METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL JOURNAL OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT
REVISTA INTERNACIONAL DE ESTADISTICA

Vol. L - n. 3-4 **31-XII-1992**

PIETRO MULIERE - SONIA PETRONE

Generalized Lorenz curve and monotone dependence orderings

PIETRO MULIERE (*) - SONIA PETRONE (*)

Generalized Lorenz curve and monotone dependence orderings

1. Introduction

The intuitive meaning of monotone dependence for a bivariate random variable (X, Y) is that large values of Y correspond stochastically to large values of X (positive dependence) or, in the opposite case, large values of Y correspond to small values of X (negative dependence). This concept of monotone dependence has played a fundamental role in many recent new ideas in statistics.

A number of families of bivariate distributions with a natural interpretation of monotone dependence have been introduced. Some early concepts that have been recently developed include quadrant dependence (Lehmann, 1966), association (Esary, Proschan and Walkup, 1967), concordance (Tchen, 1980; Kimeldorf and Sampson, 1978). Different notions of monotone dependence are defined through the conditional distribution of $Y \mid X = x$ or through the regression function $E(Y \mid X = x)$. Examples are regression dependence (Tukey, 1958; Lehmann, 1966), total positivity of order two (Karlin, 1968), monotone regression (Shea, 1979). A weaker type of monotonic dependence based on expectations of random variables $(Y \mid X > x)$

(*) Dipartimento di Economia Politica e Metodi Quantitativi, Università di Pavia - Via S. Felice 5 - 27100 Pavia.

was introduced by Yanagimoto (1973) and Kowalczyk and Pleszczynska (1977).

Literature concerning monotone dependence can be found, for instance, in Schriever (1986); Block et *al.* (1990).

In order to compare two bivariate distributions having the same pair of marginals to determine whether one distribution is more positively dependent then the other, several partial orderings on the class of bivariate distributions with fixed marginals have been introduced. Some specific examples of such orderings are: Tchen's (1980) more concordant ordering; Rinott and Pollak's (1980) covariance ordering. A general concept of a positive dependence ordering is studied by Kimeldorf and Sampson (1987); see also Scarsini (1984).

Is this paper we introduce a new characterization of monotone dependence in drawing a comparison between the generalized Lorenz curve of the regression function $E(Y \mid X)$ and the Lorenz curve of Y. This notion seems appropriate when we ask the relation to be invariant under increasing transformations of X but sensibile to increasing transformations of Y. The importance of this property follows from the fact that there exist practical situations in which metric is relevant just in the case of one variable.

A partial ordering of monotone dependence on the class of nonnegative bivariate r.v.'s with given marginals is consequently defined. Relations with other orderings are examined. Finally, it is suggested a real valued parameter that adequately reflects the properties of the defined ordering.

The motivation of the proposed partial ordering arises in the study of economic problems (e.g. income distribution, taxation and welfare programs) but we believe it to be useful in several applications such as selection, discrimination problems, statistical quality control, psychological screening.

2. Generalized Lorenz curve: definition and properties

Let X be a nonnegative random variable (r.v.) with distribution function (d.f.) F_x and finite expectation $E(X) > 0$. We consider the following definition of the Lorenz curve of X (Pietra, 1915, p. 782, see also: Frosini, 1987, p. 195; Giorgi, 1992, pp. 12, 149; Gatswirth, 1971, p. 1037):

$$L_X(p) = \frac{1}{E(X)} \int_0^p F_x^{-1}(z)\, dz\,, \qquad 0 \le p \le 1$$

where: $F_x^{-1}(z) = \inf\{x : F_x(x) \ge z\}\,,\quad 0 \le z \le 1$.

Let g be a real function on $\mathbf{R}$. Following an idea of Mahalanobis (1960) and Kakwani (1977), we can define the generalized Lorenz curve of the r.v. $g(X)$.

Definition 2.1 (Kakwani, 1977, p. 720). Let $g(x)$ be a continuous function of x such that its first derivative exists and $g(x) \ge 0$. If $0 < E(g(X)) < +\infty$, we define:

$$L_{g(X)}(p) = \frac{1}{E(g(X))} \int_0^{x_p} g(t)\, dF_x(t) = \frac{1}{E(g(X))} \int_0^p g(F_x^{-1}(z))\, dz\,,$$

$0 \le p \le 1$, where $x_p = F_x^{-1}(p)$.

Let us now consider a bivariate nonnegative r.v. (X, Y), with marginals F_x and F_y respectively, and suppose that $E(X)$ and $E(Y^2)$ are positive and finite. Blitz and Brittain (1964) introduced the correlation curve of Y on X, which corresponds to the generalized Lorenz curve of the regression function $m(X) = E(Y \mid X)$. Taguchi (1981) more specifically analyzes the correlation curve, using the notion of mean difference, essentially in the case of linear regression function.

