ASSESSING THE STATISTICAL COMPETENCIES OF PRIMARY SCHOOL STUDENT TEACHERS

Godino, Juan D.
Departamento de Didáctica de la Matemática
Universidad de Granada
18071 Granada (España)
E.mail: jgodino@ugr.es

Roa, Rafael
Departamento de Didáctica de la Matemática
Universidad de Granada
18071 Granada (España)
E.mail: rroa@ugr.es

Wilhelmi, Miguel R.
Departamento de Matemáticas e Informática
Universidad Pública de Navarra
31071 Pamplona (Navarra, España)
E. mail: miguelr.wilhelmi@unavarra.es

Abstract:

In this paper we analyse the competencies of future primary school teachers in solving data analysis projects. To study their probabilistic intuitions we posed a project, which involved carrying out an experiment, collecting and analysing data to 55 student teachers. We conclude that these student teachers know the elementary statistical concepts and procedures they should teach; however they scarcely interpret the results to get a conclusion about the problem investigated in the project. Even when the future teachers show a conceptual and procedural competence, they lack interpretative and application skills. We suggest that, a condition for a meaningful statistical education in primary school is taking into account these shortages in the training programme directed to teachers.

1. Introduction

Nowadays Statistics is part of mathematics curricula for primary and secondary school classes in many countries. The reasons to include statistics teaching have been repeatedly highlighted over the past 20 years (Gal, 2002), and include the usefulness of statistics for daily life, its instrumental role in other disciplines, the need for a basic statistical knowledge in many professions and its role in developing a critical reasoning. A theme on "Data analysis, chance and probability" is included in the mathematics Official Curriculum for Primary Education in Spain. The mastering of the basic statistics content and its pedagogy is required if we want primary school teachers be able to carry out an efficient teaching of the topic and therefore should be included in the training of these teachers.

In this paper we describe an experience directed to train primary school student teachers in statistics and its didactics that was carried out in the Department of Didactic of Mathematics at the University of Granada (Spain). In this experience a didactic unit based on a data analysis project was used to contextualise the elementary statistical contents included in the Spanish curriculum for primary education; at the same time the project served to develop the pedagogical content knowledge related to those contents. Statistical projects are conceived to introduce an exploratory and participating dynamic in the classroom, in agreement with the recent recommendations on teaching statistics (Nolan and Speed, 1999; Connor, Davies and Holmes, 2006). In fact, it would be desirable that the future teachers choose their own projects to work on them in teams of two o three students (Batanero and Godino, 2002).

Below, we firstly describe the educational context and the teaching methodology used in training the future primary school teachers. Then we briefly describe the project, "Check your intuitions about chance". In section 4 we report results of assessing the students' statistical knowledge and competences. We conclude with some final reflections as regards the need to improving the training program of future primary school teachers in Spain and highlight the potential of data analysis projects as a tool to develop statistical knowledge.

2. Educational setting and methodology

We worked on the data analysis project described in section 3 with a group of 55 primary school student teachers that took part in a course of "Mathematical curriculum" taught in the Faculty of Education at the University of Granada (Spain). These students had previously taken a Mathematics and its Didactics course the previous year with only 20 teaching hours devoted to statistics and related pedagogy. The course "Mathematics curriculum" is mainly oriented to train the future teachers in the pedagogical content knowledge. The project was presented to the students in a 90 minutes session, where they were asked to invent a random sequence as well as to carry out the simulation experiment of throwing a fair coin 20 times in order to collect data for the experiment. After each student performed the experiment and took his/he data, the lecturer introduced the notion of run, number of runs in 20 throwing and length of the longest run. Each student was asked to compute these data, as well as the number of heads in 20 throwing for both the invented and the real sequence in the experiment. Each student data were recorded in a sheet that was photocopied and given to the whole class. Students were asked to analyse this small data set, get a conclusion about whether intuitions about random sequences in the classroom were sound or not and produce a written report . The students knew (from the previous course) some statistical procedures and concepts; a discussion in the classroom was organised to see which of them were suitable to analyse the data and get a conclusion about people's intuition.

