

A partial ordering of dependence for contingency tables

Francesca Greselin^{*}

Michele Zenga^{*}

Summary:

For two-way contingency tables, whose row and column variables have nominal categories, we introduce a new partial order according to the strength of dependence within the class of all contingency tables with given margins. The order holds among tables which share the same configuration of signs in the deviation from independence (contingencies) and takes into account the absolute value of these contingencies. Our order, which seems to possess a more direct and intuitive interpretation relative to similar ones available in the literature, is characterized by a sequence of frequency transfers over a closed path, which involve an even number of cells and leave the marginal distributions unchanged. We derive necessary and sufficient conditions for maximal (minimal) tables, according to our notion of dependence. Furthermore a simple algorithm is provided to check these conditions. Finally, an associated measure of dependence is briefly mentioned.

Keywords:

Dependence ordering, stochastic ordering, contingency tables, measures of dependence, bivariate distributions with given margins.

1. Introduction

The literature on dependence orderings is quite extensive and it deals mainly with ordinal qualitative and quantitative variables. In this framework, “two

^{*} Dipartimento di Metodi Quantitativi per le Scienze Economiche e Aziendali, Università degli Studi di Milano Bicocca, Piazza dell'Ateneo Nuovo 1 - 20126 Milano
(e-mail: francesca.greselin@unimib.it, michele.zenga@unimib.it)

The present work reflects the common thinking of the authors, even if, more specifically, M. Zenga wrote the Introduction, the beginning of Section 3 and the Conclusions, while F. Greselin wrote the other parts of the work.

random variables X and Y are said to be positively dependent if either random variable being large probabilistically indicates that the other random variable is large” (Kimeldorf and Sampson, 1987). This concept of positive dependence has played a fundamental role in many recent new ideas in statistics. Far from attempting an exhaustive review, we will briefly recall here that several notions of positive dependence were proposed by Lehmann (1966), Esary, Proshan and Walkup (1967), Karlin (1968), Shaked (1977), among many others. Shaked and Shantikumar (1994) provided an extensive review of all these, and presented most theoretical results derived from the study of their respective implications.

Successively, a new approach to the study of positive dependence has begun. The basic idea is to compare two bivariate distributions having the same pair of margins in order to determine whether one distribution is more positively dependent than the other one. Thus it is to attempt a partial order of the distributions according to their degree of positive dependence. Tchen’s more concordant ordering (1980), Rinott and Pollak’s covariance ordering (1980) and Shaked and Tong’s ordering for multivariate r.v. (1985) are the better known examples of such orderings.

To date, in the literature relatively few works consider dependence orderings for nominal categorical variables, as we intend to study in the present paper.

In this context, an interesting approach to dependence orderings is presented in Forcina and Giovagnoli (1987), where the authors proposed the “asymmetric dependence pre-ordering”, the “weak interdependence pre-ordering” and the “strong interdependence pre-ordering”. Their orderings are based on linear transformations on matrices, which produce a loss of information (hence a loss of dependence) in the rows or columns of the transformed matrix. Unfortunately, if the basic idea is still to compare two bivariate distributions having the same pairs of margins, these proposals are not useful: they yield the modification of marginal distributions and the transformation of integer joint frequencies into real ones.

Earlier, Joe (1985) proposed to compare two distributions by the Lorenz ordering of the vectorialized $r \times c$ entries. More recently, Scarsini (1990) slightly modified Joe’s proposal, also taking into account the given margins. In section 2 we briefly recall their approaches and we analyze how appropriate they can be in assessing dependence.

Hence, the purpose of this paper is to introduce a new dependence ordering between bivariate distributions with fixed row and column sums and a special kind of frequency transfer – among the cells of a double table – which generates a decrease in dependence.

Section 3 presents the ‘directional dependence ordering’: this definition is natural enough and in a sense more intuitive than earlier ones considered in the literature: it takes into account the signs and the absolute values of the

$r \times c$ contingencies. The new (partial) ordering is characterized in terms of cyclic frequency transfers. The notion of cycle frequency transfer generalizes, over a closed set of pair of cells, the transformation suggested by Diaconis and Sturmfels (1998) and carried out over 4 cells, in order to generate a Markov chain in the class of all bivariate distributions with given margins.

In section 4, necessary and sufficient conditions are given to characterize maximal (or minimal) elements in the chosen set, endowed with the directional dependence ordering.

Finally, in section 5, a new measure of dependence is proposed and some of its properties are sketched.

2. Dependence orderings for categorical variables, in the class of bivariate distributions with given margins

In this section, after introducing some notations, we recall Joe's and Scarsini's proposals and we discuss their relevance in comparing bivariate distributions with respect to (w.r.t.) their degree of dependence.

Let the n statistical units of a given population be classified according to the qualitative variables A and B , both categorical, with a finite number of unordered categories, denoted by $a_1, \dots, a_j, \dots, a_c$ and $b_1, \dots, b_i, \dots, b_r$ respectively.

With regard to the typology of statistical data (Kendall and Stuart (1979), Leti (1983), Zanella (1988)), the following three cases are usually examined:

- a) both A and B margins are fixed;
- b) only the marginal frequencies of one variable are fixed;
- c) there are no constraints on the marginal frequencies (in other words, only the total population size, say n , is fixed).

In this paper we refer to case a).

Let \mathcal{F} be the class of all $r \times c$ contingency tables with non negative integer entries n_{ij} , whose row sums $n_{i\cdot}$ and column totals $n_{\cdot j}$ are considered fixed.

