Stochastic orderings between distributions and their sample spacings – II

Baha-Eldin Khaledi, Subhash Kochar *

Indian Statistical Institute, Delhi Centre, 7, SJS Sansanwal Marg, New Delhi 110016, India

Received October 1998; received in revised form December 1998

Abstract

Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ denote the order statistics of a random sample of size n from a probability distribution with distribution function F. Similarly, let $Y_{1:m} \leqslant Y_{2:m} \leqslant \cdots \leqslant Y_{m:m}$ denote the order statistics of an independent random sample of size m from another distribution with distribution function G. We assume that F and G are absolutely continuous with common support $(0, \infty)$. The corresponding normalized spacings are defined by $U_{i:n} \equiv (n-i+1)(X_{i:n} - X_{i-1:n})$ and $V_{j:m} \equiv (m-j+1)(Y_{j:m} - Y_{j-1:m})$, for $i = 1, \ldots, n$ and $j = 1, \ldots, m$, where $X_{0:n} = Y_{0:n} \equiv 0$. It is proved that if X is smaller than Y in the hazard rate order sense and if either F or G is a decreasing failure rate (DFR) distribution, then $U_{i:n}$ is stochastically smaller than $V_{j:m}$ for $i \leqslant j$ and $n - i \geqslant m - j$. If instead, we assume that X is smaller than Y in the likelihood ratio order and if either F or G is DFR, then this result can be strengthened from stochastic ordering to hazard rate ordering. Finally, under a stronger assumption on the shapes of the distributions that either F or G has log-convex density, it is proved that X being smaller than Y in the likelihood ratio order implies that $U_{i:n}$ is smaller than $V_{j:m}$ in the sense of likelihood ratio ordering for $i \leqslant j$ and $n - i = m - j$.

MSC: primary 62N05; secondary 62E10

Keywords: Likelihood ratio ordering; Hazard rate ordering; Stochastic ordering; Dispersive ordering; Normalized spacings

1. Introduction

In this note we study the connections between various types of stochastic orderings between two probability distributions and their normalized spacings when random samples of possibly different sizes are drawn from them. There are several notions of stochastic orderings of varying degree of strength and they have been discussed in detail in Shaked and Shanthikumar (1994). We briefly review some of these here.

Let X and Y be two random variables with distribution functions F and G, survival functions $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$; and density functions f and g, respectively. We say that X is *stochastically smaller* than Y (denoted by $X \leqslant_{st} Y$) if $\overline{F}(x) \leqslant \overline{G}(x)$ for all x. This is equivalent to $E(\phi(X)) \leqslant E(\phi(Y))$ for all nondecreasing functions $\phi : \mathscr{R} \to \mathscr{R}$ for which the expectations exist. X is said to be smaller than Y in the sense of hazard rate ordering (denoted by $X \leqslant_{\rm hr} Y$) if $\overline{G}(x)/\overline{F}(x)$ is nondecreasing in x for all x such that $\overline{F}(x) > 0$. In case F and G are absolutely continuous, this is equivalent to $r_G(x) \leqslant r_F(x)$ for all x, where $r_F = f/\overline{F}$ and $r_G = g/\overline{G}$ are the hazard (or failure) rates of F and G, respectively. If $g(x)/f(x)$ is nondecreasing in x, then we say that X is smaller than Y in the sense of *likelihood ratio ordering* ($X \leqslant_{\rm lr} Y$). We have the following chain of implications among the above orderings,

$$X \leqslant_{\rm lr} Y \Rightarrow X \leqslant_{\rm hr} Y \Rightarrow X \leqslant_{\rm st} Y.$$

The above notions of stochastic dominance among univariate random variables can be extended to the multivariate case. A random vector $\boldsymbol{X} = (X_1, \dots, X_n)$ is smaller than another random vector $\boldsymbol{Y} = (Y_1, \dots, Y_n)$ in the multivariate stochastic order (and written as $\boldsymbol{X} \overset{\rm st}{\preccurlyeq} \boldsymbol{Y}$) if $E[\phi(\boldsymbol{X})] \leqslant E[\phi(\boldsymbol{Y})]$, for all nondecreasing functions ϕ for which the expectations exist.

