***S**tatistica & **A**pplicazioni*
Vol. III, n. 2, 2005
pp. 59-93

A new index to measure association between categorical and ordinal variables

Angelo Marcello Mineo[§]
Mariantonietta Ruggieri[§]

Summary: *In this paper a new index to analyse the dependence between categorical variables is presented and is compared to other measures of association, mainly based on the X^2 Pearson's statistic. The new index is compared to well-known measures of cograduation, as well. To restrict our comparisons, the domain we consider includes all the square contingency tables belonging to the same Fréchet class, that is all the contingency tables in which the marginal frequencies are given as fixed in both characters. Anyway, the new index is good even if there are no constrain on the marginal distributions, but only a constrain on n, the total of the observations.*

Keywords*: Fréchet class, association, cograduation, contingencies.*

1. Introduction

The study of the relationships between categorical variables is traditionally carried out by means of the information provided by a contingency table.

The first relevant contributions to the association analysis related to qualitative variables are due to Pearson (1904) and Yule (1900, 1912). According to Pearson the association could be measured by indices usually used for continuous variables, such as the linear correlation index, since he assumed that the variables involved in a contingency table have a continuous bivariate probability distribution. According to Yule such hypothesis is useless and wrong, especially for variables that are unquestionably discrete.

In fact, both the points of view are valid in different cases. For some nominal variables is extremely dubious considering continuous distributions. Nevertheless, such idea is widely used to build particular models and to make inference for contingency tables (Agresti, 2002).

The interest about the categorical data analysis resumed in the '50 years. Among others, Goodman and Kruskal (1954, 1959) proposed new indices and applied inferential procedures to association measures.

[§] Dipartimento di Scienze Statistiche e Matematiche "S. Vianelli"- University of Palermo – Viale delle Scienze, 90128 Palermo (e-mail: elio.mineo@dssm.unipa.it, ruggieri@dssm.unipa.it)

From the ‘70 years, useful tools to face multivariate analysis of categorical variables have been defined; among all, the main ones are represented by the log-linear models and the logit models (see Agresti, 2002, chapter 16).

In this paper particular relevance is given to the approach on the study of the association based on the total ordering of a set of tables according to their grade of connexion (Greselin-Zenga, 2001, 2002). In section 2 the difference between *distributive* independence and *absolute* or *complete* dependence is recalled. In section 4 the most important indices of association and cograduation used in literature are showed. In section 5 a new index of association that tries to overcome some of the limits presented by the indices described in section 4 is proposed. In section 6 comparisons between the indices seen in section 4 and the new index are showed. In order to make a lower number of comparisons, we choose to consider a set of square tables with the same marginal frequencies, constituting a Fréchet class, but usually the proposed index works well also for a set of tables with the same n. Moreover, the Fréchet class contains a sufficient amount of tables to make a significant number of comparisons among the considered indices. The Fréchet class is introduced in section 3.

2. Distributive independence

Let A and B be two categorical variables, with categories respectively given by $\{a_1, a_2,\dots, a_r\}$ and $\{b_1, b_2,\dots, b_c\}$. Let n be the units of a statistical population where A and B are collected. The study of association is usually carried on contingency tables like that shown in table 1, where n_{ij} denotes the joint frequency of the observations (a_i, b_j); $n_{i.}$ and $n_{.j}$ represent, respectively, the marginal frequencies of A and B.

Table 1. *A general contingency table*

A/B	b_1	b_2	…	b_j	…	b_c	Totals
a_1	n_{11}	n_{12}	…	n_{1j}	…	n_{1c}	$n_{1.}$
a_2	n_{21}	n_{22}	…	n_{2j}	…	n_{2c}	$n_{2.}$
…	…	…	…	…	…	…	…
a_i	n_{i1}	n_{i2}	…	n_{ij}	…	n_{ic}	$n_{i.}$
…	…	…	…	…	…	…	…
a_r	n_{r1}	n_{r2}	…	n_{rj}	…	n_{rc}	$n_{r.}$
Totals	$n_{.1}$	$n_{.2}$	…	$n_{.j}$	…	$n_{.c}$	n

It is well known that when each of the relative conditional distributions (of A given B or, symmetrically, of B given A) is equal to the corresponding marginal one, so that the relative conditional distributions are all equal among them, then between the characters A and B there is *distributive inde-*

pendence. Therefore, the hypothesis of independence can be expressed with the following equivalent relations:

$$\frac{n_{ij}}{n_{.j}} = \frac{n_{i.}}{n} \qquad \forall\ (i, j)$$

or

$$\frac{n_{ij}}{n_{i.}} = \frac{n_{.j}}{n} \qquad \forall\ (i, j).$$

Then, a generic value of the joint frequencies, under the hypothesis of independence, can be indicated as

$$\hat{n}_{ij} = \frac{n_{i.}n_{.j}}{n} \qquad \forall\ (i, j)$$

and takes the name of *theoretical frequency of independence*.

The difference between an observed frequency and the corresponding *theoretical frequency of independence*

$$c_{ij} = n_{ij} - \hat{n}_{ij}$$

is called *contingency* and plays a fundamental role in studies on association. Indeed, in case of independence it will be

$$c_{ij} = 0 \qquad \forall\ (i, j),$$

so an appropriate function of the distance between the observed frequencies n_{ij} and the corresponding *theoretical frequencies of independence* $\hat{n}_{ij}$ can be used as a measure of association between the two categorical variables A and B. The most of the indices proposed in literature to study the association of two characters are just based on the *contingencies* c_{ij}, or on the *relative contingencies* ρ_{ij}:

$$\rho_{ij} = \frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}}.$$

A situation of maximum dependence is not uniquely characterized, depending on $r \neq c$ or $r = c$. It may be unilateral, if $r \neq c$, or bilateral, if $r = c$. We have maximum dependence when to each modality of a character corresponds only one modality of the other character.

The following examples show respectively as (see table 2):

a) the character B depends on A perfectly, but the character A does not depend on B: in any row there is only one joint frequency different from 0 ($r > c$);
b) the character A depends on B perfectly, but the character B does not depend on A: in any column there is only one joint frequency different from ($r < c$);
c) the two characters A and B are perfectly associated: in any row and in any column there is only one frequency different from 0 ($r = c$).

Table 2. *Examples of perfect dependence between two variables A and B: in (a) B depends on A perfectly, in (b) A depends on B perfectly and in (c) A and B are perfectly associated*

A/B	$\mathbf{b_1}$	$\mathbf{b_2}$	$\mathbf{b_3}$	**Totals**
$\mathbf{a_1}$	n_{11}	0	0	$\boldsymbol{n_{1.}}$
$\mathbf{a_2}$	0	0	n_{23}	$\boldsymbol{n_{2.}}$
$\mathbf{a_3}$	0	0	n_{33}	$\boldsymbol{n_{3.}}$
$\mathbf{a_4}$	0	n_{42}	0	$\boldsymbol{n_{4.}}$
Totals	$\boldsymbol{n_{.1}}$	$\boldsymbol{n_{.2}}$	$\boldsymbol{n_{.3}}$	$\boldsymbol{n}$

(a)

A/B	$\mathbf{b_1}$	$\mathbf{b_2}$	$\mathbf{b_3}$	$\mathbf{b_4}$	**Totals**
$\mathbf{a_1}$	0	n_{12}	0	0	$\boldsymbol{n_{1.}}$
$\mathbf{a_2}$	0	0	n_{23}	n_{24}	$\boldsymbol{n_{2.}}$
$\mathbf{a_3}$	n_{31}	0	0	0	$\boldsymbol{n_{3.}}$
Totals	$\boldsymbol{n_{.1}}$	$\boldsymbol{n_{.2}}$	$\boldsymbol{n_{.3}}$	$\boldsymbol{n_{.4}}$	$\boldsymbol{n}$

(b)

A/B	$\mathbf{b_1}$	$\mathbf{b_2}$	$\mathbf{b_3}$	**Totals**
$\mathbf{a_1}$	0	n_{12}	0	$\boldsymbol{n_{1.}}$
$\mathbf{a_2}$	0	0	n_{23}	$\boldsymbol{n_{2.}}$
$\mathbf{a_3}$	n_{31}	0	0	$\boldsymbol{n_{3.}}$
Totals	$\boldsymbol{n_{.1}}$	$\boldsymbol{n_{.2}}$	$\boldsymbol{n_{.3}}$	$\boldsymbol{n}$

(c)

The point (c) includes the two following extreme situations, in which an ***absolute*** **association** or **dissociation** occurs (Kendall and Stuart, 1999); in particular:

1. there is ***absolute*** **positive association** (or simply ***absolute*** **association**) if all the frequencies n_{ij} are on the main diagonal, while the frequencies in the other cells are null;

2. there is ***absolute*** **negative association** (or simply ***absolute*** **dissociation**) if all the frequencies n_{ij} are on the secondary diagonal, while the frequencies in the other cells are null.

The cases 1. and 2. can be obtained just for square tables ($r \times r$); for rectangular tables ($r \times c$) the meaning of the difference between positive association and negative association is not straightforward.

We can not have *absolute* association if a constrain on the marginal totals, not only on n, is considered. Other situations can occur, i.e.:

3. the most of the frequencies n_{ij} are on the main diagonal, compatibly with the observed marginal totals, and the frequencies in the other cells could not be 0 (*complete* positive association);
4. the most of the frequencies n_{ij} are on the secondary diagonal, always considering the marginal totals as fixed, and the frequencies in the other cells could not be 0 (*complete* negative association).

Then, the difference between *absolute* and *complete* association depends on the dominion of the tables we are considering.

A good index of association should vary on the range [-1,1], assuming value +1 only when there is the case 1., value -1 only when there is the case 2. and value 0 only in case of *distributive independence*.

In literature, the measures of association do not always distinguish the two different situations of positive and negative association; many of them can assume only positive values. Moreover, if they vary in the range [-1,+1], often they assume exactly values +1 and –1 not only in case 1. and 2, but also in case 3. and 4. Sometimes, they assume value 0 even if the two characters are not independent, as we shall see in the next sections.

3. The Fréchet class

In the statistical literature, there are three different ways to solve problems concerning the association (Leti 1983, Zanella 1988):

- the marginal distribution is fixed for only one character;
- the marginal distributions are fixed for both the characters A and B;
- there are no constrain on the marginal distributions and only the total population n is fixed.

