FAMILIES OF POSITIVELY DEPENDENT RANDOM VARIABLES

TAKEMI YANAGIMOTO

(Received Oct. 12, 1971; revised May 18, 1972)

Summary

Possible definitions of positive dependence of random variables are systematically and exhaustively examined and previous results on this notion are improved. Three approaches to the definitions are proposed. They are similar to those given by the author and Sibuya in defining "stochastically larger component of a random vector." Unbiasedness of rank tests of independence is treated.

1. Introduction

A random variable (X, Y) is positively dependent if, roughly speaking, larger values of one component correspond stochastically to larger values of the other. Some definitions of this notion have been introduced and studied in [3], [7]–[10] and [12]. The purpose of this paper is to present systematic ways of defining of positive dependence of different degrees of strictness, to examine the hierarchical structure of these definitions, and thus to improve previous results on the relations among the definitions. As an example distributions of explicit form are classified by our definitions.

We define families of two-dimensional random variables (X, Y), or equivalently of distribution functions $F(x, y)$. A family of positively dependent variables will be denoted by $\mathcal{P}(\cdot)$ with one or two parameters inside the parentheses. Corresponding to each family $\mathcal{P}(\cdot)$, a family of negatively dependent variables denoted by $\mathcal{N}(\cdot)$ can be defined always dually to $\mathcal{P}(\cdot)$.

We present three approaches, that is we introduce three classes of $\mathcal{P}(\cdot)$'s. In the first one we consider four two-dimensional intervals, which are finite or infinite in one or two directions, and apply the definition of positive association in two-by-two tables. In the second approach we consider conditional distributions, if one component, under the condition that the other has a certain value, is stochastically larger than the former component, under the condition that the latter has a smaller value, then the random variable is defined to be positively dependent. In the last approach we consider are a termed transformation of a random vector in order to define positive dependence.

These approaches are closely related to each other, and one of our objectives is to study relations among them. The first approach gives rather stronger definitions and the second weaker ones. The last approach enables us to state a proposition on unbiasedness of certain tests. Our approaches are very similar to those in the author's previous joint paper on the notion of "stochastically larger" [13]. In the first and second approaches we use the notations $\mathcal{R}_3$, $\mathcal{R}_2$, $\mathcal{R}_2'$, $\mathcal{R}_2''$, $\mathcal{R}_1$, $\mathcal{R}_0$ and $\mathcal{R}_E$ defined in [13], and their properties.

In Sections 2, 3 and 4 the three classes of $\mathcal{P}(\cdot)$'s are defined and their properties are examined respectively. In Section 4 we also study properties of test statistics for independence. Counter examples showing the absence of certain implications are presented in Appendix.

2. Families $\mathcal{P}(i, j)$

In this section we define families of positively dependent distributions $\mathcal{P}(i, j)$, $i, j \in I$, where $I=\{1, 2', 2'', 2, 3\}$. These are related to the stochastic order relationships $\mathcal{R}_1$, $\mathcal{R}_2'$, $\mathcal{R}_2''$, $\mathcal{R}_2$ and $\mathcal{R}_3$, and constitute a hierarchical structure: for $i>i'$ and $j\geqq j'$ or $i\geqq i'$ and $j>j'$, the dependence in $\mathcal{P}(i, j)$ is more strict than in $\mathcal{P}(i', j')$.

Throughout the paper (X, Y) is a random variable with its distribution function $F(x, y)$. Its marginal distribution are denoted by $F_1(x)=F(x, \infty)$ and $F_2(y)=F(\infty, y)$, and its conditional distributions by $F(x|y)$ and $F(y|x)$. The probability function corresponding to $F(x, y)$ is denoted by $P_F(\cdot)$, or simply by $P(\cdot)$.

Firstly we define families $\mathcal{P}(3, 3)$, $\mathcal{P}(3, 2')$, $\mathcal{P}(3, 2'')$, $\mathcal{P}(3, 1)$ $\mathcal{P}(2'', 1)$ and $\mathcal{P}(1, 1)$.

DEFINITION 2.1. (i) (X, Y) (or its distribution function $F(x, y)$) is said to be in $\mathcal{P}(3, 3)$, if $P_F((a_1, a_2](b_1, b_2])\times P_F((a_2, a_3](b_2, b_3])\geqq P_F((a_1, a_2](b_2, b_3])\times P_F((a_2, a_3](b_1, b_2])$ for all $a_1<a_2<a_3$ and $b_1<b_2<b_3$.

(ii) (X, Y) is said to be in $\mathcal{P}(3, 2')$, if $P((a_1, a_2](b_1, b_2])\times P((a_2, a_3](b_2, \infty))\geqq P((a_1, a_2](b_2, \infty))\times P((a_2, a_3](b_1, b_2])$ for all $a_1<a_2<a_3$ and $b_1<b_2$.

(iii) (X, Y) is said to be in $\mathcal{P}(3, 2'')$, if $P((a_1, a_2](-\infty, b_1])\times P((a_2, a_3](b_1, b_2])\geqq P((a_1, a_2](b_1, b_2])\times P((a_2, a_3](-\infty, b_1])$ for all $a_1<a_2<a_3$ and $b_1<b_2$.

(iv) (X, Y) is said to be in $\mathcal{P}(3, 1)$, if $P((a_1, a_2](-\infty, b])\times P((a_2, a_3](b, \infty))\geqq P((a_1, a_2](b, \infty))\times P((a_2, a_3](-\infty, b])$ for all $a_1<a_2<a_3$ and b.

(v) (X, Y) is said to be in $\mathcal{P}(2'', 1)$, if $P((-\infty, a_1](-\infty, b])\times P((a_1, a_2](b, \infty))\geqq P((-\infty, a_1](b, \infty))\times P((a_1, a_2](-\infty, b])$ for all $a_1<a_2$ and b.

(vi) (X, Y) is said to be in $\mathcal{P}(1, 1)$, if $P((-\infty, a](-\infty, b])\times P((a, \infty)(b, \infty))\geqq P((-\infty, a](b, \infty))\times P((a, \infty)(-\infty, b])$ for all a and b.

