Using Informatic Methods for Quantitative Analysis of the Knowledge-Based Economy

Lect. Simona GHITA PhD, prof. Vergil VOINEAGU Ph.D, prof. Emilia ȚIȚAN PhD,
Lect. Cristina BOBOC PhD, assist. Daniela TODOSE, PhD Candidate
Statistics and Econometrics Department, Academy of Economic Studies Bucharest

During the European Council Reunion in Lisbon, in March 2000, it has been established an ambitious goal: that Europe must become, in 10 years at most, the most competitive knowledge-based economy, the action-plan for reaching this objective having three main elements: informational society, the education and the research. In this paper we present some of the statistical methods and informatic solutions for analyzing a knowledge-based economy, using non-parametric statistical tests, in order to reveal the dimension of research-development activity in European countries.
Keywords: *knowledge-based economy, knowledge management, research-development activity, non-parametric test, R&D personnel, R&D expenditure.*

The knowledge economy and the knowledge management

The international cooperation in scientific field between the European Union member countries and the non-member countries has more and more significance (the "European Research Area" concept). This is a way of solving major problems, through intensifying the common efforts, in research activity. These ample projects will be financed, in order to increase the competitivity of European economy, creating real research networks.

It is obvious that mankind is on its way between a technology-based era and a knowledge-based era; it has been created the "knowledge-based society" concept and the "knowledge-based economy" concept. The knowledge based era has the following characteristics:

- the intensive technology developments;
- technologies became more and more accessible;
- the spectacular increasing of the performance/price ratio, in information technology and communication domain;
- technology is based on science in a greater proportion;
- integration of the research-development activity in education, training activities and in technology-innovation transfer activities;
- the limitation of the science-development and utilization caused by the insufficiency of the human resources.

Despite the ambiguity associated with the concept of the Knowledge Economy (KE), a working definition provides some insights into its Human Resources Development dimensions. The definition provided by APEC Human Resources Development Working Group is:

"The Knowledge Economy is that strategic combination of organizational changes, policy, settings and capacity building based on the innovation and creativity promoted by expanded international trade and global networked information technologies, which achieves sustainable economic growth and social wellbeing."

The intention of this definition is to encompass five interrelated features of the Knowledge Economy, beyond the technical dimensions of Information and Communication Technology (ICT):

- The organizational changes, particularly in management thinking and work organization that follow the adoption of advanced technologies.
- The capacity building required to equip people to respond effectively to these technologies. People must have the capacity to manage the KE.
- Human resources development in relation to the KE, is a strategic issue, extending far beyond training and education provision.

- The significance of policymaking and implementation in supporting the introduction of these technologies (particularly in relation to fiscal policy and trade and investment openness).
- The relationship between economic growth and social welfare that emerges in the KE.

In the Knowledge Economy, the *Knowledge Management* became more and more popular in the practices of private and public organizations, as long as it is capable to manage the risks in the new technological and economical context.

The economy became more and more immaterial. In an organizational context, data represents facts or values of results, and relations between data and other relations have the capacity to represent information. Patterns of relations of data and information and other patterns have the capacity to represent knowledge. For the representation to be of any utility it must be understood, and when understood the representation is information or knowledge to the one that understands. The following associations can be made:

- **Information** relates to description, definition, or perspective and entails an understanding of the relations between data (*what, who, when, where*).
- **Knowledge** comprises strategy, practice, method, or approach and embodies both a consistency and completeness of relations which, to an extent, creates its own context (*how*).
- **Wisdom** embodies principle, insight, moral, or archetype and arises when one understands the foundational principles responsible for the patterns representing knowledge being what they are (*why*).

