AN INTEGRATIVE WORK IN STATISTICAL CLASSES

Diana R Kohan[1], Teresita Terán[2] and Marisa Battisti[1]
National University of Entre Ríos, Faculty of Engineering, Argentina
National University of Rosario, Faculty of Veterinary Sciences, Argentina
dikohan@ingenieria.uner.edu.ar

Statistics is dictated in careers such as Engineering, Bioengineering, Veterinary, Economic Sciences. It is often observed that student present difficulties when trying to apply contents of Statistics to real problems. We proposed to incorporate an integrative work from a real database during the course, in which the students must analyse topics already developed in class to carry out an analysis whose stages are parts of a whole in an application problem. As in the digital era is necessary the insertion of university culture, the task is to provide students with a first contact with different software available for data analysis. It is intended to enrich the formative and cognitive process of the student, allowing them to acquire skills inherent to statistical reasoning with critical spirit.

INTRODUCTION

Statistics is dictated in careers such as Engineering, Medicine, Veterinary, Economic Sciences, among others. As a teaching team with wide experience in this subject, it is often observed that, at the end of the course, students have difficulties when trying to apply the contents of Statistics to real problems in the areas that concern them.

As an attempt to solve this issue, we proposed a practical, integrative work whose stages are parts of a whole in an application problem, in which students should analyze topics already developed in class.

Our society needs the solid insertion of the university culture in the digital age (Castañeda et al., 2010), therefore it is believed that to appeal to different forms of teaching using new technological resources can be beneficial for students.

In this sense, the proposal consists in the application of contents within the framework of a real statistical problem, in a way that students are given a first contact with some of the various statistical software packages to carry out data analysis (Godino, 1995), teaching them to draw reliable and statistically valid conclusions. The integrative work has the purpose of enrich the students' learning process, contributing to their statistical training.

THEORETICAL FRAMEWORK

The idea behind this type of proposal is to provide students with a context and an environment (Castañeda et al., 2010) so that they can achieve an understanding and learn the basic concepts of Statistics. This discipline is understood as a way of thinking and not as a way of doing, because it helps to solve real problems (Brousseau, 1986). In fact, teaching must start with situations in which students must develop and plan ways of solving problems, analyzing data, checking initial hypotheses and undertaking the decision-making process accordingly.

We agree with Godino (1995), who posits that the starting point of statistics should be students coming face to face with real data, as it is believed that the right selection of problems can introduce the student to situations of uncertainty about which you should make inferences and decide from a database.

The possibility and the need to use the computer in the processes of teaching statistics at undergraduate level are known, being incorporated as a didactic resource, so that student understands it as an instrument of calculation and graphic representation in the exploration of a set of data (Godino, 1995).

The use of computers in the classroom offers a number of benefits, as Batanero (2001) points out. It is clear that the nature of computers themselves, their dynamics and speed, and the increasing availability of a vast range of software are features that bring every aspect of statistical processes closer to the learners since, in Batanero's words, the latter might experience and explore every aspect of statistical processes, ranging from the sample planning or experimental design, to

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

data collection, management, simulation, and analysis, so that they can interpret and communicate results.

The use of computational tools facilitates statistical analysis, by eliminating the need for manual calculation in large volumes of data; thus, it is assumed that the implementation of statistical software will largely allow students to spend more time on the interpretation of the results obtained.

There are diverse types of software packages that may be useful in teaching Statistics, such as SPSS, R Commander, Infostat, Statgraphics, etc., and generic software, such as the spreadsheet (Batanero, 2001). From this point of view, the need for teachers to be trained in the use of ICTs in Statistics to enable them to provide meaningful learning environments to their students (Marelli et al., 2016) is clear.

It is considered relevant to conceive and manage a training process that encourages the development of various transversal competences, which allow students, as future professionals, to solve those problems that they will face in their work environment. In this conception, university teachers are called to think about methodologies to critically and reflexively educate young people, and it is a challenge to design different learning situations (Buscá et al., 2010).

