"A Message From the President"
What Are the Open Problems in Bayesian Statistics?
ISBA Bulletin, *18(1)*, 1-4, 2011.

Michael I. Jordan
ISBA President, 2011
jordan@stat.berkeley.edu

From time to time, I am approached by young students who are considering a career in statistics and who ask "What are the open problems in statistics?" While I'm often tempted to respond that "we don't tend to think that way in statistics," the nature of the question tends to imply a student with mathematical training of the kind that I usually look for in a prospective student, and so I do my best to give a thoughtful response and cast our activities in terms of "open problems."

This scenario came to my mind as I sat down to contemplate writing this column. It had already occurred to me that one of the consolations of being President of ISBA (or President of any society) is that one can ask others to do the actual work that's attributed to the President of the society. It was but a short step to realize that this might also apply to the writing of my column. Indeed, inspired by the notion of "crowd-sourcing" that is all the rage, I realized that as ISBA President I had been given an unparalleled opportunity for "statistician-sourcing." People might respond to the ISBA President in ways that they might not respond to humble old me. And so I thought that I would seize the opportunity, assemble a distinguished panel of statisticians and see what their views on the "open problems of statistics" might be. I imagined that this might be of interest beyond my recruiting scenario.

My polling methodology is rather open to critique, I am afraid. In particular, the individuals assembled are a highly non-random sample—they are a set of people who have the misfortune of being in the intersection of two sets: (a) highly-respected senior statisticians and (b) entries in my email address book.

The question that I asked was "What do you view as the top two or three open problems in Bayesian statistics?" The focus on Bayes is due to the ISBA context of course, but I also think that frequentist statisticians are more accustomed to thinking in terms of "open problems," and I wanted to make the question challenging (given that I didn't have to answer it myself). Here is the distinguished group that I wrote to: Andrew Barron, Susie Bayarri, Jim Berger, José Bernardo, Peter Bickel, Larry Brown, Brad Carlin, George Casella, Ming-Hui Chen, Merlise Clyde, Phil Dawid, Persi Diaconis, David Draper, David Dunson, Brad Efron, Steve Fienberg, Peter Green, Alan Gelfand, Andrew Gelman, Ed George, Malay

Ghosh, Nils Hjort, Peter Hoff, Jay Kadane, Rob Kass, Jun Liu, Steve MacEachern, Xiao-Li Meng, Peter Mueller, Tony O'Hagan, Luis Pericchi, Sonia Petrone, Fernando Quintana, Adrian Raftery, Sylvia Richardson, Thomas Richardson, Christian Robert, Judith Rousseau, Fabrizio Ruggeri, Mark Schervish, David Spiegelhalter, Terry Speed, Steve Stigler, Aad van der Vaart, Stephen Walker, Larry Wasserman, Mike West, and Wing Wong. I (amazingly) had a response rate not too far from 100%, and the responses were invigorating.

I note parenthetically that one person didn't answer my question but instead conducted his own mini-poll of colleagues the results of which he transmitted to me; impressed by this skill in delegation of responsibility, I intend to nominate this person in the next ISBA presidential election.

I turn to the results of my poll. I have organized the results into categories, with examples of open problems listed within each category. In several cases I have used quotes from individuals when I felt that a paraphrase would be less clear than the original text. I organize my results as a "top-five list."

5. *Nonparametrics and semiparametrics.* Bayesian nonparametrics is viewed by some of my respondents as a class of methods looking for a problem, and so the main open problem in Bayesian nonparametrics is (for some people) that of finding a characterization of classes of problems for which these tools are worth the trouble.

   But the success stories in frequentist nonparametrics are alluring to many in my group of respondents, and the concrete open problems raised for nonparametrics by the group are generally frequentist in character. From Andrew Barron: "Suppose in an i.i.d. sampling model that the parameter value of the distribution from which the data are sampled has the property that the prior probability of Kullback neighborhoods of that value are given positive probability. Then, from that condition alone, does it follow that the risk of the Bayes procedure at that parameter value will converge to zero?" Wing Wong: "Can we construct priors on a very large parameter space (e.g., the space of all densities) so that a 'marginal inference' of a function of the parameter can be viewed as 'optimal' in some sense? Must the prior depend on the function?" Larry Wasserman: "Find a full nonparametric prior on a function space such that the $(1-\alpha)$ posterior probability region has frequentist coverage (approximately/asymptotically) equal to $(1-\alpha)$."

   Many of the problems listed in the other categories below were also raised in the nonparametric context. Indeed, problems surrounding prior specification and identifiability were viewed as particularly virulent in the nonparametric setting. David Dunson: "Nonparametric Bayes models involve infinitely many parameters and priors are typically chosen for convenience with hyperparameters set at seemingly reasonable values with no proper objective or subjective justification." And Stephen Walker: "Despite a lot of recent work on Bayesian nonparametric regression I am far from convinced that the current presented models will stand the test of time. The models are too big and too unidentifiable."

Finally, it was noted by several people that one of the appealing applications of frequentist nonparametrics is to semiparametric inference, where the nonparametric component of the model is a nuisance parameter. These people felt that it would be desirable to flesh out the (frequentist) theory of Bayesian semiparametrics. For example, Thomas Richardson asked for "Bayesian approaches to dealing with mis-specification, e.g., when will a $(1-\alpha)$ posterior credible region for a parameter have $(1-\alpha)$ frequentist coverage even if some ('nuisance') parts of the likelihood are mis-specified?".

