Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 16 (1997), No. 1, 191–200

Characterization of the Exponential Distribution by Properties of the Difference $X_{k+s:n} - X_{k:n}$ of Order Statistics

H.-J. Rossberg, M. Riedel and B. Ramachandran

In memoriam Paul Günther (1926 - 1996)

Abstract. Let $X_1, X_2, ..., X_n$ be independent and identically distributed random variables subject to a continuous distribution function F, let $X_{1:n}, X_{2:n}, ..., X_{n:n}$ be the corresponding order statistics, and write

$$P(X_{k+s:n} - X_{k:n} \geq x) = P(X_{s:n-k} \geq x) \qquad (x \geq 0) \tag{0}$$

where n, k and s are fixed integers with $k + s \leq n$. It is an old question if condition (0) implies that F is of exponential type. In [8] we showed among others that condition (0) can be greatly relaxed; namely, it can be replaced by asymptotic relations (either as $x \to \infty$ or $x \downarrow 0$) to derive this very result. Using a theorem on integrated Cauchy functional equations and in essential way a result of [8] we find now a more elegant and deeper theorem on this subject. The case of lattice distributions is also considered and some new problems are stated.

Keywords: *Exponential distributions, integrated Cauchy functional equations*

AMS subject classification: Primary 60 E 05, secondary 62 E 10, 45 A 05

1. Introduction

Let $X_1, X_2, ..., X_n$ be independent and identically distributed random variables with common distribution function F, and let $X_{1:n}, X_{2:n}, ..., X_{n:n}$ be the corresponding order statistics. Throughout the present paper we keep integers k, n and s with $1 \leq k < n$ and $k + s \leq n$ fixed and put

$$d_s(x) = P(X_{k+s:n} - X_{k:n} \geq x) - P(X_{s:n-k} \geq x) \qquad (x \geq 0) \tag{1.1}$$

and

$$\delta_s(x, \rho) = P(X_{k+s:n} - X_{k:n} \geq x) - e^{-\rho(n-k)x} \qquad (x \geq 0). \tag{1.2}$$

Then we can formulate the two following problems.

H.-J. Rossberg: Univ. of Leipzig, Inst. Math., Augustusplatz 10, D – 04109 Leipzig
M. Riedel: Univ. of Leipzig, Inst. Math., Augustusplatz 10, D – 04109 Leipzig
B. Ramachandran: Indian Stat. Inst., 7 S.J.S. Sansanwal Marg, New Delhi – 110016, India

ISSN 0232-2064 / $ 2.50

Problem 1. Does

$$d_s(x) = 0 \qquad (x \geq 0) \tag{1.3}$$

for some $s \geq 1$ imply that F is of exponential type on $(0, +\infty)$?

Problem 2. Does

$$\delta_s(x, \rho) = 0 \qquad (x \geq 0) \tag{1.4}$$

for some $s \geq 1$ imply that F is of exponential type on $(0, +\infty)$ with parameter ρ?

Essential steps in solving Problems 1 and 2 are the following ones.

Case $s = 1$: Rossberg [9] tackled both problems for continuous F. He applied two analytic function theory methods that are radically different. Also the results have different characters: It is only in the solution for Problem 2 that an impleasant assumption on the zeros of f – the characteristic function of F – proves to be necessary. Ramachandran [6] developed the first method. His result could later be derived by means of the integrated Cauchy functional equation (1.9) (see [7: Theorem 2.5.5]; the result is reformulated in the Appendix of the present paper).

Case $s > 1$: Ahsanullah [2] showed that (1.3) together with F having increasing failure rate is a characterizing property of the exponential distribution function. A very lucky idea came even still earlier: Ahsanullah [1] assumed (1.3) for two different values $1 \leq s_1 < s_2 \leq n - k$ and derived the exponential distribution function under the assumption that F has a Lebesgue density, which is strictly positive on $(0, +\infty)$ and under the additional implicit condition that F satisfies either

$$\overline{F}(x + y) \leq \overline{F}(x)\,\overline{F}(y) \tag{1.5}$$

or

$$\overline{F}(x + y) \geq \overline{F}(x)\,\overline{F}(y) \tag{1.6}$$

where $\overline{F} = 1 - F$.

