^-} h(x) = d-1$
(c) E(h(X)) exists and finite.

It is easy to see that several wellknown distributions (like exponentiated Weibull, exponentiated Pareto, Burr, Power,…etc) arise from the above family by suitable choices for the function h(x), the values of the parameters d and c and the domain (a ,b) .

2. MAIN RESULTS

A recurrence relation is a relation in which the function under consideration, S_n , is defined in terms of a smaller value of n. Recurrence relations play a vital role in statistics. In fact, a recurrence relation together with some initial conditions define a unique function. On the other hand, recurrence relations can be used to reduce the number of operations required to obtain a general form for the function under consideration. This has motivated several authors to use this concept to identify some probability distributions (see, e.g., Al-Hussaini et al. [5], Ahmad [2], Lin [15], Khan et. al. [14] and Fakhry [9]).

Corresponding author: *Ali A. A-Rahman**
Associate Professor, Institute of Statistical Studies and Research, Cairo University, Egypt

The following Theorem identifies the distribution defined by (1.1) using a recurrence relation between conditional moments of h^m (X), m = 1, 2 …given X<y.

Theorem 2.1. Let X be a continuous random variable with distribution function F(x), and density function f(x) such that F(a) = 0 and F(b) =1 and F(.) has continuous first order derivative on (a, b) with $\acute{F}$ (x) > 0 for all x. Then X has the distribution defined by (1.1) if and only if for any finite number y ∈ (a, b) and any natural number m= 1, 2 … the following relation is satisfied:

$$E(h^m (X) |X < y) = \frac{c}{c+m} h^m(y) + \frac{m\, d}{c+m} E (h^{m-1}(X)|X < y), \tag{2.1}$$

where h (X) is defined as before.

Proof. Necessity

By definition

$$E(h^m(X)|X<y) = \frac{\int_a^y h^m(x)\, dF(x)}{F(y)}$$

Integrating by parts, and noting that F (a) = 0, one gets:

$$E(h^m (X) | X < y) = h^m(y) - \frac{m}{F(y)} \int_a^y \acute{h}(x)\, h^{m-1}(x)\, F(x)\, d x \tag{2.2}$$

It is easy to see that:

$$\acute{h}(x) = \frac{-(d-h(x))f(x)}{c\, F(x)}, \tag{2.3}$$

the integral on the right can be written as follows:

$$I = \int_a^y \acute{h}(x)\, h^{m-1}(x)\, F(x)\, d x = \frac{-1}{c} \int_a^y (d-h(x))\, h^{m-1}(x)\, f(x)\, dx$$

$$= \frac{1}{c} \int_a^y h^m(x)\, f(x)\, d x - \frac{d}{c} \int_a^y h^{m-1}(x)\, f(x)\, d x$$

Substituting this result in equation (2.2), one gets:

$$E(h^m(X)|X<y)= h^m(y) - \frac{m}{c} E(h^m(X)|X<y) + \frac{md}{c} E(h^{m-1}(X)|X< y).$$

Solving this equation for $E(h^m(X)|X<y)$, we get:

$$E (h^m(X) |X < y) = \frac{c}{c+m} h^m(y) + \frac{m\, d}{c+m} E (h^{m-1}(X) | X < y).$$

Sufficiency.

Equation (2.1) can be written as an equation of the unknown function F(y) as follows:

$$\int_a^y h^m(x)\, f(x)\, d x = \frac{c}{c+m} h^m(y)\, F(y) + \frac{m\, d}{c+m} \int_a^y h^{m-1}(x)\, f(x)\, d x$$

Differentiating both sides with respect to y, dividing both sides by h^{m-1} (y), we get:

$$h(y)\, f(y) = \frac{c}{c+m} h(y)\, \acute{F}(y) + \frac{cm}{c+m} \acute{h}(y)\, F(y) + \frac{md}{c+m} f(y)$$

Recalling that f(y) = $\acute{F}(y)$, cancelling out $\frac{c}{c+m}$ h(y) f(y) from both sides, multiplying the result by $\frac{c+m}{m}$, adding to both sides –[h(y)f(y) + c $\acute{h}(y)F(y)$],and multiplying the result by $[(d-h(y))f(y)]^{-1}$, we get:

$$\frac{f(y)}{F(y)} = \frac{-c\, \acute{h}(y)}{d-h(y)}$$

Integrating both sides with respect to y from x to b and using the fact that F (b) =1, we get:

F(x) = $[\,d - h(x)]^{c}$,

Remarks (2.1).

