

Statistics Education as a Field for Research and Practice

Carmen Batanero
University of Granada, Spain
ICME 10, Regular Lecture

Abstract

In this paper, we summarise the contributions from different disciplines to the current state of statistics education and describe examples of recent research. We restrict our analysis to the teaching of statistics at University level, although the situation we picture here can be extended to school statistics. We finally suggest some potential areas for priority research. Our objective is to provide information to those interested in the topic as well as to contribute to the development of research and practice in statistics education at an international level.

Introduction

As argued by Hacking (1990), the most decisive conceptual event of the twentieth century has been the discovery that the world is not deterministic. It is then not surprising that nowadays statistics is widely taught at different levels, due to its usefulness for the daily life, its instrumental role in other disciplines, and its relevance in developing critical reasoning. *“One of the most notable achievements of western societies in the last few decades has been the extension of modern education, including mathematics, to a very substantial proportion of the population”... “It is within this context that the movement for statistics education has taken root”* (Vere-Jones, 1995, p.13).

Didactic materials, educational software, journals, meetings and conferences on teaching statistics have grown spectacularly in the past few years to fulfil the need to train a large number of students. This interest, however, does not just concern mathematicians, and considering statistics education as a sub-field of mathematics education would be too simple. The strong specificity of statistics education is reflected in the philosophical, ethical and procedural questions that are still being debated within statistics and its applications. Furthermore, it is clear that contributions to this field are not just received from mathematics teachers, but from statisticians, psychologists and teachers in other fields, who use statistics as an instrument. In this paper, we summarise these contributions. Given the limitation of space, we restrict our analysis to the teaching of statistics at University level, although the situation we picture here can be extended to school statistics.

Teaching Statistics at University level

Statistics is one of the most widely taught topics at University level, and many institutions have added in the past years the requirement of a data-oriented, or quantitative-literacy course in their core curriculum (Garfield, Chance & Snell, 2001). At this level, statistics is studied mainly as a tool to solve problems in other fields such as education, geography or medicine. These service courses often emphasize the teaching of formulas for calculating statistics (e.g. correlation coefficients or confidence intervals) without much concern towards the data context or interpretative activities. In other cases, the courses are over-mathematised for these students, which often involve meeting concepts of advanced stochastic thinking without any prior or concurrent experience of advanced algebra or calculus. Moreover, although many statistics students are able to manipulate definitions and algorithms with apparent competence, they often lack understanding of the connections among the important concepts of the discipline (Schau & Mattern, 1997) and they do not know what statistical procedure to apply when they face a real problem of data analysis

(Quilici & Mayer, 1996).

Fortunately, increasingly easy access to powerful computing facilities has saved time previously devoted to laborious calculations and encouraged less formal, more intuitive approaches to statistics (Biehler, 1997; Ben-Zvi, 2000). Consequently, changes are recommended in statistics teaching at University level, both in course content and in the teaching approach: “*Note carefully that is not at all clear that statistical skills in the traditional sense are required*”... “*We should ask whether traditional introductions to statistics for general students are too narrow*” Moore (1997, p. 124).

To add complexity to the situation, statistics is taught at University level by lecturers with a variety of backgrounds, the majority of whom are statisticians, but that also includes economists, health care professionals, engineers, psychologists or educators, and very rarely mathematicians or mathematics educators. Education is for these lecturers only a secondary research field. This explains the fact that, until very recently, research in advanced stochastic teaching and learning has not attracted mathematics educators and thus has had a small presence at, for example, the annual Psychology of Mathematics Education Conferences. However, the situation is starting to change, mainly due to the influence of research in psychology and to international activity, that we summarise below.

Research on Advanced Stochastics

Influence from Psychology

Psychological research has shown that many adults, even those that we can consider to be statistically well trained, tend to make erroneous stochastic judgements in their day-to-day life. These errors also extend to problems solved in a teaching setting and have been widely documented in relation to concepts such as randomness, compound probability, association in contingency tables, conditional probabilities, Bayes problems and sampling (Nisbett & Ross, 1980; Kahneman, Slovic, & Tversky, 1982; Scholz, 1991).

Starting from Simon’s theory of bounded rationality, Kahneman and his collaborators developed the *heuristics and biases* programme (Kahneman, Slovic & Tversky; 1982), which was the dominant paradigm for these studies in the early eighties and is still very influential. It assumes that people do not follow the normative mathematical rules that guide formal scientific inference when they make a decision under uncertainty and that, instead, they use more simple judgmental heuristics. Heuristics reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations and are in general useful; however they sometimes cause serious and systematic errors and are resistant to change. For example in the *representativeness* heuristics, people tend to estimate the likelihood for an event based on how well it represents some aspects of the parent population. An associated fallacy that has been termed *belief in the law of small numbers* is the belief that even small samples should exactly reflect all the characteristics in the population distribution. A huge amount of research in stochastic education at both undergraduate and advanced level has been based on the idea of heuristics (see a summary in Shaughnessy, 1992). Recent research on heuristics at advanced level includes assessment (Hirsch & O’Donnell, 2001; Garfield, 2003), studying the effect of teaching experiments on the use of heuristics (e.g., Pfannkuch & Brown, 1996; Barragués, 2002), explaining misconceptions related to concepts such as randomness (Falk & Konold, 1997) in terms of heuristics or providing alternative explanations for incorrect responses to some of the tasks proposed in research about heuristics.

