isid/ms/2005/02
February 28, 2005
http://www.isid.ac.in/~statmath/eprints

Dependence Orderings for Generalized Order Statistics

Baha-Eldin Khaledi
Subhash Kochar

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Dependence Orderings for Generalized Order Statistics

Baha-Eldin Khaledi *
Department of Statistics
College of Sciences
Kermanshah, Iran
e-mail: bkhaledi@hotmail.com

Subhash Kochar
Indian Statistical Institute
7, SJS Sansanwal Marg
New Delhi-110016, India
e-mail : kochar@isid.ac.in

Abstract

Generalized order statistics (g**OS**s) unify the study of order statistics, record values, k-records, Pfeifer's records and several other cases of ordered random variables. In this paper we consider the problem of comparing the degree of dependence between a pair of g**OS**s thus extending the recent work of Averous, Genest and Kochar (2005). It is noticed that as in the case of ordinary order statistics, copula of g**OS**s is independent of the parent distribution. For this comparison we consider the notion of *more regression dependence or more stochastic increasing*. It follows that under some conditions, for $i < j$, the dependence of the jth generalized order statistic on the ith generalized order statistic decreases as i and j draw apart. We also obtain a close form expression for the Kendall's coefficient of concordance between a pair of record values.

Key words : Dispersive ordering; Pure birth process; Exponential distribution; Kendall's tau; Monotone regression dependence; Stochastic increasingness; Record values.

1 Introduction

Order statistics and record values play an important role in statistics, in general, and in Reliability Theory and Life Testing, in particular. Their distributional and stochastic properties have been studied extensively but separately in the literature. However, they can be considered as special cases of generalized order statistics (g**OS**s) (cf. Kamps, 1995) which in addition cover particular sequential order statistics, kth record values, Pfeifer's record model, k_n record from nonidentical distributions, and ordered random variables which arise from truncated distributions. It is well known that a sequence of record values can be viewed as a sequence of the occurrence times of a certain non-homogeneous Poisson process. It is also connected to the failure times of a minimal repair process. There is a close connection between Pfeifer's records and the occurrence times of a pure birth process (cf. Pfeifer, 1982b).

*Corresponding author. Work done while visiting Indian Statistical Institute, Delhi Center

As said above many interesting stochastic orderings results for order statistics and spacings on one hand; and for record values and record increments on the other hand, have been obtained separately by many investigators without realizing that perhaps they can be unified under the umbrella of g**OS**s. Kamps (1995) in the last chapter of his book studied some reliability properties of generalized order statistics. Franco, Ruiz and Ruiz (2002) obtained some stochastic ordering results for spacings of generalized order statistics.

Recently Averous, Genest and Kochar (2005) have studied the dependence properties of order statistics of a random sample from a continuous distribution. To compare the degree of association between two such pairs of ordered random variables, they considered a notion of relative monotone regression dependence (or stochastic increasingness). Using this concept, they proved that for $i < j$, the dependence of the jth order statistic on the ith order statistic decreases as i and j draw apart. In this paper we study dependence properties of a pair of generalized order statistics and as a consequence these results will be applicable to order statistics, record values, occurrence times of a pure birth process and all those models which are covered under g**OS**s.

The organization of the paper is as follows. In Section 2, we introduce g**OS**s and state the main theorem which describes the conditions under which a pair of g**OS**s is more dependent than another pair in the sense of *more SI* ordering. It is seen that the work of Averous, Genest and Kochar (2005) can be extended to the g**OS**s. In Section 3 we point out a close connection that exists between the concepts of dispersive ordering and that of *more SI* ordering. The proofs of the various results are given in this section. In the last section, we obtain a close form expression for the value of the Kendall's τ between a pair of record values.

2 Main results

First we give the definition of the joint distribution of n generalized order statistics (cf. Kamps, 1995, p. 49).

