Statistical Science
2017, Vol. 32, No. 1, 29–32
DOI: 10.1214/16-STS591
Main article DOI: 10.1214/16-STS576

Prior Specification Is Engineering, Not Mathematics

James G. Scott

In their thought-provoking paper, Drs. Simpson et al. argue that "the current practice of prior specification is not in a good shape." I agree, and offer some reasons why this is so, rooted in the culture and practice of Bayesian statistics as it stands today.

The Bayesian collaborator on a scientific project is often put in the position of asking, with all appropriate tact, why a particular prior has been chosen and whether something else might actually be a bit wiser. The experience is, I imagine, like working at a tattoo parlor: it widens your perspective about what kinds of poor choices are even possible.

For the R package maintainer, this experience must be magnified ten- or a hundred-fold. I suspect that Drs. Simpson et al. never would have imagined some of the things that people do with priors, until they undertook the job of writing and supporting an R package that does Bayesian inference for a wide class of models. I appreciate very much the authors' effort here to share their wisdom from the front lines of prior specification, and to formulate some general principles arising from this hard-won practical experience. I will organize my discussion of their article, which is both thought-provoking and excellent, around two broad questions that surround the practical art of prior specification.

DOES THE AUTHORS' PROPOSAL ADDRESS THE PROBLEM?

There is tremendous value in the authors' discussion of criteria for good default priors. Here, they identify many common mistakes, which to my eye have a common theme: choices that make the prior rather more informative than you intended. In particular, my vocabulary has been enriched by the concept of "forced overfitting," in which a default prior has the unintended consequence of rewarding a needlessly complex model. This is most obvious in the case of a variance component for random effects in a hierarchical model, a prototypical kind of nuisance parameter.

However, while it does not diminish my appreciation of the paper, I am not convinced about the "PC prior" formalism itself. For multivariate parameters, in particular, I have not yet been convinced that this formalism, or any other, is adequate to the task of answering the questions of prior choice that I have confronted in my recent scientific collaborations. More generally, I am leery of transferring intuitions gleaned from the scalar case to the high-dimensional case.

I will give a simple example. In Tansey et al. (2017), we describe an application in which the goal is to estimate the background radiation intensity across a wide spatial area. The details are unimportant here, but the essence is this: we have an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that describes spatial adjacencies among locations, and a parameter $\{\theta^{(s)} : s \in \mathcal{V}\}$ at each node in the graph, parametrizing the background radiation at that location. To estimate θ, we used a prior that penalizes first-differences across edges in the graph:

$$p(\theta) \propto \prod_{(r,s)\in\mathcal{E}} p\big(\theta^{(r)} - \theta^{(s)} \mid \tau\big),$$

where τ is a precision parameter. If $p(\cdot)$ is a Gaussian distribution, then this is a traditional Gaussian Markov random field (specifically, an intrinsic CAR prior). This fits in the class of random-effects models described in the authors' Section 3.3, and we could therefore have used equation (3.3) as a prior for τ in a Gaussian CAR model.

But instead, we chose a Laplace prior for these first differences. Why? For several reasons. First, the Laplace prior leads to a nonlinear spatial smoother that adapts to different degrees of smoothness in different regions of the graph, which our situation called for. The Gaussian CAR prior, on the other hand, leads to linear shrinkage, which has important consequences

James G. Scott is Associate Professor of Statistics, Department of Information, Risk, and Operations Management and Department of Statistics and Data Sciences, University of Texas at Austin, 2110 Speedway, Box B6500, Austin, Texas 78712, USA (e-mail: james.scott@mccombs.utexas.edu).