Monotone coupling and stochastic ordering of order statistics

1 author:

Taizhong Hu
University of Science and Technology of China

[*System Science and Mathematical Sciences* (English Series), **8**(1995), 209-214]

Monotone coupling and stochastic ordering of order statistics*

Taizhong Hu

University of Science and Technology of China

Abstract Let $X_i, Y_j, i, j = 1, \ldots, n$, be independent nonnegative r.v.'s with $X_i \sim F_i(t) = 1-\exp(-R_i(t))$ and $Y_j \sim G_j(t) = 1-\exp(-R_j^*(t))$, $j = 1, \ldots, n$. Two kinds of model are considered, and the monotone coupling of order statistics $X_{(1)}, \ldots, X_{(n)}$ and $Y_{(1)}, \ldots, Y_{(n)}$ is established under certain conditions ; that is, it is possible to construct on a common probability space r.v.'s X_i', Y_i', $i = 1, \ldots, n$, such that for each i, $Y_{(i)}' \leq X_{(i)}'$ a.s. and $(X_1', \ldots, X_n') =^d (X_1, X_2, \cdots, X_n)$, $(Y_1', \ldots, Y_n') =^d (Y_1, Y_2, \cdots, Y_n)$. The result of Model 1 generalizes the results of Proschan and Sethuraman (1976), Ball (1985) and Barbour et al (1991). We also obtain a result about stochastic ordering of order statistics. All these results are useful in studying epidemic model and reliability theory.

Key words: monotone coupling, order statistics, epidemic.
AMS Subject Classification (1991): 62N05.

1 Introduction

Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be independent random variables (r.v.'s). A monotone coupling of the X and Y order statistics means that there exist r.v.'s $X_1', X_2', \ldots, X_n'; Y_1', Y_2', \ldots, Y_n$ such that

$$(X_1', X_2', \ldots, X_n') =^d (X_1, X_2, \ldots, X_n),$$

$$(Y_1', Y_2', \ldots, Y_n') =^d (Y_1, Y_2, \ldots, Y_n),$$

$$Y_{(j)}' \leq X_{(j)}', \ j = 1, \ldots, n.$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ are the order statistics of $X_1, X_2, \ldots, X_n$ (with the corresponding notations for the $X', Y, Y'-$sample). Monotone coupling is a strong notion which trivially implies

*Supported by the NNSF of China

stochastic ordering. A r.v. U is said to be stochastically larger than a r.v. V, denoted $U \succ^{st} V$, if $P(U > t) \geq P(V > t)$ holds for all $t > 0$. A random vector $\mathbf{U} = (U_1, U_2, \cdots, U_n)$ is said to be stochastically larger than a random vector $\mathbf{V} = (V_1, V_2, \cdots, V_n)$, denoted $\mathbf{U} \succ^{st} \mathbf{V}$, if $g(\mathbf{U}) \succ^{st} g(\mathbf{V})$ for every increasing function g.

Let $X_i, i = 1, \ldots, n$, be independent exponential r.v.'s with parameters $\lambda_i, i = 1, \ldots, n$, and if $Y_i,\ i = 1, \ldots, n$, are independent exponential r.v.'s with common parameters equal to $\sum_{i=1}^{n} \lambda_i / n$, then there exists a monotone coupling of the order statistics . This result is due to Proschan and Sethuraman [1], and independently to Ball [2]. Ball used this result to compare the behavior of the heterogeneous epidemic model with a suitable chosen homogeneous epidemic model. Barbour, Lindvall and Rogers [3] , hereafter refered to BLR [3], proved a extension to a more general class of distributions for which the failure rate function $r(x)$ is decreasing, and $xr(x)$ is increasing.

The present paper is concerned with the comparison, in greater generality, between the order statistics of two heterogeneous sets of independent r.v.'s. This is of interest partly because it is a natural extension of coupling techniques to prove some stochastic ordering of order statistics for statistical analysis and inference, and also because it allows us to compare the behavior of the more complicated epidemic models and to find some applications in reliability theory.

Now we recall a definition of majorization and some facts, which will be used in the following sections.

Definition 1. Let $\mathbf{a} = (a_1, a_2, \cdots, a_n)$, $\mathbf{b} = (b_1, b_2, \cdots, b_n)$ be two n-dimensional vectors and $a_{[1]} \geq a_{[2]} \geq \cdots \geq a_{[n]},\ b_{[1]} \geq b_{[2]} \geq \cdots \geq b_{[n]}$ denote their ordered components.

