Ann. Inst. Statist. Math.
31 (1979), Part A, 67-84

SOME CONCEPTS OF POSITIVE DEPENDENCE FOR BIVARIATE INTERCHANGEABLE DISTRIBUTIONS

MOSHE SHAKED*

(Received Sept. 13, 1977; revised Feb. 19, 1979)

Abstract

In this work we consider some familiar and some new concepts of positive dependence for interchangeable bivariate distributions. By characterizing distributions which are positively dependent according to some of these concepts, we indicate real situations in which these concepts arise naturally. For the various families of positively dependent distributions we prove some closure properties and demonstrate all the possible logical relations. Some inequalities are shown and applied to determine whether under- (or over-) estimates, of various probabilistic quantities, occur when a positively dependent distribution is assumed (falsely) to be the product of its marginals (that is, when two positively dependent random variables are assumed, falsely, to be independent). Specific applications in reliability theory, statistical mechanics and reversible Markov processes are discussed.

1. Introduction

Some interchangeable bivariate random vectors (X_1, X_2) (that is, random vectors with permutation invariant distributions) which satisfy

(1.1) $\mathrm{P}\{X_1 \in I,\ X_2 \in I\} \geqq \mathrm{P}\{X_1 \in I\}\,\mathrm{P}\{X_2 \in I\}$ for every interval I,

are known in the literature, (see e.g. Hewett and Bulgren [9], Jensen [13], Tong [34] and Sidak [31]). Jensen [14] observed that many of the random vectors that satisfy (1.1) satisfy also the more general inequality

* This work was partially supported by National Science Foundation GP-30707X1. It is part of the author's Ph.D. dissertation prepared at the University of Rochester and supervised by A. W. Marshall.

Key words: Positively dependent interchangeable distributions, closure properties, positive definite kernels, conditionally positive definite kernels, Bochner's theorem, random environment, reversible Markov processes.

AMS classification numbers: Primary 60E05, Secondary 62H05.

(1.2) $$P\{X_1 \in B,\ X_2 \in B\} \geqq P\{X_1 \in B\} P\{X_2 \in B\} \quad \text{for every Borel set B in R}.$$

Jensen called random vectors that satisfy (1.2) positively dependent. We will not adopt this definition because, as will be shown later (Example 2.3), there exist random vectors that satisfy (1.2) for every B and still have a correlation coefficient which is arbitrarily close to -1.

Dykstra, et al. [2], Sidak [31] and Shaked [27], [29] have discussed bivariate distributions which can be represented as mixtures of independent bivariate distributions with equal marginals. Such distributions are called positively dependent by a mixture (PDM), (see Section 2). If (X_1, X_2) has a PDM distribution then (note, whenever expectations or integrals are written they are assumed to exist)

(1.3) $$\text{Cov}\,(h(X_1),\ h(X_2)) \geqq 0 \quad \text{for every measurable function } h:\ R \rightarrow R,$$

hence (X_1, X_2) also satisfies (1.2).

In Section 2 we define some families of bivariate positively dependent interchangeable distributions through (1.1), (1.2) and (1.3). The usefulness of the various definitions that we introduce stems from the fact that many interchangeable bivariate distributions that arise naturally in some practical and theoretical situations (see Lancaster [18], Jensen [14], Shaked [27], [29]) are positively dependent according to some of our definitions. The purpose of this paper is to develop methods that can help us in identifying positively dependent distributions (Sections 2 and 3) and to prove some inequalities that are satisfied by them (Section 4). The inequalities can be used to determine whether over- (or under-) estimates occur when one acts as if positively dependent random variables are independent. Some applications are discussed in Section 5.

2. Definitions and interrelations

2.1. *Definitions*

An interchangeable random vector (X_1, X_2) or its distribution F is said to be:

(i) *diagonal square dependent* (DSD) if (1.1) holds,

(ii) *generalized diagonal square dependent* (GDSD) if (1.2) holds (Jensen [14] considered such random vectors).

(iii) *positively dependent by mixture* (PDM) if F admits the representation

(2.1) $$F(x_1, x_2) = \int_{\Omega} F^{(\omega)}(x_1) F^{(\omega)}(x_2) dG(\omega)$$

where Ω is a Borel set in R^m, G is a probability measure on Ω and $F^{(\omega)}(\cdot)$, which is Borel measurable in ω, is a univariate distribution function for every $\omega \in \Omega$ (Shaked [29]).

(iv) *positively dependent by expansion* (PDE) if F admits the expansion

$$dF(x_1, x_2) = d\tilde{F}(x_1)d\tilde{F}(x_2)\left[1 + \sum_{i=1}^{\infty} a_i \varphi_i(x_1)\varphi_i(x_2)\right], \qquad \text{a.e.} \tag{2.2}$$

where $\tilde{F}$ is the univariate marginal of F, $\{\varphi_i\}$ is a set of functions satisfying

$$\int_{-\infty}^{\infty} \varphi_i(x)d\tilde{F}(x) = 0, \qquad i = 1, 2, \cdots \tag{2.3}$$

and a_i are nonnegative real numbers (Lancaster [18] considered such distributions).

(v) *positive definite dependent* (PDD) if F is a positive definite (p.d) kernel on $S \times S$ where S is the support of X_1.

Concerning Definition (iv) we remark that for most of the known expansions of PDE distributions, the set of functions $\{\varphi_i\}_{i=1}^{\infty}$ of (2.2) satisfies, in addition to (2.3), the orthogonality conditions $\int_{-\infty}^{\infty} \varphi_i(x)\varphi_j(x) \cdot d\tilde{F}(x) = \delta_{ij}$, $i, j = 1, 2, \cdots$ (see Lancaster [18], Jensen [14] and references there).

2.2. *Two characterizations*

The following results characterize two of the families of distributions which were just defined. They can be used to identify some positively dependent distributions. Also the following propositions will be used later in the paper.

PROPOSITION 2.1 (Shaked [29]). A bivariate distribution is PDM if, and only if, it is the joint distribution of $g(U_1, W)$ and $g(U_2, W)$ for some i.i.d. random variables U_1 and U_2, a random vector W which is independent of the U_i's and a Borel measurable function g.

PROPOSITION 2.2. The random vector (X_1, X_2) is PDD if, and only if, (1.3) holds for every measurable function h.

PROOF. Assume F is PDD. Let h be a function such that $\mathrm{E}|h(X_1) \cdot h(X_2)| < \infty$. For every integer n define $h_n(x) = h(x)$ if $|x| \leqq n$, 0 otherwise. Using integration by parts on $\mathrm{E}\, h_n(X_1)h_n(X_2)$ and then using the dominated convergence theorem and the fact that the distribution of (X_1, X_2) is a covariance function it can be shown that

$$\mathrm{E}\, h(X_1)h(X_2) \geqq 0 \qquad \text{for every measurable } h\,. \tag{2.4}$$

Replacing in (2.4) $h(x)$ by $h(x) - \mathrm{E}\, h(X_1)$ we obtain (1.3).