In this section we examine some properties of the generalized Lorenz curve of $m(X) = E(Y \mid X)$. For the sake of simplicity, attention is restricted to the class $\pi(F_x, F_y)$ of bivariate nonnegative r.v.'s (X, Y) with continuous marginal d.f.'s F_x and F_y. We shall also assume that for all $(X, Y) \in \pi(F_x, F_y)$ the regression function $m(x) = E(Y \mid X = x)$ is continuous with finite first derivative $m'(x)$.

From definition 2.1, the generalized Lorenz curve of $[m(X) = E(Y \mid X)]$ is defined as:

$$L_{E(Y|X)}(p) = \frac{1}{E(E(Y \mid X))} \int_0^{x_p} m(t)\, dF_x(t) = \frac{1}{E(Y)} \int_0^p m(F_x^{-1}(z))\, dz$$

$0 \le p \le 1$.

PROPERTY 1.

i) $L_{E(Y|X)}(p)$ passes through the points (0,0) and (1,1).
ii) $L_{E(Y|X)}(p)$ is increasing if and only if $m(x) > 0$ for all x.
iii) $L_{E(Y|X)}(p)$ is concave if and only if $m(x)$ is nondecreasing for all x.

Proof. (i) From definition. (ii) follows immediately observing that: $\dfrac{\partial L_{E(Y|X)}(p)}{\partial p} = \dfrac{m(x_p)}{E(Y)}$. (iii) $\dfrac{\partial^2 L_{E(Y|X)}(p)}{\partial p^2} = \dfrac{m'(x_p)}{E(Y) f_x(x_p)}$, where f_x is the density function of F_x. Since $f_x(x_p) > 0$, the sign of the second derivative is that of $m'(x)$.

A second criterion by Kakwani (1977, Corollary 1) states that $L_{E(Y|X)}(p)$ is above (below) the egalitarian line if the elasticity $\eta(x) = \dfrac{x\, m'(x)}{m(x)}$ is less (greater) than zero for all $x \geq 0$. Kakwani (1977) also pointed out that $L_{E(Y|X)}(p)$ is not the same thing as the Lorenz curve for $E(Y|X)$. Both are identical if $m(x)$ is strictly monotonic and has a continuous derivative $m'(x) > 0$ for all x.

PROPERTY 2. $L_Y(p) \leq L_{E(Y|X)}(p) \leq L^*_Y(p)$, where:

$$L^*_Y(p) = \frac{1}{E(Y)} \int_{1-p}^{1} F_y^{-1}(z)\, dz \qquad 0 \leq p \leq 1$$

Proof. In order to show that: $L_Y(p) \leq L_{E(Y|X)}(p)$, note that

$$L_{E(Y|X)}(p) = \frac{1}{E(Y)} E(Y|X \leq x_p)\, F_x(x_p),$$

and:

$$L_Y(p) = \frac{1}{E(Y)} E(Y \mid Y \leq y_p)\, F_y(y_p).$$

Now, $(Y \mid X \leq x_p)$ is stochastically larger than $(Y \mid Y \leq y_p)$, that is:

$$P\{Y \leq y \mid X \leq x_p\} \leq P\{Y \leq y \mid Y \leq y_p\} \quad \text{for all } y \in \mathbf{R},$$

and the result follows.

Analogously, in order to prove: $L_{E(Y|X)}(p) \leq L_{Y^*}(p)$, note that:

$$L^*_Y(p) = \frac{1}{E(Y)} E(Y \mid Y > y_{1-p})\,[1 - F_y(y_{1-p})]$$

Now, $(Y \mid Y > y_{1-p})$ is stochastically larger than $(Y \mid X \leq x_p)$ and the result follows.

A notation $\doteq$ will be used to denote the equivalence of random variables, defined as: $X \doteq Y$ if the probability of the event $\{X = Y\}$ is equal to 1.

PROPERTY 3. $L_{E(Y|X)}(p) = p$ for all $p \in (0, 1)$ if and only if $E(Y \mid X) \doteq E(Y)$.

Proof. $L_{E(Y|X)}(p) = p$ for all $p \in (0,1)$ can be rewritten as:

$$\int_0^{x_p} E(Y \mid X = x)\, dF_x(x) = \int_0^{x_p} E(Y)\, dF_x(x) \quad \text{for all } p \in (0,1),$$

and this is equivalent to $E(Y \mid X) \doteq E(Y)$.

PROPERTY 4.

i) $L_{E(Y|X)}(p) = L_Y(p)$ for all $p \in (0,1)$ iff $m(x)$ is increasing and such that $E(Y \mid X) \doteq Y$.

ii) $L_{E(Y|X)}(p) = L^*_Y(p)$ for all $p \in (0,1)$ iff $m(x)$ is decreasing and such that $E(Y \mid X) \doteq Y$.