1) $\mathbf{a}$ is said to majorize $\mathbf{b}$ (in symbols, $\mathbf{a} \succ^{m} \mathbf{b}$, or $\mathbf{b} \prec^{m} \mathbf{a}$) if

$$\sum_{i=1}^{m} a_{[i]} \geq \sum_{i=1}^{m} b_{[i]} \qquad \text{for } \ m = 1, \ldots, n-1$$

and $\sum_{i=1}^{n} a_{[i]} = \sum_{i=1}^{n} b_{[i]}$.

2) $\mathbf{b}$ is said to be weakly supermajorized by $\mathbf{a}$ (in symbols, $\mathbf{b} \prec^{w} \mathbf{a}$) if

$$\sum_{i=m}^{n} a_{[i]} \leq \sum_{i=m}^{n} b_{[i]},\ m = 1, 2, \ldots, n.$$

The notion of majorization defines a partial ordering of the diversity of the components of vectors. For a comprehensive treatment of majorization, see Marshall and Olkin [4]. In particular, we have the following facts.

Fact 1. $\mathbf{a} \succ^{m} \mathbf{b}$ if and only if there exists a finite number of vectors $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_N$ such that

$$\mathbf{a} = \mathbf{c}_1 \succ^{m} \mathbf{c}_2 \succ^{m} \cdots \succ^{m} \mathbf{c}_N = \mathbf{b}$$

and such that for all j, $\mathbf{c}_j$ and $\mathbf{c}_{j+1}$ differ in two coordinates only.

Fact 2. $\mathbf{a} \succ^{m} \mathbf{b}$ if and only if $\sum_{i=1}^{n} g(a_i) \geq (\leq) \sum_{i=1}^{n} g(b_i)$ holds for all convex (concave) function g.

Fact 3. $\mathbf{b} \prec^w \mathbf{a}$ if and only if there exists a real vector $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ such that $\mathbf{b} \prec^m \mathbf{c}$, and $\mathbf{c} \geq \mathbf{a}$ (i.e., $c_i \geq a_i,\ i = 1, 2, \ldots, n$).

Two kinds of model are considered in this paper. The first model, considered by Proschan and Sethuraman, Ball and BLR, is an accelerating life model:

Let $V_1, V_2, \ldots, V_n$ be independent identically distributed nonnegative r.v.'s with a commom c.d.f. F. Define two samples $X_i = V_i/\lambda_i,\ Y_i = V_i/\mu_i,\ i = 1, 2, \ldots, n$, where the $\lambda_i, \mu_i' s$ are positive constants. We obtain a monotone coupling of the order statistics (as Theorem 1), under the condition that the failure rate $r(x)$ is decreasing, $xr(x)$ is increasing, and $(\mu_1, \mu_2, \ldots, \mu_n) \prec^w (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

The second model is a proportional hazard model:

Let $X_i,\ i = 1, \ldots, n$, be independent nonnegative r.v.'s with failure rate $\lambda_i r(x),\ i = 1, \ldots, n$, and let $Y_i, i = 1, \ldots, n$, be independent nonnegative r.v.'s with failure rate $\mu_i r(x),\ i = 1, \ldots, n$. We also have a monotone coupling of the order statistics (as Theorem 2), only if $(\mu_1, \mu_2, \ldots, \mu_n) \prec^w (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Finally in Section 3 we give a weak result about stochastic ordering of order statistics under the weak condition.

2 Monotone coupling of order statistics

Lemma 1. Let $\xi_1, \xi_2, \eta_1, \eta_2$ be independent nonnegative r.v.'s with distributions

$$\begin{aligned} P(\xi_i > t) &= \overline{F}(\lambda_i t), \\ P(\eta_i > t) &= \overline{G}(\mu_i t), \quad i = 1, 2. \end{aligned}$$

where $\lambda_1, \lambda_2, \mu_1, \mu_2$ are positive constants and $(\lambda_1, \lambda_2) \succ^m (\mu_1, \mu_2)$. Assume that

$$\overline{F}(t) = \exp(-\int_0^t r(s)ds),$$

where $r(x)$ is decreasing, and $xr(x)$ is increasing. Then on some probability space there exist r.v.'s $\xi_1', \xi_2', \eta_1', \eta_2'$ such that

$$(\xi_1', \xi_2') =^d (\xi_1, \xi_2), \quad (\eta_1', \eta_2') =^d (\eta_1, \eta_2) \tag{1}$$

and

$$\eta_{(1)}' \leq \xi_{(1)}', \quad \eta_{(2)}' \leq \xi_{(2)}' \text{ a.s..} \tag{2}$$

