Can elearning platform use augment the statistical learning? Some evidence from Italy

Emma Zavarrone*
IULM University, Milan, Italy - emma.zavarrone@iulm.it

Corrado Crocetta
University of Foggia, Foggia, Italy - corrado.crocetta@gmail.com

Maria Gabriella Grassia
University “Federico II”, Naples, Italy - mgrassia@unina.it

Abstract

The study of statistics, in degree programs not specifically designed, is one of the most difficult obstacles to overcome in the students' perspective. Italian students have a superficial knowledge of mathematics and they tend to assimilate the statistics to the mathematics so they develop prejudices towards statistics too. In the last time the PISA-OCSE surveys indicate a slight improvement in the understanding of mathematics at the level of high school, however, we are still far from the European average. The statistical literacy can be considered as latent construct measured through observed repeated variables linked to the passed exams. In detail, the analysis uses the scores generated from e-learning tool (MathXL) in 10 repeated occasions for each student conjointly to demographic and social aspects. The analysis of the statistical literacy growth carries on three cohortes of students in an Italian humanistic university for the academic years 2012-2013, 2013-2014 and 2014-2015 (respectively \( n_{2012-2013} = 378 \), \( n_{2013-2014} = 402 \) and \( n_{2014-2015} = 450 \)) in students perspective. The statistical learning growth is investigated both through nonlinear latent growth modeling and latent state trait analysis. This choice comes from the rejection of linear trajectories previously tested and confirmed in the literature. The results for former two academic years highlight a nonlinear growth with an effect of voto di diploma. The same results have been confirmed by latent state trait analysis. For the last academic year, the results can be available after ending of the course on April 2015.

Keywords: latent growth model; latent state trait analysis; Gompertz curve.

1. Introduction

Over the last two decades, the lively debate around the definition of statistical literacy did not concern only the unambiguous definition of the term (Wallman, 1993; Watson, 199; Garfield, 1999; Snell, 1999; Gal, 2000; Rumsey, 2002; Garfield and BenZvi, 2008). As Milo Schield states (2012), the statistical literacy is not as common as statistical reasoning or statistical thinking but it has a 50th year history among statisticians (Schield, 2012). Firstly, the importance of a consistence definition of statistical literacy was discussed during the First International Workshop on Statistical Reasoning, Thinking, and Literacy (Israel, July 1999) whose authors main contributes were collected in the popular publication (Ben-Zvi, Garfield, 2004; Garfield, delMas, and Chance, 2003). According to scholars (Watson, 1997; Gal, 2000; Watson and Moritz, 2000b; Jones, 2000, Watson and Callingham, 2003) a definition of statistical literacy goals was formalized. Garfield, delMas, and Chance (2003) suggested a three-tiered framework based on the difference among: statistical literacy (the ability to organize, to manage, to describe and to represent statistical data and the knowledge of statistical concept), statistical reasoning (the way people reason with statistical ideas, make sense of statistical information and explain statistical process) and statistical thinking (the ability to understand the process of statistical investigation and to critique and evaluate results of a problem solved or a statistical study). During the last years, the statistical literacy assessment has become a multidisciplinary area of interest due to the daily life implications of statistical data knowledge. As Cigerenzer (2003, 2008) states: Numbers are public, but the public is not generally numerate. Statistical literacy is a necessary precondition for an educated citizenship in a technological democracy. In this paper following Garfield et al. (2003) approach, the statistical literacy has been considered as a latent variable that grows with nonlinear rate than linear ones. The Section 2 deals with the methodology, the Section 3 presents the data and in the last Section principal
results will be shown.

2. Methodology
In order to promote statistical literacy among publics a great number of researches about its assessments have been produced. Researches about students statistical literacy attitudes (I Gal, L Ginsburg, 1994; Gal, Ginsburg, and Schau, 1997; Gal, 2002; J Garfield, Ben-Zvi, 2007; Martinez-Dawson, R, 2001; Schau, 2003; Carmichael et.al., 2009; Carmichael, 2010; Calderia and Mourio, 2010) point out that middle school students attitudes towards statistics varied between neutral and positive while in the university contest it can be detected a positive relationship between constructivist learning environment (Mvududu, 2003), collaborative projects (North, Zewotir and Gal, 2014) and attitudes towards statistics. Few researchers have explored the statistical literacy growth and would seem that does not exist an analysis of statistical literacy with latent curve growth approach. The development of change models has interested many scholars, we group them into three main approaches:

- **LCM, Latent Curve Model:** the manifested change is the consequence of a latent variable change:

  \[ Y = \Lambda \eta + \varepsilon \]

  \( Y \) is the vector of repeated measure, \( \Lambda \) is the matrix of factor loadings, \( \eta \) is the vector of latent factors and \( \varepsilon \) is the vector of residuals. Autoregressive latent models and factor mixture models can be considered as development of this approach.

