STATISTICS AND SCIENCE, OBJECTIVITY AND TRUTH: COMMENTS ON DENNIS 2000

Aaron M. Ellison
Center for Environmental Literacy
and
Department of Biological Sciences
Mount Holyoke College
South Hadley, MA 01075-6418

phone: 413-538-2110
fax: 413-538-2548
e-mail: aellison@mtholyoke.edu

I comment on three substantive issues raised by Dennis (2000): (1) the scientific method; (2) differences between Bayesian and frequentist statistics; and (3) statistical education for ecologists. I draw inspiration for this essay from the writings of Neil Postman (e.g., Postman and Weingartner 1969, Postman 1995), Paul Feyerabend (1975, 1978, 1987) and David Orr (1992), who, being firmly rooted in the postmodernist camp, provide valuable insights about education, science, and society.[1]

Scientists construct convincing explanations and acquire reliable knowledge. We do this through rigorous application of "the scientific method", a device for eliminating or reducing points of reasoned skepticism. All scientists share a common vision that objective truth is "out there", but because there is a diversity of reasons that we become scientists, we approach that truth in many ways (Feyerabend 1975, 1978). In the end, however, the "collective process of empirical investigation, weeding out of untenable notions, and careful checking of working hypotheses" is independent of frequentist or Bayesian (or other) methods (Galileo had neither to use); that this process leads to *progress* is a core belief of modernism (see footnote 1). Straight-jacketing ourselves to one method, as advocated by Dennis, recalls an observation by Heisenberg (1958): "[w]e have to remember that what we observe is not nature itself, but nature exposed to our methods of questioning", and may slow that progress, if in fact, it exists. "Tobacco-company science" can benefit by having only one acceptable method of scientific inquiry, and many

---

[1]Dennis' critique of Bayesian inference is motivated in large measure by a strong aversion to the lapping waves of "popular postmodernism" on the shoals of science, and the impact of feminist theory in the sciences. In reality, Dennis is concerned about the distinction between *relativism* and *objectivism* within the scientific method (see Feyerabend 1987 for an historically-informed, philosophical review). *Postmodernism* is a reaction to *modernism* and denies an unbridled faith in progress (see Cahoone 1996). Dennis further muddies the waters by lumping *deconstructionism*, a method of literary criticism invented by Paul de Man (1989) and Jacques Derrida (1967) with postmodernism. Although many postmodernists are also relativists (see Gross and Levitt 1994) and engage in literary (or scientific) deconstruction, it is possible to be an objective scientist within a postmodern framework, just as it is possible to be a relativist within a modernist framework, and neither need engage in deconstructionism. Einstein and Heisenberg were successful relativist-modernists, whereas I consider Bayesians to be objectivist-postmodernists.

political decisions, to our collective detriment and in the face of more than adequate scientific evidence, hinge on the distinction between $P = 0.05$ and $P = 0.06$.

It is worth noting that "science" (and the science of ecology) is not equivalent to truth. Rather (to extend a concept developed by Postman 1995), science is a specialized language and method that we employ to learn about truth (objective reality) and to explain errors (arising from non-scientific, objectively false versions of reality). Similarly, statistics is not science, it is a specialized language and tool that we employ to speak about science and to describe scientific results. Thus, statistics is a language used to interpret another language (science) to learn about truth and explain error. As a plethora of human languages enriches our social lives, so a multiplicity of statistical languages can enrich our understanding of science and bring us closer to understanding reality.

"The" scientific method

Ecologists are preoccupied with the hypothetico-deductive, falsificationist method laid out by Popper (1959), viewed through the lens of statistics developed by Fisher (1922), and Neyman and Pearson (1928), and placed on a pedestal by Platt (1964), Strong (1980), Simberloff (1980), and Peters (1991). When published, the papers by Strong (1980) and Simberloff (1980) were useful and important antidotes to decades of fuzzy thought and story-telling masquerading as ecological science (although there is much useful data in those old papers), but subsequent ossification of hypothetico-deductivist methods in ecology shares much with the spread of gospel by religious disciples blindly following their master. Advances in modern science, including ecology, owe much to the professed use of the hypothetico-deductive, falsificationist method, but in many (most?) instances, it is applied post-hoc to ideas conceived inductively and tested with experiments deliberately designed to address the favored, "alternative" hypothesis. We should not forget that induction, which Popper attempted to toss into the dustbin of scientific history, is still widely used in “real” mathematics, and undergirds most (ecological) modeling activities (Nicholls et al. 1999 is a recent example with real-world policy implications), activities which Dennis does not consider to be science. This is indeed odd, as models are convincing explanations based on reliable knowledge, and often generate falsifiable hypotheses. Feyerabend (1975, 1978), Howson and Urbach (1993), and Schrader-Frechette and McCoy (1993) provide useful antidotes to knee-jerk hypothetico-deductivism, and are recommended reading for all ecologists who want to become more than just producers and curators of ANOVA tables.

