A DECISION APPROACH TO ORDERING STOCHASTIC DEPENDENCE

By T. Bromek and T. Kowalczyk

Polish Academy of Sciences

Editors' Note: This paper is being published posthumously and in dedicated remembrance of Tadeusz Bromek, who died in an automobile accident on August 23, 1988 in Warsaw. He was very active in the Polish statistical community as well as the international statistical community.

An ordering of global dependence is defined on the basis of a natural ordering of pairs of distributions describing two classes of objects. Its properties are investigated; the links with orderings of multinormal and 2×2 distributions are shown.

1. Introduction. Traditionally, two types of stochastic dependence of components of a vector $\mathbf{X}$ have been distinguished in statistical literature, namely monotone and global dependence. Orderings for monotone dependence were considered by many authors; an overview was given by Yanagimoto (1990). Kimeldorf and Sampson (1987) introduced an axiomatic approach to the matter of orders of monotone dependence. The abundance of formalizations for orderings of monotone dependence contrasts with the silence concerning orderings of global dependence (see Dabrowska (1985)). It seems that a good starting point could be two-class discriminant analysis, with one class reserved for the distribution of the vector $\mathbf{X}$ and the other class reserved for the respective product of marginal distributions of $\mathbf{X}$. Thus, a natural ordering of pairs of distributions describing two classes of objects (Niewiadomska-Bugaj (1987)), called prognostic ordering and denoted $\leq_p$, may be a base to define an ordering of global dependence, called global ordering and denoted $\leq_g$.

In Section 2 we recall the definition of $\leq_p$ and prove some of its new properties. Section 3 contains the definition and properties of $\leq_g$.

2. Prognostic Order $\leq_p$. Consider a two-class discriminant problem corresponding to a pair $(\mathbf{Z}_1, \mathbf{Z}_2)$ where $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are random vectors supported on $\mathcal{Z} \subset R^k$. Distribution P_i of $\mathbf{Z}_i$ describes the ith class of the considered population $(i = 1, 2)$. Let a classification rule for $(\mathbf{Z}_1, \mathbf{Z}_2)$ be a Borel measurable function

AMS 1980 subject classifications. 62H20.

Key words and phrases. Global dependence, monotone dependence, likelihood ratio, two-class discriminant analysis.

$\partial : \mathcal{Z} \to [0,1]$, where $\partial(\mathbf{z})$ is the probability of classifying an object with features vector $\mathbf{z}$ to the 1st class. Performance of ∂ is measured by the error rates $a_{12}(\partial)$ and $a_{21}(\partial)$:

$$a_{12}(\partial) = \int_{\mathcal{Z}} (1 - \partial(\mathbf{z}))dP_1, \quad a_{21}(\partial) = \int_{\mathcal{Z}} \partial(\mathbf{z})dP_2.$$

For two pairs of random vectors $(\mathbf{Z}_1, \mathbf{Z}_2)$, $(\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)$, there is defined an ordering with respect to their discriminant powers:

> $(\mathbf{Z}_1, \mathbf{Z}_2) \leq_p (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)$ iff for every decision rule ∂ for $(\mathbf{Z}_1, \mathbf{Z}_2)$ there exists a decision rule $\tilde{\partial}$ for $(\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)$ such that $a_{12}(\tilde{\partial}) \leq a_{12}(\partial)$ and $a_{21}(\tilde{\partial}) \leq a_{21}(\partial)$.

It is clear that $\leq_p$ is a preorder. Thus, a relation $\underset{p}{\approx}$ defined by:

$$(\mathbf{Z}_1, \mathbf{Z}_2) \underset{p}{\approx} (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2) \text{ iff } (\mathbf{Z}_1, \mathbf{Z}_2) \leq_p (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2) \text{ and } (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2) \leq_p (\mathbf{Z}_1, \mathbf{Z}_2)$$

is an equivalence.

THEOREM 1.