All other families $\mathcal{P}(i, j)$ for $i, j\in\{1, 2', 2'', 3\}$ can be defined analogously. $\mathcal{P}(2, j)$, $\mathcal{P}(i, 2)$ and $\mathcal{P}(2, 2)$ for $i, j\in I$ are defined by $\mathcal{P}(2', j)\cap\mathcal{P}(2'', j)$, $\mathcal{P}(i, 2')\cap\mathcal{P}(i, 2'')$ and $\mathcal{P}(2', 2')\cap\mathcal{P}(2', 2'')\cap\mathcal{P}(2'', 2')\cap\mathcal{P}(2'', 2'')$, respectively. $\mathcal{N}(i, j)$, $i, j\in I$ is defined by inverting the inequalities in Definition 2.1.

We denote a component Y (or X) under the condition that $a_1<X\leqq a_2$ (or $b_1<Y\leqq b_2$) by $Y|_{a_1<X\leqq a_2}$ (or $X|_{b_1<Y\leqq b_2}$). The following Proposition 2.1 shows that $\mathcal{P}(i, j)$'s are also characterized by comparisons of conditioned variables. For simplicity we state propositions only on the $\mathcal{P}(i, j)$'s mensioned in Definition 2.1. Similar statements hold for other combinations i and j.

PROPOSITION 2.1. (i) $(X, Y)\in\mathcal{P}(3, 3)$ iff $Y|_{a_2<X\leqq a_3}\succ Y|_{a_1<X\leqq a_2}(\mathcal{R}_3)$ for all $a_1<a_2<a_3$, and also iff $X|_{b_2<Y\leqq b_3}\succ X|_{b_1<Y\leqq b_2}(\mathcal{R}_3)$ for all $b_1<b_2<b_3$.

(ii) $(X, Y)\in\mathcal{P}(3, 2')$ iff $Y|_{a_2<X\leqq a_3}\succ Y|_{a_1<X\leqq a_2}(\mathcal{R}_2')$ for all $a_1<a_2<a_3$, and iff $X|_{b_2<Y}\succ X|_{b_1<Y\leqq b_2}(\mathcal{R}_3)$ for all $b_1<b_2$.

(iii) $(X, Y)\in\mathcal{P}(3, 2'')$ iff $Y|_{a_2<X\leqq a_3}\succ Y|_{a_1<X\leqq a_2}(\mathcal{R}_2'')$ for all $a_1<a_2<a_3$; and iff $X|_{b_1<Y\leqq b_2}\succ X|_{Y\leqq b_1}(\mathcal{R}_3)$ for all $b_1<b_2$.

(iv) $(X, Y)\in\mathcal{P}(3, 1)$ iff $Y|_{a_2<X\leqq a_3}\succ Y|_{a_1<X\leqq a_2}(\mathcal{R}_1)$ for all $a_1<a_2<a_3$, and iff $X|_{Y>b}\succ X|_{Y\leqq b}(\mathcal{R}_3)$ for all b.

(v) $(X,Y)\in\mathcal{P}(2'', 1)$ iff $Y|_{a_1<X\leqq a_2}\succ Y|_{X\leqq a_1}(\mathcal{R}_1)$ for all $a_1<a_2$, and iff $X|_{Y>b}\succ X|_{Y\leqq b}(\mathcal{R}_2'')$ for all b.

(vi) $(X, Y)\in\mathcal{P}(1, 1)$ iff $Y|_{X>a}\succ Y|_{X\leqq a}(\mathcal{R}_1)$ for all a, and iff $X|_{Y>b}\succ X|_{Y\leqq b}(\mathcal{R}_1)$ for all b.

It is postulated in the above that the inequality $Y|_{X\in A}\succ Y|_{X\in A'}(\mathcal{R}_k)$ is satisfied if either probability of A or A' is zero. Thus the symbols $\mathcal{P}(i, j)$ correspond closely to the symbols $\mathcal{R}_k$. The following Proposition 2.2 shows the hierarchical structure of strictness of $\mathcal{P}(i, j)$'s. Fig. 1 illustrates just a part of it, since there are too many families to be shown. $\mathcal{P}(i, j)\to\mathcal{P}(i', j')$ is equivalent to $\mathcal{P}(i, j)\subset\mathcal{P}(i', j')$.

PROPOSITION 2.2. Define the partial ordering in $I=\{1, 2', 2'', 2, 3\}$ by $3>2>2'>1$ and $2>2''>1$, then $\mathcal{P}(i, j)\subset\mathcal{P}(i', j')$ iff $i\geqq i'$ and $j\geqq j'$.

PROOF. That $i\geqq i'$ and $j\geqq j'$ imply $\mathcal{P}(i, j)\subset\mathcal{P}(i', j')$ follows immediately from Definition 2.1. We show that $(X, Y)\in\mathcal{P}(i, j)$ does not imply $(X, Y)\in\mathcal{P}(i', j')$ if j is not larger than nor equal to j'. We see in [13] that there exist distribution functions $F(y)$ and $G(y)$, such that $F(y)\succ G(y)(\mathcal{R}_j)$ but not $F(y)\succ G(y)(\mathcal{R}_{j'})$. Let X have two positive mass at only two points, x_0 and x_1 $(x_0<x_1)$, and let $F(y)$ be the conditional distribution function of Y given $X=x_0$ and $G(y)$ the conditional

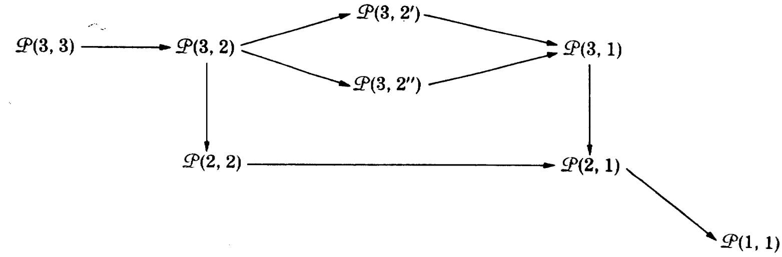

Fig. 1. Implication scheme (i).

distribution function of Y given $X=x_1$. Then Proposition 2.1 implies that $(X, Y) \in \mathcal{P}(i, j)$ but $(X, Y) \notin \mathcal{P}(i', j')$.