The language of science: frequentist and Bayesian statistical inference

I have discussed elsewhere in greater detail some of the contrasts between frequentist and Bayesian inference (Ellison 1996; see Howson and Urbach 1993 for a readable review, and the papers collected by Earman 1983 for a thorough vetting of the philosophical arguments concerning hypothetico-deductivist and Bayesian methods). I only emphasize here that frequentist statistics answer the question “how probable are my data, given my (null) hypothesis?” (i.e., $P(\text{data} \mid H)$), whereas Bayesian inference answers the question, “how probable is my hypothesis (null or alternative(s)), given my data?” (i.e., $P(H \mid \text{data})$). Far from asserting that frequentist methods are “an anachronistic yoke impeding ecological progress”, I suggested that Bayesian inference had utility both in ecological research (where it would allow us to use effectively *unbiased* results from previous research, as in the “macroecological” approach advocated by Brown and Maurer [1989, Brown 1999], among others), and more importantly, in expressing degrees of uncertainty in a policy- and decision-making framework. My interest in this topic initially grew out of teaching a course in decision theory, where Bayesian inference is used extensively (e.g., Berger 1985, Smith 1988, Chechile and Carlisle 1991).

Fundamentally, we scientists seek to understand the world around us. We do this by drawing objective conclusions from observations and experiments that also allow for accurate predictions of future events. Frequentist statistics are very useful for drawing objective conclusions from observations and experiments, but are not well-suited for prediction and inference. Bayesian methods are useful for prediction and inference, given objective, *unbiased* data (Bayesian *priors*) from careful observations, controlled experiments, appropriate null hypotheses, and, when necessary, formal elicitation of expert opinion (unlike, in all respects, the "cooked" example Dennis presents[2]). For example, whereas clinical trials of new drugs are grounded firmly in frequentist methods, medical diagnosis and consequent improvements in healthcare and longevity owe much to the principles of "specificity" and "sensitivity" that emerge directly from Bayes theorem (e.g., Rosner 2000). Like clinicians who are responsible for the health and well-being of their patients, ecologists would be well-served by using all available tools to understand the mechanics of ecological systems (frequentist analysis) and to predict their responses to future conditions (model selection with or without Bayesian *or* frequentist inference; see for example Burnham and Anderson 1998).

Until recently, ecologists, like statisticians, have had only asymptotically accurate, frequentist methods available in their statistical toolkit. Increasing computational speed allowed for the development and use of exact statistical tests, many of which do not require knowledge of underlying distributions.

---

[2]Dennis' example examining copper concentrations is cooked in at least two ways. First, the "expert opinion" of the investigator employed by the mining company is clearly biased. Second, and statistically more importantly, Dennis has changed the point-null hypothesis from that presented in the frequentist example ($\mu_0 = 48$) to that presented in the Bayesian example ($\mu_0 = N(20,4)$). Dennis equates the expert prior with the null hypothesis, when in fact, the Bayesian would also be testing the same null hypothesis as the frequentist. There are other, known problems with testing point null hypotheses using Bayesian inference (see Lee 1997: 124*ff*), notably that it is impossible to put a continuous prior probability density on a point estimate (Lee 1997: 126). However, it is possible to put a prior probability density on a very small interval around the point estimate, as in μ , $(\mu_0 - g, \mu_0 + g)$. The data, and a properly identified prior on a point-null hypothesis would clearly have helped to identify the mining scientist as biased.