We shall be assuming throughout this paper that all distributions under study are absolutely continuous with common support $(0, \infty)$. Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ denote the order statistics of a random sample $X_1, X_2, \dots, X_n$ from a life distribution with distribution function F. Similarly, let $Y_{1:m} \leqslant Y_{2:m} \leqslant \cdots \leqslant Y_{m:m}$ denote the order statistics of an independent random sample $Y_1, Y_2, \dots, Y_m$ from another life distribution with distribution function G. The corresponding normalized spacings are defined by $U_{i:n} \equiv (n-i+1)(X_{i:n} - X_{i-1:n})$, $i = 1, 2, \dots, n$ and $V_{j:m} \equiv (m-j+1)(Y_{j:m} - Y_{j-1:m})$, $j = 1, 2 \dots, m$. Here $X_{0:n} = Y_{0:m} \equiv 0$. Under certain shape restrictions and stochastic orderings between F and G, Kochar (1998) considered the problem of stochastically comparing the spacings of two samples *in case* $n = m$. In particular, it was shown that if $X \leqslant_{\rm hr} Y$ and if either X or Y has decreasing failure rate (DFR) distribution, then $\boldsymbol{U} \overset{\rm st}{\preccurlyeq} \boldsymbol{V}$. Comparatively stronger results on spacings are obtained if one assumes that F and G are likelihood ratio ordered and at least one of these distributions has log-convex density.

In this note we extend the above results to the case when the sample sizes are not necessarily equal. We prove in Section 2 that if $X \leqslant_{\rm hr} Y$ and if either X or Y has DFR distribution, then $U_{i:n}$ is stochastically smaller than $V_{j:m}$ for $i \leqslant j$ and $n - i \geqslant m - j$. If instead, we assume that X is smaller than Y in the likelihood ratio order and if either F or G is DFR, then this result can be strengthened from stochastic ordering to hazard rate ordering. Finally, if we make a stronger assumption on the shapes of the distributions that either X or Y has log-convex density and X is smaller than Y in the likelihood ratio order, then it is proved that $U_{i:n}$ is smaller than $V_{j:m}$ in the sense of likelihood ratio ordering for $i \leqslant j$ and $n - i = m - j$.

2. Main results

We shall be using the following lemma for proving the various results in this section.

Lemma 2.1. (a) $X \leqslant_{\rm st} Y \Rightarrow X_{i:n} \leqslant_{\rm st} Y_{j:m}$ *for* $i \leqslant j$ *and* $n - i \geqslant m - j$.

(b) *Let* $U_{u,i:n} \overset{\rm dist}{\equiv} \{(n-i+1)(X_{i:n} - X_{i-1:n}) | X_{i-1:n} = u\}$ *and* $V_{u,j:m} \overset{\rm dist}{\equiv} \{(m-j+1)(Y_{j:m} - Y_{j-1:m}) | Y_{j-1:m} = u\}$. *Let either* F *or* G *be DFR. Then*

$$X \leqslant_{\rm hr} Y \Rightarrow U_{u,i:n} \leqslant_{\rm hr} V_{u,j:m}$$

for $n - i \geqslant m - j$ *and* $u \geqslant 0$.

Proof. (a) It is easy to prove (see also Raqab and Amin, 1996) that for $i \leqslant j$ and $n - i \geqslant m - j$,

$$X_{i:n} \leqslant_{\rm lr} X_{j:m}. \tag{2.1}$$

It is well known that $X \leqslant_{\rm st} Y$ implies $X_{j:m} \leqslant_{\rm st} Y_{j:m}$. The required result follows from this and (2.1).

(b) The survival function of $U_{u,i:n}$ at x is

$$\overline{F}_{u,i:n}(x) = \left[\frac{\overline{F}((x/(n-i+1)) + u)}{\overline{F}(u)}\right]^{(n-i+1)},$$

and that of $V_{u,j:m}$ is

$$\overline{G}_{u,j:m}(x) = \left[\frac{\overline{G}((x/m-j+1)) + u)}{\overline{G}(u)}\right]^{(m-j+1)}.$$

We show that for $n-i \geqslant m-j$, the hazard rate of $U_{u,i:n}$ is uniformly greater than that of $V_{u,j:m}$ for all $u \geqslant 0$. Since F is DFR and $r_F(x) \geqslant r_G(x)$, we have for $n-i \geqslant m-j$ and $u \geqslant 0$,

$$\begin{aligned} r_{U_{u,i:n}}(x) &= r_F\left(\frac{x}{n-i+1} + u\right) \\ &\geqslant r_F\left(\frac{x}{m-j+1} + u\right) \\ &\geqslant r_G\left(\frac{x}{m-j+1} + u\right) \\ &= r_{V_{u,j:m}}(x) \end{aligned}$$

for all $x \geqslant 0$. That is,

$$U_{u,i:n} \leqslant_{\text{hr}} V_{u,j:m} \quad \text{for } u \geqslant 0. \qquad \square$$