PROPOSITION 2.3. For any $i, j \in I=\{1, 2', 2'', 2, 3\}$ it holds that
(i) $(X, Y) \in \mathcal{P}(i, j)$ iff $(Y, X) \in \mathcal{P}(j, i)$,
(ii) $(X, Y) \in \mathcal{P}(i, j)$ iff $(-X, -Y) \in \mathcal{P}(\sigma(i), \sigma(j))$, where $\sigma(i)=i$ for $i \in I-\{2', 2''\}$. $\sigma(2')=2''$ and $\sigma(2'')=2'$,
(iii) $(X, Y) \in \mathcal{P}(i, j)$ iff $(X, -Y) \in \mathcal{N}(i, \sigma(j))$, and $(-X, Y) \in \mathcal{N}(\sigma(i), j)$,
(iv) for r and s be nondecreasing functions, $(X, Y) \in \mathcal{P}(i, j)$ implies $(r(X), s(Y)) \in \mathcal{P}(i, j)$,
(v) $(X, Y) \in \mathcal{P}(i, j)$ and $(X, Y) \in \mathcal{N}(i, j)$ imply (X, Y) is independent.

PROPOSITION 2.4. (i) When $F(x, y)$ has the probability density function $f(x, y)$, $F(x, y) \in \mathcal{P}(3, 3)$ iff $f(x, y)f(x', y') \geqq f(x, y')f(x', y)$ for almost all $x<x'$, $y<y'$. In this case $F(x, y)$ was called in [10] to be positively likelihood ratio dependent.

(ii) $F(x, y) \in \mathcal{P}(3, 1)$ iff $F(y|x) \geqq F(y|x')$ for all y and almost all $x<x'$. In this case $F(x, y)$ was called in [10], [11] to be positively regression dependent.

(iii) $F(x, y) \in \mathcal{P}(2'', 1)$ iff $F(x, y)/F_1(x)$ is nonincreasing in x for all y. In this case $F(x, y)$ was called in [4] to be left tail decreasing.

(iv) $F(x, y) \in \mathcal{P}(1, 1)$ iff $F(x, y) \geqq F_1(x)F_2(y)$ for all x and y. In this case $F(x, y)$ was called in [10] to be positively quadrant dependent.

Example 2.1. A wide family of distributions denoted by

$$H_\alpha(x, y)=F(x)G(y)\{1+\alpha A(F(x))B(G(y))\}$$

was given by Farlie [5].

For convenience suppose that $F(x)$ and $G(y)$ are continuous distribution functions, and that A and B are differentiable. A sufficient condition for $H_\alpha(x, y)$ to be a distribution function is that $B(1)=0$ and $|P(x)Q(y)|\cdot|\alpha|<1$ for all x, y, where $P(x)=d(xA(x))/dx$ and $Q(y)=d(yB(y))/dy$. It follows from Proposition 2.3 (iv) that classification of the dependence of $H_\alpha(x, y)$ does not depend on what are $F(x)$ and $G(y)$. We may restrict to the case that $F(x)$ and $G(y)$ are the distribution function of uniform distribution $U(0,1)$. Then the density function of $H_\alpha(x, y)$, $h_\alpha(x,y)$, and the conditional distribution function, $H_\alpha(y|x)$, are given by

$$h_\alpha(x, y)=1+\alpha P(x)Q(x)$$

and

$$H_\alpha(y|x)=y\{1+\alpha P(x)B(y)\} ,$$

respectively. After direct calculations we get that for $\alpha\geqq 0$ the following four statements hold.

(i) $H_\alpha(x, y)\in \mathcal{P}(1,1)$ iff $B(y)(A(x)-A(0))\geqq 0$ for all $1>x$, $y>0$.

(ii) $H_\alpha(x, y)\in \mathcal{P}(2'',1)$ iff $B(y)(A(x)-A(x'))\geqq 0$ for all $1>x>x'>0$ and $1>y>0$.

(iii) $H_\alpha(x, y)\in \mathcal{P}(3,1)$ iff $B(y)(P(x)-P(x'))\leqq 0$ for all $1>x>x'>0$ and $1>y>0$.

(iv) $H_\alpha(x, y)\in \mathcal{P}(3,3)$ iff $(P(x)-P(x'))(Q(y)-Q(y'))\geqq 0$ for all $1>x>x'>0$ and $1>y>y'>0$.

Example 2.2. Let $X_{(2)}\geqq X_{(1)}$ be order statistics of a sample of size 2 from a continuous distribution function $F(x)$. We know that a necessary and sufficient condition for $(X_{(1)}, X_{(2)}-X_{(1)})$ to be independent is that $F(x)$ is exponential. We can extend this fact to the case of negatively dependence. In fact the conditional distribution function of $(X_{(1)}, X_{(2)}-X_{(1)})$ given $X_{(1)}=x$, $F(y|x)$, is given by for any y and almost all x

$$1-F(y|x)=(1-F(y+x))/(1-F(x)) .$$

From this the following three conditions are shown to be equivalent.

(i) $-\log(1-F(x))$ is convex,

(ii) $(X_{(1)}, X_{(2)}-X_{(1)})\in \mathcal{N}(3,1)$

and

(iii) $(X_{(1)}, X_{(2)}-X_{(1)})\in \mathcal{N}(3,2')$.

The dual statements also hold. $F(x)$ was called in [1] to have increasing hazard ratio, if the condition (i) holds and $F(0)=0$. Here we remark that the condition (ii) can hold even if $F(x)$ is discrete, and that it gives another definition in wide sense of the concept of increasing hazard ratio.

Example 2.3. Let $F_U(x, y)=\min\{F_1(x), F_2(y)\}$ and $F_L(x, y)=\max\{0, 1-F_1(x)-F_2(y)\}$. Then for any x, y $F_U(x, y)\geqq F(x, y)\geqq F_L(x, y)$, if $F(x, y)$ is continuous. Let

$$F_\alpha(x, y)=\begin{cases}(1-\alpha)F_1(x)F_2(y)+\alpha F_U(x, y) & \text{for } 1\geqq\alpha\geqq 0\\ (1+\alpha)F_1(x)F_2(y)-\alpha F_L(x, y) & \text{for } -1\leqq\alpha<0 .\end{cases}$$

This family was introduced in [6]. For $\alpha \geqq 0$ it holds that $F_\alpha(x, y) \in \mathcal{P}(3, 1)$, $\mathcal{P}(2'', 2'')$ and $\mathcal{P}(2', 2')$, but not that $F_\alpha(x, y) \in \mathcal{P}(2', 2'')$ nor $\mathcal{P}(2'', 2')$.