Proof. Firstly, we compute

$$\begin{aligned} P(\xi_{(1)} > t) &= \overline{F}(\lambda_1 t)\overline{F}(\lambda_2 t) \\ &= \exp(-R(\lambda_1 t) - R(\lambda_2 t)) \\ &\geq \exp(-R(\mu_1 t) - R(\mu_2 t)) \\ &= P(\eta_{(1)} > t), \end{aligned}$$

exploiting the concavity of $R(t) = \int_0^t r(s)ds$ and Fact 2.

Next, we take any $0 \le v \le w$ such that

$$R(\lambda_1 w) + R(\lambda_2 w) = R(\mu_1 v) + R(\mu_2 v) \tag{3}$$

so that $P(\eta_{(1)} > v) = P(\xi_{(1)} > w)$, and show that for $t \ge w$

$$P(\xi_{(2)} > t | \xi_{(1)} = w) \ge P(\eta_{(2)} > t | \eta_{(1)} = v). \tag{4}$$

Without loss of generality, we may assume that

$$\lambda_1 < \mu_1 < \mu_2 < \lambda_2. \tag{5}$$

For convenience, we define

$$\theta_1 = \frac{\lambda_1 r(\lambda_1 w)}{\lambda_1 r(\lambda_1 w) + \lambda_2 r(\lambda_2 w)}, \qquad \theta_2 = 1 - \theta_1;$$

$$\varphi_1 = \frac{\mu_1 r(\mu_1 v)}{\mu_1 r(\mu_1 v) + \mu_2 r(\mu_2 v)}, \qquad \varphi_2 = 1 - \varphi_1.$$

$$\begin{aligned} a_1 &= -R(\lambda_1 t) + R(\lambda_1 w); \\ a_2 &= -R(\lambda_2 t) + R(\lambda_2 w); \\ b_1 &= -R(\mu_1 t) + R(\mu_1 v); \\ b_2 &= -R(\mu_2 t) + R(\mu_2 v). \end{aligned}$$

Noting that (3) and (5) imply

$$\lambda_1 w < \mu_1 v < \mu_2 v < \lambda_2 w,$$

and hence

$$\frac{1}{2} \le \varphi_2 \le \theta_2 \le 1. \tag{6}$$

Using the above notation, we get that

$$\begin{aligned} LHS := {} & P(\xi_{(2)} > t | \xi_{(1)} = w) \\ = {} & \theta_1 e^{a_2} + \theta_2 e^{a_1} \\ = {} & \frac{1}{2}(e^{a_1} + e^{a_2}) + (\theta_2 - \frac{1}{2})(e^{a_1} - e^{a_2}), \end{aligned}$$

and

$$\begin{aligned} RHS := {} & P(\eta_{(2)} > t | \eta_{(1)} = v) \\ = {} & \frac{1}{2}(e^{b_1} + e^{b_2}) + (\varphi_2 - \frac{1}{2})(e^{b_1} - e^{b_2}). \end{aligned}$$

In order to prove (4), we first need to show that

$$a_1 + a_2 \ge b_1 + b_2, \tag{7}$$

$$a_1 \geq a_2, \qquad b_1 \geq b_2 \quad \text{and } a_1 \geq b_1. \tag{8}$$

Indeed, observing that $R(\cdot)$ is concave function and applying Fact 2, the inequality (7) follows immediately. Next

$$a_1 - b_1 \geq (R(\mu_1 t) - R(\lambda_1 t)) - (R(\mu_1 w) - R(\lambda_1 w)) \geq 0$$

because, under the hypothesie that $xr(x)$ is increasing, we have that $R(\mu_1 t) - R(\lambda_1 t)$ is increasing in t. Similarly we can show that $a_1 \geq a_2$ and $b_1 \geq b_2$, then (7) and (8) hold.

Since (7) and (8) can be divided into three cases, we now verify (4) by considering the three cases.

