Comment

Wing Hung Wong

The authors have presented a clear and elegant exposition of the MCMC methodology, illustrated by three substantial applications. Their descriptions of the background of the applications and insightful discussions of the modelling and computational issues will be helpful to all seriously interested in Bayesian computation.

A QUESTION ON THE CHOICE OF PRIORS

There is quite a bit of arbitrariness in the choice of the prior models. For instance, in the prostate cancer example, the scale parameters are assumed to have independent proper gamma distributions. Thus, for each scale parameter one needs to introduce two free constants to describe the gamma prior. Why is it necessary to have this extra level of randomness? On the other hand, the parameter δ in the pairwise-difference prior (6.1) in the nuclear medicine imaging example is treated as a free constant and given the value 2. It seems to me that the role of this latter parameter is quite similar to the scale parameters in the prostate cancer example, namely, to control the strength of local regularity in space or time. Why should it be given a fixed value in this case?

COMMENTS ON NUCLEAR MEDICINE IMAGING

(a) Would the authors please discuss why it is controversial to use Bayesian modelling in measuring uncertainty in image analysis? I am very interested in further elaborations of their position on this issue.

(b) In Section 6.1, it was remarked that the "point spread function" is often known from calibration experiments. Is this the case for the actual study in Section 6.4? The "raw data" presented there consist of a 256×256 image where the photon counts in individual pixels vary between 0 and 93. The direct use of the Poisson model of Section 6.1 would require us to assume, in effect, that there are 256×256 independent counting elements. In actuality, the counting elements in a traditional gamma camera are photomultiplier tubes whose diameters typically are of the order 1–3 cm. Each scintillation event would generate many thousands of light photons collected by several nearby photomultiplier tubes, and the location of the scintillation event is "computed" by the circuitry based on the relative strength of the signals from the several tubes. In principle, the signals from the individual tubes are available and the "computation" of the position of the scintillation event would then become a statistical inference problem! In many cases, it may be reasonable, as a first approximation, to use a Gaussian point spread function with a suitable standard deviation to represent the uncertainty in this measurement of the scintillation position. This depends on the thickness of the scintillating crystal, collimator design and the sizes of the photomultiplier tubes, and I do not necessarily disagree with the authors' treatment in this example. I merely wish to point out that statisticians should not automatically leave the issue of the point spread function to the medical physicists. This is particularly true in more sophisticated imaging modalities such as SPECT and PET. For example, for the 510-keV gamma photons in PET, the effect of Compton scattering would contribute much more significantly to the blurring. Since part of the scattering occurs inside the body, it is not possible to determine the exact effect of this by calibration experiments.

SEQUENTIAL BUILDUP BY MARKOV CHAIN MONTE CARLO

In Section 7, the authors presented a useful update on promising recent developments on the construction of efficient Monte Carlo algorithms. I will supplement their discussion by venturing to outline an idea which I hope will be helpful in this regard. Let us first consider the method of simulated tempering (Marinari and Parisi, 1992) in more detail. Let $f(z)$ be an unnormalized density on a space Z, that is, $f(z)$ is nonnegative but needs not integrate to 1. To sample from $f(\ )$, Marinari and Parisi propose to create a Markov chain with an enlarged state vector (k, z), where z takes value in Z and k ranges from 1 to m. For any k, z is updated according to a transition kernel which has an invariant density proportional to the $1/T_k$ power of $f(\ )$. For example, the update may be one complete Gibbs sampling scan over the components of z. After each update of z, k may be moved to the next

Wing Hung Wong is Professor, Department of Statistics, Chinese University of Hong Kong, Shatin, NT, Hong Kong.