Batanero, C. (Ed.), Training Researchers in the Use of Statistics, 113-116.

Printed in Granada, Spain.

THEODORE CHADJIPADELIS

DISCUSSION

Wang's interesting paper about teaching and using stochastic processes shows that estimating and using transition probabilities is essential, especially in political and social sciences. Predicting the behaviour of a structure (e.g. an educational system, the World Wide Web, a traffic network, an academic organisation) by using stochastic processes and operational research methods gives the opportunity to evaluate the effectiveness of the structure and to control its future evolution.

The papers by McLean and Iversen deal with the need of understanding and explaining the real world. All of us (statisticians and researchers) try to understand and explain the world using theories and data. Statistics is just one of several methods dealing with the above issue. As statisticians, we use data to estimate and predict, and we assign probabilities to express a complex non-deterministic world.

The Greek philosopher Aristotle said that there are three kinds of events:

- Events for which we know the outcome (that is, events with probability one);
- Events for which we can predict the outcome (that means that we can compute a frequency distribution without using data, working with probabilities);
- Events, about which we can say nothing (we only can estimate a frequency distribution using data, working with statistics).

In any case before reaching a conclusion you need a reasonable explanation for every model or outcome in the frame of the original problem. As Kruskal (1979) said,

> *"A scientist confronted with empirical observations goes from them to some sort of inference, decision, action, or conclusion. The end point of this process may be the confirmation or denial of some complicated theory; it may be a decision about the next experiment to carry out. An end point is typically accompanied by a statement, or at least by a feeling, of how sure the scientist is of his new ground. These inferential leaps are, of course, never made only in the light of the immediate observations. There is always a body of background knowledge and intuition, in part explicit and in part tacit."* (Kruskal, 1979, p. 84).

These three papers deal with the need to use "statistics" and "probability" to understand and describe the "real world". All of us agree that, statistics is not a set of rules and recipes for the analysis of data. It is not reduced to the use of complicated computer programs and nice graphs. It requires a good knowledge of the observed phenomenon, the planning of observation, good knowledge of the data gathered, and description and examination of the hypotheses about the parameters of the analysed phenomenon.

There is an old Chinese proverb: "If you give a man a fish, he will have a single meal; if you teach him how to fish, he will eat all his life". The latter concept means that we design courses that encourage students to think creatively and imaginatively

about their scientific research problems and the role of modern theories of statistical data analysis and modelling. Compare this with courses that usually describe the mathematical solutions of various routine statistical analysis problems without much proof, and are based on probability models for observed data whose validity is not usually checked.

But should we make (or try to make) a "statistician" out of a researcher? Or should we make researchers understand that they need a statistician as a helper, a colleague or even as a leader? Potential information can only be useful if it is generated. One reason why some Japanese industries achieve high levels of productivity is that employees are provided with statistical tools for generating, analysing, and acting on their own information.

As a member of the board of the Greek Statistical Institute I have been involved in many training staff seminars for researchers and university lecturers. Although we try to "educate" them in statistics, talking about assumptions, requirements, statistical and stochastical thinking, they always ask for the analysis of their specific data set, and look for examples that are similar to their own problems. We should educate the statisticians in a more practical sense:

> *"The brilliant minds of mathematical statistics would do well to leave the construction of abstract admissible decision functions, cease to ride martingales into the teeth of zero-one laws and join the few of us who are attempting to stem the tide of confusion"* (Hunter, 1981, p. 113).

Hunter (1981) also criticised statistics education in Mathematics departments saying:

> *"The statisticians' training, narrow and technical, is the orderly climb up a staircase of mathematical problems that each have only one right answer. Later steps rest on earlier ones. Progress is always up. Teachers watch the climbing techniques of the fledgling statisticians, and help them master the steps, one at a time. Statisticians' work, for which this training is supposed to equip them, is the disorderly climbing of rugged hills, outdoors, in fair weather and foul. The path is anything but clear. A promising path can get lost in tangled undergrowth or a patch of dense forest. Or else: a path branches in several directions and there is not enough time or money to explore all of them to determine which is the best to follow"* (Hunter, 1981, pp. 113-114).