Ahsanullah's problem was treated by Gather [4] without the additional condition and under the weaker assumption that F is continuous and strictly increasing for $x > 0$. From her proof it follows also that the answer to Problem 1 with just one value $s > 1$ is positive provided that the distribution function F satisfies (1.5) or (1.6).

A new idea was elaborated in [8] where we assumed $1 \leq s \leq n - k$ and noticed the following:

(i) Strict equality in (1.3) and (1.4) is not needed in any point. Instead we can modify Problems 1 and 2 imposing an asymptotic condition (either as $x \to \infty$ or $x \to 0$) on d_s or δ_s to be sufficient, provided that F shares some general properties (see, for instance, Propositions 1.1 and 1.2).

(ii) Contrary to the fact that in [9] very different methods had to be found it is now the same elementary method that allows to treat both modified problems corresponding to (1.3) and (1.4).

As examples we quote the following two known results.

Proposition 1.1 (see [8]). *Let the distribution function F with $F(0) = 0$ have a continuous and bounded density p on $[0, +\infty)$. Then F is of exponential type provided that*

$$d_s(x) = o(F^s(x)) \qquad (x \to 0). \tag{1.7}$$

The counterpart where $x \to +\infty$ runs as follows.

Proposition 1.2 (see [8]). *Let F be a continuous distribution function and assume that we have regular variation, i.e. for every $u > 0$ there exists*

$$\lim_{x\to\infty} \frac{\overline{F}(x+u)}{\overline{F}(x)} = Q(u). \tag{1.8}$$

Then F is of exponential type on $(0, +\infty)$ provided that

$$d_s(x) = o\left(\overline{F}(x)^{n-k-s+1}\right) \qquad (x \to +\infty) \tag{1.9}$$

and $\overline{F}(u) - Q(u)$ does not change the sign for $u \geq 0$.

To our opinion the formulation of Proposition 1.2 is not satisfactory since assumption (1.8) is comparatively strong; it immediately implies that Q is of exponential type. But in spite of that we need Proposition 1.2 as a tool in the present note. We underline that our Theorem 2.2 is not a generalization of Proposition 1.2 because the speed of convergence in assumption (2.1) is somewhat higher and we assume $s = 1$. Note that Theorem 2.2 is by far more elegant since assumption (1.8) is completely omitted; it is also essentially deeper because in the proof we need also a remarkable theorem on "integrated Cauchy functional equations" before we can apply Proposition 1.2 to conclude that F is of exponential type.

To finish the introduction we remind the reader of the following fact. If the function F is continuous, then the Markov property of order statistics holds (see [3]). Then condition (1.3) for $s = 1$ can easily be rewritten as

$$\binom{n}{k} \int_{-\infty}^{+\infty} \overline{F}^{n-k}(x+u)\, dF^k(u) = F^{n-k}(x) \qquad (x \geq 0) \tag{1.10}$$

whence $F(0) = 0$ trivially follows. Putting

$$\overline{F}^{n-k} = f \qquad \text{and} \qquad \sigma(x) = \binom{n}{k} F^k(x) \quad (x \geq 0)$$

we obtain the integrated Cauchy functional equation

$$f(x) = \int_0^{+\infty} f(x+u)\, d\sigma(u) \qquad (x \geq 0) \tag{1.11}$$

whose theory is presented in [7]. In Section 2 we will use this equation modified by an error term.

2. Results for $s = 1$

Our first statement is a stability theorem. It is nothing but an immediate application of Theorem 4.4.1 in connection with Remark (1) of [7]. By a stroke of good fortune it is possible to combine it with Proposition 1.2 so that we easily obtain Theorem 2.2. Note that it is the assumption "F strictly increasing on $(0,\eta)$" that greatly simplifies formula (2.2).