The new index, we will introduce in section 5, has as dominion all the square tables with n fixed; nevertheless, to show its properties and to compare it with other association indices (section 6), it is not necessary to consider a so large dominion. Thus, in this paper we consider the second constrain, i.e. we refer to the case where the marginal totals are fixed for both variables A and B.

Now we can define a Fréchet class: let $n_{i.}$ ($i = 1, 2,\dots, r$) and $n_{.j}$ ($j = 1, 2,\dots, c$) be the assigned marginal frequencies. All the possible tables having the same marginal distributions constitute the Fréchet class $T(n_{i.}; n_{.j})$.

We will denote with $T\{n_{ij}\}$ the generic table belonging to the same Fréchet class $\mathbf{T}(n_{i.}; n_{.j})$.

It is necessary to point out that:

1) to each table $T\{n_{ij}\} \in \mathbf{T}(n_{i.}; n_{.j})$ is always associated the same table $T\{\hat{n}_{ij}\}$ containing the independence frequencies $\hat{n}_{ij}$;
2) the independence table $T\{\hat{n}_{ij}\}$ does not belong necessarily to the Fréchet class $\mathbf{T}(n_{i.}; n_{.j})$, since, as it is known, the independence frequencies $\hat{n}_{ij}$ can not be integer numbers, even if $T\{\hat{n}_{ij}\}$ has the same marginal distributions of the Fréchet class $\mathbf{T}(n_{i.}; n_{.j})$;
3) the Fréchet class is obviously included in the set of tables with n fixed.

4. Measures of association

4.1 *Indices of association for (2×2) tables*

Let's consider the simplest case, i.e. the case of (2×2) tables. In literature, several indices have been proposed to measure the association of two dichotomic variables, but they present some drawbacks. The most important indices used so far are the Q and Y Yule's statistics and the V Pearson's statistic. It is worth remembering the harsh controversy arisen between the two statisticians in order to validate the superiority of each index.

The particularity of a (2×2) table is that, if the marginals are fixed, the knowledge of just one frequency n_{ij} is sufficient to determine the other ones. Then, an association analysis can be based on only one frequency. In particular, if the observed frequency n_{11} is compared with the corresponding theoretical frequency $\hat{n}_{11} = \dfrac{n_{1.}n_{.1}}{n}$, it can be said that:

1. if $n_{11} = \hat{n}_{11}$, then there is independence between A and B;
2. if $n_{11} > \hat{n}_{11}$, there is (positive) association between A and B;
3. if $n_{11} < \hat{n}_{11}$, there is dissociation, or negative association, between A and B.

So the contingency

$$c_{11} = n_{11} - \hat{n}_{11}$$

can be written also as

$$c_{11} = \frac{n_{11}n_{22} - n_{12}n_{21}}{n}.$$

In this case, we can fix two conditions to construct an association index:

a) the index has to be function of c_{11};
b) the index has to be increasing (positively in association case, negatively in dissociation case) when c_{11} increases.

As it can be easily showed, both the indices Q and Y, proposed by Yule (1900, 1912):

$$Q = \frac{n_{11}n_{22} - n_{12}n_{21}}{n_{11}n_{22} + n_{12}n_{21}} \tag{1}$$

$$Y = \frac{(n_{11}n_{22})^{1/2} - (n_{12}n_{21})^{1/2}}{(n_{11}n_{22})^{1/2} + (n_{12}n_{21})^{1/2}} \tag{2}$$

and the Pearson's index V:

$$V = \frac{n_{11}n_{22} - n_{12}n_{21}}{\sqrt{(n_{11} + n_{12})(n_{11} + n_{21})(n_{12} + n_{22})(n_{21} + n_{22})}} \tag{3}$$

satisfy the two above conditions, but Q and Y can be not considered as real "measures", because they assume value –1 and +1, respectively, not always in correspondence of an *absolute* negative or positive association. In fact, their values are:

1. zero, when the hypothesis of absence of association holds;
2. 1 when $n_{12} \times n_{21} = 0$;
3. -1 when $n_{11} \times n_{22} = 0$.

So both the Q and Y indices result equal to ± 1 if only one of the two frequencies lying on the main or secondary diagonal is zero.

V works better. It is equal to 1 when both the frequencies on the secondary diagonal are zero and –1 when $n_{11} = n_{22} = 0$. Moreover, it has an important and desirable property, i.e. it has increasing or decreasing values on a linear scale. In other words, as we shall show in section 6.1, when the V index varies from –1 to +1, because of the moving of a frequency unit inside the cells of the table, it increases its value of a constant quantity; in this case we have to suppose that each table, that is possible to obtain, has fixed marginals. So, the shown indices measure different aspects of association. For this reason, it appears of great importance an appropriate choice of the association index, otherwise conflicting conclusions could be reached, because different indices calculated on the same table could give different results.

4.2 Indices of association for (r×c) tables

In literature, many association indices are based on the squared contingencies or on the absolute value of contingencies.

For ($r \times c$) tables Pearson (1904) suggested to use the known X^2 index:

$$X^2 = \sum_{i,j} \frac{c_{ij}^2}{\hat{n}_{ij}} = n \sum_{i,j} \frac{c_{ij}^2}{n_{i.} n_{.j}} = n \left(\sum_{i,j} \frac{n_{ij}^2}{n_{i.} n_{.j}} - 1 \right) \tag{4}$$

Nevertheless, the X^2 index is not useful to measure association or dissociation, because it can assume only positive values. Moreover, its superior limit increases indefinitely as n increases. To avoid the influence of n, Pearson proposed to divide X^2 just by n:

$$\Phi^2 = \frac{X^2}{n} \tag{5}$$

Then, by trying to limit its possible value in the range [0, 1], Pearson proposed the *index of contingency* P:

$$\mathrm{P} = \left(\frac{X^2}{n + X^2} \right)^{1/2} = \left(\frac{\Phi^2}{\Phi^2 + 1} \right)^{1/2} \tag{6}$$

but P never reaches the limit value +1; its superior limit depends on r and c. For square tables, if an *absolute* association exists, it depends on the r categories (Kendall and Stuart, 1999):

$$\max \mathrm{P} = \left(\frac{r-1}{r} \right)^{1/2} ;$$

in fact, in this case it is max $X^2 = n(r-1)$.

For this reason, Tschuprow (1939) proposed a new index:

$$T = \left(\frac{X^2}{n[(r-1)(c-1)]^{1/2}} \right)^{1/2} \tag{7}$$

which in case of *absolute* association has value +1 if $r = c$, but if $r \neq c$ has not value +1. For square tables T is equal to the C Cramer's index (1946):

$$C = \left(\frac{X^2}{n \min[(r-1)(c-1)]} \right)^{1/2} \tag{8}$$

Such index is normalized, so it reaches value +1 whatever the number of rows and columns of the contingency table, by allowing the comparison between tables with different size.

Another known index, proposed by Mortara (1922), is given by the weighted arithmetic mean of $|\rho_{ij}|$, with weights $\hat{n}_{ij}$, multiplied for 1/2:

$$M_1 = \frac{1}{2n} \sum_{i,j} |\rho_{ij}| \cdot \hat{n}_{ij} = \frac{1}{2n} \sum_{i,j} |c_{ij}| = \frac{1}{2} \sum_{i,j} |f_{ij} - f_{i.} f_{.j}| \tag{9}$$

with $f_{ij} = \frac{n_{ij}}{n}$, $f_{i.} = \frac{n_{i.}}{n}$, $f_{.j} = \frac{n_{.j}}{n}$; instead, Pearson proposed an index based on the weighted quadratic mean of $|\rho_{ij}|$, with weights $\hat{n}_{ij}$:

$$M_2 = \left[\frac{1}{n} \sum_{i,j} |\rho_{ij}|^2 \cdot \hat{n}_{ij} \right]^{1/2} = \left[\frac{1}{n} \sum_{i,j} \frac{c_{ij}^2}{\hat{n}_{ij}} \right]^{1/2} = \left[\frac{X^2}{n} \right]^{1/2} \tag{10}$$

Obviously, for (2×2) tables is $M_2 = T = C$.

The index M_1 is almost normalized, but it never has value +1. In fact,

- if $r>c$, then

$$M_1 \leq 1 - \sum_j \left(\frac{n_{.j}}{n} \right)^2 ,$$

where the superior limit can be reached only if the character B depends perfectly on A;

- if $r<c$, then

$$M_1 \le 1 - \sum_i \left(\frac{n_{i.}}{n} \right)^2 ,$$

where the superior limit can be reached only if the character A depends perfectly on B (Leti, 1983).

It can be observed that the above superior limits correspond to the Gini's heterogeneity indices computed on the marginal distributions of B and A, respectively; so, their maximum will be (1-1/c) and (1-1/r). Then, the following indices assume value in [0, 1] (Frosini, 1993; Zenga, 1988):

- M_1/(1-1/c), if the marginal distribution of B is fixed;
- M_1/(1-1/r), if the marginal distribution of A is fixed;
- M_1/min[(1-1/r),(1-1/c)], if the total n is fixed and the marginal frequencies of the character with less modalities are all equal (obviously, if $r = c$, all the marginal frequencies have to be equal).

It is the case to emphasize as no one of the indices T, C and M_1, satisfies the condition of increasing negatively when a situation of dissociation occurs.

4.3 Measures of concordance/discordance for ordinal variables

It is necessary to point out that when the variables are ordinal, all the indices listed before are strictly linked to the way each variable is ordered. For example, in a (2×2) table, by changing rows and columns, the indices' values do not change, but their signs change. Therefore, the results obtained with the previous indices often appear ambiguous, even after the proposed transformations. In particular, when their value is far from zero, i.e. when there is not independence, it is not univocally determined the degree of association.

When A and B are two ordinal characters, there is association in this case: if the level of A increases, the level of B increases or decreases; so, the relationship between A and B is monotone and the degree of such relationship can be described by measures similar to the correlation ones.

The measures we shall analyse classify each pair of observations as concordant or discordant. If we suppose that the indices of two observations are (i, j) and (i', j'), then a pair is:

- concordant, when ($i<i'$) and ($j<j'$) or ($i>i'$) and ($j>j'$);
- discordant, when ($i<i'$) and ($j>j'$) or ($i>i'$) and ($j<j'$);
- tied, when the observations have the same classifications on A and/or B, i.e. ($i=i'$) and/or ($j=j'$).

A method to obtain a measure of association could be to treat the two characters with a metric scale. In this way, to measure the concordance, it could be sufficient calculating the common correlation indices; the problem is how we can assign scores to these variables. The choice of the scores is absolutely arbitrary, so in correspondence of different ways to assign the scores, different values of correlation indices could be obtained. For this reason, the association measures have not to depend on the used score system.

