DISTANCE LEARNING: NEW FRONTIERS FOR SOLVING OLD PROBLEMS

GALMACCI, Gianfranco
Dipartimento di Economia, Finanza e Statistica
Italy

MILITO, Anna Maria
Dipartimento di Metodi Quantitativi per le Scienze Umane
Italy

After a discussion on the use of technology for teaching purposes, the authors present an experiment carried out in Italy and based on the use of e-learning to face the problem of guiding large classes of students for lab activities in basic courses of Statistics. Even though this solution has been adopted to solve a local problem, it seems to have a wide range of potential applications.

INTRODUCTION

Over the last decades Statistics has increasingly been used to exchange social information so much so that it has been called a *new language*, although it is yet a difficult language. Mass media (newspapers, TV, radio, Internet) has shown a huge transformation using, in their own language, words and concepts belonging to the statistical field. Information is generally, and frequently, presented through tables, graphs, statistical reports; exit polls and projections on anticipated election results are released, along with opinion polls whose results are presented in daily and weekly newspapers and on popular television programs (Galmacci & Milito, 2000).

The point is that those who use statistical language for social communication often do not know it very well and consequently they communicate wrong information. A high Thus a training course on basic statistics concepts could help these people to use and communicate statistics correctly.

The need to provide statistical knowledge at every level has become very evident. In order to make statistics a part of every citizen's culture, it is necessary to give him a statistical education, at least a basic one. It should start from school, to go on at University and after, in some fields, even inside the work world.

The international scientific community has for a long time been debating about statistical teaching. Just regarding different ways of teaching statistics, much consideration has been dedicated to the selection of effective teaching strategies, tending to prefer constructivist pedagogic models based on a data oriented approach. Such models generate an active learning process and seem to guarantee a bigger coherence between teaching and learning (Milito & Marsala, 2002).

The idea of the present project arises from authors' previous research activities that brought to the following question: can e-learning, as an integrative methodology of traditional teaching, contribute to the teaching of ways to communicate statistics? In particular, will e-learning be able by itself to bring students to focus on how to report results in a simple, clear and effective way, to be understood by everybody? The large number of academic e-learning projects and the wide use of such tools in public and private companies for internal and customer trainings seem to suggest a positive answer. We have adopted this approach to face an old and great problem in Italian universities - very large class sizes. (frequently more than 250 students). In these situations it is impossible to provide assisted lab activities for all of them, even though such activities represent the best way to achieve the course objectives. The experiment here described has been based on a set of teaching modules, designed for e-learning platforms, for motivating students to the use of the basic Excel facilities for simple data analysis. Some tests have been used to assess students' knowledge as well as their abilities for data analysis and for reporting and communicating results.

Even though the Italian situation is different from those of most of the other countries, there hare many other situations where an e-learning environment with these characteristics can play an important role. Of course, a limit for e-learning is the so called *digital gap*, that is the different situations that can limit the use of communication technologies. This gap takes place both between industrialized and developing countries and between different strata of population inside industrialized countries. It is a complex theme because it concerns several contexts of human life, i.e. social, cultural, economical and political.

Many obstacles must be overcome before the Internet becomes part of the infrastructures present in each country. Some examples are the lack of necessary IT structures, the high cost of telephone lines, the scarce presence of activity related to computers and data processing training, the geographical spreading of connections concentrated inside big cities and completely absent in rural areas (Roversi, 2004). Training aspect is strictly related to the aim concerning digital gap overcoming. Even in our context, also with reference to university instruction, this gap is present still today among students and that makes some problems rise up.

COMPUTER MEDIATED COMMUNICATION

Over the last ten years computer mediated communication has opened a great number of unbelievable possibilities and has quickly transferred people to a new era dominated by technology. In this context, the Shannon classical linear model, i.e. messages transferred from a sender to a receiver with an ending feedback, required to be updated. A generally accepted definition that describes the new complex reality is known as the *network paradigm* (Rice & Love, 1987). According to this paradigm, communication is seen as a negotiation process among actors operating inside a structured situation.