Assume now that (1.3) holds, then (2.4) holds since $\mathrm{E}\, h(X_1)h(X_2) \geqq (\mathrm{E}\, h(X_1))(\mathrm{E}\, h(X_2)) = (\mathrm{E}\, h(X_1))^2 \geqq 0$. This fact implies, using Fubini's Theorem, that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{h}(x_1)\tilde{h}(x_2)F(x_1, x_2)dx_1dx_2 \geqq 0 \qquad \text{for all measurable } \tilde{h}\,. \tag{2.5}$$

Let $a_1, \cdots, a_n$ and $t_1, \cdots, t_n$ be arbitrary $2n$ real numbers such that $t_i \neq t_j$ when $i \neq j$. Let $\delta > 0$ and define $h_\delta(x) = \sum_{i=1}^{n} a_i I_{[t_i, t_i+\delta]}(x)$ where I_A is the indicator function of the set A. Then, using the fact that F is a distribution function, we have $\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j F(t_i, t_j) = \lim_{\delta \to 0} \delta^{-2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_\delta(x_1)h_\delta \cdot (x_2)F(x_1, x_2)dx_1dx_2 \geqq 0$, that is, F is a p.d. kernel.

2.3. *Interrelations*

Knowledge of the logical implications among the various families of Definition 2.1 can be useful in identifying positively dependent distributions. We prove the implications and counter-implications that are summarized in Fig. 2.1. The sign $\not\Longrightarrow$ means that there exists a random vector (X_1, X_2) that satisfies the property near the tail of the arrow but does not satisfy the one near the head of the arrow. All the random vectors (X_1, X_2) in Fig. 2.1 are assumed to be interchangeable. Note that all the relations of implication and counter-implication between any two of the families in Fig. 2.1 are determined by the implications and the counter-implications of Fig. 2.1.

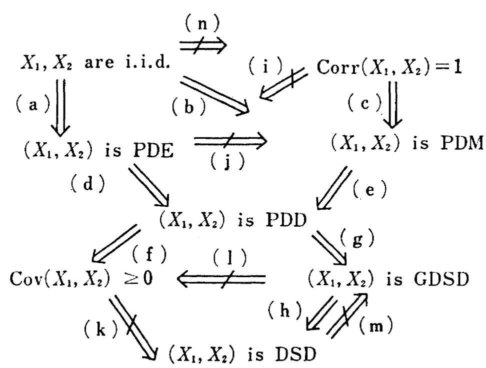

Fig. 2.1

Proofs of the implications of Fig. 2.1

The implications (a), (b), (f), (g) and (h) are obvious, (e) is proved by Dykstra et al. [2].

Proof of (c)

If (X_1, X_2) is an interchangeable random vector and $\operatorname{Corr}(X_1, X_2) = 1$ then the distribution F of (X_1, X_2) must have all its mass on the line $x_1 = x_2$. This implies that given X_1, conditionally, X_1 and X_2 are (degenerate) i.i.d. with probability one.

Proof of (d)

Let h be a Borel measurable function. Assume that (X_1, X_2) is PDE. Then by using (2.2) we have $\mathrm{E}\, h(X_1)h(X_2) = \mathrm{E}\, h(X_1)\, \mathrm{E}\, h(X_2) + \sum_{i=1}^{\infty} a_i \left(\int_{-\infty}^{\infty} h(x)\varphi_i(x) d\tilde{F}(x) \right)^2 \geqq 0$. By Proposition 2.2 the proof is complete, (this is a generalization of a theorem of Jensen [14]).

The following examples provide proofs of (i), (j) and (l) of Fig. 2.1. Counterexamples for (k), (m) and (n) are easy to construct.

Example 2.1 (Proof of (i)). If (X_1, X_2) is a bivariate normal random vector with zero means, unit variances and unit correlation coefficient then it is not PDE. In fact every bivariate distribution whose total mass is concentrated on the 45° line and has there at least uncountable points of increase is not PDE.

Example 2.2 (Proof of (j)). Let $b_1, \cdots, b_5$ be five distinct numbers and let (X_1, X_2) be a discrete random vector with $p_{ij} = \mathrm{P}(X_1 = b_i,\ X_2 = b_j)$ as given in the following table (all the probabilities here are multiplied by 42):

X_1 \ X_2	b_1	b_2	b_3	b_4	b_5
b_1	4	2	0	0	2
b_2	2	4	2	0	0
b_3	0	2	4	3	0
b_4	0	0	3	4	2
b_5	2	0	0	2	4

Denote $P = (p_{ij})_{i,j=1}^{5}$ and $q_i = \mathrm{P}(X_1 = b_i)$. We assert that (X_1, X_2) is PDE. To see this note that the matrix P is non-negative definite (Hall and Newman [8]) hence also the matrix $R = (p_{ij}/q_i^{1/2} q_j^{1/2})_{i,j=1}^{5}$ is non-negative definite. By writing down the spectral decomposition of R, noting that 1 is an eigenvalue of R associated with the eigenvector $(q_1^{1/2}, \cdots, q_5^{1/2})$, the decomposition (2.2) is established (with $a_0 = 1$, $a_i \geqq 0$, $i = 1, 2, 3, 4$, because they are eigenvalues of a non-negative definite matrix, and $a_i = 0$ for $i > 4$). The fact that (X_1, X_2) is not PDM follows from the discussion of Hall and Newman [8] concerning the matrix P.

Example 2.3 (Proof of (l)). Let (X_n, Y_n), $n = 1, 2, \cdots$ be discrete

$X_n \backslash Y_n$	-1	0	1
1	$\frac{2}{n+4}-\frac{4}{(n+4)^2}$	0	$\frac{4}{(n+4)^2}$
0	0	$\frac{n}{n+4}$	0
-1	$\frac{4}{(n+4)^2}$	0	$\frac{2}{n+4}-\frac{4}{(n+4)^2}$

random vectors with the following joint probabilities:
The vector (X_n, Y_n) is GDSD, but $\mathrm{Corr}(X_n, Y_n)=-n(n+4)^{-1}<0$. Note that $\mathrm{Corr}(X_n, Y_n)\to -1$ as $n\to\infty$.

Some concepts of positive dependence for general (i.e., not necessarily interchangeable) bivariate random vectors have been introduced recently (Lehmann [19], Esary et al. [5], [6] and Yanagimoto [36]). The strongest of these concepts, i.e., the one that implies all the others is the concept of *positive likelihood ratio dependence* (PLRD). The weakest among them is the concept of *positive quadrant dependence* (PQD). We will show now that, restricting (X_1, X_2) to be an interchangeable random vector, then

(i) (X_1, X_2) is PDM and PDE $\not\Rightarrow (X_1, X_2)$ is PQD,

(ii) (X_1, X_2) is PLRD $\not\Rightarrow (X_1, X_2)$ is DSD.

