Review of "Principles of Uncertainty by J.B. Kadane"

Michael Evans
Department of Statistics, University of Toronto

This book could be considered as an introduction to the fields of probability and statistics. The book develops the subject from the basics to the point where a student learns how to be involved with significant applications. The book is nonstandard in a number of ways, however, and some of these we believe would preclude it from being used in this way for most undergraduate, mathematically oriented introductory courses. For example, the mathematical prerequisites for the book are not high, as many of the mathematical concepts used are developed within the book itself, but a student would have to be at least willing to spend the time on this as opposed to focusing on just probability and statistics. For example, the book contains a fairly complete development of integration including the Riemann integral, Riemann-Stieltjes integral and the McShane integral (an alternative approach to the Lebesgue integral), and the spectral and singular value decompositions of linear algebra are developed from first principles. Certainly this could be appropriate and highly useful for a student, but some of these topics would be better learned in courses that specialized in, for example, real analysis, measure theory and linear algebra. For a senior student, however, who wanted a rigorous introduction to the field, and did not have the time or inclination to take the appropriate mathematics courses, the approach taken in the book could work very well.

Overall this is a very notable book. First it is a joy to read. The prose is very clear. Mathematical aspects of the subject are developed rigorously with no hand waving and the mathematics is combined with conceptual discussion in just the right balance. On almost every page I found insightful comments and developments that prompted thinking about the material, often from a new perspective. Copious and eclectic references are provided to a wide literature. Second the book sets out to accomplish a very worthy and necessary goal for the subject: to develop a full theory of statistical reasoning. For a variety of reasons most statistics texts fail in this regard but there is no denying that this book accomplishes this goal. From the perspective of a student this is very satisfying as a coherent and full theory is presented with no loose ends or appeals to intuition as opposed to logical reasoning. For both of these reasons I would recommend this book to any serious student of statistics.

The completeness of the book, however, brings us to a discussion of the way in which this book is very nonstandard. This is concerned with a controversial aspect of the subject, namely, there is little agreement among statisticians about an appropriate theory of statistical reasoning. Certainly the author understands this and clearly states the point-of-view taken and why. This is laudable because this central question for the field must be addressed otherwise we are left with no way to determine when a statistical analysis, and thus the conclusions or inferences drawn from such an analysis, can be deemed correct. Here correct means that we have obeyed the rules as prescribed by a theory. It is clearly not acceptable to say that a statistical methodology is acceptable simply because it is commonly used. The P-value is used in much of empirical science and yet it has long been understood that the foundational logic behind this tool is at best weak, see for example Ziliak and McCloskey (2009).

Of course, just having a complete theory of statistical reasoning is not enough. This theory must have broad support and for this it must be seen to deal adequately and unambiguously with the problems the field is supposed to address. Statisticians have broadly speaking distributed themselves along a spectrum where at one end there are those who claim objectivity for their approach and at the other extreme are those who claim that subjectivity is a necessary aspect of any coherent theory of statistical reasoning. This book reflects the approach taken by those at the extreme subjective end. The objective statistician complains that subjectivity cannot be part of any *scientific* approach to statistics while the subjectivist complains about the failure of the others to provide a complete theory that is free of logical flaws. While this debate remains open there are attempts, for example, see Bernardo, Berger and Sun (2009), at resolution.

Any approach to statistical reasoning is in practice based on ingredients chosen by someone, whether that be the analyst or a committee of researchers, etc. For example, there is the choice of a statistical model which comprises a set of probability distribution one of which was *supposed* to have generated the observed data. In the vast majority of applications, however, there is nothing to suggest that a particular statistical model is the correct one. Rather we rely on individual judgement to make a reasonable subjective choice. As the statistical model is a part of every approach to a theory of statistical reasoning, the claim that there is a theory of statistics that is purely objective is manifestly false. A Bayesian statistician adds another ingredient, namely, the prior probability distribution which reflects beliefs about which probability distributions in the model are more or less likely to be true. As with the model, this is a subjective choice. The Bayesian adds this ingredient primarily because of dissatisfaction with the attempts at theories of inference that have been developed based upon a statistical model alone. Some Bayesians are still dissatisfied with these ingredients for the problem and add yet another, namely, a utility function which reflects preferences among a set of possible consequences that obtain dependent upon actions taken and the true distribution in the model that generated the data. A complete theory is now available as we proceed by the principle of maximized expected utility to choose that data dependent action that maximizes the expected utility. Following the approach advocated in this book a statistician is lead to a choice of model, prior and utility based upon considering bets they will make or offer about unknowns.