3. Families $\mathcal{P}(\cdot, 0)$ and $\mathcal{P}(\cdot, E)$

In [13] we know other weaker order relationships $\mathcal{R}_0$ and $\mathcal{R}_E$, so we may define families $\mathcal{P}(i, 0)$ and $\mathcal{P}(i, E)$ along the line of Proposition 2.2.

DEFINITION 3.1. (i) $(X, Y) \in \mathcal{P}(4, 0)$ if $Y|_{a_3<X\leqq a_4} \succ Y|_{a_1<X\leqq a_2} (\mathcal{R}_0)$ for all $a_1<a_2<a_3<a_4$.

(ii) $(X, Y) \in \mathcal{P}(3, 0)$ if $Y|_{a_2<X\leqq a_3} \succ Y|_{a_1<X\leqq a_2} (\mathcal{R}_0)$ for all $a_1<a_2<a_3$.

Other families $\mathcal{P}(i, 0)$, $\mathcal{P}(0, j)$, $\mathcal{N}(i, 0)$ and $\mathcal{N}(0, j)$ are defined similarly for $i, j \in I$.

(iii) $(X, Y) \in \mathcal{P}(0, 0)$ if (X, Y) has nonnegative Kendall measure of association, that is $P_F((X-X')(Y-Y')>0) \geqq P_F((X-X')(Y-Y')<0)$, where the random variables (X, Y) and (X', Y') are independent and have the same distribution.

We could define $\mathcal{P}(4, i)$'s, $i \in I=\{1, 2', 2'', 2, 3\}$, but they coincide with $\mathcal{P}(3, i)$'s, respectively. $\mathcal{P}(4, 0)$ is really needed since the order relation $\mathcal{R}_0$ does not satisfy the transitivity, (see Example A.1 in Appendix).

PROPOSITION 3.1. Let (X, Y) and (X', Y') be identically and independently distributed (i.i.d.) random variables, and let (X_l, X_s) be defined by $(X, X')|_{Y>Y'}$. Then (i) $(X, Y) \in \mathcal{P}(4, 0)$ iff $X_l \succ X_s (\mathcal{R}_{3A})$, (ii) $(X, Y) \in \mathcal{P}(3, 0)$ iff $X_l \succ X_s (\mathcal{R}_3)$, (iii) $(X, Y) \in \mathcal{P}(2', 0)$ iff $X_l \succ X_s (\mathcal{R}_2')$ and (iv) $(X, Y) \in \mathcal{P}(0, 0)$ iff $X_l \succ X_s (\mathcal{R}_0)$. And similar results also hold for the other families of $\mathcal{P}(\cdot, 0)$, $\mathcal{P}(0, \cdot)$, $\mathcal{N}(\cdot, 0)$ and $\mathcal{N}(0, \cdot)$.

Tne following proposition gives the hierarchical structure of strictness of $\mathcal{P}(i, j)$'s defined in Definitions 2.1 and 3.1. Let $I'=\{0, 1, 2', 2'', 2, 3\}$. We extend the partial ordering in I to that in I' by adding an order relation $1>0$.

PROPOSITION 3.2. (i) It does not necessarily hold that $\mathcal{P}(1, 0) \subset \mathcal{P}(0, 0)$ nor dually that $\mathcal{P}(0, 1) \subset \mathcal{P}(0, 0)$.

(ii) $\mathcal{P}(i, j) \subset \mathcal{P}(i', j')$ for $i, j, i', j' \in I'$ iff $i \geqq i'$ and $j \geqq j'$ except for the above case.

(iii) $\mathcal{P}(4, 0) \subset \mathcal{P}(i, j)$ for $i, j \in I'$ iff $j=0$, and $\mathcal{P}(4, 0) \supset \mathcal{P}(i, j)$ for $i, j \in I'$ iff $i=3$ and $j \neq 0$. Dual statements also hold.

PROOF. The implications are shown by Proposition 3.1 and the hierarchical structure of $\mathcal{R}$'s. The absence of implications in the cases

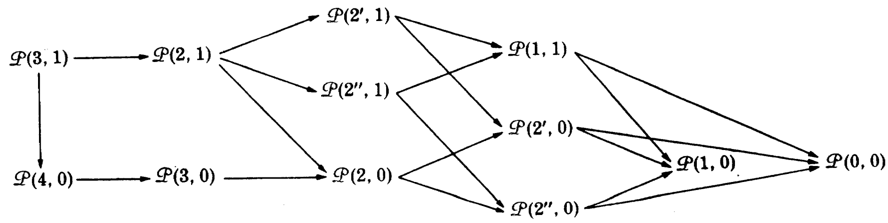

Fig. 2. Implication scheme (ii).

of (ii) and (iii) except for $\mathcal{P}(4,0)\not\subset\mathcal{P}(3,0)$, is shown by constructing analogous counter examples as in the proof of Proposition 2.2. To prove that $\mathcal{P}(0,j)\not\subset\mathcal{P}(0,j')$ for $j'>j$ and $j',j\in I'$, it is sufficient to show that there exists $F(y)$ and $G(y)$, such that $F(y)\succ G(y)\,(\mathcal{R}_j)$ but not $F(y)\succ G(y)\,(\mathcal{R}_{j'})$, which is assured in [13]. The remaining cases are shown by counter examples, Examples A.1 and A.2 in Appendix.

PROPOSITION 3.3. Paraller results for $\mathcal{P}(i,0)$, $i\in I'\cup\{4\}$, corresponding to Proposition 2.3 (i), (ii) and (iii) are valid, assuming $\sigma(0)=0$ and $\sigma(4)=4$. But the other ones do not hold.

Before introducing families $\mathcal{P}(\cdot,E)$ we shall define 'generalized covariance' which will be used in the definition of $\mathcal{P}(E,E)$.

DEFINITION 3.2. The generalized covariance of $F(x,y)$ is defined by

$$\int\int[F(x,y)-F_1(x)F_2(y)]dxdy\ ,$$

which is allowed to be positive or negative infinite. (X,Y) is said to have positive generalized covariance, if the generalized covariance is nonnegative or positive infinite.

PROPOSITION 3.4 ([8], [10]). Suppose $\mathrm{Cov}(X,Y)$ exists, then the generalized covariance also exists, and they coincide with each other.

DEFINITION 3.3. (i) $(X,Y)\in\mathcal{P}(3,E)$, if $Y|_{a_2<X\leqq a_3}\succ Y|_{a_1<X\leqq a_2}\,(\mathcal{R}_E)$ for all $-\infty\leqq a_1<a_2<a_3\leqq\infty$.

