Review of *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction*, by Frank Lad. Wiley: New York, 1996.

This spirited book aspires to recast the teaching of mathematical statistics in the pure subjectivist mold of Bruno de Finetti. It is largely modeled after de Finetti's *Theory of Probability*, which appeared in Italian in 1970 and then in English in two volumes in 1974 and 1975. More about statistics and less about probability, it omits stochastic processes in favor of material on proper scoring rules, normal mixtures, and Bayesian regression that has been developed since de Finetti wrote, and it has all the panoply of a textbook—extended examples and numerous exercises—that *Theory of Probability* lacked. But like that treatise, it devotes many pages to quantity, prevision, and coherence before introducing standard probability distributions as a footnote to exchangeability, and it even outdoes de Finetti's far-ranging opinionatedness and stubborn consistency.

In the pure subjectivist view, a probability is always someone's degree of belief, not a property of a random phenomenon in nature. De Finetti squared this subjectivism with modern empiricism by insisting that *degree of belief* be made operational: a person can have probabilities only for events or variables that are going to be observed, and these probabilities must be certified by the person's willingness to bet on how the observations come out. Most statisticians who call themselves Bayesians agree with de Finetti in principle but compromise in practice. In order to communicate with their fellow statisticians and especially with users of statistics, these practical Bayesians are willing to speak about phenomena being governed by objective probabilities and about estimating parameters in statistical laws that express these objective probabilities. They insist only that the estimation be conducted by Bayesian means: in order to estimate the parameter θ in a statistical law p_θ for an observable x, we are supposed to assess our prior subjective probabilities $p(\theta)$ and to use them together with the observed x and the probabilities $p_\theta(x)$ to compute posterior probabilities for θ and perhaps predictive probabilities for future x.

Most of the Bayesian books now on the market take the compromise with objectivism as their starting point. Christian P. Robert, for example, opens his advanced textbook, *The Bayesian Choice: A Decision-Theoretic Motivation* with "The main purpose of statistical theory is to derive from observations of a random phenomenon an *inference* about the probability distribution underlying this phenomenon." Others begin by apologizing ex-

tensively for their eventual acquiescence. José M. Bernardo and Adrian F. M. Smith spend 200 hundred pages of their *Bayesian Theory* explaining subjectivism and de Finetti's representation theorem before settling in to the usual $p_\theta(x)$.

But for Frank Lad there is no compromise. His subjectivism is pure at the beginning, in the middle, and at the end. For him the only task of statistics is the subjective prediction of future observations, and the very concepts of statistical estimation and hypothesis testing must be scrapped, to be replaced by the operational evaluation of subjective predictions by means of proper scoring rules.

Lad's claim to be the purest of the pure is buttressed by his outright rejection of the idea of calibration. You are *calibrated* as a probability predictor if the subjective probabilities you give pan out approximately as relative frequencies; for example, of all the events to which you assign a probability of .95, approximately 95% should happen. Many Bayesians think that calibration is a reasonable and coherent goal for a probability predictor, thus making another compromise with objectivism and frequentism. Not Lad. He regards calibration as a snare and a delusion, and he advances a remarkable argument in support of his position. A probability predictor, he claims, is always calibrated. Consider the events to which you assign probability one-half. For every event in your list, you must also include its complement, for when you assign probability one-half to one of the two you also assign probability one-half to the other. But since exactly one of the two will have happened, exactly half of the events to which you assign probability one-half will happen. You can't go wrong. This argument does not work so easily for probabilities other than one-half, but Lad forces it to work by supposing that you give subjective quantiles for unknown quantities. For example, you might give quartiles $q_{.25}, q_{.5}, q_{.75}$ for x, and when you do this you are giving probability one-fourth to four events, exactly one of which will happen:

$$P(x \leq q_{.25}) = P(q_{.25} \leq x \leq q_{.5}) = P(q_{.5} \leq x \leq q_{.75})) = P(q_{.75} \leq x) = \frac{1}{4}.$$

Many readers will find this argument too silly to merit refutation. They will take for granted that a probability predictor should be evaluated by someone else, who chooses the events for which the predictor must give probabilities. Being calibrated should mean performing well against many different challenges—much in the spirit of Richard von Mises's conception of probability as relative frequency invariant under subsequence selection. Lad will

have none of this.

We cannot understand Lad's point of view without reference to the basic tenets of his philosophy of science. This philosophy holds that science can talk only about operationally defined measurements. In order to introduce the idea of probability and expected value into science, we must accept the thesis that your expected values (or your *previsions*, as Lad and de Finetti prefer to say) are measurements of your knowledge that are made together with measurements of your values. We measure your uncertain knowledge of a quantity X by asking you to choose among gambles, and your prevision of X is the resulting measurement, $P(X)$. The introspection on your part required by the measurement may involve considerable effort, and if you limit this effort, we may obtain only a prevision interval, $[P_l(X), P_u(X)]$. But in principle your knowledge can be measured exactly; as Lad says (p. 66), "Presumably any prevision interval can be shortened via further reflection and sharpened judgment if you decide it is worth the effort." The fact that your values are implicated in the measurement means that we must reject the doctrine that the scientific method is value free. And the fact that your prevision is defined exclusively in terms of introspection means that what happens in the future is utterly irrelevant to its meaning and validity.

