Kamer Ainur AIVAZ
Marioara MIREA
Faculty of Economics Sciences,
"Ovidius" University of Constanta, Romania

Case studies

A STUDY REGARDING THE MAJOR GEOGRAPHICAL DISPARITIES IN THE ROMANIAN NUMBER OF SCHOOLS

Keywords
Disparities
The factorial analysis of the Correspondents (FAC),
The contribution of the axis to the inertia of a point
The contribution of the point to the inertia of an axis

JEL Classification
I21,C10, C40, C80

Abstract

According to a new report of the European Commission, the European area in which one lives can have a considerable influence on the education and on the life perspectives.
Starting with the results that are included in this report, the current paper aims to move the analysis to the Romanian space, studying the discrepancies between macro regions (NW, Center, NE, SE, South-Muntenia, Bucharest – Ilfov, SW, West) to the principal educational indicators. The research is going to use the FAC method (The Factorial Analysis of the Correspondents) which is a descriptive method of multidimensional analysis of the data based on the calculus of the χ^2 distances between the points. Through this method we are going to test not only the independence of the variables, but their association as well.

Introduction

The report, which is named "Mind the Gap – education inequality across EU regions", shows significant disparities of the educational opportunities and of the related results between the member states and within them. It seems that there is a big difference between the north and the south regarding the achieved educational level - the highest shares of the persons with a low qualification level that have passed the secondary school or lower can be mainly found in the regions from the Southern Europe, especially in Portugal and Spain. By contrast, the regions with the lowest shared of the persons with a low qualification level are mainly found in the United Kingdom, Belgium, and Sweden and in the Netherlands. The geographical inequalities in education persist despite the arrangements undertaken by the member states in order to promote the equality in education and in the professional training. The report - the first one of this kind – urges the member states of the European Union to undertake more efforts in order to reduce these inequalities.

"All the European citizens should benefit of education and high quality training, regardless of the place in which they live. It is the moment to respect the assumed commitments. The fight against the geographical inequalities in regards to education is a prior condition for the social cohesion and for the balanced regional development. The European structural funds can and should be used in order to aid the elimination of the inequalities" declared Androulla Vassiliou, the European commissioner for education, culture, multilingualism and youth.

Starting with the results that are included in this report, the current paper aims to move the analysis to the Romanian space, studying the discrepancies between macro regions (NW, Center, NE, SE, South-Muntenia, Bucharest – Ilfov, SW, West) to one of the principal educational indicators, number of schools. The research is going to use the FAC method (The Factorial Analysis of the Correspondents) which is a descriptive method of multidimensional analysis of the data based on the calculus of the χ^2 distancesbetween the points. Through this method weare going to test not only the independence of the variables, but their association as well.(Kachigan, 1982)

Methodology

The factorial correspondence analysis (FCA) is a descriptive method of multidimensional data analysis that applies to the study of the connections between two non-numeric variables. In order to calculate the distance between two points, the χ2 distance is used in the FCA.

The FCA method was developed by Paul Benzecri (Benzecri, 1992), who highlighted the algebraic properties of this method and emphasized both the use of the FCA method for testing the independence of variables, but also for the decrease of the associations between them.

Based on a large table, the FCA discloses a system of factorial axes which synthesizes the initial information in a graphical form easy to interpret. The factorial axes are "ranked" in a descending order, according to their importance in explaining the total variance of the cloud of points.

Basically, the FCA seeks to find the unrelated axes on which to project the point line cloud and the point column cloud. For this purpose, there are calculated the eigenvalues (a_k) of the matrix of inertia and the coordinates of the eigenvectors associated to each a_k value.

$$\sum_{k=1}^{k} a_k = \frac{i^2}{n}$$

The calculation of the line points and column points coordinates on the factorial axes are performed using the coordinates of the eigenvectors associated with the eigenvalues of the matrix of inertia. In order to assess the distance between two points, the χ2 distance is used. (Greene, 2003)

After processing the data in SPSS, using

the factorial correspondence analysis, we obtainthe statistical indicators calculated for the line points and for the column points, and the graphical representation of these points in the factorial axes system.
The presentation of the FCA methodology will be carried out considering the group of units in a contingency table (Table 1) that reveals the number of school units by level of education and region.

