Stochastic comparisons of order statistics

Song-Ho Kim

Iowa State University
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Stochastic comparisons of order statistics

Kim, Song-Ho, Ph.D.

Iowa State University, 1988
Stochastic comparisons of order statistics

by

Song-Ho Kim

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Statistics
Major: Statistics (Statistical Theory)

Approved:

Signature was redacted for privacy.

In Charge of Major Work
Signature was redacted for privacy.

For the Major Department
Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1988
TABLE OF CONTENTS

1. INTRODUCTION
   1.1 Introduction
   1.2 Scope of the Present Investigation
   1.3 Stochastic Order Relations
   1.4 Positive Dependence Concepts
      1.4.1 Bivariate case
      1.4.2 Multivariate case
   1.5 Miscellaneous Definitions

2. STOCHASTIC ORDERINGS AND EXPECTATIONS
   2.1 Introduction
   2.2 Expectations - One Sample
   2.3 Expectations - Two Sample

3. DEPENDENCE STRUCTURE OF ORDER STATISTICS
   3.1 Introduction
   3.2 Dependence Structure-Order Statistics
   3.3 Covariance-Order Statistics
   3.4 Covariance-Spacings

4. NONSTANDARD SITUATIONS
   4.1 Introduction
   4.2 I.n.i.d. Case
   4.3 Dependent Case

5. BIBLIOGRAPHY

6. ACKNOWLEDGMENTS
1.

1. INTRODUCTION

1.1 Introduction

Given a random sample, $X_1, X_2, \ldots, X_n$, we can arrange the $X$'s in ascending order of magnitude and then write

$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}.$$ 

We call $X_{r:n}$ the $r^{th}$ order statistic $(r = 1, 2, \ldots, n)$. If the distribution function of $X$ is $F(x)$, then $F_{r:n}(x)$ or $F_{(r)}(x)$ denotes the distribution function of the $r^{th}$ order statistic. Define the $k^{th}$ spacing $D_k$ of the order statistics $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ as $D_k = X_{k:n} - X_{k-1:n}$ for $k = 2, \ldots, n$ and $D_1 = X_{1:n} - b$ when $X$ is bounded below by finite number, $b$. When $(X_i, Y_i) (i = 1, 2, \ldots, n)$ are $n$ independent random variables having a common bivariate distribution, the $Y$-variate $Y_{r:n}$ associated with $X_{r:n}$ is called the concomitant of the $r^{th}$ order statistic (David (1973)).

If $X_1, X_2, \ldots, X_n$ is a sample of size $n$ from a life distribution $F$, then the order statistics may be interpreted as the successive failure times of the components of a system. With this interpretation the $r^{th}$ order statistic is the failure time of a $k$-out-of-$n$ system of identical components, where $k = n - r + 1$. (A system of $n$ components is called a $k$-out-of-$n$ system if it functions if and only if at least $k$ components function.) Hence, $P[X_{r:n} > t]$ is the reliability of a $k$-out-of-$n$ system at time $t$. The special cases $k = n$ and $k = 1$ correspond respectively to series and parallel systems.
Consider the situation of two k-out-of-n systems of independently failing identical components whose lifetimes X and Y have c.d.f.s F and G. Then, it is well known that if the survival probability of an X-component is greater than that of a Y-component (i.e., \( P[X > t] \geq P[Y > t] \) for all \( t \)), then \( P[X_{r1:n} > t] \geq P[Y_{r1:n} > t] \) for all \( t \) and \( r = 1, 2, \ldots, n \). In other words, if X is stochastically larger than Y, then \( X_{r1:n} \) is stochastically larger than \( Y_{r1:n} \) for \( r = 1, 2, \ldots, n \). This implies \( E[X_{r1:n}] \geq E[Y_{r1:n}] \) for \( r = 1, 2, \ldots, n \) if \( E[X] < \infty \) and \( E[Y] < \infty \). Ross (1983, p. 256) relaxed the i.i.d. assumption. Hence, if each component life of a k-out-of-n system, \( X_i \), is stochastically larger than \( Y_i \), the corresponding component life of the other k-out-of-n system, then the system reliability of the k-out-of-n system with component lifetimes \( X_1, X_2, \ldots, X_n \) is larger than that of the k-out-of-n system with component lifetimes \( Y_1, Y_2, \ldots, Y_n \).

Stochastic ordering is a very strong kind of ordering. Consequently, many other weaker orderings have been studied. A major aim of this dissertation is to investigate the properties of order statistics under these orderings. For nonnegative random variables the results can always be interpreted in terms of k-out-of-n systems. However, we will also consider orderings for unbounded variates including symmetrically distributed variates. The results may then be relevant in selection procedures, outlier detection, and tests of spread.

Order statistics are dependent because of the inequality relations among them. Bickel (1967) showed that \( \text{cov}(X_{i1:n}, X_{j1:n}) \geq 0 \),
provided \( \mathbb{E}[X_{i:n}^2] + \mathbb{E}[X_{j:n}^2] < \infty \) when \( X_1, X_2, \ldots, X_n \) are a random sample from a population with c.d.f. \( F(x) \) and p.d.f. \( f(x) \), the latter being continuous and strictly positive on \( \{x|0 < F(x) < 1\} \). Esary et al. (1967) noted that \( X_{i:n}, X_{2:n}, \ldots, X_{n:n} \), being increasing functions of \( X_1, X_2, \ldots, X_n \), are associated (see Definition 1.12 (f)) when \( X_1, X_2, \ldots, X_n \) are mutually independent random variables. This implies that \( \text{cov}(X_{i:n}, X_{j:n}) \geq 0 \) for \( 1 \leq i \leq j \leq n \), where \( X_1, X_2, \ldots, X_n \) are independent or associated random variables, not necessarily identically distributed. We will show, however, that \( \text{cov}(X_{i:n}, X_{j:n}) \) can be negative if \( X_1, X_2, \ldots, X_n \) are sufficiently negatively dependent.

More generally, we investigate the dependence structure of order statistics as expressed through orderings of their covariances. For example, is it necessarily true that in random samples the covariance of two order statistics \( X_{i:n}, X_{j:n} \) decreases as \( i \) and \( j \) draw apart? Such questions, first studied by Tukey (1958), are closely related to "regression dependence." Under what conditions does \( X_{j:n} - X_{i:n} \) increase (or decrease) stochastically with \( X_{h:n} \), where \( h \leq i < j \)? The answer is clearly of interest in life testing. We will correct and extend Tukey's pioneering results. Related results on spacings \( X_{i:n} - X_{i-1:n} \) will also be obtained.

1.2 Scope of the Present Investigation

Section 1.3 defines and reviews various stochastic order relations, including stochastic ordering, convex ordering, s-ordering, \( * \)-shaped ordering, r-ordering, dispersive ordering, majorization
ordering, Lorenz ordering, sign-change ordering, and likelihood ratio ordering. Also the interrelationships among these orderings are presented. Section 1.4 includes the definition of many notions of positive dependence and the presentation of their hierarchical order. Section 1.5 contains miscellaneous definitions needed later.

In Chapter 2, we first investigate the stochastic comparison of order statistics from exchangeable random variables in a single sample. Corresponding to the stochastic ordering between $X$ and $Y$, we consider the case

$$F(t) \preceq G(t)$$

respectively for $t < 0$ (1.2.1)

when $X$ and $Y$ are random variables symmetric about 0, with c.d.f.'s $F$ and $G$. We consider the case, $F \preceq G$ (see Definition 1.5 and Remark 1.5.2) with $f(0) \geq g(0) > 0$, which is a stronger assumption than (1.2.1). Also examples are provided in each case.

In Chapter 3, using positive dependence concepts, we prove that if $X_1, X_2, ..., X_n$ are exchangeable and multivariate totally positive of order two (MTP^2) (see Definition 1.13 and Remark 1.13), then the order statistics are MTP^2. Application of this result to concomitants of order statistics will be considered. Also we investigate the dependence structure of order statistics as expressed through orderings of their covariances under various assumptions. This includes a special example involving a Polya frequency function of order two (PF^2) (see Definition 1.19 and Remark 1.19.3).

In Chapter 4, we first treat the stochastic comparisons of order
statistics from independent but nonidentically distributed (i.n.i.d.) variates as an extension of Chapter 3. We then consider order statistics under positive dependence, negative dependence, and exchangeability.

1.3 Stochastic Order Relations

In this section, we introduce some stochastic order relations between random variables. First, we consider the concept of one random variable being stochastically larger than another.

**Definition 1.1.** The random variable $X$ is stochastically larger than the random variable $Y$, written $X \geq_{st} Y$, if

$$\Pr\{X > a\} \geq \Pr\{Y > a\} \text{ for all } a. \quad (1.3.1)$$

**Remark 1.1.** If the first moments of $X$ and $Y$ exist, then (1.3.1) implies $E[X] \geq E[Y]$.

**Definition 1.2.** [Birnbaum (1948)] $Y$ is more peaked than $X$, if

$$\Pr\{|Y| > t\} \leq \Pr\{|X| > t\} \text{ for all } t \geq 0. \quad (1.3.2)$$

**Remark 1.2.** If $X$ and $Y$ are nonnegative random variables, then (1.3.2) reduces to (1.3.1). Also, if $X(Y)$ has symmetric distribution function $F(G)$ about 0, (1.3.2) is equivalent to $G(t) \geq F(t)$ for all $t \geq 0$.

In order to compare relative skewness, Van Zwet (1964) defined convex ordering. This ordering played a prominent role in initiating a broad range of research on reliability theory.
Definition 1.3. [Van Zwet (1964)] Let $F$ and $G$ be continuous distributions, with $G$ strictly increasing on its support, an interval. Then, $F$ is convex with respect to $G$ (written $F \leq G$) if $G^{-1}F(x)$ is a convex function of $x$ on the support of $F$.

Remark 1.3. Let $X(Y)$ have distribution function $F(G)$. Van Zwet (1964) showed that if $F \leq G$, then \[ \frac{\mu_{2k+1}(X)}{\sigma^{2k+1}(X)} \leq \frac{\mu_{2k+1}(Y)}{\sigma^{2k+1}(Y)}, \] where $k = 1, 2, \ldots$, the standardized odd central moments. In reliability theory, if $F$ is a life distribution, $G(X) = 1 - e^{-X}$, and $F \leq G$, then $F$ is said to have increasing failure rate (IFR). Van Zwet (1964) also introduced another ordering, $s$-ordering, that is restricted to the class of symmetric distributions. The main purpose of this ordering is the comparison of relative heaviness of tail among symmetric distributions.

Definition 1.4. [Van Zwet (1964)] Let $F$ and $G$ be continuous symmetric distributions, with $G$ strictly increasing on its support, an interval. Then, $F \leq G$ if $G^{-1}F(x)$ is concave-convex in $x$ on the support of $F$.

Van Zwet (1964) proved that if $F \leq G$, then \[ \frac{\mu_{2k}(X)}{\sigma^{2k}(X)} \leq \frac{\mu_{2k}(Y)}{\sigma^{2k}(Y)}, \] $k = 2, 3, \ldots$.

Remark 1.4. Since it is not true that a coherent system of independent IFR components is necessarily IFR, Barlow and Proschan (1966) defined the class of distributions with increasing failure rate average (IFRA). That is, $F$ has IFRA if \( -\frac{1}{t} \log F(t) \) is increasing in $t \geq 0$. Then, Barlow and Proschan (1975) showed that a coherent system itself has an
IFRA life distribution if each of its independent components has an IFRA life distribution. It can be shown that IFR implies IFRA. More generally, Barlow and Proschan (1966) introduced star-shaped ordering, a weaker ordering than convex ordering.

**Definition 1.5.** [Barlow and Proschan (1966)] Let \( F \) and \( G \) be continuous distributions, with \( G \) strictly increasing on its support, and \( F(0) = G(0) = 0 \). Then, \( F \) is star-shaped with respect to \( G \) (written \( F \preceq G \)) if 
\[ \frac{G^{-1}F(x)}{x} \text{ is increasing for } x \geq 0. \]

**Remark 1.5.1.** If \( F \) is a life distribution, \( G(x) = 1 - e^{-x} \) and \( F \preceq G \), then \( F \) is IFRA. If \( F \preceq G \), Barlow and Proschan (1975) show that (a) \( \bar{F}(x) \) crosses \( G(0)x \) at most once, and from above, as \( x \) increases from 0 to \( \infty \), for each \( \theta > 0 \), (b) if, in addition, \( F \) and \( G \) have the same mean, then a single crossing does occur, and \( F \) has smaller variance than \( G \).

**Remark 1.5.2.** Corresponding to \( s \)-ordering and its relation to convex ordering, Lawrence (1975) introduced \( r \)-ordering: \( F \preceq G \) if and only if 
\[ \frac{1}{x} G^{-1}F(x) \text{ is increasing (decreasing) for } x \geq 0 \text{ (} x < 0 \), and \( F(0) = G(0) = 1/2 \). And Lawrence (1975) showed that \( F \preceq G \Rightarrow F \preceq G \).

In order to compare relative dispersiveness, Lewis and Thompson (1981) defined dispersive ordering.

**Definition 1.6.** [Lewis and Thompson (1981)] If any two quantiles of \( G \) are more widely separated than the corresponding quantiles of \( F \), then \( F \prec G \) (i.e., \( F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha) \) for any \( 0 < \alpha < \beta < 1 \)).
Remark 1.6. Deshpande and Kochar (1983) showed that \( F < G \) is equivalent to \( G^{-1}F(t) - t \) is increasing in \( t \) (i.e., \( \frac{d}{dt}[G^{-1}F(t)-t] > 0 \) \( \iff \) \( f[F^{-1}(q)] \geq g[G^{-1}(q)] \) for \( 0 < q < 1 \) with \( F(x) = q \)). Actually, the pioneering work in this ordering was done by Saunders and Moran (1978). They showed that with density \( f^r(x) = \Gamma^{-1}(r)x^{-r-1}e^{-x}, x > 0, \)
\( (r_1) < (r_2) \) \( F \) \( < \) \( F \) when \( r_1 < r_2 \). Also, Shaked (1982a) showed that when \( F \) and \( G \) are two distribution functions which are strictly increasing and continuous on their support \([0,\infty)\), then \( F < G \) iff (a) \( F(x) - G(x) > 0 \) for all \( x \in [0,\infty) \) and (b) for every \( c \) the distribution functions \( X+c \) and \( Y \) cross at most once and if there is a sign change, \( F_X+c - G_Y \) changes sign from - to +. By (a), \( F < G \Rightarrow F < G \), when \( F \) and \( G \) are distributions from nonnegative random variables.

Definition 1.7. A vector \( b = (b_1, \ldots, b_n) \) majorizes the vector \( a = (a_1, \ldots, a_n) \) if \( \sum_{i=k}^{n} b_{(i)} \geq \sum_{i=k}^{n} a_{(i)} \) for \( k = 2, \ldots, n \) and \( \sum_{i=1}^{n} a_{(i)} = \sum_{i=1}^{n} b_{(i)} \), where the \( b_{(i)} \)'s and \( a_{(i)} \)'s are the components of \( b \) and \( a \), respectively, in ascending order. Write \( b \geq a \). A real-valued function \( \phi \) defined on \( \mathbb{R}^n \) is said to be Schur-convex (Schur-concave) if \( b \geq a \Rightarrow \phi(b) \geq \phi(a) \).

Motivated by this, Boland and Proschan (1986) defined majorization ordering.
Definition 1.8. [Boland and Proschan (1986)] Let \( X \) and \( Y \) be non-negative random variables with c.d.f.s \( F \) and \( G \). \( G \preceq F \) (\( \preceq \) for majorization) if \( \int_{x}^{\infty} [1-G(t)]dt \geq \int_{x}^{\infty} [1-F(t)]dt \) for all \( x \geq 0 \) and \( \int_{0}^{\infty} [1-F(t)]dt = \int_{0}^{\infty} [1-G(t)]dt = \mu < \infty \).

