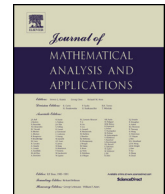




Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa



Characterizations of the logistic and related distributions

Chin-Yuan Hu^a, Gwo Dong Lin^{b,*}

^a Department of Business Education, National Changhua University of Education, Changhua 50058, Taiwan

^b Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan

ARTICLE INFO

Article history:

Received 1 April 2017

Available online xxxx

Submitted by V. Pozdnyakov

Keywords:

Characterization

Order statistics

Stochastic order

Logistic distribution

Exponential distribution

Pareto type II distribution

ABSTRACT

It is known that few characterization results of the logistic distribution were available before, although it is similar in shape to the normal one whose characteristic properties have been well investigated. Fortunately, in the last decade, several authors have made great progress in this topic. Some interesting characterization results of the logistic distribution have been developed recently. In this paper, we further provide some new results by the distributional equalities in terms of order statistics of the underlying distribution and the random exponential shifts. The characterization of the closely related Pareto type II distribution is also investigated.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

The logistic distribution is similar to a normal distribution in shape [16] and has an explicit closed form, so it has some advantages in practical applications. As remarked by Kotz [13], few characterizations of the logistic distribution were available before, but recently, some interesting results have been developed. In this paper, we will further provide some more new results by properties of order statistics.

We first introduce some notations. Let X obey the distribution F , denoted by $X \sim F$. Let $\{X_j\}_{j=1}^n$ be a random sample of size n from distribution F and denote the corresponding order statistics by $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$. The distribution function of $X_{k,n}$ is denoted by $F_{k,n}$. It is known that $F_{k,n}$ is the composition of $B_{k,n-k+1}$ and F (see, e.g., [11]), where $B_{\alpha,\beta}$ is the beta distribution with parameters $\alpha, \beta > 0$, namely,

$$F_{k,n}(x) = B_{k,n-k+1}(F(x)) = k \binom{n}{k} \int_0^{F(x)} t^{k-1} (1-t)^{n-k} dt, \quad x \in \mathbb{R} \equiv (-\infty, \infty), \quad (1)$$

* Corresponding author.

E-mail addresses: buhuua@gmail.com (C.-Y. Hu), gdlin@stat.sinica.edu.tw (G.D. Lin).

$$B_{\alpha,\beta}(u) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^u t^{\alpha-1}(1-t)^{\beta-1} dt, \quad u \in [0, 1]. \quad (2)$$

On the other hand, for $Y \sim G$, we say that X is less than or equal to Y in the usual stochastic order, denoted by $X \leq_{st} Y$, if $\overline{F} \leq \overline{G}$, where $\overline{F}(x) = 1 - F(x) = \Pr(X > x)$.

Let us start with an interesting simple example. Clearly, for general distribution F , we have $X_{1,2} \leq_{st} X$ because $F_{1,2} = 1 - \overline{F}^2 \geq F$ by (1). One possibility to adjust this “inequality” is to choose a nonnegative random variable Z , independent of X and X_j ’s, such that

$$X \stackrel{d}{=} X_{1,2} + Z \quad (3)$$

or

$$X - Z \stackrel{d}{=} X_{1,2}, \quad (4)$$

where $\stackrel{d}{=}$ means equality in distribution. One might think that the solutions of the distributional equations (3) and (4) are the same, but this is not true in general, because the characteristic function of Z is not equal to the reciprocal of that of $-Z$, namely, $E[e^{itZ}] \neq (E[e^{-itZ}])^{-1}$, $t \in \mathbb{R}$, in general. For example, if Z has the standard exponential distribution \mathcal{E} , then the solution of (3) is a logistic distribution $F(x) = 1/[1 + e^{-(x-\mu)}]$, $x \in \mathbb{R}$, where $\mu \in \mathbb{R}$ is a constant, while the solution of (4) is a negative (or reversed) exponential distribution $F(x) = e^{(x-\mu)/2}$, $x \leq \mu$, where $\mu \in \mathbb{R}$ is a constant (see, e.g., [9,14,1], and note that the smoothness conditions on F therein are redundant due to Lemmas 1–3 below).

Throughout the paper, let U and ξ obey the uniform distribution \mathcal{U} on $[0, 1]$ and the standard exponential distribution \mathcal{E} , respectively. Moreover, let $\{U_j\}_{j=1}^n$ and $\{U'_j\}_{j=1}^n$ be two random samples of size n from \mathcal{U} , and let $\{\xi_j\}_{j=1}^n$ and $\{\xi'_j\}_{j=1}^n$ be two random samples of size n from \mathcal{E} . All the above random variables X , U , ξ , X_j , U_j , U'_j , ξ_j and ξ'_j , $j = 1, 2, \dots, n$, are assumed to be independent from now on.

Mimicking the above characterization approaches (3) and (4), several authors have considered the general stochastic inequality $X_{k,n} \leq_{st} X_{k+1,n}$ and solved the distributional equations (a) $X_{k+1,n} \stackrel{d}{=} X_{k,n} + a\xi$ and (b) $X_{k,n} \stackrel{d}{=} X_{k+1,n} - b\xi$, or, more generally, (c) $X_{k,n} + a\xi_1 \stackrel{d}{=} X_{k+1,n} - b\xi_2$, where a and b are nonnegative constants. In particular, the equality

$$X_{k,n} + \frac{1}{n-k}\xi_1 \stackrel{d}{=} X_{k+1,n} - \frac{1}{k}\xi_2$$

also characterizes the logistic distribution. (See [4,19,2,3] for equations arising from $X_{k,n} \leq_{st} X_{k+r,n}$ with $1 \leq r \leq n-k$.) Besides, the distributional equations arising from (a) $X \leq_{st} X_{n,n}$, (b) $X_{1,n} \leq_{st} X_{n,n}$ and (c) $X_{m,m} \leq_{st} X_{n,n}$, where $m < n$, were investigated by Zykov and Nevzorov [20], Ananjevskii and Nevzorov [5] as well as Berred and Nevzorov [6], respectively.