Data and results
The analysed indicators were:
• Number of schools by levels of education and regions;
• School population by levels and regions.
(www.tempoonline.com)
After processing the data in SPSS, using the factorial correspondence analysis, we obtain the statistical indicators calculated for the line points, and for the column points, and the graphical representation of these points in the factorial axis system. (Baltagi, 2008)
The presentation of the FCA methodology will be carried out considering the group of units in a contingency table that shows the level of the two indicators by levels of education and regions.
The table of row profiles (output Row Profiles) and the table column profiles (output Column Profiles)
The table of row profiles (Table 2) includes, as terms, the $n_{j/i}$ partial relative frequencies calculated as the ratio of the n_{ij} absolute partial frequencies and the $n_{i.}$ marginal absolute frequencies, for each value of the variable X: $f_{i/j} = \frac{n_{ij}}{n_{i.}}$
A row profile shows, for each x_i value, the share of the statistical units per the y_i values of the variable Y (Kachigan, 1982). For our data, the values in this row profile table shows the distribution of the number of schools by levels of education and regions.
Thus, in Romania, we notice the following situation on the number of schools:
• Of the total number of schools, 55.2% are secondary education schools, 23% are high schools, 19.5% are preschool education institutions, 1.5 % are higher education institutions, 0.7% are primary education institutions;
• Bucharest- Ilfov has the greatest number of preschool institutions, well above the average of 19.5%, and the greatest number of higher education institutions, i.e. 5.8 %, compared to an average of 1.5%;
• High schools and secondary education institutions are quite homogeneously represented in the Romanian regions.
A statistical association between two regions (two row profiles) shows a similar distribution of the number of schools by levels of education.
A column profile shows (Table 3), for each y_i value, the share of the statistical units on the x_i values of the variable X (the column profiles for the data presented in Table 1).The values in this table show the distribution of schools by regions and by levels of education:
• Bucharest- Ilfov has the greatest share in the preschool and higher education category, i.e. 16.3% and 32.4%, followed by the North -West;
• the West Region holds the first place for primary education, while the North-east holds the first place for secondary education and the North-west - for high schools;
• All these shares highlight a quite important gap for this indicator.
A statistical association between two column profiles (types of educational level) shows a similar distribution by regions.
The value of the χ2 statistics
In the Summary output, in the Chi- Square column, there is shown the calculated value of the χ2 test statistics used in testing the hypothesis of the independence between variables (Maddala, 2001).
Testing the hypothesis of the independence between variables involves the formulation of the following hypotheses:
• The null hypothesis H_0: the hypothesis of the independence between the statistical variables (there are no connection between

the statistical variables);
• The alternative hypothesis H_1: the hypothesis of dependence (there are connections between the statistical variables). (Verbeek, 2005)
In the output above, the calculated value of the test statistic is Sig.=0,000<0,05, which indicates that the hypothesis H_0 is rejected. We can guarantee by a 95% probability that there are connections between the variables considered.The description of these connections is achieved by interpreting the FCA results for the row profiles and for the column profiles.

The eigenvalues and the variance explained by each factorial axis

In the FCA, the maximum number of factorial axes is min (m-1; p-1)=min(7, 4)=4. As in principal component analysis, the largest eigenvalue shows the variance of the first factorial axis and the sum of the eigenvalues measures the total inertia of the point cloud. For the output shown in the Summary Table, the largest eigenvalue is $^{a}_{1}$=0,049. The inertia (variance) explained by each factorial axis is shown in the column *Proportion of Inertia*. The first factorial axis explains 77.2% of the total variance. The choice of the number of factorial axes is made, in the FCA, according the Benzecri criterion: there are chosen those factorial axes which explain at least 70% of the total variance. In our case, therefore, we need a single factorial axis.

