FASCICULI MATHEMATICI

Nr 28 1998

DEDICATED TO PROFESSOR DOBIESŁAW BOBROWSKI
ON THE OCCASION OF HIS 70^{th} BIRTHDAY

UDO KAMPS

SUBRANGES OF GENERALIZED ORDER STATISTICS FROM EXPONENTIAL DISTRIBUTIONS

ABSTRACT. Identical distributions of a subrange and a corresponding order statistic, as well as related moment equations, are known to be characteristic properties of exponential distributions. Results for ordinary order statistics and record values are particular cases of characterization theorems based on subranges of generalized order statistics.

KEY WORDS: order statistics, record values, generalized order statistics, characterizations, exponential distributions, subranges, spacings.

1. INTRODUCTION

Many characterizations of exponential distributions by means of distributional properties of order statistics and record values are known.

Let $X_{1,n} \leq \ldots \leq X_{n,n}$ denote the order statistics based on i.i.d. random variables $X_1, \ldots, X_n$, $n \geq 2$, with distribution function F. If F is the distribution function of an exponential distribution with parameter λ, i.e.,

$$F(x) = 1 - e^{-\lambda x}, \quad x > 0, \quad \lambda > 0 \quad (F \equiv \text{Exp}(\lambda) \text{ for short}),$$

then subranges or spacings are distributed as a corresponding order statistic from F:

$$X_{s,n} - X_{r,n} \sim X_{s-r,n-r} \tag{1}$$

for all integers r, s and n with $1 \leq r < s \leq n$.

Ahsanullah (1984, 1991), Iwińska (1985, 1986), Gather (1988), and Gajek and Gather (1989) deal with characterizations of exponential distributions via property (1) or the moment equation

$$EX_{s,n} - EX_{r,n} = EX_{s-r,n-r} \tag{2}$$

as well as via the corresponding conditions for record values. A detailed survey of characterizations of distributions by means of identically distributed functions of order statistics is given in Gather et al. (1996).

In Kamps (1995), generalized order statistics are introduced as a unified approach to ordinary order statistics and record values and to a variety of other models of ordered random variables such as sequential order statistics, k-th record values and Pfeifer's record model.

In this paper, we derive characterizations of exponential distributions via distributional properties of subranges of generalized order statistics, including the above mentioned results for ordinary order statistics and record values as particular cases.

2. Generalized order statistics

Let $X(1,n,m,k),\dots,X(n,n,m,k)$ be generalized order statistics based on an absolutely continuous distribution function F with density function f, which means that the joint density function of the above quantities is given by

$$f^{X(1,n,m,k),\dots,X(n,n,m,k)}(x_1,\dots,x_n)=$$
$$=k\left(\prod_{j=1}^{n-1}\gamma_j\right)\left(\prod_{i=1}^{n-1}(1-F(x_i))^m f(x_i)\right)(1-F(x_n))^{k-1}f(x_n),$$
$$F^{-1}(0+)<x_1\le\dots\le x_n<F^{-1}(1),$$

with $n\in N$, $k>0$, $m\in R$ such that $\gamma_r=k+(n-r)(m+1)>0$ for all $1\le r\le n$.

In the case $m=0$ and $k=1$ the model reduces to the joint density of ordinary order statistics, and in the case $m=-1$ and $k\in N$ we obtain the joint density of the first n k-th record values based on a sequence $X_1,X_2,\dots$ of i.i.d. random variables with distribution function F.

The marginal density function of the r-th generalized order statistic is given by

$$f^{X(r,n,m,k)}(x)=\frac{c_{r-1}}{(r-1)!}(1-F(x))^{\gamma_r-1}f(x)g_m^{r-1}(F(x)) \tag{3}$$

(see Kamps 1995, p. 14), and the density function of the subrange $W_{r,s,n}=$ $=X(s,n,m,k)-X(r,n,m,k)$, $1\le r<s\le n$, has the following representation

$$f^{W_{r,s,n}}(w)=\frac{c_{s-1}}{(r-1)!(s-r-1)!}\int_{-\infty}^{\infty}(1-F(y))^m f(y)g_m^{r-1}(F(y))\cdot \tag{4}$$
$$\left[h_m(F(y+w))-h_m(F(y))\right]^{s-r-1}(1-F(y+w))^{\gamma_s-1}f(y+w)dy$$

with $$c_{r-1}=\prod_{j=1}^{r}\gamma_j,\qquad 1\le r\le n,$$

$$h_m(x)=\int(1-x)^m dx=\begin{cases}-\frac{1}{m+1}(1-x)^{m+1}, & m\ne-1,\\ \log\frac{1}{1-x}, & m=-1,\end{cases}$$

$$g_m(x)=h_m(x)-h_m(0), \quad x\in[0,1) \text{ (see Kamps 1995, p. 14).}$$

In Lemma 2.1, we will derive representations of the survival functions of $W_{r,s,n}$ and $X(r,n,m,k)$ in terms of incomplete beta and gamma functions, which will be used in the proofs of the characterizations presented in Section 3. For order statistics, the expressions can be found in Gather (1988) and David (1981).

