Competences Offered to Statisticians by the Italian Universities and Required by the Job Market

Enrica Aureli*, Domenica Fioredistella Iezzi**

** Department of Statistics, Probability and Applied Statistics, University "La Sapienza" of Rome, Italy*
*** Department of Economical and Financial Studies and Quantitative Methods, University of Rome "Tor Vergata" , Italy*

Summary. The development of the European higher education system, within the frame of the Bologna process, requires that higher education be more and more integrated with the wider economic strategies (Lisbon objectives) and that employability be taken into account by the study programmes. Higher education institutions are invited not only to answer to the needs of the national labour markets but also to incorporate into their perspectives the European one. The Italian university reform provides for a first triennial cycle with educational and training purposes that give graduates an immediate access to the labour market. The labour market requires, on the other hand, graduates to have not only a theoretical education, but also professional competences. This paper aims to build cognitive maps of the study programmes in Statistics and rank the competences required by the job market to graduates in Statistics. We apply for that various methods of multivariate and textual analysis.
Keywords: Competences; Statistics; Programme statements; K-means cluster analysis; Multiple correspondence analysis; Rasch analysis; Textual analysis.

1. Competences and the Italian University reform

The reform of the Italian university system (decree of 3 November 1999, no. 509) has radically transformed the formative objectives of the study programmes (Aureli & Iezzi, 2004, 2005; Iezzi, 2005). Since 2001, the Italian university programmes are organised on 3 cycles: Three-year cycle (CDL1 = Bachelor); Two-year cycle (CDL2 = Master); Third cycle of variable length (CDL3 = Ph.D.).

The CDL1 aims at guaranteeing undergraduate students with an adequate command of general scientific concepts and methods, as well as professional skills. General admission requirement is the Italian high school certificate awarded to those who pass a national exam, after completing 13 years of schooling; foreign comparable certifications may be accepted as well. Admission to individual degree programmes may be subject to specific requirements. First level degrees are awarded to students who have earned 180 credits.

The CDL2 includes highly qualified activity in specific areas. Access to the second level is through the Italian CDL1 or a foreign comparable degree. This degree is awarded to students who have earned a global amount of 300 credits, including a maximum of 180 of the first level that have been recognised for access to the CDL2. The drawing up of an original dissertation is compulsory.

A limited number of programmes regulated by EU directives (in dentistry, human medicine, veterinary medicine, pharmacy) share the following different features: access is free to those who have an Italian high school graduation diploma or a foreign comparable certification. Admission is always subject to an admission exam; normal length is 5 years (human medicine takes 6 years).

The postgraduate studies consist of the following typologies of degree programmes:

A) Ph.D. programmes; B) 2nd level specialisation programmes;
C) 2nd level university master degree programmes.

Competence, or *competency* (Spencer & Spencer, 1993) is “an underlying characteristic of an individual that is causally related to criterion-referenced effective and/or superior performance”. In our case, a competence may be defined "an enduring characteristics of an individual that causes and foresees his/her work-hunting and professional behaviour".

Competences are a mix of knowledge, skills and attitudes that make it possible graduates to be successful in their interactions with others at work, school, home, and in the community at large. To match the university educational offers and the job market requisites we assume the so-called ‘iceberg model’ of competences (Spencer & Spencer, 1993, see Table 1).

The concept of competence became influential initially in business organisations, more specifically in the field of recruiting and selecting new employees. The McClelland’s (1973) goal of “testing for competence rather than for intelligence” is the starting point of the competence movement (Barrett & De-

Table 1. Iceberg model of competences

Measurement	Performance	Skills Knowledge	Behaviour
Hidden	Competence	Self-concept Stable personal attributes	Competence Convictions/ beliefs Identity
Guess	Ability	Motives & intentions	Spirituality

pinet, 1991). McClelland argued that commonly used I.Q. and personality tests were poor predictors of successful performance, and that competence assessment had to be developed instead. He suggested the Behaviour Event Interview (BEI), an interview method that combines Flanagan's Critical Incident Method (Dailey, 1971; Boyatzis, 1982) with the Thematic Apperception Test (McClelland, 1989).

The BEI is often applied to discover differences between a person, who has been selected by knowledgeable judges as outstanding, and a reference person. The underlying assumption is that it is less difficult to deciding who is competent than what makes them competent (McClelland, 1998). Starting from several studies conducted in various types of organisations, Spencer & Spencer (1993) wrote a dictionary of competences that distinguishes superior from average performers in middle to upper-level jobs.

Consoli & Benadusi (1999) point out that the concept of competence emerges together with that of learning, formative credits, project organisation, continuous evaluation. There is an ongoing national and international debate (Ajello, 2002; Boyatzis, 1982; Capaldo *et al*., 1996; Carretta *et al*., 1992; Civelli & Manara 1997; Cerase, 2002; Consoli, 1998; Consoli & Benadusi, 1999; Fabbris, 2004; Iezzi, 2003; Zan, 1988) on competences required by the job market from new graduates.

The aim of this paper is to verify points of contact between the university offer and the demand of the job market. The structure of the paper is as follows: in Section 2 we present the data at hand; in Section 3 the results of programme statement analysis; in Section 4 the ranking of competences required by job market from new graduate in Statistics, and in Section 5 the conclusions of our analysis.

We will use *partial credit* method (Wright & Masters, 1982) to rank the competences required from graduates. Partial credit is a Rasch model that transforms categorical data in continuous data, providing a complete solution to almost every measurement problem encountered in social science studies.

2. The data

Our study examines the Italian university system and the new role of competences in CDL1. In this way, we can match the competences achievable through university studies and those required by the labour market.

The functional aspect of this new system favours the development of skills rather than just knowledge as it happened in Italy in the days before the reform. That is why we tried to verify if and how the university educational offer is shaped according to the labour market perspectives, and which are the skills a graduate should posses to get a job.

We realise our analysis in two steps: in the first, we examine the reformed

university system; in the second, we identify the requests of the market. The point of contact between the first and the second phase is the graduate.