(i). For any Borel measurable function $f : \mathcal{Z} \to R^k$

$$1^\circ \ (f(\mathbf{Z}_1), f(\mathbf{Z}_2)) \leq_p (\mathbf{Z}_1, \mathbf{Z}_2);$$

if f *is an injection, then*

$$2^\circ \ (f(\mathbf{Z}_1), f(\mathbf{Z}_2)) \underset{p}{\approx} (\mathbf{Z}_1, \mathbf{Z}_2);$$

(ii). $(\mathbf{Z}_1, \mathbf{Z}_2)$ *is a minimal element of* $\leq_p$ *iff* $\mathbf{Z}_1$ *and* $\mathbf{Z}_2$ *are distributed identically;*

(iii). $(\mathbf{Z}_1, \mathbf{Z}_2)$ *is a maximal element of* $\leq_p$ *iff there exist a set* $A \subset \mathcal{Z}$ *such that* $P_1(A) = 1$ *and* $P_2(A) = 0$.

PROOF.

(i). To any classification rule $\bar{\partial}$ for $(f(\mathbf{Z}_1), f(\mathbf{Z}_2))$ there corresponds a classification rule $\partial = \bar{\partial} \circ f$ for $(\mathbf{Z}_1, \mathbf{Z}_2)$ with error rates a_{12} and a_{21}, respectively equal to those of $\bar{\partial}$. Thus 1° holds. Then, applying to $(f(\mathbf{Z}_1), f(\mathbf{Z}_2))$ the inverse function $f^{-1} : f(\mathcal{Z}) \to \mathcal{Z}$ (which is Borel measurable since f is the injection), we get $(\mathbf{Z}_1, \mathbf{Z}_2) \leq_p (f(\mathbf{Z}_1), f(\mathbf{Z}_2))$ which implies equivalence 2°.

(ii). Obviously, for any pair $(\mathbf{Z}_1, \mathbf{Z}_2)$ and any classification rule ∂ for that pair,

$$a_{12}(\partial) + a_{21}(\partial) = 1 \quad \text{iff} \quad \int_{\mathcal{Z}} \partial(\mathbf{z}) dP_1 = \int_{\mathcal{Z}} \partial(\mathbf{z}) dP_2.$$

Therefore $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are distributed identically iff, for any classification rule ∂ applied to $(\mathbf{Z}_1, \mathbf{Z}_2)$, $a_{12}(\partial) + a_{21}(\partial) = 1$. On the other hand, for any $(\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)$ and any constant rule $\partial(\tilde{\mathbf{z}}) = \mathcal{L}$, $0 \leq \mathcal{L} \leq 1$, $a_{12}(\partial) = \mathcal{L}$ and $a_{21}(\partial) = 1 - \mathcal{L}$; hence $(\mathbf{Z}_1, \mathbf{Z}_2) \leq_p (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)$ iff $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are distributed identically.

(iii). Let $A \subset \mathcal{Z}$ satisfy $P_1(A) = 1$, $P_2(A) = 0$, and let ∂ be a rule such that $\partial(\mathbf{z}) = 1$ if $\mathbf{z} \in A$ and $\partial(\mathbf{z}) = 0$ if $z \in \mathcal{Z} \setminus A$. Then $a_{12}(\partial) = a_{21}(\partial) = 0$, and hence $(\mathbf{Z}_1, \mathbf{Z}_2)$ is a maximal element of $\leq_p$.

Conversely, if for $(\mathbf{Z}_1, \mathbf{Z}_2)$ there exists a rule ∂ such that $a_{12}(\partial) = a_{21}(\partial) = 0$, then for a set $A = \{z \in \mathcal{Z}; \partial(\mathbf{z}) > 0\}$ we have $P_2(A) = 0$, $P_1(A) \geq \int_A \partial(\mathbf{z}) dP_1 = \int_{\mathcal{Z}} \partial(\mathbf{z}) dP_1 = 1 - a_{21}(\partial) = 1$. ||

It follows from the Neyman-Pearson Lemma that this set consists of threshold rules based on the likelihood ratio $h = f_2/f_1$, where f_i is a density function of $\mathbf{Z}_i$ with respect to some measure ν (we set $h(\mathbf{z}) = \infty$ if $f_1(\mathbf{z}) = 0$). These rules are defined by

$$\partial(\mathbf{z}) = \begin{cases} 1 & \text{if } h(\mathbf{z}) < \gamma \\ s & \text{if } h(\mathbf{z}) = \gamma \\ 0 & \text{if } h(\mathbf{z}) > \gamma, \end{cases}$$

for $\gamma > 0$ and $s \in [0, 1]$.

