AMS 206 – Homework Assignment #5
due Wednesday, February 22

Required Problems

1. Consider a manufacturing process that is filling 2-liter soda bottles. In general, processes such as this can be considered to give normally distributed output, perhaps with known standard deviation, and here we will assume that we know $\sigma = .05$. We also know that there is a hard upper bound, in that not more than 2.08 liters will fit into the 2-liter bottle. We might also assume that there is a lower bound, because bottles with too little liquid (say 1.2 liters) won't move along the conveyor belt correctly and will require stopping the machines. Thus we take the likelihood to be:

$$f(\mathbf{y}|\mu) \propto \exp\left\{-\frac{1}{2*(.05)^2}\sum_{i=1}^{n}(y_i-\mu)^2\right\}\prod_{i=1}^{n} I_{\{1.2\le y_i \le 2.08\}}.$$

   (If you think about it, you can ignore the range restrictions in the likelihood in this problem.) Suppose we want to put a vague prior on μ, so we make it uniform between 1.2 and 2.08 liters. Use the `bottles` data from the class website and fit this model with Metropolis-Hastings.

   (a) Compute your overall acceptance probability, and adjust your proposal distribution until your acceptance probability is between 0.2 and 0.5.

   (b) Show your trace plot. If you have burn-in, include plots both with and with-out the burn-in.

   (c) Compare the posterior mean and variance from your MCMC sample to the maximum likelihood estimates $\hat{\mu} = \bar{y}$ and $s_\mu^2 = \frac{s^2}{n} = \frac{\sum_{i=1}^n (y_i-\bar{y})^2}{n(n-1)}$.

   (d) Does it appear that the range restriction has any effect on the posterior? Does it appear that the prior is truly vague?

2. Using the `cookies` data, a Poisson likelihood, and a Gamma(9,1) prior for the mean, use Metropolis-Hastings to create a sample from the posterior for the mean number of chips per cookie. Try to have at least 5000 samples after burn-in. (Hint: to avoid numerical problems when computing the probability of accepting a proposal, group terms before evaluating exponentiation, i.e., you should have a lambda.star to the something in the numerator, and a lambda[i] to the same thing in the denominator, so evaluate this as (lambda.star/lambda[i]) to that power; similarly, put everything into a single exp.)

   (a) Compute your overall acceptance probability, and adjust your proposal distribution until your acceptance probability is between 0.2 and 0.5.

   (b) Show your trace plot. If you have burn-in, include plots both with and with-out the burn-in.

   (c) Compare the posterior mean and variance from your MCMC sample to the theoretical values.

Optional Problem

3. In class, Gibbs sampling for normal data with unknown mean and variance was presented, but the complete conditional for τ required the evaluation of $\sum_i (y_i - \mu_{t+1})^2$ at each iteration, which could be a lot of work for a larger dataset. Re-write the complete conditional for τ so that it only depends on hyperparameters and the sufficient statistics $\bar{y}$ and $\sum_i (y_i - \bar{y})^2$, which can be computed once before starting MCMC, rather than re-computed at each iteration. Run a Gibbs sampler for the alcohol content in the `beer` data using this complete conditional.