Among the most used measures of concordance/discordance, we want to remember the indices τ, Γ and *d* proposed by Goodman and Kruskal (1954, 1959), Kendall (1938) and Somers (1962), respectively:

$$\Gamma = \frac{N_c - N_d}{N_c + N_d} \tag{11}$$

$$\tau = \frac{N_c - N_d}{\sqrt{(N_c + N_d + T_a)(N_c + N_d + T_b)}} \tag{12}$$

$$d_{ba} = \frac{N_c - N_d}{N_c + N_d + T_b} \tag{13}$$

$$d_{ab} = \frac{N_c - N_d}{N_c + N_d + T_a} \tag{14}$$

For the (13) and (14), the used notation, d_{ba} and d_{ab}, is according to the response variable (B in the first case, A in the second one). Moreover, it is:

- N_c the number of concordant pairs;
- N_d the number of discordant pairs;
- T_a and T_b the number of tied pairs for the variable A and B, respectively.

The following relation holds: $\tau^2 = d_{ba}*d_{ab}$; moreover, if the marginal distributions are equal, $\tau = d_{ba}=d_{ab}$.

For square tables, the indices τ and *d* work better than Γ, because they assume exactly values +1 or -1, only if *absolute* positive or negative association occurs, while Γ can assume values –1 or +1 even if *complete* positive or negative association occurs. Nevertheless, it is evident that they can assume value 0 even if between the two characters there is not distributive independence: this happens when $N_c=N_d$.

5. A new index of association for square table

When (r×r) tables are considered, we propose a new index of association having the following good properties: it is function of the contingencies $c_{ij} = n_{ij} - \hat{n}_{ij}$; it takes a positive sign if association occurs and a negative sign if dissociation does; it assumes exactly the value -1 in case of *absolute* dissociation and the value +1 in case of *absolute* association, but only if the variability of the frequencies on diagonals is minimum. This index is:

$$M = \pm \frac{\sum_{i,j} |c_{ij}|}{\max \sum_{i,j} |c_{ij}|} = \pm \frac{\sum_{i,j} \left| n_{ij} - \frac{n_{i.} n_{.j}}{n} \right|}{\max \sum_{i,j} \left| n_{ij} - \frac{n_{i.} n_{.j}}{n} \right|} = \pm \frac{\sum_{i,j} |n_{ij} n - n_{i.} n_{.j}|}{\max \sum_{i,j} |n_{ij} n - n_{i.} n_{.j}|} \quad (15)$$

where $\sum_{i,j} |c_{ij}|$ represents the sum of the absolute values regarding all the differences $c_{ij} = n_{ij} - \frac{n_{i.} n_{.j}}{n}$, while $\max \sum_{i,j} |c_{ij}|$, fixed n, is the maximum value that $\sum_{i,j} |c_{ij}|$ can assume. Obviously, $\sum_{i,j} |c_{ij}|$ reaches its maximum value in a frequency table where the association is maximum, that is in a table of *absolute* association, in which the frequencies n_{ij} are all on the main diagonal. But, if n is fixed, the frequencies can be distributed on the diagonal in different ways. In particular, the more the frequencies are equally distributed on the diagonal, the more the intensity of association increases. Thus, $\sum_{i,j} |c_{ij}|$ assumes its maximum value when the variability among the frequencies on the diagonal is the lowest.

Likewise, $\sum_{i,j} |c_{ij}|$ reaches its maximum value in a frequency table where the dissociation is maximum, that is in a table of *absolute* dissociation, in which the frequencies n_{ij} are all equally distributed on the secondary diagonal.

To understand better what we mean, let's consider some examples.

The following table shows, for different values of n (n = 5, 6, 7, 8), the different ways in which n can be divided among the cells of the main diagonal (the same is by considering the secondary diagonal). For simplicity, we suppose that $r = c = 3$. So, any row of the following table represents the possible values of the main diagonal of a (3×3) table for the considered n. For each case, the standard deviation σ among the frequencies, the $\sum_{i,j} |c_{ij}|$ and the X^2 Pearson's statistic are reported. The values of σ, $\sum_{i,j} |c_{ij}|$ and X^2 do not change if the rows (or the columns) of the square table are inverted. Thus, if we fix the values of the frequencies inside each cell, it is unnecessary to consider all the permutations, because is not important the order of the objects on the diagonal:

Cells on the diagonal			σ	$\Sigma_{i,j}\|c_{ij}\|$	X^2
n=5					
1	1	3	0,94	5,60	10,00
1	2	2	0,47	6,40	10,00
n=6					
1	1	4	1,41	6,00	12,00
1	2	3	0,82	7,33	12,00
2	2	2	0,00	8,00	12,00
n=7					
1	1	5	1,89	6,29	14,00
1	2	4	1,25	8,00	14,00
1	3	3	0,94	8,57	14,00
2	2	3	0,47	9,14	14,00
n=8					
1	1	6	2,36	6,50	16,00
1	2	5	1,70	8,50	16,00
1	3	4	1,25	9,50	16,00
2	2	4	0,94	10,00	16,00
2	3	3	0,47	10,50	16,00

As it can be noted, the sum $\sum_{i,j}\left|c_{ij}\right|$ is maximum when σ is minimum, so that to build the table where the association is maximum it is necessary dividing *n* equally on the diagonal as far as possible. On the contrary, X^2 achieves its maximum in any case, whatever the frequencies are shared on the diagonal. It is easy to show that also the indices normalized in [-1,1], seen in section 4, reach their extremes always when the frequencies lie on the secondary or main diagonal, respectively, even if the variability between the frequencies is not a minimum.

The sign of the *M* index is easy to determine for dichotomic tables. Indeed, in this case, as we shall show in section 5.1, it depends on the sign of c_{11}. In particular, it is positive if the sign of c_{11} is positive and it is negative if the sign of c_{11} is negative. For square frequency tables with $r \neq 2$, the way to determine the sign of *M* is described in details in section 5.2.

The *M* index provides a good measure of the association/dissociation among the *r* categories of two categorized variables: it changes between –1 and 0 in presence of dissociation and between 0 and +1 in presence of association, as it will result evident later on.

For all these reasons, the M index seems to be more coherent than the indices of association T (7) and C (8). In fact, T and C not only do not distinguish the association by the dissociation, but they are obtained by dividing X^2 to its maximum, where this maximum is always the value that X^2 assumes in a whatever table where the frequencies occur all on the main diagonal, but these frequencies not always have the lowest variability.

For analogous reasons, the M index overcomes the limits of the index of cograduation τ (12) and d (13, 14) as well, since its numerator is divided only by the maximum value achievable in a table where the dissociation, or alternatively the association, is *absolute*. Moreover, M is zero only if distributive independence occurs, while τ and d are zero even if N_c=N_d = 0 (see section 4.3).

5.1 *The case of (2×2) tables*

When the two observed characters are dichotomic, as we have already said in section 4.1, the knowledge of the marginals and of the frequency n_{11} is equivalent to the knowledge of the other frequencies, so that computing $|c_{11}| = \left| n_{11} - \frac{n_{1.}n_{.1}}{n} \right|$ on the observed table and on the table where the association/dissociation is maximum is sufficient to determine the M index, that, for this kind of tables, has this simple expression:

$$M = \pm \frac{|c_{11}|}{\max|c_{11}|} = \pm \frac{|n_{11}n - n_{1.}n_{.1}|}{\max|n_{11}n - n_{1.}n_{.1}|}$$

Indeed, the $|c_{ij}|$ are all equals, so $\sum_{i,j} |c_{ij}| = 4|c_{11}|$; the same thing happens for the quantity $\max|c_{ij}|$, in fact $\max \sum_{i,j} |c_{ij}| = 4 \max|c_{11}|$.

Really, we can avoid to specify the sign by considering c_{11} instead of $|c_{11}|$:

$$M = \frac{c_{11}}{\max|c_{11}|} \tag{16}$$

In fact:

a) if c_{11} = 0, then there is independence between A and B;

b) if c_{11} >0, there is association or positive association between A and B;

c) if $c_{11} < 0$, there is dissociation or negative association between A and B.
To determine the maximum of $|c_{11}|$, the maximum association/dissociation table is necessary. Then, we have to compute the module of n by r = 2: if the result is zero, we attribute n/r to each cell on the main/secondary diagonal, otherwise if the result is 1, this can be attributed indifferently to one of the two cells.

5.2 Computing the new index for (r×r) tables

To measure the association in (r×r) frequency tables, it is necessary to use the (15). In this case, it is necessary to determine first the sign of M on the observed frequency table to see if association or dissociation occurs. The way to determine the sign is given by the difference between the sum of the frequencies lying on the main diagonal $\sum_{i=1}^{r} n_{ii}$ and the sum of the frequencies lying on the secondary diagonal $\sum_{i=1}^{r} n_{i(r-i+1)}$. If this difference is positive, then a positive sign is chosen, otherwise a negative sign is considered. The difference between $\sum_{i=1}^{r} n_{ii}$ and $\sum_{i=1}^{r} n_{i(r-i+1)}$ could be zero: in this case, the difference between the sum of the contingencies on the main diagonal $\sum_{i=1}^{r} c_{ii}$ and the sum of the contingencies on the secondary diagonal $\sum_{i=1}^{r} c_{i(r-i+1)}$ can be computed. If also this difference is zero, but this rarely happens, then the difference between the number of concordant pairs N_c and the number of discordant pairs N_d (see section 4.3) can be considered. Finally, if $N_c = N_d = 0$, the indices of cograduation we have seen in section 4.3 are wrongly zero, even if there is not a distributive independence, as we shall see in section 6.2. In this case, we choose the positive sign, if the sum of the positive contingencies is more than the sum of the negative ones and vice versa.

To determine the table of maximum association/dissociation is sufficient distributing uniformly n on the main/secondary diagonal. If this is not possible, that is if n/r is not an integer, then its smaller integer value is attributed to each frequency on the diagonal. The module of n and r can not be greater than *(r-1)*, so it will be again shared, unit by unit, to *(r-1)* diagonal cells, where the *(r-1)* cells are chosen arbitrarily. In fact, however we choose the *(r-1)* cells for adding a unit of the module, the sum of the absolute contingencies $\sum_{i,j} |c_{ij}|$ will be unchanged. This sum will be that one where the variability of the frequencies is a minimum, so it is the highest sum that we can compute for all the tables with fixed n, that is where no constrain on the marginal distributions is considered, but only on n.