In order to simplify the notation, let the function H_m be defined by

$$H_m(z)=\begin{cases} I_z(\frac{k}{m+1}+n-s,s-r), & m>-1, \\ I_z(-\frac{k}{m+1}-n+r+1,s-r), & m<-1, \\ \frac{1}{(s-r-1)!}\Gamma(s-r,-k\log z), & m=-1, \end{cases} \tag{5}$$

with $$I_z(a,b)=\frac{1}{B(a,b)}\int_0^z t^{a-1}(1-t)^{b-1}dt, \qquad z\in(0,1), \qquad a,b>0,$$

and $$\Gamma(a,z)=\int_z^\infty t^{a-1}e^{-t}dt, \qquad z>0, \qquad a>0.$$

Moreover, for $x,y>0$, let

$$a_m(x,y)=\begin{cases} (1-F(x+y))^{m+1}/(1-F(y))^{m+1}, & m>-1, \\ (1-F(y))^{m+1}/(1-F(x+y))^{m+1}, & m<-1, \\ (1-F(x+y))/(1-F(y)), & m=-1, \end{cases} \tag{6}$$

and $$b_m(x)=\begin{cases} (1-F(x))^{m+1}, m>-1, \\ (1-F(x))^{-(m+1)}, m<-1, \\ 1-F(x), \qquad m=-1. \end{cases}$$

Lemma 2.1. The survival functions of $W_{r,s,n}$ and $X(s-r,n-r,m,k)$, $1\le r<$ $<s\le n$, are given by

$$1-F^{W_{r,s,n}}(x)=\int_0^\infty H_m(a_m(x,y))dF^{X(r,n,m,k)}(y)$$

and $$1-F^{X(s-r,\,n-r,\,m,k)}(x)=H_m(b_m(x))..$$

Proof. Applying formula (4) we find

$$1-F^{W_{r,s,n}}(x)=\int_x^\infty F^{W_{r,s,n}}(w)dw=$$

$$= \frac{c_{s-1}}{(r-1)!(s-r-1)!}\int_0^\infty (1-F(y))^m f(y) g_m^{r-1}(F(y)) \cdot$$

$$\cdot \int_{x+y}^\infty [h_m(F(w)) - h_m(F(y))]^{s-r-1}(1-F(w))^{\gamma_s - 1} f(w) dw dy.$$

Some elementary calculations with respect to $m > -1$, $m < -1$ and $m = -1$ (and the substitutions $z = 1 - h_m(F(w))/h_m(F(y))$, $z = 1 - h_m(F(y))/h_m(F(w))$, and $z = \log((1-F(y))/(1-F(w)))$, respectively) lead to the desired representation, since

$$\begin{aligned}\Gamma(\tfrac{k}{m+1}+n-s)/\Gamma(\tfrac{k}{m+1}+n-r) &= \\ &= (-1)^{s-r}\Gamma(-\tfrac{k}{m+1}-n+r+1)/\Gamma(-\tfrac{k}{m+1}-n+s+1) = \\ &= (m+1)^{s-r}\Big/\prod_{j=r+1}^{s}\gamma_j = (m+1)^{s-r}\frac{c_{r-1}}{c_{s-1}}.\end{aligned}$$

On the other hand, the representation of the survival function of $X(s-r, n-r, m, k)$ in the case $m > -1$ was established in Nasri-Roudsari (1996).

Observing that, for $m < -1$, $X(s-r, n-r, m, k)$ and $X(s-r, n-r, -m-2, k')$ with $k' = k + (2n - r - s - 1)(m+1)$ are identically distributed, we derive the desired representation via transformation applying the above result in the case $m > -1$. For $m = -1$, the assertion is directly seen by using the ordinary expression of the survival function of $X(s-r, n-r, -1, k)$ (cf. Kamps 1995, p. 15).