Theorem 2.1. *Let $X \leqslant_{\text{hr}} Y$ and either F or G be DFR, then*

$$U_{i:n} \leqslant_{\text{st}} V_{j:m} \quad \text{for } i \leqslant j \quad \text{and} \quad n-i \geqslant m-j. \tag{2.2}$$

Proof. Let us assume that F is DFR. As shown in Kochar and Kirmani (1995), the survival function of $V_{j:m}$ (denoted by $\overline{G}^*_{j:m}(x)$) is

$$\overline{G}^*_{j:m}(x) = \int_0^\infty \left[\frac{\overline{G}((x/m-j+1)) + u)}{\overline{G}(u)}\right]^{(m-j+1)} g_{j-1:m}(u)\,\mathrm{d}u, \tag{2.3}$$

where $g_{j-1:m}$ is the density of $Y_{j-1:m}$. Since hazard rate ordering implies stochastic ordering, Lemma 2.1(b) implies that for $n-i \geqslant m-j$,

$$\overline{G}^*_{j:m}(x) \geqslant \int_0^\infty \left[\frac{\overline{F}((x/n-i+1)) + u)}{\overline{F}(u)}\right]^{(n-i+1)} g_{j-1:m}(u)\,\mathrm{d}u, \tag{2.4}$$

for all $x \geqslant 0$.

Now the function

$$\left[\frac{\overline{F}((x/n-i+1)) + u)}{\overline{F}(u)}\right]^{(n-i+1)}$$

is nondecreasing in u since F is assumed to be DFR. Also by Lemma 2.1(a), $X_{i:n} \leqslant_{\text{st}} Y_{j:m}$ for $i \leqslant j$ and $n-i \geqslant m-j$. Using these results, we get

$$\overline{G}^*_{j:m}(x) \geqslant \int_0^\infty \left[\frac{\overline{F}((x/n-i+1)) + u)}{\overline{F}(u)}\right]^{(n-i+1)} f_{i-1:n}(u)\,\mathrm{d}u \quad \text{for } x \geqslant 0. \tag{2.5}$$

The quantity on the RHS of the inequality (2.5) is the survival function of $U_{i:n}$. This proves the required result. □

In the next theorem we assume likelihood ratio ordering between X and Y instead of hazard rate ordering and establish that under the condition that either F or G is DFR, normalized spacings from the two samples are hazard rate ordered for the above choices of i and j. To prove this result we shall need the following lemma of Kochar and Kirmani (1995).

Lemma 2.2. *Let $\psi_1(x,y)$ and $\psi_2(x,y)$ be positive real-valued functions such that*

(i) *for $y_1 \leqslant y_2$,*

$$\frac{\psi_2(x,y_2)}{\psi_2(x,y_1)} \quad \text{is nondecreasing in } x,$$

(ii) *for $y_1 \leqslant y_2$,*

$$\frac{\psi_1(x,y_2)}{\psi_2(x,y_1)} \quad \text{is nondecreasing in } x,$$

(iii) *for each fixed x,*

$$\frac{\psi_1(x,y)}{\psi_2(x,y)} \quad \text{is nondecreasing in } y.$$

Then for functions ψ_1 and ψ_2 satisfying the above conditions, $Z_1 \leqslant_{\mathrm{lr}} Z_2$ implies

$$\frac{E[\psi_1(x,Z_2)]}{E[\psi_2(x,Z_1)]} \quad \text{is nondecreasing in } x, \tag{2.6}$$

provided that the expectations exist.

