Stochastic Orders and Decision under Risk
IMS Lecture Notes – Monograph Series (1991)

STOCHASTIC ORDERS AND THEIR APPLICATION TO A UNIFIED APPROACH TO VARIOUS CONCEPTS OF DEPENDENCE AND ASSOCIATION

BY REINHARD BERGMANN

Rheinfelden

Multivariate stochastic partial orderings are studied, especially in the context of probability inequalities. Extensions of stochastic orderings to the multi-component case for general product spaces are developed. They provide a sound basis for a unified representation of dependence and association notions.

1. Introduction. Stochastic orderings have found a wide field of application in probability, statistics, and statistical decision theory, see Stoyan (1983), Mosler and Scarsini (1991), as comprehensive references. In probability theory, they are useful in deducing probability inequalities, comparing stochastic models, establishing bounds and inequalities in reliability and queueing theory, in statistics for example in hypothesis testing, simultaneous comparisons, multiple decision problems, and in economics in decisions under risk, particularly in multi-attribute utility theory.

The approach in this paper is mainly to define various stochastic orderings starting from interesting multivariate probability inequalities. To characterize the stochastic orderings, several quite different equivalent conditions are given. The stochastic orderings are associated with inequalities between expectations of functions with respect to the corresponding distributions or random variables. A very interesting and important problem is to find the class of functions which implies the inequality between the expectations. For this issue solutions are given.

We prefer the presentation in terms of random variables rather than of distribution functions or probability measures in order to facilitate intuitive handling of the inequalities. Consider a partially ordered measurable space $(E, \leq)$ and random variables X, Y with values in E.

AMS 1980 Subject Classification: Primary 60E15; Secondary 62H20

Key words and phrases: Multivariate stochastic orders, probability inequalities, dependence, association.

A common description of multivariate probability inequalities is

$$P(X \in A) \leq P(Y \in A), \qquad A \in \mathcal{A}, \tag{1.1}$$

where $\mathcal{A}$ is some interesting class of sets and P a probability measure on measurable space $(E, \mathcal{M})$, $\mathcal{M}$ being a σ-algebra of subsets of E.

Now, let $\mathcal{F}$ denote a class of functions for which the expectations in the following inequality exist,

$$Ef(X) \leq Ef(Y), \qquad f \in \mathcal{F}. \tag{1.2}$$

Then, the central problem addressed in this paper can be described as "For a given class of sets $\mathcal{A}$ which by probability inequality (1.1) defines a stochastic order find the class of functions $\mathcal{F}$ which implies the expectation inequality (1.2)" and vice versa: "For given $\mathcal{F}$ of (1.2) find the class $\mathcal{A}$ so that (1.1) holds."

DEFINITION 1.1. A real valued function f defined on E is increasing (decreasing) if for $x, y \in E$, $x \leq y$ implies $f(x) \leq (\geq) f(y)$. A set $A \subset E$ is increasing (decreasing) if the indicator function of A, denoted by I_A, is increasing (decreasing). Or, equivalently, if for $x \in A$ $x \leq (\geq) y$ implies $y \in A$.

Denote by $\mathcal{F}_i$ the class of all increasing functions and by $\mathcal{A}_i$ the class of all increasing sets. Then a familiar notion is to say that X is stochastically smaller than Y, in symbols $X \leq_d Y$, if (1.2) holds for all increasing functions $f \in \mathcal{F}_i$. The notation $\leq_d$ comes from the univariate case of real valued random variables where $\leq_d$ is defined by $P(X \geq t) \leq P(Y \geq t)$ for all real t. It is well known that $X \leq_d Y$ for a general space E is equivalent to probability inequality (1.1) for the class of increasing sets $\mathcal{A}_i$.

In Section 2 we give definitions and characterizations of multivariate stochastic orderings for further interesting classes of functions $\mathcal{F}$ and sets $\mathcal{A}$. In most cases, this is done by weakening the condition of the stochastic distribution ordering in requiring that (1.1) holds for all sets A which belong to a subcollection of the collection of all increasing sets. We shall consider the class of all convex sets which are additionally increasing or symmetric or G-invariant with respect to a subgroup G of all orthogonal groups, respectively. In this context the Schur-convex ordering of Nevius, Proschan, and Sethuraman (1977) may be included in the concept of G-increasing orderings. In terms of all these stochastic partial orderings, a formulation of monotonicity properties of probability inequalities can be given.

In Section 3 we extend the above definitions to the multi-component case by defining product partial orderings on general n-dimensional product spaces.

These orders allow comparing the strength of dependence structures and are particularly useful in describing qualitative dependence properties.

Section 4 gives an application of the product stochastic orderings to a unified concept of dependence and association in a generalized sense.

The final Section 5 concerns remarks about the comparison of the dependence structure of sets of random variables and stochastic processes.

2. Multivariate Stochastic Orderings. The approach behind the idea to construct multivariate stochastic orderings for random variables X, Y defined on a partially ordered measurable space $(E, \leq)$ runs as follows. First, the random variables are mapped on the real axis by a function $f : E \to \mathbb{R}^1$. Then, the real valued random variables $f(X)$ and $f(Y)$ are compared by univariate stochastic orderings, for example, by the most familiar distribution ordering $\leq_d$, the convex or concave orderings $\leq_c$, $\leq_{cv}$ (see Stoyan 1983).

The crucial step is to choose appropriate mapping functions and a suitable compatible univariate ordering. Compatibility means that the properties of f must be compatible with those of the functions for which the one-dimensional ordering implies the expectation inequality. It is not interesting to choose a univariate ordering as weak as possible because the class of functions for which the expectation inequality holds then becomes too small. An important purpose of a useful multivariate stochastic ordering is to establish its equivalence to the expectation inequality for a class of functions which join certain properties with the original functions f. The compatibility is guaranteed if the functions composed from f and the class of functions for which the univariate ordering implies the expectation inequality preserve these original properties. In a different approach, using cones of functions, Marshall (1991) studies questions of generating stochastic orderings which preserve certain properties.

2.1. Unimodal Stochastic Orderings. Motivated by Anderson's (1955) paper, various so-called unimodal orderings have been used, see Ahmed, Leon, Proschan (1981), Eaton (1987). In this subsection we define two stochastic versions of unimodal partial orderings.

DEFINITION 2.1. A real valued function f defined on a linear space E is unimodal (reverse unimodal) if the set $\{x : f(x) \geq (\leq) u\}$ is convex for all real u.

DEFINITION 2.2. A real valued function f defined on a linear space E is symmetric if $f(x) = f(-x)$ for all x. A set $A \subset E$ is symmetric if $A = -A$.

The following simple properties and relationships are true for unimodal and convex functions.

PROPOSITION 2.1. (i) *f is reverse unimodal if and only if $-f$ is unimodal.*

(ii) *Each concave function f is unimodal.*

(iii) *The indicator function I_A of a convex set A is unimodal.*

The converse statement of (ii) is not true; (iii) gives a simple but important counterexample since the indicator function I_A of convex set A is unimodal but not concave. Thus the class of unimodal functions is really larger than those of the concave functions.

The following characterization of a concave function by a convex set is well known.

PROPOSITION 2.2. *Let $E = \mathbb{R}^k$ and f be a real valued function on $\mathbb{R}^k$. Then f is concave if and only if the set $H = \{(x,u) \in \mathbb{R}^k \times \mathbb{R}\colon f(x) \geq u\}$ is convex (as a subset of $\mathbb{R}^{k+1}$).*

REMARK. H is not necessarily an increasing set.

The following first definition of a stochastic unimodal partial ordering with respect to the class of symmetric and convex sets is known in the literature as "peakedness" ordering; see Birnbaum (1948), Sherman (1955), Olkin and Tong (1988), and Dharmadhikari and Joag-dev (1988).

DEFINITION 2.3. Two random variables X and Y with values in E are said to be ordered with respect to convex and symmetric sets, $X \leq_{cs} Y$, if

(a) $P(X \in A) \leq P(Y \in A)$ for all sets $A \in \mathcal{A}_{cs}$, the class of convex, symmetric sets.

THEOREM 1. *$X \leq_{cs} Y$ is equivalent to each of the following conditions.*

(b) *$P(f(X) \geq t) \leq P(f(Y) \geq t)$ for all t and functions $f \in \mathcal{F}_{us}$, the class of all unimodal, symmetric functions; i.e. $f(X) \leq_d f(Y)$ for all unimodal symmetric functions f.*

(c) *$Ef(X) \leq Ef(Y)$ for all unimodal, symmetric functions $f \in \mathcal{F}_{us}$ provided the expectations exist.*

(b′) *$P(f(X) < t) \leq P(f(Y) < t)$, for all t; i.e. $f(Y) \leq_d f(X)$ for all reverse unimodal, symmetric functions f.*

PROOF. If $X \leq_{cs} Y$ for f unimodal and symmetric, then the set $\{x : f(x) \geq u\}$ is convex and symmetric so that (b) holds.

Since the indicator function I_A of a convex and symmetric set is unimodal and symmetric (see Proposition 2.1.(iii)), (b) implies (a), i.e. $X \leq_{cs} Y$.