The results of column points

For each category of the variable Y, there are calculated the coordinates on the factorial axes, the contributions of points to the inertia of an axis and the contributions of axes to the inertia of a point.
• The coordinates on the factorial axes (column *Score in Dimension*)
The coordinates of the column points on the factorial axes show their position in the space represented by the axes. A positive coordinate on the first factorial axis indicates that the column point will be projected on the right side and a negative coordinate shows its projection on the left side of the axis (Everitt, Dunn, 2001).
The values in Table 5 show that the point "South East" has a positive coordinate on the first factorial axis (0.099) and a negative coordinate on the second factorial axis
(-0.042); the point "Northeast" has a positive coordinate on the first factorial axis (0.319) and also a positive coordinate on the second factorial axis (0.048); the point "Northwest" has a negative coordinate on the first factorial axis (-0.104) and also a negative coordinate on the second factorial axis
(-0.122); the point "Bucharest- Ilfov" has a negative coordinate on the first factorial axis (-1,226) and a positive coordinate on the second factorial axis (0.435).
• The contribution of points to the inertia of an axis (column *Contribution of point to Inertia of Dimension*)
These values show the contribution of the category (levels of education) to the dispersion of the factorial axis. The points with higher contributions on an axis are those points that contribute to the formation of the axis. These are called explanatory points for the formation of the respective axis. The first factorial axis is explained at a rate of 58.4% by the point "Bucharest- Ilfov", at a rate of 14.3% by the point "South – Muntenia" etc.
• The contribution of the factorial axis to the inertia of a point (column *Contribution of Dimension to Inertia of Point*)
This value expresses the contribution of a factorial axis in explaining the dispersion of a point. There are highlighted the following points explained by the factorial axes. For the values shown in Table 5, we note that the inertia of the point "Bucharest – Ilfov" is explained at a rate of 93.5% by the first factorial axis and at a rate of 5.7 % by the second factorial axis.
Figure 1 suggestively shows the associations and the differences between the regions and the level of education.
Thus:

• The Bucharest-Ilfov region is situated at a large distance from the other regions of the country in terms of the number of schools;
• Large discrepancies exist between the primary education and the other levels of education;
• Visible similarities can be seen between the regions of South-Muntenia and Northeast and between the North-west and South-west regions, their levels being close to the average at the country level.

Conclusions

The challenges facing the European Union and, implicitly, Romania are now numerous, interdependent and complex. The social consequences of the global financial and economic crisis are felt in all Member States. In this context, the education and training systems must change their priorities so as to ensure that all European citizens have the knowledge, skills and competencies to meet the challenges and demands of jobs and modern life.

In 2010, the Member States and the European Commission agreed to include education and training as the key elements of the EU 2020 Strategy for a smart and sustainable economic growth, favorable to the inclusion in the current decade. This is supported by the Strategic Framework for education and training (*Education and Training 2020*) and by the four of its long-term objectives. The Strategic Framework is the basis of the European cooperation in education and training, thus playing an important role for the wider objectives of Europe 2020.

The existence of effective and harmonized indicators is essential to monitor the progress made in achieving these objectives.

In the context of quality assurance, the public access to school education units is an important factor for achieving the strategic objective of increasing the quality and efficiency of education. A central point of interest of this evaluation may be represented by the overall education system, the schools representing an extremely important variable.

Moreover, in most countries, the investments in education have remained largely unchanged over the last decade, until 2008, before the economic crisis. In response to the crisis, some governments have taken specific measures to ensure that the existing funding levels are not modified in order to ensure the continuous operation of the system and to protect the reforms implemented over the last decade.