In this paper we will solve the distributional equations arising from stochastic inequalities: (i) $X_{k,n} \leq_{st} X_{k,n-1}$, (ii) $X_{k,n-1} \leq_{st} X_{k+1,n}$, (iii) $X_{k,n} \leq_{st} X_{k,k}$, (iv) $X_{1,k} \leq_{st} X_{n-k+1,n}$, (v) $X_{k,n} \leq_{st} X_{k,n-m}$, and (vi) $X_{m-k,n-k} \leq_{st} X_{m,n}$.

To do this, some useful lemmas are given in the next section. The main characterization results are stated and proved in Sections 3 and 4. For simplicity, we first deal with the closely related Pareto type II distribution in Section 3, and then the logistic distribution in Section 4. Finally, we pose an open problem in Section 5.

2. Lemmas

We need some lemmas in the sequel. Lemma 1(i) was given without proof in [15, p. 38], but has been ignored in the literature. We provide a proof here for completeness.

Lemma 1.

- (i) Let Y and Z be two independent random variables. If Y has an absolutely continuous distribution, then so does $Y + Z$, regardless of the distribution of Z .
- (ii) If, in addition to the assumptions in (i), both Y and Z are positive random variables, then the product YZ has an absolutely continuous distribution.

Proof. Let F , G and H be the distributions of $Y + Z$, Y and Z , respectively. Then

$$F(x) = \int_{-\infty}^{\infty} G(x-z)dH(z), \quad x \in \mathbb{R}.$$

Since G is absolutely continuous, we have that for each $\varepsilon > 0$, there exists a $\delta > 0$ such that $\sum_{j=1}^n [G(b_j) - G(a_j)] < \varepsilon$ if $\sum_{j=1}^n (b_j - a_j) < \delta$, where $a_j < b_j \leq a_{j+1} < b_{j+1}$, $j = 1, 2, \dots, n-1$. This in turn implies that for the above $\{(a_j, b_j)\}_{j=1}^n$,

$$\sum_{j=1}^n [F(b_j) - F(a_j)] = \int_{-\infty}^{\infty} \sum_{j=1}^n [G(b_j - z) - G(a_j - z)]dH(z) < \int_{-\infty}^{\infty} \varepsilon dH(z) = \varepsilon.$$

Hence, part (i) is proved. To prove part (ii), we recall first that both the logarithmic and exponential functions are absolutely continuous, and that the composition preserves the property of absolute continuity. Then consider $\log(YZ) = \log Y + \log Z$ and use part (i) to conclude that $\log(YZ)$ has an absolutely continuous distribution, and hence so does YZ . This completes the proof. \square

It is known that the inverse function of an absolutely continuous function with positive derivative *almost everywhere* is not necessarily absolutely continuous. However, we have the following useful result.

Lemma 2. Let F be an absolutely continuous distribution on $[0, 1]$ and $F'(x) = f(x) > 0$ on $(0, 1)$. Then the inverse function of F is itself an absolutely continuous distribution.

Proof. By the assumptions, F is a strictly increasing and continuous function from $[0, 1]$ to $[0, 1]$, so is its inverse function F^{-1} . This implies that F^{-1} is a continuous distribution on $[0, 1]$. Moreover, $\frac{d}{dt} F^{-1}(t) = 1/f(F^{-1}(t))$ is a positive measurable function on $(0, 1)$ (see, e.g., [18, pp. 8–9]). By changing variables $x = F^{-1}(t)$, we have $\int_0^1 [\frac{d}{dt} F^{-1}(t)]dt = \int_0^1 [1/f(x)] \cdot f(x)dx = 1$. Therefore, the distribution function F^{-1} has no singular part and is absolutely continuous. The proof is complete. \square

Lemma 3. If the distribution $F_{k,n}$ of order statistic $X_{k,n}$ is absolutely continuous, then so is the underlying distribution F of X .

Proof. Recall (see (1)) that $F_{k,n}(x) = B_{k,n-k+1}(F(x))$, $x \in \mathbb{R}$, where $B_{k,n-k+1}$, defined in (2), is the beta distribution $B_{\alpha,\beta}$ with parameters $\alpha = k$ and $\beta = n - k + 1$, and has a positive continuous density function on $(0, 1)$. Therefore, $F = B_{k,n-k+1}^{-1} \circ F_{k,n}$ is absolutely continuous by Lemma 2. The proof is complete. \square

Lemma 4. Let $1 \leq k < n$. Then we have the following identities:

- (i) $F_{k,n} - F_{k+1,n} = \binom{n}{k} F^k \overline{F}^{n-k},$
- (ii) $F_{k,n} - F_{k,n-1} = \binom{n-1}{k-1} F^k \overline{F}^{n-k},$ and
- (iii) $F_{k,n-1} - F_{k+1,n} = \binom{n-1}{k} F^k \overline{F}^{n-k}.$

Proof. For parts (i) and (ii), see, e.g., [8] as well as [7, p. 23], while part (iii) follows from parts (i) and (ii) and the identity:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \quad \square$$

Denote the left and the right extremities of F by ℓ_F and r_F , respectively. It is known that if the absolutely continuous F satisfies the functional equation $F'(x) = F(x)(1 - F(x))$, $x \in (\ell_F, r_F)$, then F is a logistic distribution. In the next lemma we extend this result.

Lemma 5. Let $r, \theta > 0$, $a \in [0, 1]$, and let F be an absolutely continuous distribution function satisfying $x^a F'(x) = \theta F(x)(1 - F^r(x))$ for $x \in (\ell_F, r_F)$.

- (i) If $a = 1$, then $\ell_F = 0, r_F = +\infty$ and

$$F(x) = \left(\frac{\lambda x^{r\theta}}{1 + \lambda x^{r\theta}} \right)^{1/r}, \quad 0 \leq x < \infty,$$

where λ is a positive constant.

- (ii) If $a \in [0, 1)$, then $\ell_F = -\infty, r_F = +\infty$ and

$$F(x) = \left(\frac{\lambda \exp(\frac{r\theta}{1-a} x^{1-a})}{1 + \lambda \exp(\frac{r\theta}{1-a} x^{1-a})} \right)^{1/r}, \quad -\infty < x < \infty,$$

where λ is a positive constant.