Further increases in the power of desktop computers, and improved methods of numerical analysis and approximation, have meant that Bayesian software can now be implemented more easily (e.g., Pole et al. 1994, Cook and Broemeling 1995, Albert 1996). Not surprisingly given these tools, a careful examination of assumptions of, and similarities and differences between Bayesian and frequentist inference has led to a greater diversity of approaches in the major statistical journals. Contrary to Dennis' assertion, practitioners[3] of frequentist and Bayesian inference interact and discuss areas of overlap and disagreement (e.g., Pratt 1965, Smith 1984, Efron 1986, 1998, Casella and Berger 1987, Lewis and Berry 1994, Samaniego and Reneau 1994, Hill 1995, Robbins 1995, Moore 1997, Gopalan and Berry 1998). It makes sense for ecologists to adopt the same degree of pluralism. I suggest that ecologists adopt good statistical practice and use whatever method is appropriate to the question ($P(\text{data} \mid H)$ or $P(H \mid \text{data})$?) — frequentist, exact, likelihood, information-theoretic, or Bayesian.

Statistical education for ecologists

Dennis argues that statistics is a post-calculus subject, and therefore all ecologists should have one year of calculus prior to entry into graduate school, wherein they should take a standard "math-stat" course in statistical theory. In principle I agree, but I suspect that this blanket prescription would not accomplish very much. "Real" undergraduate calculus generally is poor preparation for anything other than the follow-on, advanced calculus courses and undergraduates do not readily transfer their calculus skills to other courses (e.g., ecology, physics, statistics).

Statistics (and mathematics) is neither truth nor science, rather, it is a specialized language that has been developed to speak about science. Statistical education (I explicitly avoid the term "training",

[3]Neither should be called adherents to "frequentism" or "Bayesianism". These terms conjure up visions of religion and politics (pick your favorite "ism"), neither of which has a place in objective modernist or postmodernist science.

which, as Orr [1992] has pointed out, is what we do to animals) therefore should be about understanding the language of statistics, and understanding what limitations that language places on our understanding and interpretation of scientific results. Of course, to develop that understanding, it helps to know what statistics are, how they are used, how they are interpreted, and why society reveres them when it tells us what we want to know, and disparages them otherwise (contrast the front section of any issue of *USA Today* with Huff's classic book *How to lie with statistics* [1954]). Substitute "ecologists" for "society" in the last sentence, and we have the kernel of statistical education for ecologists. This has been broadly recognized by statisticians, who have developed curricula addressing these goals not only narrowly within statistics classes, but also across disciplinary boundaries, at both high school and university levels (e.g., Mosteller 1988, Cobb and Moore 1997, Nolan and Speed 1999, Roberts et al. 1999). Ecologists need to catch up to this trend, not form a rear-guard protecting the gates of a decaying castle. We also need this education so that we know how to provide unbiased, objective information to the individuals and groups we have charged with making ecological and environmental policy decisions. Otherwise, we will have nothing left to study.

Acknowledgments

This essay was written during a week's worth of field work in southern Oregon and northern California, supported by a grant from the Packard Foundation. The *Darlingtonia* fens of the Siskiyou and Klamath Mountains and the coastal redwood forests reminded me why I am a scientist and an ecologist.

References

Albert, J. H. 1996. Bayesian computation using Minitab. Wadsworth Publishing Company, Belmont.

Berger, J. O. 1985. Statistical decision theory and Bayesian analysis, second edition. Springer-Verlag, New York.

Brown, J. H. 1999. Macroecology: progress and prospect. Oikos 87: 3-14.

Brown, J. H., and B. A. Maurer. 1989. Macroecology: the division of food and space among species on continents. Science 243: 1145-1150.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York.

Cahoone, L. E., editor. 1996. From modernism to postmodernism: an anthology. Blackwell Publishers, Malden.

Casella, G., and R. L. Berger. 1987. Reconciling Bayesian and frequentist evidence in the one-sided testing problem. Journal of the American Statistical Association 82: 106-111.

Chechile, R. A., and S. Carlisle. 1991. Environmental decision making : a multidisciplinary perspective. Van Nostrand Reinhold, New York.

Cobb, G. W., and D. S. Moore. 1997. Mathematics, statistics, and teaching. American Mathematical Monthly 104: 801-823.

Cook, P., and L. D. Broemeling. 1995. Bayesian statistics using Mathematica. The American Statistician 49: 70-76.

De Man, P. 1989. Critical writings 1953-1978, edited by L. Waters. University of Minnesota Press, Minneapolis.

Derrida, J. 1967. L’écriture et la différence. Editions du Seuil, Paris. (Translated [1978] as Writing and Difference. University of Chicago Press, Chicago.)

Earman, J., editor. 1983. Testing scientific theories. University of Minnesota Press, Minneapolis.

Efron, B. 1986. Why isn't everyone a Bayesian? The American Statistician 40: 1-11.

