JOURNAL OF MULTIVARIATE ANALYSIS **56**, 75–89 (1996)
ARTICLE NO. 0005

Bivariate Dependence Properties of Order Statistics

PHILIP J. BOLAND*

University College Dublin Belfield, Dublin 4, Ireland

MYLES HOLLANDER†

The Florida State University

KUMAR JOAG-DEV

University of Illinois

AND

SUBHASH KOCHAR

Indian Statistical Institute, New Delhi-110016, India

If $X_1, ..., X_n$ are random variables we denote by $X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)}$ their respective order statistics. In the case where the random variables are independent and identically distributed, one may demonstrate very strong notions of dependence between any two order statistics $X_{(i)}$ and $X_{(j)}$. If in particular the random variables are independent with a common density or mass function, then $X_{(i)}$ and $X_{(j)}$ are TP_2 dependent for any i and j. In this paper we consider the situation in which the random variables $X_1, ..., X_n$ are independent but otherwise arbitrarily distributed. We show that for any $i<j$ and t fixed, $P[X_{(j)} > t \mid X_{(i)} > s]$ is an increasing function of s. This is a stronger form of dependence between $X_{(i)}$ and $X_{(j)}$ than that of association, but we also show that among the hierarchy of notions of bivariate dependence this is the strongest possible under these circumstances. It is also shown that in this situation, $P[X_{(j)} > t \mid X_{(i)} > s]$ is a decreasing function of $i = 1, ..., n$ for

Received February 8, 1994; revised May 1995.

AMS subject classification number: 62G30, 60E15, 62D05.

Keywords and phrases: order statistics, associated random variables, right corner set increasing, right tail increasing, stochastically increasing, TP_2 property, positively quadrant dependent, reliability, counting processes, exchangeability, finite sampling problem, k out of n system.

* Research supported in part by AFOSR Grant No. 91-0048.

† Research supported by AFOSR Grant No. 91-0048.

0047-259X/96 $12.00

any fixed $s<t$. We give various applications of these results in reliability theory, counting processes, and estimation of conditional probabilities. We also consider the situation where $X_1, ..., X_n$ represent a random sample of size n drawn without replacement from a linearly ordered finite population. In this case it is shown that $X_{(i)}$ and $X_{(j)}$ are TP_2 dependent for any i and j, and the implications are discussed.

1. Introduction

There are several well-known notions of positive dependence between two random variables. There is a large literature on this topic with important contributions by Lehmann [16], Esary and Proschan [9], Harris [10], Barlow and Proschan [4], and Shaked [19], amongst others. The reader would also do well to see Jogdeo [11] and Block, Sampson, and Savits [6] for further discussions. Perhaps the strongest notion of dependence between random variables S and T is that of TP_2 dependence, a concept developed extensively by Karlin [12]. S and T are **$\mathbf{TP_2}$ dependent** if their joint density $f(s, t)$ is totally positive of order 2 in s and t, or more precisely if

$$\begin{vmatrix} f(s_1, t_1) & f(s_1, t_2) \\ f(s_2, t_1) & f(s_2, t_2) \end{vmatrix} \geqslant 0$$

whenever $s_1 < s_2$, $t_1 < t_2$. The random variables S and T are **right corner set increasing (RCSI)** if for fixed s and t, $P[S>s, T>t \mid S>s', T>t']$ is increasing in s' and t'. We will say that a function is increasing (decreasing) if it is nondecreasing (nonincreasing). Shaked [19] and Kochar and Deshpande [15] have noted that S and T are RCSI if and only if $r(t \mid S>s)$ is decreasing in s for every fixed t, where $r(t \mid S>s)$ denotes the hazard rate of the conditional distribution of T given $S>s$. We say that T is **stochastically increasing** in S if $P(T>t \mid S=s)$ is increasing in s for all t, and write **SI(T | S)**. Shaked [19] has shown that SI$(T \mid S)$ if and only if $R(t \mid S=s)$ is decreasing in s, where R represents a cumulative hazard (failure rate) function. Lehmann [16] uses the term **positively regression dependent** to describe SI. If $P(T>t \mid S=s)$ is decreasing we say that T is stochastically decreasing in S (written SD$(T \mid S)$) or that T is negatively regression dependent on S.

We say that T is **right tail increasing** in S if $P[T>t \mid S>s]$ is increasing in s for all t, and denote this relationship by **RTI(T | S)**. If S and T are continuous lifetimes, then T is right tail increasing in S if $r(s \mid T>t) \leqslant r(s \mid T>0) = r_S(s)$ for all $s>0$ and for each fixed t. We say T is **left tail increasing** in S if $P[T<t \mid S<s]$ is decreasing in s for all t, and we

denote this by **LTD(T | S)**. Barlow and Proschan [4] define **LTD(T | S)** to mean $P[T \leqslant t \mid S \leqslant s]$ is decreasing in s for all t, and, of course, this is equivalent to our definition for continuous random variables. If $P(T > t \mid S > s)$ is decreasing in s we say that T is right tail decreasing in S and write RTD($T \mid S$). Similarly LTI($T \mid S$) means T is left tail increasing in S. The random variables S and T are **associated** (written A(S, T)) if $\mathrm{Cov}[\Gamma(S, T), \Delta(S, T)] \geqslant 0$ for all pairs of increasing binary functions Γ and Δ. We say S and T are **positively quadrant dependent** if

$$P[S \leqslant s, T \leqslant T] \geqslant P[S \leqslant s] P[T \leqslant t]$$

for all s, t, and we write **PQD(S, T)**. The various implications between these notions of positive dependence, at least for continuous random variables, are summarised by Fig. 1 from Barlow and Proschan [4] (see also Chap. 5 of Tong [21]).

