SWITCHING ORDER STATISTICS THROUGH RANDOM POWER CONTRACTIONS

JACEK WESOŁOWSKI^{1*} AND MOHAMMAD AHSANULLAH²

Politechnika Warszawska and Rider University

Summary

This paper investigates a new random contraction scheme which complements the lengthbiasing and convolution contraction schemes considered in the literature. A random power contraction is used with order statistics, leading to new and elegant characterizations of the power distribution. In view of Rossberg's counter-example of a non-exponential law with exponentially distributed spacings of order statistics, possibly the most appealing consequence of the result is a characterization of the exponential distribution via an independent exponential shift of order statistics.

Key words: exponential distribution; order statistics; power distribution; random contraction; random shifting.

1. Introduction

Let U_1, \ldots, U_n be independent identically distributed (iid) random variables (rvs) with a common distribution uniform on [0, 1]. Let (V_1, \ldots, V_n) and (U_1, \ldots, U_n) be iid random vectors. Then, as observed by Nevzorov (2001 Chapter 3),

$$U_{k:n} \stackrel{\mathrm{d}}{=} U_{k:m} V_{m+1:n} \,, \tag{1}$$

where $X_{k:n}$ denotes the *k*th order statistic from the sample (X_1, \ldots, X_n) .

This observation falls in the general random contraction setting, which can be described in the following way: let U be an rv with a distribution concentrated on [0, 1], and let X be a positive rv which is independent of U. Then the distribution of XU is a random contraction of the distribution of X. Assume that Y is an rv such that the distributions of X and Y are somehow related, and consider the equation

$$Y \stackrel{d}{=} XU \,. \tag{2}$$

Such schemes have been studied in the literature mainly in the context of identifiability and identification of the distribution of X at least in two cases:

(1) length-biasing, i.e. the distribution functions (cdfs) F_X of X and F_Y of Y are related by $E(Y^r)F_X(x) = \int_0^x y^r dF_Y(y)$, for any x > 0, while it is assumed that X and Y are positive and $E(Y^r) < \infty$ for some real r. Recent contributions in this area include Pakes (1996, 1997) and Pakes, Sapatinas & Fosam (1996).

Received September 2002; revised March 2003; accepted March 2003.

^{*}Author to whom correspondence should be addressed.

¹ Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Warszawa, Poland. e-mail: wesolo@alpha.mini.pw.edu.pl

² Dept of Management Science, Rider University, Lawrenceville, New Jersey, USA.

^{© 2004} Australian Statistical Publishing Association Inc. Published by Blackwell Publishing Asia Pty Ltd.

(2) convolutions, i.e. $X \stackrel{d}{=} Y_1 + \cdots + Y_m$, where Y_1, \ldots, Y_m are iid rvs with the same distribution as *Y*. Within such a scheme, given the distribution of *U* the law of the Y_j has been characterized for instance by Kotz & Steutel (1988), Milne & Yeo (1989), Yeo & Milne (1991) and Pakes (1994, 1995).

Often in these schemes the contraction distribution of U has been chosen to be the α -power distribution pow(1, α), where in general the pow(a, α) distribution, with positive parameters a and α is defined by the probability density function (pdf)

$$f(x) = \frac{\alpha}{a} \left(\frac{x}{a}\right)^{\alpha - 1} \mathbf{I}_{(0,a)}(x) \,.$$

In the order statistics scheme defined by (1) with m + 1 = n and $U = V_{n:n}$ we have that $U \stackrel{d}{=} pow(1, n)$. Observe also that if X_1, \ldots, X_n are iid rvs with $pow(1, \alpha)$ distribution, and (Z_1, \ldots, Z_n) and (X_1, \ldots, X_n) are iid, then (1) can be extended to

$$X_{k:n} \stackrel{\mathrm{d}}{=} X_{k:m} Z_{m+1:n} ,$$

which follows from (1) by the representation $X_{k:n} \stackrel{d}{=} U_{k:n}^{1/\alpha}$. If m+1 = n then the contracting variable $Z_{n:n} \stackrel{d}{=} pow(1, n\alpha)$.

