

Lorenz ordering of order statistics from log-logistic and related distributions

Christian Kleiber

Fachbereich Statistik, Universität Dortmund,

44221 Dortmund, Germany.

E-mail: Kleiber@statistik.uni-dortmund.de

February 2002

Abstract

This paper obtains Lorenz ordering relationships among order statistics from log-logistic samples of possibly different sizes. Some results extend other families including the Lomax, Burr III and Burr XII distributions.

Key words: k -out-of- n systems, order statistics, Lorenz order, (generalized) log-logistic distributions.

AMS 2000 classification: 60E15, 60K10.

1 Introduction

Order statistics are important in many branches of statistics and applied probability. For example, in reliability theory they arise in connection with the lifetime of a k -out-of- n system. It is then natural to study aspects of size and/or variability in the distributions of the order statistics $X_{i:n}$, here defined in the ascending order by $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$. These may be investigated via functionals of the distributions, e.g. the moments, or globally using stochastic orders. Since life lengths are nonnegative by definition, the Lorenz variability order originating from economics is a natural tool for studying variability in a global sense. It is also known as the harmonic-new-better-than-used-in-expectation (HNBUE) order in a reliability framework (Kochar, 1989).

Following Arnold (1987), let \mathcal{L} denote the set of non-negative random variables with finite and positive first moment. For $X \in \mathcal{L}$ with c.d.f. F and quantile function $F^{-1}(t) = \sup\{x \mid F(x) \leq t\}$, $t \in [0, 1]$, the Lorenz curve is defined by

$$L(u) = \frac{1}{E(X)} \int_0^u F^{-1}(t) dt, \quad u \in [0, 1].$$

For $X_1, X_2 \in \mathcal{L}$, the random variable X_1 is said to be more variable than X_2 in the Lorenz sense if $L_1(u) \leq L_2(u)$ for all $u \in [0, 1]$. This is denoted as $X_1 \geq_L X_2$.

Lorenz ordering of order statistics has been studied for various lifetime parent distributions, including the exponential, uniform, Pareto, and power function distributions. See Arnold and Villaseñor (1998) for a recent survey of the field as well as further references, and Wilfling (1996b) for additional results on the power function distribution. However, the only distribution for which a complete characterization of the Lorenz ordering among its order statistics is available is the exponential distribution (Arnold and Nagaraja, 1991). This paper settles the problem for a second distribution by providing a necessary and sufficient condition for the Lorenz ordering of two log-logistic order statistics from random samples of possibly different sizes. The characterization is obtained by combining results from the statistics and econometrics literatures. Some results extend to generalized log-logistic distributions, including the Burr III and Burr XII distributions and related models.

2 Lorenz ordering of order statistics from a two-parameter log-logistic distribution

The c.d.f. of the two-parameter log-logistic distribution is given by

$$F(x) = 1 - [1 + (x/b)^a]^{-1}, \quad x > 0.$$

Here b is a scale and a is a shape parameter. If a random variable X follows a log-logistic distribution this is denoted as $X \sim LL(a, b)$. It is worth noting that the log-logistic distribution is one of the few distributions for which p.d.f., c.d.f. and quantile function are all available in simple closed form. Furthermore, its hazard rate allows for a variety of shapes, it can be monotonically decreasing as well as first increasing, then decreasing, an important aspect for modelling purposes.

Log-logistic order statistics have been studied by Ragab and Green (1984), Ali and Khan (1987), Balakrishnan and Malik (1987, 1988) and Al-Shboul and Khan (1989). For a $LL(a, b)$ parent distribution the p.d.f. of $X_{i:n}$ is given by

$$\begin{aligned} f_{i:n}(x) &= i \binom{n}{i} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x) \\ &= \frac{n! a x^{ai-1}}{(i-1)! (n-i)! b^{ai} [1 + (x/b)^a]^{n+1}}, \quad x > 0, \end{aligned}$$

