Ann. Inst. Statist. Math.
Vol. 44, No. 3, 501–517 (1992)

FURTHER DEVELOPMENTS ON SOME DEPENDENCE ORDERINGS FOR CONTINUOUS BIVARIATE DISTRIBUTIONS*

Z. FANG AND H. JOE

Department of Statistics, University of British Columbia, Vancouver, B.C., Canada V6T 1Z2

(Received February 2, 1990; revised July 22, 1991)

Abstract. The dependence orderings, "more associated" and "more regression dependent", due to Schriever (1986, *Order Dependence*, Centre for Mathematics and Computer Sciences, Amsterdam; 1987, *Ann. Statist.*, **15**, 1208–1214) and Yanagimoto and Okamoto (1969, *Ann. Inst. Statist. Math.*, **21**, 489–505) respectively, are studied in detail for continuous bivariate distributions. Equivalent forms of the orderings under some conditions are given so that the orderings are more easily checkable for some bivariate distributions. For several parametric bivariate families, the dependence orderings are shown to be equivalent to an ordering of the parameter. A study of functionals that are increasing with respect to the "more associated ordering" leads to inequalities, measures of dependence as well as a way of checking that this ordering does not hold for two distributions.

Key words and phrases: Dependence ordering, regression dependence, concordance, copula.

1. Introduction

There have been several recent papers on bivariate and multivariate dependence orderings, for example, Schriever (1986, 1987), Kimeldorf and Sampson (1987), Metry and Sampson (1988), Block *et al.* (1990), Joe (1990*b*). There is also recent work on families of multivariate distributions, for example, Marshall and Olkin (1988). One reason for this work on multivariate dependence is to obtain properties of multivariate distributions; these properties are useful for deciding on appropriate models for multivariate data. In this paper, we link the work on the "more associated" and "more regression dependent" orderings with families of continuous bivariate distributions. The importance of research in this area is emphasized in Kimeldorf and Sampson (1987). Our results and examples add substantially to the understanding of dependence orderings.

* This research has been supported by NSERC Canada grants and a Scientific Grant of the University of Science and Technology of China.

The "more associated" partial ordering is due to Schriever (1986, 1987) and is studied in more detail in Block *et al.* (1990) for bivariate empirical distributions. The "more regression dependent" partial orderings are due to Yanagimoto and Okamoto (1969). The definitions and alternative forms of these orderings are given in Section 2 as well as new results that make the orderings more easily checkable for some bivariate distributions. In Section 3, it is shown that these orderings are equivalent to an ordering on the parameter(s) for several parametric bivariate families. With orderings, it is natural to consider functionals that are increasing with respect to it (cf. Kimeldorf and Sampson (1989) and the references cited above). Some functionals are derived in Section 4 using the orderings for empirical distributions in Block *et al.* (1990). The link between the (discrete) empirical distributions and continuous distributions adds to the understanding of the orderings, and the functionals provide a way of checking whether two given bivariate distributions are ordered. Sections 3 and 4 can be considered as separate sequels to Section 2, although the functionals in Section 4 lead to inequalities for the bivariate families in Section 3. In both these sections, we have made more extensive studies than previous authors. Furthermore, new statistical measures of dependence are derived in Section 4 along with the functionals. The corresponding tests of independence versus positive dependence are then more powerful for bivariate distributions that are "more associated" or "more regression dependent".

2. Definitions of orderings and new results

We will mainly be using H for a continuous bivariate cumulative distribution function (cdf), and F and G for univariate margins, with subscripts or superscripts (sometimes a prime symbol) as needed to distinguish distributions. Univariate quantile functions will have a superscript -1 on a univariate cdf. We use the term "increasing" in place or "non-decreasing" and "decreasing" in place or "non-increasing". For random variables or vectors, we use the symbol $\sim$ to mean "distributed as" and the symbol $\overset{d}{=}$ for "equal in distribution" or "stochastically equal".

To study the dependence structure of continuous bivariate distributions, the effect of the univariate margins can be separated out because of the following well-known result (see, for example, Sklar (1959)). If $H(x,y)$ is a continuous cdf with univariate margins $F(x)$ and $G(y)$, then $C(u,v) = H(F^{-1}(u), G^{-1}(v))$ is a copula or a bivariate distribution with uniform $(0,1)$ margins and $C(F_0(x), G_0(y))$ is a bivariate distribution with univariate margins F_0, G_0 where F_0, G_0 are arbitrary continuous univariate distribution functions. That is, the bivariate structure is in C, "independent" of the univariate margins. Therefore, for dependence comparisons of cdf's $H(x,y)$ and $H'(x,y)$, we could assume without loss of generality that $F(x) = F'(x)$ and $G(y) = G'(y)$, where F, G and F', G' are respectively the univariate margins of H and H'.

We now state definitions and prove results. Definition 2.1 is an adaptation of the definition of Schriever (1987) to continuous H, H'.

DEFINITION 2.1. Suppose (X,Y) and (X',Y') are pairs of continuous random variables such that $(X,Y) \sim H$, $(X',Y') \sim H'$, $X \overset{d}{=} X'$, $Y \overset{d}{=} Y'$. Then

(X', Y') is said to be *more associated* than (X, Y), denoted by $(X, Y) \prec^a (X', Y')$ or $H \prec^a H'$, if there exist functions ϕ, ψ such that for x_1, x_2 in the support of X and y_1, y_2 in the support of Y,

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

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

$$(2.3) \quad (X', Y') \stackrel{d}{=} (\phi(X, Y), \psi(X, Y)).$$

DEFINITION 2.2. (a) If in Definition 2.1 $\phi(x, y) = x$, then (X', Y') is said to be *more regression-*1 *dependent* than (X, Y), denoted by $(X, Y) \prec^{r1} (X', Y')$ or $H \prec^{r1} H'$. (b) If in Definition 2.1 $\psi(x, y) = y$, then (X', Y') is said to be *more regression-*2 *dependent* than (X, Y), denoted by $(X, Y) \prec^{r2} (X', Y')$ or $H \prec^{r2} H'$.

Note that (2.1) implies that ϕ and ψ are increasing functions (with respect to all arguments). It is stated in Schriever (1986, 1987) and not difficult to check that $\prec^a$, $\prec^{r1}$, $\prec^{r2}$ are partial orderings. The orderings in Definition 2.2, stated in a different form, are called monotone regression dependence orderings in Yanagimoto and Okamoto (1969). Schriever (1986, 1987) mentions the equivalence without proofs or conditions. With some continuity assumptions, the above definitions can be put in alternative forms that are easier to use; these are the main results of this section.

Let $G(y \mid x)$ be the conditional cdf of Y given $X = x$ and let $G^{-1}(u \mid x) = \inf\{y : G(y \mid x) > u\}$, $0 < u < 1$, be the conditional quantile function; this quantile function is the right continuous version and not the usual left continuous version. Similarly define $F(x \mid y)$ as the conditional cdf of X given $Y = y$ and let its right continuous inverse be $F^{-1}(u \mid y)$. Before stating the definition of Yanagimoto and Okamoto (1969), we give an alternative form of Definition 2.2(a). There is a correspondence between results for $\prec^{r1}$ and $\prec^{r2}$ by interchanging pairs of random variables, so we will not state equivalent results for $\prec^{r2}$ unless necessary.

LEMMA 2.1. *Suppose $X = X'$, $X \sim F$, $(X, Y) \sim H$, $(X', Y') \sim H'$, and that $G(y \mid x)$, $G'(y \mid x)$ are continuous in y for all x. Then $Y' \stackrel{d}{=} \psi(X, Y)$ where $\psi(x, y) = G'^{-1}(G(y \mid x) \mid x)$.*

PROOF. Given $X = x$, $Y \sim G(\cdot \mid x)$. Since $G(y \mid x)$ is continuous in y, $G(Y \mid x)$ is uniform conditional on $X = x$. Also if U is a uniform random variable on $(0, 1)$, then conditional on $X = x$, $G'^{-1}(U \mid x)$ has distribution $G'(y \mid x)$ by the continuity of this function in y. Therefore $G'^{-1}(G(Y \mid X) \mid X) \stackrel{d}{=} Y'$. □

Remark. With a little bit more effort, the conclusion of Lemma 2.1 can be shown to be still true if $G(y \mid x)$ and $G'(y \mid x)$ are not continuous in y for all y (when x is fixed), provided $G'^{-1}(u \mid x)$ is constant over an interval $(G(y_0^- \mid x), G(y_0 \mid x))$ for any discontinuity point of y_0 of $G(\cdot \mid x)$.

