Computer Practical 3
Markov Chains and Monte Carlo

1. A warm-up. In a simplified model of the game of Monopoly, we consider the motion of the piece around a loop of 40 spaces. We can model this as a Markov chain on the integers $0, \dots, 39$ in which the transition kernel adds the result of two six-sided dice to the current state modulo 40 to obtain the new state.

   (a) Implement a piece of R code which simulates this Markov chain.

   (b) Run the code for a large number of iterations, say $100,000$, and plot a histogram of the states visited.

   (c) Based on the output of the chain, would you conjecture that there is an invariant distribution for this Markov chain? If so, what?

   (d) Write the transition kernel down mathematically.

   (e) Show that the Markov kernel you have written down is invariant with respect to any distribution conjectured in part (c).

2. Gibbs Sampling: recall the Poisson changepoint model discussed in lectures, and on p26 of the supporting notes, and think about the following closely related model.

   Observations $y_1, \dots, y_n$ comprise a sequence of M iid $\mathsf{N}(\mu_1, 1)$ random variables followed by a second sequence of $n - M$ iid $\mathsf{N}(\mu_2, 1)$ random variables. M, μ_1 and μ_2 are unknown.

   The prior distribution over M is a discrete uniform distribution on $\{1, \dots, n-1\}$ (there is at least one observation of each component). The prior distribution over μ_i $(i = 1, 2)$ is $\mathsf{N}\left(0, 10^2\right)$. The three parameters are treated as being a priori independent.

   (a) Write down the joint density of $y_1, \dots, y_n, \mu_1, \mu_2$ and M and obtain the posterior distribution of μ_1, μ_2 and M, up to proportionality, in as simple a form as you can.

   (b) Find the "full conditional" distributions of μ_1, μ_2 and M. (i.e. the conditional distributions of each of these variables given all other variables).

   (c) Implement a Gibbs sampler which makes use of these full conditional distributions in order to target the posterior distribution identified in part (b).

   (d) Simulate some data from the model for various parameter values and test your Gibbs sampler.

   (e) How might you extend this algorithm if instead of a changepoint model you had a mixture model in which every observation is drawn from a mixture, i.e.:

   $$Y_1, \dots, Y_n \overset{\text{iid}}{\sim} p\mathsf{N}(\cdot; \mu_1, 1) + (1-p)\mathsf{N}(\cdot; \mu_2, 1)$$

   (so the likelihood is $\prod_{i=1}^{n} [p\mathsf{N}(y_i; \mu_1, 1) + (1-p)\mathsf{N}(y_i; \mu_2, 1)]$ in which p, μ_1 and μ_2 are unknown (and M is no longer a parameter of the model)?

   Consider the following things:

   i. The prior distribution over p.

   ii. Any other variables you may need to introduce.

   iii. The resulting algorithm.

   If you have time, implement the resulting algorithm and apply it to some simulated data.