Because every change in teaching practices involves re-thinking the evaluative practice, it is called the *formative assessment*. The objective is to improve the teaching-learning processes, which will enable students to learn more and the teaching staff to improve their pedagogical practice (López Pastor, 2012). Through it, it is possible to make value judgments that describe and interpret the evolution of knowledge and learning, counting on indicators that qualify in quantitative and qualitative ways (Peñaloza Figueroa & Vargas Pérez, 2006). Adhering to López Pastor, we agree that: "a fundamental aspect is the use of evaluation as a strategy to improve and favor learning, rather than as a simple certification of success or failure in them" (López Pastor, 2012). We share the idea that this formative character in the evaluation process lies in other concepts such as alternative assessment, authentic evaluation and learning-oriented assessment.

METHODOLOGY

Generally, a course of Statistics is developed in a semester and a Practical Integrative Work (PIW) is developed during the course, the tacks being available at the beginning of the term in the campus that the course has available on the virtual platform Moodle. These will be based on a pre-existing database with results of clinical analysis of a large number of patients.

This is an extra-curricular work that students do in groups of no more than four members, all having had a first introductory class to the statistical programs and face-to-face discussion, and others through the virtual platform.

The initial class is centered on the managing of input of data to various software and the handling of the main commands for its treatment and the elaboration of graphs.

Each group should analyze the variables assigned to them on the database based on other variables such as, for example, age and gender. Students will perform their grouped practices in three stages.

At all stages of this work, students are expected to use software such as a spreadsheet, R Commander, SPSS, Epidat, Infostat, or similar. Likewise, it should not be overlooked that the focus is placed not merely on the technical aspects of the different statistical programs, but on the correct report of the results of the research.

Stages

- In the *first stage*, concepts of population and sample are introduced, explaining sampling in finite populations, random sample, and different methods of selecting a sample based on high volume data. These techniques allow students to study the behavior of a variable to be able to develop conclusions about the population of interest.

  Simultaneously, students are directed to the study of descriptive statistics so that they can extract information of characteristics, present the results, showing the different statistical graphs and different measures of central tendency and of dispersion, emphasizing all the interpretations.

- In the second stage, probabilistic distributions are approached. Under this theme, and assuming that the probability distribution of some of the variables under study follows a known empirical distribution, such as Binomial, Poisson, Gaussian, among others, students are induced to calculate probabilities that quantify in some way, the possibilities of finding a certain behavior of the variable under study. Following this line, questions are developed to locate students in experimental situations such as the Bernoulli process, applying the contents of distributions such as Binomial, Negative Binomial, or Geometric distribution. With the already presented concepts of population and sample, it is propitious to approach the notions of parameters and statistics, and even more, the sample distributions of statistics as the sample mean, placing students in situations that lead them to determine the probability of the average of a certain variable belongs to a certain interval. Then, it is necessary to guide students to estimate parameters, mentioning point estimates and confidence intervals, doing the practice with software to verify what happen with the estimation error when the confidence coefficient increases or decreases, or decide what size the sample should be in order to obtain a certain magnitude of error.
- The *last stage* refers to the contrast of statistical hypotheses. Having worked with random samples, along this final stage students are able to answer questions such as: Is such a parameter equal to a certain known numerical value? Are the means the same in different genders? The available data come from a Gaussian population? Is there a linear association between two of the studied variables? Taking back the assumptions made for the calculation of probabilities, the student must verify if the collected data come from the hypothesized specific theoretical distribution, using the corresponding statistical contrasts, particularly the test of goodness of fit.

Evaluation of the proposal

The PIW is evaluated throughout the semester in two instances. The resolutions corresponding to the first and second stage must be presented in the form of a written report for correction, one for each working group, on dates scheduled in a timely manner. The evaluation methodology for each one is a checklist integrated by a series of criteria that will show the teacher if each group has reached the comprehension of the contents.

The criteria to be examined include the ability of the members to analyze each question formulated, to evaluate the potentials solutions they propose, the degree of linkage they establish between the instructions, and the concepts and properties exhibited in the theoretical classes, the order and clarity in the formulation of the presented results, and the justification of the procedures carried out accompanied by the complete interpretation of their findings in computer.