THEOREM 2.1. *With the suppositions of Lemma 2.1, if $H \prec^{r1} H'$, then $\psi(x, y) = G'^{-1}(G(y \mid x) \mid x)$ can be taken to be the unique function satisfying the condition in Definition 2.2 and hence it is increasing in x.*

PROOF. A function ψ such that $Y' = \psi(X, Y)$ must be increasing by Definition 2.2. $\psi(x, y)$ must be strictly increasing for y in the support of $G(\cdot \mid x)$ for all x, because if for a given x, $\psi(x, \cdot)$ has a constant region with positive probability, then $Y' = \psi(x, Y)$ has a mass at some point or $G'(\cdot \mid x)$ is not continuous. It follows that if $G(\cdot \mid x)$ is constant on an interval (y_1, y_2) then $G'(\cdot \mid x)$ is constant on $(\psi(x, y_1), \psi(x, y_2))$, so that, given $X = x$, $\psi(x, Y) \stackrel{d}{=} Y' \sim G'(\cdot \mid x)$ and $\Pr(Y \leq y \mid X = x) = P(\psi(x, Y) \leq \psi(x, y) \mid X = x)$. Therefore, $G(y \mid x) = G'(\psi(x, y) \mid x)$, which in turn implies $\psi(x, y) \leq G'^{-1}(G(y \mid x) \mid x)$, with equality whenever $G'(\cdot \mid x)$ is strictly increasing at $\psi(x, y)$. Hence, given $X = x$, $\psi(x, Y) = G'^{-1}(G(Y \mid x) \mid x)$ if $G'(\cdot \mid x)$ is strictly increasing at $\psi(x, Y)$. The set of points where $G'(\cdot \mid x)$ is strictly increasing has probability one with respect to $G'(\cdot \mid x)$. So $\psi(x, Y) = G'^{-1}(G(Y \mid x) \mid x)$ with probability one. The conclusion follows. □

DEFINITION 2.3. (Yanagimoto and Okamoto (1969)) Suppose H, H' are continuous bivariate distributions with the same pair of univariate margins. H' is *more monotone regression dependent* than H (written $H \prec^{mr} H'$) if for any $x_1 < x_2$ and u, v in $(0, 1)$,

$$G^{-1}(u \mid x_2) \geq G^{-1}(v \mid x_1) \quad \Rightarrow \quad G'^{-1}(u \mid x_2) \geq G'^{-1}(v \mid x_1). \tag{2.4}$$

PROPOSITION 2.1. *Assume $G(y \mid x)$ and $G'(y \mid x)$ are continuous in y for all x. Then $H \prec^{mr} H'$ if and only if*

$$G(y \mid x_1) \geq G'(y' \mid x_1) \quad \Rightarrow \quad G(y \mid x_2) \geq G'(y' \mid x_2) \tag{2.5}$$

for any $x_1 < x_2$ and any y, y' with y in the support of $G(\cdot \mid x_1)$ and y' in the support of $G(\cdot \mid x_2)$.

PROOF. Yanagimoto and Okamoto (1969) have this result with the added condition that $G(y \mid x)$ and $G'(y \mid x)$ are strictly increasing in y for all x. A careful check of their proof shows that this is not needed. □

THEOREM 2.2. *Assume $G(y \mid x)$ and $G'(y \mid x)$ are continuous in y and x. Then $H \prec^{mr} H'$ if and only if $H \prec^{r1} H'$.*

PROOF. Let $(X, Y) \sim H$ and $(X', Y') \sim H'$. $H \prec^{r1} H'$ implies $(X', Y') \stackrel{d}{=} (X, \psi(X, Y))$ with $\psi(x, y)$ increasing in x and y. Suppose the left inequality of (2.5) holds for some x_1, y, y'. Then $\Pr(Y \leq y \mid X = x_1) \geq \Pr(\psi(X, Y) \leq y' \mid X = x_1) = \Pr(\psi(x_1, Y) \leq y' \mid X = x_1) \geq \Pr(\psi(x_2, Y) \leq y' | X = x_1)$ for all $x_2 \geq x_1$. Taking a limit as x_1 increases to x_2 leads to the right inequality of (2.5). Therefore $H \prec^{mr} H'$.

For the converse, assume that (2.5) holds for $x_1 < x_2$. Let $\psi(x, y) = G'^{-1}(G(y \mid x) \mid x)$. By Lemma 2.1 and Theorem 2.1, it suffices to show that $\psi(x, y)$ is increasing in x. Fix $x_1 < x_2$ and fix y. Let y' be the largest value satisfying $G(y \mid x_1) = G'(y' \mid x_1)$. Then (2.5) implies that $\psi(x_2, y) = G'^{-1}(G(y \mid x_2) \mid x_2) \geq G'^{-1}(G'(y' \mid x_2) \mid x_2) \geq y' = G'^{-1}(G'(y' \mid x_1) \mid x_1) = G'^{-1}(G(y \mid x_1) \mid x_1) = \psi(x_1, y)$. □

Next we go on to results for the $\prec^a$ ordering. Assuming ϕ and ψ in Definition 2.1 to be continuous, we obtain an equivalent definition.

THEOREM 2.3. *Let $\phi(x,y)$ and $\psi(x,y)$ be continuous functions such that $(X,Y) \sim H$, $(X',Y') \sim H'$ and $X' \stackrel{d}{=} \phi(X,Y)$, $Y' \stackrel{d}{=} \psi(X,Y)$. Then $H \prec^a H'$ if and only if ϕ and ψ are increasing in both arguments and*

$$\phi_1\psi_2 - \phi_2\psi_1 \geq 0 \quad \forall x, y \tag{2.6}$$

in the support of H, where $\phi_1(x,y)$, $\phi_2(x,y)$ are respectively the right partial derivatives with respect to the first and second arguments at (x,y), and ψ_1, ψ_2 are similarly defined.

PROOF. It suffices to show that (2.2) and (2.6) are equivalent assuming (2.1). Note that if $x' = \phi(x,y)$ and $y' = \psi(x,y)$, then the lines $y = y_0$ and $x = x_0$ get transformed to the increasing curves $C_1 = (\phi(x,y_0), \psi(x,y_0))$ and $C_2 = (\phi(x_0,y), \psi(x_0,y))$ respectively. Condition (2.6) means that the orientation of the curves does not change compared with the original lines, that is, the right-hand slope of C_2 at (x_0,y_0), which is $\psi_2(x_0,y_0)/\phi_2(x_0,y_0)$, is greater than or equal to the right-hand slope of C_1 at (x_0,y_0), which is $\psi_1(x_0,y_0)/\phi_1(x_0,y_0)$. This and the remainder of the proof can be easier seen via a few diagrams, which the reader is invited to supply.

Suppose (2.6) holds. For increasing functions, $\phi(x,y)$ and $\psi(x,y)$,

$$\phi(x_1,y_1) < \phi(x_2,y_2) \quad \text{and} \quad \psi(x_1,y_1) > \psi(x_2,y_2) \tag{2.7}$$

imply either

$$x_1 < x_2 \quad \text{and} \quad y_1 > y_2 \quad \text{or} \quad x_1 > x_2 \quad \text{and} \quad y_1 < y_2. \tag{2.8}$$

However assuming (2.6), (2.7) is never consistent with (2.8), so that (2.2) holds.

Now suppose that (2.6) does not hold, so that for some (x_0,y_0), $\phi_1(x_0,y_0) \cdot \psi_2(x_0,y_0) - \phi_2(x_0,y_0)\psi_1(x_0,y_0) < 0$. This means that the right-hand slope of the curve C_2 at (x_0,y_0) is less than the right-hand slope of the curve C_1 at (x_0,y_0). Then there exist $\delta, \epsilon > 0$ such that $\phi(x_0+\delta, y_0) < \phi(x_0, y_0+\epsilon)$ and $\psi(x_0+\delta, y_0) > \psi(x_0, y_0+\epsilon)$. Hence (2.2) does not hold. □

Remark. If (X',Y') is obtained from (X,Y) via a linear transform, then ϕ, ψ increasing and condition (2.6) mean that the matrix of the transform is nonnegative and has a nonnegative determinant.

If in the above theorem, we assume the stronger condition $\phi_1\psi_2 - \phi_2\psi_1 > 0$, which implies $\phi_1 > 0$ and $\psi_2 > 0$, then the transform from (x,y) to $(x',y') = (\phi(x,y), \psi(x,y))$ is one-to-one and the inverse transform has the Jacobian matrix

$$\begin{bmatrix} \phi_1 & \phi_2 \\ \psi_1 & \psi_2 \end{bmatrix}^{-1} = (\phi_1\psi_2 - \phi_2\psi_1)^{-1} \begin{bmatrix} \psi_2 & -\phi_2 \\ -\psi_1 & \phi_1 \end{bmatrix}.$$

That is, if the inverse transform is $(x,y) = (\eta(x',y'), \zeta(x',y'))$, then $\eta_1\zeta_2 - \eta_2\zeta_1 > 0$, $\eta_2 \le 0$, $\zeta_1 \le 0$.

This leads to the following more symmetric definition of a "more associated" ordering, in which there are more conditions on ϕ, ψ than in Definition 2.1. Without the conditions, results would contain too much technical details or not be true.

