STRENGTHENING THE UNDERSTANDING OF SAMPLING DISTRIBUTION AND ERRORS ASSOCIATED TO THE HYPOTHESIS TESTING IN STUDENTS OF AGRICULTURAL AND ENVIRONMENTAL SCIENCES

María Virginia López, María del Carmen Fabrizio and María Cristina Plencovich
Facultad de Agronomía, Universidad de Buenos Aires, Argentina
mvlopez@agro.uba.ar

The concept of sampling distribution is not easily understood by university students of agricultural and environmental sciences. Although students can define correctly the errors involved in hypothesis testing, they show lack of a deep understanding of the subject when they have to draw conclusions. To strengthen the learning of these concepts, a practical assignment was planned in the second course of Statistics. Starting from an experimental situation, students had to design their own experiments by randomizing four varieties of canola in twenty plots. In the following class, they received the results of the «virtual sowing», with data provided by the teaching assistant. The data were simulated from normal probability distributions. Students analyzed their own data and the variability of results was commented. It was also discussed that a type II error does not imply a mistake. This study discusses the results of the teaching and learning experience.

INTRODUCTION

In our experience of teaching Statistics for university non-mathematical students, we have been able to state that there are some concepts which are difficult to grasp in depth (López et al., 2009). Despite students being able to use a method of multiple comparisons of means, it is hard for them to understand that the sample mean has a known probability distribution that enables to make inferences even from a unique experiment provided it is correctly randomised. We have been working hard in the past years looking for techniques that help students to understand these concepts intuitively. Many authors have worked on the same issue using simulation exercises designed to convey the concepts of sampling distributions and sampling variability. Vaughan (2003) classified these exercises in two categories. The first category refers to a real experimental setting, in which students have to take a randomised sample which directly involves them (Rossman & Chance, 1999; Gourgey, 2000). The second category refers to an exercise in which a computing software creates pseudo-randomised observations, allowing students to see the resulting sampling distribution of various summary or test statistics (Schwarz & Sutherland, 1997; Anderson-Cook, 1999; delMas, Garfield, & Chance, 1999).

The first category has as a drawback the fact that the exercise is generally limited to binomial or multinomial sampling from finite populations, while the sample size and number of samples are inherently constrained by the available time. Despite the fact that the second category allows a number of samples of any size to be "drawn" from a much broader collection of underlying distributions, students play a passive role, basically "watching" the demonstration as they experiment with different sample sizes and alternative source populations.

This study presents an alternative experience. Students were assigned different treatments to experimental units in a problem, according to their own randomisation pattern, obtaining results corresponding to simulated probablity distributions. This experience was underpinned by the idea that when getting different results in the analysis by different randomisation, the concept of probability distribution of estimators could be incorporated together with the idea of committing errors in decision-making in a hypothesis test.

Students in the University of Buenos Aires of Agricultural or Environmental Sciences Programs have to take two Statistics courses in their studies: General Statistics and Statistical Models. The former has a workload of 80 hours, encompassing descriptive statistics, probability and inferential statistics, including comparison of two means, simple linear regression and contingency table analysis. The latter, a 48-hour course, has the focus mainly on the design of agronomic experiments. Students learn basic concepts of one and two-way analysis of variance and multiple regression. In both courses classes are theoretical and practical. Theoretical classes are crowded, whereas practical classes have less than 35 students.

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php [© 2010 ISI/IASE]

This study seeks to describe a learning experience in which the students of the Statistical Models Course had an active role in the design and analysis of an agricultural experience to strengthen the concept of estimator variability.

METHODS

The experience was carried out by 31 students of the Statistical Models Course. Each student was in charge of a virtual trial and had to analyse its results. In the first practical class, they were handed a sheet with a sketch (see Appendix) of an experimental field with 20 numbered plots. Each of them had to randomise four canola cultivars, using random number tables or personal calculators. This process was monitored by the teacher assistant. At the end of the class, they handed in their sheets. No design was repeated. Teachers had previously lectured about the principles of the experimental design and reviewed the concepts of variability and basic statistical concepts- such as estimation, hypothesis testing and Type I and II errors- and stressed the need for a correct randomisation in order to get valid results. In the following practical class, teachers gave students the results of their "virtual sowing" in kg/ha. Results were obtained by data simulation of 4 normal distributions, having the same variance, but only two cultivars with the same mean (Table 1).