Ellison, A. M. 1987. Density-dependent dynamics of *Salicornia europaea* monocultures. Ecology 68: 737-741.

Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications 6: 1036-1046.

Enquist, B. J., J. H. Brown, and G. B. West. 1998. Allometric scaling of plant energetics and population density. Nature 395: 163-165.

Feyerabend, P. 1975. Against method. Verso, London.

Feyerabend, P. 1978. Science in a free society. Verso, London.

Feyerabend, P. 1987. Farewell to reason. Verso, London.

Fisher, R. A. 1922. On the mathematical foundations of a theoretical statistics. Philosophical Transactions of the Royal Society of London A222: 309-368.

Gopalan, R., and D. A. Berry. 1998. Bayesian multiple comparisons using Dirichlet process priors. Journal of the American Statistical Association 93: 1130-1139.

Gross, P. R., and N. Levitt. Higher superstision: the academic left and its quarrels with science. Johns Hopkins University Press, Baltimore.

Heisenberg, W. 1958. Physics & philosophy : the revolution in modern science. Harper and Row, New York.

Hill, B. M. 1995. Posterior distribution of percentiles: Bayes' theorem for sampling from a population. Journal of the American Statistical Association 677-691.

Howson, C., and P. Urbach. 1993. Scientific reasoning: the Bayesian approach, second edition. Open Court Publishing Company, Chicago.

Huff, D. 1954. How to lie with statistics. W. W. Norton, New York.

Lewis, R. J., and D. A. Berry. 1994. Group sequential clinical trials: a classical evaluation of Bayesian decision-theoretic designs. Journal of the American Statistical Association 89: 1528-1534.

Moore, D. S. 1997. Bayes for beginners? some reasons to hesitate. The American Statistician 51: 254-261.

Mosteller, F. 1988. Broadening the scope of statistics and statistical education. The American Statistician 42: 93-99.

Neyman, J., and E. S. Pearson. 1928. On the use and the interpretation of certain test criteria for purposes of statistical inference. Biometrika 20: 175-240 (part I) and 263-294 (part II).

Nicholls, R. J., F. M. J. Hoozemans, and M. Marchand. 1999. Increasing flood risk and wetland losses due to global sea-level rise: regional and global analyses. Global Environmental Change 9: S69-S87.

Nolan, D., and T. P. Speed. 1999. Teaching statistics theory through applications. The American Statistician 53: 370-375.

Orr, D. W. 1992. Ecological literacy : education and the transition to a postmodern world. State University of New York Press, Albany.

Peters, R. H. 1991. A critique for ecology. Cambridge University Press, Cambridge.

Pole, A., M. West, and J. Harrison. 1994. Applied Bayesian forecasting and time series analysis. Chapman and Hall, New York.

Popper, K. 1959. The logic of scientific discovery. Basic Books, New York.

Postman, N. 1995. The end of education. Vintage Books, New York.

Postman, N., and C. Weingartner. 1969. Teaching as a subversive activity. Dell Publishing Company, New York.

Platt, J. R. 1964. Strong inference. Science 146: 347-353.

Pratt, J. W. 1965. Bayesian interpretation of standard inference statements. Journal of the Royal Statistical Society 27: 169-203.

Robbins, H. 1995. The empirical Bayes approach to statistical decision problems. Annals of Mathematical Statistics 35: 1-20.

Roberts, R., Scheaffer, R., and A. Watkins. 1999. Advanced placement statistics — past, present, and future. The American Statistician 53: 307-321.

Rosner, B. 2000. Fundamentals of biostatistics, fifth edition. Duxbury, Pacific Grove.

Samaniego, F. J., and D. M. Reneau. 1994. Toward a reconciliation of the Bayesian and frequentist approaches to point estimation. Journal of the American Statistical Association 89: 947-957.

Schrader-Frechette, K. S., and E. D. McCoy. 1993. Method in ecology: strategies for conservation. Cambridge University Press, Cambridge.

Simberloff, D. 1980. A succession of paradigms in ecology: essentialism to materialism and probabalism. Synthese 43: 3-39.

Smith, A. F. M. 1984. Present position and potential developments: some personal views: Bayesian statistics. Journal of the Royal Statistical Society A 147: 245-259.

Smith, J. Q. 1988. Decision analysis: a Bayesian approach. Chapman and Hall, London.

Strong, D. R. 1980. Null hypotheses in ecology. Synthese 43: 271-285.