The *abstract-rules* framework (Nisbett & Ross, 1980) assumes that people are able to

acquire a correct statistical reasoning, and that they develop intuitive versions of abstract statistical rules, such as the Law of large numbers, which are well adapted to deal with a wide range of problems, but fail when they are applied beyond that range. These rules are used to solve statistical problems, when we recognise some cues in the problematic situation. However, the rule systems we use naturally (and that can be taught) are pragmatic and are induced in the process of solving recurrent everyday problems. With respect to training in statistical reasoning the suggestion is that “*many of the inferential principles central to the education we are proposing can be appreciated fully only if one has been exposed to some elementary statistics and probability theory*” (Nisbett & Ross, 1980, p. 281). Problem solving is improved, according to these authors, when the sample space is clearly defined, people recognise the role of chance in the experiment, and the context forces the subject to think statistically.

A more recent theoretical framework is the *adaptive algorithms* approach (Cosmides & Tooby, 1996; Gigerenzer, 1994). Adaptive algorithms serve to solve adaptive problems (such as finding food, avoiding predation or communicating) and take a long time to be shaped, due to natural selection. Since their natural environments shape adaptive algorithms they are more effective when the tasks are presented in a format close to how data are perceived and remembered in ordinary life. According to this theory people should have little difficulty in solving statistical tasks if the data are presented in a natural format of frequencies (absolute frequencies) instead of using rates or percentages (Sedlmeier, 1999). Frequency representations lead to simpler algorithms that give immediate accessible solution to many statistical problems, while most people would be unable to use the complex algorithm required when information is given by fractions of percentages (Gigerenzer, 1994).

Sedlmeier (1999) analyses and summarises recent teaching experiments carried out by psychologists that follow one of these theories. The results of these experiments suggest that statistical training is effective if students are taught to translate statistical tasks to an adequate format, including tree diagrams and absolute frequencies. However, learning is assessed through participant’s performance to tasks that are very close to those used in the training, so that it is difficult to evaluate the extent to which students would be able to transfer this knowledge to a wider type of problems. While this research provides relevant empirical information and potential theoretical explanations for students’ difficulties in advanced stochastics, there is still a large amount of work to be done by mathematics and statistics educators to integrate these results and use them to design and evaluate teaching sequences in natural settings where students are expected to meet wider curricular requirements.

The emergence of statistics education

Another strong influence comes from the field of statistics, where interest in education arose since the creation of the Education Committee by the International Statistical Institute (ISI) at a time where underlying concern for both the ISI and the United Nations was the need for better statistical information from the developing countries (Vere-Jones, 1995).

The ISI started to pay more attention to teaching statistics in schools since the mid seventies, where socio-economic conditions in developed countries, frequent use of quantitative information in newspapers and more widespread use of personal computers lead to increasing demand of statistics education for the general citizen. The International Conferences on Teaching Statistics (ICOTS) were started in 1982 by the ISI to bring together statistics teachers at all levels, disciplines and countries and have continued every

four years. These conferences were complemented with a series of Round Table Conferences focussed on specific themes. In 1991 the International Association for Statistics Education (IASE) was created as a separate section of the ISI and took over the organisation of the ICOTS, starting in 1994 and the Round Tables conferences on specific topics: Introducing Data Analysis in the Schools (Pereira- Mendoza, 1993), Role of Technology in Teaching and Learning Statistics (Garfield & Burrill, 1997), Training Researchers in the Use of Statistics (Batanero, 2001) and Curricular Development in Statistics (in 2004, the proceedings are still in preparation).

The journals *Teaching Statistics* first published in 1979 and *Journal of Statistics Education*, started in 1993, soon became main tools to improve statistics education all over the world. This activity was complemented with stochastics working or discussion groups at the International Conferences on Mathematics Education, the Psychology of Mathematics Education Conferences, and in the European, Latin American, Australasian and other regional conferences in Mathematics Education as well as by the Stochastic Thinking Reasoning and Literacy Research Forum started in 1999. The research approach of mathematics educators is far different from that taken by psychologists. On one hand the mathematical and epistemological analyses reveal that the complexity of concepts, tasks and students' responses investigated by the psychologists is often greater than the psychological research assumes, and suggests the need for re-analysing them from a mathematical perspective. On the other hand, cantering on isolated types of tasks does not always reveal in depth the students' understanding of a concept, since responses are sometimes very dependent on the task variables. Theoretical constructs taken from mathematics education also contribute to a different perspective of the same phenomena.

Moreover, although many different theoretical concepts are taken from psychology or mathematics education, some statisticians are trying to develop specific frameworks to describe statistical thinking. The most influential example is the theoretical framework developed by Wild and Pfannkuch (1999) to "*investigate the complex thought processes involved in solving real world problems using statistics with a view to improving such problem solving*" (p. 224). The authors developed a complex four-dimensional framework including four components. The first component is termed the *statistical investigation cycle PPDAC (Problem, Planning, Data, Analysis, Conclusion)*; secondly there is an *interrogative cycle*, the third component in the model describes the *types of thinking*, present in statistical problem solving and finally the model describes a series of *dispositions*, such as curiosity, imagination or scepticism.