Remark 1.8.1. Ross (1983, p. 280) says \( Y \) is more variable than \( X \) if \( \int_{x}^{\infty} [1-G(t)]dt \geq \int_{x}^{\infty} [1-F(t)]dt \) for all \( x \geq 0 \) provided \( X \) and \( Y \) are non-negative random variables. A useful characterization of \( G \preceq F \) is that

\[
G \preceq F \iff \int_{0}^{\infty} \psi(t)dG(t) \geq \int_{0}^{\infty} \psi(t)dF(t)
\]

holds for all increasing convex functions \( \psi \), provided the integrals exist. Also, obviously \( G \preceq F \) implies \( G \gtrsim F \). It can be shown that \( G \preceq F \implies G \succeq F \) (Boland and Proschan (1986)).

Remark 1.8.2. Actually the above majorization ordering is a natural development from the idea of majorization. Using this idea, Boland and Proschan (1986) obtained the following inequalities: Assume that \( G_i \succeq F_i \) for \( i = 1, 2, \ldots, n \),

a) \( \int_{x}^{\infty} P[Y_{n1n} + \ldots + Y_{kn} > t]dt \geq \int_{x}^{\infty} P[X_{n1n} + \ldots + X_{kn} > t]dt \)

for all \( x \geq 0 \) and \( k = 1, 2, \ldots, n \),

b) \( (EY_{1n}, \ldots, EY_{n1n}) \succeq (EX_{1n}, \ldots, EX_{n1n}) \).

In this majorization context, Arnold (1987) introduced Lorenz ordering, which is a more theoretical construction than Boland and Proschan's (1986) ordering. However, Lorenz ordering is in fact exactly
the same as majorization ordering when two variables have equal means.

**Definition 1.9.** [Arnold (1987, p. 33)] \( X \preceq Y \) (i.e., \( X \) does not exhibit more inequality in the Lorenz sense than does \( Y \)), if \( L_X(u) \geq L_Y(u) \) for every \( u \in [0,1] \), where \( L_X(u) = \frac{\int_0^u F^{-1}(y) dy}{\int_0^1 F^{-1}(y) dy} \) and \( F^{-1}(y) = \sup \{ x : F(x) \leq y \} \), \( 0 < y < 1 \).

**Remark 1.9.1.** The assumption of Definition 1.9 is that \( X \) and \( Y \) be nonnegative random variables with finite means. If the expected values of \( X \) and \( Y \) are equal and positive, then \( X \preceq Y \) iff \( E[h(X)] \leq E[h(Y)] \) for every continuous convex function. Whitt (1980) showed that \( E[h(X)] \leq E[h(Y)] \) for any increasing convex function \( \iff \) \( E[u(X)] \leq E[u(Y)] \) for any convex function provided \( E[X] = E[Y] \). Hence, \( X \preceq Y \iff E[h(X)] \leq E[h(Y)] \iff E[u(X)] \leq E[u(Y)] \iff X \preceq Y \), provided \( E[X] = E[Y] \) and \( E[X] > 0 \). Furthermore, Lorenz ordering is easily shown to be a weaker ordering than dispersive ordering. But Lorenz ordering is more general than majorization ordering in that it permits distributions with different means. We get easily; \( X \preceq Y \iff \)

\[
E[g(X)] \leq E[g(Y)] \text{ for every continuous convex } g.
\]

Arnold (1987) showed that in a class of nonnegative random variables with positive finite expectations, \( X \preceq Y \Rightarrow X \preceq Y \). (Note that in this case it does not matter whether \( E[X] \) and \( E[Y] \) are different or not.)

**Remark 1.9.2.** And similar to the coupling argument, Arnold (1987)
showed that when $X$ and $Y$ are nonnegative random variables
with $E[X] = E[Y]$, $X \leq Y$ iff $\exists$ jointly distributed random variables
$X', Z'$, such that $X \overset{d}{=} X'$ and $Y \overset{d}{=} E[X'|Z']$, where $X \overset{d}{=} Y$ implies $X$ and $Y$
have the same distributions.

Since if $X \leq Y$, then $F_X(\lambda x) - F_Y(\lambda x)$ has at most one sign change
(from $-\rightarrow +$) as $x$ ranges from 0 to $\infty$, we can define intermediate
ordering between *-shaped ordering and Lorenz ordering.

**Definition 1.10.** [Arnold (1987, p. 79)] We will say that $X$ is sign-change
ordered with respect to $Y$ and write $X \leq_Y Y$, if
\[
\frac{F_X^{-1}(v)}{E(X)} - \frac{F_Y^{-1}(v)}{E(Y)} \text{ has at most one sign change (from + to -) as } v \text{ ranges from 0 to 1.}
\]

**Remark 1.10.** It is easy to prove $X \leq_Y Y \Rightarrow X \leq_L Y$ (see Theorem 6.4
of Arnold (1987)). Actually the beauty of Lorenz ordering or majorization
ordering is from the sign change property.

There is another ordering which is very useful by Karlin (1968).

**Definition 1.11.** [Karlin (1968)] Let $X$ and $Y$ denote continuous
random variables having respective densities $f$ and $g$. We say that $X$
is larger than $Y$ in the sense of likelihood ratio, and write $X >_\text{LR} Y$
if $\frac{f(x)}{g(x)} \leq \frac{f(y)}{g(y)}$ for all $x \leq y$.

**Remark 1.11.** It can be shown that $X >_\text{LR} Y \implies X >_\text{st} Y$. Also, Ross (1983,
p. 260) showed that $X \geq Y \Rightarrow \frac{f_X(t)}{1-F_X(t)} \leq \frac{f_Y(t)}{1-F_Y(t)}$ for all $t \geq 0$, provided
X and Y are nonnegative random variables. (Note \( \frac{f(t)}{1-F(t)} \) is the failure rate at time t.) Chan et al. (1983) used an equivalent form of \( X \geq Y \) to define another convex ordering.

**Nonnegative Case**

1. Convex Ordering → *-Shaped Ordering → Majorization Ordering or Lorenz Ordering
2. Dispersive Ordering → Stochastic Ordering

(1): By Remark 1.4.
(2): By Theorem 2.3, if \( f(0) \geq g(0) > 0 \).
(3): By Remark 1.9.1, if \( \mu(F) = \mu(G) \).
(4): By Remark 1.6.

**Symmetric Case**

6. s-Ordering → r-Ordering → Dispersive Ordering

(6): By Remark 1.5.2.
(7): By Theorem 2.7, if \( f(0) \geq g(0) > 0 \).

**General Case**

Likelihood Ratio Ordering → Stochastic Ordering

(8): By Remark 1.11.

Figure 1.1. The hierarchy of stochastic relationships
1.4 Positive Dependence Concepts

Lehmann (1966) initiated the systematic study of types of dependence in the bivariate case. In the context of reliability theory, Esary et al. (1967) introduced association to obtain bounds related to coherent systems. Barlow and Proschan (1975) provides an excellent review of results on positive dependence. Multivariate generalizations of the notions of positive dependence were initiated by Harris (1970). Block and Ting (1981) obtained multivariate versions of the various notions of positive dependence. Also, Shaked (1982b) obtained some related results. Karlin and Rinott (1980a, b) summarized previous results and gave a clear development of multivariate dependence concepts.

In this section, we will deal with positive dependence concepts in two parts: bivariate and multivariate.

1.4.1 Bivariate case

The following definitions are mainly from Barlow and Proschan (1975).

Definition 1.12. Given random variables X and Y, we say the following;

(a) X and Y are positively quadrant dependent if

$$P[X \leq x, Y \leq y] \geq P[X \leq x]P[Y \leq y]$$

for all x and y.

We write PPD(X,Y).

(b) Y is left tail decreasing in X if

$$P[Y \leq y | X \leq x]$$


is decreasing in $x$ for all $y$.

We write $\text{LTD}(Y|X)$.

(c) $Y$ is right tail increasing in $X$ if

$$P(Y > y | X > x)$$

is increasing in $x$ for all $y$.

We write $\text{RTI}(Y|X)$.

(d) $Y$ is stochastically increasing in $X$ if

$$P(Y > y | X = x)$$

is increasing in $x$ for all $y$.

We write $\text{SI}(Y|X)$.

(e) Let $X$, $Y$ have joint probability density $f(x,y)$. Then, $f(x,y)$ is totally positive of order $n$ if

$$\begin{vmatrix} f(x_1, y_1) & \ldots & f(x_1, y_r) \\ \vdots & \ddots & \vdots \\ f(x_r, y_1) & \ldots & f(x_r, y_r) \end{vmatrix} \geq 0$$

for all $x_1 < \ldots < x_r$, $y_1 < y_2 < \ldots < y_r$, $r = 1, 2, \ldots, n$ in the domain of $X$ and $Y$. We write $\text{TP}_n(X,Y)$. A function which is totally positive for all finite orders is said to be totally positive (TP) function.

(f) Random variables $X_1, X_2, \ldots, X_n$ are associated if

$$\text{cov}(f(x), g(x)) \geq 0$$
for all pairs of increasing functions \( f, g \).

\( A(X,Y) \) implies that \( X \) and \( Y \) are associated random variables.

**Remark 1.12.** Lehmann (1966) proved that

\[
P\{x < X, Y < y \} > P\{x < x\}P\{Y < y\} \quad \text{for all } x \text{ and } y
\]

\[
\iff P\{x > X, Y > y\} > P\{x > x\}P\{Y > y\} \quad \text{for all } x \text{ and } y
\]

\[
\iff P\{x > X, Y < y\} > P\{x > x\}P\{Y < y\} \quad \text{for all } x \text{ and } y
\]

\[
\iff P\{x < X, Y > y\} < P\{x < x\}P\{Y > y\} \quad \text{for all } x \text{ and } y.
\]

Instead of (d), \( SI(Y|X) \), Lehmann (1966) says that \( Y \) is *positively regression dependent* on \( X \). Also, Esary et al. (1967) showed that

\( PQD(X, Y) \iff \text{cov}[h_1(X), h_2(Y)] \geq 0 \) for all nondecreasing functions, \( h_1 \) and \( h_2 \).

**Definition 1.13.** [Harris (1970)] Random variables \( X \) and \( Y \) are said to be *right corner set increasing* if \( P\{X > x, Y > y | X > x', Y > y'\} \) is increasing in \( x' \) and \( y' \) for each fixed \( x, y \). We write \( RCSI(X,Y) \).

These bivariate notions of positive dependence are arranged in a hierarchy in the following figure (Barlow and Proschan (1975)).
Figure 1.2. Hierarchy of positive dependence in the bivariate case
1.4.2. **Multivariate case**

Multivariate versions of positive dependence will now be discussed. Let \( f(x_1, x_2, \ldots, x_n) \) be a probability density on \( B \), where \( B = \bigcap_{i=1}^{n} B_i \) and \( B_i \) is a Borel subset of \( \mathbb{R} \). Define \( X_{A'Y} = (x_{A'Y_1}, \ldots, x_{A'Y_n}) \) and \( X_{Y'} = (x_{1Y'_1}, \ldots, x_{nY'_n}) \), where \( A\cap = \min(a, b) \), \( A\cup = \max(a, b) \).

**Definition 1.14.** [Karlin and Rinott (1980a)] A random vector \( X = (X_1, \ldots, X_n) \) which has probability density function (or probability function), \( f \) on \( B \) is called **multivariate totally positive of order two** (\( \text{MTP}_2 \)) if \( f(x)f(y) < f(xy)f(x_{A'Y}) \), for every \( x, y \in B \). (Karlin (1968) called this \( \text{TP}_2 \) when \( n = 2 \).)

**Remark 1.14.** Any measurable function \( g \) is called a \( \text{MTP}_2 \) function on \( \chi \) if \( g(x)g(y) \leq g(xy)g(x_{A'Y}) \) for every \( x, y \in \chi \), where \( \chi \) is a partially ordered space. Karlin and Rinott (1980a) developed the fundamental properties of \( \text{MTP}_2 \) functions as follows:

(a) If \( f(x_1, x_2, \ldots, x_n) \) is \( \text{MTP}_2 \), then any marginal is also \( \text{MTP}_2 \).

(b) If \( f \) and \( g \) are \( \text{MTP}_2 \), then \( fg \) is also \( \text{MTP}_2 \).

(c) Independent random variables have a joint \( \text{MTP}_2 \) p.d.f. (or p.f.).

(d) If the random vector \( (X_1', \ldots, X_n') \) has a joint \( \text{MTP}_2 \) p.d.f. (or p.f.) and \( \phi_1, \phi_2, \ldots, \phi_n \) are all increasing (or all decreasing) functions, then the random vector \( (\phi_1(X_1'), \ldots, \phi_n(X_n')) \) has a joint \( \text{MTP}_2 \) p.d.f. (or p.f.).

**Definition 1.15.** [Barlow and Proschan (1975, p. 149)] A random vector \( \tilde{X} = (X_1, \ldots, X_n) \) with probability density function (or probability
function) \( f \) on \( B \) is called \textit{totally positive of order two in pairs} (\( \text{TP}_2 \) in pairs) if each pair of arguments is \( \text{TP}_2 \) when the remaining variables are fixed.

\begin{remark}
Kemperman (1977) showed that \( \text{TP}_2 \) in pairs implies \( \text{MTP}_2 \) and he gave a counter example of \( \text{MTP}_2 \Rightarrow \text{TP}_2 \) in pairs. However, if we assume \( f(x) > 0 \) for all \( x \in \mathbb{R} \), then \( \text{MTP}_2 \Leftrightarrow \text{TP}_2 \) in pairs.
\end{remark}

\begin{definition}
[Barlow and Proschan (1975, p. 146)]
(a) A random variable \( Y \) is \textit{stochastically increasing in random variables} \( X_1, X_2, \ldots, X_k \) if \( P(Y > y | X_1 = x_1, \ldots, X_k = x_k) \) is increasing in \( x_1, x_2, \ldots, x_k \). We write \( Y \uparrow \text{st} \) in \( X_1, X_2, \ldots, X_k \).
(b) Random variables \( X_1, X_2, \ldots, X_n \) are \textit{conditionally increasing in sequence} if \( X_i \) is stochastically increasing in \( x_1, x_2, \ldots, x_{i-1} \) for \( i = 2, 3, \ldots, n \).
\end{definition}

\begin{remark}
Barlow and Proschan (1975, p. 149) showed that \( \text{TP}_2 \) in pairs \( \Rightarrow \) conditionally increasing in sequence \( \Rightarrow \) associatedness. Esary et al. (1967) showed that associatedness is preserved under (a) taking subsets, (b) forming unions of independent sets, (c) forming sets of nondecreasing functions, (d) taking limits in distribution. Also, it can be shown that \( X_1, \ldots, X_n \) are associated iff
\[
P(x \in A \cap B) \geq P(x \in A)P(x \in B) \tag{1.4.1}
\]
whenever \( A \) and \( B \) are open upper sets (\( U \) is an upper set if \( x \in U \) and \( y \geq x \) imply \( y \in U \)).
Definition 1.17. [Shaked (1982b)] If for every $\mathbf{x} = (x_1, x_2, \ldots, x_n)$

\[ P(\mathbf{x} > \mathbf{y}) \geq \prod_{i=1}^{n} P(x_i > y_i), \quad (1.4.2) \]

then we say that $\mathbf{x} = (x_1, \ldots, x_n)$ is positively upper orthant dependent (PUOD), and if for every $\mathbf{y}$

\[ P(\mathbf{x} < \mathbf{y}) \geq \prod_{i=1}^{n} P(x_i < y_i), \quad (1.4.3) \]

then we say that $\mathbf{x} = (x_1, \ldots, x_n)$ is positively lower orthant dependent (PLOD).

Remark 1.17. When $n = 2$, $(X_1, X_2)$ is PUOD iff $(X_1, X_2)$ is PLOD by

Remark 1.12. Using (1.4.1), Shaked (1982b) discussed a general theory of concepts of positive dependence, which are weaker than association but stronger than orthant dependence. Dykstra et al. (1973) showed that if the $X_i$'s are associated, then they are PUOD and PLOD.

Joag-Dev (1983) defined a weaker condition than association for discussing the characterization of independence via uncorrelatedness.