(Arnold, 1983, Ragab and Green, 1984). Arnold (1983, p. 60) observes that this is the density of a certain generalized Pareto distribution which he calls the Feller-Pareto distribution. The general form of the Feller-Pareto p.d.f. is

$$f(x) = \frac{ax^{ap-1}}{b^{ap} B(p, q) [1 + (x/b)^a]^{p+q}}, \quad x > 0, \quad (1)$$

where all four parameters a, b, p, q are positive (Arnold, 1983, includes an additional location parameter). The distribution is known under several other names: in econometrics it is called a generalized beta distribution of the second kind (McDonald, 1984), hereafter: GB2, whereas in actuarial science the term transformed beta distribution is used (Venter, 1983). In this paper, (1) is referred to as the density of a GB2 distribution. Note that $LL(a, b) \equiv GB2(a, b, 1, 1)$, so that

the GB2 distribution may be considered as a generalized log-logistic distribution. Thus, for a log-logistic parent distribution

$$X_{i:n} \sim GB2(a, b, i, n - i + 1). \quad (2)$$

The following Theorem provides a complete characterization of the Lorenz ordering among order statistics from two samples from the same log-logistic distribution.

Theorem 1 *If $X \sim LL(a, b)$, with $a > 1$, then the following are equivalent:*

(a) $X_{i:n} \geq_L X_{j:m}$.

(b) $0 \leq j - i \leq m - n$.

Proof: The assumption $a > 1$ assures the existence of the mean and therefore of the Lorenz curve. Since the Lorenz order is scale free the parameter b plays no role in Lorenz ordering results and may be set equal to one without loss of generality.

Sufficiency of the condition follows from Lorenz ordering results for the GB2 distribution, namely for $X_i \sim GB2(a_i, b_i, p_i, q_i)$, $i = 1, 2$,

$$a_1 \leq a_2, \quad a_1 p_1 \leq a_2 p_2, \quad \text{and} \quad a_1 q_1 \leq a_2 q_2 \quad \text{together imply} \quad X_1 \geq_L X_2. \quad (3)$$

(Kleiber, 1999). The necessary conditions for $X_1 \geq_L X_2$ are

$$a_1 p_1 \leq a_2 p_2, \quad \text{and} \quad a_1 q_1 \leq a_2 q_2, \quad (4)$$

see Wilfling (1996a) for a direct proof and Kleiber (2000) for a more general approach using regularly varying functions. A comparison of these conditions with (1) and (2) yields the result. ■

Theorem 1 says that a log-logistic order statistic with a smaller index coming from a smaller sample is more variable than a log-logistic order statistic with a larger index coming from a larger sample.

The log-logistic distribution is only the second distribution for which a complete characterization of Lorenz ordering among order statistics from two samples (from the same distribution) is

available, the first being the exponential distribution which was studied by Arnold and Nagaraja (1991).

Theorem 1 may be generalized to log-logistic distributions with $a \leq 1$, in particular to the one-parameter log-logistic distribution with $a = 1$. In these cases certain order statistics cannot be considered because their mean – and consequently their Lorenz curve – does not exist. For a general GB2 distribution, $E(X) < \infty$ iff $aq > 1$, hence $a(n-i+1) > 1$ is the required condition in the present context. In particular, for the one-parameter ($a = 1$) log-logistic distribution only sample maxima must be excluded to assure $n - i + 1 > 1$, for the remaining order statistics the condition $0 \leq j - i \leq m - n$ is again necessary and sufficient for $X_{i:n} \geq_L X_{j:m}$.

It is interesting to look at some special cases of Theorem 1: For minima,

$$X_{1:n} \geq_L X_{1:n+1}$$

for maxima,

$$X_{n:n} \geq_L X_{n+1:n+1}$$

and for medians

$$X_{n+1:2n+1} \geq_L X_{n+2:2n+3},$$

hence all three statistics are decreasing in the Lorenz sense with increasing sample size. This may be contrasted with previously considered parent distributions: for the Pareto distribution with c.d.f. $F(x) = 1 - (x/x_0)^{-\alpha}$, $x \geq x_0$, $x_0, \alpha > 0$, it is known that maxima are increasing while minima are decreasing, whereas for the power function distribution with c.d.f. $F(x) = (x/x_0)^\gamma$, $0 \leq x \leq x_0$, $x_0, \gamma > 0$, the situation is reversed (Arnold and Villaseñor, 1998).