In order to compare the relative degree of dependence represented by two bivariate distributions \mathbf{T} and \mathbf{T}' of \mathcal{F} , Joe (1985) suggested that the column vectors $\text{vec}(\mathbf{T})$ and $\text{vec}(\mathbf{T}')$, respectively obtained by piling up the columns of \mathbf{T} and \mathbf{T}' , are compared by using the classical majorization ordering of Marshall and Olkin (1979). Recalling that if $x=(x_1, \dots, x_p)$ and $y=(y_1, \dots, y_p)$ are two vectors of dimension p ($p=r \times c$, in the present case), then x majorizes y

and we write $x \succ_M y$, if $\sum_{i=1}^k x_{[i]} \geq \sum_{i=1}^k y_{[i]}$, for $k=1, \dots, p-1$ where $x_{[1]} \geq \dots \geq x_{[p]}$
 and $y_{[1]} \geq \dots \geq y_{[p]}$, and $\sum_{i=1}^p x_{[i]} = \sum_{i=1}^p y_{[i]}$.

In other words, after rearranging in two $r \times c$ vectors the internal frequencies of each table, Joe's proposal is to consider their ordering, based on the Lorenz curve.

In order to understand the rationale of Joe's proposal, let us observe that, in the case of independence, each row in a bivariate table is a distribution similar¹ to that of the marginal row (and the same holds for columns); moreover, if the margins are uniform, i.e. for $n_{i\cdot} = n/r$ and $n_{\cdot j} = n/c$ $\forall i \in \{1, \dots, r\}$ and $\forall j \in \{1, \dots, c\}$, all entries n_{ij} coincide. Hence, roughly speaking, a decrease in the concentration (variability) of the entries n_{ij} can be related to a decrease in dependence. But Joe did not consider the margins, so that the ordering \preccurlyeq_j presents an odd behavior w.r.t. the independence table, as remarked in section 4.

A related concept investigated by Scarsini (1990), developing an idea suggested by Cifarelli and Regazzini (1986), consists of comparing, not \mathbf{T} with \mathbf{T}' , but rather $\mathbf{T}/\hat{\mathbf{T}}$ with $\mathbf{T}'/\hat{\mathbf{T}}$ in the majorization ordering. Here, matrix division is understood to be componentwise, and $\hat{\mathbf{T}} = \{\hat{n}_{ij}\}$ is the independence table.

Scarsini's ordering, denoted by \preccurlyeq_s , takes into account the given margins which were not explicitly involved in Joe's proposal.

Let us note that the two orderings \preccurlyeq_j and \preccurlyeq_s coincide when the margins of \mathbf{T} and \mathbf{T}' are uniform.

The orderings \preccurlyeq_j and \preccurlyeq_s possess the property of invariance under the transposition and permutation of rows and/or columns. In the above mentioned work, Joe obtained necessary and sufficient conditions to characterize maximal and minimal matrices either for margins chosen in \mathbb{N} or in \mathbb{R}^+ . Scarsini showed that there is a unique minimum in \preccurlyeq_s , i.e. the independence table. For the general case $r \times c$ there is not a maximum w.r.t.

¹ We say that two distributions are *similar* if they have the same relative frequencies (Gini 1914-15). The analysis of dependence based on the similarity of the conditioned univariate distributions represented by the rows (analogously by the columns) is due to Salvemini (1939), then it was refined by Castellano (1960) and further developed also by Leti (1983) and Zanella (1988). References to the earlier works are available in Naddeo (1987).

\leq_s , except in the case of equal margins, giving the diagonal table as the unique maximum.

The orderings of Joe and Scarsini appear to be very useful in comparing the degree of concentration in two tables $\mathbf{T}=\{n_{ij}\}$ and $\mathbf{T}'=\{n'_{ij}\}$ in a given class \mathcal{F} , i.e., in measuring how the n units are more or less concentrated in the $r \times c$ cells of a table. However, they do not compare each n_{ij} with the corresponding joint frequency n'_{ij} in \mathbf{T}' , but rather the components of one specific rearrangement of $\text{vec}(n_{ij})$ with another (possibly different!) arrangement of $\text{vec}(n'_{ij})$. This is not meaningful in a dependence context, where the pair (i,j) corresponds to a specific choice of the i -th category of B and the j -th category of A so that n_{ij} and $\rho^*_{ij}=n_{ij}/\hat{n}_{ij}$ are conceptually related to the couple of indexes.

3. Directional dependence ordering and cycle frequency transfers

In this section, the concept of directional dependence ordering is introduced and analysed. Then, the notions of a cycle in a table and of a cyclic frequency transfer that decreases the directional dependence are defined. Some examples are given. This section ends with Theorem 3.5 that shows the strong relation between these definitions.

Definition 3.1: Directional dependence (partial) ordering

Let $\mathbf{T}=\{n_{ij}\}$ and $\mathbf{T}'=\{n'_{ij}\}$ be two tables in \mathcal{F} and let $c_{ij}=n_{ij}-\hat{n}_{ij}$ (respectively c'_{ij}) be their contingencies. Between the variables A and B there is more directional dependence in \mathbf{T} than in \mathbf{T}' , if and only if, $\forall (i,j) \ i:1,\dots,r; \ j:1,\dots,c$:

- i) $c'_{ij} \times c_{ij} \geq 0$;
- ii) $|c'_{ij}| \leq |c_{ij}|$.

The following notation:

$$\mathbf{T}' \leq_{DD} \mathbf{T}$$

indicates that \mathbf{T}' precedes \mathbf{T} in the directional dependence ordering.