A consequence of all the above activity was the publication of a big and varied number of teaching experiences and suggestions. Research in advanced stochastics is also starting to progressively grow under the influence of the IASE and of several funded projects and initiatives by the American Statistical Association (e.g. the Undergraduate Statistics Education Initiative (USEI) and the Consortium for the Advancement of Undergraduate Statistics Education (CAUSE)). Although this research follows a variety of approaches, some research agendas (e.g. Shaughnessy, 1992; Shaughnessy, Garfield & Greer, 1996; Batanero, Garfield, Ottaviani & Truran, 2000), as well as the creation of *Statistics Education Research Journal* with the specific purpose to promote research show that researchers are now interested in linking the isolated pieces of research and in developing general principles and theoretical knowledge in order to conduct quality research in statistical education.

Recent Research in Specific Concepts

To follow I have described some examples of this type of research concerning particular advanced stochastics topics. Additional research suggests that misconceptions and learning difficulties as regards statistical concepts taught at secondary school level still persist at University level. Some examples are randomness (Falk & Konold, 1997), probability (Pfannkuch & Brown, 1996; Hirsch & O'Donnell, 2001), average and variation (Noss, Hoyles & Pozzi, 1999) and comparing samples (Estepa, Batanero, & Sanchez, 1999). In other cases research has focussed on describing the effect of new teaching methodologies in improving the general performance in the introductory statistics course at university level (Keeler & Steinhorst, 1995; Schuyten, Dekeyser & Goemine, 1999 and different papers at ICOTS conferences), students' perception of statistics (Reid & Petocz, 2002), their ability to identify which statistical procedure is adequate to answer a research question (Quilici, & Mayer, 1996; Gadner & Hudson, 1999) or developing test instruments to simultaneously assess a variety of misconceptions (Garfield, 1998, 2003; Hirsch & O'Donnell, 2001).

Association and correlation

The perception of co variation between stimuli, behaviour and outcomes is a main component of human adaptable behaviour, and this explains the interest towards association from clinical, social and development psychologists. Inhelder and Piaget (1955) considered that the evolutionary development of the concepts of correlation and probability are related, and that understanding correlation requires prior comprehension of proportionality, probability, and combinatorics. Inhelder and Piaget's research on the development of the idea of correlation was carried out with adolescents (12-15 year olds), and after them, most research in psychology has focussed on undergraduates and suggests that adult reasoning on association is, as a rule, very poor (Nisbett & Ross, 1980). In these experiments subjects are presented two-way tables consisting of two rows and two columns where a sample of subjects is classified according to the presence/absence of two qualitative attributes, such as eye and hair colour (dark or fair hair, dark or blue eyes). Subjects are asked to judge if there is relationship between the two attributes (i.e. judgement of correlation) in these 2x2 contingency tables and to justify the procedure to reach their conclusion. In Table 1 we describe the data in this type of problem, where a, b, c and d represent absolute frequencies.

Table 1. Data in a 2x2 Contingency Table

	B	Not B	Total
A	a	b	a+b
Not A	c	d	c+d
Total	a+c	b+d	A+b+c+d

Inhelder and Piaget found that some adolescents who are able to compute single probabilities only analyse the cases in cell [a] in Table 1 (presence-presence of the two characteristics A and B). When they admit that the cases in cell [d] (absence-absence) are also related to the existence of correlation, they do not understand that cells [a] and [d] have the same meaning concerning the association, and they only compare [a] with [b] or [c] with [d] instead. The correct strategy should use the comparison of two probabilities $P(B|A)$ and $P(B|NotA)$, and, according to Piaget and Inhelder, recognition of this fact only happens at 15 years of age. Following Piaget and Inhelder, several psychologists have studied the judgement of association in adults, using various kinds of tasks. Association judgement is very poor in general and is influenced by the data format, strength of correlation and previous theories or expectations (see Beyth-Maron, 1982, for a survey).

Estepa and his collaborators started from these results, and, after a mathematical and epistemological analysis of the concept of correlation, designed a questionnaire including a variety of open-ended tasks to give an alternative explanation of pre-university students' strategies and judgments in terms of *misconceptions* regarding correlation (Estepa, 1993; Batanero, Estepa, Godino and Green, 1996). In the *determinist* conception of association, some students expect that correlated variables should be linked by a mathematical function; students with *unidirectional* conception of association only perceive direct association (positive sign of the correlation coefficient), and interpret inverse association (negative sign) as independence; *causal* conception consists of assuming that correlation always involves a cause and effect relationship between the variables; and in the *local* conception subjects base their judgment on only part of the data (e.g. they only use one cell in 2x2 contingency tables). These conceptions have been confirmed in later research (Morris, 1997; 1999).

Estepa and colleagues organised two teaching experiments based on computers and found general improvement in students' strategies and conceptions after teaching, although the unidirectional and causal misconceptions concerning statistical association were harder to eradicate. Using qualitative methods, such as observation, interviews and the analysis of the students' interaction with the computer, they documented specific *acts of understanding* of correlation (Estepa, 1993; Batanero, Godino, & Estepa, 1998). For example, *the study of the association between two variables has to be made in terms of relative frequencies*; however, in the first teaching session the students tried to solve the problems in terms of absolute frequencies. Although the lecturer pointed out this mistake to them at the end of that session, the same incorrect procedure appeared recurrently from the same students through several sessions, until the students eventually overcame this difficulty.

