Stats 295, Homework 2

Due date: October 25

Programming exercises can be completed in any language of your choice.

1. Ehrenfest model of diffusion

   (a) Implement simulation of the Ehrenfest diffusion model and plot one realization of the chain with the total number of particles N=100 and 300 time steps.

   (b) Use the ergodic theorem to approximate the stationary variance of the number of particles and compare your estimate with the analytical result.

2. Recall that we used the following Metropolis-Hastings algorithm to update α and β in the beta-binomial hierarchical model. To propose new values of a positive parameter we multiply its current value by $e^{\lambda(U-0.5)}$, where $U \sim U[0,1]$ and λ is a tuning constant. Prove that the proposal density is

$$q(y_{\text{new}} \mid y_{\text{cur}}) = \frac{1}{\lambda y_{\text{new}}}.$$

3. Consider a two state continuous-time Markov SIS model, where the disease status X_t cycles between the two states: 1=susceptible, 2=infected. Let the infection rate be λ_1 and clearance rate be λ_2. Suppose that an individual is susceptible at time 0 ($X_0 = 1$) and infected at time T ($X_T = 2$). We don't know anything else about the disease status of this individual during the interval $[0, T]$. If T is small enough, it is reasonable to assume that the individual was infected only once during this time interval. We would like to obtain the distribution of the time of infection I, conditional on the information we have.

   (a) Implement a Metropolis-Hastings sampler to draw realizations from the distribution

$$\Pr(I \mid X_0 = 1, X_t = 2, N_t = 1) \propto \Pr(0 < t < I : X_t = 1, I \le t < T : X_t = 2),$$

   where N_t is the number of infections. Since X_t is a continuous-time Markov chain, the last probability (it is actually a density) can be written as

$$\Pr(0 < t < I : X_t = 1, I \le t < T : X_t = 2) = \underbrace{\lambda_1 e^{-\lambda_1 I}}_{\text{density of time until infection}} \times \overbrace{e^{-\lambda_2 (T-I)}}^{\text{prob of staying infected}}.$$

   To make things concrete, set $\lambda_1 = 0.1$, $\lambda_2 = 0.5$ and $T = 1.0$. For your proposal distribution, use a uniform random walk with reflective boundaries 0 and T. In other words, given a current value of the infection time t_c, generate $u = \text{Unif}_{[t_c - \delta, t_c + \delta]}$ ($2\delta < T$) and then make a proposal value

$$t_p = \begin{cases} u & \text{if } 0 < u < T, \\ 2T - u & \text{if } u > T, \\ -u & \text{if } u < 0. \end{cases}$$

   This is a symmetric proposal.

(b) Run your MCMC for 1000 iterations and plot the histogram of the posterior distribution of the infection time.

(c) Try a couple of sets of values for λ_1 and λ_2 and examine the effect of these changes on the posterior distribution of the infection time. Comment on the effect of the relationship between λ_1 and λ_2 on the shape of the infection time posterior density/histogram.