Statistical Science
2011, Vol. 26, No. 1, 10–11
DOI: 10.1214/11-STS337C
Main article DOI: 10.1214/10-STS337

Bayesian Statistical Pragmatism

Andrew Gelman

I agree with Rob Kass' point that we can and should make use of statistical methods developed under different philosophies, and I am happy to take the opportunity to elaborate on some of his arguments.

FOUNDATIONS OF PROBABILITY

Kass describes probability theory as anchored upon physical randomization (coin flips, die rolls and the like) but being useful more generally as a mathematical model. I completely agree but would also add another anchoring point: *calibration.* Calibration of probability assessments is an objective, not subjective process, although some subjectivity (or scientific judgment) is necessarily involved in the choice of events used in the calibration. In that way, Bayesian probability calibration is closely connected to frequentist probability statements, in that both are conditional on "reference sets" of comparable events. We discuss these issues further in Chapter 1 of *Bayesian Data Analysis*, featuring examples from sports betting and record linkage.

CONFIDENCE INTERVALS AND HYPOTHESIS TESTS

I agree with Kass that confidence and statistical significance are "valuable inferential tools." They are treated differently in classical and Bayesian statistics, however. In the Neyman–Pearson theory of inference, confidence and statistical significance are two sides of the same coin, with a confidence interval being the set of parameter values not rejected by a significance test. Unfortunately, this approach falls apart (or, at the very least, is extremely difficult) in problems with high-dimensional parameter spaces that are characteristic of my own applied work in social science and environmental health.

In a modern Bayesian approach, confidence intervals and hypothesis testing are both important but are *not* isomorphic; they represent two different steps of inference. Confidence statements, or posterior intervals, are summaries of inference about parameters conditional on an assumed model. Hypothesis testing—or, more generally, model checking—is the process of comparing observed data to replications under the model if it were true. Statistically significance in a hypothesis test corresponds to some aspect of the data which would be unexpected under the model. For Bayesians, as for other statistical researchers, both these steps of inferences are important: we want to make use of the mathematics of probability to make conditionally valid statements about unobserved quantities, and we also want to make use of this same probability theory to reveal areas in which our models do not fit the data.

SAMPLING

Kass discusses the role of sampling as a model for understanding statistical inference. But sampling is more than a metaphor; it is crucial in many aspects of statistics. This is evident in analysis of public opinion and health, where analyses rely on random-sample national surveys, and in environmental statistics, where continuous physical variables are studied using space-time samples. But even in areas where sampling is less apparent, it can be important. Consider medical experiments, where the object invariably is inference for the general population, not merely for the patients in the study. Similarly, the goal of Kass and his colleagues in their neuroscience research is to learn about general aspects of human and animal brains, not merely to study the particular creatures on which they have data. Ultimately, *sample* is just another word for *subset*, and in both Bayesian and classical inference, appropriate generalization from sample to population depends on a model for the sampling or selection process. I have no problem with Kass' use of sampling as a framework for inference, and I think this will work even better if he emphasizes the generalization from real samples to real populations—not just mathematical constructs—that are central to so much of our applied inferences.

SUBJECTIVITY AND BELIEF

The only two statements in Kass' article that I clearly disagree with are the following two claims: "the only solid foundation for Bayesianism is subjective," and

Andrew Gelman is Professor, Department of Statistics and Department of Political Science, Columbia University, New York, New York 10027, USA (e-mail: gelman@stat.columbia.edu).