Proposition 2.1. *Let a continuous distribution function $F < 1$ be strictly increasing on $[0,\eta)$, with $\eta > 0$, and assume, moreover, $F(0) = 0$ and*

$$\int_{-\infty}^{+\infty} \overline{F}^{n-k}(x+u)\,d\sigma(u) = \overline{F}^{n-k}(x)(1 + O(e^{-\varepsilon x})) \qquad (x \to +\infty) \tag{2.1}$$

for some $\varepsilon > 0$. Then there exists a unique $\alpha > 0$ such that

$$\int_{0}^{+\infty} e^{-\alpha y}\,d\sigma(y) = 1$$

and

$$\overline{F}^{n-k}(x) = e^{-\alpha x}(1 + O(e^{-\varepsilon x})) \qquad (x \to +\infty). \tag{2.2}$$

Now we see at once that under the above suppositions condition (1.8) is satisfied with $Q^{n-k}(u) = e^{-\alpha u}$. Hence we can apply Proposition 1.2 and obtain our main result, the desired characterization theorem.

Theorem 2.2. *Let the above suppositions be true and assume either*

$$\overline{F}^{n-k}(x) \geq e^{-\alpha x} \qquad (x \geq 0)$$

or

$$\overline{F}^{n-k}(x) \leq e^{-\alpha x} \qquad (x \geq 0).$$

Then F is of exponential type.

Problem 2.3. The above considerations seem to be of interest also in view of the ideas which they stimulate. For instance:

(i) Is it possible to modify Theorem 2.2 (or the corresponding Theorem 4.4.1 of [7]) such that in the assumption $d_1(x) \to 0$ not as $x \to +\infty$ but as $x \downarrow 0$ with a certain speed?

(ii) Problem 2 has not been treated by other authors in the long period from 1972 - 1994. Will it be possible to improve the results of [8] concerning Gather's problem? The question makes sense since this paper shows the intimate connection of Problems 1 and 2 and their modifications.

3. Results for $1 < s \le n - k$

We are now able to carry the above results over to the case $1 < s \le n - k$. Our considerations are similar and so are the results but we need the additional assumption (3.3). It will prove useful to introduce the measures

$$\sigma_r(x) = \binom{n}{k} \int_0^x \overline{F}^r(u)\, dF^k(u) \qquad (0 \le r < n - k)$$

with

$$\sigma_r(\infty) = \frac{\binom{n}{k}}{\binom{r+k}{k}} > 1.$$

To begin with we take from [8] the representation

$$\begin{aligned} d_s(x) = (-1)^{s-1} \sum_{r=0}^{s-1} (-1)^r \binom{n-k}{r} \binom{n-k-1-r}{s-r-1} \\ \times \left[\overline{F}^{n-k-r}(x) - a_r \int_0^{+\infty} \overline{F}^{n-k-r}(x+u)\, dK_r(u) \right] \end{aligned} \tag{3.1}$$

where

$$K_r(u) = \binom{r+k}{r} \int_0^{F(u)} (1-w)^r dw^k \qquad (0 \le r \le n-k-1,\, n \ge 0)$$

are distribution functions and

$$a_r = \frac{\binom{n}{k}}{\binom{r+k}{r}} \ge 1 \qquad (r \ge 0).$$

By means of the above measures σ_r (note that $d\sigma_r(u) = a_r\, dK_r(u)$) this formula can be written as

$$\begin{aligned} d_s(x) &= \binom{n-k}{s-1} \left[\overline{F}^{n-k-s+1}(x) - \int_0^{+\infty} \overline{F}^{n-k-s+1}(x+u)\, d\sigma_{s-1}(u) \right] \\ &\quad + O(\overline{F}^{n-k-s+2}(x)) \\ &= \binom{n-k}{s-1} \Bigg[\overline{F}^{n-k-s+1}(x)(1 + O(\overline{F}(x)) \\ &\quad - \int_0^{+\infty} \overline{F}^{n-k-s+1}(x+u)\, d\sigma_{s-1}(u) \Bigg] \qquad (x \to +\infty). \end{aligned} \tag{3.2}$$

