Exercises for Causal Thinking (Math-352)

September 25, 2023

Exercise Sheet 1

Exercise 1. Determine whether each of these questions are phrased as causal questions or not (yes or no).

1. Does the Moderna vaccine reduce the risk of severe COVID-19 infection?

2. Do women with breast cancer survive longer than men with prostate cancer?

3. Is the life expectancy in Switzerland longer than the life expectancy in Italy?

4. Does drinking 0.5 L beer compared to 0.5 L Coca Cola at 19h00 affect the quality of sleep?

5. Would drinking a cup of coffee 2 hours before your exam improve your performance?

Exercise 2. Suppose you were asked by an administrator at EPFL to evaluate whether drinking 2 espressos versus no coffee before the exam of MATH-352 Causal Thinking gives better exam results among the students. How would you design a study to investigate this question?

Exercise 3. What is wrong with the following claims?

(a) "Data show that income and marriage have a high positive correlation. Therefore, your earnings will increase if you get married."

(b) "Data show that as the number of fires increases, so does the number of fire fighters. Therefore, to cut down on fires, you should reduce the number of fire fighters."

(c) "Data show that people who hurry tend to be late to their meetings. Don't hurry, or you'll be late."

Exercise 4 (Baseball). A baseball batter Tim has a better batting average than his teammate Frank. However, someone notices that Frank has a better batting average than Tim against both right-handed and left-handed pitchers. How can this happen? (Present your answer in a table.)

Exercise 5 (Warm up with the Monty Hall problem)**.** In a game show, contestants must choose to open one of three doors. A prize is hidden behind one of the doors, but only the game show host knows which door leads to the prize. After the contestant makes her choice, the game show host opens one of the two doors that does not contain the prize. The contestant is now given the choice of opening her initially chosen door versus the other remaining door. Which door should she choose?

Exercise 6. Determine, for each of the following causal stories, whether you should use the aggregate or the segregated data to determine the true effect.

(a) There are two treatments used on kidney stones: Treatment A and Treatment B. Doctors are more likely to use Treatment A on large (and therefore, more severe) stones and more likely to use Treatment B on small stones. Should a patient who doesn't know the size of his or her stone examine the general population data, or the stone size-specific data when determining which treatment will be more effective?

(b) There are two doctors in a small town. Each has performed 100 surgeries in his career, which are of two types: one very difficult surgery and one very easy surgery. The first doctor performs the easy surgery much more often than the difficult surgery and the second doctor performs the difficult surgery more often than the easy surgery. You need surgery, but you do not know whether your case is easy or difficult. Should you consult the success rate of each doctor over all cases, or should you consult their success rates for the easy and difficult cases separately, to maximize the chance of a successful surgery?