The Annals of Statistics
1987, Vol. 15, No. 3, 1208–1214

AN ORDERING FOR POSITIVE DEPENDENCE[1]

BY B. F. SCHRIEVER

Nederlandse Philips Bedrijven

In this paper we introduce a partial ordering for positive dependent bivariate distributions. Our main result shows that tests of independence based on rank statistics such as Spearman's rho, Kendall's tau, Fisher–Yates' normal score statistic, van der Waerden's statistic and the quadrant statistic become more powerful under increasing positive dependence. In other words, these measures of positive dependence preserve the ordering stochastically in samples whenever it is present between underlying distributions.

1. Introduction. In this paper we introduce a (partial) ordering for bivariate distributions, called "more associated," which expresses the strength of positive dependence. The ordering makes precise an intuitive notion of one bivariate distribution being more positive-dependent than another. Suggestions for such orderings given in the literature are, for one reason or another, not completely satisfactory in our opinion. For example, Whitt (1982) discusses orderings for multivariate distributions which are related to the so-called HPK inequality [see also, Eaton (1982), Formula (3.6)]. These orderings seem to be quite strong because they do not order standard bivariate normal distributions according to the correlation parameter. A much weaker ordering, called "more concordant" in Section 2, is independently introduced in Cambanis, Simons and Stout (1976), in Tchen (1976) and also in Yanagimoto and Okamoto (1969). Although this ordering has the desirable property that most well-known measures of positive dependence preserve the ordering (in populations), it is unsatisfactory because this order-preserving property does not carry over to samples in a useful sense. The ordering of Kowalczyk and Pleszczyńska (1977) is still weaker than the ordering "more concordant." Yanagimoto and Okamoto (1969) also introduce a second ordering, which in fact is a special case of the ordering we shall introduce.

The examples in Section 2 illustrate that our ordering arises naturally in several models and families of bivariate distributions. Furthermore, it is proved in Section 2 that the ordering "more associated" is stronger than the ordering "more concordant." Therefore, it follows (cf. Section 3) that most well-known rank measures of positive dependence preserve the ordering "more associated" in populations. The main results of this paper are presented in Section 4 and show that when the ordering "more associated" is present between two underlying distributions (with continuous marginals), then the measures of positive dependence preserve the ordering stochastically in samples from these underlying

Received December 1985; revised December 1986.

[1]Work prepared while the author was a research fellow at the Vrije Universiteit, Amsterdam.

AMS 1980 *subject classifications*. Primary 62H99; secondary 62H20.

Key words and phrases. Positive bivariate dependence, more associated, more concordant, rank tests of independence.

distributions. This implies that tests based on these measures have power functions which are monotone in the ordering.

Schriever (1987) derives monotonicity results, similar to those of Section 4, in other nonparametric testing problems (one-sided two-sample problem and k-sample trend problem). His approach is based on partial orderings over pairs of permutations.

The ordering "more associated" can easily be generalized to the multivariate case such that sampling properties similar to those of Section 4 continue to hold.

2. Definitions and basic properties. Consider two arbitrary (empirical or underlying) bivariate distribution functions $H^{(1)}$ and $H^{(2)}$. For $k = 1, 2$, let $(X^{(k)}, Y^{(k)})$ be a pair of random variables with distribution function $H^{(k)}$ and let the marginal distribution functions of $X^{(k)}$ and $Y^{(k)}$ be denoted by $F^{(k)}$ and $G^{(k)}$, respectively. The *support* $\mathscr{X}^{(k)}$ of the distribution $F^{(k)}$ of $X^{(k)}$ is defined as $\mathscr{X}^{(k)} = \{x \in (-\infty, +\infty]\colon\ F^{(k)}(\xi) < F^{(k)}(x) \text{ for all } \xi < x\}$. This definition is unusual, but it is convenient since it makes $F^{(k)}$ strictly increasing on $\mathscr{X}^{(k)}$. (A real-valued function f defined on $\mathscr{X} \subseteq \mathbb{R}$ is said to be *increasing* on $\mathscr{X}$ when $z_1 < z_2\ (z_1, z_2 \in \mathscr{X}) \Rightarrow f(z_1) \le f(z_2)$ and is said to be *strictly increasing* on $\mathscr{X}$ if strict inequality is implied.) The support $\mathscr{Y}^{(k)}$ of the marginal distribution $G^{(k)}$ is similarly defined. Definitions and properties below are frequently formulated in terms of random variables, but they only concern their distributions. We use the symbol $\sim$ as short for "is distributed as."

