Statistics 5114 Extra Credit: MCMC Assignment

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 assignment, you will implement a very basic sampler. The following is a simple 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

Let $X = (x_1, \ldots x_n)$ and let $x_i \sim N(\mu = 200, \phi = \frac{1}{2})$, where $\phi = 1/\sigma^2$. Under standard reference priors ($p(\mu, \phi) \propto 1/\phi$), the posterior density is (proportionality constants dropped):

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

Part 1

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 represents $N = \{10, 30, 100\}$ simulated data points from the above model. Initialize the sampler at $\mu_0 = 0$ and $\phi_0 = 5$.

Part 3

Provide a nice (up to 2 pages) write up detailing all of the above, as well as the following: 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). Also provide contour plots of the joint samples from $(\mu, \phi|X)$. Note: you'll have summaries for 3 simulations with $N = \{10, 30, 100\}$.