STATISTICS IN APPLIED MATH PROJECT DEVELOPMENT DONE BY GRADUATE ENGINEERING STUDENTS

Patricia E Balderas Cañas
Universidad Nacional Autónoma de México
Facultad de Ingeniería
empatbal@servidor.unam.mx

ABSTRACT

The main objective of this naturalistic study (Lincoln & Guba, 1985) was to analyze the statistics used by some graduated students to reach some answers, as they developed a team class project. The study suggested that the skills and math knowledge of the students were pertinent to analyze the selected problem and to provide some answers to the set questions from their inquiry. It was remarkable that students restricted the problem as they realized that were not able to get some missing information in the period of time considered. In that sense, the students' knowledge was part of the instructional mode derived from human and empirical inquiry (Keeves & Lakomski, 1999; Wild & Pfannkuch, 1999). Those project developments turned into high cognitive demands for the students, and the main students' reasoning gain was identified by the links they done between statistics, linear programming and/or forecasting concepts in the final report.

THEORETICAL ELEMENTS

Posing problem is a key question of students' learning in order to get conceptual mathematics understanding (Balderas, 2006). In spite of "… some statistics students are able to manipulate definitions and algorithms with apparently mastery, they lack of understanding links between main concepts of this discipline and they do not know how and when a particular procedure is pertinent as they do data analysis" (Schau and Mattern, 1997; cited in Batanero, 2004).

Conceptual and procedural understandings (Hiebert and Lefevre, 1986) are present as students decide how and when use concepts and procedures. Reaching both understandings require from students to make pertinent links between involved concepts of statistics and other areas. Intuitions, heuristics and strategies are necessary in problem solving and posing, many of these are documented in elsewhere (for example Grouws, 1992; Campos and Balderas, 2001). Knowledge dispositions (Prawat, 1989) emerge as the students discuss some relevant ideas when they develop class projects. For example, relative frequencies are related with proportional thinking, and this last is part of the conceptual core of statistics and calculus (Balderas, 2005).

I assume that students must comprehend or make sense from the experience, and learn from it. Indeed, the students should transfer their knowledge and understandings in many situations in different areas like applied statistics. Data collection and the learning process from that process ideally take place in making decisions. Wild and Pfannkuch (1999) designed a framework to statistics thinking that emerged from research with professionals and users of statistics, as well as professional adviser done by statisticians. This framework is no exclusive of statistics and those authors asume thinking quality could improve if we get a better understanding of statistics (*id*); indeed, they explain the researcher thinking by means of four dimensions that operate at the same time: dimension one, the research cicle, with a main step, the planning, dimension two, where emerge several types of thinking, followed by dimension three, to critize plans intended, that is a interrogative cicle, followed by an exceptisim step of dimension four, named dispositions.

For the purposes of this work I point out that, according with Wild and Pfannkuch (*op.cit*.), it is required to have data, learn from data and make decisions under uncertainty. Those is, the statistics thinking is firstly a knowledge strongly associated to the context and with synthetize abilities, it demands pertinent communicative abilities to extract vital information from the client ideas, chiefs, administrative personnel, patients, etc. But, it is necessary overcome the just mechanical application of statistical techniques, so learn from the statistics requires all the actions involved in the four dimensions cited. In fact, dimension four that it concerns with

dispositions show us the challenge faced by statistics educators, between others me, to reach the students think statistically. At the beginning, to make the students identify problems that could be focused by means of statistics, then students should analize the problems to get a better and deph comprenhension of them, then students turn into perseverant researchers and opened minded to new ideas, and of course, that they restrict their impulse to do instantaneous conclusions.

METHODOLOGICAL ELEMENTS

This naturalistic study followed Lincoln and Guba (1985) recommendations to make interpretations of students' activities by means of final reports of class projects. So, the study was a case study with six exemplars (the six small students' teams). The main objective of this study was to analyze the statistics used by those students to reach some answers, as they developed each project. The main statistical topics used by the students are listed in the Annex.

Participants of the study were graduate engineering students enrolled in applied math courses (fall terms of 2005 and 2007); they work in small teams, three or less members, and participated in this study as volunteers. Each team identified and selected a problem, related with their labor or daily life environment, to conduit a class project in order to apply as many concepts and procedures as possible that they were previously discussed in the class. The students set some questions connected with such problems, and planned some inquiry strategies with some of the resources discussed with the whole class. Those resources mainly came from some applications of linear algebra, linear programming, decision theory, or forecast, discussed along the course.

