ASSESSING STATISTICAL REASONING THROUGH PROJECT WORK

Rui Pimenta
ESTSP- Instituto Politécnico do Porto, Portugal
rep@estsp.ipp.pt

New technologies involve a reformulation of contents and methodology used for teaching statistics. Developing students' statistical reasoning becomes an important task for teachers of applied statistics. This is particularly true in the field of health sciences. In this work the statistical reasoning ability acquired by health sciences students was evaluated in the context of their final undergraduate project.

INTRODUCTION

Nowadays statistics teaching is generalized in training of several professional groups, since it is considered a tool for these students both to allow them better understanding of other topics promoting critical evaluation of research works as well as giving them a methodological basis for their future research. A usual problem at advanced statistical teaching is the gap between the understanding of concepts and the technical capacity to apply them (Biehler, 1997). Nowadays, this gap is getting smaller, due to technology, which allows students to explore all steps in statistical research.

Although understanding is now easier due to simulation support, some new aims are also introduced, and there's a danger of only teaching software instead of teaching statistics. This problem led the Faculty at the Escola Superior de Tecnologia da Saúde of the Instituto Politécnico do Porto (ESTSP-IPP) to reformulate their teaching, in the health sciences area. Working with technology and projects, which are preferably interdisciplinary, and included a compulsory research project at the end of the undergraduate programme was promoted to achieve this aim. At the same time, the institution encounters another challenge, which involves the students' training in subjects such as experimental design, research methodology and data analysis.

STATISTICAL REASONING

Statistics teaching with technology support suppresses the need for learning algorithms and procedures, which are now provided by statistical software, easy to learn without a deep conceptual understanding. T*utorials, statistical advisors* and other tools diminish the relevance of conceptual learning. It is also possible to carry out many procedures without full knowledge of all concepts involved, since software is used as a *black box* that gives results, with no necessity of knowing how they are obtained. For example we can perform a factor analysis, without knowing how to invert a matrix, what is a determinant or an eigenvalue. The main difficulty in education is transmitting the inherent elements of *statistical reasoning*.

This expression is used with an increasing frequency to talk about the specific reasoning, which recognizes the presence of variability of multiple phenomenon's, and which uses analytic methods to reduce and control this variability, make decisions or carry out predictions. This kind of reasoning includes according to Wild and Pfannkuch (1999) five basic components:

- *Recognizing the need for data:* The basis of statistical method is the hypothesis that many real life situations can only be understood through the appropriate analysis of previously collected data. Personal experience or anecdotal evidence is unreliable and produce misjudgements or wrong decisions.
- *Transnumeration*: The authors use this concept to indicate the comprehension that appears when the data representation is changed. When dealing with a real system as a modelling perspective, three types of transnumeration may occur: (1) from measurement that grasp real world qualities or characteristics, (2) when transforming raw data to tabular or graphic representation, which uncover tendencies or patterns; (3) when communicating, the meaning of data in a way which is understandable to other people.
- *Perception of variation*: An appropriate data collection and judgments from the same, is based on the understanding of the variation that exists and is transmitted in the data as well as of the uncertainty originated by the unexplained variation. Statistic tries to predict causes of

variation, and learn from the context.

- *Reasoning with statistical models*: Any statistical instrument, even a simple graphic, a regression line or a statistic summary is a model, since it is used to represent reality. It is important to differentiate the data from the model and at the same time relate the model with the data.
- *Integrating statistic in the context***:** This is the main component of statistical reasoning that results from the implications and conjectures originated by statistical knowledge in a given context. Statistical thinking is only meaningful when it becomes part of the context. Applying statistical methods without considering the context is a complete nonsense. For example, the conclusions we get in health sciences may not have applications in other areas.

The above considerations led us to investigate the assessment of statistical reasoning of undergraduate students, in order to use the results to improve our teaching practice (Pimenta, 2005). Below we present, some results of this assessment; we analyzed some statistical reasoning elements acquired by students, who finished the second cycle in Physical therapy, through the assessment of their use of statistics in their final projects.