Using very simple words, the new technology has opened a very great space that allows any kind of expansion in terms of relationships involving a great number of persons. Of course, in this space education has immediately found its own location. In facts, the integration of multimedial tools inside pedagogical and social practice has an important role in order to help individuals, companies and schools in adapting themselves to the new social, economic, cultural environment that characterize the *information society*.

Key points to wonder about, concerning communication research and didactical practice evaluation, deal with the way computer mediated communication resources are used and what are their effects. However, the evaluation process requires several data: basic information, information about the system and the duty, about activity models, about human resources and students reflections and, above all, about results. The assessment must determine whether a resource is useful, as well as providing an understanding of how and why it becomes so useful, in order to apply it inside other contexts.

The didactic usage of integrated wide band communication arises from specific needs of some training contexts, first of all university that, because of scarce resources, don't succeed in fulfilling an instruction enquiry more and more growing in quantity and quality. For instance, at the University of Perugia e-learning platforms have been in use since the beginning of 2000, mainly as teaching support in place of previous web pages. This usage began in the Department of Statistics, but in a short time many other Departments (mainly in the scientific area) experimented with similar platforms to utilize the tools and the better organization available inside these environments.

Previous reasons and experiences have motivated a new project, set up at the University of Palermo, to provide a new kind of support for students attending an introductory course of Statistics within the degree course on Science and Techniques of Growth and Education Psychology. The main goal of the project is to enable students of very crowded classes (i.e. students that cannot have suitable facilities available) to analyze simple data sets and to report and "communicate" the basic statistical characteristics of these data.

TEACHING AND IT: A BIT OF HISTORY

Computers have always been seen as a potential support for teaching, even before the multimedial era. For instance, Suppes (1981) reports on various experiments carried out during the '70s for setting up "official" university courses completely handled by computers, however everything was limited by the lack of graphical and audio interfaces. Personal computers and Unix workstations with their new multimedia facilities gave a strong boost to new teaching tools testing. Starting from end of the '80s many books included sooner diskettes and later CDs trying to stimulate the learning process (as well as the sales, thanks to these novelties!).

Statistics has not been deeply involved in this early process: technology has been mainly used to provide books with data sets or to produce tapes with course lectures for distance learning (AMS, Open University, etc).

During the second half of the '80s a new computer network, ARPAnet (the former name of the Internet), became popular among academic and research institutions. The growth of this network and of some multimedia technologies (both hardware and software) made it possible to experiment with new teaching approaches. Hypertext technology and new powerful graphic tools allowed management of the information spread out on the network in a complete new way (Gopher was one of the most meaningful example) and allowed reconsideration of the role of computers as teaching support. However, the new deal for networking has been driven by the World Wide Web, especially after the first graphic browser, Mosaic, was released in a public version by NCSA (US National Center for Supercomputing Applications) on September 27, 1993.

Starting from that date, almost every technological approach has been reconsidered in a wider way to take into account the newest horizons; in fact, someone dates the e-learning birthday at the beginning of 1994. However, in our opinion, it is not important to fix any birthday, partly because everything depends on what we consider as e-learning. We have a large variety of definitions, some coming from social sciences, some from computer science, some referring to specific techniques and tools, and some oriented to interpret the term in a very broad sense. What is relevant to understand here is how distance learning has become possible and if and how this teaching approach has come to maturity.

The new WWW facilities enabled the establishment of many different environments to support teachers or just to allow interested people (students as well as teachers not yet skilled on statistics) to better understand, or to deepen, some topics. On-line course material, electronic books, virtual labs, applets with simulated examples, animations, have quickly populated the Web, as well as lot of data sets (Galmacci, 2001). At the same time, technology made possible video-conferences using private lines or large band network connections, with low cost software and hardware. All these resources widely stimulated experiments to explore new teaching approaches and to adapt them to the students' learning processes.

