MULTIPLE CORRELATIONS ANALYSIS WITHIN TEXTILE INDUSTRY FIRMS FROM ROMANIA

Podasca Raluca
Ph.D student at the Bucharest University of Economic Studies,
Faculty of Management, Bucharest, Romania
Email: raluca.podasca@yahoo.com

ABSTRACT

In order to study the connections between economic phenomena and processes is necessary first to know their objective form of manifestation. Economic and social phenomena are not usually uniquely determined, being the result of the influence of many causes. In this connection system, the dependency relationships don't have all the same importance, the action of some of them compensating each other. In the statistical analysis of the relationship of dependence between phenomena, it is important the issue of measuring the relationship between two or more characteristics used in studies of economic and social mass phenomena. Thus, it should be determined whether there is a relation of dependence between the characteristics, and if so, the dependence should be expressed through a correlation indicator which may express the degree in which one phenomena contributes to other phenomena.

Key words*: correlation coefficient, multiple correlations, textile industry*

1. Introduction

A large number of main and secondary factors, of essential and nonessential factors, act on socio- economic phenomena, which are found in connectivity. Statistics, using a variety procedures and methods available can study the concrete manifestation of these connections, can express them quantitatively and measure the intensity with which they occur. Starting from the fact that statistics study mass phenomena within which act statistical laws, whose main feature is the fact that they manifest as a tendency, it is required the interpretation of causal relationships still as a tendency. Meanwhile, statistical practice proves that in the process of producing mass socio- economic phenomena, not all phenomena manifest with the same intensity and in the same sense. If the phenomena studied are complex, the number of factors is larger and causal relationships are more difficult to identify and measure. The complex interdependency relations of the economy frequently make some qualitative factors be connected, which does not always allow a direct measurement. In addition, within the social phenomena, besides the objective factors, essential and nonessential, is added the influence of subjective factors that can change the tendencies of interdependence relations expression.

Statistics has many methods to study the dependencies between two or more variables. Among them are the regression and the correlation analysis. These methods study the dependence between a resultant variable and one or more independent variables. The resultant feature is called dependent feature (Y), endogenous or effect feature. The independent variables are called exogenous or causal features. Regression indicates the analytical form in which a variable depends on another variable and the correlation indicates the degree of the dependence between the variables.

2. The classification of statistic correlations

The variety of interdependency relations encountered in economic phenomena requires identification, selection and prioritization of influence factors. Practically, identifying connections between phenomena can be achieved only after a qualitative multilateral analysis. In this analysis, statistics proves to be an interdisciplinary science that uses its models in relation to the particularities of the phenomena which form the object of study consistent with the results obtained by specialized sciences studying the same field. With this analysis there are identified the factors and their degree of importance.

The forms of interdependences are varied and sometimes difficult to explain. According to Isaic-Maniu (2004) is necessary to classify statistic correlations according to certain criteria.

a) By the number of independent features studied:
   - Single correlation when studying the correlation of a dependent characteristic (y) and an independent one (x), for example the relationship between the number of sales (x) and the volume of sales of goods (y);
   - Multiple correlation when studying the correlation of a dependent characteristic (y) and two or more independent features (x_1, x_2, ..., x_n) such dependence of the volume of sales of goods (y), the number of sales (x_1) and commercial area (x_2);
b) By the direction of the link:
   - Direct correlation when both characteristics change accordingly;
   - Reverse correlation when dependent feature change opposite the independent one;
c) By the analytical expression of the connection:
   - Linear correlation which can be expressed by the function of grade 1;
   - Nonlinear (curved) correlation which can be expressed using nonlinear functions (parabola, hyperbola, exponential function, etc.);

d) By the time when connection occur:
   - Concurrent of synchronous connections when the outcome variable and the causal variable change at the same time;
   - Asynchronous connections when the variation of the resulting feature occurs after a period of time after the variation of the causal feature;

e) By the approach method:
   - Simple methods (elementary)
   - Analytical methods.

The elementary methods to study the statistic connections are:

- the independent parallel series method;
- the grouping method;
- the correlation table method;
- the graphic method.

The independent parallel series method involves ordering the values of the independent variable X (ascending or descending) seeking the way in which dependent variable Y is arranged. If the X and Y features are both ordered approximately ascending then we appreciate that the two variables are directly linked. If the X feature is ordered ascending and the Y feature is ordered approximately descending then we appreciate that the two variables are liked reverse. If Y feature hasn't an ordering trend (ascending or descending) we can appreciate that the two variables are not connected.

