

# Profiling and Labour Market Accessibility for the Graduates in Economics at Naples University

Simona Balbi, Maria Gabriella Grassia<sup>1</sup>

*Department of Mathematics and Statistics, University of Naples “Federico II”, Italy*

**Summary.** In this paper, after defining a pseudo-panel of groups observed at subsequent times, we propose a strategy for the construction of a set of association rules related to different survey occasions. First, we measure the similarity between systems built at different times for understanding the stability of the phenomenon. We apply a procedure developed for symbolic data analysis for this purpose. The procedure consists of two phases: the definition of the pseudo-panel and that of a system of rules referred to the semantic marking technique. Then, the agreement between the systems is measured. We applied such a strategy for studying the labour market accessibility for graduate in Economics, the University of Naples “Federico II”, and the market evolution during an eight-year time span.

**Keywords:** Semantic marking technique; Pseudo-panel; Association rule; Symbolic objects.

## 1. The pseudo-panel definition

The Faculty of Economics at the University of Naples “Federico II” has been carrying out for over twenty years recurrent sample surveys on its graduates in order to evaluate their labour market accessibility. We will examine the last three surveys for evaluating the evolution of the phenomenon and apply a coherent policy.

The questionnaire and the survey methods for these surveys are constant in time. Therefore, it is possible to examine the evolution of the graduates’ be-

---

<sup>1</sup> This paper is the result of the joint effort of the two authors. S. Balbi was responsible for the final editing of Sections 1, 6 and 7, whereas M.G. Grassia was responsible for the other Sections.

haviours and destiny by constructing a *pseudo-panel* formed by *cohorts*, that is to say, sets of individuals identified according to characteristics that do not vary in time according to the studied phenomenon.

The analyses will be carried out on higher order units, formed by aggregating the elementary units, which are present at each survey occasion. For instance, it is possible to study if the selection devices used for graduate women who have obtained the maximum final score have changed in a given period.

The literature on symbolic objects has produced statistical methods for the analysis of complex structures. The complexity relates both to the characteristics of the units and the membership relationship linking each elementary unit to its own object (Section 2).

In the following, we put forward a strategy that, taking an advantage from the tools developed within the analysis of symbolic objects, makes a pseudo-panel approach feasible in the described context. At each survey occasion, the proposed strategy defines the constitutive elements of a pseudo-panel in terms of association between descriptors.

We propose a data-driven strategy suitable to set up a pseudo-panel according to the data association structure. The rules (called *implication rules* or *logical rules* “if-then”) will be referred to the survey waves and associated to measures of the rule authenticity.

The comparison between the rules may give a measure of the structural stability of the phenomenon. The rules refer to a symbolic data analysis frame, the *symbolic marking*, and the comparison between rules will be carried out by means of a similarity measure between symbolic objects.

## 2. The symbolic objects

A symbolic object,  $s$ , is defined as a triplet:

$$s = (a, R, d)$$

where:  $d=(d_1, \dots, d_j, \dots, d_p)$  is a set of values on  $p$  descriptors,  $Y=(Y_1, \dots, Y_j, \dots, Y_p)$ , of the object,

$a$  is a *recognition function*,

$R=(R_1, \dots, R_j, \dots, R_p)$  is the type of relation applied for the comparison between the description provided at a conceptual level, in *intention*, from  $d$ , and an observation.

The descriptors of a symbolic object can be on a nominal, continuous or discrete scale and can have several categories for each object. The Boolean function  $a$  has categories *true* and *false* and identifies those elements which belong to the  $d$  description set and that are the *extension* of the  $s$  object,  $ext(s)$ .