Theorem 2.2. *Let $X \leqslant_{\mathrm{lr}} Y$ and let either X or Y be DFR. Then*

$$U_{i:n} \leqslant_{\mathrm{hr}} V_{j:m} \quad \text{for } i \leqslant j \text{ and } n-i \geqslant m-j.$$

Proof. Assume that F is DFR. We have to show that for $i \leqslant j$ and $n-i \geqslant m-j$, $\overline{G}^*_{j:m}(x)/\overline{F}^*_{i:n}(x)$ is nondecreasing in x. As in Kochar and Kirmani (1995), this ratio can be expressed as

$$\frac{C(j:m)\int_0^\infty [\overline{G}(x/(m-j+1)+u)]^{m-j+1}\,\mathrm{d}G^{j-1}(u)}{C(i:n)\int_0^\infty [\overline{F}(x/(n-i+1)+u)]^{n-i+1}\,\mathrm{d}F^{i-1}(u)}, \tag{2.7}$$

which can be further written as

$$\frac{E[\psi_1(x/(m-j+1),Y_{j-1:j-1})]}{E[\psi_2(x/(n-i+1),X_{i-1:i-1})]}, \tag{2.8}$$

where

$$\psi_1(x,y) = C(j:m)\overline{G}(x/(m-j+1)+y)]^{m-j+1},$$

$$\psi_2(x,y) = C(i:n)\overline{F}(x/(n-i+1)+y)]^{n-i+1}$$

and

$$C(i:n) = \frac{n!}{(i-1)!\,(n-i+1)!}. \tag{2.9}$$

It is shown below that the functions ψ_1 and ψ_2 satisfy all the conditions of Lemma 2.2.

(a) The ratio

$$\frac{\psi_2(x,y_2)}{\psi_2(x,y_1)} = \frac{\overline{F}(x/(n-i+1)+y_2)]^{n-i+1}}{\overline{F}(x/(n-i+1)+y_1)]^{n-i+1}}$$

is nondecreasing in x for $y_1 \leqslant y_2$, since F is DFR. Thus the condition (i) of Lemma 2.2 is satisfied.

(b) We show that for $y_1 \leqslant y_2$,

$$\frac{\psi_1(x,y_2)}{\psi_2(x,y_1)} = \frac{C(j:m)\overline{G}(x/(m-j+1)+y_2)]^{m-j+1}}{C(i:n)\overline{F}(x/(n-i+1)+y_1)]^{n-i+1}}$$

is nondecreasing in x. Now for $n-i \geqslant m-j$,

$$\frac{\mathrm{d}\log}{\mathrm{d}x}\left\{\frac{\psi_1(x,y_2)}{\psi_2(x,y_1)}\right\} = r_F\left(\frac{x}{n-i+1}+y_1\right) - r_G\left(\frac{x}{m-j+1}+y_2\right)$$

$$\geqslant r_F\left(\frac{x}{m-j+1}+y_1\right) - r_G\left(\frac{x}{m-j+1}+y_2\right)$$

$$(\text{since } F \text{ is DFR and } n-i+1 \geqslant m-j+1)$$

$$\geqslant 0,$$

since $X \leqslant_{\text{hr}} Y$.

Thus the condition (ii) of Lemma 2.2 is satisfied.

(c) One can show on the same lines that under the stated assumptions, condition (iii) of the above lemma is satisfied for $n-i \geqslant m-j$.

It is easy to verify that $X \leqslant_{\text{lr}} Y$ implies

$$X_{i-1:i-1} \leqslant_{\text{lr}} Y_{j-1:j-1} \quad \text{for } i \leqslant j.$$

The required result now follows from Lemma 2.2. □

By taking $F = G$ in the above theorem, we get the following result on spacings from a DFR distribution.

Corollary 2.1. *In case of random samples from a DFR distribution*

$$U_{i:n+1} \leqslant_{\text{hr}} U_{i:n} \leqslant_{\text{hr}} U_{i+1:n+1} \quad \textit{for } i = 1,\ldots,n.$$

This corollary implies that the successive normalized spacings of a random sample from a DFR distribution are increasing according to hazard rate ordering, a result proved earlier by Kochar and Kirmani (1995).

One of the basic criteria for comparing variability in two probability distributions is that of dispersive ordering.

Definition 2.1. X is less dispersed than Y ($X \leqslant_{\text{disp}} Y$) if

$$F^{-1}(v) - F^{-1}(u) \leqslant G^{-1}(v) - G^{-1}(u), \quad \forall 0 < u \leqslant v < 1. \tag{2.10}$$

This means that the difference between any two quantiles of F is smaller than the difference between the corresponding quantiles of G. A consequence of $X \leqslant_{\text{disp}} Y$ is that $|X_1 - X_2| \leqslant_{\text{st}} |Y_1 - Y_2|$ and which in turn implies $\text{var}(X) \leqslant \text{var}(Y)$ as well as $E[|X_1 - X_2|] \leqslant E[|Y_1 - Y_2|]$, where X_1, X_2 (Y_1, Y_2) are two independent copies of $X(Y)$. For details see Shaked and Shanthikumar (1994, Section 2.B).