Students were given an assignment (home work) to complete the analysis and report during the following week; they should use the statistical graphs and procedures that they considered better suited to answer the project questions. They were free to work in teams and use any computing software available to them, but each student should write his/her own individual report. We asked the students to include any statistical table and graphical representation they needed to analyse the data and also explain in detail their conclusions about the people' s intuitions on chance and the ability to simulate (invent) a random sequence resulting of throwing a coin 20 times.

After the project was finished and the results were discussed in the classroom, this project was also used by the lecturer to discuss with the future teachers the pedagogical content knowledge involved in teaching and learning of basic statistics in primary school in the following days. Nevertheless, due to lack of space, in this paper we focus our attention in the results of assessing what statistical knowledge was developed by the teachers in the activity, although in the course we tried to articulate the content knowledge and pedagogical content knowledge in teachers' s training.

3. Data analysis project: "Check your intuitions about chance"

A didactical unit, which was based on a statistical project and was designed to develop the theme of "data handling" in 6th level of primary education, was presented to the future teachers. The first step was to propose the future teachers to carry out themselves the experiment involved in the project and to analyse the data obtained in the whole class. Secondly, we organised a collective discussion of the statistical and didactical features of this project, the teaching of statistics through project work and the extent to which the project was useful to teach the statistical content in the 6th level. The sequence of activities in the project was as follows.

1) Initial instructions and collective discussion. We propose the future teachers the following questions: Do you have a good intuition about chance? Are you able to write apparent random results of throwing a coin 20 times (without really throwing the coin, just inventing the results) in such a way that other people could think you really threw the coin?

2) Individual experiments and collecting data. Future teachers were encouraged to try themselves the experiment and invent a random sequence. At the same time they were asked to perform the experiment in order to compare the invented (simulated) and real (random) sequence of results: Let's check your intuitions regarding the random results. Below you have two grids; write down 20 simulated throwing of the coin in the first grid. Flips 20 times a coin and complete the second grid. Write H for head and T for tail.

Simulated throwing																			
Real throwing																			

3) Classroom discussion, new questions and activities. After each student performs the experiment they are asked how can we to distinguish a simulated sequence from a real random sequence. After a while some students might suggest to compare the number of heads and tails in the two sequences, and argue that we should have 10 tails and 10 heads, given that the events have equal probabilities. The lecture might pose new questions similar to the followings: If the sequence is random, should we get exactly 10 heads and 10 tails? What if we get 11 heads and 9 tails? Do you think in this case the sequence is not random? What about 18 heads and 2 tails? Would it be useful to compare the number of heads and tails in the real and simulated experiments for the whole class?

The session continues collecting the data about number of heads, number of runs and length of the longest run for the whole classroom. Students are given a copy of the data for the whole classroom and are asked to carry out a statistical analysis to get conclusions on peoples' s probabilistic intuitions about these kind of stochastic processes.

4. Assessing the future primary school teachers' statistical knowledge

After the students presented their reports, a qualitative analysis of responses serve to generate some variables that were analysed with SPSS, including cluster analysis (k-means) of variables to identify typologies in Students' performance as regards statistics. Since the students of our sample had received a previous training in elementary statistic and probability during their first year of their carrier (for about 20 hours) they were supposed to know frequency tables, bar charts, histograms, averages and dispersion measures. In the following we describe four typologies identified in the cluster analysis that we interpret as cognitive configurations (Godino, Batanero and Font, 2007), as regards the students' statistical knowledge and competence.

Type 1 (33 students, 60%): These students do not recognise the individual experiments carry out by each student in the class as part of the sample that should be globally analysed to conclude on the people' s intuitions. Part of them only analyses their own data (compare their results in the two sequences) and expect a coincidence. In some cases they only analyse a variable (usually the number of heads), for the real or simulated experiment, but they do not compare the same variables in the two samples (real and simulated sequences). They prepare absolute frequency tables, but either do not elaborate a graphical representation or prepare an incorrect one (for example plotting results in the two sequences for each student in a paired line graph); they usually compute the mean, but do not get any conclusion about differences in the distribution or about the probabilistic intuitions in the group of students.

Type 2 (10 students, 18,2%): These students analyse the three pair of variables (number of heads, number of runs, length of the greater run, for the simulated and real cases). They produce tables, graphical summaries (bar charts of line graphs), and statistical summaries independently for each variable. They do not compare the simulated and real distributions and only describe partial aspects of the distributions; e.g. the averages.