(ii) $(X,Y)\in\mathcal{P}(2',E)$ if $Y|_{a_2<X}\succ Y|_{a_1<X\leqq a_2}\,(\mathcal{R}_E)$ for all $-\infty\leqq a_1<a_2\leqq\infty$.

(iii) $(X,Y)\in\mathcal{P}(1,E)$ if $Y|_{a>x}\succ Y|_{X\leqq a}\,(\mathcal{R}_E)$ for all a.

(iv) $(X,Y)\in\mathcal{P}(E,E)$ if (X,Y) has positive generalized covariance.

(v) $(X,Y)\in\mathcal{P}(0,E)$, if $Y_l\succ Y_s\,(\mathcal{R}_E)$, where (Y_l,Y_s) is defined similarly as (X_l,X_s) in Proposition 3.1.

Other families of $\mathcal{P}(i,E)$, $\mathcal{P}(E,i)$, $\mathcal{N}(i,E)$ and $\mathcal{N}(E,i)$ can be defined similarly for $i\in I'$.

PROPOSITION 3.5. (i) $(X,Y)\in\mathcal{P}(1,E)$ implies $(X,Y)\in\mathcal{P}(E,E)$, provided the generalized covariance exists.

(ii) $(X, Y) \in \mathcal{P}(1, E)$ implies $(X, Y) \in \mathcal{P}(0, E)$, provided $\iint [F(x-0, y)+F(x, y)-(F_1(x-0)+F_1(x))F_2(y)]dF_1(x)dy$ exists.

PROOF. Supplemental conditions are needed to apply Fubini's theorem. Let x be an arbitrary number, then the distribution functions of $Y|_{X>x}$ and $Y|_{X\leqq x}$ are $(F_2(y)-F(x, y))/(1-F_1(x))$ and $F(x, y)/F_1(x)$ respectively. From $(X, Y) \in \mathcal{P}(1, E)$ it follows after some calculations that

$$\int [F(x, y)-F_1(x)F_2(y)]dy \geqq 0$$

which proves (i). Next the generalized mean difference of $Y|_{X\geqq x}$ and $Y|_{X<x}$ is also nonnegative as is that of $Y|_{X>x}$ and $Y|_{X\leqq x}$. Hence it follows that

$$\int [F(x-0, y)-F_1(x-0)F_2(y)]dy \geqq 0 \ .$$

and

$$\int [F(x, y)+F(x-0, y)-(F_1(x)+F_1(x-0))F_2(y)]dy \geqq 0 \ ,$$

which implies

$$\int \left\{\int [F(x, y)+F(x-0, y)-(F_1(x)+F_1(x-0))F_2(y)]dy\right\} dF_1(x) \geqq 0 \ .$$

On the other hand the generalized mean difference of Y_l and Y_s is given by

$$\begin{aligned}&\int \left\{\int F(x-0, y)/c\,dF_1(x) - \int [F_2(y)-F(x, y)]/c\,dF_1(x)\right\} dy \\ &\quad = 1/c \int \left\{\int [F(x, y)+F(x-0, y)-(F_1(x)+F_1(x-0))F_2(y)]dF_1(x)\right\} dy \ ,\end{aligned}$$

where $c=P(X>X')$, since $\int [F_1(x)-F_1(x-0)]dF_1(x)=1$. It is nonnegative, when Fubini's theorem is applicable. This completes the proof.

Let $I''=I' \cup \{E\}$. The partial ordering in I'', an extension of that in I', is given by adding the order relation $1>E$.

PROPOSITION 3.6. (i) $\mathcal{P}(i, j) \subset \mathcal{P}(i', j')$ for $i, j, i', j' \in I''$ iff $i \geqq i'$ and $j \geqq j'$ except for the cases in Propositions 3.2 and 3.5.

(ii) $\mathcal{P}(4, 0) \supset \mathcal{P}(i, j)$ for $i, j \in I''$ iff $i=3$, $j \neq 0$ and $j \neq E$. For $i, j \in I''$ $\mathcal{P}(4, 0) \subset \mathcal{P}(i, j)$ implies $\mathcal{P}(3, 0) \subset \mathcal{P}(i, j)$ unless $(i, j)=(4, 0)$.

PROOF. We need to show only the cases where the new parameter E is involved. The implications are obtained by the hierarchical struc-

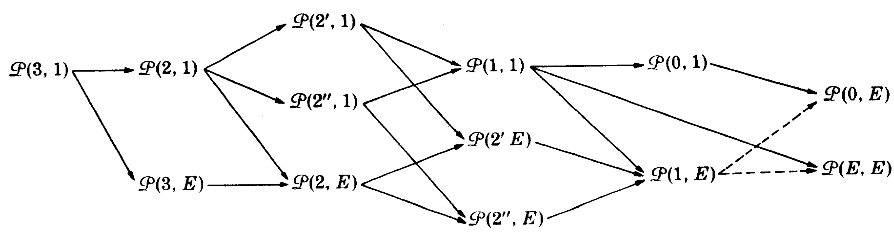

Fig. 3. Implication scheme (iii).

ture of $\mathcal{R}$'s. The absence of implications that for $i, i', j \in I'$ $\mathcal{P}(i', E) \not\subset \mathcal{P}(i, j)$, is obtained by constructing analogous examples in the proof of Proposition 2.2. For remaining cases it is shown by the counter examples, Examples A.3, A.4, A.5, A.6, A.7 and A.8 in Appendix.

PROPOSITION 3.7. Parallel results for $\mathcal{P}(\cdot, \cdot)$ in Definition 3.3 corresponding to Proposition 2.3 (i), (ii) and (iii) are valid, defining $\sigma(E) = E$. But the other ones do not hold.

Example 3.1. Let $X_{(1)} \geqq \cdots \geqq X_{(n)}$ be order statistics of a random sample $X_1, \cdots, X_n$ from the distribution $F(x)$. Under certain regularity conditions Bickel [2] and Lehmann [10] showed $\mathrm{Cov}(X_{(k)}, X_{(k')}) \geqq 0$ for any $1 \leqq k, k' \leqq n$. Actually Bickel proved, under the condition that $E(X_{(k)}^2)$ and $E(X_{(k')}^2)$ exist, that $(X_{(k)}, X_{(k')}) \in \mathcal{P}(3, E)$ for $k < k'$, and that $\mathrm{Cov}(X_{(k)}, X_{(k')}) \geqq 0$. Lehmann showed, under the condition that $F(x)$ is continuous, that $(X_{(k)}, X_{(k')}) \in \mathcal{P}(3, 3)$, and that $(X_{(k)}, X_{(k')}) \in \mathcal{P}(3, 1)$.

