EXERCISES 8

Exercise 8.1. Consider the following unadjusted Langevin algorithm:

$$X_{n+1} = X_n + \gamma \nabla \log p_\star(X_n) + \sqrt{2\gamma\beta^{-1}} W_{n+1},$$

where $W_{n+1} \sim \mathcal{N}(0, 1)$ and $\gamma > 0$ is the step size. Assume that $p_\star$ is a Gaussian:

$$p_\star(x) = \mathcal{N}(x; \mu, \sigma^2),$$

where $\mu \in \mathbb{R}^d$ and $\sigma^2 > 0$. Derive the limiting distribution $p_\star^{\gamma,\beta}(x)$ of X_n as $n \to \infty$. Discuss any constraints on γ for convergence and the limit $\beta \to \infty$.

Exercise 8.2. Let

$$p_\star(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2},$$

where $\nu > 0$ is a parameter. This is called *Student's t-distribution*. Derive the gradient of $\log p_\star(x)$ and implement the unadjusted Langevin algorithm and Metropolis-adjusted Langevin algorithm for this distribution. Plot the histogram of the samples for different values of ν and γ and discuss any interesting result you observe.

Exercise 8.3. Assume that you are given conditional distributions

$$p(x|y) = \text{Exp}(y)$$
$$p(y|x) = \text{Exp}(x),$$

and you have implemented the Gibbs sampler using these conditionals.

1. Write the Gibbs sampler algorithm (on paper) assuming only accessing uniform random numbers.
2. Now show that these conditionals are consistent with a target of the form $p(x, y) \propto \exp(-xy)$.
3. Show that the "target density" is not a valid density. What issue does this highlight about Gibbs sampling implementations?
4. Finally, assume that we have the joint density $p(x, y)$ in the same form as you derived in Part 2, but constrained to $(x, y) \in (0, 1) \times (0, 1)$. In this case, derive the Gibbs sampler by deriving full conditionals.

Exercise 8.4. Consider a generic target $p(x)$ and a proposal $q(x)$. Consider two algorithms: (i) The rejection sampler with proposal $q(x)$ and (ii) the independent MH sampler with proposal $q(x)$. Assume that M is chosen optimally for the rejection sampler. Prove that, for a given x, the MH acceptance ratio is always higher than the rejection sampler acceptance ratio, i.e., the MH algorithm accepts more often than the rejection sampler.