Proof. i) If $m(x)$ is increasing and $E(Y \mid X) \doteq Y$, the result immediately follows.

Suppose that: $L_{E(Y|X)}(p) = L_Y(p)$, for every $p \in (0,1)$. Then, by

Property 1, $m(x)$ is increasing and $L_{E(Y|X)}(p)$ is the Lorenz curve of $E(Y \mid X)$, so that from the assumed equality follows that $E(Y \mid X)$ and Y are identically distributed. This implies that $E(Y|X) \doteq Y$ (see: Dabrowska (1985, p. 71). ii) can be proved analogously.

PROPERTY 5. If $m(x)$ is nondecreasing for all x then [Cov $(X, Y) \geq 0$] and $L_{E(Y|X)}(p) \leq p$. The inequalities are reversed if $m(x)$ is nonincreasing. In both cases, Cov $(X, Y) = 0$ and $L_{E(Y|X)}(p) = p$ hold if and only if $E(Y \mid X) \doteq E(Y)$.

Proof. From Property 1 and Lemma 2.1 of Shea (1979, p. 1122).

The following properties 6 and 7 can be easily proved.

PROPERTY 6. If $E(Y \mid X = x) = \alpha + \beta x$, then

$$L_{E(Y|X)}(p) = p - \beta \frac{E(X)}{E(Y)} (p - L_X(p)).$$

We shall say that a function $h: \mathbf{R} \to \mathbf{R}$ is F_x-increasing if for all $s, t \in \mathbf{R}$: $F_x(s) < F_x(t)$ implies $h(s) < h(t)$.

PROPERTY 7. $L_{E(Y|X)}(p)$ is invariant under F_x-increasing transformations of X.

PROPERTY 8. Let (X, Y) and $(X', Y') \in \pi(F_x, F_y)$.
Then $L_{E(Y|X)}(p) = L_{E(Y'|X')}(p)$ for all $p \in (0,1)$ if and only if $E(Y \mid X)$ and $E(Y' \mid X')$ have the same distribution.

Proof. $L_{E(Y|X)}(p) = L_{E(Y'|X')}(p)$ for all p is equivalent to:

$$\int_0^{x_p} E(Y \mid X = x)\, dF_x(x) = \int_0^{x_p} E(Y' \mid X' = x)\, dF_x(x) \text{ for every } p \in (0,1)$$

which holds iff $E(Y \mid X)$ and $E(Y' \mid X')$ have the same distribution.

PROPERTY 9. Let (X, Y) and $(X', Y') \in \pi(F_x, F_y)$. Then $L_{E(Y|X)}$ will lie above (below) $L_{E(Y'|X')}$ if the elasticity of $E(Y \mid X = x)$ is less (greater) than the elasticity of $E(Y' \mid X' = x)$.

Proof. See Kakwani (1977, p. 720).

3. GENERALIZED LORENZ CURVE OF $E(Y \mid X)$ AND POSITIVE (NEGATIVE) DEPENDENCE

The Lorenz curve has been widely used as a tool for ordering distributions. Let $X \sim F_x$, $Y \sim F_y$ be two nonnegative r.v.'s with finite positive expectations. The distributions F_x and F_y or equivalently the r.v.'s X and Y are ordered by Lorenz ordering if the Lorenz curve of X is nowhere below the Lorenz curve of Y:

$$X \leq_L Y \text{ if } L_X(p) \geq L_Y(p) \text{ for every } p \in (0,1)$$

(see Marshall and Olkin, 1979, p. 5).

We propose to use the generalized Lorenz curve for ordering bivariate distributions according to monotone dependence. Our aim is to measure monotone dependence in drawing a comparison between the generalized Lorenz curve of $E(Y \mid X)$ and the Lorenz curve of Y. Roughly speaking, the more is the generalized Lorenz curve of $E(Y|X)$ similar to (or far from) the Lorenz curve of Y, the stronger is the positive (negative) dependence of Y on X.

We now introduce a monotone dependence structure based on $L_{E(Y|X)}$. In what follows $\boldsymbol{L}$ denotes the family of all bivariate r.v.'s (X, Y) with monotonic dependence, $\boldsymbol{L}^+$ and $\boldsymbol{L}^-$ corresponding to positive and negative dependence.

DEFINITION 3.1 The set of all ordered pairs (X, Y) satisfying:

$$L_{E(Y|X)} \leq p \text{ for all } p \in (0,1)$$

will be denoted by $\boldsymbol{L}^+$. It will be denoted by $\boldsymbol{L}^-$ if the inequality is reversed.