Case 1: Suppose that $b_2 \leq a_2 \leq b_1 \leq a_1 \leq 0$. Using (6) we then have

$$\begin{aligned} LHS - RHS &= \frac{1}{2}(e^{a_1} + e^{a_2}) - \frac{1}{2}(e^{b_1} + e^{b_2}) \\ &\quad + (\theta_2 - \frac{1}{2})(e^{a_1} - e^{a_2}) - (\varphi_2 - \frac{1}{2})(e^{b_1} - e^{b_2}) \\ &\geq e^{a_2} - e^{b_2} + \varphi_2(e^{a_1} - e^{a_2} - e^{b_1} + e^{b_2}) \\ &= (1 - \varphi_2)(e^{a_2} - e^{b_2}) + \varphi_2(e^{a_1} - e^{b_1}) \geq 0. \end{aligned}$$

Case 2: Suppose that $b_2 \leq b_1 \leq a_2 \leq a_1 \leq 0$, then

$$\begin{aligned} LHS - RHS &= e^{a_2} - e^{b_1} + \theta_2(e^{a_1} - e^{a_2}) + (1 - \varphi_2)(e^{b_1} - e^{b_2}) \\ &\geq 0 \end{aligned}$$

Case 3: Suppose that $a_2 \leq b_2 \leq b_1 \leq a_1 \leq 0$. Using (6) and (7) we get

$$\begin{aligned} LHS - RHS &\geq (\theta_2 - \frac{1}{2})(e^{a_1} - e^{a_2}) - (\varphi_2 - 1)(e^{b_1} - e^{b_2}) \\ &\geq (\varphi_2 - \frac{1}{2})[(e^{a_1} - e^{b_1}) + (e^{b_2} - e^{a_2})] \\ &\geq 0. \end{aligned}$$

Combining the above three cases yields the inequality (4).

Let $F_{(1)}, G_{(1)}$ denote the cumulative distribution functions (c.d.f.) of $\xi_{(1)}$ and $\eta_{(1)}$, respectively, and also let $F_{(2)|(1)}$ and $G_{(2)|(1)}$ denote the conditional c.d.f.'s of $\xi_{(2)}|\xi_{(1)}$ and $\eta_{(2)}|\eta_{(1)}$, respectively. Now construct $\xi'_{(i)}, \eta'_{(i)}$, $i = 1, 2$, from four independent $U[0,1]$ r.v.'s U_1, U_2, U_3 and U_4 in the following obvious manner:

$$\xi'_{(1)} = F^{-1}_{(1)}(U_1), \qquad\qquad \eta'_{(1)} = G^{-1}_{(1)}(U_1);$$

$$\xi'_{(2)} = F^{-1}_{(2)|(1)}(U_2|\xi'_{(1)}), \qquad \eta'_{(2)} = G^{-1}_{(2)|(1)}(U_2|\eta'_{(1)}),$$

where $F^{-1}_{(1)}, F^{-1}_{(2)|(1)}, G^{-1}_{(1)}$ and $G^{-1}_{(2)|(1)}$ are respectively the inverse functions of $F_{(1)}, F_{(2)|(1)}, G_{(1)}$ and $G_{(2)|(1)}$. Since $F_{(1)}(t) \leq G_{(1)}(t)$ for all $t > 0$, then $\eta'_{(1)} \leq \xi'_{(1)}$. Noting that $w = \xi'_{(1)}$ and $v = \eta'_{(1)}$ satisfy (3) and (4), we get that $\eta'_{(2)} \leq \xi'_{(2)}$. Now to get $\xi'_i, i = 1, 2$, we set

$$\xi'_1 = \xi'_{(1)} I_{(U_3 \leq g(\xi'_{(1)}, \xi'_{(2)}))} + \xi'_{(2)} I_{(U_3 > g(\xi'_{(1)}, \xi'_{(2)}))};$$

$$\xi_2' = \xi_{(1)}' I_{(U_3 > g(\xi_{(1)}', \xi_{(2)}'))} + \xi_{(2)}' I_{(U_3 \le g(\xi_{(1)}', \xi_{(2)}'))}.$$

where

$$g(x, y) = P(\xi_1 < \xi_2 | \xi_{(1)} = x, \xi_{(2)} = y).$$

disposing of $\eta_i' s$ in a similar fashion.

We can easily prove that the $\xi_i', \eta_i', i = 1, 2$, constructed above, satisfy the condition (1) and (2), and this completes the proof of the lemma.

