Homework 3

Problem 1

Consider the multivariate normal model with known variance-covariance matrix Σ and unknown mean μ. Derive the posterior distribution of μ (show all the algebraic steps to get to the final distribution).

Problems 2

Importance sampling Devise and implement an importance sampler for estimate the expected value of a mixture of two beta distributions (e.g., 0.3*beta(5,2)+0.7*beta(2,8)). Use the same sample to estimate the probability that this random variable is included in the interval [0.45, 0.55].

Problem 3

Prove that the following algorithm is equivalent to the Accept-Reject algorithm:

1. Generate $X \sim g$;
2. Generate $U \mid X = x \sim U_{[0,Mg(x)]}$;
3. Accept $Y = X$ if $U \leq f(x)$;
4. Return to 1. otherwise

This is equivalent to prove that $P(Y < y) = P(X < x \mid U < f(x))$.

Problem 4

Consider a univariate normal model with mean μ and variance τ. Suppose we use a Beta(2,2) prior for μ (somehow we know μ is between zero and one) and a log-normal(1,10) prior for τ (recall that if a random variable X is log-normal(m, v) then log$X$ is N(m, v)). Assume a priori that μ and τ are independent. Implement a Metropolis-Hastings algorithm to evaluate the posterior distribution of μ and τ. Remember that you have to jointly accept or reject μ and τ. Also compute the posterior probability that μ is bigger than 0.5.

Here are the data:

2.3656491 2.4952035 1.0837817 0.7586751 0.8780483 1.2765341
1.4598699 0.1801679 -1.0093589 1.4870201 -0.1193149 0.2578262

Note, this is a lot like the bioassay example we did in class.

Problem 5

Exercises from Hoff's book.

6.2 Mixture model: The file `glucose.dat` contains the plasma glucose concentration of 532 females from a study on diabetes (see Exercise 7.6).

a) Make a histogram or kernel density estimate of the data. Describe how this empirical distribution deviates from the shape of a normal distribution.

b) Consider the following mixture model for these data: For each study participant there is an unobserved group membership variable X_i which is equal to 1 or 2 with probability p and $1-p$. If $X_i = 1$ then $Y_i \sim \text{normal}(\theta_1, \sigma_1^2)$, and if $X_i = 2$ then $Y_i \sim \text{normal}(\theta_2, \sigma_2^2)$. Let $p \sim \text{beta}(a, b)$, $\theta_j \sim \text{normal}(\mu_0, \tau_0^2)$ and $1/\sigma_j \sim \text{gamma}(\nu_0/2, \nu_0\sigma_0^2/2)$ for both $j = 1$ and $j = 2$. Obtain the full conditional distributions of $(X_1, \ldots, X_n)$, p, θ_1, θ_2, σ_1^2 and σ_2^2.

c) Setting $a = b = 1$, $\mu_0 = 120$, $\tau_0^2 = 200$, $\sigma_0^2 = 1000$ and $\nu_0 = 10$, implement the Gibbs sampler for at least 10,000 iterations. Let $\theta_{(1)}^{(s)} = \min\{\theta_1^{(s)}, \theta_2^{(s)}\}$ and $\theta_{(2)}^{(s)} = \max\{\theta_1^{(s)}, \theta_2^{(s)}\}$. Compute and plot the autocorrelation functions of $\theta_{(1)}^{(s)}$ and $\theta_{(2)}^{(s)}$, as well as their effective sample sizes.

d) For each iteration s of the Gibbs sampler, sample a value $x \sim \text{binary}(p^{(s)})$, then sample $\tilde{Y}^{(s)} \sim \text{normal}(\theta_x^{(s)}, \sigma_x^{2(s)})$. Plot a histogram or kernel density estimate for the empirical distribution of $\tilde{Y}^{(1)}, \ldots, \tilde{Y}^{(S)}$, and compare to the distribution in part a). Discuss the adequacy of this two-component mixture model for the glucose data.

Problem 6

Exercises from Hoff's book.

7.1 Jeffreys' prior: For the multivariate normal model, Jeffreys' rule for generating a prior distribution on $(\boldsymbol{\theta}, \Sigma)$ gives $p_J(\boldsymbol{\theta}, \Sigma) \propto |\Sigma|^{-(p+2)/2}$.

a) Explain why the function p_J cannot actually be a probability density for $(\boldsymbol{\theta}, \Sigma)$.

b) Let $p_J(\boldsymbol{\theta}, \Sigma | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ be the probability density that is proportional to $p_J(\boldsymbol{\theta}, \Sigma) \times p(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n | \boldsymbol{\theta}, \Sigma)$. Obtain the form of $p_J(\boldsymbol{\theta}, \Sigma | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$, $p_J(\boldsymbol{\theta} | \Sigma, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ and $p_J(\Sigma | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$.