Table 1. Parameters of the Normal Distributions in Kg/ha used for simulation

Cultivar	Mean	Standard Deviation
Zafiro	3005	309.63
Master	3005	309.63
Iciola	2492	309.63
SRM C 2654	2231	309.63

This class was held at the Computing Centre. Students had to organize, enter and analyse data, working first with descriptive measures and then with analysis of variance and multiple comparison tests and assessing model assumptions. Results were discussed in class.

RESULTS AND DISCUSSION

Students were very interested in carrying out this experience and were able to randomise correctly. In the Computing Centre, students participated intensively and reported their results individually. In five out of 31 random patterns, the hypothesis that the mean was the same for all populations was not rejected. When students were informed about the true means, they grasped the idea that the non-rejection of the hypothesis was an error. Some of the students asked what to do in order not "to make mistakes" and also wondered why they had made such mistakes. These comments helped to discuss the real meaning of possible errors in decision-making based on hypothesis testing. Furthermore, the different results of the multiple comparison test (Tukey test) showed that, even if no mistakes are made in calculations or randomisation, results are obtained- which may not be actually true- varying according to different randomisation. Thus, students observed that in some data sets Zafiro average yield was higher than Master, whereas in other ones the order was changed. The concept of probability distribution of estimators- that had been studied by students in a theoretical way- was deeply understood through their classmates´ results. In this case, it was relevant to speak about power and analyse why no students had been able to detect significant differences between Iciola and SRM C 2654 cultivars. This gave way to a discussion about how to increase the power in trials by increasing the number of repetitions if it is desirable to detect small differences in means, as well as decreasing the experimental error by identifying and controlling possible variation sources. Different numbers of repetitions were calculated for different power levels.

REFERENCES

Anderson-Cook, C. M. (1999). An In-Class Demonstration to Help Students Understand Confidence Intervals. *Journal of Statistics Education*, *7*(3). Online: www.amstat.org/publications/jse/secure/v7n3/anderson-cook.cfm.

delMas, R. C., Garfield, J., & Chance, B. L. (1999). A Model of Classroom Research in Action: Developing Simulation Activities to Improve Students' Statistical Reasoning. *Journal of Statistics Education*, *7*(3). Online: www.amstat.org/publications/jse/secure/v7n3/delmas.cfm.

Gourgey, A. F. (2000). A Classroom Simulation Based on Political Polling To Help Students Understand Sampling Distributions. *Journal of Statistics Education*, *8*(3). (www.amstat.org/publications/jse/secure/v8n3/gourgey.cfm).

López, M. V., Fabrizio, M. C., & Plencovich, M. C. (2009). Evaluation of design and variability concepts among students of agriculture. In *Assessment Methods in Statistical Education: An International Perspective* (VISA Project). Wiley & Sons. In press.

Rossman, A., & Chance, B. (1999). Teaching the Reasoning of Statistical Inference: A Top Ten List. *The College Mathematics Journal*, *30*(4), 297-305.

Schwarz, C. J., & Sutherland, J. (1997). An On-Line Workshop Using a Simple Capture-Recapture Experiment to Illustrate the Concepts of a Sampling Distribution. *Journal of Statistics Education*, *5*(1). Online: www.amstat.org/publications/jse/secure/v5n1/schwarz.html.

Vaughan, T. S. (2003). Teaching Statistical Concepts With Student-Specific Datasets. *Journal of Statistics Education*, *14*(1). Online: www.amstat.org/publications/jse/v11n1/vaughan.html.

APPENDIX

Last and First Names: ……..

"Canola" (rapeseed) or "00" rapeseed represents an option to diversify winter-spring agricultural production for farmers in the central area of Santa Fe province. It has agro-climatic characteristics similar to wheat and a base price of about T U$S 225/ton. Besides, its sowing dates are earlier than or similar to wheat. If cultivars are selected and managed properly, they may leave room for different rotation sequences in the field. Due to biofuel production, canola production has been increased in the past years since it complies with European quality standards.

In order to quantify the differences in the cultivar potential yield, you are hired to carry out a trial in a field in Santa Fe. You have decided to compare 4 cultivars (T1: Zafiro, T2: Master; T3: Iciola 41; T4: SRM C 2654) sown in a plot of uniform fertility and moisture. The experimental field with 20 plots is shown below:

1	2	3	4	5
10	9	8	7	6
11	12	13	14	15
20	19	18	17	16

a) Distribute the cultivars in the plots using the concepts you have learned in Statistical Models.
b) Wait for the harvest and then analyse the data. Draw some conclusions and recommend the most convenient cultivar/s for the area.