PROPOSITION 3.1 If $(X, Y) \in L^+ [L^-]$, then Cov $(X, Y) \geq 0$ $[\leq 0]$.

Cov $(X, Y) = 0$ holds iff $E(Y \mid X) \doteq E(Y)$.

Proof. If $(X, Y) \in L^+$, then $L_{E(Y|X)}(p) \leq p$ for every p. This implies that:

$$\int_0^{x_p} [E(Y \mid X = x) - E(Y)] \, dF_x(x) \leq 0, \quad p \in (0,1).$$

Consider first:

$$\mathrm{Cov}(X, Y) = \int_0^{\infty} x [E(Y \mid X = x) - E(Y)] \, dF_x(x).$$

For any point x_p, from the assumption we have:

$$\mathrm{Cov}(X, Y) = \int_0^{x_p} x [E(Y \mid X = x) - E(Y)] \, dF_x(x) +$$

$$+ \int_{x_p}^{\infty} x [E(Y \mid X = x) - E(Y)] \, dF_x(x) \geq$$

$$\geq x_p \int_0^{x_p} [E(Y \mid X = x) - E(Y)] \, dF_x(x) +$$

$$+ x_p \int_{x_p}^{\infty} [E(Y \mid X = x) - E(Y)] \, dF_x(x) =$$

$$= x_p \int_0^{\infty} [E(Y \mid X = x) - E(Y)] \, dF_x(x) = 0.$$

The implication: $(X, Y) \in L^- \Rightarrow \mathrm{Cov}(X, Y) \leq 0$ can be proved analogously.

Cov $(X, Y) = 0 \Leftrightarrow E(Y \mid X) = E(Y)$ is a consequence of Property 3.

Let us now order the family $L^+ [L^-]$.

DEFINITION 3.2. For each (X, Y) and (X', Y') belonging to $L^+ [L^-]$:

$$(X, Y) \leq_{L^+} (X', Y') \qquad [\leq_{L^-}]$$

if $F_x = F_{x'}$, $F_y = F_{y'}$, and $L_{E(Y|X)}(p) \geq L_{E(Y'|X')}(p)$ $[\leq]$ for all $p \in (0, 1)$.

Clearly, the set of all minimal elements with respect to $\leq_{L^+}$ $[\leq_{L^-}]$ is equal to the set of the r.v.'s (X, Y) such that $E(Y \mid X) \doteq E(Y)$ (see Property 3). The set of all maximal elements with respect to $\leq_{L^+}$ $[\leq_{L^-}]$ consists of all bivariate r.v.'s (X, Y) whose joint d.f. is equal to:

$$F^+(x, y) = \min \{F_x(x), F_y(y)\}$$
$$[F^-(x, y) = \max \{0, F_x(x) + F_y(y) - 1\}], (x, y) \in \mathbf{R}^2$$

(see Property 4). The d.f. F^+ $[F^-]$ is usually referred to as the upper (lower) Fréchet bound of the family of all bivariate distribution functions with marginals F_x and F_y.

PROPOSITION 3.2. Let (X, Y) and $(X', Y') \in L^+$. If

$(X, Y) \leq_{L^+} (X', Y')$ then $\mathrm{Cov}(X, Y) \leq \mathrm{Cov}(X', Y')$

Proof. The result can be obtained by paraphrasing the proof of Proposition 3.1.

From Property 9 of section 2 it follows that: $(X, Y) \leq_{L^+} (X', Y')$ if the elasticity of $E(Y \mid X = x)$ is less than the elasticity of $E(Y' \mid X' = x)$.

In what follows we examine the properties of L^+-ordering with respect to other orderings of positive dependence. The case of negative dependence is obviously analogous.

DEFINITION 3.3. Let X and Y be two random variables.

a) (Lehmann, 1966, p. 1138). X and Y are positively quadrant dependent (QD^+) if $(Y \mid X > x)$ is stochastically larger than Y:

$$(Y \mid X > x) \geq_{\mathrm{st}} Y$$

b) (Esary, Proschan and Walkup, 1967, p. 1466). X and Y are associated ($A(X, Y)$) if:

$$\mathrm{Cov}\,(f(X, Y)\,;\; g(X, Y)) \geq 0$$

for all increasing functions f and g.

c) (Kowalczyk and Pleszczynska, 1977, p. 1221). X and Y are positively dependent in expectation (EQD^+) if:

$$E(Y \mid X > x) \geq E(Y), \quad \text{for all } x \in \mathbf{R}.$$

The implications among these notions of bivariate positive dependence properties are:

$$A(X, Y) \Rightarrow QD^+ \Rightarrow EQD^+.$$

Let us now order the family of EQD^+ r.v.'s.