Conditions i) and ii) appear to be appropriate in comparing and assessing the nature (the configuration of the contingency signs cell by cell) and the strength (the absolute values) of the mutual dependence of A and B, as the following tables \mathbf{T}_1 and \mathbf{T}_2 show:

B\A	a ₁	a ₂	a ₃	
b ₁	11	7	2	20
b ₂	7	7	2	16
b ₃	4	5	2	11
b ₄	3	1	3	7
	25	20	9	54

B\A	a ₁	a ₂	a ₃	
b ₁	12	6	2	20
b ₂	7	8	1	16
b ₃	3	5	3	11
b ₄	3	1	3	7
	25	20	9	54

B\A	a ₁	a ₂	a ₃	
b ₁	9.3	7.4	3.3	20
b ₂	7.4	5.9	2.7	16
b ₃	5.1	4.1	1.8	11
b ₄	3.2	2.6	1.2	7
	25	20	9	54

$\mathbf{T}_1\{n'_{ij}\} \quad \ll_{DD} \quad \mathbf{T}_2\{n_{ij}\} \quad \hat{\mathbf{T}}\{\hat{n}_{ij}\}$

In \mathbf{T}_1 and in \mathbf{T}_2 the signs of the contingencies agree, cell by cell, as condition i) of def. 3.1 requires. Where the sign is positive as, for example, in cell (1,1), the inequality $n_{11} \geq n'_{11} > \hat{n}_{11}$ holds, i.e. $c_{11} = 2.7 \geq c'_{11} = 1.7$. There is a sort of ‘attraction’² between the first category of variable A and the first category of B, in that they appear together in a number of units higher than the expected joint frequency in case of independence, and this ‘attraction’ is stronger in \mathbf{T}_2 than in \mathbf{T}_1 . Analogously, with a negative sign of the contingency, for example in cell (2,3), the inequality $n_{23} \leq n'_{23} < \hat{n}_{23}$ holds, i.e. $|c_{23}| = 1.7 \geq |c'_{23}| = 0.7$. The second category of B and the third category of A appear together in few units, less than those that we expect if A and B were independent, and this ‘repulsion’ is stronger in \mathbf{T}_2 than in \mathbf{T}_1 .

The qualifier ‘directional’ in def. 3.1 corresponds to a specific configuration of the contingency signs in the table:

B\A	a ₁	a ₂	a ₃	
b ₁	+	-	-	20
b ₂	-	+	-	16
b ₃	-	+	+	11
b ₄	-	-	+	7
	25	20	9	54

By the very nature of variables A and B, the labelling of their respective categories is immaterial, so the following property of the directional dependence ordering is meaningful and required:

² Benini (1901) coined the terms ‘attraction’ and ‘repulsion’ to indicate this kind of relation in assessing dependence.

THEOREM 3.2:

The directional dependence ordering is invariant under the transposition and permutation of rows and/or columns. That is, if $\mathbf{S} \ll_{DD} \mathbf{T}$ then:

$$\mathbf{S}^T \ll_{DD} \mathbf{T}^T \quad \text{and} \quad \mathbf{P}_1 \mathbf{S} \mathbf{P}_2 \ll_{DD} \mathbf{P}_1 \mathbf{T} \mathbf{P}_2$$

for all $r \times r$ permutation matrices \mathbf{P}_1 and all $c \times c$ permutation matrices \mathbf{P}_2 .

PROOF:

The thesis comes directly from def. 3.1. ♦

Before introducing the frequency transfers, we need the definition of a cycle in \mathbf{T} :

Definition 3.3: Cycle

An ordered set of $2 \times k$ cells: $\{n_{i_1 j_1}, n_{i_1 j_2}, n_{i_2 j_2}, n_{i_2 j_3}, \dots, n_{i_k j_k}, n_{i_k j_{k+1}}\}$

in $\mathbf{T} = \{n_{ij}\}$ is a cycle if and only if, with $k : 2 \leq k \leq \lfloor (r \times c) / 2 \rfloor$:

- i) $i_1, i_2, \dots, i_k \in \{1, 2, \dots, r\}$;
- ii) $j_1, j_2, \dots, j_k \in \{1, 2, \dots, c\}$;
- iii) $j_{k+1} = j_1$;
- iv) $n_{i_m j_m} \geq \hat{n}_{i_m j_m} + 1$; $n_{i_m j_{m+1}} \leq \hat{n}_{i_m j_{m+1}} - 1 \quad \forall m = 1, 2, \dots, k$.

In other words a cycle in \mathbf{T} is a closed path among the cells of \mathbf{T} that picks up an ordered set of even entries:

- each having $|c_{ij}| \geq 1$;
- paired in the same row or column, with opposite signs of contingencies.

We call length of the cycle the number $2 \times k$ of the involved cells.

Definition 3.4: Cyclic frequency transfers which decrease directional dependence

Let $\mathbf{T}\{n_{ij}\} \in \mathcal{F}(n_{i\cdot}; n_{\cdot j})$, and let $\{n_{i_1 j_1}, n_{i_1 j_2}, n_{i_2 j_2}, n_{i_2 j_3}, \dots, n_{i_k j_k}, n_{i_k j_{k+1}}\}$ be a cycle in \mathbf{T} . Define $\mathbf{T}'\{n'_{ij}\} \in \mathcal{F}(n_{i\cdot}; n_{\cdot j})$ with the positions:

- i) $n'_{i_m j_m} = n_{i_m j_m} - 1$ for $m = 1, 2, \dots, k$;
- ii) $n'_{i_m j_{m+1}} = n_{i_m j_{m+1}} + 1$ for $m = 1, 2, \dots, k$;
- iii) $n'_{ij} = n_{ij}$ otherwise;

The transformation $\mathbf{T} \rightarrow \mathbf{T}'$ is called a **decreasing cyclic frequency transfer**, for $\mathbf{T}' \ll_{DD} \mathbf{T}$.