Proof. Define the increasing function $G(x) = (1 - F^r(x))^{-1} - 1$ from (ℓ_F, r_F) onto $(0, \infty)$. Then $G'(x) = rF^{r-1}(x)F'(x)(1 - F^r(x))^{-2}$, and hence

$$x^a G'(x) = r\theta G(x), \quad x \in (\ell_F, r_F).$$

(a) If $a = 1$, solving the above equation leads to $G(x) = \lambda x^{r\theta}$, $x \in (\ell_F, r_F)$, for some constant $\lambda > 0$. On the other hand, we have, by the definition of G , that

$$F(x) = \left(\frac{G(x)}{1 + G(x)} \right)^{1/r} = \left(\frac{\lambda x^{r\theta}}{1 + \lambda x^{r\theta}} \right)^{1/r}, \quad x \in (\ell_F, r_F),$$

and hence, $\ell_F = 0$ and $r_F = +\infty$, because F is a distribution function. This proves part (i).

(b) If $a \in [0, 1)$, we have instead $G(x) = \lambda \exp(\frac{x^a}{1-a})$, $x \in (\ell_F, r_F)$, for some constant $\lambda > 0$. The required result then follows from both the definition of the function G and the fact that F is a distribution function. The proof is complete. \square

Some equalities (in distribution) of the next lemma are essentially due to [17, Lecture 3], but we provide here an alternative and possibly simpler proof.

Lemma 6. Let $\xi_{k,n}$ ($U_{k,n}$, resp.) be the k -th smallest order statistic of a random sample of size n from the standard exponential distribution \mathcal{E} (the uniform distribution \mathcal{U} , resp.). Then the following statements are true.

- (i) The Laplace transform of $\xi_{k,n}$ is $L_{\xi_{k,n}}(s) = \frac{n-k+1}{n-k+1+s} \cdot \frac{n-k+2}{n-k+2+s} \cdots \frac{n}{n+s}$, $s \geq 0$.
- (ii) $\xi_{k,n} \stackrel{d}{=} \sum_{j=1}^k \frac{1}{n-j+1} \xi_{k-j+1}$, where $1 \leq k \leq n$.
- (iii) $\xi_{m,n} \stackrel{d}{=} \xi_{m-k,n-k} + \xi'_{k,n}$, where $1 \leq k < m \leq n$.
- (iv) $U_{n-k+1,n} \stackrel{d}{=} \prod_{j=1}^k U_j^{1/(n-j+1)}$, where $1 \leq k \leq n$.
- (v) $U_{k,n} \stackrel{d}{=} U_{k,m-1} \cdot U'_{m,n}$, where $1 \leq k < m \leq n$.

Proof. Recall that the distribution of $\xi_{k,n}$ is $F_{\xi_{k,n}}(x) = k \binom{n}{k} \int_0^{\mathcal{E}(x)} t^{k-1} (1-t)^{n-k} dt$, $x \geq 0$, where $\mathcal{E}(x) = F_{\xi}(x) = 1 - e^{-x}$, $x \geq 0$. Then the Laplace transform of $\xi_{k,n}$ is

$$L_{\xi_{k,n}}(s) = E[e^{-s\xi_{k,n}}] = \int_0^{\infty} e^{-sx} dF_{\xi_{k,n}}(x) = k \binom{n}{k} \int_0^{\infty} e^{-(n-k+1+s)x} (1-e^{-x})^{k-1} dx, \quad s \geq 0.$$

By integration by parts, it follows from the above integral that

$$\begin{aligned} L_{\xi_{k,n}}(s) &= k \binom{n}{k} \frac{k-1}{n-k+1+s} \int_0^{\infty} e^{-(n-k+2+s)x} (1-e^{-x})^{k-2} dx = \cdots \\ &= k \binom{n}{k} \frac{k-1}{n-k+1+s} \cdot \frac{k-2}{n-k+2+s} \cdots \frac{1}{n-1+s} \cdot \int_0^{\infty} e^{-(n+s)x} dx \\ &= \frac{n-k+1}{n-k+1+s} \cdot \frac{n-k+2}{n-k+2+s} \cdots \frac{n-1}{n-1+s} \cdot \frac{n}{n+s}, \quad s \geq 0. \end{aligned}$$

This proves part (i), which in turn implies parts (ii) and (iii) by the fact that $E[e^{-s(\xi/k)}] = k/(k+s)$, $s \geq 0$. Part (iv) follows from part (ii) because $U_j \stackrel{d}{=} \exp(-\xi_{k-j+1})$ and the order statistic $U_{n-k+1,n} \stackrel{d}{=} \exp(-\xi_{k,n})$. To prove part (v), we have $U_{n-m+1,n} \stackrel{d}{=} U_{n-m+1,n-k} \cdot U'_{n-k+1,n}$ by using part (iii), and then reset $k = n - m + 1$. The proof is complete. \square

Lemma 7. Let Y obey the Pareto type II (or log-logistic) distribution $G(y) = y/(1+y)$, $y \geq 0$. Let $\{Y_j\}_{j=1}^n$, independent of U and $\{U_j\}_{j=1}^n$, be a random sample of size n from G . Then we have the following equalities in distribution:

- (i) $1/Y \stackrel{d}{=} Y$ and in general, $1/Y_{k,n} \stackrel{d}{=} Y_{n-k+1,n}$, where $1 \leq k \leq n$.
- (ii) $Y_{k,n-1} \stackrel{d}{=} Y_{k,n}/U^{1/(n-k)}$, where $1 \leq k \leq n-1$.

- (iii) $Y_{k,n-m} \stackrel{d}{=} Y_{k,n}/U_{n-k-m+1,n-k}$, where $1 \leq k \leq n-m$.
- (iv) $Y_{k,n-1} \stackrel{d}{=} Y_{k+1,n} \cdot U^{1/k}$, where $1 \leq k \leq n-1$.
- (v) $Y_{m-k,n-k} \stackrel{d}{=} Y_{m,n} \cdot U_{m-k,m-1}$, where $2 \leq k+1 \leq m \leq n$.

