PRACTICE OF SCIENCE

Weighing the Evidence

Marc Mangel

A better understanding of evidence will always be important in science. At my university, for example, the Center for Informal Learning and Schools (*1*) brings together students and post-docs from the natural and social sciences. Topics they continually discuss include what constitutes data, what is evidence, and how is evidence used to draw conclusions.

The goal of *The Nature of Scientific Evidence* is to help answer those questions. To do so, Mark Taper (an ecologist at Montana State University) and Subhash Lele (a statistician at the University of Alberta) have drawn authors from the fields of ecology, statistics, and philosophy. The choice of ecology as the illustrative science is a good one, because ecology has strong traditions in both the discovery of new knowledge and the application of that knowledge to important problems of society. The chapters are grouped in five sections: "Scientific Process"; "Logics of Evidence"; "Realities of Nature"; "Science, Opinion, and Evidence"; and "Models, Realities, and Evidence." Each chapter is followed by commentary, typically from two individuals, and a rejoinder by the author. The volume is aimed toward students as well as established scientists, statisticians, and philosophers. It reaches its target: there is something in it for everyone.

As might be expected of an edited volume, the technical level of the chapters varies considerably. Some of the mathematically easiest material (introductory in nature, where statistical symbols are explained) appears late in the book. There are many big ideas, but they are scattered around and one needs to work through (or at least, as Solly Zuckerman reputedly said, hum through) the technical details in order to get to them. As in most edited volumes, there is too much repetition. However, the book is a rare find: a source that could be used in graduate seminars in statistics, philosophy, or biology if the chapters are suitably chosen. It is brimming with ideas.

The Nature of Scientific Evidence
Statistical, Philosophical, and Empirical Considerations
Mark L. Taper and Subhash R. Lele, Eds.

University of Chicago Press, Chicago, 2004. 585 pp. $85, £59.50. ISBN 0-226-78955-1. Paper, $30, £21. ISBN 0-226-78957-8.

A number of broad themes, including the definition of the scientific process and the structure of scientific knowledge, weave throughout the text. Evidence enters science in a variety of ways. Contributor Richard Royall has noted elsewhere (*2*) that given a set of observations, one may ask: What should I believe? What should I do? And how should I interpret the observations as evidence? The third question can be split and distilled to, What do the observations tell me about the truth of the hypothesis being considered? and What do they tell me about the hypothesis's predictive accuracy? The topic of whether we are seeking the truth (which, it appears, statisticians are more likely to believe) or increasingly better understanding of reality (which scientists are more likely to believe) also appears in many chapters.

The operational issues center on the conflict between frequentist and Bayesian approaches to statistics, approaches that differ primarily in their notions of the relationship between data and hypotheses and in their treatments of prior information. The authors, like most statisticians and statistically oriented scientists, assume that it is necessary to choose between these two paradigms, instead of asking for the virtues of each that will help us gain a better understanding of the world. I would guess that most scientists are facultative Bayesians. They find frequentist statistics troubling because it often does not tell scientists what they really want to know; the trouble with Bayesian statistics is that the prior is described as "subjective," with all that this charged word connotes. At times the writing is vituperative (making it kind of fun to read), with Bayesians and frequentists attacking one another with well-known saws. But saws only cut, not build. One contributor notes that "statistics today is a conceptual and theoretical mess," and the treatment of Bayesian and frequentist statistics in the volume does little to help solve that problem. Thus, the two schools of statistical thought, while battling among themselves (in a kind of last-statistician-standing showdown), let scientists down. Perhaps not unexpectedly, John Hammersley once wrote, "Scientists have learned to expect everything from mathematicians short of actual help" (*3*).

Philosophers can help scientists understand the operation of science. In her chapter, Deborah Mayo summarizes the philosophical foundation of frequentist statistics in a précis of her earlier book (*4*), but the volume lacks a comparable treatment of the philosophical foundations of Bayesian statistics [as in (*5*)]. A philosophy of evidence should be inclusive. In addition to providing statements of evidence, it should show how to generalize, model, and use or discard data as well as how to deal with contradictory data. Unfortunately, the contributors offer little discussion of how one makes the trade-off between a firm philosophical foundation and gaining a deeper understanding of the natural world, especially for nonreductionist questions. The chapters that are the most oriented toward scientific problems are also those that have the least statistics.

The volume would have been improved by the inclusion of one illustrative problem, treated by both methods, that shows the weaknesses, strengths, and commonalities of each approach to evidence. However, frequentists and Bayesians agree on the general conclusions that evidence is comparative and that data may support one hypothesis over another (their argument is in quantifying that support) but the support for a single hypothesis cannot be quantified. At the philosophical level, most contributors would agree that although we will never know the truth, we might reach increasingly better understanding of nature. Scientists tend not to reject a theory that has some explanatory power or predictive power, even if it fails in other cases, when there is no alternative theory available (they'd have nothing to do). At the scientific level, most would agree that we need to carefully choose a model—or models, so that we can compare multiple models with the data—and know what assumptions are being made, so that we can separate the useful models from the others. At the statistical level, especially in these days of easy computing, we need to really understand the statistics that we are using.

The Nature of Scientific Evidence is far from perfect, but the volume is valuable and important. It deserves a read by everyone.

The reviewer is in the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA. E-mail: msmangel@ams.ucsc.edu

References and Notes

1. The center is a partnership among the University of California Santa Cruz; King's College London; and the Exploratorium, San Francisco; www.exploratorium.edu/cils.
2. R. M. Royall, *Statistical Evidence: A Likelihood Paradigm* (Chapman and Hall, London, 1997).
3. J. M. Hammersley, *Bull. Inst. Math. Appl.* **10**, 235 (1974).
4. D. G. Mayo, *Error and the Growth of Experimental Knowledge* (Univ. Chicago Press, Chicago, 1996).
5. C. Howson, P. Urbach, *Scientific Reasoning: The Bayesian Approach* (Open Court, La Salle, IL, 1989).