STOCHASTIC PROPERTIES OF SPACINGS IN A SINGLE-OUTLIER EXPONENTIAL MODEL

BAHA-ELDIN KHALEDI

Department of Mathematics
University of Arizona
Tucson, Arizona 85721
E-mail: bkhaledi@math.arizona.edu

SUBHASH KOCHAR

Indian Statistical Institute
New Delhi, 110016, India
E-mail: kochar@.isid.ac.in

Let $X_1, \ldots, X_n$ be independent exponential random variables with possibly different scale parameters. Kochar and Korwar [*J. Multivar. Anal.* 57 (1996)] conjectured that, in this case, the successive normalized spacings are increasing according to hazard rate ordering. In this article, we prove this conjecture in the case of a single-outlier exponential model when all except one of the parameters are identical. We also prove that the spacings are more dispersed and larger in the sense of hazard rate ordering when the vector of scale parameters is more dispersed in the sense of majorization.

1. INTRODUCTION

Many authors have studied the stochastic properties of order statistics and spacings from restricted families of distributions. Barlow and Proschan [3] proved that in the case of a random sample from a decreasing failure rate distribution, the successive normalized spacings are stochastically larger. Kochar and Kirmani [10] strengthened this result from stochastic ordering to hazard rate ordering. The corresponding problem, when the random variables are not identically distributed, has also been studied by many researchers, including Pledger and Proschan [15], Shaked and Tong

[17], Kochar and Korwar [11], Kochar and Rojo [12], and Nappo and Spizzichino [14] among others. For a review of this topic, see Kochar [9].

Before we go into their details, let us quickly review some important notions of stochastic orderings of various kinds. We shall denote the density function, the survival function, and the hazard rate function of a random variable X by f_X, $\bar{F}_X$, and r_{F_X}, respectively. A random variable X is said to be *stochastically* larger than another random variable Y (denoted by $X \geq_{st} Y$) if $\bar{F}_X(x) \geq \bar{F}_Y(x)$ for all x. A stronger notion of stochastic dominance is that of *hazard rate* ordering. X is said to be larger than Y in *hazard rate* ordering (denoted by $X \geq_{hr} Y$) if $\bar{F}_X(x)/\bar{F}_Y(x)$ is nondecreasing in x. Finally, X is said to be larger than Y in *likelihood ratio* ordering (denoted by $X \geq_{lr} Y$) if $f_X(x)/f_Y(x)$ is nondecreasing in x. For more details on stochastic orderings, see Shaked and Shanthikumar [16, Chap. 1]. In case X and Y have a common left end point of their supports, we have the following chain of implications among the above stochastic orders: $X \geq_{lr} Y \Rightarrow X \geq_{hr} Y \Rightarrow X \geq_{st} Y$.

A random variable X is said to be more dispersed than another random variable Y (denoted by $X \geq_{disp} Y$) if $F_X^{-1}(\beta) - F_X^{-1}(\alpha) \geq F_Y^{-1}(\beta) - F_Y^{-1}(\alpha)$ whenever $0 < \alpha \leq \beta < 1$, where F_X^{-1} and F_Y^{-1} are the right continuous inverses of the distribution functions F_X and F_Y of X and Y, respectively. One of the consequences of the dispersive ordering is that, in this case, the variances of the corresponding random variables are ordered. For other properties of dispersive ordering see Shaked and Shanthikumar [16, Sect. 2.B].

At this point, let us also give the definitions of majorization and Schur convexity. Let $\{x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}\}$ denote the increasing arrangement of the components of the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. The vector $\mathbf{y}$ is said to majorize the vector $\mathbf{x}$ (written $\mathbf{x} \overset{m}{\preceq} \mathbf{y}$) if $\sum_{i=1}^{j} y_{(i)} \leq \sum_{i=1}^{j} x_{(i)}$ for $j = 1, \ldots, n - 1$ and $\sum_{i=1}^{n} y_{(i)} = \sum_{i=1}^{n} x_{(i)}$. Functions that preserve the majorization ordering are said to be Schur-convex. See Marshall and Olkin [13, Chap. 3] for properties of such functions.