The results used in the proof of Theorem 1 also apply to the one-sample case, for which the following negative result is then immediate:

Corollary 2 *Assume $X \sim LL(a, b)$ with $a > 1$. Then $X_{i:n}$ and $X_{j:n}$ are never comparable in the Lorenz sense.*

This is in marked contrast to previous results for other families such as the uniform, Pareto, power function and exponential distributions. For these models, there always exists some pair

$X_{i:n}, X_{j:n}$ of order statistics that can be Lorenz ordered. The negative result in the present case may be explained by noting that the log-logistic density is regularly varying at both infinity and at the origin, a property that is inherited by the densities of its order statistics. Specifically, for $X_{i:n}, X_{j:n}$ with $i \leq j$ the lower tail of the distribution of $X_{i:n}$ is heavier than the lower tail of the distribution of $X_{j:n}$, whereas for the upper tail the situation is reversed. It is known that such constellations preclude Lorenz ordering, see Kleiber (2000) for further details.

3 Extensions

The results from the previous section may be generalized in several directions. First, for two samples from different log-logistic distributions the following result is available:

Corollary 3 *Assume $X \sim LL(a_1, b_1), Y \sim LL(a_2, b_2)$, with $a_i > 1, i = 1, 2$.*

- (a) *If a sample of size n from $LL(a_1, b_1)$ and a second sample of size m from $LL(a_2, b_2)$ are available, then $a_1 \leq a_2, a_1 i \leq a_2 j$ and $a_1(n - i + 1) \leq a_2(m - j + 1)$ imply $X_{i:n} \geq_L Y_{j:m}$.*
- (b) *$X_{i:n} \geq_L Y_{j:m}$ implies $a_1 i \leq a_2 j$ and $a_1(n - i + 1) \leq a_2(m - j + 1)$.*

Proof: Part (a) follows directly from (3) whereas part (b) is a consequence of (4). ■

In particular, for $a_1 \leq a_2$ the maxima are ordered, specifically

$$X_{n:n} \geq_L Y_{n+1:n+1},$$

and for the minima the analogous relation

$$X_{1:n} \geq_L Y_{1:n+1}$$

is obtained. However, a complete characterization of the Lorenz order for order statistics from different log-logistic distributions is currently unavailable because condition (3) is only sufficient but not necessary.

A second extension is to more general distributions. The log-logistic distribution is a special case of both the Burr III and the Burr XII distributions which in turn are special cases of the GB2

distribution. The Burr III distribution possesses the c.d.f.

$$F(x) = [1 + (x/b)^{-a}]^{-p}, \quad x > 0, \quad (5)$$

where $a, b, p > 0$, whereas the Burr XII distribution is given by

$$F(x) = 1 - [1 + (x/b)^a]^{-q}, \quad x > 0, \quad (6)$$

where $a, b, q > 0$. The log-logistic distribution is obtained for $p = 1$ and $q = 1$, respectively, moreover $\text{Burr III}(a, b, p) \equiv \text{GB2}(a, b, p, 1)$ and $\text{Burr XII}(a, b, q) \equiv \text{GB2}(a, b, 1, q)$. Note that

$$X \sim \text{Burr III}(a, b, p) \iff \frac{1}{X} \sim \text{Burr XII}(a, 1/b, p). \quad (7)$$

It is not difficult to see that for $X \sim \text{Burr III}(a, b, p)$

$$X_{n:n} \sim \text{Burr III}(a, b, np)$$

and for $X \sim \text{Burr XII}(a, b, q)$

$$X_{1:n} \sim \text{Burr XII}(a, b, nq).$$

The following corollary is therefore another direct consequence of Theorem 1:

Corollary 4 (a) For $X \sim \text{Burr III}(a_1, b_1, p_1)$, $Y \sim \text{Burr III}(a_2, b_2, p_2)$, $a_i > 1$, $i = 1, 2$,

$$X_{n:n} \geq_L Y_{m:m} \iff a_1 \leq a_2 \text{ and } a_1 p_1 n \leq a_2 p_2 m.$$

In particular, for two samples from the same distribution

$$X_{n:n} \geq_L X_{n+1:n+1}.$$

(b) For $X \sim \text{Burr XII}(a_1, b_1, q_1)$, $Y \sim \text{Burr XII}(a_2, b_2, q_2)$, $a_i q_i > 1$, $i = 1, 2$,

$$X_{1:n} \geq_L Y_{1:m} \iff a_1 \leq a_2 \text{ and } a_1 q_1 n \leq a_2 q_2 m.$$

In particular, for two samples from the same distribution

$$X_{1:n} \geq_L X_{1:n+1}.$$

Smallest and largest order statistics are of special interest in reliability applications because they describe the lifetime of series and parallel systems, respectively. Unfortunately, apart from the maxima (minima) no other order statistic from a Burr III (Burr XII) distribution follows a GB2 distribution.