The grouping method involves the distribution of units in homogeneous groups according to an independent feature. There are centralized the numerical data of resultant characteristic for each group, calculating averages and relative indicators. By comparing the independent characteristic's variation with the indicators calculated for the resulting feature it can be appreciated the existence and form of the link between the two variables.

The correlation table is double entry table that shows a group of units of a collectivity based on two characteristics: one dependent and one independent. It is mainly used in a large number of observations. Correlation table is an important tool for the interpretation of the relationship between phenomena, but it cannot determine the form or the intensity of the connection.

Table no. 1. The correlation table

X \ Y	y_1	y_2	…	y_m	Sum
x_1	n_{11}	n_{12}	…	n_{1m}	$n_{1.}$
x_2	n_{21}	n_{22}	…	n_{2m}	$n_{2.}$
…	…	….	…	…	…
x_k	n_{k1}	n_{k2}	…	n_{km}	$n_{k.}$
Sum	$n_{.1}$	$n_{.2}$	…	$n_{.m}$	n

To obtain the correlation graphic, it is used the rectangular axis system, namely quadrant I, wherein the values of X and Y are positive. The independent variable X is represented on the Ox axis and the dependent variable Y is represented on the Oy axis. If we have a large number of statistical observations, we obtain a large number of points on the graph. Therefore, the correlation graph is also called the graph of point clouds. In this way we can identify the existence and direction of the link on one side and the form of the link on the other side.

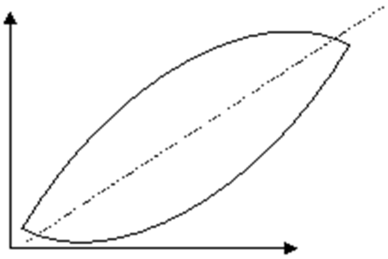

Fig. no.1. Direct linear correlation

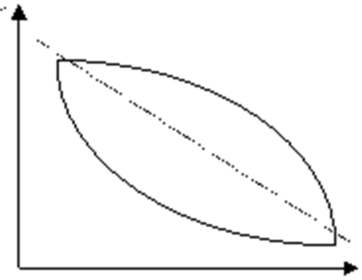

Fig. no.2. Reverse linear correlation

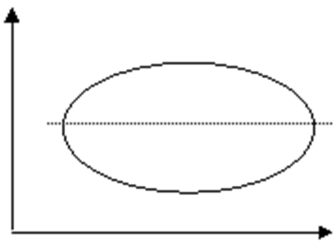

Fig. no.3. The lack of connection

Methods previously presented (elementary methods) are used as starting point of statistical research correlation between phenomena. With their help it is made a systematization of information needed for applying analytical methods for calculating and interpreting the phenomena.

The possibility of using analytical methods depends on the specific nature of the phenomena studied, on the nature of the information available and on the number of factors studied. If we have one factorial feature (X) and one result feature (Y) it will be applied the simple correlation method. Instead, if we register more factorial features (X_1, X_2,…,X_n) and one result feature (Y) it will be applied the multiple correlation method. According to Pagano (1990), correlation and regression are very much related. They both involve the relationship between two variables. Regression however is primarily concerned with using the relationship for prediction, whereas correlation is concerned primarily with finding out whether a relationship exists and with determining its intensity and direction.

3. Simple and multiple correlations

To highlight statistic connections and measure their trend, are used equations of estimation corresponding to analytic functions. There functions are called regression and their graphical representations are done by regression curve. For the right choice of regression function, the series of distribution are represented through a correlation graph based on which it can be appreciated if the link is linear or curvilinear. As a tool for verifying the accuracy of the estimation function there are used also the methods of dispersion analysis.

Regression function expresses how the result feature (Y) changes if the values of the factorial feature vary (X) and the other factors have a constant action. Other features are considered nonessential and are synthesized into a single value with average character. If the graphical representation indicates a linear trend, the regression equation expressing this connection is:

$$\hat{y}_i = a + b \cdot x_i$$

Parameter "a" is the intercept and indicates the level that would have reached characteristic Y if all factors, except the registered one, had constant action on it. Parameter "b" is the regression coefficient and it geometrically represents the slope of the regression. This parameter is positive for a direct connection or negative for a reverse connection and shows how changes in average variable Y if X changes by one unit. When b=0 the two variables are independent. The determination of the parameters "a" and "b" is made using the method of least squares, which determines the minimum of the function:

$$f(a, b) = \sum (y_i - \hat{y}_i)^2$$

The great popularity of this method among statisticians and economists is due to the fact that is easy to apply and because of the quality of the estimators, which generally meet a number of useful qualities for analyzes and prognoses. This method leads to a high degree of determination and does not require too many costs.

