AN ORDERING OF DEPENDENCE

BY MARCO SCARSINI[1]

Università "La Sapienza"

An ordering of dependence is defined on the space of probability measures on a finite product space, with fixed marginals. The definition of this ordering involves the Lorenz curve of the likelihood ratio of a probability measure w.r.t. the product measure of the marginals. A minimal element w.r.t. this dependence ordering always exists and equals the product measure. Conditions for the existence of a maximal element are examined. The ordering is generalized to the case of infinite spaces. Comparison with some other orders of dependence is considered.

1. Introduction. The purpose of this paper is to study an ordering of dependence for pairs of random variables (r.v.'s) taking values in sets that are finite, but not necessarily numerical or (even partially) ordered. The basic idea is that, when two r.v.'s are stochastically independent, then there is no dependence. An ordering of dependence should express this fact by having its unique minimum in the case of independent r.v.'s. Independence is then a reference point and a pair of r.v.'s is more dependent than another if it is more distant than the other from the reference situation of independence, in a sense that will be specified later.

If we want to compare two r.v.'s (or, analogously, two probability measures on a finite product space) w.r.t. dependence, then it makes sense to start considering only probability measures having the same marginals. The study of dependence will be carried out by introducing a preorder on the space of probability measures on a finite product space having the same marginals, or (equivalently) on the space of all probability matrices of fixed dimensions having prescribed row and column sums. Since we are considering qualitative r.v.'s, any ordering of dependence should not depend on the order in which the different possible outcomes are considered, in other words, the ordering should be invariant w.r.t. permutations of lines (rows and/or columns) in the probability matrices.

[1]Partially supported by MPI.

AMS 1980 subject classification. Primary 62H20.

Key words and phrases. Probability measures with fixed marginals, generalized concentration curve.

I am greatly indebted to Eugenio Regazzini for giving me a copy of his unpublished papers on generalized concentration curve and dependence. I also wish to thank all the people who offered comments when this paper was presented at seminars in Stanford and Perugia and at the "Symposium on Dependence in Statistics and Probability" in Hidden Valley.

The tool that we will use is the concentration curve of a probability measure with respect to another, introduced by Cifarelli and Regazzini (1987). Our results are strictly related to some ideas of Ali and Silvey (1965a), (1965b). Their results are not expressed in terms of concentration curves, though.

2. Generalized Concentration Curve. Let P, Q be two probability measures on the power set of a finite space X. From now on, for $x \in X$, we will write $P(x)$ instead of $P(\{x\})$. Cifarelli and Regazzini (1987) have defined the concentration curve of a measure P w.r.t. Q, as follows. Let ℓ_P be the (generalized) likelihood ratio of P w.r.t. Q: for $x \in X$

$$\ell_P(x) = \begin{cases} \frac{P(x)}{Q(x)}, & \text{if } Q(x) \neq 0, \\ \infty, & \text{if } Q(x) = 0, \end{cases}$$

where it is assumed that P and Q never vanish simultaneoulsy.

It is clear that ℓ_P is a r.v. on $(X, 2^X, Q)$ with values in $[0, \infty]$, and, if $P \ll Q$, then

$$E(\ell_P) = \sum_{x \in X} \frac{P(x)}{Q(x)} Q(x) = 1.$$

If m is the distribution function of ℓ_P

$$m(t) = Q\{x \in X : \ell_P(x) \leq t\},$$

and m^{-1} is the right-continuous generalized inverse of m

$$m^{-1}(z) = \sup\{t : m(t) \leq z\},$$

then the concentration curve of P with respect to Q is the Lorenz curve of ℓ_P, that is

$$\phi_P(u) = \int_0^u m^{-1}(z)\, dz.$$

The rationale for calling ϕ_P the concentration function of P w.r.t. Q is the following. Let $A(u)$ be a set of the form

$$A(u) = \{x : \ell_P(x) \leq u\}.$$

If $Q(A(u)) = t$, then $P(A(u)) = \phi(t)$.

Therefore, for each set A containing the "poorest" points (in terms of ℓ_P), ϕ_P relates the probability mass concentrated on A by Q to the probability mass concentrated on A by P. For points outside the range of Q, a sort of randomization is performed. Actually, in the framework of hypothesis testing, ϕ is the so called α–β curve for testing the null hypothesis P versus the alternative Q, when randomization is allowed (see Lehmann (1986), pp. 76–77).

