VISUAL STATISTICS: A MULTIMEDIA STATISTICS FACILITATOR

Antonio Giusti, Alessandra Petrucci and Francesco Polverini
Dipartimento di Statistica "Giuseppe Parenti", Università di Firenze, Italy

The difficulties encountered in teaching elementary statistics to students not at ease with mathematics, is well known. Several initial basic concepts in statistics, like location, variability, correlation, etc. involve the use of very simple formulas that, however, have a forbidding aspect for people not acquainted with mathematical notation. For these people, a visual approach, based on interactive graphics, implemented on a personal computer, would be extremely helpful. The aim of the research we are carrying out is the development of a multimedia software package, implemented through a matrix language, which could be used in teaching a first course in statistics to high school or "non statistics" college students. The package, whose construction is well underway, illustrates the most important basic concepts of statistics through the use of graphic displays, animations, audio comments, etc.

INTRODUCTION

The software package we are developing here at the Department of Statistics of the University of Florence (Italy) is addressed to high school students and to university students with a non statistics major. Before discussing the package in greater detail, let us briefly describe the environment of its potential users. In particular, let us examine the Italian situation as regards the teaching of statistics in secondary schools and at the college level.

No courses called "statistics" are taught at the secondary level: statistical education should be part of a mathematical education, although mathematics teacher avoid the subject entirely. The courses in mathematics developed by the Ministry of Education for the various types of schools into which the Italian educational system is divided, always contain some notions of statistics, but the implementation of the ministry directives is poor if not completely absent. Excluding some (few) professional schools that happen to have an illuminated mathematics teacher, we can safely affirm that no statistics is taught in the Italian schools, not even data handling, charting etc., let alone sampling and statistical inference.

This situation can be attributed to the role statistics plays at the university level. If we exclude the degree in Statistics, the role played by statistics is minimal in any other degree offered by the Italian universities. In particular, high school and junior high school mathematics teachers must have, by law, a university degree in mathematics but the

curriculum for this degree, most often, does not contemplate any course in statistics; at best, teachers may choose to take a course in mathematical statistics which is taught like any other strictly mathematics course. At the end of their studies, they do not have any feeling for real data and, often, not even for inductive reasoning. It is not surprising that, as teachers, they carefully avoid anything that has to do with statistics and data handling. At the university level the situation is only slightly better; statistics courses are very few and they are often taught by non statisticians: to give an example, most of the graduates in agricultural sciences have not taken any course in statistics since none are required for their degree. The few statistics courses which are offered, even the ones taught by statisticians, are conducted as formal lectures with a preference for rigor, theorems, proofs, formulas etc. In conclusion, there prevails a very traditional, we might also say "elitist", way of teaching statistics as well as other disciplines, in particular the more mathematical ones. At the same time there has been a tremendous increase in the student population, and a consequential need for what in the English speaking countries is called *democratization of mathematics*. (Vere-Jones,1995) This is particularly necessary in a country like Italy where education is seen mostly as the acquisition of humanistic knowledge and analytical, quantitative subjects have a place only if they are offered at the more abstract and philosophical level. The job market, instead, requires concrete quantitative and computing skills that the present educational system rarely provides.

Right now the Italian educational system is undergoing an extensive reform both at the secondary and at the tertiary level. The reform should affect every aspect of the system: the recruiting of educators, the content of courses, the methodology of teaching, etc. The time is thus ripe for proposals that should have an impact on where, what, how much and how statistics will be taught in the future. We need to develop tools that help to modernize the way the subject is taught and that offer a guide for high school teachers and a practice tool for both high school and beginning university students. It is in this spirit that we are developing our product.

ON THE TEACHING OF STATISTICS

After the brief description of the environment in which we are moving let us state even more briefly our position regarding the teaching of statistics. We agree with most of the points made by Moore in a recent article (Moore, 1997), even if some of his proposals need to be carefully adapted to the specific Italian educational system. In particular, the

opinion that we should exploit the relationships among content, pedagogy and technology to change our introductory statistics courses by encouraging more active learning both at the individual and group level, by emphasizing statistical thinking and statistical concepts rather than formal derivations, by relying on real data to carry out projects and laboratory exercises exploiting all the computer technology we can for computation, find us in total agreement. The most pressing problem in Italy, however, is the already mentioned problem of teachers who have degrees in mathematics but have no feeling for applied, and often even for theoretical, statistics. The introduction of courses in statistics, separate from courses in mathematics, taught by statisticians, would solve the problem, but would take too long a time to put into practice, due to the current shortage of qualified statisticians. The short run solution would, in our opinion, be the organization of training courses in statistics for high school mathematics teachers. A useful aid for these courses and for the successive transfer of the learned concepts to the students, would be the software package we are developing.