Remember that (b) $f(X) \leq_d f(Y)$ is equivalent to $Eg(f(X)) \leq Eg(f(Y))$, $f \in \mathcal{F}_{us}$ and g increasing. Choosing for g the identical function, (b) implies (c).

Now, it is crucial that the composed function $g(f(x))$ is unimodal and symmetric for $f \in \mathcal{F}_{us}$ and g increasing. This is true since $\{x : f(x) \geq u\} = \{x : g(f(x)) \geq g(u)\}$. Therefore, (c) implies (b).

(b′) is a simple consequence of Proposition 2.1.(i). This completes the proof. ∎

REMARK. Note that if one condition of Theorem 1 holds, the inequalities in (b), (c) hold, especially for all concave and symmetric functions, and the inequalities in (b′) whenever f is convex and symmetric.

The next theorem shows that $\mathcal{F}_{us}$ is the largest class of functions which is monotone for the relation $\leq_{cs}$ with respect to the probability operator defined in (b).

THEOREM 2. *Let $X \leq_{cs} Y$ imply $P(f(X) \geq t) \geq P(f(Y) \geq t)$ for a function f and all t. Then f is unimodal and symmetric.*

PROOF. The assumption of the theorem can be rewritten as: $P(X \in A) \leq P(Y \in A)$ for all convex, symmetric sets A implies $P(X \in D_t) \geq P(Y \in D_t)$ for all t where $D_u = \{x : f(x) \geq u\}$. Therefore, the set D_u must be convex, i.e. f is unimodal. Furthermore, D_u is also symmetric, especially $-y \in D_{f(Y)}$, i.e. $f(-y) \geq f(y)$ and $y \in D_{f(-y)}$, that is $f(y) \geq f(-y)$ and therefore $f(-y) = f(y)$. ∎

Now, we state Anderson's Theorem as a monotonicity property in terms of the stochastic partial ordering $\leq_{cs}$ in $E = \mathbb{R}^k$. Some applications follow.

PROPOSITION 2.3. (Anderson 1955). *If the random variable X with values in $\mathbb{R}^k$ has a symmetric and unimodal density, then*

$$X - y \leq_{cs} X - \lambda y \text{ for all } y \in \mathbb{R}^k \text{ and any real number } \lambda,\ |\lambda| \leq 1,$$

or equivalently,

$$\begin{aligned} &X + \lambda_2 y \leq_{cs} X + \lambda_1 y, \text{for all } y \in \mathbb{R}^k \\ &\quad \text{and for any real numbers } \lambda_1, \lambda_2,\ |\lambda_1| < |\lambda_2| \ . \end{aligned}$$

That means, the nearer the values of X are concentrated about the origin the larger becomes the probability.

PROPOSITION 2.4. *When the random variable Y is independent of X and X has a unimodal symmetric density, then for all real λ_1, λ_2*

$$X + \lambda_2 Y \leq_{cs} X + \lambda_1 Y \text{ if } |\lambda_1| < |\lambda_2|.$$

The next proposition applies to the Gaussian case with zero mean.

PROPOSITION 2.5. *Let X_{Σ_i} have a normal distribution $N(0, \Sigma_i)$ on $\mathbb{R}^k$ with mean 0 and covariance matrix Σ_i, $i = 1, 2$. If $\Sigma_2 - \Sigma_1$ is positive semi-definite, then $X_{\Sigma_2} \leq_{cs} X_{\Sigma_1}$.*

This proposition has been generalized to elliptically contoured distributions by Fefferman, Jodeit, and Perlman (1972).

Now, we give a further definition of a stochastic partial ordering concerning unimodality. This is done by replacing the word "symmetric" in the definition of relation $\leq_{cs}$ by "increasing".

DEFINITION 2.4. Two random variables X and Y with values in E are said to be ordered with respect to convex and increasing sets, $X \leq_{ci} Y$, if

(a) $P(X \in A) \leq P(Y \in A)$ for all sets $A \in \mathcal{A}_{ci}$, the class of all convex, increasing sets.

THEOREM 3. *$X \leq_{ci} Y$ is equivalent to each of the following conditions.*

(b) *$P(f(X) \geq t) \leq P(f(Y) \geq t)$ for all t and functions $f \in \mathcal{F}_{ui}$, the class of all unimodal, increasing functions; i.e. $f(X) \leq_d (Y)$, $f \in \mathcal{F}_{ui}$.*

(c) *$Ef(X) \leq Ef(Y)$ for all unimodal, increasing functions $f \in \mathcal{F}_{ui}$ provided the expectations exist.*

(b′) *$P(f(X) < t) \leq P(f(Y) < t)$ for all t; i.e. $f(Y) \leq_d f(X)$ for all reverse unimodal, decreasing functions f.*

The proof is similar to that of Theorem 1 by replacing the term "symmetric" by "increasing". In particular, (b′) is equivalent to (b) since $-f$ is unimodal and increasing if and only if f is reverse unimodal and decreasing.

Note that in Theorem 3 above the inequality in (b) holds for all concave increasing functions, and the inequality in (b′) whenever f is convex and decreasing.

Furthermore, we say that the random variables X, Y are ordered with respect to convex and decreasing sets, $X \leq_{cd} Y$, if

(a) $P(X \in A) \leq P(Y \in A)$ for all sets $A \in \mathcal{A}_{cd}$, the class of convex, decreasing sets.

Then we have, parallel to Theorem 3,

THEOREM 4. *$X \leq_{cd} Y$ is equivalent to each of the following conditions.*

(b) *$P(f(X) \geq t) \leq P(f(Y) \geq t)$ for all t and functions $f \in \mathcal{F}_{ud}$, the class of all unimodal, decreasing functions; i.e. $f(X) \leq_d f(Y)$, $f \in \mathcal{F}_{ud}$.*

(c) *$Ef(X) \leq Ef(Y)$ for all unimodal, decreasing functions $f \in \mathcal{F}_{ud}$ provided the expectations exist.*

(b′) *$P(f(X) < t) \le P(f(Y) < t)$ for all t; i.e. $f(Y) \le_d f(X)$ whenever f is reverse unimodal and increasing.*

We note again that (b′) is true for convex increasing functions.

Finally, we remark that the partial ordering $\le_{ci}$ for increasing and unimodal functions was already used by Levhari et al. (1975) and Mosler (1982) in an economic context. There, a unimodal function is called quasiconcave.

2.2. Partial Orderings Invariant With Respect to an Orthogonal Group. The symmetry property above can be replaced by a more general invariance property.

We shall adopt the following notations. Let g be a transformation $E \to E$. Then for $A \subset E$

$$g(A) = \{z : z = g(x), \quad x \in A\}.$$

Let E be endowed with an inner product such that $(E, \langle\cdot,\cdot\rangle)$ is a finite dimensional inner product space, and let G be any subgroup of $\mathcal{O}(E)$, the orthogonal group of the inner product space $(E, \langle\cdot,\cdot\rangle)$. The topology on $\mathcal{O}(E)$ is the usual topology of the orthogonal group.

DEFINITION 2.5. We say that f is G-invariant if $f(x) = f(gx)$ for all $x \in E$ and all $g \in G$, and that the set A is G-invariant if $x \in A$ implies $gx \in A$ for all $g \in G$.

DEFINITION 2.6. For a partially ordered finite dimensional inner product space $(E, \le, \langle\cdot,\cdot\rangle)$ we say that the random variables X, Y are ordered with respect to convex and invariant sets, $X \le_{civ} Y$, if

(a) $P(X \in A) \le P(Y \in A)$ for all sets $A \in \mathcal{A}_{civ}$, the class of convex, G-invariant sets.

Then we can generalize Theorem 1 to

THEOREM 5. *$X \le_{civ} Y$ is equivalent to each of the following conditions.*

(b) *$P(f(X) \ge t) \le P(f(Y) \ge t)$ for all t and functions $f \in \mathcal{F}_{uiv}$, the class of all unimodal, G-invariant functions; i.e. $f(X) \le_d f(Y)$, $f \in \mathcal{F}_{uiv}$.*

(c) *$Ef(X) \le Ef(Y)$ for all unimodal, G-invariant functions $f \in \mathcal{F}_{uiv}$ if the expectations exist.*

(b′) *$P(f(X) < t) \le P(f(Y) < t)$ for all t; i.e. $f(Y) \le_d f(X)$ whenever f is reverse unimodal and G-invariant.*

The proof is omitted since it parallels that of Theorem 1.

Now in terms of $\le_{civ}$ for $E = \mathbb{R}^k$ we can state a generalization of Anderson's Theorem which is due to Mudholkar (1966).

PROPOSITION 2.6. (Mudholkar 1966). *Let X have a unimodal G-invariant density on $\mathbb{R}^k$ and $\alpha(y) = \sum_{i=1}^N \alpha_i g_i y$, $g_i \in G$, $\alpha_i \geq 0$, $\sum_{i=1}^N \alpha_i = 1$. Then*

$$X - y \leq_{civ} X - \alpha(y), \quad y \in \mathbb{R}^k;$$

and if Y is independent of X

$$X + Y \leq_{civ} X + \alpha(Y).$$

In the Gaussian case, Eaton and Perlman (1991) give the following statement.