In this paper we are interested in the following version of (2)

$$X_{k:n} \stackrel{\mathrm{d}}{=} X_{k:n-1} Z \,, \tag{3}$$

where $Z \stackrel{d}{=} pow(1, \alpha)$ is independent of the X_j . If $X_1 \stackrel{d}{=} pow(a, \alpha/n)$ for an arbitrary fixed a > 0 then (3) holds. We also consider

$$X_{k:n} \stackrel{\mathrm{d}}{=} X_{k+1:n} Z \,. \tag{4}$$

If $X_1 \stackrel{d}{=} pow(a, \alpha/k)$ then (4) holds. Converses of both results are treated in Section 2.

Section 3 is devoted to a related problem connected with the equation

$$X_{k:n-1}Z_1 \stackrel{d}{=} X_{k+1:n}Z_2,$$
 (5)

where $Z_i \stackrel{d}{=} \text{pow}(1, \alpha_i)$, i = 1, 2, are independent of the X_j . Again it follows that if $\alpha_1/n = \alpha_2/k = \alpha$, say, and $X_1 \stackrel{d}{=} \text{pow}(a, \alpha)$, for an arbitrary a > 0 then (5) holds true.

It is well known that if X_1, \ldots, X_n are iid exponential rvs, then $X_{k+1:n} - X_{k:n} \stackrel{d}{=} W$, where W has an exponential distribution, for any $k = 1, \ldots, n-1$. However Rossberg (1972) gives an example showing this property alone for a single value k does not characterize the parent distribution as exponential. The characterization holds true under some additional technical and rather unfriendly conditions; see e.g. Riedel & Rossberg (1994) or Rossberg, Riedel & Ramachandran (1997). It appears that our results on power distributions lead to characterizing the exponential law by equidistribution of $X_{k+1:n}$ and $X_{k:n} + W$, where W is an independent exponential rv — which looks rather unexpected in view of Rossberg's counter-example. For a recent discussion of related characterizations via equidistribution conditions for order statistics consult Gather, Kamps & Schweitzer (1998).

The approach we develop in this paper is applicable to characterizations of the logistic distribution based on exponential or Laplace random shifts of order statistics; see e.g. George

298

^{© 2004} Australian Statistical Publishing Association Inc.

& Mudholkar (1981) or George & Rousseau (1987). It appears that most of rather technical conditions used in these papers can be avoided; details are given by Markiewicz (2002). If $Z \stackrel{d}{=} pow(1, \alpha)$ then $W = -\log(Z) \stackrel{d}{=} exp(\alpha)$. Since $-\log$ is a decreasing func-

If $Z \stackrel{d}{=} pow(1, \alpha)$ then $W = -\log(Z) \stackrel{d}{=} exp(\alpha)$. Since $-\log$ is a decreasing function the following analogues of (3), (4) and (5) can be immediately formulated. Assume that $W \stackrel{d}{=} exp(\alpha)$ is independent of the rvs Y_1, \ldots, Y_n which are iid. If $Y_1 \stackrel{d}{=} exp(\alpha/n)$ then $Y_{k+1:n} \stackrel{d}{=} Y_{k:n-1} + W$. If $Y_1 \stackrel{d}{=} exp(\alpha/(n-k))$ then $Y_{k+1:n} \stackrel{d}{=} Y_{k:n} + W$. Let $W_1 \stackrel{d}{=} exp(\alpha)$ and $W_2 \stackrel{d}{=} exp((n-k)\alpha)$ and W_1, W_2 be independent of Y_1, \ldots, Y_n . If $Y_1 \stackrel{d}{=} exp(\alpha)$ then $Y_{k:n-1} + W_1 \stackrel{d}{=} Y_{k:n} + W_2$. The converse statements follow directly from our discussion of the power contraction scheme in Section 2.

We denote by $\{X_{k:n}, k = 1, ..., n\}$ order statistics from the iid rvs $X_1, ..., X_n$ with cdf F (f denotes the pdf if it exists). Similarly, $F_{k:n}$ ($f_{k:n}$) denotes the cdf (pdf if it exists) of respective order statistics, k = 1, ..., n.

The recurrence relations connecting cdfs and pdfs of order statistics gathered below, which are known (see e.g. David & Shu, 1978; David, 1981 p.25), are used in the proofs of main results in Sections 2 and 3.