Definition 2.1 *Let $n \in N$, $k \geq 1$, $m_1, \ldots, m_{n-1} \in I\!R$, $M_r = \sum_{j=r}^{n-1} m_j$, $1 \leq r \leq n-1$ be parameters such that*

$$\gamma_r = k + n - r + M_r \geq 1 \textit{ for all } r \in \{1, \ldots, n-1\}$$

and let $\tilde{m} = (m_1, \ldots, m_{n-1})$, if $n \geq 2$ ($\tilde{m} \in I\!R$ arbitrary, if $n = 1$).
If the random variables $U(r, n, \tilde{m}, k)$, $r = 1, \ldots, n$, possess a joint density function of the form

$$f^{U(1,n,\tilde{m},\ k),\ldots,U(n,n,\tilde{m},\ k)}(u_1, \ldots, u_n) = k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{j=1}^{n-1} (1-u_i)^{m_i} \right) (1-u_n)^{k-1}$$

on the cone $0 \leq u_1 \leq \ldots \leq u_n < 1$ of $I\!R^n$, then they are called uniform generalized order statistics.

Generalized order statistics based on some distribution function F are now defined by means of the quantile transformation

$$X(r, n, \tilde{m},\ k) = F^{-1}(U(r, n, \tilde{m},\ k)),\ r = 1, \ldots, n$$

and they are denoted by g**OS**s. As discussed in Kamps (1995), for suitable choices of the parameters these reduce to the joint distributions of order statistics from a continuous distribution, record values, Pfeifer's record values and so on.

Let (S, T) be a continuous bivariate random vector with joint distribution function H. Recall that T is said to be stochastically increasing in S if and only if, for all $s, s', t \in \mathbb{R}$,

$$s \leq s' \Longrightarrow P(T \leq t | S = s') \leq PT \leq t | S = s). \tag{2.1}$$

Let $H_{[s]}$ denote the distribution function of the conditional distribution of T given $S = s$. The above implication may then be expressed in the alternate form

$$s \leq s' \Longrightarrow H_{[s']} \circ H_{[s]}^{-1}(u) \leq u,$$

where $u \in (0, 1)$. .

Note that property (2.1) is not symmetric in S and T, but that in case these variables are independent, $H_{[s']} \circ H_{[s]}^{-1}(u) \equiv u$ for all $u \in (0, 1)$ and for all $s, s' \in \mathbb{R}$. Observe also that if $\xi_p = F^{-1}(p)$ denotes the pth quantile of the marginal distribution of S, then (2.1) is equivalent to the condition

$$0 < p \leq q < 1 \Longrightarrow H_{[\xi_q]} \circ H_{[\xi_p]}{}^{-1}(u) \leq u$$

holding true for all $u \in (0, 1)$.

To compare the relative degree of dependence between arbitrary pairs of g**OS**s we use the notion of *more stochastically increasing* dependence ordering as discussed in Averous, Genest and Kochar (2005). For $i = 1, 2$, let (S_i, T_i) be a pair of continuous random variables with joint cumulative distribution function H_i and marginals F_i and G_i.

Definition 2.2 *T_2 is said to be more stochastically increasing in S_2 than T_1 is in S_1, denoted by $(T_1 | S_1) \prec_{\mathrm{SI}} (T_2 | S_2)$ or $H_1 \prec_{\mathrm{SI}} H_2$, if and only if*

$$0 < p \leq q < 1 \Longrightarrow H_{2[\xi_{2q}]} \circ H_{2[\xi_{2p}]}{}^{-1}(u) \leq H_{1[\xi_{1q}]} \circ H_{1[\xi_{1p}]}{}^{-1}(u), \tag{2.2}$$

for all $u \in (0, 1)$, where for $i = 1, 2$, $H_{i[s]}$ denotes the conditional distribution of T_i given $S_i = s$, and $\xi_{ip} = F_i^{-1}(p)$ stands for the pth quantile of the marginal distribution of S_i.