PROPOSITION 2.7. *Suppose that X_{Σ_i} have a k-variate normal distribution $N(0, \Sigma_i)$ on $\mathbb{R}^k$ with mean vector 0 and covariance matrix Σ_i for $i = 1, 2$. If $\Sigma_2 \geq \Sigma_1$, i.e. $\Sigma_2 - \Sigma_1$ is positive semi-definite, then*

$$X_{\Sigma_2} \leq_{civ} X_{\Sigma_1}$$

provided that Σ_1 is G-invariant, i.e. $g\Sigma_1 g' = \Sigma_1$ for all $g \in G$,and G acts effectively on $\mathbb{R}^k$.

By definition, G acts effectivelly on $\mathbb{R}^k$ if $0 = \sum_{i=1}^N \alpha_i g_i x$ for all $x \in \mathbb{R}^k$ where $g_i \in G$, $\alpha_i \geq 0$, $\sum_1^N \alpha_i = 1$, that means 0 is in the convex hull of $\{gx \mid g \in G\}$ for all $x \in \mathbb{R}^k$, or, to put the last condition in another way, 0 is the minimal element with respect to the ordering $\leq_G$ induced by G (see Section 2.4): $0 \leq_G x$ for all $x \in \mathbb{R}^k$.

2.3. The Stochastic Majorization Ordering. Also the notion of stochastic majorization introduced by Nevius, Proschan, and Sethuraman fits very well with the concept presented here and outlined so far. Their results and many related ones are collected in the book of Marshall and Olkin (1979).

In this context, it is worth mentioning that the stochastic majorization order can be obtained from the distribution order $\leq_d$ which is characterized by increasing sets and functions (see Section 1) by choosing the majorization order as the partial ordering of the partially ordered space $(E, \leq)$. The reason for that is that Schur-convex functions are defined as increasing functions with respect to the majorization order.

Moreover, the stochastic majorization ordering is embedded as a special case in a general stochastic ordering of functions and sets which are increasing with respect to the order induced by a subgroup G of all orthogonal matrices; see Section 2.4.

Throughout this subsection we assume $E = \mathbb{R}^k$.

Consider two real vectors $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$ where $x_{(1)} \geq \ldots \geq x_{(k)}$ and $y_{(1)} \geq \ldots \geq y_{(k)}$ are the components of x, y rearranged in decreasing order.

DEFINITION 2.7. We say that y majorizes x and write $x \prec y$ if

$$\sum_{i=1}^{m} x_{(i)} \leq \sum_{i=1}^{m} y_{(i)}, \quad \text{for } m = 1, \cdots, k-1 \text{ and } \sum_{i=1}^{k} x_i = \sum_{i=1}^{k} y_i.$$

Sometimes the majorization order is also called "Schur-ordering" because of the next definition.

DEFINITION 2.8. A function on $\mathbb{R}^k$ which is increasing (decreasing) with respect to the majorization order $\prec$ is called Schur-convex (concave). A set $A \subset \mathbb{R}^k$ is said to be Schur-convex (concave) if its indicator function is Schur-convex (concave), i.e. if $x \in A$ and $x \prec y$ ($y \prec x$) imply $y \in A$.

DEFINITION 2.9. For two random variables X, $Y \in \mathbb{R}^k$, X is said to stochastically majorize Y, $X \leq_{sc} Y$, if

(a) $P(X \in A) \leq P(Y \in A)$ for every set $A \in \mathcal{A}_{sc}$, the class of Schur-convex sets.

The notation $\leq_{sc}$ symbolizes that alternatively we could say X and Y are ordered with respect to Schur-convex sets.

THEOREM 6. (Nevius, Proschan, Sethuraman 1977). *$X \leq_{sc} Y$ is equivalent to each of the following statements.*

(b) *$P(f(X) \geq t) \leq P(f(Y) \geq t)$ for all t and functions $f \in \mathcal{F}_{sc}$, the class of all Schur-convex functions; i.e. $f(X) \leq_d f(Y)$, $f \in \mathcal{F}_{sc}$.*

(c) *$Ef(X) \leq Ef(Y)$ for every Schur-convex $f \in \mathcal{F}_{sc}$ for which both expectations exist.*

(b′) *$P(f(X) < t) \leq P(f(Y) < t)$ for all t; i.e. $f(Y) \leq_d f(X)$ for all Schur-concave f.*

Note that an equivalent definition is obtained when the term "Schur-convex" is replaced by "Schur-concave" and the roles of X and Y are exchanged.

The following statements in terms of $\leq_{sc}$ can be given.

PROPOSITION 2.8. (Marshall and Olkin 1974). *Let X have a Schur-concave density on $\mathbb{R}^k$. Then*

$$\theta \prec \xi \textit{ implies } X - \theta \leq_{sc} X - \xi.$$

A special consequence is the following important result for location parameter families.

PROPOSITION 2.9. *Let X_θ have a density $f_\theta(x) = f(x-\theta)$ which is Schur-concave on $\mathbb{R}^k$. Then*

$$\theta \prec \xi \textit{ implies } X_\theta \leq_{sc} X_\xi.$$

Finally, we remark that a weak stochastic majorization order was defined by Nevius, Proschan, Sethuraman (1977a). It corresponds to the weak majorization ordering. This implies that in addition to Schur-concavity the functions and sets have to be increasing with respect to the componentwise order in $\mathbb{R}^k$.

2.4. Stochastic Orderings Increasing with Respect to an Orthogonal Group. Motivated by the concept of majorization and Schur functions, preorders can be defined by special groups of orthogonal transformations, see Eaton (1982, 1987).

We note that the stochastic ordering with respect to a given group (which will be defined in this section) is the ordering $\leq_d$ of Section 1 when the preorder with respect to that group is chosen as the order $\leq$ of the underlying space.

Remember Section 2.2 where we assumed that $(E, \langle\cdot,\cdot\rangle)$ was a finite dimensional inner product space and G any closed subgroup of $\mathcal{O}(E)$, the orthogonal group of the inner product space.

On the inner product space $(E, \langle\cdot,\cdot\rangle)$, G is a closed subgroup of $\mathcal{O}(E)$ and induces an ordering as follows. Given $x \in E$, $C(x)$ denotes the convex hull of the G-orbit of x, that is, $C(x)$ is the convex hull of $\{gx : g \in G\}$. Since G is closed, $C(x)$ is compact.

DEFINITION 2.10. (Eaton 1987). For x and y in E, write $x \leq_G y$ to mean that $x \in C(y)$. The relation $\leq_G$ is called the G-induced ordering on $(E, \langle\cdot,\cdot\rangle)$.

A real valued function f defined on E is decreasing (increasing) if $x \leq_G y$ implies $f(x) \geq (\leq) f(y)$. A set $A \subset E$ is increasing (decreasing) if the indicator function of A is increasing (decreasing), i.e. for $x \in A$ $x \leq_G y$ ($y \leq_G x$) implies $y \in A$.

Note that any f which is increasing or decreasing is also G-invariant, because it necessarily satisfies $f(x) = f(gx)$ for all $x \in E, g \in G$ since $x \leq_G gx$ and $gx \leq_G x$.

DEFINITION 2.11. Two random variables X and Y with values in E are said to be ordered with respect to G, $X \leq_G Y$, if

(a) $P(X \in A) \leq P(Y \in A)$ for all sets $A \in \mathcal{A}_G$, the class of all G-increasing sets.

THEOREM 7. *$X \leq_G Y$ is equivalent to each of the following conditions.*

(b) *$P(f(X) \geq t) \leq P(f(Y) \geq t)$ for all t and functions $f \in \mathcal{F}_G$, the class of G-increasing functions; i.e. $f(X) \leq_d f(Y)$, $f \in \mathcal{F}_G$.*

(c) *$Ef(X) \leq Ef(Y)$ for all G-increasing functions $f \in \mathcal{F}_G$ provided the expectations exist.*

(b′) *$P(f(X) < t) \leq P(f(Y) < t)$ for all t; i.e. $f(Y) \leq_d f(X)$ for all G-decreasing functions f.*

The proof is similar to that of Theorem 1 and therefore omitted.

For statistical applications, especially with $E = \mathbb{R}^k$, important examples of orthogonal subgroups are

(1) $\mathcal{P}_k$ – the group of $k \times k$ permutation matrices,

(2) $\mathcal{D}_k$ – the group of coordinate sign changes,

(3) $\mathcal{P}_k \cup \mathcal{D}_k$ – the group generated by $\mathcal{P}_k$ and $\mathcal{D}_k$.

REMARKS. (i) A G-invariant and unimodal function f is G-increasing. This means that the stochastic ordering $\leq_G$ is weaker than $\leq_{civ}$.

(ii) If $G = \mathcal{P}_k$ for $E = \mathbb{R}^k$ then the majorization order $x \prec y$ is equivalent to $x \leq_{\mathcal{P}_k} y$. That means that the stochastic ordering $X \leq_{\mathcal{P}_k} Y$ is equivalent to the relation $X \leq_{sc} Y$ for Schur-convex functions and sets defined in the preceding sections. Combining this result with Remark (i) above, we see in particular that the majorization relation $\leq_{sc}$ is weaker than the convex invariant ordering $\leq_{civ}$ which is invariant with respect to $\mathcal{P}_k$, the group of permutation matrices.