Let $X_1, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i for $i \in \{1, \ldots, n\}$. Let us denote the ith-order statistic and the ith spacing by $X_{i:n}$ and $D_{i:n} = X_{i:n} - X_{i-1:n}$, respectively. Here, $X_0 \equiv 0$. Let $D_{i:n}^* = (n - i + 1)D_{i:n}$ denote the ith normalized spacing. Kochar and Korwar [11] conjectured that for $i = 1, \ldots, n - 1$, $D_{i+1:n}^* \geq_{hr} D_{i:n}^*$. In Section 2, we prove their conjecture when $\lambda_1 = \cdots = \lambda_{n-1} = \lambda$ and $\lambda_n = \lambda^*$. Such a model is known as a single-outlier exponential model with parameters (λ, λ^*) and it has been studied by many researchers, including Kale and Sinha [7], Joshi [6], Barnett and Lewis [4, p. 193], Gross et al. [5], and Khaledi and Kochar [8], among others. Balakrishnan [2] obtained some recurrence relations to compute the single and product moments of order statistics when observations follow the single- as well as multiple-outlier model. Using these formulas, he studied the properties of the various estimators in outlier models. We also prove in Section 2 that if $\lambda_1^* < \lambda_2^* < \lambda_2 < \lambda_1$ and $\lambda_1^* + (n - 1)\lambda_1 = \lambda_2^* + (n - 1)\lambda_2$ (in which case $(\lambda_1, \ldots, \lambda_1, \lambda_1^*) \overset{m}{\succeq} (\lambda_2, \ldots, \lambda_2, \lambda_2^*)$), for $i \in \{1, \ldots, n\}$, $D_{i:n}^{(1)}$ is greater than $D_{i:n}^{(2)}$ according to hazard rate as well as dispersive ordering, where $D_{i:n}^{(1)}$ and $D_{i:n}^{(2)}$ respectively stand for ith spacing of single-outlier models with parameters (λ_1, λ_1^*) and (λ_2, λ_2^*).

2. HAZARD RATE ORDERING AMONG SPACINGS

Let $X_1, \ldots, X_n$ follow the single-outlier model with parameters (λ, λ^*). It is easy to see that the joint density function of $(D_{1:n}, \ldots, D_{n:n})$ is

$$f_{D_{1:n}, \ldots, D_{n:n}}(x_1, \ldots, x_n) = \sum_{\theta=1}^{n} h(\theta) \prod_{i=1}^{\theta} \alpha_i^* e^{-\alpha_i^* x_i} \prod_{i=\theta+1}^{n} \alpha_i e^{-\alpha_i x_i},$$

where $\alpha_i = (n - i + 1)\lambda$, $\alpha_i^* = (n - i)\lambda + \lambda^*$, $i = 1, \ldots, n$, and

$$h(\theta) = \frac{(n - 1)! \lambda^{n-1} \lambda^*}{\displaystyle\prod_{i=1}^{\theta} \alpha_i^* \prod_{i=\theta+1}^{n} \alpha_i}, \qquad \theta = 1, \ldots, n. \tag{2.1}$$

One can see that $\{h(\theta), \theta = 1, \ldots, n\}$ is a probability mass function of a discrete random variable Θ. For $i = 1, \ldots, n$, the marginal density function of $D_{i:n}$ can be expressed as

$$f_{D_{i:n}}(x) = H_i \alpha_i e^{-\alpha_i x} + \bar{H}_i \alpha_i^* e^{-\alpha_i^* x}, \tag{2.2}$$

where

$$H_i = \sum_{\theta=1}^{i-1} h(\theta), \qquad i = 2, \ldots, n, \quad \text{and} \quad H_i = 0. \tag{2.3}$$

Thus, the density function of $D_{i:n}$ is a mixture distribution of two exponential random variables with parameters α_i and α_i^*. In the following theorem, we prove that $D_{i+1:n}^* \geq_{hr} D_{i:n}^*$, thus proving the conjecture of Kochar and Korwar [10] in the case of the single-outlier exponential model.

THEOREM 2.1: *Let $X_1, \ldots, X_n$ follow the single-outlier exponential model with parameters (λ, λ^*). Then*

$$D_{i+1:n}^* \geq_{hr} D_{i:n}^*, \qquad i = 1, \ldots, n - 1,$$

where $D_{i:n}^* = (n - i + 1) D_{i:n}$ denotes the ith normalized spacing.

