PS #1
Education 252L
Due 1-25 at noon

When writing up your solutions, aim to write a brief narrative with respect to what you are seeing rather than simply cutting and pasting results. To re-emphasize: brevity is fine but messiness less so.

All source available here: https://github.com/ben-domingue/252L_winter2018

0. *R exercises:* Having you develop some proficiency in R over the first week is an important part of this problem set. I am not going to ask you to write a ton of your own code in this course. You will largely be interacting with my code. You'll get a feel for what that is like below. But, some knowledge of the basics will be helpful. Please go through these R exercises on data examination and working with vectors. I'm not going to grade these exercises (you don't need to write them up), but I am going to assume that you know how to do such things unless you tell me otherwise.

1. Exploration of collections of Bernoulli random variables.
   Get the code from here:
   https://github.com/ben-domingue/252L_winter2018/blob/master/ps1/ps1-bern.R
   [To do this, you can click raw and then save the resulting plain text from your browser.]
   I've embedded a number of questions in that source. Look for #Q.

2. Get data from:
   https://github.com/ben-domingue/252L_winter2018/blob/master/data/ps1-logreg.Rdata
   [To download this file, go to that page and click "Download". Once downloaded, you need to load into R. With .Rdata files, you just do something like:
   ```
   load("ps1-logreg.Rdata")
   ```
   to get them into R, assuming the R working directory is the one where the file resides.]
   There are three variables: x, y1, and y2 in this data.frame (df). The latter two are dichotomous indicators that are associated with x. Consider logistic regressions of y1 and y2 on x. Once you load() the data, you can do that via:
   ```
   glm(y1~x,df,family="binomial")->m1
   glm(y2~x,df,family="binomial")->m2
   ```
   Questions: (A) How would you compare the association between y1/y2 & x? (B) How would you interpret the regression coefficients from (say) m1? (C) Do m1 and m2 show equivalent model fit? Can you notice anything peculiar about either y1 or y2 (in terms of their association with x)?

3. Likelihood problem

The likelihood will be an object of keen interest all quarter. To get you back to speed in terms of thinking about likelihood, I've crafted this brief exercise.
https://github.com/ben-domingue/252L_winter2018/blob/master/ps1/ps1-like.R
I anticipate that this will be terra incognita for some of you. That's ok! A great opportunity to schedule a time to come by my office so that we can walk through it together.

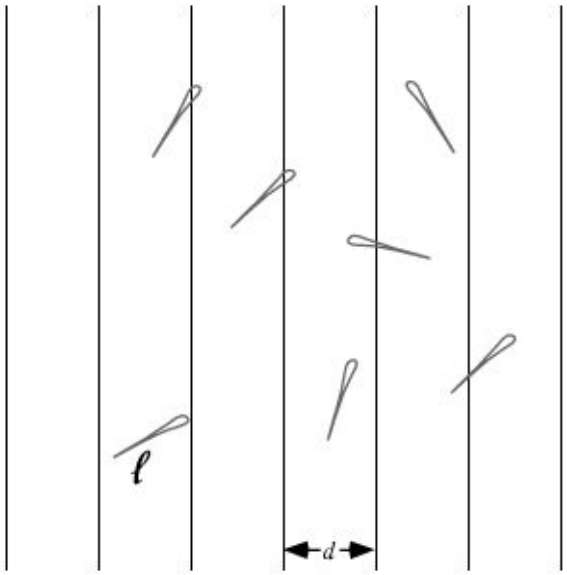
4. Item Quality
   Download item response data from here:
   https://github.com/ben-domingue/252L_winter2018/blob/master/data/emp-rasch.txt
   Consider the item statistics (p-values & item-total correlations) discussed in the Crocker & Algina text. What do you think? As a point of contrast, consider them vis-a-vis the item statistics generated by this data:
   https://github.com/ben-domingue/252L_winter2018/blob/master/data/rasch.txt

Trickier coding problem, BONUS

Buffon's needle is a classic probability problem. I'm using a simple statement of the problem from here (you can also find versions of the answer here):

Buffon's Needle Problem

DOWNLOAD
Wolfram Notebook



Buffon's needle problem asks to find the probability that a needle of length ℓ will land on a line, given a floor with equally spaced parallel lines a distance d apart. The problem was first posed by the French naturalist Buffon in 1733 (Buffon 1733, pp. 43-45), and reproduced with solution by Buffon in 1777 (Buffon 1777, pp. 100-104).

There are analytic solutions to the problem (those are fun to work out too!), but I want to challenge you to discover the answer via simulation. We'll use simulation frequently this fall and will make for a nice introduction. I won't specifically grade this exercise, but I encourage you to try it. [There are also lots of online resources on this problem.]