Definition 1.18. [Joag-Dev (1983)] Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a random vector. A will denote an arbitrary proper subset of the index set $1, 2, \ldots, n$, $\overline{A}$ its complement and $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ a vector of constants. Then, a vector $\mathbf{x}$ is said to be strongly positively orthant dependent (SPOD) if for every $\mathbf{A}$ and $\mathbf{c}$, the following three conditions hold:
and \[ P[X \leq C_i, i \in A] P[X_j \leq C_j, j \in A] \leq P[X \leq C_i, i \in A] P[X_j \leq C_j, j \in A]. \] (1.4.6)

The vector $X$, on the other hand, is said to be strongly negatively orthant dependent (SNOD) if the reverse inequalities between the left and right sides of (1.4.4), (1.4.5) and (1.4.6) hold for every $C$.

**Remark 1.18.** Joag-Dev (1983) showed that Association $\Rightarrow$ SPOD $\Rightarrow$ PUOD and PLOD.

### 1.5 Miscellaneous Definitions

Barlow et al. (1963) treat properties of distributions with monotone failure rate.

**Definition 1.19.** [Barlow et al. (1963)] Let $X$ have distribution function $F$ with p.d.f. $f$. If $\frac{f(x)}{1-F(x)}$ is nondecreasing in $x$, then we say that $F$ is an *increasing failure rate (IFR)* distribution or $X$ IFR. If $\frac{f(x)}{1-F(x)}$ is nonincreasing in $x$, then $F$ is a *decreasing failure rate (DFR)* distribution or $X$ DFR.

**Remark 1.19.** It can be shown that $F$ is IFR (DFR) iff

\[
\frac{F(x+z)-F(x)}{1-F(x)}
\]

is nondecreasing (nonincreasing) in $x$ for all $z > 0$ whenever the denominator is nonzero. Also, $\frac{f(t)}{1-F(t)}$ is called the *failure rate*
(\gamma(t)) at time t.

**Definition 1.20.** [Schoenberg (1951)] A Polya frequency function of order two (PF₂) is a nonnegative measurable function g(x) defined for all real x, such that

\[
\frac{g(x_1-y_1) g(x_1-y_2)}{g(x_2-y_1) g(x_2-y_2)} > 0
\]

whenever \( x_1 < x_2 \) and \( y_1 < y_2 \) and \( g(x) \neq 0 \) for at least two distinct values of x.

**Remark 1.20.1.** Two alternative definitions are (a) \( g(x) \) is PF₂ if \( \log g(x) \) is concave on \( -\infty < x < \infty \), and (b) \( g(x) \) is PF₂ if, for fixed \( z > 0 \), \( \frac{g(x+z)}{g(x)} \) is a nonincreasing function of x in the interval \( (a,b) \), where

\[
a = \inf_{g(y)>0} \{ y \} \quad \text{and} \quad b = \sup_{g(y)>0} \{ y \}.
\]

**Remark 1.20.2.** A PF₂ is not necessarily a p.d.f. Rao (1986) summarized the properties of PF₂ density below:

(a) A density \( f(x) \) is PF₂ iff, for all \( z \), \( \frac{F(x+z)-F(x)}{f(x)} \) is decreasing in x.

(b) If \( f(x) \) is PF₂, then \( f(x) \) is unimodal.

(c) If \( F(x) \) is IFR, then its survival function \( 1-F(x) \) is PF₂, and conversely.

(d) If \( f(x) \) is a PF₂ density of a positive random variable, then
F(x) is IFR. The converse is not true.

(e) If log f(x) is concave (convex), then F(x) is IFR (DFR).

Remark 1.20.3. The following densities are PP:

(a) Normal: \( f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} \), \(-\infty < x < \infty\).

(b) Gamma: \( g_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \), \( x \geq 0 \) for \( \alpha \geq 1 \).

(c) Weibull: \( f_{\alpha,\lambda}(x) = \alpha \lambda x^{\alpha-1} e^{-(\lambda x)^\alpha} \), \( x \geq 0 \), \( \alpha \geq 1 \).

(d) Laplace: \( f(x) = \frac{1}{2} e^{-|x|} \), \(-\infty < x < \infty\).

(e) Truncated Normal: \( f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-C}{\sigma}\right)^2} \), \( \sigma > 0 \), \(-\infty < x < C\).

where \( \sigma > 0 \), \(-\infty < \mu < \infty\), \( a = \int_{-\infty}^{C} \frac{1}{\sqrt{2\pi}\sigma} e^{-1/2\left(\frac{y-C}{\sigma}\right)^2} dy \).

Remark 1.20.4. Karlin and Rinott (1980a, p. 481) prove that if \( \mathbf{X} = (X_1, \ldots, X_n) \) be a random vector of independent components \( X_1, \ldots, X_n \), each \( X_i \), \( i = 1, 2, \ldots, n \), governed by a PP density and \( \mathbf{Y} = (Y_1, \ldots, Y_n) \) have a joint MTP density on \( \mathbb{R}^n \), then \( \mathbf{Z} = \mathbf{X} + \mathbf{Y} \) has a MTP density, provided \( \mathbf{X} \) and \( \mathbf{Y} \) are independent.

Definition 1.21. [Lewis and Thompson (1981)] A distribution \( F \) is dispersive if \( F*G < F*H \) whenever \( G < H \). (Note * is the convolution operation.)

Definition 1.22. [Ibragimov (1956)] A distribution \( F \) is strongly unimodal if for every unimodal \( G \), the convolution \( F*G \) is unimodal.
Remark 1.22. Ibragimov (1956, p. 255) proves that an absolutely continuous distribution \( F \) is strongly unimodal iff its p.d.f. \( f \) has PF sub-2 density (i.e., \( \log f(x) \) is concave on \(-\infty < x < \infty\)). Also, Lewis and Thompson (1981, p. 88) showed that an absolutely continuous \( F \) is strongly unimodal iff it is dispersive. It can be shown that strongly unimodal \( \Rightarrow \) unimodal.

Definition 1.23. [Block et al. (1985)] A random vector \( X = (X_1, X_2, \ldots, X_n) \) is said to be negatively dependent through stochastic ordering (NDS) if \( \{(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n) \mid X_i = x\} \) stochastically decreases in \( x \) for all values of \( i = 1, 2, \ldots, n \).

Definition 1.24. [Shaked (1977)] Consider \( n \)-variate distribution functions which admit the representation

\[
F(x_1, x_2, \ldots, x_n) = \int_{\Omega} \prod_{i=1}^{n} F(w_i)(x_i) d\tau(w)
\]

where \( \{F(w), w \in \Omega\} \) is a family of univariate d.f.s, \( \Omega \) is a subset of a finite dimensional Euclidean space and \( \tau \) is a d.f. on \( \Omega \). Such d.f.s, which are mixtures of independent \( n \)-variate d.f.s with equal marginals, will be called positively dependent by mixture (PDM).

Definition 1.25. [Karlin and Rinott (1980b)] A random vector \( X = (X_1, \ldots, X_n) \) which has probability density function (or probability function), \( f \) on \( \mathcal{B} \) is called multivariate reverse rule of order two (MRR2) if
Remark 1.25. But if \( f(x_1, \ldots, x_n) \) is MRR, then any marginal is not necessarily MRR. In view of this anomaly, we introduce the following definition.

Definition 1.26. [Karlin and Rinott (1980b)] An MRR density \( f(x) \) of the random vector \( X = (X_1, X_2, \ldots, X_n) \) is said to be strongly MRR (S-MRR) if for any set of PF functions \( \{ \phi_i \} \), each resulting marginal

\[
g(x_{v_1}, x_{v_2}, \ldots, x_{v_k}) = \int \cdots \int f(x_1, \ldots, x_n) \phi_1(x_{j_1}) \phi_2(x_{j_2}) \cdots \phi_{n-k}(x_{j_{n-k}}) \, dx_{j_1} \cdots dx_{j_{n-k}}
\]

is MRR in the variables \( x_{v_1}, x_{v_2}, \ldots, x_{v_k} \) where \( (v_1, \ldots, v_k) \) and \( (j_1, \ldots, j_{n-k}) \) are complementary sets of indices.

Remark 1.26. Examples include:

(a) the multinomial,
(b) multivariate hypergeometric,
(c) multivariate Hahn,
(d) Dirichlet family of densities on a simplex,
(e) $N(\mathbf{0}, \Sigma)$, where $\Sigma = \Lambda - \|\alpha_i \alpha_j\|_{ij=1}^n$ with $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, $\lambda_i > 0$, $\alpha_i > 0$, $i = 1, 2, \ldots, n$. Hence, equicorrelated multivariate case is $\text{S-MRR}_2$ when $-\frac{1}{n-1} \leq \rho < 0$.

**Definition 1.27.** [Shaked and Tong (1985)] Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ be two exchangeable random vectors. It follows that $X_1 \overset{st}{=} X_2 \overset{st}{=} \cdots \overset{st}{=} X_n$ and $Y_1 \overset{st}{=} Y_2 \overset{st}{=} \cdots \overset{st}{=} Y_n$, where $\overset{st}{=} \overset{st}{=}$ means equal in distribution. Assume $X_1 \overset{st}{=} Y_1$.

(a) $\mathbf{X} \overset{A}{>}_{\text{st}} \mathbf{X}$ (that is, $\mathbf{X}$ is more dispersed than $\mathbf{X}$ in the sense $A$) if

$$
\left| \sum_{i=1}^n C_i X_i \right| \geq \left| \sum_{i=1}^n C_i Y_i \right| \quad \text{where} \quad \sum_{i=1}^n C_i = 0.
$$

(b) $\mathbf{X} \overset{B}{>}_{\text{st}} \mathbf{Y}$ if for all $t \in \mathbb{R}$,

$$
(F_{\mathbf{X}}(t), \ldots, F_{\mathbf{X}}(t)) \overset{B}{\geq} (F_{\mathbf{Y}}(t), \ldots, F_{\mathbf{Y}}(t)).
$$

(c) $\mathbf{X} \overset{C}{>}_{\text{st}} \mathbf{Y}$ if

$$
(E_{\mathbf{X}}(t), \ldots, E_{\mathbf{X}}(t)) \overset{C}{\geq} (E_{\mathbf{Y}}(t), \ldots, E_{\mathbf{Y}}(t)).
$$

Note that $\mathbf{X} \overset{A}{>}_{\text{st}} \mathbf{Y}$ and $\mathbf{X} \overset{B}{>}_{\text{st}} \mathbf{Y}$ are labelled differently than in Shaked and Tong's definitions.
2. STOCHASTIC ORDERINGS AND EXPECTATIONS

2.1 Introduction

If the distribution function of $X$ is $F(x)$, then $F_{X_{1:n}}(x)$ or $F_X(x)$ denotes the distribution function of $X_{1:n}$ or $X(x)$. The mean or expected value of the random variable $X_{1:n}$ will be written as

$$E[X_{1:n}] = \int_{-\infty}^{\infty} x dF_{X_{1:n}}(x).$$

Then,

$$|E(X_{1:n})| = n^{\frac{n-1}{2}} \int_{-\infty}^{\infty} x [F(x)]^{\frac{1}{2}} (1-F(x))^{\frac{n-1}{2}} dF(x) \leq n^{\frac{n-1}{2}} \int_{-\infty}^{\infty} |x| dF(x),$$

which shows that $|E(X_{1:n})|$ exists provided $E|X|$ exists. In this chapter, we will assume that $E|X|$ exists whenever we treat the expected values of the order statistics.

Let $X_1, X_2, \ldots, X_n$ ($Y_1, Y_2, \ldots, Y_n$) be random samples with c.d.f.s $F_G$. Now for the two sample case, we have

$$E[Y_{1:n}] - E[X_{1:n}] = \int_{-\infty}^{\infty} y dG_{X_{1:n}}(y) - \int_{-\infty}^{\infty} x dF_{X_{1:n}}(x)$$

$$= \int_{-\infty}^{\infty} y \frac{1}{B(r, n-r+1)} [G(y)]^{r-1} [1-G(y)]^{n-r} dG(y) - E[X_{1:n}]$$

$$= \int_{-\infty}^{\infty} G^{-1} F(x) \frac{1}{B(r, n-r+1)} [F(x)]^{r-1} [1-F(x)]^{n-r} dF(x) - E[X_{1:n}]$$

(since $Y = G^{-1} F(X)$)
\[ = \int_{-\infty}^{\infty} [G^{-1}(F(x)) - x] dP_{r_{1:n}}(x) . \quad (2.1.1) \]

Hence, for the comparison of expected values of order statistics from two different populations, the function \( G^{-1}(F(x)) - x \) will play a prominent role in the above context. In this chapter, we will first consider the stochastic comparison of the order statistics from exchangeable random variables in the one sample case. Secondly, using various stochastic orderings, we will investigate the comparisons of the expected values of the order statistics and of the spacings from two populations.

2.2 Expectations - One Sample

Many authors have studied recurrence relations among expected values of order statistics (see, e.g., Chapter 3, David (1981)). Here we will discuss the comparison of order statistics from one sample. We need the following additional notation:

\[ X^*_r, \text{ the } r^{th} \text{ order statistic in any subset of } n \text{ exchangeable variates drawn from a larger set } X_1, X_2, \ldots, X_n, \text{ (n' > n)}, \]

each with marginal c.d.f. \( F(x) \).

Thus, \( X^*_r \) and \( X^*_r \) have the same marginal distribution. However, unless \( X^*_s \) is based on the same subset as \( X^*_r \), the joint distribution of \( X^*_r \) and \( X^*_s \) is not the same as that of \( X^*_r \) and \( X^*_s \). In particular, for \( r < s \),
\[ P(X_{r:n} > X_{s:n}) = 0 \]

but \[ P(X_{r:n}^* > X_{s:n}^*) > 0. \]

**Theorem 2.1.** Let \( X_1, X_2, \ldots, X_n \) be exchangeable random variables. Then, for \( n_1 < n_2 < \ldots < n_k \leq n \),

\[
X_{i:n}^* \leq X_{i+1:n}^* \leq \ldots \leq X_{n:n}^*, \tag{2.2.1}
\]

and

\[
X_{i:n_1}^* \geq X_{i:n_2}^* \geq \ldots \geq X_{i:n_k}^*. \tag{2.2.2}
\]

\[
X_{n-k+1:n_k}^* \geq X_{n-k+1:n_k-1}^* \geq \ldots \geq X_{1:n_1}^*. \tag{2.2.3}
\]

**Proof.** Since \( X_1, X_2, \ldots, X_n \) are exchangeable, \( X_{i:n}^* = X_{i:n} \) for \( i = 1, 2, \ldots, n \). But \( P(X_{i:n}^* \leq X_{j:n}^*) = 1 \) for \( i < j \). Hence, (2.2.1) follows by coupling (Ross (1983, p. 255)). Inequalities (2.2.2) and (2.2.3) follow similarly since, e.g.,

\[
P(X_{i:n_1}^* \geq X_{i:n_2}^*) = 1. \]

**Theorem 2.2.** Let \( X_1, X_2, \ldots, X_n \) be a random sample from a continuous population with c.d.f. \( F \). Let \( n_1 < n_2 \). Then,

\[
X_{r:n_1}^* \geq X_{s:n_2}^* \text{ for } r \geq s
\]

and
where \(1 \leq r \leq n_1\) and \(1 \leq s \leq n_2\).

**Proof.** Let \(l(x; r, s, n_1, n_2) = \frac{f_{r:n_1}(x)}{f_{r:n_2}(x)}\), where \(f_{r:n_1}(x)\) is p.d.f. of the \(r\)th order statistic from \(X_1, X_2, \ldots, X_{n_1}\). Since

\[
l(x; r, s, n_1, n_2) = \frac{n_1! (s-1)! (n_2-s)!}{n_2! (r-1)! (n_1-r)!} \frac{\frac{f^{r-s}(x)}{[1-F(x)]}}{n_2-n_1+x-s'},
\]

for \(r \geq s\), \(l(x; r, s, n_1, n_2)\) is increasing with respect to \(x\) and for \(r \leq s-n_2+n_1\), \(l(x; r, s, n_1, n_2)\) is decreasing in \(x\). \(\Box\)

**Remark 2.1.** Theorems 2.1 and 2.2 include as a special case of the corresponding results for the unstarred order statistics.

