Activity 01 - The Lady Tasting Tea

STS 2300

Updated: 2024-08-08

**Background**: In England in the 1920s, a group of scientists were enjoying afternoon tea when a biologist, [Dr. Muriel Bristol](https://en.wikipedia.org/wiki/Muriel_Bristol), proclaimed that she thought the tea tasted different depending on whether milk was poured into the cup before or after the tea. Some of the other scientists were skeptical of her claim and wanted to test it. They wondered how the tea could possibly taste any different. Several of the guests decided that they would use their scientific and statistical knowledge to design an experiment that would test whether Dr. Bristol could in fact tell if the milk was poured in first or last.

**Note**: This activity is inspired by real events described in The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century by David Salsburg.

1. Partner up with 1-2 people near you and introduce yourselves if you haven’t yet. Discuss and take notes with some ideas for the following two questions:
   * How would you design an experiment to test whether Dr. Bristol could really tell the difference between the two methods for adding milk? Provide an explanation for how you would conduct this experiment.
   * Once you conducted the experiment, what sort of result would convince you that she could tell the difference between the two methods for adding milk? What sort of result would convince you that she couldn’t tell the difference? Are there any results that would leave you unsure whether she could or could not tell the difference? Explain.

Feel free to add some of your ideas / notes in the space below:

**Wait here once you finish until we have a chance to discuss everyone’s approaches as a class**

1. The people that day decided to pour *eight* cups of tea. In some of the cups, the milk was added before the tea and in some of the cups the milk was added after the tea. Dr. Bristol was then asked to taste each cup and to identify when the milk was added. How many of the eight cups do you think Dr. Bristol would correctly identify *if she couldn’t actually tell when the milk was added* (i.e., if she was just guessing)?

1. Since she would have a 50/50 chance of being right if she were guessing, we could flip a coin to “simulate” how well she might do. We are going to use R to do this! Open the corresponding .R file for Activity 1. Run the following three lines of code and see if you can figure out what each line is doing. We will discuss it together and you can add notes about each line of code.

coin <- c("heads", "tails")  
sample(coin, size = 1)  
sample(coin, size = 8, replace = TRUE)

1. How many of the eight cups did Dr. Bristol get correct when she guessed at random (third line of your code)? Is this the same thing that other people got? Why or why not?

1. Try repeating the third line nine more times (for a total of ten, including your previous answer). Write down the number of heads each time below.

| 1st | 2nd | 3rd | 4th | 5th | 6th | 7th | 8th | 9th | 10th |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |

1. Recall that you just simulated results from a world where Dr. Bristol had been guessing at random. In reality, the people in the room didn’t know if she was guessing at random or not. Based on what you’ve seen, would her correctly identifying 5 of the 8 cups convince you she *wasn’t* guessing at random? Why or why not?

1. What if she had correctly identified 7 of the 8 cups? Would this result be enough to make you think she wasn’t guessing at random? Why or why not?

1. How many cups do you think Dr. Bristol actually got correct on that day?

**Wait here once you finish until we have a chance to discuss everyone’s approaches as a class**

1. The “bonus code” simulates this process 1,000 times instead of 10. Why do you think that might be important to do?

1. What do we learn when simulating this experiment 1,000 times?

The full code from the .R file is as follows. Let’s revisit this after learning some more about R in Notes 01. We can then come back here and take some notes.

# Create a coin object that contains "heads" and "tails"  
  
coin <- c("heads", "tails")  
  
  
# "Flip" the coin once  
  
sample(coin, size = 1)  
  
  
# "Flip" the coin eight times  
  
sample(coin,   
 size = 8,   
 replace = TRUE)  
  
  
  
  
# Bonus code to simulate the process 1000 times  
  
coin\_sim <- replicate(1000, sample(coin,   
 size = 8,   
 replace = TRUE))  
  
sim\_heads <- apply(coin\_sim,   
 MARGIN = 2,   
 FUN = function(x) { sum(x == "heads")} )  
  
table(sim\_heads)  
  
prop.table(table(sim\_heads))  
  
hist(sim\_heads, breaks = seq(-0.5, 8.5, 1))