4. *Priors.* Not surprisingly, priors were on the minds of many. Elicitation remains a major source of open problems. Tony O'Hagan avers: "When it comes to eliciting distributions for two or more uncertain quantities we are working more or less in the dark." Mike West pointed to the fact that many scientific fields express their prior knowledge in terms of "scientifically predictive models," and using these models in a statistical setting involves the quintessentially Bayesian tasks of understanding assumptions and conducting detailed sensitivity analyses. Aad van der Vaart turned objective Bayes on its head and pointed to a lack of theory for "situations where one wants the prior to come through in the posterior" as opposed to "merely providing a Bayesian approach to smoothing." And Sonia Petrone noted that we often wish to model data that arise from human behavior and human beliefs, and in such settings the modeling of human beliefs thus arises (implicitly at least) in both the likelihood and the prior, and there should be some consistency in our approaches to these specifications.

3. *Bayesian/frequentist relationships.* As already mentioned in the nonparametrics section, many respondents expressed a desire to further hammer out Bayesian/frequentist relationships. This was most commonly evinced in the context of high-dimensional models and data, where not only are subjective approaches to specification of priors difficult to implement but priors of convenience can be (highly) misleading. Open problems discussed here often are couched as statements about frequentist coverage of Bayesian procedures. More broadly, Brad Efron reminds us that "two connecting technologies are empirical Bayes and the bootstrap." Some respondents pined for non-asymptotic theory that might reveal more fully the putative advantages of Bayesian methods; e.g., David Dunson: "Often, the frequentist optimal rate is obtained by procedures that clearly do much worse in finite samples than Bayesian approaches." Finally, some respondents, whose names I will not reveal for their own protection, asked whether there might be a sense in which it is worthwhile to give up some Bayesian coherence in return for some of the advantages of the frequentist paradigm, including simplicity of implementation and computational tractability.

2. *Computation and statistics.* It was interesting to see some disagreement on the subject

of computation, with some people feeling that MCMC has tamed the issue, and with others (the majority by my count) opining that many open problems remain. E.g., Alan Gelfand: "Arguably the biggest challenge is in computation. If MCMC is no longer viable for the problems people want to address, then what is the role of INLA, of variational methods, of ABC approaches?".

Several respondents asked for a more thorough integration of computational science and statistical science, noting that the set of inferences that one can reach in any given situation are jointly a function of the model, the prior, the data and the computational resources, and wishing for more explicit management of the tradeoffs among these quantities. Indeed, Rob Kass raised the possibility of a notion of "inferential solvability," where some problems are understood to be beyond hope (e.g., model selection in regression where "for modest amounts of data subject to nontrivial noise it is impossible to get useful confidence intervals about regression coefficients when there are large numbers of variables whose presence or absence in the model is unspecified a priori") and where there are other problems ("certain functionals for which useful confidence intervals exist") for which there is hope. Terry Speed raised the intriguing possibility of a connection between the notion of "inference being possible" when (and only when) simulation from a model is possible (and this may well be the subject of a future column; not mine, but Terry's).

Several respondents, while apologizing for a certain vagueness, expressed a feeling that a large amount of data does not necessarily imply a large amount of computation; rather, that somehow the inferential strength present in large data should transfer to the algorithm and make it possible to make do with fewer computational steps to achieve a satisfactory (approximate) inferential solution.

Other respondents were concerned with interactions between model complexity and algorithmic complexity; for example Jun Liu referred to a notion of "weak identifiability" in complex latent variable models where even though parameters might be identifiable via a proper posterior the inference algorithm might run aground (e.g., MCMC failing to mix).

1. *Model selection and hypothesis testing.* I have placed this topic as number one not only for the large numbers of respondents mentioning it, but also for the urgency that was transmitted. From Jim Berger: "We just don't have any agreed upon methods, and the problem is especially important because the Bayesian and frequentist methods can differ so much. This is also crucially important because science is choking on the multiplicity problem, and Bayesian model selection is likely the way forward to its solution." George Casella is concerned about lack of theory for inference after selection: "We now do model selection but Bayesians don't seem to worry about the properties of basing inference on the selected model. What if it is wrong? What are the

consequences of setting up credible regions for a certain parameter β_1 when you have selected the wrong model? Can we have procedures with some sort of guarantee?". And many people feel that prior specification for model selection is still wide open.

There are also open problems at the foundations of model selection. José Bernardo: "My favorite problem is to reach some form of agreement on hypothesis testing and model selection. There are two rather different Bayesian attitudes: to compute a posterior probability for the hypotheses (which needs a sharp prior, very different from those commonly used for estimation) or to use decision analysis to minimize an expected loss (which may be done with conventional, possibly noninformative, priors)." David Draper agrees for the need for more work on decision-theoretic foundations in model selection (and he adds that he views Bayesian decision theory for group decision-making as entirely open). Christian Robert holds out for some radical new framework.

On a more practical note, many people noted the lack of off-the-shelf methods for model criticism and diagnostics. Steve MacEachern: "Our current diagnostics are in a sorry state." And David Spiegelhalter: "How best to make checks for prior/data conflict an integral part of Bayesian analysis?" And the last word on the matter goes to Andrew Gelman: "For model checking, a key open problem is developing graphical tools for understanding and comparing models. Graphics is not just for raw data; rather, complex Bayesian models give opportunity for better and more effective exploratory data analysis."

And thus ends my statistician-sourced column, which I've quite enjoyed "writing." I will forgo drawing any grander conclusions at this point, for at least two reasons: (1) I am past my deadline and am being pursued by the Editor of the Bulletin, and (2) I am well over my page limit. I do wish to take the opportunity, however, to solicit reactions from the larger community. I'd enjoy hearing from anyone who feels that my panel of experts has missed a fundamental "open problem" or otherwise wishes to comment on the material presented here. My email is `jordan@stat.berkeley.edu`. With any luck I'll get enough responses to fill my second column.