**Remark 2.2.** Since \(X \geq Y\) implies \(X \geq Y\), we can replace stochastic LR St ordering by likelihood ordering in Theorem 2.1 if \(X_1, X_2, \ldots, X_n\) are a random sample.

### 2.3 Expectations - Two Sample

Let \(X(Y)\) be a random variable with c.d.f. \(F(G)\). Since

\[
F_{r:n_1}(x) = \frac{1}{B(r, n-r+1)} \int_0^{F(x)} t^{r-1} (1-t)^{n-r-1} dt,
\]

\(F(x) \leq G(x)\ \forall x\) implies \(F_{r:n_1}(x) \leq G_{r:n_1}(x)\ \forall x\) (i.e., \(F \geq G\) \(\Rightarrow\) \(F \geq G\) \(\Rightarrow\))
Also, it is well known that \( F_{i:n} \geq G_{i:n} \) implies \( E[X_{i:n}] \geq E[Y_{i:n}] \) for \( i = 1, 2, \ldots, n \), provided the expectations exist (i.e., \( F \geq G \Rightarrow E[X_{i:n}] \geq E[Y_{i:n}] \)). From this simple result, stochastic relationships are closely related to the comparison of the expected values of order statistics from two different populations.

In fact, Ross (1983, p. 256) shows that for i.n.i.d. variates (independent and nonidentically distributed variates) \( X_1, X_2, \ldots, X_n \) and \( Y_1, Y_2, \ldots, Y_n \) with \( X_i \leq Y_i \) \( (i = 1, 2, \ldots, n) \), one has for any increasing function \( \phi \), that

\[
\phi(X_1, X_2, \ldots, X_n) \leq \phi(Y_1, Y_2, \ldots, Y_n).
\]

It follows that \( X_{i:n} \leq Y_{i:n} \) for \( i = 1, 2, \ldots, n \).

Doksum (1974) calls the function \( G^{-1}F(x) - x \) in (2.2.1) the shift function, since \( X_i \), when shifted by \( G^{-1}F(x) - x \), has the same distribution as \( Y_i \), i.e., \( G^{-1}F(x) \equiv Y_i \). Note that convexity of \( G^{-1}F(x) - x \) is equivalent to \( F \leq G \) (i.e., \( G^{-1}F(x) \) is a convex function). Similarly, for nonnegative random variables,

\[
\frac{G^{-1}F(x) - x}{x} \uparrow \text{w.r.t. } x \iff F \leq G \text{ (i.e., } \frac{G^{-1}F(x)}{x} \uparrow \text{w.r.t. } x).\]

Also, since \( \binom{n-r}{r-1} F^{r-1}(x) (1-F(x))^{n-r} \) is totally positive function in \( r \) and \( x \) (see Definition 1.12 (e)), the relation
shows that the number of sign changes in $E[Y_{r:n} - X_{r:n}]$ with $r$ is no greater than the number of sign changes in $G^{-1}F(x) - x$ as $x \to \infty$, by the variation diminishing property of totally positive functions (Karlin (1968, p. 21); Boland and Proschan (1986)). For example, if no sign change occurs (i.e., $G^{-1}(x) - x \geq 0$ for all $x$), then $E[Y_{r:n}] \geq E[X_{r:n}]$ for $r = 1, 2, \ldots, n$.

Oja (1981) discusses the shift function using the notions of convexity of order $k$. Barlow and Proschan (1966) obtain various stochastic inequalities for linear combinations of order statistics under the assumptions of $F \leq G$ and $F \leq G$. Also, Oja (1981) found a further stochastic inequality for linear combinations of order statistics under $F \leq G$.

From a property of *-shaped ordering (i.e., Remark 1.5.1), first we consider the case where $F$ and $G$ have different means because this case includes the case where $F$ and $G$ do not cross.

**Lemma 2.1.** Let $F$ and $G$ be absolutely continuous distributions, with $G$ strictly increasing on its support, and $F(0) = G(0) = 0$. If $F \leq G$ and $f(0) > g(0) > 0$, then $F(x) > G(x)$ for all $x$ (i.e., $F \leq G$).

**Proof.** $F \leq G \iff \frac{G^{-1}F(x)}{x} \uparrow w.r.t. x > 0$. Let $h(x) = \frac{G^{-1}F(x)}{x}$. Then, there are three possibilities:

(a) $h(x) < 1$ all finite $x$
(b) \( h(x) = 1 \) for some finite \( x \)

(c) \( h(x) > 1 \) all finite \( x \).

The corresponding c.d.f.s must look as follows:

---

Since \( f(0) > g(0) > 0 \), only (c) is possible.

Thus, \( F \preceq G \), \( f(0) > g(0) > 0 \) and \( \mu(F) = \mu(G) \) is impossible. Hence, the case of \( F \preceq G \) and \( f(0) > g(0) > 0 \) can occur only for different means.
Theorem 2.3. [Deshpande and Kochar (1983)] Let \( F \) and \( G \) be absolutely continuous such that \( F(0) = G(0) = 0 \) and let the corresponding density functions be such that \( f(0) \geq g(0) > 0 \). Then, \( F \leq G \) implies \( F \leq G \) (i.e., for any \( 0 < \alpha < \beta < 1 \), \( F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha) \)).

\[ \begin{align*}
\text{Proof.} \quad F \leq G & \iff G^{-1}F(x) \uparrow \text{w.r.t. } x \\
& \iff \frac{f(x)}{g(G^{-1}F(x))} \geq \frac{G^{-1}F(x)}{x} \quad \text{for } x > 0 \quad \text{(by differentiation)}
\end{align*} \]

Since

\[ \lim_{x \to 0^+} \frac{G^{-1}F(x)}{x} = \frac{f(0)}{g(0)} \geq 1 \quad \text{(by L'Hospital's rule)}, \]

\[ \Rightarrow \frac{f(x)}{g[G^{-1}F(x)]} \geq 1 \quad \text{for } x > 0 \]

\[ \Rightarrow g[F^{-1}(\alpha)] \geq g[G^{-1}(\alpha)] \quad \text{for } 0 < \alpha < 1 \quad \text{with } F(x) = \alpha \]

\[ \iff F < G \quad \text{(Remark 1.6)}, \]

Remark 2.3.

\[ G^{-1}F(x) \uparrow \text{w.r.t. } x \iff G^{-1}F(x) \uparrow \leq G^{-1}F(y) - y \quad \text{(for any } x < y) \]

\[ \iff y - x \leq G^{-1}F(y) - G^{-1}F(x) \]

\[ \iff F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha) \quad \text{with} \]

\[ F(x) = \alpha \text{ and } F(y) = \beta \quad \text{for } 0 < \alpha < \beta < 1, \]

\[ G^{-1}F(x) \uparrow \text{w.r.t. } x \iff F < G \quad \text{(i.e., } \frac{d}{dt}[G^{-1}F(t) - t] > 0 \iff \text{disp} \]

\[ F < G). \quad \text{This was proved by Deshpande and Kochar (1983). Hence, disp} \]
the dispersiveness ordering is related to the comparison of the expected values of the order statistics from two populations.

**Theorem 2.4.** [Oja (1981)] Let $F$ and $G$ be absolutely continuous distribution functions from random variables $X$ and $Y$. If $F < G$, then $[X_{s:n} - X_{r:n}]_{st} < [Y_{s:n} - Y_{r:n}]_{st}$ for any $1 \leq r < s \leq n$. 

**Proof.** By Remark 2.3 (i.e., $F < G \iff G^{-1}_F(x) \trianglerighteq w.r.t. x$), for any $1 \leq r < s \leq n$,

$$G^{-1}_F(X_{s:n}) - G^{-1}_F(X_{r:n}) \geq X_{s:n} - X_{r:n}.$$ 

Therefore, $P[G^{-1}_F(X_{s:n}) - G^{-1}_F(X_{r:n}) > a] \geq P[X_{s:n} - X_{r:n} > a] \forall a$,

since $X_{s:n} - X_{r:n} > a$ implies $G^{-1}_F(X_{s:n}) - G^{-1}_F(X_{r:n}) > a$ for all $a$.

Then, since $G^{-1}_F(X_{s:n}) \trianglerighteq Y_{s:n}$, the result follows immediately. $lacksquare$

**Remark 2.4.1.** Hence, if any two quantiles of $G$ are more widely separated than the corresponding quantiles of $F$, then the spacings of the $Y_{i:n}$ are stochastically larger than the corresponding spacings of the $X_{i:n}$.

**Remark 2.4.2.** Since $G^{-1}_F(t) = G^{-1}_{r:n} F_{r:n}(t)$ for any $r = 1,2,...,n$,

$$F < G \iff \frac{d}{dt} [G^{-1}_F(t) - t] > 0$$

$$\iff \frac{d}{dt} G^{-1}_{r:n} F_{r:n}(t) - t] > 0$$

$$\iff F_{r:n} \trianglerighteq G_{r:n}$$
We have considered the case when $F$ and $G$ do not cross. Next, we discuss the case of exactly one crossing. Marshall and Proschan (1970) treated the single crossing of $F(x)$ and $G(x)$ as $x; 0 \rightarrow \infty$ with respect to *-shaped ordering.

**Theorem 2.5.** [Marshall and Proschan (1970)] Let $X_i^*(Y_i^*)$, the non-negative random variables, have distribution $F_i^*(G_i^*)$ with common mean $\mu_i^*$, and $F_i^* < G_i^*$, $i = 1, 2, \ldots, n$. Let $X_1, X_2, \ldots, X_n$ be associated and $Y_1, Y_2, \ldots, Y_n$ mutually independent. Then,

$$E[X_{1n}] \geq E[Y_{1n}] \tag{2.3.1}$$

$$E[X_{n1}] \leq E[Y_{n1}] \tag{2.3.2}$$

**Remark 2.5.** If $F \leq G$ and $\mu(F) = \mu(G)$, then by Remark 1.5.1 (b), $F(x)$ crosses $G(\theta x)$ exactly once, and from above, as $x$ increases from 0 to $\infty$, for each $\theta > 0$. Then, by Karlin's (1968) variation diminishing property, the sign change of $E[X_{r1n} - Y_{r1n}]$ in $r = 1, 2, \ldots, n$ is at most once, but (2.3.1) and (2.3.2) show that there exists $c$ such that

$$E[X_{r1n}] \geq E[Y_{r1n}] \text{ for } r < c$$

and

$$E[X_{r1n}] \leq E[Y_{r1n}] \text{ for } r > c.$$  

In the case of exactly one crossing of $F(x)$ and $G(x)$ as $x \rightarrow \infty$, we may consider the case such that there exists $\theta \geq G^{-1}F(x) - x \leq 0$
respectively for \( x \geq a \). Let us define \( F \leq_1 G \) if there exists \( a \) such that \( G^{-1}(x-a) \leq 0 \), respectively for \( x \leq a \). (Note \( \leq_1 \) is a special case of Oja's (1981, Definition 4.2,)) Then, Oja (1981) showed that if \( F \leq_1 G \), then \( E[C(X)] \leq E[C(Y)] \) for any convex function \( C \) where \( X \sim F \) and \( Y \sim G \) provided \( \mu(F) = \mu(G) \) (Oja (1981, Theorem 4.3, p. 159)). As stated in Remark 1.8.1, \( F \leq G \) (i.e., majorization ordering, Boland and Proschan (1986)) \( \iff \) \( E[C(X)] \leq E[C(Y)] \) for any convex function \( C \) provided \( \mu(F) = \mu(G) \). Hence, \( F \leq_1 G \) implies \( F \leq_1 G \) provided \( \mu(F) = \mu(G) \) and \( F \) and \( G \) are distribution functions of nonnegative continuous random variables. Also, Oja (1981) showed that \( F < G \rightarrow F \leq_1 G \) (Theorem 4.2, p. 158) (i.e., \( F < G \Rightarrow F \leq_1 G \Rightarrow F < G \) provided \( \mu(F) = \mu(G) \)). (Note Oja's (1981) \( < \) is different from our \( < \).) Also, \( F \leq G \iff \sum L \leq G \) provided \( \mu(F) = \mu(G) > 0 \). As mentioned in Remark 1.8.2,

\[
F \leq G \Rightarrow (a) \int_x^{\infty} P[X_{n,1} + \cdots + X_{k,1} > t] dt \leq \int_x^{\infty} P[Y_{n,1} + \cdots + Y_{k,1} > t] dt
\]

for all \( x \geq 0 \) and \( k = 1, 2, \ldots, n \).

(b) \( E[X_{1,1}; \ldots, E[X_{n,1}]] \leq (E[Y_{1,1}; \ldots, E[Y_{n,1}]] \).

Now (b) implies \( E[X_{n,1} ; E[Y_{n,1} ] \leq E[Y_{1,1} ] \) and \( E[X_{1,1} ] \geq E[Y_{1,1} ] \) but \( E[Y_{1,1}] - E[X_{1,1}] \) may theoretically at least undergo many sign changes as \( i, 1 \to n \). Recall that if \( F \leq G \) and \( \mu(F) = \mu(G) \), then there is exactly one sign change.

Furthermore, Oja (1981) compares the shift function (i.e., \( G^{-1}(x-a) - x \)) with the linear function \( ax + b \) (Oja (1981, Definition 5.3,
Intuitively, the expected values of order statistics depend on skewness, and peakedness or heaviness of tail. Birnbaum (1948) introduced the peakedness ordering: \( X \) is more peaked than \( Y \) if \( P(|X| > t) < P(|Y| > t) \) for all \( t > 0 \). This is just \( |X| < |Y| \). Also, this is equivalent to \( F(t) > G(t) \) for all \( t > 0 \) provided \( X \) and \( Y \) are random variables symmetric about zero, with c.d.f.s \( F \) and \( G \).

**Theorem 2.6.** Let \( X \) and \( Y \) be random variables symmetric about 0, with c.d.f.s \( F \) and \( G \). If \( X \) is more peaked than \( Y \) (i.e., \( G(x) \leq F(x) \) for all \( x > 0 \)), then

\[
E[Y_{x1n}] \geq E[X_{x1n}] \quad \text{for} \quad \frac{1}{2}(n+1) \leq x \leq n.
\]

**Proof.** From David (1981, p. 38),

\[
E[Y_{x1n}] - E[X_{x1n}] = \int_{0}^{\infty} [G_{n-r+11n}(x) - F_{n-r+11n}(x)] - G_{x1n}(x) + F_{x1n}(x) \, dx.
\]

Now it suffices to show that

\[
F_{x1n}(x) - G_{x1n}(x) \geq F_{n-r+11n}(x) - F_{n-r+11n}(x) \quad \text{for} \quad x > 0, r \geq \frac{1}{2}(n+1).
\]

\[
\text{LHS} = I_{F(x)}(r,n-r+1) - I_{G(x)}(r,n-r+1), \quad \text{where} \quad I \text{ denotes on Incomplete Beta function}
\]

\[
= \int_{0}^{1} \frac{t^{r-1}(1-t)^{n-r}}{B(r,n-r+1)} \, dt.
\]
Similarly,

\[ \text{RHS} = \int_0^F(x) \frac{t^{n-r}(1-t)^{r-1}}{G(x) B(r, n-r+1)} \, dt. \]

Since \( t^{r-1}(1-t)^{n-r} = \left( \frac{t}{1-t} \right)^{2r-1-n} \geq 1 \) for \( r \geq \frac{1}{2}(n+1) \) and \( t \geq \frac{1}{2} \), we have

\[ E[Y_{r:n}] \geq E[X_{r:n}] \quad \text{for} \quad \frac{1}{2}(n+1) \leq r \leq n. \]

**Remark 2.6.** Since

\[ F_{r:n}(x) \leq G_{r:n}(x) \iff F(x) \leq G(x) \quad r = 1, 2, \ldots, n \]

i.e., iff \( x \leq 0 \)

it follows that neither of \( X_{r:n}, Y_{r:n} \) is stochastically larger than the other.