(1) If we put m =1 in Theorem (2.1), we obtain Ouyang's result [19].
(2) If we put m = 1, h(X) =$\frac{-Z(X)}{Z(b)-\frac{g(k)}{1-n(k)}}$, d = $-\frac{g(k)/[(1-n(k)]}{Z(b)-\frac{g(k)}{1-nk)}}$, c = $\frac{n(k)}{1-n(k)}$, where g(·) and n(·) are finite real valued functions of k and Z(X) is a differentiable function such that:

$\lim_{x\to a^+} Z(x)$= $\frac{g(k)}{1-n(k)}$ and $\lim_{x\to b^-} Z(x)$ = Z(b) then we get Talwalker's result [21]

$$F(x) = \left[\frac{\frac{g(k)}{1-n(k)} - Z(x)}{\frac{g(k)}{1-n(k)} - Z(b)}\right]^{n(k)/[1-n(k)]} , \quad x \in (a, b),$$

Iff

E(Z(X)|X<y) = n(k)Z(y) + g(k).

The following Theorem identifies the distribution (1.1) using a recurrence relation between moments of some function of the k^{th} and $(k-1)^{th}$order statistics.

Theorem (2.2). Let X be an absolutely continuous random variable with cumulative distribution function F(·),survival function G(·) and density function f(·). Let $X_1, X_2, \ldots, X_n$ be a random sample from F(·). Denote by $Y_1 < Y_2 < \ldots < Y_n$ the corresponding ordered sample. Then under the same conditions posed on the function h(·) , the random variable X has the distribution defined by equation (1.1) iff for any natural number m, the following recurrence relation is satisfied:

$$E\,(h^m(Y_{k+1})) = \frac{m+kc}{kc}\, E(h^m(Y_k)) - \frac{md}{kc}\, E(h^{m-1}(Y_k)), \qquad k=1,2,\ldots,n-1 \tag{2.4}$$

Proof. Necessity

The density function of the k^{th} order statistic is given by:

$f_n(y_k) = \alpha_{k:n}$ f(y) $F^{k-1}(y)\ G^{n-k}$ (y), where $\alpha_{k:n} = \frac{n!}{(k-1)!(n-k)!}$

Then by definition we have:

$$E(h^m(y_k)) = \alpha_{k:n} \int_a^b h^m(y)\, f(y)\, F^{k-1}(y)\, G^{n-k}(y)\, dy$$
$$= \frac{\alpha_{k:n}}{k} \int_a^b h^m(y)\ [G(y)]^{n-k} d\,[F(y)]^k\ )$$

Integrating by parts, recalling that f(y) = - $\grave{G}$(y) and noting that F(a) = G(b) =0, one gets:

$$E(h^m(y_k)) = \frac{-m\,\alpha_{k:n}}{k} \int_a^b h^{m-1}(y)\,[G(y)]^{n-k}\,[F(y)]^k \grave{h}(y)dy + \frac{n-k}{k}\ \alpha_{k:n} \int_a^b h^m(y)\, f(y)\, [F(y)]^k [G(y)]^{n-k-1}dy$$

Making use of equation (2.3) to eliminate $\grave{h}$(y), we can write E ($h^m(Y_k)$) as follows:

$$E\,(h^m(y_k)) = \frac{m\,\alpha_{k:n}}{kc} \{-\int_a^b h^m(y)\, f(y)\, [F(y)]^{k-1}(y)\, [G(y)]^{n-k}(y)\, dy + d\int_a^b h^{m-1}(y)\, f(y)[F(y)]^{k-1}\, [G(y)]^{n-k}dy\} + \frac{(n-k)\alpha_{k:n}}{k} \int_a^b h^m(y)\, f(y)\, [F(y)]^k [G(y)]^{n-k-1} d\,y$$

$$= \frac{-m\,\alpha_{k:n}}{kc} \int_a^b h^m(y)\, f(y)\, [G(y)]^{n-k}\, [F(y)]^{k-1} d\,y + \frac{m\,d\,\alpha_{k:n}}{kc} \int_a^b h^{m-1}(y)\, f(y)\, [G(y)]^{n-k}\, [F(y)]^{k-1}dy + \frac{(n-k)\alpha_{k:n}}{k} \int_a^b h^m(y)\, f(y)\, [G(y)]^{n-(k+1)}\, [F(y)]^{(k+1)-1}\, d\,y$$

$$= \frac{-m\,\alpha_{k:n}}{kc}\, \frac{E(h^m(Y_K))}{\alpha_{k:n}} + \frac{md\ \alpha_{k:n}}{kc}\, \frac{E(h^{m-1}(Y_k))}{\alpha_{k:n}} + \frac{(n-k)\alpha_{k:n}}{k}\, \frac{E(h^m(Y_{k+1}))}{\alpha_{k+1:n}}$$