In the first phase, we will examine the stated intentions in the programme's arrangements (Ministry decree 4th August 2000). The reform provides for each disciplinary class with a "declaratory", i.e. a programme statement. We will use the 41 statements of "Statistics". Each statement is divided into two parts: the aims of this discipline and the job opportunities for its graduates.

The programmes of the Statistics class represent a small part in comparison with the amount of programmes activated after the reform (1.5% of the triennial programmes). All programmes were organised by five universities: Milan "Bicocca", Padua, Bologna, Rome "La Sapienza", and Messina. We considered the ministerial arrangement as a reference for the 41 programmes.

In 2004, we interviewed 137 people who graduated in Statistics at the University "La Sapienza" between March 2000 and March 2001 through a CATI – *Computer Assisted Telephone Interviewing* survey (Aureli & Ottaviani, 2004). We collected data on graduates' knowledge, technical and "soft" skills and confronted this with their work needs. Knowledge was composed of questions on "*how is it important for your current job*" each one of the 13 following disciplines: Insurance, Demography, Law, Economics, Finance, Operational Research, Computer Science, Mathematics, Probability, Sociology, Economical statistics, and Social statistics.

3. A model to analyse the statements

The first phase of the analysis was the transformation of the verbatim information into a data matrix. We obtained a matrix of 42 rows, 41 CDL1s and the ministerial arrangement. A pre-processing step was required to eliminate meaningless words (Bolasco, 1999). After that, we applied lexical analysis and identified *m* matrixes one for each of two themes: "programme aims" and "job opportunities".

We have built two different data matrices because we wanted to test each of two analytical strategies:

i) a matrix **T** of $[(n+1) \times p)]$ order; where n is the number of statements, 1 represents the model and p is the number of meaningful words;
ii) a matrix **S** of $[(n+1) \times q)]$ order, where n is the number of statements, 1 is the model and q the number of chosen keywords ($p>q$).

For the first strategy, we applied principal component categorical (PCC) to identify and show the association structures on the table. This statistical technique is useful when it is not possible to explain the relationships among objects because of the number of variables.

The principal factors obtained by PCC showed *m* text maps and *m* keywords. We used an integrated approach to classify the texts by starting the re-

sults of PCC. We searched the best partition of texts and the words from initial configuration of group centroids.

For the second analytical strategy, we applied multiple correspondence analysis (MCA) to build a map. The **S** matrix was transformed in *m* blocks of matriz $\mathbf{Z} = [Z_1, ..., Z_m]$ with full disjunctive codes. The dimension of **Z** is $[(n+1) \times q)]$. The application of MCA detected the latent factors (Greenacre, 1984).

Even in this case, we applied an integrated approach to classify the texts. We used three different algorithms to identify the number of groups (single linkage, centroid, and Ward' method). We identified three groups with Ward's method and four with single linkage and centroid methods.

Since the university reform started, the Italian Ministry of Education, University and Research organised a web site for reporting all the activities of the Italian Universities (http://offertaformativa.miur.it/corsi/). We searched the information on first and second study levels and acquired the 41 statements related to Statistics.

The wealth of words depends on topic, the descriptive analysis showed a major wealth for "aims" than for "job opportunities" (Table 2). A detailed list of aims means a good specificity of a programme that justifies various educational offers in a given faculty.

Table 2. Descriptive analysis of graphical forms

Key topic	*no. units*	*no. words*	*graphical forms*	*% graphical form*	*most frequent words*
Aim (Model)	1	175	111	63	TO POSSESS
Job opportunities (Model)	1	193	131	68	ACTIVITY
Aim (CDL1)	41	9718	1465	15	STATISTICS
Job opportunities (CDL1)	41	4598	1041	23	FIRM

The more frequent word is "to possess" for job opportunities topic and "activity" for aims. Those words are also in declaratory statements, but the more common graphical form is "firm" for job opportunities and "statistics" for aims.

We noticed that a change of outlook is starting in university statements, both for knowledge taking and technical and "soft" skill training.

The ministerial purpose is the possession of analytical skills: "graduates must possess adequate knowledge and tools to analyse data". The statements promote the professional role of Universities and define a graduate in Statistics as a problem solver. This young professional is able to analyse and manage complex systems (Table 3).

An exploratory analysis of the model shows that job opportunity keywords focus on educational activity, study and learning, while the statements introduce professional aspects. In the last ones, we find a direct connection to job

Table 3. Keyword of study programme aims

MODEL	Number of words	STATEMENTS	Number of words
TO POSSESS	7	STATISTICAL	148
SUITABLE	3	DATA	83
KNOWLEDGE	3	ANALYSIS	78
STATISTICS	3	MANAGEMENT	72
TREATMENT	2	STATISTICS	61
DATA	2	TOOLS	57
STATISTICAL	2	COMPETENCES	54
TOOLS	2	KNOWLEDGE	52
DISCIPLINES	2	STATISTICIAN	42
COMPETENCES	2	SYSTEMS	41

* A star marks the words used both by the model and by statements.

Table 4. Keyword of job opportunities in statements and model

MODEL	Number of words	STATEMENTS	Number of words
ACTIVITY	6	FIRMS	58
STATISTICIAN	4	ACTIVITY	41
EDUCATIONAL	3	STATISTICS	36
STUDY	2	STATISTIC	35
KNOWLEDGES	2	MANAGEMENT	34
TO UNDERSTAND	2	ANALYSIS	32
CASE	2	PROFESSIONAL	28
CAPABILITY	2	CONTROL	25
FIELD	2	SYSTEMS	21
LEARNING	2	SERVICES	19

* A star marks the words used both by model and by statements

market: firms, activity, management, analysis, professional, control, systems and services (Table 4).

The ministerial arrangement is oriented to curriculum vitae, while the statements are focused on job context and the projection of present students into future graduates.

Maps, obtained with ACM, mark latent factors of aims and job opportunities.

