Int. J. Contemp. Math. Sciences, Vol. 7, 2012, no. 44, 2167 – 2177

On Characterization of a Certain Family of Distributions Based on Some Recurrence Relations

Ali A. A-Rahman

Institute of Statistical Studies and Research
Department of Mathematical Statistics
Cairo University, Egypt
draliahmed2001@yahoo.com

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($h^m(Y_k)$), E($h^m(Y_{k-1})$) 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 Modified Weibull, Weibull, Rayleigh, exponential, Linear failure rate, 1^{st}type Pearsonian distributions, Burr, Pareto power and uniform distributions are obtained as special cases.

Keywords: Characterization, left truncated moments, order statistics, recurrence relations, Modified Weibull, Rayleigh, exponential, Linear failure rate, 1^{st}type Pearsonian distribution, Burr, Pareto, Power, beta, uniform distributions.

1. Introduction

There is no denying that over the past few decades there has been an increasing interest in characterizing different types of probability distributions. Some excellent references are, e.g., Azlarov and Volodin [23] , Galambos and Kotz [12] , Kagan , Linnik and Rao [5] and Mchlachlan and Peel[11], among others.

Different methods have been used to identify several types of distributions. Osaki and Li [21], Ahmed [6],Dimaki and Xekalaki [8], and Chou and Huang [7], among others, have used the concept of left truncated moments to characterize some probability distributions like exponential, Pareto, power, Poisson, negative binomial, beta and binomial. In fact, characterizations by left truncated moments are very important in practice since, for example , in reliability studies some measuring devices may be unable to record values smaller 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 [3], Wu and Ouyang [13], Ahsanullah and Nevzorov [16], Asadi et al [17] and Govindarajulu [24] , among others.

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

(1.1) $$F(x) = 1 - (h(x) + d)^c , \qquad x \in (a, b)$$

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) $\lim_{x \to a^+} h(x) = 1 - d$ and (b) $\lim_{x \to b^-} h(x) = -d$.

(c) E (h(X)) exists and finite.

It is easy to see that several well known distributions (like Weibull, Burr, Pareto,…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

Recurrence relations are very interesting in their own right. They are used to reduce the number of operations required to obtain a general form for the function under consideration. Moreover, several authors have used the concept of recurrence relations to identify some probability distributions see, e.g., Al-Hussaini et. al.[10],Ahmad [2], Lin [9], Khan et al. [4] and Fakhry [18].

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), survival function G(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 $\grave{F}$ (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

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

is satisfied, where h (x) is defined as before.

Proof. Necessity.

By definition

$$E(h^m(X) |X > y) = \frac{-\int_y^b h^m(x)dG(x)}{G(y)} .$$

Integrating by parts, noting that G(b)=0, one gets:

(2.2) $$E(h^m (X) | X > y) = h^m(y) + \frac{m}{G(y)} \int_y^b \grave{h}(x)\, h^{m-1}(x) G(x)\, d x$$

It is easy to see that:

(2.3) $$\grave{h}(x) = \frac{-(h(x)+d)f(x)}{c\,G(x)} ,$$

the integral on the right can be written as follows:

$$I = \int_y^b \grave{h}(x)\, h^{m-1}(x) G(x)\, d x = \frac{-1}{c} \int_y^b (h(x)+d)\, h^{m-1}(x) f(x)\, dx$$

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

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 G(y) as follows:

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

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

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

Recalling that f(y) = - $\acute{G}$(y), cancelling out $\frac{c}{c+m}$ h(y) f(y) from both sides, then multiplying the result by $\frac{c+m}{m}$,the last equation can be written as follows:

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

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

$$G(x) = (\,h(x) + d)^c \;, \qquad x \in (a,b)$$

Remarks (2.1).

(1) If we put m =1 in Theorem (2.1), we obtain Ouyang's result [14].

(2) If we put m = 1, h(x) = $\frac{Z(x)}{Z(a)+\frac{g(k)}{n(k)-1}}$, d = $\frac{g(k)/[n(k)-1]}{Z(a)+\frac{g(k)}{n(k)-1}}$, c = $\frac{n(k)}{1-n(k)}$, where n(·) and g(·) are finite real valued functions of k then we get Talwalker's result [22]

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

iff

$$E(Z(X) \mid X > y\,) = n(k)\, Z(y) + g(k)$$

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.

