HW 02: More Bayes, Data, and Sampling

> ❗ Due date
>
> This assignment is due on **Wednesday, September 10 before class (hard deadline so I can post the solution in time for you to study for midterm)**. To be considered on time, the following must be done by the due date:
>
> - Final `.pdf` file submitted on Gradescope

Learning goals

In this assignment, you will...

- practice the Bayes' theorem again
- practice determining types of data
- relates random sampling to probability concept

Conceptual exercises

Instructions

The conceptual exercises focus on explaining concepts and showing results mathematically. Show your work for each question. I expect you not to skip any steps and carefully understand the thinking process when working on these exercises.

You may write the answers and associated work for conceptual exercises by hand or type them in your Quarto document.

Exercise 1: The Game Show Host or "Monty Hall" Problem (14 points)

Watch a clip from a movie titled 21 (2008) in this Youtube video. The same problem is also posted in *Parade Magazine* in 1990 :

"Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, 'Do you want to pick door No. 2?' Is it to your advantage to switch your choice?"

Answer these question to help you make the decision:

1. Before the host open door no. 3:
    - what is the probability that the car is in door No. 1?
    - what is the probability that the car is in door No. 2?
    - what is the probability that the car is in door No. 3?
2. May be it doesn't seems too obvious, but the host will never open a door concealing the car. Now, compute the following conditional probability:
    - Probability that the host open door No. 3 given that the car is in door No.1
    - Probability that the host open door No. 3 given that the car is in door No.2
    - Probability that the host open door No. 3 given that the car is in door No.3
3. Using the law of total probability, can you compute the probability that the host open door No. 3?
4. Compute the following posterior probability using Bayes' theorem:
    - Probabiliy that the car is in door No.1 given that the host open door No. 3
    - Probabiliy that the car is in door No.2 given that the host open door No. 3
    - Probabiliy that the car is in door No.3 given that the host open door No. 3
5. Compare the prior probability (in point 1) and the posterior probability (in point 4), is it better to keep door No. 1 or switch to door No. 2?

Exercise 2: Data and Sampling (6 + 9 points)

1. For each of the following examples, indicate whether the data are cross-sectional, time-series, or panel data.
    a. An avid runner records the number of miles that she runs every day for 100 straight days.
    b. A random sample of 100 older adults, aged 65 and older, were surveyed in January 2023 about their health status and medical expenditures.
    c. A financial analyst randomly picks 20 companies that are listed on the New York Stock Exchange and gathers data on their profits and sales for the year 2022.

d. A Canadian economist gathers annual data on each of Canada's 13 administrative divisions, including the population and the unemployment rate for each division, for each year between 2000 and 2020.
e. A random sample of 100 college students is asked whether they have received an influenza vaccine in the last year.
f. A random sample of 100 college seniors is asked for their semester grade point average for each of their first six semesters at the university.

2. Consider drawing a simple random sample of $n = 10$ observations from a population consisting of 100 units.

   a. How many possible ways are there to draw the simple random sample?
   b. What is the probability that a given observation from the population is in the simple random sample that is drawn?
   c. What is the probability that any two given observations from the population are in the simple random sample that is drawn?

Grading (30 points)

Component	Points
Ex 1	14
Ex 2	15
Formatting	1

"Formatting" grade is to assess whether or not you mark specific page of the pdf to a specific question to help the TA when grading.