6. Comparison between the proposed index and other measures of association and cograduation

6.1 *Comparing association indices for (2×2) tables*

Let's consider an example. A sample of 41 psychiatric patients are classified in psychopathic and neurotic (see table 3).

Table 3. *Contingency table related to a sample of 41 psychiatric patients*

Suicide tendency	**Kind of patient** Psychotic	Neurotic	*Totals*
Present	2	7	9
Absent	18	14	32
Totals	20	21	41

$T\{n_{ij}\} = T_3$

The corresponding independence table is shown in table 4.

Table 4. *Independence contingency table for the table 3*

Suicide tendency	**Kind of patient** Psychotic	Neurotic	*Totals*
Present	4.39	4.61	9
Absent	15.61	16.39	32
Totals	20	21	41

$T\{\hat{n}_{ij}\}$

The Fréchet class including $T\{n_{ij}\}$ is formed by all the tables having marginal totals of row equal to **r** = (9, 32) and marginal totals of column equal to **c** = (20, 21). If r = c = 2, as in this case, such tables can be generated by giving to n_{11} the following values (Greselin, 2003):

$$\max(0, n_{.1} - n_{2.}) \leq n_{11} \leq \min(n_{1.}, n_{.1});$$

thus, it is

$$\max(0, 20 - 32) \leq n_{11} \leq \min(9, 20) \qquad \Rightarrow \qquad 0 \leq n_{11} \leq 9$$

The cardinality of the Fréchet class is given by:

$$\#_{2,2}\,(\mathbf{r}, \mathbf{c}) = \min(n_{1.}, n_{.1}) - \max(0, n_{.1} - n_{2.}) + 1 =$$

$$= \min(9, 20) - \max(0, 20\text{-}32) + 1 = 10.$$

As we have already said, a main feature of these tables is that, if the marginal totals are fixed, the knowledge of only one frequency, in this case of n_{11}, permits to determine the frequencies of the other cells.

Indeed, $n_{12} = n_{1.} - n_{11}$, $n_{21} = n_{.1} - n_{11}$, $n_{22} = n_{2.} - n_{12}$.

Then, by means of an algorithm allowing unitary increments and decrements of frequencies, the tables of Fréchet class can be determined. We have started from the table T_1, where n_{11} assumes its minimum value, and by transfers of frequencies we have found all the tables:

0	←9	9
20→	12	32
20	21	41

T_1

1	←8	9
19→	13	32
20	21	41

T_2

2	←7	9
18→	14	32
20	21	41

T_3

3	←6	9
17→	15	32
20	21	41

T_4

4	←5	9
16→	16	32
20	21	41

T_5

5	←4	9
15→	17	32
20	21	41

T_6

6	←3	9
14→	18	32
20	21	41

T_7

7	←2	9
13→	19	32
20	21	41

T_8

8	←1	9
12→	20	32
20	21	41

T_9

9	←0	9
11→	21	32
20	21	41

T_{10}

The arrows show the frequency transfers allowing the transformation from a table to the next one.

As it can be noted, the independence table $T\{\hat{n}_{ij}\}$ does not belong to the tables of the Fréchet class; this is obvious, because the $\hat{n}_{ij}$ are not integers.

The corresponding tables of the contingencies $c_{ij} = n_{ij} - \hat{n}_{ij}$, associated to the Fréchet class, are:

-4,39	4,39	0

-3,39	3,39	0

-2,39	2,39	0

4,39	-4,39	0
0	0	0

T_1

3,39	-3,39	0
0	0	0

T_2

2,39	-2,39	0
0	0	0

T_3

-1,39	1,39	0
1,39	-1,39	0
0	0	0

T_4

-0,39	0,39	0
0,39	-0,39	0
0	0	0

T_5

0,61	-0,61	0
-0,61	0,61	0
0	0	0

T_6

1,61	-1,61	0
-1,61	1,61	0
0	0	0

T_7

2,61	-2,61	0
-2,61	2,61	0
0	0	0

T_8

3,61	-3,61	0
-3,61	3,61	0
0	0	0

T_9

4,61	-4,61	0
-4,61	4,61	0
0	0	0

T_{10}

By means of the sign of c_{11}, it is possible to understand if association or dissociation occurs. So, for the tables T_1, T_2, T_3, T_4, T_5, it is $c_{11}<0$; a situation of dissociation is present and such dissociation decreases from T_1 to T_5, therefore T_1 is the table with maximum dissociation. For the tables T_6, T_7, T_8, T_9, T_{10}, it is $c_{11}>0$; a situation of association is present and such association increases from T_6 to T_{10}, therefore T_{10} is the table with maximum association.

We have computed on each table of the Fréchet class:

- the Yule's indices Q and Y;
- the Pearson's indices V, X^2, ϕ^2, P, M_2;
- the index T of Tschuprow and C of Cramer;
- the M_1 Mortara's index;
- the indices of cograduation, Γ, τ, d_{ba}, d_{ab};
- the new index M.

It is easy to show that in this case it is $C=T=M_2$.

To compute M we have obtained the table of maximum association (dissociation) by dividing n for the values of the main (secondary) diagonal. The module of n by r = 2 is 1, in this case. Then, the result is still attributed to one of the two cells on the diagonal. The tables below report the maximum association and dissociation tables, with the respective contingency tables:

21	0	21
0	20	20
21	20	41

10,24	-10,24	0
-10,24	10,24	0
0	0	0

Max association c_{ij}

0	21	21
20	0	20
20	21	41

-10,24	10,24	0
10,24	-10,24	0
0	0	0

Max dissociation c_{ij}

Thus, the denominator of M will be always 10.24 (see section 5.1), while its numerator is the element c_{11} (with sign) of the observed table. The association and cograduation indices computed on the Fréchet class tables are reported on table 5.

Table 5. *Association and cograduation indices for the considered Fréchet class tables*

table	Q	Y	V	X^2	ϕ^2	P	C	M_1	M
T_1	-1,000	-1,000	-0,518	10,982	0,268	0,460	0,518	0,214	-0,429
T_2	-0,842	-0,548	-0,400	6,549	0,160	0,371	0,400	0,165	-0,331
T_3	-0,636	-0,359	-0,282	3,255	0,079	0,271	0,282	0,117	-0,233
T_4	-0,388	-0,202	-0,164	1,101	0,027	0,162	0,164	0,068	-0,136
T_5	-0,111	-0,056	-0,046	0,087	0,002	0,046	0,046	0,019	-0,038
T_6	0,172	0,087	0,072	0,212	0,005	0,072	0,072	0,030	0,060
T_7	0,440	0,232	0,190	1,477	0,036	0,186	0,190	0,079	0,157
T_8	0,673	0,387	0,308	3,881	0,095	0,294	0,308	0,127	0,255
T_9	0,860	0,570	0,426	7,425	0,181	0,392	0,426	0,176	0,352
T_{10}	1,000	1,000	0,543	12,108	0,295	0,478	0,543	0,225	0,450

table	τ	Γ	d_{ba}	d_{ab}
T_1	-0,518	-1,000	-0,625	-0,429
T_2	-0,400	-0,840	-0,483	-0,331
T_3	-0,282	-0,640	-0,340	-0,233
T_4	-0,164	-0,390	-0,198	-0,136
T_5	-0,046	-0,110	-0,056	-0,038
T_6	0,072	0,170	0,087	0,060
T_7	0,190	0,440	0,229	0,157
T_8	0,308	0,670	0,372	0,255
T_9	0,426	0,860	0,514	0,352
T_{10}	0,543	1,000	0,656	0,450

As it can be noted, the indices Q, Y and Γ assume value –1 (+1) in T_1 (T_{10}), but for these tables, even if the dissociation (association) is maximum compared to the other tables, there is not an absolute dissociation (association), i.e. the frequencies do not lie all on the secondary (main) diagonal. The V, τ, d_{ba}, d_{ab} and M indices work better. They assume positive (negative) values when positive (negative) association occurs with higher (lower) values in T_{10} (T_1), but never achieve the extremes +1, -1 for the considered example.

It is interesting to point out that V and τ are equal for (2×2) tables, while the equality between M and d_{ab} seems only a case, as we can verify if another Frèchet class is considered. That will be better evident in the next section, when other examples are showed. Moreover, if all the (2×2) square tables with n = 41 are considered, M is the only index that reaches the value +1 (-1) only when the frequencies on the main (secondary) diagonal have the lowest variability, while the V and τ indices take value +1 (-1) always when the frequencies lie on the two diagonals, with no consideration of their variability.

With regard to the T, C and M_2 indices, that are equal for (2×2) tables, as we expected, they are better than X^2, ϕ^2, P, M_1 because they are normalized in [0, 1], although they achieve the value +1 when any *absolute* association occurs. In this class they reach the highest value in T_{10}, where the association is greater than in the other tables. The graphics below (fig.1, fig.2) show the behaviour of the analyzed indices; the indices V, τ, d_{ba}, d_{ab} and M have the property to vary constantly on a linear scale, when a frequency moves from a cell to another one, increasing the association.