PROOF: We prove the result for $\lambda^* > \lambda$. The proof for the case $\lambda^* < \lambda$ follows using the same kind of arguments. From (2.2), we find that the survival function of $D_{i:n}^*$ is $\bar{F}_{D_{i:n}^*}(x) = H_i e^{-\lambda x} + \bar{H}_i e^{-\eta_i x}$, where $\eta_i = ((n - i)\lambda + \lambda^*)/(n - i + 1)$. To prove the theorem, we have to show that for any $i \in \{1, \ldots, n - 1\}$,

$$g(x) = \frac{\bar{F}_{D_{i+1:n}^*}(x)}{\bar{F}_{D_{i:n}^*}(x)}$$

is nondecreasing in x. The numerator of $g'(x)$, the derivative of $g(x)$, is

$$A(x) = [H_i e^{-\lambda x} + \bar{H}_i e^{-\eta_i x}][-\lambda H_{i+1} e^{-\lambda x} - \eta_{i+1} \bar{H}_{i+1} e^{-\eta_{i+1} x}]$$
$$+ [H_{i+1} e^{-\lambda x} + \bar{H}_{i+1} e^{-\eta_{i+1} x}][\lambda H_i e^{-\lambda x} + \eta_i \bar{H}_i e^{-\eta_i x}]$$

$$= (\lambda^* - \lambda) \left\{ \frac{\bar{H}_i H_{i+1}}{n - i + 1} e^{-(\eta_i + \lambda)x} - \frac{\bar{H}_{i+1} H_i}{n - i} e^{-(\eta_{i+1} + \lambda)x} \right.$$
$$\left. - \frac{\bar{H}_i \bar{H}_{i+1}}{(n - i + 1)(n - i)} e^{-(\eta_i + \eta_{i+1})x} \right\}$$

$$\geq (\lambda^* - \lambda) \left\{ \left(\frac{\bar{H}_i H_{i+1}}{n - i + 1} - \frac{\bar{H}_{i+1} H_i}{n - i} \right) e^{-(\eta_{i+1} + \lambda)x} \right.$$
$$\left. - \frac{\bar{H}_i \bar{H}_{i+1}}{(n - i + 1)(n - i)} e^{-(\eta_i + \eta_{i+1})x} \right\} \qquad (2.4)$$

$$= \frac{(\lambda^* - \lambda)}{(n - i)(n - i + 1)} \left\{ \{(n - i)\bar{H}_i - (n - i + 1)\bar{H}_{i+1} + \bar{H}_i \bar{H}_{i+1} \} \right.$$
$$\left. \times e^{-(\eta_{i+1} + \lambda)x} - \bar{H}_i \bar{H}_{i+1} e^{-(\eta_i + \eta_{i+1})x} \right\}. \qquad (2.5)$$

The inequality in (2.4) follows, since $\lambda^* > \lambda$ implies $\eta_{i+1} > \eta_i$.

Again, $\lambda^* > \lambda$ implies $\lambda < \eta_i$, which, in turn, implies $e^{-(\lambda + \eta_{i+1})x} > e^{-(\eta_i + \eta_{i+1})x}$ for every $x \geq 0$. Also, for $\lambda^* > \lambda$,

$$\{(n - i)\bar{H}_i - (n - i + 1)\bar{H}_{i+1}\} = (n - i)h(i) - \bar{H}_{i+1}$$
$$\geq 0, \qquad (2.6)$$

since for $\lambda^* > \lambda$, $h(j)$ is a decreasing function of j. Under these results in (2.5), we find that $A(x)$ and hence $g'(x)$ is nonnegative for $x \geq 0$. This proves the required result. ∎

Remark: Kochar and Kirmani [10] proved a similar result when $X_1, \ldots, X_n$ was a random sample from a DFR distribution.