Now, let us extend this set of rules admitting $\gamma = 0$ and $\gamma = +\infty$, and let

$$\begin{aligned} C_{(\mathbf{Z}_1, \mathbf{Z}_2)} &= \{P_1(h(\mathbf{z}) > \gamma) + (1 - s)P_1(h(\mathbf{z}) = \gamma), P_2(h(\mathbf{z}) < \gamma) + sP_2(h(\mathbf{z}) = \gamma); \\ &\quad\; 0 \leq \gamma \leq \infty,\ 0 \leq s \leq 1\}. \end{aligned}$$

It is easy to see that $C_{(\mathbf{Z}_1, \mathbf{Z}_2)}$ is a curve joining points (0,1) and (1,0) which is continuous, convex, and nonincreasing. It will be called the divergence curve for $(\mathbf{Z}_1, \mathbf{Z}_2)$. Obviously, $C_{(\mathbf{Z}_1, \mathbf{Z}_2)}$ is the set of errors $(a_{12}(\partial), a_{21}(\partial))$ for threshold rules from the extended set of rules with minimal error rates.

We will say that $C_{(\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)} \leq C_{(\mathbf{Z}_1, \mathbf{Z}_2)}$ iff for any $(x_1, x_2) \in C_{(\mathbf{Z}_1, \mathbf{Z}_2)}$ there exists $(x_1, \tilde{x}_2) \in C_{(\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)}$ such that $x_2 \geq \tilde{x}_2$.

The following is an equivalent definition of $\leq_g$:

$$(\mathbf{Z}_1, \mathbf{Z}_2) \leq_p (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2) \text{ iff } C_{(\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2)} \leq C_{(\mathbf{Z}_1, \mathbf{Z}_2)}.$$

3. Global Dependence Order $\leq_g$. Given a random vector X, consider a discriminant problem corresponding to a pair $({}^{\perp}\mathbf{X}, \mathbf{X})$, where ${}^{\perp}\mathbf{X}$ is a random vector distributed according to the product of marginal distributions of $\mathbf{X}$. For any pair of random vectors $\mathbf{X}$ and $\mathbf{Y}$, we define the following ordering $\leq_g$ of global dependence:

$$\mathbf{X} \leq_g \mathbf{Y} \text{ iff } ({}^{\perp}\mathbf{X}, \mathbf{X}) \leq_p ({}^{\perp}\mathbf{Y}, \mathbf{Y}).$$

THEOREM 2.

(i). $\leq_g$ *is a preorder.*

(ii). For any random vectors $\mathbf{X}(n-\dim)$ *and* $\mathbf{Y}(k-\dim)$ *supported on* $\mathcal{X}$ *and* $\mathcal{Y}$, *respectively,* 1° *for Borel measurable functions* $f : \mathcal{X} \to R^n$, $g : \mathcal{Y} \to R^k$ *such that* $f(x_1, \ldots, x_n) = (f_1(x_1), \ldots, f_n(x_n))$, $g(y_1, \ldots, y_k) = (g_1(y_1), \ldots, g_k(y_k))$, *where* f_i, g_j *are injections,*

$$\mathbf{X} \leq_g \mathbf{Y} \textit{ iff } f(\mathbf{X}) \leq_g g(\mathbf{Y}), \tag{1}$$

2° *for any n-elements and k-elements permutations* $\Pi^{(n)}$ *and* $\Pi^{(k)}$

$$\mathbf{X} \leq_g \mathbf{Y} \textit{ iff } \Pi^{(n)}(\mathbf{X}) \leq_g \Pi^{(k)}(\mathbf{Y}). \tag{2}$$

(iii). $\mathbf{X}$ *is a minimal element of* $\leq_g$ *iff* $\mathbf{X} =^{\perp} \mathbf{X}$.