**Corollary 2.6.** For \( \frac{1}{2}(n+1) \leq s \leq n \) and \( 1 \leq r \leq \frac{1}{2}(n+1) \), \( E[Y_{s:n} - Y_{r:n}] \geq E[X_{s:n} - X_{r:n}] \) under the same assumptions as Theorem 2.6.

The proof follows immediately from symmetry considerations. ■

**Example 2.1.** The T distribution with \( m \) degrees of freedom is symmetrically distributed about 0 with p.d.f.,

\[ f_m(t) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi} \Gamma\left(\frac{m}{2}\right)} \left(1 + \frac{t^2}{m}\right)^{-\frac{(m+1)}{2}} \quad -\infty < t < \infty. \]
For $m = 1$, the density of $T$ reduces to Cauchy density, having no absolute moments of any integral order. As $m \to \infty$, the density function, $f_m$, approaches that of a standard normal variate. If $F_{m+1}(x) > F_m(x)$ for all $x > 0$ (i.e., $T_{m+1}$ is more peaked than $T_m$), then by the Theorem 2.6 and Corollary 2.6, we have

$$E[T_m(s)] \geq E[T_{m+1}(s)] \text{ for } \frac{1}{2}(n+1) \leq s \leq n$$

and

$$E[T_m(s) - T_m(x)] \geq E[T_{m+1}(s) - T_{m+1}(x)] \text{ for } 1 \leq r \leq \frac{1}{2}(n+1)$$

and $\frac{1}{2}(n+1) \leq s \leq n$.

Before we prove $F_{m+1}(x) > F_m(x)$ for all $x > 0$, we need the following two lemmas.

**Lemma 2.2.** $f_{m+1}(0) > f_m(0)$.

**Proof.**

$$f_m(0) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{\pi m} \Gamma\left(\frac{m}{2}\right)} = \frac{1}{\sqrt{\pi} B\left(\frac{1}{2}, \frac{m}{2}\right)}$$

Similarly, $f_{m+1}(0) = \frac{1}{\sqrt{\pi} B\left(\frac{1}{2}, \frac{m+1}{2}\right)}$.

It suffices to prove that

$$\frac{1}{f_m(0)} - \frac{1}{f_{m+1}(0)} > 0 \text{ (i.e., } \sqrt{m} B\left(\frac{1}{2}, \frac{m}{2}\right) - \sqrt{m+1} B\left(\frac{1}{2}, \frac{m+1}{2}\right) > 0).$$

$$\sqrt{m} B\left(\frac{1}{2}, \frac{m}{2}\right) - \sqrt{m+1} B\left(\frac{1}{2}, \frac{m+1}{2}\right)$$

$$= \sqrt{m} \int_0^1 x^{\frac{1}{2} - 1} (1-x)^{\frac{m}{2} - 1} dx - \sqrt{m+1} \int_0^1 x^{\frac{1}{2} - 1} (1-x)^{\frac{m+1}{2} - 1} dx$$
\[
\int_0^1 x^{m-1} (1-x)^{\frac{1}{2} - 1} \left( \frac{1}{\sqrt{m}} - \frac{1}{\sqrt{m+1}}(1-x)^{\frac{1}{2}} \right) dx.
\]

Since \( \sqrt{m} - \sqrt{m+1}(1-x)^{\frac{1}{2}} \) is strictly convex on \((0,1)\), by Jensen's inequality the integral is greater than

\[
\sqrt{m} - \sqrt{m+1}(1-E(x))^{\frac{1}{2}} = 0 \quad \text{(by } E(x) = \frac{1}{m+1}).
\]

Hence, \( f_{m+1}(0) > f_m(0) \) for any \( m = 1, 2, \ldots \)

Lemma 2.3. \( f_{m+1}(1) > f_m(1) \).

Proof. It suffices to prove that

\[
\frac{1}{f_m(1)} - \frac{1}{f_{m+1}(1)} > 0.
\]

\[
\frac{1}{f_m(1)} - \frac{1}{f_{m+1}(1)} = \sqrt{m} B\left(\frac{1}{2}, \frac{m}{m+1}\right) \left(1 + \frac{1}{m}\right)^{\frac{1}{2}}
- \sqrt{m+1} B\left(\frac{1}{2}, \frac{m+1}{m+2}\right) \left(1 + \frac{1}{m+1}\right)^{\frac{1}{2}})
\]

\[
= \int_0^1 x^{\frac{1}{2} - 1} (1-x)^{\frac{1}{2} - \frac{1}{m} - \frac{1}{m+1}} \left[\sqrt{m} + \frac{1}{2}(m+1)
- \sqrt{m+1}(1-x)^{\frac{1}{2}} \left(1 + \frac{1}{m+1}\right) \right] dx.
\]

(Since \( \sqrt{m} + \frac{1}{2} \frac{1}{m}(m+1) \) is strictly convex on \((0,1)\), by Jensen's inequality),

\[
> \sqrt{m} + \frac{1}{2} \frac{1}{m}(m+1) - \sqrt{m+1}(1-E(x))^\frac{1}{2} \left(1 + \frac{1}{m+1}\right) \frac{1}{2}(m+2)
\]
\[
\sqrt{m\left(1 + \frac{1}{m}\right)^2} - \sqrt{m\left(1 + \frac{1}{m+1}\right)^2}.
\]

Now it suffices to prove \((1 + \frac{1}{x})^{\frac{1}{2}}x + \frac{1}{2}\) is a decreasing function on \(1 \leq x < \infty\).

\[
\frac{d}{dx}\left[(1 + \frac{1}{x})^{\frac{1}{2}}x + \frac{1}{2}\right] = \frac{d}{dx}\left[(\frac{1}{2}x^{\frac{1}{2}} + \frac{1}{2})\log(1 + \frac{1}{x})\right](1 + \frac{1}{x})^{\frac{1}{2}}x + \frac{1}{2}
\]

\[
= \left[\frac{1}{2}\log(1 + \frac{1}{x}) - (\frac{1}{2}x^{\frac{1}{2}} + \frac{1}{2})\frac{x^{-2}}{1 + \frac{1}{x}}\right](1 + \frac{1}{x})^{\frac{1}{2}}x + \frac{1}{2}
\]

\[
= \frac{1}{2}[\log(1 + \frac{1}{x}) - \frac{1}{x}](1 + \frac{1}{x})^{\frac{1}{2}}x + \frac{1}{2}
\]

\(< 0 \text{ (by } \log(1 + \frac{1}{x}) - \frac{1}{x} < 0 \text{ for any } 1 \leq x < \infty)\)

\[\therefore F_{m+1}(1) > F_m(1) \text{ for any } m = 1, 2, 3, \ldots \]

Claim. \(F_{m+1}(x) \geq F_m(x) \text{ for all } x \geq 0.\)

Proof.

\[
\frac{f_m(t)}{f_{m+1}(t)} = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m+1}\Gamma\left(\frac{m}{2}\right)} \frac{(1 + t^{\frac{1}{2}})^{\frac{1}{2}(m+1)}}{(1 + t^{\frac{1}{2}})^{\frac{1}{2}(m+2)}} - \infty < t < \infty.
\]
By the symmetric property of $t$ distribution, we consider only positive values of $t$,

$$
= c \frac{(m+1+t^2)^{\frac{1}{2}(m+2)}}{(m+t^2)^{\frac{1}{2}(m+1)}}
$$

where $c = \frac{\Gamma\left(\frac{m+2}{2}\right) \frac{1}{2}(m+1)}{\sqrt{m\pi} \Gamma\left(\frac{m}{2}\right) \frac{1}{2}(m+2)}$

$$
\frac{d}{dt}[f_m(t)]
\frac{1}{(m+t^2)^{\frac{1}{2}(m+1)}} \frac{d}{dt}[\frac{1}{(m+t^2)^{\frac{1}{2}(m+1)}}] - \frac{d}{dt}[\frac{1}{(m+t^2)^{\frac{1}{2}(m+1)}}] \frac{1}{(m+t^2)^{\frac{1}{2}(m+1)}}
$$

\[\text{Numerator} = t(m+2)(m+t^2)^{\frac{1}{2}(m+2)}(m+1+t^2)^{\frac{1}{2}m}
- t(m+1)(m+t^2)^{\frac{1}{2}(m-1)}(m+1+t^2)^{\frac{1}{2}(m+2)}
= t(m+1+t^2)^{\frac{1}{2}m}(m+t^2)^{\frac{1}{2}(m-1)}[(m+2)(m+t^2) - (m+1)(m+t^2+1)]
= t(t^2-1)(m+t^2)^{\frac{1}{2}(m-1)}(m+t^2+1)^{\frac{1}{2}m}.
$$

Hence,
Figure 2.2. The shape of \( t \)-distribution with \( m \) and \( m+1 \) degrees of freedom

From Lemma 2.2, Lemma 2.3 and (2.3.3), we have \( F_{m+1}(x) \geq F_m(x) \) for \( x \geq 0 \) and for \( m = 1, 2, \ldots \).

Using Tiku and Kumra (1985), we can see these results illustrated in Table 2.1.

**Theorem 2.7.** Let \( X \) and \( Y \) be random variables symmetric about 0, with c.d.f.s \( F \) and \( G \). Then, if \( F \leq G \) and \( f(0) \geq g(0) > 0 \), then \( \frac{X_{s,n} - X_{r,n}}{f_{s,n} - f_{r,n}} \leq \frac{Y_{s,n} - Y_{r,n}}{f_{s,n} - f_{r,n}} \) for any \( 1 \leq r < s \leq n \).
Table 2.1. Expected values of order statistics of t-distributions from $m = 3$ degrees of freedom to $m = 19$ and standard normal distribution in sample size, 10 ($n = 10$) (Tiku and Kumra (1985))

<table>
<thead>
<tr>
<th>$m$</th>
<th>$E[T_{6;10}]$</th>
<th>$E[T_{7;10}]$</th>
<th>$E[T_{8;10}]$</th>
<th>$E[T_{9;10}]$</th>
<th>$E[T_{10;10}]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.1395858</td>
<td>0.3432024</td>
<td>0.7862594</td>
<td>1.2980512</td>
<td>2.5283165</td>
</tr>
<tr>
<td>4</td>
<td>0.1350104</td>
<td>0.4181056</td>
<td>0.7490249</td>
<td>1.2073913</td>
<td>2.1719769</td>
</tr>
<tr>
<td>5</td>
<td>0.1323781</td>
<td>0.4089456</td>
<td>0.7283123</td>
<td>1.1590206</td>
<td>2.0028543</td>
</tr>
<tr>
<td>6</td>
<td>0.1306693</td>
<td>0.4030378</td>
<td>0.7151333</td>
<td>1.1289959</td>
<td>1.9046147</td>
</tr>
<tr>
<td>7</td>
<td>0.1294709</td>
<td>0.3989131</td>
<td>0.7060142</td>
<td>1.108583</td>
<td>1.8405509</td>
</tr>
<tr>
<td>8</td>
<td>0.1285844</td>
<td>0.3958706</td>
<td>0.6993311</td>
<td>1.0937537</td>
<td>1.7955165</td>
</tr>
<tr>
<td>9</td>
<td>0.1279020</td>
<td>0.3935343</td>
<td>0.6942235</td>
<td>1.0825376</td>
<td>1.7621470</td>
</tr>
<tr>
<td>10</td>
<td>0.1273606</td>
<td>0.3916840</td>
<td>0.6901936</td>
<td>1.0737472</td>
<td>1.7364377</td>
</tr>
<tr>
<td>11</td>
<td>0.1269207</td>
<td>0.3901824</td>
<td>0.6869328</td>
<td>1.0666731</td>
<td>1.7160262</td>
</tr>
<tr>
<td>12</td>
<td>0.1265561</td>
<td>0.3889395</td>
<td>0.6842404</td>
<td>1.0608576</td>
<td>1.6994301</td>
</tr>
<tr>
<td>13</td>
<td>0.1262493</td>
<td>0.3878938</td>
<td>0.6819797</td>
<td>1.0559925</td>
<td>1.6856720</td>
</tr>
<tr>
<td>14</td>
<td>0.1259869</td>
<td>0.3870019</td>
<td>0.6800546</td>
<td>1.0518625</td>
<td>1.6740820</td>
</tr>
<tr>
<td>15</td>
<td>0.1257604</td>
<td>0.3862320</td>
<td>0.6783956</td>
<td>1.0483127</td>
<td>1.6641855</td>
</tr>
<tr>
<td>16</td>
<td>0.1255630</td>
<td>0.3855608</td>
<td>0.6769510</td>
<td>1.0452289</td>
<td>1.6556368</td>
</tr>
<tr>
<td>17</td>
<td>0.1253892</td>
<td>0.3849705</td>
<td>0.6756820</td>
<td>1.0425251</td>
<td>1.6481783</td>
</tr>
<tr>
<td>18</td>
<td>0.1252351</td>
<td>0.3844473</td>
<td>0.6745581</td>
<td>1.0401351</td>
<td>1.6416140</td>
</tr>
<tr>
<td>19</td>
<td>0.1250975</td>
<td>0.3839803</td>
<td>0.6735561</td>
<td>1.0380073</td>
<td>1.6357924</td>
</tr>
<tr>
<td>$\infty$</td>
<td>0.1226678</td>
<td>0.3757647</td>
<td>0.6560591</td>
<td>1.0013570</td>
<td>1.5387527</td>
</tr>
</tbody>
</table>
Proof. From the argument of Theorem 2.3, if \( f(0) \geq g(0) > 0 \) and

\[
\frac{G^{-1}(x)}{x} \uparrow \text{w.r.t.} \ x > 0, \text{ then } F^{-1}(-\beta) - F^{-1}(-\alpha) \leq G^{-1}(-\beta) - G^{-1}(-\alpha) \text{ for any } \frac{1}{2} < \alpha < \beta < 1.
\]

For any \( 0 < \beta' < \alpha' < \frac{1}{2} \), \( 3 \) \( \alpha \) and \( \beta' \div \frac{1}{2} < \alpha < \beta < 1 \) (i.e., \( \alpha = 1-\alpha' \) and \( \beta = 1-\beta' \)), \( F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha) \),

\[
F^{-1}(\beta)-F^{-1}(\alpha) = F^{-1}(\alpha')-F^{-1}(\beta') \text{ and } G^{-1}(\beta)-G^{-1}(\alpha) = G^{-1}(\alpha')-G^{-1}(\beta').
\]

Hence, for any \( 0 < \beta' < \alpha' < \frac{1}{2} \), \( F^{-1}(\alpha')-F^{-1}(\beta') = F^{-1}(\beta)-F^{-1}(\alpha) \leq G^{-1}(\beta)-G^{-1}(\alpha) = G^{-1}(\alpha')-G^{-1}(\beta') \). Hence, for any \( 0 < \beta' < \alpha' < \frac{1}{2} \),

\[
F^{-1}(\alpha')-F^{-1}(\beta') \leq G^{-1}(\alpha')-G^{-1}(\beta'). \text{ It suffices to prove that for any } 0 < \alpha' < \frac{1}{2} < \beta < 1, F^{-1}(\beta)-F^{-1}(\alpha') \leq G^{-1}(\beta)-G^{-1}(\alpha'). \text{ But } F^{-1}(\beta)-F^{-1}(\alpha') = F^{-1}(\beta)-F^{-1}(\beta') + F^{-1}(\alpha') \leq G^{-1}(\beta)-G^{-1}(\beta') + G^{-1}(\alpha') = G^{-1}(\beta)-G^{-1}(\alpha'), \text{ Hence, } F^{-1}(\beta)-F^{-1}(\alpha) \leq G^{-1}(\beta)-G^{-1}(\alpha) \text{ for any } 0 < \alpha < \beta < 1. \]

Simply applying Theorem 2.4, we have

\[
[x_{s:n} - x_{r:n}] \leq [x_{s:n} - x_{r:n}] \quad \text{for any } 1 \leq r < s \leq n.
\]
Example 2.2. Van Zwet (1964) (see Definition 1.4) showed that any symmetric U-shaped density \( \leq \) uniform \( \leq \) normal \( \leq \) logistic \( \leq \) Laplace \( \leq \) Cauchy (see Van Zwet (1964, p. 70-73)). Hence, if we change scale in order to meet the assumption, \( f(0) \geq g(0) > 0 \), then U-shaped density < uniform < normal < logistic < Laplace < Cauchy. (Note: By Theorem 2.4, these relationships imply the stochastic ordering among the spacings.)
3. DEPENDENCE STRUCTURE OF ORDER STATISTICS

3.1 Introduction

Lehmann (1966) initiated the systematic study of types of dependence in the bivariate case. In the context of reliability theory, Esary et al. (1967) introduced association to obtain bounds related to coherent systems (see Section 1.4). Barlow and Proschan (1975) provide an excellent review of results. Multivariate generalizations of the notions of positive dependence were initiated by Harris (1970). Block and Ting (1981) obtained multivariate versions of the various notions of positive dependence. Also, Shaked (1982b) obtained some related results. Karlin and Rinott (1980a,b) summarized previous results and gave a clear development of multivariate dependence concepts.