DEFINITION 3.4. For each (X, Y) and (X', Y') which are EQD^+:

$$(X, Y) \leq_{EQD^+} (X', Y')$$

if $F_x = F_{x'}$, $F_y = F_{y'}$ and for all $x \in \mathbf{R}$: $E\,(Y \mid X > x) \leq [E\,(Y' \mid X' > x)]$.

We have introduced in Definition 3.2 an ordering of monotonic dependence based on $L_{E(Y|X)}$, which results to be equivalent to $\leq_{EQD^+}$ as is shown in the following

PROPOSITION 3.3. For each (X, Y) and (X', Y') belonging to L^+:

$$(X, Y) \leq_{L^+} (X', Y') \text{ iff } (X, Y) \leq_{EQD^+} (X', Y').$$

Proof. $(X, Y) \leq_{L^+} (X', Y')$ is equivalent to:

$L_{E(Y|X)}(p) \geq L_{E(Y'|X')}(p)$ for every $p \in (0,1)$, that is:

$$\frac{1}{E(Y)} E(Y \mid X \leq x_p) F_x(x_p) \geq \frac{1}{E(Y')} E(Y' \mid X' \leq x_p) F_{x'}(x_p), \; p \in (0,1).$$

Observing that:

$$E(Y \mid X \leq t) = \frac{1}{P(X \leq t)} [E(Y) - E(Y \mid X > t)\, P(X > t)],$$

the previous inequality holds for every $p \in (0,1)$ if and only if:

$$\frac{1}{E(Y)} \{E(Y) - E(Y \mid X > x_p)(1-p)\} \geq \frac{1}{E(Y')} \{E(Y') - E(Y' \mid X' > x_p)(1-p)\}$$

that is:

$$E(Y \mid X > x) \leq E(Y' \mid X' > x), \quad \text{for all } x.$$

4. L-DEPENDENCE FUNCTIONS

Kowalczyk and Pleszczynska (1977) and Kowalczyk (1977) propose to associate with any (X, Y) with finite expectations and continuous marginals a function $\mu_{yx}: (0,1) \to [-1, +1]$ defined by:

$$\mu_{yx}(p) = \mu_{yx}^{+}(p) \quad \text{if} \quad E(Y \mid X > x_p) - E(Y) \geq 0$$
$$= \mu_{yx}^{-}(p) \quad \text{if} \quad E(Y \mid X > x_p) - E(Y) \leq 0$$

where:

$$\mu_{yx}^{+}(p) = \frac{E(Y \mid X > x_p) - E(Y)}{E(Y \mid Y > y_p) - E(Y)}$$

and

$$\mu_{yx}^{-}(p) = \frac{E(Y \mid X > x_p) - E(Y)}{E(Y) - E(Y \mid Y < y_{1-p})}$$

The function μ_{yx} is called monotonic dependence function of Y on X.

In fact, given two r.v.'s (X, Y) and (X', Y') with marginal distributions respectively equal, it holds:

$$(X, Y) \leq_{EQD^+} (X', Y') \text{ if and only if } \mu_{yx}^{+}(p) \leq \mu_{y'x'}^{+}(p)$$

for all $p \in (0,1)$.

We can also introduce a monotone dependence function c_{YX} based on the normalized differences between the points on the generalized Lorenz curve $L_{E(Y|X)}$ and the corresponding points on the egalitarian line:

$$c_{YX}(p) = c_{YX}^{+}(p) \quad \text{if } L_{E(Y|X)}(p) \leq p$$
$$= c_{YX}^{-}(p) \quad \text{if } L_{E(Y|X)}(p) > p, \quad 0 < p < 1$$

where:

$$c_{YX}^{+}(p) = \frac{p - L_{E(Y|X)}(p)}{p - L_Y(p)}$$

and:

$$c_{YX}^{-}(p) = \frac{p - L_{E(Y|X)}(p)}{L_{Y^*}(p) - p}.$$

The two functions μ_{yx} and c_{YX} are clearly related. Indeed we have the following:

PROPOSITION 4.1. For all $p \in (0,1)$, $\mu_{yx}(p) = c_{YX}(p)$.

Proof. The result follows observing that:

$$E(Y \mid X \leq x_p)\, P(X \leq x_p) = E(Y) - E(Y \mid X > x_p)\, P(X > x_p).$$

It is immediate to note that from Proposition 3.3, it follows that: $(X, Y) \leq_{L^+} (X', Y')$ iff $c_{YX}^{+}(p) \leq c_{Y'X'}^{+}(p)$ for all $p \in (0,1)$.