Generalizing Sukhatme's (1937) result, it is shown in Kamps (1995, p. 16) that the normalized spacings

$$D(1, n, m, k) = \gamma_1 X(1, n, m, k)$$

and $$D(r, n, m, k) = \gamma_r (X(r, n, m, k) - X(r-1, n, m, k)), \quad 2 \le r \le n$$

based on an exponential distribution with parameter λ, are independent and identically distributed according to $\mathrm{Exp}(\lambda)$.

As a consequence, it is easily seen that for $F \equiv \mathrm{Exp}(\lambda)$ an extension of relation (1) to generalized order statistics holds true.

Lemma 2.2. Let $F \equiv \mathrm{Exp}(\lambda)$. *Then*

(7) $$X(s, n, m, k) - X(r, n, m, k) \sim X(s-r, n-r, m, k)$$

for all integers r, s and n with $1 \le r < s \le n$.

Proof. Let $Y_1,\ldots,Y_n$ be i.i.d. random variables with $Y_1 \sim \text{Exp}(\lambda)$. We have

$$X(s,n,m,k) - X(r,n,m,k) = \sum_{j=r+1}^{s}(X(j,n,m,k) - X(j-1,n,m,k)) =$$

$$= \sum_{j=r+1}^{s} \frac{1}{\gamma_j} D(j,n,m,k) \sim \sum_{j=r+1}^{s} \frac{1}{\gamma_j} Y_j \sim \sum_{j=r+1}^{s} \frac{1}{\gamma_j} D(j-r,n-r,m,k) =$$

$$= \sum_{j=1}^{s-r} \frac{1}{\gamma_{j+r}} D(j,n-r,m,k) = X(s-r,n-r,m,k).$$

Hence, in characterization results based on (7) it remains to show that the respective properties determine exponential distributions uniquely.

3. CHARACTERIZATION THEOREMS

In the case of order statistics, Gather (1988) shows that exponential distributions are characterized if relation (1) is fulfilled for two distinct values of s. This assertion was previously stated by Ahsanullah (1975), but in the proof the NBU/NWU property of the underlying distribution is implicitly assumed, as pointed out by Gather (1988). Dimaki and Xekalaki (1993) characterize Pareto distributions via a corresponding transformation of relation (1) by analogy with Ahsanullah's (1975) proof. Implicitly, they also use an additional condition. Using the representations of $W_{r,s,n}$ and $X(s-r,n-r,m,k)$ established in Lemma 2.1, Gather's (1988) result can be extended directly to generalized order statistics, and thus it is seen to hold true, e.g., for sequential order statistics, record values and k-th record values. It should be noted that for the simplicity of presentation, we assume that the underlying distribution function F is absolutely continuous. In Gather's (1988) result for order statistics, merely continuity of F is required.

Theorem 3.1. Let F be absolutely continuous with density function f, $F(0) = 0$, and let F be strictly increasing on $(0, \infty)$.

Then $F \equiv \text{Exp}(\lambda)$ for some $\lambda > 0$ iff there exist integers r, s_1, s_2 and n, $1 \le r < < s_1 < s_2 \le n$, such that (7) holds true for $s = s_1$ and $s = s_2$.

Proof. Let the function H_m be denoted by $H_m^{(j)}$, when $s = s_j$ in (5), $j = 1,2$. Applying Lemma 2.1, the equality of the survival functions of $W_{r,s,n}$ and $X(s-r,n-r,m,k)$ is equivalent to

(8) $$\int_0^\infty H_m(a_m(x,y))dF^{X(r,n,m,k)}(y) = H_m(b_m(x)) \quad \text{for all } x > 0.$$

Now we can follow the argumentation of Gather (1988), using the facts that $H_m^{(j)}$ is continuous and strictly increasing, $j = 1,2$, and that $H_m^{(1)} \circ \left(H_m^{(2)}\right)^{-1}$ is strictly convex on $(0,1)$.

This amounts to concluding that

$$1 - F(x+y) = (1-F(x))(1-F(y)) \quad \text{for all} \quad x, y > 0$$

which is the characterizing functional equation of exponential distributions.

Under NBU/NWU assumption, the relations (1) and (2) for ordinary order statistics as well as the corresponding relations for record values are known to be characteristic properties of exponential distributions. Theorem 3.2 presents an extension to generalized order statistics.

Theorem 3.2. Let F be absolutely continuous with density function f, $F(0) = 0$, *let F be strictly increasing on* $(0,\infty)$ *and NBU or NWU.*

Then $F \equiv \mathrm{Exp}(\lambda)$ *for some* $\lambda > 0$ *iff there exist integers r, s and n,* $1 \le r < s \le n$, *such that*

(*i*) $$X(s,n,m,k) - X(r,n,m,k) \sim X(s-r, n-r, m, k)$$

or

(*ii*) $$EX(s,n,m,k) - EX(r,n,m,k) = EX(s-r, n-r, m, k)$$

assuming that the expected values are finite.

