DISCUSSION

M. Clyde

Duke University

I would like to begin by thanking the authors for their many contributions to Bayesian model selection and for providing an excellent summary of the growing body of literature on Bayesian model selection and model uncertainty. The development of computational tools such as the Gibbs sampler and Markov chain Monte Carlo approaches, has led to an explosion in Bayesian approaches for model selection. On the surface, Bayesian model averaging (BMA) and model selection is straightforward to implement: one specifies the distribution of the data, and the prior probabilities of models and model specific parameters; Bayes theorem provides the rest. As the authors point out the two major challenges confronting the practical implementation of Bayesian model selection are choosing prior distributions and calculating posterior distributions. In my experience, I have found that this is especially true in high dimensional problems, such as wavelet regression or non-parametric models, where subjective elicitation of prior distributions is practically infeasible and enumeration of the number of potential models is intractable (Clyde et al. 1998; Clyde and George 2000).

Choice of Prior Distributions

The specification of prior distributions is often broken down into two parts: (1) elicitation of distributions for parameters specific to each model, such as the distribution for regression coefficients in linear models, $p(\beta|\gamma, \sigma^2)$, and (2) selection of a prior distribution over models $p(\gamma)$. For high dimensional problems one cannot specify the prior probability of each γ directly, and practical implementations of Bayesian selection have usually made prior assumptions that the presence or absence of a variable is independent of the presence or absence of other variables. As a special case of this, the uniform prior distribution over models is appealing in that posterior probabilities of models depend only on the likelihood. However, this prior distribution may not be sensible for model averaging when there are groups of similar variables and does not provide the proper "dilution" of posterior mass over similar models (see Clyde 1999; Hoeting et al. 1999), and discussion therein (George 1999; 1999a). In this regard, uniform and independent prior distributions must be used carefully with highly correlated explanatory variables and special consideration should be given to constructing the model space. Even with

Merlise Clyde is Associate Professor, Institute of Statistics and Decision Sciences, Duke University, Durham NC 27708-0251, U.S.A. email: clyde@stat.Duke.EDU.