DEFINITION 2.1. The pair $(X^{(2)}, Y^{(2)})$ is said to be *more associated* than $(X^{(1)}, Y^{(1)})$, denoted by $(X^{(2)}, Y^{(2)}) \ge_a (X^{(1)}, Y^{(1)})$ or by $H^{(2)} \ge_a H^{(1)}$, if there exist functions $\phi\colon \mathscr{X}^{(1)} \times \mathscr{Y}^{(1)} \to \mathscr{X}^{(2)}$ and $\psi\colon \mathscr{X}^{(1)} \times \mathscr{Y}^{(1)} \to \mathscr{Y}^{(2)}$ such that for all $x_1, x_2 \in \mathscr{X}^{(1)}$ and $y_1, y_2 \in \mathscr{Y}^{(1)}$

$$(2.1)\quad x_1 \le x_2, \quad y_1 \le y_2 \Rightarrow \phi(x_1, y_1) \le \phi(x_2, y_2), \quad \psi(x_1, y_1) \le \psi(x_2, y_2),$$

$$(2.2)\quad \phi(x_1, y_1) < \phi(x_2, y_2), \quad \psi(x_1, y_1) > \psi(x_2, y_2) \Rightarrow x_1 < x_2, \quad y_1 > y_2,$$

$$(2.3)\qquad (X^{(2)}, Y^{(2)}) \sim \big(\phi(X^{(1)}, Y^{(1)}),\ \psi(X^{(1)}, Y^{(1)})\big).$$

Of course, (2.1) means that ϕ and ψ are increasing in both arguments and (2.2) excludes reflection about the diagonal of the mapping (ϕ, ψ). Reflection is excluded in order to make the ordering applicable in situations which are not symmetric in the X and Y variable. [In case (2.2) is dropped, $(X, Y) \ge_a (Y, X) \ge_a (X, Y)$ and, hence, any function which is not symmetric in both variables does not preserve the ordering.]

In the special case that the function ϕ in Definition 2.1 satisfies $\phi(x, y) = x$ for all $x \in \mathscr{X}^{(1)}$ and $y \in \mathscr{Y}^{(1)}$, $(X^{(2)}, Y^{(2)})$ is said to be *more regression dependent* $(\ge_r)$ than $(X^{(1)}, Y^{(1)})$. In fact, this ordering is considered by Yanagimoto and Okamoto (1969) under the additional assumption that the distributions of both pairs have the same continuous marginals.

The ordering $\ge_a$ is called "more associated" because for any pair $(X, Y) \sim H$ with marginal distributions F and G, $H \ge_a FG$ implies that (X, Y) are associated in the sense of Esary, Proschan and Walkup (1967). The converse is, however, not true. It is proved in Schriever (1986) and Yanagimoto and Okamoto (1969) that $H \geq_r FG$ iff Y is regression dependent on X in the sense of Lehmann (1966), provided F and G are continuous.

It is seen from Definition 2.1 that the ordering $\geq_a$ is invariant under strictly increasing transformations of the marginals in the following sense. Define an equivalence relation $=_a$ by $(X^{(2)}, Y^{(2)}) =_a (X^{(1)}, Y^{(1)})$ iff there exist strictly increasing functions $\nu: \mathscr{X}^{(1)} \to \mathscr{X}^{(2)}$ and $\mu: \mathscr{Y}^{(1)} \to \mathscr{Y}^{(2)}$ such that $(X^{(2)}, Y^{(2)}) \sim (\nu(X^{(1)}), \mu(Y^{(1)}))$. Next, let $(X^{(2)}, Y^{(2)}) =_a (U^{(2)}, V^{(2)})$ and $(X^{(1)}, Y^{(1)}) =_a (U^{(1)}, V^{(1)})$. Then the invariance property states that

$$(2.4) \qquad (X^{(2)}, Y^{(2)}) \geq_a (X^{(1)}, Y^{(1)}) \quad \text{iff } (U^{(2)}, V^{(2)}) \geq_a (U^{(1)}, V^{(1)}).$$

Note that the relation $=_a$ is reflexive, symmetric and transitive and therefore indeed is an equivalence relation.