More detailed analysis of the undergraduates' students understanding of correlation and regression was carried out by Sánchez (1999) who found difficulties in translating between the different representations of correlation (verbal description, table, scatter plot and correlation coefficient), and in computing and interpreting the two regression lines. Later Estepa and Sánchez (2001) described students' difficulties in relating the ideas of linear regression, correlation coefficient and covariance lack of distinction between interdependence and unilateral dependence, problems in adequately choosing the dependent and independent variables; and excessive emphasis on linear dependence.

Probability distributions

One main difficulty in teaching a statistics course at university level is making the transition from data analysis to statistical inference. In order to make this transition, students are introduced to probability distributions, with most of the emphasis being placed on the normal distribution. Wilensky (1995, 1997) examined student behaviour when solving problems involving the normal distribution. He defined *epistemological anxiety* as the feeling of confusion and indecision that students experience when faced with the different paths for solving a problem. In interviews with students and professionals with statistical knowledge, Wilensky asked them to solve a problem by using computer simulation. Although most subjects in his research could solve problems related to the normal distribution, they were unable to justify the use of the normal distribution instead of another concept or distribution, and showed a high epistemological anxiety.

Batanero, Tauber and Sánchez (2004) organised a teaching experiment based on the use of computers, which was based on a theoretical framework on the meaning and understanding of mathematical objects (Godino & Batanero, 1998, Godino, 2002). In this model, the

meaning (understanding) of the normal distribution (or other mathematical object) is conceived as a complex system, which contains different types of elements, including problems and situations from which the object emerges; representations of data and concepts, procedures and strategies to solve the problem; definitions and properties of the concepts, arguments and proofs. The setting of this study was an elective, introductory statistics course offered by the Faculty of Education, University of Granada. Taking the course were 117 students (divided into 4 groups), most of whom were majoring in Pedagogy or Business Studies. The instruction for the topic of normal distributions was designed to take into account the different elements of meaning as just described, which were identified by a previous analysis of University textbooks and the changes involved in the same by the use of computer. This helped to define the *institutional meaning* intended for the normal distribution in the teaching experience, which was observed and analysed in detail by the authors, including the written responses of students to selected tasks carried out throughout the teaching.

At the end of the course students were given written questionnaires as well as some open-ended tasks to be solved with the use of computers. In working on these tasks students were provided with a new data file which included two variables for which the normal distribution provided a good fit and other variables for which this was not the case. Each student worked individually with statistical software and produced a written report that included all the tables and graphs needed to support their responses. Students were encouraged to give detailed reasoning. The researchers expected the students to analyse the different variables in the data file, to use different software options and approaches to check the properties of the different variables, and hence determine which variable would best approximate a normal distribution. However, only a moderate percentage (49%) of students selected one of the variables that fitted a normal distribution. Qualitative analyses of responses to the questionnaire, qualitative analysis of written protocols as well as interviews with a small number of students, were used to describe the main tendencies in the students' *personal meaning* for the normal distribution after the instruction, and to identify main agreements and differences with the intended institutional meaning. This included properties assigned by the students to the normal distribution and related concepts, performance in computing algorithms and procedures, including the use of software, understanding of representations, and ability to develop a sound argumentation, as well as errors and semiotic conflicts related to all these elements.

Sampling distributions

Much of inferential reasoning combines ideas about samples and sampling distributions, a theme where many erroneous judgment and inadequate heuristics were discovered in early psychological research following the pioneer work by Kahneman, Slovic and Tversky (1982). More recent research has focussed on students' understanding and tried to improve this understanding in different teaching experiments, often based on the use of computers. Sampling distributions are complex to understand as they require the integration of conceptual knowledge about probability, variability, sampling, random variables and the normal distribution. Although all these contents are taught at the introductory course, and many university students are able to apply each of these concepts, they often lack the ability to integrate all these ideas and apply them in inferential reasoning.

Well, Pollatsek and Boyce (1990) discovered that university students seem to understand that the means of larger samples are more likely to resemble the population mean. However, they do not understand the implications of this fact for the variability of the sample mean. Recognition that the estimates of a population parameter will vary and that

this variation will conform to a predictable pattern was not understood by university students in the research of Rubin, Bruce and Tenney (1991) either. Interviews and cluster analysis helped Finch (1998) to identify different types of understanding of the law of large numbers, including students who reverted the relationships (assuming the variance of estimation increases with the sample size), understanding the decreasing of variance with sample size as explained with changes in the distribution and giving multiple meanings to the Law of large numbers. Lipson (1994) analysed the potential offered by computers to teach sampling distributions and in a later study (Lipson, 1977) she used concept maps as pre-test and post test and found that students' participation in computer simulation activities was associated with a growth in their understanding of sampling distribution, including the idea that the sampling distribution is characterised by shape, centre and spread, and that the sample statistics can be used to estimate the population parameter.