A natural evolution of these efforts has been the e-learning software platform, which appeared in the first half of the '90s. Such platforms, i.e. well organized environments to manage the various tasks for distance teaching and learning, soon became very popular among public and private organization to set up low cost internal training courses or to provide customer technical trainings.

Statistics was not strictly dealt with in this specific process (business corporations were mainly interested in promoting their own products), however during the last decade many universities have experimented with these tools for statistics courses (see Aydýnlý, Härdle & Rönz, 2003; Stephenson, 2002), often using free software e-learning platforms.

However, one of the most relevant initiatives in this context was promoted by the US government that "… in 1997 brought together government, industry and academia to launch the ADL (Advanced Distributed Learning) Initiative. Through this collaborative process, the ADL Initiative became the Department of Defense (DoD) wide strategy for using learning and information technologies to modernize education and training and to promote cooperation for developing e-learning standardization …" (*http://www.adlnet.org/*). One of ADL tasks is the development of specifications and guidelines for courses: SCORM (Sharable Content Object Reference Model) is the most relevant issue, "a reference model that defines the interrelationship

of course components, data models and protocols so that learning content objects are sharable across systems that conform with the same model".

Recently a new term has been introduced, m-learning, that points out a new technological step: the mobile era. Mobile computing is based on wireless communications (WiFi and mobile phones) and allows a much more flexible use of the network resources, regardless of the "client location". Then, on-line information can be available everywhere and can be used as a handbook for quick flying assistance.

E-LEARNING SUPPORT FOR LAB ACTIVITIES: A PROPOSAL

The idea to set up an e-learning environment for lab activities described in this paper, as previously stated, was developed to help students of an introductory course of statistics to learn and practice simple data analysis and software tools by themselves. The Science of Education Faculty of the University of Palermo requires all students to attend classes on Basic Statistics, regardless of the specific degree course. Unfortunately, the global number of Faculty students is very high (several thousands) and consequently Statistics classes have many students (generally more than 200). In this situation it is impossible to provide lab facilities for all of them (not enough computers and no time to rotate small groups for assisted activities).

The role of a course on Statistics in this Faculty is mainly concerned with giving students the main elements and keywords to communicate statistical concepts, starting from data that generally describe social facts (population characteristics and behavior, psychological aspects, consumer trends, exit polls, etc.). To attain this goal, lab activities for data analysis cannot be left aside. Therefore, the only possibility to get student started with rudimentary computational methods for statistical analyses is to provide good tutorials that can guide them, step by step, to handle data, to understand the results and to report them in a clear and correct speech.

At the beginning of our project we had to make some preliminary decisions: first, the choice of statistical software and secondly, the choice of an e-learning platform. The first choice was almost that are easy to use, statistically correct and intuitive, all of them are very expensive. Furthermore, in our opinion, it is difficult to imagine professionals that need statistics from time to time, willing to invest money for such software when they have a spreadsheet already available. Thus the need for a spreadsheet-based software was evident, We chose Excel, but it can be easily extended to Lotus 123 or StarOffice, OpenOffice, etc. The second choice, of an e-learning platform, required much more care. After a comparison of several platforms (one of them commercial), we selected *moodle*, a SCORM compatible package freely available at *http://moodle.org* and one of the most popular among academics.

Data used for this purpose consist of variables from a questionnaire submitted to first-year students of the degree course on Science and Techniques of Growth and Education Psychology to collect information concerning their family profile and the reason of their choice to come to this Faculty (Galmacci & Milito, 2002). The statistics students' task was to understand the main characteristics of this "population" and to report results, ensuring that concepts were presented plainly, clearly and properly.

E-learning platforms (and particularly *moodle*) facilitate the organization of material in different ways to reflect both the kind of support to be provided and the teaching approach. In our case, the environment has been organized by paragraphs, one for each step of planned activities, that presents, explains and shows how to carry out each phase of the analysis. Moodle provides several tools to stimulate, support and control the activities (chats, forums, wikis), to submit exercises (assignments, quizzes, surveys, etc) and to assess them (rating can use scales based on the theory of "separate and connected knowing").