5. MEASURES OF *L*-DEPENDENCE

It is evident that the ordering $\leq_{L^+}$ is only partial: two bivariate r.v.'s are not comparable (under $\leq_{L^+}$) whenever their generalized Lorenz curves intersect. In order to compare pairs of r.v.'s that are not ordered by $\leq_{L^+}$ ordering, it is wise to choose (partial or total) orders that are finer than $\leq_{L^+}$ ordering. An order $\leq_A$ is finer than another partial order $\leq_B$ if $(X, Y) \leq_B (X', Y')$ implies $(X, Y) \leq_A (X', Y')$ i.e., $\leq_A$ orders all bivariate distributions that $\leq_B$ orders. An order is total if it orders every pair of bivariate distributions.

A measure I of L-dependence is a functional of the bivariate r.v. that induces a total order $\leq_I$ defined by:

$$(X, Y) \leq_I (X', Y') \text{ if } I(X, Y) \leq I(X', Y')$$

which is finer than $\leq_{L^+}$ and $\geq_{L^-}$.

A measure of L-dependence is given by the area between the egalitarian line and the generalized Lorenz curve $L_{E(Y|X)}$:

$$A_{YX} = \frac{1}{2} - \int_0^1 L_{E(Y|X)}(p)\, dp$$

Normalizing the area A_{YX} a relative measure of L-dependence, which takes values in $[-1, +1]$, can be deduced.

PROPOSITION 5.1. Let $A_Y = \int_0^1 [p - L_Y(p)]\, dp$. Then:

i) $C_{YX} = A_{YX} / A_Y$ is a relative measure of L-dependence, that is:

$-1 \le c_{YX} \le +1$;

ii) $|C_{YX}| = 1$ if and only if $E(Y \mid X) \doteq Y$. More specifically,
$C_{YX} = +1$ if $m(x) = E(Y \mid X = x)$ is a nondecreasing function for all x, and
$C_{YX} = -1$ if $m(x)$ is nonincreasing;

iii) $C_{YX} = 0$ if $E(Y \mid X) \doteq E(Y)$;

iv) If $E(Y \mid X = x) = \alpha + \beta x$, then

$$C_{YX} = \beta \frac{E(X)}{E(Y)} \frac{A_X}{A_Y} = \rho \frac{\sigma_X}{\sigma_Y} \frac{E(X)}{E(Y)} \frac{A_X}{A_Y},$$

where ρ is the linear correlation coefficient.

Proof. i) - ii) - iii) follow from Properties 2, 4 and 3 of section 2 respectively. iv) is immediate.

An expressive relation can be stated between C_{YX} and the covariance: Cov $(Y, F_x(X))$. It is well known (see Stuart, 1954) that the mean difference Δ_Y of a random variable Y with d.f. F_y can be defined as:

$$\Delta_Y = 4\, \mathrm{Cov}(Y, F_y(Y)),$$

so that the concentration area A_Y can be expressed as:

$$A_Y = \frac{\mathrm{Cov}(Y, F_y(Y))}{E(Y)}$$

being: $A_Y = \Delta_Y / (4\, E(Y))$. Given (X, Y) with marginals d.f.'s F_x and F_y, we have:

$$A_{YX} = \frac{\mathrm{Cov}(Y, F_x(X))}{E(Y)}.$$

Consequently, the monotone dependence measure C_{YX} becomes:

$$C_{YX} = \frac{\mathrm{Cov}(Y, F_x(X))}{\mathrm{Cov}(Y, F_y(Y))}.$$

An alternative approach to derive the measure C_{YX} is proposed by Schechtman and Yitzhaki (1987) and Yitzhaki and Olkin (1991).

6. AN EXAMPLE

In studying the dependence structure of a family of fixed marginal bivariate distributions indexed by a real parameter it is natural to inquire whether the bivariate distributions are more positively dependent according to C_{YX} as the parameter increases.

We consider the following example. For any pair F_x and F_y of continuous univariate d.f.'s, the following family of bivariate distributions was studied by Farlie (1960), Gumbel (1958) and Morgenstern (1956):

$$F(x, y; \alpha) = F_x(x) F_y(y) \{1 + \alpha (1 - F_x(x)) (1 - F_y(y))\}$$

$-1 \le \alpha \le +1$.