Theorem 1. Let $X_1, X_2, \dots, X_n; Y_1, Y_2, \dots, Y_n$ be independent nonnegative r.v.'s with distributions

$$\begin{aligned} P(X_i > t) &= \overline{F}(\lambda_i t) \equiv \exp(-\int_0^{\lambda_i t} r(s) ds), \\ P(Y_i > t) &= \overline{F}(\mu_i t), \qquad i = 1, 2, \dots, n, \end{aligned} \tag{2.9}$$

where $\lambda_1, \dots, \lambda_n; \mu_1, \dots, \mu_n$ are positive constants and

$$(\mu_1, \mu_2, \dots, \mu_n) \prec^m (\lambda_1, \lambda_2, \dots, \lambda_n). \tag{10}$$

Assume that the failure rate $r(x)$ is decreasing, and $xr(x)$ is increasing. Then on some probability space $(\Omega, \mathcal{F}, P)$ there exist r.v.'s $X_1', \dots, X_n'; Y_1', \dots, Y_n'$ such that

$$\begin{aligned} &(X_1', X_2', \dots, X_n') =^d (X_1, X_2, \cdots, X_n), \\ &(Y_1', Y_2', \dots, Y_n') =^d (Y_1, Y_2, \dots, Y_n) \\ &Y_{(i)}' \le X_{(i)}' \quad \text{a.s. for } i = 1, \dots, n. \end{aligned} \tag{2.11}$$

Proof. The proof proceeds by repeated application of Lemma 1. Start with an independent sequence $X_i,\ i = 1, \dots, n$, with distributions given by (9), and assume that the probability space $(\Omega, \mathcal{F}, P)$ supports an infinite sequence of independent $U[0, 1]$ r.v.'s $U_i,\ i = 1, 2, \dots$. By Fact 1, there exist a finite number of real vectors $\mathbf{c}_j = (c_{j1}, c_{j2}, \dots, c_{jn}),\ j = 1, \dots, N$, such that

$$(\lambda_1, \dots, \lambda_n) = \mathbf{c}_1 \succ^m \mathbf{c}_2 \succ^m \cdots \succ^m \mathbf{c}_N = (\mu_1, \dots, \mu_n)$$

and also such that for any j, $\mathbf{c}_j$ and $\mathbf{c}_{j+1}$ differ in two coordinates only. We now describe the first step of the procedure. Without loss of generality assume that $\mathbf{c}_1$ and $\mathbf{c}_2$ differ in the first two coordinates only, and then consolidate X_1 and X_2. Make two independent $U[0, 1]$ r.v.'s $W_1,\ W_2$ by the recipe

$$W_1 = F_{(1)}(\tilde{X}_1), \qquad W_2 = F_{(2)|(1)}(\tilde{X}_2 | \tilde{X}_1),$$

where $F_{(1)}$ is the c.d.f. of $\tilde{X}_1 \equiv \min(X_1, X_2)$ and $F_{(2)|(1)}(\cdot | x)$ is the conditional c.d.f. of $\tilde{X}_2 \equiv \max(X_1, X_2)$ given $\tilde{X}_1$. Set $\xi_{2i} = X_i, i = 3, \dots, n$, and apply the method of Lemma 1 to construct two r.v.'s ξ_{21}, ξ_{22} on the probability space $(\Omega, \mathcal{F}, P)$ (see BLR [3]), such that

$$\min(\xi_{21}, \xi_{22}) \le \tilde{X}_1, \qquad \max(\xi_{21}, \xi_{22}) \le \tilde{X}_2.$$

Hence we have changed $X_j, j=1,\dots,n$ into $\xi_{2j}, j=1,\dots,n$, on $(\Omega,\mathcal{F},P)$ in such a way that

1) $\xi_{2j}, j=1,\dots,n$, are independent;

2) $\xi_{2j} \sim F_{2j}(t) = 1-\exp(-R(c_{2i}t)),\ i=1,\dots,n$;

3) $\xi_{2(j)} \le X_{(j)}$ a.s. for $j=1,\dots,n$,

where $\xi_{2(j)}, j=1,\dots,n$, are the order statistcs of $\xi_{2j},\ j=1,\dots,n$. 3) means the change from $X_j,\ j=1,\dots,n$ to $\xi_{2j},\ j=1,\dots,n$, does not increase the order statistics. We now just go on doing this from independent r.v.'s $\xi_{2j},\ j=1,\dots,n$. After $N-1$ steps we change $\xi_{N-1,j},\ j=1,\dots,n$ into $\xi_{Nj},\ j=1,\dots,n$, on the probability space $(\Omega,\mathcal{F},P)$ such that

1) $\xi_{Nj}, j=1,\dots,n$, are independent;

2) $\xi_{Nj} \sim F_{Nj}(t) = 1-\exp(-R(c_{Nj}t)),\ j=1,\dots,n$;

3) $\xi_{N(j)} \le \xi_{N-1(j)}$ a.s. for $j=1,\dots,n$.