So given that all approaches to statistics are inherently subjective why don't we just go all the way and accept the prescription for the theory as offered by this text? First, it is worth noting that there is one part of a statistical analysis that can indeed be objective, namely, not dependent on the statistician. At least if it is collected correctly through random sampling, the data is objective.

In fact one can argue that this is the primary reason why random sampling is such an important tool in statistical practice as it insures that the subjective biases of the statistician do not influence the data observed. In general it seems completely unreasonable to claim that, because subjectivity is an inherent part of statistics, we must just accept a statistical analysis based on the model, prior and utility as correct and only reject it if we personally don't accept these ingredients. Because of the subjective nature of the ingredients, this means that all such statistical analyses must be suspect, just as we should be suspicious of any analysis based on data that is subjectively chosen.

Furthermore, there is no need for such an extreme position. For example, as already noted, the model is a subjective choice. So how do statisticians cope with this? We ask the question: is the (objective) data surprising for each probability distribution in the model? If so, then we have evidence that the model chosen was indeed a poor choice. Otherwise we can at least proceed as if the model makes some sense. We'll never know if it is correct or not, and so this *model checking* does not completely deal with the subjectivity of the choice, but at least it gives us a way to rationally assess whether or not our subjective choice is seriously in error. Model checking has long been noted to be just good statistical practice but it is not mentioned in this text. Model choice is discussed but any collection of models is in itself a model which then needs to be checked. If the model is accepted (and note this does not mean we absolutely believe it is correct), then we consider the reasonableness of the prior. For example, we can check to see whether or not the data is pointing to the true probability distribution in the model being a surprising value for the prior. If we are placing most of our prior belief on values that are surprising in light of the data, then we have to worry about the effect the contradiction, between our subjective beliefs and the objective data, is having on our conclusions. This is called checking for prior-data conflict. If the prior passes this check it is not in any sense correct or objective, but again it gives us a way to rationally assess whether or not our subjective choice is seriously in error, see, for example, Evans and Jang (2012).

There are typically many other choices made. If we can come up with ways of checking the reasonableness of these choices against the data, then this seems like an appropriate approach. But what if we can't? Then clearly we should restrict ourselves to a theory that prescribes inferences that are invariant to such choices. For example, it is not at all clear how we would check a parameterization chosen for a model against the data. But there are inference methods in Bayesian problems that are invariant to the parameterization. Also, it is not at all clear how one checks a utility function. A utility function, however, is generally not necessary, as we can prescribe rules for inferences that don't require one. While there are undoubtedly many contexts where personal utilities are relevant (for example, personal investment decisions) the concept of utility is at least problematical due to a variety of paradoxes (as noted in this text) and many find it completely irrelevant for scientific problems where the goal is to characterize evidence as free from personal biases as possible.

Undoubtedly those who accept the approach to statistical reasoning described in this text will reject the notion of checking the model or checking the prior because the methods we must use to do this are not Bayesian (there is one model and one prior) and thus are not *coherent*. But what does this coherence mean? Effectively it seems based on the betting formulation of probability and the avoidance of sure loss as described in Chapter 1. But why is this the "correct" way to think about probability? Certainly it is reasonable but it seems perfectly fine to think of other ways of assigning probabilities, like relative frequencies, so that the axioms of probability are satisfied. When developing a theory of statistical reasoning it seems better to take probability as a primitive concept that is simply measuring degree of belief however it was assigned. Trying to force probability to only correspond to long run frequencies or to only correspond to bets seems artificial in both cases. In the end we need a theory of statistical reasoning that is not only complete, in the sense the ingredients and rules lead unambiguously to inferences, but also allows us to check that the ingredients chosen are at least sensible in light of objective data.

Some may argue that this is still allowing for too much subjectivity as two statisticians can choose different models and/or different priors for the same situation. But this isn't a problem unless both sets of ingredients pass their checks and the conclusions drawn are substantially different. In this case statistics tells us what to do, namely, we need to collect more (objective) data to resolve the disagreement.

I highly recommend this book to all those interested in statistics. That statistical reasoning is inherently subjective seems indisputable. Acknowledging this and recognizing the benefits of subjectivity through wise choices, but also recognizing the dangers of subjectivity and protecting ourselves from these, is perhaps a way for the field to move past what seems like an almost eternal division in the field. The author has done a great service to the field of statistics by presenting this material so clearly and completely.

References

Berger, J.O, Bernardo, J.M. and Sun, D. (2009) The formal definition of reference priors. Ann. Statist. Volume 37, Number 2 , 905-938.

Evans, M. and Jang, G. H. (2011) Weak informativity and the information in one prior relative to another . Statistical Science, Vol. 26, 3, 423-439.

Ziliak, S.T. and McCloskey, D.N. (2009) The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice and Lives. The University of Michigan Press, Ann Arbour.