The following well known result will be used later.

PROPOSITION 2.1. *(Strassen (1965)) Let S, T be two r.v.'s such that $E(S) = E(T)$, and let $\psi_S(\psi_T)$ be the Lorenz curve of $S(T)$. Then $\psi_S(u) \le \psi_T(u)$ $\forall u \in [0,1]$ iff $\exists S', Z'$ defined on a common probability space, such that $S \overset{\mathcal{L}}{=} S'$ and $T \overset{\mathcal{L}}{=} E(S'|Z')$.*

3. Definition of the Dependence Ordering. Let X, Y be finite sets and let Q_X, Q_Y be two probability measures on X and Y, respectively, such that $Q_X(x) > 0$ $\forall x \in X$ and $Q_Y(y) > 0$ $\forall y \in Y$. Let $\mathcal{P}(Q_X, Q_Y)$ be the set of all probability measures on $X \times Y$, whose marginals are Q_X, Q_Y. Let μ be the product measure of Q_X and Q_Y. For each $P \in \mathcal{P}(Q_X, Q_Y)$, we can define a likelihood ratio ℓ_P of P w.r.t. the product measure μ.

Since every probability measure on a finite product space can be represented by a matrix, we will use the same symbol P for the probability measure and the corresponding matrix, namely

$$P = \{p_{ij}\}, \qquad p_{ij} = P(x_i, y_j) \qquad x_i \in X, y_j \in Y.$$

Let $P_1, P_2 \in \mathcal{P}(Q_X, Q_Y)$. Using an idea introduced by Cifarelli and Regazzini (1986) to study measures of dependence, a dependence ordering $\overset{\mathrm{D}}{\succeq}$ is defined as follows

$$P_1 \overset{\mathrm{D}}{\succeq} P_2 \quad \text{iff} \quad \phi_{P_1}(u) \le \phi_{P_2}(u) \quad \forall u \in [0,1].$$

The relation $\overset{\mathrm{D}}{\succeq}$ is a preorder (reflexive and transitive).

The rationale for the ordering $\overset{\mathrm{D}}{\succeq}$ is the following: If X, Y are independent, i.e. $P = \mu$, then $\ell_P \equiv 1$. The more X, Y are dependent, the more P differs from μ, the more ℓ_P is spread out, the lower ϕ_P is. In a situation of strong dependence between X and Y, many pairs (x, y) will have "small" probabilities, and some pairs will have "high" probabilities, where small and high is measured in terms of the corresponding mass concentrated by the product measure μ. Therefore ℓ_P assumes values far from one with high μ-probability, and ϕ_P tends to be low.

From now on, for the sake of brevity, we will use the symbol $\mathcal{P}(Q_X, Q_Y)$ to indicate the ordered space $(\mathcal{P}(Q_X, Q_Y), \overset{\mathrm{D}}{\succeq})$, unless otherwise stated.

PROPOSITION 3.1. *Let $P \in \mathcal{P}(Q_X, Q_Y)$ and let $R = \Pi_1 P \Pi_2$, where Π_1, Π_2 are permutation matrices. Then $\ell_P \overset{\mathcal{L}}{=} \ell_R$.*

PROOF. Let $L_P = \{\ell_P(x, y)\}$. Then $L_R = \Pi_1 L_P \Pi_2$. Hence the result. ||

PROPOSITION 3.2. *Let $P \in \mathcal{P}(Q_X, Q_Y)$ and let P^* be obtained by pooling two lines of P. Then*

$$\phi_{P^*}(u) \geq \phi_P(u) \qquad \forall u \in [0,1],$$

and

$$\phi_{P^*}(u) = \phi_P(u) \qquad \forall u \in [0,1]$$

iff the two lines are proportional.

