On stochastic orderings between distributions and their sample spacings

Subhash C. Kochar [*,1]

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

Abstract

Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ denote the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from a probability distribution with distribution function F. Similarly, let $Y_{1:n} \leqslant Y_{2:n} \leqslant \cdots \leqslant Y_{n:n}$ denote the order statistics of an independent random sample $Y_1, Y_2, \ldots, Y_n$ from G. The corresponding spacings are defined by $U_{i:n} \equiv X_{i:n} - X_{i-1:n}$ and $V_{i:n} \equiv Y_{i:n} - Y_{i-1:n}$, for $i = 1, 2, \ldots, n$, 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 DFR (decreasing failure rate) distribution, then the vector of $U_{i:n}$'s is stochastically smaller than the vector of $V_{i:n}$'s. If instead, we assume that X is smaller than Y in the likelihood ratio order and if either F or G is DFR, then $U_{i:n}$ is smaller than $V_{i:n}$ in the hazard rate sense for $1 \leqslant i \leqslant n$. Finally, if we make a stronger assumption on the shapes of the distributions that either X or Y has log-convex density, then the random vector of $U_{i:n}$'s is smaller than the corresponding random vector of $V_{i:n}$'s in the sense of multivariate likelihood ratio ordering.

MSC: primary 62N05; secondary 62E10

Keywords: Likelihood ratio ordering; Hazard rate ordering; Stochastic ordering; Multivariate stochastic ordering; Multivariate likelihood ratio ordering; Dispersive ordering

1. Introduction

In the case of nonnegative (skewed) random variables, the notions of stochastic orderings and variability orderings are intimately connected. For example, two exponential distributions with different hazard rates are ordered stochastically as well as according to variability ordering. There are several notions of stochastic ordering as well as of variability ordering of different degrees of strength. The differences in the variabilities in probability distributions are reflected in their samples in the form of differences in the lengths of the

corresponding sample spacings when random samples of the same size are drawn from them. If one probability distribution is more dispersed than the other then the sample spacings for that distribution will be comparatively larger in some stochastic sense. We shall make this statement more precise later. Note that almost all well-known measures of dispersion, including sample range, sample variance and Gini's mean difference are functions of sample spacings.

Spacings are of great importance in statistics and life testing. A large number of goodness-of-fit tests are based on functions of sample spacings. In the life testing context, imagine n items put on test. Then the spacings represent times between consecutive failures.

In this note we obtain connections between various types of stochastic orderings between two probability distributions and their corresponding sample spacings when random samples of the same size are drawn from them. First we review some well-known notions of stochastic orders. These can be found at one place in the book by Shaked and Shanthikumar (1994).

Let X and Y be two random variables with distribution functions F and G; and survival functions $\overline{F}$ and $\overline{G}$, respectively. Let F^{-1} and G^{-1} be the right continuous inverses of F and G, defined by $F^{-1}(u) = \sup\{x : F(x) \leq u\}$ and $G^{-1}(u) = \sup\{x : G(x) \leq u\}$, $u \in [0,1]$. We shall denote by f and g the densities of X and Y, respectively. Throughout this paper the term *increasing* is used for *monotone nondecreasing* and *decreasing* for *monotone nonincreasing*.

Definition 1.1. X is said to be stochastically smaller than Y (denoted by $X \leq_{st} Y$) if

$$\overline{F}(x) \leq \overline{G}(x) \quad \text{for all } x. \tag{1.1}$$

It is well known that Eq. (1.1) is equivalent to

$$E[\phi(X)] \leq E[\phi(Y)] \quad \text{for all increasing functions } \phi : \mathscr{R} \to \mathscr{R}, \tag{1.2}$$

for which the expectations exist.

Definition 1.2. X is said to be smaller than Y in the sense of hazard rate ordering (denoted by $X \leq_{hr} Y$) if

$$\frac{\overline{F}(x)}{\overline{G}(x)} \quad \text{is decreasing in } x. \tag{1.3}$$

In the continuous case this is equivalent to

$$r_G(x) \leq r_F(x) \quad \text{for all } x, \tag{1.4}$$

where $r_F = f/\overline{F}$ and $r_G = g/\overline{G}$ are the hazard (or failure) rates of F and G, respectively.