Solving the last equation for E ($h^m(y_{k+1})$), we get:

$$E(h^m(y_{k+1})) = \frac{(m+kc)}{kc} E(h^m(y_k)) - \frac{m\,d}{kc} E(h^{m-1}(y_k))$$

Sufficiency

Equation (2.4) can be written in integral form as follows:

$$\frac{n!}{k!(n-k-1)!}\int_a^b h^m(y)\, f(y)\, [F(y)]^k [G(y)]^{n-k-1} dy = \frac{n!\,(m+kc)}{(k-1)!(n-k)!kc}\int_a^b h^m(y) f(y)\, [F(y)]^{k-1}[G(y)]^{n-k} dy - \frac{n!\; md}{(k-1)!(n-k)!kc}\int_a^b h^{m-1}(y) f(y) [F(y)]^{k-1}[G(y)]^{n-k} dy \qquad (2.5)$$

Consider the integral on the left side, noting that f(y) = - $\acute{G}$(y), we get:

$$I= \int_a^b h^m(y) f(y)\, [F(y)]^k [G(y]^{n-k-1} dy = -\int_a^b h^m(y)\, [F(y)]^k \; \frac{d\,[G(y)]^{n-k}}{n-k}$$

Integrating by parts, recalling that f(y) = $\acute{F}$(y) and making use of the facts F(a) = G(b)=0, we get:

$$I = \frac{m}{n-k}\int_a^b h^{m-1}(y)\; \acute{h}(y)\; [G(y)]^{n-k}[F(y)]^k dy + \frac{k}{n-k}\int_a^b h^m(y) f(y) [F(y)]^{k-1}[G(y)]^{n-k} dy$$

Substituting this result in equation (2.5), multiplying both sides by $\frac{(n-k)!}{n!}$ (k-1)! , cancelling out

$\int_a^b h^m(y)\, f(y)\, [F(y)]^{k-1}\; [G(y)]^{n-k} dy$ from both sides then multiplying the result by $\frac{kc}{m}$, we get :

$$c\int_a^b h^{m-1}(y)\, \acute{h}(y)\, [G(y)]^{n-k}\; [F(y)]^k dy = \int_a^b h^m(y)\, f(y)\, [G(y)]^{n-k}\; [F(y)]^{k-1} dy - d\int_a^b h^{m-1}(y) f(y)\, [F(y)]^{k-1}[G(y)]^{n-k}\, dy$$

Therefore,

$$\int_a^b h^{m-1}(y)\, [F(y)]^{k-1}[G(y)]^{n-k}\left[\big(d-h(y)\big)f(y) + c\acute{h}(y)F(y)\right]dy = 0$$

Using the Munt- Szasz theorem (see Boas [8]), one gets:

(d-h(y)) f(y) +c $\acute{h}$(y) F(y) =0

Adding to both sides – [c $\acute{h}$(y) F(y)], then dividing both sides by [d –h(y)] F(y), one gets:

$$\frac{f(y)}{F(y)} = \frac{-c\,\acute{h}(y)}{d-h(y)}$$

Integrating both sides from x to b and using the fact that F (b) =1, we get:

F(x) = $[d-h(x)]^c$

The proof is complete.

Remarks (2.2).
(1) Set k=n-1 in equation (2.4), we have a recurrence relation concerning the maximum.
(2) Set n=2r+1, and k= r+1, we obtain a recurrence relation concerning the median.
(3) Set k =1 in equation (2.4), we get a recurrence relation concerning the minimum.

The next Theorem gives a recurrence relation between conditional moments of $h^m(Y_k)$ given Y_k < t

Theorem (2.3).Let X be an absolutely continuous random variable with cumulative distribution function F(·),and density function f(·). Let $X_1, X_2, \ldots, X_n$ be a random sample from F(·). Denote by $Y_1 < Y_2 < \ldots < Y_n$ the corresponding ordered sample. Then under the same conditions posed on the function h(·), the random variable X has the distribution defined by equation (1.1) iff for any natural number m, the following recurrence relation is satisfied :

$$E(h^m(Y_k)|\, Y_k < t) = \frac{kc}{m+kc}\; h^m(t) + \frac{md}{kc+m}\; E(h^{m-1}(Y_k)|Y_k < t)\,, \qquad k=1,2,\ldots,n \qquad (2.6)$$

Proof. Necessity.