DEFINITION 2.4. Let $(X,Y) \sim H$, $(X',Y') \sim H'$, where H, H' are continuous bivariate distributions and $X \stackrel{d}{=} X'$, $Y \stackrel{d}{=} Y'$. Let S, S' be connected subsets of $\mathcal{R}^2$ that include the supports of H, H' respectively. Then, $H \prec^A H'$ if

(a) there exist continuous increasing functions $\phi(x,y)$, $\psi(x,y)$, from S to S', such that $(X',Y') \stackrel{d}{=} (\phi(X,Y),\ \psi(X,Y))$ and (i) $\phi_1\psi_2 - \phi_2\psi_1 > 0$ or (ii) $\phi_1 > 0$, $\psi_2 = 0$, $\phi_2 = 0$, $\psi_1 \ge 0$, or

(b) there exist continuous functions $\eta(x',y')$, $\zeta(x',y')$, from S' to S, such that $(X,Y) \stackrel{d}{=} (\eta(X',Y'), \zeta(X',Y'))$, $\eta(x',y')$ is increasing in x' for fixed y' and decreasing in y' for fixed x', $\zeta(x',y')$ is decreasing in x' for fixed y' and increasing in y' for fixed x', and (i) $\eta_1\zeta_2 - \eta_2\zeta_1 > 0$ or (ii) $\eta_1 > 0$, $\zeta_2 = 0$, $\eta_2 = 0$, $\zeta_1 \le 0$.

DEFINITION 2.5. With the same assumptions as in Definition 2.4, $H \prec^{R1} H'$ if $\phi(x,y) = x$ or $\eta(x,y) = x$ and $H \prec^{R2} H'$ if $\psi(x,y) = y$ or $\zeta(x,y) = y$.

The conditions (ii) of parts (a) and (b) of Definition 2.4 are to take care of the Fréchet upper and lower bounds ($H^+(x,y) = \min[F(x), G(y)]$ and $H^-(x,y) = \max[F(x)+G(y)-1, 0]$ respectively). Note that for the special case $x' = x$ for the $\prec^{R1}$ ordering, if the conditions of Theorem 2.1 hold, then $\zeta(x,y') = G^{-1}(G'(y' \mid x) \mid x)$, and ζ is increasing in y' and decreasing in x.

Kimeldorf and Sampson (1987) define a positive dependence ordering (PDO) as one satisfying 10 properties (P0) to (P9). The extended "more associated" ordering $\prec^A$ satisfies 9 out of 10 properties of a PDO, whereas $\prec^a$ satisfies fewer of the 10 properties. Without proof, we state the following results.

PROPOSITION 2.2. *The ordering $\prec^A$ satisfies properties* (P0) *to* (P5), (P7) *and* (P8) *of a PDO.* (P6) *is satisfied if the class of functions a is restricted to those that are continuous and strictly increasing.*

PROPOSITION 2.3. *The orderings $\prec^{R1}$ and $\prec^{R2}$ both satisfy properties* (P0) *to* (P5), *and* (P7) *of a PDO.* (P6) *is satisfied if the class of functions is restricted to those that are continuous and strictly increasing.* (P8) *is not satisfied but $(X,Y) \prec^{R1} (X',Y')$ if and only if $(Y,X) \prec^{R2} (Y',X')$. In addition,* (P9) *is almost satisfied for $\prec^{R1}$ (or $\prec^{R2}$): if H_n, H'_n, H, H' are all continuously differentiable and strictly increasing, $H_n \prec^{R1} H'_n$, and $H_n \to H$, $H'_n \to H'$ in distribution as $n \to \infty$, then $H \prec^{r1} H'$.*

The property (P1) is the concordance ordering of Yanagimoto and Okamoto (1969) and Tchen (1980). We use the notation $H \prec^c H'$ if $H(x,\infty) = H'(x,\infty)$,

$H(\infty, y) = H'(\infty, y)$ and (P1) is satisfied, that is, $H(x,y) \le H'(x,y)$ for all x, y. All of the orderings in this section imply the concordance ordering.

In Block *et al.* (1990), one result for bivariate empirical distributions is that the $\prec^A$ ordering can be handled through a "bridge" distribution by the $\prec^{R1}$ and $\prec^{R2}$ orderings. This is extended to bivariate continuous distributions below and it provides a better understanding of the $\prec^A$ ordering.

THEOREM 2.4. (a) *If $H \prec^A H'$ then there exists H^* such that $H \prec^{R2} H^*$ and $H^* \prec^{R1} H'$ and there exists H_* such that $H \prec^{R1} H_*$ and $H_* \prec^{R2} H'$.*

(b) *Suppose $H \prec^a H'$ with ϕ and ψ in Definition 2.1 being continuous. If $\phi(\cdot, y)$ is strictly increasing for all y or if $\phi_1(x,y) = 0$ implies $\psi_1(x,y) = 0$, then there exists H^* such that $H \prec^{r2} H^*$ and $H^* \prec^{r1} H'$. If $\psi(x, \cdot)$ is strictly increasing for all x or if $\psi_2(x,y) = 0$ implies $\phi_2(x,y) = 0$, then there exists H_* such that $H \prec^{r1} H_*$ and $H_* \prec^{r2} H'$.*

PROOF. (a) We will prove the first conclusion as the second is similar. Suppose $(X,Y) \sim H$, $(X',Y') \sim H'$ and $H \prec^A H'$. First assume that H is not the Fréchet lower bound and that H' is not the Fréchet upper bound. Then there exist continuous increasing functions ϕ, ψ such that $(X',Y') \stackrel{d}{=} (\phi(X,Y), \psi(X,Y))$ and the mapping is one-to-one and has a positive Jacobian. Let $(X^*, Y^*) = (\phi(X,Y), Y)$ and denote its cdf by H^*. Then $H \prec^{R2} H^*$. Let $\chi(u,y) = \inf\{t : \phi(t,y) > u\}$; that is, $\chi(\cdot, y)$ is the functional inverse of the function $\phi(\cdot, y)$ with y fixed; since $\phi(\cdot, y)$ is continuous and strictly increasing, $\chi(\cdot, y)$ has these same properties and $x = \chi(\phi(x,y), y)$. Therefore, $\psi(X,Y) \stackrel{d}{=} \psi(\chi(\phi(X,Y),Y),Y)$. Let $\psi^*(s,t) = \psi(\chi(s,t),t)$. Then

$$\begin{aligned}(X',Y') &\stackrel{d}{=} (\phi(X,Y), \psi(X,Y)) \stackrel{d}{=} (X^*, \psi(\chi(X^*,Y),Y)) \\ &\stackrel{d}{=} (X^*, \psi^*(X^*,Y)) \stackrel{d}{=} (X^*, \psi^*(X^*,Y^*)).\end{aligned}$$

Hence $H^* \prec^{R1} H'$ if $\psi^*(s,t)$ is increasing in s and t. Since $\chi(s,t)$ is increasing in s, so is $\psi^*(s,t)$. The right partial derivative of $\psi^*(s,t)$ with respect to t is $\psi_2^* = \psi_1\chi_2 + \psi_2 = \psi_1(-\phi_2/\phi_1) + \psi_2 > 0$. The identity $\chi_2 = -\phi_2/\phi_1$ follows from the identities $\phi(\chi(u,y),y) = u$ and $\chi(\phi(x,y),y) = x$.

If H' is the Fréchet upper bound or H is the Fréchet lower bound, then $H \prec^{R1} H'$ and $H \prec^{R2} H'$, by Proposition 2.3.

(b) Again because of symmetry we will prove just the first part. The above proof is almost valid with the following additional technical details. $Y \stackrel{d}{=} \psi(\chi(\phi(X,Y),Y),Y)$ holds as before with $\chi(\cdot, y)$ being the right continuous inverse of $\phi(\cdot, y)$. χ has a right derivative everywhere, $\phi(\chi(u,y),y) = u$ is still valid and $\chi(\phi(x,y),y) = x$ is satisfied unless $\phi_1(x,y) = 0$. Now, $\psi_2^* = \psi_1(-\phi_2/\phi_1) + \psi_2 \ge 0$ if $\phi_1 > 0$ and $\psi_2^* = \psi_1\chi_2 + \psi_2 = \psi_2 \ge 0$ if $\phi_1 = 0$. Therefore $H^* \prec^{r1} H'$. $\square$

3. Bivariate parametric families and examples

One objective of this section is to show that some frequently-mentioned parametric bivariate families are ordered by the $\prec^{R1}$ (or $\prec^{R2}$ or $\prec^{A}$) ordering and hence the $\prec^{c}$ ordering. At the same time, the use of the results in Section 2 is illustrated. A second objective is to have examples and counterexamples to show that certain results in Section 2 cannot be strengthened or that some conditions cannot be relaxed.