Definition 1.3. X is said to be smaller than Y in the sense of likelihood ratio ordering (denoted by $X \leq_{lr} Y$) if

$$\frac{f(x)}{g(x)} \quad \text{is decreasing in } x. \tag{1.5}$$

We have the following chain of implications among these partial orderings of distributions:

$$X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y.$$

The above notions of stochastic dominance among univariate random variables can be extended to the multivariate case.

Definition 1.4. A random vector $X = (X_1, \ldots, X_n)$ is smaller than another random vector $Y = (Y_1, \ldots, Y_n)$ in the multivariate stochastic order (and written as $X \overset{st}{\leqslant} Y$) if

$$E[\phi(X)] \leqslant E[\phi(Y)] \quad \text{for all increasing functions } \phi \tag{1.6}$$

for which the expectations exist.

Karlin and Rinott (1980) introduced and studied the concept of multivariate likelihood ratio ordering. Let f and g denote the density functions of X and Y, respectively.

Definition 1.5. A random vector $X = (X_1, \ldots, X_n)$ is smaller than another random vector $Y = (Y_1, \ldots, Y_n)$ in the multivariate likelihood ratio order (written as $X \overset{lr}{\leqslant} Y$) if

$$f(x)g(y) \leqslant f(x \wedge y)g(x \vee y) \quad \text{for every } x \text{ and } y \text{ in } \mathscr{R}^n, \tag{1.7}$$

where

$$x \wedge y = (min(x_1, y_1), \ldots, min(x_n, y_n))$$

and

$$x \vee y = (max(x_1, y_1), \ldots, max(x_n, y_n)).$$

It is known that multivariate likelihood ratio ordering implies multivariate stochastic ordering, but the converse is not true. Also if two random vectors are ordered according to multivariate stochastic ordering or multivariate likelihood ratio ordering, then their corresponding subsets are also ordered accordingly. It should be noted that unless the components of random vectors are independent, component-wise stochastic (likelihood ratio) ordering between two random vectors may not imply multivariate stochastic (likelihood ratio) ordering between them. See Chapters 1 and 4 of Shaked and Shanthikumar (1994) for more details on various kinds of stochastic orderings and their inter-relationships.

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

Definition 1.6. X is less dispersed than Y ($X \overset{disp}{\leqslant} Y$) if

$$F^{-1}(v) - F^{-1}(u) \leqslant G^{-1}(v) - G^{-1}(u), \quad \forall 0 < u \leqslant v < 1. \tag{1.8}$$

This means that the difference between any two quantiles of F is smaller than the difference between the corresponding quantiles of G. It is easy to see that in the continuous case, the relation $X \overset{disp}{\leqslant} Y$ can be equivalently expressed as $r_G(G^{-1}(u)) \leqslant r_F(F^{-1}(u))$, for all $0 \leqslant u \leqslant 1$. A consequence of $X \overset{disp}{\leqslant} Y$ is that $|X_1 - X_2| \leqslant_{st} |Y_1 - Y_2|$ and which in turn implies $var(X) \leqslant var(Y)$ 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 Section 2.B of Shaked and Shanthikumar (1994).

Bagai and Kochar (1986) proved the following result on connections between hazard rate ordering and dispersive ordering under some restrictions on the shapes of the distributions.

Theorem 1.1. (a) If $X \leqslant_{hr} Y$ and either F or G is DFR (decreasing failure rate), then $X \overset{disp}{\leqslant} Y$,

(b) if $X \overset{disp}{\leqslant} Y$ and either F or G is IFR (increasing failure rate), then $X \leqslant_{hr} Y$.

Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ denote the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from a distribution with distribution function F. Similarly, let $Y_{1:n} \leqslant Y_{2:n} \leqslant \cdots \leqslant Y_{n:n}$ denote the order statistics of an independent random sample $Y_1, Y_2, \ldots, Y_n$ from a distribution with distribution function G. The corresponding spacings are defined by $U_{i:n} \equiv X_{i:n} - X_{i-1:n}$ and $V_{i:n} \equiv Y_{i:n} - Y_{i-1:n}$, for $i = 1, 2, \ldots, n$, where $X_{0:n} = Y_{0:n} \equiv 0$. We use U and V to denote the vectors of spacings of the X-sample and the Y-sample, respectively.

Bartoszewicz (1986) in his Lemma 3(c) has shown that $X \overset{\text{disp}}{\leqslant} Y \Rightarrow U \overset{\text{st}}{\leqslant} V$. This observation along with the result contained in Theorem 1.1 (a) leads to the following theorem.