For any k = 1, ..., n - 1 and any n = 2, 3, ...

(i)
$$F_{k:n} - F_{k:n-1} = {\binom{n-1}{k-1}} F^k (1-F)^{n-k},$$
 (6)

and if the pdf exists it has the form

$$n(F_{k:n} - F_{k:n-1})f = Ff_{k:n}.$$
(7)

(ii) Also, $F_{k:n} - F_{k+1:n} = \binom{n}{k} F^k (1-F)^{n-k}$, (8)

and if the pdf exists it has the form

$$k(F_{k:n} - F_{k+1:n})f = Ff_{k:n}.$$
(9)

(iii) If the pdf exists then

$$n(1-F)f_{k;n-1} = (n-k)f_{k;n}, \qquad (10)$$

$$k(1-F)f_{k+1:n} = (n-k)Ff_{k:n}.$$
(11)

2. One-sided contractions

This section gives two characterizations of the power distribution. Both are based on properties of switching order statistics by one-sided contraction, while the contracting rv has a power distribution in the interval [0, 1]. Such results, while giving a new insight into the structure of the order statistics distribution, also complement characterizations within other random contraction schemes considered in the literature, where random contraction is applied to length-biased distributions and convolutions.

Theorem 1. Let $U \stackrel{d}{=} pow(1, \alpha)$ for some $\alpha > 0$ be independent of X_1, \ldots, X_n , which are positive iid rvs. If

$$X_{k:n} \stackrel{\mathrm{d}}{=} X_{k:n-1} U \tag{12}$$

for an arbitrary but fixed $k \in \{1, ..., n-1\}$ then a > 0 exists such that $X_1 \stackrel{d}{=} pow(a, \alpha/n)$.

© 2004 Australian Statistical Publishing Association Inc.

Proof. By (12), for any x > 0,

$$F_{k:n}(x) = \int_0^1 F_{k:n-1}(x/u) \alpha u^{\alpha-1} \, du$$

Denote $a = \sup\{x > 0: F(x) < 1\} \le \infty$. Then (12) implies $\inf\{x > 0: F(x) > 0\} = 0$. Substituting t = x/u in the above integral we obtain

$$F_{k:n}(x) = \alpha \int_0^{x/a} u^{\alpha - 1} du + \alpha \int_{x/a}^1 F_{k:n-1}(x/u) u^{\alpha - 1} du$$

= $(x/a)^{\alpha} + \alpha x^{\alpha} \int_x^a F_{k:n-1}(t) t^{-\alpha - 1} dt$,

where it is understood that $(x/a)^{\alpha} = 0$ in the case $a = \infty$. Observe that the last expression is differentiable in x. Hence it follows that $F_{k:n}$ is also differentiable. Consequently the density f of X_1 and the densities of all order statistics exist. Upon differentiation we get

$$f_{k:n}(x) = \frac{\alpha x^{\alpha - 1}}{a^{\alpha}} + \alpha^2 x^{\alpha - 1} \int_x^a F_{k:n-1}(t) t^{-\alpha - 1} dt - \alpha x^{-1} F_{k:n-1}(x)$$

Substituting for the integral from the previous equation we arrive at

$$xf_{k:n}(x) = \alpha \left(F_{k:n}(x) - F_{k:n-1}(x) \right),$$
(13)

holding for any $x \in (0, a)$. By (6) for any $x \in (0, a)$, the difference $F_{k:n}(x) - F_{k:n-1}(x)$ is non-zero, since $\inf\{x > 0: F(x) > 0\} = 0$. Consequently the density f is always positive in (0, a) and (13) via (7) leads to

$$\frac{F'(x)}{F(x)} = \frac{\alpha}{nx}$$

whence $F(x) = Kx^{\alpha/n}$ for any $x \in (0, a)$, which yields $a < \infty$ and $K = a^{-\alpha/n}$.

Remark 1. For U in the above result we can take $U_{n:n}$, where U_1, \ldots, U_n is a random sample from the uniform (or, more generally, pow $(1, \alpha)$) distribution. Then (12) implies that X_1 is uniform on (0, a) (or, more generally, pow (a, α)).

By the duality between the power and exponential distribution mentioned in Section 1 we have the following consequence of Theorem 1.

