ASSESSING UNDERGRADUATE STUDENTS OF A STATISTICS COURSE IN ENVIRONMENTAL SCIENCE

CAPILLA, Carmen
Universidad Politécnica
Spain

This paper presents a case study of assessing an applied statistics course. The assessment includes 10 computer lab activities, that students may voluntary complete during the fall semester. An individual voluntary project is also incorporated as part of the assessment. At the end of the semester there is a final examination about the course contents. These assessment methods and the difficulties to implement them are discussed. Students have heterogeneous backgrounds. Some of them show fear and anxiety towards the subject. This interferes with the implementation of the assessment and implies discipline problems.

INTRODUCTION

The objective of this paper is to present a case study of assessing applied statistics learning in undergraduate environmental science studies at the Universidad Politécnica de Valencia (Spain). The students are in their seventh term and their background in statistics before enrolling for the course, ranges from no statistical experience whatsoever to basic skills acquired in an introductory statistics course (descriptive analysis and probability). The majority of the students have completed the first cycle (three academic years) of Agricultural Technical Engineering. They have a very limited mathematical background, and they often show anxiety and fear towards anything related to mathematics. On the other hand the introductory statistics of their first cycle is taught as part of a Mathematics subject, only a few hours are dedicated to introducing statistics and is taught by the Applied Mathematics Department of our university.

At the beginning of the course, students are asked about their previous statistical knowledge. In the introductory statistics course of the first cycle that some of them have already completed, the teaching approach is more theoretical. Other students have also completed in their first cycle a 4.5 credit course on design of experiments. In this course they use the software Statgraphics, but the majority of exercises and examples are not related with Agriculture Engineering. The consequence of all of this is that their previous attitude towards statistics is in general negative. This carries out problems of discipline. For example, we have to include in the course some control systems to prevent students from cheating. On the other hand a high percentage of them work outside of the university. Although the course is taught in the afternoon, they can not regularly attend the lectures and computer lab classes. This year only 58.6% (total number enrolled students n=66) students have been attending the classes. The other students usually never appear or come to the university.

STUDENTS PREVIOUS STATISTICAL KNOWLEDGE

We have been trying to evaluate students' reasoning about basic statistical concepts since the year 2005, after attending the 5th Iberoamerican Conference on Mathematical Education (Porto, Portugal, July 2005). During this conference we attended the discussion group on statistics education. In this group, Estrada (2005) presented a paper about the application of the Statistical Reasoning Assessment Test (SRAT) described in Garfield (2003), in other context. In order to assess the students' previous statistical knowledge, a multiple-choice test is given at the start of the course. The selected questions from SRAT are 1, 4, 5, 14, 15, 16 and 17, and are related with basic statistical concepts (see Table 1 below). The questionnaire provides scores to facilitate an evaluation of students' previous reasoning on several basic statistical methods and concepts. We are also using this test with computer science engineering students.

The test results reveal that some students have difficulties and misconceptions when reasoning about the basic statistical concepts, which they ought to know following the completion of an introductory course. Because of that and the limitation of time, only a brief review on these concepts is given. This review consists of 5 hour lecture and 6 hour computer lab work, using Statgraphics and SPSS. More details are given in next section about the computer lab classes.

Table 1 gives the average correct and incorrect test scores at the beginning of the semester and two months later. The correct score results are worse two months after the review, except for the interpretation of two-way tables. The incorrect scores are slightly better for misconceptions involving averages, for the misconception "good samples have to represent a high percentage of the population", and the law of small numbers. This indicates that the brief review is not effective in trying to improve students' background at the beginning of the course. A course program change would be necessary including a module on statistical descriptive analysis (statistical measures and graphs), and more hours should be devoted to teach these methods. The actual keywords of the statistics course, published in the official national bulletin that approved the official environmental science degree, did not include statistical descriptive methods and probability. The actual descriptors are: Probability distributions, hypothesis test, sampling, analysis of variance, correlation and regression, and introduction to multivariate analysis.

Table 1
Average correct and incorrect scores before the review of basic concepts, and two months later. Percentages of the different types of misconceptions about the mean are given. Students were enrolled in the statistics course of the Environmental Science degree (year 2006).