The multivariate normal distribution is widely used, probably more for convenience than for physical reasons. The families here that are $\prec^{R1}$ ordered enjoy dependence properties similar to the bivariate normal family, and because of this, could be competing models to the bivariate normal family for bivariate data, such as lifetime or reliability data.

Example 3.1. (Clayton (1978), Cook and Johnson (1981) and Oakes (1982)) This family has had several applications. We write it in the form

$$H(x,y;\theta)=[F(x)^{-\theta}+G(y)^{-\theta}-1]^{-1/\theta},\qquad \theta\ge 0. \tag{3.1}$$

The case $\theta=0$ corresponds to independence and the other limit $\theta=\infty$ corresponds to the Fréchet upper bound. If $0<\theta_1<\theta_2<\infty$, then

$$\begin{aligned}\psi(x,y;\theta_1,\theta_2)&=G^{-1}(G(y\mid x;\theta_1)\mid x;\theta_2)\\&=G^{-1}(\{[\{(F(x)^{-\theta_1}+G(y)^{-\theta_1}-1)^{-1-1/\theta_1}\\&\qquad\cdot F(x)^{-1-\theta_1}\}^{-\theta_2/(\theta_2+1)}-1]F(x)^{-\theta_2}+1\}^{-1/\theta_2}).\end{aligned}$$

A sketch of the proof that $\psi(x,y;\theta_1,\theta_2)$ is increasing in x is as follows: let $\lambda=\theta_2/\theta_1$, $\alpha=\theta_2(\theta_1+1)/[\theta_1(\theta_2+1)]$, $w=G(y)^{-\theta_1}-1$, $z=F(x)^{-\theta_1}$; then $r(z)=[(1+w/z)^{\alpha}-1]z^{\lambda}$ is increasing in z for fixed w, since $dr(z)/dz\ge 0$. By Theorem 2.1, $H(\cdot;\theta)$ is increasing with respect to $\prec^{R1}$ as θ increases. Since $H(F^{-1}(u),G^{-1}(v);\theta)=H(G^{-1}(v),F^{-1}(u);\theta)$, the last statement is also valid with $\prec^{R1}$ replaced by $\prec^{R2}$.

Example 3.2. (Frank (1979), Genest (1987)) We write Frank's family in a different parametrization than previous authors:

$$H(x,y;\theta)=-\frac{1}{\theta}\log\left[1-\frac{(1-e^{-\theta F(x)})(1-e^{-\theta G(y)})}{(1-e^{-\theta})}\right],\qquad -\infty\le\theta\le\infty. \tag{3.2}$$

The case $\theta=0$ corresponds to independence and the limits $\theta=-\infty,\infty$ correspond to the Fréchet lower and upper bounds respectively. If $-\infty<\theta_1<\theta_2<\infty$, $\theta_1,\theta_2\ne 1$, then

$$\begin{aligned}\psi(x,y;\theta_1,\theta_2)&=G^{-1}(G(y\mid x;\theta_1)\mid x;\theta_2)\\&=G^{-1}\left(-\frac{1}{\theta_2}\log\left\{1-\frac{1-e^{-\theta_2}}{(u^{-1}-1)e^{-\theta_2 F(x)}+1}\right\}\right),\end{aligned}$$

where $u = G(y \mid x;\theta_1) = [1 - e^{-\theta_1} - (1 - e^{-\theta_1 F(x)})(1 - e^{-\theta_1 G(y)})]^{-1} e^{-\theta_1 F(x)}(1 - e^{-\theta_1 G(y)})$. $\psi(x, y; \theta_1, \theta_2)$ is increasing in x since $(u^{-1} - 1)e^{-\theta_2 F(x)} = e^{(\theta_1 - \theta_2)F(x)}(1 - e^{-\theta_1 G(y)})^{-1}(e^{-\theta_1 G(y)} - e^{-\theta_1})$ is decreasing in x for fixed y. Again, by Theorem 2.1, $H(x, y; \theta)$ is increasing in the $\prec^{R1}$ ordering as θ increases.

Example 3.3. (Gumbel (1961)) This family and also (3.1) lead to families of bivariate extreme value distributions (Joe (1990*a*)). Let

$$H(x, y; \theta) = \exp\{-[(-\log F(x))^\theta + (-\log G(y))^\theta]^{1/\theta}\}, \qquad \theta \geq 1. \tag{3.3}$$

The case $\theta = 0$ corresponds to independence and the other limit $\theta = \infty$ corresponds to the Fréchet upper bound. If $1 < \theta_1 < \theta_2 < \infty$, and y are fixed, then $\psi(x, y; \theta_1, \theta_2) = G^{-1}(G(y \mid x; \theta_1) \mid x; \theta_2) = y_2(x)$, where $y_2 = y_2(x)$ or $w = w(z) = w(z, \theta_2) = w(\theta_2)$ is the root of

$$\begin{aligned}(\theta_2^{-1} - 1)\log(z^{\theta_2} + w) - (z^{\theta_2} + w)^{1/\theta_2} - (\theta_1^{-1} - 1)\log(z^{\theta_1} + v)\\ + (z^{\theta_1} + v)^{1/\theta_1} + (\theta_2 - \theta_1)\log z = \beta(z, w(z)) = 0,\end{aligned}$$

with $z = -\log F(x)$, $w = (-\log G(y_2(x)))^{\theta_2}$, $v = (-\log G(y))^{\theta_1}$. This example is more complex than the previous two because $G^{-1}(y \mid x; \theta)$ does not have an explicit form. A sketch of the proof that $\psi(x, y; \theta_1, \theta_2)$ is increasing in x is as follows: It suffices to show that $\partial w/\partial z \geq 0$; from $\partial\beta/\partial z = 0$,

$$\begin{aligned}&\frac{\partial w(z)}{\partial z}\frac{z}{\theta_2}\{(z^{\theta_2} + w)^{-1}[(z^{\theta_2} + w)^{1/\theta_2} + \theta_2 - 1]\}\\ &\quad = -\frac{z^{\theta_2}}{z^{\theta_2} + w}[(z^{\theta_2} + w)^{1/\theta_2} + \theta_2 - 1]\\ &\qquad - \frac{z^{\theta_1}}{z^{\theta_1} + v}[(z^{\theta_1} + v)^{1/\theta_1} + \theta_1 - 1] + (\theta_2 - \theta_1).\end{aligned}$$

Since $w \to v$ as $\theta_2 \to \theta_1$, it is enough to show for θ_1, v, z fixed that $-z^{\theta_2}[(z^{\theta_2} + w(\theta_2))^{1/\theta_2} + \theta_2 - 1]/(z^{\theta_2} + w(\theta_2)) + \theta_2$ is increasing in $\theta_2 > \theta_1$. Let $t = t(\theta) = (z^\theta + w)^{1/\theta}$, $s = (z^{\theta_1} + v)^{1/\theta_1}$ and $r(\theta) = -(z/t)^\theta(t + \theta - 1) + \theta$, $\theta \geq \theta_1$. Then $\partial r/\partial\theta = 1 - (z/t)^\theta - (z/t)^\theta c(t)\log(z/t)$, where $c(t) = t - (\theta - 1)/(t + \theta - 1)$ is increasing in t and $c(0) = -1$. By definition, $u = z/t \in [0, 1]$, so $\partial r/\partial\theta \geq 0$ if $a(u) = 1 - u^\theta - cu^\theta \log u \geq 0$ for all $0 \leq u \leq 1$, $\theta \geq 1$, $c \geq -1$. Note that $a(0) = 1$, $a(1) = 0$ and $(da/du)(u_0) = 0$ implies $-\log u_0 = (c + \theta)/(c\theta)$ is the only root; $-\log u_0 > 0$ only if $c > 0$. Therefore, for $c > 0$, $a(u)$ increases to a peak at u_0 and then decreases; for $c = 0$, $a(u) = 1 - u^\theta$ decreases, and for $-1 \leq c < 0$, $a(u)$ also decreases.

Example 3.4. Let $C(u, v; \theta)$ be a family of copulas and F, G be strictly increasing and continuous univariate cdf's. Then $C(F, G; \theta)$ is a family of bivariate cdf's with margins F, G and $C(1 - F, 1 - G, \theta)$ is a family of survival functions with margins F, G. That is,

$$K(x, y; \theta) = F(x) + G(y) - 1 + C(1 - F(x), 1 - G(y); \theta) \tag{3.4}$$

is another family of bivariate cdf's based on the family $C(u, v; \theta)$. If $C(F, G; \theta)$ is ordered by $\prec^{R1}$, then (3.4) is also ordered by $\prec^{R1}$. Similar results hold with $\prec^{R1}$ replaced by $\prec^{R2}$ or $\prec^{A}$. This result applies to the copulas implicit in (3.1), (3.2) and (3.3). The proof is straightforward and is omitted.