METHOD, RESULTS AND DISCUSSION

We carried out a content analysis of the projects presented by 43 students, when finishing their degree (2nd Cycle), all of who had attended a course of experimental design in our School. Half the students had started to work as soon as they finished the bachelor's degree (1st cycle of studies), of which had a varied professional experience (2-20 years) before finishing the second cycle. The majority of the participants in the study (65.1%; n=28) were women, 13 of them (30.2%) were married, and the average age was 27 (maximum of 47).

The favourite topics in the projects were quality of life, intervention in Physical therapy, assessing treatments, elderly people, and lesions inherent to a certain professional group. The average number of pages per project was 24 (ranging between 14 and 33) and the average number of variables 3.9 (between 1 and 17). The statistical analysis was usually done with SPSS (*Statistical Package for the Social Sciences*), although some students resorted to *Excel*. Nine of the projects did not include statistics, because either it was a qualitative study or a theory based project. Consequently, the results of our evaluation are restricted, because a quarter of the students' were not evaluated and we have a small sample size. In spite of this we consider this study provides a method to analyze students' projects that can be applied to other situations and it also provides some general conclusions about how to improve students training in statistical reasoning.

Below, we present the results concerning each of the five components of statistical reasoning described by Wild and Pfannkuch (1999), in the 34 projects that included some statistical analysis. For each of these components we defined a series of indicators, and analyzed its frequency. Examples of each one of the described categories can be analyzed in detail in Pimenta (2005).

- *Recognizing the Need for Data* All the projects analyzed showed that the students recognized the collecting data needed to solve their problems (Table 1). Most students described the social variables characterizing the sample with the aim of recognizing the study generalization and limitations. The majority of cases also resorted to a relevant sampling method. Only a quarter of students prepared their own questionnaire, although this was unnecessary in most of the projects.

Table 1: Frequency (and percentage) of recognizing the need for data

	Yes	No	Not applied
Recognize need for data	34(100)	0(0.0)	0(0.0)
Correct characterization of sample	29(85.3)	5(14.7)	0(0.0)
Appropriate collection of data	32(94.1)	2(5.9)	0(0.0)
Development of a questionnaire	8(23.5)	3(8.8)	23(67.6)

- *Transnumeration* We considered that the transnumeration had occurred, if the students transformed the raw data into a statistical table or graphic, or if they presented a summary of data, which was appropriate to the kind of problem, well done and well interpreted (Table 2). All these procedures served to draw conclusions that were not visible in the original data. Most of our students used tables as well as statistics summaries and graphics, though they did not always interpret their meaning correctly.

Table 2: Frequency (and percentage) of transnumeration

	Yes	No	Not applied
Correct interpretation of a table	21(61.8)	13(38.2)	0(0.0)
Summarizing key aspects with statistics	25(73.5)	9(26.5)	0(0.0)
Graphics correctly produced	21(61.8)	0(0.0)	13(38.2)
Graphics correctly interpreted	18(52.9)	3(8.0)	13(38.2)
Measures of central tendency well interpreted	24(70.6)	10(29.4)	0(0.0)
Measures of spread well interpreted	19(55.9)	0(0.0)	15(44.1)

- *Perception of Variation* The majority of students recognized the presence of variability in their data and all of them showed a numeric comprehension (Table 3). However, few of them were clearly conscious of the uncertainty in their conclusions or in the data collected. Such fact suggests the predominance of a deterministic conception and the need to reinforce their experience of randomness and their training in dealing with random phenomena.

Table 3: Frequency (and percentage) of perception of variation

	Yes	No	Not applied
Perception of variation	31(91.2)	3(8.8)	0(0.0)
Perception of uncertainty	2(5.9)	28(82.4)	4(11.8)
Numerical perception	34(100.0)	0(0.0)	0(0.0)

- *Reasoning with Statistical Models* Few students used explicitly the idea of statistical model or differentiated clearly the empirical data, from the mathematical model or the difference between the conclusions obtained through a model (which has a deterministic character) and the uncertainty associated to the real situation. Probably, this is due to the difficulty associated with a perception of uncertainty, as well as the abstraction of the modelling idea (Table 4).

