Note: Homework 3A will consist of problems 1, 2, 6, 7, and 8. The remaining problems will constitute Homework 3B and will be due at a later date.

1. **Parts (a) and (b) are mandatory for graduate students and extra credit for undergrads.** The Gibbs sampler can take a long time to converge if the target distribution is multimodal. Suppose we are trying to sample from a density for a parameter θ that is the mixture of three normal densities: θ has density $f(\theta) = 0.45f_1(\theta|\mu = -3, \sigma^2 = 1/3) + 0.10f_2(\theta|\mu = 0, \sigma^2 = 1/3) + 0.45f_3(\theta|\mu = 3, \sigma^2 = 1/3)$, where $f_j(\cdot|\mu, \sigma^2)$ here represents the normal density with mean μ and variance σ^2. The marginal distribution of θ is plotted on the last page.

   Note that we could easily sample from this distribution using ordinary Monte Carlo methods, but consider using Gibbs sampling to sample from it.

   (a) [**Extra credit for undergrads.**] If the indicator $\delta \in \{1, 2, 3\}$, argue that the full conditional distribution for θ is

   $$\theta|\delta \sim N(\mu_\delta, \sigma_\delta^2).$$

   What are μ_δ and σ_δ^2 here, for $\delta = 1, 2, 3$?

   (b) [**Extra credit for undergrads.**] Use Bayes' theorem to show that the full conditional for δ is

   $$P[\delta = k|\theta] = \frac{P[\delta = k] \times \text{dnorm}(\theta, \mu_k, \sigma_k)}{\sum_{m=1}^{3} P[\delta = m] \times \text{dnorm}(\theta, \mu_m, \sigma_m)}, \text{ for } k \in \{1, 2, 3\}$$

   where `dnorm` represents the normal density function (R shorthand).

   (c) Write a Gibbs sampling algorithm (you can use the helpful R commands given on the course web page) to sample from the joint density of (θ, δ). Begin the chain with the initial values $\delta^{[0]} = 2$ and $\theta^{[0]} = 0$, and generate 1000 values of θ. Give a plot of a relative frequency histogram of the θ values – using a command like `hist(theta.values, freq=F)` – and comment on how it compares to the true marginal density of θ plotted on the last page. (You could try repeating part (c) a few times to get a sense for the variability in the Gibbs sampler results, as well.)

   (d) Repeat part (c), but generate 40000 values of θ. Again give a plot of a relative frequency histogram of the θ values and comment on how it compares to the true marginal density of θ plotted on the last page. (You could try repeating part (d) a few times to get a sense for the variability in these Gibbs sampler results, as well.)

   (e) Try trace plots for the θ values and plots of the autocorrelation functions for parts (c) and (d). Comment on what these MCMC diagnostics tell you.

2. **Parts (a) and (d) are mandatory for graduate students and extra credit for undergrads. Part(c) is extra credit for everyone.** Two separate test developers

have created different IQ tests. Each of these tests is designed so that scores follow a normal distribution with population mean 100 and standard deviation 15. Hence any set of test scores can be standardized to so that the standardized scores follow a $N(0,1)$ distribution. We wish to investigate the coefficient of correlation ρ between test-takers' scores on the two tests. Consider a random sample of n test-takers, who each take both IQ tests. Let $X_1, \ldots, X_n$ be the test-takers' standardized scores on the first test, and let $Y_1, \ldots, Y_n$ be the corresponding standardized scores on the second test. The likelihood is

$$L(\rho|\mathbf{x},\mathbf{y}) = [2\pi]^{-n}[1-\rho^2]^{-n/2}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\sum x_i^2 - 2\rho\sum x_i y_i + \sum y_i^2\right]\right\}.$$

(a) [**Extra credit for undergrads.**] Suppose we use the prior $p(\rho) = 1$, $0 < \rho < 1$, for ρ. Show that the posterior for ρ is

$$p(\rho|\mathbf{x},\mathbf{y}) \propto [(1-\rho^2)]^{-n/2}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\sum x_i^2 - 2\rho\sum x_i y_i + \sum y_i^2\right]\right\}$$

(b) We gather data on 13 test-takers. Their standardized scores on the two tests are:

```
X_i:  0.92  0.42  3.62  0.89 -0.69  0.45 -0.11 -0.14 -0.47  1.09 -0.34  0.62  0.27
Y_i:  0.26  1.65  2.10  0.62 -1.16  1.29 -0.82 -0.36 -0.29  0.86  0.19  1.25  0.33
```

Plugging in the necessary summary statistics, write and simplify the posterior (up to a constant of proportionality). [Hint: Use the formula from part (a), and your posterior density should be an expression that depends on ρ.]

(c) [**Part (c) is Extra credit for EVERYONE.**] We will use the following proposal density to generate values of ρ:

(i) Given current value $\rho^{[t]}$, sample $\rho^* \sim \text{Uniform}(\rho^{[t]} - 0.2, \rho^{[t]} + 0.2)$.

(ii) If the sampled $\rho^* < 0$, then set $\rho^* = |\rho^*|$.

(iii) If the sampled $\rho^* > 1$, then set $\rho^* = 2 - \rho^*$.

Argue that this is a symmetric proposal density.

(d) [**Extra credit for undergrads.**] Explain carefully and completely how the provided R code on the course web page performs the Metropolis-Hastings algorithm.