References:

[1] Baltagi,B.H.(2008). *Econometrics*, Springer.
[2] Benzecri,J.P., (1992).*Correspondence Analysis Handbook*, New-York.
[3] Everitt,B., Dunn, G., (2001) *Applied Multivariate Data Analysis*, Arnold, London.
[4] Greene,W.H., (2003). *Econometric Analysis*, Prentice Hall.
[5] Kachigan,S.,(1982). *Multivariate statistical analysis*, Radius Press, New York
[6] Maddala,G.S.,(2001). *Introduction to Econometrics*, Wiley.
[7]Verbeek,M.,(2005) *A Guide to Modern Econometrics*, Wiley.
[8]www.tempoonline.com

Table 1.
The number of schools by levels of education and regions

Correspondence Table						
Regions	Level of education					
	pre school	primary (elementary) education	secondary education (middle school)	high school	Higher education	Active Margin
North-West	199	9	508	263	16	995
Centre	239	2	530	213	13	997
North-East	156	3	694	260	15	1128
South-East	173	7	515	209	7	911
South Muntenia	145	1	673	210	4	1033
Bucharest-Ilfov	223	5	206	134	35	603
South-West Oltenia	96	3	468	160	4	731
West	136	22	277	166	14	615
Active Margin	1367	52	3871	1615	108	7013

Source: Table processed in the SPSS programme, based on the data from Statistical Yearbook of Romania, NIS, Bucharest (*Anuarul statistic al României,INS, Bucureşti*), 2013

Table 2.
The table of row profiles

Row Profiles						
Regions	Level of education					
	Pre school	primary (elementary) education	secondary education (middle school)	high school	Higher education	Active Margin
North-West	.200	.009	.511	.264	.016	1.000
Centre	.240	.002	.532	.214	.013	1.000
North-East	.138	.003	.615	.230	.013	1.000
South-East	.190	.008	.565	.229	.008	1.000
South Muntenia	.140	.001	.652	.203	.004	1.000
Bucharest-Ilfov	.370	.008	.342	.222	.058	1.000
South-West Oltenia	.131	.004	.640	.219	.005	1.000
West	.221	.036	.450	.270	.023	1.000
Mass	.195	.007	.552	.230	.015	

Source: Table processed in the SPSS programme, based on the data from Statistical Yearbook of Romania, NIS, Bucharest (*Anuarul statistic al României,INS, Bucureşti*), 2013

Table 3.
The table of column profiles

Column Profiles						
Regions	Level of education					
	pre school	primary (elementary) education	secondary education (middle school)	high school	Higher education	Mass
North-West	.146	.173	.131	.163	.148	.142
Centre	.175	.038	.137	.132	.120	.142
North-East	.114	.058	.179	.161	.139	.161
South-East	.127	.135	.133	.129	.065	.130
South Muntenia	.106	.019	.174	.130	.037	.147
Bucharest-Ilfov	.163	.096	.053	.083	.324	.086
South-West Oltenia	.070	.058	.121	.099	.037	.104
West	.099	.423	.072	.103	.130	.088
Active Margin	1.000	1.000	1.000	1.000	1.000	

Source: Table processed in the SPSS programme, based on the data from Statistical Yearbook of Romania, NIS, Bucharest (*Anuarul statistic al României,INS, Bucureşti*), 2013

Table4.
The calculated value of the χ2 statistics, the eigenvalues and the inertia explained by each factorial axis

Summary								
Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia		Confidence Singular Value	
					Accounted for	Cumulative	Standard Deviation	Correlation 2
1	.221	.049			.772	.772	.013	.100
2	.107	.012			.182	.954	.018	
3	.044	.002			.030	.984		
4	.032	.001			.016	1.000		
Total		.063	444.624	.000[a]	1.000	1.000		
a. 28 degrees of freedom								