The third stage is evaluated in a final colloquium during the last week of the course, in which all the groups participate. Here, the objectives of the evaluation are the mastery of statistical terminology and the ability to communicate the drafted conclusions in a socialization setting.

This PIW is conceived as a regularization requirement to enrich the learning process of students, contributing to their statistical training.

Currently, to regularize the course students must also meet a minimum percentage of attendance at theoretical and practical classes, and the approval of two theoretical-practical exams with a percentage of at least 50 percent in each test, with the possibility of retaking only one of them.

RESULTS

The experience was implemented in both semesters of 2017 in the Bioengineering career of the Faculty of Engineering at the National University of Entre Ríos. The results showed that there was an increase in the percentage of students who regularized the course with respect to all those who attended (not including those students who enrolled and did not attend any class) as shown in table 1.

Table 1. Percentage of students that regularized the course in 2016 and 2017

	Semester	
Year	*First*	*Second*
2016	62,5	71
2017	100	82

From the written reports it was appreciated that although it was requested to obtain random samples using different selection methods, these were not always extracted with methods that guaranteed the randomness. Other times, the results were not obtained by using statistical software, but the sections were resolved manually. In all these situations, the groups were asked to correct and resubmit the paper. In the stage of the integrating colloquium, it was observed that although students were familiar with different statistical software packages, they showed some difficulties when interpreting the results obtained. However, as this instance was a socialization of results and self-evaluation, a favorable change in the performance of students in the second partial test was noted.

DISCUSSION

During the development of the proposal, students acquired skills inherent in the communication of statistical results, qualities that as future professionals may safely apply in research of their professional field.

The evaluation of the stages allowed for a continuous monitoring of the learning processes of students, being able to detect difficulties in the appropriation of the knowledge.

The instance of the integrating colloquium allowed an exchange of point of views between the groups, contributing a space for the clarification of the own concepts of the Statistical Inference, contents that were evaluated in the second partial test.

Students overcame the obstacles detected in initial stages, such as sampling, probabilistic distributions, among others.

This work attempts to enrich students' formative and cognitive processes, allowing them to acquire skills inherent to statistical reasoning and promoting a critical spirit, which is a desirable competence in any future professional.

REFERENCES

Batanero, C. (2001). *Didáctica de la Estadística*. Grupo de Investigación en Educación Estadística.

Brousseau, G. (1986). Fondements et méthodes de la didactiques des mathématiques. *Recherches en Didactique des Mathématiques, 7*(2), 33-115.

Buscá, F., Pintor, P., Martínez L. & Peire, T. (2010). Sistemas y procedimientos de Evaluación Formativa en docencia universitaria: resultados de 34 casos aplicados durante el curso académico 2007-2008. *Estudios sobre Educación, 18,* 255-276.

Castañeda, M. B., Cabrera, A., Navarro, Y. & De Vries, W. (2010). *Procesamiento de datos y análisis estadísticos utilizando SPSS*. Edipucrs.

Godino, J. (1995). ¿Qué aportan los ordenadores a la enseñanza y aprendizaje de la Estadística? *UNO, 5*,45-56.

López Pastor, V. M. (2012). Evaluación formativa y compartida en la universidad: clarificación de conceptos y propuestas de intervención desde la Red Interuniversitaria de Evaluación Formativa. *Psychology, Society, & Education*, *4*(1), 117-130.

Marelli, C. & Fernández, J. M. (2016). Importancia del software estadístico en la enseñanza y aprendizaje en la Universidad de Carabobo (Venezuela). Aula de encuentro, Vol. 1, No. 16, pp. 89-102.

Peñaloza Figueroa, J. L. & Vargas Pérez, C. G. (2006). ¿Qué debe cambiar en el aprendizaje de la estadística en las ciencias del comportamiento? In *Actas de las XIV Jornadas de ASEPUMA y II Encuentro Internacional*.