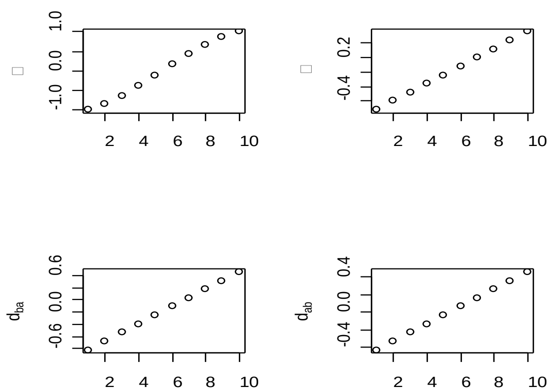

Figure 1: *Graphics reporting the cograduation indices computed on the Fréchet class [$n_{1.}$=9, $n_{2.}$=32, $n_{.1}$=20, $n_{.2}$=21]*

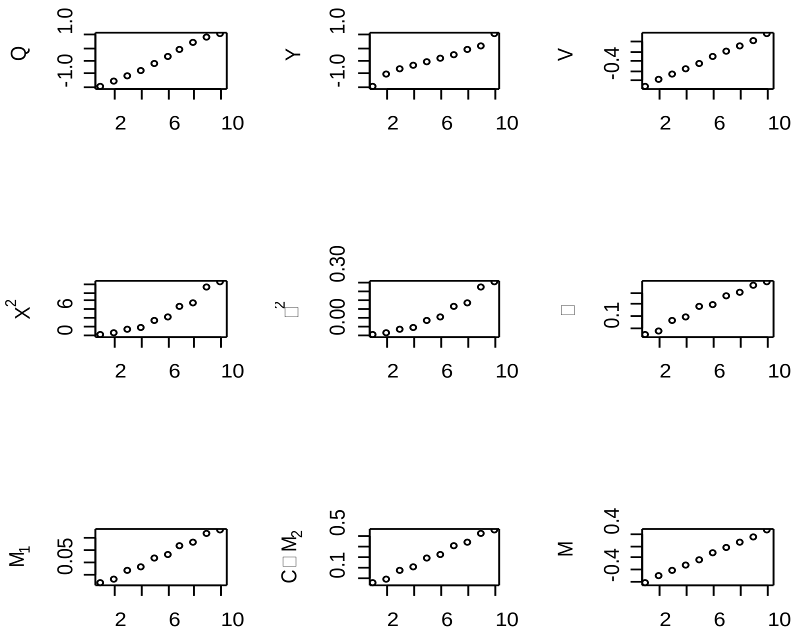

Figure 2: *Graphics reporting the association indices computed on the Fréchet class* $[n_{1.}=9, n_{2.}=32, n_{.1}=20, n_{.2}=21]$

6.2 *Comparing association indices for (r×r) tables*

In the following examples (3×3) tables are considered. We have chosen some particular cases collecting all the most significant situations that could occur, that is:

1. the marginals of the observed table are all equal;
2. the marginal distributions of the observed table are equal with equal extreme frequencies;
3. the marginal distributions of the observed table are equal with any frequencies;
4. the marginal distributions of the observed table are different.

For each observed table the Fréchet class has been generated according to an algorithm we have developed by using the free software R (R Development Core Team, 2004). The problem of determining the number of tables with equal marginals it has been dealt by many authors that have considered different kind of tables (Leti 1970, Greselin 2003).

We have computed on each table of a Fréchet class the following indices:

- the Pearson's indices X^2, ϕ^2, P, M_2;

- the normalized X^2 index (X^2/max X^2), that we indicate by NX^2 and the C Cramer's index;
- the M_1 Mortara's index;
- the indices of cograduation Γ, τ, d_{ba}, d_{ab};
- the new index M.

EXAMPLE 1: Fréchet class [$n_{1.}=n_{2.}=n_{3.}=n_{.1}=n_{.2}=n_{.3}=3$]
The computed indices for this Fréchet class, with cardinality 55, are reported in Appendix A, table 1. To compute the denominator of M we have divided n = 9 by r = 3. In this case, we have only to assign the result of n/r to each cell in the main/secondary diagonal; below is also reported the table of independence:

T_{20} Maximum association

3	0	0
0	3	0
0	0	3

T_1 Maximum dissociation

0	0	3
0	3	0
3	0	0

T_{30} Distributive independence

1	1	1
1	1	1
1	1	1

The considered Fréchet class includes the following tables as well:

T_{17}

3	0	0
0	0	3
0	3	0

T_{55}

0	3	0
3	0	0
0	0	3

T_4

0	0	3
3	0	0
0	3	0

T_{52}

0	3	0
0	0	3
3	0	0

The above tables are the most significant cases. Let's compare the computed indices. As we expected, the X^2, ϕ^2, P and M_2 indices assume value 0 in T_{30}, achieving their maximum value for the tables T_1, T_4, T_{17}, T_{20}, T_{52} and T_{55}, where perfect association occurs. For the last tables the indices NX^2, C and M_1 assume value 1, since they are normalized. For this example we can consider M_1 normalized, i.e. we can divide M_1 for (1-1/r), because the marginal frequencies are all equal (see section 4.2).

The indices Γ, τ, d_{ba}, d_{ab} indices (here $\tau = d_{ba} = d_{ab}$, because the marginal distributions are equal) and the new M index assume value 0 in T_{30}, +1 in T_{20} (*absolute* positive association) and -1 in T_1 (*absolute* negative association). The new M index assumes value +1 also in T_{17} and T_{55} and value -1 also in T_4 and T_{52}. For these tables seem to be more informative the indices Γ, τ, d_{ba} and d_{ab}, that assume respectively values 0.33 and –0.33, as is correct, since only a pair of modalities is associated/dissociated in these cases. Nevertheless, the considered class includes only tables with given marginals. If we consider all the tables where only n is fixed, diagonal tables where the frequencies have not the lowest variability occur. For all these tables the indices of cograduation assume always values –1 or +1, according to the sign of the association, while M does not, so in this situation it seems better. Moreover, it is possible that the Γ, τ, d_{ba} and d_{ab} indices assume wrongly value 0 also when there is not distributive independence, as we are going to see in the next examples.

In Appendix B we report the graphical behaviour of all the computed indices. In fig.1b the indices sorted on the whole Fréchet class are showed, while in fig.2b we do not consider the repeated values.

EXAMPLE 2: Fréchet class [$n_{1.} = n_{.1} = n_{3.} = n_{.3} = 2$, $n_{2.} = n_{.2} = 3$]
The computed indices for this Fréchet class, with cardinality 25, are reported in Appendix A, table 2. To compute the denominator of M, we have that the result of the module of $n = 7$ by $r = 3$ is different from 0. Thus, to construct the tables where the association/dissociation is maximum, we have to approximate n/r to its smallest integer and assign it to each cell in the main/secondary diagonal. The rest, that is 1, could be assigned to one of the 3 cells indifferently, but since a Fréchet class is considered, we attribute it to the cell (2, 2), so that the marginals remain unchanged. Below the tables of maximum association/dissociation and the independence table are reported:

2	0	0
0	3	0
0	0	2

T_{10}
Maximum association

0	0	2
0	3	0
2	0	0

T_1
Maximum dissociation

0.57	0.86	0.57
0.86	1.29	0.86
0.57	0.86	0.57

Distributive independence

As it can be noted in the distributive independence table, no integer element appears, then it is immediately evident that this table does not belong to the considered Fréchet class. In fact, the most of the computed indices never assume value 0.

The X^2, ϕ^2, P, M_1 and M_2 indices reach their maximum value for the tables T_1 and T_{10}, where perfect association occurs. For the same tables the indices NX^2 and C assume value 1. The Γ, τ, d_{ba}, d_{ab} and M indices assume value +1 in T_{10} (*absolute* positive association) and -1 in T_1 (*absolute* negative association). Also for this example $\tau = d_{ba} = d_{ab}$, because the marginal distributions are equal. For the following tables, it is $N_c = N_d$ (see section 4.3):

1	0	1
0	3	0
1	0	1

T_5

1	0	1
1	1	1
0	2	0

T_6

0	1	1
2	1	0
0	1	1

T_{15}

1	1	0
0	1	2
1	1	0

T_{16}

0	2	0
1	1	1
1	0	1

T_{23}

The Γ, τ, d_{ba} and d_{ab} indices assume wrongly value 0, while M does not. So the new M index seems to be more informative because in the tables above there is not independence.

Also in this case, if only n is fixed, the Γ, τ, d_{ba} and d_{ab} indices would assume values –1 or +1 even for absolute dissociation/association tables, where the variability of the diagonal frequencies is not the lowest.

In fig.3b of the Appendix B the graphical behaviour of all the sorted indices is reported. In fig.4b their repeated values are not considered.

EXAMPLE 3: Fréchet class [$n_{1.} = n_{.1} = 3$, $n_{2.} = n_{.2} = 2$, $n_{3.} = n_{.3} = 4$]
The computed indices for this Fréchet class, with cardinality 39, are reported in Appendix A, table 3. To determine the denominator of M, the table with maximum association/dissociation is necessary. By executing the module operation of $n = 9$ by $r = 3$, it is obtained 0, so we attribute n/r to each cell on the diagonal. The tables of maximum association and dissociation, that are:

3	0	0
0	3	0
0	0	3

Maximum association

0	0	3
0	3	0
3	0	0

Maximum dissociation

do not belong to the considered Fréchet class, because the constrains on the marginals are not respected. For this reason M can not reach the value +1 and –1 in table of the class.

The X^2, ϕ^2, P, M_1 and M_2 indices reach their maximum value for the table T_{19}, where the indices NX^2, C, Γ, τ, d_{ba} and d_{ab} assume value 1, as we expected. Indeed, in this table:

3	0	0
0	2	0
0	0	4

T_{19}

there is a perfect positive association; but the intensity of the association is not the maximum, if only a constrain on n is given.

So, it is not correct that in such situation the above indices achieve their maximum value. In this case, M works better than the other indices, because it does not achieve the value 1 in T_{19}, but only when the frequencies n_{11}, n_{22}, n_{33} are equally distributed.

It is still the case to point out that again the Γ, τ, d_{ba} and d_{ab} indices (for this example $\tau = d_{ba} = d_{ab}$, because the marginal distributions are equal) are 0 in T_{10} and T_{26}, where there is not independence. Indeed, the considered Fréchet class does not include the independence table:

1	0.67	1.33
0.67	0.44	0.89
1.33	0.89	1.78

Distributive independence

For all these reasons it is evident that M is the index attaining the best results. In fig.5b of the Appendix B the graphical behaviour of all the sorted indices in the considered Fréchet class is reported. In fig.6b their values are considered only once.

EXAMPLE 4: Fréchet class [$n_{1.}=2$, $n_{2.} = 4$, $n_{3.} = 3$, $n_{.1} = 5$, $n_{.2}=1$, $n_{.3} = 3$]
The computed indices for this Fréchet class, with cardinality 24, are reported in Appendix A, table 4. As in the previous examples, to determine the table with maximum association/dissociation and then the denominator of M, it is necessary to divide $n = 9$ by $r = 3$. Since a rest 0 is obtained, the result of n/r is given to each cell on the table diagonal. The tables of maximum association and dissociation:

3	0	0
0	3	0
0	0	3

Maximum association

0	0	3
0	3	0
3	0	0

Maximum dissociation

do not belong to the considered Fréchet class. In fact, no table with absolute association or dissociation belongs to it. Inside the Fréchet class, the table where we have the maximum association is T_{24}. For this table the X^2, ϕ^2, P, M_1, M_2, NX^2, C and M indices reach their maximum value, while the Γ, τ, d_{ba} and d_{ab} indices assume the maximum value in the table T_{17}.

