Coincidence Analysis
Principal components, correspondence, network analysis

M. Escobar M. (modesto@usal.es)

Universidad de Salamanca

10th October 2013
Stata is a cool statistics software which has been well accepted by the communities of

- Economist (and econometricians)
- Biomedical researchers

Recently, psychologists have also been adopting it, because of the implementation of SEM. However, social scientists have been less willing to adapt it. To increase the use of Stata by them, it should improve in key areas:

- Increase in table capacities and allow multiple response analysis.
- Offer more user-friendly graph options.
- Implement data-mining including CART, CHAID, and neural networks.
- Introduce social-network analysis
The aims of this presentation are

- To connect social networks with *coincidence analysis*, which is a statistical framework to study concurrence of events in large sets of scenarios.
- To announce an ado program that is able to perform this new statistical framework with Stata.
- We applied previous versions of this analysis to
  - Content analysis of media (Linz Archive)
  - Content analysis of textbooks (Representation of science in scholar texts)
  - Multiresponse analysis in questionnaires (Readers of newspapers)
- I will present today an analysis of characters in picture albums.
Coincidence analysis
Definition

- Coincidence analysis is a set of techniques whose object is to detect which characters, subjects, objects, attributes or events tend to appear at the same time in different delimited spaces.
- These delimited spaces are called $n$ scenarios, and are considered as units of analysis ($i$).
- In each scenario a number of $J$ incidences $X_j$ may occur (1) or may not (0) occur.
- We call incidence matrix $(\mathbf{X})$ an $n \times J$ matrix composed by 0 and 1, according to the incidence or not of every $X_j$.
- In order to make comparative analysis of coincidences, these scenarios may be classified in $H$ sets.
Two concrete incidences \((x_{ij} \text{ and } x_{ik})\) are defined as coincident if they occur in the same scenario:

\[
(x_{ij} = 1 \land x_{ik} = 1) \implies f_{ijk} = 1
\]

From the incidence matrix \((X)\), the coincidence matrix \((F)\) can be obtained by

\[
F = X'X
\]

where each element \(f_{jk}\) represents the number of scenarios where \(X_j\) and \(X_k\) are both 1, that is to say, the two characters, subjects, objects, attributes or events coincide.

As may be imagined, there is a special case \((f_{jj})\), which represents the number of incidences of \(X_j\) in the \(n\) scenarios.
By definition two incidences \((X_j \text{ and } X_k)\) are independent in a collection of \(n\) scenarios if and only if

\[
f_{jk} = \frac{f_{jj} f_{kk}}{n}
\]

where \(f_{jj}\) and \(f_{kk}\) are the number of scenarios amongst \(n\) where each of them can be observed.
Therefore, in a broad sense two incidences are coincident (or dependent) if they meet the following criterion:

\[ f_{jk} > \frac{f_{jj} f_{kk}}{n} \]

However, as long as it is usual to work with samples of scenarios, the Haberman residuals (\( r_{jk} \)), with normal distribution, may be used:

\[ r_{jk} = \frac{f_{jk} - \frac{f_{jj} f_{kk}}{n}}{\sqrt{\frac{1-f_{jj}}{n} \frac{1-f_{kk}}{n}}} \]
Adjacency

Definition

- Two incidents can be considered adjacent according to the following rule:

\[ p(r_{jk} \leq 0) < c \land j \neq k \]

- Therefore, a \( J \times J \) matrix \( A \) may be elaborated with 0 valued diagonal elements and 1 in the case where \( r_{jk} \) is significantly below the level \( c \). Other elements should also be 0.