Let $X_1, \ldots, X_n$ be independent exponential random variables with unequal parameters. Pledger and Proschan [15] proved that for $i \in \{1, \ldots, n\}$, $D_{i:n}$ is stochastically larger when the parameters are unequal than when they are all equal. This prompted them to examine the question of whether the survival function of $D_{i:n}$ is Schur-convex in $(\lambda_1, \ldots, \lambda_n)$. They came up with a counterexample to show that this is not true in general. Kochar and Korwar [11] proved that in the special case of second spacing, whereas the survival function of $D_{2:n}$ is Schur-convex in $(\lambda_1, \ldots, \lambda_n)$, its hazard rate is not Schur-concave. They proved, however, that the hazard rate of $D_{2:2}$ is Schur-concave. We now examine this question for the single-outlier model with parameters (λ, λ^*). In the rest of this section, we assume that $\lambda^* < \lambda$. We will treat it as a part of the model. We prove later in this section that if $\lambda_1^* < \lambda_2^* < \lambda_2 < \lambda_1$ and $\lambda_1^* + (n - 1)\lambda_1 = \lambda_2^* + (n - 1)\lambda_2$ (in which case

$(\lambda_1,\ldots,\lambda_1,\lambda_1^*) \overset{m}{\geqslant} (\lambda_2,\ldots,\lambda_2,\lambda_2^*))$, $D_{i:n}^{(1)}$ is greater than $D_{i:n}^{(2)}$ according to the hazard rate as well as the dispersive ordering for $i \in \{1,\ldots,n\}$. To prove it, we need the following lemmas.

LEMMA 2.1: *Let $X_1,\ldots,X_n$ follow the single-outlier exponential model with parameters (λ,λ^*). Then,*

$$\lambda^* < \lambda \Leftrightarrow H_i \leq \frac{i-1}{n} \quad \text{for } i = 1,\ldots,n, \tag{2.7}$$

where H_i is given by (2.3). The inequality in (2.7) is reversed for $\lambda^* > \lambda$.

PROOF: $\lambda^* < \lambda$ implies that the function $h(j)$ in (2.1) is increasing in $j, j = 1,\ldots,n$. Note that

$$(h(1),h(2),\ldots,h(n)) \overset{m}{\geqslant} (1/n,\ldots,1/n).$$

The required result follows from the definition of majorization. ∎

LEMMA 2.2: *Let $X_1,\ldots,X_n$ follow the single-outlier exponential model with parameters (λ_1,λ_1^*). Let $Y_1,\ldots,Y_n$ be another set of random variables following the single-outlier exponential model with parameters (λ_2,λ_2^*):*

(i) *If $\lambda_1^* < \lambda_2^* < \lambda_2 < \lambda_1$, then $\Theta_1 \geq_{lr} \Theta_2$.*
(ii) *If $\lambda_1 < \lambda_2 < \lambda_2^* < \lambda_1^*$, then $\Theta_1 \leq_{lr} \Theta_2$.*

Θ_1 and Θ_2 correspond to random variable Θ with probability mass function $h(j)$ in (2.1) for X_i's and Y_i's, respectively.

PROOF: (i) We prove that for $\theta = 1,\ldots,n-1$,

$$\frac{h_2(\theta+1)}{h_1(\theta+1)} \leq \frac{h_2(\theta)}{h_1(\theta)},$$

where h_1 and h_2 correspond to h in (2.1) for X_i's and Y_i's, respectively. This inequality holds for $\theta = 1,\ldots,n-1$ if and only if

$$\frac{(n-\theta-1)\lambda_1 + \lambda_1^*}{(n-\theta-1)\lambda_2 + \lambda_2^*} \leq \frac{\lambda_1}{\lambda_2}. \tag{2.8}$$

Since $\lambda_1^* < \lambda_2^*$ and $\lambda_2 < \lambda_1$, it is easy to see that (2.8) is true.

(ii) In this case, the inequality in (2.8) is reversed, which, in turn, implies that $\Theta_1 \leq_{lr} \Theta_2$. This proves the result. ∎

THEOREM 2.2: *Let $X_1,\ldots,X_n$ follow the single-outlier exponential model with parameter (λ_1,λ_1^*) and let $Y_1,\ldots,Y_n$ be another set of random variables following the single-outlier exponential model with parameters (λ_2,λ_2^*). If*

$$\lambda_1^* < \lambda_2^* < \lambda_2 < \lambda_1 \quad \text{and} \quad \lambda_1^* + (n-1)\lambda_1 = \lambda_2^* + (n-1)\lambda_2, \tag{2.9}$$

then

$$D_{i:n}^{(1)} \geq_{\mathrm{hr}} D_{i:n}^{(2)}.$$

Remark: Under (2.9), $(\lambda_1^*, \lambda_1, \ldots, \lambda_1) \overset{m}{\geq} (\lambda_2^*, \lambda_2, \ldots, \lambda_2)$.