These two assertions state in fact that there is not an implication relationship between the chain of concepts of Fig. 2.1 and the chain of concepts of Esary and Proschan [5], Yanagimoto [36], or of Shaked [30].

Proof of (i)

Let Y_1 and Y_2 be i.i.d. such that $P(Y_1=-1)=1/2-(3/44)^{1/2}$, $P(Y_1=1)=1-P(Y_1=-1)$ and let W be a random variable, independent of Y_1 and Y_2 such that $P(W=-1)=P(W=1)=11/32$ and $P(W=0)=10/32$. Then the joint distribution of $X_1=|Y_1-W|$ and $X_2=|Y_2-W|$ is PDM and PDE but $P(X_1<1, X_2<2)<P(X_1<1)P(X_2<2)$ hence (X_1, X_2) is not PQD.

Proof of (ii)

Let (X_1, X_2) be a discrete random vector with the following joint probabilities (all probabilities here are multiplied by 62):

$X_1 \backslash X_2$	a_1	a_2	a_3
a_3	8	6	9
a_2	6	4	6
a_1	9	6	8

$(a_1<a_2<a_3)$

It is easily seen to be PLRD but $P(X_1=a_2, X_2=a_2)<P(X_1=a_2)P(X_2=a_2)$, hence (X_1, X_2) is not DSD.

3. Closure properties

In this section we prove some closure properties of the families of distributions that were introduced in Section 2. They can be used to identify positively dependent random vectors and to generate new distributions that belong to the family, from known ones. The results are summarized in Table 3.1.

THEOREM 3.1. *Let* (X_1, Y_1) *and* (X_2, Y_2) *be independent random vectors. If* (X_i, Y_i), $i=1,2$ *are, respectively, PDM, PDE, PDD, then* $(h(X_1, X_2), h(Y_1, Y_2))$ *is, respectively, PDM, PDE, PDD, for every Borel measurable* $h: R^2 \to R$.

PROOF. If (X_i, Y_i), $i=1,2$ are PDM then the assertion of the theorem is easily proved using Proposition 2.1.

Assume now that (X_i, Y_i), $i=1,2$ are PDE, that is, the distribution F_i of (X_i, Y_i) admits the expansion $dF_i(x, y) = d\tilde{F}_i(x)d\tilde{F}_i(y)\left[1+\sum_{k=1}^{\infty} a_k^{(i)}\varphi_k^{(i)}(x)\varphi_k^{(i)}(y)\right]$ a.e., $i=1,2$. Let h be a measurable function and define $A(x)=\{(x_1, x_2): h(x_1, x_2)\leq x\}$. Then the distribution G of $(h(X_1, X_2), h(Y_1, Y_2))$ is $G(x, y)=\iint_{(x_1,x_2)\in A(x)}\iint_{(y_1,y_2)\in A(y)} dF_1(x_1, y_1)dF_2(x_2, y_2)=\tilde{G}(x)\tilde{G}(y)+\sum_{i=1}^{2}\sum_{k=1}^{\infty} a_k^{(i)}\Psi_k^{(i)}(x)\Psi_k^{(i)}(y)+\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} a_k^{(1)}a_l^{(2)}\chi_{k,l}(x)\chi_{k,l}(y)$ where $\tilde{G}(x)=\iint_{A(x)} d\tilde{F}_1(x_1)\cdot d\tilde{F}_2(x_2)$, $\Psi_k^{(i)}(x)=\iint_{A(x)}\varphi_k^{(i)}(x_i)d\tilde{F}_i(x_i)d\tilde{F}_{3-i}(x_{3-i})$, and $\chi_{k,l}(x)=\iint_{A(x)}\varphi_k^{(1)}(x_1)\varphi_l^{(2)}(x_2)\cdot d\tilde{F}_1(x_1)d\tilde{F}_2(x_2)$. Note that for every $i=1,2$ and $k=1,2,\cdots$ the signed measure induced by $\Psi_k^{(i)}$ is absolutely continuous with respect to $\tilde{G}$, hence there exist $\tilde{\Psi}_k^{(i)}$—the Radon Nikodim derivative of $\Psi_k^{(i)}$ with respect to $\tilde{G}$, such that, $\Psi_k^{(i)}(x)=\int_{-\infty}^{x}\tilde{\Psi}_k^{(i)}(x_0)d\tilde{G}(x_0)$. Similarly, for every k and l there exists $\tilde{\chi}_{k,l}$ such that $\chi_{k,l}(x)=\int_{-\infty}^{\infty}\tilde{\chi}_{k,l}(x_0)d\tilde{G}(x_0)$ and dG can be written as

$$dG(x, y)=d\tilde{G}(x)d\tilde{G}(y)\left[1+\sum_{i=1}^{2}\sum_{k=1}^{\infty} a_k^{(i)}\tilde{\Psi}_k^{(i)}(x)\tilde{\Psi}_k^{(i)}(y)+\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} a_k^{(1)}a_l^{(2)}\tilde{\chi}_{k,l}(x)\tilde{\chi}_{k,l}(y)\right], \quad \text{a.e.} \tag{3.1}$$

Representation (3.1) is the same as (2.2) but to complete the proof we have to verify the orthogonality conditions (2.3). Indeed, for $i=1,2$, $k\geq 1$, $\int_{-\infty}^{\infty}\tilde{\Psi}_k^{(i)}(x)d\tilde{G}(x)=\lim_{x\to\infty}\Psi_k^{(i)}(x)=\left(\int_{-\infty}^{\infty} d\tilde{F}_{3-i}(x_{3-i})\right)\left(\int_{-\infty}^{\infty}\varphi_k^{(i)}(x_i)d\tilde{F}_1(x_i)\right)=$

$1\cdot 0=0$. Similarly $\int_{-\infty}^{\infty}\tilde{\chi}_{k,l}(x)d\tilde{G}(x)=0,\ k\geqq 1,\ l\geqq 1.$

Assume now that (X_i, Y_i) are PDD, $i=1,2$ and let I_i denote the support of X_i, $i=1,2$. Then for $i=1,2$, the distribution F_i of (X_i, Y_i) is p.d. on $I_i\times I_i$. Hence there exist two stochastic processes $\{Z_i(t),\ t\in I_i\}$, $i=1,2$, that we can assume to be independent, such that $F_i(x,y)$ $\mathrm{E}\,Z_i(x)Z_i(y)$, $(x,y)\in I_i\times I_i$. Let $h(x,y)$ be a real measurable function defined on $I_1\times I_2$ and denote $B(x)=\{(x_1,x_2)\colon h(x_1,x_2)\leqq x\}$. Then the distribution G of $(h(X_1,X_2), h(Y_1,Y_2))$ is

$$G(x,y)=\mathrm{E}\int_{I_2}\int_{I_2}\int_{I_1}\int_{I_1} I_{B(x)}(x_1,x_2)I_{B(y)}(y_1,y_2)dZ_1(x_1)dZ_1(y_1)dZ_2(x_2)dZ_2(y_2)\ ,$$

where I_A is the indicator function of the set A (for validity of this result, see Loeve [20], p. 472). Hence, for any signed measure H, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x,y)dH(x)dH(y)=\mathrm{E}\left[\int_{I_1}\int_{I_2}\int_{-\infty}^{\infty} I_{A(x)}(x_1,x_2)dH(x)dZ_1(x_1)dZ_2(x_2)\right]^2\geqq 0.$ That is, G is p.d. on $R\times R$.

