POSING PROBLEMS AS A STATISTICS TEACHING ACTIVITY AT ENGINEERING UNDERGRADUATE AND GRADUATE LEVELS

Patricia E. Balderas Cañas and Ana María Vázquez Vargas
Universidad Nacional Autónoma de México, México
empatbal@servidor.unam.mx

In this work we pose some problems related with data analysis, test of hypothesis, and making decisions. Our objective in posing stochastic problems is promoting connections among some important stochastic concepts, in order that students would be able to apply statistical procedures when they face real problems of data analysis. We are concerned with design of some key situations and problems for conducting statistics teaching to promote statistics learning trough case analysis and project development.

INTRODUCTION

A major demand from the engineering world is making decision from data collection and that is a motive to teach statistics courses as a service subject in many courses. From labor world there is a great requirement for graduate engineers to have statistical tools literacy, so this point cannot be missed from statistics courses (Wilson, 2002; Vardeman, 2002) Our concern is based on Schau and Mattern (1997; cited in Batanero, 2004) recognition of a student's lack understanding of connections among important stochastic concepts.

Below, we first introduce some theoretical elements, then we analyze the relevance of posing open problems and working with project and case studies to overcome some common learning difficulties in statistics and to relate statistical and calculus concepts. Next we describe some problems we have posed to engineering undergraduate and graduate students and that are useful to this purpose and finally include some reflections.

THEORETICAL ELEMENTS

Focusing on conceptual understanding in mathematics learning categorized by Hiebert and Lefevre (1986), and in particular at the procedural level, students would know how and when use an algorithm, and which requires students' acquisition of necessary connections to other concepts of the same or other area. Relative frequencies for example, involved proportional thinking that is studied in statistics and calculus courses (Balderas, 2005). So, a curriculum design for engineers must provide links between calculus and statistics courses, between other disciplines. And, the elementary courses set the foundations to have access to higher statistics concepts.

To promote student understanding at a generalization and abstraction levels, so that he or she would be able to acquire and apply a strong conceptual system, we think that posing problems is a key issue due to many reasons. Firstly, because "… although many statistics students are able to manipulate definitions and algorithms with apparent competence, they often lack understanding of the connections among the important concepts of the discipline they do not know what statistical procedure to apply when they face a real problem of data analysis…" (Schau and Mattern, 1997; cited in Batanero, 2004) Secondly, there is an increasingly easy access to powerful computing facilities that frees time previously devoted to laborious calculations and encourages less formal, more intuitive approaches to statistics (Biehler, 1997; Ben-Zvi, 2000). Thirdly, new course content, new teaching approach and new assessment are required due to this exchange of approaches. And, between other reasons, a recent answer to Moore's question on what skills in the traditional sense should be required (1997) done by Wild and Pfannkuch (1999). But, in spite of Batanero's recommendation (2004) to change the course content and teaching approach at University level, this is slowly taking place, and it seems neither it is produced as soon as work world needs, nor early graduate engineers have a strong statistics background. So, we are interested in designing problems close to real and professional situations that could be analyzed in these courses under the umbrella of case study and project development.

RELEVANCE OF PROBLEM SOLVING AND CASE STUDIES IN OVERCOMING LEARNING DIFFICULTIES

We realize that students have some learning difficulties, for example on linear regression, correlation coefficient and covariance. On this respect, Sánchez (1999), found some difficulties in translating between the different representations of correlation (verbal description, table, scatter plot and correlation coefficient), and Estepa and Sánchez (2001) described students' difficulties "in relating the ideas of linear regression, correlation coefficient and covariance lack of distinction between interdependence and unilateral dependence, problems in adequately choosing the dependent and independent variables; and excessive emphasis on linear dependence."

A key question for this work was suggested by Wilensky (1995 and 1997) when they asked for "…making the transition from data analysis to statistical inference". On that concern, we accept that in order to make that transition, students are usually introduced to probability distributions, with most emphasis on the normal distribution (Batanero, 2004).

Indeed, we remark that there are many misunderstandings, particularly with the level of significance, α, which is defined as the probability of rejecting a null hypothesis, given that it is true (Batanero, 2004). Accordingly with Batanero "the most common misinterpretation of this concept consists of switching the two terms in the conditional probability,…, interpreting the level of significance as the probability that the null hypothesis is true, once the decision to reject it has been taken…" (*idem*).

We think that case studies and project development would allow overcoming those misunderstandings, better than the closed items commonly settled in class, due to the use of a wide variety of representations, and because every team would choose the ways to communicate, analyze and solve questions that emerge from discussion. With this, we do not want say that closed items are not necessary, but a reduction of that emphasis and a wider focus should be better. We know that isolated types of tasks rarely reveal depth students' understanding (Estepa and Sánchez, 2001). Case studies and project development in opposition to isolated types of tasks require posing problems from a holistic research approach (Keeves and Lakomski, 1999). Posing problem in statistics education should be done with relation to other disciplines, work world tasks and research areas.

Finally, Wilensky (1995 and 1997) defined epistemological anxiety as "the feeling of confusion and indecision that students experience when faced with the different paths for solving a problem." In interviews with students and professionals with statistical knowledge Wilensky asked his students to solve a problem by using computer simulation. Although most subjects in his research could solve problems related to the normal distribution, they were unable to justify the use of the normal distribution instead of another concept or distribution, and showed a high epistemological anxiety. When students work cooperatively as they develop their project, epistemological anxiety is also reduced, and it is still present in graduate students when they perform data analysis and make statistical inferences.