We give a complete explanation from our point of view. Let $U_1, \cdots, U_n$ be i.i.d. random variables with uniform distribution $U(0, 1)$, and let $U_{(1)} \geqq \cdots \geqq U_{(n)}$ be their order statistics. Then by direct calculation it follows that $(U_{(k)}, U_{(k')}) \in \mathcal{P}(3, 3)$ for any $1 \leqq k, k' \leqq n$. Since $F^{-1}(u) = \inf\{x \mid F(x) \geqq u\}$ is nondecreasing in u $(0 < u < 1)$, $(F^{-1}(U_{(k)}), F^{-1}(U_{(k')})) \in \mathcal{P}(3, 3)$, which implies $(X_{(k)}, X_{(k')}) \in \mathcal{P}(3, 3)$, a fortiori $(X_{(k)}, X_{(k')}) \in \mathcal{P}(E, E)$. Thus it follows $\mathrm{Cov}(X_{(k)}, X_{(k')}) \geqq 0$ by Proposition 3.4, provided $\mathrm{Cov}(X_{(k)}, X_{(k')})$ exists.

Other families can be defined similarly. Here we give one of them, which defines the family of random variables with the nonnegative quadrant measure of association [9].

DEFINITION 3.4. For the two-dimensional random variable (X, Y) we write $X \succ Y (\mathcal{R}_{1/2})$, if the median of X is not less than that of Y.

4. Families $\mathcal{P}(C)$ and $\mathcal{P}(A)$

In this section we present the third approach based on transformations in order to define two families $\mathcal{P}(C)$ and $\mathcal{P}(A)$. Families of negatively dependent random variables $\mathcal{N}(C)$ and $\mathcal{N}(A)$ can be defined dually.

Let U and V be the independent random variables each having a uniform distribution $U(0, 1)$. R^2 is the two dimensional Euclidian space and $(0, 1)^2$ is the unit square in R^2.

DEFINITION 4.1. (i) A function $r=(r_1, r_2) : (0, 1)^2 \rightarrow R^2$ is nondecreasing, if $u \leqq u'$ and $v \leqq v'$ imply that $r_1(u, v) \leqq r_1(u', v')$ and $r_2(u, v) \leqq r_2(u', v')$.

(ii) A function r is coherent, if for a suitable choice of numbers c_1 and c_2 in the set $\{-1, 1\}$, $c_1(u'-u) \geqq 0$ and $c_2(v'-v) \geqq 0$ imply that $(r_1(u', v')-r_1(u, v))(r_2(u', v')-r_2(u, v)) \geqq 0$. A nondecreasing function is coherent.

DEFINITION 4.2. (X, Y) is said to be positively coherent dependent, if there exists a coherent function $r(\cdot, \cdot)$ such that (X, Y) and $r(U, V)$ have a common distribution function $F(x, y)$. This is denoted by $(X, Y) \in \mathcal{P}(C)$.

PROPOSITION 4.1. $\mathcal{P}(3, 1) \subset \mathcal{P}(C)$, that is, if $(X, Y) \in \mathcal{P}(3, 1)$, then there exists a nondecreasing function $r(u, v)=(r_1(u), r_2(u, v))$ such that (X, Y) and $r(U, V)$ have a common distribution function $F(x, y)$.

PROOF. Since $(X, Y) \in \mathcal{P}(3, 1)$, conditional distribution functions satisfy the condition that $(u'-u)(F(y \mid F_1^{-1}(u'))-F(y \mid F_1^{-1}(u))) \leqq 0$ for all y and all $u, u' \in A$ for which A $(\subset(0, 1))$ has Lebesgue measure one. For $0<u<1$ let $\bar{F}(y \mid F_1^{-1}(u))=\inf\{F(y \mid F_1^{-1}(u')) \mid u' \in I,\ u' \geqq u\}$, then they also are conditional distribution functions of $F(x, y)$ given $X=F_1^{-1}(u)$. $r_1(u)$ and $r_2(u, v)$ are defined by $F_1^{-1}(u)$ and $\bar{F}^{-1}(v \mid F_1^{-1}(u))$ respectively. Then $r=(r_1, r_2)$ fulfills the required conditions.

Here we discuss the unbiasedness of tests of independence against positive dependence. Varying with the strictness of positive dependence tests may or may not be unbiased. In [10] Lehmann showed that the test of independence against $(X, Y) \in \mathcal{P}(3, 1)$ with its rejection region $T \geqq c$ for a nondecreasing rank statistic T and a suitable number c is unbiased. In [12] Blomqvist's test of independence against $\mathcal{P}(1, 1)$ was shown to be unbiased. The following proposition gives another relation.

We denote Kendall's rank correlation, Spearman's rank correlation and Blomqvist's rank correlation by T_K, T_S and T_B respectively: if (X_i, Y_i), $i=1, \cdots, n$ is a sample, then

$$T_K=\sum_{i \neq j} \operatorname{sgn}(X_i-X_j)(Y_i-Y_j)/n(n-1),$$

$$T_S=3 \sum_{i \neq j, j'} \operatorname{sgn}(X_i-X_j)(Y_i-Y_{j'})/n(n^2-1)$$

and

$$T_B=\left[1/4\sum_i(\operatorname{sgn}(X_i-X_0)+1)(\operatorname{sgn}(Y_i-Y_0)+1)\right]\Big/2[n/2]\ ,$$

where X_0 and Y_0 are sample medians and $[x]$ means the greatest integer less than or equal to x.

PROPOSITION 4.2. Suppose that distributions are continuous. Let T be one of T_K, T_S and T_B, then, for an arbitrary c and $F(x, y)\in\mathcal{P}(C)$, $P_F(T>c)\geqq P_0(T>c)$, where P_0 denotes the product probability measure of the two marginal distributions of $F(x, y)$. In other words a test of independence against $(X, Y)\in\mathcal{P}(C)$ with its rejection region $T>c$ for any c, is unbiased.