Example 3.5. (subclass of bivariate stable distributions, Press (1972), Paulauskas (1976)) This example differs from the previous ones in that the dependence is indexed by two parameters and the $\prec^{A}$ ordering corresponds to a partial ordering among these two parameters. Let $\alpha \in (0, 2]$ be a fixed index of the stable law. Let V, W be independent standard symmetric stable random variables with index α (characteristic function is $\chi(s) = \exp\{-|s|^{\alpha}\}$). Let $\beta, \gamma \in [(.5)^{1/\alpha}, 1]$ and let $\bar{u} = (1 - u^{\alpha})^{1/\alpha}$ if $u \in [0, 1]$. Let $(X, Y) = (X_{\beta}, Y_{\gamma}) = (\beta V + \bar{\beta} W, \bar{\gamma} V + \gamma W)$. The cdf of (X, Y) does not have a closed form but the characteristic function is $\chi(s, t; \beta, \gamma) = \exp\{-[|s\beta + t\bar{\beta}|^{\alpha} + |s\bar{\gamma} + t\gamma|^{\alpha}]\}$. This family reduces to a one-parameter family only if $\alpha = 2$ (bivariate normal). $(\beta, \gamma) = (1, 1)$ corresponds to independence and $(\beta, \gamma) = ((.5)^{1/\alpha}, (.5)^{1/\alpha})$ corresponds to the Fréchet upper bound. If $1 \geq \beta_1 \geq \beta_2 > (.5)^{1/\alpha}$ and $1 \geq \gamma_1 \geq \gamma_2 > (.5)^{1/\alpha}$, then

$$\begin{bmatrix} X_{\beta_2} \\ Y_{\gamma_2} \end{bmatrix} = (\beta_1\gamma_1 - \bar{\beta}_1\bar{\gamma}_1)^{-1} \begin{bmatrix} \beta_2\gamma_1 - \bar{\beta}_2\bar{\gamma}_1 & \beta_1\bar{\beta}_2 - \beta_2\bar{\beta}_1 \\ \gamma_1\bar{\gamma}_2 - \gamma_2\bar{\gamma}_1 & \beta_1\gamma_2 - \bar{\beta}_1\bar{\gamma}_2 \end{bmatrix} \begin{bmatrix} X_{\beta_1} \\ Y_{\gamma_1} \end{bmatrix}.$$

All of the entries of the matrix of transformation are nonnegative and the determinant is positive so that by Definition 2.4, $(X_{\beta_1}, Y_{\gamma_1}) \prec^{A} (X_{\beta_2}, Y_{\gamma_2})$. The measure of association given in Press (1972) and Paulauskas (1976) does not separate these bivariate distributions as well because it is one-dimensional.

Example 3.6. (linear combinations) In this example, we have two stable random variables that are each linear combinations of more than 2 symmetric stable random variables (with the same index α). The following theorem will be used.

THEOREM 3.1. *Let $\boldsymbol{Z} = (Z_1, \ldots, Z_k)^T$, where $k \geq 2$ and the superscript T denotes a transpose. Suppose that the support of $\boldsymbol{Z}$ includes a rectangle in $\mathcal{R}^k$. Let $(X', Y')^T = A\boldsymbol{Z}$ and $(X, Y)^T = B\boldsymbol{Z}$, where A, B are $2 \times k$ matrices and A, B, $\boldsymbol{Z}$ are such that $X \overset{d}{=} X'$, $Y \overset{d}{=} Y'$. If there exists a nonnegative 2×2 matrix C with positive determinant such that $A = CB$, then $(X, Y) \prec^{A} (X', Y')$. If B has full row rank, then candidates for C have the form AB^-, where B^- is a generalized inverse (BB^- equals the identity matrix of order 2).*

PROOF. $(X', Y')^T = A\boldsymbol{Z} = CB\boldsymbol{Z} = C(X, Y)^T$ implies $A = CB$ since $A\boldsymbol{Z} = CB\boldsymbol{Z}$ for all $\boldsymbol{Z}$ in a neighbourhood of some point in $\mathcal{R}^k$ implies $A = CB$. The $\prec^{A}$ ordering is satisfied only if C is nonnegative and $|C| > 0$. If $BB^- = I_2$, then $AB^- = CBB^- = C$. For such a B^-, $BB^-B = B$, so that by a result on p. 24 of Rao (1973), $AB^-B = A$. □

Example 3.6. (continued) Suppose Z_1, Z_2, Z_3 are independent standard symmetric stable random variables of index α. Let $(X_{\theta}, Y_{\theta})^T = A_{\theta}\boldsymbol{Z}$, $0 < \theta <$

0.5, where $A_\theta = a(\theta)\begin{bmatrix} 1 & 1-\theta & \theta \\ 1 & \theta & 1-\theta \end{bmatrix}$, $a(\theta) = [1+\theta^\alpha+(1-\theta)^\alpha]^{-1}$. Let $0<\theta_1<\theta_2<0.5$. Then $(X_{\theta_2},Y_{\theta_2})^T = C(X_{\theta_1},Y_{\theta_1})^T$, with $C = a(\theta_2)[a(\theta_1)(1-2\theta_1)]^{-1}\begin{bmatrix} 1-\theta_1-\theta_2 & \theta_2-\theta_1 \\ \theta_2-\theta_1 & 1-\theta_1-\theta_2 \end{bmatrix}$; this is based on the choice of $A_\theta^- = [a(\theta)(1-2\theta)]^{-1}\begin{bmatrix} 0 & 1-\theta & -\theta \\ 0 & -\theta & 1-\theta \end{bmatrix}^T$. Since the matrix C is nonnegative and has determinant being a positive constant times $(1-\theta_1-\theta_2)^2-(\theta_2-\theta_1)^2 = (1-2\theta_1)(1-2\theta_2)$, $(X_{\theta_1},Y_{\theta_1}) \prec^A (X_{\theta_2},Y_{\theta_2})$.

Schriever (1986) has essentially this example with $A_\theta = a(\theta)\begin{bmatrix} 1 & 0 & \theta \\ 0 & 1 & \theta \end{bmatrix}$, $a(\theta) = [1+\theta^\alpha]^{-1}$. In this case, there is no solution C to $A_{\theta_2} = CA_{\theta_1}$ for $\theta_2>\theta_1>0$, and the resulting (X_θ,Y_θ) are not $\prec^A$ ordered.

Example 3.7. (bivariate exponential distributions with singular component) This example consists of a family that is $\prec^{r1}$ ordered but not $\prec^{R1}$ ordered because the mapping from (x,y) to (x',y') via the function ψ in Definition 2.1 is not one-to-one. The representation in Lemma 2.1 and Theorem 2.1 is still valid even though the conditional distributions $G(y \mid x)$ are not continuous in y for a fixed x; see the remark following Lemma 2.1. In this example, $\bar H$ and $\bar G(\cdot \mid x)$ are survival functions.

Let X, Y be two independent exponential random variables with mean 1. For $0\le\lambda\le 1$, let $(X_\lambda,Y_\lambda) = (X,\min\{X/\lambda, Y/(1-\lambda)\})$. The case $\lambda=0$ corresponds to independence and the case $\lambda=1$ corresponds to the Fréchet upper bound. The survival function is

$$\bar H(x,y;\lambda) = \exp\{-\max[x,\lambda y]-(1-\lambda)y\}, \qquad x,y\ge 0.$$

For $0\le\lambda_1<\lambda_2\le 1$, $(X_{\lambda_2},Y_{\lambda_2}) = (X_{\lambda_1},\min\{X_{\lambda_1}/\lambda_2,(1-\lambda_1)Y_{\lambda_1}/(1-\lambda_2)\}) = (X_{\lambda_1},\psi(X_{\lambda_1},Y_{\lambda_1};\lambda_1,\lambda_2))$, with $\psi(x,y;\lambda_1,\lambda_2)$ increasing in x and y, so that $H(\cdot;\lambda_1) \prec^{r1} H(\cdot;\lambda_2)$. Also it can be directly verified that $\psi(x,y;\lambda_1,\lambda_2) = \bar G^{-1}(\bar G(y \mid x;\lambda_1) \mid x;\lambda_2)$.

In the next example, we show what can happen when the continuity assumptions on the conditional distributions in Lemma 2.1 and Theorem 2.1 are not satisfied. The following lemma is needed.