The first factor juxtaposes theory and application words. In fact, we observe that on the left hand side there are knowledge, skills and discipline, and on the right one tools, data and systems. The second factor juxtaposes knowledge and

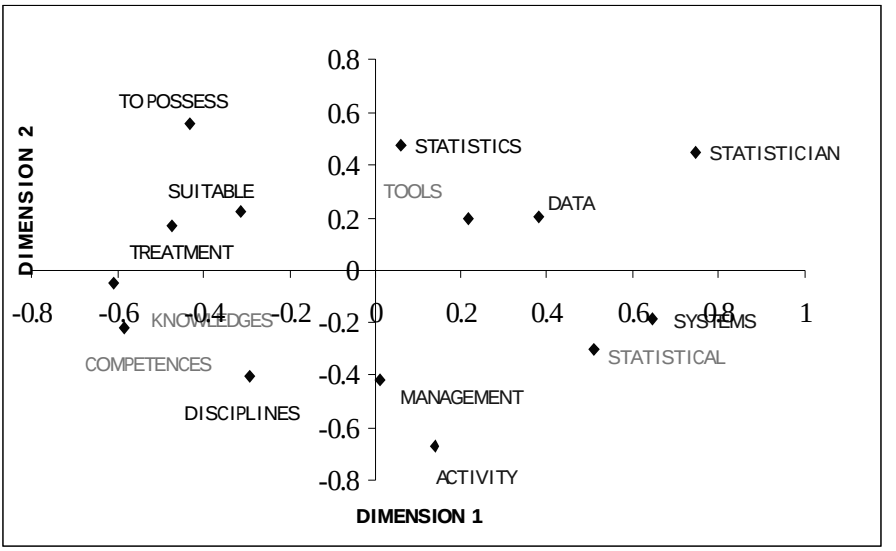


Figure 1. Keyword map of declaratory statement aims

ability; we observe that the verb "to possess" is in the opposite position to management and analysis (Figure 1).

The map of job opportunities shows one principal latent factor that marks prerequisites for work: learning, capability, analysis and management in the private and public sectors. In Figure 2 the two sectors are in an antithetical position. About 60% of the university statements focus on knowledge.

The employment issues classify the CDL1 into four groups that move along a continuum: from public administration oriented programmes, where the words "services", "study" and "educational" prevail, to market-oriented ones, where the words "firm", "activity", and "control" dominate.

Inside this continuum, we have crossed public and private-oriented matter. In the public-private oriented group, we have words like "knowledge", "professional" and "management"; in the private-public oriented group, we found words like "system", "capacity" and "analysis" (Table 5).

Based on MCA, we classified the CDL1. Programme aims divided CDL1 into four groups: theoretical, professional, theoretical-professional, professional-theoretical (Table 6).

Within the theoretical group, knowledge dominates all the aims. In the opposite position, we find the professional group that prefers to teach the tools of the subject matter. The groups may be disposed along a *continuum*: it starts with a theoretical approach ("to possess", "adequate", "treatment") and ends up to a professional planning ("tools", "data" and "statistics". Along this continuum, we have a theoretical-professional and a professional-theoretical approach in which "knowledge", "skill", "ability" and "disciplines" link together.

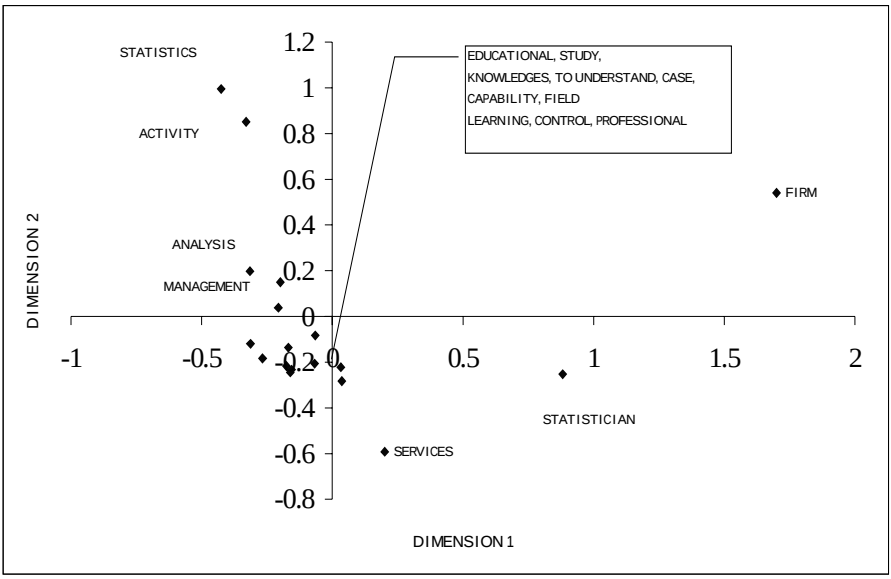


Figure 2. Keyword map of job opportunities

Table 5. Classification of CDL1 based on job opportunities

Key topic	*GROUP*			
	1	***2***	***3***	***4***
CDL1 characteristics	Public oriented	Public-private oriented	Private-public oriented	Private oriented
Keywords	Services Study Educational	Professional Management Knowledge	Systems Capacity Analysis	Firm Activity Control

Table 6. Statement classification of CDL1 based on programme aims

Key topic	*GROUP*			
	1	**2**	**3**	**4**
CDL1 characteristics	Theoretical (cdl=knowledge)	Theoretical-professional (knowledge+skill)	Professional-theoretical (skill+knowledge)	Professional (cdl= skill)
Keywords	To possess Adequate Treatment	Knowledge Competences Discipline	Systems Management Analysis	Tools Data Statistics
Number of CDL1s	7	14	14	7

Moreover, there is cohesion between objectives and job opportunities of the classes of CDL1 (last row of Table 6). The lexical analysis of the *corpus* highlights that it prevails programmes in which theoretical and practical aspects are balanced. However, if we divide the *corpus* into "aims" and "job opportunities", in the second we can find a clear partition between traditional and modern vision of University.