PROOF. Without loss of generality, let x_1, x_2 be the two lines that are pooled, to form a new line x, say. Then

$$\begin{aligned}
\ell_{P^*}(x,y) &= \frac{P^*(x,y)}{Q_X(x)Q_Y(y)} \\
&= \frac{Q_X(x_1)}{Q_X(x_1)+Q_X(x_2)} \frac{P(x_1,y)}{Q_X(x_1)Q_Y(y)} \\
&\quad \frac{Q_X(x_2)}{Q_X(x_1)+Q_X(x_2)} \frac{P(x_2,y)}{Q_X(x_2)Q_Y(y)} \\
&= \alpha \ell_P(x_1,y) + (1-\alpha)\ell_P(x_2,y),
\end{aligned}$$

with

$$\alpha = \frac{Q_X(x_1)}{Q_X(x_1)+Q_X(x_2)}.$$

If we apply Proposition 1.1, we obtain $\phi_{P^*}(u) \geq \phi_P(u)$, $\forall u \in [0,1]$.
Of course, if the two lines are proportional, then

$$\ell_P(x_1,y) = \ell_P(x_2,y) = \ell_{P^*}(x,y);$$

therefore, $\phi_P = \phi_{P^*}$. ||

Related results can be found in Ali and Silvey (1965a), (1965b).
The space $\mathcal{P}(Q_X, Q_Y)$ has a minimum.

PROPOSITION 3.3. *Let $P^* \in \mathcal{P}(Q_X, Q_Y)$. Then $P \overset{D}{\succeq} P^* \quad \forall P \in \mathcal{P}(Q_X, Q_Y)$ iff $P^* = \mu$.*

PROOF. This is an immediate consequence of the fact that $\phi(u) = u \; \forall u \in [0,1]$ iff $P \equiv Q$ (Cifarelli and Regazzini (1987)). ||

PROPOSITION 3.4. *Let $P^* \in \mathcal{P}(Q_X, Q_Y)$ and $Q_X = Q_Y$. Then*

$$P^* \overset{D}{\succeq} P \quad \forall P \in \mathcal{P}(Q_X, Q_Y) \tag{1}$$

iff

$$P^*(x,x) = Q_X(x) = Q_Y(x) \qquad \forall x \in X \tag{2}$$

(modulo marginal preserving permutations of lines in P^*).

PROOF. (2) $\Longrightarrow$ (1): If we order the x's according to the value of Q_X increasingly, we have that

$$\max_{\substack{(x,y)\in X^2 \\ P\in\mathcal{P}(Q_X,Q_Y)}} \ell_P(x,y) = \ell_{P^*}(x_1,x_1) = 1/Q_X(x_1)$$

where $P^*(x_1,x_1) = Q_X(x_1)$. Now, if $P^*(x_1,x_1) = Q_X(x_1)$, then

$$P^*(x_1,y) = P^*(y,x_1) = 0 \qquad \forall\ y \neq x_1.$$

Iterating the procedure, we have

$$\max_{\substack{(x,y)\in(X\setminus\{x_1\})^2 \\ P\in\mathcal{P}(Q_X,Q_Y)}} \ell_P(x,y) = \ell_{P^*}(x_2,x_2) = 1/Q_X(x_2)$$

where $P^*(x_2,x_2) = Q_X(x_2)$, and so on.

Therefore $P^* \overset{\mathrm{D}}{\succeq} P \quad \forall P \in \mathcal{P}(Q_X,Q_Y)$, when P^* satisfies (2).

(1) $\Longrightarrow$ (2): Assume that $\tilde{P} \overset{\mathrm{D}}{\succeq} P \quad \forall P \in \mathcal{P}(Q_X,Q_Y)$, and $\tilde{P} \neq P^*$ (modulo permutations of lines), where P^* satisfies (2). Then $\tilde{P} \overset{\mathrm{D}}{\succeq} P^*$. Now, since $P^* \overset{\mathrm{D}}{\succeq} P$ $\forall P \in \mathcal{P}(Q_X,Q_Y)$, then $\phi_{\tilde{P}} = \phi_{P^*}$. But, then

$$\phi_{\tilde{P}}(1 - Q_X^2(x_1)) = \phi_{P^*}(1 - Q_X^2(x_1)) = 1 - Q_X(x_1),$$

which is possible only if

$$\tilde{P}(x_1,x_1) = P^*(x_1,x_1) = Q_X(x_1).$$

Furthermore,

$$\phi(1 - Q_X^2(x_1) - Q_X^2(x_2)) = \phi_{P^*}(1 - Q_X^2(x_1) - Q_X^2(x_2)) = 1 - Q_X(x_1) - Q_X(x_2),$$

which is possible only if $\tilde{P}(x_2,x_2) = P^*(x_2,x_2) = Q_X(x_2)$, etc. Iteration of the argument gives $\tilde{P} = P^*$. $\quad ||$

Proposition 3.4 shows that, when $Q_X = Q_Y$, then there exists a maximum on $\mathcal{P}(Q_X,Q_Y)$. This maximum is unique modulo permutations of lines. This is not the only case in which a maximum exists, as the following proposition shows.