COROLLARY 3.1. *The families of PDM, PDE and PDD distributions are closed under convolutions.*

Next we have the following closure theorem:

THEOREM 3.2. *A mixture of PDM[PDD] distributions is PDM [PDD].*

PROOF. Write the distribution of (X_1, X_2) as

$$F(x_1,x_2)=\int_S F_s(x_1,x_2)dG(s) \tag{3.2}$$

where for every $s\in S\subset R$, $F_s(x_1,x_2)$ is an interchangeable bivariate distribution function, Borel measurable in the variable s, and G is a probability measure defined on subsets of S.

For PDM distributions the assertion of the theorem is trivial. If $F_s(x_1,x_2)$ is PDD for every $s\in S$ then, by noting that a mixture of p.d. kernels is p.d., one sees that (X_1, X_2) is PDD.

Note that a mixture of PDE distributions is not necessarily PDE. The bivariate symmetric normal distribution with zero means and unit correlation coefficient is clearly a mixture of PDE distributions, but it is not a PDE distribution (see Example 2.1).

The next theorem deals with closure under passage to the limit (in distribution). The part of the theorem that deals with PDM distributions is proved in Shaked [29]. The proof of the rest uses standard techniques and will be omitted (for details see Shaked [26]).

THEOREM 3.3. *Let* $\{F_n\}_{n=1}^{\infty}$ *be a sequence of bivariate interchangeable distributions with* F *being its limit (in distribution) when* $n\to\infty$. *If, for every* n, F_n *is PDD[PDM] then* F *is PDD[PDM].*

Note that the class of PDE distributions is not closed under passages to the limit. Bivariate symmetric normal random vector with correlation coefficient $(n-1)/n$ is PDE but the limiting distribution of such random vectors is not PDE (see Example 2.1). However, by combining Implication (a) of Fig. 2.1 with Theorem 3.3 we obtain the following corollary which shows the positive dependence of the limit of a sequence of PDE distributions.

COROLLARY 3.2. *The limit in distribution of a sequence of PDE distributions is PDD.*

The closure properties that were discussed are summarized in the following table:

Table 3.1. Closure properties

Family of distributions	Mixtures	Closed under: Transformations in the sense of Theorem 3.1	Closed under: Passages to the limit (in distribution)
PDM	Yes	Yes	Yes
PDE	No	Yes	No
PDD	Yes	Yes	Yes

4. Some inequalities

In many applications of probability two random variables X_1 and X_2 are assumed to be independent even when they are not so. This is done when the joint distribution of X_1 and X_2 is unknown or when it is difficult to deal with it analytically. Theorems 4.1 and 4.1′ discuss the bias that is caused in some cases when independence is assumed while in fact (X_1, X_2) is PDM or PDE. The following definition is needed: Let I be a subset of R. A symmetric kernel K defined on $I\times I$ (i.e. $K(x, y)=K(y, x)$ for all $x, y\in I$) is said to be *conditionally positive definite* (c.p.d.) on $I\times I$ if for any positive integer n and for every choice of $x_1,\cdots, x_n$ in I and real numbers $a_1,\cdots, a_n$ it holds that

$$(4.1)\qquad \sum_{i=1}^{n}\sum_{j=1}^{n} K(x_i, x_j)a_ia_j\geqq 0 \qquad \text{whenever } \sum_{i=1}^{n} a_i=0\,,$$

(such kernels were discussed by Parthasarathy and Schmidt [22], Horn [10], [11], [12] and Johansen [15].) It can be shown (Shaked [26]), using a result of Parthasarathy and Schmidt [22], p. 3 and a standard approximation technique that K is c.p.d. on $I\times I$ if and only if

$$\int_I \int_I K(x, y)dH(x)dH(y) \geqq 0 \qquad \text{whenever } \int_I dH(x)=0 . \tag{4.2}$$

THEOREM 4.1. *Let* (X_1, X_2) *be a PDM vector and let* Y_1 *and* Y_2 *be i.i.d. random variables which are distributed as* X_1. *Then for every kernel* K *which is c.p.d. on* $S\times S$, *where* S *is the support of* X_i, *the inequality*

$$\mathrm{E}\, K(X_1, X_2) \geqq \mathrm{E}\, K(Y_1, Y_2) \tag{4.3}$$

holds whenever the expectations exist.

PROOF. Representation (2.1), the fact that for every $\omega \in \Omega$, $\int_S dA_\omega(x) =0$ [where $dA_\omega(x)=dF^{(\omega)}(x)-\int_\Omega dF^{(\omega')}(x)dG(\omega')$, see (2.1)], the assumption that K is c.p.d. on $S\times S$ and inequality (4.2) imply $\mathrm{E}\, K(X_1, X_2)-\mathrm{E}\, K(Y_1, Y_2)\geqq 0$.

THEOREM 4.1′. *Let* (X_1, X_2) *be a PDE vector and let* Y_1 *and* Y_2 *be i.i.d. random variables which are distributed as* X_1. *Then for every kernel* K *which is c.p.d. on* $S\times S$, *where* S *is the support of* X_1, *the inequality* (4.3) *holds whenever the expectations exist.*

PROOF. Representation (2.2), the nonnegativity of a_i of (2.2), the orthogonality relations (2.3) and the assumption that K is c.p.d. on $S\times S$ imply $\mathrm{E}\, K(X_1, X_2)-\mathrm{E}\, K(Y_1, Y_2)\geqq 0$.

The family of the c.p.d. kernels is the most general one for validity of (4.3) for all PDM (or PDE) vectors as the following theorems show.