The example given above in table \mathbf{T}_2 shows a cycle and it concerns 2×3 cells in \mathbf{T}_2 : namely $\{n_{11}, n_{12}, n_{22}, n_{23}, n_{33}, n_{31}\}$. One can devise a decreasing cyclic frequency transfer which transforms \mathbf{T}_2 into \mathbf{T}_1 , without changing the margins. Moreover, the cells losing a unit (i.e. n_{11}, n_{22}, n_{33}) are such that

$n_{ij} \geq \hat{n}_{ij} + 1$, whereas the cells gaining a unit (i.e. n_{12} , n_{23} , n_{31}) are such that $n_{ij} \leq \hat{n}_{ij} - 1$. As a result of transfers, the modified frequencies n'_{ij} lay in an intermediate position between the initial n_{ij} and those of independency \hat{n}_{ij} . The signs of the contingencies have not changed, while their absolute value has decreased. Hence, these transfers draw the table closer to the situation of independence without modifying the direction of dependence or the margins.

Obviously, one may consider \mathbf{T} as the transformation of \mathbf{T}' , by an opposite, increasing cyclic frequency transfer.

The intimate connection between the cyclic frequency transfer and the directional dependence ordering is openly expressed by:

THEOREM 3.5³:

Let $\mathbf{S}, \mathbf{T} \in \mathcal{F}$, such that $\mathbf{T} \preccurlyeq_{DD} \mathbf{S}$ and $\mathbf{S} \neq \mathbf{T}$. Then, by a finite series of cyclic frequency transfers, \mathbf{S} can be transformed into \mathbf{T} .

PROOF:

Let us first prove that in \mathbf{S} there is a decreasing cycle. Let d_1 be the positive integer defined by:

$$d_1 = \sum_i \sum_j |c_{ij}^S - c_{ij}^T|,$$

where c_{ij}^X denotes the contingencies of table \mathbf{X} (for $\mathbf{X} = \mathbf{S}, \mathbf{T}$). Let (i_1, j_1) be a pair of row and column indexes such that $n_{i_1 j_1}^S > n_{i_1 j_1}^T \geq \hat{n}_{i_1 j_1}$ i.e.: $c_{i_1 j_1}^S > c_{i_1 j_1}^T \geq 0$. Given that $\sum_{j=1}^c c_{i_1 j}^T = 0$ and given the ranking between \mathbf{T} and \mathbf{S} , we now show that there is a column index $j_2 \neq j_1$ such that: $c_{i_1 j_2}^S < c_{i_1 j_2}^T \leq 0$.

If such j_2 does not exist, then:

$$0 = \sum_{j=1}^c c_{i_1 j}^S = \sum_{\substack{j=1 \\ j \neq j_1}}^c c_{i_1 j}^S + c_{i_1 j_1}^S > \sum_{\substack{j=1 \\ j \neq j_1}}^c c_{i_1 j}^S + c_{i_1 j_1}^T \geq \sum_{\substack{j=1 \\ j \neq j_1}}^c c_{i_1 j}^T + c_{i_1 j_1}^T = 0$$

thus giving a contradiction.

This argument can be repeated in column j_2 , looking for a row, say i_2 , in which $c_{i_2 j_2}^S > c_{i_2 j_2}^T \geq 0$. After a finite number of these correspondences, say p , having $r \times c$ cells, this cell collection will eventually reach a cell (i_k, j_k)

³ Theorem 3.5 states the analogue of Muirhead (1903), Hardy, Littlewood and Polya (1929, 1934, 1952) result, for \mathbf{x}, \mathbf{y} vectors in \mathbb{R}^n : "If $\mathbf{x} \preccurlyeq_M \mathbf{y}$ then \mathbf{x} can be derived from \mathbf{y} by a finite series of \mathbf{T} -transforms: $\mathbf{x} \rightarrow \mathbf{xT}$ ".

already visited (with $k: 1 \leq k \leq p$), thus locating a decreasing cycle in \mathbf{S} : $\{n_{i_k j_k}, n_{i_k j_{k+1}}, \dots, n_{i_p j_p}, n_{i_p j_{p+1}}\}$. Observe that: $2 \leq p \leq \lfloor r \times c / 2 \rfloor$.

Let \mathbf{S}_1 be the matrix obtained from \mathbf{S} by the decreasing frequency transfer carried out on the above cycle, and let $d_2 = \sum_i \sum_j |c_{ij}^{\mathbf{S}_1} - c_{ij}^{\mathbf{S}}|$. A cyclic frequency transfer involves at least four cells, so $d_2 \leq d_1 - 4$.

Now, if $\mathbf{S}_1 \preceq_{\text{DD}} \mathbf{S}$ and $\mathbf{S}_1 \neq \mathbf{S}$, a cycle can be identified in \mathbf{S}_1 . A finite sequence of m such transforms (with $m \leq d_1/4$), being the sequence d_1, d_2, \dots , a strictly decreasing series of non negative integers, yields $d_m = 0$. Therefore, the sequence $\mathbf{S}, \mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_m$ gives $\mathbf{S}_m = \mathbf{T}$. ♦

4. The role and characterization of extremes

In this section, we obtain results for minimal and maximal matrices in a given class \mathcal{F} , w.r.t. the directional dependence ordering (DDO). Let us recall the well-known definition for an extremal⁴ element in a partially ordered set:

Definition 4.1: *Maximal (minimal) element in a poset*

Let \mathbf{S} be a matrix in $\mathcal{F}(n_i; n_j)$. \mathbf{S} is a **maximal (minimal)** matrix in \mathcal{F} w.r.t. DDO if and only if there is no $\mathbf{T} \in \mathcal{F}$, $\mathbf{T} \neq \mathbf{S}$ such that $\mathbf{S} \preceq_{\text{DD}} \mathbf{T}$ ($\mathbf{S} \succeq_{\text{DD}} \mathbf{T}$).

