Review 2

1. **What to Buy?**

   You are deciding which of two books to purchase.

   Book 1 has the ratings:

Rating	1	2	3	4	5
Count	0	0	0	1	4

   Book 2 has the ratings:

Rating	1	2	3	4	5
Count	0	1	2	20	200

   Assume your rating of each book is drawn from the same distribution as the folks who rated the book in the past. What is the probability you would prefer book 2 to book 1?

2. **Information Gain in Texas Holdem**

   Texas Holdem is played with a standard 52-card deck. You start with two private hole cards, and throughout the hand five community cards are revealed in three stages:

   - Pre-Flop: no community cards have been shown

   - Flop: the first three community cards are shown

   - Turn: the fourth community card is shown

   - River: the fifth community card is shown

   By the end of the hand, the seven total cards (your two hole cards plus the five community cards) determine your final hand. Treat the set of possible final hands as a random variable.

   At each stage: Pre-Flop, Flop, Turn and River, compute the entropy of the distribution over all possible final hands.

3. **Quant Interview Cup-Cakes**

   A box has 100 cupcakes. Either 51 have blue frosting and the rest have pink frosting, or 49 have blue and the rest have pink. Both scenarios are equally likely.

   a. You look at one cupcake. It is blue. What is the probability that the majority were blue?

   b. You look at three cupcakes. Two are blue. What is the probability that the majority were blue? Note: each time you are given a cupcake you put it back before the next is selected.

c. A hypergeometric is a random variable for the number of successes if you remove items **without** replacement from a fixed population. If $X \sim \text{Hypergeom}(t, k, n)$, the PMF is

$$P(X = x) = \frac{\binom{k}{x}\binom{t-k}{n-x}}{\binom{t}{n}}$$

Where:

- t is the population size,
- k is the number of "success" items in the population,
- n is the number of draws (without replacement),
- X counts the number of successes observed in those n draws.

Three cupcakes are drown without replacement and two are blue. What is the probability that the majority were blue?

4. **What word did they say?**

A child was meant to read a word "tiger". We have an audio recording of them saying the word. How can you build a principled estimate for if they got the question correct?

5. **Medical Summary?**

We have medical records from 50 patients. We have 4 candidate summaries for patient 1. Which of the candidate summaries is the best?