THEOREM 4.2. *Let* S *be a Borel set and let* $g(x_1, x_2)$ *be a symmetric kernel, defined on* $S\times S$, *which is not c.p.d. on* $S\times S$. *Then there exist a PDM vector* (X_1, X_2) *and i.i.d. random variables* Y_1 *and* Y_2 *such that* Y_1 *is distributed as* X_1, *the support of* Y_1 *is contained in* S *and*

$$\mathrm{E}\, g(X_1, X_2) < \mathrm{E}\, g(Y_1, Y_2) . \tag{4.4}$$

PROOF. Let $B(\cdot)\not\equiv 0$ be a function that satisfies $\int_S dB(x)=0$ and $\int_S \int_S g(x_1, x_2)dB(x_1)dB(x_2)<0$. Denote $\alpha=\int_A dB(x)=-\int_{S-A} dB(x)$ where $A=\{x\colon dB(x)>0\}$. Define $dF^{(1)}(x)=\begin{cases}\alpha^{-1}dB(x) & \text{if } x\in A\\ 0 & \text{if } x\in S-A\end{cases}$ and $dF^{(2)}(x) =\begin{cases}0 & \text{if } x\in A\\ -\alpha^{-1}dB(x) & \text{if } x\in S-A\end{cases}$. Clearly $dF^{(1)}$ and $dF^{(2)}$ determine probability measures. Let (X_1, X_2) be a PDM random vector with joint distribution

$$F(x_1, x_2)=\frac{1}{2}F^{(1)}(x_1)F^{(1)}(x_2)+\frac{1}{2}F^{(2)}(x_1)F^{(2)}(x_2)\ , \qquad (x_1, x_2)\in S\times S\ , \tag{4.5}$$

and let Y_1 and Y_2 be i.i.d. with common distribution $\tilde{F}$. Then

$$\begin{aligned}
&\mathrm{E}\, g(X_1, X_2)-\mathrm{E}\, g(Y_1, Y_2)\\
&=\int_S\int_S g(x_1, x_2)\Big[\frac{1}{2}dF^{(1)}(x_1)dF^{(1)}(x_2)+\frac{1}{2}dF^{(2)}(x_1)dF^{(2)}(x_2)\\
&\qquad -\Big(\frac{1}{2}dF^{(1)}(x_1)+\frac{1}{2}dF^{(2)}(x_1)\Big)\Big(\frac{1}{2}dF^{(1)}(x_2)+\frac{1}{2}dF^{(2)}(x_2)\Big)\Big]\\
&=\frac{1}{4}\int_S\int_S g(x_1, x_2)[dF^{(1)}(x_1)dF^{(1)}(x_2)+dF^{(2)}(x_1)dF^{(2)}(x_2)\\
&\qquad -dF^{(1)}(x_1)dF^{(2)}(x_2)-dF^{(1)}(x_2)dF^{(2)}(x_1)]\\
&=\frac{\alpha^{-2}}{4}\int_S\int_S g(x_1, x_2)dB(x_1)dB(x_2)<0\ .
\end{aligned}$$

THEOREM 4.2′. *Let S be a Borel set and let $g(x_1, x_2)$ be a symmetric kernel, defined on $S\times S$ which is not c.p.d. on $S\times S$. Then there exists a PDE random vector (X_1, X_2) and i.i.d. random variables Y_1 and Y_2 such that Y_1 is distributed as X_1, the support of Y_1 is contained in S and* (4.4) *holds.*

PROOF. Let $B(\cdot)$ be as in the proof of Theorem 4.2. Let (X_1, X_2) and Y_1 and Y_2 be distributed as in the proof of Theorem 4.2. Clearly (4.4) holds. To complete the proof we need to show that F of (4.5) is PDE. Note that $dF(x_1, x_2)=d\tilde{F}(x_1)d\tilde{F}(x_2)[1+\varphi(x_1)\varphi(x_2)]$, where $\varphi(x)=[d\tilde{F}(x)]^{-1}\cdot(1/2)(dF^{(1)}(x)-dF^{(2)}(x))$. Clearly $\int_S\varphi(x)d\tilde{F}(x)=0$, hence F is PDE.

In the following examples we apply Theorems 4.1 and 4.1′ to specific kernels $K(\cdot,\cdot)$. We assume that (X_1, X_2) is a PDM or PDE vector, Y_1 and Y_2 are i.i.d. and that Y_1 is distributed as X_1. We assume also that the expectations that we write exist. Applications are discussed in Section 5.

Example 4.1. Assume that the support of X_1 is $[0, \infty)$.

(i) For $\alpha\leqq 0$ or $1\leqq\alpha\leqq 2$, $\mathrm{E}\,(X_1+X_2)^\alpha\geqq\mathrm{E}\,(Y_1+Y_2)^\alpha$.

(ii) For $0\leqq\alpha\leqq 1$, $\mathrm{E}\,(X_1+X_2)^\alpha\leqq\mathrm{E}\,(Y_1+Y_2)^\alpha$.

To prove (i) and (ii) recall that if $w(x)$ is a Laplace transform of a non-negative measure and it converges for $a<x<b$ then $w(x+y)$ is p.d. on $(a/2, b/2)\times(a/2, b/2)$ (Widder [35], p. 273). For $\alpha\leqq 0$ and $x>0$, $x^\alpha=$

$\int_0^\infty e^{-xt}t^{-\alpha-1}/\Gamma(-\alpha)dt$. Hence $(x+y)^\alpha$ is p.d. on $(0,\infty)\times(0,\infty)$. For $1\leqq\alpha\leqq 2$ the kernel $(x+y)^{\alpha-2}$ is p.d. on $(0,\infty)\times(0,\infty)$ as was shown above, hence $\int_0^x\int_0^y (x'+y')^{\alpha-2}dx'dy'=(1/\alpha(\alpha-1))[(x+y)^\alpha-(x^\alpha+y^\alpha)]$ is p.d. on $(0,\infty)\times(0,\infty)$. Using this fact, it is easily seen that $(x+y)^\alpha$ is c.p.d. on $(0,\infty)\times(0,\infty)$ and the proof of (i) is complete. For $0\leqq\alpha\leqq 1$ and for every $c>0$, $\exp(-cx^\alpha)$ is a Laplace transform of a probability measure on $(0,\infty)$ (Feller [7], p. 448), hence by the theorem of Parthasarathy and Schmidt [22], p. 3, $-(x+y)^\alpha$ is c.p.d. on $(0,\infty)\times(0,\infty)$. This proves (ii).

Note that for $\alpha>2$ the kernel $(x+y)^\alpha$ is not c.p.d. on $(0,\infty)\times(0,\infty)$ because if it were c.p.d. then $(\partial^2/\partial x\partial y)(x+y)^\alpha=\alpha(\alpha-1)(x+y)^{\alpha-2}$ is p.d. on $(0,\infty)\times(0,\infty)$, (Horn [11]). But this would imply that $\begin{vmatrix}(1+1)^\beta & (1+2)^\beta\\ (1+2)^\beta & (2+2)^\beta\end{vmatrix}\geqq 0$ for $\beta=\alpha-2>0$ which is false.

Possible applications in reliability theory of the inequalities of the previous and the next examples are discussed in Section 5.

Example 4.2.

(i) For $-1\leqq\alpha\leqq 0$ $\mathrm{E}\,|X_1-X_2|^\alpha\geqq\mathrm{E}\,|Y_1-Y_2|^\alpha$.

(ii) For $0\leqq\alpha\leqq 2$ $\mathrm{E}\,|X_1-X_2|^\alpha\leqq\mathrm{E}\,|Y_1-Y_2|^\alpha$.