	Beginning N=34	Two months later N = 22
Correct reasoning skills		
Understands how to select an appropriate average	1.43	1.3
Understands sampling variability	0.24	0.23
Distinguishes between correlation and causation	0.82	0.73
Correctly interprets two-way tables	1.12	1.36
Misconceptions		
Misconceptions involving means	1.06	1
a) Averages are the most common number	17%	13%
b) Fails to take outliers into consideration when computing the mean	5.7%	0%
c) Compares groups based on their averages	77%	78.3%
d) Confuses mean with median	2.9%	0%
Good samples have to represent a high percentage of the population	0.59	0.55
Law of small numbers	1.47	1.18
Correlation implies causation	0.65	0.91

Figure 1 below shows the results of the total correct scores by gender and studies two months after the brief review of these concepts. Although the differences are non-significant, 2006 year students, who attended the computer lab classes and had already completed an agricultural technical engineer degree, have the lowest correct scores.

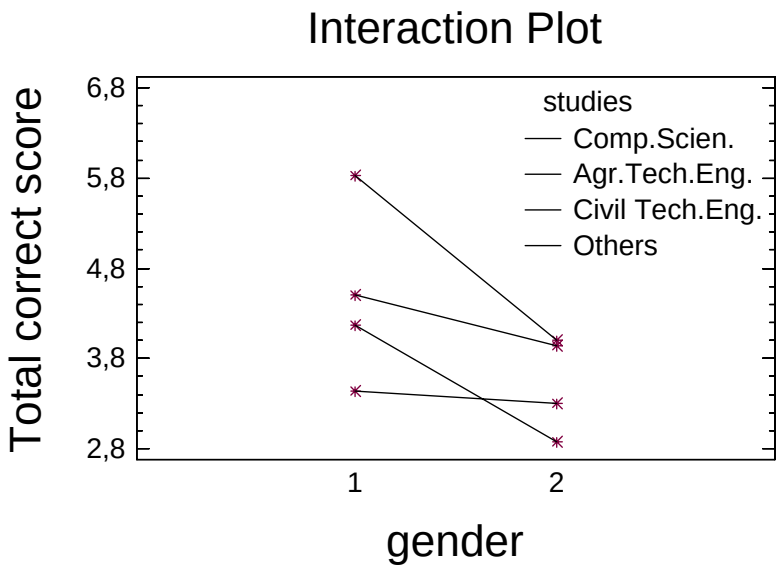

Figure 1. Total correct scores of the SRAT by gender and studies

COURSE ASSESSMENT

The educational literature recommends the use of a variety of performance assessment methods (Gal & Garfield, 1997). Following these recommendations, the environmental science statistics course assesses performance on the basis of marks obtained in laboratory tasks, an individual project and a final exam. Before the year 2005 a midterm exam was also included as part of the assessment.

On-going students evaluation with computer lab activities

The students' written answers to computer laboratory activities provide an ongoing evaluation of what they have learned regarding the main concepts studied in the lectures. They may work in the computer lab classroom in groups of two or three people. In the year 2005, we told them that team work was not compulsory and the majority of the students decided to solve the computer lab tasks on their own. This year team work has been compulsory again.

The attendance at the computer lab activities accounts for 10% of the final mark. The evaluation of these activities reports accounts for 20% of the final mark. Since 2005 the attendance at the computer lab classes is voluntary. Before that year the attendance was compulsory, and this implied a higher percentage of attendance: e.g. year 2002 62.7% (total number of students enrolled for the course n=80), year 2003 76.4%(n=80), in comparison with year 2005 56.5%(n=62) year 2006 58.6%(n=66). When the computer classes were compulsory, some students tried to cheat and there were some problems of discipline. This damaged the computer lab atmosphere and other students that have a positive attitude towards the subject. On the other hand, with voluntary attendance some students are present at the computer lab classes just because they received a mark for it, and the majority of them without attending the theory lectures or studying the theory by themselves.

When the computer lab classes began to be voluntary in the year 2005, a separate mark was given for the attendance (10%) and for the report of the lab tasks (20%). A control system consisting in signing the entrance and exit hour was used, but some students lied about these numbers. Therefore the instructor has to personally watch the numbers they write when they sign the attendance control sheet.