It is easy to prove that if $0 \le \alpha_1 < \alpha_2 < 1$, then:

$$F(x, y; \alpha_1) \le_{L^+} F(x, y; \alpha_2)$$

It is sufficient to consider the case when F_x and F_y are uniform distributions on (0,1), in which case:

$$f(x, y; \alpha) = 1 + \alpha (2x - 1)(2y - 1)$$

is the density corresponding to $F(x, y; \alpha)$ and

$$E(Y \mid X = x) = \frac{1}{6}(3 - \alpha + 2\alpha x).$$

The regression function is linear, so that from Proposition 5.1:

$$C_{YX} = \rho \frac{\sigma_X}{\sigma_Y} \frac{E(X)}{E(Y)} \frac{A_X}{A_Y} = \rho = \alpha / 3.$$

7. Some remarks

Extension to the discrete case is possible. Let (X, Y) be a nonnegative bivariate r.v. and suppose that (X, Y) takes values (x_i, y_j) with probability $p(x_i, y_j)$, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, h$. Let p_x be the marginal probability function of X. The generalized Lorenz curve of $E(Y \mid X)$ is defined as:

$$L_{E(Y|X)}(p) = \frac{1}{E(Y)} \left\{ \sum_{x_i < x_p} E(Y \mid X = x_i)\, p_x(x_i) + E(Y \mid X = x_p)\,(p - F_X(x_p^-)) \right\}$$

where $F_X(x_p^-) = \lim_{x \to x_p^-} F_X(x)$. Therefore, $L_{E(Y|X)}(p)$ coincides with the line connecting the $(k+1)$ points:

$$(0, 0); \left(\sum_{j=1}^{i} p_X(x_j),\ \sum_{j=1}^{i} \frac{E(Y \mid x_j)\, p_X(x_j)}{E(Y)} \right),\ i = 1, \ldots, k$$

In this case, the properties of the generalized Lorenz curve of the regression function presented in section 2 must be reformulated considering the following inequalities:

$$L_Y(p) \le L^+_{E(Y|X)}(p) \le L_{E(Y|X)}(p) \le L^-_{E(Y|X)}(p) \le L^*_Y(p)$$

where: $L^+_{E(Y|X)}$ ($L^-_{E(Y|X)}$) is the generalized Lorenz curve of the regression function when the distribution of (X, Y) is the upper bound F^+

(respectively, the lower bound F^-) of the Fréchet class generated by F_X and F_Y.

The equality $L_Y(p) = L^+_{E(Y|X)}(p)$ ($L^*_Y(p) = L^-_{E(Y|X)}(p)$) holds for all $p \in (0,1)$ if and only if F^+ (F^-) is such that $Y \doteq g(X)$ for a nondecreasing (nonincreasing) function g.

This suggests to define a relative measure of L-dependence as:

$$C_{YX} = \frac{A_{YX}}{A^+_{YX}} = \frac{\mathrm{Cov}(Y, F_X(X))}{\mathrm{Cov}^+(Y, F_X(X))} \qquad \text{if} \quad A_{YX} \ge 0$$

$$= -\frac{A_{YX}}{A^-_{YX}} = -\frac{\mathrm{Cov}(Y, F_X(X))}{\mathrm{Cov}^-(Y, F_X(X))} \qquad \text{if} \quad A_{YX} < 0$$

where

$$A^+_{YX} = \int_0^1 [p - L^+_{E(Y|X)}(p)]\, dp \quad \text{and} \quad A^-_{YX} = \int_0^1 [p - L^-_{E(Y|X)}(p)]\, dp.$$

The monotone dependence function c_{YX} and the index C_{YX} proposed in this work can be used in many practical problems. Some hints for possible applications in economics can be found in: Kakwani (1980); Plotnick (1981); Nygård and Sandström (1981); Muliere (1986), where economic problems as taxation or welfare programs are studied with alternative methods.

Acknowledgement

The authors wish to thank the referee for his helpful comments.

REFERENCES

Blitz, R.C. and Brittain, J.A. (1964). An Extension of the Lorenz Diagram to the Correlation of Two Variables, *Metron*, XXIII, pp. 137-143.

Block, H.W., Sampson, A.R. and Savits, T.H. Eds (1990). *Topics in Statistical Dependence*, Institute of Mathematical Statistics, Lecture Notes, Vol. 16, Hayward.

Dabrowska, D. (1985). Descriptive Parameters of Location, Dispersion and Stochastic Dependence, *Statistics*, 16, pp. 63-88.

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

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

FROSINI, B.V. (1987). *Lezioni di Statistica*, (II Ediz.), Vita e Pensiero, Milano.

GASTWIRTH, J.L. (1971). A general definition of the Lorenz curve, *Econometrica*, 39, pp. 1037-1039.

GIORGI, G.M. (1992). *Il rapporto di concentrazione di Gini: genesi, evoluzione ed una bibliografia commentata*, Editrice Ticci, Siena.

GUMBEL, E.J. (1958). Distributions à plusieurs variables dont les marges sont donnés, *Comptes Rendus de l'Acad. Sci. Paris*, 246, pp. 2717-2720.

KAKWANI, N.C. (1977). Applications of Lorenz curves in Economic Analysis, *Econometrica*, 45, pp. 719-727.

KAKWANI, N.C. (1980). *Income Inequality and Poverty*, Oxford University Press.