PROOF. It is sufficient to show that for a nondecreasing r, such that $r(U, V)$ has the distribution function $F(x, y)$, $T(r(u_1, v_1),\cdots, r(u_n, v_n))\geqq T((u_1, v_1),\cdots,(u_n, v_n))$ for almost all $((u_1, v_1),\cdots,(u_n, v_n))$. Using the fact that for almost all $((u, v), (u', v'))$ $(u-u')(v-v')>0$ implies $(r_1(u, v)-r_1(u', v'))(r_2(u, v)-r_2(u', v'))>0$, this is obtained by direct calculation.

Finally we define the family $\mathcal{P}(A)$, which was introduced in [3] with the symbol $\mathcal{A}$.

DEFINITION 4.3. (X, Y) is called to be associated, if for any nondecreasing function $r: R^2\to R^2$ $\operatorname{Cov}(r(X, Y))\geqq 0$ if it exists, which is denoted by $(X, Y)\in\mathcal{P}(A)$.

PROPOSITION 4.3. (i) $\mathcal{P}(3, 1)\subset\mathcal{P}(C)\subset\mathcal{P}(A)$, and both converse statements do not hold.

(ii) $\mathcal{P}(2', 1)\subset\mathcal{P}(A)\subset\mathcal{P}(1, 1)$, which implies that $\mathcal{P}(2'', 1)\cup\mathcal{P}(1, 2')\cup\mathcal{P}(1, 2'')\subset\mathcal{P}(A)$. And both converse statements do not hold.

(iii) Let $\mathcal{P}(i, j)$ be a family defined in Sections 2 and 3. $\mathcal{P}(C)\subset\mathcal{P}(i, j)$ implies $\mathcal{P}(1, 1)\subset\mathcal{P}(i, j)$.

(iv) It does not hold that $\mathcal{P}(2', 2')\cup\mathcal{P}(2'', 2'')\subset\mathcal{P}(C)$.

Remark. The problem, whether $\mathcal{P}(2', 2'')$ or $\mathcal{P}(2, 2)$ is included by $\mathcal{P}(C)$ or not, is open.

PROOF. (i) follows from [3] and Proposition 4.1. (ii) is a result in [4]. (iii) and (iv) follow from Examples A.6 and A.9.

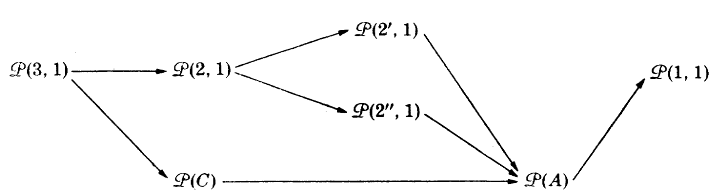

Fig. 4. Implication scheme (iv).

Here we give two typical distributions, one of which is in $\mathcal{P}(C)$ and the other not in $\mathcal{P}(C)$.

Example 4.1. Let $r(u, v)=(r_1(u, v), r_2(u, v))$ be

$$r_1(u, v)=\begin{cases} u/2 & \text{for } u<1/2 \text{ and } v<1/2 \\ u/2+1/2 & \text{for } u>1/2 \text{ and } v>1/2 \\ u/2+1/4 & \text{otherwise}, \end{cases}$$

and $r_2(u, v)\equiv r_1(v, u)$. Then (X, Y) $(=r(U, V))$ is distributed uniformly on the shadowed areas in Fig. 5. It holds that (X, Y) is in $\mathcal{P}(C)$ but not in $\mathcal{P}(3, 1)\cup\mathcal{P}(2', 2')\cup\mathcal{P}(2', 2'')\cup\mathcal{P}(2'', 2')\cup\mathcal{P}(2'', 2'')$.

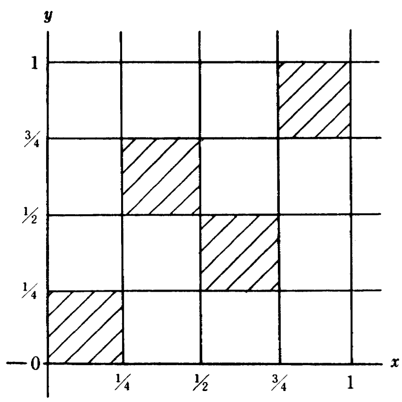

Fig. 5.

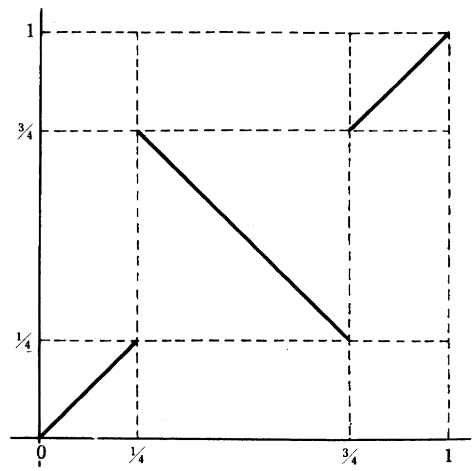

Fig. 6.

Example 4.2. Let (X, Y) be a random variable distributed uniformly on the thick lines in Fig. 6. (X, Y) is not in $\mathcal{P}(C)$. In fact let $(X_1, Y_1), \cdots, (X_4, Y_4)$ be i.i.d. random variables with the same distribution as (X, Y). Then

$$\begin{aligned} &P_F(Y_4>Y_3>Y_2>Y_1 \mid X_1>X_2>X_3>X_4) \\ &\qquad =1-P_F(T_K>-1)=1/16>1/4!=1-P_0(T_K>-1), \end{aligned}$$

which contradicts to Proposition 4.2, if $(X, Y)\in\mathcal{P}(C)$. On the other hand $(X, Y)\in\mathcal{P}(A)$.

Acknowledgement

The author wishes to express his hearty thanks to Mr. M. Sibuya, IBM Japan, and Prof. K. Isii, Osaka University, for their valuable suggestions and advices. His hearty thanks extend to the referee and Prof. M. Okamoto for their useful comments on the original draft.

Appendix. Counter examples to implication relationships

To complete the proofs of Propositions 3.2, 3.6 and 4.3, counter examples are presented here. Only brief interpretation is given for simplicity. For example it is interpretted in Example A.3 only that $\mathcal{P}(2,1)\not\subset\mathcal{P}(3,E)$, which implies that $\mathcal{P}(3,E)$ contains none of $\mathcal{P}(2',1)$, $\mathcal{P}(2'',1)$ and $\mathcal{P}(2,E)$, that dually $\mathcal{P}(1,2)\not\subset\mathcal{P}(E,3)$ and so on.