Let $X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$ denote the order statistics of the random variables $X_1, ..., X_n$. There is an extensive literature on the dependence structure of the order statistics $X_{(i)}$ and $X_{(j)}$. Tukey [22] showed (see also Kim and David [14]) that if the X_i's are independent with common distribution function which is "subexponential" in both tails, then the covariance of $X_{(i)}$ and $X_{(j)}$ decreases as i and j draw apart. Bickel [5] proved that in the independently distributed case with common density, $\mathrm{Cov}(X_{(i)}, X_{(j)}) \geqslant 0$ for all i and j. Of course, when $X_1, ..., X_n$ are independent (but not necessarily identically distributed) the order statistics $(X_{(1)}, ..., X_{(n)})$ are associated. This yields many useful product inequalities for order statistics of independent random variables and, in particular, implies that in this case $\mathrm{Cov}(X_{(i)}, X_{(j)}) \geqslant 0$ for all i and j (see also Lehmann

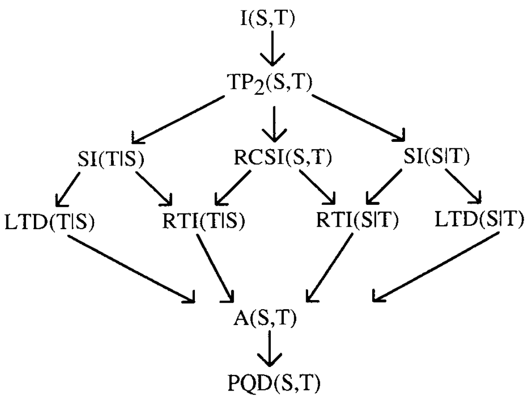

FIG. 1. Implications among notions of bivariate dependence.

[16]). Recent developments concerning the stochastic comparison of order statistics are reviewed in Kim, Proschan, and Sethuraman [13] (where comparisons of individual order statistics and vectors of order statistics from underlying heterogeneous distributions are treated with the use of majorization and Schur functions) and in the extensive text "Stochastic Orders and Their Applications" by Shaked and Shanthikumar [20]. Other interesting results concerning the log concavity property of the sequence of distribution functions of the ith-order statistics from underlying heterogeneous distributions are treated in the series of papers by Bapat and Beg [3], Sathe and Bendre [18], and Balasubramanian and Balakrishnan [2].

It seems natural to expect some degree of positive dependence between the order statistics $X_{(i)}$ and $X_{(j)}$ based on the random variables $X_1, ..., X_n$. In the independent identically distributed case (with a common density or mass function) one may readily verify that the order statistics $X_{(i)}$ and $X_{(j)}$ are TP_2 dependent for any i and j. To see this in the continuous case note that for $i<j$, the joint density of $X_{(i)}$ and $X_{(j)}$ is given by

$$f_{(i),(j)}(s, t) = \begin{cases} c(i, j, n) F^{i-1}(s)[F(t)-F(s)]^{j-i-1}[\bar{F}(t)]^{n-j} f(s) f(t), & s \leqslant t, \\ 0, & \text{otherwise}, \end{cases}$$

where $c(i, j, n)$ is a normalizing constant. Algebraically it is easy to show that $f_{(i),(j)}(s_1, t_1) f_{(i),(j)}(s_2, t_2) \geqslant f_{(i),(j)}(s_1, t_2)\, f_{(i),(j)}(s_2, t_1)$ for any $s_1 < s_2$ and $t_1 < t_2$ is equivalent to

$$[F(t_1)-F(s_1)][F(t_2)-F(s_2)] \geqslant [F(t_2)-F(s_1)][F(t_1)-F(s_2)],$$

which is always true. Hence for the case where $X_1, ..., X_n$ are independent identically distributed and continuous (or discrete), we have that $X_{(i)}$ and $X_{(j)}$ are positively dependent in any of the senses discussed above. Of course in the independent identically distributed case the order statistics have the Markov property implying another type of dependence (see Chap. 1 of David [7] or Chap. 1 of Arnold *et al.* [1]). These results suggest that perhaps similar notions of dependence might hold for the order statistics $X_{(i)}$ and $X_{(j)}$ based on a set of independent (but not necessarily identically distributed) random variables $X_1, ..., X_n$.

In Section 2 we investigate the RCSI and SI properties of the order statistics $X_{(i)}$ and $X_{(j)}$ in this situation. We give an example where $n=2$ and X_1, X_2 are exponentially distributed with different hazard rates yet $X_{(1)}$ and $X_{(2)}$ are not RCSI. Another example shows that $X_{(2)}$ is not necessarily SI in $X_{(1)}$, although we also prove that $X_{(i)}$ is SI in $X_{(1)}$ for any $i=2, ..., n$

whenever $X_1, \ldots, X_n$ are independent with proportional hazards on a common support. In Section 3 we prove the very general result that when $X_1, \ldots, X_n$ are independent but with **arbitrary** distributions, then $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ for any $1 \leqslant i < j \leqslant n$. A duality argument implies therefore that $\mathrm{LTD}(X_{(i)} \mid X_{(j)})$ holds for any $1 \leqslant i < j \leqslant n$. In the process of proving the RTI property of these order statistics, we also show that $P(X_{(j)} > t \mid X_{(i)} > s)$ is a decreasing function of $i = 1, \ldots, n$ for any fixed $s \leqslant t$. Applications of these results are given in reliability theory, counting processes, and estimation of conditional probabilities.