Although the ordering "more associated" is defined for arbitrary bivariate distributions, it turns out to behave more regularly on so-called similarity classes of bivariate distributions. A class of bivariate distributions $\mathscr{H}$ is called a *similarity class* if for any two pairs, $(X^{(2)}, Y^{(2)}) \sim H^{(2)} \in \mathscr{H}$ and $(X^{(1)}, Y^{(1)}) \sim H^{(1)} \in \mathscr{H}$, there exist strictly increasing functions $\nu: \mathscr{X}^{(1)} \to \mathscr{X}^{(2)}$ and $\mu: \mathscr{Y}^{(1)} \to \mathscr{Y}^{(2)}$ such that $X^{(2)} \sim \nu(X^{(1)})$ and $Y^{(2)} \sim \mu(Y^{(1)})$, that is, the marginals of distributions in $\mathscr{H}$ are equal up to strictly increasing transformations. Examples of similarity classes are $\mathscr{H} =$ {bivariate distribution functions with continuous marginal distribution functions}, $\mathscr{H} =$ {empirical bivariate distribution functions based on N observations without ties} and $\mathscr{H} =$ {bivariate distributions with given marginals}. It is not hard to verify [using (2.4) and the following proposition] that $\geq_a$ is a partial ordering on any similarity class of the equivalence classes defined by $=_a$.

Proposition 2.1 shows that the ordering "more associated" is stronger than the ordering "more concordant" of Cambanis, Simons and Stout (1976), Tchen (1976) and Yanagimoto and Okamoto (1969).

DEFINITION 2.2. The pair $(X^{(2)}, Y^{(2)})$ is said to be *more concordant* than $(X^{(1)}, Y^{(1)})$, denoted by $(X^{(2)}, Y^{(2)}) \geq_c (X^{(1)}, Y^{(1)})$ or by $H^{(2)} \geq_c H^{(1)}$, when $F^{(2)}(x) = F^{(1)}(x)$, $G^{(2)}(y) = G^{(1)}(y)$ and $H^{(2)}(x, y) \geq H^{(1)}(x, y)$ for all $x, y \in \mathbb{R}$.

Obviously, the class of bivariate distributions which is "more concordant" than independence equals the class of distributions which are called (positively) quadrant-dependent by Lehmann (1966). The ordering $\geq_c$ is a partial ordering on the class of all bivariate distributions.

PROPOSITION 2.1. $H^{(2)} \geq_a H^{(1)}$, $F^{(2)} = F^{(1)}$, $G^{(2)} = G^{(1)} \Rightarrow H^{(2)} \geq_c H^{(1)}$.

PROOF. Let $\phi: \mathscr{X}^{(1)} \times \mathscr{Y}^{(1)} \to \mathscr{X}^{(1)}$ and $\psi: \mathscr{X}^{(1)} \times \mathscr{Y}^{(1)} \to \mathscr{Y}^{(1)}$ be functions such that (2.1), (2.2) and (2.3) hold. Choose $x \in \mathscr{X}^{(1)}$ and $y \in \mathscr{Y}^{(1)}$. First, consider the case that $\phi(x, y) \leq x$ and $\psi(x, y) \leq y$. Then by (2.1) and (2.3), it follows that $H^{(1)}(x, y) = P\{X^{(1)} \leq x, \ Y^{(1)} \leq y\} \leq P\{X^{(2)} \leq \phi(x, y), \ Y^{(2)} \leq$