The normalized spacings from a DFR distribution have DFR distributions (cf. Barlow and Proschan, 1966). Bagai and Kochar (1986) proved that if $X \leqslant_{\text{hr}} Y$ and either F or G is DFR, then $X \leqslant_{\text{disp}} Y$. Combining these results with the above theorem we get the following corollaries.

Corollary 2.2. *Let $X \leqslant_{\rm lr} Y$ and either F or G be DFR, then*

$$U_{i:n} \leqslant_{\rm disp} V_{j:m} \quad \text{for } i \leqslant j \text{ and } n-i \geqslant m-j.$$

Corollary 2.3. *In case of random samples from a DFR distribution*

$$U_{i:n+1} \leqslant_{\rm disp} U_{i:n} \leqslant_{\rm disp} U_{i+1:n+1} \quad \text{for } i=1,\ldots,n;$$

and as a consequence

$$\mathrm{var}(U_{i:n+1}) \leqslant \mathrm{var}(U_{i:n}) \leqslant \mathrm{var}(U_{i+1:n+1}) \quad \text{for } i=1,\ldots,n.$$

Kochar (1998) proved that if $X \leqslant_{\rm lr} Y$ and either F or G has a log-convex density, then $U_{i:n} \leqslant_{\rm lr} V_{i:n}$ for $1 \leqslant i \leqslant n$. The natural question is whether this result can be extended to the case when the sample sizes are not equal. However, we have only a partial result in this case as stated below.

Theorem 2.3. *Let $X \leqslant_{\rm lr} Y$ and let either F or G have a log-convex density, then $U_{i:n} \leqslant_{\rm lr} V_{j:m}$ for $i \leqslant j$ provided $n-i = m-j$.*

Proof. We have to prove that under the given conditions

$$\frac{g^*_{j:m}(x)}{f^*_{i:n}(x)} = \frac{C(j,m)E\{g(x/(m-j+1)+Y_{j-1:j-1})\overline{G}^{m-j}(x/(m-j+1)+Y_{j-1:j-1})\}}{C(i,n)E\{f(x/(n-i+1)+X_{i-1:i-1})\overline{F}^{n-i}(x/(n-i+1)+X_{i-1:i-1})\}}, \tag{2.11}$$

is nondecreasing in x. Here $C(i:n)$ is as defined in (2.9). Define

$$\psi_1(x,y) = g(x/(m-j+1)+y)\overline{G}^{m-j}(x/(m-j+1)+y)$$

and

$$\psi_2(x,y) = f(x/(n-i+1)+y)\overline{F}^{n-i}(x/(n-i+1)+y).$$

Replacing the DFR property by the log-convexity of f and using the same kind of arguments as in the proof of Theorem 2.2, we get the required result. □

Acknowledgements

The authors are grateful to the referee for his helpful comments and careful reading of the manuscript. One of the authors, B. Khaledi thanks the Indian Council for Cultural Relations, New Delhi, India, Razi University, Kermanshah, Iran and the Ministry of Culture and Higher Education of the Islamic Republic of Iran, Tehran, Iran for arranging scholarships which enabled him to study at the Indian Statistical Institute, New Delhi, India.

References

Bagai, I., Kochar, S.C., 1986. On tail ordering and comparison of failure rates. Commun. Statist. Theor. Meth. 15, 1377–1388.

Barlow, R.E., Proschan, F., 1966. Inequalities for linear combinations of order statistics from restricted families. Ann. Math. Statist. 37, 1574–1592.

Kochar, S.C., 1998. On stochastic ordering between distributions and their sample spacings. Statist. Probab. Lett., in press.

Kochar, S.C., Kirmani, S.N.U.A., 1995. Some results on normalized spacings from restricted families of distributions. J. Statist. Planning Inference 46, 47–57.

Raqab, M.Z., Amin, W.A., 1996. Some ordering results on order statistics and record values. IAPQR Trans. 21 (1), 1–8.

Shaked, M., Shanthikumar, J.G., 1994. Stochastic Orders and their Applications. Academic Press, San Diego, CA.