(iv). For $\mathbf{X}$ *with continuous marginal distribution: if the distribution of* $\mathbf{X}$ *is degenerate, then* $\mathbf{X}$ *is a maximal element of the preorder* $\leq_g$.

(v). For normally distributed k-dimensional random vectors $\mathbf{X}$ *and* $\mathbf{Y}$ *with identical sets of orthogonal eigenvectors for the correlation matrices of* $\mathbf{X}$ *and* $\mathbf{Y}$, *and for each pair of eigenvalues* β_i, $\bar{\beta}_i$ *of correlation matrices of* $\mathbf{X}$ *and* $\mathbf{Y}$, *respectively:*

$$\textit{if } \bar{\beta}_i > \beta_i > 1 \textit{ or } 1 > \beta_i > \bar{\beta}_i \;\; i = 1, \ldots, k \textit{ then } \mathbf{X} \leq_g \mathbf{Y}.$$

PROOF. (ii). It follows from Th.1 (i) that for an injection $f : \mathcal{X} \to R^n$, we have $({}^{\perp}\mathbf{X}, \mathbf{X}) \underset{p}{\approx} (f({}^{\perp}\mathbf{X}), f(\mathbf{X}))$. On the other hand for a function $f : \mathcal{X} \to R^n$ independently transforming vector components, we have $f({}^{\perp}\mathbf{X}) \sim^{\perp} f(\mathbf{X})$. Therefore, $({}^{\perp}\mathbf{X}, \mathbf{X}) \underset{p}{\approx} ({}^{\perp}f(\mathbf{X}), f(\mathbf{X}))$. Analogously, $({}^{\perp}\mathbf{Y}, \mathbf{Y}) \underset{p}{\approx} ({}^{\perp}g(\mathbf{Y}), g(\mathbf{Y}))$. Thus (1) holds due to the definition and transitivity of $\leq_g$. The proof of (2) is analogous since for any n-element permutation Π, $\Pi({}^{\perp}\mathbf{X}) =^{\perp} \Pi(\mathbf{X})$.

Proofs of (iii) and (iv) follow immediately from Th.1 (ii) and (iii), respectively.

(v). By (ii), we may restrict consideration to the vectors with standardized marginals. We shall show that, under the assumptions of (v), the error rate $a_{21}(\partial)$ of any threshold rule ∂ for $({}^{\perp}\mathbf{X}, \mathbf{X})$ would diminish and $a_{12}(\partial)$ would not change when ∂ was applied for $({}^{\perp}\mathbf{Y}, \mathbf{Y})$.

Let Σ and $\bar{\Sigma}$ denote the correlation matrices for $\mathbf{X}$ and $\mathbf{Y}$, respectively. The likelihood ratio of $\mathbf{X}$ against ${}^{\perp}\mathbf{X}$ is

$$h(\mathbf{x}) = \mid \Sigma \mid^{-1/2} \exp((-1/2)\mathbf{x}'(\Sigma^{-1} - I)\mathbf{x}).$$

Let ∂ be a threshold rule such that $\partial(\mathbf{x}) = 1$ if $h(\mathbf{x}) < \gamma$ and $\partial(\mathbf{x}) = 0$ if $h(\mathbf{x}) > \gamma$. Let $a_{ij}(\partial)$ denote the error rates of ∂ for $({}^{\perp}\mathbf{X}, \mathbf{X})$ and $\bar{a}_{ij}(\partial)$ be the error rates of ∂ for $({}^{\perp}\mathbf{Y}, \mathbf{Y})$. Then