Now set

$$X'_j = X_j, \qquad Y'_j = \xi_{Nj},\ j=1,\dots,n.$$

Then $X'_j,\ Y'_j,\ j=1,\dots,n$, are just what we desire to construct.

Remark: (i) The result of Theorem 1 for the special case $\mu_1=\mu_2=\cdots\mu_n=\sum_{i=1}^n \mu_i/n$ was given in BLR [3] . Their proof is not so straightforward. They used an approximating method. Using Fact 2, our technique of the proof here is simpler and more powerful.

(ii) Noting Fact 3, the condition (10) in Theorem 1 can be replaced by

$$(\mu_1,\mu_2,\dots,\mu_n) \prec^w (\lambda_1,\lambda_2,\dots,\lambda_n).$$

(The author thanks Prof. Harry Joe for this observation.)

Proschan and Sethuraman [1] considered proportional hazard model. They obtained a result that the vector of the order statistics of one sample is stochastically smaller than that of the other. Here we can prove the following stronger result in a similar fashion to Theorem 1 with obvious modifications, whose proof is omitted. It should be noted that there is no condition imposed on the failure rate $r(t)$.

Theorem 2. Let $X_1,X_2,\dots,X_n;Y_1,Y_2,\dots,Y_n$ be independent nonnegative r.v.'s with distributions

$$\begin{aligned} P(X_i>t) &= \overline{F}_i(t) = \exp(-\lambda_i R(t)); \\ P(Y_i>t) &= \overline{G}_i(t) = \exp(-\mu_i R(t)),\ i=1,\dots,n \end{aligned}$$

where $\lambda_1,\dots,\lambda_n;\mu_1,\dots,\mu_n$ are positive constants and $R(t)=\int_0^t r(s)ds$. Assume that

$$(\mu_1,\mu_2,\dots,\mu_n) \prec^w (\lambda_1,\lambda_2,\dots,\lambda_n).$$

Then the monotone coupling of order statistics in Theorem 1 holds.

We have proved Theorem 1 under the condition that F has a decreasing failure rate $r(x)$, and $xr(x)$ is increasing. But in certain situation, F has a incresing failure rate, does the monotone coupling of order statistics in Theorem 1 still hold? Can the result in Theorem 1 be sharpened in certain way? The answer is affirmative. We delineate it as the following corollary without proof (see BLR [3]).

Corollary 1. Let $X_j, Y_j,\ j = 1, \ldots, n$, be defined in Theorem 1. Assume that $xr(x)$ is increasing, and $x^{1-\alpha}r(x)$ is decreasing for some $\alpha > 0$. If

$$(\mu_1^\alpha, \mu_2^\alpha, \ldots, \mu_n^\alpha) \prec^w (\lambda_1^\alpha, \lambda_2^\alpha, \ldots, \lambda_n^\alpha),$$

then the monotone coupling of order statistics in Theorem 1 holds.

3 Stochastic ordering of order statistics

Before we give a result about stochastic ordering of order statistics under weak condition, we first give a definition of uniform majorization for two real vectors of functions.

Definition 2. Let $\mathbf{f}(t) = (f_1(t), \ldots, f_n(t)), \mathbf{g}(t) = (g_1(t), \ldots, g_n(t)),\ t \in \Lambda$, denote two real vectors of functions. $\mathbf{f}(t)$ is said to majorize $\mathbf{g}(t)$ uniformly in Λ, in symbols, $\mathbf{f}(t) \succ^{u.m.} \mathbf{g}(t),\ t \in \Lambda$, if there exist a finite number of real vectors of functions $\mathbf{h}_j(t),\ j = 1, \ldots, N$, such that

$$\mathbf{f}(t) = \mathbf{h}_1(t) \succ^m \mathbf{h}_2(t) \succ^m \cdots \succ^m \mathbf{h}_N(t) = \mathbf{g}(t) \text{ for all } t \in \Lambda$$

and also such that for all j and $t \in \Lambda$, $n-2$ components of $\mathbf{h}_j(t) - \mathbf{h}_{j+1}(t)$ are zero.