Before 2005 students had to present the report at the end of the practice. As they had not previously study the theory, they lacked time to answer. Therefore, since 2005 we decided to change the deadline to present the report with the final conclusions to 15 days after the day of the practice. This has also been done in the year 2006. The experience during these years is that cheating increased, is more difficult to detect, and that the students seem to be studying less than before. Therefore we are considering for the future course assessment that students present the lab report at the end of the class.

In Table 2, the contents of this year practices are given. Each computer lab class lasts two hours. Statgraphics and SPSS are used. SPSS was introduced in the course in the year 2004 (only two descriptive methods practices), and since 2005 it is used in all practices. These two packages are the only statistical software available with campus license in our university network. We are considering that students only apply SPSS in future years, because Statgraphics have some errors and is less complete than SPSS. We are also considering to introduce in the computer lab classes the R free software. Before doing this we would like to know the results of teaching applied statistics (our students are not statisticians) in other university.

Table 2
Computer lab contents during the fall semester 2006

Computer lab tasks	Contents
1	Two-way tables
2	Univariate descriptive methods (sample description parameters, histograms and box plots)
3	Bivariate descriptive methods (scatterplots, covariance, correlation, regression line)
4	Binomial and Poisson distributions
5	Probability papers, normal distribution
6	Hypothesis tests and confidence intervals for the mean and standard deviation
7	Comparison of two means and variances
8	Analysis of Variance
9	Linear Regression
10	Introduction to time series descriptive analysis and multivariate analysis

Real case environmental studies are used in the exercises of the computer lab classes. They may work in teams of two or three people. Some exercises consist in the analysis of atmospheric pollution and climatology data observed in Valencia urban area. This allows to make connections with other subjects of the environmental science program. In particular in the following semester they study Atmospheric Pollution with the Associate Professor Aragon, with whom we also have some research cooperation. We also use the references Piegorsch and Bailer (1997), Berthouex and Brown (2002) and Wilks (2006) for some of the computer lab exercises.

The group scores that students teams obtained this year in the lab tasks (scaled from 0 to 0.2), have been analyzed using a linear model. The analysis indicate that there are significant differences between teams and lab tasks, and the interactions between these effects is also significant. Figure 2 shows the means and 95% Tukey HSD intervals for the 10 lab tasks in the year 2006. Although the theoretical concepts involved in the three first lab tasks are easier, the obtained scores are the lowest. There was an upward score trend until the computer lab task #8 (Analysis of Variance). Some students have already studied some probability distributions, inference and analysis of variance techniques in the first cycle (contents of lab tasks #4-8). The correlation between the scores in practices #3 (scatterplot, correlation, simple regression line) and #5 (normal probability plot) is the highest (0.8701). Students scores' in the last lab class (#10) is the lowest. In this last practice students work on descriptive analysis of time series analysis and an introduction to multivariate analysis (principal component analysis). The theory involved in these modules is more complex.

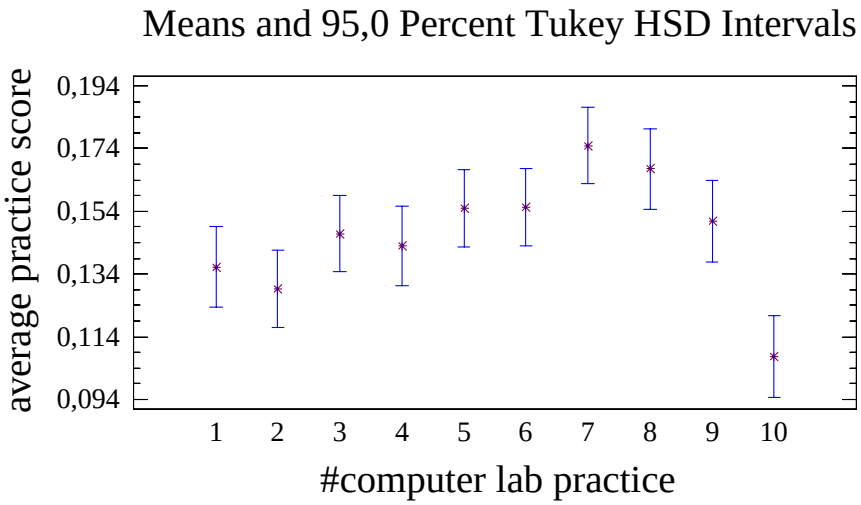

Figure 2. Means and 95% Tukey HSD intervals for the 10 lab tasks in the year 2006.