Statistical literacy and thinking is another issue. It is useful to the general public in understanding and criticising what is written in the press, what is seen on TV, what is presented by the authorities. We should educate the general public in order to become critical citizens. But this is the story of statistical education in compulsory education.

Let me tell you some conclusions I have drawn from the papers at the conference:

- Learn as much as you reasonably can about the general subject matter field and the specific environment in which the data were collected;
- As part of this effort, statisticians need to probe, be curious, and ask good "non-statistical" questions;
- Correlation measured from an observational study does not imply causation; Confusing correlation and causation is particularly troublesome in the social sciences;
- The real problem is often different from the one initially posed;
- An empirical approach is sometimes better than a theoretical one;
- Scientific logic is our business. Statisticians can often be most helpful by getting

perspective on all aspects of a particular problem and then contributing ideas related to scientific method (Hooke, 1980);
- Try to understand what is really going on;
- Valuable data are sometimes non-numerical;
- Scientific inference is broader than statistical inference.

And finally, let me tell you an old story about the practising statistician by Salsburg (1973). David Salsburg asked himself the question, what is it really like to be a practising statistician. Below we reproduce his answer:

> *"The statistician is first called into consultation during the design of a scientific experiment. At this point, the texts tell us the statistician is supposed to estimate minimal sample sizes and prepare a BIBD that produces all kind of clever contrasts for testing.*
> *I do not do this. Instead, I spend my time asking stupid questions. I know that when the experiment is finished I will have to analyse the data. I must protect myself from impending chaos. With such a fear behind me, I ask such questions as whether it is possible to observe something every 15 minutes or whether this thing they have given a name can, in fact, be observed at all.*
> *I ask them what can be wrong. Frequently, I am the only one at the conference with a non-deterministic outlook. The others conceive only three or four clear-cut outcomes. I think about the in-between outcomes, the two correlated variables that happen to go different ways, the test tube someone is bound to drop, the patient who revives from death's door on placebo.*
> *I know that when it is all over with, the man who must make some kind of decision about the results will ask me to compare two means or to show him a linear regression. I try to make sure that the design will produce two comparable means regardless of how many test tubes are dropped and that, somewhere, there will be a somewhat controllable variable manipulating a somewhat responsible variable.*
> *The bulk of my time, however, is spent in trying to make sense out of data… When I see data it is frequently because the results have not made sense to the client… I feel very uneasy with a client who nods blandly and takes back my numbers for his report. I feel better if he argues with me. After all, I do not know an isatine derivative from an isonitroso-acetlyamine, but I hope to God he does…*
> *I suspect that at least 50 per cent of all the data accumulated today never gets more than a cursory look, and I doubt if 5 per cent of it gets examined effectively. All that money, all that anguish, all that pain will have gone for nothing and even be spent again and again in unknowing duplication unless these floods of data are converted to usable information. This is, par excellence, the place for the well-trained statistician…"* (Salsburg, 1973, p. 152).

REFERENCES

Hooke, R. (1980). Getting people to use statistics properly. *The American Statistician*, 34, 39-42.

Hunter, W. G. (1981). The practice of statistics: The real world is an idea whose time has come. *The American Statistician,* 35, 72-76.

Hunter, W. G. (1981). Six statistical tales. *The Statistician, 30,* 107-117.

Kruskal, W. H. (1979). Comments on Dr Braham's article. *Journal of the American Statistical Association,* 74, 84-86.

Iversen, G. (2001). Should we even teach statistics? A Bayesian perspective. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 103-112). Granada: International Association for Statistical Education and International Statistical Institute.

McLean, A. (2001). Statistics on the catwalk. The importance of models in training researchers

in statistics. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 87-101). Granada: International Association for Statistical Education and International Statistical Institute.

Salsburg, D. S. (1973). Sufficiency and the waste of information. *American Statistician*, 27, 152-154.

Wang, A. L. (2001). How much stochastics processes do we teach and to whom? In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 73-85). Granada: International Association for Statistical Education and International Statistical Institute.

Theodore Chadjipadelis
Aristotle University of Thessaloniki,
Department of Education,
54006 Thessaloniki, Greece.
E-mail: chadji@eled.auth.gr