$$a_{12}(\partial) = 1 - \int_{D_\gamma} \ldots \int (2\pi)^{-n/2} \exp((-1/2)\mathbf{x}'\mathbf{x}) dx_1 \ldots dx_k,$$
$$a_{21}(\partial) = \int_{D_\gamma} \ldots \int (2\pi)^{-n/2} \mid \Sigma \mid^{-1/2} \exp((-1/2)\mathbf{x}'\Sigma^{-1}\mathbf{x}) dx_1 \ldots dx_k$$

and

$$\bar{a}_{12}(\partial) = 1 - \int_{D_\gamma} \ldots \int (2\pi)^{-n/2} \exp((-1/2)\mathbf{y}'\mathbf{y}) dy_1 \ldots dy_k$$
$$\bar{a}_{21}(\partial) = \int_{D_\gamma} \ldots \int (2\pi)^{-n/2} \mid \bar{\Sigma} \mid^{-1/2} \exp((-1/2)\mathbf{y}'\bar{\Sigma}^{-1}\mathbf{y}) dy_1 \ldots dy_k,$$

where

$$D_\gamma = \{\mathbf{x} : \mid \Sigma \mid^{-1/2} \exp((-1/2)\mathbf{x}'(\Sigma^{-1} - I)\mathbf{x} < \gamma\}.$$

Putting

$$\mathbf{y} = \bar{\Sigma}^{1/2}\Sigma^{-1/2}\mathbf{x}, \quad \text{we get}$$
$$\bar{a}_{21}(\partial) = \int_{\bar{D}_\gamma} \ldots \int (2\pi)^{-n/2} \mid \Sigma \mid^{-1/2} \exp((-1/2)\mathbf{x}'\Sigma^{-1}\mathbf{x}) dx_1 \ldots dx_k,$$

where

$$\bar{D}_\gamma = \{\mathbf{x} : \mid \Sigma \mid^{-1} \exp((-1/2)\mathbf{x}'\Sigma^{-1/2}\bar{\Sigma}^{1/2}(\Sigma^{-1} - I)\bar{\Sigma}^{1/2}\Sigma^{-1/2}\mathbf{x} < \gamma\}.$$

We shall show that $\bar{D}_\gamma \subset D_\gamma$. Let $\mathbf{x}^1, \ldots, \mathbf{x}^k$ be the orthonormal set of eigenvectors, common for Σ and $\bar{\Sigma}$. Substituting $\mathbf{x} = \Sigma_{i=1}^k \zeta_i \mathbf{x}^i$ to D_γ and $\bar{D}_\gamma$ we get:

$$\mathbf{x} \in D_\gamma \quad \text{iff} \quad (2\pi)^{-n/2} \mid \Sigma \mid^{-1} \exp((-1/2)\Sigma_{i=1}^k (\beta_i^{-1} - 1)\zeta_i^2 < \gamma,$$
$$\mathbf{x} \in \bar{D}_\gamma \quad \text{iff} \quad (2\pi)^{-n/2} \mid \Sigma \mid^{-1} \exp((-1/2)\Sigma_{i=1}^k (\bar{\beta}_i/\beta_i)(\beta_i^{-1} - 1)\zeta_i^2 < \gamma.$$

Under the assumptions of (v), $\bar{D}_\gamma \subset D_\gamma$ and $\bar{a}_{21}(\partial) \leq a_{21}(\partial)$. Thus, $({}^{\perp}\mathbf{X}, \mathbf{X}) \leq_p ({}^{\perp}\mathbf{Y}, \mathbf{Y})$ since $\bar{a}_{12}(\partial) = a_{12}(\partial)$. $\quad \|$

COROLLARY 1. *Let* $\mathbf{X}, \mathbf{Y}$ *be* $k -$ dim *normally distributed random vectors and let* ζ_{ij} *and* $\bar{\zeta}_{ij}$ *be the elements of the correlation matrices of* $\mathbf{X}$ *and* $\mathbf{Y}$.

(i). For $k = 2$, $\mathbf{X} \leq_g \mathbf{Y}$ *iff* $\mid \zeta_{12} \mid \leq \mid \bar{\zeta}_{12} \mid$.

