Statistics 5444: MCMC Assignment

For each homework assignment, turn in at the beginning of class on the indicated due date. Late assignments will only be accepted with special permission. Write each problem up *very* neatly (LaTeX is preferred). Show all of your work.

The Metropolis-Hastings sampler is a key device in Markov Chain Monte Carlo (MCMC), and can be used to solve many complex inference problems. In this HW, you will implement a very basic sampler. The following is a simple 1-D example of a Metropolis Sampler for sampling from a density of the form $\pi(\theta)$, where θ is an unknown parameter. To sample correct values of θ, follow the steps:

1. Set θ to some fixed value (initialize)
2. $\theta^* \sim g(\theta^*|\theta)$ (proposal function)
3. $\alpha = \min(1, \frac{\pi(\theta^*)}{\pi(\theta)} \frac{g(\theta|\theta^*)}{g(\theta^*|\theta)})$ (acceptance ratio)
4. wtih probability α, set $\theta = \theta^*$.
5. Store θ, and repeat starting at step 2.

Notes:
1) An example of a symmetric proposal is $g(\theta^*|\theta) = \frac{1}{\sqrt{2\pi}\psi} e^{-(\theta^*-\theta)/2\psi^2}$. ψ in this example is the step size.
2) MCMC samplers *eventually* sample from the target distribution.

Problem 1

Show that the above algorithm eventually draws samples from $\pi(.)$. That is show that $\pi(.)$ is the stationary distribution for the Markov chain defined by the Metropolis-Hastings algorithm.

Problem 2

Let $X = (x_1, \dots x_n)$ and let $x_i \sim N(\mu = 200, \phi = \frac{1}{2})$, where $\phi = 1/\sigma^2$

Part 1

Under standard reference priors ($p(\mu, \phi) \propto 1/\phi$), the posterior density is (proportionality constants dropped):

$$p(\mu, \phi|X) \propto \phi^{n/2-1} \exp\left(-\frac{\phi}{2}\sum_{i=1}^{n}(x_i - \mu)^2\right).$$

Write down pseudo code for the Metropolis-Hastings Sampler. Be specific about your choice of proposal functions.

Part 2

Implement a Metropolis-Hastings sampler for sampling from the distribution for $(\mu, \phi|X)$, where X is a 100 simulated data points from the above model. Initialize the sampler at $\mu_0 = 0$ and $\phi = 5$. Show the trace plots for both μ and ϕ. Report the burn-in time and draw histograms for both of the marginal posteriors (after burn-in).

Part 3

Consider the K component mixture model:

$$x_i \sim \sum_{k=1}^{K} \pi_k N(x|\theta_k, \sigma_k^2) \qquad (i = 1, \ldots, n),$$

where $N(x|\theta_k, \sigma_k^2)$ represents a Gaussian distribution with mean θ_k, and variance σ_k^2, with $\sum_k \pi_k = 1$.
For $K = \{2, 5, 20, 50\}$, implement the Gibbs sampling procedure discussed in class (you will need to choose values for your parameters when you simulate data, as well as n).

Part 4

Provide a 2 page report summarizing the performance of your Gibbs sampler under varying values of K, n, and your parameters. This is an open ended simulation study, so there are many choices for you to make. The point is that you describe the sampling procedure under varying situations, and conclude with your overall experience with both implementing the sampler and its overall performance.