Garfield, del Mas and Chance (1997), developed a simulation software and instructional and assessment materials to guide students' exploration and discovery of sampling distributions. The software allowed students to specify and change the shape of the distribution for a random variable in a theoretical population, simulate sampling for different sampling sizes of that population, and experimentally get the empirical sampling distributions for different parameters. They also designed diagnostic test-items to assess students' understanding of elemental sampling concepts. Later (delMas, Garfield & Chance, 1999) the authors carried out a teaching experiment with three groups of students at the University introductory level, and used the data collected to improve the teaching approach, materials and assessment instruments. They were surprised that after the first series of experiments many of their students still showed serious misconceptions about sampling distribution. For example, the variability they expected in the distribution was not consistent with the sample size and they did not understand that the sampling distribution would resemble a normal distribution with increasing sample sizes. Much better results were obtained when they tried a model of learning through conceptual change where 141 students were asked to test their predictions and confront their misconceptions, following a constructivist model of learning. The authors concluded that students with misconceptions have to experience contradictory evidence and take note of this contradiction with their previous expectations before they change their views about random phenomena.

Some Priority Areas For Research In Advanced Stochastics

The above overview suggests that probability and statistics pose important challenges to research in the area of advanced mathematical thinking. Existing research has been carried out by different scientific communities, not just by mathematics educators, and for this reason the sources of information are very widely spread and not always easily available. At the same time, the diversity of research problems and approaches is very wide, something which is also the case for undergraduate stochastics. *"At first sight statistical education might seem a rather narrow topic"... "a closer look, however, reveals that just the opposite may be closer to the truth"... "the last 50 years have seen statistical education grow from a narrow focus on training professional stand to a movement that stretches downward into the primary and even the kindergarten programme and outwards, through the training for a wide range of academic and technical disciplines"* (Vere-Jones, 1995, p. 16).

Given the changes implied by technology on the practice and meaning of statistics, Shaughnessy, Garfield and Greer (1996) suggested we should focus our research on the influence of technology on the teaching and learning of statistics, including case studies to document students' learning as well as analysing teachers' professional development. More

specifically Rossman (1997) identified correlation, sampling distribution and confidence intervals as three priority particular research areas, since they are fundamental ideas whose understanding can be enhanced with the help of technology. Their suggestions are still applicable to the particular case of advanced stochastics, where the wide current number of statistical procedures taught at University level and the scarce previous research makes it difficult to select a few priority areas for didactical research. To follow I have complemented these research agendas with a few suggestions.

Hypothesis testing and Bayesian inference

Misconceptions in statistics does not just affect students or people with poor statistical training, Psychological and educational research has shown widespread misconceptions among scientists who use statistical inference in their daily work, including those who received a previous statistical training (see Harlow, Mulaik & Steiger, 1997 for a summary of misuses and abuses of statistical inference). These errors concern particularly the test of hypotheses and lead to a paradoxical situation, where, on one hand, a significant result is required to get a paper published in many journals and, on the other hand, significant results are misinterpreted in these publications (Falk & Greenbaum, 1995; Lecoutre & Lecoutre, 2001).

A particularly misunderstood concept is the *level of significance*, α , which is defined as the probability of rejecting a null hypothesis, given that it is true. The most common misinterpretation of this concept consists of switching the two terms in the conditional probability, that is, interpreting the level of significance as the probability that the null hypothesis is true, once the decision to reject it has been taken. For example, Birnbaum (1982) reported that his students found the following definition reasonable: "A level of significance of 5% means that, on average, 5 out of every 100 times we reject the null hypothesis, we will be wrong". Falk (1986) found that most of her students believed that α was the probability of being wrong when rejecting the null hypothesis at a significance level α . Similar results were found by Vallecillos (1994) in 436 University students from different backgrounds (statistics, medicine, psychology, engineering and business studies) and Krauss and Wassner (2002) in University lecturers involved in the teaching of research methods. More specifically they found that 4 out of every 5 *methodology instructors* have misconceptions about the concept of significance, just like their students. The level of significance is not the only concept misunderstood in significance testing, but there is also confusion between the roles of the null and alternative hypotheses (Vallecillos, 1994, 1995) as well as between the statistical alternative hypothesis and the research hypothesis (Chow, 1996).

The wide research on students' and professionals' misunderstanding and misuse of statistical tests has not been followed by related efforts in designing and evaluating teaching experiments oriented to help students and researchers overcome these difficulties. Recommendations to substitute or complement statistical tests with confidence intervals (e.g. Wilkinson, 1999) do not take into account the fact that their appealing feature is based on a fundamental misunderstanding (Lecoutre, 1998), which consists of thinking of the parameters as random variables and assuming that confidence intervals contain the parameters with a specified probability.

Bayes' thinking is more intuitive than the frequentist probability for students and reflects their everyday thinking about uncertainty better. It is for this reason that some researchers (e.g. Lecoutre & Lecoutre, 2001) are suggesting changing the practice and teaching of

statistics towards Bayesian methods. However, although there is an increasing number of publications about how best introduce Bayesian concepts to students coming from non scientific specialities (e. g. Berry, 1997, Albert & Rossman, 2001) reported results from experiments or research focussed on teaching Bayesian statistics are very limited. Moreover, some experiences reported suggest that students can make mistakes in interpreting their inferential results or in specifying their prior distributions (Albert, 2000). Given the relevance of a correct understanding and application of statistical inference to improve empirical research (including research in mathematics education) this is a possible area where research is an urgent need. Both the teaching of classical and inferential statistics are clear priority areas for further research in advanced stochastics.

