Stat 532 **Name:** ______________________
HW 5
Due 10/31/16

Please prepare your solutions using LaTeX or another word processing software.

1. (15 points) For this question we are going to take a deeper look at the impact of our prior distributions and the number of observed data points. Generate fifteen observations from a standard normal distribution and use independent priors on the parameters from the normal distribution: $p(\theta) \sim N(\mu_0, \tau_0^2)$ and $p(\sigma^2) \sim IG(\nu_0/2, \nu_0\sigma_0^2/2)$. This will require running the a Gibbs sampler to estimate the posterior distribution.

   First plot the marginal posterior distributions for θ under the flat prior $p(\theta, \sigma^2) \propto 1$. Note that this is proportional to the likelihood function. Next, rerun the Gibbs sampler using at least 5 different values for μ_0, τ_0^2, where some of the values should be close to the truth and some should be very different. The same principles apply for σ^2, but fix $p(\sigma^2) \propto \frac{1}{\sigma^2}$ so that we can see the difference for θ. Make sure to look at a few situations where τ_0^2 is very small and very large. Explain what you have learned from this question and create a figure that contains the prior, likelihood, and posterior on a single graph.

2. (10 points) Assume you are faced with a modeling scenario for a multivariate normal response where you do not have a good sense of the covariance structure. Select and justify a prior distribution for Σ of (Σ^{-1}).

3. (10 points) Verify that for a multivariate normal distribution, the variance and expectation can be extracted from the kernel, $\exp\left[-\frac{1}{2}\left(\tilde{\theta}^t A\tilde{\theta} - \tilde{\theta}^T B + \dots\right)\right]$, where the variance is A^{-1} and the expectation is $A^{-1}B$.

4. (20 points) Simulate data from a multivariate normal distribution, with $p = 3$. Choose a mean vector $\tilde{\mu}$ and covariance matrix Σ. Run a Gibbs sampler and verify that your sampler returns the true values. Comment on the results.

5. This result is an important tool for modeling data from "heavy-tailed" distributions. Let $y \sim N(\theta, \sigma^2 \times \gamma)$. Thus the sampling model for y can be written as:

$$p(y|\theta, \sigma^2\gamma) \propto (1/\sigma^2\gamma)^{1/2} \exp\left(-\frac{1}{2}\frac{(y-\theta)^2}{\sigma^2\gamma}\right) \qquad (1)$$

   (a) (15 points) Let $\sigma^2 \sim IG(a, b)$. Find $p(y|\theta, \gamma) = \int p(y|\theta, \gamma, \sigma^2)p(\sigma^2)d\sigma^2$

(b) (10 points) State what a and b must be for $p(y|\theta, \gamma)$ to have a Cauchy distribution and state the *location* and *scale* parameters. (Note: a Cauchy distribution is a t distribution with 1 degree of freedom).

(c) (10 points) Recall that if $p(y) = \int p(y|\lambda)p(\lambda)d\lambda$, then we can generate $p(y)$ via the following algorithm:

$$\begin{aligned} \text{sample } \lambda^* &\sim p(\lambda) \\ \text{sample } y^* &\sim p(y|\lambda^*). \end{aligned}$$

The resulting y^* is a sample from $p(y)$.

Use this method to simulate 1000 Cauchy random variables, using the result you obtained from part 1. Plot a histogram and state the minimum and maximum of your 1000 sample draws.

(d) (5 points) How might this result be useful in the context of modeling over-dispersed data?

6. (5 points) Select and state your paper choice for the project. There is a discussion board on D2L to make sure we don't have duplicate selections. Do you have any specific questions or concerns about the project?