Proof. It is easy to check part (i). To prove the remaining parts, recall that the distribution of $Y_{k,n}$ is $G_{k,n}(y) = k \binom{n}{k} \int_0^{G(y)} t^{k-1} (1-t)^{n-k} dt$, $y \geq 0$. Then we have $H(y) \equiv \Pr(Y_{k,n}/U^{1/(n-k)} \leq y) = \int_0^1 G_{k,n}(yu^{1/(n-k)}) du$. By changing variables,

$$H(y) = \int_0^1 G_{k,n}(yt) dt^{n-k} = (n-k) \int_0^1 G_{k,n}(yt) t^{n-k-1} dt, \quad y \geq 0.$$

Therefore, $G_{k,n-1} = H$ iff, by differentiation,

$$\left(\frac{1}{1+y}\right)^n = \int_0^1 n \left(\frac{1}{1+yt}\right)^{n+1} t^{n-1} dt, \quad y \geq 0,$$

which is, however, a special case of the well-known identity

$$\left(\frac{1}{1+y}\right)^{\beta_1} = \frac{\Gamma(\beta_1 + \beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)} \int_0^1 \left(\frac{1}{1+yt}\right)^{\beta_1 + \beta_2} t^{\beta_1-1} (1-t)^{\beta_2-1} dt, \quad y \geq 0$$

(see [10, p. 314]). This proves part (ii). Part (iii) follows from part (ii) by iteration and Lemma 6(iv), while part (iv) follows from either parts (i) and (ii) (letting $k_* = n-k$) or Lemma 8(iii) below because $Y \stackrel{d}{=} \exp(X)$ and $U \stackrel{d}{=} \exp(-\xi)$. Finally, we prove part (v) by using part (iv), iteration and Lemma 6(iv) again. The proof is complete. \square

Lemma 8. Let X obey the standard logistic distribution $F(x) = 1/[1 + \exp(-x)]$, $x \in \mathbb{R}$. Then we have the following equalities in distribution:

- (i) $X_{k,n-1} \stackrel{d}{=} X_{k,n} + \frac{1}{n-k}\xi$, where $1 \leq k \leq n-1$.
- (ii) $X_{k,n-m} \stackrel{d}{=} X_{k,n} + \xi_{m,n-k}$, where $1 \leq k \leq n-m$.
- (iii) $X_{k,n-1} \stackrel{d}{=} X_{k+1,n} - \frac{1}{k}\xi$, where $1 \leq k \leq n-1$.
- (iv) $X_{m-k,n-k} \stackrel{d}{=} X_{m,n} - \xi_{k,m-1}$, where $2 \leq k+1 \leq m \leq n$.

Proof. The results follow from Lemma 7 by noting that (a) $X \stackrel{d}{=} \log Y$, (b) $\xi \stackrel{d}{=} -\log U$ and (c) $-\log U_{k,n} \stackrel{d}{=} \xi_{n-k+1,n}$ for all $1 \leq k \leq n$. The proof is complete. \square

For the proof of the next lemma, see [14, Lemma 5].

Lemma 9. Let f and g be two functions real analytic and strictly monotone in $[0, \infty)$. Assume that for each $n \geq 1$, the n -th derivatives $f^{(n)}$ and $g^{(n)}$ are strictly monotone in some interval $[0, \delta_n)$. Let $\{x_n\}_{n=1}^\infty$ be a sequence of positive real numbers converging to zero. If $f(x_n) = g(x_n)$, $n = 1, 2, \dots$, then $f = g$.

3. Characterizations of the Pareto type II distribution

We start with the Pareto type II distribution which is easier to handle, and recall that the uniform order statistic $U_{k,n} \sim B_{k,n-k+1}$.

Theorem 1. Let $Y \sim G$ be a positive random variable and let $1 \leq k \leq n-1$ be fixed integers. Let Y_1, Y_2, \dots, Y_n be n independent copies of Y , and let U , independent of $\{Y_i\}_{i=1}^n$, be a random variable with uniform distribution on $[0, 1]$. Then the distributional equality

$$Y_{k,n-1} \stackrel{d}{=} Y_{k,n}/U^{1/(n-k)} \quad (5)$$

holds iff G is a Pareto distribution $G(y) = \lambda y/(1 + \lambda y)$, $y \geq 0$, where $\lambda > 0$ is a constant.

Proof. The sufficiency part follows from Lemma 7(ii) because $(\lambda Y)_{\ell,m} \stackrel{d}{=} \lambda Y_{\ell,m}$ for $\lambda > 0$ and for all $1 \leq \ell \leq m$. To prove the necessity part, we note, by Lemma 1(ii), that the distribution $G_{k,n-1}$ of $Y_{k,n-1}$ is absolutely continuous, and so is G by Lemmas 2 and 3. Rewrite (5) as

$$G_{k,n-1}(y) = \int_0^1 G_{k,n}(yu^{1/(n-k)}) du, \quad y \geq 0.$$

By changing variables $t = yu^{1/(n-k)}$, we have

$$G_{k,n-1}(y) = (n-k)y^{-(n-k)} \int_0^y G_{k,n}(t)t^{n-k-1} dt, \quad y > 0.$$

Taking differentiation leads to

$$yG'_{k,n-1}(y) = (n-k)[G_{k,n}(y) - G_{k,n-1}(y)], \quad y > 0. \quad (6)$$

With the help of (1) and Lemma 4(ii), (6) is equivalent to

$$yG'(y) = G(y)[1 - G(y)], \quad y \in (\ell_G, r_G).$$

Finally, Lemma 5(i) with $r = \theta = 1$ completes the proof. \square

Corollary 1. Under the same assumptions of Theorem 1, the distributional equality $Y_{k,n-1} \stackrel{d}{=} Y_{k,n}/U^{1/\alpha}$ holds for some $\alpha > 0$, iff G is a Pareto distribution with $\bar{G}(y) = 1/[1 + \lambda y^{\alpha/(n-k)}]$, $y \geq 0$, where λ is a positive constant.

Proof. Consider $Y'_i = Y_i^{\alpha/(n-k)}$ for $i = 1, 2, \dots$. Then $Y'_{k,n} = Y_{k,n}^{\alpha/(n-k)}$, and apply Theorem 1 to the case: $Y'_{k,n-1} \stackrel{d}{=} Y'_{k,n}/U^{1/(n-k)}$. \square

Corollary 2. Under the same assumptions of Theorem 1, the distributional equality

$$Y_{k,n-1} \stackrel{d}{=} Y_{k+1,n} \cdot U^{1/k}$$

holds iff G is a Pareto distribution $G(y) = \lambda y/(1 + \lambda y)$, $y \geq 0$, where $\lambda > 0$ is a constant.