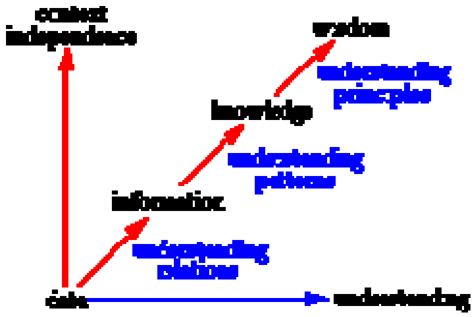

Fig.1. Knowledge economy

In an organizational context, data represents facts or values of results, and relations between data and other relations have the capacity to represent information. Patterns of relations of data and information and other patterns have the capacity to represent knowledge. For the representation to be of any utility it must be understood, and when understood the representation is information or knowledge to the one that understands. Yet, what is the real value of information and knowledge, and what does it mean to manage it?

Resolving non-parametric tests with software statistical product STATISTICA for Windows

We present several statistical techniques and their applications with software statistical product STATISTICA for Windows, tests that may be used as alternatives to parametric tests, when the variables are measured on an ordinal, interval or ratio scale, and the assumption of normality is questionable. These tests can be used instead of Student's t-ratio.

STATISTICA is a true powerful analytics platform with more than two decades of history in delivering successful business results, a global user base of more than 600.000 users. STATISTICA Data Analysis Software Product provides the most comprehensive array of data analysis, data management, data visualization and data mining procedures. Its techniques include the widest selection of predictive modeling, clustering, classifications and exploratory techniques in one software platform.

a) ***Mann-Whitney U Test procedure*** provides a powerful nonparametric statistical method for independent two-group designs. It is commonly used when we wish to avoid the assumptions regarding the normal distribution of the populations and that the standard deviations are equal.

We consider two groups of European Union member countries: the older E.U. member countries and the recent E.U. member countries. The data we are analyzing refer to the proportion of the employment in the research-development (R&D) activity in the total labour-force (%). We test if the population distributions from which the two samples were drawn are the same.

Research and experimental development (R&D) comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications.

Proportion of research-development (R&D) personnel of the labour force: R&D personnel include all persons employed directly on R&D sector, plus persons supplying direct services to R&D, such as managers, administrative staff and office staff. Head count data measure the total number of R&D personnel who are mainly or partly employed on R&D. R&D personnel in head-count are expressed as a percentage of the labour force (comprises of population aged 15 and over who are employed or unemployed but not inactive). The calculated Mann-Whitney test value, computed with STATISTICA for Windows is U=12, which means that, for a significance level of p<0,001 the two groups of European Union member countries have a significant different average proportions of the R&D personnel in the total labour-force.

Mann-Whitney U Test (mu.sta)

By variable STATUS
Group 1: 101-Old Group 2: 103-New

variable	Rank Sum Old	Rank Sum New	U	Z	p-level
RD_PERS	209,0000	67,00000	12,00000	3,286921	,001014

b) ***Spearman's and Kendall's Rank Correlation Coefficients*** provides a procedure for studying if there exist an association between two variables measured on an ordinal, interval or ratio scale and which are not normally distributed. The two variables we selected are: the share of R&D personnel in total labor force and the percentage of R&D expenditures in GDP.

Proportion of R&D expenditure in GDP: R&D expenditure include all expenditures for R&D performed within the business enterprise sector on the national territory during a given period, regardless of the source of funds. They are shown as a percentage of Gross Domestic Product (GDP).

Applying the STATISTICA product we found a Spearman's Rank Correlation Coefficient of 0,926 and a Kendall's Rank Correlation Coefficient of 0,788, which allow us to conclude that there is a rather strong direct dependence between the two variables included in the study.

