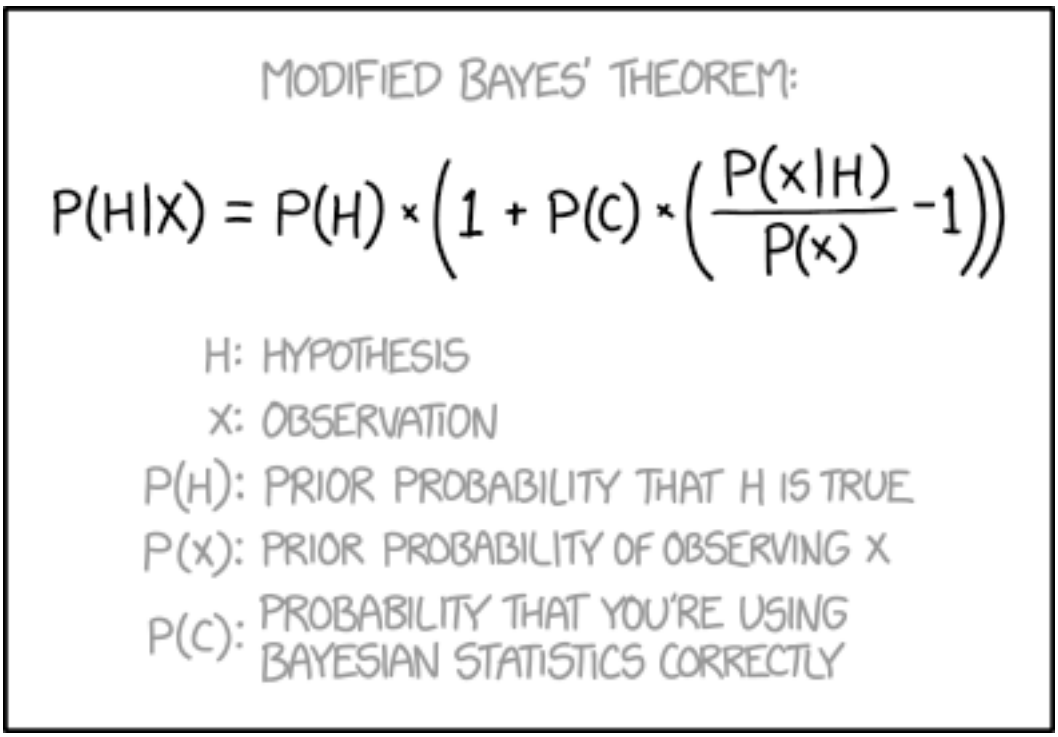
Day 8 - Probability II

Introduction

The purpose of this activity is to continue to familiarize ourselves with probability, introducing the concept of **conditional probability** through two delightful "paradoxes" in statistics.



Funny Bayes' joke

A question about coins

Two statisticians, Christian and Alex, are sitting in a coffee shop discussing probability. Alex pulls out two fair coins, flips them, and observes the result without showing Christian. Alex then tells Christian that at least one of the coins landed heads up, and asks Christian,

"What is the probability that the other coin is also heads up?"

Christian, a master probabilist, responds 1/2.

1. Do you agree with Christian? Explain your reasoning.

   > Any answer is okay! The correct answer is that Christian is an idiot, though.

2. Suppose we wanted to *estimate* the probability in question. How could we do that?

   > We need to replicate this process a bunch of times! Flip two coins, restrict to outcomes where at least one lands heads up, and see how often (in the long run) the other one also lands heads up.

3. In your groups, flip two coins and record the results. **Note that the order of the coins matters!** You should identify one coin as "coin1" and the other as "coin2", and track their results separately. That is, "HT" should be recorded separately from "TH."

   Repeat this process until you have each flipped a pair of coins ten times. Count up the total number of times you observed each combination of coin flips (HH, HT, TH, and TT) and use that to fill in the table below.

   > **!** Important
   >
   > Ensure that after you have filled in the code below, you set the code chunk option `eval = F` to `eval = T` so that the table prints when render!

```
library(tidyverse)
group_flips <- tibble(
  coin1 = rep(c("H", "T"), 2),
  coin = c(rep("H", each = 2), rep("T", each = 2)),
  count = c(#, #, #, #) # order: HH, TH, HT, TT
)
group_flips
```

4. How could you use the above table to *estimate* the probability of interest?

> **Tip**
>
> What rows of the table represent possible events if at least one of the coins is known to have landed heads up? What rows of the table represent the case where the other (unknown) coin landed heads up?

> Remove the TT row, and divide HH by the sum of HH, HT, TH. (TT is the only outcome we *know* cannot occur based on the information we are given)

5. Use the table from question 3 to estimate the probability of interest. Show your work.

> This will depend specifically on your results!

6. Does the estimated probability match what you expected? Explain.

> I hope so! It matches mine. :)

7. Does approximately 40 simulations feel like enough to get a reliable estimate of the probability of interest? Explain.