For the following statistical applications of the stochastic ordering $\leq_G$ we assume $E = \mathbb{R}^k$.

PROPOSITION 2.10. (Eaton 1982). *Suppose X_θ is $N_p(0, \Sigma_0 + \theta\theta')$ distributed for $\theta \in \mathbb{R}^p$ where Σ_0 is a fixed $p \times p$ positive definite matrix, and G is a reflection group such that $g\Sigma_0 g' = \Sigma_0$ for all $g \in G$, i.e. Σ_0 is G-invariant. Then*

$$X_\theta \leq_G X_\xi \ \textit{if}\ \theta \leq_G \xi.$$

EXAMPLE 2.1. Suppose $\Sigma_0 = D$ is diagonal and take $G = \mathcal{D}_k$, the group of coordinate sign changes. Then for X_θ distributed as $N(0, D + \theta\theta')$

$$X_\theta \leq_{\mathcal{D}_k} X_\xi \ \text{if}\ \theta \leq_{\mathcal{D}_k} \xi.$$

That means, X_θ increases stochastically with respect to $\leq_{\mathcal{D}_k}$ if θ $\mathcal{D}_k$-increases, i.e. for $\mathcal{D}_k$-increasing sets A, $P(X_\theta \in A)$ is a function of $|\theta_1|, \dots, |\theta_p|$ and is increasing in $|\theta_i|$ for $i = 1, \dots, p$. Especially, $P(|X_i| > a_i, i = 1, \dots, p)$ is increasing in $|\theta_i|$, $i = 1, \dots, p$.

EXAMPLE 2.2. Take $G = \mathcal{P}_k$, the group of permutation matrices, so that $g\Sigma_0 g' = \Sigma_0$ for all $g \in G$ where Σ_0 has intraclass correlation structure, that is, $\Sigma_0 = \sigma^2 R$ where the p diagonal elements of R are all unity and the off diagonal elements of R are all ρ, $-1/(p-1) < \rho < 1$. When X_θ is $N(0, \Sigma(\theta))$ where $\Sigma(\theta)) = \Sigma_0 + \theta\theta'$, then

$$X_\theta \leq_{\mathcal{P}_k} X_\xi \text{ for } \theta \leq_{\mathcal{P}_k} \xi,$$

that means, for any $\mathcal{P}_k$-increasing function (i.e. Schur-convex function) f, $Ef(X_\theta)$ is $\mathcal{P}_k$-increasing as a function of θ.

EXAMPLE 2.3. Suppose $\Sigma_0 = \sigma^2 I_p$ and let $G = \mathcal{P}_k \cup \mathcal{D}_k$. When X_θ is normal as in the previous examples, then

$$X_\theta \leq_G X_\xi \text{ for } \theta \leq_G \xi.$$

3. Stochastic Orderings for General Product Spaces. The aim of this section is the investigation of dependence structures which play a crucial role in many problems of probability theory and statistics. Our concern is to develop a tool which allows us to describe and compare the strength of dependence in a qualitative way. For this purpose we use the concept of one-component or marginal stochastic orderings given in the preceding sections and extend it to stochastic product partial orderings defined on product spaces.

Consider a partially ordered linear space $(E, \leq)$ and random vectors $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ each defined on the product space $\times_{i=1}^n E$. In this case we write $(E^n, \leq)$ where $\leq$ means that each component is ordered with respect to $\leq$.

Although we confine ourselves to the case where $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ are random vectors in $(E^n, \leq)$ it is worth mentioning that all forthcoming considerations are true for a general product $\times_{i=1}^n E_i$ of partially ordered linear spaces E_i.

In the case $E = \mathbb{R}^1$, i.e. when the partially ordered space is the reals, Bergmann (1978) (see Stoyan (1983), p. 27) defined the following three relations between random vectors X, Y in $\mathbb{R}^n$.

$$\begin{aligned}
&X \leq_K Y \text{ if } P(X \geq x) \leq P(Y \geq x) \text{ for all } x \in \mathbb{R}^n, \\
&X \leq_D Y \text{ if } P(X < x) \leq P(Y < x) \text{ for all } x \in \mathbb{R}^n, \text{ and} \\
&X \leq_{KC} Y \text{ if} \int_t^\infty P(X \geq x)dx \leq \int_t^\infty P(Y \geq x)dx \\
&\qquad \text{for all } t = (t_1, \ldots, t_n) \in \mathbb{R}^n.
\end{aligned}$$

A characterization of these relations by inequalities between expectations is given in the following statement.

PROPOSITION 3.1. *For random vectors* $X = (X_1, \ldots, X_n)$ *and* $Y = (Y_1, \ldots, Y_n)$ *on* $\mathbb{R}^n$,

$$\begin{aligned} &X \le_K Y (X \le_D Y) \text{ if and only if} \\ &\qquad E(f_1(X_1) \ldots f_n(X_n)) \le E(f_1(Y_1) \ldots f_n(Y_n)) \qquad (3.1) \\ &\qquad \text{for all increasing (decreasing), positive functions } f_1, \ldots, f_n. \end{aligned}$$

$X \le_{KC} Y$ *if and only if (3.1) holds for all increasing, positive, and convex functions* $f_1, \ldots, f_n$.

Furthermore, for $X = (X_1, \ldots, X_n)$, let $X_{(j)}$ denote the $(n-1)$-variate random vector with the jth component X_j omitted.

PROPOSITION 3.2. *Let the n-variate random vectors X and Y have equal* $(n-1)$*-variate distribution functions* $X_{(j)} =_d Y_{(j)}$ *for each* $j = 1, \ldots, n$. *Then* $X \le_K Y$ *is equivalent to the inequality (3.1) for all increasing functions* $f_1, \ldots, f_n$

In the preceding sections stochastic orders for random variables with values in general partially ordered spaces $(E, \le)$ were defined. The idea used there was to map the E-valued random variables by real functions into the real axis and then to apply a known stochastic order to the real valued random variables obtained, especially the partial order $\le_d$. The crucial problem is to choose a class of mapping functions to the reals and a class of functions (for which the expectation inequality holds) corresponding to the stochastic order on the reals. This means that both classes of functions must be compatible in the sense that the composed function is in the class of mapping functions.

Bearing this concept in mind, we give the definitions of several stochastic product partial orderings. We mainly use the relation $\le_K$ between n-dimensional distribution functions because of the preservation character of the inherent increasing functions.

Let $X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_n)$ be random vectors belonging to $(E^n, <)$.

DEFINITION 3.1. Two n-variate E^n-valued random vectors X, Y are said to be ordered with respect to

$$\begin{aligned} X \le_{KP} Y \text{ iff } &(f_1(X_1), \ldots, f_n(X_n)) \le_K (f_1(Y_1), \ldots, f_n(Y_n)) \\ &\text{for all increasing functions } f_1, \ldots, f_n \in \mathcal{F}_i. \end{aligned} \qquad (3.2)$$

$$X \leq_{CSP} Y \text{ iff (3.2) holds for all functions } f_1, \dots, f_n \in \mathcal{F}_{us}, \text{ the class of all unimodal and symmetric functions.}$$

$$X \leq_{CIP} Y \text{ iff (3.2) is true for all functions } f_1, \dots, f_n \in \mathcal{F}_{ui}, \text{ the class of all unimodal and increasing functions.}$$

$$X \leq_{DP} Y \text{ iff } (f_1(X_1), \dots, f_n(X_n)) \leq_D (f_1(Y_1), \dots, f_n(Y_n)) \text{ for all increasing functions } f_1, \dots, f_n \in \mathcal{F}_i. \tag{3.3}$$

$$X \leq_{CDP} Y \text{ iff (3.3) for all increasing, unimodal } f_1, \dots, f_n \in \mathcal{F}_{ui}.$$

$$X \leq_{KCP} Y \text{ iff } (f_1(X_1) \dots f_n(X_n)) \leq_{KC} (f_1(Y_1) \dots f_n(Y_n)) \text{ for all increasing and convex functions } f_1, \dots, f_n \in \mathcal{F}_{ci}.$$

The index "P" of the orders indicates that it is a product partial order. "C" comes from convex, "S" from symmetric, "I" from increasing, "D" from distribution.

A characterization of the six product partial orderings in terms of probability inequalities for sets with certain properties and of inequalities between expectations for certain classes of functions is given in the next theorem.

THEOREM 8.

(i) *$X \leq_{KP} Y$ is equivalent to each of the following three conditions:*

(a) $$P(X_1 \in A_1, \dots, X_n \in A_n) \leq P(Y_1 \in A_1, \dots, Y_n \in A_n) \tag{3.4}$$

for all sets $A_1, \dots, A_n \in \mathcal{A}_i$, the class of all increasing sets.

(b) $$P(f_1(X_1) \geq t_1, \dots, f_n(X_n) \geq t_n) \leq P(f_1(Y_1) \geq t_1, \dots, f_n(Y_n) \geq t_n) \ \forall t_1, \dots, t_n \in \mathbb{R}^1 \tag{3.5}$$

whenever $f_1, \dots, f_n \in \mathcal{F}_i$, the class of all increasing functions $f_i : E \to \mathbb{R}^1$.