Corollary 1. Let $W \stackrel{d}{=} \exp(\alpha)$ for some $\alpha > 0$ be independent of Y_1, \ldots, Y_n , which are iid *rvs.* If $Y_{k+1:n} \stackrel{d}{=} Y_{k:n-1} + W$ for an arbitrary and fixed $k \in \{1, \ldots, n-1\}$ then there exists $\gamma \in \mathbb{R}$ such that $Y_1 + \gamma \stackrel{d}{=} \exp(\alpha/n)$.

Now we are ready to consider the condition dual to (12).

Theorem 2. Let $U \stackrel{d}{=} pow(1, \alpha)$ for some $\alpha > 0$ be independent of X_1, \ldots, X_n , which are positive, iid rvs. If

$$X_{k:n} \stackrel{\mathrm{d}}{=} X_{k+1:n} U \tag{14}$$

for an arbitrary but fixed $k \in \{1, ..., n-1\}$ then a > 0 exists such that $X_1 \stackrel{d}{=} pow(a, \alpha/k)$.

© 2004 Australian Statistical Publishing Association Inc.

300

Proof. As in the proof of Theorem 1 denote $a = \sup\{x > 0: F(x) < 1\} \le \infty$ and observe that $\inf\{x > 0: F(x) > 0\} = 0$. By (14) we arrive at

$$F_{k:n}(x) = \int_0^1 F_{k+1:n}\left(\frac{x}{u}\right) \alpha u^{\alpha-1} \, du = \left(\frac{x}{a}\right)^{\alpha} + \alpha x^{\alpha} \int_x^a F_{k+1:n}(t) t^{-\alpha-1} \, dt$$

for any $x \in (0, a)$. Again the form of the above equation implies existence of densities. Then

$$xf_{k:n}(x) = \alpha \left(F_{k:n}(x) - F_{k+1:n}(x) \right) \qquad \left(x \in (0, a) \right).$$
(15)

By (8), $F_{k:n}(x) - F_{k+1:n}$ and, consequently, $f_{k:n}(x)$ are non-zero, and thus f(x) is also non-zero on (0, a). Thus by (9) we get from (15) for any $x \in (0, a)$

$$\frac{F'(x)}{F(x)} = \frac{\alpha}{kx} \, ,$$

leading to $F(x) = Kx^{\alpha/k}$ in (0, a), implying $a < \infty$ and $K = a^{-\alpha/k}$ and the hypothesis.

Remark 2. In Theorem 2, taking $U_{k:k}$ for U as a special case where U_1, \ldots, U_k is a random sample from the U(0, 1) (or pow(1, α)) distribution, results in the U(0, *a*) (or pow(*a*, α)) distribution for X_1 .

As for Theorem 1, an exponential analogue of Theorem 2 follows directly by the standard transformation leading from exponential to power distributions.

Corollary 2. Let $W \stackrel{d}{=} \exp(\alpha)$ for some $\alpha > 0$ be independent of Y_1, \ldots, Y_n which are iid *rvs.* If $Y_{k+1:n} \stackrel{d}{=} Y_{k:n} + W$ for an arbitrary but fixed $k \in \{1, \ldots, n-1\}$ then there exists $\gamma \in \mathbb{R}$ such that $Y_1 + \gamma \stackrel{d}{=} \exp(\alpha/(n-k))$.

We say again that the above result looks somewhat unexpected in light of Rossberg's (1972) counter-example — it is known that a closely related condition, $Y_{k+1:n} - Y_{k:n} \stackrel{d}{=} W$ does not characterize the exponential law.

3. Two-sided contractions

In this section we consider equality in distribution of random power contractions of two order statistics, i.e. we equate in distribution right-hand sides of (12) and (14) with different contraction variables. However, here we have to restrict our considerations to absolutely continuous distributions. An exponential analogue follows along the same lines as the respective results in the previous section.

Theorem 3. Let U_i be an rv with the power distribution $pow(1, \alpha_i)$, i = 1, 2, and let U_1 , U_2 be independent of X_1, \ldots, X_n , which are positive absolutely continuous iid rvs such that $inf\{x > 0: F(x) > 0\} = 0$. Assume that

$$X_{k:n-1}U_1 \stackrel{d}{=} X_{k+1:n}U_2$$
(16)

for an arbitrary but fixed $k \in \{1, ..., n-1\}$. Then $\alpha_1/n = \alpha_2/k = \alpha$, say, and there exists a > 0 such that $X_1 \stackrel{d}{=} pow(a, \alpha)$.