Theorem 3. Suppose that $X_1, \ldots, X_n; Y_1, \ldots, Y_n$ are independent nonnegative r.v.'s, and that

$$\begin{aligned} P(X_i > t) &= \exp(-R_i(t)) \\ P(Y_i > t) &= \exp(-R_i^*(t)), \quad i = 1, \ldots, n. \end{aligned}$$

If $(R_1(t), R_2(t), \ldots, R_n(t)) \succ^{u.m.} (R_1^*(t), R_2^*(t), \ldots, R_n^*(t))$, then

$$X_{(k)} \succ^{st} Y_{(k)} \text{ for } k = 1, \ldots, n.$$

Proof. For any n, $X_{(1)}$ and $Y_{(1)}$ have the same distribution. Also by fact 2

$$\begin{aligned} P(X_{(2)} \geq t) &= \sum_{i=1}^n (1 - \exp(-R_i(t)) \exp(-\sum_{j \neq i} R_j(t)) + \exp(-\sum_{i=1}^n R_i(t)) \\ &= \exp(-\sum_{i=1}^n R_i(t))[-(n-1) + \sum_{i=1}^n \exp(R_i(t))] \\ &\geq \exp(-\sum_{i=1}^n R_i^*(t))[-(n-1) + \sum_{i=1}^n \exp(R_i^*(t))] \\ &= P(Y_{(2)} \geq t). \end{aligned}$$

Now suppose that $n > 2$. By definition, there exist a finite number of real vectors of functions $\mathbf{C}_j(t) = (C_{j1}(t), C_{j2}(t), \ldots, C_{jn}(t)),\ j = 1, \ldots, N$, such that

$$\mathbf{R}(t) = C_1(t) \succ^m C_2(t) \succ^m \cdots \succ^m C_N(t) = R^*(t), \quad t \geq 0,$$

where $\mathbf{R}(t) = (R_1(t), \ldots, R_n(t))$ and $\mathbf{R}^*(t) = (R_1^*(t), \ldots, R_n^*(t))$, and also such that for all $t \geq 0$ and any $j,\ n-2$ components of $C_j(t) - C_{j+1}(t)$ are zero. Suppose that there exist a sequence of independent r.v.'s $X_{ij},\ j = 1, \ldots, n,\ i = 1, \ldots, N$, such that

$$P(X_{ij} > t) = \exp(-C_{ij}(t)),$$

and

$$(X_{11}, X_{12}, \ldots, X_{1n}) =^d (X_1, X_2, \ldots, X_n),$$
$$(X_{N1}, X_{N2}, \ldots, X_{Nn}) =^d (Y_1, Y_2, \ldots, Y_n).$$

Without loss of generality, assume that $C_{1j}(t) = C_{2j}(t),\ j = 1, \ldots, n-2$, then $(C_{1,n-1}(t), C_{1n}(t)) \succ^m (C_{2,n-1}(t), C_{2n}(t))$. Let $\tilde{X}_{1(1)} < \tilde{X}_{1(2)} < \cdots < \tilde{X}_{1(n-2)}$ be the order statistics of the sample $X_{11}, \ldots, X_{1,n-2}$, and $X_{1(1)} < X_{1(2)} < \cdots < X_{1(n)}$ be the order statistics of the sample enlarged by the inclusion of $X_{1,n-1}$ and X_{1n}. So we have that for any $t > 0$ and $k \geq 3$

$$P(X_{1(k)} \leq t) = \begin{cases} P(\tilde{X}_{1(k-1)} \leq t)[P(X_{1n} > t, X_{1,n-1} \leq t) + P(X_{1n} \leq t, X_{1,n-1} > t)] & \\ \qquad +P(\tilde{X}_{1(k)} \leq t)P(X_{1,n-1} > t, X_{1n} > t) & \\ \qquad\quad +P(\tilde{X}_{1(k-2)} \leq t)P(X_{1,n-1} \leq t, X_{1n} \leq t), & \text{if } k \leq n-2; \\ P(\tilde{X}_{1(n-2)} \leq t)[P(X_{1,n-1} \leq t, X_{1n} > t) + P(X_{1,n-1} > t, X_{1n} \leq t)] & \\ \qquad +P(\tilde{X}_{1(n-3)} \leq t)P(X_{1,n-1} \leq t, X_{1n} \leq t), & \text{if } k = n-1; \\ P(\tilde{X}_{1(n-2)} \leq t)P(X_{1,n-1} \leq t, X_{1n} \leq t), & \text{if } k = n. \end{cases}$$

