

LAX, P. D. (Chairman). (1982). *Report of the Panel on Large-scale Computing in Science and Engineering*. National Science Foundation, Washington, D. C.

RHEINBOLDT, W. C. (Chairman). (1985). *Future Directions in Com-*

*putational Mathematics, Algorithms, and Scientific Software*. Report of the Panel on Future Directions in Computational Mathematics, Algorithms, and Scientific Software. SIAM, Philadelphia.

# Comment

Jessica Utts

My discussion will be divided into two parts. The first part consists of a treatise on the responsibility which accompanies the use of computers in statistical research. I offer several recommendations to complement those in the article.

The second part is a short description of a setup which works fairly well at the University of California at Davis and was not mentioned in the report. It might be of interest to other statistics departments.

## 1. SCIENCE FICTION OR FUTURE FACT?

There has been a science fiction novel living in my head for the past 10 years or so. It started when I was a graduate student studying robustness and I realized that most users would think of the computational aspects of robust procedures as a black box. This story occurs 30 to 40 years in the future. There are no more statisticians. There are statistical clerks, and every university department has at least one. Research is done by collecting data and giving it to the statistical clerk, who takes it from there. The clerk feeds the data into the computer and out pops the appropriate model, estimate, or whatever, complete with the associated significance or confidence levels. These are sent to journals, along with a post hoc explanation for the results of any of the tests which turned out to be "significant." Everyone is quite happy with this arrangement. No one knows how the computer generates these answers, but everyone knows that if the computer produced them, they must be right. All sorts of interesting (and not so interesting) hypotheses are being proved this way, and when they don't agree with common sense, everyone knows that common sense must be wrong.

In the current version of the story, something finally goes wrong. I haven't worked out the details, but it is a result which contradicts common sense so much that someone (a fresh young scientist, of course) actually has the audacity to question what is happening in the

computer. In order to determine what the computer should be doing, a team of scholars attempts to decipher the statistical literature. To their dismay, they find that the literature is unreadable to them. Finally, they locate a few old statisticians who have long since retired, and with their help they piece together the story. It seems that when the computer software was being developed, most statisticians didn't pay much attention. The packages which were eventually implemented were written by people who were good at selling, but who didn't really understand the concepts involved. A few statisticians tried to protest, but since they were advocating the use of their own services, no one took them seriously. After all, the journals were much more likely to publish the computerized version of the results, so why bother with the more cautious and complicated interpretations the statisticians were trying to sell?

Of course I will never write this novel, but if things continue on their present course I may very well watch it unfold from science fiction into future fact. There are even those who believe that it is already well under way. One of our graduate students told me that a recent cocktail party response to his statement that he was studying statistics was "aren't you afraid of being replaced by a computer?"

So am I against the use of computers in statistical research? Of course not. In fact, I embrace these developments. After all, the world is a complex, non-normal, non independent and identically distributed place and complex models are much more likely to accurately describe reality. Tools like the bootstrap, high resolution graphics, and interactive data analysis programs are important and useful for applied statisticians.

What I advocate is that we as research statisticians begin to play a greater role in determining that our work is properly applied. Our techniques are simultaneously becoming more complex and more automated. They are less and less likely to be understood by nonstatisticians. I was concerned when people started using calculators which give regression coefficients without producing plots. But the potential for misuse

---

*Jessica Utts is Associate Professor, Division of Statistics, University of California, Davis, California 95616.*