^{© 2004} Australian Statistical Publishing Association Inc.

Proof. Denote, as earlier, $a = \sup\{x > 0: F(x) < 1\} \le \infty$. Then (16) is equivalent to

$$\int_0^1 F_{k:n-1}\left(\frac{x}{u}\right) \alpha_1 u^{\alpha_1 - 1} \, du = \int_0^1 F_{k+1:n}\left(\frac{x}{u}\right) \alpha_2 u^{\alpha_2 - 1} \, du \qquad \left(x \in (0, a)\right),$$

and with similarity to the previous section, it follows that for any $x \in (0, a)$

$$\left(\frac{x}{a}\right)^{\alpha_1} + \alpha_1 x^{\alpha_1} \int_x^a F_{k:n-1}(t) t^{-\alpha_1 - 1} dt = \left(\frac{x}{a}\right)^{\alpha_2} + \alpha_2 x^{\alpha_2} \int_x^a F_{k+1:n}(t) t^{-\alpha_2 - 1} dt.$$

Consider first the case $\alpha_1 = \alpha_2$. Then the above equation implies $F_{k:n-1}(x) = F_{k+1:n}(x)$, for all $x \in (0, a)$. But this is impossible in view of (6) and (8). So, $\alpha_1 \neq \alpha_2$.

Taking the derivative with respect to x we get

$$x^{-\alpha_2} \left(\alpha_2 F_{k+1:n}(x) - \alpha_1 F_{k:n-1}(x) \right) = (\alpha_2 - \alpha_1) \left(a^{-\alpha_2} + \alpha_2 \int_x^a F_{k+1:n}(t) t^{-\alpha_2 - 1} dt \right).$$

Differentiating again we obtain, after some simple algebra,

$$x(\alpha_2 f_{k+1:n}(x) - \alpha_1 f_{k:n-1}(x)) = \alpha_1 \alpha_2 (F_{k+1:n}(x) - F_{k:n-1}(x)) \qquad (x \in (0,a)),$$

and by (6) and (8) the expressions on both sides of this equation are non-zero. But, by (10) and (11), the expression on the left-hand side equals

$$x\frac{n-k}{1-F(x)}\Big(\frac{\alpha_2 F(x)}{k}-\frac{\alpha_1}{n}\Big)f_{k:n}(x)\,,$$

so $f_{k:n}$ is non-zero, and hence f also is non-zero. Then using (7) and (9) we arrive at

$$\frac{f(x)}{F(x)(1-F(x))} \left(c_1 - c_2 F(x) \right) = \frac{c_1 c_2}{x} \qquad \left(x \in (0, a) \right),$$

where $c_1 = \alpha_1/n$ and $c_2 = \alpha_2/k$. The above equation implies $c_1 \ge c_2$. Also it can be rewritten as a simple differential equation in (0, a),

$$c_1 \frac{F'(x)}{F(x)} + (c_1 - c_2) \frac{F'(x)}{1 - F(x)} = \frac{c_1 c_2}{x}$$

Hence it follows that

$$F(x)^{c_1} \left(1 - F(x)\right)^{c_1 - c_2} = K x^{c_1 c_2}$$
(17)

for any $x \in (0, a)$, where K is a non-zero real constant. This implies $a < \infty$. Observe now that for $x \to a$ the left-hand side of (17) tends to zero if $c_1 > c_2$ while the right-hand side tends to $Ka^{c_1c_2} > 0$. Hence $c_1 = c_2 = \alpha$, say, and (17) then has the form $F(x) = K_1 x^{\alpha}$, $x \in (0, a)$, with $K_1 = a^{-\alpha}$.

Remark 3. Similarly to the results of Section 2, there is an interesting case in Theorem 3 if we take $U_{n:n}$ and $U_{k:k}$ for U_1 and U_2 , respectively, where U_1, \ldots, U_n is a sample from the U(0, 1) (or pow(1, α)) distribution. Then it follows that $X_1 \stackrel{d}{=} U(0, a)$ (or pow(a, α)).