Take the easy case of $k = n-1,\ n$ first. By Fact 2 we have (with $C_{ij}(t)$ abbreviated to C_{ij})

$$\begin{aligned} P(X_{1n} \leq t) &= P(\tilde{X}_{2(n-2)} \leq t)[1 + \exp(-C_{1,n-1} - C_{1n}) - \exp(-C_{1,n-1}) - \exp(-C_{1n})] \\ &\leq P(\tilde{X}_{2(n-2)} \leq t)(1 - \exp(-C_{2,n-1}))(1 - \exp(-C_{2n})) \\ &= P(X_{2(n)} \leq t) \\ P(X_{1(n-1)} \leq t) &= [P(\tilde{X}_{2(n-3)} \leq t) - 2P(\tilde{X}_{2(n-2)} \leq t)] \exp(-C_{1,n-1} - C_{1n}) \\ &\qquad +[P(\tilde{X}_{2(n-2)} \leq t) - P(\tilde{X}_{2(n-3)} \leq t)][\exp(-C_{1,n-1}) + \exp(-C_{1n})] \\ &\qquad\quad +P(\tilde{X}_{2(n-3)} \leq t) \\ &\leq [P(\tilde{X}_{2(n-3)} \leq t) - 2P(\tilde{X}_{2(n-2)} \leq t)] \exp(-C_{2,n-1} - C_{2n}) \\ &\qquad +[P(\tilde{X}_{2(n-2)} \leq t) - P(\tilde{X}_{2(n-3)} \leq t)]\ [\exp(-C_{2,n-1}) + \exp(-C_{2n})] \\ &\qquad\quad +P(\tilde{X}_{2(n-3)} \leq t) \\ &= P(X_{2(n-1)} \leq t), \end{aligned}$$

where the last inequality follows from Fact 2 and $P(\tilde{X}_{2(n-2)} \leq t) \leq P(\tilde{X}_{2(n-3)} \leq t)$.

Next we turn to the case $k \leq n-2$. Using Fact 2 and

$$P(\tilde{X}_{2(k-1)} \leq t) \leq P(\tilde{X}_{2(k-2)} \leq t),$$

it follows that

$$\begin{aligned}
P(X_{1(k)} \le t) &= [P(\tilde{X}_{2(k)} \le t) + P(\tilde{X}_{2(k-2)} \le t) - 2P(\tilde{X}_{2(k-1)} \le t)] \exp(-C_{1,n-1} - C_{1n}) \\
&\qquad +[P(\tilde{X}_{2(k-1)} \le t) - P(\tilde{X}_{2(k-2)} \le t)][\exp(-C_{1,n-1}) + \exp(-C_{1n})] \\
&\qquad\quad +P(\tilde{X}_{2(k-2)} \le t) \\
&\le [P(\tilde{X}_{2(k)} \le t) + P(\tilde{X}_{2(k-2)} \le t) - 2P(\tilde{X}_{2(k-1)} \le t)] \exp(-C_{2,n-1} - C_{2n}) \\
&\qquad +[P(\tilde{X}_{2(k-1)} \le t) - P(\tilde{X}_{2(k-2)} \le t)][\exp(-C_{2,n-1}) + \exp(-C_{2n})] \\
&\qquad\quad +P(\tilde{X}_{2(k-2)} \le t) \\
&= P(X_{2(k)} \le t).
\end{aligned}$$

So we get

$$X_{1(j)} \succ^{st} X_{2(j)}, \quad j = 1, \ldots, n.$$

Similarly, we can go on doing this to prove that for $i = 1, \ldots, N-1$,

$$X_{i(j)} \succ^{st} X_{i+1(j)}, \quad j = 1, \ldots, n.$$

This completes the proof of the theorem. ∎

Monotone coupling of order statistics in Theorem 3 is in general impossible, and a counterexample can be found in BLR [3].

References

[1] Proschan, F. & Sethuraman, J., Stochastic comparison of order statistics from heterogeneous populations, with applications in reliability., *J. Multivariate Anal.* **6**(1976), 608-616.

[2] Ball, F., Deterministic and stochastic epidemic with several kinds of susceptibles., *Adv. Appl. Prob.*, **17**(1985), 1-22.

[3] Barbour, A.D., Lindvall, T., & Rogers, L.C.G., Stochastic ordering of order statistics., *J. Appl. Prob.*, **28**(1991), 278-286.

[4] Marshall, A.W. & Olkin, I., *Inequalities: theory of majorization and its applications,* Academic Press, New York. 1979.

[tex file: mon15.tex]