> No!

8. The R code below simulates the coin flipping experiment 1,000,000 times, and calculates the probability of interest.

```
library(tidyverse)

# simulated flipping two coins 1e6 times
flips <- sample(c("H", "T"), size = 1e6*2, replace = T)

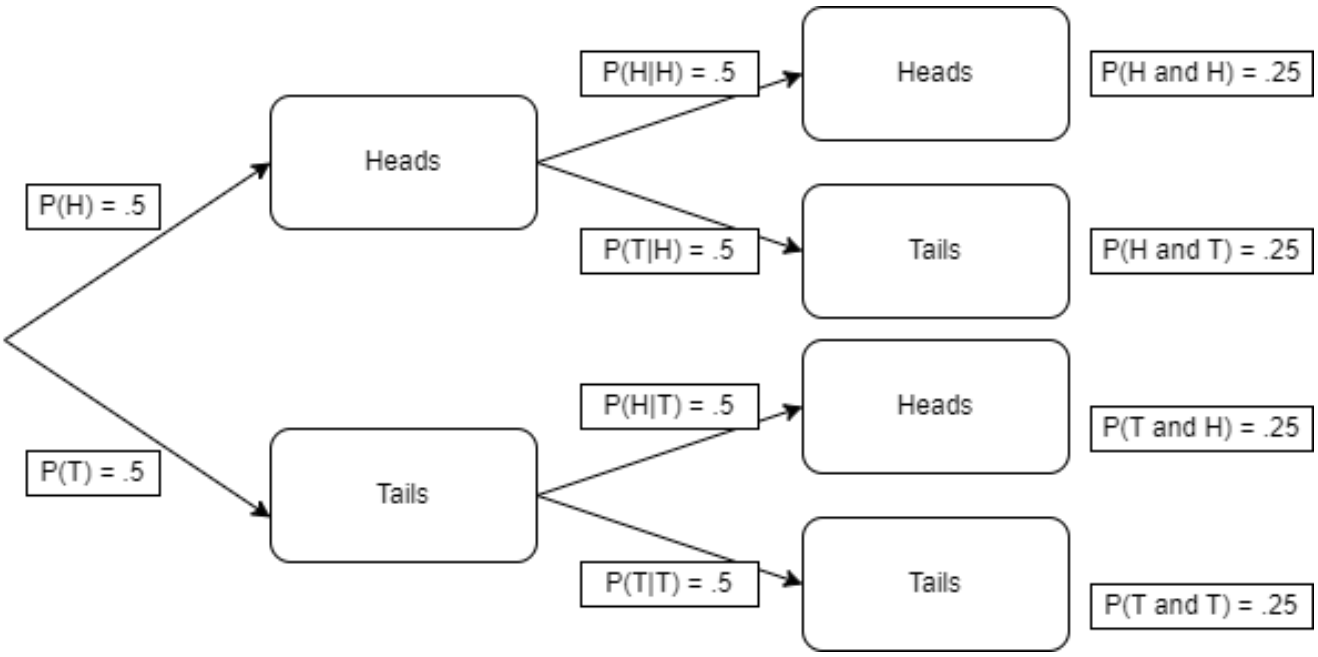
# summarize flips
as_tibble(matrix(flips, ncol = 2, byrow = T)) %>%
  rename(coin1 = V1, coin2 = V2) %>%
  # filter to rows with either coin1 = H, coin2 = H, or both
  filter(coin1 == "H" | coin2 == "H") %>%
  # what proportion were both heads?
  summarize(
    heads = mean(coin2 == "H" & coin1 == "H")
  )
```

```
# A tibble: 1 x 1
  heads
  <dbl>
1 0.334
```

Are you convinced based on 1,000,000 simulations that the estimated probability is reliable?

> Yeah that seems reasonable.

9. Use the tree diagram below to describe how the probability could be obtained without using simulation.

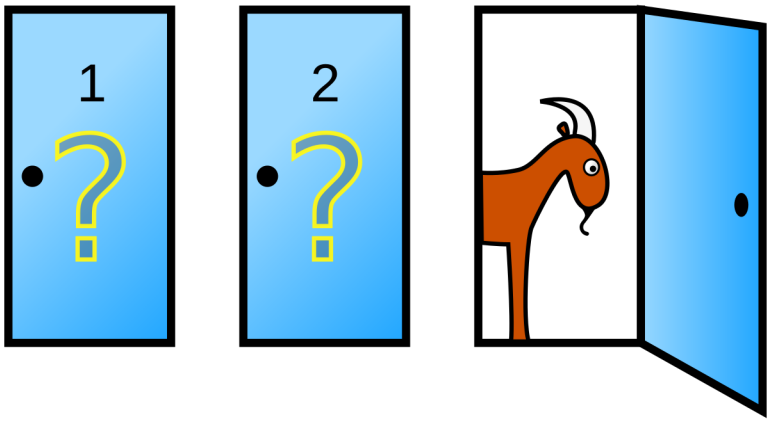


> .25 / .75

The Monty Hall problem

In 1963, a game show called *Let's Make a Deal* first aired in the United States. The show featured a charismatic host, named Monty Hall, who offered audience members, called "traders", a deal: they can either take home an item of modest value, or trade it for an unknown item. Inspired by the game show, Craig F. Whitaker wrote to Marilyn vos Savant's "Ask Marilyn" column in *Parade* magazine in 1990, with the following question:

> 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?