Proof. The sufficiency part is a consequence of Lemma 7(iv). To prove the necessity part, denote $Y^* = 1/Y$. Then $Y_{\ell,m}^* = (1/Y)_{\ell,m} \stackrel{d}{=} 1/Y_{m-\ell+1,m}$ for all $1 \leq \ell \leq m$. By assumptions, we have the equality $1/Y_{k,n-1} \stackrel{d}{=} 1/Y_{k+1,n} \cdot 1/U^{1/k}$, or, equivalently, $Y_{n-k,n-1}^* \stackrel{d}{=} Y_{n-k,n}^* \cdot 1/U^{1/(n-(n-k))}$. It follows from Theorem 1 (letting $k_* = n - k$) that Y^* has a Pareto distribution $G_*(y) = \lambda_* y / (1 + \lambda_* y)$, $y \geq 0$, for some constant $\lambda_* > 0$. This in turn implies that Y has the Pareto distribution $G(y) = \lambda y / (1 + \lambda y)$, $y \geq 0$, where $\lambda = 1/\lambda_*$. We can prove the last claim directly, or by using Lemma 7(i), because $\lambda_* Y^* \stackrel{d}{=} 1/(\lambda_* Y^*) = Y/\lambda_*$ having the standard Pareto type II distribution. The proof is complete. \square

Theorem 2. Let $Y \sim G$ be a positive random variable and let $1 \leq k \leq n-1$ be fixed integers. Let Y_1, Y_2, \dots, Y_n be n independent copies of Y , and let B , independent of $\{Y_i\}_{i=1}^n$, be a random variable having beta distribution $B_{\alpha,\beta}$ with parameters $\alpha = 1$ and $\beta = n - k$, that is, $F_B(u) = 1 - (1 - u)^{n-k}$, $u \in [0, 1]$. Assume further that $\lim_{y \rightarrow 0^+} G(y)/y = \lambda > 0$. Then the distributional equality

$$Y_{k,k} \stackrel{d}{=} Y_{k,n}/B \quad (7)$$

holds iff G is the Pareto distribution $G(y) = \lambda y / (1 + \lambda y)$, $y \geq 0$.

Proof. The sufficiency part follows from Lemma 7(iii) with $n - m = k$ and the fact $B \stackrel{d}{=} U_{1,n-k}$. To prove the necessity part, we note first that G is absolutely continuous by Lemmas 1–3, and then rewrite (7) as the functional equation:

$$G^k(y) = \int_0^1 \int_0^{G(yu)} k \binom{n}{k} t^{k-1} (1-t)^{n-k} dt dF_B(u), \quad y \geq 0. \quad (8)$$

Now, it suffices to prove that the solution of equation (7) is unique under the smoothness condition on the distribution. Namely, if the absolutely continuous distribution F on $(0, \infty)$ satisfies $\lim_{y \rightarrow 0^+} F(y)/y = \lambda > 0$ and

$$F^k(y) = \int_0^1 \int_0^{F(yu)} k \binom{n}{k} t^{k-1} (1-t)^{n-k} dt dF_B(u), \quad y \geq 0, \quad (9)$$

then we will prove that $F = G$. From (8) and (9) it follows that

$$|F^k(y) - G^k(y)| \leq \frac{1}{E[B^k]} \int_0^1 |F^k(yu) - G^k(yu)| dF_B(u), \quad y \geq 0, \quad (10)$$

where $E[B^k] = 1/\binom{n}{k}$. Define the bounded function

$$g(y) = \left| \frac{F^k(y) - G^k(y)}{y^k} \right|, \quad y > 0, \quad \text{and} \quad g(0) = \lim_{y \rightarrow 0^+} g(y) = 0,$$

and the increasing function

$$h(y) = \sup_{0 \leq t \leq y} g(t), \quad y > 0, \quad \text{and} \quad h(0) = \lim_{y \rightarrow 0^+} h(y) = 0.$$

By (10), we see that

$$g(y) \leq \int_0^1 g(uy) dH(u), \quad y \geq 0, \quad (11)$$

where $H(u) = (1/E[B^k]) \int_0^u t^k dF_B(t)$, $u \in [0, 1]$. Now, by (11) and the definition of the increasing function h , we have

$$h(y) \leq \int_0^1 h(uy) dH(u) \leq h(y) \int_0^1 dH(u) = h(y), \quad y \geq 0.$$

This in turn implies that h is a constant function and hence $h(y) = 0$, $y \geq 0$, because $h(0) = h(0^+) = 0$. Consequently, $g(y) = 0$, $y \geq 0$, and $F = G$. The proof is complete. \square

The next result is the counterpart of Theorem 2 for the minimum order statistics.

Corollary 3. *Under the same setting in Theorem 2 with the condition on G replaced by $\lim_{y \rightarrow 0^+} \overline{G}(1/y)/y = 1/\lambda > 0$ (equivalently, $\lim_{y \rightarrow +\infty} y\overline{G}(y) = 1/\lambda > 0$), the distributional equality $Y_{1,k} \stackrel{d}{=} Y_{n-k+1,n} \cdot B$ holds iff G is the Pareto distribution $G(y) = \lambda y/(1 + \lambda y)$, $y \geq 0$.*

Proof. Use Lemma 7(v), Theorem 2 and the fact $(1/Y)_{\ell,m} \stackrel{d}{=} 1/Y_{m-\ell+1,m}$ for all $1 \leq \ell \leq m$. \square

We now further extend Theorem 2 under some stronger smoothness conditions.

Theorem 3. *Let $Y \sim G$ be a positive random variable and let n, m, k be three fixed positive integers with $1 \leq k \leq n - m$. Let Y_1, Y_2, \dots, Y_n be n independent copies of Y , and let B_1 , independent of $\{Y_i\}_{i=1}^n$, be a random variable having beta distribution $B_{\alpha,\beta}$ with parameters $\alpha = n - m - k + 1$ and $\beta = m$. Assume further that the distribution function G satisfies the following conditions:*

- (i) G is real analytic and strictly increasing in $[0, \infty)$ and for each $i \geq 1$, its i -th derivative $G^{(i)}$ is strictly monotone in some interval $[0, \delta_i)$.
- (ii) $\lim_{y \rightarrow 0^+} [G^k(y) - (\lambda y)^k]/(\lambda y)^{k+1} = -k$ for some positive constant λ .