Theorem 1.2. *Let* $X \leqslant_{\text{hr}} Y$ *and let either* F *or* G *be DFR. Then*

$$U \overset{\text{st}}{\leqslant} V. \tag{1.9}$$

Corollary 1.1. *Under the conditions of Theorem 1.2*
  (a)

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

In particular,

$$X_{n:n} - X_{1:n} \leqslant_{\text{st}} Y_{n:n} - Y_{1:n}.$$

  (b)

$$s_X^2 \leqslant_{\text{st}} s_Y^2,$$

where s_X^2 *and* s_Y^2 *are the sample variances of the two samples.*
  (c)

$$\eta_X \leqslant_{\text{st}} \eta_Y$$

where

$$\eta_X = \left[\binom{n}{2} \right]^{-1} \sum\sum_{i<j} |X_{j:n} - X_{i:n}|$$

is the Gini's mean difference for the X-*sample. Similarly we define* η_Y.

Proof. (a) The result follows by adding the corresponding components of the random vectors U and V from $i + 1$ to j and using the above theorem.
  (b) Note that the sample variance can be expressed as

$$s_X^2 = [n(n-1)]^{-1} \sum\sum_{i<j} (X_{j:n} - X_{i:n})^2$$

$$= [n(n-1)]^{-1} \sum\sum_{i<j} (U_{j:n} + U_{j-1:n} + \cdots + U_{i+1:n})^2$$

which is an increasing function of U. Since increasing functions of stochastically ordered random vectors are stochastically ordered, the required result follows from the above theorem.
  (c) The proof follows from the previous theorem and the fact that, as in part (b), the Gini's mean difference can be expressed in the form of an increasing function of the vector of spacings. $\square$

The essence of the above results is that the differences in the variabilities in two probability distributions are reflected in their samples in the form of stochastic orderings between the corresponding sample spacings.

David and Groeneveld (1982) have used expected lengths of spacings as a measure of local variability in a distribution. However, they have considered a one-sample problem.

We pursue this topic further in this note and show that the results of Theorem 1.2 can be strengthened under some stronger conditions on the underlying distributions. These results might be useful in studying the properties of estimates of measures of dispersion.

2. Main results

In the next theorem, we assume likelihood ratio ordering between X and Y and strengthen the results of Theorem 1.2 from stochastic ordering to hazard rate ordering.

Theorem 2.1. Let $X \leqslant_{lr} Y$ and let either F or G be DFR. Then

$$U_{i:n} \leqslant_{hr} V_{i:n} \quad \text{for } 1 \leqslant i \leqslant n. \tag{2.1}$$

To prove this result we shall need the following lemma from Kochar and Kirmani (1995):

Lemma 2.1. 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, $Z_1 \leqslant_{lr} Z_2$ implies

$$\frac{E[\psi_1(x, Z_2)]}{E[\psi_2(x, Z_1)]} \quad \text{is increasing in } x, \tag{2.2}$$

provided that the expectations exist.

Proof of Theorem 2.1. We shall prove Eq. (2.1) assuming that F is DFR. As shown in Kochar and Kirmani (1995), the survival function of $V_{i:n}$ is

$$\overline{H}_{G:i}(x) = C(i:n) \int_0^\infty [\overline{G}(x+u)]^{n-i+1} \, dG^{i-1}(u)$$

$$= C(i:n) E[\psi_1(x, Y_{i-1:i-1})], \tag{2.3}$$

where,

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

$\psi_1(x, y) = \overline{G}^{n-i+1}(x+y)$ and $Y_{i-1:i-1}$ is the maximum of $(i-1)$ i.i.d. Y_i's.

Similarly, the survival function of $U_{i:n}$ is

$$\overline{H}_{F:i}(x) = C(i:n)E[\psi_2(x, X_{i-1:i-1})], \tag{2.4}$$

where, $\psi_2(x, y) = \overline{F}^{n-i+1}(x + y)$ and $X_{i-1:i-1}$ is the maximum of $(i-1)$ i.i.d. X_i's.