(c) $$E(f_1(X_1) \dots f_n(X_n)) \leq E(f_1(Y_1) \dots f_n(Y_n)) \tag{3.1}$$

for all functions $f_1, \dots, f_n \in \mathcal{F}_{i+}$, the class of all increasing and positive functions provided the expectations exist.

(ii) *$X \leq_{CSP} Y$ if and only if one of the following conditions holds:*

(a) *(3.4) holds for all sets $A_1, \dots, A_n \in \mathcal{A}_{cs}$, the class of convex, symmetric sets.*

(b) *(3.5) holds for all unimodal, symmetric functions $f_1, \dots, f_n \in \mathcal{F}_{us}$.*

(c) *(3.1) holds whenever $f_1, \dots, f_n \in \mathcal{F}_{us+}$, the class of all unimodal symmetric positive functions.*

(iii) *$X \leq_{CIP} Y$ if and only if one of the following conditions holds:*

(a) *(3.4) holds for all sets* $A_1, \dots, A_n \in \mathcal{A}_{ci}$*, the class of convex, increasing sets.*

(b) *(3.5) holds for all unimodal, increasing functions* $f_1, \dots, f_n \in \mathcal{F}_{ui}$.

(c) *(3.1) holds whenever* $f_1, \dots, f_n \in \mathcal{F}_{ui+}$*, the class of all unimodal, increasing positive functions.*

(iv) $X \leq_{DP} Y$ *if and only if one of the following conditions holds:*

(a) *(3.4) holds for all sets* $A_1, \dots, A_n \in \mathcal{A}_d$*, the class of decreasing sets.*

(b′) $P(f_1(X_1) < t_1, \dots, f_n(X_n) < t_n) \leq P(f_1(Y_1) < t_1, \dots, f_n(Y_n) < t_n)\ \forall t_1, \dots, t_n \in \mathbb{R}^1$ (3.5′)

for all increasing $f_1, \dots, f_n \in \mathcal{F}_i$.

(c) *(3.1) holds whenever* $f_1, \cdots, f_n \in \mathcal{F}_{d+}$*, the class of all increasing, positive functions.*

(v) $X \leq_{CDP} Y$ *if and only if one of the following conditions holds:*

(a) *(3.4) holds for all sets* $A_1, \dots, A_n \in \mathcal{A}_{cd}$*, the class of convex, decreasing sets.*

(b′) *(3.5′) holds for all reverse unimodal, increasing functions* $f_1, \dots, f_n \in \mathcal{F}_{rui}$.

(c) *(3.1) holds whenever* $f_1, \dots, f_n \in \mathcal{F}_{ud+}$*, the class of all unimodal, decreasing, positive functions.*

(vi) $X \leq_{KCP} Y$ *if and only if one of the following conditions holds:*

(c′) $E((f_1(X_1) - t_1)_+ \cdot \dots \cdot (f_n(X_n) - t_n)_+) \leq E((f_1(Y_1) - t_1)_+ \cdot \dots \cdot (f_n(Y_n) - t_n)_+)\ \forall\, t_1, \dots, t_n \in \mathbb{R}^1$ *for all convex, increasing functions* $f_1, \dots f_n \in \mathcal{F}_{ci}$ *where* $z_+ = \max(z, 0)$.

(b″) $\int_{t_1}^{\infty} \dots \int_{t_n}^{\infty} P(f_1(X_1) \geq x_1, \dots, f_n(X_n) \geq x_n) dx_n \dots dx_1 \leq \int_{t_1}^{\infty} \dots \int_{t_n}^{\infty} P(f_1(Y_1) \geq x_1, \dots, f_n(Y_n) \geq x_n) dx_n \dots dx_1$ *whenever* $f_1, \dots, f_n \in \mathcal{F}_{ci}$.

(c) *(3.1)(4.1) holds whenever* $f_1, \dots, f_n \in \mathcal{F}_{ci+}$ *the class of all convex, increasing, positive functions.*

The proof of Theorem 8 is obtained by combining Theorem 1 with Proposition 3.1 above and is therefore not given here.

REMARK concerning the statements of Theorem 8 for the relations $\leq_{CIP}$ and $\leq_{CDP}$. We note that a concave or log-concave function is unimodal and a convex or log-convex function is reverse unimodal. This has the consequence that the class of unimodal and increasing functions $\mathcal{F}_{ui}$ contains all increasing concave or log-concave functions and the class of reverse unimodal increasing functions $\mathcal{F}_{rui}$ contains all increasing convex or log-convex functions. Therefore, our definitions generalize those of Shaked (1982) who used the more restricted class of concave increasing functions.

A still more general product partial order can be defined by replacing the property of symmetry by a more general invariance property.

For this purpose, consider a closed subgroup G of the orthogonal group $\mathcal{O}(E)$ in a partially ordered inner product space $(E, \leq, \langle\cdot,\cdot\rangle)$.

DEFINITION 3.2. The n-variate E^n-valued random vectors X, Y are said to be ordered with respect to $X \leq_{CIVP} Y$ if

$$(f_1(X_1), \ldots, f_n(X_n)) \leq_K (f_1(Y_1), \ldots, f_n(Y_n))$$
$$\text{for all unimodal and } G\text{-invariant functions } f_1, \ldots, f_n \in \mathcal{F}_{uiv}.$$

THEOREM 9. *$X \leq_{CIVP} Y$ is equivalent to each of the following three conditions:*

(a) *$P(X_1 \in A_1, \ldots, X_n \in A_n) \leq P(Y_1 \in A_1, \ldots, Y_n \in A_n)$ for all convex, G-invariant sets $A_1, \ldots, A_n \in \mathcal{A}_{civ}$.*

(b) *$P(f_1(X_1) \geq t_1, \ldots, f_n(X_n) \geq t_n) \leq P(f_1(Y_1) \geq t_1, \ldots, f_n(Y_n) \geq t_n)$ where $t_1, \ldots, t_n \in \mathbb{R}^1$ for all unimodal G-invariant functions $f_1, \ldots, f_n \in \mathcal{F}_{uiv}$.*

(c) *$E(f_1(X_1) \cdot \ldots \cdot f_n(X_n)) \leq E(f_1(Y_1) \cdot \ldots \cdot f_n(Y_n))$ whenever $f_1, \ldots, f_n \in \mathcal{F}_{uiv}$, the class of unimodal G-invariant and positive functions for which the expectations exist.*

Since the proof parallels that of Theorem 8, it is omitted.

For the relation in Theorem 9 and for the first three relations in Theorem 8 further equivalent conditions can be added.

For this reason, we mention that for $\mathbb{R}^n$ Rüschendorf (1980) and Mosler (1982) established the equivalence of $X \leq_K Y$ to the expectation inequality for a more general class of functions than considered in (3.1). Namely

$$X \leq_K Y \text{ if and only if } Ek(X) \leq Ek(Y)$$

(provided the expectations are finite) whenever $k \in \mathcal{F}_\Delta$, the class of all functions $k : \mathbb{R}^n \to \mathbb{R}^1$ which are right continuous and $k(\ldots, x_{i_1}, \ldots, x_{i_k}, \ldots)$ is Δ-monotone on $\mathbb{R}^{n-k}$ for all $0 \leq k \leq n-1$ and all $x_{i_1}, \ldots, x_{i_k}$. (An $(n-k)$-variate function is called Δ-monotone if the multivariate difference operator is nonnegative. In the case of an absolutely continuous function this is equivalent to the nonnegativity of the mixed partial derivative of order $n-k$.)

Then, for all stochastic product partial orderings defined by means of $\leq_K$ a fourth condition can be given by applying Rüschendorf's result.

THEOREM 8–9′.

(i) $X \leq_{KP} Y$ *iff*

$$Ek(f_1(X_1), \ldots, f_n(X_n)) \leq Ek(f_1(Y_1), \ldots, f_n(Y_n)) \tag{3.6}$$

for all increasing functions $f_1, \ldots, f_n \in \mathcal{F}_i$ *and all* Δ*-monotone functions* $k \in \mathcal{F}_\Delta$ *for which the expectations are finite.*

(ii) $X \leq_{CSP} Y$ *iff (3.6) holds for all* $k \in \mathcal{F}_\Delta$ *and for all unimodal, symmetric* $f_1, \ldots, f_n \in \mathcal{F}_{us}$.

(iii) $X \leq_{CIP} Y$ *iff (3.6) holds for all* $k \in \mathcal{F}_\Delta$ *and for all unimodal, increasing* $f_1, \ldots, f_n \in \mathcal{F}_{ui}$.

(iv) $X \leq_{CIVP} Y$ *iff (3.6) holds for all* $k \in \mathcal{F}_\Delta$ *and for all unimodal,* G*-invariant* $f_1, \ldots, f_n \in \mathcal{F}_{uiv}$.