Proposition 3.1. *Let a continuous distribution function $F < 1$ be strictly increasing on $[0, \eta)$ $(\eta > 0)$ with $F(0) = 0$. Assume*

$$\overline{F}(x) = O(e^{-\varepsilon x}) \qquad (x \to +\infty) \tag{3.3}$$

for some $\varepsilon > 0$ and, further,

$$d_s(x) = O(\overline{F}^{n-k-s+1}(x)\, e^{-\varepsilon x}) \qquad (x \to +\infty). \tag{3.4}$$

Then there exists a unique $\alpha_s > 0$ such that

$$\int_0^{+\infty} e^{-\alpha_s y} d\sigma_{s-1}(y) = 1$$

and

$$\overline{F}^{n-k-s+1}(x) = e^{-\alpha_s x}\big(1 + O(e^{-\varepsilon x})\big) \qquad (x \to +\infty). \tag{3.5}$$

Proof. Using (3.2) and assumptions (3.3) and (3.4) we get

$$\int_0^{+\infty} \overline{F}^{n-k-s+1}(x+u)\, d\sigma_{s-1}(u) = \overline{F}^{n-k-s+1}(x)\big(1 + O(e^{-\varepsilon x})\big) \qquad (x \to +\infty).$$

Accordingly, Theorem 4.4.1 of [7] permits to draw the conclusion (3.5) ∎

By the same arguments as in Section 2 we obtain now the result corresponding to Theorem 2.2.

Theorem 3.2. *Let the above suppositions be true and assume further that*

$$\overline{F}^{n-k-s+1}(x) \geq e^{-\alpha_s x} \qquad (x > 0).$$

Then F is of exponential type.

4. Lattice distribution functions as solutions

We assume now that $F < 1$ is a lattice right continuous distribution function. In this case we need the above formula (3.1) for $s = 1$; but it was derived under the assumption of continuity. We show therefore first that it is true also for lattice distribution functions.

We start with a remark on the tail function of the difference of two order statistics in lattice case. Consider a linear transformation

$$Y_j = aX_j + b \quad (j = 1, 2, \ldots, n;\ 0 < a \in \mathbb{R})$$

of the random variables X_j. Then we have for the corresponding order statistics $X_{j:n}$ the same transformation, i.e.

$$Y_{j:n} = aX_{j:n} + b \qquad (j = 1, 2, \ldots, n).$$

Hence we may restrict ourselves to lattice distribution functions concentrated on the set $\mathbb{Z}$ of integers.

Proposition 4.1. *Let F be a lattice distribution function concentrated on $\mathbb{Z}$. Then, for all real $x \geq 0$,*

$$P(X_{k+1:n} - X_{k:n} > x) = \binom{n}{k} \sum_{i=-\infty}^{+\infty} \overline{F}^{n-k}(x+i)\big(F^k(i) - F^k(i-)\big) \tag{4.1}$$

which can be rewritten as the Lebesgue-Stieltjes integral

$$P(X_{k+1:n} - X_{k:n} > x) = \binom{n}{k} \int_{-\infty}^{+\infty} \overline{F}^{n-k}(x+u)\, F^k(du).$$

Proof. We note that F and $\overline{F}$ are constant over every open interval of the form $(m, m+1)$ with $m \in \mathbb{N}$. This is also true for the left-hand side of (4.1). Hence it suffices to prove (4.1) for all non-negative integers $x = m$. For this purpose, using the general joint distribution function of $(X_{k:n}, X_{k+s:n})$ given in [5] for the special case $s = 1$, we obtain

$$\begin{aligned} P(X_{k+1:n} - X_{k:n} = j) &= \sum_{i=-\infty}^{+\infty} P(X_{k:n} = i, X_{k+1:n} = i+j) \\ &= \binom{n}{k} \int_{A_{i,j}} du^k\, d(1-v)^{n-k} \end{aligned}$$

where we denoted

$$A_{i,j} = \Big\{(u,v) :\, u < v,\ F(i-) < u < F(i),\ F((i+j)-) < v < F(i+j\Big\}.$$