Using the foregoing positive dependence concepts (see Section 1.4), we will investigate in this chapter the dependence structure of order statistics and furthermore the covariance structure of order statistics.

Define the inverse of a cumulative distribution function (c.d.f.) \( F \) as \( F^{-1}(u) = \inf\{x \in \mathbb{R} | F(x) \geq u\} \) where \( \mathbb{R} \) is the set of real numbers and \( 0 < u < 1 \).

Throughout this chapter, we will use "increasing" in place of "nondecreasing" and "decreasing" in place of "nonincreasing."

3.2 Dependence Structure—Order Statistics

Bickel (1967) showed \( \text{cov}(X_{i:n}, X_{j:n}) \geq 0 \), provided \( \text{EX}_{1:n}^2 + \text{EX}_{j:n}^2 < \infty \) whenever \( X_1, X_2, \ldots, X_n \) are a random sample from a population with c.d.f.
F(x) and p.d.f. f(x), the latter being continuous and strictly positive on \([x|0 < F(x) < 1]\). Esary et al. (1967) noted that \(X_1, X_2, \ldots, X_n\) being increasing functions of \(X_1, X_2, \ldots, X_n\), are associated when \(X_1, X_2, \ldots, X_n\) are mutually independent random variables. This implies that \(\text{cov}(X_{i:n}, X_{j:n}) \geq 0\) for \(1 \leq i \leq j \leq n\), where \(X_1, X_2, \ldots, X_n\) are independent random variables, not necessarily identically distributed. The assumptions of continuity and strict positivity on \([x|0 < F(x) < 1]\) are not needed. Also, Lehmann (1966) mentioned that \((X_{i:n}, X_{j:n})\) is TP_2, when \(X_1, X_2, \ldots, X_n\) is a random sample from a continuous population p.d.f. (As indicated in Figure 1.2, TP_2 \(\Rightarrow\) positive covariance.)

**Theorem 3.1.** Let \(X_1, X_2, \ldots, X_n\) be a random sample of i.i.d. random variables, \(X_1\) having a probability density function \(f\). Then, the joint density of the order statistics \(X_{1:n}, X_{2:n}, \ldots, X_{n:n}\) is MTP_2.

**Proof.**

\[
\prod_{i=1}^{n} f_{X_{1:n}, \ldots, X_{n:n}}(x_1, x_2, \ldots, x_n) = n! \prod_{i=1}^{n} g(x_1, x_2, \ldots, x_n) f(x_i)
\]

where

\[
g(x_1, \ldots, x_n) = \begin{cases} 
1 & x_1 < x_2 < \ldots < x_n \\
0 & \text{otherwise}
\end{cases}
\]

If we prove \(g(x_1, x_2, \ldots, x_n)\) is MTP_2, then it follows that the product \(\prod_{i=1}^{n} f(x_i)\) is MTP_2 (by Remark 1.14(b) and (c)). Suppose \(g\) is not a MTP_2 function. Then, \(g(x)g(y) = 1\) and \(g(xy)g(xy) = 0\) holds for some \(x\) and \(y\). But \(g(x)g(y) = 1 \iff g(x) = 1\) and
Corollary 3.1. Let $X_1, X_2, \ldots, X_n$ be a random sample of i.i.d. discrete random variables $X_i$ having a probability function (p.f.) $f$. Then, the joint p.f. of the order statistics $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ is MTP$_2$.

Proof. 

For $x_1, x_2, \ldots, x_n$, 

$$f_{X_{1:n}, \ldots, X_{n:n}}(x_1, x_2, \ldots, x_n) = \frac{n!}{x_1! x_2! \cdots x_k!} g(x_1, \ldots, x_n) \prod_{i=1}^{n} f(x_i)$$

where

$$g(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & x_1 \leq x_2 \leq \ldots \leq x_n \\ 0 & \text{otherwise} \end{cases}$$

and $x_{1:n} = x_{2:n} = \ldots = x_{k:n} < x_{k+1:n} = \ldots = x_{1+2-1:n} < x_{1+2:n} \leq x_{n:n}$ and $\Sigma r_i = n$.

This proof is the same as for Theorem 3.1. □

Remark 3.1.1. Hence, order statistics from either continuous or discrete populations are MTP$_2$. Thus, by Remark 1.14(a) $(X_{1:n}, X_{j:n})$ is TP$_2$ for $1 \leq i < j \leq n$. This implies $\text{cov}[f(X_{1:n}, X_{j:n}), g(X_{1:n}, X_{j:n})] \geq 0$, where $f$ and $g$ are both increasing (or decreasing) functions. Also, if $f_{X_{1:n}, \ldots, X_{n:n}}(x_1, \ldots, x_n) > 0$ for all $x$, then $X_{1:n}, \ldots, X_{n:n}$ is TP$_2$ in pairs by Remark 1.15.

Remark 3.1.2. Theorem 3.1, without reference to the discrete case, was proved by Karlin and Rinott (1980a).
Using these powerful dependence concepts, we will investigate
the dependence structure of the concomitants of order statistics.
We may express our stochastic model in the form

\[ Y_i = g(X_i, Z_i), \quad i = 1, 2, \ldots, n \]

where all \( X \)'s and \( Z \)'s are mutually independent. Then, for \( r = 1, 2, \ldots, n \),

\[ Y_{r \upharpoonright n} = g(X_{r \upharpoonright n}, Z_{[r]}), \quad (3.2.1) \]

where \( X_{r \upharpoonright n} \) and \( Z_{[r]} \) are independent and \( Z_{[1]}, \ldots, Z_{[n]} \) are mutually
independent.

**Theorem 3.2.** When \( g \) is an increasing (or a decreasing) function on
\( \mathbb{R} \), then \( Y_{[1:n]}, \ldots, Y_{[n:n]} \) are associated.

**Proof.** Since \( (X_{1:n}, X_{2:n}, \ldots, X_{n:n}) \) and \( (Z_{[1]}, \ldots, Z_{[n]}) \) are independent
sets of associated random variables under the model (3.2.1), their
union is also associated (Remark 1.16(b)). Since any increasing (or
decreasing) functions of associated random variables are associated
(Remark 1.16(c)), \( Y_{[1:n]}, Y_{[2:n]}, \ldots, Y_{[n:n]} \) are associated, provided \( g \)
is an increasing (or decreasing) function on \( \mathbb{R} \). \( \blacksquare \)

A slightly stronger result holds for an additive model such as

\[ Y_{r \upharpoonright n} = g(X_{r \upharpoonright n}) + Z_{[r]}, \quad r = 1, 2, \ldots, n. \quad (3.2.2) \]

If \( g \) is a monotone function, increasing or decreasing, then \( g(X_{1:n}) \),
\( g(X_{2:n}), \ldots, g(X_{n:n}) \) are associated and hence, \( Y_{[1:n]}, \ldots, Y_{[n:n]} \)
are associated. An important example is given by the linear model

\[ Y_i = \mu_Y + \rho \sigma_X(x_i - \mu_X) + \epsilon_i, \]

for which \( Y_{[r:n]} = \mu_Y + \frac{\sigma_Y}{\sigma_X} (x_{[r:n]} - \mu_X) + \epsilon_{[r]}, \) \( r = 1, 2, \ldots, n, \) where \( E(X) = \mu_X, \) \( E(Y) = \mu_Y, \) \( \text{var}(X) = \sigma_X^2, \) \( \text{var}(Y) = \sigma_Y^2 \) and \( \text{corr}(X,Y) = \rho. \)

**Theorem 3.3.** Let \( g \) be monotone function under the model (3.2.2), and let \( (Z_{[1]}, Z_{[2]}, \ldots, Z_{[n]}) \) be a random vector of independent components, each \( Z_{[i]}, i = 1, 2, \ldots, n, \) governed by a Polya frequency function of order two (PF2) (see Definition 1.20). Then, \( Y_{[l:n]}, \ldots, Y_{[n:n]} \) have a MTP2 density.

**Proof.** Since \( (X_{[1:n]}, X_{[2:n]}, \ldots, X_{[n:n]}) \) has MTP2 density, \( [g(x_{[1:n]}), \ldots, g(x_{[n:n]})] \) is MTP2 provided \( g \) is a monotone function (see Remark 1.14(d)). Hence, results follow by Remark 1.20.4. ■

**Theorem 3.4.** Let \( (X_i, Y_i) (i = 1, 2, \ldots, n) \) be \( n \) independent random variables having a common bivariate distribution. If \( Y \) is stochastically increasing in \( X, \) then \( Y_{[x:n]} \) is stochastically increasing in \( x_{[x:n]} \).

**Proof.** The result follows at once from the fact that

\[ f_{Y_{[x:n]}|x_{[x:n]} = x}(y) = f_{Y|X=x}(y) \]

so that

\[ P[Y_{[x:n]} > y|x_{[x:n]} = x] = P[Y > y|x = x], \] ■
Finally, we consider relaxation of the independence of Theorem 3.1 and Corollary 3.1.

**Theorem 3.5.** If \((X_1, X_2, \ldots, X_n)\) is exchangeable and MTP\(_2\), then \((X_1, X_2, \ldots, X_n)\) is MTP\(_2\).

**Proof.** By the exchangeability of \(X_1, \ldots, X_n\), all orderings of \(X_1, X_2, \ldots, X_n\) have the same probabilities, namely \(\frac{1}{n!}\). Hence,

\[
f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = n! g(x_1, \ldots, x_n) f_{X_1, \ldots, X_n}(x_1, x_2, \ldots, x_n)
\]

where

\[
g(x_1, \ldots, x_n) = \begin{cases} 1 & x_1 \leq \ldots \leq x_n \\ 0 & \text{otherwise} \end{cases}
\]

and \(f_{X_1, \ldots, X_n}\) is p.d.f. of \((X_1, X_2, \ldots, X_n)\). Since \(g(x_1, \ldots, x_n)\) is MTP\(_2\), it follows that the product \(g(x_1, \ldots, x_n) f_{X_1, \ldots, X_n}(x_1, \ldots, x_n)\) is MTP\(_2\) (see Remark 1.14(b)).

**Remark 3.4.** This theorem includes the result of Theorem 3.1 and Corollary 3.1 since i.i.d. variates are exchangeable and MTP\(_2\).

### 3.3 Covariance-Order Statistics

We investigated the dependence structure of order statistics in Section 3.2. Using these results, we consider the structure of
covariances of order statistics. Problems involving comparisons among covariances of order statistics have been studied by Tukey (1958).

Note that if $P_{2}(X,Y)$, then $Y$ is positively regression dependent on $X$, which implies $\text{cov}(X,Y) \geq 0$ (see Figure 1.2). Hence, for any $1 \leq r < s \leq n$, $X_{s:n}$ is positively regression dependent on $X_{r:n}$, since $P_{2}(X_{r:n}, X_{s:n})$.

Theorem 3.6. Let $F$ be the c.d.f. of $X$. If $F$ is IFR (see Definition 1.20), then $X_{s:n} - X_{r:n}$ is negatively regression dependent on $X_{r:n}$ for $1 \leq r < s \leq n$.

Proof. We have to show that $P[X_{s:n} - X_{r:n} \leq z | X_{r:n} = x]$ is increasing in $x$ for any $z > 0$. Now for a random sample of $n$ from a continuous parent, the conditional distribution of $X_{s:n}$ given $X_{r:n} = x$ ($s > r$), is just the distribution of the $(s-r)^{th}$ order statistics in a sample of $n-r$ drawn from $\frac{f(y)}{1-F(x)} \ (y \geq x)$, i.e., from the parent distribution truncated on the left at $x$ (see, e.g., David (1981, p. 20)).

Hence,

$$
\Pr[X_{s:n} - X_{r:n} \leq z | X_{r:n} = x] = \Pr[X_{s:n} \leq x+z | X_{r:n} = x]
$$

$$
= \sum_{i=s-r}^{n-r} \binom{n-r}{i} \left( \frac{F(x+z) - F(x)}{1-F(x)} \right)^{i-1} \left( \frac{1-F(x+z)}{1-F(x)} \right)^{n-r-i}
$$

$$
= \frac{1}{B(s-r,n-r+1)} \int_{0}^{F(x+z)-F(x)} t^{s-r-1}(1-t)^{n-r} dt.
$$

(3.3.1)
Then by Remark 1.19, (3.3.1) is increasing in $x$ for any $z > 0$ if $F$ is IFR. ■

**Remark 3.6.1.** Similarly, it follows that if $F$ is DFR (see Definition 1.19), then $X_{s:n} - X_{r:n}$ is positively regression dependent on $X_{r:n}$.

**Remark 3.6.2.** Tukey (1958) stated Theorem 3.6 incorrectly, writing decreasing instead of nondecreasing. Furthermore, Lehmann (1966) repeated Tukey's incorrect statement.

**Lemma 3.2.** [Tukey (1958)] If $W$ is negatively regression dependent on $Z$ and $Z$ is positively regression dependent on $Y$, and the distribution of $W$ given $Z$ is unaffected by $Y$, then $W$ is negatively regression dependent on $Y$.

**Proof.** By hypothesis, for any $y' < y''$,

$$P[Z \leq z | Y = y'] \geq P[Z \leq z | Y = y'']$$

$$\Rightarrow F_{Z|Y=y'}^{-1}(u) \leq F_{Z|Y=y''}^{-1}(u), \text{ where } 0 < u < 1. \quad (3.3.2)$$

Also by hypothesis, for any $z' < z''$,

$$P[W \leq w | Z = z'] \leq P[W \leq w | Z = z''].$$

Since the distribution of $W$ given $Z$ is unaffected by $Y$, we have

$$P[W \leq w | Z_{Y=y'}, = z'] \leq P[W \leq w | Z_{Y=y''} = z''].$$
so that by (3.3.2), \( P[w \leq w|z|y^*] = F_{Z|Y=y^*}^{-1}(u) \leq P[w \leq w|z|y^\infty] = F_{Z|Y=y^\infty}^{-1}(u) \). Writing

\[
H(w|y') = P[w \leq w|y = y'] ,
\]

\[
H(w|y') = \int_0^1 P[w \leq w|z|y = y'] = F_{Z|Y=y'}^{-1}(u)du
\leq \int_0^1 P[w \leq w|z|y = y^\infty] = F_{Z|Y=y^\infty}^{-1}(u)du
= H(w|y^\infty) .
\]

Correspondingly, if \( W \) is positively regression dependent on \( Z \), and \( Z \) is positively regression dependent on \( Y \), and the distribution of \( W \) given \( Z \) is unaffected by \( Y \), then \( W \) is positively regression dependent on \( Y \).

The following theorem was stated incorrectly in Tukey (1958).

**Theorem 3.7.** Let \( F \) be the distribution function of \( X \). If \( X \) is IFR, then \( X_{s:n} - X_{r:n} \) is negatively regression dependent on \( X_{k:n} \) for any \( k \) satisfying \( 1 \leq k < r < s \leq n \) (this implies \( \text{cov}(X_{k:n}, X_{s:n}) \leq \text{cov}(X_{k:n}, X_{r:n}) \)).