In other words, \mathbf{S} is an extremal table w.r.t. DDO if and only if there is no other table \mathbf{T} in \mathcal{F} , having the same contingency signs as \mathbf{S} and with all contingencies ‘heavier or equal’ (or ‘lighter or equal’) than \mathbf{S} .

Now, we can rephrase the characterizing proposition:

THEOREM 4.2:

Necessary and sufficient condition for a maximal (minimal) matrix in \mathcal{F} w.r.t. DDO.

A matrix $\mathbf{S} = \{n_{ij}\}$ is maximal (minimal) in \mathcal{F} w.r.t. DDO if and only if \mathbf{S} has no increasing (decreasing) cycles.

The proof is straightforward, hence it is omitted.

⁴ See, for example, Rosen (2000), page 721.

Theorem 4.2, directly connected to the concept of cyclic frequency transfers, is unable to give a practical method for distinguishing extreme tables. Before we can achieve an operative characterization for maximal tables, we need two definitions:

Definition 4.3: *Reduced submatrix*

Let S be an $r \times c$ matrix, and S^- the submatrix obtained by deleting a line from S . S^- is called the reduced submatrix of S .

A *line* of a matrix designates either a row or a column of the matrix. So, if the deleted line in S is for example row i , then S^- is an $(r-1) \times c$ matrix (if the deleted line is a column, changes are made accordingly).

Definition 4.4: *DDmax-saturated line*

A row or a column in a matrix is a DDmax-saturated line if and only if one of the following two conditions holds:

- i) each cell (i,j) in the line with a negative contingency sign has the minimum frequency⁵ $\tilde{n}_{ij} = \max(0, n_{i.} + n_{.j} - n)$;

or:

- ii) all contingencies share the same sign⁶.

Each of the two conditions in def 4.4 assures that no increasing cycle contains a pair of cells in that line.

THEOREM 4.5:

Necessary and sufficient conditions for a maximal matrix in \mathcal{F} w.r.t. DDO.

An $r \times c$ matrix S is maximal in \mathcal{F} w.r.t. DDO if and only if by deleting successively a DDmax-saturated line in S and in its reduced submatrices one arrives at a uni-dimensional array.

PROOF:

Suppose first that a finite series of dropped DDmax-saturated lines in S leads to reduce S to a vector. In S only case i) occurs: we identify a line, for instance row k , in which for each negative contingency sign in cell (k,j) there is the respective minimum frequency \tilde{n}_{kj} . Hence there is no increasing cycle in S which involves a pair of cells in row k .

⁵ As it is well known, because of the given margins, the joint frequencies n_{ij} must obey the inequalities: $\max(0, n_{i.} + n_{.j} - n) \leq n_{ij} \leq \min(n_{i.}, n_{.j})$ for $i=1, \dots, r$ and $j=1, \dots, c$.

⁶ Condition ii) can not hold in the initial matrix, but it can appear as soon as we deal with one of its reduced submatrices.

Let us turn our attention to the reduced submatrix S^- , obtained from S by deleting row k : we repeat the same process in matrix S^- . At this point, also ii) of def. 5.2 can occur. In any event, having identified a DDmax-saturated line, there is no increasing cycle in S^- involving such a line; consequently the line is dropped from S^- , obtaining $(S^-)^-$. This sequence stops upon reaching a table $1 \times t$ (or $s \times 1$), in which –obviously– no increasing cycle is obtainable, hence S is a maximal table.

Now, let us suppose S maximal and in all lines of all reduced submatrices of S neither i) nor ii) of def 3.4 holds. Then, in any line there are at least two cells with opposite sign of the contingencies and a cell with negative contingency sign and non minimum frequency. Let us choose a line in S , for example row k , thus in row k there are two frequencies n_{kp} and n_{kq} , such that $\min(n_{k\cdot}, n_{\cdot p}) > n_{kp} > \hat{n}_{kp}$ and $\tilde{n}_{kq} < n_{kq} < \hat{n}_{kq}$. In turning our attention to column p , we identify one row t in which: $\tilde{n}_{tp} < n_{tp} < \hat{n}_{tp}$. A finite sequence of these correspondences builds an increasing cycle that eventually closes on an already visited cell, leading to a contradiction. ♦

Maximum dependence of A on B is not so clearly defined in the literature, but we can adopt the following definitions⁷:

Definition 4.6: *Complete dependence of A on B ('massima dipendenza unilaterale di A da B ' in the Italian literature)*

In table $T = \{n_{ij}\}$ there is the complete dependence of variable A on variable B if $\forall i: 1, \dots, r$: a unique j exists such that $n_{ij} \neq 0$.

In other words, for each category of B there is a unique category of A which can be associated with it.

Definition 4.7: *Absolute dependence between A and B ('massima dipendenza bilaterale tra A e B ' in the Italian literature)*

In table $T = \{n_{ij}\}$ there is the absolute dependence between variables A and B if Def 4.6 holds for A on B and for B on A .

COROLLARY 4.8:

All matrices which show complete or absolute dependence are maximal w.r.t. DDO.

⁷ For the maximum dependence (association) we will rephrase Kendall and Stuart (1979, page 560), referring to tetrachoric tables and briefly recalled here: "Considering a population classified according to the presence or absence of two attributes A and B , we say that *association* is *complete* if all A 's are B 's. *Absolute association* arises when all A 's are B 's and all B 's are A 's."