If in nature the nonlinearity is predominant, in economic and social life we encounter the same situation. On medium and especially on long term the behavior of effect variable is rather nonlinear in relation with causal variable. Thus, marketing specialists frequently refer to sales evolution that describes in a nonlinear manner the product's life on market. The economic growth phenomenon illustrated by the evolution of synthetic indicators (Gross Domestic Product, Net Domestic Product) is characterized rather by increases followed by decreases than to rigid linear growth. Therefore, to describe a large number of economic processes, nonlinear models are needed. According to Pecican (2009), these kinds of models are classified as follows:

1. Nonlinear models in relation to variables but linear in relation to the parameters to be estimated. These models can be converted into linear representations through appropriate transformations. Parameter estimation problem is simplified in relation to the ability of the econometrician to realize the linearization of the model.
2. Nonlinear models both in relation to variables and parameters. Such a model may represent the link between demand, price and income. In this second category of nonlinear models, residual variable may create problems impeding the linearization of the model. Therefore, we can consider an alternative model that allows linearization (if it does not betray the reality described by the data) or we can use alternative estimation methods.

Often, the variation way of the dependent variable is the result of many interaction factors. The influence degree of the factors varies from one unit to another, and the overall influence of a factor is found through another factor that is interdependent. In case of mass phenomena, the resultative variable may be a function of several variables:

$$y = f(x_1, x_2, \ldots, x_n)$$

Whether the relationship between each factor and outcome variable is linear, the estimation equation is:

$$y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \ldots + a_n \cdot x_n$$

According to Guyon (2001), to calculate the values of these parameters is necessary to determine the system of normal equations:

$$n \cdot a_0 + a_1 \cdot \sum x_{1i} + a_2 \cdot \sum x_{2i} + \ldots + a_n \cdot \sum x_{ni} = \sum y_i$$
$$a_0 \cdot \sum x_{1i} + a_1 \cdot \sum x_{1i}^2 + a_{2\acute{e}} \sum x_{1i} \cdot x_{2i} + \ldots + a_n \cdot \sum x_{1i} \cdot x_{ni} = \sum x_{1i} \cdot y_i$$
$$a_0 \cdot \sum x_{2i} + a_1 \cdot \sum x_{1i} \cdot x_{2i} + a_2 \cdot \sum x_{2i}^2 + \ldots + a_n \cdot \sum x_{2i} \cdot x_{ni} = \sum x_{2i} \cdot y_i$$
.....
$$a_0 \cdot \sum x_{ni} + a_1 \cdot \sum x_{1i} \cdot x_{ni} + a_2 \cdot \sum x_{2i} \cdot x_{ni} + \ldots + a_n \cdot \sum x_{ni}^2 = \sum x_{ni} \cdot y_i$$

The system can be solved using one of the following methods: Cramer's method, matrix method, Gauss-Jordan's method.

4. Case study: multiple correlations analysis within textile industry firms from Romania

Formens company was founded in 1999 in Botosani, Romania, by two engineers with textile industry background. Its proximity with Ukraine, the Republic of Moldova and Iasi University (specialized in textiles since 1934) locates the company in a strategic textile hub recognized for over a century. The combination of French and Romanian know-how made the company leader of the suit industry in France. Their adaptability to the market demand as well as their anticipation of their clients' needs distinguished them on two different fronts:

- on the "Ready to Wear" market for international brands; starting with the development of entire collections up to the production of the suits, jackets, trousers and other casual pieces;
- on the "Made to Measure" market with fused, half canvas, full canvas constructed suits as well as made to measure shirts.

Formens acquired the leading position for made to measure suits on several countries in Europe, including France.