- From \( A \) the \( J \times J \) distance matrix \( D \), with geodesics (shortest paths between elements), can be obtained.
Analysis

Five alternatives

- Principal component analysis (Pearson, 1901) with tetrachoric correlations (Everitt 1910).
- Correspondence analysis (Benzecri 1973), using matrix $X$ as input and obtaining only column coordinates (incidents).
- Network (Moreno 1934) and coincidence (Escobar 2009) analyses, based on Haberman residuals of $F$.
- Other analysis, based on Haberman residuals, can be used: multidimensional scaling (MDS) and cluster analysis.
Principal components
Definition

- "Principal components analysis (PCA) is a statistical technique used for data reduction. The leading eigenvectors from the eigen decomposition of the correlation or covariance matrix variables describe a series of uncorrelated linear combinations of the variables that contain most of the variance. In addition to data reduction, the eigenvectors from a PCA are often inspected to learn more about the underlying structure of the data" (Stata-MV, 2013: 573)

- In a coincidence analysis it can be used to obtain the factors that structure the appearance of events, subjects or objects.
The source of PCA is a correlation matrix. However, our material are $J \times J$ dichotomous variables. A way to adequate this nature of nominal variable to correlation measures is using tetrachoric correlations, assuming “a latent bivariate normal distribution ($X_j, X_k$) for each pair of variables” ... this correlation matrix “can be used to perform a factor analysis or a principal component analysis of binary variables using factormat or pcmat commands” (Stata-R, 2013: 2366).

So, in the coincidence analysis, a tetrachoric correlation matrix must be performed using the $X$ matrix, and its results ($r(Rho)$) should be used as input for the principal components analysis.
Correspondence analysis

Definition

- “Correspondence analysis (CA) offers a geometric representation of the rows and columns of a two-way frequency table that is helpful in understanding the similarities between the categories of variables and the association between the variables.” ... “In some respects, CA can be thought of as an analogue to principal components for nominal variables” (Stata-MV, 2013:32)
- With this analysis, similar categories (incidences according to its pattern of concurrence) can be represented together in an euclidean space of \( \min((R - 1)(C - 1)) \) dimensions.
We could use \( F \) as a contingency table introduced in the CA. However, that is not appropriate as long as the \( X_j \) are not independent as the categories of \( X \) and \( Y \) variables are in this sort of analysis.

So, instead of \( F \), it is possible to introduce the matrix \( X \), and consider only column coordinates for the analysis to perform a CA of indicator matrix. (See help mkmat)
Social network analysis

Definition

- Social network analysis (SNA) is “a set of methods for the analysis of social structures, methods that specifically allow an investigation of the relational aspects of these structures” (Scott, 2000: 38).
- Its main mathematical source is topology, especially graph theory.
- A graph $G$ is a collection of points or vertices $x_1, x_2, ..., x_n$ (denoted by the set $X$), and a collection of lines $a_1, a_2, ..., a_m$ (denoted by the set $A$) joining all or some of these points. The graph $G$ is then fully described and denoted by the doublet $(X, A)$. (Christofides, 1975)
This doublet \((X, A)\) can be represented just by a \(n \times n\) matrix \(A\) with whose elements \(a_{ij}\) represent if the point \(i\) is connected to the point \(j\).

There are, however, two other ways of storing graphs:

- Adjacency list where \(n\) rows are points, and columns are only neighbors points
- Edge list where rows are \(m\) connections with a first column indicating the origin point and a second indicating the target point. (Mihura, 2011:7)
Ostensibly, Stata doesn’t have tools for network analysis.

Although there are no tools for SNA in Stata, some advanced users have begun to write some routines. I wish to highlight the following works:

- Corten (2010) wrote a routine to visualize social networks [netplot]
- Mihura (2012) created routines (SGL) to calculate networks centrality measures, including two Stata commands [netsis and netsummarize]
- Recently, White (2013) presented a suite of Stata programs for network meta-analysis which includes the network _graphs of Anna Chaimani in the UK users group meeting. And Grund (2013) announced a presentation on plotting and analyzing social networks in the Nordic and Baltic Stata Users Group.
Multidimensional scaling (MDS) is “dimension-reduction and visualization technique. Dissimilarities (for instance, Euclidean distances between observations in a high-dimensional space are represented in a lower-dimensional space (typically two dimensions) so that the Euclidian distance in the lower-dimensional space approximates the dissimilarities in the higher-dimensional space” (Stata-MV, 2013: 450).