In particular, Γ assumes value 1, but in T_{17} there is not an *absolute* positive association. Moreover, Γ assumes the value –1 in T_1, but such table does not have an *absolute* negative association. In this case, τ, d_{ba} and d_{ab}, that are not equal because the marginal distributions are different, assume their minimum value. Instead, M achieves its lowest value in T_{11}. Again, the Γ, τ, d_{ba} and d_{ab} indices are 0 where there is not independence (T_{12}, T_{20}, T_{22}). Also in this case, the Fréchet class does not include the independence table:

1.11	0.22	0.67
2.22	0.44	1.33
1.67	0.33	1

Distributive independence

In fig.7b of the Appendix B we report the graphical behaviour of the indices sorted inside the Fréchet class, while in fig.8b such values are not repeated more than once.

7. Conclusions

In this paper a new index of association, to measure the dependence between two categorical variables when frequency tables are observed, has been presented. It is showed as the new index works well for (2×2) tables and more generally for (r×r) contingency tables, with any marginal distributions. It represents a real "measure" varying in all the points of a scale from –1 to 0 and from 0 to +1 when an unit of frequency moves inside the cells of a table. It has the desirable property to take value +1 when there is *absolute* association, -1 when there is *absolute* dissociation and assumes value 0 only in presence of distributive independence. We have compared the new index with the most used indices of association and cograduation, well-known in literature, showing the goodness of the new index. Some significant examples report the value of all the indices computed in a whole Fréchet class, by considering all the possible tables with given margins.

In a next paper, we intend to generalize the proposed index to (r×c) tables, by investigating its properties. Our purpose is dealing with inferential aspects, too.

References

Agresti A. (2002). *Categorical data analysis*. Wiley, New York.

Castellano V. (1981). *Istituzioni di Statistica*. Edited by B.M. Cesare, Roma, Sistema.

Cramer H. (1946). *Mathematical methods of Statistics*. Princeton University Press, Princeton.

Frosini B.V. (1993). *Connessione, regressione, correlazione*. Celuc libri, Milano.

Goodman L.A., Kruskal W.H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, **49**, 732-764.

Goodman L.A., Kruskal W.H. (1959). Measures of association for cross classifications II: Further discussion and references. *Journal of the American Statistical Association*, **58**, 123-163.

Greselin F. (2003). Counting and enumerating frequency tables with given margins. *Statistica & Applicazioni*, **I**, 87-104.

Greselin F., Zenga M. (2001). Partial and total orderings on tables in the Fréchet class. *Research report n° 38/2001, Department of quantitative methods for Economics, in revision (Research report n° 79/2004), University of Milan "Bicocca"*.

Greselin F., Zenga M. (2002). "*Measures of association in the Fréchet class*". Atti della XLI Riunione Scientifica della SIS, Sessioni Spontanee, Cleup, Padova, 33-36.

Kendall M. (1938). A new measure of rank correlation. *Biometrika*, **36**, 177-193.

Kendall M., Stuart A. (1999). *The advanced theory of Statistics*. Vol.2A, C. Griffin, London.

Leti G. (1970). La distribuzione delle tabelle della classe di Fréchet. *Metron*, **XXVIII**, 86-121.

Leti G. (1983). *Statistica descrittiva*. Il Mulino, Bologna.

Mortara G. (1922). *Lezioni di Statistica metodologica*. Città di Castello, Società tipografica "Leonardo da Vinci".

Pearson K. (1904). *On the theory of contingency and its relation to association and normal correlation*. Drapers' Co. Memoirs, Biometric Series, London.

R Development Core Team, (2004). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

Somers R. (1962). A new asymmetric measure of association for ordinal variables. *American Sociological Review*, **27**, 799-811.

Tschuprow A.A. (1939). *Principles of the mathematical theory of correlation*. William Hodge & Co, London.

Yule G. (1900). On the association of attributes in Statistics. *Philosophical Transaction of the Royal Society of London (A)*, **194**, 257-319.

Yule G. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, **75**, 579-642.

Zanella A. (1988). *Lezioni di Statistica. Parte II: Struttura di dati in due o più dimensioni. La Connessione*. Vita e pensiero, Milano.

Zenga M. (1988). *Introduzione alla Statistica descrittiva*. Vita e pensiero, Milano.

APPENDIX A

Table 1a. *Association and cograduation indices computed on the Fréchet class* $[n_{1.}=n_{2.}=n_{3.}=n_{.1}=n_{.2}=n_{.3}=3]$

table	X^2	ϕ^2	P	NX^2	C	M_1	M_2	M	τ	Γ	d_{ba}	d_{ab}
T_1	18,00	2,00	0,82	1,00	1,00	1,00	1,41	-1,00	-1,00	-1,00	-1,00	-1,00
T_2	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,78	-0,91	-0,78	-0,78
T_3	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,56	-0,65	-0,56	-0,56
T_4	18,00	2,00	0,82	1,00	1,00	1,00	1,41	-1,00	-0,33	-0,33	-0,33	-0,33
T_5	6,00	0,67	0,63	0,33	0,58	0,50	0,82	-0,50	-0,48	-0,62	-0,48	-0,48
T_6	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,33	-0,39	-0,33	-0,33
T_7	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,30	-0,40	-0,30	-0,30
T_8	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,15	-0,20	-0,15	-0,15
T_9	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,11	-0,13	-0,11	-0,11
T_{10}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	0,50	0,04	0,05	0,04	0,04
T_{11}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	-0,50	-0,04	-0,05	-0,04	-0,04
T_{12}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,15	0,20	0,15	0,15
T_{13}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,33	0,39	0,33	0,33
T_{14}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,11	0,13	0,11	0,11
T_{15}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,30	0,40	0,30	0,30
T_{16}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	0,50	0,48	0,62	0,48	0,48
T_{17}	18,00	2,00	0,82	1,00	1,00	1,00	1,41	1,00	0,33	0,33	0,33	0,33
T_{18}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,56	0,65	0,56	0,56
T_{19}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,78	0,91	0,78	0,78
T_{20}	18,00	2,00	0,82	1,00	1,00	1,00	1,41	1,00	1,00	1,00	1,00	1,00
T_{21}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,78	-0,91	-0,78	-0,78
T_{22}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,59	-0,80	-0,59	-0,59
T_{23}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	-0,50	-0,48	-0,62	-0,48	-0,48
T_{24}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	-0,50	-0,41	-0,52	-0,41	-0,41
T_{25}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,30	-0,40	-0,30	-0,30
T_{26}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,11	-0,13	-0,11	-0,11
T_{27}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,30	-0,40	-0,30	-0,30
T_{28}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,15	-0,20	-0,15	-0,15
T_{29}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,15	-0,20	-0,15	-0,15
T_{30}	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
T_{31}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,15	0,20	0,15	0,15
T_{32}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,15	0,20	0,15	0,15
T_{33}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,30	0,40	0,30	0,30
T_{34}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,11	0,13	0,11	0,11
T_{35}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,30	0,40	0,30	0,30
T_{36}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	0,50	0,48	0,62	0,48	0,48
T_{37}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	0,50	0,41	0,52	0,41	0,41
T_{38}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,59	0,80	0,59	0,59
T_{39}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,78	0,91	0,78	0,78
T_{40}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,56	-0,65	-0,56	-0,56
T_{41}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	-0,50	-0,41	-0,52	-0,41	-0,41
T_{42}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,30	-0,40	-0,30	-0,30

T_{43}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	-0,33	-0,15	-0,20	-0,15	-0,15
T_{44}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	-0,50	-0,04	-0,05	-0,04	-0,04
T_{45}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,11	0,13	0,11	0,11
T_{46}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,11	-0,13	-0,11	-0,11
T_{47}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	0,50	0,04	0,05	0,04	0,04
T_{48}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,15	0,20	0,15	0,15
T_{49}	4,00	0,44	0,55	0,22	0,47	0,33	0,67	0,33	0,30	0,40	0,30	0,30
T_{50}	6,00	0,67	0,63	0,33	0,58	0,50	0,82	0,50	0,41	0,52	0,41	0,41
T_{51}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,56	0,65	0,56	0,56
T_{52}	18,00	2,00	0,82	1,00	1,00	1,00	1,41	-1,00	-0,33	-0,33	-0,33	-0,33
T_{53}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	-0,67	-0,11	-0,13	-0,11	-0,11
T_{54}	10,00	1,11	0,73	0,56	0,75	0,67	1,05	0,67	0,11	0,13	0,11	0,11
T_{55}	18,00	2,00	0,82	1,00	1,00	1,00	1,41	1,00	0,33	0,33	0,33	0,33

Table 2a. *Association and cograduation indices computed on the Fréchet class* $[n_{1.}=n_{.1}=n_{3.}=n_{.3}=2,\ n_{2.}=n_{.2}=3]$

table	X^2	ϕ^2	P	NX^2	C	M_1	M_2	M	τ	Γ	d_{ba}	d_{ab}
T_1	14,00	2,00	0,82	1,00	1,00	0,65	1,41	-1,00	-1,00	-1,00	-1,00	-1,00
T_2	7,20	1,03	0,71	0,51	0,72	0,41	1,01	-0,63	-0,69	-0,85	-0,69	-0,69
T_3	10,10	1,44	0,77	0,72	0,85	0,53	1,20	-0,81	-0,38	-0,43	-0,38	-0,38
T_4	3,70	0,53	0,59	0,26	0,51	0,33	0,73	-0,50	-0,25	-0,33	-0,25	-0,25
T_5	7,00	1,00	0,71	0,50	0,71	0,49	1,00	0,75	0,00	0,00	0,00	0,00
T_6	4,30	0,61	0,62	0,31	0,55	0,33	0,78	0,50	0,00	0,00	0,00	0,00
T_7	3,70	0,53	0,59	0,26	0,51	0,33	0,73	0,50	0,25	0,33	0,25	0,25
T_8	10,10	1,44	0,77	0,72	0,85	0,53	1,20	0,81	0,38	0,43	0,38	0,38
T_9	7,20	1,03	0,71	0,51	0,72	0,41	1,01	0,63	0,69	0,85	0,69	0,69
T_{10}	14,00	2,00	0,82	1,00	1,00	0,65	1,41	1,00	1,00	1,00	1,00	1,00
T_{11}	7,20	1,03	0,71	0,51	0,72	0,41	1,01	-0,63	-0,69	-0,85	-0,69	-0,69
T_{12}	1,90	0,28	0,47	0,14	0,37	0,20	0,53	-0,31	-0,44	-0,64	-0,44	-0,44
T_{13}	3,70	0,53	0,59	0,26	0,51	0,33	0,73	-0,50	-0,25	-0,33	-0,25	-0,25
T_{14}	6,40	0,92	0,69	0,46	0,68	0,43	0,96	-0,66	-0,19	-0,23	-0,19	-0,19
T_{15}	4,30	0,61	0,62	0,31	0,55	0,33	0,78	0,50	0,00	0,00	0,00	0,00
T_{16}	4,30	0,61	0,62	0,31	0,55	0,33	0,78	0,50	0,00	0,00	0,00	0,00
T_{17}	3,70	0,53	0,59	0,26	0,51	0,33	0,73	0,50	0,25	0,33	0,25	0,25
T_{18}	6,40	0,92	0,69	0,46	0,68	0,43	0,96	0,66	0,19	0,23	0,19	0,19
T_{19}	1,90	0,28	0,47	0,14	0,37	0,20	0,53	0,31	0,44	0,64	0,44	0,44
T_{20}	7,20	1,03	0,71	0,51	0,72	0,41	1,01	0,63	0,69	0,85	0,69	0,69
T_{21}	10,10	1,44	0,77	0,72	0,85	0,53	1,20	-0,81	-0,38	-0,43	-0,38	-0,38
T_{22}	6,40	0,92	0,69	0,46	0,68	0,43	0,96	-0,66	-0,19	-0,23	-0,19	-0,19
T_{23}	4,30	0,61	0,62	0,31	0,55	0,33	0,78	0,50	0,00	0,00	0,00	0,00
T_{24}	6,40	0,92	0,69	0,46	0,68	0,43	0,96	0,66	0,19	0,23	0,19	0,19
T_{25}	10,10	1,44	0,77	0,72	0,85	0,53	1,20	0,81	0,38	0,43	0,38	0,38