Different orientation for different students

On the other hand, the influences from different philosophical views about randomness, probability and statistical inference are still reflected in stochastics teaching and research: *“in the field of probability there continues an ongoing fierce debate on the foundations, even though for pragmatic reasons only, this debate has calmed down in recent times”* (Borovcnik & Peard, 1996, p.239). These conceptions influence changes in the role given to probability within the curriculum, from being the central core to trying to teach statistics without resort to probability (focus on exploratory data analysis only) or favouring classical, Bayesian, computer-intensive (resampling methods) or mathematical-abstract approaches to inference. However, there is no consensus about which of these approaches is better suited for a particular type of student or which are best ways to introduce a given approach. We need to find which approach is needed for each type of student and which are the best ways to introduce a given approach.

In spite of our efforts intuitions do not progress with teaching and misconceptions remain after formal training in statistics. We should concentrate on clarifying why current teaching of statistics does not improve stochastics intuition and think of some alternative ways, for example including some elements of psychology or philosophy in this teaching. *“In any case statistics should be taught in conjunction with material on intuitive strategies and inferential errors”... “It seems to us that this would have the advantages both of clarifying the underlying principles of statistics and probability and of facilitating an appreciation of their applications to concrete judgmental tasks”* (Nisbett & Ross, 1980, p.281).

Web based teaching

Advances in technology and increasing student enrolment numbers have led many universities to offer on-line courses although few studies have compared online and traditional methods of teaching of advanced statistics and results are inconclusive. While Hilton & Christensen (2002) evaluated the impact of incorporating multimedia presentations into the traditional lecture format and found that it did not improve student's learning or attitudes, Utts et al., (2003) found that students in a traditional and mixed setting (using Internet and traditional teaching) had similar performance. At the IASE Satellite Conference on Statistics Education and the Internet, most presentations focussed on analysing Internet resources for teaching statistics or presenting examples of these resources.

Theoretical frameworks that take into account the new meanings of statistical concepts in virtual environments and research methods that are useful to analyse the enormous amount of information generated in the interaction between students and teacher in such an environment are an urgent need. Given the increasing access of students to virtual learning and the extensive and varied resources available there is a need to analyse how the meaning

of statistics concepts is affected by these resources and how can we incorporate them in the teaching both in group and individual learning.

The nature of advanced stochastics

Whilst in other areas of mathematics the boundaries between elementary and advanced thinking are reasonably well defined, such a distinction is much more fuzzy in the case of stochastics, where current secondary school curricula in many countries include ideas about association and inference. Also, situations which require mature stochastic thinking for correct interpretation, such as voting, investment, research planning or quality control increasingly form part of the information to which many citizens and professionals are exposed. So, while advanced mathematical thinking tends to be used only in a formal way and after some systematic training, advanced stochastic thinking is now being formally or informally used by many people with little formal training in either mathematics or stochastics.

Moreover some apparently simple concepts, such as randomness, independence and variation are complex because each of them describes a separate continuum, are interconnected and usually are taken for granted in teaching (Gal, in press). For example, judging whether a sequence of outcomes is likely to have been randomly generated involves being able to measure a number of parameters including relative frequency, length and distribution of runs, variation found in subsets, estimating probabilities, etc. Furthermore, since no random sequence of outcomes exactly fits the expected patterns suggested by probability theory, judging randomness in a particular situation also requires an understanding of the logic of hypothesis testing and the features of sampling distributions. These are both advanced stochastic ideas where students have many difficulties, and whose understanding, in turn, actually rests on the idea of randomness (and we have thus a circular situation).

We therefore need to reflect on the exact amount of formalism needed to teach these concepts and what the best way to use computers to help us in this teaching is. In this sense, research in statistics and probability can be paradigmatic for finding ways and approaches to introduce advanced mathematical ideas to wider audiences, and to rethink the very meaning of what is advanced mathematical thinking.

References

- Albert, J. (2000). Using a sample survey project to assess the teaching of statistical inference, *Journal of Statistical Education*, 8. On line: [<http://www.amstat.org/publications/jse/secure/v8n1/albert.cfm>]
- Albert, J. H., & Rossman, A. (2001). *Workshop statistics. Discovery with data. A bayesian approach*. Bowling Green, OH: Key College Publishing.
- Batanero, C. (2001) (Ed.). *Training researchers in the use of statistics. Proceedings of the 1996 IASE Round Table Conference*. Granada, Spain: International Statistical Institute.
- Batanero, C., Estepa, A., Godino, J. D. & Green, D. R. (1996) Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27(2), 151-169.
- Batanero, C., Tauber, L., & Sánchez, L. (2004). Students' reasoning about the normal distribution. In J. B. Garfield & D. Ben-Zvi (Eds.). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht, Netherlands: Kluwer.
- Batanero, C., Garfield, J. B., Ottaviani, M. G., & Truran, J. (2000). Research into statistical education: Some priority questions. *Statistics Education Research Newsletter* 1(2), with discussion in *SERN* 2(1) and response in *SERN* 2(2). On line:

[<http://www.ugr.es/~batanero/Newsletters/Newsjan01.PDF>]