4. A model for competence ranking

We converted the ordinal data into values applying a Rasch partial credit model (RPCM) in order to analyse the data on knowledge, technical and soft skills collected on ordinal scales. RPCM is an extension of the Rasch one-parameter 'item response' model proposed by Rasch (1992) and extended by Wright & Masters (1982), Wright & Mok (2000) and Linacre (2004).

We measured distance and determined the rate of change to define scale units and measures using verbal scale of four items (much, enough, little, not at all). Rasch methods are to obtain objective, fundamental, linear measures from stochastic observations of ordered category responses.

For a given item with *m* score categories, the probability a statistical unit scores *x* on item *i* is given by:

$$P_{xni} = \frac{\exp \sum_{j=0}^{x} (B_n - D_{ij})}{\sum_{k=0}^{mi} \exp \sum_{j=0}^{x} (B_n - D_{ij})},$$

where x=0, 1, 2, …, *m*-1, B_n is the individual competence level and D_{ij} is a step difficulty, and $\sum_{j=0}^{0} (B_n - D_{ij}) \equiv 0$.

The RPCM estimates the probability that a graduate scores *x* on the *m* steps of question *i* as a function of his or her competence level B_n and step difficulty D_{ij} out of the *m* steps in prompt *i*.

We used WINSTEPS program to implement Rasch's JMLE (UCON) (unconditional maximum likelihood, joint maximum likelihood). The JMLE (UCON) method was iterated to obtain more precise estimates, standard errors and fit statistics. We used proportional curve fitting for finding improved estimates.

We examined technical skills through the assessment of 19 items aggregated in 5 classes of topics:

1) writing documents;
2) looking for sources and methods;
3) analysing data;
4) organising and checking trials;
5) speaking a foreign language.

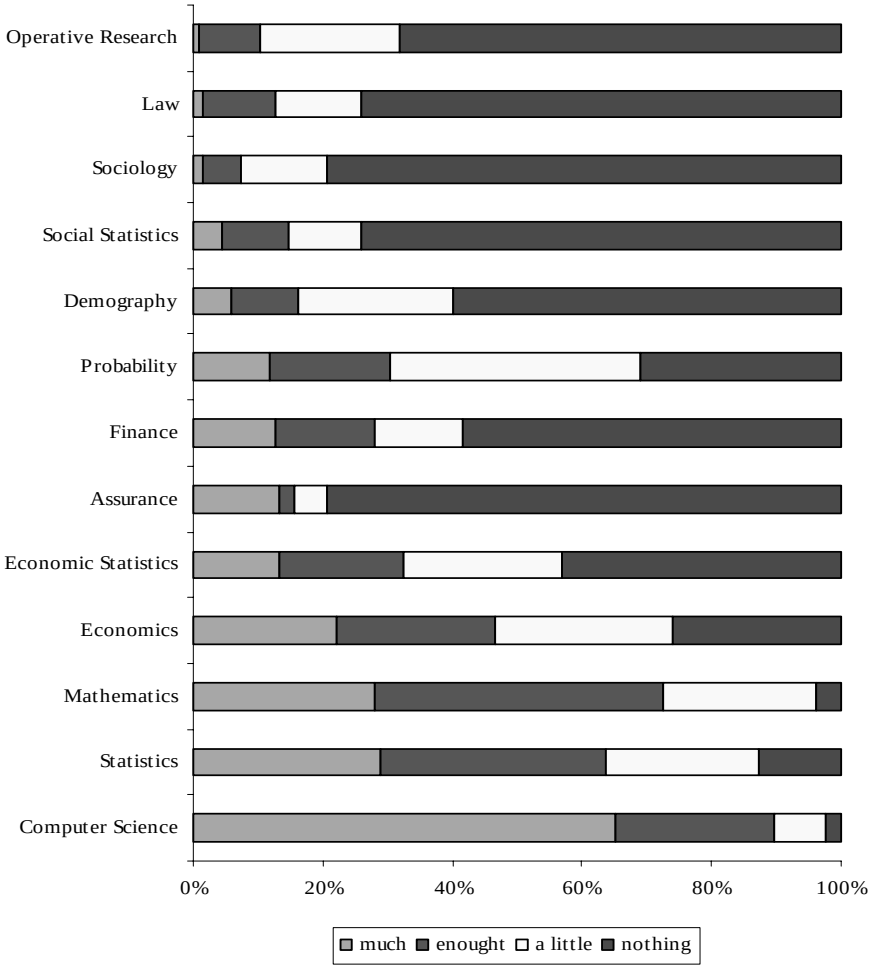


Figure 3. Frequency distribution of the knowledge dimension

We verified soft abilities through the examination of 18 items separated in 5 classes of individual propensity: 1) decision taking; 2) communicating with others; 3) working in a team; 4) setting and solving problems; 5) inclination to long life learning. We have to highlight that there is no clear cut between 'technical' and 'soft' competences.

In knowledge dimension (Figure 3), 60% of graduates' responses (Sociology, Economical Statistics, Social Statistics, Operational Research, Insurance, Law) are concentrated on the negative side.

Technical skills (Figure 4) are more variable than the other dimensions, with the exception of the use of personal computer, which is widespread (Figure 5). Soft ability dimension items are positively oriented (Figure 5).