(2.4) $$E\left(h^m(Y_k)\right) = \frac{c(n-k+1)}{c(n-k+1)+m} E(h^m(Y_{k-1})) - \frac{md}{c(n-k+1)+m} E(h^{m-1}(Y_k)),$$
k=2,3,…,n

Proof. Necessity

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

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

Then by definition we have:

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

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

$$E(h^m(y_k)) = \frac{m\,\alpha_{k:n}}{n-k+1}\int_a^b h^{m-1}(y)\, G^{n-k+1}(y)\, F^{k-1}(y)\, \grave{h}(y)dy +$$
$$+\frac{(k-1)}{n-k+1}\ \alpha_{k:n}\int_a^b h^m(y) f(y) F^{k-2}(y) G^{n-k+1}(y)\, dy$$

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

$$E(h^m(y_k)) = \frac{-m\,\alpha_{k:n}}{c(n-k+1)}\left\{\int_a^b h^m(y) f(y) F^{k-1}(y) G^{n-k}(y)\, dy + d\int_a^b h^{m-1}(y) f(y)F^{k-1}(y)\ G^{n-k}(y)dy\right\}+$$
$$\frac{(k-1)\alpha_{k:n}}{n-k+1}\int_a^b h^m(y) f(y) F^{k-2}(y) G^{n-k+1}(y)dy =$$
$$= \frac{-m\,\alpha_{k:n}}{c(n-k+1)}\ \frac{(k-1)!(n-k)!}{n!}\ E(h^m(y_k)) - \frac{m\,d\,\alpha_{k:n}}{c(n-k+1)}\ \frac{E(h^{m-1}(y_k))}{\alpha_{k:n}} +$$
$$\frac{(k-1)\alpha_{k:n}}{n-k+1}\ \frac{(k-2)!E(h^m(y_{k-1}))}{n!}(n-k+1)!$$

Solving the last equation for E ($h^m(y_k)$) , we get:

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

Sufficiency

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

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

$$C\,(n-k+1)\frac{n!}{(k-2)!(n-k+1)!}\int_a^b h^m(y)f(y)\,F^{k-2}(y)\,G^{n-k+1}(y)dy\; -$$
$$md\,\frac{n!}{(k-1)!(n-k)!}\int_a^b h^{m-1}(y)f(y)F^{k-1}(y)G^{n-k}(y)d\,y$$

Consider the 1st integral on the right side

$$I= \int_a^b h^m(y)f(y)\,F^{k-2}(y)\,G^{n-k+1}(y)\,dy = \int_a^b h^m(y)\,G^{n-k+1}(y)\,\frac{d\,F^{k-1}(y)}{k-1}$$

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

$$I = \frac{-m}{k-1}\int_a^b h^{m-1}(y)\;\grave{h}(y)\,G^{n-k+1}(y)\,F^{k-1}(y)dy +$$
$$\left(\frac{n-k+1}{k-1}\right)\int_a^b h^m(y)f(y)F^{k-1}(y)G^{n-k}(y)dy$$

Substituting this result in equation (2.5), cancelling out

c (n-k+1) $\int_a^b h^m(y)\,f(y)\,F^{k-1}(y)\,G^{n-k}(y)\,dy$ from both sides, then multiplying the result by $\frac{1}{m}$, one gets:

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

Therefore, we have:

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

Using the Munt- Szasz theorem (see Boas [20]), one gets:

$$(h(y)+d)\,f(y) + c\,G(y)\,\grave{h}(y) = 0$$

Therefore, $\frac{-f(y)}{G(y)} = \frac{c\,\grave{h}(y)}{h(y)+d}$

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

$$G(x) = [h(x)+d]^c, \qquad x \in (a,b)$$

The proof is complete.

Remarks(2.2).

(1) Set k=n in equation (2.4), we have a recurrence relation concerning the maximum.
(2) Set n=2r+1, and k= r+1 in equation (2.4), we obtain a recurrence relation concerning the median.
(3) Set k =2 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(·), survival function G(·),and density function f(·). Let X_1, X_2, …, 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

(2.6) $$E(h^m(Y_k) \mid Y_k > t) = \frac{(n-k+1)c}{(n-k+1)c+m} h^m(t) - \frac{md}{(n-k+1)c+m} E(h^{m-1}(Y_k) \mid Y_k > t),$$

k=1,2,…,n

Proof. Necessity.