I selected the six projects that appear in the annex. These projects allowed students to address some questions, and then design some methodologies to reach some answers. Several ideas were discussed in class as the projects were developed. Professor asked them for final reports to assign the class grade, so the grades relied on project organization, pertinence and amount of concepts and procedures included in the subject syllabus.

TEACHING DETAILS

Applied math course content was mainly oriented by students' preferences and the teacher was a developer of students' conceptual change, so the teaching goal was to reach a better development of person and her conceptions (Light and Cox, 2001, 28-43).

Teaching prompted from the students to be responsible of their own learning and the subject content were built by them too. So, a working hypothesis was that student's conceptual change was possible if the knowledge was socially constructed.

Several class sesions were implemented with computer assisted systems (Maple®9, Excell®, and TreePlan worksheets, between others). At the beginning of course, the main commands of each computer resource were illustrated in class.

The students involved in the study selected the applied math content (concepts and methods) to pose and solve some engineering problems.

Project and corresponding reports were presented by students in a special format to allow the evaluation of the following aspects, and grading them up to the maximun percent indicated between the parentheses: problem possing (15%), question possing (15%), objectives possing (15%), study design (20%), development of the study (20%), and formulation of conclusions (15%).

For this teaching approach, the grade of the final project report is part of student's learning and performance; which are broad areas of educational research for their own right, and they are discussed elsewhere (Keeler, 1997; Schau and Mattern, 1997).

RESULTS AND DISCUSSION

The students that developed Project 1 observed that was very frequently the huge vehicle traffic at some toll bridge in a main Mexican high way. They noticed that this traffic trouble could be associated with de amount of toll bars, which were functioning in each direction and each time period. They asked how to minimize labor cost due to the amount of toll bars, but optimizing this

amount to satisfy the demand as the traffic flu could require, and to reduce the accumulation of vehicle traffic in those toll bars.

These students correctly minimized labor cost and optimized the amount of toll bars. Final report of the project made by the team, provided elements to assume that the students studied the problem, and this study induced math activities (procedures) to provide an acceptable solution to some of the questions set by the team. That solution involved some statistics and linear algebra concepts, properties and procedures.

Most of the students correctly used statistical arguments, but they were no so estable; because some students used to change the random behavior of some variables as deterministic variables, and they did not realize that it was necessary conduit testing hyphotesis of the average values assumed (projects 3 and 6). I consider that those problems are part of a wider problem: how the researchers use statistics (Batanero, 2001).

Students used three sources of data: students' measurements (projects 4 and 5), institutional data (projects 1, 2, 3 and 6), and the Internet Web (projects 1, 2, and 3).

Projects 1, 2, 3 and 6 where discussed before the course as a result of work activities of some members of those teams. However, all teams reformulated few questions to use broadly statistics and linear algebra, as a result of teaching requeriments. And, of course the information was anonymously handled. Teams that developed projects 4 and 5, each one designed an experiment to collect data with the advice of teacher.

Thus, all teams looked for and had data; they learned from data and made decisions under uncertainty.

CONCLUSIONS

The study suggested that the skills and math knowledge of the students were pertinent to analyze the selected problem and to provide some answers to the set questions from their inquiry. It was remarkable that students restricted the problem as they realized that were not able to get some missing information in the period of time considered (project 1 and 5). In that sense, the students' knowledge was part of the instructional mode derived from human and empirical inquiry (Keeves & Lakomski, 1999; Wild & Pfannkuch, 1999). Those project developments turned into a high cognitive demands for the students, and the main students' reasoning gain was identified by the links they done between statistics, linear programming and/or forecasting concepts in the final report.

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.

Balderas, P. (2006) "Posing and solving linear algebra and statistics problems with Maple". In Josef Böhm (*ed.*) *Proceedings of Technology and its Integration into Mathematics Education DES-TIME-2006*. Linz, Austria: Teachware. ISBN 3-901769-74-9.

Batanero, C. (*ed.*) (2001). *Training Researchers in the Use of Statistics.* International Association for Statistical Education and International Statistical Institute. Granada: Universidad de Granada, España.

Batanero, C. (2004). *Statistics Education as a Field for Research and Practice*. Regular Lecture, ICME 10, Copenhagen.

Campos, M.A. and Balderas, P. (2001) "Las representaciones como fundamento de una didáctica de las matemáticas". Pontificia Universidad de Chile. *Pensamiento Educativo*, (27), 169-194.

Grouws, D. (1992). *Handbook of Research on Mathematics Teaching and Learning*. New York: NCTM.