In Section 4 we discuss the dependence of $X_{(i)}$ and $X_{(j)}$ when $X_1, \ldots, X_n$ are identically distributed but not independent. An elementary example shows that even when $n = 2$ it does not necessarily follow that $\mathrm{RTI}(X_{(2)} \mid X_{(1)})$. We also consider however the important special case when $(X_1, \ldots, X_n)$ represents a random sample taken without replacement from a finite linearly ordered population (with possible replications). We show that in this case $X_{(i)}$ and $X_{(j)}$ are TP_2 dependent, and give applications to the finite sampling problem.

2. RCSI and SI Properties of Order Statistics

The concept of RCSI dependence was introduced by Harris [10] in order to define multivariate increasing failure (hazard) rate functions. The following example shows, however, that in general $X_{(i)}$ and $X_{(j)}$ are not RCSI (and, hence, implicitly not TP_2 dependent).

Example 2.1. Let $n = 2$, and suppose X_1, X_2 are independent and exponentially distributed with means 1 and 0.5, respectively. Then $X_{(1)}$ and $X_{(2)}$ have a joint survival function given by

$$\begin{aligned}\bar{F}_{(1),(2)}(s, t) &= \bar{F}_1(s)\bar{F}_2(t) + \bar{F}_1(t)\bar{F}_2(s) - \bar{F}_1(t)\bar{F}_2(t)\\ &= e^{-(\lambda_1 s + \lambda_2 t)} + e^{-(\lambda_1 t + \lambda_2 s)} - e^{-(\lambda_1 + \lambda_2)t} \qquad \text{for} \quad s < t.\end{aligned}$$

Now $X_{(1)}$ and $X_{(2)}$ are RCSI implies in particular that $\mathrm{Prob}[X_{(1)} > s, X_{(2)} > t \mid X_{(1)} > s, X_{(2)} > t'] = \bar{F}_{(1),(2)}(s, t)/\bar{F}_{(1),(2)}(s, t') \uparrow$ in s for any $t' < t$. However,

$$\bar{F}_{(1),(2)}(0, 5)/\bar{F}_{(1),(2)}(0, 2) = 0.0449 > 0.0411 = \bar{F}_{(1),(2)}(1, 5)/\bar{F}_{(1),(2)}(1, 2),$$

and, hence, $(X_{(1)}, X_{(2)})$ are not RCSI, even in the exponentially distributed case. ∎

Although $X_{(2)}$ and $X_{(1)}$ are not RCSI in Example 2.1, the following result shows that in this case is $\mathrm{SI}(X_{(2)} \mid X_{(1)})$.

THEOREM 2.2. *Let $X_1, ..., X_n$ be independent random variables with differentiable densities and proportional hazard functions on an interval I (where the support of $X_i = I$ for each i). Then $X_{(i)}$ is SI in $X_{(1)}$.*

Proof. Let f_i and r_i denote, respectively, the density function and the hazard rate function of X_i for $i = 1, ..., n$. From the proportional hazards assumption, there exist positive constants α_i such that $r_i(x) = \alpha_i r_1(x)$ for $i = 1, ..., n$ ($\alpha_1 = 1$).

For $s < t$ in the interval of common support, let us denote by $H_{ij}(s, t)$ the probability that at least $n - j + 1$ of the X's in $\{X_1, ..., X_{i-1}, X_{i+1}, ..., X_n\}$ are greater than t and the others lie in the interval $(s, t]$. Then we have

$$\begin{aligned} P[X_{(j)} > t \mid X_{(1)} = s] &= \frac{\sum_{i=1}^n f_i(s)\, H_{ij}(s, t)}{\sum_{i=1}^n f_i(s) \prod_{k \neq i} \bar{F}_k(s)} \\ &= \frac{\sum_{i=1}^n \alpha_i \bar{F}_i(s)\, H_{ij}(s, t)}{(\sum_{i=1}^n \alpha_i) \prod_{k=1}^n \bar{F}_k(s)} \\ &= \frac{1}{\sum_{i=1}^n \alpha_i} \sum_{i=1}^n \alpha_i C_{ij}(s, t), \end{aligned}$$

where $C_{ij}(s, t)$ is the probability that at least $(n - j + 1)$ out of $(n - 1)$ events occur where the probability of the kth event is $p_k = \bar{F}_k(t)/\bar{F}_k(s)$, which is nondecreasing in s. Hence $C_{ij}(s, t)$ is a nondecreasing function of s for $s < t$ and the required result follows. ▌

A similar argument would show that if $X_1, ..., X_n$ have the property that $f_1(s)/F_1(s), ..., f_n(s)/F_n(s)$ are proportional then $X_{(i)}$ is SI in $X_{(n)}$. We now give an example illustrating that, in general, $X_{(2)}$ is not SI in $X_{(1)}$.

EXAMPLE 2.3. Let X_1 and X_2 be independent random variables which are uniformly distributed on $A_1 = (0, 3) \cup (10, 13)$ and $A_2 = (2, 5) \cup (12, 15)$, respectively. Then

$$P(X_{(2)} > 13 \mid X_{(1)} = 1) = \tfrac{1}{3} > P(X_{(2)} > 13 \mid X_{(1)} = 4) = 0$$

and, hence, $X_{(2)}$ is not SI in $X_{(1)}$.

Despite the partial results on RCSI and SI for order statistics presented in this section, we will show in Section 3 that $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ holds, in general, for any $i < j$ when $X_1, ..., X_n$ are independent random variables.