The conditional density function of the k^{th} order statistic $Y_k \mid Y_k > t$ (see, Ahsanullah [15]) is given by:

(2.7) $$f_n(Y_k \mid Y_k > t) = \frac{n-k+1}{[G(t)]^{n-k+1}} f(y)[G(y)]^{n-k} \qquad y \in (t,b)$$

Therefore

$$E(h^m(Y_k) \mid Y_k > t) = \frac{n-k+1}{[G(t)]^{n-k+1}} \int_t^b h^m(y) f(y) [G(y)]^{n-k} dy =$$

$$= \frac{-1}{[G(t)]^{n-k+1}} \int_t^b h^m(y) d[G(y)]^{n-k+1}$$

Integrating by parts, recalling that G(b)=0, we get:

$$E(h^m(Y_k) \mid Y_k > t) = h^m(t) + \frac{m}{[G(t)]^{n-k+1}} \int_t^b h^{m-1}(y) \grave{h}(y) [G(y)]^{n-k+1} dy$$

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

$$E(h^m(Y_k \mid Y_k > t) = h^m(t) - \frac{m}{c[G(t)]^{n-k+1}} \int_t^b h^m(y) f(y) [G(y)]^{n-k} dy - \frac{md}{c[G(t)]^{n-k+1}} \int_t^b h^{m-1}(y) f(y) [G(y)]^{n-k} dy$$

$$= h^m(t) - \frac{m}{(n-k+1)c} E(h^m(Y_k) \mid Y_k > t) - \frac{md}{(n-k+1)c} E(h^{m-1}(Y_k) \mid Y_k > t)$$

Solving the last equation for $E(h^m(Y_k) \mid Y_k > t)$, one gets:

$$E(h^m(Y_k) \mid Y_k > t) = \frac{(n-k+1)c}{m+(n-k+1)c} h^m(t) - \frac{md}{m+(n-k+1)c} E(h^{m-1}(Y_k) \mid Y_k > t)$$

Sufficiency.

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

$$(n\text{-}k+1)\int_t^b h^m(y)f(y)[G(y)]^{n-k}dy = \frac{(n-k+1)c}{m+(n-k+1)c}h^m(t)[G(t)]^{n-k+1} - \frac{md(n-k+1)}{m+(n-k+1)c}\int_t^b h^{m-1}(y)f(y)[G(y)]^{n-k}dy$$

Differentiating both sides with respect to t, recalling that f(t) = - $\acute{G}$ (t) ,cancelling out (n- k+1) h^{m-1}(t) $[G(t)]^{n-k}$ from both sides, one gets:

$$-h(t)f(t) = \frac{-(n-k+1)c}{m+(n-k+1)c}h(t)f(t) + \frac{cm}{m+(n-k+1)c}\acute{h}(t)G(t) + \frac{md}{m+(n-k+1)c}f(t)$$

Cancelling out $\frac{-(n-k+1)c\, h(t)f(t)}{m+(n-k+1)c}$ from both sides, multiplying the result by $\frac{m+(n-k+1)c}{m}$, one gets:

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

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

$$G(x) = [h(x)+d]^c \qquad x\in(a,b)$$

This completes the proof.

Remarks (2.3).

(1) If we put k=n in equation (2.7), we get :

$$E(h^m(Y_n)|Y_n > t) = E(h^m(X|X > y),$$

Therefore, We can say that theorem (2.3) generalizes theorem (2.1).

(2) If we put k=1 in equation (2.6), we obtain a recurrence relation for the minimum.

(3) If we put n=2r+1 and k= r+1 in equation (2.6), we obtain a recurrence relation for the median.

General Comments.

In all of the previous 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) = X^α, α > 0,d=1, c=-b, a=0 and b = ∞ , we obtain recurrence relations concerning Burr distribution with positive parameters α and b. For α = 1, the Pareto distribution of the second type is obtained.