PROOF: Without loss of generality, let us assume that $\lambda_1^* + (n-1)\lambda_1 = 1$.
From (2.2), the survival functions of $D_{i:n}^{(1)}$ and $D_{i:n}^{(2)}$ are

$$\bar{F}_{D_{i:n}^{(1)}}(x) = P_i e^{-\alpha_{i1}x} + \bar{P}_i e^{-\alpha_{i1}^* x},$$

$$\bar{F}_{D_{i:n}^{(2)}}(x) = Q_i e^{-\alpha_{i2}x} + \bar{Q}_i e^{-\alpha_{i2}^* x},$$

where P_i and Q_i correspond to H_i in (2.2) for $D_{i:n}^{(1)}$ and $D_{i:n}^{(2)}$, respectively, and $\alpha_{i1} = (n-i+1)\lambda_1$, $\alpha_{i1}^* = (n-i)\lambda_1 + \lambda_1^*$, $\alpha_{i2} = (n-i+1)\lambda_2$, and $\alpha_{i2}^* = (n-i)\lambda_2 + \lambda_2^*$.
We have to show that

$$\phi(x) = \frac{\bar{F}_{D_{i:n}^{(1)}}(x)}{\bar{F}_{D_{i:n}^{(2)}}(x)}$$

is nondecreasing in x. After some simplifications, the numerator of $\phi'(x)$, the derivative of $\phi(x)$, is

$$g(x) = -(\alpha_{i1} - \alpha_{i2})P_i Q_i e^{-(\alpha_{i1}+\alpha_{i2})x} + (\alpha_{i2}^* - \alpha_{i1}^*)\bar{P}_i \bar{Q}_i e^{-(\alpha_{i1}^*+\alpha_{i2}^*)x}$$
$$- (\alpha_{i1}^* - \alpha_{i2})Q_i \bar{P}_i e^{-(\alpha_{i2}+\alpha_{i1}^*)x} + (\alpha_{i2}^* - \alpha_{i1})\bar{Q}_i P_i e^{-(\alpha_{i1}+\alpha_{i2}^*)x}. \quad (2.10)$$

Using the assumption $\lambda_1^* < \lambda_2^* < \lambda_2 < \lambda_1$ and the fact that $\lambda_i^* + (n-1)\lambda_i = 1$, $i = 1,2$, it follows that $\alpha_{i1} + \alpha_{i2}^* < \alpha_{i1} + \alpha_{i2}, \alpha_{i1} + \alpha_{i2}^* > \alpha_{i1}^* + \alpha_{i2}^*, \alpha_{i1} + \alpha_{i2}^* > \alpha_{i1}^* + \alpha_{i2}$ and all $(\alpha_{i1} - \alpha_{i2}), (\alpha_{i2}^* - \alpha_{i1}^*)$, and $(\alpha_{i2} - \alpha_{i1}^*)$ are nonnegative. Using these observations in (2.10), we see

$$g(x) \geq e^{-(\alpha_{i1}+\alpha_{i2}^*)x}\{-(\alpha_{i1} - \alpha_{i2})P_i Q_i + (\alpha_{i2}^* - \alpha_{i1}^*)\bar{P}_i \bar{Q}_i$$
$$- (\alpha_{i1} - \alpha_{i2}^*)\bar{Q}_i P_i + (\alpha_{i2} - \alpha_{i1}^*)Q_i \bar{P}_i\}$$

$$= \frac{e^{-(\alpha_{i1}+\alpha_{i2}^*)x}}{n-1}\{Q_i - P_i - (nQ_i - (i-1))\lambda_2^* + (nP_i - (i-1))\lambda_1^*\}$$

$$\geq \frac{e^{-(\alpha_{i1}+\alpha_{i2}^*)x}}{n-1}\{Q_i - P_i - n(Q_i - P_i)\lambda_2^*\}$$

$$= \frac{e^{-(\alpha_{i1}+\alpha_{i2}^*)x}}{n-1}(Q_i - P_i)(1 - n\lambda_2^*)$$

$$\geq 0. \quad (2.11)$$

The second inequality in (2.11) follows, since by Lemma 2.1, $P_i \leq (i-1)/n$ and $\lambda_1^* < \lambda_2^*$. From Lemma 2.2, it follows that $Q_i \geq P_i$, since likelihood ratio ordering

implies usual stochastic ordering. This observation, along with the fact that $\lambda_2^* \leq 1/n$, implies the last inequality in (2.10). ∎