Proof. If *F* is NBU (NWU) then

$$H_m(a_m(x,y)) \underset{(\ge)}{\le} H_m(b_m(x)) \quad \text{for all} \quad x, y > 0.$$

(*i*) Since relation (8) is equivalent to

$$(c(x) =) \int_0^\infty (H_m(a_m(x,y)) - H_m(b_m(x))) dF^{X(r,n,m,k)}(y) = 0$$

for all $x > 0$ we conclude that

$$H_m(a_m(x,y)) - H_m(b_m(x)) = 0 \quad \text{for all} \quad x, y > 0$$

which implies

$$a_m(x,y) = b_m(x) \quad \text{for all} \quad x, y > 0$$

and thus the assertion.

(*ii*) The equality of expectations leads to $\int_0^\infty c(x)dx = 0$, which completes the proof.

In particular, Theorem 3.2 includes well-known results for order statistics and record values. As characterizations of exponential distributions by order statistics, part i) generalizes a result of Iwiñska (1985) and Theorem 2.1 in Ahsanullah (1984) which is proved under IFR/DFR assumption. Part ii) includes Theorem 2.3 in Gajek and Gather (1989) for order statistics and their remark on record values, as well as Theorem 1 in Iwiñska (1986) for order statistics and Theorem 2 in Iwiñska (1986) for record values (cf. Ahsanullah 1991, Theorem 2.2). In Kamps and Gather (1997) part ii) is shown in the particular case $s = r + 1$.

REFERENCES

[1] M. Ahsanullah, A characterization of the exponential distribution. In: G.P. Patil et al. (Eds.), *Statistical Distributions in Scientific Work*, Reidel, Dordrecht, 3(1975), 131-135.

[2] M. Ahsanullah, A characterization of the exponential distribution by higher order gap. *Metrika* 31(1984), 323-326.

[3] M. Ahsanullah, Some characteristic properties of the record values from the exponential distribution. Sankhya 53 B, (1991), 403-408.

[4] H.A. David, *Order Statistics*, 2nd ed. Wiley, New York, (1981).

[5] C. Dimaki, E. Xekalaki, Characterizations of the Pareto distribution based on order statistics. In: V.V. Kalashnikov and V.M. Zolotarev (Eds.), *Stability Problems for Stochastic Models*. Springer, Berlin, (1993), 1-16.

[6] L. Gajek, U. Gather, Characterizations of the exponential distribution by failure rate- and moment properties of order statistics. In: J. Hüsler and R.D. Reiss (Eds.), *Extreme Value Theory*, Springer, Berlin, (1989), 114-124.

[7] U. Gather, On a characterization of the exponential distribution by properties of order statistics, *Statist. Prob.* Letters 7(1988), 93-96.

[8] U. Gather, U. Kamps, N. Schweitzer, *Characterizations of distributions via identically distributed functions of order statistics,* To appear in: C.R. Rao and N. Balakrishnan (Eds.), Handbook of Statistics-16, Order Statistics and Their Applications, North-Holland, Amsterdam, (1997).

[9] M. Iwiñska, On a characterization of the exponential distribution by order statistics. In: *Numerical Methods and Their Applications*, Proc. 8th Sess. oznan *Circle Zesz. Nauk. Ser I*, Akad. Ekon. Poznan 132(1985), 51-54.

[10] M. Iwińska, On the characterizations of the exponential distribution by order statistics and record values. *Fasciculi Mathematici* 16(1986), 101-107.

[11] U. Kamps, A concept of generalized order statistics, *J. Statist. Plann. Inf.* 48(1995), 1-23.

[12] U. Kamps, U. Gather, Characteristic properties of generalized order statistics from exponential distributions. *Applicationes Mathematicae*, (1997), to appear.

[13] D. Nasri-Roudsari, (1996) Extreme value theory of generalized order statistics, *J. Statist. Plann. Inf.* 55(1996), 281-297.

[14] P.V. Sukhatme, Tests of significance for samples of the χ^2-population with two degrees of freedom, *Ann. Eugenics* 8(1937), 52-56.

(Institute of Statistics, Aachen University of Technology, 52056 Aachen, and Department of Statistics, University of Dortmund, 44221 Dortmund, Germany)

Received on 23.09.1996 and, in revised form, on 09.04.1997.