PROPOSITION 3.5. *Consider $\mathcal{P}(Q_X,Q_Y)$, with Q_X, Q_Y uniform on X, Y, respectively. Let card$(X) = M$, card$(Y) = N$, $N > M$. Let P^* be defined as follows.*

Partition P^* into n square matrices of the following form: P_1 is the largest NW square sub-matrix of $P_0 \stackrel{def}{=} P^*$. $R_n =$ is the matrix obtained by deleting the lines of P_n from P_{n-1}, where P_n is the largest NW square submatrix of R_{n-1}. Let the elements of each matrix P_n be zero outside the main diagonal. There exists only one possible configuration of this type such that $P^* \in \mathcal{P}(Q_X, Q_Y)$. Furthermore $P^* \stackrel{D}{\succeq} P \quad \forall P \in \mathcal{P}(Q_X, Q_Y)$. If $P \equiv P^*$, then P is obtained from P^* via permutation of lines.

PROOF.

$$\max_{\substack{(x,y)\in X\times Y \\ P\in\mathcal{P}(Q_X,Q_Y)}} \ell_P(x,y) = \ell_{P^*}(x_1,y_1) = Q_Y(y_1)$$

If $P^*(x_1, y_1) = Q_Y(y_1)$, then $P(x, y_1) = 0 \quad \forall x \neq x_1$. Under this constraint,

$$\max_{\substack{(x,y)\in X\times \{Y\setminus y_1\} \\ P\in\mathcal{P}(Q_X,Q_Y)}} \ell_P(x,y) = \ell_{P^*}(x_2,y_2) = Q_Y(y_2)$$

and so on, for $y_1, \ldots, y_M$.

Given these constraints, the argument can be repeated for the matrices R_1, ..., R_n, and the result follows. ||

4. Tetrachoric Tables. In the case of 2×2 tables, a (unique) maximum exists.

THEOREM 4.1. *Let $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$.*

$$\begin{array}{lll} Q_X(x_1) = \alpha & Q_X(x_2) = \beta = 1-\alpha & \alpha \le \beta \\ Q_Y(y_1) = \gamma & Q_Y(y_2) = \delta = 1-\gamma & \gamma \le \delta \end{array}$$

If P^* has the following form

$$\begin{array}{ll} P^*(x_1,y_1) = \alpha \wedge \gamma & P^*(x_1,y_2) = (\alpha-\gamma)_+ \\ P^*(x_2,y_1) = (\gamma-\alpha)_+ & P^*(x_2,y_2) = \beta \wedge \delta, \end{array}$$

then P^* is the unique maximum w.r.t. $\stackrel{D}{\succeq}$.

PROOF. Without loss of generality, assume $\delta \le \beta$ (whence $\alpha \le \gamma$). Then

$$\begin{array}{ll} \phi_{P^*}(x) = 0 & 0 \le x \le \alpha\delta \\ = \frac{\gamma-\alpha}{\beta\gamma}x - \frac{\gamma-\alpha}{\beta\gamma}\alpha\delta & \alpha\delta \le x \le \alpha\delta + \beta\gamma \\ = \frac{x}{\gamma} - \frac{\delta}{\gamma}(\alpha+\gamma) & \alpha\delta + \beta\gamma \le x \le 1 - \beta\delta \\ = \frac{x}{\beta} - \frac{1}{\beta} + 1 & 1 - \beta\delta \le x \le 1. \end{array}$$