We have to prove that under the given conditions,

$$\frac{\overline{H}_{G:i}(x)}{\overline{H}_{F:i}(x)} = \frac{E[\psi_1(x, Y_{i-1:i-1})]}{E[\psi_1(x, X_{i-1:i-1})]} \quad \text{is increasing in } x. \tag{2.5}$$

It is easy to see that $X \leqslant_{lr} Y$ implies that $X_{j:j} \leqslant_{lr} Y_{j:j}$ for $j = 1, 2, \ldots, n$. By identifying $X_{i-1:i-1}$ with Z_1 and $Y_{i-1:i-1}$ with Z_2, we notice that the required result will follow if conditions (i), (ii) and (iii) of the previous lemma are satisfied.

Let us verify them one by one.

Condition (i) is satisfied since

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

is increasing in x for $y_1 \leqslant y_2$ as F is assumed to be DFR.

Let us now verify condition (ii).

$$\frac{\psi_1(x, y_2)}{\psi_2(x, y_1)} = \left[\frac{\overline{G}(x + y_2)}{\overline{F}(x + y_1)}\right]^{n-i+1}$$

will be increasing in x for $y_1 \leqslant y_2$ if and only if

$$\log\left[\frac{\psi_1(x, y_2)}{\psi_2(x, y_1)}\right] = (n - i + 1)[\log\overline{G}(x + y_2) - \log\overline{F}(x + y_1)]$$

is increasing in x for $y_1 \leqslant y_2$. Differentiating both sides with respect to x, we see that this will be true if and only if

$$(n - i + 1)[r_F(x + y_1) - r_G(x + y_2)] \geqslant 0 \quad \text{for all } x \text{ and for } y_1 \leqslant y_2. \tag{2.6}$$

Now Eq. (2.6) is true since

$$r_G(x + y_2) \leqslant r_F(x + y_2) \leqslant r_F(x + y_1),$$

for $y_1 \leqslant y_2$ as $r_G \leqslant r_F$ and F is DFR.

By using the same kind of arguments as above, it can be shown that for fixed x,

$$\frac{\psi_1(x, y)}{\psi_2(x, y)} = \left[\frac{\overline{G}(x + y)}{\overline{F}(x + y)}\right]^{n-i+1}$$

is increasing in y.

Hence the result. $\square$

It is known that the spacings of a random sample from a DFR distribution are DFR (cf. Barlow and Proschan, 1966). Using this result in conjunction with Theorem 1.1, we get the following corollary.

Corollary 2.1. *Under the conditions of Theorem* 2.1

$$U_{i:n} \overset{\text{disp}}{\leqslant} V_{i:n} \quad \text{for } i = 1, 2, \ldots, n.$$

In the next theorem we assume instead that either F or G has log-convex density, a condition stronger than DFR property, and establish multivariate likelihood ratio ordering between the vectors of spacings U and V. We shall need the concept of *supermodular functions* as defined below.

Definition 2.1. A real function ϕ on $\mathscr{R}^n$ is called supermodular if

$$\phi(x) + \phi(y) \leqslant \phi(x \vee y) + \phi(x \wedge y) \quad \text{for all } x, y \in \mathscr{R}^n.$$

It is known that a function ϕ is supermodular if and only if all its second derivatives are nonnegative (cf. Shaked and Shanthikumar, 1997).

Theorem 2.2. *Let $X_1, \ldots, X_n$ be a random sample from F and let $Y_1, \ldots, Y_n$ be an independent random sample from G. Let $X \leqslant_{lr} Y$ and let either F or G have log-convex density. Then*

$$U \overset{lr}{\leqslant} V. \tag{2.7}$$

Proof. Let us assume that f is log-convex. The proof is similar for the case when g is log-convex. As in Kochar and Kirmani (1995), the joint density of $U = (U_{1:n}, \ldots, U_{n:n})$ is

$$h_F(u_1, \ldots, u_n) = n! \prod_{j=1}^{n} f\left(\sum_{i=1}^{j} u_i\right), \quad u_i \geqslant 0, \ i = 1, \ldots, n,$$

and that of $V = (V_{1:n}, \ldots, V_{n:n})$ is

$$h_G(v_1, \ldots, v_n) = n! \prod_{j=1}^{n} g\left(\sum_{i=1}^{j} v_i\right), \quad v_i \geqslant 0, \ i = 1, \ldots, n.$$

To prove Eq. (2.7), we have to show that under the given conditions,

$$h_F(u)h_G(v) \leqslant h_F(u \wedge v)h_G(u \vee v) \quad \text{for all } u, v \in \mathscr{R}^n. \tag{2.8}$$

Since under $X \leqslant_{lr} Y$, g/f is nondecreasing, it follows that

$$\frac{h_G(u \vee v)}{h_F(u \vee v)} \geqslant \frac{h_G(v)}{h_F(v)} \quad \text{for all } u, v \in \mathscr{R}^n,$$

which in turn implies that

$$h_G(u \vee v)h_F(u \wedge v) \geqslant h_F(u \vee v)h_F(u \wedge v)\frac{h_G(v)}{h_F(v)}, \tag{2.9}$$

for all $u, v \in \mathscr{R}^n$.