For the relations $\leq_{KP}$, $\leq_{CSP}$, $\leq_{CIP}$ we now consider the case of identical $(n-1)$-variate marginal distributions. Then, in Theorems 8 and 9 the restriction to positive functions ("product"-expectation inequality) can be dropped by applying Proposition 3.2. This is of particular interest in the case of bivariate random vectors. It becomes especially important when we compare with the independent case; here the assumption of equal marginal distributions holds in a natural way.

THEOREM 10. *Let the* n*-variate random vectors* $X, Y \in E^n$ *have equal* $(n-1)$*-variate distributions,* $X_{(j)} =_d Y_{(j)}$ *for each* $j = 1, \ldots, n$. *Then*

(i) $X \leq_{KP} Y$ *iff*

$$E(f_1(X_1) \cdot \ldots \cdot f_n(X_n)) \leq E(f_1(Y_1) \cdot \ldots \cdot f_n(Y_n)) \tag{3.1}$$

for all increasing functions $f_1, \ldots, f_n \in \mathcal{F}_i$.

(ii) $X \leq_{CSP} Y$ *iff (3.1) holds for all unimodal, symmetric functions* $f_1, \ldots, f_n \in \mathcal{F}_{us}$.

(iii) $X \leq_{CIP} Y$ *iff (3.1) holds for all unimodal, increasing functions* $f_1, \ldots, f_n \in \mathcal{F}_{ui}$.

(iv) $X \leq_{CIVP} Y$ *iff (3.1) holds for all unimodal,* G*-invariant functions* $f_1, \ldots, f_n \in \mathcal{F}_{uiv}$.

We should like to close this section with two remarks which are of particular importance since they show the generality of the definitions of our product partial orderings.

First we point out that the partial ordering $\leq$ of the space $(E, \leq)$ has a very large range of applications. Especially if we consider a partially ordered inner product space $(E, \leq, \langle\cdot,\cdot\rangle)$ we can again choose the partial ordering induced by an orthogonal subgroup G, i.e. $\leq_G$. Then, the application of the product ordering $\leq_{KP}$ of increasing functions and sets, respectively, is of particular interest. If we take $E = \mathbb{R}^k$ and $G = \mathcal{P}_k$, the group of all $k \times k$ permutation matrices, we obtain stochastic majorization product orderings (since Schur-convex functions are defined as increasing functions via the majorization preorder).

We resume these facts in the following definition and theorem.

Let the n-variate random vectors $X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_n)$ be defined on the n-dimensional product space $\mathbb{R}^k \times \ldots \times \mathbb{R}^k$, i.e. each component X_i, Y_i has values in $\mathbb{R}^k$ for $i = 1, \ldots, n$.

DEFINITION 3.3. X is said to majorize stochastically Y in the product space, in symbols $X \leq_{SCP} Y$, if

$$(f_1(X_1), \ldots, f_n(X_n)) \leq_K (f_1(Y_1), \ldots, f_n(Y_n))$$
$$\text{for all Schur-convex functions } f_1, \ldots, f_n \in \mathcal{F}_{sc}.$$

THEOREM 11. *$X \leq_{SCP} Y$ is equivalent to each of the following conditions:*

(a) $P(X_1 \in A_1, \ldots, X_n \in A_n) \leq P(Y_1 \in A_1, \ldots, Y_n \in A_n)$ *for all Schur-convex sets* $A_1, \ldots, A_n \in \mathcal{A}_{sc}$.

(b) $P(f_1(X_1) > t_1, \ldots, f_n(X_n) > t_n) \leq P(f_1(Y_1) > t_1, \ldots, f_n(Y_n) > t_n)$ *for all* $t_1, \ldots, t_n \in \mathbb{R}^1$ *and all Schur-convex functions* $f_1, \ldots, f_n \in \mathcal{F}_{sc}$.

(c) $E(f_1(X_1) \cdot \ldots \cdot f_n(X_n)) \leq E(f_1(Y_1) \cdot \ldots \cdot f_n(Y_n))$ *for all Schur-convex, positive functions* $f_1, \ldots, f_n \in \mathcal{F}_{sc+}$ *for which the expectations exist.*

In the case of equal $(n-1)$-component distribution functions (c) the above is true for all Schur-convex functions $f_1, \ldots, f_n$ without the assumption of nonnegativity.

A further interesting example is to take $G = \mathcal{D}_k$ as the matrix group of coordinate sign changes on $\mathbb{R}^k$. Then we obtain the absolute value ordering, see Eaton (1982), (1987, p. 157). The stochastic absolute value ordering for product spaces is useful in the investigation of absolute value association problems, see Ahmed, Leon, and Proschan (1981), and Jogdeo (1977).

The last two examples give rise to our second remark. They show the generality of the definitions of the product partial orderings in a further sense. What we want to point out is that for the space E again a product space may be chosen. For $E = \mathbb{R}^k$ with $\leq$ the usual componentwise order, the stochastic ordering $\leq_{KP}$ for $n = 2$ leads to the well known notion of association in the sense of Esary, Proschan, and Walkup (1967). On the other hand, if E is a product space of general partially ordered spaces $(E_i, \leq^i)$, $i = 1, \ldots, k$, then for $n = 2$ we get the concept of generalized association of Ahmed, Leon, and Proschan (1981).

4. Concepts of Dependence and Association. There are many papers dealing with various concepts of dependence and association. This is done by comparison with the independent case. Since independence is a basic assumption in many statistical procedures, a crucial question is whether the validity of the results is affected by certain departures from the independence assumption.

Our approach to a unified concept of dependence and association is to compare the dependent vector X with its independent version Y by product partial orderings of the preceding section.

Our framework includes the well known dependence and association concepts: positive or negative upper and lower orthant dependence and the positive dependence notions of Shaked (1982) as well as the notions of association of Esary, Proschan, and Walkup (1967) and of generalized association of Ahmed, Leon, and Proschan (1981).

Let the n-component random vector $X = (X_1, \ldots, X_n)$ have values in E^n for a partially ordered space $(E, \leq)$ and $\mathcal{F}$ a class of functions defined on E.

Definition 4.1. The random vector $X = (X_1, \ldots, X_n)$ is said to be positive (negative) $\mathcal{F}$-dependent if

$$E(f_1(X_1) \cdot \ldots \cdot f_n(X_n)) \geq (\leq) \prod_{i=1}^{n} Ef_i(X_i)$$

whenever $f_1, \ldots, f_n \in \mathcal{F}$ and the expectations exist.

Consider the classes $\mathcal{F}_{i+}$ of increasing, positive, $\mathcal{F}_{d+}$ of decreasing, positive, $\mathcal{F}_{us+}$ of unimodal, symmetric, positive, $\mathcal{F}_{ui+}$ of unimodal, increasing, positive functions, $\mathcal{F}_{ud+}$ of unimodal, decreasing, positive, $\mathcal{F}_{ci+}$ of convex, increasing, positive, $\mathcal{F}_{uiv+}$ of unimodal, G-invariant, positive functions. From Theorems 8 and 9 (replacing there Y by an independent version X we immediately get the following characterizations and equivalent statements in terms of probability inequalities of the $\mathcal{F}$-independence definition.

Theorem 12. *X is positive (negative) dependent with respect to*

(i) $\mathcal{F}_{i+}$ *iff*

$$P(X_1 \in A_1, \ldots, X_n \in A_n) \geq (\leq) P(X_1 \in A_1) \cdot \ldots \cdot P(X_n \in A_n) \qquad (4.1)$$

for all increasing sets $A_1, \ldots, A_n \in \mathcal{A}_i$.

(ii) $\mathcal{F}_{d+}$ *iff (4.1) for all decreasing sets* $A_1, \ldots, A_n \in \mathcal{A}_d$.

(iii) $\mathcal{F}_{us+}$ *iff (4.1) for all convex, symmetric sets* $A_1, \ldots, A_n \in \mathcal{A}_{cs}$.

(iv) $\mathcal{F}_{ui+}$ *iff (4.1) for all convex, increasing sets* $A_1, \ldots, A_n \in \mathcal{A}_{ci}$.

(v) $\mathcal{F}_{ud+}$ *iff (4.1) for all convex, decreasing sets* $A_1, \cdots, A_n \in \mathcal{A}_{cd}$.

(vi) $\mathcal{F}_{uiv+}$ *iff (4.1) for all convex, G-invariant sets* $A_1, \ldots, A_n \in \mathcal{A}_{civ}$.

(vii) $\mathcal{F}_{ci+}$ *iff*

$$\int_{t_1}^{\infty} \cdots \int_{t_n}^{\infty} P(f_1(X_1) > x_1, \ldots, f_n(X_n) > x_n) dx_n \ldots dx_1 \geq (\leq)$$
$$\prod_{i=1}^{n} \int_{t_i}^{\infty} P(f_i(X_i) > x_i) dx_i \;\; \text{whenever} \;\; f_1, \ldots, f_n \in \mathcal{F}_{ci}.$$

Of special interest is the two-component case $n = 2$ of a two-dimensional product space. Then the marginal distributions are equal and Theorem 10 applies. Now, the $\mathcal{F}$-dependence of $X = (X_1, X_2)$ can be expressed in an alternative way in terms of the covariance, namely

$$\begin{array}{c} \operatorname{cov}(f_1(X_1), f_2(X_2)) \geq (\leq)\ 0 \;\; \text{whenever} \;\; f_1, f_2 \in \mathcal{F} \\ \text{provided that the expectations exist.} \end{array} \qquad (4.2)$$

Proposition 4.1. *For a bivariate random vector* $X = (X_1, X_2)$ *with values in* $E \times E$ *the statements of Theorem 12 are true for all dependence classes of functions* $\mathcal{F}_i$, $\mathcal{F}_d$, $\mathcal{F}_{us}$, $\mathcal{F}_{ui}$, $\mathcal{F}_{ud}$, $\mathcal{F}_{uiv}$, $\mathcal{F}_{ci}$, *that means, without the assumption of positiveness.*

It is easy to see that the following proposition holds.