To prove (i) and (ii) recall that if Ψ is a real characteristic function then $\Psi(x-y)$ is p.d. on $R\times R$ (Bochner's theorem). If Ψ is a real characteristic function of an infinitely divisible distribution then $\log\Psi(x-y)$ is c.p.d. on $R\times R$ (see Johansen [15]). For $0<\alpha\leqq 2$, $\exp\{-|u|^\alpha\}$ is a characteristic function of an infinitely divisible distribution hence $-|x-y|^\alpha$, $0<\alpha\leqq 2$, is c.p.d. on $R\times R$. This proves (ii). If Ψ is a real characteristic function, then for every $0\leqq\beta<1$, $(1-\beta\Psi(u))^{-1}$ is a characteristic function and $(1-\beta\Psi(x-y))^{-1}$ is p.d. on $R\times R$. Let $a>0$ and $0<p<1$ then (by using e.g., Polya Criterion, Feller [7], p. 509) one can verify that $\Psi(u)=1-a^{-1}|u|^p$ if $|u|\leqq a$, 0 otherwise, is a characteristic function and hence $\rho(x,y)=[1-\beta(1-a^{-1}|x-y|^p)]^{-1}$ if $|x-y|\leqq a$, 1 otherwise, is p.d. on $R\times R$. Hence for $0\leqq\beta<1$, $-1<\alpha<0$ and $a>0$,

$$\begin{aligned}&\mathrm{E}\,[1-\beta(1-a^{-1}|X_1-X_2|^{-\alpha})^{-1}]I_{[-a,a]}(|X_1-X_2|)+\mathrm{P}\,(|X_1-X_2|>a)\\ &\quad\geqq\mathrm{E}\,[1-\beta(1-a^{-1}|Y_1-Y_2|^{-\alpha})^{-1}]I_{[-a,a]}(|Y_1-Y_2|)+\mathrm{P}\,(|Y_1-Y_2|>a)\,,\end{aligned}$$

where I_A is the indicator function of the set A. Dividing both sides by a and letting $\beta\to 1$ we obtain

$$\begin{aligned}&\mathrm{E}\,|X_1-X_2|^\alpha I_{[-a,a]}(|X_1-X_2|)+a^{-1}\,\mathrm{P}\,(|X_1-X_2|>a)\\ &\quad\geqq\mathrm{E}\,|Y_1-Y_2|^\alpha I_{[-a,a]}(|Y_1-Y_2|)+a^{-1}\,\mathrm{P}\,(|Y_1-Y_2|>a)\,.\end{aligned}$$

Now by letting $a\to\infty$, (i) is proved for $-1<\alpha\leqq 0$. That (i) holds also for $\alpha=-1$ can be seen by letting $\alpha\to-1$.

Example 4.3.
(i) $\mathrm{E}\min(X_1, X_2)\geqq\mathrm{E}\min(Y_1, Y_2)$
(ii) $\mathrm{E}\max(X_1, X_2)\leqq\mathrm{E}\max(Y_1, Y_2)$.

The kernel $K(x, y)=\min(x, y)$ is a covariance function of a stochastic process (Prabhu [23], p. 37), hence it is p.d. on $R\times R$. This proves (i). The second inequality follows from the first and the fact that $\mathrm{E}(X_1+X_2)=\mathrm{E}(Y_1+Y_2)$. We note that a more general result for PDM (but not PDE) random vectors is obtained in Shaked [29] by a different method.

Example 4.4. Assume $\mathrm{P}(0\leqq X_1\leqq 1)=1$.
(i) $\mathrm{E}\min(X_1, X_2)-\mathrm{Cov}(X_1, X_2)\geqq\mathrm{E}\min(Y_1, Y_2)$
(ii) $\mathrm{E}\max(X_1, X_2)+\mathrm{Cov}(X_1, X_2)\leqq\mathrm{E}\max(Y_1, Y_2)$.

The inequality (i) follows by simple arithmetics from the fact that $\min(x, y)-xy$ is p.d. on $[0, 1]\times[0, 1]$ (Sukhatme [32], p. 1921). The second inequality follows from the first.

Example 4.5. Let $f(x_1, x_2)$ be a density of a PDM or PDE random vector and denote its marginal by $\tilde{f}$. Then

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^k(x_1, x_2)dx_1dx_2\geqq\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{f}^k(x_1)\tilde{f}^k(x_2)dx_1dx_2\,, \quad k=2, 3,\cdots . \tag{4.6}$$

To see it, note that by Theorems 4.1 and 4.1′, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^k(x_1, x_2)dx_1dx_2$ $\geqq\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^{k-1}(x_1, x_2)\tilde{f}(x_1)\tilde{f}(x_2)dx_1dx_2\geqq\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^{k-2}(x_1, x_2)\tilde{f}^2(x_1)\tilde{f}^2(x_2)dx_1dx_2\geqq$ $\cdots\geqq\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{f}^k(x_1)\tilde{f}^k(x_2)dx_1dx_2$. The meaning of (4.6) in statistical mechanics is discussed in Section 5.

Example 4.6. Let $f(x_1, x_2)$ be the density of a PDM or PDE random vector (x_1, x_2). Let $\tilde{f}$ be the marginal density then, as can be verified from Theorems 4.1 and 4.1′, $\mathrm{E}(f(X_1, X_2)/\tilde{f}(X_1)\tilde{f}(X_2))\geqq 1$, that is the expected likelihood ratio under the hypothesis of PDM or PDE (versus the hypothesis of independence) is not smaller than 1. Statistical applications of this result are not yet known.

5. Applications

Reliability theory

Assume two identical components with lifelengths X_1 and X_2 operate in a random environment (this is the case, e.g. when the user or climate in place of operation cannot be predicted at the time of production). Then their joint distribution is PDM (Shaked [29]). Assume also that death of one component is undetected until the second component dies, too. The loss incurred during the operation, t, of only one component can be assumed to be proportional to t^α, $0 \leqq \alpha$. Then the expected loss is proportional to $\mathrm{E}\,|X_1 - X_2|^\alpha$. From Example 4.2 (ii) it follows that if $\alpha \leqq 2$ then the expected loss computed under the assumption of independence overestimates the actual loss. A different situation arises if only one component functions and upon failure of the first the second standby component starts functioning in the same environment. The income (gain) in this case can be proportional to t^α where t is the total time of operation. Then the expected income is proportional to $\mathrm{E}\,(X_1+X_2)^\alpha$ and from Example 4.1 we see that when $0 \leqq \alpha \leqq 1$ it is overestimated and when $1 \leqq \alpha \leqq 2$ it is underestimated. We also see that for $\alpha > 2$ it cannot be determined without additional assumptions whether the income is over- (or under-) estimated.