(ii). For $k = 3$, *if* $\zeta_{12}\zeta_{23}\zeta_{31} = 0$ *and* $\bar{\zeta}_{ij} = \eta\zeta_{ij}$, *for* $1 \leq \eta \leq (\zeta_{12}^2 + \zeta_{23}^2 + \zeta_{13}^2)^{-1/2}$, $i \neq j, i, j = 1, 2, 3$, *then* $\mathbf{X} \leq_g \mathbf{Y}$.

Now, we will consider pairs of binary random variables $\mathbf{X} = (X^{(1)}, X^{(2)})$. A 2×2 distribution of $\mathbf{X}$ is specified by $P = (p_{ij})$, p_{ij} being the probability that $X^{(1)} = i$, $X^{(2)} = j$ $(i, j = 1, 2)$. The set of 2×2 distributions will be denoted $\mathcal{P}_{2\times 2}$.

A distribution $P \in \mathcal{P}_{2\times 2}$ is either positive dependent $(p_{11}p_{22} \geq p_{12}p_{21})$ or negative dependent $(p_{11}p_{22} \leq p_{12}p_{21})$. A natural monotone ordering $\leq_m$ of $\mathcal{P}_{2\times 2}$ is given by:

$$P \leq_m P' \text{ if } p_{11} \leq p'_{11},\ p_{22} \leq p'_{22},\ p_{12} \geq p'_{12},\ p_{21} \geq p'_{21}$$

and P is positive dependent, or

$$p_{11} \geq p'_{11},\ p_{22} \geq p'_{22},\ p_{12} \leq p'_{12},\ p_{21} \leq p'_{21}$$

and P is negative dependent.
We shall show that $P \leq_m P'$ implies $P \leq_g P'$ (but the opposite implication is not true).

LEMMA 1. *Let $Q = (q_{ij})$, $R = (r_{ij})$ belong to $\mathcal{P}_{2\times 2}$ and let $\mathcal{E}_{ii} = r_{ii} - q_{ii}$ for $i = 1,2$, $\mathcal{E}_{ij} = q_{ij} - r_{ij}$ for $i \neq j, i,j = 1,2$.*

If $q_{11}q_{22} \geq q_{12}q_{21}$ ($q_{11}q_{22} \leq q_{12}q_{21}$); $\mathcal{E}_{ij} \geq 0$, $i,j = 1,2$ ($\mathcal{E}_{ij} \leq 0$, $i,j = 1,2$); and $\min(\mathcal{E}_{ij}; i,j = 1,2) = 0$ ($\max(\mathcal{E}_{ij}; i,j = 1,2) = 0$), then

$$\frac{r_{ij}}{r_{i.}r_{.j}} \leq \frac{q_{ij}}{q_{i.}q_{.j}},\ \frac{r_{ii}}{r_{i.}r_{.i}} \geq \frac{q_{ii}}{q_{i.}q_{.i}} \quad i \neq j, i,j = 1,2. \tag{3}$$

PROOF. Replacing in (3) r_{ij} by $q_{ij} - \mathcal{E}_{ij}$ for $i \neq j$, $i,j = 1,2$, and r_{ii} by $q_{ii} + \mathcal{E}_{ii}$ for $i = 1,2$, and taking into account that $\mathcal{E}_{11} + \mathcal{E}_{22} = \mathcal{E}_{12} + \mathcal{E}_{21}$, we obtain inequalities equivalent to (3), which obviously hold. ||

THEOREM 3. *If $P \leq_m P'$ then $P \leq_g P'$.*

PROOF. Let $\mathbf{X} \sim P \in \mathcal{P}_{2\times 2}$. Instead of $C_{(\perp\mathbf{X},\mathbf{X})}$ we will write $C_{(P)}$. Then the divergence curve $C_{(P)}$ is convex and piece-wise linear; it joins points (0,1) and (1,0) and consists of four segments with slopes equal to $(-p_{ij}/p_{i.}p_{.j})$, ordered nondecreasingly.