This is the classical method to represent nodes in network analysis (Corten, 2010). Its coordinates are used to position nodes in a bi-dimensional representation.
Cluster analysis

Definition

- Cluster analysis is aimed to determinate the natural groupings of observations (Stata-MV, 2013: 94).
- There are numerous procedures to aggregate cases: single, complete, average, median, Ward, ...
- In the coincidence analysis, clustering could be useful to rank the concurrence of events, using the Haberman residuals ($r_{jk}$) or the distance matrix ($D$) as inputs to cluster.
The aim is to analyze the set of characters in the collection.

The first step is to quantify the number of pictures of every character.

However, it is not only important how many times they appear, but also with whom.

These ideas are based on the interactionist theory of the self in G. H. Mead.

The pictures are going to be considered as scenarios.

The characters are going to be considered as incidences (variables). Do they or don’t they appear?
Turina’s Archive consists of 1,438 photographs from the family album, plus over 1,800 other photos stored in folders, and a collection of postcards acquired by Turina himself. The file set adds up to a total of 5,271 graphic materials.

The photos and Turina’s Archive documentation come from the Spanish Library of Contemporary Music and Theatre (Juan March Foundation in Madrid).
Joaquín Turina Pérez was born in Seville in December 1882. He studied music from an early age and completed his training in Madrid and Paris, where he met artists such as Isaac Albéniz and Manuel de Falla. He returned to Madrid at the beginning of World War I. In 1910 he was responsible for the management of the theater Eslava in Madrid and from 1919 he served as the conductor of the Teatro Real. In 1931 he became Professor of Composition at Conservatory of Madrid and in 1935 was appointed as a member of the Real Academia de Bellas Artes de San Fernando. He died in 1949 leaving behind musicals like Fantastic Dances and Fancy Clock. He also published academic works like A Treatise on Musical Composition (1946).
Pictures of Turina
Turina (1882-1949)
Turina’s nuclear family

Turina-Garzón’s family

Coincidence
M. Escobar

Introduction
Definitions
Coincidence
Dependence
Adjacency

Analysis
PCA
CA
SNA
MDS
Cluster

Application
Data
Albums
Matrix

coin

Results
Figures
Graphs

Remarks
Final
Turina’s nuclear family
Examples of codification

Coincidence
M. Escobar

Introduction
Definitions
Coincidence
Dependence
Adjacency

Analysis
PCA
CA
SNA
MDS
Cluster

Application
Data
Albums
Matrix

coin

Results
Figures
Graphs

Remarks
Final
The $\mathbf{X}$ matrix is constructed with $i$ rows representing pictures, and the $j$ columns representing people:

$$
\mathbf{X} = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 0
\end{bmatrix}
$$
coin
What is it?

- coin is an ado program in its development phase, which is capable of performing coincidence analysis.
- Its input is a dataset with scenarios as rows and events as columns.
- Its outputs are:
  - Different matrices (frequencies, percentages, residuals (3), edges and adjacencies)
  - Several network graphs (circle, mds, pca, ca) and dendrograms (single, average, waverage, complete, wards, median, centroid)
  - Measures of centrality
- Its syntax is simple, but flexible. Many options (output, bonferroni, p value, minimum, special event, graph control and options, ...).
### Frequencies of coincidences