Table no.2. Economic indicators of Formens company

Year	Productivity (million RON/employee)	Turnover (million RON)	Profit (million RON)
2008	0.11	56.8	4.8
2009	0.13	68.8	5.0
2010	0.17	96.3	7.7
2011	0.21	121.8	12.0
2012	0.23	144.5	15.0
2013	0.24	166.1	13.2

Source: Ministry of Public Finance of Romania

The resultative variable (Y) is the profit and the causal variables are productivity (X_1) and turnover (X_2).

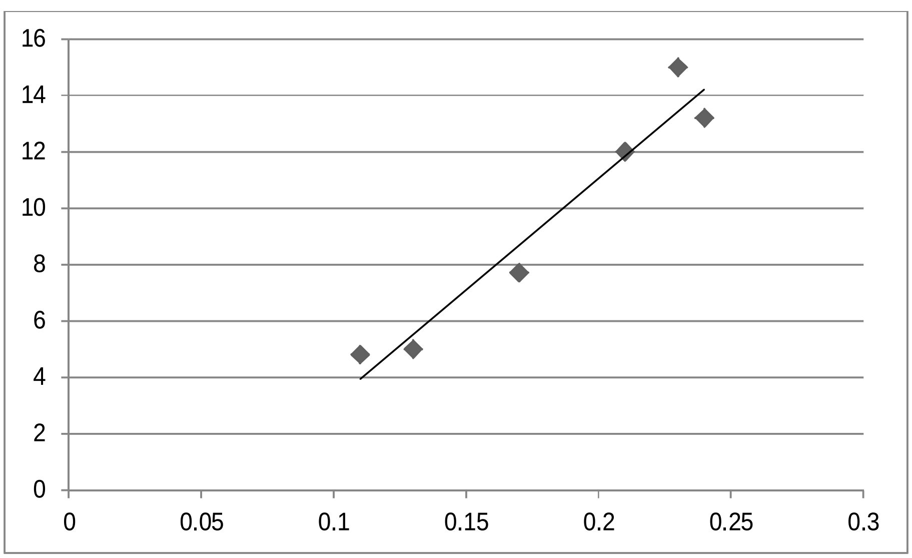

Fig.no.4. The graphical representation of the connection between productivity and profit

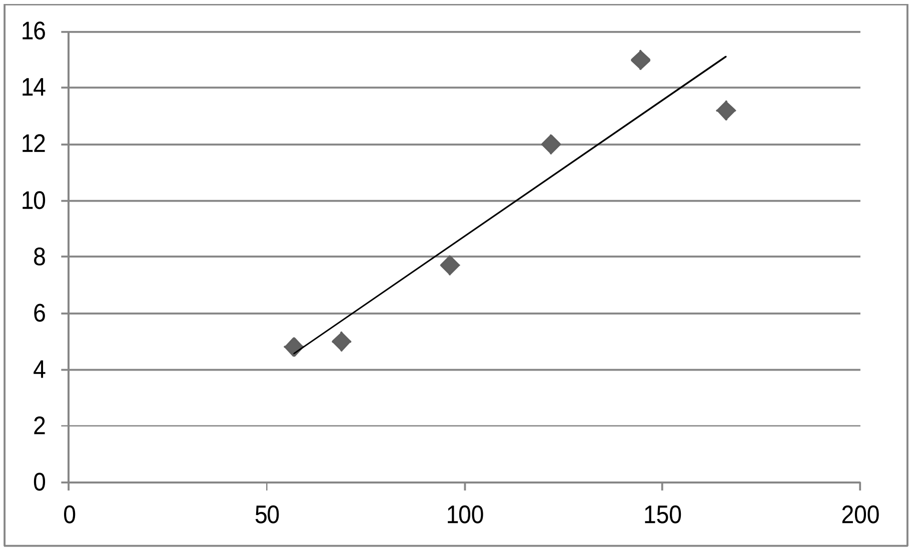

Fig.no.5. The graphical representation of the connection between turnover and profit

Drawn graphs indicate the existence of linear connections between X_1 and Y and between X_2 and Y. As a result, the equation characterizing the connection between variables, correspond to the multiple linear model:

$$\hat{y}_i = a_0 + a_1 \cdot x_{1i} + a_2 \cdot x_{2i}$$

The normal equation system becomes:

$$6 \cdot a_0 + a_1 . \sum x_{1i} + a_2 . \sum x_{2i} = \sum y_i$$
$$a_0 . \sum x_{1i} + a_1 . \sum x_{1i}^2 + a_2 . \sum x_{1i} . x_{2i} = \sum x_{1i} . y_i$$
$$a_0 . \sum x_{2i} + a_1 . \sum x_{1i} . x_{2i} + a_2 . \sum x_{2i}^2 = \sum x_{2i} . y_i$$

