# STATS 116: Homework 2

## Due: Thursday, July 13, 2023 at 10:00 pm PDT on Gradescope

There are 7 problems on this assignment, each worth 8 points, although subparts within a problem may not be equally weighted. Credit will be assigned primarily based on reasoning and work, not the final answer. You do not need to simplify arithmetic expressions unless otherwise noted. While you may discuss the problems on this assignment other students, you must write up your own solutions. As per the syllabus, you may occasionally use the Internet or other public resources to clarify concepts with citation when this information is used as part of your own solution to a homework problem. However, you may not search for direct solutions to any problems assigned for homework or exams. For example, you can ask ChatGPT to clarify a particular concept from lecture that may be related to a problem, but you cannot feed it any part of a course assignment or a substantively similar version.

1. Jane Villanueva is taking a pregnancy test because she was artificially inseminated. Suppose 20 percent of artificially inseminated women get pregnant, although it is known that 25 percent of artificially inseminated women test positive on this pregnancy test. The false positive rate of the test among artificially inseminated women is 10 percent (i.e. 10 percent of non-pregnant artificially inseminated women test positive on the test).

   **(a)** Compute the false negative rate of the test: that is, the probability that an artificially inseminated pregnant woman tests negative.

   **(b)** What is the probability that Jane is pregnant, given that she takes the test twice and gets a positive result both times? Assume the two test results are conditionally independent given Jane's pregnancy status.

2. Harrison flips a fair coin repeatedly, recording the sequences of flips he observes.

   **(a)** What is the probability he observes the sequence HH for the first time before observing the sequence HT for the first time?

   **(b)** What is the probability he observes the sequence HHT for the first time before observing the sequence HTH?

3. (BH 2.10) Fred is working on a major project. In planning the project, two milestones are set up, with dates by which they should be accomplished. This serves as a way to track Fred's progress. Let $A_1$ be the event that Fred completes the first milestone on time, $A_2$ be the event that he completes the second milestone on time, and $A_3$ be the event that he completes the project on time. Suppose

that $P(A_{j+1} \mid A_j) = 0.8$ but $P(A_{j+1} \mid A_j^c) = 0.3$ for $j = 1, 2$, since if Fred falls behind on his schedule it will be hard for him to get caught up. Also, assume that the second milestone supersedes the first, in the sense that once we know whether he is on time in completing the second milestone, it no longer matters what happened with the first milestone. We can express this by saying that $A_1$ and $A_3$ are conditionally independent given $A_2$ and they're also conditionally independent given $A_2^c$.

**(a)** Find the probability that Fred will finish the project on time, given that he completes the first milestone on time. Also find the probability that Fred will finish the project on time, given that he is late for the first milestone.

**(b)** Suppose that $P(A_1) = 0.75$. Find the probability that Fred will finish the project on time.

4. (BH 2.36) Suppose that in the population of college applicants, being good at baseball is independent of having a good math score on a certain standardized test (with respect to some measure of "good"). A certain college has a simple admissions procedure: admit an applicant if and only if the applicant is good at baseball or has a good math score on the test.

**(a)** Give an intuitive explanation of why it makes sense that among students that the college admits, having a good math score is negatively associated with being good at baseball, i.e., conditioning on having a good math score decreases the chance of being good at baseball.

**(b)** Show that if $A$ and $B$ are independent and $C = A \cup B$, then $A$ and $B$ are conditionally dependent given $C$ (as long as $P(A \cap B) > 0$ and $P(A \cup B) < 1$), with

$$P(A \mid B, C) < P(A \mid C)$$

This phenomenon is known as Berkson's paradox, especially in the context of admissions to a school, hospital, etc.

5. (BH 2.40) Consider the Monty Hall problem, except that Monty enjoys opening door 2 more than he enjoys opening door 3, and if he has a choice between opening these two doors, he opens door 2 with probability $p$, where $0.5 \leq p \leq 1$. To recap: there are three doors, behind one of which there is a car (which you want), and behind the other two of which there are goats (which you don't want). Initially, all possibilities are equally likely for where the car is. You choose a door, which for concreteness we assume is door 1. Monty Hall then opens a door to reveal a goat, and offers you the option of switching. Assume that Monty Hall knows which door has the car, will always open a goat door and offer the option of switching, and as above assume that if Monty Hall has a choice between opening door 2 and door 3, he chooses door 2 with probability $p$ (with $0.5 \leq p \leq 1$).

**(a)** Find the unconditional probability that the strategy of always switching succeeds (unconditional in the sense that we do not condition on which of doors 2 or 3 Monty opens).

**(b)** Find the probability that the strategy of always switching succeeds, given that Monty opens door 2.

**(c)** Find the probability that the strategy of always switching succeeds, given that Monty opens door 3.

6. (BH 3.1) People are arriving at a party one at a time. While waiting for more people to arrive they entertain themselves by comparing their birthdays. Let $X$ be the number of people needed to obtain a birthday match, i.e. before person $X$ arrives no two people have the same birthday, but when person $X$ arrives there is a match. Find the PMF of $X$.

7. Suppose $X$ and $Y$ are discrete random variables with finite support that have the same distribution.

   (a) Is it possible that $P(X = Y) = 0$? Give an example, or show that it's not possible.

   (b) Is it possible that $P(X > Y) = 1$? Give an example, or show that it's not possible.