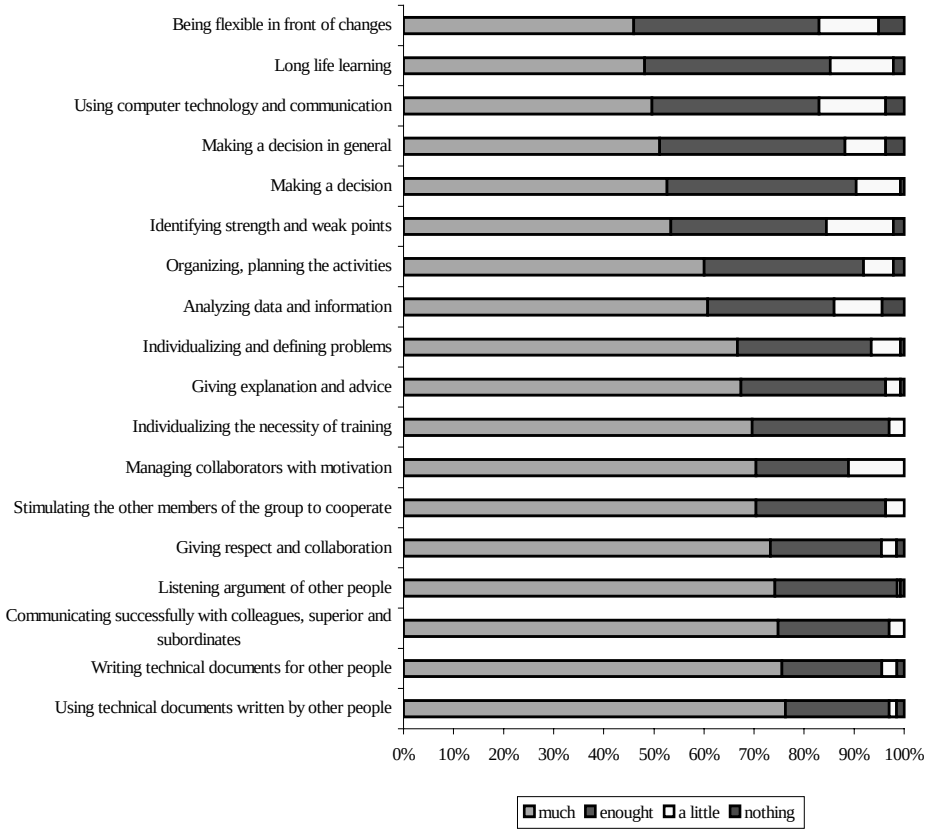


Figure 4. Frequency distribution of the 'technical' skills

We used Cronbach's Alpha (Cronbach, 1951) to measure if and how a set of items for each 'competence dimension' (knowledge, technical and soft skills) describe a uni-dimensional latent construct. The resulting reliability is good for the three subsets; in fact, the alpha value equals 0.76 for knowledge items, 0.86 for skill ones and 0.82 for the ability ones.

The knowledge item map (Figure 6) partially confirms the results of the descriptive analysis: the *continuum*'s top includes the least used disciplines (Sociology and Law) and the bottom includes the widely used ones (Computer Science). We can observe some changes of position.

During the initial stages of employment, freshly graduates in Statistics make use of the basic subjects, such as Mathematics, Statistics and Economics, and a little specialist knowledge.

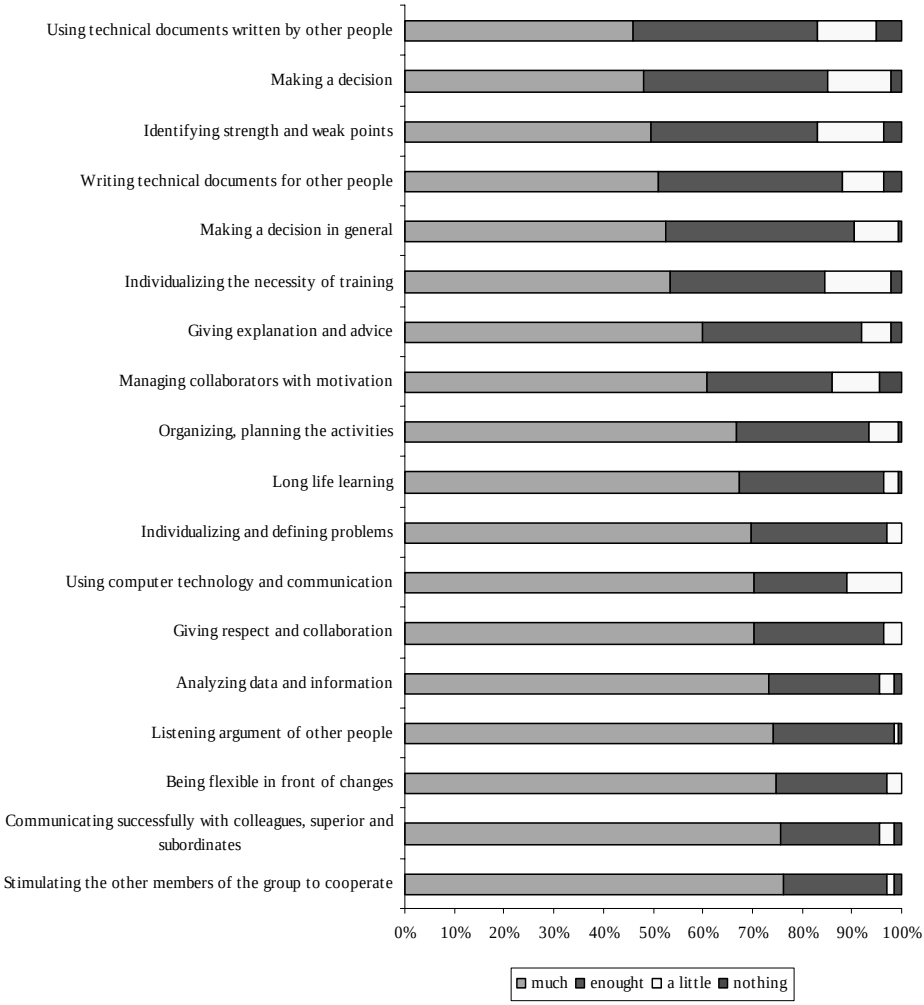


Figure 5. Frequency distribution of the soft abilities

It is worth noticing that some positions, such as Probability, moved and we would not have been able to observe it if we stopped at a descriptive analysis. Moreover, some distances among the disciplines differ.

