L&S 39F: Data science and the mind (Fall 2005)

Homework 2

October 6 (Due *beginning* of session on Tue, Oct 20)

*This instruction sheet is ideally accessed electronically - the cyan boxes below provide clickable links. All coding should be completed in a Python Jupyter Notebook (or Python). Please type your solutions into a *single* PDF that includes both figures and written explanations, with your own Python code enclosed in Appendix. Email this PDF to the course instructor with the title "L&S39F HW2 (YOUR NAME)". You should turn in your individual solutions, although collaborative discussions are encouraged (acknowledge your collaborator(s) in the write-up). Late work receives a deduction of 1 point per delayed hour.*

In this assignment, you will model human behavior in a cognitive arithmetic task. The purpose is to illustrate how the approximate number sense may give rise to behavior. Specifically, you will be simulating the comparison task from work by Pica et al. (2004) in the session reading materials. To do so, you will need to 1) download *Code* and *Data* under *Hw2* in the NB column from the bcourses syllabus; 2) copy and paste the content of "hw2_starter_code.txt" into a new Python notebook.

Background — Running the starter code, you will have specified 3 values of weber fraction that roughly correspond to the levels of numerosity in three populations: children (wf_1), Mundurukú speakers (wf_2), and adults in industrialized societies (wf_3). Your job is to predict behavior, particularly the accuracy at which these three populations can perform the following comparison task under two conditions. In the first simulated condition, n_1 and n_2 represent the numbers of dots displayed on the left and right sides of a screen respectively, and a hypothetical participant for each population is asked to judge whether the two sides contain an equal number of dots. In this case, n_1 is varied from 1 to 100 incrementally with a step of 1, and $n_2 - n_1 = 1$ (i.e. $n_1 = 1, n_2 = 2$ for trial 1, $n_1 = 2, n_2 = 3$ for trial 2,...). In the second condition, n_1 and m_2 play similar roles except that $m_2 - n_1 = 10$. You will need to specify these 3 variables for your simulations: n_1, n_2, m_2.

Model — Pages $S5-6$ in the supplementary material of Pica et al. (2004) suggests one way to model human behavior in this comparison task. Read these pages - you will need their equation for $p_{comparison}$ for answering the questions. Note *erfc* is a function that would estimate the probability of error for comparing two numbers - it has the property that its output value bears an inverse relationship with its input argument. This function has been imported from *scipy* in the starter code provided, so please use it directly.

Q1. Generate Figure 1 (entitled "$n_2 - n_1 = 1$") that plots probability of error against n_1 for the first condition. Your figure should include three colored curves [1 point], one for each population in question; it should have a x-axis that runs from 1 to 100, representing the numerical magnitude of n_1; it should also have a y-axis limit that is made *identical* to that in Q2 [1 point]. Place legends in the bottom right corner that specify the colors you used for the three populations in question [1 point].

Q2. Generate Figure 2 (entitled "$m_2 - n_1 = 10$") similarly but for the second condition. [2 point]

Q3. Describe how weber fraction affects error probabilities from what you have observed. [1 point] Explain why error probabilities should increase as n_1 increases in value under both conditions. Your answer should relate explicitly to $p_{comparison}$ and the property of *erfc*. [1 point] Explain why the terminal ($n_1 \to \infty$) error probability for the three curves (e.g. in Figure 1) should approach 0.5. [1 point]

Q4. Plot the average error rates from the *Data* file you have donloaded for corresponding values of n_1 on Figures 1 (condition 1) [1 point] and 2 (condition 2) [1 point]. Does the model account well for your data?