(e) Using the code on the course web page, sample from the posterior distribution of ρ. Perform diagnostics to check convergence and check autocorrelation, and perform remedial action if needed. Summarize the posterior distribution of ρ, including giving an estimated density plot, point estimate, and 95% interval estimate for ρ.

3. In class (for the Prussian cavalry example) we derived the posterior predictive distribution in the case where $Y_1, \ldots, Y_n \sim$ i.i.d. Poisson(λ), with the Gamma prior $\lambda \sim G(2,4)$. Instead of deriving the posterior predictive distribution analytically, we could have sampled from it using Monte Carlo methods.

   (a) Randomly sample $\lambda^{[1]}, \ldots, \lambda^{[J]}$ from a Gamma(shape $= \sum y_i + 2$, rate $= n + 4$) distribution. Using these, sample $Y^{[1]}, \ldots, Y^{[J]}$ from Poisson($\lambda^{[j]}$) distributions, $j = 1, \ldots, J$.

   (b) Plot the approximate posterior predictive distribution, similarly as our in-class example. How does it appear to compare to the observed Prussian army data distribution?

4. In a NASA experiment, 14 male rats were sent into space. When they returned, the red blood cell mass (in ml) of each rat was measured. In addition, 14 other male rats were kept on earth during the same period of time. Those rats also had their red blood cell mass measured. Assume the red blood cell masses for the two groups can be modeled with a normal distribution, with equal variances across the two groups. The data are:

   ```
   Space rats: 8.59 8.64 7.43 7.21 6.87 7.89 9.79 6.85 7.00 8.80 9.30 8.03 6.39 7.24
   Earth rats: 8.65 6.99 8.40 9.66 7.62 7.44 8.55 8.70 7.33 8.58 9.88 9.94 7.14 9.04
   ```

   (a) Suppose the research question of interest was to test whether the mean red blood cell mass differed for the two groups. Answer this question based on a Bayesian hypothesis test. Clearly state your prior specifications (use $\mu_\Delta = 0, \sigma_\Delta^2 = 1/5$). Give a posterior probability for each hypothesis being true.

   (b) Suppose the research question of interest was to test whether the mean red blood cell mass for the space group was lower than the mean red blood cell mass for the control group. Answer this question based on a Bayesian hypothesis test. Clearly state your prior specifications. (The researcher believed *a priori* that the rat population as a whole might have average red blood cell mass somewhere around 7 ml, but was not at all sure about the effect of the space travel.) Give a posterior probability for each hypothesis being true.

5. A physician is interested in determining whether the mean systolic blood pressure of a certain set of patients is less than 130. She takes a random sample of 17 patients and measured their systolic blood pressure. Assume the measurements follow a $N(\mu, \sigma^2)$ distribution, with μ unknown and known $\sigma^2 = 225$. (You can use normal-normal results from Chapter 5 to obtain the posterior distribution for μ.) The physician says *a priori* that she is 95% sure that the true mean systolic blood pressure is between 120 and 140. The data are:

   118 140 90 150 128 112 134 140 112 126 112 148 124 130 142 105 125

   (a) Conduct a Bayesian hypothesis test of $H_0 : \mu \geq 130$ vs. $H_a : \mu < 130$, basing your conclusions on the posterior distribution for μ. Clearly state your prior specifications.
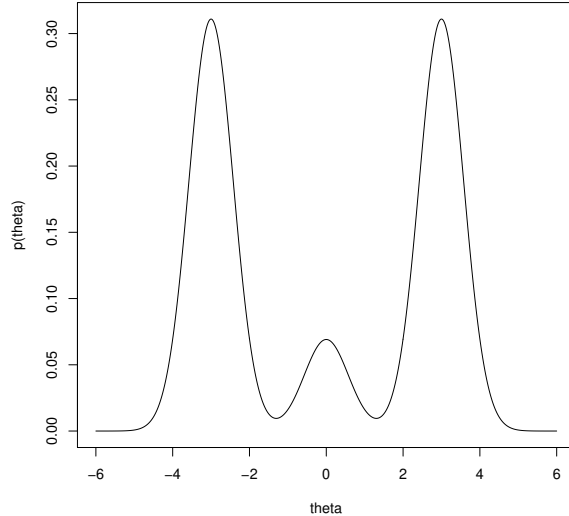
Figure 1: True distribution for θ.

(b) Conduct a classical t-test using $\alpha = 0.05$. Are the substantive conclusions any different from those in part (a)?

6. Do Problem 6.3 from the *Bayes Rules!* textbook. [Read Section 6.3 for more details about "mixing slowly."]

7. Do Problem 7.6 from the *Bayes Rules!* textbook. [Hint: Use R functions like `rnorm` and `runif` to draw the requested value, and **(just for the purposes of this exercise!)**, remember to enter `set.seed(84735)` before each random draw, so that the random number that you draw will match the random number that I draw.]

8. Do Problem 7.7 from the *Bayes Rules!* textbook. [Hint: Use R functions like `dnorm`, `dunif`, and `dexp` to calculate the value of the proposal density for the specified current value and proposed value of the parameter. These proposal density values are the q parts that go in the numerator and denominator of the Metropolis-Hastings acceptance ratio probability.]

9. Do Problem 8.8 from the *Bayes Rules!* textbook. [You do not have to provide the sketches of the intervals on the posterior pdf. And you can use the `hpd` function in the `TeachingDemos` package to get the HPD intervals.]

10. Do Problem 8.9 from the *Bayes Rules!* textbook.