Obviously, (2.2) implies that T_2 is stochastically increasing in S_2 if S_1 and T_1 are independent. It also implies that if T_1 is stochastically increasing in S_1, then so is T_2 in S_2; and conversely, if T_2 is stochastically decreasing in S_2, then so is T_1 in S_1. As observed in Averous, Genest and Kochar (2005), the above definition of more SI ordering depends on the joint distributions of the underlying random variables only through their copulas. Also

$$(T_1 | S_1) \prec_{\mathrm{SI}} (T_2 | S_2) \Rightarrow C_1(u, v) \leq C_2(u, v), \tag{2.3}$$

where C_i is the copula of (S_i, T_i), $i = 1, 2$, which in turn implies that

$$\kappa(S_1, T_1) \le \kappa(S_2, T_2)$$

where $\kappa(S, T)$ represents Spearman's rho, Kendall's tau, Gini's coefficient, or indeed any other copula-based measure of concordance satisfying the axioms of Scarsini (1984). In the special case where $F_1 = F_2$ and $G_1 = G_2$, it also follows from (2.3) that the pairs (S_1, T_1) and (S_2, T_2) are ordered by Pearson's correlation coefficient, namely

$$\mathrm{corr}(S_1, T_1) \le \mathrm{corr}(S_2, T_2).$$

Note that the copula of a pair of g**OS**s is independent of the parent distribution F. For comparing two different g**OS**s we use the following pre-ordering on $I\!R^+$.

Definition 2.3 *A vector* $\mathbf{x}$ *in* $I\!R^+$ *is said to be p-larger than another vector* $\mathbf{y}$ *also in* $I\!R^+$ *(written* $\mathbf{x} \overset{p}{\succeq} \mathbf{y}$*) if* $\prod_{i=1}^{j} x_{(i)} \le \prod_{i=1}^{j} y_{(i)}, j = 1, \ldots, n$*, where* $x_{(1)} \le \ldots \le x_{(n)}$ *and* $y_{(1)} \le \ldots y_{(n)}$ *are the increasing arrangements of the components of* $\mathbf{x}$ *and* $\mathbf{y}$*, respectively.*

Now we state the main theorem of the this paper whose proof is given in Section 3.

Theorem 2.1 *Let* $(X\,(r, n, \tilde{m}, k)\,,\ r = 1, \ldots, n)$ *and* $\left(X'(r', n', \tilde{m}', k')\ r = 1, \ldots, n\right)$ *be the g***OS***s based on distributions* F *and* G*, respectively. Let* $\gamma_r = k + n - r + \sum_{h=r}^{n-1} m_h$ *and* $\gamma_r' = k' + n' - r + \sum_{h=r}^{n-1} m_h'$*. Then for* $i \le j$ *and* $i' \le j'$*,*

$$\left(X'(j', n', \tilde{m}', k') \mid X'(i', n', \tilde{m}', k')\right) \prec_{\mathrm{SI}} \left(X(j, n, \tilde{m}, k) \mid X(i, n, \tilde{m}, k)\right),$$

provided the following conditions are satisfied,

(a1) $i \ge i'$ *and* $j - i \le j' - i'$*,*

(a2) $(\gamma_{\ell_1}, \ldots, \gamma_{\ell_{i'}}) \overset{p}{\succeq} (\gamma_1', \ldots, \gamma_{i'}')$
for some set $\{\ell_1, \ldots, \ell_{i'}\} \subset \{1, \ldots, i\}$*.*

(a3) $(\gamma_{k_1}', \ldots, \gamma_{k_{j-i}}') \overset{p}{\succeq} (\gamma_{i+1}, \ldots, \gamma_j)$
for some set $\{k_1, \ldots, k_{j-i}\} \subset \{i' + 1, \ldots, j'\}$

It is well known that for specific sets of parameters, n, k and m_i, $i = 1, \ldots, n-1$, the g**OS**s reduce to the well known ordered random variables. Now we find sufficient conditions on the parameters of the various sub-models of g**OS**s for which Theorem 2.1 holds.

(A) Order Statistics from i.i.d random variables. For $n \ge 1$, let $X_{i:n}$ denote the *ith* order statistic based on a random sample $X_1, \ldots, X_n$ from a continuous distribution with cdf F. This is a special case of g**OS**s with $m_1 = \ldots = m_{n-1} = 0$ and $k = 1$. In this case $\gamma_r = n - r + 1$, $r = 1, \ldots, n-1$. Let $m_i = m_i' = 0$, $i = 1, \ldots, n-1$ and $k = k' = 1$. With these settings we see that the conditions (a2) and (a3) are satisfied when $n - i \le n' - i'$

and $n-j \geq n^{'}-j^{'}$. That is, for $i \geq i^{'}$, $j-i \leq j^{'}-i^{'}$, $n-i \leq n^{'}-i^{'}$ and $n-j \geq n^{'}-j^{'}$, we have

$$\left(X'_{j':n'} \mid X'_{i':n'}\right) \prec_{\mathrm{SI}} \left(X_{j:n} \mid X_{i:n}\right),$$

as proved recently by Averous, Genest and Kochar (2005). In the special case of one-sample problem when $n = n'$, we have the following results.