Mechanical statistics

Rao [24], p. 142, introduced, for $0 < \alpha$, $\alpha \neq 1$, the quantity $(1-\alpha)^{-1} \cdot \log \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^\alpha(x_1, x_2)dx_1dx_2$ as a measure of closeness of the density f to a uniform distribution (large values indicating a higher degree of closeness). Intuitively it is clear that assumptions of independence will cause this measure to decrease. It is shown in Example 4.5 that for integer $\alpha \geqq 2$ this is, indeed, the case.

Reversible Markov processes

A (strictly) stationary Markov process $\{Z(t),\ -\infty < t < \infty\}$ is said to be *reversible in time* if

$$\mathrm{P}\,(Z(t-h) \in A \,|\, Z(t)=z) = \mathrm{P}\,(Z(t+h) \in A \,|\, Z(t)=z) \tag{5.1}$$

for all t, $h > 0$, z and Borel sets A (see Keilson [17] and refernces there). Define $X_i = Z(t_i)$, $i=1, 2$, for some $t_1 < t_2$. Sarmanov [25] showed that if (X_1, X_2) has a density then (X_1, X_2) is PDE. Eagleson [3] proved that if $Z(t)$ is a chain then (X_1, X_2) is PDE provided (X_1, X_2) is φ^2-bounded (for definition see Lancaster [18]). It can be shown that, without any restrictions concerning the range of $Z(t)$ or its underlying distribution, (X_1, X_2) is PDM. This follows from the independence of X_1 and X_2 given $Z((t_1+t_2)/2)$ and from (5.1). Thus, by Example 4.3 we see that $\mathrm{E} \max (X_1, X_2)$ is over-estimated if independence is assumed. Estimation of the expected value of $|X_1 - X_2|^\alpha$ may be needed in some applications, the inequalities of Example 4.2 then may be used.

Positive dependence of some well known distributions

We end the paper by applying the results of Sections 2 and 3 to show some examples of positively dependent interchangeable distributions. More examples can be found in Shaked [29].

Example 5.1 (Bivariate geometric, Esary and Marshall [4]). Let U_i have the probability function $\mathrm{P}(U_i=k)=(1-\theta)\theta^{k-1}$; $k=1,2,\cdots$; $i=1,2$; $0\leqq\theta<1$, and let W have the probability function $\mathrm{P}(W=k)=(1-\theta_0)\theta_0^{k-1}$; $k=1,2,\cdots$; $0\leqq\theta_0<1$, and assume that U_1, U_2 and W are independent. Then $X_i=\min(U_i, W)$, $i=1,2$ have a bivariate geometric distribution in the narrow sense (BVG-N) and (X_1, X_2) is PDM by Proposition 2.1. The joint distribution of (X_1, X_2) is determined by

$$\mathrm{P}(X_1>k_1,\ X_2>k_2)=\theta^{k_1+k_2}\theta_0^{\max(k_1,k_2)}\,;\qquad k_1, k_2=0,1,2,\cdots.$$

Esary and Marshall define also a wider family of bivariate distributions with geometric marginals (BVG-W). The joint distribution of (Y_1, Y_2) which has an exchangeable BVG-W distribution is determined by

$$\mathrm{P}(Y_1>k_1,\ Y_2>k_2)=p^{k_1+k_2}p_0^{\max(k_1,k_2)}\,;\qquad k_1, k_2=0,1,2,\cdots, \tag{5.2}$$

where p and p_0 satisfy, by definition, $0\leqq p<1$, $pp_0<1$ and $0\leqq p_0(2p-p^2)\leqq 1$. Esary and Marshall show that there exist exchangeable BVG-W distributions which are not BVG-N. We will show now that the only PDM BVG-W distributions are BVG-N. Assume that (Y_1, Y_2) with probabilities determined by (5.2) is PDM. We have to show that $p_0\leqq 1$. By implications (e) and (f) of Fig. 2.1, $\mathrm{Cov}(Y_1, Y_2)\geqq 0$. Some computation shows that $\mathrm{Cov}(Y_1, Y_2)=p^2p_0(1-p_0)/((1-pp_0)^2(1-p^2p_0))$. Hence $p_0\leqq 1$.

Example 5.2 (Bivariate binomial). Let (X_i, Y_i), $i=1,2,\cdots,n$ be i.i.d. random vectors with the following joint probabilities:

X_i \ Y_i	0	1
1	$P_{1\cdot}-P_{11}$	P_{11}
0	$1-2P_{1\cdot}+P_{11}$	$P_{1\cdot}-P_{11}$

Then $(X, Y)\equiv\left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right)$ is said to have an exchangeable bivariate binomial distribution. Aitken and Gonin [1] showed that if $P_{11}\geqq P_{1\cdot}^2$, then (X, Y) is PDE. Clearly under this condition (X_i, Y_i) is PDM, $i=1,\cdots,n$. Application of Theorem 3.1 $n-1$ times shows that (X, Y) is PDM.

Example 5.3 (Bivariate distributions with desired marginals, Method I (Shaked [28])). Let $\Psi(u)$ be a probability generating function of a non-negative integer-valued random variable, and let $F(x)$ be a univariate distribution. Then,

$$G(x_1, x_2) = \Psi(F(x_1)F(x_2)) \tag{5.3}$$

is an exchangeable bivariate distribution with marginals $\Psi(F(x))$. Similarly $H(x_1, x_2) = 1 - \bar{H}(x_1, -\infty) - \bar{H}(-\infty, x_2) + \bar{H}(x_1, x_2)$ is an exchangeable bivariate distribution where

$$\bar{H}(x_1, x_2) = \Psi((1-F(x_1))(1-F(x_2))) . \tag{5.4}$$

The marginals of H are $1-\Psi(1-F(x))$. The distributions G and H of (5.3) and (5.4) remain well defined if $\Psi(u)$ is of the form

$$\Psi(u) = \int_0^\infty u^x d\varphi(x) \tag{5.5}$$

where φ is a probability measure on $[0, \infty)$. Let $\tilde{F}$ be a given univariate distribution and assume that it can be expressed as $\tilde{F}(x) = \Psi(F(x))$ [or as $\tilde{F}(x) = 1 - \Psi(1-F(x))$] for some nontrivial distribution F, and Ψ of the form (5.5). Then (5.3) [(5.4)] define a bivariate distribution with $\tilde{F}$ as its marginal.

It is easy to verify that G and H of (5.3) and (5.4) are PDM.

Note that the bivariate Burr's distribution (Takahasi [33]), the bivariate logistic distribution (Malik and Abraham [21]) and the bivariate extreme value distribution (Johnson and Kotz [16], p. 254) are special cases of (5.3) and (5.4) (for verifications see Shaked [28]).