Preparing the materials, we assumed that some student could not be friendly yet with spreadsheets, moreover, we dealt with the basic typical computational aspects of statistical analyses as well as with the meanings of the indicators. Topics involved in this first release include those of a first course on Statistics, as the data matrix has allowed, namely:

- Data organization in a spreadsheet
- Data recoding (when it is necessary to use the same scale)

- Missing values management
- Class distributions (simple and bivariate)
- Graphical representations
- Mean, median and standard deviation
- Conditional dependence in two ways tables
- Regression and correlation

Each paragraph (in practice, each module) is based on examples showing how to handle data, how to apply formulas using the spreadsheet language (i.e. for recoding data or handling missing values), how to build two ways tables, how to obtain correct graphical representations, how to draw a box-blot, etc. Moreover, examples are discussed to give students the keys for interpreting the statistical meanings and to communicate them. Special care has been devoted to signal the statistical inaccuracies still present in some of the Excel computations (Knüsel, 2002) in order to avoid common mistakes. Since *moodle* allows handling almost any kind of data and applications, the material includes a talk on the use and abuse of statistics in the Italian media (a video-stream provided by the CIRDIS site at *http://cirdis.stat.unipg.it*).

The project has been developed just before the beginning of the course so that, currently, it must be considered a prototype. Students have only been "invited" to use the e-learning environment, without any duty. This test is very important for many reasons: participation reflects a real interest and it is important to monitor how many students will complete the whole work schedule or when and where they stop. Results of this first experiment are not available yet; at the moment, besides the assessments, we are collecting system logs to check students' contacts (frequency, length, time, etc) and activities. Currently, all the material is in Italian (*http://e-learning.unipg.it*) however, after the first experimental period, it will become available in English and opened for guest users.

REFERENCES

Aydýnlý, G., Härdle, W., & Rönz, B. (2003). E-learning/e-teaching of statistics: a new challenge. *IASE Satellite Conference on "Statistics & the Internet"*, Berlin (*http://www.stat.auckland.ac.nz/~iase/publications/6/Haerdle.pdf*).

Galmacci, G., & Milito, A.M. (2000). Statistical education for communicating in modern societies. *Conference on Regional and Urban Statistics. Proceedings SCORUS 2000* (pp. 242-247).

Galmacci, G. (2001). In C. Batanero (Ed), The impact of the internet on research's training, proceedings of the IASE Round Table on "Training Researchers in the Use of Statistics" (pp. 159-169). Tokyo (*http://www.stat.auckland.ac.nz/~iase/publications.php?show=9*).

Galmacci, G., & Milito, A.M. (2002). The effects of some teaching techniques on learning statistics. *International Conference on Teaching Statistics - ICOTS 6*. Cape Town, South Africa.

Knüsel, L. (2002). On the reliability of Microsoft Excel XP for statistical purposes. *Computational Statistics and Data Analysis*, 39(1), 109-110.

Milito, A.M., & Marsala, M.R. (2002). Insegnare e apprendere la statistica a scuola. *Quaderni di ricerca del Dipartimento di Metodi Quantitativi per le Scienze Umane*, Palermo.

Rice, R.R., & Love, G. (1987). Electronic emotion: a content and network analysis of a computer-mediated communication network. *Communication Sciences*, 14, 85-105.

Roversi, A. (2004). Introduzione alla comunicazione mediata dal computer. *Il Mulino*

Stephenson, W.R. (2002). Experiencing statistics at a distance. *International Conference on Teaching Statistics - ICOTS 6*. Cape Town, South Africa.

Suppes, P. (1981). *University-level computer-assisted instruction at Stanford: 1968-1980*. Stanford University, Institute for Mathematical Studies in the Social Sciences.