(a) $(X_{k:n}| \, X_{i:n}) \prec_{\mathrm{SI}} (X_{j:n}| \, X_{i:n})$ *for all* $1 \leq i < j < k \leq n$;

(b) $(X_{j:n}| \, X_{i:n}) \prec_{\mathrm{SI}} (X_{j+1:n+1}| \, X_{i+1:n+1})$ *for all* $1 \leq i < j \leq n$;

(c) $(X_{n+1:n+1}| \, X_{1:n+1}) \prec_{\mathrm{SI}} (X_{n:n}| \, X_{1:n})$ *for every integer* $n \geq 2$.

(B) ***k*-Records.** Let $\{X_i, i \geq 1\}$ be a sequence of i.i.d random variables from a continuous distribution F and let k be a positive integer. The random variables $L^{(k)}(n)$ given by $L^{(k)}(1) = 1$,

$$L^{(k)}(n+1) = min\{j \in N; X_{j:j+k-1} > X_{L^{(k)}(n):L^{(k)}((n)+k-1)}\},\ n \geq 1,$$

are called the nth k-record times and the quantities $X_{L^{(k)}(n):L^{(k)}((n)+k-1)}$ which we denote by $R(n:k)$ are termed the nth k-records (cf. Kamps, 1995, p.34). The joint density of the first n k-records corresponding to a sequence of independent random variables from a continuous distribution F is a special case of the joint density of first n g**OS**s with $m_1 = \ldots = m_{n-1} = -1$. In this case $\gamma_r = k$, $r = 1, \ldots, n-1$. Now let $m_i = m_i^{'} = -1$, $i = 1, \ldots, n-1$ and $k = k^{'}$. Using the above setting it follows that the conditions (a2) and (a3) of Theorem 2.1 are satisfied. Therefore, for $i \geq i^{'}$, $j-i \leq j^{'}-i^{'}$, we have

$$\left(R'(j^{'}:k) \mid R'(i^{'}:k)\right) \prec_{\mathrm{SI}} \left(R(j:k) \mid R(i:k)\right),$$

where $R(j:k)$,$j \geq 1$ and $R'(j^{'}:k)$, $j' \geq 1$ stand for the jth and j'th k-records. This means that for $i < j$, the dependence of the jth k-record on the ith k-record decreases as i and j draw apart.

(C) **Two Stage Progressive Type II Censoring**. Let $X_1, \ldots, X_v$ be a random sample from a continuous distribution F. Let these be the lifetimes of v items put on test at time $t = 0$. At the time of the r_1th failure, n_1 functioning items are randomly selected and removed from the test. The test terminates when further r_2 items have failed. The $n = r_1 + r_2$ observations $X_{1:v} \leq \ldots \leq X_{n:v}$ are called order statistics arising in progressive type II censoring with two stages. This is a special case of g**OS**s with $m_1 = \ldots = m_{r_1-1} = m_{r_1+1} = \ldots = m_{n-1} = 0$, $m_{r_1} = n_1$ and $k = v - n_1 - n + 1$. In this case $\gamma_r = v - r + 1$, $r = 1, \ldots, r_1$ and $\gamma_r = v - n_1 - r + 1$, $r = r_1 + 1, \ldots, n-1$. Let $m_i = m_i' = 0$, $i = 1, \ldots, r_1 - 1, r_1 + 1, \ldots, n-1$, $m_{r_1} = m'_{r_1} = n_1$, $k = v - n_1 - n + 1$ and $k^{'} = v' - n_1 - n + 1$. With these settings we see that the conditions (a2) and (a3) are satisfied when $v-i \leq v^{'}-i^{'}$ and $v-j \geq v^{'}-j^{'}$. That is, for $i \geq i^{'}$, $j-i \leq j^{'}-i^{'}$, $v-i \leq v^{'}-i^{'}$ and $v-j \geq v^{'}-j^{'}$, we have

$$\left(X'_{j':v'} \mid X'_{i':v'}\right) \prec_{\mathrm{SI}} \left(X_{j:v} \mid X_{i:v}\right).$$

As discussed in Kamps (1995), there are many other models like Pfeifer's records, sequential order statistics, order statistics with non-integral sample size etc which can also be expressed as special cases of g**OS**s.

