Probability in the Engineering and Informational Sciences, **27**, 2013, 451–454.
doi:10.1017/S026996481300017X

COMMENTS ON "ORDERING PROPERTIES OF ORDER STATISTICS FROM HETEROGENEOUS POPULATIONS: A REVIEW WITH AN EMPHASIS ON SOME RECENT DEVELOPMENTS"

MAOCHAO XU
Department of Mathematics
Illinois State University, Normal, IL, USA
E-mail: mxu2@ilstu.edu

Professors Balakrishnan and Zhao have written an excellent survey on the recent developments of stochastic comparisons of order statistics, which cover almost every aspect of ordering properties of order statistics from both continuous and discrete heterogeneous populations. My discussion will be limited to the skewness of order statistics and order statistics from heterogeneous populations with different shape parameters.

1. SKEWNESS OF ORDER STATISTICS

It is known in the literature that two common and popular measures to compare skewness of distributions are *convex transform order* ($\leq_{\mathrm{c}}$) and *star order* ($\leq_\star$); see, for example, Marshall and Olkin [8]. Assume that $X_1, \ldots, X_n$ are independent samples, and $Y_1, \ldots, Y_n$ are the other independent samples. If X_i's have the same distribution as X, and Y_i's have the same distribution as Y, then (cf. Thm 5.7, Barlow and Proschan [2]),

$$X \leq_{\mathrm{c}(\star)} Y \Longrightarrow X_{k:n} \leq_{\mathrm{c}(\star)} Y_{k:n},$$

which means that more skewed the population is, the more skewed are order statistics from this population. Then, a natural question arises as to how the heterogeneity affects the skewness of order statistics? Intuitively, order statistics from more heterogeneous populations are more skewed. However, this is not true in general as shown in Example 2.4 of *Balakrishnan and Zhao*! Such results are only true under certain conditions. For example, it is shown in Kochar and Xu [5] that the largest order statistics from heterogeneous exponential samples are more skewed than the one from homogeneous exponential samples in the sense of convex transform order without any restriction on the parameters. As a consequence, it holds that

$$cv(X_{n:n}) \geq \left(\sqrt{\sum_{i=1}^{n} \frac{1}{i^2}}\right) \Bigg/ \left(\sum_{i=1}^{n} \frac{1}{i}\right), \tag{1}$$

 0269-9648/13 $25.00

where $X_{n:n}$ is the largest order statistic from any heterogeneous exponential sample with size n, and cv means the coefficient of variation. In the literature, the discussions are only limited to heterogeneous exponential random variables. In fact, it has been shown in Kochar and Xu [7] that for some distributions in the PHR family, order statistics from those heterogeneous populations are more skewed than that from homogeneous populations. For example, the largest order statistics from heterogeneous Weibull distributions with different scale parameters are more skewed than that from homogeneous Weibull distributions.

One may wonder why should we care about the skewness of order statistics? The research in this direction, in fact, is of great interest from both theoretical and practical point of view. I briefly discuss the significance of this research in the following. More details may be found in Kochar and Xu [7].

(a) *Detecting the heterogeneity in the population.*
The research on the skewness of order statistics is useful in detecting the heterogeneity of population with limited information. For example, suppose a "black box"parallel system is composted of independent exponential components (which is a common assumption in engineering). The available observations are only the lifetimes of the "black box". Then an interesting question is whether the types of composing components are the same based on the available data? One "dirty"but quick way to check this answer is to look at the sample coefficient of variation. If it is far larger from the value in Eq. (1), we then reject the homogeneity assumption. The formal test statistic can also be developed based on Eq. (1).

(b) *Unifying and simplifying the study on stochastic comparisons.*
Under the restriction of some skewness order, stochastic comparisons of order statistics may be equivalent for different stochastic orders. For example, under the star ordering, stochastic comparisons of order statistics based on stochastic order, hazard rate order or dispersive order may be equivalent; excess wealth order may be just equivalent to expect value order, (cf. Kochar and Xu [5]). This is not surprising since magnitude and dispersion of order statistic are determined by the shape of its density function, which is closely related to the skewness. Moreover, under some skewness order restriction, the proof of stochastic comparison may be greatly simplified. For example, Kochar and Xu [6] proved that for two non-negative random variables X and Y with distribution functions F and G, respectively, if $X \leq_\star Y$, then

$$X \leq_{\mathrm{st}} Y \Longleftrightarrow \lim_{x \to 0^+} F(x)/G(x) \geq 1.$$

This result reveals that under the star order restriction, stochastic order between two random variables is determined by the magnitudes of distribution functions near origins. This fact can be used to simplify the proof.

There are also applications of skewness of order statistics in the statistical inference. For example, we may improve the estimates of distribution functions of order statistics under the star order restriction. One may refer to Barlow et al. [1] and Kochar and Xu [7] for more details.