In fact, the students were working with models, even when they were unconscious of its character, More precisely 97.1% of the projects, applied descriptive statistics procedures, 8 (22%) the non parametric Wilcoxon test, 7 (20%) the *t*-test for paired samples, 5(14.3%) the Chi-square test. Some more resorted to the Kruskal-Wallis or Mann-Whitney tests, or to Pearson or Spearman correlations. In same cases, students applied procedures, which had not been taught in experimental design subject, for example, intra and inter classes' correlations and the Cronbach's α. However, as it can be seen in Table 4, there were serious difficulties in applying statistical procedures, setting correctly an hypothesis test, or respecting the method assumptions. These types of difficulties were previously described in professionals of health sciences (e.g. in Altman, 2002; Altman, Goodman and Schroter, 2002).

Table 4: Frequency (and percentage) of reasoning with statistical methods

	Yes	No	Not applied
Reasoning with statistical models	1(2.9)	33(97.1)	0(0.0)
Respect the method assumptions	19 (54.3)	11(31.4)	5(14.3)
Correct use of hypothesis tests	16(47.1)	12(35.3)	6(17.6)
Correct establishment of hypotheses	26 (85.7)	5 (14.3)	34 (100)

- *Integrating Statistics in the Context* Finally, we analyzed if the students were able to interpret the statistical results in the context of their research project (Table 5). All the projects did some extend of integration. Statistics was justified and structured in the method section, as well as in the presentation and discussion of results. All the projects were correctly, documented with the description of the problem and variables, problematic and references.

Table 5: Frequency (and percentage) of integrating statistics in the context

	Yes	No	Not applied
Integration of statistics in the context	34(100.0)	0(0.0)	0(0.0)
Graphics appropriated to the variable type	20(58.9)	1(2.9)	13(38.2)
Graphics appropriated to the problem	21(61.8)	0(0.0)	13(38.2)
Tendency central measures appropriated to the problem	29(85.3)	5(14.7)	0(0.0)
Tendency central measures appropriated to variables	29(85.3)	5(14.7)	0(0.0)
Spread measures appropriated to the problem	25(73.5)	0(0.0)	9(26.5)
Spread measures appropriated to the variables	25(73.5)	0(0.0)	9(26.5)

In general, graphics and statistical measures were appropriated to both types of variables and problems, in particular in the use of descriptive statistics. Statistical inference was more problematic. Therefore, according to our data these students knew how to choose the appropriated descriptive method to its problem, and interpreted it in the context.

IMPLICATIONS FOR STATISTICS TEACHING

The study carried out, suggest these students achieved a good level of statistical reasoning, which allowed them to recognize the need for data, understand the variation in the same, carry out transnumeration procedures and integrate statistics in the research context, in a descriptive study. The students did not reach a similar level, when working with statistical methods, especially at inferential level; probably this fact is related to their poor perception of uncertainty in their conclusions. Our pretension of converting future professionals into statistician's amateurs is too optimistic, since a reasonable and efficient application of statistics requires a wide knowledge and abilities. We neither need to prepare professionals in calculus level nor in the graphic representation, because nowadays computers solve these problems. Possibly we should change the aims of teaching statistics to professional, from transmitting a more or less passive knowledge to transmitting active statistic literacy (Gal, 2002; Batanero, 2003) that refers to two interrelated components: "*(a) people's ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant (b) their ability to discuss or communicate their reactions to such statistical information* " (Gal, 2002, pp. 2-3).

Current technology places the challenge of restructuring statistics teaching, in a way that we achieve the aim of transmitting both statistical reasoning and literacy. Encouraging students to do their own projects is a possible way, since in our study we found a remarkable wealth of statistical ideas in these projects.

REFERENCES

Altman, D. (2002). Poor-quality medical research. *Journal of the American Medical Association*, 287(21), 2765-2767.

Altman, D., Goodman, S. and Schroter, S. (2002). How statistical expertise is used in medical research. *Journal of American Medical Association*, 287(21), 2817-2820.

Batanero, C. (2003). Los retos de la cultura estadística. *Jornadas Interamericanas de Enseñanza de la Estad*ística, Buenos Aires.

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review*, 65(2), 1; 67-190.

Gal, I (2002). Adult's statistical literacy. Meanings, components, responsibilities. *International Statistical Review,* 70(1), 1-25.

Pimenta, R. (2005). *O uso da Estatística no Projecto de Investigação em Fisioterapi.* Unpublished Master's Thesis. Universidad de Santiago de Compostela, Spain.

Wild, C. and Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (with discussion). *International Statistical Review,* 67(3), 223-265.