Vectors in R: Analyzing the Monty Hall Problem

Preparation

The Monty Hall Problem is a famous probability puzzle that challenges our intuition about statistics and decision-making. Named after Monty Hall, the host of the American television game show "Let's Make a Deal," this problem presents a scenario where a contestant must choose between three doors. Behind one door is a prize, while the other two conceal goats. After the contestant makes an initial choice, the host, who knows what's behind the doors, opens another door to reveal a goat. The contestant is then given the option to stick with their original choice or switch to the remaining unopened door. The question is: which is the best strategy?

Before we dive into working with R, let's gain an illustrative understanding of the problem. Watch the video below from the 1994 German television show "Geh aufs Ganze," where a contestant faces this very dilemma.

- Video "Geh aufs Ganze" (1994)

To learn about vectors in R, watch the lecture on representing data from Harvard's CS50 Introduction to Programming with R, as well as the short video on vectors:

- Video CS50R - Representing Data
- Video CS50R - Short on Vectors

Now you should be prepared to solve the following exercises using R and RStudio.

Exercise 1: Play and record

To get an even better understanding of the problem, let's first play the game a couple of times and record our observations using R. Will then look at the results using R and vectors.

1. Play "Let's Make a Deal" several times with a partner. Use cups or similar containers to simulate the doors. One person should act as the game show host, the other as the contestant. For each game, record where the prize was hidden, which cup the contestant chose, and whether the contestant decided to switch or stay.

2. Transform your game observations into vectors in R. Consider what data representation would be most suitable using only vectors.

3. Using R, determine how many times the contestant initially selected the correct door.

4. Using R, calculate the contestant's overall winning percentage.

5. Using R, calculate the percentage of times the contestant stayed with their initial cup choice.

6. Calculate the winning percentage for instances when the contestant switched versus when they stayed with their initial choice.

7. What insights can you draw from this data and analysis regarding the optimal strategy? How might you deepen your understanding of the Monty Hall problem?

Exercise 2: Historical data

To gain a more comprehensive understanding, let's examine a dataset you've obtained from the famous TV show "Let's Make a Deal."

1. Load the data from their respective files into vectors in R.

2. Perform the same calculations as before:

   a. Calculate the winning percentage for cases where the contestant stayed

   b. Calculate the winning percentage for cases where the contestant switched

   c. Determine the percentage of staying versus switching decisions