PROBLEMS PROPOSED

According to statistics education research (Keeler, 1997) and peer discussion with statistics educators at undergraduate and graduate levels, we start from the recognition of the role, importance and misconceptions of Null-Hypothesis Significance-Test Procedure that have discussed by Chow (1996) and has been the focus for research and teaching strategies design (Batanero, 2000, 119-176, 320-384). So the problems we propose for statistics teaching will involved test of hypothesis and are based on case analysis and project development. Our objective with this approach is to promote that students connect important stochastic concepts that allow them to apply statistical procedures when they face real problems of data analysis.

Some problems were designed to address class projects for undergraduate engineering students in statistics course. Those problems should be developed by small teams (with 3 or 4 members), in class as a short task (as problem 1) or as extended task to be done out of class (as problems 2, 3, and 4). These problems are opened, thus we expect that students could use and relate many statistical concepts and procedures, such as descriptive analysis, random variables and distributions, sampling distributions, test of hypothesis, decision-making, correlation, simulation, etc. Problems 3 and 4 were designed for graduated students.

To design and pose these types of problems we analyzed some materials disseminated in lectures, proceedings and journals of statistics and statistics education. A first version of the third problem was used in the fall semester of 2005. One four-graduate student's team developed an inquiry of how many buses should be scheduled to attend the demand in a specific Mexican bus stop, in a workday from 8 to 10 am. The team, set a hypothesis, then collected proper data during eight workday-weeks, then did data analysis and proponed a forecasting time series model on the basis of six weeks-data, and validated the model with the rest of data. However, they did not spontaneously tested the mean absolute deviation that was part of model acceptation criteria (Levine, Stephan, Krehbiel and Berenson, 2002)

Problem 1 A supervisor of a filling process, in a main soda company, wants to know whether 300 ml bottles have been full filled, or no, and make a decision to stop the filling process.

Problem 2 A head of traffic local department in a tourist city have to decide whether to split the location of some bus stops, since they are now located in the same point, which cause space conflicts at pick hours. How should he proceed to make a decision?

Problem 3 How can we plan the amount of buses and the correspondent schedule to cover at least 80% of the largest demand, at a particular bus route on every workday?

Problem 4 There are four valves in a water treatment process, and we want to prevent some crash times so we need know in advance how to reduce crash time occurrences.

In the spring semester of 2006, we analyzed some students' answers to those problems, that we will present in the Conference.

SOME CONCLUSIVE REMARKS

We want to mention that some anxious or uncomfortable feelings when handle stochastic data seems to be a result of deterministic phenomena study as the major focus in scholar courses (Meletiou-Mavrotheris, 2002; Meletiòu-Navrotheris and Lee, 2002). Then, since no random sequence of outcomes exactly fits the expected patterns suggested by probability theory, judging randomness in a particular situation also requires an understanding of the logic of test of hypothesizing and the features of sampling distributions (Batanero, 2004). We propose that curriculum design on statistics should include recommendations on formalism, rigor and pertinence level of core concepts and methodological foundations to a better use of technology and didactical approaches.

REFERENCES

Balderas, P. (2005). *Visualización y uso de tecnología en el aula*. Memorias de la XI Jornada de investigación en el aula. SAEM THALES. Departamento de Didáctica de la Matemática. Universidad de Granada, Noviembre, 17

Batanero, C. (2004). Statistics education as a field for research and practice. Regular Lecture, ICME 10, Copenhagen.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning*, 21(1 and 2), 127-155.

Bieheler, R. (1997) Software for learning and for doing statistics. *International Statistical Review*, 65(2), 167-190.

Chow, L. S. (1996). *Statistical Significance: Rationale, Validity and Utility*. London: Sage.

Estepa, A. and Sánchez, F. T. (2001). Empirical research on the understanding of association and implications for the training of researchers. In C. Batanero (Ed.), *Training Researchers In The Use of Statistics*, (pp. 37-51). Granada, Spain: International Association for Statistical Education and International Statistical Institute.

Hiebert, J. and Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: The Case of Mathematics*, (pp. 1-27). London: Lawrence Erlbaum Associates.

Keeler, C. M. (1997). Portfolio assessment in graduate level statistics courses. In I. Gal and J.B. Garfield (Eds.), *The Assessment Challenge in Statistics Education*, (pp. 165-178). Amsterdam: IOS Press and the International Statistical Institute.

Keeves, J. P. and Lakomski, G. (Eds.) (1999). *Issues in Educational Research*. Oxford: Pergamon.

Levine, D. M., Stephan, D., Krehbiel, T. C. and Berenson, M. L. (2002). *Statistics for Managers Using Microsoft® Excel*. New Jersey: Prentice Hall.

Meletiou-Mavrotheris, M. (2002). On the formalist view of mathematics: impact on statistics instruction and learning. *Proceedings of III European Research in Mathematics Education*. http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG5/TG5_list.html

Meletiòu-Navrotheris, M. and Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistic Educations Research Journal*, 1(2), 22-37.

Mills, J. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education*, 10(1), www.amstat.org/publications/jse/v10n1/mills.html

More, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, 635, 123-165.

Sánchez, F. T. (1999). *Significado de la Correlación y Regresión para los Estudiantes Universitarios* [Meanings of Correlation and Regression for Undergraduates]. Unpublished PhD. University of Granada.

Schau, C. and Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. In I. Gal and J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education*, (pp. 91-104). Amsterdam: IOS Press and the International Statistical Institute.

Vardeman, S. (2002). Providing "real" context in statistical quality control courses for engineers. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.

Wild, C. J. and Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3).

Wilensky, U. (1995). Learning probability through building computational models. In D. Carraher and L. Meira (Eds.), *Proceedings of the 19th PME Conference*, Vol 3, (pp. 152-159). Recife, Brazil: Organizing Committee.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, 33, 171-202.

Wilson, R. (2002). What does this have to do with us? Teaching statistics to engineers. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.