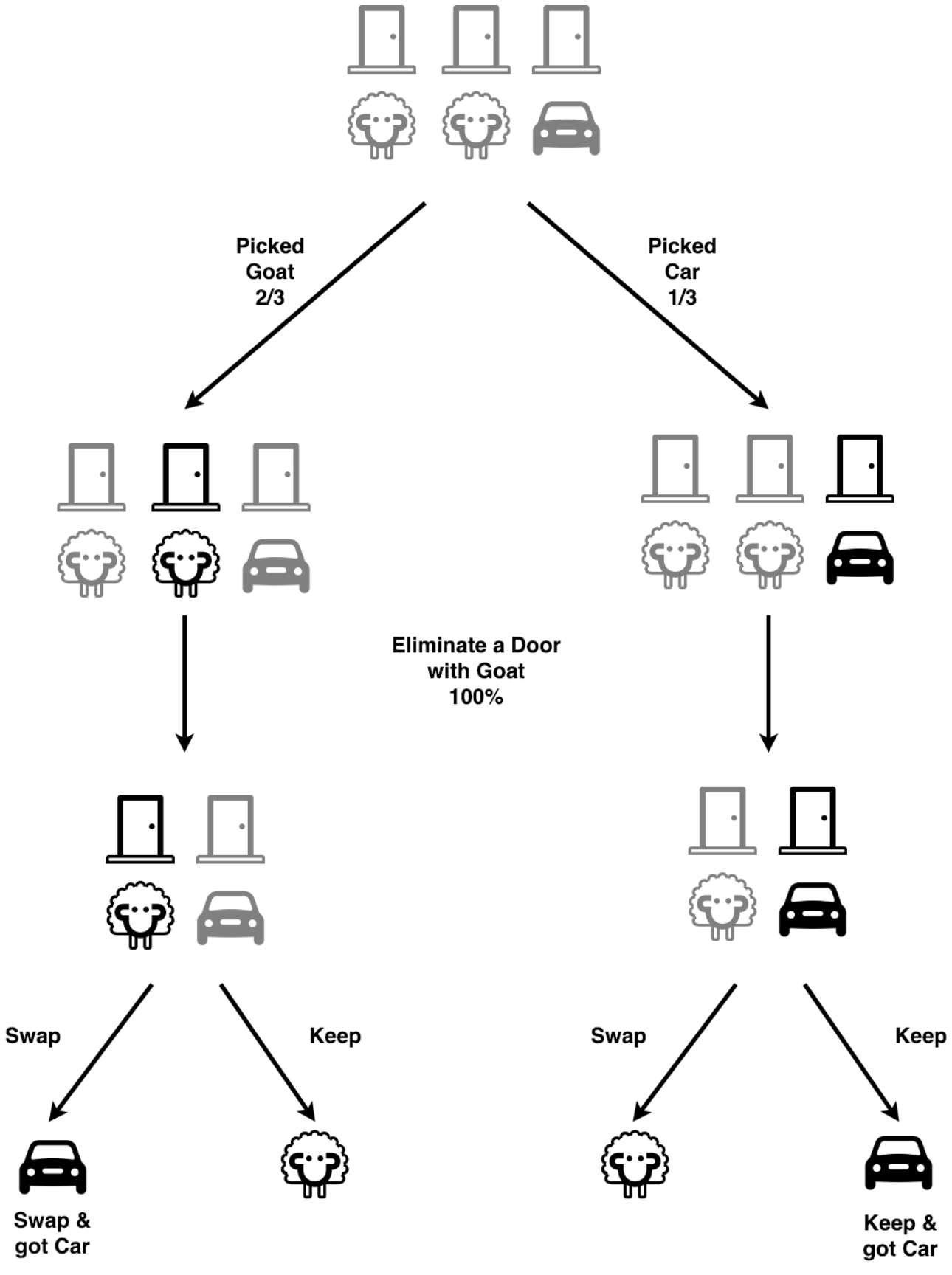
Source: Wikipedia

10. How would you respond to Craig? Is it to the player's advantage to switch their selection? Or does it matter?

    It is to their advantage! Lot of different ways to think about this one - we'll discuss in class!

11. Use the tree diagram below to determine the probability of winning the car if you switch doors. Based on this probability, should you switch?

    2/3 - and yes!



Source: Medium