Table 3a. *Association and cograduation indices computed on the Fréchet class* $[n_{1.}=n_{.1}=3,\ n_{2.}=n_{.2}=2,\ n_{3.}=n_{.3}=4]$

table	X^2	ϕ^2	P	NX^2	C	M_1	M_2	M	τ	Γ	d_{ba}	d_{ab}
T_1	9,00	1,00	0,71	0,50	0,71	0,46	1,00	-0,69	-0,85	-1,00	-0,85	-0,85
T_2	14,10	1,56	0,78	0,78	0,88	0,54	1,25	-0,81	-0,73	-0,83	-0,73	-0,73
T_3	7,90	0,88	0,68	0,44	0,66	0,43	0,94	-0,65	-0,65	-0,81	-0,65	-0,65
T_4	6,20	0,69	0,64	0,34	0,59	0,37	0,83	-0,56	-0,54	-0,70	-0,54	-0,54
T_5	9,60	1,06	0,72	0,53	0,73	0,46	1,03	-0,69	-0,35	-0,43	-0,35	-0,35
T_6	7,00	0,78	0,66	0,39	0,62	0,40	0,88	-0,59	-0,46	-0,60	-0,46	-0,46
T_7	3,10	0,34	0,50	0,17	0,41	0,23	0,58	-0,35	-0,31	-0,44	-0,31	-0,31
T_8	9,30	1,03	0,71	0,51	0,72	0,35	1,01	-0,52	-0,15	-0,20	-0,15	-0,15
T_9	3,40	0,38	0,53	0,19	0,44	0,25	0,62	-0,37	-0,15	-0,22	-0,15	-0,15
T_{10}	2,90	0,32	0,49	0,16	0,40	0,21	0,57	-0,31	0,00	0,00	0,00	0,00
T_{11}	7,80	0,86	0,68	0,43	0,66	0,37	0,93	0,56	0,15	0,20	0,15	0,15
T_{12}	6,10	0,67	0,63	0,34	0,58	0,36	0,82	0,54	0,04	0,05	0,04	0,04
T_{13}	3,30	0,36	0,52	0,18	0,42	0,22	0,60	0,33	0,23	0,33	0,23	0,23
T_{14}	10,60	1,17	0,73	0,59	0,77	0,42	1,08	0,63	0,42	0,52	0,42	0,42
T_{15}	5,10	0,57	0,60	0,29	0,53	0,31	0,75	0,46	0,35	0,47	0,35	0,35
T_{16}	5,70	0,63	0,62	0,32	0,56	0,36	0,79	0,54	0,54	0,70	0,54	0,54
T_{17}	11,30	1,25	0,75	0,63	0,79	0,49	1,12	0,74	0,54	0,64	0,54	0,54
T_{18}	9,60	1,06	0,72	0,53	0,73	0,44	1,03	0,67	0,77	0,91	0,77	0,77
T_{19}	18,00	2,00	0,82	1,00	1,00	0,64	1,41	0,96	1,00	1,00	1,00	1,00
T_{20}	7,90	0,88	0,68	0,44	0,66	0,43	0,94	-0,65	-0,65	-0,81	-0,65	-0,65
T_{21}	6,20	0,69	0,64	0,34	0,59	0,37	0,83	-0,56	-0,54	-0,70	-0,54	-0,54
T_{22}	2,80	0,31	0,49	0,16	0,40	0,25	0,56	-0,37	-0,39	-0,56	-0,39	-0,39
T_{23}	4,50	0,50	0,58	0,25	0,50	0,31	0,71	-0,46	-0,27	-0,37	-0,27	-0,27
T_{24}	5,60	0,63	0,62	0,31	0,56	0,30	0,79	-0,44	-0,12	-0,16	-0,12	-0,12
T_{25}	3,40	0,38	0,53	0,19	0,44	0,25	0,62	-0,37	-0,15	-0,22	-0,15	-0,15
T_{26}	2,90	0,32	0,49	0,16	0,40	0,21	0,57	-0,31	0,00	0,00	0,00	0,00
T_{27}	1,00	0,11	0,32	0,06	0,24	0,12	0,33	0,19	0,12	0,18	0,12	0,12
T_{28}	3,80	0,42	0,55	0,21	0,46	0,27	0,65	0,41	0,27	0,37	0,27	0,27
T_{29}	6,40	0,72	0,65	0,36	0,60	0,33	0,85	0,50	0,39	0,50	0,39	0,39
T_{30}	5,10	0,57	0,60	0,29	0,53	0,31	0,75	0,46	0,35	0,47	0,35	0,35
T_{31}	5,70	0,63	0,62	0,32	0,56	0,36	0,79	0,54	0,54	0,70	0,54	0,54
T_{32}	5,30	0,59	0,61	0,30	0,54	0,35	0,77	0,52	0,62	0,80	0,62	0,62
T_{33}	9,30	1,03	0,71	0,51	0,72	0,49	1,01	0,74	0,81	0,91	0,81	0,81
T_{34}	9,60	1,06	0,72	0,53	0,73	0,46	1,03	-0,69	-0,35	-0,43	-0,35	-0,35
T_{35}	5,60	0,63	0,62	0,31	0,56	0,30	0,79	-0,44	-0,12	-0,16	-0,12	-0,12
T_{36}	9,60	1,06	0,72	0,53	0,73	0,43	1,03	0,65	0,12	0,14	0,12	0,12
T_{37}	7,80	0,86	0,68	0,43	0,66	0,37	0,93	0,56	0,15	0,20	0,15	0,15
T_{38}	6,40	0,72	0,65	0,36	0,60	0,33	0,85	0,50	0,39	0,50	0,39	0,39
T_{39}	13,00	1,44	0,77	0,72	0,85	0,54	1,20	0,81	0,62	0,67	0,62	0,62

Table 4a. *Association and cograduation indices computed on the Fréchet class* $[n_{1.}=2, n_{2.}=4, n_{3.}=3, n_{.1}=5, n_{.2}=1, n_{.3}=3]$

table	X^2	ϕ^2	P	NX^2	C	M_1	M_2	M	τ	Γ	d_{ba}	d_{ab}
T_1	7,20	0,80	0,67	0,40	0,63	0,36	0,89	-0,54	-0,74	-1,00	-0,69	-0,78
T_2	7,20	0,80	0,67	0,40	0,63	0,35	0,89	-0,52	-0,49	-0,67	-0,46	-0,52
T_3	6,70	0,74	0,65	0,37	0,61	0,33	0,86	-0,50	-0,41	-0,56	-0,39	-0,44
T_4	8,80	0,98	0,70	0,49	0,70	0,42	0,99	-0,63	-0,16	-0,20	-0,15	-0,17
T_5	4,50	0,50	0,58	0,25	0,50	0,32	0,71	-0,48	-0,45	-0,65	-0,42	-0,48
T_6	3,60	0,40	0,53	0,20	0,45	0,22	0,63	-0,33	-0,25	-0,38	-0,23	-0,26
T_7	1,60	0,18	0,39	0,09	0,30	0,14	0,42	-0,20	-0,12	-0,20	-0,12	-0,13
T_8	2,80	0,31	0,49	0,16	0,39	0,20	0,56	0,30	0,08	0,13	0,08	0,09
T_9	4,30	0,48	0,57	0,24	0,49	0,30	0,69	0,44	0,20	0,29	0,19	0,22
T_{10}	7,60	0,84	0,68	0,42	0,65	0,42	0,92	0,63	0,41	0,50	0,39	0,44
T_{11}	9,00	1,00	0,71	0,50	0,71	0,49	1,00	-0,74	-0,16	-0,20	-0,15	-0,17
T_{12}	7,20	0,80	0,67	0,40	0,63	0,40	0,89	0,59	0,00	0,00	0,00	0,00
T_{13}	3,70	0,41	0,54	0,21	0,45	0,27	0,64	0,41	0,16	0,25	0,15	0,17
T_{14}	4,00	0,44	0,55	0,22	0,47	0,25	0,67	0,37	0,33	0,50	0,31	0,35
T_{15}	4,00	0,44	0,55	0,22	0,47	0,27	0,67	0,41	0,49	0,75	0,46	0,52
T_{16}	6,40	0,71	0,64	0,36	0,60	0,37	0,84	0,56	0,65	0,89	0,62	0,70
T_{17}	9,90	1,10	0,72	0,55	0,74	0,47	1,05	0,70	0,82	1,00	0,77	0,87
T_{18}	7,20	0,80	0,67	0,40	0,63	0,35	0,89	-0,52	-0,57	-0,78	-0,54	-0,61
T_{19}	5,20	0,58	0,61	0,29	0,54	0,25	0,76	-0,37	-0,29	-0,41	-0,27	-0,30
T_{20}	8,80	0,98	0,70	0,49	0,70	0,43	0,99	0,65	0,00	0,00	0,00	0,00
T_{21}	9,00	1,00	0,71	0,50	0,71	0,42	1,00	-0,63	-0,29	-0,37	-0,27	-0,30
T_{22}	4,60	0,51	0,58	0,26	0,51	0,20	0,71	0,30	0,00	0,00	0,00	0,00
T_{23}	5,80	0,64	0,63	0,32	0,57	0,28	0,80	0,43	0,29	0,41	0,27	0,30
T_{24}	12,60	1,40	0,76	0,70	0,84	0,51	1,18	0,76	0,57	0,64	0,54	0,61