#### coin Turina-Valle, f

<table>
<thead>
<tr>
<th></th>
<th>Turina</th>
<th>Garzon</th>
<th>Joaquin</th>
<th>Maria</th>
<th>Concha</th>
<th>Jose Luis</th>
<th>Obdulia</th>
<th>Valle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joaquin Turina</td>
<td>291</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia Garzón</td>
<td>42</td>
<td>216</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joaquín</td>
<td>42</td>
<td>71</td>
<td>262</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>María</td>
<td>25</td>
<td>62</td>
<td>124</td>
<td>222</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concha</td>
<td>20</td>
<td>39</td>
<td>68</td>
<td>100</td>
<td>134</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>José Luis</td>
<td>18</td>
<td>30</td>
<td>40</td>
<td>60</td>
<td>64</td>
<td>101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia</td>
<td>13</td>
<td>27</td>
<td>33</td>
<td>54</td>
<td>60</td>
<td>58</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>Josefa Valle</td>
<td>2</td>
<td>9</td>
<td>15</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Turina</td>
<td>Garzon</td>
<td>Joaquin</td>
<td>Maria</td>
<td>Concha</td>
<td>Jose Luis</td>
<td>Obdulia</td>
<td>Valle</td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>Joaquin Turina</td>
<td>30.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia Garzón</td>
<td>-4.5</td>
<td>30.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joaquín</td>
<td>-6.5</td>
<td>1.6</td>
<td>30.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>María</td>
<td>-7.6</td>
<td>1.7</td>
<td>10.3</td>
<td>30.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concha</td>
<td>-4.5</td>
<td>1.6</td>
<td>6.1</td>
<td>14.7</td>
<td>30.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>José Luis</td>
<td>-3.2</td>
<td>1.5</td>
<td>2.6</td>
<td>8.7</td>
<td>14.7</td>
<td>30.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia</td>
<td>-3.5</td>
<td>1.8</td>
<td>2.1</td>
<td>8.8</td>
<td>15.2</td>
<td>17.5</td>
<td>30.2</td>
<td></td>
</tr>
<tr>
<td>Josefa Valle</td>
<td>-2.2</td>
<td>2.1</td>
<td>4.4</td>
<td>2.5</td>
<td>0</td>
<td>-1.6</td>
<td>-1.5</td>
<td>30.2</td>
</tr>
</tbody>
</table>
### List of significant Haberman residuals

<table>
<thead>
<tr>
<th>Haberman</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.5</td>
<td>José Luis &lt;-&gt; Obdulia</td>
</tr>
<tr>
<td>15.2</td>
<td>Concha &lt;-&gt; Obdulia</td>
</tr>
<tr>
<td>14.7</td>
<td>María &lt;-&gt; Concha</td>
</tr>
<tr>
<td>14.7</td>
<td>Concha &lt;-&gt; José Luis</td>
</tr>
<tr>
<td>10.3</td>
<td>Joaquín &lt;-&gt; María</td>
</tr>
<tr>
<td>8.8</td>
<td>María &lt;-&gt; Obdulia</td>
</tr>
<tr>
<td>8.7</td>
<td>María &lt;-&gt; José Luis</td>
</tr>
<tr>
<td>6.1</td>
<td>Joaquín &lt;-&gt; Concha</td>
</tr>
<tr>
<td>4.4</td>
<td>Joaquín &lt;-&gt; Josefa Valle</td>
</tr>
<tr>
<td>2.6</td>
<td>Joaquín &lt;-&gt; José Luis</td>
</tr>
<tr>
<td>2.5</td>
<td>María &lt;-&gt; Josefa Valle</td>
</tr>
<tr>
<td>2.1</td>
<td>Obdulia Garzón &lt;-&gt; Josefa Valle</td>
</tr>
<tr>
<td>2.1</td>
<td>Joaquín &lt;-&gt; Obdulia</td>
</tr>
<tr>
<td>1.8</td>
<td>Obdulia Garzón &lt;-&gt; Obdulia</td>
</tr>
<tr>
<td>1.7</td>
<td>Obdulia Garzón &lt;-&gt; María</td>
</tr>
</tbody>
</table>
### Adjacencies matrix

coin Turina-Valle, a

<table>
<thead>
<tr>
<th></th>
<th>Turina</th>
<th>Garzon</th>
<th>Joaquin</th>
<th>María</th>
<th>Concha</th>
<th>JoseLuis</th>
<th>Obdulia</th>
<th>Valle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joaquin Turina</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia Garzón</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joaquín</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>María</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concha</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>José Luis</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Obdulía</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Josefa Valle</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Tetrachoric matrix