Spearman Rank Order Correlations (spearman.sta)

MD pairwise deleted

Pair of Variables	Valid N	Spearman R	t(N-2)	p-level
%RD_PERS & %RD_EXP	24	,926288	11,52999	,000000

Kendall Tau Correlations (spearman.sta)

MD pairwise deleted

Pair of Variables	Valid N	Kendall Tau	Z	p-level
%RD_PERS & %RD_EXP	24	,787660	5,392333	,000000

c) χ^2 ***test procedure*** provides a classification of variables on various categories and computes the χ^2 test value. The procedure compares the observed frequencies with the expected, theoretical frequencies, for each category, in order to test if all categories contain the same proportion values or each category contains a certain user specified proportion value. To be specific, we tested if there is a significant difference between the average proportion of R&D expenditures in GDP, for European countries with different status (in relation with European Union: old E.U. member countries, new E.U. member states, non-E.U. member states, candidate E.U. member states). In order to respond to this question, we grouped the European countries after two variables: one categorical (nominal) variable: country status in relation with European Union and the other, a numerical variable: the percentage of R&D expenditure in GDP. The data consists of frequency counts, that are placed in a 4x4 table (table no. 1 – observed frequencies and table no. 2 – expected frequencies):

Observed frequencies

Table no. 1

Country status	Percentage of R&D expenditures in GDP (%)				Total
	0-1	1-2	2-3	3-4	
Old E.U. member states	2	7	4	2	15
New E.U. member states	8	2	-	-	10
E.U. candidates countries	3	1	-	-	4
Non-E.U. member countries	-	1	1	1	3
Total	13	11	5	3	32

Expected frequencies

Table no. 2

Country status	Percentage of R&D expenditures in GDP (%)			
	0-1	1-2	2-3	3-4
Old E.U. member states	6,09	5,16	2,34	1,41
New E.U. member states	4,06	3,44	1,56	0,94
E.U. candidates countries	1,62	1,38	0,63	0,38
Non-E.U. member countries	1,22	1,03	0,47	0,28

Observed vs. Expected Frequencies (new.sta)

Continue... Chi-Square = 19,22500 df = 15 p < ,203800

Case	observed FRECV_OB	expected FRECV_TH	O - E	(O-E)**2 /E
C: 1	2,00000	6,09000	-4,09000	2,74681
C: 2	8,00000	5,15000	2,85000	1,57718
C: 3	3,00000	2,34000	,66000	,18615
C: 4	0,00000	1,41000	-1,41000	1,41000
C: 5	7,00000	4,06000	2,94000	2,12897
C: 6	2,00000	3,44000	-1,44000	,60279
C: 7	1,00000	1,56000	-,56000	,20103
C: 8	1,00000	,94000	,06000	,00383
C: 9	4,00000	1,62000	2,38000	3,49654
C: 10	0,00000	1,38000	-1,38000	1,38000
C: 11	0,00000	,63000	-,63000	,63000
C: 12	1,00000	,38000	,62000	1,01158
C: 13	2,00000	1,22000	,78000	,49869
C: 14	0,00000	1,03000	-1,03000	1,03000
C: 15	0,00000	,47000	-,47000	,47000
C: 16	1,00000	,28000	,72000	1,85143
Sum	32,00000	32,00000	,00000	19,22500

The computed χ^2 test value is 19,225, so we can conclude that the proportion of the R&D expenditures in GDP depends on the European country status, for a significance level p<0,2038.

Computed Pearson's Contingency Coefficient value is:

$$CC = \frac{\sqrt{\frac{\chi^2_{comp}}{N + \chi^2_{comp}}}}{\sqrt{\frac{L-1}{L}}} = 0{,}69$$

Where χ^2_{comp} is the computed χ^2 test value, obtained from analysis, *N* is the total number of observations in the *RxC* table, and *L* is the smaller value of *R* or *C*. We found the Pearson's Contingency Coefficient value 0,69, which means that there is a medium-intensity association between the proportion of the R&D expenditures in GDP and the European country status.

References

[1] Haworth, Nigel - „*Human Resource Development Strategies for Knowledge Economy"*;

[2] Titan, Emilia, - "*Statistica. Teorie si aplicatii in sectorul tertiar*", Editura Meteor Press, Bucuresti, 2003;

[3] *** - „*Vers les sociétés du savoir"*- Raport mondial UNESCO;

[4] EUROSTAT;

[5] http://www.statsoft.com