

Computer Intensive Statistics: APTS 2025–26 Computer Practical

2

Markov Chains and Monte Carlo

Andi Wang (andi.wang@warwick.ac.uk),

July 2026

The first problem sheet contained several problems and if you want to continue working on those that is fine; this sheet just contains a couple of simple questions to give you a chance to try some Markov chain-based problems.

1. *A warm-up which also appeared in the preliminary material; if you've never implemented something like this before then this might be a useful preliminary step.*

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 with a state space consisting of 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) Check whether the Markov kernel you have written down is invariant with respect to any distribution conjectured in part (c).
2. *An actual Gibbs Sampler.*

Recall the Poisson changepoint model discussed in lectures, and further detailed in the supporting notes, and think about the following closely related model: Observations $y_1, \dots, y_n$ comprise a sequence of M iid $N(\mu_1, 1)$ random variables followed by a second sequence of $n - M$ iid $N(\mu_2, 1)$ 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 $N(0, 10^2)$. 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 making 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.:

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

(The likelihood is now $\prod_{i=1}^n [p\mathbf{N}(y_i; \mu_1, 1) + (1-p)\mathbf{N}(y_i; \mu_2, 1)]$, with p , μ_1 , and μ_2 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.

3. *Exercise using string diagrams.* Consider the following variant of the two-component Gibbs sampler (sometimes referred to as the Synchronous Gibbs Sampler¹):

Given current position (x_n, y_n) and the two conditionals $f_{X|Y}$ and $f_{Y|X}$, sample $x_{n+1} \sim f_{X|Y}(\cdot|y_n)$ and sample $y_{n+1} \sim f_{Y|X}(\cdot|x_n)$.

Notably y_{n+1} is simulated given x_n , *not* x_{n+1} .

- (a) Try implementing this on a bivariate Gaussian. You should find this does *not* have the full f as its invariant distribution in general. In fact, its invariant distribution is the product of the two marginals of f .
- (b) Represent the synchronous Gibbs update kernel $P : \mathcal{X} \otimes \mathcal{Y} \rightarrow \mathcal{X} \otimes \mathcal{Y}$ as a string diagram.
- (c) Prove directly using string diagrams that indeed the product of the two marginals is the invariant distribution of P . (In fact, the algorithm is even reversible.)

¹See e.g. Gonzalez, Low, Gretton, Guestrin (2011). Parallel Gibbs sampling: From colored fields to thin junction trees, JMLR.