Table no. 3. Processed data

No.	y_i	x_{1i}	x_{2i}	$(x_{1i})^2$	$(x_{2i})^2$	$x_{1i} \cdot y_i$	$x_{2i} \cdot y_i$	$x_{1i} \cdot x_{2i}$	y_i^2
1	4.8	0.11	56.8	0.0121	3226.24	0.528	272.64	6.248	23.04
2	5	0.13	68.8	0.0169	4733.44	0.65	344	8.944	25
3	7.7	0.17	96.3	0.0289	9273.69	1.309	741.51	16.371	59.29
4	12	0.21	121.8	0.0441	14835.24	2.52	1461.6	25.578	144
5	15	0.23	144.5	0.0529	20880.25	3.45	2167.5	33.235	225
6	13.2	0.24	166.1	0.0576	27589.21	3.168	2192.52	39.864	174.24
Sum	57.7	1.09	654.3	0.2125	80538.07	11.625	7179.77	130.24	650.57

Replacing data in the table, the normal equation system becomes:

$$6 \cdot a_0 + 1.09 \cdot a_1 + 654.3 \cdot a_2 = 57.7$$
$$1.09 \cdot a_0 + 0.2125 \cdot a_1 + 130.24 \cdot a_2 = 11.625$$
$$654.3 \cdot a_0 + 130.24 \cdot a_1 + 80538.07 \cdot a_2 = 7179.77$$

$$\Delta = \begin{vmatrix} 6 & 1.09 & 654.3 \\ 1.09 & 0.2125 & 130.24 \\ 654.3 & 130.24 & 80538.07 \end{vmatrix}$$

$$\Delta_0 = \begin{vmatrix} 57.7 & 1.09 & 654.3 \\ 11.625 & 0.2125 & 130.24 \\ 7179.77 & 130.24 & 80538.07 \end{vmatrix}$$

$$\Delta_1 = \begin{vmatrix} 6 & 57.7 & 654.3 \\ 1.09 & 11.625 & 130.24 \\ 654.3 & 7179.77 & 80538.07 \end{vmatrix}$$

$$\Delta_2 = \begin{vmatrix} 6 & 1.09 & 57.7 \\ 1.09 & 0.2125 & 11.625 \\ 654.3 & 130.24 & 7179.77 \end{vmatrix}$$

$\Delta = 21.908; \Delta_0 = -132.649; \Delta_1 = 2412.554; \Delta_2 = -0.871$

$a_0 = -6.055; a_1 = 110.122; a_2 = -0.040$

Replacing the parameter values it is obtained the following linear regression equation:

$\hat{y}_i = -6.055 + 110.122 \cdot x_{1i} - 0.04 \cdot x_{2i}$

Correlation analysis between variables can also be characterized by the partial correlation coefficient which measures the intensity of the connection between two variables in the assumption that all other variables remain constant. The values of the partial correlation coefficients ($r_{Y/X1} = 0.97$ and $r_{Y/X2} = 0.93$) indicate that between productivity and profit there is a direct, very strong connection and also that between turnover and profit there is a direct, strong connection. Multiple correlation coefficient between the two independent variables and the result variable indicates that between productivity, turnover and profit there is a strong, direct correlation ($r_{Y/X1,X2}=0.86$).

Rosko Textil is a textile manufacturing company in Arad County, Romania. The firm has its own factory built in free zone Arad- Curtici. Its sole shareholder is the Dutch company Branded Apparel that is controlled by the Americans from Sun Capital Partners. 100% of their products are exported in the European Union countries.