<table>
<thead>
<tr>
<th>Coincidence</th>
<th>M.Escobar</th>
<th>Introduction</th>
<th>Definitions</th>
<th>Coincidence</th>
<th>Dependence</th>
<th>Adjacency</th>
<th>Analysis</th>
<th>PCA</th>
<th>CA</th>
<th>SNA</th>
<th>MDS</th>
<th>Cluster</th>
<th>Application</th>
<th>Data</th>
<th>Albums</th>
<th>Matrix</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Turina</th>
<th>Garzon</th>
<th>Joaquin</th>
<th>Maria</th>
<th>Concha</th>
<th>Jose Luis</th>
<th>Obdulia</th>
<th>Valle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joaquín Turina</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia Garzón</td>
<td>-0.28</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joaquín</td>
<td>-0.39</td>
<td>0.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>María</td>
<td>-0.49</td>
<td>0.12</td>
<td>0.55</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concha</td>
<td>-0.34</td>
<td>0.12</td>
<td>0.38</td>
<td>0.76</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>José Luis</td>
<td>-0.21</td>
<td>0.08</td>
<td>0.14</td>
<td>0.46</td>
<td>0.65</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia</td>
<td>-0.25</td>
<td>0.10</td>
<td>0.12</td>
<td>0.47</td>
<td>0.67</td>
<td>0.87</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Josefa Valle</td>
<td>-0.25</td>
<td>0.17</td>
<td>0.37</td>
<td>0.17</td>
<td>-0.10</td>
<td>-0.71</td>
<td>-0.70</td>
<td>1.00</td>
</tr>
</tbody>
</table>
### Distances matrix

coin Turina-Valle, d

<table>
<thead>
<tr>
<th></th>
<th>Turina</th>
<th>Garzon</th>
<th>Joaquin</th>
<th>María</th>
<th>Concha</th>
<th>JoseLuis</th>
<th>Obdulia</th>
<th>Valle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joaquín Turina</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia Garzón</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joaquín</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>María</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concha</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>José Luis</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obdulia</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Josefa Valle</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
### Centrality measures

**coin Turina-Valle, c alpha(.10)**

<table>
<thead>
<tr>
<th></th>
<th>Degree</th>
<th>Closeness</th>
<th>Betweeness</th>
<th>Eigenvector</th>
<th>Power(0.1)</th>
<th>Power(-0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joaquín Turina</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Obdulia Garzón</td>
<td>0.38</td>
<td>0.67</td>
<td>0.02</td>
<td>0.26</td>
<td>1.55</td>
<td>0.80</td>
</tr>
<tr>
<td>Joaquín</td>
<td>0.63</td>
<td>0.86</td>
<td>0.09</td>
<td>0.42</td>
<td>1.90</td>
<td>0.65</td>
</tr>
<tr>
<td>María</td>
<td>0.75</td>
<td>1.00</td>
<td>0.18</td>
<td>0.47</td>
<td>2.04</td>
<td>0.56</td>
</tr>
<tr>
<td>Concha</td>
<td>0.50</td>
<td>0.75</td>
<td>0.00</td>
<td>0.38</td>
<td>1.76</td>
<td>0.74</td>
</tr>
<tr>
<td>José Luis</td>
<td>0.50</td>
<td>0.75</td>
<td>0.00</td>
<td>0.38</td>
<td>1.76</td>
<td>0.74</td>
</tr>
<tr>
<td>Obdulia</td>
<td>0.63</td>
<td>0.86</td>
<td>0.09</td>
<td>0.42</td>
<td>1.90</td>
<td>0.65</td>
</tr>
<tr>
<td>Josefa Valle</td>
<td>0.38</td>
<td>0.67</td>
<td>0.02</td>
<td>0.26</td>
<td>1.55</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Graph comparisons
Principal components, Correspondence, MDS and circle.

Turina (PCA)

Turina (CA)

Turina (MDS)

Turina (Circle)
Graph comparisons
Single, complete, average, Ward dendrograms

Clusters (single)

Clusters (complete)

Clusters (average)

Clusters (wards)

Distance

Clusters (single)
Clusters (complete)
Clusters (average)
Clusters (wards)
Characters in Turina’s Album
Family and colleagues
Remarks
About coincidence analysis

I’ve proposed a manner of analyzing coincidences using the existing statistical tools. I think that the novelty of coincidence analysis is putting together different techniques in order to represent reality. This may also be useful in comparing different kinds of analysis with dichotomous variables. We think that the above approach could be extensively used with the aid of the coin program.
Thanks for your attention and comments!