3 Auxiliary results and proofs

In this section we prove some auxiliary results to prove our Theorem 2.1. As we will see, there is a close connection between the concepts of dispersive ordering and *more SI ordering*.

Definition 3.1 *A random variable X with distribution function F is said to be less dispersed than another variable Y with distribution G, written as $X \leq_{disp} Y$ or $F \leq_{disp} G$, if and only if*

$$F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha)$$

for all $0 < \alpha \leq \beta < 1$.

It is easy to see that the $F \leq_{disp} G$ is equivalent to

$$F\{F^{-1}(u) - c\} \leq G\{G^{-1}(u) - c\} \text{ for every } c \geq 0 \text{ and } u \in (0,1).$$

For general information about dispersive ordering and its properties, refer to Section 2.B of Shaked and Shanthikumar (1994). The next proposition establishes a close connection between dispersive ordering and *more SI ordering.*

Proposition 3.1 *Let X_i and Y_i be independent random variables with distribution functions F_i and G_i, respectively for $i = 1, 2$. Then*

$$X_2 \leq_{disp} X_1 \text{ and } Y_1 \leq_{disp} Y_2 \Rightarrow (X_2+Y_2)|X_2 \prec_{\mathrm{SI}} (X_1+Y_1)|X_1$$

Proof. Let ξ_{ip} denote the pth quantile of F_i, $\ i = 1, 2$. Since X_i and Y_i are independent for $i = 1, 2$, $H_{i[\xi_{ip}]}(z) = P[X_i + Y_i \leq z | X_i = \xi_{ip}] = G_i(z - \xi_{ip})$ and $H^{-1}_{i[\xi_{ip}]}(u) = G_i^{-1}(u) + \xi_{ip}$. This gives

$$H_{i[\xi_{iq}]} \circ {H_{i[\xi_{ip}]}}^{-1}(u) = G_i[G_i^{-1}(u) - (\xi_{iq} - \xi_{ip})].$$

Since $X_2 \leq_{disp} X_1$,

$$\xi_{2q} - \xi_{2p} \leq \xi_{1q} - \xi_{1p} \text{ for } 0 < p \leq q < 1 \tag{3.1}$$

In order to prove Proposition 3.1 one needs only show that one has for $0 < p \leq q < 1$,

$$H_{1[\xi_{1q}]} \circ {H_{1[\xi_{1p}]}}^{-1}(u) \leq H_{2[\xi_{2q}]} \circ {H_{2[\xi_{2p}]}}^{-1}(u),$$

that is,

$$G_1[G_1^{-1}(u) - (\xi_{1q} - \xi_{1p})] \le G_2[G_2^{-1}(u) - (\xi_{2q} - \xi_{2p})]. \qquad (3.2)$$

Since $Y_1 \le_{disp} Y_2$, by taking $c = \xi_{1q} - \xi_{1p} \ge 0$ it follows from the definition of dispersive ordering that

$$G_1[G_1^{-1}(u) - (\xi_{1q} - \xi_{1p})] \le G_2[G_2^{-1}(u) - (\xi_{1q} - \xi_{1p})]$$

Now (3.2) follows from it and (3.1) since $X_2 \le_{disp} X_1$ and G_2 is nondecreasing.

∎

We shall be using the following known results to prove Theorem 2.1 in this section.