3. Right Tail Increasing Property of Order Statistics

We now consider the dependence relation RTI of $X_{(j)}$ and $X_{(i)}$. Note that

$$\begin{aligned}
&\mathrm{RTI}(X_{(j)} \mid X_{(i)}) \\
&\quad \Leftrightarrow P[X_{(j)} > t \mid X_{(i)} > s] \uparrow s \qquad \text{for any fixed } t \\
&\quad \Leftrightarrow P[-[X_{(j)}] < -t \mid -[X_{(i)}] < -s] \uparrow s \qquad \text{for any fixed } t \\
&\quad \Leftrightarrow P[[-X]_{(n-j+1)} < -t \mid [-X]_{(n-i+1)} < s] \downarrow s \qquad \text{for any fixed } t.
\end{aligned}$$

Hence $X_{(j)} \mathrm{RTI} X_{(i)} \Leftrightarrow [-X]_{(n-j+1)} \mathrm{LTD} [-X]_{(n-i+1)}$, where $[-X]_{(k)}$ represents the kth-order statistic of $\{-X_1, ..., -X_n\}$. Therefore $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ holds in general (that is, for any set of n independent random variables), if and only if $\mathrm{LTD}(X_{(n-j+1)} \mid X_{(n-i+1)})$ holds in general. We will see in this section that $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ holds in general for all $1 \leqslant i < j \leqslant n$ and, hence equivalently, that $\mathrm{LTD}(X_{(i)} \mid X_{(j)})$ for all $1 \leqslant i < j \leqslant n$.

If $X_1, ..., X_n$ are independent random variables (with possibly different distributions) we denote by $N(u)$ the counting variable representing the number of observations in $\{X_1, ..., X_n\}$ which are less than or equal to u. It is easy to see that the events $[X_{(j)} > t]$ and $[N(t) < j]$ are identical. For $i < j$ we want to show that $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ or that $P[X_{(j)} > t \mid X_{(i)} > s] \uparrow$ in s for any fixed t. Since $i < j$, we need only consider $s < t$, and, therefore, this is equivalent to showing that

$$P[N(t) < j \mid N(s_2) < i] \geqslant P[N(t) < j \mid N(s_1) < i] \tag{3.1}$$

for every $s_1 < s_2 < t$. For the sake of brevity we let $N_1 = N(s_1)$, $N_2 = N(s_2) - N(s_1)$, $N_3 = N(t) - N(s_2)$, $N_4 = n - N(t)$, and, hence, the inequality (3.1) becomes

$$P[N_4 > n-j \mid N_1 + N_2 < i] \geqslant P[N_4 > n-j \mid N_1 < i]. \tag{3.2}$$

This inequality may be restated in terms of what one may call a **generalized multinomial distribution** (G-multinomial). Here we have n independent trials, and each trial consists of putting a "ball" into one of four boxes. The probabilities of placement into the four boxes may vary from one trial to the next (corresponding to the different distributions of $X_1, ..., X_n$), and N_l (for $l = 1, 2, 3$, and 4) will denote the total number of balls placed in box l after the n trials. Using this notation for the generalized multinomial distribution, we may interpret (3.2) in the following way: The probability that the number of balls in box 4 exceeds $n-j$, knowing that the total number in boxes 1 and 2 is less than i **is greater than or equal to** the probability that the number of balls in box 4 exceeds $n-j$, knowing only that the total number in box 1 is less than i. Given the negative dependence between the

N_l's ($l=1, 2, 3$, and 4), this result seems most plausible. The following lemma will be useful in completing the proof.

LEMMA 3.1. *Let (N_1, N_2, N_3, N_4) have the generalized multinomial distribution as above. Then*

(a) $\mathrm{SD}(N_4 \mid N_1)$, *or equivalently, $P[N_4<k \mid N_1=l]$ is nondecreasing in l for any $k=1, ..., n$ and*

(b) $\mathrm{LTI}(N_4 \mid N_1)$, *or equivalently, $P[N_4 \geqslant k \mid N_1<l]$ is decreasing in l for any $k=0, 1, ..., n$.*

Proof. Let U_i be the indicator of the event that the ith ball goes to Box 1. Then U_i, $i=1, ..., n$, are independent random variables each having a logconcave (also known as Polya frequency of type 2) density function. According to a theorem of Efron [8], if h is a real function which is nondecreasing in each U_i then

$$E[h(U_1, ..., U_n) \mid \Sigma U_i] \text{ is nondecreasing in } \Sigma U_i.$$

For fixed k, $1 \leqslant k \leqslant n$, define $h(U_1, ..., U_n)=P[N_4<k \mid U_1, ..., U_n]$. When $U_1, ..., U_n$ are given, the distribution of N_4 can be viewed as that obtained by performing independent trials with the balls for which $U_l=0$. For a particular j, if U_j is increased from 0 to 1, keeping U_i fixed for $i \neq j$, it is seen that h will increase since the balls used in the independent trials would be the same as before except for the j^{th} ball. Thus by Efron's theorem

$$E[h(U_1, ..., U_n) \mid \Sigma U_i]=P[N_4<k \mid \Sigma U_i] \text{ is nondecreasing in } \Sigma U_i.$$

But $N_1=\Sigma U_i$ and, hence, the assertion (a) follows.