The conditional density function of the k^{th} order statistic $Y_k| Y_k$ <t (see, Ahsanullah [3]) is given by:

$$f_n(Y_k|Y_k<t) = \frac{k}{[F(t)]^k} f(y)[F(y)]^{k-1}, \quad y\in(a, t) \tag{2.7}$$

Therefore

$$E\,(h^m(Y_k)|Y_k<t) = \frac{k}{[F(t)]^k} \int_a^t h^m(y)\, f(y)\, [F(y)]^{k-1} dy =$$
$$= \frac{1}{[F(t)]^k} \int_a^t h^m(y)\, d[F(y)]^k$$

Integrating by parts, noting that F (a) = 0, we get:

$$E(h^m(Y_k)|Y_k<t) = h^m(t) - \frac{m}{[F(t)]^k} \int_a^t h^{m-1}(y)\, \acute{h}(y)\, [F(y)]^k dy$$

Using equation (2.3) to eliminate $\acute{h}$(y) from the 2^{nd} term, one gets:

$$E\,(h^m(Y_k)|Y_k< t) = h^m(t) - \frac{m}{c[F(t)]^k}\int_a^t h^m(y)\, f(y)\, [F(y)]^{k-1}dy + \frac{md}{c[F(t)]^k}\int_a^t h^{m-1}(y)\, f(y)[F(y)]^{k-1}dy$$
$$= h^m(t) - \frac{m}{kc} E(h^m(Y_k)|Y_k<t) + \frac{md}{kc} E(h^{m-1}(Y_k)|Y_k<t)$$

Solving the last equation for E($h^m(Y_k)|Y_k$< t), one gets:

$$E(h^m(Y_k)|Y_k<t) = \frac{kc}{m+kc}\, h^m(t) + \frac{md}{m+kc}\, E(h^{m-1}(Y_k)|Y_k<t)$$

Sufficiency.

Equation (2.6) can be written in the following integral form:

$$k\,(m+ k\, c) \int_a^t h^m(y)f(y)[F(y)]^{k-1}dy = k\, c\, h^m(t)[F(t)]^k + m\, d\, k \int_a^t h^{m-1}(y)f(y)[F(y)]^{k-1}dy$$

Differentiating both sides with respect to t, dividing both sides by k $h^{m-1}(t)\, F^{k-1}(t)$, we get:

(m+kc) h(t) f(t)= ckh(t) f(t) + cm $\acute{h}$(t)F(t)+ mdf(t)

Cancelling out ck h(t) f(t) from both sides, adding –m h(t) f(t) – cm $\acute{h}$(t) F(t) to both sides, and multiplying the result by $\frac{1}{m[d-h(t)]F(t)}$, one gets:

$$\frac{-c\acute{h}(t)}{d-h(t)} = \frac{f(t)}{Ft)}$$

Integrating both sides with respect to t from x to b , and recalling that F(b) = 1, we get:

F(x) = $[d - h(x)]^c$

The proof is complete.

Remarks (2.3).

(1) If we set k =1, we obtain a recurrence relation for the minimum. Moreover, using equation (2.7), we have:

$E(h^m(Y_1)|Y_1<t) = E(h^m(X)|X<t)$

Therefore, we can say that Theorem (2.3) generalizes Theorem (2.1).

(2) If we put k=n, we obtain a recurrence relation for the maximum.
(3) If we put n=2r+1 and k= r+1, we obtain a recurrence relation for the median.

GENERAL COMMENTS

In all of the foregoing theorems, several results can be picked out for some wellknown distributions by suitable choices for the function h(X), the values of the parameters d and c and the domain (a, b) as follows:

(1) If we set h(X)= exp – $[\frac{X}{\beta}]^{\alpha}$, d=1,c=$\theta$,a=0 and b=∞, we obtain recurrence relations concerning the exponentiated Weibull distribution with positive parameters α,β and θ(see,e.g,. Nassar and Eissa [18]). For θ=1, we have Weibull distribution. For θ=1 and α=1, we have the exponential distribution.
(2) If we set h(X)= exp – [αX+βX^{λ}], d=1,c=1,a=0 and b=∞ , we obtain recurrence relations concerning the modified Weibull distribution with parameters α,β ≥0,where α + β >0 , and λ >0 .For λ=2, we have the linear failure rate distribution.(see, e.g., Zaindin and Sarhan [23]).
(3) If we set h(X) = - exp – $[\frac{\beta}{X}]$, d=0, c=θ > 0,a=0 and b=∞, we obtain recurrence relations concerning the inverse Weibull distribution with positive parameters β and θ. For c=2, we have inverse Rayleigh distribution.
(4) If we set h(X)= $[\frac{\beta}{X}]^{\alpha}$, d=1, c=θ >0,a=β, b=∞ , we obtain recurrence relations concerning exponentiated Pareto of the first type with positive parameters β, α and θ(see, e.g., Massom and Woo [16].For θ=1, we have Pareto distribution of the first type.
(5) If we set h(X) = $[1+X]^{-\alpha}$, α>0,d=1,c=θ>0, a=0 and b=∞ , we obtain recurrence relations concerning the exponentiated Pareto of the second type (see, e.g., Abu- Zinadah [1]). For θ=1, we have the Pareto distribution of the second type.
(6) If we set h(X)= -X, d=0, c>0, a=0 and b=1, we obtain recurrence relations concerning the power distribution with parameter c. For c=1, we have the uniform distribution.
(7) If we set h(X) = $[1-X]^{\theta}$, c=1, d=1,a=0 and b=1, we obtain recurrence relations concerning beta distribution with parameters 1, θ.
(8) If we set h(x) = $[1+x^{\alpha}]^{-\theta}$, α, θ > 0, d=1, c=1, a=0 and b =∞ , we obtain recurrence relations concerning Burr distribution. For α = 1, the Pareto distribution is obtained.