Theorem 3.1 *(Khaledi and Kochar, 2004). Let $X_{\lambda_1}, \dots, X_{\lambda_n}$ be independent random variables such that X_{λ_i} has gamma distribution with shape parameter $a \ge 1$ and scale parameter λ_i, for $i = 1, \dots, n$. Then, $\boldsymbol{\lambda} \overset{p}{\succeq} \boldsymbol{\lambda}'$ implies*

$$\sum_{k=1}^{n} X_{\lambda_k} \ge_{disp} \sum_{k=1}^{n} X_{\lambda'_k}.$$

Lemma 3.1 *(Lewis and Thompson, 1981). The random variable X satisfies $X \le_{disp} X + Y$ for any random variable Y independent of X if and only if X has a logconcave density.*

Theorem 3.2 *(cf. Kamps, 1995, p.81). Let $X(r, n, \tilde{m}, k)$, $r = 1, \dots, n$ be the g**OS**s based on the distribution function F with $F(x) = 1 - e^{-x}$, $x \ge 0$. Let*

$$Y_1 = \gamma_1 X(1, n, \tilde{m}, k) \text{ and } Y_j = \gamma_j (X(j, n, \tilde{m}, k) - X(j-1, n, \tilde{m}, k)),\ j = 2, \dots, n,$$

where $\gamma_j = k + n - j + \sum_{i=j}^{n-1} m_i$. Then the random variables $Y_1, \dots, Y_n$ are stochastically independent and identically distributed according to distribution F.

Moreover, for $r = 2, \dots, n$ we have the representation

$$X(r, n, \tilde{m}, k) \overset{st}{=} \sum_{j=1}^{r} X_{\gamma_j},$$

where X_{γ_j} has exponential distribution with hazard rate γ_j, $j = 1, \dots, r$.

To prove the main result in this section we use the following lemma which may be of independent interest.

Lemma 3.2 *Let $X_{\gamma_1}, \dots, X_{\gamma_n}$ be independent random variables such that X_{γ_k} has gamma distribution with shape parameter $a \ge 1$ and scale parameter γ_k, for $k = 1, \dots, n$ and let $X_{\gamma'_1}, \dots, X_{\gamma'_{n'}}$ be another set of independent random variables such that $X_{\gamma'_k}$ has gamma distribution with shape parameter $a \ge 1$ and scale parameter γ'_k, for $k = 1, \dots, n'$. Then if the conditions $(a1) - a(3)$ of Theorem 2.1 are satisfied, then for $i \le j$, $i' \le j'$,*

$$\sum_{k=1}^{j'} X_{\gamma'_k} \mid \sum_{k=1}^{i'} X_{\gamma'_k} \prec_{\mathrm{SI}} \sum_{k=1}^{j} X_{\gamma_k} \mid \sum_{k=1}^{i} X_{\gamma_k}.$$

Proof : Using Proposition 3.1, it is enough to show that under the assumed conditions

$$(A) \qquad \sum_{\nu=1}^{i} X_{\gamma_\nu} \geq_{disp} \sum_{\nu=1}^{i'} X_{\gamma_\nu'}$$

and

$$(B) \qquad \sum_{\nu=i+1}^{j} X_{\gamma_\nu} \leq_{disp} \sum_{\nu=i'+1}^{j'} X_{\gamma'_\nu}.$$

For $i \geq i'$, we have

$$\begin{aligned} \sum_{\nu=1}^{i} X_{\gamma_\nu} &= \sum_{\nu=1}^{i'} X_{\gamma_{\ell_\nu}} + \sum_{\nu \notin \{\ell_1, \ldots, \ell_{i'}\}} X_{\gamma_\nu} \\ &\geq_{disp} \sum_{\nu=1}^{i'} X_{\gamma_{\ell_\nu}} \\ &\geq_{disp} \sum_{\nu=1}^{i'} X_{\gamma'_\nu}, \end{aligned}$$

since the density function of a gamma random variable with shape parameter $a \geq 1$ is logconcave and a convolution of independent random variables with logconcave densities is logconcave, the first inequality follows from Lemma 3.1. The second inequality follows from Theorem 3.1 under the condition (a2). This completes the proof of (A).

The proof of (B) follows on the same lines under the condition (a3).

■

Proof of Theorem 2.1. It is clear from the definition of the joint distribution of g**OS**s that their copula is independent of the parent distribution. Hence without loss of generality we can assume that the distributions F and G both are standard exponential. It follows from Theorem 3.2 that

$$X(j,n,\tilde{m},k) \stackrel{st}{=} \sum_{h=1}^{j} X_{\gamma_h} \quad \text{and } X(j,n,\tilde{m},k) \mid X(i,n,\tilde{m},k) \stackrel{st}{=} \sum_{h=1}^{j} X_{\gamma_h} \mid \sum_{h=1}^{i} X_{\gamma_h},$$

where X_{γ_h} has exponential distribution with hazard rate γ_h, $h = 1, \ldots, j$ and X_{γ_h}'s are independent. Now the required result follows from Lemma 3.2.

