# Computationally-Intensive Statistical Methods

## Assessment Exercises

The work provided here is intended to take students up to half a week to complete. Students should talk to their supervisors to find out whether or not their department requires this work as part of any formal accreditation process. It is anticipated that departments will decide on the appropriate level of assessment locally, and perhaps select some of these exercises.

Ex 1 There is another type of bootstrap-based confidence interval not discussed in the lectures, called *ABC*. Use the references to find out about it, and write up both a description of the method and a comparison of its strengths and weaknesses, compared to the methods in the lecture notes.

Ex 2 Practical 9 analysed only part of the Australian AIDS survival data. Extend the analysis to the complete dataset.

Ex 3 Venables and Ripley (2002, §14.5) discuss a time series of beaver body temperature ('beaver 2') that has a clear changepoint at about the time the beaver leaves its lodge. Venables and Ripley assume that the changepoint is exactly at the time of leaving the lodge, but this is just an assumption.

(a) Suppose that the temperatures are IID $N(\gamma_1, \sigma^2)$ prior to time $T$ (a random multiple of 10 mins) and $N(\gamma_2, \sigma^2)$ thereafter. Set up a suitable Bayesian model, create an MCMC scheme to sample from it, run it, and consider the evidence for the temperature rise coinciding with leaving the lodge.

(b) As Venables and Ripley show, the observations are strongly autocorrelated. Modify your sampling scheme to allow for an AR(1) process.

Ex 4 Consider a probit model for LD50 for the following data

| dose | animals | deaths |
|---|---|---|
| 50 | 5 | 1 |
| 100 | 10 | 4 |
| 200 | 5 | 2 |
| 300 | 10 | 6 |
| 2000 | 5 | 4 |

with a suitable vague prior. Find a suitable HPD region for LD50. How does it compare to a logit model and to a classical GLM analysis?