**Proof.** From Theorem 3.6, if \( F \) is IFR, then \( X_{s:n} - X_{r:n} \) is negatively regression dependent on \( X_{r:n} \) for \( r < s \). And for \( k < r \), \( X_{r:n} \) is positively regression dependent on \( X_{k:n} \). From the Markov property of order statistics (see, e.g., David (1981, p. 20)),

\[
f_{X_{s:n}|X_{r:n}=x}(r), X_{k:n}=x(k) = f_{X_{s:n}|X_{r:n}=x}(y) .
\]

(3.3.3)
Now by the direct application of Lemma 3.2, if X is IFR, then \( X_{s1:n} - X_{r1:n} \)
is negatively regression dependent on \( X_{k1:n} \) for any \( 1 \leq k < r < s \leq n \).

**Remark 3.7.1.** Similarly, if X is DPR, then \( X_{s1:n} - X_{r1:n} \) is positively regression dependent on \( X_{k1:n} \) for any \( 1 \leq k < r < s \leq n \).

**Corollary 3.7.** Let \( F \) be a continuous distribution function of \( X \). If \( X \) and \(-X\) are IFR, then the covariance of any two order statistics is less than the variance of either, and the covariance between order statistics \( X_{r1:n} X_{s1:n} \) is monotone in \( r \) and \( s \) separately, decreasing as \( r \) and \( s \) separate from one another.

**Proof.** By Theorem 3.7, if \( X \) is IFR, then, for \( r \) fixed, the covariance between the order statistics \( X_{r1:n} X_{s1:n} \) is monotone in \( s \), decreasing as \( s \) increases and the covariance of \( X_{r1:n} X_{s1:n} \) is less than the variance of \( X_{r1:n} \). Let \( Y = -X \). If \( Y \) is IFR, by Theorem 3.7, \( Y_{s1:n} - Y_{r1:n} \) is negatively regression dependent on \( Y_{k1:n} \) for any \( k \) satisfying \( 1 \leq k < r < s \leq n \). Since \( Y_{i1:n} = -X_{n-1+i1:n} \) if \(-X\) is IFR, then

\[
X_{n-r1:n} - X_{n-s1:n}
\]
is negatively regression dependent on \( -X_{n-k1:n} \) for

\[
X_{S1:n} - X_{r1:n}
\]
is positively regression dependent on \( X_{k1:n} \), where

\[
s' = n-r+1, \quad r' = n-s+1, \quad \text{and} \quad k' = n-k+1.
\]

On dropping the primes, it follows that \( \text{cov}(X_{r1:n}, X_{k1:n}) \leq \text{cov}(X_{s1:n}, X_{k1:n}) \) for \( r < s \leq k \).

**Remark 3.7.2.** However, if \( X \) is DFR (i.e., \( \frac{f(x)}{1-F(x)} \downarrow \) in \( x \uparrow \)), then \( f(x) \) is a decreasing function in \( x \) since \( 1-F(x) \) is a decreasing function in \( x \). Hence, \( \frac{f(x)}{F(x)} \) is a decreasing function in \( x \), which means that \(-X\) is IFR. So the case of \( X \) DFR and \(-X\) DPR does not exist.
Remark 3.7.3. But if $X$ is IFR and $-X$ DFR, then, for $r$ fixed, the covariance between the order statistics $X_{r:n}$, $X_{s:n}$ is monotone in $s$, decreasing as $s$ increases and the covariance of $X_{r:n}$, $X_{s:n}$ is less than the variance of $X_{r:n}$. But for $s$ fixed, the covariance between the order statistics $X_{r:n}$, $X_{s:n}$ is monotone in $r$, increasing as $r$ decreases and the covariance of $X_{r:n}$, $X_{s:n}$ is greater than the variance of $X_{s:n}$. Similar results hold for the case of $X$ DFR and $-X$ IFR.

Numerical illustrations are provided by the Gamma distribution, $f(x) = \frac{1}{\Gamma(r)} x^{r-1} e^{-x}$, for which, for $r > 1$, $X$ and $-X$ are IFR (Table 3.1), and by the Pareto distribution for which $X$ is DFR but $-X$ is IFR (Table 3.2).

Polya-type distributions, which were introduced by Schoenberg (1951), have been applied extensively in several domains of mathematics, statistics, economics, and mechanics. Rao (1986) summarized previous results. Here we introduce a property of $PF_2$ density (see Definition 1.20, Remark 1.20.1, Remark 1.20.2, and Remark 1.20.3).

Theorem 3.8. Let $F$ be the continuous distribution function of $X$ with p.d.f. $f$. If $f$ is $PF_2$ (i.e., strongly unimodal, see Definition 1.22 and Remark 1.22), then the covariance of any two order statistics is less than the variance of either, and the covariance between order statistics $X_{r:n}$, $X_{s:n}$ is monotone in $r$ and $s$ separately, decreasing as $r$ and $s$ separate from one another.
Table 3.1. Variances and covariances of the $\xi$th and $\eta$th order statistics in a sample of size 7 from $f(x) = e^{-x}$ for $x > 0$
i.e., Gamma (1,2) (Prescott (1974))

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\xi$</th>
<th>$m$</th>
<th>cov($X_{\xi:n}, X_{\eta:n}$)</th>
<th>$\xi$</th>
<th>$m$</th>
<th>cov($X_{\xi:n}, X_{\eta:n}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0.1202</td>
<td>3</td>
<td>4</td>
<td>0.2325</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0.0962</td>
<td>3</td>
<td>5</td>
<td>0.2155</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0.0841</td>
<td>3</td>
<td>6</td>
<td>0.2012</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
<td>0.0763</td>
<td>3</td>
<td>7</td>
<td>0.1870</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
<td>0.0705</td>
<td>4</td>
<td>4</td>
<td>0.3584</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>6</td>
<td>0.0657</td>
<td>4</td>
<td>5</td>
<td>0.3331</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>7</td>
<td>0.0610</td>
<td>4</td>
<td>6</td>
<td>0.3115</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.1819</td>
<td>4</td>
<td>7</td>
<td>0.2899</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0.1598</td>
<td>5</td>
<td>5</td>
<td>0.5306</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>0.1454</td>
<td>5</td>
<td>6</td>
<td>0.4975</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>5</td>
<td>0.1345</td>
<td>5</td>
<td>7</td>
<td>0.4641</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>6</td>
<td>0.1254</td>
<td>6</td>
<td>6</td>
<td>0.8952</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>7</td>
<td>0.1165</td>
<td>6</td>
<td>7</td>
<td>0.8384</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.2549</td>
<td>7</td>
<td>7</td>
<td>2.2467</td>
</tr>
</tbody>
</table>
Table 3.2. Variances and covariances of the $z^{th}$ and $m^{th}$ order statistics in a sample of size 7 from $f(x) = 3a^3x^{-4}$, $a > 0$, $x \geq a$ (Malik (1966))

<table>
<thead>
<tr>
<th>n</th>
<th>$z$</th>
<th>m</th>
<th>$\text{cov}(X_{z_{1:n}}'X_{m_{1:n}})$</th>
<th>$z$</th>
<th>m</th>
<th>$\text{cov}(X_{z_{1:n}}'X_{m_{1:n}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0.0028</td>
<td>3</td>
<td>4</td>
<td>0.0173</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.0029</td>
<td>3</td>
<td>5</td>
<td>0.0194</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0.0031</td>
<td>3</td>
<td>6</td>
<td>0.0233</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0.0034</td>
<td>3</td>
<td>7</td>
<td>0.0349</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>0.0038</td>
<td>4</td>
<td>4</td>
<td>0.0331</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>0.0046</td>
<td>4</td>
<td>5</td>
<td>0.0372</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>0.0069</td>
<td>4</td>
<td>6</td>
<td>0.0446</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.0074</td>
<td>4</td>
<td>7</td>
<td>0.0669</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.0079</td>
<td>5</td>
<td>5</td>
<td>0.0764</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>0.0086</td>
<td>5</td>
<td>6</td>
<td>0.0917</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.0097</td>
<td>5</td>
<td>7</td>
<td>0.1375</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>0.0117</td>
<td>6</td>
<td>6</td>
<td>0.2429</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>0.0175</td>
<td>6</td>
<td>7</td>
<td>0.3643</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.0158</td>
<td>7</td>
<td>7</td>
<td>3.0367</td>
<td></td>
</tr>
</tbody>
</table>
Proof. By Corollary 3.7, it suffices to prove that if \( f \) is \( PF_2 \), then \( X \) and \(-X\) are IFR. By Remark 1.19, \(-X\) is IFR if and only if \( \frac{f(x)}{F(x)} \) is decreasing in \( x \) by the dual to the failure rate. Now choosing \( z = \infty \) and \( z = -\infty \) in (1.5.1), we conclude that \( PF_2 \) density implies both \( X \) IFR and \(-X\) IFR. \( \square \)

Remark 3.8. Examples of \( PF_2 \) densities in Remark 1.20.3 have the interesting covariance structure of order statistics given by Theorem 3.8.

3.4 Covariance-Spacings

If \( X \) is distributed exponentially, then spacings \( D_1, D_2, \ldots, D_n \) are mutually independent and each \( D_k \) is distributed exponentially (see, e.g., David (1981, p. 21)).

Theorem 3.9. Let \( F \) be the distribution function of \( X \). If \( F \) is IFR, then \( D_s \) is negatively regression dependent on \( D_r \) (this implies \( \text{cov}(D_r, D_s) \leq 0 \) for any \( 1 \leq r < s \leq n \)).

Proof. \( P[D_s \leq w|D_r = x] = P[X_{s;n} - X_{s-1;n} \leq w|X_{r;n} - X_{r-1;n} = x] \)

\[
= \int_{-\infty}^{\infty} P[X_{s;n} - X_{s-1;n} \leq w|X_{r;n} = x+y, X_{r-1;n} = y] dF_{r-1;n}(y)
\]

\[
= \int_{-\infty}^{\infty} P[X_{s;n} - X_{s-1;n} \leq w|X_{r;n} = x+y]dF_{r-1;n}(y)
\]

(by (3.3.3)).

If \( F \) is IFR, then by Theorem 3.7, for any \( 0 < x_1 < x_2 \) and any fixed \( y \),
(3.4.1)\]

From (3.4.1), \( P[D_S \leq w | D_x = x] \) is increasing in \( x \) for any \( w \), i.e., \( D_S \) is negatively regression dependent on \( D_x \).

**Remark 3.9.1.** Similarly, it follows that if \( F \) is DFR, \( D_S \) is positively regression dependent on \( D_x \), which implies \( \text{cov}(D_x, D_S) > 0 \).

**Remark 3.9.2.** In life testing, a statistic that plays a central role is the total time on test. Assume \( n \) items are placed on test at time \( 0 \) and that successive failures are observed at times \( X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n} \). If we stop at the \( r \)th failure, then the total time on test is

\[
\tau(t) = \sum_{i=1}^{r} (n-i+1)D_i.
\]

If \( F \) is distributed exponentially, then

\[
\text{var}(\tau(t)) = \sum_{i=1}^{r} (n-i+1)^2 \text{var}(D_i).
\]

Theorem 3.9 and Remark 3.9.1 shows that

\[
\text{var}_{IFR}(\tau(t)) \leq \text{var}_{Exp}(\tau(t)) \leq \text{var}_{DFR}(\tau(t)).
\]

Numerical illustrations of Theorem 3.9 are given in Tables 3.3 and 3.4.
Table 3.3. Covariance of $D_r$ and $D_s$ in a sample of size 7 from Gamma (1.2) (Prescott (1974))

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>cov($D_r$, $D_s$)</th>
<th>6</th>
<th>7</th>
<th>cov($D_r$, $D_s$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>-0.0240</td>
<td>6</td>
<td>7</td>
<td>-0.0234</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-0.0100</td>
<td>1</td>
<td>4</td>
<td>-0.0078</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-0.0080</td>
<td>2</td>
<td>6</td>
<td>-0.0043</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>-0.0083</td>
<td>3</td>
<td>7</td>
<td>-0.0053</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.4. Covariance of $D_r$ and $D_s$ in a sample of size 7 from $f(x) = 3a^3x^{-4}$, $a > 0$, $x > a$ and $f(x) = 0$ otherwise (Malik (1966))

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>cov($D_r$, $D_s$)</th>
<th>6</th>
<th>7</th>
<th>cov($D_r$, $D_s$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>0.0001</td>
<td>6</td>
<td>7</td>
<td>0.0756</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0.0003</td>
<td>1</td>
<td>4</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>0.0008</td>
<td>2</td>
<td>6</td>
<td>0.0012</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0.0020</td>
<td>3</td>
<td>7</td>
<td>0.0058</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>0.0079</td>
<td>4</td>
<td>6</td>
<td>0.0035</td>
<td></td>
</tr>
</tbody>
</table>
4. NONSTANDARD SITUATIONS

4.1 Introduction

In Chapters 2 and 3, we have focused on stochastic comparisons in random samples from univariate populations. Recently, David (1985) reviewed and systematized some aspects of the treatment of order statistics when these arise from non-i.i.d. variates \(X_1, X_2, \ldots, X_n\), where \(n\) is fixed.

Section 4.2 deals with stochastic comparisons of order statistics from independent but nonidentically distributed (i.n.i.d.) variates. Section 4.3 deals with positively dependent variates, negatively dependent variates, and exchangeable variates.

4.2 i.n.i.d. Case

Marshall and Proschan (1970) show that for i.n.i.d. nonnegative variates \(X_1, \ldots, X_n\) and \(Y_1, \ldots, Y_n\) with c.d.f.s \(F_i, G_i\) for which \(F_i \prec G_i\) with pairwise common means \(\mu_i, i = 1, \ldots, n\), one has

\[
E[X_{1:n}] \geq E[Y_{1:n}], \tag{4.2.1}
\]

\[
E[X_{n:n}] \leq E[Y_{n:n}], \tag{4.2.2}
\]

(see Theorem 2.5). Since \(X_{i:n}\) \((X_{n:n})\) represent the lifetimes of series (parallel) systems with \(n\) independent components, the \(i\)th having lifetime \(X_i\) \((i = 1, 2, \ldots, n)\), we see that the \(X\)-series system has larger expected life than the \(Y\)-series system, and that the result is reversed for parallel systems. Furthermore, if \(F_i \prec G_i\) for \(i = 1, 2, \ldots, n\), then
The following result holds for any i.n.i.d. r.v.s $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$.

**Theorem 4.1.** [Marshall and Olkin (1979, p. 351)] Let $F_i(x) = P[X_i > x]$, $G_i(x) = P[Y_i > x]$, $i = 1, 2, \ldots, n$. If $(G_1(x), \ldots, G_n(x)) \preceq (F_1(x), \ldots, F_n(x))$, $-\infty < x < \infty$, then $Y_{1:n} \geq X_{1:n}$ and $Y_{1:n} \leq X_{1:n}$.

(This implies (4.2.1) and (4.2.2) if the expectations exist.)

Sen (1970) obtained the above result in the special case that

$$G_1(x) = G_2(x) = \ldots = G_n(x) = \frac{\sum_{i=1}^{n} F_i(x)}{n} \quad \text{for all } x.$$ 

**Theorem 4.2.** [Marshall and Olkin (1970, p. 325)] Let $X_1, X_2, \ldots, X_n$ ($Y_1, Y_2, \ldots, Y_n$) be i.n.i.d. variates. If $E[\phi(X)] \leq E[\phi(Y)]$ for all increasing (decreasing) Schur-convex functions, then $X_{1:n} \leq Y_{1:n}$.