Hence, we get for $m \geq 0$

$$\begin{aligned} P(X_{k+1:n} - X_{k:n} > m) &= \sum_{j=m+1}^{+\infty} P(X_{k+1:n} - X_{k:n} = j) \\ &= \binom{n}{k} \sum_{i=-\infty}^{+\infty} \int_{F(i-)}^{F(i)} du^k \sum_{j=m+1}^{+\infty} \int_{F((i+j)-)}^{F(i+j)} d(1-v)^{n-k} \\ &= -\binom{n}{k} \sum_{i=-\infty}^{1} \big(F^k(i) - F^k(i-)\big) \int_{F(i+m)}^{+\infty} d(1-v)^{n-k} \end{aligned}$$

and this is equivalent to the assertion ∎

Now we assume (2.1) with $\sigma = \binom{n}{k} F^k$. Then Theorem 4.4.1 of [7] tells us that the step function $\overline{F}^{n-k}$ satisfies for a certain $a > 0$

$$\overline{F}^{n-k}(x) = p(x)\, e^{-ax}\big(1 + k(x)\big) = c_l > 0 \qquad (l < x < l+1,\ l \in \mathbb{N}_0) \tag{4.2}$$

with some sequence $c_l \downarrow 0$ and some function k such that

$$k(x) = O(e^{-\varepsilon x}) \qquad (x \to +\infty) \tag{4.3}$$

where p has period 1. In other words, the function p fulfils

$$p(x) = c_l \frac{e^{ax}}{1 + k(x)} \qquad (l < x < l + 1)$$

and

$$p(x+1) = c_{l+1} \frac{e^{a(x+1)}}{1 + k(x+1)} \qquad (l < x < l + 1).$$

Hence

$$\frac{c_{l+1}}{c_l} = e^{-a} \frac{(1 + k(x+1))}{1 + k(x)} \qquad (l < x < l + 1)$$

follows from periodicity so that $\frac{c_{l+1}}{c_l} \to e^{-a} < 1$ as $l \to +\infty$. For $k = 0$ we obtain $c_{l+1} = e^{-a} c_l$ $(l > 0)$ with

$$\sum_{l=0}^{+\infty} c_l = c_0 \left(1 + e^{-a} + \ldots\right) = \frac{c_0}{1 - e^{-a}}$$

in accordance with Theorem 2.5.5 of [7] a correct version of which is given in the Appendix below; it tells us that a mixture of the degenerate and a geometric type distribution function is a solution of (1.10) in the lattice case.

We wish now to derive $k = 0$ from appropriate conditions. For this purpose we check the proof of the above Proposition 1.2 given in [8] step by step and find that the following counterpart to this proposition holds.

Proposition 4.3. *Let $F < 1$ be lattice supported on $0, 1, \ldots$ with $F(-0) = 0$. Assume that (1.8) is true and that $\overline{F}(u) - Q(u)$ does not change the sign for $u \geq 0$. Then (1.9) implies that there exists $\alpha > 0$ such that*

$$e^{u\alpha} \overline{F}(u) = 1 \qquad (u \in \mathbb{N}_0). \tag{4.4}$$

Since (1.8) is an obvious consequence of (4.2), (4.3) and the periodicity of p we obtain now the following desired counterpart to Theorem 2.2.

Theorem 4.2. *Let $F < 1$ be lattice supported on $\mathbb{N}_0$ with $F(-0) = 0$. If it satisfies (2.1), then there exists $\alpha > 0$ such that (4.2) and (4.3) are true. In case that, moreover,*

$$\overline{F}(x) \geq e^{-\alpha x} \qquad (x \geq 0)$$

we have the condition (4.4).

5. Appendix

We turn to a reformulation of Theorem 2.5.5 in [7] which is not quite in order since order statistics do not necessarily have the Markov property, if F is discontinuous (see [3]). Denote the degenerate distribution function concentrated at t by δ_t.