$\psi(x, y)\} \le P\{X^{(2)} \le x, Y^{(2)} \le y\} = H^{(2)}(x, y)$. Next, consider the case that $\phi(x, y) \le x$ and $\psi(x, y) > y$. Then by (2.2) and (2.3), $P\{X^{(1)} > x, Y^{(1)} < y\} \ge P\{X^{(2)} > \phi(x, y), Y^{(2)} < \psi(x, y)\} \ge P\{X^{(2)} > x, Y^{(2)} \le y\}$. Since $X^{(2)} \sim X^{(1)}$ and $Y^{(2)} \sim Y^{(1)}$, it also follows for this case that $H^{(1)}(x, y) \le H^{(2)}(x, y)$. For the remaining two cases $\phi(x, y) > x$, $\psi(x, y) > y$ and $\phi(x, y) > x$, $\psi(x, y) \le y$, the result $H^{(1)}(x, y) \le H^{(2)}(x, y)$ follows in a similar way. Hence, $H^{(2)} \ge_c H^{(1)}$. □

We close this section with two examples of the ordering $\ge_a$.

EXAMPLE 2.1. This example gives more insight in the way probability mass is transformed by $\ge_a$. Consider the case where $(X^{(2)}, Y^{(2)})$ is a linear transform, with positive coefficients of $(X^{(1)}, Y^{(1)})$. Since the ordering $\ge_a$ is invariant under location and scale transformations, there is no loss of generality to assume that

$$(X^{(2)}, Y^{(2)}) \sim \left((1-\alpha)X^{(1)} + \alpha Y^{(1)}, (1-\beta)X^{(1)} + \beta Y^{(1)}\right), \tag{2.5}$$

for $0 \le \alpha \le \beta \le 1$. It is seen that the function $(\phi(x, y), \psi(x, y)) = ((1-\alpha)x + \alpha y, (1-\beta)x + \beta y)$ maps the mass of the distribution $H^{(1)}$ which lies in the square with vertices (e, e), $(-e, e)$, $(-e, -e)$, $(e, -e)$ onto the rhombus with vertices (e, e), $((2\alpha - 1)e, (2\beta - 1)e)$, $(-e, -e)$, $((1 - 2\alpha)e, (1 - 2\beta)e)$. Clearly, the mass of $H^{(2)}$ is more concentrated around the line $y = x$ than that of $H^{(1)}$.

Special cases of the transform (2.5) give rise to one-parameter families of bivariate distributions $\{H^{(\alpha)}\}_{\alpha \in A}$, e.g.,

$$(X^{(\alpha)}, Y^{(\alpha)}) \sim ((1-\alpha)X + \alpha Y, Y), \quad \text{for } \alpha \in A = [0,1]$$

and

$$(X^{(\alpha)}, Y^{(\alpha)}) \sim ((1-\alpha)X + \alpha Y, \alpha X + (1-\alpha)Y), \quad \text{for } \alpha \in A = [0, \tfrac{1}{2}].$$

In both families we have $H^{(\alpha_2)} \ge_a H^{(\alpha_1)}$ iff $\alpha_2 \ge \alpha_1$, provided $\alpha_1, \alpha_2 \in A$.

EXAMPLE 2.2. Let $(X^{(1)}, Y^{(1)})$ and $(X^{(2)}, Y^{(2)})$ have bivariate normal distributions with correlation parameters ρ_1 and ρ_2, respectively, then $(X^{(2)}, Y^{(2)}) \ge_a (X^{(1)}, Y^{(1)})$ iff $-1 < \rho_1 \le \rho_2 \le 1$. Moreover, the ordering $\ge_a$ is equivalent to the ordering of the correlation parameter in any elliptical family.

3. Measures of dependence preserving the ordering. In this section we give examples of measures of positive dependence which preserve the ordering $\ge_a$ in populations. Measures of interest can be written as real-valued functions of (underlying or empirical) distribution functions. Such a measure Φ is said to preserve $\ge_a$ when

$$H^{(2)} \ge_a H^{(1)} \Rightarrow \Phi(H^{(2)}) \ge \Phi(H^{(1)}). \tag{3.1}$$

Since the ordering $\ge_a$ is invariant under strictly increasing transformations of the marginals, measures which preserve $\ge_a$ must also be invariant under such transformations. Rank measures have this invariance property. Furthermore, it follows from Proposition 2.1 that any invariant measure which preserves $\ge_c$

also preserves $\geq_a$ on similarity classes, i.e., (3.1) holds provided $H^{(2)}$ and $H^{(1)}$ are in the same similarity class.