- Batanero, C., Godino, J., & Estepa, A. (1998). Building the meaning of statistical association through data analysis activities. In A. Olivier, & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 221-236). Stellenbosh, South Africa: University of Stellenbosh.
- Barragués, J. I. (2002). *La enseñanza de la probabilidad en primer ciclo de universidad. Análisis de dificultades y propuesta alternativa de orientación constructivista* [Teaching probability at introductory University level. Analysis of difficulties and alternative constructivist proposal]. Unpublished Ph.D. University of the Basque Country, San Sebastian, Spain.
- Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning* 2(1&2), 127-155.
- Berry, D. A. (1995). *Basic statistics: A Bayesian perspective*. Wadsworth: Belmont.
- Beyth-Marom, R. (1982). Perception of correlation reexamined. *Memory and Cognition*, 10(6), 511-519.
- Biehler, R. (1997). Software for learning and for doing statistics». *International Statistical Review*, 65(2), 167-190.
- Birnbaum, I. (1982). Interpreting statistical significance. *Teaching Statistics*, 4, 24-27.
- Borovcnik, M., & Peard, R. (1996). Probability. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook in mathematics education* (Part 1, pp. 239-288). Dordrecht: Kluwer.
- Chow, L. S. (1996). *Statistical significance: Rationale, validity and utility*. London: Sage.
- Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58, 1-73.
- delMas, R. C., Garfield, J. B., & Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical thinking. *Journal of Statistics Education*, 7(3), 1-16 [On Line]: <http://www.amstat.org/publications/jse>
- Estepa, A. (1993). *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores* [Students' preconceptions on statistics association and its evolution after a teaching experiment based on computers]. Unpublished Ph.D. University of Granada, Spain.
- Estepa, A., Batanero, C., & Sanchez, F. T. (1999). Students' intuitive strategies in judging association when comparing two samples. *Hiroshima Journal of Mathematics Education*, 7, 17-30
- Estepa, A., & Sánchez-Cobo, F. (2001). Empirical research on the understanding of association and implications for the training of researchers. In C. Batanero (Ed.), *Training researchers in the use of statistics* (pp. 37-51). Granada, Spain: International Statistical Institute.
- Falk, R. (1986). Misconceptions of statistical significance. *Journal of Structural Learning*, 9, 83-96.
- Falk, R., & Greenbaum, C. W. (1995). Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory and Psychology*, 5 (1), 75-98.
- Falk, F. & Konold, C. (1997). Making sense of randomness. Implicit encoding as a basis for judgment. *Psychological Review*, 104, 310-318.
- Finch, S. (1998). Explaining the law of large numbers. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (v.2, pp. 731-736). Singapore: International Statistical Institute.
- Fischbein, E., & Grossman, A. (1997). Schemata and intuitions in combinatorial reasoning. *Educational Studies in Mathematics*, 34, 27-47.

- Gadner, P. L. & Hudson, I. (1999). University students' ability to apply statistical procedures. *Journal of Statistics Education*, 7(1). On line [<http://www.amstat.org/publications/jse/v4n1>].
- Gal, I. (In press). Towards "probability literacy" for all citizens: building blocks and instructional dilemmas. In G. Jones (Ed.), *Exploring probability in school. Challenges for teaching and learning*. Dordrecht: Kluwer.
- Garfield, J. B. (1998). The statistical reasoning assessment. Development and validation of a research tool. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (v.2, pp. 781-786). Singapore: IASE.
- Garfield, J. B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, 1(2), 22-38.
- Garfield, J. & Burrill, G. (editors) (1997). *Research on the Role of Technology in Teaching and Learning Statistics*, Proceedings of the 1996 IASE Round Table Conference. Voorburg, The Netherlands: International Statistical Institute.
- Garfield, J., Chance, B., & Snell, J. L. (2001). Technology in college statistics courses. In D. Holton et al. (Eds.); *The teaching and learning of mathematics at university level: An ICMI study* (pp. 357-370). Dordrecht: Kluwer.
- Garfield, J. B., del Mas, B., & Chance, B. (1997). Assessing the effects of a computer microworld on statistical reasoning. Paper presented at the *Join Meetings of the American Statistical Association*, Anaheim, CA, August, 1997. On line [www.gen.umn.edu/faculty_staff/delmas/asa_1997_paper/asa_paper.htm].
- Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice-versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129-161). Chichester: Wiley.
- Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática [An ontological and semiotic approach to mathematical cognition]. *Recherches en Didactique des Mathématiques*, 22(2/3), 237-284.
- Godino, J. D., & Batanero, C. (1998). Clarifying the meaning of mathematical objects as a priority area of research in mathematics education. In A. Sierpiska & J. Kilpatrick (Eds.), *Mathematics Education as a research domain: A search for identity* (pp. 177-195). Dordrecht: Kluwer.
- Hacking, I. (1990). *The taming of chance*. Cambridge, MA: Cambridge University Press.
- Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (1997). *What if there were no significance tests?* Mahwah, NJ: Lawrence Erlbaum Associates.
- Hilton, S. C. & Christensen, B. H. (2002). Evaluation of the impact of multimedia lectures on student learning and attitudes. In B. Phillips (Ed.), *Proceedings of the sixth International Conference on the Teaching of Statistics*. Cape Town: International Statistical Institute. CD ROM.h
- Hirsch, L., & O'Donnell, A. M. (2001). Representativeness in statistical reasoning: identifying and assessing misconceptions. *Journal of Statistics Education*, 9(2). On line [<http://www.amstat.org/publications/jse/v9n2>].
- Inhelder, B., & Piaget, J. (1955). *De la logique de l'enfant à la logique de l'adolescent* [From the child's logic to the adolescent's logic]. Paris: Presses Universitaires de France.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Keeler, C. M. & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, 3(2). On line [<http://www.amstat.org/publications/jse/v4n1>].