Then the distributional equality

$$Y_{k,n-m} \stackrel{d}{=} Y_{k,n}/B_1 \quad (12)$$

holds iff G is the Pareto distribution $G(y) = \lambda y/(1 + \lambda y)$, $y \geq 0$.

Proof. The sufficiency part follows from Lemma 7(iii) and the fact $B_1 \stackrel{d}{=} U_{n-m-k+1,n-k}$. To prove the necessity part, we note first that G is absolutely continuous as before, and then we rewrite (12) as the functional equation:

$$G_{k,n-m}(y) = \int_0^1 G_{k,n}(yu) dF_{B_1}(u), \quad y \geq 0. \quad (13)$$

Now, it suffices to prove that the solution of equation (12) is unique under the smoothness condition on the distribution. Namely, if the absolutely continuous distribution F on $(0, \infty)$ satisfies the above conditions (i) and (ii) and

$$F_{k,n-m}(y) = \int_0^1 F_{k,n}(yu) dF_{B_1}(u), \quad y \geq 0, \quad (14)$$

then we will prove that $F = G$.

Define the increasing function $H(y) = \max\{F(y), G(y)\}$, $y \geq 0$. From (1) it follows that for any $a > 0$ and $0 \leq y \leq a$,

$$\begin{aligned} |F_{k,n-m}(y) - G_{k,n-m}(y)| &= k \binom{n-m}{k} \left| \int_{G(y)}^{F(y)} t^{k-1} (1-t)^{n-m-k} dt \right| \\ &\geq k \binom{n-m}{k} (1-H(y))^{n-m-k} \frac{1}{k} |F^k(y) - G^k(y)| \\ &\geq \binom{n-m}{k} (1-H(a))^{n-m-k} |F^k(y) - G^k(y)|. \end{aligned} \quad (15)$$

On the other hand, we have

$$|F_{k,n}(y) - G_{k,n}(y)| \leq \binom{n}{k} |F^k(y) - G^k(y)|, \quad y \geq 0. \quad (16)$$

Combing (12)–(16) leads to

$$\begin{aligned} \binom{n-m}{k} (1-H(a))^{n-m-k} |F^k(y) - G^k(y)| &\leq |F_{k,n-m}(y) - G_{k,n-m}(y)| \\ &\leq \int_0^1 |F_{k,n}(yu) - G_{k,n}(yu)| dF_{B_1}(y) \leq \binom{n}{k} \int_0^1 |F^k(yu) - G^k(yu)| dF_{B_1}(y). \end{aligned} \quad (17)$$

Define the bounded increasing function

$$h(y) = \sup_{0 < t \leq y} \left| \frac{F^k(t) - G^k(t)}{t^{k+1}} \right|, \quad y > 0, \quad \text{and} \quad h(0) = \lim_{y \rightarrow 0^+} h(y) = 0.$$

Then from the inequality (17) it follows that for any $a > 0$,

$$\begin{aligned} \binom{n-m}{k} (1-H(a))^{n-m-k} h(y) &\leq \binom{n}{k} \int_0^1 h(yu) u^{k+1} dF_{B_1}(y) \\ &\leq \binom{n}{k} h(y) \int_0^1 u^{k+1} dF_{B_1}(u) = \binom{n}{k} h(y) E[B_1^{k+1}], \quad 0 \leq y \leq a. \end{aligned} \quad (18)$$

Recall that $E[B_1^k] = \binom{n-m}{k} / \binom{n}{k}$. Then rewrite the inequality (18) as follows:

$$E[B_1^k](1 - H(a))^{n-m-k}h(y) \leq h(y)E[B_1^{k+1}], \quad 0 \leq y \leq a, \quad a > 0. \quad (19)$$

We claim that there exists a $y_0 > 0$ such that $h(y_0) = 0$. Otherwise, we have, by (19),

$$E[B_1^k](1 - H(a))^{n-m-k} \leq E[B_1^{k+1}], \quad \forall a > 0,$$

which in turn implies, by letting $a \rightarrow 0^+$, that $E[B_1^k] \leq E[B_1^{k+1}]$, a contradiction. Therefore, $h(y_0) = 0$ for some $y_0 > 0$ and hence $F(y) = G(y)$ for $y \in [0, y_0]$. By Lemma 9 and the assumptions on F and G , we conclude that $F = G$. The proof is complete. \square

Corollary 4. Let $Y \sim G$ be a positive random variable and let n, m, k be three fixed positive integers with $k + 1 \leq m \leq n$. Let Y_1, Y_2, \dots, Y_n be n independent copies of Y , and let B_2 , independent of $\{Y_i\}_{i=1}^n$, be a random variable having beta distribution $B_{\alpha, \beta}$ with parameters $\alpha = m - k$ and $\beta = k$. Assume further that the distribution function G_* of $1/Y$ satisfies the following conditions:

- (i) G_* is real analytic and strictly increasing in $[0, \infty)$ and for each $i \geq 1$, its i -th derivative $G_*^{(i)}$ is strictly monotone in some interval $[0, \delta_i)$.
- (ii) $\lim_{y \rightarrow 0^+} [G_*^{k_*}(y) - (y/\lambda)^{k_*}] / (y/\lambda)^{k_*+1} = -k_*$ for some positive constant λ , where $k_* = n - m + 1$.

Then the distributional equality $Y_{m-k, n-k} \stackrel{d}{=} Y_{m, n} \cdot B_2$ holds iff G is the Pareto distribution $G(y) = \lambda y / (1 + \lambda y)$, $y \geq 0$.

Proof. Use Lemma 7(v), Theorem 3 and the fact $(1/Y)_{\ell, m} \stackrel{d}{=} 1/Y_{m-\ell+1, m}$ for all $1 \leq \ell \leq m$. \square

In summary, for a positive random variable $Y \sim G$, we have the following characteristic properties of the Pareto distribution $G(y) = \lambda y / (1 + \lambda y)$, $y \geq 0$, where λ is a positive constant (compare with Lemma 7).

