Statistical Science
2011, Vol. 26, No. 2, 175–178
DOI: 10.1214/10-STS318B
Main article DOI: 10.1214/10-STS318

Discussion of "Calibrated Bayes, for Statistics in General, and Missing Data in Particular" by R. J. A. Little

Michael D. Larsen

I would like to thank Rod Little for a thought-provoking and well-presented paper on the "calibrated Bayes" approach to statistics. The author makes a strong case for the advantages of Bayesian methods and multiple imputation when dealing with missing data: the ability to fill in the data while accounting for the missing information in the inference is highly desirable. The article expounds the idea of a calibrated Bayesian approach to statistical problems in general and to missing data issues in particular. It would certainly be interesting to see an expanded treatment of how to implement calibration in the Bayesian context. Does this primarily mean selecting and transforming variables and models to get a good fit to the data? Does it also mean running more analyses to check sensitivity to missing data and model/variable assumptions? What about hierarchical models (e.g., Bayarri and Castellanos, 2007)? Advances in (MCMC) algorithms, computing power and (free) software on the web have made Bayesian approaches feasible for a much broader group of statisticians and other researchers. Indeed, a significant portion of the article summarizes and illustrates some techniques. There is a need for more "how to be calibrated" guidance, including computing tools and textbook examples, for applied Bayesians in practice.

One example from recent work comes to mind. In this example, a frequentist analysis is going to be reported, but there are missing data. Multiple imputation in this context is useful for building confidence in the results, because it is possible to compare and contrast results under different missing data assumptions. In an additional analysis of data from the Diabetes Prevention Program (Knowler et al., 2002), parent's age at death was being used as a predictor of the onset of diabetes in a population of adult pre-diabetics. Parents who live a long time generally are a good predictor of health of their children; the premature death of a parent does not augur well for offspring. But nearly $1/3$ of the parents were still alive at the beginning of the study (when parental age at death was captured). Not surprisingly, these parents were less likely to have had a cardiovascular event in the past than were the other parents. Their adult children tended to be younger than the other study participants. In the analysis using parent's age at death as a predictor variable, should data from the $1/3$ of the subjects be discarded from the analysis?

An attempt was made to model time until death for the parents who were living at study entry. Several variables were predictive of parental longevity. It was, then, possible to multiply impute age at death under some models, and then conduct the primary analysis utilizing multiple imputation combining rules. In the end, the results did not change much from the analysis based on only the complete cases—other than being younger, the patients with living parents did not differ much on average from the others. Even if a Bayesian analysis is not ultimately reported in detail, use of a multiple imputation procedure did seem to lead credence to the frequentist-procedure results; that fact can be stated very succinctly in a medical journal article. Statistical practice would move closer to "calibrated Bayes" if checks such as the one described here became standard and expected instead of novel.

If the analysis in the example described above had been substantially different from the complete case answer, then more work (i.e., statistical modeling and model checking on the available data) would have been needed to understand why. One might then discover something important in the data that would not be apparent for either analysis alone. Today, one could imagine that substantially more effort would have been needed to get an alternative Bayesian analysis accepted in many journals as the primary analysis instead of the complete case analysis. Statistics in practice would be closer to "calibrated Bayes" if well done Bayesian

Michael D. Larsen is Associate Professor of Statistics, Biostatistics Center, George Washington University, 6110 Executive Boulevard, Suite 750 Rockville, Maryland 20852, USA (e-mail: mlarsen@bsc.gwu.edu).