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*. London: Lawrence Erlbaum Associates, Publishers, 1-27.

Keeler, C. M. (1997). “Portfolio Assessment in Graduate Level Statistics Courses”. In I. Gal y J.B. Garfield (*eds*.), *The Assessment Challenge in Statistics Education* (165-178). Amsterdam: IOS Press.

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

Wild, C.J. and Pfannkuch, M. 1999. “Statistical Thinking in Empirical Enquiry”. *International Statistical Review*, 67 (3), 223-265.

Light, G. and Cox, R. 2001. *Learning and Teaching in Higher Education: The Reflective Professional*. London: Sage Publications.

Lincoln, Y and Guba, E (1985) Naturalistic Inquiry. Newbury Park: Sage Publications.

Prawat, R.S. (1989). “Promoting access to knowledge, strategy, and disposition in students: a research synthesis”. *Review of Educational Research, 59* (1), 1-41.

Schau, C. and Mattern, N. (1997). Assessing students’ connected understanding of statistical relationships. In I. Gal y J.B. Garfield (*eds*.), *The Assessment Challenge in Statistics Education* (91-104). Amsterdam: IOS Press.

Wild, C.J. and Pfannkuch, M. (1999). “Statistical Thinking in Empirical Enquiry”. *International Statistical Review*, 67 (3), 223-265.

ANNEX

Details of the six analized projects

Project 1. Optimizing labor cost and service rate in a main toll bridge close to Mexico City (2005, fall semester)

Question: How could it be minimize personel costs, but satisfaying service demand?

Statistics concept/procedures: Data analysis, hypothesis testing, linear regression.

Statistical conclusions: Proportions provided by CAPUFE were 90% confidents. Parameters of linear programming model must be adjusted according to season times and wheter is workday or weekend.

Project 2. Analyzing chocolate production in the South of Mexico (2005, fall semester)

Question: Is it profitable the chocolate production for a particular company in the south of Mexico?

Statistical objective: Based on data chocolate production of 33 months, students made a forcast of production level for the month 34.

Statistics concept/procedures used: Data analysis, linear regression, hypothesis testing, seasonal factors, time series methods, decision trees.

Statistical conclusions: The best forecast production was 8,931.76 tons that was calculated by a movile averages method over 12 periods because it was more stable than others. (The criterium was to select the method with lower MAD, mean absolute deviation)

Project 3. Analyzing ATM service in Mexico City (2005, fall semester)

Question: How could it be calculated an average service rate of an ATM?

Statistics concept/procedures used: Data analysis, quantification of seasonal factors and trend of service rates, linear regression, quequeing lines.

Statistical conclusions: Due to missing historical information students pointed out the weakness of their study. Students recommended designing a computer program in order to get a better register of arrivals at an ATM. And, include seasonal aspect, like holydays, in the study.

Project 4. Analyzing Metro service in Mexico City (2007, fall semester)

Questions: What is the average of the daily ammount of people that use colective transport system (metro) in East terminal of the nineth line in Mexico City, from 6:00 to 9:00 hours? Why is there passager agglomeration on those hours?

Statistics concept/procedures used: Poisson distributions, relative frequencies, arrival rates.

Statistical conclusions: Daily average was 27074 people, from 6:00 to 9:00 hours. Train capacity 1479 passangers, flu rate at East terminal 7,52 users/second, requerided time to full fill the train 195.11 seconds, time period between two trains 3:15 minutes. So, if the time interval between two trains is 3:15 minutes, 1466 users could abord the train, and the agglomeration should be reduced. If time interval between two trains is longer than 3:15 minutes, some people would not abord the train and the agglomeration could be increased shorthly.

Project 5. Analyzing computer lab breaks in a main Mexican university (2007, fall semester)

Question: What is the average service break per term in a computer lab at a main university?

Statistics concept/procedures used: Data analysis.

Statistical conclusions: Due to missing historical information students pointed out the weakness of their study.

Project 6. Analyzing fridge sales in a main Mexican company (2007, fall semester)

Questions: How are the variabily and trend of commercial fridge sale profits? What is the forecast of average sale profit for the most saled products?

Statistics concept/procedures used: Data analysis.

Statistical conclusions: Variability of net profit from 243,202 to 1,146,852 (MXN), expected average 695,027 (MXN), and Standard deviation of 216,364 (MXN). Politics of the Company were a variability of net profit from 392,574 to 997,480 (MXN). Recommendations generate a line base that considers more variables that could affect net profits and look for correlations and trends along the time.