1. $Y_{k, n-1} \stackrel{d}{=} Y_{k, n} / U^{1/(n-k)}$.
2. $Y_{k, n-1} \stackrel{d}{=} Y_{k+1, n} \cdot U^{1/k}$.
3. $Y_{k, k} \stackrel{d}{=} Y_{k, n} / B$ (equivalently, $Y_{k, k} \stackrel{d}{=} Y_{k, n} / U_{1, n-k}$).
4. $Y_{1, k} \stackrel{d}{=} Y_{n-k+1, n} \cdot B$ (equivalently, $Y_{1, k} \stackrel{d}{=} Y_{n-k+1, n} \cdot U_{1, n-k}$).
5. $Y_{k, n-m} \stackrel{d}{=} Y_{k, n} / B_1$ (equivalently, $Y_{k, n-m} \stackrel{d}{=} Y_{k, n} / U_{n-m-k+1, n-k}$).
6. $Y_{m-k, n-k} \stackrel{d}{=} Y_{m, n} \cdot B_2$ (equivalently, $Y_{m-k, n-k} \stackrel{d}{=} Y_{m, n} \cdot U_{m-k, m-1}$).

Here, the random variables $U \sim \mathcal{U}$, $B \sim B_{1, n-k}$, $B_1 \sim B_{n-m-k+1, m}$, $B_2 \sim B_{m-k, k}$, and on the RHS of each equality, the two random variables are independent.

4. Characterizations of the logistic distribution

We are now ready to provide characterization results of the logistic distribution.

Theorem 4. Let $X \sim F$ and let $1 \leq k \leq n - 1$ be fixed integers. Then the distributional equality

$$X_{k, n-1} \stackrel{d}{=} X_{k, n} + \frac{1}{n-k} \xi$$

holds iff F is a logistic distribution $F(x) = 1/[1 + e^{-(x-\mu)}]$, $x \in \mathbb{R}$, where $\mu \in \mathbb{R}$ is a constant.

Proof. Take $Y_i = \exp(X_i)$, $U = \exp(-\xi)$ and $\lambda = e^{-\mu}$. Then the result follows immediately from Theorem 1. \square

Corollary 5. Let $X \sim F$, $\alpha > 0$ and let $1 \leq k \leq n-1$ be fixed integers. Then the distributional equality $X_{k,n-1} \stackrel{d}{=} X_{k,n} + \frac{1}{\alpha}\xi$ holds iff F is a logistic distribution $F(x) = 1/\{1 + e^{-[\alpha/(n-k)](x-\mu)}\}$, $x \in \mathbb{R}$, where $\mu \in \mathbb{R}$ is a constant.

Corollary 6. Let $X \sim F$, $\alpha > 0$ and let $1 \leq k \leq n-1$ be fixed integers. Then the distributional equality

$$X_{k,n-1} \stackrel{d}{=} X_{k+1,n} - \frac{1}{\alpha}\xi \quad (20)$$

holds iff F is a logistic distribution $F(x) = 1/[1 + e^{-(\alpha/k)(x-\mu)}]$, $x \in \mathbb{R}$, where $\mu \in \mathbb{R}$ is a constant.

Proof. Use Corollary 5 and the fact that $X_{\ell,m} \stackrel{d}{=} -(-X)_{m-\ell+1,m}$ for all $1 \leq \ell \leq m$. \square

The counterpart of (20), namely, $X_{k+1,n} \stackrel{d}{=} X_{k,n-1} + \frac{1}{\alpha}\xi$, and the two-sided case: $X_{k,n} + a\xi_1 \stackrel{d}{=} X_{k,n-1} + b\xi_2$ (see Corollary 5), where $\alpha, a, b > 0$, were investigated by Wesolowski and Ahsanullah [19]. All the solutions of these two equations are exponential distributions.

The next result improves and extends Theorem 6 of [14] by an approach different from the previous method of intensively monotone operator [12].

Theorem 5. Let $1 \leq k \leq n-1$ be fixed integers. Assume that $X \sim F$ satisfies $\lim_{x \rightarrow -\infty} F(x)/e^x = e^{-\mu}$ for some constant $\mu \in \mathbb{R}$. Then the distributional equality

$$X_{k,k} \stackrel{d}{=} X_{k,n} + \xi_{n-k,n-k}$$

holds iff F is the logistic distribution $F(x) = 1/[1 + e^{-(x-\mu)}]$, $x \in \mathbb{R}$.

Proof. The sufficiency part follows from Lemma 8(ii), while the necessity part is a consequence of Theorem 2. \square

The next result is the counterpart of Theorem 5 for the minimum order statistics.

Corollary 7. Let $1 \leq k \leq n-1$ be fixed integers. Assume that $X \sim F$ satisfies $\lim_{x \rightarrow -\infty} \overline{F}(-x)/e^x = e^\mu$ for some constant $\mu \in \mathbb{R}$. Then the distributional equality

$$X_{1,k} \stackrel{d}{=} X_{n-k+1,n} - \xi_{n-k,n-k}$$

holds iff F is the logistic distribution $F(x) = 1/[1 + e^{-(x-\mu)}]$, $x \in \mathbb{R}$.

Proof. Use Lemma 8(iv), Theorem 5 and the fact $X_{\ell,m} \stackrel{d}{=} -(-X)_{m-\ell+1,m}$ for all $1 \leq \ell \leq m$. \square

Using Theorem 3 and Lemma 8(ii), we further extend Theorem 5 to the following.

Theorem 6. Let n, m, k be three fixed positive integers with $1 \leq k \leq n-m$ and let $X \sim F$ satisfy $\lim_{x \rightarrow -\infty} [e^{-k(x-\mu)} F^k(x) - 1]/e^{x-\mu} = -k$ for some constant $\mu \in \mathbb{R}$. Assume further that the distribution function G of $\exp(X_1)$ is real analytic and strictly increasing in $[0, \infty)$ and that for each $i \geq 1$, its i -th derivative $G^{(i)}$ is strictly monotone in some interval $[0, \delta_i)$. Then the distributional equality

$$X_{k,n-m} \stackrel{d}{=} X_{k,n} + \xi_{m,n-k}$$

holds iff F is the logistic distribution $F(x) = 1/[1 + e^{-(x-\mu)}]$, $x \in \mathbb{R}$.