Examples 4.9:

B\A	a ₁	a ₂	a ₃	
b ₁	25	0	0	25
b ₂	0	15	0	15
b ₃	0	0	9	9
b ₄	0	5	0	5
	25	20	9	54

T₃

B\A	a ₁	a ₂	a ₃	
b ₁	25	0	0	25
b ₂	0	20	0	20
b ₃	0	0	9	9
	25	20	9	54

T₄

T₃ shows the complete dependence of A from B, while in T₄ we observe the absolute dependence. Each row in the two tables is a DDmax-saturated line, i.e. no increasing cyclic frequency transfer can be devised in T₃ or in T₄, and this assures their maximality w.r.t. DDO.

Furthermore, if variables A and B were nominal and ordered, it would be meaningful to identify the cograduation and contragraduation matrices⁸.

COROLLARY 4.11:

The cograduation and contragraduation matrices are maximal w.r.t. DDO.

Examples 4.12:

B\A	a ₁	a ₂	a ₃	
b ₁	0	11	9	20
b ₂	7	9	0	16
b ₃	11	0	0	11
b ₄	7	0	0	7
	25	20	9	54

T₅

B\A	a ₁	a ₂	a ₃	
b ₁	20	0	0	20
b ₂	5	11	0	16
b ₃	0	9	2	11
b ₄	0	0	7	7
	25	20	9	54

T₆

⁸ The names ‘cograduation table’ and ‘contragraduation table’ were coined by Salvemini (1939) to designate two tables singled out by Fréchet: the table which in each cell shows the maximum cumulative frequencies among all possible bivariate distributions with the same margins (cograduation), and that which similarly shows the minimum cumulative frequencies (contragraduation). References to Salvemini’s work are available in Naddeo (1987).

One can easily verify that a process of eliminations of DDmax-saturated lines in T_5 (cograduation) and T_6 (contragraduation) leads to reduce each table to a vector.

In recalling Theorem 3.2, we can generalize Corollary 4.11:

COROLLARY 4.13:

All tables which are obtainable by permuting rows and/or columns from a cograduation matrix or from a contragraduation matrix are maximal w.r.t. DDO.

Joe, in the above cited work, obtains necessary conditions for maximality in \preceq_J (see Theorem 4) and the equivalent Lemma 3.1 in Jurkat and Ryser (1967) gives the characterization of extremal matrices in \mathcal{F} , regarded as a convex set. They, in fact, propose a constructive algorithm for writing all matrices of Corollary 4.13. Hence all maximal tables in Joe's ordering are maximal w.r.t. DDO.

However, there are maximal tables in DDO that are not included in Corollary 4.13. As a counterexample, let us observe table T_7 :

T_7

B\A	a ₁	a ₂	a ₃	
b ₁	20	0	0	20
b ₂	1	15	0	16
b ₃	4	5	2	11
b ₄	0	0	7	7
	25	20	9	54

Looking at its submatrix:

1	15
4	5

the submatrix contains four lines with all positive elements, but no cycle can be built in it, for in each column the contingency signs agree. Hence T_7 is maximal in DDO.

Remarks:

Given one dependence direction, there can be one or more maximal tables w.r.t. DDO: T_5 and T_7 are maximal, and $T_5 \not\preceq T_7$ nor $T_5 \succeq T_7$.

In focusing now on the characterization of minimal tables, depending on the specific values of these marginal frequencies and the grand total $n = n_{1.} + \dots + n_{r.} = n_{.1} + \dots + n_{.c}$ the independence table $\hat{T} = \{\hat{n}_{ij}\}$ may or may not belong to \mathcal{F} .

If $\hat{\mathbf{T}} = \{\hat{n}_{ij}\} \in \mathcal{F}$ then $\hat{\mathbf{T}}$ is the lower bound for DDO. This means that $\hat{\mathbf{T}}$ is the unique minimal matrix $\hat{\mathbf{T}}$, i.e.: $\mathbf{T} \in \mathcal{F}$, \mathbf{T} is comparable with $\hat{\mathbf{T}}$ and $\hat{\mathbf{T}} \leq_{\text{DD}} \mathbf{T}$.

The following counterexample shows that this natural comparability with the independence table doesn't hold in Joe's ordering, leading to an anomalous situation:

B\A	a ₁	a ₂	a ₃	
b ₁	4	4	4	12
b ₂	4	5	15	24
b ₃	4	15	17	36
	12	24	36	72

\mathbf{T}_8

B\A	a ₁	a ₂	a ₃	
b ₁	2	4	6	12
b ₂	4	8	12	24
b ₃	6	12	18	36
	12	24	36	72

$\hat{\mathbf{T}}$

$\hat{\mathbf{T}} \in \mathcal{F}$ but \mathbf{T}_8 is not comparable with $\hat{\mathbf{T}}$ in Joe's ordering!

Before concluding this section concerning extremal tables in DDO, we remark that, as for maximal tables, one can easily define the DDmin-saturated line and state a characterizing proposition for recognizing minimal tables.

5. A measure of dependence derived from the directional dependence ordering

Goodman and Kruskal (1979) offer a fundamental survey on measures of dependence between two statistical variables; as well as Haberman (1982) and Kendall and Stuart (1979), they believe that several measures are possible, reflecting different goals: measures of prediction, symmetric measures, etc.

In the present work we assumed the symmetric role of variables A and B and we showed that the DD-ordering permits discrimination among the large subset of tables which do not represent extreme dependence situations, describing what kind of dependence between A and B they represent (a specific dependence direction) and possibly comparing pairs of tables possessing this directional dependence to a greater or lesser extent.