Theorem A.1. *Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables subject to a non-degenerate distribution function F. Suppose, for some $k < n$, that $X_{k+1:n} - X_{k:n}$ has the same distribution function as $X_{1:n-k}$. Then we have the following assertions:*

(i) *If F is continuous, it is necessarily an exponential distribution function.*

(ii) *If F is a lattice distribution function, then either*

(a) *it is a two-point distribution function concentrated at 0 and some $a > 0$, i.e.*

$$F = p\delta_0 + (1-p)\delta_a \qquad (0 < p < 1)$$

or

(b) *it is a mixture of the form*

$$F = p\delta_0 + (1-p)\sum_{n=0}^{+\infty} \frac{\theta^n}{1-\theta}\delta_{na} \qquad (0 < p < 1,\ 0 < a,\ 0 < \theta < 1).$$

Proof. Our assumption is that, for all x,

$$P(X_{k+1:n} - X_{k:n} > x) = \overline{F}^{n-k}(x). \tag{A.1}$$

Hence $F(0-) = 0$, and the X_j are necessarily non-negative random variables. Case (i) was settled in [9]. Therefore, we focus attention on the case that F is a lattice distribution function. By Proposition 4.1 we have the formula

$$P(X_{k+1:n} - X_{k:n} > x) = \binom{n}{k}\int_0^{+\infty} \overline{F}^{n-k}(x+u)\,dF^k(u). \tag{A.2}$$

Hence (A.1) is equivalent to

$$\binom{n}{k}\int_0^{+\infty} \overline{F}^{n-k}(x+u)\,F^k(du) = \overline{F}^{n-k}(x). \tag{A.3}$$

Consider the case $\binom{n}{k}F^k(0) > 1$. Then (A.3) implies that we must have $\overline{F}(x) = 0$ for $x > 0$, ruled out by assumption that $F \neq \delta_0$. Hence $\binom{n}{k}F^k(0) \le 1$.

Suppose firstly that $\binom{n}{k}F^k(0) = 1$. Then it turns out that F is of the type (ii)/(a), as shown in [6]. Secondly, consider the case $\binom{n}{k}F^k(0) < 1$. Then (A.3) is solvable using the information we have on solutions of the integrated Cauchy functional equations on $[0, +\infty)$ (as presented, for instance, in [7: Chapter 2]) and we conclude that F must be of the form (ii)/(b) ∎

References

[1] Ahsanullah, M.: *A characterization of the exponential distribution.* In: Statistical Distributions in Scientific Work, Vol. 3 (eds.: G. P. Patil et al.). Dordrecht - Boston: Reidel 1975, pp. 1131 - 135.

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

[3] Arnold, B. C., Becker, A., Gather, U. and H. Zahedi: *On the Markov property of order statistics.* J. Statist. Plan. Inference 9 (1984), 147 - 154.

[4] Gather, U.: *On a characterization of the exponential distribution by properties of order statistics.* Stat. & Prob. Letters 7 (1989), 93 - 96.

[5] Khatri, C. G.: *Distributions of order statistics for discrete case.* Ann. Inst. Stat. Math. 14 (1962), 167 - 171.

[6] Ramachandran, B.: *An integral equation in probability theory and its implications.* In: Statistics and Probability. Essays in Honor of C. R. Rao (eds.: G. Kallianpur et al.). Amsterdam: North-Holland 1982, pp. 609 - 616.

[7] Ramachandran, B. and K. Lau: *Functional Equations in Probability Theory.* New York - London: Academic Press 1991.

[8] Riedel, M. and H.-J. Rossberg: *Characterization of the exponential distribution function by properties of the difference $X_{k+s:n} - X_{k:n}$ of order statistics.* Metrika 41 (1994), 1 - 19.

[9] Rossberg, H.-J.: *Characterization of the exponential and the Pareto distributions by means of some properties of the distributions which differences and quotients of order statistics are subject to.* Math. Operationsf. Stat. 3 (1972), 207 - 216.

Received 13.05.1996