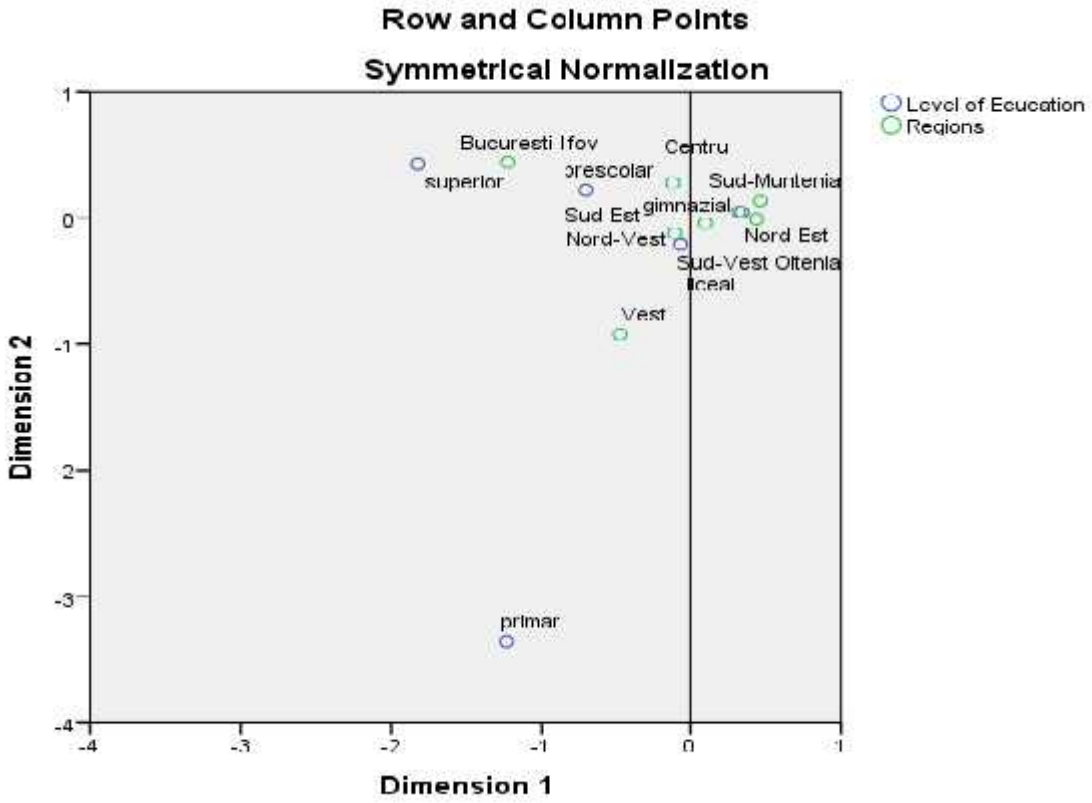
Source: Table processed in the SPSS programme, based on the data from Statistical Yearbook of Romania, NIS, Bucharest (*Anuarul statistic al României,INS, Bucureşti*), 2013

Table5.
The results on column points (training levels)

Overview Row Points[a]									
Regions	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	Total
North-West	.142	-.104	-.122	.001	.007	.020	.279	.185	.464
Centre	.142	-.117	.278	.002	.009	.102	.184	.500	.683
North-East	.161	.319	.048	.004	.074	.003	.831	.009	.840
South-East	.130	.099	-.042	.001	.006	.002	.497	.044	.541
South-Muntenia	.147	.463	.136	.007	.143	.026	.938	.039	.977
Bucharest-Ilfov	.086	-1.226	.435	.031	.584	.152	.935	.057	.992
South-West Oltenia	.104	.439	-.010	.005	.091	.000	.984	.000	.984
West	.088	-.469	-.922	.012	.087	.695	.345	.647	.992
Active Total	1.000			.063	1.000	1.000			
a. Symmetrical normalization									

Source: Table processed in the SPSS programme, based on the data from Statistical Yearbook of Romania, NIS, Bucharest (*Anuarul statistic al României,INS, Bucureşti*), 2013

Figure 1. The representation of the categories of statistical variables in the system of the first two factorial axes



Source: The author's processing of the data provided by NIS