The preceding Corollary does however cover several special cases of the Burr III and Burr XII distributions: For example, the Lomax or Pareto (II) distribution is a Burr XII $(1, b, q)$ distribution. Also, a Burr XII distribution with $a = q$ is known as a paralogistic distribution in actuarial science and a Burr III distribution with $a = p$ is therefore also referred to as an inverse paralogistic distribution (Klugman, Panjer and Willmot, 1998). Consequently, Lorenz ordering results for maxima or minima from such parent distributions follow directly from the more general results of Corollary 4. For brevity, details are omitted.

In view of (7) it would seem that Corollary 4 also yields Lorenz ordering results for the inverse (or inverted) Lomax distribution, a Burr III $(1, b, p)$ distribution. However, for a Burr III distribution the condition for the existence of the first moment is $a > 1$, so that the Lorenz curve does not even exist for inverse Lomax maxima.

Acknowledgements

Research support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 475, is acknowledged gratefully.

References

- Ali, M.M., and Khan, A.H. (1987). On order statistics from the log-logistic distribution, *Journal of Statistical Planning and Inference*, 17, 103-108.
- Al-Shboul, Q.M., and Khan, A.H. (1989): Moments of order statistics from doubly truncated log-logistic distribution. *Pakistan Journal of Statistics*, B 5, 299-307.
- Arnold, B.C. (1983). *Pareto Distributions*, Fairland, MD: International Co-operative Publishing House.
- Arnold, B.C. (1987). *Majorization and the Lorenz Order*, Lecture Notes in Statistics 43, Berlin - New York: Springer.
- Arnold, B.C. and Nagaraja, H.N. (1991): Lorenz ordering of exponential order statistics, *Statistics and Probability Letters* 11, 485-490.
- Arnold, B.C. and Villaseñor, J.A. (1998): Lorenz ordering of order statistics and record values, in: N. Balakrishnan and C.R. Rao (eds.), *Handbook of Statistics, Vol. 16*, Amsterdam: Elsevier, 75-87.
- Balakrishnan, N., and Malik, H.J. (1987): Moments of order statistics from truncated log-logistic distributions, *Journal of Statistical Planning and Inference*, 17, 251-267.
- Balakrishnan, N., and Malik, H.J. (1988): Recurrence relations and identities for moments of order statistics, II: Specific continuous distributions, *Communications in Statistics - Theory and Methods*, 17, 2657-2694.
- Kleiber, C. (1999). On the Lorenz order within parametric families of income distributions. *Sankhyā*, B 61, 514-517.
- Kleiber, C. (2000). *Halbordnungen von Einkommensverteilungen*. Angewandte Statistik und Ökonometrie Vol. 47, Göttingen: Vandenhoeck & Ruprecht.
- Klugman, S.A., Panjer, H.H., and Willmot, G.E. (1998). *Loss Models*. New York: John Wiley.

- Kochar, S.C. (1989): On extensions of DRML and related partial orderings, *Communications in Statistics - Stochastic Models* 5, 235-245.
- McDonald, J.B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52, 647-663.
- Ragab, A., and Green, J. (1984). On order statistics from the log-logistic distribution and their properties, *Communications in Statistics - Theory and Methods*, 13, 2713-2724.
- Venter, G. (1983). Transformed beta and gamma distributions and aggregate losses. *Proceedings of the Casualty Actuarial Society*, 70, 156-193.
- Wilfling, B. (1996a). Lorenz ordering of generalized beta-II income distributions. *Journal of Econometrics*, 71, 381-388.
- Wilfling, B. (1996b): Lorenz ordering of power-function order statistics, *Statistics and Probability Letters* 30, 313-319.