REFERENCES

(1) Abou-Zinadah, H. (2010). A study on Mixture of exponentiated Pareto and exponential distributions. Journal of Applied Sciences Research, 6,358-376.
(2) Ahmad, A. (1998).Recurrence relations for single and product moments of record values from Burr type XII distribution and a characterization. J. Applied Statistical Science, 7, 7-15.
(3) Ahsanullah, M.(2009).On some characterizations of univariate distributions based on truncated moments of order statistics. Pakistan J. of Statistics, 25, 83-91.
(4) Ahsanullah,M. and Nevzorov,V. (2001). Orderd Random Variables. Nova, New York.
(5) Al-Hussaini, E., Abdel-Baset, A. and Al-Kashif, M.(2005). Recurrence relations for moment and conditional moment generating functions of generalized order statistics.Metrika, 61, 199-220.
(6) Asadi, M., Rao, C. and Shanbhag,D. (2001).Some unified characterization results on generalized Pareto distributions.J. Statistical Planning Inference, 93, 29-50.
(7) Azlarov,T. and Volodin, N.(1986). Characterization Problems Associated with The Exponential Distribution Springer Verlag Berlin.
(8) Boos, R.(1954).Entire Functions. Academic Press.
(9) Fakhry, E.(1993).Some characteristic properties of a certain family of probability distributions. The Annual Conference, Institute Of Statistical Studies And Research, Cairo University, Egypt,28, 139-147.
(10) Galambos,J. and Kotz, S.(1978).Characterization of Probability Distributions Springer- Verlag Berlin.
(11) Govindarajulu, Z. (2001). Characterization of double exponential distribution using moments of order statistics. Communication of Statistics. Theory and Methods, 30, 2355-2372.
(12) Gupta, P. (1985).Some Characterizations of distributions by truncated moments. Statistics, 16,445-473.
(13) Kagan, A., Linnik, K., and Rao, C. (1973). Characterization Problems in Mathematical Statistics. John Wiley and Sons. New York.
(14) Khan, A., Yaqub, M. and Parvez, S. (1983).Recurrence relations between moments of order statistics, Naval Res. Logist. Quart., 30.419-441.
(15) Lin, G. (1988). Characterization of distributions via relationships between two moments of order statistics. J. Statistical planning inference, 19, 73-80.
(16) Masoom, M. and Woo, J. (2010). Estimation of tail-probability and reliability in exponentiated Pareto case. Pakistan J. of Statistics, 26, 39-47.
(17) Mclachlan, G. and Peel, D. (2002).Finite Mixture Models. Wiley, New York.
(18) Nassar, M. and Eissa, F. (2003). On the exponentiated Weibull distribution Communications in Statistics, Theory and Methods, 32, 1317-1336.
(19) Ouyang, L. (1983). On characterization of distributions by conditional expectations. Tamkang J. Management Science, 4, 13-21.

(20) Pakes, A., Fakhry, E., Mahmoud, M. and Ahmad, A. (1996). Characterizations by regressions of order statistics. J. Applied Statistical Science, 3, 11-23.
(21) Talwalker, S. (1977). A note on characterization by conditional expectation. Metrika, 24,129-136.
(22) Wu, J. and Ouyang,L. (1996).On Characterizing distributions by conditional expectations of functions of order statistics.Metrika,34,135-147.
(23) Zaindin, M. and Sarhan, A. (2009). Parameters estimation of the Modified Weibull distribution. Applied Mathematical Sciences, 3, 541-550.

Source of support: Nil, Conflict of interest: None Declared