Sta 601/360: Lab 7

The Log-normal distribution is often used to model positive right-skewed random variables, with pdf given by

$$f_X(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}}\, e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}},\ x>0.$$

The log-normal distribution has mean

$$e^{\mu+\sigma^2/2}$$

and variance

$$\left(e^{\sigma^2}-1\right)e^{2\mu+\sigma^2}.$$

One example of such a variable is air pollution measurements, which are taken in micrograms per cubic meter. Pollution levels must be positive, and tend to be right-skewed with a heavy right tail.

The file data.txt contains 100 measurements of fine particulate air pollution from the St. Louis metropolitan area, courtesy of the EPA.

1. Construct a model using a lognormal likelihood for the data, with weak priors of your choice for the two parameters μ and σ^2.

2. Perform Metropolis Hastings (either within a Gibbs sampler or with bivariate proposals, up to you) to estimate the two parameters. Provide trace plots as evidence that your sampler has converged and explored the space reasonably well.

3. Provide 95% confidence intervals for the mean and variance of the pollution levels (remember that μ is **not** the mean and σ^2 is **not** the variance!).

There is an easier way to do this particular analysis (can you see what it is?), but I'm making you do it the hard way because we're going to build on this in next week's lab, where the easy way won't work.