Table no. 4. Economic indicators of Rosko Textil company

Year	Productivity (million RON/employee)	Turnover (million RON)	Profit (million RON)
2008	0.41	240.7	13.0
2009	0.31	251.9	18.7
2010	0.28	311.7	18.0
2011	0.28	332.5	21.5
2012	0.29	348.7	18.5
2013	0.26	318.8	17.0

Source: Ministry of Public Finance of Romania

Table no.5. Processed data

No.	y_i	x_{1i}	x_{2i}	$(x_{1i})^2$	$(x_{2i})^2$	$x_{1i} \cdot y_i$	$x_{2i} \cdot y_i$	$x_{1i} \cdot x_{2i}$	y_i^2
1	13	0.41	240.7	0.1681	57936.49	5.33	3129.1	98.687	169
2	18.7	0.31	251.9	0.0961	63453.61	5.797	4710.53	78.089	349.69
3	18	0.28	311.7	0.0784	97156.89	5.04	5610.6	87.276	324
4	21.5	0.28	332.5	0.0784	110556.25	6.02	7148.75	93.1	462.25
5	18.5	0.29	348.7	0.0841	121591.69	5.365	6450.95	101.123	342.25
6	17	0.26	318.8	0.0676	101633.44	4.42	5419.6	82.888	289
Sum	106.7	1.83	1804.3	0.5727	552328.37	31.972	32469.53	541.163	1936.19

Replacing data in the table, the normal equation system becomes:

$6 \cdot a_0 + 1.83 \cdot a_1 + 1804.3 \cdot a_2 = 106.7$
$1.83 \cdot a_0 + 0.5727 \cdot a_1 + 541.163 \cdot a_2 = 31.972$
$1804.3 \cdot a_0 + 541.163 \cdot a_1 + 552328.37 \cdot a_2 = 32469.53$

$$\Delta = \begin{vmatrix} 6 & 1.83 & 1804.3 \\ 1.83 & 0.5727 & 541.163 \\ 1804.3 & 541.163 & 552328.37 \end{vmatrix}$$

$$\Delta_0 = \begin{vmatrix} 106.7 & 1.83 & 1804.3 \\ 31.972 & 0.5727 & 541.163 \\ 32469.53 & 541.163 & 552328.37 \end{vmatrix}$$

$$\Delta_1 = \begin{vmatrix} 6 & 106.7 & 1804.3 \\ 1.83 & 31.972 & 541.163 \\ 1804.3 & 32469.53 & 552328.37 \end{vmatrix}$$

$$\Delta_2 = \begin{vmatrix} 6 & 1.83 & 106.7 \\ 1.83 & 0.5727 & 31.972 \\ 1804.3 & 541.163 & 32469.53 \end{vmatrix}$$

$\Delta = 348.593$; $\Delta_0 = 8194.434$; $\Delta_1 = -12389.956$; $\Delta_2 = 2.071$

$a_0 = 23.507$; $a_1 = -35.543$; $a_2 = 0.006$

Replacing the parameter values it is obtained the following linear regression equation:

$$\hat{y}_i = 23.507 - 35.543 \cdot x_{1i} + 0.006 \cdot x_{2i}$$

The values of the partial correlation coefficients ($r_{Y/X1} = -0.76$ and $r_{Y/X2} = 0.62$) indicate that between productivity and profit there is a indirect, strong connection and that between turnover and profit there is a direct, medium connection.
Multiple correlation coefficient between the two independent variables and the result variable indicates that between productivity, turnover and profit there is a medium, indirect correlation ($r_{Y/X1,X2} = -0.56$).

5. Conclusions

Both economic and social phenomena met lately a more complex evolution, which can hardly be explained by simple causal relationships. Interconnections between independent factors, which manifest directly or indirectly, entail an increase in the degree of difficulty in identifying causal variables and the intensity of the influence they exert over effect variables. To remove there impediments, it can be used multiple regression method which analyze the relationship between a dependent variable and two or more independent features.

Regression function is a statistical hypothesis, which expresses the trend of the connection between features, considering as variable only the registered factors. But the empirical values of the dependent variable are formed under the influence of all factors, including factors registered, so between empirical values and estimated values may appear errors due to the influence of other factors with random character. Although the regression equations allow us to appreciate the connection between the two features, this is only the first step towards measuring the correlation between phenomena.

References

1. Guyon, X.(2001). *Statistique et econometrie*. Paris: Ellipses.
2. Isaic- Maniu, A., & Mitrut, C., & Voineagu, V. (2004). *Statistics*. Bucharest: Universitary Publishing House.
3. Ministry of Public Finance of Romania. (2014). *Information about economic agents*. Retrieved from http://www.mfinante.ro/infocodfiscal.html
4. Pagano, R. (1990). *Understanding Statistics in the Behavioral Sciences*. U.S.A.: West Publishing Company.
5. Pecican, E. S. (2009). *Econometrics for Economists*. Bucharest: Economical Publishing House.