COGSCI 131 – Assignment 9
DUE: April 16th at class start

In this assignment, you will implement the Metropolis algorithm to sample from a posterior distribution on a psychological parameter, based on data on bcourses.

Assume that a Weber ratio W has prior,

$$P(W) = e^{-W}$$

A likelihood function has been provided that compute the probability of answering correctly (the behavioral data) given W and the behavioral stimuli (n_1 and n_2).

Recall that in the Metropolis-Hastings algorithm, you start with an initial W, you propose a change to a new value W' (e.g. by adding a sample from a Normal(0,.1) distribution to W) and you accept this change always if it is better, and otherwise with probability P(W'|D)/P(W|D). NOTE: Please pay attention to which terms do and do not have logs in this assignment.

1. [5pts, HELP] We will implement functions that compute the logarithm of the prior and the logarithm posterior, as opposed to the unlogged versions. Explain why in 1-2 sentences.

2. [5pts, HELP] Explain in a 1-2 sentences why the form of the acceptance ratio P(W'|D)/P(W|D) is convenient for us – that is, why it saves us from doing an integral.

3. [5pts, HELP] Write functions to compute something proportional to the log prior and log posterior in this model. You will prevent later frustration by ensuring that your prior correctly handles cases when W<0 (what should it return?).

4. [25pts, HELP] Implement the Metropolis algorithm, starting from a random W, and plot (a) the posterior score of W over the first 300 samples; (b) the value of W over the first 300 samples, and (c) a histogram of the samples of W over the first 10,000 samples after 1000 samples of "burn in."

5. [15pts, SOLO] Use your sampler to determine the probability that W is in the interval [0.60, 0.65]. What is that probability?

6. [15pts, SOLO] Run your sampler on the prior (assuming no data) and plot histograms of the prior samples and the posterior samples (from Q4c) in the same graph. What relationship between these distributions does your plot show, and what does it mean?

7. [25pts, SOLO] **Extra credit**: Do you get "better" estimates with this algorithm if you run one chain for 1000 steps or 100 chains for 10 steps each and just concatenate their samples? To answer this, you should come up with both (a) a simulation that tries both of these options (you'll want to run more than one time!) and (b) a measure of how "good" a run was (i.e. how accurate its samples were). Write a sentence or two justifying your measure of goodness, create a visualization of the two options, andexplain in a sentence how the visualization answers the question.