We have already seen (see Fig. 1) that the bivariate notion of SI implies that of LTD and, therefore, in a similar way $\mathrm{SD} \Rightarrow \mathrm{LTI}$. Hence, part (b) follows from (a). ∎

COROLLARY 3.2. *Let $X_1, ..., X_n$ be independent random variables. Then for any $s<t$ and j given*

(a) *$P[X_{(j)}>t \mid X_{(i)}>s]$ is a decreasing function of i and*

(b) *$P[X_{(j)}<s \mid X_{(i)}<t]$ is an increasing function of i.*

Proof. (a) Letting $s_1=s$ we have from part (b) of Lemma 3.1 that

$$P[N_4>n-j \mid N_1<i]=P[N(t)<j \mid N(s)<i]=P[X_{(j)}>t \mid X_{(i)}>s]$$

is decreasing in i. (b) follows in a similar fashion. ∎

We will now establish the validity of (3.2). In the sequence of n trials we arbitrarily select one trial (without loss of generality the last). For each $l=1, 2, 3, 4$ we decompose $N_l = N'_l + N''_l$, where N'_l refers to the first $(n-1)$ trials and N''_l to the last trial. Wet let $p_l = P[N''_l = 1]$ for $l = 1, 2, 3, 4$. Then with the above notation we have the following.

LEMMA 3.3. *For any pair of integers* (k, l), $k, l \geqslant 1$,

$$P[N'_4 + N''_4 \geqslant k \mid N'_1 + N''_1 + N''_2 < l] \geqslant P[N'_4 + N''_4 \geqslant k \mid N'_1 + N''_1 < l],$$

or equivalently,

$$P[N'_4 + N''_4 \geqslant k, N'_1 + N''_1 + N''_2 < l] P[N'_1 + N''_1 < l] \geqslant P[N'_4 + N''_4 \geqslant k, N'_1 + N''_1 < l] P[N'_1 + N''_1 + N''_2 < l]. \tag{3.3}$$

Proof. We use the notation $A(k, l) = P[N'_4 \geqslant k, N'_1 < l]$ and $B(l) = P[N'_1 < l]$. From part (b) of Lemma 3.1, we have

$$A(k, l-1)B(l) - A(k, l)B(l-1) \geqslant 0. \tag{3.4}$$

By assigning the value 1 successively to each of N''_1, N''_2, N''_3, and N''_4 on both sides of (3.3) we get

$$\{p_4 A(k-1, l) + (p_1 + p_2)A(k, l-1) + (1 - p_1 - p_2 - p_4)A(k, l)\} \times \{(1 - p_1)B(l) + p_1 B(l-1)\} \tag{3.5}$$

on the left hand side and

$$\{p_4 A(k-1, l) + p_1 A(k, l-1) + (1 - p_1 - p_4)A(k, l)\} \times \{(1 - p_1 - p_2)B(l) + (p_1 + p_2)B(l-1)\} \tag{3.6}$$

on the right side.

After carrying out the multiplications and subtracting (3.6) from (3.5), the resulting expression contains only two types of terms: those having factor $p_2 p_4$ and others having factor p_2. The expression obtained by gathering the terms having factor $p_2 p_4$ is

$$\{A(k-1, l)B(l) - A(k-1, l)B(l-1) + A(k, l)B(l-1) - A(k, l)B(l)\},$$

which reduces to

$$A(k-1, l)P[N'_1 = l-1] - A(k, l)P[N'_1 = l-1] = P[N'_4 = k-1, N'_1 < l]P[N'_1 = l-1],$$

and, hence, is nonnegative. The expression with factor p_2 coincides with the left side of (3.4) and is already known to be nonnegative. This establishes (3.3) and, hence, Lemma 3.3. ▮

THEOREM 3.4. *Let $X_1, ..., X_n$ be independently distributed random variables. Then for any $i<j$,* $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ *and* $\mathrm{LTD}(X_{(i)} \mid X_{(j)})$.

Proof. The result of Lemma 3.3 may be interpreted as follows. For a given generalized multinomial (N_1, N_2, N_3, N_4) if the probability vector $(p_{i1}, p_{i2}, p_{i3}, p_{i4})$ of outcomes for the ith trial is replaced by the vector $(p_{i1}+p_{i2}, 0, p_{i3}, p_{i4})$ then the distribution of N_4 given $N_1 < l$ becomes stochastically larger. Doing this successively for all "trials" yields the desired inequality (3.2). ▮

Having established that $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ holds in general when $i<j$, one might naturally ask if one may also show that $\mathrm{RTI}(X_{(i)} \mid X_{(j)})$? The following example shows that this is not true in general.

EXAMPLE 3.5. Let X_1 take the values $\frac{1}{2}$ or 1 with probability $\frac{1}{2}$, and let X_2 be uniform on [0,1]. Assuming X_1 and X_2 are independent,

$$P\left(X_{(1)} > \frac{1}{4} \,\middle|\, X_{(2)} > s\right) = \frac{7-4s}{8-4s} \downarrow s \qquad \text{for} \quad s \in \left[\frac{1}{2}, 1\right].$$

Hence $X_{(1)}$ is not RTI in $X_{(2)}$.

We now give some applications of Corollary 3.2 and Theorem 3.4.