Figure 3 shows the teams' average scores in the year 2006 (#5 is the only student that completed the first cycle in Biology at the University of Valencia, and #6 is an Erasmus student from Germany, who had some difficulties with the Spanish language). Figure 4 plots the scores time series of several students teams (year 2006).

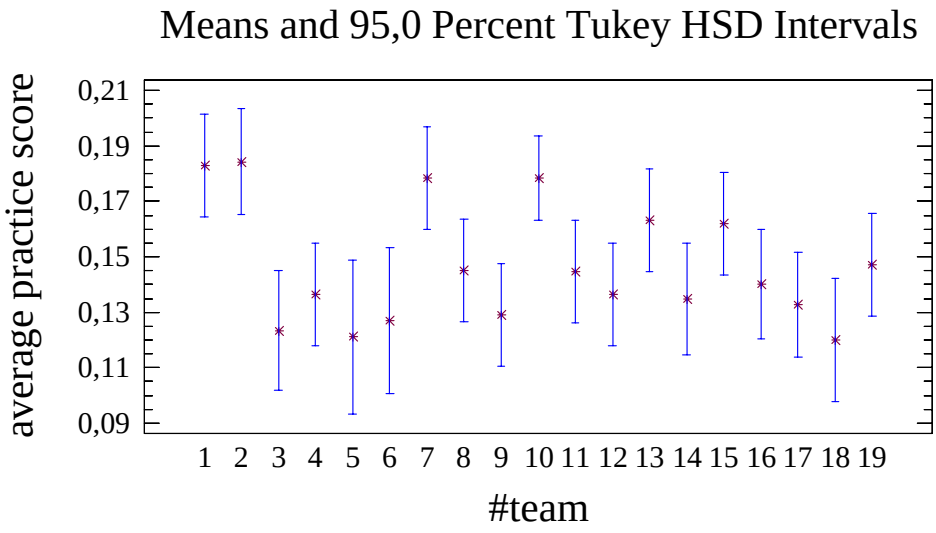

Figure 3. Teams average practice scores in 2006.

An example of computer lab activity is given in next page (Exercise 4, Practice #2). This activity may theoretically help students to understand that the comparison of groups have to take into account both the sampling means and spread.

The first three computer lab classes were about descriptive methods. The results of these practices have been related with the scores obtained in questions 1, 5, 15, 16, and 17 of the SRAT, two months after the review and the computer lab tasks about the reasoning skills evaluated in these questions. The results are given in Table 3. A very frequent misconception has been to compare groups based on their averages (78.3%). The average students' scores in the computer lab tasks (#2 and 7) with exercises related with this concept, have been higher for those who correctly answered this question.

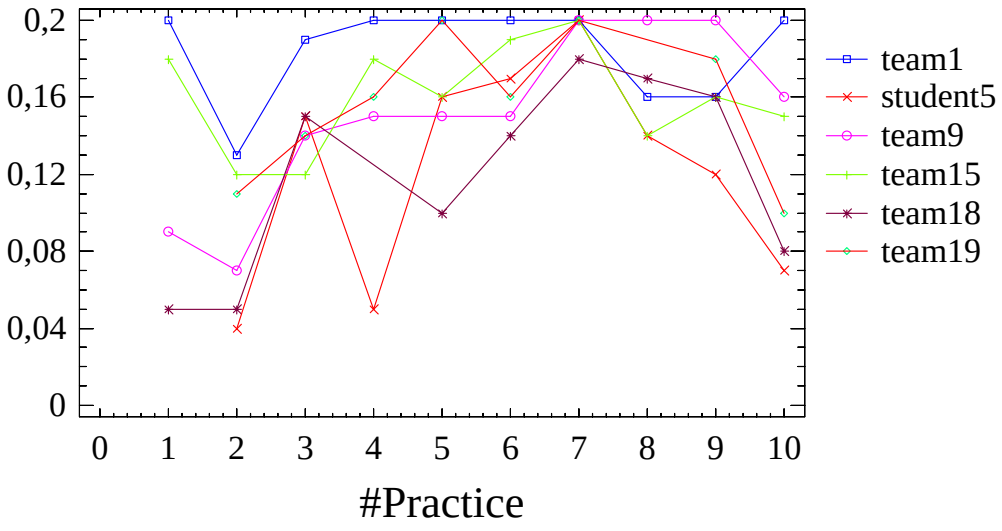

Figure 4. Time series of scores of several teams.