For any $P \in \mathcal{P}(Q_X, Q_Y)$, we have, for $i = 1, 2$,

$$\ell_P(x_i, y_1) \geq 1 \Longleftrightarrow \ell_P(x_i, y_2) \leq 1,$$

and, for $j = 1, 2$,

$$\ell_P(x_1, y_j) \geq 1 \Longleftrightarrow \ell_P(x_2, y_j) \leq 1.$$

Therefore, only two possible arrangements are possible:

$$\ell_P(x_2, y_2) \leq \ell_P(x_1, y_1) \leq \ell_P(x_2, y_1) \leq \ell_P(x_1, y_2) \tag{3}$$

or

$$\ell_P(x_1, y_2) \leq \ell_P(x_2, y_1) \leq \ell_P(x_1, y_1) \leq \ell_P(x_2, y_2). \tag{4}$$

Moreover, if we indicate $p_{ij} = P(x_i, y_j)$, $i, j = 1, 2$,

$$\begin{gathered} \alpha + \gamma - 1 \leq p_{11} \leq \alpha \\ 0 \leq p_{12} \leq 1 - \gamma \\ \gamma - \alpha \leq p_{21} \leq 1 - \alpha \\ 0 \leq p_{22} \leq 1 - \gamma. \end{gathered}$$

If P is such that (4) holds, then $\phi_P(x)$ is a broken line with corners in $x = \alpha\delta$, $\alpha\delta + \beta\gamma$, $1 - \beta\delta$.

$$\begin{gathered} \phi_P(\alpha\delta) = p_{12} \geq 0 = \phi_{P^*}(\alpha\delta) \\ \phi_P(\alpha\delta + \beta\gamma) = \gamma - \alpha + 2p_{12} \geq \gamma - \alpha = \phi_{P^*}(\alpha\delta + \beta\gamma) \\ \phi_P(1 - \beta\delta) = 1 - p_{22} \geq \gamma = \phi_{P^*}(1 - \beta\delta) \end{gathered}$$

If P is such that (3) holds, then $\phi_P(x)$ is a broken line with corners in $x = \beta\delta$, $\beta\delta + \alpha\gamma$, $1 - \alpha\delta$

$$\begin{gathered} \phi_P(\beta\delta) = p_{22} \geq 0 = \phi_{P^*}(\beta\delta) \\ \phi_P(\beta\delta + \alpha\gamma) = p_{11} + p_{22} \geq \alpha + \gamma - 1 \geq \phi_{P^*}(\beta\delta + \alpha\gamma) = (\gamma - \alpha)\tfrac{1-\beta\gamma}{\beta\gamma} \\ \phi_P(1 - \alpha\delta) = 1 - p_{12} \geq \gamma \geq \phi_{P^*}(1 - \alpha\delta) = \tfrac{1-\alpha\delta-\delta(\alpha+\gamma)}{\gamma}. \end{gathered}$$

Since $\phi_P(x) \geq \phi_{P^*}(x)$ for every x, where ϕ_P has a corner, and since ϕ_P, ϕ_{P^*} are Lorenz curves (i.e. increasing, convex, with $\phi_P(0) = \phi_{P^*}(0) = 0$, $\phi_P(1) = \phi_{P^*}(1) = 1$), then $\phi_P(x) \geq \phi_{P^*}(x)\ \forall\ x \in [0,1]$. Therefore, P^* is the maximum w.r.t. $\overset{\mathrm{D}}{\succeq}$. ||

It could be proven that, in the case of $2 \times n$ matrices, a maximum exists, but, in general, it is not unique (not even modulo permutations). No such maximum exists in general.

5. Comparison with Different Orderings. Joe (1985) proposed an ordering of dependence for contingency tables with fixed row and column sums. We indicate this ordering by $\overset{\mathrm{J}}{\succeq}$. Let $P_1,\ P_2 \in \mathcal{P}(Q_X, Q_Y)$. $P_1 \overset{\mathrm{J}}{\succeq} P_2$ iff $\mathrm{vec}(P_1) \overset{\mathrm{m}}{\succeq} \mathrm{vec}(P_2)$, where vec$(P)$ is the vector obtained by piling the columns of P, and $\overset{\mathrm{m}}{\succeq}$ is the usual majorization ordering (see Marshall and Olkin (1979)).

Joe's ordering has a major drawback as an ordering of dependence: it does not have a unique minimum corresponding to the product measure. This is due to the fact that the values in the probability matrices are not weighed according to their marginals. The ordering $\overset{\mathrm{J}}{\succeq}$ performs well in this respect only when the marginals are uniform. This suggests the connection between the orderings $\overset{\mathrm{D}}{\succeq}$ and $\overset{\mathrm{J}}{\succeq}$, at least when Q_X and Q_Y assume only rational values, which is not a restrictive assumption for contingency tables.