(2) If we Set h(X) = exp -[$\alpha X + \beta X^\lambda$], where λ>0, and α,$\beta \geq 0$ such that α+β>0, c=1,d=0, a=0, and b=∞, we obtain recurrence relations concerning the Modified Weibull distribution with parameters α,β and λ (see,e.g., Zaindin and Sarhan [19]).For β=0,we have the

exponential distribution. For α=0 and λ=2, we have Rayleigh distribution. For λ=2, we have the Linear failure rate distribution with parameters α and β. For α=0, we have Weibull distribution with positive parameters β and λ.

(3) If we set h(X)= $\frac{-X}{r-\mu}$, d = $\frac{r}{r-\mu}$, c=$\theta > 0$, a= μ and b=r, we obtain recurrence relations concerning the first type personian distributions with parameters μ,r and θ.

(4)If we Set h(X) = 1- X^{α} , d =0, c =1,a =0,and b=1,we obtain recurrence relations concerning the Power distribution with parameter α > 0. For α= 1, we have the uniform distribution.

(5)If we Set h(X) = 1-X, d =0,a =0 and b = 1, we obtain recurrence relations concerning beta distribution with parameters 1, c.

References

(1) A.A.Ahmad (1996).Characterizations by regressions of order statistics. J. Applied Statistical Science,3,11-23.

(2) A.A.Ahmad (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) A.G.Pakes, M.E.Fakhry, M.R.Mahmoud and A.A.Ahmad (1996). Characterizations by regressions of order statistics. J. Applied Statistical Science,3,11-23.

(4) A.H. Khan,M. Yaqub, and S. Parvez (1983).Recurrence relations between moments of order statistics,Naval Res. Logist. Quart., 30.419-441.

(5) A. Kagan, V. Linnik, and C. Rao, (1973). Characterization Problems in Mathematical Statistics. John Wiley and Sons. New York.

(6) A.N.Ahmed (1991). Characterization of beta, binomial, and Poisson distributions. IEEE Transactions on Reliability,40, 290-295.

(7) C. Chou, and W. Haung (2003).Characterizations of the gamma distribution via conditional moments. Sankhya, Ser.A.,65,271-283.

(8) C.Dimaki and E.Xekalaki (1996). Towards a unification of certain characterizations by conditional expectations. Ann. Inst. Statistical Mathematics,48,157-168.

(9) D.G.Lin (1988). Characterization of distributions via relationships between two moments of order statistics. J. Statistical planning inference,19,73-80.

(10) E.K. Al-Hussaini, A.A. Ahmed, and M.A. Al-Kashif (2005). Recurrence relations for moment and conditional moment generating functions of generalized order statistics.Metrika,61,199-220.

(11) G. Mclachlan and D. Peel (2002).Finite Mixture Models. Wiley, New York.

(12) J. Galambos and S. Kotz (1978).Characterization of Probability Distributions Springer- Verlag Berlin.

(13) J. Wu and L .Y. Ouyang (1996).On Characterizatizing distributions by conditional expectations of functions of order statistics.Metrika,34, 135-147.

(14) L .Y. Ouyang (1983). On characterization of distributions by conditional expectations. Tamkang J. Management Science,4,13-21.

(15) M. Ahsanullah (2009).On some characterizations of univariate distributions based on truncated moments of order statistics. Pakistan J. of Statistics, 25,83-91.

(16) M. Ahsanullah and V. Nevzorov (2001). Orderd Random Variables. Nova, New York.

(17) M. Asadi, C. Rao and D.N. Shanbhag (2001).Some unified characterization results on generalized Pareto distributions. J. Statistical Planning Inference, 93, 29-50.

(18) M.E. Fakhry (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.

(19) M. Zaindin and A.M. Sarhan (2009). Parameters estimation of the Modified Weibull distribution. Applied Mathematical Science, 3, ,541-550.

(20) P.R. Boos (1954).Entire Functions. Academic Press.

(21) S. Osaki and X. Li (1989).Characterization of gamma and negative binomial distributions. IEEE. Transactions on Reliability,R73,379-382.

(22) S. Talwalker (1977). A note on characterization by conditional expectation. Metrika,24,129-136.

(23) T.A. Azlarov and N. A. Volodin (1986). Characterization Problems Associated with The Exponential Distribution Springer Verlag Berlin.

(24) Z. Govindarajulu (2001). Characterization of double exponential distribution using moments of order statistics. Communication of Statistics. Theory and Methods, 30, 2355-2372.

Received: July, 2012