EXERCISE 4 (PRACTICE #2)- DESCRIPTIVE ANÁLISIS OF OZONE CONCENTRATION AT THREE LOCATIONS OF VALENCIA, USING SPSS

Friday ozone concentrations (ppb) are available at three stations of the automatic network to monitor atmospheric pollution in Valencia (Data source: Environment Laboratory, Valencia Town Hall, 1996).
Data are in the file OZONO.SAV. To open the file:

File...Open...Data
Name: OZONO

OZONO variable has 277 values of this pollutant concentration and the variable ESTACIÓN the codes of the sampling area (1, 2 or 3).

4.1- Compare using descriptive methods the central tendency, the spread, skewness and kurtosis .of ozone concentration at the three stations. Use sample measures and histograms and box plots. In the answer include the quantitative results and the description of the plots that justify your answer.

You may use the option:

Analyze...Descriptive statistics...Exploration
Dependents: OZONO
Factors: ESTACION

Assessment using individual projects

This year individual projects have been introduced as a voluntary evaluation activity. Individual projects can be useful to assess students' understanding of the application of statistical methods to environmental problems, project design, and communication skills. Projects include a written assignment that accounts for 15% of the final mark and a 15-minute oral presentation that accounts for 5%. However only 4 students (females) decided to present a project proposal this year. At the end of the semester (deadline to submit the project), only two students presented it. When the other students have been asked about their negative attitude towards this voluntary activity, they have claimed that they are involved in many works of other subjects of the environmental science program, and they lack time to do the project. We think that they also lack motivation towards statistics, as they show through their attitude in relation to their class

attendance, their reluctance to study the theory before the lab classes, and their poor previous mathematical and statistical background.

Table 3.
Average scores observed in the computer lab tasks #1, 2, 3 and 7 as a function of the answers to some SRAT items.

	Average	Standard deviation
Correctly interprets two-way Tables (SRAT item #5)	Comp.lab.task #1	
No (n=6)	0.13	0.07
Yes (n=15)	0.14	0.05
Misconceptions: averages are the most common number (SRAT items #1 and #17)	Comp.lab.task #2	
No (n=18)	0.15	0.05
Yes (n=3)	0.10	0.08
Misconceptions: compares groups based on their averages (SRAT item #15)	Comp.lab.task #2	
No (n=5)	0.17	0.04
Yes (n=16)	0.13	0.06
(SRAT item #15)	Comp.lab.task #7	
No (n=5)	0.20	0.00
Yes (n=18)	0.17	0.04
Distinguishes between correlation and causation (SRAT item #16)	Comp.lab.task #3	
No (n=14)	0.15	0.03
Yes (n=9)	0.16	0.03

For the guidance and evaluation of the individual projects, we have followed the guidelines given in the web page of the project ARTIST (https//app.gen.umn.edu/artist, delMas et al. 2006). The presented projects were about descriptive analysis of monthly atmospheric pollution at Viveros station (Valencia), year 2005, and descriptive and trend analysis of the daily mean traffic density in the road N-232, Morella.

Assessment using a final examination

At the end of the semester, students take an exam that covers the entire content of the course, excluding the review. Therefore the exam includes questions related with probability distributions, test hypothesis, analysis of variance, regression, and introduction to time series analysis and multivariate analysis. In this exam, we try to emphasize statistical analysis and interpretation skills. Some of the questions require students to interpret output from the statistical software used in the computer lab classes. Some examples of questions in the final examination in January 2007 have submitted to the workshop.