As before, Theorem 6 and Lemma 8(iv) together lead to the following.

Corollary 8. Let n, m, k be three fixed positive integers with $k + 1 \leq m \leq n$ and let $X \sim F$ satisfy $\lim_{x \rightarrow -\infty} [e^{-k_*(x+\mu)} (\overline{F}(-x))^{k_*} - 1] / e^{x+\mu} = -k_*$ for some constant $\mu \in \mathbb{R}$, where $k_* = n - m + 1$. Assume further that the distribution function G_* of $\exp(-X_1)$ is real analytic and strictly increasing in $[0, \infty)$ and that for each $i \geq 1$, its i -th derivative $G_*^{(i)}$ is strictly monotone in some interval $[0, \delta_i)$. Then the distributional equality

$$X_{m-k,n-k} \stackrel{d}{=} X_{m,n} - \xi_{k,m-1} \quad (21)$$

holds iff F is the logistic distribution $F(x) = 1/[1 + e^{-(x-\mu)}]$, $x \in \mathbb{R}$.

In summary, for a random variable $X \sim F$, we have the following characteristic properties of the logistic distribution $F(x) = 1/[1 + e^{-(x-\mu)}]$, $x \in \mathbb{R}$ (compare with Lemma 8). Here, $\mu \in \mathbb{R}$, $\xi \sim \mathcal{E}$, and on the RHS of each equality, the two random variables are independent.

1. $X_{k,n-1} \stackrel{d}{=} X_{k,n} + \frac{1}{n-k} \xi$.
2. $X_{k,n-1} \stackrel{d}{=} X_{k+1,n} - \frac{1}{k} \xi$.
3. $X_{k,k} \stackrel{d}{=} X_{k,n} + \xi_{n-k,n-k}$.
4. $X_{1,k} \stackrel{d}{=} X_{n-k+1,n} - \xi_{n-k,n-k}$.
5. $X_{k,n-m} \stackrel{d}{=} X_{k,n} + \xi_{m,n-k}$.
6. $X_{m-k,n-k} \stackrel{d}{=} X_{m,n} - \xi_{k,m-1}$.

5. Open problem

Finally, we would like to pose the following problem in which part (i) is the counterpart of (21) for exponential distribution, and in part (ii), the first two cases, $m = k + 1, k + 2$, have been solved by AlZaid and Ahsanullah [4] and Ahsanullah et al. [2].

Problem. Let $X \sim F$ and let $1 \leq k < m \leq n$ be fixed integers. Then solve the general distributional equations: (i) $X_{m,n} \stackrel{d}{=} X_{m-k,n-k} + \xi_{k,n}$ and (ii) $X_{m,n} \stackrel{d}{=} X_{k,n} + \xi_{m-k,n-k}$.

Acknowledgments

The authors would like to thank the Editor-in-Chief and the Referee for helpful comments and constructive suggestions which improve the presentation of the paper.

References

- [1] M. Ahsanullah, A. Berred, V.B. Nevzorov, On characterizations of the exponential distributions, J. Appl. Statist. Sci. 19 (2011) 37–43.
- [2] M. Ahsanullah, V.B. Nevzorov, G.P. Yanev, Characterizations of distributions via order statistics with random exponential shifts, J. Appl. Statist. Sci. 18 (2010) 297–305.
- [3] M. Ahsanullah, G.P. Yanev, C. Onica, Characterizations of logistic distribution through order statistics with independent exponential shifts, Econ. Qual. Control 27 (2012) 85–96.

- [4] A.A. AlZaid, M. Ahsanullah, A characterization of the Gumbel distribution based on record values, *Comm. Statist. Theory Methods* 32 (2003) 2101–2108.
- [5] S.M. Ananjevskii, V.B. Nevzorov, On families of distributions characterized by certain properties of ordered random variables, *Vestnik St. Petersburg Univ. Math.* 49 (2016) 197–203.
- [6] A. Berred, V.B. Nevzorov, Characterizations of families of distributions, which include the logistic one, by properties of order statistics, *J. Math. Sci.* 188 (2013) 673–676.
- [7] H.A. David, H.N. Nagaraja, *Order Statistics*, 3rd edn., Wiley, New Jersey, 2003.
- [8] H.A. David, V.S. Shu, Robustness of location estimators in presence of an outlier, in: H.A. David (Ed.), *Contributions to Survey Sampling and Applied Statistics*, Academic Press, New York, 1978, pp. 235–250.
- [9] E.O. George, G.S. Mudholkar, On the logistic and exponential laws, *Sankhyā, Ser. A* 44 (1982) 291–293.
- [10] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products*, 8th edn., Elsevier, New York, 2014.
- [11] J.S. Hwang, G.D. Lin, Characterizations of distributions by linear combinations of moments of order statistics, *Bull. Inst. Math. Acad. Sin.* 12 (1984) 179–202.
- [12] A.V. Kakosyan, L.B. Klebanov, J.A. Melamed, *Characterization of Distributions by the Method of Intensively Monotone Operators*, Springer, New York, 1984.
- [13] S. Kotz, Characterizations of statistical distributions: a supplement to recent surveys, *Int. Stat. Rev.* 42 (1974) 39–65.
- [14] G.D. Lin, C.-Y. Hu, On characterizations of the logistic distribution, *J. Statist. Plann. Inference* 138 (2008) 1147–1156.
- [15] E. Lukacs, *Characteristic Functions*, 2nd edn., Hafner Pub. Co., New York, 1970.
- [16] G.S. Mudholkar, E.O. George, A remark on the shape of the logistic distribution, *Biometrika* 65 (1978) 667–668.
- [17] V.B. Nevzorov, *Records: Mathematical Theory*, Translations of Mathematical Monographs, vol. 194, Amer. Math. Soc., Rhode Island, 2001.
- [18] G.R. Shorack, J.A. Wellner, *Empirical Processes with Applications to Statistics*, Wiley, New York, 1986.
- [19] J. Wesolowski, M. Ahsanullah, Switching order statistics through random power contractions, *Aust. N. Z. J. Stat.* 46 (2004) 297–303.
- [20] V.O. Zykov, V.B. Nevzorov, Some characterizations of families of distributions, including logistic or exponential ones, by properties of order statistics, *J. Math. Sci.* 176 (2011) 203–206.