Who is this book for? Lad answers expansively. He envisions using the book for three distinct courses. In his first course (for second-year students in his New Zealand university, corresponding perhaps to juniors majoring in statistics in an American university), he covers the introductory chapter on history and philosophy and then the material on prevision and coherence, leading into an introduction to exchangeability and distribution theory. In his second course (for his third-year students), he goes into depth on distribution theory, emphasizing normal mixtures. Finally, in his graduate seminar, he emphasizes proper scoring rules and Bayesian regression.

I would be reluctant to use the book as a primary text for any of these audiences. Although Lad's historical and philosophical introduction is admirable in its breadth, its judgments are so eccentric that it would need to be extensively supplemented. While I appreciate the care he has lavished on his treatment of the intricacies of coherence and exchangeability, I would worry whether this material opens as many doors for the beginning statistics student as the decision-theoretic topics that one would more often find in a theoretical Bayesian course. The material Lad suggests for a second course has greater appeal to the practitioner, but Lad himself suggests that much of it has "been rendered obsolete" by the development of Monte Carlo meth-

ods for Bayesian computation and by the development of Bayesian expert systems (p. 423). As for the graduate seminar, I would happily use Lad's exposition of proper scoring rules, but I would want to supplement it extensively with material such as the work of Phil Dawid, which Lad rejects as too close to frequentism.

Although I would not use it as a textbook, Lad's book will sit on my shelf alongside de Finetti's treatise. It systematically updates de Finetti, pulling together much of the Bayesian work of the past several decades. My main disappointment with it is that so stubbornly continues the construction of a purer-than-white Bayesian sect, determined not to see any kernel of truth in the objectivist and causal conceptions that dominate statistical practice. The notion of objective statistical law is surely too easily accepted by mainstream statistical theory and practice. I believe that it is incomplete so long as it excludes every element of subjectivity. And it may make more sense for statistical work in agriculture and medicine than for work in Lad's specialty, economics. But the objective and frequency aspects of probability are ineluctable, and any philosopher of probability and statistics should want to offer a positive account of them.

What disturbs me most are those points where the author links the objectivist philosophy he opposes to crimes we all deplore. On the first page of the preface, he says that the twentieth-century fascination with theories of objective probability is connected to an outlook that "has helped to create false hopes for science and, in some cases, to allow the development of abusive authority." We later learn that this is a swipe at the Soviet version of Marxist materialism. This linking of the philosophy of probability with politics is both flimsy and irresponsible. In light of de Finetti's misadventures with fascism, it takes on an ugly cast. Does subjectivism not have political crimes to its own credit? Are we to debate the philosophy of probability by counting victims of competing mass murderers?

On the whole, Lad maintains a remarkably positive tone, and he frequently praises the work of other scholars, even those with whom he disagrees (including this reviewer!). But he occasionally lets his anger show. While preparing this review, I read this passage, from the last chapter of the book (p. 451), to my wife:

> experienced readers of the objectivist statistical persuasion may well have bristled at the unchallenged direction this text has followed, well aware of its implications for your own respected re-

> search programs. While not relished, your disconcertion has been intended. May it match the ill-ease felt by subjectivist statisticians who attempt to engage in the professional arena of scientific statistical analysis, still dominated by directed activities of searching for true, unobservable, randomness-generated structures.

My wife obliged me by censuring this bombast, but in a way that put me in the same tub; the author, she told me, seems to have a highly inflated idea of the place in the world of himself and the brethren he is quarreling with.

Yes, the importance of the tempest in our teapot may not quite match the emotion we invest in it. But perhaps our debates have a larger literary purpose. Lad prefaces each of his chapters with a passage from a novel by Samuel Beckett. Upon consulting an encyclopedia to refresh my memory, I learned that Beckett "spun fable after fable of persons trapped by perfectly logical, demoralizing absurdity." There we are—the unconverted as Lad sees us and Lad as we see him—each of us trapped in the absurdity of his perfect logic.

Christian P. Robert, *The Bayesian Choice: A Decision-Theoretic Motivation*. Springer: New York. 1994.

José M. Bernardo and Adrian F. M. Smith, *Bayesian Theory*. Wiley: New York. 1994.

Bruno de Finetti, *Theory of Probability*. Wiley: Chichester. Translated from the Italian by Antoino Machi and Adrian Smith. Volume 1, 1974. Volume 2, 1975.

Glenn Shafer, Faculty of Management, Rutgers University
gshafer@andromeda.rutgers.edu
September 5, 1998