Let $-\mathcal{F} = \{f : -f \in \mathcal{F}\}$.

Proposition 4.2. *Let* $X = (X_1, X_2)$. *X is positive (negative)* $\mathcal{F}$*-dependent iff X is positive (negative)* $-\mathcal{F}$*-dependent.*

This means for example that the $\mathcal{F}_{ui}$-dependence also includes all reverse unimodal and decreasing functions and that the $\mathcal{F}_i$-increasing dependence and the $\mathcal{F}_d$-decreasing dependence are equivalent.

There is an easy way to prove a proposition corresponding to Proposition 4.2 in terms of classes of sets $\mathcal{A}$ in view of Theorem 12.

We say that X is positive (negative) $\mathcal{A}$-dependent if the probability inequality (4.1) of Theorem 12 holds for all sets $A_1, \ldots, A_n \in \mathcal{A}$.

For the set $A \subset E$ let $\mathcal{A}^c = \{A : A^c \in \mathcal{A}\}$ where A^c denotes the complement of A.

PROPOSITION 4.3. *$X = (X_1, X_2)$ is $\mathcal{A}$-dependent if and only if X is $\mathcal{A}^c$-dependent.*

A restatement of Theorem 12 gives

PROPOSITION 4.4. *X is $\mathcal{F}$-dependent if and only if X is $\mathcal{A}$-dependent. Here $\mathcal{F}$ is any of the classes of functions $\mathcal{F}_i$, $\mathcal{F}_d$, $\mathcal{F}_{us}$, $\mathcal{F}_{ui}$, $\mathcal{F}_{ud}$, $\mathcal{F}_{uiv}$, and $\mathcal{A}$ the corresponding class $\mathcal{A}_i$, $\mathcal{A}_d$, $\mathcal{A}_{cs}$, $\mathcal{A}_{ci}$, $\mathcal{A}_{cd}$, $\mathcal{A}_{civ}$.*

The result stated in the following example concerns the Gaussian case; for a reference see Tong (1980).

EXAMPLE 4.1. Suppose that $Y = (Y_1, Y_2)$ has an $(n_1 + n_2)$-dimensional normal distribution $N_{n_1+n_2}\,(0, \Sigma)$ with $Y_i \in \mathbb{R}^{n_i}$ for $i = 1, 2$ and where

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

If rank $\Sigma_{12} = 1$ then Y is positive $\mathcal{A}_{cs}$-dependent, which is equivalent to the fact that Y is positive $\mathcal{F}_{us}$-dependent. A corollary of this result is

EXAMPLE 4.2. Suppose $X = (X_1, \ldots, X_n)$ has an n-dimensional normal distribution $N_n(0, \Sigma)$. Then X is positive $\mathcal{A}_{cs}$-dependent; or, equivalently, X is positive $\mathcal{F}_{us+}$-dependent.

The concept of association is related to a single random vector Z. In our context the notion of association is obtained by considering a random vector with two identical components $X_1 = X_2 = Z$.

DEFINITION 4.2. Let the random vector Z have values in the partially ordered space $(E, \leq)$. Z is said to be positive (negative) $\mathcal{F}$-associated if $X = (Z, Z)$ is positive (negative) $\mathcal{F}$-dependent; or, equivalently, because of (4.2) if

$$\begin{aligned} &\operatorname{cov}(f_1(Z), f_2(Z)) \geq (\leq) 0 \\ &\qquad \text{whenever } f_1, f_2 \in \mathcal{F} \text{ and the covariance exists.} \end{aligned} \tag{4.3}$$

We call Z positive (negative) $\mathcal{A}$-associated if $X = (Z, Z)$ is positive (negative) $\mathcal{A}$-dependent, that is if

$$P(Z \in A_1, Z \in A_2) \geq (\leq) P(Z \in A_1) P(Z \in A_2) \quad \text{for all sets } A_2, A_2 \in \mathcal{A}.$$

Then we have

PROPOSITION 4.5. *For any of the classes $\mathcal{F}_i$, $\mathcal{F}_d$, $\mathcal{F}_{us}$, $\mathcal{F}_{ui}$, $\mathcal{F}_{ud}$, $\mathcal{F}_{uiv}$ Z is $\mathcal{F}$-associated if and only if Z is $\mathcal{A}$-associated where $\mathcal{A}$ is the corresponding class $\mathcal{A}_i$, $\mathcal{A}_d$, $\mathcal{A}_{cs}$, $\mathcal{A}_{ci}$, $\mathcal{A}_{cd}$, $\mathcal{A}_{civ}$.*

The next example gives a result concerning the Gaussian case, see Eaton (1987) for a reference.

EXAMPLE 4.3. Suppose $X = (X_1, \ldots, X_k)$ has a k-dimensional normal distribution $N_k(0, \Sigma)$ with zero mean. If each element of $\Sigma = (\sigma_{ij})$ is nonnegative then X is positive $\mathcal{A}_i$-associated (which is equivalent to the positive $\mathcal{F}_i$-association of X).

If we consider the k-dimensional Euclidean space $E = \mathbb{R}^k$ the positive association of $Z = (Z_1, \ldots, Z_k)$ with respect to the class of all increasing functions $\mathcal{F}_i$ coincides with the association notion of Esary, Proschan, and Walkup (1967).

On the other hand, the remarks at the end of Section 3 concerning the generality of the definitions of the product partial orderings apply to the dependence and association concepts, too. Especially, if E is chosen to be the k-dimensional product space of partially ordered spaces $(E_j, \leq^j)$, $j = 1, \ldots, k$, endowed with the componentwise ordering, and the Z_j are random components of $Z = (Z_1, \ldots, Z_k)$ with values in E_j, then the positive $\mathcal{F}_i$-association of Z is the same as the generalized association of Ahmed, Leon, and Proschan (1981).

A further interesting example is that of Schur-convex dependence and association. Let $E = \mathbb{R}^k$. Then the real valued random vector Z is positive (negative) Schur-convex associated with respect to $\mathcal{F}_{sc}$ if and only if $\text{cov}(f_1(Z), f_2(Z)) \geq (\leq) 0$ for all Schur-convex $f_1, f_2 \in \mathcal{F}_{sc}$. This is equivalent to saying that Z is positive (negative) $\mathcal{A}_{sc}$-associated which means $P(Z \in A_1, Z \in A_2) \geq (\leq) P(Z \in A_1)P(Z \in A_2)$ for all Schur-convex sets $A_1, A_2 \in \mathcal{A}_{sc}$.

We close this section with two remarks. First we note that the notions of dependence and association with respect to classes of functions $\mathcal{F}$ are more general than those relative to classes of sets $\mathcal{A}$. This is the case since the expectation inequality for functions may not correspond to a probability inequality (4.1); an example is the class of convex and increasing functions $\mathcal{F}_{ci}$.

Second, we refer to Mosler (1982) who used a related approach to define some dependence and association concepts.

5. Remarks Concerning Applications of Comparisons of Dependence Structures. The product partial orderings and dependence and association notions of Sections 3 and 4 may be applied to compare dependence structures of stochastic processes and sets of random variables, especially with respect to the strength of their dependence. Two principal cases have to be distinguished, namely, internal comparisons within a set of random variables (or within a stochastic process in order to assess the development of the degree of dependence between progressive time points) and, on the other hand, external comparisons between two sets of random variables which is, for example, the purpose of canonical correlation analysis, or between two stochastic processes in order to obtain qualitative statements.

(A) Internal comparisons.

(i) Internal monotonicity of a stochastic process.

Denote by $\leq_R$ any of the relations of $R_\leq = \{\leq_{KP}, \leq_{CSP}, \leq_{CIP}, \leq_{DP}, \leq_{CDP}, \leq_{CIVP}, \leq_{SCP}\}$ and by $\mathcal{F}$ any of the function classes of $\mathcal{F}' = \{\mathcal{F}_i, \mathcal{F}_{us}, \mathcal{F}_{ui}, \mathcal{F}_d, \mathcal{F}_{ud}, \mathcal{F}_{uiv}, \mathcal{F}_{sc}\}$ and by $\mathcal{A}$ any of the classes of sets $\mathcal{A}' = \{\mathcal{A}_i, \mathcal{A}_{cs}, \mathcal{A}_{ci}, \mathcal{A}_d, \mathcal{A}_{cd}, \mathcal{A}_{civ}, \mathcal{A}_{sc}\}$.