Now Eq. (2.8) will hold if we can prove that under the assumed conditions

$$h_F(u)h_F(v) \leqslant h_F(u \wedge v)h_F(u \vee v), \text{ for all } u, v \in \mathscr{R}^n,$$

or equivalently if

$$\sum_{j=1}^{n} \log f\left(\sum_{1=1}^{j} u_i\right) + \sum_{j=1}^{n} \log f\left(\sum_{1=1}^{j} v_i\right)$$

$$\leqslant \sum_{j=1}^{n} \log f\left(\sum_{1=1}^{j}(u_i \vee v_i)\right) + \sum_{j=1}^{n} \log f\left(\sum_{1=1}^{j}(u_i \wedge v_i)\right) \tag{2.10}$$

for all $u, v \in \mathscr{R}^n$. Let

$$\psi(x_1, \ldots, x_n) = \sum_{i=1}^{n} \log f(x_1 + \cdots + x_i).$$

It is easy to see that the function ψ is supermodular since f is log-convex. Hence

$$\psi(x) + \psi(y) \leqslant \psi(x \vee y) + \psi(x \wedge y) \tag{2.11}$$

for all $x, y \in \mathscr{R}^n$, thus proving Eq. (2.10).

This completes the proof of the theorem. $\quad\square$

3. Examples

Example 3.1. Let $\overline{F}(x) = \exp[-\{x + \theta(x + e^{-x} - 1)\}]$, $\theta > 0$, be the Makeham distribution with hazard rate $r_F(x) = 1 + \theta(1 - e^{-x})$ and let $\overline{G}(x) = \exp\{-x\}$, be the exponential distribution with hazard rate $r_G(x) = 1$. Then for $\theta > 0$, $X \leqslant_{\mathrm{hr}} Y$ and Y is DFR. It follows from Theorem 1.2 that in this case $X \overset{\mathrm{disp}}{\leqslant} Y$. Corollary 1.1 is now applicable and it gives convenient bounds on the moments of measures of dispersion for the X-sample in terms of those from the exponential distribution with mean 1.

However, it can be seen that in this case $X \not\leqslant_{\mathrm{lr}} Y$.

Example 3.2. Let Y be a random variable whose distribution is a mixture of two exponential distributions with density function

$$\lambda_1 \lambda_2 / (\lambda_1 + \lambda_2) [e^{-\lambda_1 x} + e^{-\lambda_2 x}], \quad x > 0,$$

where $\lambda_1 > 0$, $\lambda_2 > 0$ are the parameters. Let X have exponential distribution with parameter $\overline{\lambda} = (\lambda_1 + \lambda_2)/2$. Then it is easy to prove that $X \leqslant_{\mathrm{lr}} Y$. Since Y has a log-convex density, the conditions of Theorem 2.2 are satisfied and as a result $U \overset{\mathrm{lr}}{\leqslant} V$. This result gives lower bounds on the moments of measures of dispersion in sampling from a mixture of two exponential distributions in terms of that from an exponential distribution with parameter $\overline{\lambda}$.

Acknowledgements

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.

Bartoszewicz, J., 1986. Dispersive ordering and the total time on test transformations. Statist. Probab. Lett. 4, 285–288.

David, H.A., Groeneveld, R.A., 1982. Measures of local variation in a distribution: expected length of spacings and variances of order statistics. Biometrika 69, 227–232.

Karlin, S., Rinott, Y., 1980. Classes of orderings of measures and related correlation inequalities I. Multivariate totally positive distributions. J. Mult. Anal. 10, 467–498.

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

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

Shaked, M., Shanthikumar, J.G., 1997. Supermodular stochastic order and positive dependence of random variables. J. Mult. Anal. 61, 86–101.