Bagai and Kochar [1] proved that if $X \geq_{hr} Y$ and either F or G is DFR (decreasing failure rate), then $X \geq_{disp} Y$. It is known that spacings of independent exponential random variables have DFR distributions (cf. Kochar and Korwar [11]). Combining these observations, we have proved the following corollary.

COROLLARY 2.1: *Under the assumptions of Theorem 2.2,*

$$D_{i:n}^{(1)} \geq_{disp} D_{i:n}^{(2)}.$$

A consequence of Corollary 2.1 is that $\text{Var}(D_{i:n}^{(1)}) \geq \text{Var}(D_{i:n}^{(2)})$, $i = 1, \ldots, n$.

References

1. Bagai, I. & Kochar, S.C. (1986). On tail ordering and comparison of failure rates. *Communications in Statistics: Theory and Method* 15: 1377–1388.
2. Balakrishnan, N. (1994). Order statistics from non-identical exponential random variables and some applications. *Computational Statistics Data Analysis* 18: 203–253.
3. Barlow, R.E. & Proschan, F. (1966). Inequalities for linear combinations of order statistics from restricted families. *Annals of Mathematical Statistics* 37: 1574–1592.
4. Barnett, V. & Lewis, T. (1994). *Outlier in statistical data.* Chichester: Wiley.
5. Gross, A.J., Hunt, H.H., & Odeh, R.E. (1986). The correlation coefficient between the smallest and largest observations when $(n − 1)$ of the n observations are i.i.d. exponentially distributed. *Communications in Statistics: Theory and Method* 15: 1113–1123.
6. Joshi, P.C. (1972). Efficient estimation of the mean of an exponential distribution when an outlier is present. *Technometrics* 14: 137–143.
7. Kale, B.K. & Sinha, S.K. (1971). Estimation of expected life in the presence of an outlier observation. *Technometrics* 13: 755–759.
8. Khaledi, B. & Kochar, S.C. (2000). Dependence among spacings. *Probability in the Engineering and Informational Sciences* 14: 461–472.
9. Kochar, S.C. (1998). Stochastic comparisons of spacings and order statistics. In A.P. Basu, S.K. Basu, & S. Mukhopadhyay (eds.), *Frontiers in reliability.* World Scientific: Singapore, pp. 201–216.
10. Kochar, S.C. & Kirmani, S.N.U.A. (1995). Some results on normalized spacings from restricted families of distributions. *Journal of Statistical Planning and Inference* 46: 47–57.
11. Kochar, S.C. & Korwar, R. (1996). Stochastic orders for spacings of heterogeneous exponential random variables. *Journal of Multivariate Analysis* 57: 69–83.
12. Kochar, S.C. & Rojo, J. (1996). Some new results on stochastic comparisons of spacings from heterogeneous exponential distributions. *Journal of Multivariate Analysis* 59: 272–281.
13. Marshall, A.W. & Olkin, I. (1979). *Inequalities: Theory of majorization and its applications.* New York: Academic Press.

14. Nappo, G. & Spizzichino, F. (1998). Ordering properties of the TTT-plot of lifetimes with Schur joint densities. *Statistics and Probability Letters* 39: 195–203.

15. Pledger, P. & Proschan, F. (1971). Comparisons of order statistics and of spacings from heterogeneous distributions. In J.S. Rustagi (ed.), *Optimizing methods in statistics.* New York: Academic Press, pp. 89–113.

16. Shaked, M. & Shanthikumar, J.G. (1994). *Stochastic orders and their applications.* San Diego, CA: Academic Press.

17. Shaked, M. & Tong, Y.L. (1984). Stochastic ordering of spacings from dependent random variables. *Inequalities in statistics and probability*, IMS Lecture Notes, Vol. 5. Hayward, CA: Institute of Mathematical Statistics, pp. 141–149.