Sta 601/360: Lab 9

Today the EPA has another dataset for you! This time, unfortunately, their technician was lazy and didn't take measurements every day–and he also forgot to record the day each measurement was taken. The first few rows of data.txt (on Sakai) look like this:

```
         pm
1 4.8885880
2 0.2031895
3 1.5695845
4 0.4603080
5 1.9514177
6 3.8712842
```

1. Write down a model similar to the one you fit last week, except this time the indicator for whether a given measurement is from a weekday or weekend is also a random variable which you need to do inference on. Assume, *a priori*, that weekdays and weekends are approximately proportionately represented in the sample–for example, use something like a Beta(5,2) prior on the probability that a given measurement is from a weekday.

   As in the previous lab, the pollution concentration is Log-Normal with parameters (μ_1, σ_1^2) for weekdays, and (μ_2, σ_2^2) for weekends, and the prior should have probability 1 that $\mu_1 > \mu_2$.

2. Perform MCMC to obtain a minimum of 10,000 post-burnin draws. Provide at least one trace plot showing that the sampler has converged.

3. Provide a plot of the sampled posterior values of μ_1 and μ_2 against each other, μ_1 on the X-axis and μ_2 on the Y-axis.

4. Provide a point estimate and 95% posterior credible interval on the number of days in the sample which are weekdays. What is the probability that the technician is coming in less often on weekends than on weekdays? (i.e., what is the probability that the proportion of weekend days is less than 2/7?)

5. Provide point estimates and 95% posterior credible intervals for each of the four parameters. How do these compare to your estimates and intervals from the previous lab?