LEMMA 3.1. *Let Y, Y' be random variables such that $Y'=\psi(Y)$ for an increasing function ψ. Suppose $Y\sim G$ and $Y'\sim G'$. If ψ is strictly increasing at z, then $G'^{-1}(G(z)^-)\le\psi(z)\le G'^{-1}(G(z)^+)$. If ψ is flat at z, say on the interval (z_1,z_2) containing z, then $G'^{-1}(G(z_1)^-)\le\psi(z_1^+)=\psi(z_2^-)\le G'^{-1}(G(z_2)^+)$.*

PROOF. $\Pr(Y\le z)\le\Pr(Y'\le\psi(z))\le\Pr(Y\le\psi^{-1}(\psi(z)^+))$, where ψ^{-1} is the inverse of ψ. Therefore $G(z)\le G'(\psi(z))\le G(\psi^{-1}(\psi(z)^+))$ and $G'^{-1}(G(z)^-)\le\psi(z)\le G'^{-1}(G(\psi^{-1}(\psi(z)^+))^+)$. □

Example 3.8. (bivariate exponential, Marshall and Olkin (1967)) A subfamily of the Marshall-Olkin bivariate exponential survival functions is

$$\bar{H}(x,y;\lambda) = \exp\{-(1-\lambda)(x+y) - \lambda \max[x,y]\}, \qquad x,y \geq 0, \quad 0 \leq \lambda \leq 1. \tag{3.5}$$

$\lambda = 0$ corresponds to independence and $\lambda = 1$ corresponds to the Fréchet upper bound. It is straightforward to show that $H(\cdot;\lambda)$ is increasing with respect to $\prec^c$ as λ increases. We will show, for $0 < \lambda_1 < \lambda_2 < 1$, that $H(\cdot;\lambda_1)$ and $H(\cdot;\lambda_2)$ are not $\prec^{r1}$ ordered. Let $(X,Y) \sim H(\cdot;\lambda_1)$ and let $(X',Y') \sim H(\cdot;\lambda_2)$ with $X = X'$.

The conditional survival function given x is

$$1 - G(y \mid x;\lambda) = \begin{cases} e^{-(1-\lambda)y}, & \text{if } \ y < x, \\ (1-\lambda)e^{-y+\lambda x}, & \text{if } \ y \geq x. \end{cases}$$

This has a jump discontinuity at x for all $\lambda > 0$ and $(G(x^- \mid x;\lambda), G(x \mid x;\lambda)) = (1 - e^{-(1-\lambda)x}, 1 - (1-\lambda)e^{-(1-\lambda)x}) = (L(\lambda), U(\lambda))$, say. $L(\lambda)$ is decreasing in λ, and $dU(\lambda)/d\lambda$ is positive if $x < (1-\lambda)^{-1}$ and negative if $x > (1-\lambda)^{-1}$. For fixed $0 < \lambda_1 < \lambda_2 < 1$, $(L(\lambda_1), U(\lambda_1))$ is not nested in $(L(\lambda_2), U(\lambda_2))$ if $e^{(\lambda_2-\lambda_1)x} > (1-\lambda_1)/(1-\lambda_2)$. For a fixed x in this region, we will show that there is no increasing function $\psi(x,y)$ such that conditional on $X = x$, $Y' \overset{d}{=} \psi(x,Y)$. Let $G(y) = G(y \mid x;\lambda_1)$ and let $G'(y) = G(y \mid x;\lambda_2)$. Suppose $Y \sim G$, $Y' \sim G'$ and $Y' = \psi(Y)$, where ψ is increasing (reference to x has been suppressed). Since Y, Y' have support on $[0,\infty)$, each with a mass only at x, $\psi(x) = x$. $\psi(y)$ must be strictly increasing outside of a neighbourhood $[x_1, x_2]$ of x, where x_1, x_2 satisfy $G(x_2) - G(x_1^-) = G'(x) - G'(x^-)$. By Lemma 3.1, $\psi(y) = G'^{-1}(G(y))$, for $y \notin [x_1, x_2]$. However $G'^{-1}(G(x_2)) \geq G'^{-1}(G(x)) > x$ which implies that ψ does not take values from $(x, G'^{-1}(G(x_2)))$. This contradicts the assumption that $Y' = \psi(Y)$ and ψ is increasing.

Example 3.9. (Morgenstern (1956), Farlie (1960)) This last example illustrates the use of Proposition 2.1 and Theorem 2.2 to show the $\prec^{r1}$ ordering. The family below is probably not very statistically useful because the distributions are perturbations of a bivariate distribution representing independence, so its range of "dependence" is smaller than the families in (3.1), (3.2) and (3.3). For functions $A(x)$, $B(y)$ on the interval $[0,1]$ with continuous first derivatives and satisfying $A(1) = B(1) = 0$, consider the family

$$H(x,y;\alpha) = xy[1 + \alpha A(x)B(y)], \qquad 0 \leq x,y \leq 1, \quad \alpha_- \leq \alpha \leq \alpha_+, \tag{3.6}$$

where α_-, α_+ are defined by the requirement that $|P(x)Q(y)||\alpha| \leq 1$, with $P(x) = (d/dx)(xA(x))$ and $Q(y) = (d/dy)(yB(y))$. The conditional distributions given x and y are respectively $G(y \mid x;\alpha) = y(1 + \alpha P(x)B(y))$ and $F(x \mid y;\alpha) = x(1+\alpha A(x)Q(y))$. Analytic forms for the inverses of these are not always possible. The following proposition gives condition under which (3.6) is $\prec^{r1}$ or $\prec^{r2}$ ordered.

PROPOSITION 3.1. *Let* $\alpha_- \leq \alpha_1 < \alpha_2 \leq \alpha_+$. (a) $H(\cdot;\alpha_1) \prec^{r1} H(\cdot;\alpha_2)$ *if* $P(x)$ *is decreasing and* $B(y)$ *is nonnegative or* $P(x)$ *is increasing and* $B(y)$ *is* *nonpositive. If these conditions are not satisfied, then* $H(\cdot;\alpha)$ *and* $H(\cdot;0)$ *are not* $\prec^{r1}$ *ordered if* $\alpha > 0$. (b) $H(\cdot;\alpha_1) \prec^{r2} H(\cdot;\alpha_2)$ *if* $Q(x)$ *is decreasing and* $A(y)$ *is nonnegative or* $Q(x)$ *is increasing and* $A(y)$ *is nonpositive. If these conditions are not satisfied, then* $H(\cdot;\alpha)$ *and* $H(\cdot;0)$ *are not* $\prec^{r2}$ *ordered if* $\alpha > 0$.

PROOF. Because of the symmetry, we prove only (a) under the condition $P(x)$ decreasing and $B(y)$ nonnegative. Let $0 \leq y, y' \leq 1$, $\Delta = \alpha_1 y B(y) - \alpha_2 y' B(y')$ and $\Gamma(x) = y'B(y')P(x)(\alpha_1 - \alpha_2)$. It is straightforward to show that

$$G(y \mid x;\alpha_1) - G(y' \mid x;\alpha_2) = y - y' + P(x)\Delta \tag{3.7}$$

and

$$G(y \mid x;\alpha_1) - G(y' \mid x;\alpha_2) = G(y \mid x;\alpha_1) - G(y' \mid x;\alpha_1) + \Gamma(x). \tag{3.8}$$

We check (2.5) in two cases, that is, assuming y, y', x_1 are such that $G(y \mid x_1;\alpha_1) \geq G(y' \mid x_1;\alpha_2)$, we want to show that $G(y \mid x_2;\alpha_1) \geq G(y' \mid x_2;\alpha_2)$ for $x_2 > x_1$.

Case (1) $y \geq y'$: If $P(x_2) < 0$, then $\Gamma(x_2) \geq 0$ and this implies $G(y \mid x_2;\alpha_1) - G(y' \mid x_2;\alpha_2) \geq 0$ by (3.8). If $P(x_2) \geq 0$ and $\Delta > 0$, $G(y \mid x_2;\alpha_1) - G(y' \mid x_2;\alpha_2) \geq 0$ by (3.7). If $P(x_2) \geq 0$ and $\Delta \leq 0$, the monotonicity of P implies $P(x_2)\Delta \geq P(x_1)\Delta$ and this leads to

$$\begin{aligned} G(y \mid x_2;\alpha_1) - G(y' \mid x_2;\alpha_2) &= y - y' + P(x_2)\Delta \geq y - y' + P(x_1)\Delta \\ &= G(y \mid x_1;\alpha_1) - G(y \mid x_1;\alpha_2) \geq 0. \end{aligned} \tag{3.9}$$

Case (2) $y < y'$: (3.8) $\Rightarrow \Gamma(x_1) \geq 0 \Rightarrow P(x_1) \leq 0 \Rightarrow P(x_2) \leq P(x_1) \leq 0 \Rightarrow \Delta \leq 0$ by (3.7) $\Rightarrow P(x_2)\Delta \geq P(x_1)\Delta$. Now (3.9) obtains again.

For the last part of (a), $H(x,y;0) \prec^{r1} H(x,y;\alpha)$ if and only if $G^{-1}(y \mid x;\alpha)$ is increasing in x or $G(y \mid x;\alpha)$ is decreasing in x. If $B(y)$ is nonnegative, the latter condition holds if and only if $P(x)$ is decreasing. □

Example 3.9. (continued) If $A(x) = x(1-x)$ and $B(y) = 1-y$. Then $P(x) = x^2 - x^3$ is not monotone and the family (3.6) is not $\prec^{r1}$ ordered. However $Q(y) = 1 - 2y$ is decreasing so (3.6) is $\prec^{r2}$ ordered and hence $\prec^{a}$ ordered as α increases.