THEOREM 5.1. *Let Q_X, Q_Y assume only rational values. Let $P_1, P_2 \in \mathcal{P}(Q_X, Q_Y)$. Split rows and columns of P_1, P_2 in such a way that their marginals become uniform. (This is always possible, given rationality of the values of Q_X, Q_Y). Call $\tilde{P}_1, \tilde{P}_2 \in \mathcal{P}(\tilde{U}_n, \tilde{U}_m)$ these new matrices. Then $P_1 \overset{D}{\succeq} P_2$ iff $\tilde{P}_1 \overset{J}{\succeq} \tilde{P}_2$.*

PROOF. If the marginals are uniform, then $\ell_{\tilde{P}_1}$ assumes each value $\ell_{\tilde{P}_1}(x_i, y_j)$ with probability $(mn)^{-1}$. Analogously for $\ell_{\tilde{P}_2}$. Furthermore

$$\ell_{\tilde{P}_h}(i,j) = mnP_h(i,j) \quad h = 1, 2.$$

Therefore, if $\overset{\mathrm{L}}{\succeq}$ is the Lorenz ordering, we have

$$\ell_{\tilde{P}_1} \overset{\mathrm{L}}{\succeq} \ell_{\tilde{P}_2} \text{ iff } \ell_{\tilde{P}_1} \overset{\mathrm{m}}{\succeq} \ell_{\tilde{P}_2} \text{ iff } \mathrm{vec}(\tilde{P}_1) \overset{\mathrm{m}}{\succeq} \mathrm{vec}(\tilde{P}_2) \text{ iff } \tilde{P}_1 \overset{\mathrm{J}}{\succeq} \tilde{P}_2.$$

Since $\phi_{P_h} = \phi_{\tilde{P}_h}$, for $h = 1, 2$, then the result follows. ||

Theorem 5.1 shows that considering the likelihood ratio as the keypoint to study dependence is the same as transforming the matrix so that the marginals are uniform, and then vectoralize it. The idea of rendering the marginals uniform, when studying dependence, has been applied to the study of concordance (positive quadrant dependence).

6. Concordance and Dependence. Consider two linearly ordered measurable spaces $(X, \mathcal{X})$, $(Y, \mathcal{Y})$, and the class $\mathcal{P}(Q_X, Y_Y)$ of probability measures on $(X \times Y, \mathcal{X} \otimes \mathcal{Y})$ with marginals Q_X, Q_Y. For $P_1, P_2 \in \mathcal{P}(Q_X, Q_Y)$, P_1 is said more concordant (or more positive quadrant dependent) than P_2 ($P_1 \overset{\mathrm{c}}{\succeq} P_2$) iff

$$P_1\{(\xi, v) : \xi > x; v > y\} \geq P_2\{(\xi, v) : \xi > x; v > y\} \qquad \forall (x, y) \in X \times Y$$

(see for instance Yanagimoto and Okamoto (1969), Tchen (1980), Scarsini (1984), Kimeldorf and Sampson (1987)).

The class $\mathcal{P}(Q_X, Q_Y)$ has a minimum P_c^- and a maximum P_c^+ w.r.t. $\overset{\mathrm{c}}{\succeq}$, which are referred to as Fréchet bounds.

$$P_c^+\{(\xi,v):\xi>x;v>y\}=\min(Q_X\{\xi:\xi>x\},Q_Y\{v:v>y\}),$$
$$P_c^-\{(\xi,v):\xi>x;v>y\}=\max(Q_X\{\xi:\xi>x\}+Q_Y\{v:v>y\}-1,0).$$

In general P^+ and P_c^+ differ and there does not exist any permutation of rows and columns that make the two coincide.

EXAMPLE.

P^+

	0	1	2	
0	1/4			1/4
1		1/4		1/4
2			1/4	1/4
3	1/12	1/12	1/12	1/4
	1/3	1/3	1/3	

P_c^+

	0	1	2	
0	1/4			1/4
1	1/12	1/6		1/4
2		1/6	1/12	1/4
3			1/4	1/4
	1/3	1/3	1/3	

No permutation of lines leads from P^+ to P_c^+ or vice versa.