EXAMPLE 3.6. *Application to counting processes.* Let $X_1, ..., X_n$ be n independent lifetimes, and for any s let $N(s)$ be the number of lifetime observations $\leqslant s$. Then $N(s) \leqslant i-1 \Leftrightarrow X_{(i)} > s$. Corollary 3.2 and Theorem 3.4 imply that for any j and t, $P[N(t) < j \mid N(s) < i]$ is an increasing function of $s<t$ (for any fixed $i<j$) and a decreasing function of $i = 1, ..., n$ (for any fixed $s<t$). In particular suppose $X_1, ..., X_{30}$ represent the independent lifetimes of 30 patients in a clinical study and we are interested in the probability that at least 10 of them survive 5 years (equivalently, that $N(5) \leqslant 20$). Then $P[N(5) \leqslant 20 \mid N(s) \leqslant i]$ is increasing over $s \in [0, 5]$ for any fixed $i<20$ and decreasing as a function of $i \in \{1, 2, ..., 20\}$ for any fixed $s<5$. In terms of censoring s might represent the time at which observations in the study are censored, or in other cases censoring might take place after the i^{th} observed lifetime.

EXAMPLE 3.7. *Application in reliability.* A system of n components which functions if and only if at least k of the n components function is called a ***k* out of *n* system**. Parallel systems and series systems are elementary examples and, generally speaking, k out of n systems form the building blocks of many more complex systems. Suppose we are interested in the probability of a k out of n system with independent components surviving a given mission time t_0 and that we periodically inspect the system at times $s_1, s_2, ..., s_m$. Suppose that over these successive inspection times we are able to observe that for some $i > k$, $X_{(n-i+1)} > s_l$ for $l = 1, ..., m$. In such a situation we note that the number of working components i at any inspection point exceeds k, and, hence, the mission time is more and more likely to be achieved since $P[X_{(n-k+1)} > t_0 \mid X_{(n-i+1)} > s_l]$ is increasing in $l = 1, ..., m$.

EXAMPLE 3.8. *Application to an estimation problem.* Let $\mathbf{X} = (X_1, ..., X_n)$ be a vector of independent random variables, but where the distributions of the X_i's might vary. There are situations (see the previous example) where it might be desirable to estimate $P[X_{(j)} > t \mid X_{(i)} > s]$ as a function of s for some (or several-fixed values of t and $i < j$. Theorem 3.4 states that in fact $P[X_{(j)} > t \mid X_{(i)} > s]$ is an increasing function of s. This result is useful therefore in identifying inadmissible estimators of $g(s) = P[X_{(j)} > t \mid X_{(i)} > s]$. Let us suppose, for example, that we observe k independent copies of $\mathbf{X}$, say $\mathbf{X}^1, ..., \mathbf{X}^k$, and from that sample we wish to estimate $P[X_{(j)} > t \mid X_{(i)} > s]$. In fact let us consider the following particular realizations where $n = 6 = k$:

$$\mathbf{X}^1 = (1.4, 1.3, 3.8, 8.2, 6.0, 9.6)$$

$$\mathbf{X}^2 = (2.2, 2.1, 1.2, 6.3, 9.2, 13.8)$$

$$\mathbf{X}^3 = (1.4, 1.3, 1.2, 7.5, 8.6, 6.8)$$

$$\mathbf{X}^4 = (1.6, 1.7, 1.9, 1.8, 9.0, 8.0)$$

$$\mathbf{X}^5 = (1.7, 1.9, 1.8, 5.1, 10.0, 9.0)$$

$$\mathbf{X}^6 = (1.9, 1.8, 2.0, 4.1, 12.0, 13.0).$$

Then a natural estimator of $P[X_{(4)} > 4 \mid X_{(1)} > 1]$ is

$$\hat{P}[X_{(4)} > 4 \mid X_{(1)} > 1] = \frac{\text{number of } \mathbf{X}\text{'s for which } X_{(4)} > 4 \text{ and } X_{(1)} > 1}{\text{number of } \mathbf{X}\text{'s for which } X_{(1)} > 1} = \frac{5}{6}.$$

Similarly, a natural estimator of $P[X_{(4)} > 4 \mid X_{(1)} > 1.5]$ is

$$\hat{P}[X_{(4)} > 4 \mid X_{(1)} > 1.5] = \tfrac{2}{3}.$$

Hence, although in theory $P[X_{(4)} > 4 \mid X_{(1)} > 1] \leqslant P[X_{(4)} > 4 \mid X_{(1)} > 1.5]$, the "natural" estimators of these quantities display the reverse order. Hence the "natural" estimation scheme here is inadmissible. This gives rise therefore to the question: How should the estimators be adjusted so that they satisfy the monotonicity properties of the parameters being estimated? Although we have not answered this question, our result helps to identify inadmissible schemes.

4. Order Statistics from a Linearly Ordered Finite Population

It is natural to ask to what extent one may relax the independence assumption in Theorem 3.4 and still have the property that $X_{(j)}$ is right tail increasing in $X_{(i)}$ for $i<j$. Even if one restricts attention to the case where $X_1, ..., X_n$ have a permutation invariant distribution function, this is not the case and one may easily construct counterexamples. This may also be seen by making use of an observation of Kim and David [14]. They noted that if X_1 and X_2 are nonconstant random variables but where X_1+X_2 is constant, then $0=V(X_1+X_2)=V(X_{(1)}+X_{(2)})=V(X_{(1)})+V(X_{(2)})+2\mathrm{Cov}(X_{(1)}, X_{(2)})$ and hence, $\mathrm{Cov}(X_{(1)}, X_{(2)})<0$. Now $\mathrm{RTI}(X_{(2)} \mid X_{(1)}) \Rightarrow X_{(2)}$ and $X_{(1)}$ are associated $\Rightarrow \mathrm{Cov}(X_{(2)}, X_{(1)}) \geqslant 0$, and, therefore, here we cannot have $\mathrm{RTI}(X_{(2)} \mid X_{(1)})$.