■

It is known that more SI ordering implies more PQD ordering (copulas are ordered) and it is also known that the Spearman's rho, Kendall's tau, or Gini's coefficient of association can be expressed as a functional of copula which preserves the ordering of copula in the same direction (cf. Joe, 1997). This leads us to the following corollary.

Corollary 3.1 *Under the conditions of Theorem 2.1,*

$$\kappa\left(X'(i',n',\tilde{m}',k'), X'(j',n',\tilde{m}',k')\right) \leq \kappa\left(X(i,n,\tilde{m},k), X(j,n,\tilde{m},k)\right)$$

where $\kappa(S,T)$ stands for any measure of concordance between S and T in the sense of Scarsini (1984), e.g., Spearman's rho, Kendall's tau, or Gini's coefficient of association

4 Kendall's τ for record values

In order to further understand the implications of Corollary 3.1 we find a closed form formula for the Kendall's coefficient of measure of concordance τ between any two records corresponding to a sequence of i.i.d random variables from an arbitrary distribution F.

Theorem 4.1 *Let $\{X_i, i \geq 0\}$ be a sequence of independent and identically distributed random variables from a continuous distribution F. Then the Kendall's coefficient of concordance τ between the records R_m and R_n is*

$$\tau(R_m, R_n) = 1 - 4 \sum_{j=m+1}^{n} \sum_{i=0}^{n-j} \frac{1}{2^{n+i+j+1}} \binom{m+j}{j} \binom{n-m+i-1}{i}.$$

PROOF : Since the copula and hence τ for a pair of records is independent of the parent distribution, without loss of generality we assume that F is standard exponential. To derive this formula we shall use the following identities :

$$\frac{1}{\beta(a, n-a+1)} \int_0^p t^{a-1}(1-t)^{n-a} dt = \sum_{j=a}^{n} \binom{n}{j} p^j (1-p)^{n-j}, \tag{4.1}$$

for $0 \leq p \leq 1$, $a = 1, \ldots, n$, $n = 1, 2, \ldots$ and

$$\int_x^{+\infty} \frac{1}{\Gamma(n)} t^{n-1} e^{-t} dt = e^{-x} \sum_{i=0}^{n-1} \frac{x^i}{i!}, \tag{4.2}$$

where $\beta(a, b)$ and $\Gamma(a)$ stand, respectively for the beta and the gamma functions.

Let $R_m^{'}$ and $R_n^{'}$ be the records corresponding to a sequence $\left\{X_i^{'}, i \geq 0\right\}$ of i.i.d random variables with common distribution function as standard exponential. We assume that this sequence is independent of the sequence $\{X_i, i \geq 0\}$.

The joint density function of (R_m, R_n) for $m \leq n$ is

$$f_{R_m, R_n}(x, y) = \frac{1}{m!(n-m-1)!} x^m (y-x)^{n-m-1} e^{-y}, \text{ for } 0 < x \leq y < \infty.$$

By definition, the Kendall's τ is

$$\tau(R_m, R_n) = 1 - 4p,$$

where

$$\begin{aligned} p &= P(R_m < R_m^{'}, R_n > R_n^{'}) \\ &= \int_0^{+\infty} \int_x^{+\infty} P(R_m < x, R_n > y) \frac{1}{m!(n-m-1)!} x^m (y-x)^{n-m-1} e^{-y} dy\, dx \end{aligned} \tag{4.3}$$

Now first we compute $h(x,y) = P(R_m < x, R_n > y)$.