Computer Science is far away from all other subjects; while Insurance, Social Statistics, Law, Operational Research are close to each other.

Technical skill dimension (Figure 7) is consistent with the knowledge scaling: working with a personal computer is the most important skill used at work by young statisticians, far above any other subject. All other subjects are close to each other: writing technical documents for other people, using technical documents written by other people, using data to evaluate.

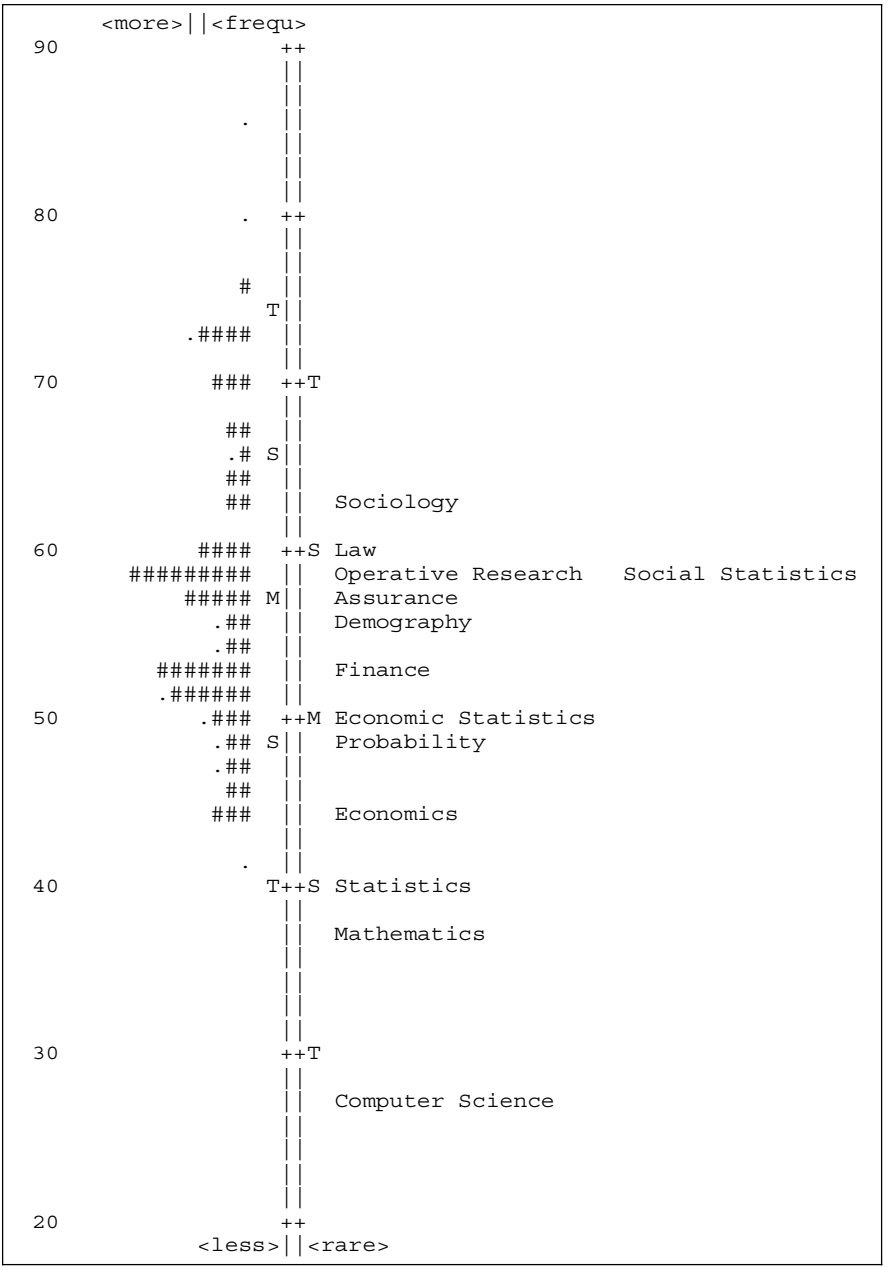
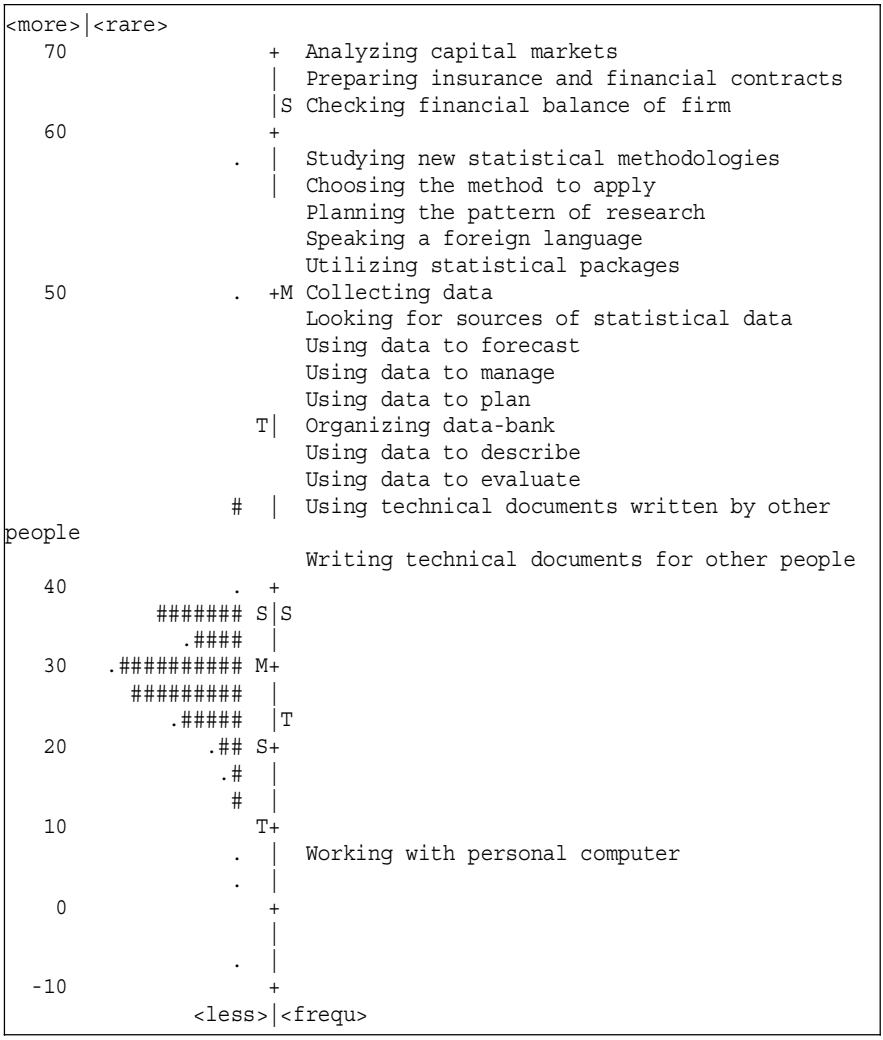