**Lemma 4.1.** [Ross (1983, p. 356)] If $X_1, X_2, \ldots, X_n$ ($Y_1, Y_2, \ldots, Y_n$) be i.n.i.d. variates with $X_i \leq Y_i$ $\ (i = 1, 2, \ldots, n)$, then one has for any increasing function $\phi$, that

$$\phi(X_1, X_2, \ldots, X_n) \leq \phi(Y_1, Y_2, \ldots, Y_n).$$

It follows that $X_{r:n} \leq Y_{r:n}$, $r = 1, 2, \ldots, n$ (see Section 2.3).
Theorem 4.3. Let \( X_i(Y_i) \), the life length of components \( i \), have absolutely continuous distribution \( F_i(G_i) \) with \( F_i(0) = G_i(0) = 0 \) and \( f_i(0) \geq g_i(0) > 0 \) for each \( i = 1, 2, \ldots, n \). Let \( X_1, X_2, \ldots, X_n \) (\( Y_1, Y_2, \ldots, Y_n \)) be independent. Then, if \( F_i \leq G_i \) for each \( i = 1, 2, \ldots, n \), then \( X_{1:n} \approx Y_{1:n} \), \( r = 1, 2, \ldots, n \) (this implies \( E[X_{1:n}] \leq E[Y_{1:n}] \)).

Proof. By Theorem 2.3, \( F_i < G_i \), \( i = 1, 2, \ldots, n \). Then, since \( F_i \) implies \( F_i < G_i \) when \( F_i \) and \( G_i \) are distributions of non-negative random variables, we can apply Lemma 4.1.

Lemma 4.2. For any \( a_i \) and \( b_i \) such that \( 0 < b_i < a_i < \frac{1}{2} \) for each \( i = 1, 2, \ldots, n \),

\[
\sum_{i=1}^{k} \left( \prod_{i=1}^{k} \frac{1}{2-a_i} \right) + \sum_{i=1}^{k} \left( \prod_{i=1}^{k} \frac{1}{2+a_i} \right) - \sum_{i=1}^{k} \left( \prod_{i=1}^{k} \frac{1}{2-b_i} \right) + \sum_{i=1}^{k} \left( \prod_{i=1}^{k} \frac{1}{2+b_i} \right) \geq 0, \quad k = 1, 2, \ldots.
\]

Proof. The expansion of

\[
\prod_{i=1}^{k} \left( \frac{1}{2-a_i} \right) + \prod_{i=1}^{k} \left( \frac{1}{2+a_i} \right)
\]

is \( 2^{-(k-1)} \) plus a sum of positive terms in \( a_1, \ldots, a_k \). Each of these terms is at least as large as the corresponding term in \( b_1, \ldots, b_k \).

The result follows.

Theorem 4.4. Let \( X_i(Y_i) \) be random variables symmetric about 0, with c.d.f.s \( F_i \) and \( G_i \), for each \( i = 1, 2, \ldots, n \). Let \( X_1, X_2, \ldots, X_n \) (\( Y_1, Y_2, \ldots, Y_n \)) be independent. Then, if \( X_i \) is more peaked than \( Y_i \)
for each \( i = 1, 2, \ldots, n \) (i.e., \( F_i(x) \geq G_i(x) \) for all \( x > 0 \) and 
\( i = 1, 2, \ldots, n \)), then subject to the existence of the expectations

\[
E[Y_{n:n}] \geq E[X_{n:n}],
\]

\[
E[Y_{1:n}] \leq E[X_{1:n}].
\]

**Proof.** Since for any r.v. \( X \) with c.d.f. \( F \) and finite expectation,

\[
EX = \int_{-\infty}^{\infty} \left(1-F(x)\right)dx - \int_{-\infty}^{\infty} F(x)dx,
\]

we have

\[
E[X_{1:n}] = \int_{-\infty}^{\infty} P[e_X > t]dt - \int_{-\infty}^{\infty} P[e_X < t]dt
\]

\[
= \int_{-\infty}^{\infty} \prod F_i(t)dt - \int_{-\infty}^{\infty} \left[1 - \prod F_i(t)\right]dt
\]

\[
= \int_{-\infty}^{\infty} \left[\prod F_i(t) + \prod F_i(t)-1\right]dt
\]

\[
\geq \int_{-\infty}^{\infty} \left[\prod G_i(t) + \prod G_i(t)-1\right]dt = E[Y_{1:n}],
\]

where the inequality follows from Lemma 4.2 with \( a_i = F_i(t) - \frac{1}{2} \) and

\( b_i = G_i(t) - \frac{1}{2} \). □
4.3 Dependent Case

From Theorem 3.5, if \((X_1, X_2, \ldots, X_n)\) is exchangeable and MTP, then 
\[
\text{cov}(X_{i:n}, X_{j:n}) \geq 0 \quad \text{for any } 1 \leq i, j \leq n.
\]
Also if \((X_1, X_2, \ldots, X_n)\) is associated, then 
\[
\text{cov}(X_{i:n}, X_{j:n}) \geq 0,
\]
since the associatedness is preserved under formation of sets of nondecreasing functions (see Remark 1.15). Pitt (1982) shows that if \((X_1, X_2, \ldots, X_n)\) is multinomially distributed with \(\Sigma = \sum_{i,j=1}^n \sigma_{ij} I_{ij}\) such that \(\sigma_{ij} \geq 0 \) for any \(i \neq j\), then \((X_1, X_2, \ldots, X_n)\) is associated. Hence, the covariance of any two order statistics from multinormal populations with nonnegative covariances of any two random variables is nonnegative. However, we will give an example which shows that \(\text{cov}(X_{i:n}, X_{j:n})\) can be negative if \(X_1, X_2, \ldots, X_n\) are sufficiently negatively dependent.

**Example 4.1.** Thigpen (1961) introduced the following transformation to generate equicorrelated standard variates

\[
Y_i = (1 - p)^{1/2}(x_i - \bar{x} - ax_0), \quad i = 1, 2, \ldots, n \tag{4.3.1}
\]

for \(-\frac{1}{n-1} \leq p < 1\), where \(a^2 = \frac{1+2(p-1)}{n(1-p)}\) and \(x_0, x_1, \ldots, x_n\) are independent standardized variates. If the \(X_s\) are normal, the \(Y_i\) are identically distributed equicorrelated multinormal variates. From (4.3.1), we have

\[
Y_{x_{i:n}} = (1 - p)^{1/2}(x_{i:n} - \bar{x} - ax_0), \quad r = 1, 2, \ldots, n, \tag{4.3.2}
\]

Since \((X_{x_{i:n}}, \bar{x})\) and \(\bar{x}\) are independent and 
\[
\text{cov}(X_{x_{i:n}}, \bar{x}) = \frac{1}{n},
\]
from (4.3.2) we have
\[ \text{cov}(Y_{r;n}, Y_{s;n}) = (1-p)\text{cov}(X_{r;n}, X_{s;n}) + \rho. \]  

(4.3.3)

For example, when \( n = 5 \), if \( \rho > -0.1763903 \), then \( \text{cov}(Y_{2;5}, Y_{4;5}) \geq 0 \).

Hence, if \(-0.25 < \rho < -0.1763903\), then \( \text{cov}(Y_{2;5}, Y_{4;5}) < 0 \). \( \blacksquare \)

Remark 4.1.1. From Remark 1.25(e), if \( -\frac{1}{n-1} \leq \rho < 0 \), then \((Y_1, Y_2, \ldots, Y_5)\)' is S-MRR2. But if \(-0.1763903 < \rho < 0\), then \( \text{cov}(Y_{2;5}', Y_{4;5}') > 0 \), which shows that \((Y_{1;5}', Y_{2;5}', \ldots, Y_{5;5}')\) is not S-MRR2.

Remark 4.1.2. From (4.3.3), the value \( \rho_c(r,s;n) \) of \( \rho \) for which \( \text{cov}(Y_{r;n}', Y_{s;n}') = 0 \), is given by

\[ \rho_c = \frac{-\text{cov}(X_{r;n}', X_{s;n}')}{1 - \text{cov}(X_{r;n}', X_{s;n}')} . \]

For \( r \) and \( s \) fixed and sufficiently large \( n \), one has \( \rho_c < -1/(n-1) \), which means that \( \text{cov}(Y_{r;n}', Y_{s;n}') > 0 \) for all permissible \( \rho \). This situation is illustrated in Table 4.1 which shows that negative covariances of order statistics do not exist below the underlined entry. Let \( \text{cov}(X_{r;n}', X_{s;n}') = \sigma_{rs;n} \).

| Table 4.1. Cov\((X_{r;n}', X_{s;n}')\) and corresponding values of \( \rho_c(r,s;n) \) |
|---|---|---|---|---|---|---|
| \( n \) | \( \sigma_{24;n} \) | \( \rho_c \) | \( n \) | \( \sigma_{14;n} \) | \( \rho_c \) | \( n \) | \( \sigma_{15;n} \) | \( \rho_c \) |
| 5 | \(-0.1763913\) | \(-0.1182830\) | 5 | \(0.0742153\) | \(-0.0801648\) |
| 6 | \(-0.1623367\) | \(-0.1141185\) | 6 | \(0.0773638\) | \(-0.0838508\) |
| 7 | \(-0.1503904\) | \(-0.1092462\) | 7 | \(0.0765598\) | \(-0.0829072\) |
| 8 | \(-0.1409533\) | \(-0.1046343\) | 8 | \(0.0747650\) | \(-0.0808065\) |
Table 4.1. (continued)

<table>
<thead>
<tr>
<th></th>
<th>( c_{24; n} )</th>
<th>( \rho_c )</th>
<th></th>
<th>( c_{11; n} )</th>
<th>( \rho_c )</th>
<th></th>
<th>( c_{15; n} )</th>
<th>( \rho_c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0.1170057</td>
<td>-0.1325101</td>
<td>9</td>
<td>0.0913071</td>
<td>-0.1004818</td>
<td>9</td>
<td>0.0727422</td>
<td>-0.0784487</td>
</tr>
<tr>
<td>10</td>
<td>0.117016</td>
<td>-0.1257478</td>
<td>10</td>
<td>0.0882494</td>
<td>-0.0967912</td>
<td>10</td>
<td>0.0707414</td>
<td>-0.0761266</td>
</tr>
<tr>
<td>19</td>
<td>0.0852931</td>
<td>-0.0952464</td>
<td>12</td>
<td>0.0830687</td>
<td>-0.0905942</td>
<td>16</td>
<td>0.0613087</td>
<td>-0.0653130</td>
</tr>
<tr>
<td>20</td>
<td>0.0835758</td>
<td>-0.0911977</td>
<td>13</td>
<td>0.0808650</td>
<td>-0.0879795</td>
<td>17</td>
<td>0.0601272</td>
<td>-0.0639737</td>
</tr>
</tbody>
</table>

From Remark 1.16, if \( (X_1, X_2, \ldots, X_n) \) is associated, then we have

\[
P(X_1 > x_1, \ldots, X_n > x_n) \geq \prod_{i=1}^{n} P(X_i > x_i) \tag{4.3.4}
\]

and

\[
P(X_1 \leq x_1, \ldots, X_n \leq x_n) \geq \prod_{i=1}^{n} P(X_i \leq x_i) \tag{4.3.5}
\]

Let \( (Y_1, Y_2, \ldots, Y_n) \) be independent random variables with the same univariate marginal distribution functions as \( (X_1, X_2, \ldots, X_n) \). Then (4.3.4) and (4.3.5) imply that

\[
X_{1:n} \geq_{st} Y_{1:n} \tag{4.3.6}
\]

and

\[
X_{n:n} \leq_{st} Y_{n:n} \tag{4.3.7}
\]

Also, Block et al. (1985) show that if \( (X_1, X_2, \ldots, X_n) \) is NDS (see Definition 1.23), then the inequalities (4.3.4) and (4.3.5) are reversed and consequently also (4.3.6) and (4.3.7).
From (4.3.6) and (4.3.7), distribution functions of \(X_{1:n}\) and \(X_{n:n}\) could be approximated by distribution functions of \(Y_{1:n}\) and \(Y_{n:n}\).

Glaz and Johnson (1984) suggested sharper bounds than \(\prod_{i=1}^{n} P[X_i \leq x_i]\), namely, for the latter, \(r_2 = P[X_1 \leq x_1] \prod_{i=2}^{n} P[X_i \leq x_i | X_{i-1} \leq x_{i-1}]\) and \(r_3 = P[X_1 < x_1, X_2 < x_2] \prod_{i=3}^{n} P[X_i < x_i | X_j < x_j; j = i-2, i-1]\).

Bhattacharyya (1970) shows that if \(X_1, X_2, \ldots\) form an exchangeable sequence, then (4.3.6) and (4.3.7) hold for any subset of \(n\) \(X\)'s. Also if \((X_1, \ldots, X_n)\) is PDM (see Definition 1.24), then \((X_1, X_2, \ldots, X_n)\), being exchangeable as well, (4.3.6) and (4.3.7) are satisfied. Furthermore, Shaked (1977) shows that if \((X_1, X_2, \ldots, X_n)\) and \((Y_1, Y_2, \ldots, Y_n)\) are, respectively, PDM and i.i.d. r.v.'s having the same univariate marginal, then for any \(x \in R\),

\[
(F_{X_{1:n}}(x), \ldots, F_{X_{n:n}}(x)) \preceq (F_{Y_{1:n}}(x), \ldots, F_{Y_{n:n}}(x)). \quad (4.3.8)
\]

Note that (4.3.8) implies (4.3.6) and (4.3.7). Marshall and Olkin (1979, p. 350) show that (4.3.8) is equivalent to

\[
(Eg(X_{1:n}), \ldots, Eg(X_{n:n})) \preceq (Eg(Y_{1:n}), \ldots, Eg(Y_{n:n})). \quad (4.3.9)
\]

for all monotone functions \(g\) such that the expectations exist.

Comparisons so far have been made between dependent random vectors and vectors having independent components. Realizing that vectors having independent components represent a special case among the class of all dependent random vectors, Shaked and Tong (1985) make a study of comparisons of dependent vectors according to the strength of
dependence. Specifically, they discuss the partial orderings of exchangeable random variables by positive dependence (see Definition 1.27). Note that if \( \chi \) is more positively dependent than \( \chi \) (or \( \chi \) is more dispersed than \( \chi \)), then we can imagine the \( Y_i \)'s to "hang together" more than the \( X_i \)'s. This observation motivates the ordering \( \chi > \chi \) (see Definition 1.27(a)). It also suggests \( \chi > \chi \) and \( \chi > \chi \) in the sense that stronger dependence of exchangeable random vectors, \( \chi \), is associated with tighter hanging together of the distributions and the expectations of the order statistics formed from \( \chi \). Such considerations show that these order statistics play an important role in assessing the strength of dependence between two random vectors. Shaked and Tong (1985) prove also that (1.5.2) \( \Rightarrow \) (1.5.4) and (1.5.3) \( \Rightarrow \) (1.5.4).

When \( \chi \) is distributed by \( N(0, \Sigma) \), where \( \Sigma \) is a correlation matrix, Slepian (1962) obtains the following result: If \( \rho_{ij} \geq \tau_{ij} \) for all \( i,j \) with two positive semidefinite correlation matrices \( \Sigma = \Sigma = (\rho_{ij}) \) and \( \Sigma = \Sigma = (\tau_{ij}) \) respectively, then

\[
P_{\Sigma = \Sigma} [X_1 > x_1, \ldots, X_n > x_n] \geq P_{\Sigma = \Sigma} [X_1 > x_1, \ldots, X_n > x_n] \quad (4.3.10)
\]

and

\[
P_{\Sigma = \Sigma} [X_1 \leq x_1, \ldots, X_n \leq x_n] \geq P_{\Sigma = \Sigma} [X_1 \leq x_1, \ldots, X_n \leq x_n] \quad (4.3.11)
\]

hold for all \( \chi = (x_1, \ldots, x_n)' \). Furthermore, the inequalities are strict if \( \Sigma, \Sigma \) are positive definite and if the strict inequality \( \rho_{ij} > \tau_{ij} \) holds for some \( i,j \).


I wish to express my sincere gratitude and appreciation to Professor H. A. David for suggesting this area of research and for giving constant guidance during the course of this study. Without his attention, I would have been, certainly, unable to do this work.

I wish to acknowledge the financial support given by the U. S. Army Research Office during my graduate study at Iowa State University.

I wish to thank Sharon Shepard for her patience and skill in the typing of this manuscript.

Also I wish to express my deepest appreciation to my wife, Gyung-Hi, for her understanding, encouragement, and moral support through the years of my graduate study.