CHARACTERISTICS OF THE PACKAGE

In order to develop the package we had to make some important initial choices regarding its specific characteristics: the interaction mode between package, instructor and students, the statistical content of the package, the delivery mode of the content, the computer platform and the operating system, the programming language and the mode of diffusion. In the development of the package, we relied heavily (or at least we tried to rely) on Biehler's suggestions (Biehler, 1997). Using Biehler's terminology, we could classify our software as a microworld that, we hope, will overcome most of the criticisms generally levied against this type of package and that will be particularly suited for the problems we face in Italy.

The first choice regarded the relationship instructor-package-learner i.e. the choice among packages that would be utilized as:

a) an aid for the teacher during a lesson or a lab session without student interaction;
b) an aid for the teacher, but with the possibility of student to intervention for specific tasks;
c) an aid for the student for individual or group practice or for assessment without the teacher;
d) a total or partial substitute of the teacher's lessons and/or assessments.

Even admitting that any product will fulfill, more or less adequately, many of the aims listed above, we can classify our product as an aid for students' practice in learning statistics (in particular in learning statistical concepts) with a limited capabilities for doing statistics. Naturally, the product will also be used by teachers for lecturing and for lab sessions. We should keep in mind, however, that the teacher will not generally be a statistician and hence the package should also be a guide for him/her allowing for the fact that he/she is a "motivated and sophisticated student".

Regarding the second point we view the package as a dynamic tool: in this first phase we limit the content to the so-called descriptive statistics that in Italy has a very strong tradition and that is only scarcely covered by the existing statistical packages. In particular, we intend to treat extensively graphical and tabular displays of data, most of the position and variability indices, including concentration indices generally not considered in other packages, measures of association, regression analysis from a descriptive point of view paying attention to the diagnostic aspects of it, and so on. In a second phase, we will extend the package to include statistical inference.

The learning process implemented in the package is more important than the content itself. As the title suggests, we rely on an interactive visual representation of the subject that allows the student to test continuously his understanding of the concepts treated in the lectures. Furthermore, the package should be equipped with a growing set of real and ad hoc data to be used to solve case studies and to illustrate fine conceptual points. We are also planning an adequate number of exercises to be solved by the students with the possibility of self assessment.

In the realization of our tool prototype we had to make a series of technical choices: first of all the hardware platform and the operating system. Unfortunately we were constrained by the existing hardware in our school system: almost all the computers available for instructions are PC, often not too powerful, and this forced us to choose the family of Windows (3.xx, 95, NT) as operating systems.

The key choice has been, however, the programming language used to develop the microworld. Our product can be classified as an embedded microworld since we decided to develop our tool in MATLAB. The choice was motivated by the ease with which statistical routines can be implemented in a matrix language like MATLAB together with the flexibility of the graphical interface. Furthermore, we feel that this will be highly

appreciated by the teachers who will have the possibility of exiting the microworld and finding themselves in a friendly environment particularly suited for mathematical computation.

Given the choices made above, in particular the choice of a multimedial interactive tool with its huge dataset, we feel that the only efficient way to distribute our package is the implementation of the system on a CD rom. Since at the college level, LAN are often available, the system can be implemented on a server with obvious advantages with respect to single copy use.

CONCLUDING REMARKS

The software package is still in a developing stage: the multimedia tools are still under construction and have not been implemented inside the package yet. Regarding the content, a prototype for the descriptive part is running in a single user “traditional” way and we are getting some feedback from a few selected students who are trying it out. We intend to use this prototype on a descriptive statistics course starting next September. Future development of the package will be in the direction of a multi-user version implemented on a server. We are also planning a WEB version in Java language as soon as the difficulties the secondary schools have, due to lack of computer resources, in accessing the internet will be removed.

The goal of the package is, ultimately, to facilitate the spreading of statistical concepts in places where such spreading is hampered by poor knowledge of the necessary basic instruments and not by lack of interest in the discipline.

REFERENCES

Moore, D. S. (1997). New pedagogy and new content: the case of statistics. *International Statistical Review*, *65,* 123-165, (with discussion).

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review*, *65,* 167-189.

Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review*, *63,* 3-23.