7. Infinite Spaces. Generalization to the case of infinite spaces requires some care in the definition of the generalized concentration curve and poses some problems in the interpretation of the ordering.

For the definition of the generalized concentration curve, again we use the results provided by Cifarelli and Regazzini (1987). Let $(X,\mathcal{X},Q_X)$, $(Y,\mathcal{Y},Q_Y)$ be two probability spaces and let $\mathcal{X},\mathcal{Y}$ contain the singletons. Let μ be the product measure of Q_X,Q_Y on $(X\times Y,\mathcal{X}\otimes\mathcal{Y})$. If $P\in\mathcal{P}(Q_X,Q_Y)$, there exists a partition $\{N,N^c\}\subset\mathcal{X}\otimes\mathcal{Y}$ and a nonnegative real function h on $X\times Y$, such that

$$P(A)=\int_{A\cap N^c}h(\mathbf{z})\ \mu(dz)+P_S(A\cap N)\ \ \forall\ A\in\mathcal{X}\otimes\mathcal{Y}$$

$$\mu(N)=0\ \ P_S(N)=P_S(X\times Y);$$

N is unique, modulo μ–null sets.

$$\ell_P(x,y)=\begin{cases}h(x,y), & \text{if } (x,y)\in N^c;\\ \infty, & \text{if } (x,y)\in N.\end{cases}$$

As before, let m be the d.f. of ℓ_P

$$m(t) = \mu\{(x,y) \in X \times Y : \ell_P(x,y) \leq t\}.$$

Then the generalized concentration curve of P w.r.t. μ is

$$\phi_P(u) = \int_0^u m^{-1}(t)\, dt \qquad u \in [0,1)$$

and we define $\phi_P(1) = 1$. Then ϕ_P is convex and increasing on $[0,1]$ and continuous on $[0,1)$. The jump $\phi_P(1) - \phi_P(1^-)$ represents the mass of P_S, which can assume any value in $[0,1]$. We have

$$\begin{array}{lll} \phi_P(u) = u & \forall u \in [0,1] & \text{iff} \quad P = \mu \\ \phi_P(u) = 0 & \forall u \in [0,1) & \text{iff} \quad P \perp \mu. \end{array}$$

The dependence ordering $\overset{\mathrm{D}}{\succeq}$ is defined as in Section 2. For $P_1, P_2 \in \mathcal{P}(Q_X, Q_Y)$, $P_1 \overset{\mathrm{D}}{\succeq} P_2$ iff $\phi_{P_1}(u) \leq \phi_{P_2}(u)$, $\forall u \in [0,1]$.

Some phenomena should warn against an acritical use of the ordering $\overset{\mathrm{D}}{\succeq}$ in this general situation. Let $X = Y = [0,1]$, $\mathcal{X} = \mathcal{Y} = \text{Bor}([0,1])$. Let $Q_X = Q_Y =$ Lebesgue measure on $[0,1]$. Let $P_k \in \mathcal{P}(Q_X, Q_Y)$ be defined as follows: P_k concentrates its mass uniformly on the functions

$$y = x + c, \quad c = \pm(1/2)^k \quad k \in I\!N,$$

with $(x,y) \in [0,1]^2$. Since $P_k \perp \mu$, $\forall\, k \in I\!N$, it holds that $\phi_{P_k}(u) = 0\ \forall\ u \in [0,1)$. But $P_k \xrightarrow{\mathcal{L}} \mu$. This fact shows that $\overset{\mathrm{D}}{\succeq}$ indicates how concentrated P is on sets of small μ–probability, rather than how dependent P is. These two concepts basically coincide when X and Y are finite, but they differ in general.

Joe (1987) considered an ordering $\overset{\mathrm{J}}{\succeq}$ of dependence for general probability measures on product spaces. This ordering generalizes the ordering $\overset{\mathrm{J}}{\succeq}$ for contingency tables. The main difference between $\overset{\mathrm{D}}{\succeq}$ and $\overset{\mathrm{J}}{\succeq}$ is that, in defining $\overset{\mathrm{J}}{\succeq}$, Joe does not consider the likelihood ratio of P w.r.t. μ, but a density w.r.t. a generic product measure (usually Lebesgue or counting measure). This implies the drawbacks that we noticed for the discrete case in Section 5.