APPENDIX B

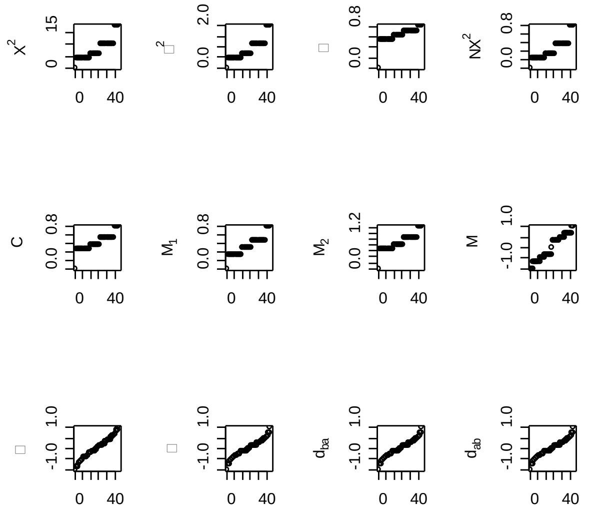

Figure 1b: *Association and cograduation indices computed on the Fréchet class $[n_{1.}=n_{2.}=n_{3.}=n_{.1}=n_{.2}=n_{.3}=3]$*

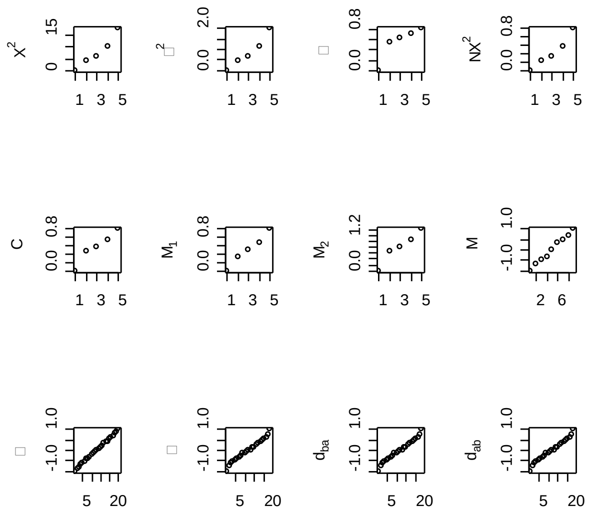

Figure 2b: *Association and cograduation indices, repeated once, computed on the Fréchet class $[n_{1.}=n_{2.}=n_{3.}=n_{.1}=n_{.2}=n_{.3}=3]$*

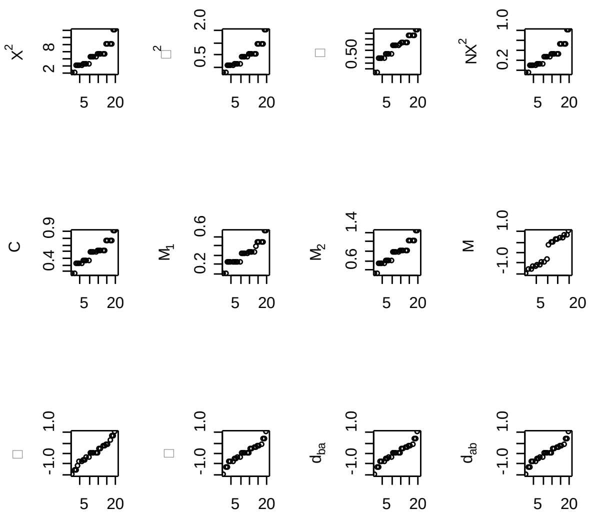

Figure 3b: *Association and cograduation indices computed on the Fréchet class* $[n_{1.}=n_{.1}=n_{3.}=n_{.3}=2,\ n_{2.}=n_{.2}=3]$

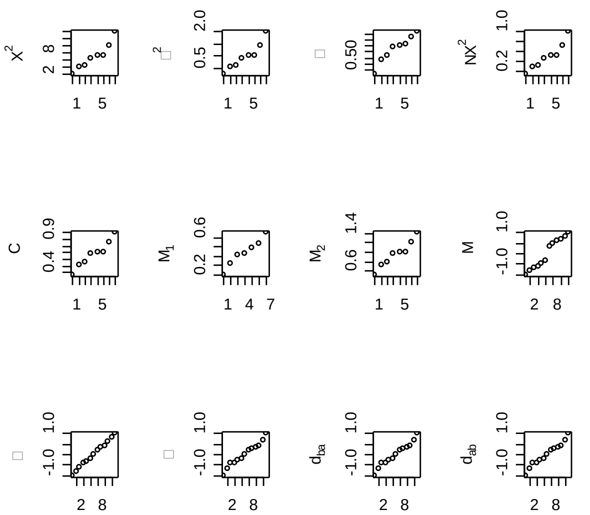

Figure 4b: *Association and cograduation indices, repeated once, computed on the Fréchet class* $[n_{1.}=n_{.1}=n_{3.}=n_{.3}=2,\ n_{2.}=n_{.2}=3]$

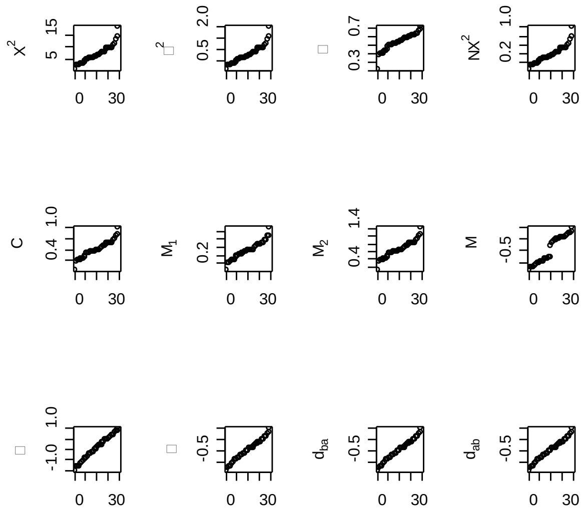

Figure 5b: *Association and cograduation indices computed on the Fréchet class* $[n_{1.}=n_{.1}=3, n_{2.}=n_{.2}=2, n_{3.}=n_{.3}=4]$

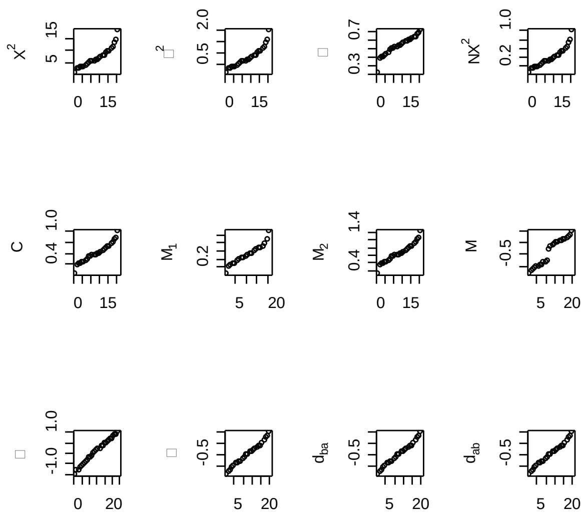

Figure 6b: *Association and cograduation indices, repeated once, computed on the Fréchet class* $[n_{1.}=n_{.1}=3, n_{2.}=n_{.2}=2, n_{3.}=n_{.3}=4]$

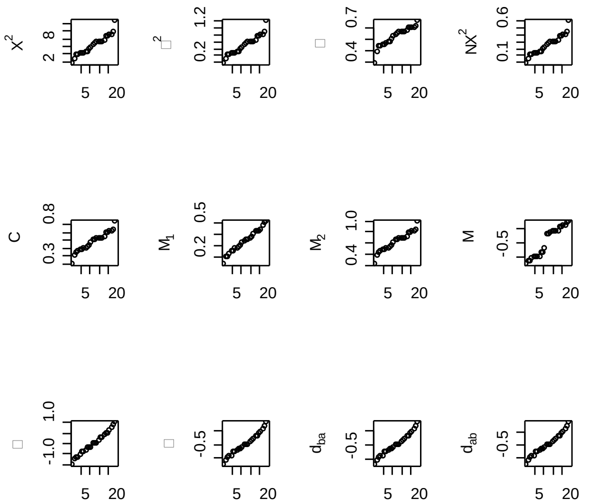

Figure 7b: *Association and cograduation indices computed on the Fréchet class* $[n_{1.}=2, n_{2.}=4, n_{3.}=3, n_{.1}=5, n_{.2}=1, n_{.3}=3]$

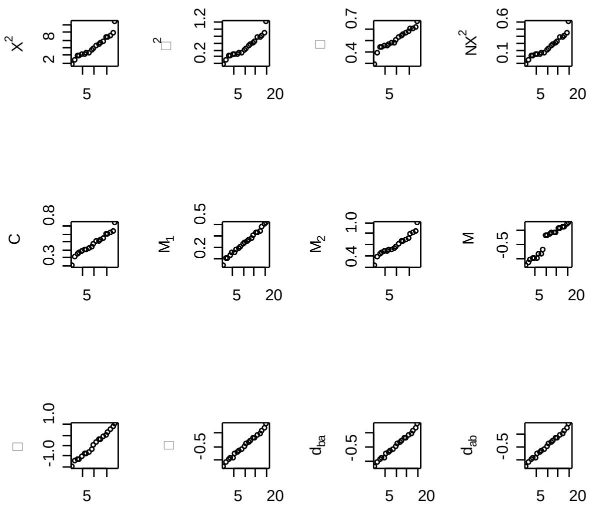

Figure 8b: *Association and cograduation indices, repeated once, computed on the Fréchet class* $[n_{1.}=2, n_{2.}=4, n_{3.}=3, n_{.1}=5, n_{.2}=1, n_{.3}=3]$