- Krauss, S. & Wassner, C. (2002). How significance tests should be presented to avoid the typical misinterpretations. In B. Phillips (Ed.), *Proceedings of ICOTS-6*. CD Rom. Ciudad del Cabo: International Association for Statistical Education.
- Lecoutre, B. (1998). Teaching bayesian methods for experimental data analysis. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (v.1, pp. 239-244). Singapore: International Statistical Institute.
- Lecoutre, B., & Lecoutre, M. P. (2001). Uses, abuses and misuses of significance tests in the scientific community: Won't the Bayesian choice be unavoidable? *International Statistical Review*, 69(3)-399-417.
- Lipson, K. (1994). Understanding the role of computer based technology in developing fundamental concepts of statistical inference. In *Proceedings of the fourth International Conference on Teaching Statistics* (v. 1, pp. 65-72). Marrakech, Morocco: The National Institute of Statistics and Applied Economics.
- Lipson, K. (1997). What do students gain from computer simulation exercises? In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 137-150). Voorburg: International Statistical Institute.
- Moore, D. S. (1997). New pedagogy and new content: the case of statistics. *International Statistical Review*, 65, 123-165.
- Morris, E. J. (1997). *An investigation of students' conceptions and procedural skills in the statistics topics correlation*. Technical report n. 230. Milton Keynes, U.K.: Centre for Information Technology, The Open University.
- Morris, E. J. (1999). *Another look at psychology students' understanding of correlation*. Technical report n. 246. Milton Keynes, U.K.: Centre for Information Technology, The Open University.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgments*. Englewood Cliffs, NJ: Prentice Hall.
- Noss, R., Hoyles, C. & Pozzi, S. (1999). How can the study of mathematics in work advance our understanding of mathematical meaning-making in general. In O. Zalavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education* (vol.3, pp. 353-360).
- Pereira-Mendoza, L. (Ed.) (1993). *Introducing data analysis into schools: Who should teach it and how?*. Proceedings of the IASE Round Table Conference. Voorburg, The Netherlands: International Statistical Institute.
- Pfannkuch, M. & Brown, C. (1996). Building and challenging students' intuitions about probability: Can we improve undergraduate learning? *Journal of Statistics Education*, 4(1). On line [<http://www.amstat.org/publications/jse/v4n1>].
- Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorise statistics word problems. *Journal of Educational Psychology*, 88(1), 144-161.
- Reid, A. & Petocz, P. (2002). Students' conceptions of statistics: A phenomenographic study. *Journal of Statistics Education*, 10(2). On line [<http://www.amstat.org/publications/jse/v10n2>].
- Rossmann, A. (1997). Workshops statistics: Using technology to promote learning by self-discovery. In J. Garfield, & G. Burrill (Eds.). *Research on the Role of Technology in Teaching and Learning Statistics*. International Statistical Institute.
- Rubin, A., Bruce, B., & Tenney, Y. (1991). Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 314-319). Voorburg, The Netherlands: International Statistical Institute.
- Sanchez, F. T. (1999). *Significado de la correlación y regresión para los estudiantes universitarios* [Meanings of correlation and regression for undergraduates]. Unpublished

PhD. University of Granada.

Schau, C. & Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 91-104). Amsterdam: IOS Press.

Scholz, R.W. (1991). Psychological research in probabilistic understanding. In R. Kapadia, & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 213-249). Dordrecht: Kluwer.

Schuyten, G., Dekeyser, H. & Goemine, K. (1999). Towards an electronic learning environment for statistics in higher education. *Education and Information Technologies*, 4(4), 409-424.

Sedlmeier, P. (1999). *Improving statistical reasoning. Theoretical models and practical implications*. Mahwah, NJ: Erlbaum.

Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation*, 18, 147-166.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. Bishop et al. (Eds.), *International handbook of mathematics education* (v.1, pp. 205-237). Dordrecht, Netherlands: Kluwer.

Utts, J., Sommer, B., Acredolo, C., Maher, M. W. & Matthews, H. R. (2003). A study comparing traditional and hybrid Internet- based instruction in introductory statistics classes. *Journal of Statistics Education*, 11(3). On line [<http://www.amstat.org/publications/jse/v11n3>].

Vallecillos, A. (1994). *Estudio teórico experimental de errores y concepciones sobre el contraste de hipótesis en estudiantes universitarios* [Theoretical-experimental study of errors and conceptions concerning statistical tests in university students]. Unpublished doctoral dissertation, University of Granada, Spain.

Vallecillos, A. (1995). Comprensión de la lógica del contraste de hipótesis en estudiantes universitarios. [University students' understanding of the logic of statistical tests]. *Recherches en Didactique des Mathématiques*, 15 (3), 53-81.

Vere-Jones, D. (1997). The coming of age of statistical education. *International Statistical Review*, 63(1), 3-23.

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

Wilensky, U. (1995). Learning probability through building computational models. In D. Carraher y L. Meira (Eds.), *Proceedings of the 19th PME Conference* (Vol. 3, pp. 152-159). Recife, Brazil: Organizing Committee.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, 33, 171-202.