4. Functionals preserving the orderings

In this section, we consider functionals that preserve the $\prec^{A}$, $\prec^{R1}$ and $\prec^{R2}$ orderings. The functionals lead to inequalities and measures of dependence, and also provide a way of showing that two distributions which are ordered by $\prec^{c}$ may not be ordered by $\prec^{A}$. Note that there is no simple way of proving or disproving that two distributions are $\prec^{A}$ ordered, when it is known that they are $\prec^{c}$ ordered but not $\prec^{R1}$ ordered or $\prec^{R2}$ ordered. We combine results from Section 2 and Block *et al.* (1990) to study a particular class of functionals on continuous bivariate distributions. Use is made of a coupling argument, which is natural from the definitions of "more associated" and "more regression dependent", but which is not possible for the concordance ordering.

Some notations are: (1) a bold-faced symbol denotes an n-vector while the same symbol without a bold-face and with a subscript indicates a component of the vector, for example, $\boldsymbol{a} = (a_1, \ldots, a_n)$, (2) $I(A)$ is the indicator function of the set A.

Let H and H' be such that $H \prec^A H'$ (or $H \prec^{R1} H'$ or $H \prec^{R2} H'$) via increasing functions $\phi(x,y)$ and $\psi(x,y)$. Let (X_i, Y_i), $i = 1, \ldots, n$, be a random sample of size n from H; since H is continuous, we can assume that the X_i's are distinct and the Y_i's are distinct. Let $(X_i', Y_i') = (\phi(X_i, Y_i), \psi(X_i, Y_i))$. Since ϕ is strictly increasing in x and ψ is strictly increasing in y, the X_i''s are distinct and so are the Y_i''s.

For n pairs (a_i, b_i), with the a_i's distinct and the b_i's distinct, a permutation $\boldsymbol{J}(\boldsymbol{a}, \boldsymbol{b})$ of $\{1, \ldots, n\}$ is defined as follows. Let $h_1, \ldots, h_n$ be such that $a_{h_1} < \cdots < a_{h_n}$. Define $J_i(\boldsymbol{a}, \boldsymbol{b})$ to be the rank of b_{h_i} among $b_1, \ldots, b_n$, with a rank of 1 meaning the smallest. For the coupled random samples, let $\boldsymbol{j} = \boldsymbol{J}(\boldsymbol{X}, \boldsymbol{Y})$ and $\boldsymbol{j}' = \boldsymbol{J}(\boldsymbol{X}', \boldsymbol{Y}')$. The orderings $\prec^A$, $\prec^{R1}$, $\prec^{R2}$ imply orderings for the permutations $\boldsymbol{j}$, $\boldsymbol{j}'$. We state the necessary definitions and results from Block *et al.* (1990).

DEFINITION 4.1. Let $\boldsymbol{i} = (i_1, \ldots, i_n)$ be a permutation of $\{1, \ldots, n\}$. For k, l distinct integers in $\{1, \ldots, n\}$, let $\Delta_{kl} = (k - l)(i_k - i_l)$. An interchange of i_k and i_l is said to be a correction of an inversion of type 1 if $\Delta_{kl} < 0$, of type 2 if $\Delta_{kl} < 0$ and $|i_k - i_l| = 1$ and of type 3 if $\Delta_{kl} < 0$ and $|k - l| = 1$.

DEFINITION 4.2. $\boldsymbol{i}'$ is said to be better ordered than $\boldsymbol{i}$ in the sense of the ordering b_t, $t = 1, 2, 3$, written $\boldsymbol{i} \prec^{b_t} \boldsymbol{i}'$, if $\boldsymbol{i} = \boldsymbol{i}'$ or $\boldsymbol{i}'$ is obtainable from $\boldsymbol{i}$ in a finite number of steps, each of which consists of correcting an inversion of type t. $\boldsymbol{i}'$ is said to be better ordered than $\boldsymbol{i}$ in the sense of the ordering b_4, written $\boldsymbol{i} \prec^{b_4} \boldsymbol{i}'$, if $\boldsymbol{i} = \boldsymbol{i}'$ or $\boldsymbol{i}'$ is obtainable from $\boldsymbol{i}$ in a finite number of steps, each of which consists of correcting an inversion of type 2 or of type 3.

The implications among these orderings is that both $\boldsymbol{i} \prec^{b_2} \boldsymbol{i}'$ and $\boldsymbol{i} \prec^{b_3} \boldsymbol{i}'$ imply $\boldsymbol{i} \prec^{b_4} \boldsymbol{i}'$ and $\boldsymbol{i} \prec^{b_4} \boldsymbol{i}'$ implies $\boldsymbol{i} \prec^{b_1} \boldsymbol{i}'$. The theorem below follows from Theorems 3.7 and 3.8 of Block *et al.* (1990) and Theorem 4.1 of Schriever (1987).

THEOREM 4.1. *Let H, H', $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{X}'$, $\boldsymbol{Y}'$, $\boldsymbol{j}$, $\boldsymbol{j}'$ be defined as above.* (a) *If $H \prec^{R1} H'$, then $\boldsymbol{j} \prec^{b_2} \boldsymbol{j}'$.* (b) *If $H \prec^{R2} H'$, then $\boldsymbol{j} \prec^{b_3} \boldsymbol{j}'$.* (c) *If $H \prec^A H'$, then $\boldsymbol{j} \prec^{b_4} \boldsymbol{j}'$.*

Now consider functionals on permutations of $\{1, \ldots, n\}$ that are increasing with respect to $\prec^{b_1}, \prec^{b_2}, \prec^{b_3}$ or $\prec^{b_4}$. Note that the relationships among the orderings mean that a functional increasing with respect to $\prec^{b_1}$ is also increasing relative to the other 3 orderings. In some cases, by Theorem 4.3, there will be corresponding functionals on continuous bivariate distributions that are increasing relative to $\prec^{R1}, \prec^{R2}$ or $\prec^A$. We will focus on $\prec^{b_4}$ and $\prec^A$. Let T_n be a functional on permutations of order n, then T_n is increasing with respect to $\prec^{b_4}$ if and only if it is increasing with respect to both $\prec^{b_2}$ and $\prec^{b_3}$. It is enough to compare T_n for $\boldsymbol{i}$ and $\boldsymbol{i}'$ which differ by a type 2 or a type 3 inversion. An obvious example is $T_{1n}(\boldsymbol{i}) = 2[n(n-1)]^{-1} \sum_{1 \le k < l \le n} [I(i_k < i_l) - I(i_k > i_l)]$. This is also increasing with respect to $\prec^{b_1}$. We are interested in T_n that are increasing with respect to $\prec^{b_4}$ but not $\prec^{b_1}$, as this can lead to a functional increasing in $\prec^A$ but not $\prec^c$. An example is given by T_{2n} in the theorem below.

THEOREM 4.2. *For $n \geq 3$, let $T_{2n}(\boldsymbol{i}) = 6[n(n-1)(n-2)]^{-1} \sum_{1\leq k<l<m\leq n}[I(i_k < i_l < i_m) - I(i_k > i_l > i_m)]$. Then $T_{2n}(\boldsymbol{i}) \leq T_{2n}(\boldsymbol{i}')$ if $\boldsymbol{i} \prec^{b_4} \boldsymbol{i}'$ but not necessary if $\boldsymbol{i} \prec^{b_1} \boldsymbol{i}'$ and $n \geq 4$. The inequality $-1 \leq T_{2n}(\boldsymbol{i}) \leq 1$ holds for all $\boldsymbol{i}$ with equality at the lower and upper bounds if $\boldsymbol{i}$ satisfies $i_k = n-k+1$ and $i_k = k$ respectively.*

PROOF. For $\prec^{b_4}$, it is straightforward to enumerate all patterns for which $\boldsymbol{i}$ and $\boldsymbol{i}'$ differ just by an inversion of type 2 or one of type 3. For $\prec^{b_1}$, consider $\boldsymbol{i} = (3,4,1,2)$ and $\boldsymbol{i}' = (3,2,1,4)$. Then $T_{2n}(\boldsymbol{i}) = 0$ and $T_{2n}(\boldsymbol{i}') = -1/4$. The last statement is obvious. □

THEOREM 4.3. *Let (X_i, Y_i), $i = 1,\ldots,n$, be a random sample of size n from the continuous distribution H. Then, as $n \to \infty$, $T_{2n}(\boldsymbol{J}(\boldsymbol{X},\boldsymbol{Y}))$ converges almost surely to* $\Pr(Y_1 < Y_2 < Y_3 \mid X_1 < X_2 < X_3) - \Pr(Y_1 > Y_2 > Y_3 \mid X_1 < X_2 < X_3) = \Pr((X_1, X_2, X_3)$ *and* (Y_1, Y_2, Y_3) *are similarly ordered*$) - \Pr((X_1, X_2, X_3)$ *and* (Y_1, Y_2, Y_3) *are oppositely ordered*$)$.