Example A.1. A random variable (X, Y) belonging to $\mathcal{P}(3,0)$ but not to $\mathcal{P}(4,0)$. Probabilities on 3×4 points:

$y \backslash x$	1	2	3
4	1/10	1/10	0
3	0	0	3/10
2	0	3/10	0
1	1/10	0	1/10

Example A.2. A random variable (X, Y) belonging to $\mathcal{P}(1,0)$ but not to $\mathcal{P}(0,0)$. Probabilities on 4×4 points:

$y \backslash x$	1	2	3	4
4	1/12	0	1/12	1/12
3	0	1/4	0	0
2	0	0	1/4	0
1	1/12	1/12	0	1/12

Example A.3. A random variable (X, Y) belonging to $\mathcal{P}(2,1)$ but not to $\mathcal{P}(3,E)$. Probabilities on 4×4 points:

$y \backslash x$	1	2	3	4
4	0	0	0	1/4
3	0	1/6	1/12	0
2	0	1/12	1/6	0
1	1/4	0	0	0

Example A.4. A random variable (X, Y) belonging to $\mathcal{P}(2',1)$ but to neither of $(\mathcal{P}(2'',0)$, $\mathcal{P}(2'',E))$. Probabilities on 3×2 points:

$y \backslash x$	1	2	3
2	1/4	0	1/4
1	1/4	1/4	0

Example A.5. A random variable (X, Y) belonging to both of $(\mathcal{P}(4, 0),\ \mathcal{P}(3, E))$ but to neither of $(\mathcal{P}(0, 1)$ and $\mathcal{P}(E, 1))$. Probabilities on 2×3 points:

$y \backslash x$	1	2
3	1/4	0
2	0	1/2
1	1/4	0

Example A.6. A random variable (X, Y) belonging to $\mathcal{P}(C)$ which is in both of $(\mathcal{P}(A),\ \mathcal{P}(1, 1))$ a fortiori but to none of $(\mathcal{P}(2', E),\ \mathcal{P}(2'', E),\ \mathcal{P}(2', 0),\ \mathcal{P}(2'', 0))$. Probabilities on 4×2 points:

$y \backslash x$	1	2	3	4
2	0	1/4	0	1/4
1	1/4	0	1/4	0

Example A.7. A random variable (X, Y) belonging to both of $(\mathcal{P}(E, E),\ \mathcal{P}(0, E))$ but to neither of $(\mathcal{P}(1, 0),\ \mathcal{P}(0, 1))$. Probabilities on 4×4 points:

$y \backslash x$	1	2	3	4
4	0	1/4	0	0
3	0	0	1/4	0
2	0	0	0	1/4
1	1/4	0	0	0

Example A.8. A random variable (X, Y) belonging to $\mathcal{P}(4, 0)$ but to neither to $(\mathcal{P}(E, E),\ \mathcal{P}(0, E))$. Probabilities on 2×5 points:

$y \backslash x$	1	2
5	0	1/3
4	1/2	0
3	0	0
2	0	0
1	0	1/6

$(-X, Y)$ belongs to both of $(\mathcal{P}(3, E),\ \mathcal{P}(E, E))$ but to neither of $(\mathcal{P}(0, 0),\ \mathcal{P}(E, 0))$.

Example A.9. A random variable (X, Y) belonging to both of $(\mathcal{P}(2', 2'),\ \mathcal{P}(2'', 2''))$ but not to $\mathcal{P}(C)$. (X, Y) distributes uniformly on the shadowed areas in Fig. 7 with probability 9/10, that is, its density

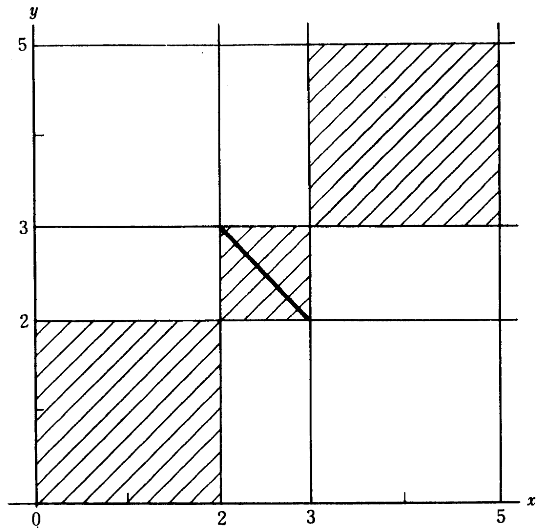

Fig. 7.

function is 1/10, and uniformly on the thick line in Fig. 7 with probability 1/10.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Barlow, R. E. and Proschan, F. (1965). *Mathematical Theory of Reliability*, John Wiley, New York.

[2] Bickel, P. J. (1966). Some contributions to the theory of order statistics, *Proc. 5th Berkeley Symp.*, 575-591.

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

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

[5] Farlie, D. J. E. (1960). The performance of some correlation coefficients for a general bivariate distribution, *Biometrika*, **47**, 307-323.

[6] Fréchet, M. (1951). Sur les tableux de correlation dont les marges sont donées, *Ann. de l'Universite de Lyon, Section A* 3e *Série, Fasc.*, **14**, 53-77.

[7] Jogdeo, K. and Patil, G. P. (1967). Characterization of independence in bivariate families with regression dependence, *Bull. Internat. Statist. Inst.*, **42** Book 1, 313-316.

[8] Konijn, H. S. (1959). Positive and negative dependence of two random variables, *Sankhya*, **21**, 269-280.

[9] Kruskal, W. H. (1958). Ordinal measures of association, *Jour. Amer. Statist. Assoc.*, **53**, 814-864.

[10] Lehmann, E. L. (1960). Some concepts of dependence, *Ann. Math. Statist.*, **37**, 1137-1153.

[11] Tukey, J. W. (1958). A problem of Berkson, and minimum variance orderly estimator, *Ann. Math. Statist.*, **29**, 588-592.

[12] Yanagimoto, T. and Okamoto, M. (1969). Partial orderings of permutations and monotonicity of a rank correlation statistic, *Ann. Inst. Statist. Math.*, **21**, 489-506.

[13] Yanagimoto, T. and Sibuya, M. (1972). Stochastically larger component of a random vector, *Ann. Inst. Statist. Math.*, **24**, 259-269.