HW6, Bayesian Modelling B 2016/17

Jonathan Rougier
School of Mathematics
University of Bristol UK

In the homeworks, questions with marks are officially 'exam-style', although you can expect any homework question to appear as an exam question, unless it is explicitly 'not examinable'.

Nothing to hand in, since the unit finishes today (3 Mar). The answers will be put on the unit webpage next Fri.

1. Have a really good look at Homework 5, to make sure you can follow all the steps of creating a statistical model and Gibbs-sampling its conditional distribution.

2. Show that the set of distributions which satisfy detailed balance with respect to P is a strict subset of the set of stationary distributions of P. [10 marks]

3. I was browsing a meta-analysis on the efficacy and safety of statin treatment (see `doi:10.1093/qjmed/hcq165`), which included the following passage:

> In order to evaluate the relative effectiveness of each study drug on CVD mortality, we used the Lu-Ades method for combining indirect evidence in mixed-treatment comparisons.[31] We estimated the posterior densities for all unknown parameters using MCMC (Markov chain Monte Carlo) for each model. Each chain used 20 000 iterations with a burn-in of 20 000, thin of 5 and updates varying between 80 and 110. We used the same seed number (SEED = 314 159, equivalent to 10 pi) for all chains. The choice of burn-in was chosen according to Gelman–Rubin approach.[32] We applied the covariate of LDL-C change and also statin dosing (high or moderate determined by the Canadian Compendium of Pharmaceuticals and Specialties),[33] using an approach developed by Cooper *et al.*[34] We assessed convergence based on trace plots and time series plots. The accuracy of the posterior estimates was done by calculating the Monte Carlo error for each parameter. As a rule of thumb, the Monte Carlo error for each parameter of interest is less than ~5% of the sample standard deviation. All results for the mixed-treatment ana-

   (a) Explain how the authors assessed the convergence of their MCMC sampler. [15 marks]

   (b) Explain how the authors assessed the accuracy of their MCMC estimates. [10 marks]

(c) Identify the information *not* given in this text which you would need to replicate their assessment of convergence and of the accuracy of their estimates. [10 marks]

4. (Not examinable) Study the handout on convergence diagnostics, and create an R function `brooks.diag` which implements this test. You may find it helpful to look at the first few lines of `gelman.diag` in the `coda` package in R. Insert your `brooks.diag` function into your `Rats.R` script, to test the convergence of the sampler.