Let $\{X_t\}_{t\in T}$ be a given stochastic process. By means of Theorem 10 we obtain the following monotonicity property of the covariances.

PROPOSITION 5.1. *For $t_1, t_2, t_3, t_4 \in T$ we have*

$$\operatorname{cov}(f(X_{t_1}), g(X_{t_2})) \leq \operatorname{cov}(f(X_{t_3}), g(X_{t_4}))$$

whenever $f, g \in \mathcal{F}$, where $\mathcal{F}$ is an element of $\mathcal{F}'$, if and only if

$$(X_{t_1}, X_{t_2}) \leq_R (X_{t_3}, X_{t_4})$$

for the relation of $\leq_R$ of $R_\leq$ corresponding to $\mathcal{F} \in \mathcal{F}'$ if and only if

$$P(X_{t_1} \in A, X_{t_2} \in B) \leq P(X_{t_3} \in A, X_{t_4} \in B)$$

whenever $A, B \in \mathcal{A}$, the corresponding element of $\mathcal{A}'$.

The first equivalence statement is also true for $\mathcal{F}_{ci}$ and the relation $\leq_{KCP}$.

REMARK. If more than two components of the stochastic process are compared then, in addition, all functions must be positive.

Comparison with the independent case leads to the notion of dependence of Section 4.

PROPOSITION 5.2. *Let $\{X_t\}_{t\in T}$ be as above and let $t, s \in T$. Then (X_t, X_s) is $\mathcal{F}$-dependent for $\mathcal{F} \in \mathcal{F}'$, that is $\operatorname{cov}(f(X_t), g(X_s)) \geq 0$ for all $f, g \in \mathcal{F}$, where $\mathcal{F}$ is an element of $\mathcal{F}'$ if and only if*

$$(X_t, X_s) \text{ is } \mathcal{A}\text{-dependent for } \mathcal{A},$$

i.e. $P(X_t \in A, X_s \in B) \geq P(X_t \in A)P(X_s \in B)$ *for all* $A, B \in \mathcal{A}$*, where* $\mathcal{A}$ *is the element of* $\mathcal{A}'$ *corresponding to* $\mathcal{F}$.

(ii) Internal association of a set of random variables.

Let $Z_1, \ldots, Z_k$ be a set of random variables, each with values in $(E, \leq)$. Then Proposition 4.5 directly applies. For $\mathcal{F} \in \mathcal{F}'$ the set $Z_1, \ldots, Z_k$ is $\mathcal{F}$-associated if and only if $Z_1, \ldots, Z_k$ are $\mathcal{A}$-associated where $\mathcal{A}$ is the corresponding element of $\mathcal{A}'$.

(B) External comparisons.

(i) External monotonicity of stochastic processes.

Let $\{X_t\}$ and $\{Y_t\}$, $t \in T$, be two stochastic processes, with values in $(E, \leq)$.

By $\mathcal{F}'_+$ we denote the family $\mathcal{F}'$ of classes of functions which are in addition nonnegative.

PROPOSITION 5.3. *For any of the relations* $\leq_R \in R_{\leq}$ *and* $t, s \in T$ $(X_t, X_s) \leq_R (Y_t, Y_s)$ *if and only if*

$$Ef(X_t)g(X_s) \leq Ef(Y_t)g(Y_s) \text{ for all } f, g \in \mathcal{F},$$

the element corresponding to $\leq_R$ *of* $\mathcal{F}'_+$*, if and only if*

$$P(X_t \in A, X_s \in B) \leq P(Y_t \in A, Y_s \in B)$$

whenever $A, B \in \mathcal{A}$*, the corresponding element of* $\mathcal{A}'$.

REMARK. This proposition is especially useful if the elements of the stochastic processes are related by a mapping which has the properties of the corresponding class of functions; see Bergmann (1978).

(ii) External comparison of sets of random variables.

A typical question in statistics (which arises in many data analysis problems) is whether two sets of random variables are correlated. For example, in canonical correlation analysis the dependence structure between the two sets is investigated by constructing linear relationships in each of the sets and then finding out the correlation between two linear functions.

The dependence notion of Section 4 permits us to consider function classes other than the linear one.

Let $Z_1, \ldots, Z_k$ and $V_1, \ldots, V_\ell$ be random variables each having values in $E(, \leq)$.

PROPOSITION 5.4. $Z = (Z_1, \ldots, Z_k)$ *and* $V = (V_1, \ldots, V_\ell)$ *are positive (negative)* $\mathcal{F}$*-dependent for* $\mathcal{F} \in \mathcal{F}'$, i.e. *cov*$(f(Z_1, \ldots, Z_k),\ g(V_1, \ldots, V_\ell)) \geq$

$(\leq)0$ *for all* $f, g \in \mathcal{F}$, *where* $\mathcal{F}$ *is an element of* $\mathcal{F}'$, *if and only if* (Z, V) *is positive (negative)* $\mathcal{A}$*-dependent for* $\mathcal{A} \in \mathcal{A}'$, *i.e.* $P(Z \in A, V \in B) \geq (\leq) P(Z \in A)P(V \in B)$ *for* $A, B \in \mathcal{A}$, *the element of* $\mathcal{A}'$ *corresponding to* $\mathcal{F}$.

REFERENCES

AHMED, A.-H. N., LEON, R., and PROSCHAN, F. (1981). Generalized association with applications in multivariate statistics. *Ann. Statist.* **9**, 168–176.

ANDERSON, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6**, 170–176.

BERGMANN, R. (1978). Some classes of semi-ordering relations for random vectors and their use for comparing covariances. *Math. Nachr.* **82**, 103–114.

BERGMANN, R. and STOYAN, D. (1978). Monotonicity properties of second order characteristics of stochastically monotone Markov chains. *Math. Nachr.* **82**, 99–102.

BIRNBAUM, Z. W. (1948). On random variables with comparable peakedness. *Ann. Math. Statist.* **19**, 76–81.

DHARMADHIKARI, S. and JOAG-DEV, K. (1988). *Unimodality, Convexity, and Applications.* Academic Press, Boston.

EATON, M. L. (1982). A review of selected topics in multivariate probability inequalities. *Ann. Statist.* **10**, 11–43.

EATON, M. L. (1987). Lectures on topics in probability inequalities. CWI Tract 35, Amsterdam.

EATON, M. L. and PERLMAN, M. D. (1991). Multivariate probability inequalities: convolution theorems, composition theorems and concentration inequalities. This volume.

ESARY, J. D., PROSCHAN, F., and WALKUP, D. W. (1967). Association of random variables with applications. *Ann. Math. Statist.* **38**, 433–441.

FEFFERMAN, C., JODEIT, M., and PERLMAN, M. D. (1972). A spherical surface measure inequality for convex sets. *Proc. Amer. Math. Soc.* **33**, 114–119.

JOGDEO, K. (1977). Association and probability inequalities. *Ann. Statist.* **5**, 495–504.

LEVHARI, D., PAROUSH, J., and PELEG, B. (1975). Efficiency analysis for multivariate distributions. *Rev. Econom. Stud.* **42**, 87–91.

MARSHALL, A. W. (1991). Multivariate stochastic orderings and generating cones of functions. This volume.

MARSHALL, A. W. and OLKIN, I. (1974). Majorization in multivariate distributions. *Ann. Statist.* **2**, 1189–1200.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

MOSLER, K. C. (1982). Entscheidungsregeln bei Risiko: Multivariate stochastische Dominanz. *Lecture Notes Econom. Math. Systems* **204**, Springer, Berlin.

MOSLER, K. C. and SCARSINI, M. (eds.) (1991). This volume.

MUDHOLKAR, G. (1966). The integral of an invariant unimodal function over an invariant convex set – an inequality and applications. *Proc. Amer. Math. Soc.* **17**, 1327–1333.

NEVIUS, S. E., PROSCHAN, F., and SETHURAMAN, J. (1977). Schur functions in statistics. II. Stochastic majorization. *Ann. Statist.* **5**, 263–273.

NEVIUS, S. E., PROSCHAN, F., and SETHURAMAN, J. (1977b). A stochastic version of weak majorization, with applications. In: Gupta, S. S. and D. S. Moore, eds., *Statistical Decision Theory and Related Topics*, Vol. II, pp. 281–296. Academic Press, New York.

OLKIN, I., and TONG, Y. L. (1988). Peakedness in multivariate distributions. In: Gupta, S. S., and J. O. Berger, eds., *Statistical Decision Theory and Related Topics*, Vol. IV. Academic Press, New York.

RÜSCHENDORF, L. (1980). Inequalities for the expectation of Δ-monotone functions. *Z. Wahrsch. Verw. Gebiete* **54**, 349–354.

SHAKED, M. (1982). A general theory of some positive dependence notions. *J. Multivariate Anal.* **12**, 199–218.

SHERMAN, S. (1955). A theorem on convex sets with applications. *Ann. Math. Statist.* **26**, 763–766.

STOYAN, D. (1983). *Comparison Methods for Queues and Other Stochastic Models.* John Wiley & Sons, Inc., Chichester.

TONG, Y. L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic Press, New York.

KRONENSTR. 18

D-1788 RHEINFELDEN, GERMANY