Finally we have the exponential version of Theorem 3.

302

^{© 2004} Australian Statistical Publishing Association Inc.

Corollary 3. Let $W_i \stackrel{d}{=} \exp(\alpha_i)$, i = 1, 2, and let W_1 , W_2 be independent of Y_1, \ldots, Y_n , which are absolutely continuous iid rvs with $\sup\{x: F(x) < 1\} = \infty$. Assume that

$$Y_{k:n-1} + W_1 \stackrel{\mathrm{d}}{=} Y_{k:n} + W_2$$

for an arbitrary but fixed $k \in \{1, ..., n-1\}$. Then $\alpha_1/n = \alpha_2/(n-k) = \alpha$, say, and there exists $\gamma \in \mathbb{R}$ such that $Y_1 + \gamma \stackrel{d}{=} \exp(\alpha)$.

References

DAVID, H.A. (1981). Order Statistics. New York: Wiley.

- DAVID, H.A. & SHU, V.S. (1978). Robustness of location estimators in presence of an outlier. In Contributions to Survey Sampling and Applied Statistics, ed. H.A. David, pp.235–250. New York: Academic Press.
- GATHER, U., KAMPS, U. & SCHWEITZER, N. (1998). Characterizations of distributions via identically distributed functions of order statistics. In *Handbook of Statistics — 16: Order Statistics and Their Applications*, eds N. Balakrishnan & C.R. Rao, pp. 257–290. Amsterdam: North-Holland.
- GEORGE, E.O. & MUDHOLKAR, G.S. (1981). Some relationships between the logistic and the exponential distributions. In *Statistical Distributions in Scientific Work, Vol. 4*, eds C. Taillie, G.P. Patil & B.A. Baldessari, pp. 401–409. Dordrecht: Reidel.
- GEORGE, E.O. & ROUSSEAU, C.C. (1987). On the logistic midrange. Ann. Inst. Statist. Math. 39, 627–635.
- KOTZ, S. & STEUTEL, F.W. (1988). Note on a characterization of exponential distributions. *Statist. Probab. Lett.* 6, 201–203.
- MARKIEWICZ, I. (2002). Characterizations Related to Distributional Properties of Order Statistics and Records. MSc Thesis, Warsaw University Techn. (In Polish)
- MILNE, R.K. & YEO, G.F. (1989). Random sum characterizations. Math. Sci. 14, 120–126.
- NEVZOROV, V.B. (2001). *Records: Mathematical Theory*. Translations of Mathematical Monographs **194**. Providence: American Mathematical Society.
- PAKES, A.G. (1994). Necessary conditions for characterization of laws via mixed sums. Ann. Inst. Statist. Math. 63, 285–310.
- PAKES, A.G. (1995). Characterization of discrete laws via mixed sums and Markov branching processes. Stochastic Process. Appl. 55, 285–300.
- PAKES, A.G. (1996). Length biasing and laws equivalent to the log-normal. J. Math. Anal. Appl. 197, 825-854.
- PAKES, A.G. (1997). Characterization by invariance under length-biasing and random scaling. J. Statist. Plann. Inference 63, 285–310.
- PAKES, A.G., SAPATINAS, T. & FOSAM, E.B. (1996). Characterizations, length-biasing, and infinite divisibility. *Statist. Papers* 37, 53–69.
- RIEDEL, M. & ROSSBERG, H-J. (1994). Characterization of the exponential distribution function by properties of the difference $X_{k+s:n} X_{k:n}$ of order statistics. *Metrika* **41**, 1–19.
- ROSSBERG, H-J. (1972). Characterization of the exponential and the Pareto distributions by means of some properties of the distributions which the differences and quotients of order statistics are subject to. *Math. Operationsforsch. Statist.* 3, 207–216.
- ROSSBERG, H-J., RIEDEL, M. & RAMACHANDRAN, B. (1997). Characterization of the exponential distribution by properties of the difference $X_{k+s:n} X_{k:n}$ of order statistics. Z. Anal. Anwendungen 16, 191–200.
- YEO, G.F. & MILNE, R.K. (1991). On characterizations of beta and gamma distributions. *Statist. Probab. Lett.* 11, 239–242.