EXAMPLE 3.1. *Linear rank statistics.* Linear rank statistics can be written in the form

$$\Phi(H) = \iint_{\mathbb{R}^2} J(F(x), G(y))\, dH(x, y), \tag{3.2}$$

where J is a real-valued function defined on $(0,1] \times (0,1]$, called the score function. It follows from Cambanis, Simons and Stout (1976) that [under some regularity conditions which are satisfied, for instance, when the integral (3.2) exists and is finite under independence] measures of the form (3.2) preserve $\geq_c$ when J is right-continuous and lattice-superadditive, i.e., $u_1 < u_2$, $v_1 < v_2 \Rightarrow J(u_1, v_1) + J(u_2, v_2) \geq J(u_1, v_2) + J(u_2, v_1)$. The score function of *Spearman's rank correlation rho, Fisher–Yates' normal score statistic, van der Waerden's statistic* and the *quadrant statistic* are all of product type $J(u, v) = f(u)g(v)$ with f and g increasing and, hence, are lattice-superadditive. Therefore, these linear rank statistics preserve $\geq_c$.

EXAMPLE 3.2. *Nonlinear rank statistics.* It is not hard to show [cf. Schriever (1986)] using the result of Cambanis, Simons and Stout (1976) that measures of the form $\Phi(H) = \iint K(H(x, y))\, dH(x, y)$ preserve $\geq_c$ when K is right-continuous, increasing and convex. An example of such a nonlinear rank statistic is *Kendall's rank correlation tau.*

EXAMPLE 3.3. *Isotonic canonical correlation.* Define the isotonic canonical correlation of a pair (X, Y) with distribution function H by $\Phi(H) = \sup \operatorname{Corr}(\nu(X), \mu(Y))$, where the supremum is taken over all increasing functions ν and μ for which the correlation exists. It is easily verified [again by using Cambanis, Simons and Stout (1976)] that this measure preserves $\geq_c$.

These examples show that all familiar rank measures of positive dependence preserve $\geq_a$ on similarity classes. Note that Pearson's product moment correlation and omnibus measures such as Pearson's chi-squared, the canonical correlation and the correlation ratio do not preserve $\geq_a$ (on similarity classes).

4. Sampling properties of the ordering. The main results formulated in this section show that the order preserving properties of the ordering $\geq_a$ given in the previous section carry over from population distributions to finite samples from these distributions.

Let $H^{(2)}$ and $H^{(1)}$ be arbitrary bivariate distribution functions and let $\hat{H}_N^{(2)}$ and $\hat{H}_N^{(1)}$ be empirical distribution functions based on N i.i.d. observations from $H^{(2)}$ and $H^{(1)}$, respectively.

THEOREM 4.1. *Let Φ be a measure which preserves $\geq_a$. Then*

$$H^{(2)} \geq_a H^{(1)} \Rightarrow P\{\Phi(\hat{H}_N^{(2)}) > t\} \geq P\{\Phi(\hat{H}_N^{(1)}) > t\}, \tag{4.1}$$

for all t and N, i.e., $\Phi(\hat{H}_N^{(2)})$ is stochastically larger than $\Phi(\hat{H}_N^{(1)})$.

PROOF. Let $(X_1^{(1)}, Y_1^{(1)}), \ldots, (X_N^{(1)}, Y_N^{(1)})$ be a sample of N i.i.d. observations from $H^{(1)}$ and let $\hat{H}_N^{(1)}$ be the corresponding empirical distribution function. Since $H^{(2)} \geq_a H^{(1)}$, there exist functions ϕ and ψ such that $(\phi(X_i^{(1)}, Y_i^{(1)}), \psi(X_i^{(1)}, Y_i^{(1)}))$ for $i = 1, \ldots, N$ is a sample of N i.i.d. observations from $H^{(2)}$. Denote its empirical distribution function by $\hat{H}_N^*$. Clearly, $\hat{H}_N^*$ has the same distribution as any other empirical distribution function $\hat{H}_N^{(2)}$ based on N i.i.d. observations from $H^{(2)}$. Furthermore, it is obvious that $P\{\hat{H}_N^* \geq_a \hat{H}_N^{(1)}\} = 1$, which implies $P\{\Phi(\hat{H}_N^*) \geq \Phi(\hat{H}_N^{(1)})\} = 1$. Hence, $P\{\Phi(\hat{H}_N^{(1)}) > t \geq \Phi(\hat{H}_N^*)\} = 0$ for all t and (4.1) follows. □