Figure 6. Map of graduates' knowledge items



(EACH ' #' IS 3 persons).

Figure 7. Map of graduates' technical skills

In the cross-occupational ability map (Figure 8), the five areas of the propensity to assume decisions, communicate with others, work in a team, solve problems and be incline to long life learning are very close to each other.

This may mean that during first years of job activity the newly hired graduates in Statistics operate more as technicians than decision makers. This result may depend on the possibility that the survey reached graduates working in research activities together with their dissertation supervisor. These precarious activities are a sort of extension of the final dissertation research activities that graduates accept waiting for job offers.

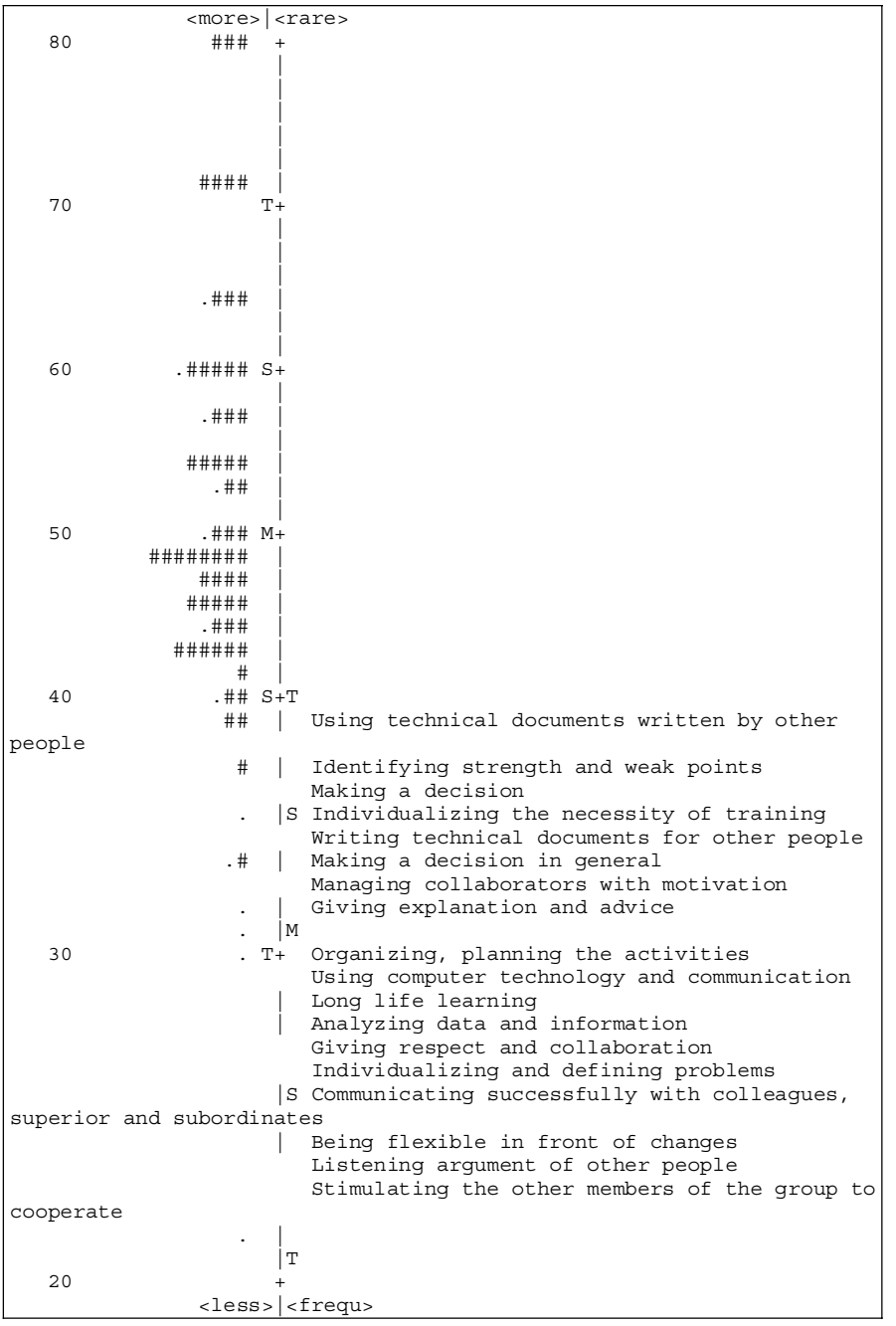


Figure 8. Cross-occupational item ability map

5. Conclusions

The aim of this study was to check if and how the university offer for higher education in Statistics had assumed the competence approach to outline the educational profiles, and if graduates' curricula and job market's requirements matched. We realised that changes for statisticians' education and employment are slowly taking place.

