STAT 237: HW1

DUE ELECTRONICALLY VIA THE RSTUDIO SERVER FRIDAY 03/04/22 BY 5PM

Imagine you are a doctor attempting to diagnose a patient who is presenting with a fever (for simplicity, let's just treat fever/no fever as a binary for this example). One of the hypotheses you're considering is that they might have the flu.

1. (Difficulty: *) Which of the expressions below represents the probability that this person has the flu given what we know about their symptoms so far?
   (a) $P(\text{flu})$
   (b) $P(\text{flu} \mid \text{fever})$
   (c) $P(\text{fever} \mid \text{flu})$
   (d) $P(\text{fever}, \text{flu})$
2. (Difficulty: **) Which of the following is best describes what is represented by the notation $P(\text{flu})$?
   (a) The probability that a person without a fever has the flu
   (b) The midpoint between the probability that a person with the flu has a fever and the probability that a person without the flu has a fever
   (c) The probability that a person has the flu if we didn't know whether they have a fever
   (d) The probability that a person has the flu once we know whether or not they have a fever
3. (Difficulty: ***) Explain the difference in meaning between $P(\text{flu}, \text{fever})$, $P(\text{fever} \mid \text{flu})$ and $P(\text{flu} \mid \text{fever})$. Which one is the smallest? How do you know?

Date: Last Revised February 27, 2022.

4. Suppose you have a deck with only three cards. Each card has two sides, and each side is either red or blue. One card has two red sides, one has two blue sides, and one has one red side and one blue side. All three cards are shuffled and placed in an envelope. Someone reaches into the envelope, pulls out a card, and places it flat on a table.
   (i) (Difficulty: ***) Sketch the "tree world" where the hypotheses correspond to the card chosen and the data corresponds to the side facing up.
   (ii) (Difficulty: **) Suppose you see a blue side facing up, and have no other information. How many paths are consistent with this observation?
   (iii) (Difficulty: ***) How likely is it that the other side of the card is also blue? Explain your reasoning.
5. Suppose scientists have discovered a gene that has two alleles known to exist in humans in roughly equal proportions. There are no known demographic differences in the relative frequency of these alleles, and the only known difference in phenotype is that people who give birth are slightly more likely to have fraternal twins if they have two copies of allele A (about 1 in 40 births) than if they have two copies of allele B (about 1 in 60 births), or if they have one copy of each allele (about 1 in 50 births).
   (i) (Difficulty: ***) Alex gives birth to fraternal twins. They had no prior pregnancies. Sketch the "tree world" for this scenario. What is the posterior probability that Alex has two copies of allele A?
   (ii) (Difficulty: ****) If Alex gives birth a second time, how likely are they to have another set of fraternal twins? (Assume for the sake of this calculation that the chance of fraternal twins depends only on this one gene)
   (iii) (Difficulty: ****) Suppose Alex does give birth a second time, this time to a single infant. Extend your tree world accordingly. What is the posterior probability now that they have two copies of allele A?
   (iv) (Difficulty: *****) Suppose Alex themself is a fraternal twin. How does this change the probabilities you calculated above? Assume that a person is equally likely to inherit either of their biological parents' copies of this gene, and that the genome of sperm is irrelevant to whether conception results in fraternal twins. (Hint: I suggest extending your trees "backwards" so that the root corresponds to the genome of Alex's biological mother)