Example 5.4 (Bivariate distributions with desired marginals, Method II (Shaked [28])). Let $\Psi(u_1, u_2)$ be a bivariate probability generating function of a non-negative random vector, or more generally, let

$$\Psi(u_1, u_2) = \int_0^\infty \int_0^\infty u_1^x u_2^y d\varphi(x, y) \tag{5.6}$$

where φ is a probability measure on $[0, \infty) \times [0, \infty)$. If $F(x)$ is a univariate distribution, then,

$$G(x_1, x_2) = \Psi(F(x_1), F(x_2)) \tag{5.7}$$

is a distribution function. Similarly $H(x_1, x_2) = 1 - \bar{H}(x_1, \infty) - \bar{H}(-\infty, x_2) + \bar{H}(x_1, x_2)$ is a distribution function where

$$\bar{H}(x_1, x_2) = \Psi(1-F(x_1), 1-F(x_2)) . \tag{5.8}$$

Note that (5.7) and (5.8) are generalizations of (5.3) and (5.4).

The distributions G and H of (5.7) and (5.8) are exchangeable when $\varphi(x, y)$ of (5.6) is exchangeable, and they are PDM if $\varphi(x, y)$ is PDM. In a way, similar to the proof of Theorem 3.1, one can verify that if φ is PDE [PDD] then G and H are PDE [PDD].

Acknowledgement

To my advisor, A. W. Marshall, I wish to express my gratefulness for his guidance and advice while I prepared most of this work. I would like also to thank the referee for helpful comments.

UNIVERSITY OF NEW MEXICO*

REFERENCES

[1] Aitken, A. C. and Gonin, H. T. (1935). On fourfold sampling with and without replacement, *Proc. Roy. Soc. Edinburgh*, **55**, 114-125.

[2] Dykstra, R. L., Hewett, J. E. and Thompson, W. A. (1973). Events which are almost independent, *Ann. Statist.*, **1**, 674-681.

[3] Eagleson, G. K. (1969). Canonical expansions of birth and death processes, *Theory Prob. Appl.*, **14**, 209-218.

[4] Esary, J. D. and Marshall, A. W. (1974). Multivariate geometric distributions generated by commutative damage process, Report NPS55EY73041A, Naval Postgraduate School, Monterey, California.

[5] Esary, J. D. and Proschan, F. (1972). Relationships among some concepts of bivariate dependence, *Ann. Math. Statist.*, **43**, 651-655.

[6] Esary, J. D., Proschan, F. and Walkup, D. W. (1967). Association of random variables with applications, *Ann. Math. Statist.*, **38**, 1466-1474.

[7] Feller, W. (1971). *An Introduction to Probability and its Applications*, 2 (2nd ed.), Wiley and Sons, New York.

[8] Hall, M. and Newman, M. (1963). Co-positive and completely positive quadratic forms, *Proc. Camb. Phil. Soc.*, **59**, 329-339.

[9] Hewett, E. H. and Bulgren, W. G. (1971). Inequalities for some multivariate F distributions with applications, *Technometrics*, **13**, 397-402.

[10] Horn, R. A. (1967). On infinitely divisible matrices, kernels and functions, *Zeit. Wahrscheinlichkeitsth.*, 8, 219-230.

[11] Horn, R. A. (1969). The theory of infinitely divisible matrices and kernels, *Trans. Amer. Math. Soc.*, **136**, 269-286.

[12] Horn, R. A. (1971). Schlicht mappings and infinitely divisible kernels, *Pacific J. Math.*, **38**, 423-430.

[13] Jensen, D. R. (1969). An inequality for a class of bivariate chi-square distributions, *J. Amer. Statist. Ass.*, **64**, 333-336.

[14] Jensen, D. R. (1971). A note of positive dependence and the structure of bivariate distributions, *SIAM J. Appl. Math.*, **20**, 749-753.

[15] Johansen, S. (1966). An application of extreme point methods to the representation of infinitely divisible distributions, *Zeit. Wahrscheinlichkeitsth.*, **5**, 304-316.

[16] Johnson, N. L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, Wiley and Sons, New York.

[17] Keilson, J. (1965). A review of transient behavior in regular diffusion and birth-death processes, II, *J. Appl. Prob.*, **2**, 405-428.

[18] Lancaster, H. O. (1958). The structure of bivariate distributions, *Ann. Math. Statist.*, **29**, 719-736.

[19] Lehmann, E. L. (1966). Some concepts of dependence, *Ann. Math. Statist.*, **37**, 1137-1153.

[20] Loeve, M. (1963). *Probability Theory*, 3rd ed., Van Nostrand, Princeton, New Jersey.

* Now at Indiana University.

[21] Malik, H. J. and Abraham, B. (1973). Multivariate logistic distributions, *Ann. Statist.*, **1**, 588-590.

[22] Parthasarathy, K. R. and Schmidt, K. (1972). *Positive Definite Kernels, Continuous Tensor Products, and Central Limit Theorems of Probability Theory*, Springer-Verlag, Berlin.

[23] Prabhu, N. U. (1965). *Stochastic Processes*, Macmillan, New York.

[24] Rao, C. R. (1965). *Linear Statistical Inference and its Applications*, Wiley and Sons, New York.

[25] Sarmanov, O. V. (1963). *Investigation of Stationary Markov Processes by the Method of Eigenfunction Expansion: Select. Translations Math. Statist. and Prob.*, **4**, 245-269.

[26] Shaked, M. (1975a). On concepts of dependence, Doctoral dissertation, University of Rochester.

[27] Shaked, M. (1975b). A note on the exchangeable generalized Farlie-Gumbel-Morgenstern distributions, *Commun. Statist.*, **4**, 711-721.

[28] Shaked, M. (1975c). *On the Distribution of the Minimum and of the Maximum of a Random Number of i.i.d. Random Variables: Statistical Distributions in Scientific Work*, Vol. I. eds. G. P. Patil, S. Kotz and J. K. Ord, Reidel, Boston, 363-380.

[29] Shaked, M. (1977a). A concept of positive dependence for exchangeable random variables, *Ann. Statist.*, **5**, 505-515.

[30] Shaked, M. (1977b). A family of concepts of dependence for bivariate distributions, *J. Amer. Statist. Ass.*, **72**, 642-650.

[31] Sidak, Z. (1973). A chain of inequalities for some types of multivariate distributions, with nine special cases, *Aplik. Mat.*, **18**, 110-118.

[32] Sukhatme, S. (1972). Fredholm determinant of a positive definite kernel of a special type and its application, *Ann. Math. Statist.*, 1914-1926.

[33] Takahasi, K. (1965). Note on the multivariate Burr's distribution, *Ann. Inst. Statist. Math.*, **17**, 257-260.

[34] Tong, Y. L. (1970). Some probability inequalities of multivariate normal and multivariate *t*, *J. Amer. Statist. Ass.*, **65**, 1243-1247.

[35] Widder, D. V. (1941). *The Laplace Transform*, Princeton University press, Princeton, New Jersey.

[36] Yanagimoto, T. (1972). Families of positively dependent random variables, *Ann. Inst. Statist. Math.*, **24**, 559-573.