Now we expect that dependence indexes - used to measure the strength of this relation - are coherent with such partial ordering.

For the central role of contingencies in DDO, any classical dependence index defined by a non-negative and strictly increasing function of the absolute value of contingencies (or of relative contingencies ρ_{ij}) is coherent with the DDO, as in Poliscchio (2002) :

- $M_1(|\rho|)$, the weighted arithmetic mean of $|\rho_{ij}| = |c_{ij}| / \hat{n}_{ij}$ with weights \hat{n}_{ij} ;
- $M_2(|\rho|)$, the weighted quadratic mean of $|\rho_{ij}| = |c_{ij}| / \hat{n}_{ij}$ with weights \hat{n}_{ij} ;
- the classic Pizzetti-Pearson statistic X^2 (Pearson, 1904);
- the global independence index of second order (Castellano, 1962);
- the total dependence index of second order (Gini, 1955).

A simple but meaningful measure of dependence coherent with DD ordering may be derived by observing that the class \mathcal{F} , chosen as the reference set, is a finite set of tables⁹.

Let $\mathcal{F}_{\prec DD}$ be the sub-class of \mathcal{F} composed by all tables that are comparable with \mathbf{T} by \prec_{DD} .

Definition 5.1: The index $I_{\prec DD}$

Given a bivariate distribution \mathbf{T} , the relative position that \mathbf{T} assumes among $\mathcal{F}_{\prec DD}$, according to the sorting induced by \prec_{DD} , is a measure of dependence $I_{\prec DD}(\mathbf{T})$, for \mathbf{T} :

$$I_{\prec DD}(\mathbf{T}) = \frac{\text{card}\{\mathbf{S} | \mathbf{S} \in \mathcal{F}_{\prec DD}; \mathbf{S} \preceq_{DD} \mathbf{T}\}}{\text{card}\{\mathbf{S} | \mathbf{S} \in \mathcal{F}_{\prec DD}\}}$$

Observe that $I_{\prec DD}$ is a relative frequency, so it is autonormalized, i.e. it assumes values in (0,1]. Its interpretation is straightforward: for example, a value of $I_{\prec DD}(\mathbf{T})=.45$ means that, among all bivariate distributions in \mathcal{F} representing the same nature of the dependence between A and B, 45% of such tables possess this directional dependence to a lesser extent w.r.t. \mathbf{T} .

Moreover $I_{\prec DD}(\mathbf{T})$ assumes its minimum value if and only if \mathbf{T} is minimal; conversely, $I_{\prec DD}(\mathbf{T}) = 1$ if and only if \mathbf{T} maximal w.r.t. DDO.

$I_{\prec DD}$ inherits all properties the DDO has, in particular it is invariant w.r.t. permutation of rows and/or columns and to transposition.

⁹ Given the class of all $r \times c$ contingency tables with integer entries n_{ij} whose row sums $n_{i\cdot}$ and column totals $n_{\cdot j}$ are considered fixed, it is possible to locate and single out all its elements (see, among many others, Leti (1970) and Greselin (2004)).

$I_{\llcorner DD}$ can be calculated by enumerating all tables in \mathcal{F} , as it was done by a C-algorithm on a Pentium PC for tables with $r \times c < 20$ and $n < 1000$: obtaining its value in seconds or in minutes. The required computing time grows exponentially as r or c or n increase, due to the cardinality \mathcal{F} . In any event, observed asymptotic properties of this index can give light to this computational limit; as well as the possibility to identify directly only all tables comparable with \mathbf{T} (without enumerating all elements in \mathcal{F}) by results coming from graph-theory. These and other issues will be argument of future work.

6. Conclusions

This paper suggests the introduction of partial dependence orderings in order to allow a deeper comprehension of some aspects of association. Partial ordering \llcorner_{DD} is also intuitive from a geometric point of view. Indeed, in the space of $r \times c$ frequencies n_{ij} , there is a point whose coordinates are the independence frequencies \hat{n}_{ij} . This point \hat{P} is the origin of contingencies c_{ij} for all tables with the same given margins. Each table is represented by a point $P = (c_{11}, \dots, c_{rs})$. In agreement with the c_{ij} signs, point P is situated in a specific orthant in the space of c_{ij} coordinates. The ranking of tables which are situated in different orthants is impossible according to DDO. Actually, the ranking induced by \llcorner_{DD} is not always feasible even among tables which lay in the same orthant: for \mathbf{T} and \mathbf{T}' belonging to the same orthant, we have that

$\mathbf{T} \llcorner_{DD} \mathbf{T}'$ if and only if, for each pair (i,j) the relation $|c_{ij}| \leq |c'_{ij}|$ holds and the contingency signs agree. Consequently, starting from the orthant in which the given table \mathbf{T} falls, with a series of transfers that increase the directional association, one would arrive at a table in the orthant with maximum association: the uniquely identified distributive independence situation – the point \hat{P} – is counterbalanced by one or more situations of maximum dependence in each orthant.

The characterization of the extremes in the class \mathcal{F} w.r.t. DDO leads us to remark that the tables of complete or absolute dependence are recognized as maximal tables. At the same time, the lack of association corresponds to the unique minimal table in DDO (provided that they are compatible with the prescribed margins). The cograduation and contragraduation matrices are maximal tables in \mathcal{F} w.r.t. DDO when the variables are ordered qualitative. Furthermore, DDO allows the identification of more maximal situations.

The idea of evaluating the degree of association in a given table T , considering its relative position in the ordered class \mathcal{F} , appears to be worthy of future developments.