KARLIN, S. (1968). *Total Positivity*, Vol. I, Stanford University Press, Stanford.

KIMELDORF, G. and SAMPSON, A.R. (1978). Monotone Dependence, *Ann. Statist.*, 6, pp. 895-903.

KIMELDORF, G. and SAMPSON, A.R. (1987). Positive dependence orderings, *Ann. Inst. Statist. Math.*, 39, pp. 113-128.

KOWALCZYK, T. (1977). General definition and sample counterparts of monotonic dependence functions of bivariate distributions, *Math. Operations forsch. u. Statist., Ser. Statist.*, 8, pp. 351-369.

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

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

MAHALANOBIS, P.C. (1960). A method of fractile graphical analysis, *Econometrica*, 28, pp. 325-351.

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

MORGENSTERN, D. (1956). Einfache Beispiele zweidimensionaler Verteilungen, *Mitteilungslatt fur Mathematische Statistik*, 11, pp. 234-235.

MULIERE, P. (1986). Alcune osservazioni sull'equità orizzontale di una tassazione, *Scritti in onore di F. Brambilla*, Edizioni di Bocconi Comunicazione, Milano, v. II, pp. 551-559.

NYGÅRD, F. and SANDSTRÖM, A. (1981). *Measuring Income Inequality*, Almqvist & Wicksel International, Stockholm.

PIETRA, G., (1915). Delle relazioni tra gli indici di variabilità (Nota I), *Atti del Reale Istituto Veneto di Scienze, Lettere ad Arti*, a.a. 1914-15, tomo 74, pt. II, pp. 775-792.

PLOTNICK, R. (1981). A measure of horizontal inequity, *The Rewiew of Economics and Statistics*, 2, pp. 283-288.

RINOTT, Y. and POLLAK, M. (1980). A stochastic ordering induced by a concept of positive dependence and monotonicity of asymptotic test sizes, *Ann. Statist.*, 8, pp. 190-198.

SCARSINI, M. (1984). On measures of concordance, *Stochastica*, 8, pp. 201-218.

SCHECHTMAN, E. and YITZHAKI, S. (1987). A measure of association based on Gini's mean difference, *Commun. Statist.* Theor. Meth., 16(1), pp. 207-231.

SHEA, G. (1979). Monotone regression and covariance structure, *Ann. Statist.*, 7, pp. 1121-1126.

SCHRIEVER, B.F. (1986). *Order Dependence*, Centre for Mathematics and Computer Science, Amsterdam.

STUART, A. (1954). The correlation between variate values and ranks in samples from a continuous distribution, *British Journal of Statistical Psychology*, pp. 37-44.

TAGUCHI, T. (1981). On a multiple Gini's coefficient and some concentration regression, *Metron*, XXXIX, pp. 69-97.

TCHEN, A.H. (1980). Inequalities for distributions with given marginals, *Ann. Probab.*, 8, pp. 814-827.

TUKEY, J. (1985). A problem of Berkson and minimum variance orderly estimators, *Ann. Math. Statist.*, 29, pp. 588-592.

YANAGIMOTO, T. (1973). Families of positively dependent random variables, *Ann. Inst. Statist. Math.*, 24, pp. 559-573.

YITZHAKI, S. and OLKIN, I. (1991). Concentration Indices and Concentration Curves, in: *Stochastic Orders and Decision under Risk* (K. Mosler and M. Scarsini eds.), Institute of Mathematical Statistics. Monograph Series. pp. 380-392.

Generalized Lorenz curve and monotone dependence orderings

SUMMARY

In this paper a partial order of monotone dependence on the class of nonnegative bivariate random variables with given marginals is defined, based on the notion of generalized Lorenz curve of the regression function. The relations with other monotone dependence orderings are examined, with particular attention to the *EQD* ordering proposed by Kowalczyk and Pleszczynska (1977). It is finally suggested a real valued parameter coherent with the introduced ordering.

La curva di Lorenz generalizzata e ordinamenti di dipendenza monotona

RIASSUNTO

Nel presente lavoro viene proposto un ordinamento parziale di dipendenza monotona per variabili casuali bidimensionali non negative e con marginali assegnate.

Tale ordinamento è basato sulla nozione di curva di Lorenz generalizzata della funzione di regressione. Vengono esaminate le relazioni con altri ordinamenti di dipendenza monotona presenti in letteratura, con particolare riferimento all'ordinamento *EQD* (Kowalczyk and Pleszczynska, 1977). Si mostra inoltre come l'utilizzo della curva di Lorenz generalizzata permetta di costruire un indice di dipendenza monotona coerente con l'ordinamento parziale introdotto.

Key words

Generalized Lorenz curve, monotone dependence, positive dependence orderings, dependence measures.