$$\begin{aligned}
h(x,y) &= \int_y^{+\infty}\int_0^x \frac{1}{m!(n-m-1)!}u^m(v-u)^{n-m-1}e^{-v}du\,dv \\
&= \int_y^{+\infty}\frac{v^{n-1}e^{-v}}{\Gamma(n+1)}\int_0^x \frac{1}{\beta(m+1,n-m)}\Big(\frac{u}{v}\Big)^m\Big(1-\frac{u}{v}\Big)^{n-m-1}du\,dv \\
&= \int_y^{+\infty}\frac{v^{n}e^{-v}}{\Gamma(n+1)}\int_0^{x/v} \frac{1}{\beta(m+1,n-m)}(z)^{(m+1)-1}(1-z)^{(n-m)-1}dz\,dv \\
&= \int_y^{+\infty}\frac{v^{n}e^{-v}}{\Gamma(n+1)}\sum_{j=m+1}^{n}\binom{n}{j}(\frac{x}{v})^j(1-\frac{x}{v})^{n-j}dv \qquad (4.4)\\
&= \sum_{j=m+1}^{n}\binom{n}{j}\frac{x^j}{\Gamma(n+1)}\int_y^{+\infty}(v-x)^{n-j}e^{-v}dv \\
&= \sum_{j=m+1}^{n}\binom{n}{j}\frac{x^j}{\Gamma(n+1)}\int_{y-x}^{+\infty}(z)^{n-j}e^{-(x+z)}dz \\
&= \sum_{j=m+1}^{n}\binom{n}{j}\frac{\Gamma(n-j+1)x^je^{-x}}{\Gamma(n+1)}\sum_{i=0}^{n-j}\frac{(y-x)^ie^{-(y-x)}}{i!} \qquad (4.5)\\
&= \sum_{j=m+1}^{n}\sum_{i=0}^{n-j}\frac{x^j(y-x)^ie^{-y}}{i!j!}
\end{aligned}$$

(4.4) and (4.5) follows respectively, from (4.1) and (4.2). Using the above expression in (4.3), We get

$$p = \sum_{j=m+1}^{n}\sum_{i=0}^{n-j}\int_0^{+\infty}\left(\int_x^{+\infty}(y-x)^{n-m+i-1}e^{-2y}dy\right)\frac{x^{m+j}}{i!j!m!(n-m-1)!}\,dx.$$

Simplifying it, we get the required result.

■

Table 1 gives the values of $\tau(R_m, R_n)$ for $1 \leq m \leq n = 7$.

Table 1: The values of $2048 \times \tau(R_i, R_j)$.

i	j = 2	3	4	5	6	7
1	1280	1024	880	784	714	660
2		1408	1168	1024	924	849
3			1488	1264	1124	1024
4				1544	1334	1199
5					1586	1388
6						1619

It is seen from the above table that for fixed i, $\tau(R_i, R_j)$ decreases with j ($\geq i$) and for fixed j, it increases with i ($\leq j$). Also for a fixed integer c, $\tau(R_i, R_{i+c})$ increases with i. It is easy to see that the conclusions of Theorem 2.1 hold in this case.

Acknowledgements

The authors are thankful to two anonymous referees and the associate editor for pointing out some mistypes in the original draft of the paper and for making many useful suggestions.

References :

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records.* John Wiley, New York.

Avérous, J., Genest, C. and Kochar, S.C. (2005). On dependence structure of order statistics. To appear in *Journal of Multivariate Analysis* .

Franco, M., Ruiz, J. M. and Ruiz, M. C. (2002). Stochastic ordering between spacings of generalized order statistics. *Probability in Engineering and Information Sciences* **16**, 471-484.

Kamps, U. (1995). *A concept of Generalized Order Statistics.* Teubner, Stuttgart.

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman & Hall, London, 1997.

Khaledi, B. and Kochar, S. C. (2004). Ordering convolutions of gamma random variables. *Sankhya* **66**, 466-473*.

Lewis, T. and Thompson, J. W. (1981). Dispersive distribution and the connection between dispersivity and strong unimodality. *Journal of Applied Probability* **18**, 76-90

Pfeifer, P. (1982a). Characterizations os exponential distributions by independent non-stationary record increments. *Journal of Applied Probability* **19**, 127-135. (Correction **19**, p. 906.)

Pfeifer, P. (1982b). The structure of elementary pure birth processes. *Journal of Applied Probability* **19**, 664-667.

Scarsini, M. (1984). On measures of concordance, *Stochastica* **8** , 201–218.

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