Acknowledgments

The authors thank the Italian MIUR for partial financial support (COFIN prot. 2002133957_004), within the project: 'Descriptive and Inferential Aspects for Categorical Data Analysis', led by professor A. Forcina, University of Perugia.

References

- Benini, R. (1901). *Principi di demografia*. Firenze, G.Barbera.
- Castellano, V. (1960). Sull'insieme delle distribuzioni doppie e triple risultanti dall'associazione d una ad una delle unità di due o tre distribuzioni semplici aventi lo stesso numero di unità. *Metron*, vol. XX, 1-4.
- Castellano, V. (1962). *Istituzioni di Statistica*. Roma, Ilardi.
- Cifarelli, D.M., Regazzini, E. (1986). Concentration function and its role in descriptive statistics, in *Proceedings of the 33rd scientific meeting of the Italian statistical society*, vol 2, 347-352, in Italian.
- Diaconis, D. and Sturmfels, B. (1998). Algebraic algorithm for sampling from conditional distributions. *The Annals of Statistics*, vol. **26**, No. 1, 363-397.
- Esary, J., Proschan, F., and Walkup, D., (1967). Association of random variables with applications. *Ann. Math. Statist.*, **38**, 1466-1474.
- Forcina, A. and Giovagnoli, A. (1987). Mathematical properties of cross-classification matrices. *Boll. U.M.I.*, VII,1-B, **2**, 523-544.
- Gini, C. (1914-15). Di una misura della dissomiglianza tra due gruppi di quantità e delle sue applicazioni allo studio delle relazioni statistiche, *Atti del Regio Istituto Veneto di Scienze Lettere ed Arti*, t. LXXIV.
- Gini, C. (1954-55). *Corso di statistica*. Roma, Veschi.
- Goodman, L.A. and Kruskal, W. (1979). *Measures of association for cross-classifications*. Springer-Verlag Inc, Berlin, New York.
- Greselin, F. (2003). Counting and enumerating rectangular arrays with given margins. *Statistica&Applicazioni*, vol. 1, n°2.
- Habermann, S.J. (1982). Measures of association, in *Encyclopedia of statistical sciences*. John Wiley & sons, vol.1, chapter 1, 130-137.

- Hardy, G.H. Littlewood J.E. and Polya G. (1929). Some simple inequalities satisfied by convex functions. *Messenger Math.* **58**, 145-152.
- Hardy, G.H. Littlewood J.E. and Polya G. (1934, 1952). *Inequalities*. 1st e 2nd ed. Cambridge Univ. Press, London and New York.
- Joe, H. (1985). An ordering of Dependence for contingency tables. *Linear algebra and its applications*, **70**, 89-103.
- Jurkat, W.B. and Ryser, H.J. (1967). Term Ranks and Permanents of Nonnegative Matrices. *Journal of algebra*, **5**, 342-357.
- Karlin, S. (1968). *Total Positivity, Vol. I*. Stanford University Press, Stanford.
- Kendall, M. and Stuart, A. (1979). *The Advanced Theory of Statistics*, vol 2: Inference and Relationship. 4th edn. New York: Macmillan.
- Kimeldorf, G., and Sampson, A. R. (1987). Positive dependence orderings. *Ann. Inst. Statist. Math.*, **39**, Part A, 113-128.
- Lehmann, E. (1966). Some concepts of dependence. *Ann. Math. Statist.*, **37**, 1137-1153.
- Leti, G. (1970). La distribuzione delle tabelle della classe di Fréchet. *Metron*, vol XXVIII, n.1-4.
- Leti, G. (1983). *Statistica descrittiva*. Il Mulino, Bologna.
- Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.
- Muirhead, R.F. (1903). Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters. *Proc. Edimburgh Math. Soc.* **21**, 144-157.
- Naddeo A. Eds, (1987). *Italian Contributions to the methodology of statistics*. Società Italiana di Statistica, Cleup Padova
- Pearson, K. (1904). Mathematical contributions to the theory of Evolution. XIII. On the theory of contingency and its relation to association and normal correlation. *Drapers' Company Res. Mem., Biometric Ser. I*.
- Polisicchio, M. (2002). Partial and Total Orderings on Contingency Tables in the Fréchet class. in *Proceedings of the 41st scientific meeting of the Italian statistical society*, vol 2, 105-112, CLEUP, Padova.
- Rinott, Y., and Pollak, M. (1980). A stochastic ordering induced by a concept of positive dependence and monotonicity of asymptotic test sizes. *Ann. Statist.*, 190-198.
- Rosen, K.H. (2000). *Handbook of discrete and combinatorial mathematics*. CRC Press, Boca Raton.
- Salvemini (1939). Sugli indici di emofilia. *Atti della I Riunione della Società Italiana di Statistica*, Pisa.

Scarsini, M. (1990). An ordering of Dependence. in: *Topics in Statistical Dependence*, Block, H.W., Sampson, A.R. Savits, T.H. (eds.), Hayward Institute of Mathematical Statistics.

Shaked, M. (1977). A family of concepts of dependence for bivariate distributions. *J. Amer. Statist. Ass.*, **72**, 642-650.

Shaked, M. and Tong, Y. L. (1985). Some partial orderings of exchangeable random variables by positive dependence. *J. Multivar. Anal.*, **17**, 333-349.

Shaked, M., and Shantikumar J. G. (1994). *Stochastic Orders and Their Applications*. San Diego: Academic Press.

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

Zanella, A. (1988). *Lezioni di Statistica, parte II: Strutture di dati in due o più dimensioni, sez. II: La Connessione*, Vita e Pensiero.