Let P, $P' \in \mathcal{P}_{2\times 2}$ and let a_k, a'_k, $k = 1,\ldots,4$, be the slopes of consecutive segments of $C_{(P)}$ and $C_{(P')}$. Clearly, $C_{(P')} \leq C_{(P)}$ if

$$a'_k \leq a_k \text{ for } k = 1,2, \quad a'_k \geq a_k \text{ for } k = 3,4 \tag{4}$$

If P is positive dependent then

$$\begin{aligned} -a_1 &= \max(p_{11}/p_{1.}p_{.1},\ p_{22}/p_{2.}p_{.2}), \\ -a_2 &= \min(p_{11}/p_{1.}p_{.1},\ p_{22}/p_{2.}p_{.2}), \\ -a_3 &= \max(p_{12}/p_{1.}p_{.2},\ p_{21}/p_{2.}p_{.1}), \\ -a_4 &= \min(p_{12}/p_{1.}p_{.2},\ p_{21}/p_{2.}p_{.1}), \end{aligned}$$

and a'_k are expressed analogously. To prove (4), it suffices to show that

$$\frac{p'_{ij}}{p'_{i.}p'_{.j}} \leq \frac{p_{ij}}{p_{i.}p_{.j}},\ \frac{p'_{ii}}{p'_{i.}p'_{.i}} \geq \frac{p_{ii}}{p_{i.}p_{.i}} \quad i \neq j, i,j = 1,2. \tag{5}$$

Denote $\mathcal{E}_{ij} = | p'_{ij} - p_{ij} |$, $i,j = 1,2$, $\mathcal{E}_o = \min(\mathcal{E}_{ij}; i,j = 1,2)$, $p^o_{ii} = p_{ii} + \mathcal{E}_o$, $p^o_{ij} = p_{ij} - \mathcal{E}_o$, $i \neq j$, $i,j = 1,2$, $P^o = (p^o_{ij})$. Then

$$P' = \begin{pmatrix} p_{11} + \mathcal{E}_{11}, & p_{12} - \mathcal{E}_{12} \\ p_{21} - \mathcal{E}_{21}, & p_{22} + \mathcal{E}_{22} \end{pmatrix} = \begin{pmatrix} p^o_{11} + (\mathcal{E}_{11} - \mathcal{E}_o), & p^o_{12} - (\mathcal{E}_{12} - \mathcal{E}_o) \\ p^o_{21} - (\mathcal{E}_{21} - \mathcal{E}_o), & p^o_{22} + (\mathcal{E}_{22} - \mathcal{E}_o) \end{pmatrix}.$$

Since $P \leq_m P^o$ and $p^o_{1.} = p_{1.}, p^o_{.1} = p_{.1}$, we see that for P and P^o the inequalities in (5) hold with (P, P') replaced by (P, P^o). By Lemma 1, since $\min\{\mathcal{E}_{ij} - \mathcal{E}_o; i,j = 1,2\} = 0$, inequalities (5) hold with (P, P') replaced by (P^o, P'). But P^o was constructed so that $P \leq_m P^o \leq_m P'$. Thus, inequalities (5) hold.

The proof for negative dependent distribution can be easily reduced to that for positive dependent one. ||

REFERENCES

DABROWSKA, D. (1985). Descriptive parameters of location, dispersion and stochastic dependence. *Math. Operationsforsch Stat.* **16** 63–88.

KIMELDORF, G. and SAMPSON, A.R. (1987). Positive dependence orderings. *Ann. Inst. Stat. Math. A* **39** 113–118.

NIEWIADOMSKA-BUGAJ, M. (1987). Standard families of two-class discriminant problems with deferred decision. *Proceedings of Conference on Discriminant Analysis, Cluster Analysis and Other Methods of Data Classification*, 223–229.

YANAGIMOTO, T. (1990). Dependence ordering in statistical models and other notions. In *Topics in Statistical Dependence* (H. Block, A.R. Sampson, and T. Savits, eds.). IMS Lecture Notes–Monograph Series, Hayward, CA.

INSTITUTE OF COMPUTER SCIENCE
POLISH ACADEMY OF SCIENCES
WARSAW, POLAND