Statistical Science
2011, Vol. 26, No. 1, 15–16
DOI: 10.1214/11-STS337D
Main article DOI: 10.1214/10-STS337
© Institute of Mathematical Statistics, 2011

Discussion of "Statistical Inference: The Big Picture" by R. E. Kass

Robert McCulloch

Kass states (page 5) that Figure 3 is not a good general description of statistical inference and that Figure 1 is more accurate. I completely agree. Kass states (page 5) that "It is important for students in introductory courses to see the subject as a coherent, principled, whole." I completely agree. Since Figure 3 represents the framework within which statistics is usually taught, at all levels, we have a serious problem. These issues have bothered me deeply for a long time. This important paper forcibly brings these matters to light and I hope it is influential.

Figure 3 represents "sampling from a finite population." There are a large number of unknown numbers and we randomly pick some of them to uncover. The "population quantities" are summaries of all of the numbers. These are the parameters. Sample quantities are summaries of the uncovered numbers. All "randomness" arises from that which we inject by randomly picking the sample. This is so obviously not a description of what we usually do in statistical modeling that I am just amazed at its persistence. "Randomness" comes from my personal need to make a decision in an uncertain context (Lindley, 1985).

As Kass does, let me take a "simple" example. I just taught an introductory statistics class. How did the need for "probabilistic thinking" come into the course in a way that the students could immediately see the need for it? The practical problem of choosing an investment portfolio for one period was considered. You could put your money in a riskless asset (government bonds) with known return or a risky asset (the market). Past data on market returns are available. Several issues need to be discussed. There is not really anything such as a "riskless asset" but compared to stocks, goverments bonds are riskless. There is a useful, but imperfect match between our model and the real world. How much do the past returns guide us representing our uncertainty about the unknown future return? Suppose you hold the portfolio for two periods. Is the return in the second period related to (independent of) the return in the first? How could returns be approximately independent? Perhaps the theory of efficient markets sheds light on this. That is how I introduced probability and (*I think*) I got away with it. I used a normal distribution to describe my uncertainty about the next return. I graphically showed that my choice of mean and variance was "somewhat consistent" with the past but emphasized that I did not have to match the past and we discussed deviations that one might want to consider. I would say that my approach was largely consistent with Kass' Figure 1. I would also say that the idea of a random sample from a finite (or infinite) population was nowhere to be seen.

Figures 1 (and 4) are not easy ideas. But not addressing them directly just makes it more confusing in the long run. What you really need in a good course are relatively simple examples which are entirely in the spirit of Figures 1 and 4. I did use coins and dice in discussing the intuitive idea of independence. However, I also emphasize that saying returns are independent is a big assumption with serious consequences.

In more complex statistical modeling where we consider many variables, many basic statistical issues must be considered. Ones of general importance that come to mind are the bias–variance trade-off in prediction and the difference between prediction given passive observations of a system versus predictions about the effect of an intervention in a system that has not been done before (correlation versus causation). These kinds of issues fit naturally into Figures 1 and 4. I assume that the term "conclusions" is meant to include predictions.

Of course, the other issue that statistical science stresses is the quantification of our uncertainty about our conclusions. Here, I still believe the Bayesian approach has real advantages. Kass knows all the arguments better than I do so I will not go through the list but I cannot resist rattling off a few. Even teaching an introductory course, the severe deficiencies of classically inspired approaches become apparent. I could not

Robert McCulloch is Professor, IROM Department, University of Texas at Austin, CBA 5.202, 1 University Station, B6500, Austin, Texas 78712-0212, USA (e-mail: robert.mcculloch1@gmail.com).