PROOF. T_{2n} is equivalent to a bounded two-sample U-statistic. See, for example, Chapter 5 of Serfling (1980). □

THEOREM 4.4. *Let (X_i, Y_i), $i = 1,2,3$, be a random sample of size 3 from the continuous distribution H. Then the functional $\tau_2(H) = \Pr(Y_1 < Y_2 < Y_3 \mid X_1 < X_2 < X_3) - \Pr(Y_1 > Y_2 > Y_3 \mid X_1 < X_2 < X_3)$ is increasing with respect to $\prec^A$ but not necessarily for $\prec^c$.*

PROOF. Let H, H' be such that $H \prec^A H'$. Let (X_i, Y_i), (X_i', Y_i'), $i = 1,\ldots,n$, be the coupled random samples of size n. By Theorem 4.1, $\boldsymbol{j} = \boldsymbol{J}(\boldsymbol{X},\boldsymbol{Y}) \prec^{b_4} \boldsymbol{j}' = \boldsymbol{J}(\boldsymbol{X}',\boldsymbol{Y}')$. Therefore, $T_{2n}(\boldsymbol{j}) \leq T_{2n}(\boldsymbol{j}')$ by Theorem 4.2. By Theorem 4.3, $\tau_2(H) \leq \tau_2(H')$.

An example of H, H' with $H \prec^c H'$, $\tau_2(H) > \tau_2(H')$ is the following continuous version of the example in the proof of Theorem 4.2. Let H have support on the union of the four squares, A_1, A_2, A_3, A_4, where $A_1 = [0,.25] \times [.5,.75]$, $A_2 = [.25,.5] \times [.75,1]$, $A_3 = [.5,.75] \times [0,.25]$ and $A_4 = [.75,1] \times [.25,.5]$; H is conditionally uniform on each of the four squares with total probabilities of 1/3, 1/6, 1/3, 1/6 respectively on A_1 to A_4. H' is similar to H but with A_2 replaced by $A_2' = [.25,.5]^2$ and A_4 replaced by $A_4' = [.75,1]^2$. It can be checked that $H \prec^c H'$ and that $\tau_2(H) = -1/8 > -5/24 = \tau_2(H')$. □

COROLLARY 4.1. *If H, H' are such that $H \prec^c H'$ and $\tau_2(H) > \tau_2(H')$, then H, H' are not $\prec^A$ ordered.*

From Theorem 4.4, inequalities can be obtained for the families in Section 3. For a family such that $H(x,y;\theta)$ is increasing relative to $\prec^A$ as θ increases,

$\tau_2(H(\cdot;\theta_1)) \le \tau_2(H(\cdot;\theta_2))$ if $\theta_1 < \theta_2$. This result would not be obtainable analytically (that is, by expressing $\tau_2(H)$ as a sum of integrals).

The use of a statistic like $T_{2n}(\boldsymbol{i})$ leading to some functional $\tau(H)$ is a way of trying to show that H, H' are not $\prec^A$ ordered when it is known that they are $\prec^c$ ordered but not $\prec^{R1}$ ordered or $\prec^{R2}$ ordered. The statistic $T_{1n}(\boldsymbol{J}(\boldsymbol{X},\boldsymbol{Y}))$ converges to Kendall's tau, a well-known measure of bivariate concordance or monotone association. $T_{2n}(\boldsymbol{J}(\boldsymbol{X},\boldsymbol{Y}))$ can also be considered as a measure of monotone association. $T_{2n}(\boldsymbol{i})$ can be generalized to $T_{p-1,n}(\boldsymbol{i})$ ($p > 3$), where all subsequences of size p of $\boldsymbol{i}$ are compared for monotonicity (with a contribution of 1 for increasing and -1 for decreasing). The proof of Theorem 4.2 carries over to this more general case. When these statistics are used for tests of independence, the power increases as the distribution becomes more associated. Also the results and applications in Chapter 4 of Schriever (1986) are valid for these statistics.

Other related statistics to T_{2n} which involve triplets are $T'_{2n}(\boldsymbol{i}) = 6[n(n-1)(n-2)]^{-1}\sum_{1\le k<l<m\le n}[I(i_k < \min\{i_l,i_m\}) - I(i_k > \max\{i_l,i_m\})]$ and $T''_{2n}(\boldsymbol{i}) = 6[n(n-1)(n-2)]^{-1}\sum_{1\le k<l<m\le n}[I(i_m > \max\{i_k,i_l\}) - I(i_m < \min\{i_k,i_l\})]$. However $T'_{2n}(\boldsymbol{i}) \le T'_{2n}(\boldsymbol{i}')$ and $T''_{2n}(\boldsymbol{i}) \le T''_{2n}(\boldsymbol{i}')$ if $\boldsymbol{i} \prec^{b_t} \boldsymbol{i}'$ for $t = 1,2,3,4$, so that these do not help to show that continuous cdf's H, H' are not $\prec^A$ ordered.

The problem of characterizing all functionals increasing with respect to $\prec^A$ does not appear solvable.

5. Discussion and future work

We have studied thoroughly the "more associated" and "more regression dependent" orderings of Schriever (1986, 1987) and Yanagimoto and Okamoto (1969) for continuous bivariate distributions, including deriving various forms of their definitions under some conditions. We have applied all of these forms (the structural form of Schriever in Definition 2.1, the condition of Yanagimoto and Okamoto in Proposition 2.1 and the representation in Theorem 2.1) to show that several families of bivariate distributions are ordered according to $\prec^A$ or $\prec^{R1}$.

With dependence concepts, the multivariate version is harder to study than the bivariate version. The geometric interpretation of Definition 2.4 suggests a multivariate "more associated" ordering, but most bivariate results do not generalize. For example, two cdf's ordering by this multivariate "more associated" ordering need not be ordered by the multivariate concordance ordering in Joe (1990*b*). Multivariate dependence orderings will however be a subject of further research.

References

Block, H. W., Chhetry, D., Fang, Z. and Sampson, A. R. (1990). Partial orders on permutations and dependence orderings on bivariate empirical distributions, *Ann. Statist.*, **18**, 1840–1850.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, **65**, 141–151.

Cook, R. D. and Johnson, M. E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data, *J. Roy. Statist. Soc. Ser. B*, **43**, 210–218.

Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution, *Biometrika*, **47**, 307–323.

Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x+y-F(x, y)$, *Aequationes Math.*, **19**, 194–226.

Genest, C. (1987). Frank's family of bivariate distributions, *Biometrika*, **74**, 549–555.

Gumbel, E. J. (1961). Bivariate logistic distributions, *J. Amer. Statist. Assoc.*, **56**, 335–349.

Joe, H. (1990*a*). Families of min-stable multivariate exponential and multivariate extreme value distributions, *Statist. Probab. Lett.*, **9**, 75–81.

Joe, H. (1990*b*). Multivariate concordance, *J. Multivariat Anal.*, **35**, 12–30.

Kimeldorf, G. and Sampson, A. R. (1987). Positive dependence orderings, *Ann. Inst. Statist. Math.*, **39**, 113–128.

Kimeldorf, G. and Sampson, A. R. (1989). A framework for positive dependence, *Ann. Inst. Statist. Math.*, **41**, 31–45.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution, *J. Amer. Statist. Assoc.*, **62**, 30–44.

Marshall, A. W. and Olkin, I. (1988). Families of multivariate distributions, *J. Amer. Statist. Assoc.*, **83**, 834–841.

Metry, M. H. and Sampson, A. R. (1988). Positive dependence concepts for multivariate empirical rank distributions, Tech. Report, Department of Mathematics and Statistics, University of Pittsburgh.

Morgenstern, D. (1956). Einfache Beispiele Zweidimensionaler Verteilungen, *Mitteilungsblatt für Mathematische Statistik*, **8**, 234–235.

Oakes, D. (1982). A model for association in bivariate survival data, *J. Roy. Statist. Soc. Ser. B*, **44**, 414–422.

Paulauskas, V. J. (1976). Some remarks on multivariate stable distributions, *J. Multivariate Anal.*, **6**, 356–368.

Press, S. J. (1972). Multivariate stable distributions, *J. Multivariate Anal.*, **2**, 444–462.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

Schriever, B. F. (1986). *Order Dependence*, CWI-Tract 20, Centre for Mathematics and Computer Sciences, Amsterdam.

Schriever, B. F. (1987). An ordering for positive dependence, *Ann. Statist.*, **15**, 1208–1214.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231.

Tchen, A. (1980). Inequalities for distributions with given marginals, *Ann. Probab*, **8**, 814–827.

Yanagimoto, T. and Okamoto, M. (1969). Partial orderings of permutations and monotonicity of a rank correlation statistic, *Ann. Inst. Statist. Math.*, **21**, 489–505.