In case the underlying distributions $H^{(2)}$ and $H^{(1)}$ have continuous marginals, the class of all empirical distribution functions based on i.i.d. observations from $H^{(2)}$ and $H^{(1)}$ is a similarity class. Then (4.1) also holds for measures Φ which preserve $\geq_a$ only on similarity classes such as Spearman's rho, Fisher–Yates' normal score statistic, van der Waerden's statistic, the quadrant statistic, Kendall's tau and the isotonic canonical correlation. This implies that tests of independence based on these statistics have a higher power against $H^{(2)}$ than against $H^{(1)}$. In particular, such tests are unbiased against all alternatives (with continuous marginals) which are more associated than independence. A similar monotonicity property for the special case $\geq_r$ is proved in a different, less transparent fashion by Yanagimoto and Okamoto (1969). Their result implies unbiasedness against regression dependent alternatives with continuous marginals [cf. Lehmann (1966)].

Another aspect of the order preserving property in samples is the following. Suppose $H^{(2)} \geq_a H^{(1)}$, then samples from $H^{(2)}$ turn out to be more frequently associated than samples from $H^{(1)}$.

THEOREM 4.2. *Let $H^{(2)} \geq_a H^{(1)}$. Then for all sample sizes N*

$$P\{\hat{H}_N^{(2)} \text{ is associated}\} \geq P\{\hat{H}_N^{(1)} \text{ is associated}\}.$$

PROOF. Let $\hat{H}_N^{(1)}$ and $\hat{H}_N^*$ be as in the proof of Theorem 4.1. By Property P4 of Esary, Proschan and Walkup (1967), $\hat{H}_N^{(1)}$is associated $\Rightarrow \hat{H}_N^*$ is associated with probability one. □

The main result (4.1) generally does not hold for the statistics considered in Section 3 when the underlying distributions $H^{(1)}$ and $H^{(2)}$ contain atoms. It is demonstrated in Schriever (1986) that the results actually carry over to the discrete case when the ordering and the statistics are appropriately modified to samples containing ties.

Acknowledgments. The author is very grateful to Dr. R. D. Gill and to Prof. Dr. J. Oosterhoff for stimulating discussions and helpful suggestions.

REFERENCES

CAMBANIS, S., SIMONS, G. and STOUT, W. (1976). Inequalities for $Ek(X, Y)$ when the marginals are fixed. *Z. Wahrsch. verw. Gebiete* **36** 285–294.

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

ESARY, J. D., PROSCHAN, F. and WALKUP, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38** 1466–1474.

KOWALCZYK, T. and PLESZCZYŃSKA, E. (1977). Monotonic dependence functions of bivariate distributions. *Ann. Statist.* **5** 1221–1227.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

SCHRIEVER, B. F. (1986). Order dependence. CWI-Tract 20. Centre for Mathematics and Computer Sciences, Amsterdam.

SCHRIEVER, B. F. (1987). Monotonicity of rank statistics in some non-parametric problems. *Statist. Neerlandica* **41**. To appear.

TCHEN, A. H.-T. (1976). Inequalities for distributions with given marginals. Technical Report 19, Dept. Statistics, Stanford Univ.

WHITT, W. (1982). Multivariate monotone likelihood ratio and uniform conditional stochastic order. *J. Appl. Probab.* **19** 695–701.

YANAGIMOTO, T. and OKAMOTO, M. (1969). Partial orderings of permutations and monotonicity of a rank correlation statistic. *Ann. Inst. Statist. Math.* **21** 489–506.

CENTRE FOR QUANTITATIVE METHODS
NEDERLANDSE PHILIPS BEDRIJVEN B. V.
BUILDING HCM-722
P.O. BOX 218
5600 MD EINDHOVEN
THE NETHERLANDS