COGSCI 131 – Assignment 5
DUE: July 29th at class start

In this assignment, you will implement the Number Game from lecture. In this consider the following Hypotheses all on numbers from 1-100:

- H1: even numbers
- H2: odd numbers
- H3: square numbers
- H4: prime numbers
- H5: multiples of 5
- H6: multiples of 10

A convenient way to represent these is to compute a Python list for each hypothesis. And we will assume equal priors on each.

1 [5pts]: Write a function that takes an argument x, a hypothesis (however you represent it) and computes a <u>size principle likelihood</u> (e.g. where the likelihood of each number in the set is equal). Write down what likelihood each hypothesis assigns to each data point in it. What does each hypothesis assign to data points not in it?

2 [20pts]: Make a plot showing the posterior predictive probability (marginalizing over hypotheses) that each number 1...100 is "in" the concept, for each of the following data sets:

- (a) No data
- (b) 50
- (c) 53
- (d) 50, 53
- (e) 16
- (f) 10, 20
- (g) 2, 4, 8
- (h) 2, 4, 8, 10

Assume that there are equal priors on each hypothesis and the size principle likelihood from Q1. Be sure that you structure your code to process the data as a list and do not "hard code" in these datasets (the likelihood function and posterior function should accept any list of data). Write a sentence for each plot about whether the model does or does not capture your intuitions about the "right" answer.

3 [20pts]: Re-make the plots from Q2 but now incorporate range-based hypotheses. To do this, assume that H1-H6 each have a prior of 1/7, and the remaining 1/7th probability is distributed equally among all intervals in the range 1-100 (There are a lot of these range-based hypotheses--how many are there?). Write a sentence for each plot about whether the including the range-based hypotheses makes them better match your own intuitions about how to generalize and why.