8. Concluding Remarks. We have introduced an ordering of dependence on the set of probability measures on a product space, with fixed marginals. We have considered only products of two spaces. The generalization to n–fold product spaces is straightforward.

The ordering $\overset{\mathrm{D}}{\succeq}$ is based on the Lorenz curve of the likelihood ratio of a probability measure w.r.t the product measure of its marginals. Consider two probability measures $P_1 \in \mathcal{P}(Q_X, Q_Y)$, $P_2 \in \mathcal{P}(Q_W, Q_Z)$, where X, Y, W, Z are finite and Q_X, Q_Y, Q_W, Q_Z assume only rational values. The argument used in Section 5 shows that, by splitting lines of P_1, P_2, it is possible to obtain $\tilde{P}_1, \tilde{P}_2 \in \mathcal{P}(U_n, U_m)$, with U_n, U_m uniform, and such that $\ell_{P_1} \stackrel{\mathcal{L}}{=} \ell_{\tilde{P}_1}$, $\ell_{P_2} \stackrel{\mathcal{L}}{=} \ell_{\tilde{P}_2}$. Comparison of $\tilde{P}_1$, $\tilde{P}_2$ w.r.t. $\stackrel{\mathrm{D}}{\succeq}$ induces an analogous comparison for P_1, P_2. This suggests to remove the constraint of considering only probability measures with fixed marginals.

It is actually possible to show that, for any probability space $(X \times Y, \mathcal{X} \otimes \mathcal{Y}, P)$, with $P \in \mathcal{P}(Q_X, Q_Y)$, there exists $([0,1]^2, \mathrm{Bor}([0,1]^2), \pi)$ such that π has uniform marginals and $\ell_P \stackrel{\mathcal{L}}{=} \ell_\pi$.

Transforming a probability measure on a product space into a probability measure on the unit square with uniform marginals (in a suitable dependence preserving way) is a common idea in the study of concordance of random variables. It is interesting to see that it appears here, given that concordance and dependence are different concepts. The linear order structure on the spaces X, Y is a basic ingredient for the definition of concordance, whereas any such structure is neglected when dealing with dependence.

Once more we want to emphasize that, while in the case of finite spaces the preorder $\stackrel{\mathrm{D}}{\succeq}$ is a bona fide dependence ordering, in the general case the intuitive rationale for the ordering fails, especially when a probability measure is not dominated by the product measure of its marginals.

REFERENCES

ALI, S.M. and SILVEY, S.D. (1965a). Association between random variables and the dispersion of a Radon-Nikodym derivative. *J. Roy. Statist. Soc.* B **27** 100–107.

ALI, S.M., and SILVEY, S.D. (1965b). A further result on the relevance of the dispersion of a Radon-Nikodym derivative to the problem of measuring association. *J. Roy. Statist. Soc.* B **27** 108–110.

CIFARELLI, D.M. and REGAZZINI, E. (1986). Sulla funzione di concentrazione e sul suo ruolo nella statistica descrittiva. In *Atti della XXXIII Riunione Scientifica della Società Italiana di Statistica*, Bari, 347–352.

CIFARELLI, D.M. and REGAZZINI, E. (1987). On a general definition of concentration function. *Sankhyā B* **49** 307–319.

JOE, H. (1985). An ordering of dependence for contingency tables. *Linear Algebra and Its Applications* **70** 89–103.

JOE, H. (1987). Majorization, randomness and dependence for multivariate distributions. *Ann. Prob.* **15** 1217–1225.

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

LEHMANN, E.L. (1986). *Testing Statistical Hypotheses.* 2nd Edition. Wiley, New York.

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

REGAZZINI, E. (1988). A few remarks on concentration comparisons. General Theory. Dipartimento di Matematica, Università di Milano.

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

STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.

TCHEN, A.H. (1980) Inequalities for distributions with given marginals. *Ann. Prob.* **8** 814–827.

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

DIPARTIMENTO DI SCIENZE ATTUARIALI
UNIVERSITA LA SAPIENZA
VIA DEL CASTRO LAURENZIANO 9
I–00161 ROMA, ITALY