Type 3 (9 students, 16,3%): Similar to type 2, but the students produce joint representations for each pair of variables (or joint tables); this facilitates the comparison of each pair of distributions, although, usually the comparison is limited to the averages, without taken into account the spread. Other students focus on the spread, and do not use the averages. They get no conclusion about the probabilistic intuitions in the group.

Type 4 (3 students, 5,5%): Similar to type 3, but the student do compare both averages and spread. They conclude that the group has a good intuition on the average number of heads and tails in a random sequence but not about the dispersion. The class has a bad intuition about runs.

In table 1 we include the frequency and percentages of students (n= 55) that used correctly or incorrectly different statistical knowledge. In general students show knowledge of statistical graphs and summaries. Some errors were produced in identifying the statistical variables (comparing individual values, instead of using the distribution of number of heads, runs or longest run in the whole group) or in inadequately plotting the data (e.g. plotting in adjacent bar graph each couple of values for each individual; plotting in the same graph the different values got in the individual experiments). However, the number of students who got a conclusion was very small, even in comparing the means or spread of two distributions.

Table 1: Frequency and percentage of statistical knowledge used by the students

	Correct		Incorrect	
Statistical knowledge	Frequency	Percentage	Frequency	Percentage
Statistical variables and values	44	80.0	11	20
Mean	16	29.1	1	1.8
Median			1	1.8
Range or standard deviation	16	29.0		
Bar charts or line diagram			13	23.6
Matched bar charts or line diagrams	10	18.2		
Identification and elaboration of frequency distributions	22	40.0	3	5.5
Conclusions on the equality/inequality of averages	16	29.1		
Conclusions on the equality/inequality of dispersions	10	18.2		
Conclusions on the probabilistic intuitions	3	5.5		

5. Final reflections

The data analysis project "*Check your intuitions about chance*" served to contextualise the elementary statistical notions and procedures included at primary school and to use them to solve a problem in a research setting. The project– situation posed has permitted to overcome the traditional "knowledge division" in textbooks and the traditional orientation toward the concepts and procedures presentation. Even in primary education teaching and learning statistics through analysis of realistic problems will give a "situational meaning" to statistical concepts and techniques. In our experience, the investigation of our intuitions about the behaviour of random sequences lead to compare frequency distributions, and thus justify the tabular, graphical and numerical data reduction (use of averages and spread measures).

Another feature of using projects in statistical education is the multivariate approach to data analysis. Making decision in random situations requires take into account, not just one variable, a unique perspective, but a multiple approach: the average number of heads and tails in the simulated and real sequences are quite similar, but not the average of run length, or the number of runs. People have a good intuition on equiprobability, but not of variation in a random process. To recognise these features of stochastic processes is crucial to overcome probabilistic biases such as the "gambler fallacy" (Tversky y Kahneman, 1982), so the project also permits future teachers to reflect on the multiple models present in a random sequence. Our research also suggest the need to improve the statistical training of future primary school teachers and the possibilities of activities similar to that we have described. This will be only possible if significant changes are introduced in the initial teachers' training syllabus assigning more teaching hours to mathematics and statistics education.

Acknowledgments: This research work has been carried out in the frame of the project: MCYT – FEDER (UE): SEJ2004-00789.

REFERENCES

Batanero C., and Godino J. D. (2002). Estocástica para maestros. En J. D. Godino (Dir.), Matemáticas y su didáctica para maestros. [http://www.ugr.es/local/jgodino/edumat-maestros]

Connor D., Davies N., and Holmes P. (2006). Using real data and technology to develop statistical thinking. In G. Burrill (Ed.), *Thinking and Reasoning with Data and Chance* (pp. 185- 194). NCTM, 68 Yearbook.

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

Godino, J. D. Batanero, C. and Font, V. (2007). The onto-semiotic approach to research in mathematics education. *Zentralblat für Didaktik der Mathematik,* 39 (1-2): 127-135.

Nolan D., and Speed T.P. (1999). Teaching statistics theory through applications. *American Statistician*, 53, 370-375.

Tversky A, and Kahneman D. (1982). Belief in the law of small numbers. In D. Kahneman, P. Slovic y A. Tversky (1982), *Judgment under uncertainty: Heuristics and biases* (pp. 23-31). New York: Cambridge University Press.