From the methodological point of view, we highlighted that, within a short document containing a homogeneous and standardized text, it is sufficient to analyse keywords. In fact, sparse matrices gave the same results, but needed much more computational burden than traditional approaches. We put forward the proposal to classify key topics and the entire *corpus* with a consensus algorithm, not to waste resources (Vichi, 1997).

From a substantive point of view, the analysis of the statements showed that, although universities have adopted the ministerial model, they modulate their statements taking into account the local market requirements, and tend to define programme statements where typology, knowledge, skill requirement and ability demand prevail in a functional way. In fact, even if the ministerial arrangement remains adherent to substantive knowledge, the universities propose programmes targeted at the development of cross-occupational competences and other job-specific competences on students.

The interviewing of newly graduated statisticians about competence use at work gave a surprising outcome: the graduates, at the very beginning of their career, apply the knowledge of basic disciplines (mathematics, probability and mathematical statistics) and in very limited cases they use the so-called 'applied statistics'. This is right the opposite of the result of similar studies and of commonsense.

Moreover, the labour market appreciates cross-occupational competences that are not explicit on graduates' curricula. This is open to a dual interpretation: either the companies prefer to shape themselves professional competences, and expect solid and flexible foundations from the educational system, or the universities should develop even students' soft skills, not only the technical-specific ones.

References

AJELLO A.M. (2002) *La competenza*, Il Mulino, Bologna.

AURELI E., IEZZI D.F. (2004) Classificazione di testi omogenei. Il problema delle matrici sparse. In: AURELI E., BOLASCO S. (eds) *Applicazioni di analisi statistica dei dati testuali*, Casa Editrice "La Sapienza", Roma: 55-68.

AURELI E., IEZZI D.F. (2005) Conoscenze, abilità e altre competenze nei profili professionali dei laureati in Scienze Statistiche. In: FABBRIS L. (ed) *Efficacia e-*

sterna della formazione universitaria: il progetto OUTCOMES, CLEUP, Padova, **7**: 129-144.

AURELI E., OTTAVIANI M.G. (2004) Dal disegno concettuale allo strumento operativo: la rilevazione delle competenze degli statistici nel mondo del lavoro, In: D'OVIDIO F. (ed) *Professioni e competenze nel lavoro dei laureati*, CLEUP, Padova, **6**: 45-68.

BARRETT G.V., DEPINET R.L. (1991) A reconsideration of testing for competence rather than intelligence, *American Psychologist*, **46 (10)**: 1012-1024.

BOLASCO S. (1999) *Analisi multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione*, Carocci, Roma.

BOYATZIS R.E. (1982) *The Competent Manager: a Model for Effective Performance*, John Wiley & Sons, New York.

CAPALDO G., VOLPE A., ZOLLO G. (1996) Competenze individuali e sviluppo organizzativo, *Sviluppo & Organizzazione*, **155**: 69-79.

CARRETTA A., DALZIEL M., MITRANI A. (1992) *Dalle risorse umane alle competenze*, Franco Angeli, Milano.

CERASE F.P. (2002) *L'analisi delle competenze nel lavoro amministrativo*, Franco Angeli, Milano.

CIVELLI F., MANARA D. (1997) *Lavorare con le competenze. Come conoscerle, gestirle, valorizzarle*, Guerini e Associati, Milano.

CONSOLI F. (1998) Le ragioni forti della metodologia delle competenze, *Scuola Democratica*, **1-2**: 88-107.

CONSOLI F., BENADUSI L. (1999) L'emergenza della metodologia delle competenze nel pensiero teorico e pratico manageriale, *Osservatorio ISFOL*, **XX(5-6)**: 30-89.

CRONBACH L.J. (1951) Coefficient alpha and the internal structure of tests, *Psychometrika*, **16**: 297-334.

DAILEY C. (1971) *Assessment of Lives*, Jossey-Bass, San Francisco.

FABBRIS L. (2004) *Capitalizzazione delle competenze dei laureati dell'Università di Padova*, Quaderno PHAROS, n. 9, CLEUP, Padova.

GREENACRE M.J. (1984) *Theory and Applications of Correspondence Analysis*, Academic Press, London.

IEZZI D.F. (2005).A new method to measure the quality on teaching evaluation of the university system: the Italian case, *Social Indicators Research*, **73(3)**: 459-477.

IEZZI D.F. (2003) L'analisi delle competenze nel lavoro amministrativo. In. CIVARDI M. (ed) *Transizione Università-lavoro. La definizione delle competenze nella pubblica amministrazione*, CLEUP, Padova: 44-58.

LINACRE J.M. (2004) *Facets Rasch Measurement Computer Program*, Winsteps.com, Chicago.

MC CLELLAND D.C. (1973) Testing for competence rather than 'intelligence', *American Psychologist*, **28**: 1–14.

MC CLELLAND D.C. (1989) *Human Motivation*, Cambridge University Press, Cambridge, UK.

MC CLELLAND D.C. (1998) Identifying competences with behavioral-event interviews, *Psychological Science*, **9**: 331-339.

RASCH G. (1992) *Probabilistic Models for some Intelligence and Attainment Tests*, Mesa Press, Chicago.

SPENCER L.M.Jr., SPENCER S.M. (1993) *Competence at Work: Models for Superior Performance*, John Wiley & Sons, New York.

VICHI M. (1988) Two-way data matrix representative syntheses of a three-way data matrix, *Statistica*, **48**: 91-106.
VICHI M. (1997) Principal Classification Analysis: a method for generating consensus dendrograms and its application to three-way data, *Computational Statistics and Data Analysis*, Springer, **27**: 311-331.
WRIGHT B. D., MASTERS G. N. (1982) *Rating Scale Analysis*, MESA Press, Chicago.
WRIGHT B.D., MOK M. (2000) Rasch models overview, *Journal of Applied Measurement*, **1**: 83-106.
ZAN S. (ed) (1988) *Logiche di azione organizzativa*, Il Mulino, Bologna.