We now consider the important example of sampling without replacement from a finite population which is linearly ordered. We begin with the situation in which there is no replication in the population, and, hence, without loss of generality the population is $\{1, ..., N\}$.

Proposition 4.1. *Let $X_1, ..., X_n$ represent the observations from a simple random sample of size n drawn without replacement from $\{1, ..., N\}$. Then for* **any** *i and j, the joint mass function $f_{(i),(j)}(s, t)$ of $(X_{(i)}, X_{(j)})$ is TP_2 in s and t. Hence, in particular* $\mathrm{SI}(X_{(j)} \mid X_{(i)})$ $(P[X_{(j)}>t \mid X_{(i)}=s] \uparrow s)$ *and* $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ $(P[X_{(j)}>t \mid X_{(i)}>s] \uparrow s)$.

Proof. Without loss of generality we assume $i<j$. Then

$$f_{(i),(j)}(s, t)=P[X_{(i)}=s, X_{(j)}=t]=\begin{cases} \dfrac{\binom{s-1}{i-1}\binom{t-s-1}{j-i-1}\binom{N-t}{n-j}}{\binom{N}{n}}, & s<t, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, the joint mass function is TP_2 in s and t if and only if $s_1<s_2$, $t_1<t_2 \Rightarrow$

$$\binom{t_1-s_1-1}{j-i-1}\binom{t_2-s_2-1}{j-i-1} \geqslant \binom{t_1-s_2-1}{j-i-1}\binom{t_2-s_1-1}{j-i-1}. \tag{4.1}$$

To demonstrate this it clearly suffices to show that in general the function $g(s, t) = \binom{t-s}{k}$ is TP_2 in s and t for any k, or, equivalently, that the function $h(t) = \binom{t}{k}$ is PF_2 (logconcave) in t for any k. Note that it is easy to check that $\binom{t+1-(s+1)}{k}\binom{t-s}{k} \geqslant \binom{t+1-s}{k}\binom{t-(s+1)}{k}$, and the TP_2 property of $g(s, t)$ follows. The interested reader should consult Karlin [12]. ∎

We note that in the sampling problem from $\{1, ..., N\}$, the order statistics have the Markov property. It is also easy to show that X_i and X_j are negatively correlated and $\mathrm{Corr}(X_i, X_j) = \mathrm{Corr}(X_1, X_2) = -1/N$. However, the above proposition implies that $X_{(i)}$ and $X_{(j)}$ are in particular associated for any i and j, and, hence, $\mathrm{Corr}(X_{(i)}, X_{(j)}) \geqslant 0$. In fact, Arnold *et al.* [1] show that $\mathrm{Corr}(X_{(i)}, X_{(j)}) = \sqrt{(i(n-j+1))/j(n-i+1)}$.

Example 4.2. LOTTO. Let $X_1, ..., X_6$ represent the numbers selected from $\{1, ..., 49\}$ in a realization of LOTTO 49. Then Proposition 4.1 implies, in particular, that $P[X_{(6)} > t \mid X_{(1)} > s]$ and $P[X_{(6)} > t \mid X_{(1)} = s]$ are increasing functions of s for any fixed t.

We now extend Proposition 4.1 to the situation when one samples from a linearly ordered population with replication.

Theorem 4.3. *Let $(X_1, ..., X_n)$ represent the observations in a simple random sample from a linearly ordered population with possible repetition. Then for any i and j, $X_{(i)}$ and $X_{(j)}$ have a TP_2 joint mass function.*

Proof. If (S, T) have TP_2 joint mass function, then it is easy to see that $(g(S), h(T))$ also has TP_2 joint mass function for any nondecreasing functions g and h. To see this suppose $s_1 < s_2$ and $t_1 < t_2$. Then

$$
\begin{aligned}
&f_{g(S),\, h(T)}(s_2, t_2) f_{g(S),\, h(T)}(s_1, t_1) \\
&\quad = \sum_{\substack{(u_2,\, v_2) \\ g(u_2) = s_2,\, h(v_2) = t_2}} f_{S,T}(u_2, v_2) \sum_{\substack{(u_1,\, v_1) \\ g(u_1) = s_1,\, h(v_1) = t_1}} f_{S,T}(u_1, v_1) \\
&\quad \geqslant \sum_{\substack{(u_1,\, v_2) \\ g(u_1) = s_1,\, h(v_2) = t_2}} f_{S,T}(u_1, v_2) \sum_{\substack{(u_2,\, v_1) \\ g(u_2) = s_2,\, h(v_1) = t_1}} f_{S,T}(u_2, v_1) \\
&\quad = f_{g(S),\, h(T)}(s_1, t_2) f_{g(S),\, h(T)}(s_2, t_1).
\end{aligned}
$$

Without loss of generality we can assume the population consists of the numbers $1, ..., M$, where there are m_i (>0) values equal to i for $i = 1, ..., M$ and $N = m_1 + \ldots + m_M$ is the size of the population. Then the order statistics $X_{(i)}$ and $X_{(j)}$ based on a sample of size n from our population (with repetitions) are nondecreasing functions of the order statistics $Y_{(i)}$ and $Y_{(j)}$ based on a random sample of size n from the population (without repetitions) $\{1, ..., N\}$ and, hence, the result follows from Theorem 4.1 and the above observation. ▮

COROLLARY 4.4. *Let $X_1, ..., X_n$ represent a simple random sample drawn without replacement from a linearly ordered finite population. Then for any i and j, $P[X_{(j)} > t \mid X_{(i)} = s]$ and $P[X_{(j)} > t \mid X_{(i)} > s]$ are increasing in s.*

EXAMPLE 4.5. A simple random sample of 100 individuals is to be selected from the residents in a city and the resulting ages observed. Suppose we are interested in the chances of at least 10 octogenarians in the sample. Then $P(X_{(91)} > 79 \mid X_{(i)} > s)$ and $P(X_{(91)} > 79 \mid X_{(i)} = s)$ are increasing functions of s for any i.