- **LCS, Latent Change Score:** the focus is the change occurred between two time splits. Considering the observed score at time \( t \) (\( Y_t \)) as the sum of a latent true score (\( \eta_t \)) plus an error (\( \varepsilon_t \)), we can model the latent true score with an autoregressive structure:

  \[ \eta_t = \eta_{t-1} + \Delta_{t-1}. \]

  \( \Delta \) is the occurred change that could not be directly measured, so it represents a latent variable (McArdle 2009).

- **LST, latent state-trait theory:** this approach is based on a second order factorial model (Geiser 2013). The first order factorial variables are represented by state-latent variables, the second order ones are the latent growth components. The first step is to estimate the presence of the latent states; subsequently the growth components. The use of contrasts represents a new choice in the estimate of the growth components.

This analysis compares the first approach with the last one.

3. Data
This study is based both on data provided by the statistical office of IULM University for the demo-educational aspects (gender, date of birth, kind of secondary schools and bachelor’s mark) and on the elearning scores and marks of the exam in Statistics in a cross sectional approach. Three different academic years have been investigated: 2012-13; 2013-14; 2014-15 and for the last year we add an analysis on collaborative relationship among the students. Table 1 supplies a picture of the population.

The passed exams in each academic year and the students with a platform score , since the year of adoption of the platform has increased both exams is a gradual accession to the platform in terms voting average. Table 2 shows an increase of the mean vote across the statistic in the last three years.

4. Results

The score obtained by the MathXL platform becomes a proxy of the amount of statistical literacy that students can acquire over the 11 weeks that is providing the course . We follow the path or accumulation of this proxy in which each line represents a student. The principal results show that the statistical learning growth cannot be explained through linear model as confirmed by low level of RMSEA (respectively \( \text{RMSEA}_{2012-2013} = 0.14 \), \( \text{RMSEA}_{2013-2014} = 0.10 \) and \( \text{RMSEA}_{2014-2015} = \text{n.a.} \)) and by the graphical representation. A nonlinear modeling could be adopt, in details we will inclined to use a Gompertz approach s than the exponential ones .

At the present, for the last year the course is not finished yet, but will be soon.

5. Conclusions
Figure 1: Statistical learning growth, 2012

Figure 2: Statistical learning growth, 2013

Figure 3: Statistical learning growth, 2014
<table>
<thead>
<tr>
<th>A.Y.</th>
<th>% women</th>
<th>Mean Mark’ Secondary Schools (st.dev.)</th>
<th>Type of secondary schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>69.3</td>
<td>76.1 (12.49)</td>
<td>Sc. 22%; Gram. 12%; Lang. 17%</td>
</tr>
<tr>
<td>2009</td>
<td>73.8</td>
<td>75.36(11.21)</td>
<td>Sc. 27%; Gram. 11%; Lang. 20%</td>
</tr>
<tr>
<td>2010</td>
<td>75.7</td>
<td>74.22(11.16)</td>
<td>Sc. 33%; Gram.. 13%; Lang. 22%</td>
</tr>
<tr>
<td>2011</td>
<td>73.6</td>
<td>74.05(9.81)</td>
<td>Sc. 39%; Gram. 16%; Lang. 13%</td>
</tr>
<tr>
<td>2012</td>
<td>74.5</td>
<td>73.57(10.30)</td>
<td>Sc. 32%; Gram. 14%; Lang. 16%</td>
</tr>
<tr>
<td>2013</td>
<td>68.8</td>
<td>72.61 (10.39)</td>
<td>Sc. 29%; Gram. 11%; Lang. 19%</td>
</tr>
<tr>
<td>2014</td>
<td>67.4</td>
<td>73.35 (10.16 )</td>
<td>Sc. 30%; Gram. 12%; Lang.. 20%</td>
</tr>
</tbody>
</table>

Sc. stands for high school science; Gram. stands for grammar school and Lang. stands for high school language

Table 1: Socio-demographic characteristics of the students

<table>
<thead>
<tr>
<th>A. Y.</th>
<th>Passed exams</th>
<th>%</th>
<th>Students in MathXL</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>404</td>
<td>17.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>443</td>
<td>19.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>280</td>
<td>12.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>212</td>
<td>9.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>294</td>
<td>12.7</td>
<td>198(67%)</td>
<td>24</td>
</tr>
<tr>
<td>2013</td>
<td>378</td>
<td>16.3</td>
<td>341(90%)</td>
<td>46</td>
</tr>
<tr>
<td>2014</td>
<td>304</td>
<td>13.1</td>
<td>273(90%)</td>
<td>33</td>
</tr>
<tr>
<td>Totale</td>
<td>2315</td>
<td>100</td>
<td>812</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2: Academic years, marks and MathXL use

The statistical learning can be estimated from an elearning scores in a longitudinal approach. The growth rate for the academic years 2012-13 and 2013-14 has a nonlinear trend, Gompertz function. However the statistical learning in a little pieces of time not able to growth.

References


North, D., Gal, I., Zewotir, T. Building Capacity for Developing Statistical Literacy in a Developing Country: Lessons Learned from an Intervention 2014 (Journal Articles; Reports - Evaluative).