FINAL ASSESSMENT RESULTS

The pass rates since 1997 are plotted on the left axis of Figure 5. It can be seen that in the year 2005 this rate has been lower. This might be the consequence of changing the practices attendance from compulsory to voluntary, which has had an impact on the attendance percentage, and of the 15 day deadline to present the lab task report. It is not possible to make a comparison of performance for those who did and did not complete the labs, because none of the latter did the exam at the end of the course.

At the end of the semester students complete a teacher`s evaluation questionnaire. The results of this questionnaire are plotted on the right axis of Figure 5. The scale for the scores of students' evaluation of the teacher, ranges from 0 to 10.

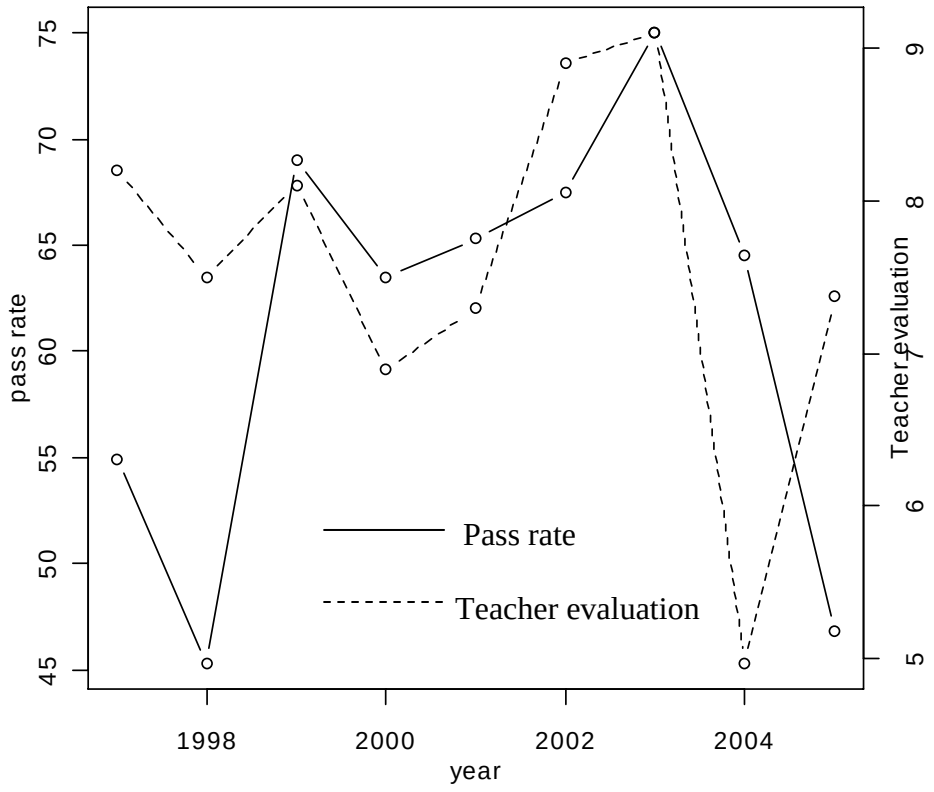

Figure 5. Course pass rates and students' evaluation of the teacher in the period 1997-2005.

REFERENCES

Berthouex, M.P. and Brown, L.C. (2002). *Statistics for Environmental Engineers*. Boca Raton: Lewis Publishers.

DelMas, R., Ooms, A. & Garfield, J. (2006). Assessing students' statistical reasoning. In Rossman, A. and Chance, B. (Ed.) *Proceedings of the Seventh International Conference on Statistical Education*, Salvador, Brasil.

Estrada, A. (2005). La educación estadística de los profesores en formación. *Conferencia Iberoamericana de Educación Matemática*, Porto, Portugal.

Gal, I. & Garfield, J. (Eds.) (1997). *The Assessment Challenge in Statistics Education.* Amsterdam: IOS press.

Garfield, J.B. (2003). Assessing statistical reasoning. *Statistical Education Research Journal*, 2, 22-32.

Piegorsch, W.W. & Bailer, A.J. (1997). *Statistics for Environmental Biology and Toxicology*. London: Chapman & Hall.

Wilks, D. (2006). *Statistical Methods in the Atmospheric Sciences*. Burlington, MA; London: Academic Press.