The finite sampling problem discussed here provides an important example where $X_1, ..., X_n$ are identically distributed (but not independent), yet $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ for any i and j. We have seen, however, that a permutation symmetric distribution function for $(X_1, ..., X_n)$ is not a sufficient condition to imply that $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$. Other examples exist where despite dependence between $X_1, ..., X_n$, it does follow that $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$. It might be worth noting that, even though the original observations might be dependent, if they are associated then so are the order statistics $(X_{(1)}, ..., X_{(n)})$.

EXAMPLE 4.6. The random vector (T_1, T_2) has the Marshall–Olkin (see Marshall and Olkin [17] or Chap. 5 of Barlow and Proschan [4]) bivariate exponential distribution if the joint survival function is given by

$$\bar{F}(t_1, t_2) = P[T_1 > t_1, T_2 > t_2] = e^{-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)}$$

for $t_1 \geqslant 0$, $t_2 \geqslant 0$. If $T_{(1)} \leqslant T_{(2)}$ are the order statistics of (T_1, T_2), then

$$P[T_{(2)} > t \mid T_{(1)} > s] = e^{-\lambda_{12}(t-s)}\{e^{-\lambda_1(t-s)} + e^{-\lambda_2(t-s)} - e^{-(\lambda_1+\lambda_2)(t-s)}\}.$$

It is routine to check that $P[T_{(2)} > t \mid T_{(1)} > s] \uparrow s$.

One might seek to establish conditions on the joint distribution of $X_1, ..., X_n$ (in, for example, the case where they have a permutation symmetric distribution function) to ensure that $\mathrm{RTI}(X_{(j)} \mid X_{(i)})$ for any $i < j$.

ACKNOWLEDGMENTS

The authors are grateful to the associate editor and referees for helpful comments.

References

1. B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja (1992). *A First Course in Order Statistics*. Wiley, New York.
2. K. Balasubramanian, and N. Balakrishnan (1993). A log-concavity property of probability of occurrence of exactly r arbitrary events. *Statist. Probab. Lett.* **16** 249–251.
3. R. B. Bapat, and M. I. Beg (1989). Order statistics for nonidentically distributed variables and permanents. *Sankhyā Ser. A* **51** 79–93.
4. R. E. Barlow, and F. Proschan (1981). *Statistical Theory of Reliability and Life Testing: Probability Models*. To Begin With, Silver Spring, MD.
5. P. J. Bickel (1967). Some contributions to the theory of order statistics. In *Fifth Berkeley Symposium on Mathematics and Statistics*, Vol. 1, pp. 575–591 (L. M. LeCam and J. Neyman, Eds.). Univ. of California Press, Berkeley, CA.
6. H. W. Block, A. R. Sampson, and T. Savits (Eds.) (1990). *Topics in Statistical Dependence*. IMS Lecture Notes, Vol. 16, Inst. Math. Statist. Hayward, CA.
7. H. A. David (1980). *Order Statistics* (2nd ed.). Wiley, New York.
8. B. Efron (1965). Increasing properties of Pólya frequency functions. *Ann. Math. Statist.* **36** 272–279.
9. J. D. Esary, and F. Proschan (1972). Relationships among some notions of bivariate dependence. *Ann. Math. Statist.* **43** 651–655.
10. R. Harris (1970). A multivariate definition for increasing hazard rate distributions. *Ann. Math. Statist.* **41** 413–417.
11. K. Jogdeo (1982). Concepts of dependence. In *Encyclopedia of Statistical Sciences*, Vol. 2, pp. 324–334 (S. Kotz and N. L. Johnson, Eds.). Wiley-Interscience, New York.
12. S. Karlin (1968). *Total Positivity*. Stanford Univ. Press, Stanford, CA.
13. J. S. Kim, F. Proschan, and J. Sethuraman (1988). Stochastic comparisons of order statistics, with applications in reliability. *Comm. Statist. Theory Methods* **17** 2151–2172.
14. S. H. Kim, and H. A. David (1990). On the dependence structure of order statistics and concomitants of order statistics. *J. Stat. Plann. Inference* **24** 363–368.
15. S. C. Kochar, and J. V. Deshpande (1986). Nonparametric tests for independence in bivariate reliability models. In *Proceedings International Symp. on Advances in Multivariate Statistical Analysis* (J. K. Ghosh and S. Dasgupta, Eds.), pp. 203–217.
16. E. L. Lehmann (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.
17. A. W. Marshall, and I. Olkin (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.
18. Y. S. Sathe, and S. M. Bendre (1991). Log-concavity of probability of occurrence of at least r independent variables. *Statist. Probab. Lett.* **11** 63–64.
19. M. Shaked (1977). A family of concepts of dependence for bivariate distributions. *J. Amer. Statist. Assoc.* **72** 642–650.
20. M. Shaked, and J. G. Shanthikumar (1994). *Stochastic Orders and Their Applications*. Academic Press, San Diego.
21. Y. L. Tong (1980). *Probability Inequalities in Multivariate Distributions*. Academic Press, New York.
22. J. W. Tukey (1958). A problem of Berkson, and minimum variance orderly estimators. *Ann. Math. Statist.* **29** 588–592.