2. ORDER STATISTICS FROM POPULATIONS WITH DIFFERENT SHAPES

In the literature, most discussion on ordering properties of order statistics assumes that the underlying populations have different scale parameters as seen from this review paper. However, order statistics from heterogeneous populations with different shape parameters may be also of interest. In the following, I will discuss order statistics from distributions with regular varying tails to make this point clear.

A distribution F is said to have regularly varying tail at ∞ with tail index $-\alpha$, denoted by $\mathrm{RV}(-\alpha)$, if its survival function is of the following form,

$$\bar{F}(t) = t^{-\alpha} L(t), \quad t \in \mathbb{R}_+,$$

where L is a slowly varying function; that is, L is a positive function on $\mathbb{R}_+$ with property

$$\lim_{t\to\infty} \frac{L(ct)}{L(t)} = 1, \quad c \in \mathbb{R}_+.$$

Equivalently, one has

$$\lim_{t\to\infty} \frac{\bar{F}(ct)}{\bar{F}(t)} = c^{-\alpha},\ c \in \mathbb{R}_+.$$

Similarly, a distribution F is said to have regularly varying tail at 0 with index α, denoted by $\mathrm{RV}_0(\alpha)$, if

$$\lim_{t\to 0} \frac{F(ct)}{F(t)} = c^{\alpha},\ c \in \mathbb{R}_+.$$

Distributions with regularly varying tails appear naturally in many fields of statistics and applied probability, including the extreme value theory, reliability theory, queuing theory, insurance, actuarial science etc. One may refer to Bingham, Goldie and Teugels [3] for an encyclopedic treatment of regular variation. Many distribution families have regular variation tail properties at ∞ or 0. For example, Burr XII, Pareto and stable survival functions have regularly varying tails at ∞; Uniform, Gamma, Weibull and Burr XII distributions have regular variation properties at 0. One may refer to Huang, Li and Xu [4] for more distributions.

The following result shown in Huang et al. [4] provides necessary conditions to compare order statistics from distributions with regular varying tail properties.

Proposition 1: *Assume X_i's are independent random variables with distribution functions F_i and survival functions $\bar{F}_i$, $i = 1, \ldots, n$, respectively, and Y_j's are other independent random variables with distribution functions G_j, and survival functions $\bar{G}_j$, $j = 1, \ldots, m$, respectively.*

(a) *If $\bar{F}_i \in \mathrm{RV}(-\alpha_i)$, $\bar{G}_j \in \mathrm{RV}(-\alpha_j^*)$, and $\alpha_i > 0, \alpha_j^* > 0$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, then*

$$X_{r:n} \le_{icx} Y_{s:m} \Longrightarrow \sum_{i=1}^{n-r+1} \alpha_{(i)} \ge \sum_{i=1}^{m-s+1} \alpha^*_{(i)}.$$

(b) *If $F_i \in \mathrm{RV}_0(\alpha_i)$, $G_j \in \mathrm{RV}_0(\alpha_j^*)$ and $\alpha_i > 0, \alpha_j^* > 0$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, then*

$$X_{r:n} \le_{icv} Y_{s:m} \Longrightarrow \sum_{i=1}^{r} \alpha_{(i)} \le \sum_{i=1}^{s} \alpha^*_{(i)}.$$

It is seen that order statistics from distributions with regular variation tails have close relations with tail indexes α's, which are shape parameters for most distributions (cf. Huang et al. [4]). However, in the literature, order statistics from populations with different shape parameters have not received much attention yet. More research in this direction is needed.

References

1. Barlow, R.E., Bartholomew, D.J., Bremner, J.M. & Brunk, H.D. (1972). *Statistical inference under order restrictions*. New York: John Wiley.
2. Barlow, R.E. & Proschan, F. (1981). *Statistical theory of reliability and life testing*. MD: Silver Spring, To Begin With.
3. Bingham, N.H., Goldie, C.M. & Teugels, J. (1987). *Regular variation*. Cambridge: Cambridge University Press.
4. Huang, Y., Li, Y. & Xu, M. (2013). Analysis of order statistics from distributions with regularly varying tails. *Communications in Statistics-Theory and Methods*, to appear.
5. Kochar, S. & Xu, M. (2009). Comparisons of parallel systems according to the convex transform order. *Journal of Applied Probability*, 46: 342–352.
6. Kochar, S. & Xu, M. (2012). Some unified results on comparing linear combinations of independent gamma random variables. *Probability in Engineering and Information Sciences*, 49: 393–404.
7. Kochar, S. & Xu, M. (2013). On the skewness of order statistics with applications. *Annals of Operations Research*, DOI: 10.1007/s10479-012-1212-4.
8. Marshall, A.W. & Olkin, I. (2007). *Life distributions*. New York: Springer.