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, 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.

     - We might give her something like 8 or 10 cups of tea. Some will add milk before tea and some will add milk after. These should be presented to her in a random order.

     - We could try to ensure things like temperature, type of tea/milk, look of the beverage, etc. are held constant (to avoid confounding variables).

   - 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.

     - We would need her to guess over 50% to convince us. Or, maybe 90% or 100%? This will partially depend on our own threshold for how much evidence it takes to convince us she has the ability she claims (related to a significance level).

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

2) 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)?

   We would expect her to get around 4 of the 8 cups (maybe a little more or less due to random chance).

3) 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)
```

   Line 1 created a coin object in R.

   Line 2 "flipped" the coin once.

   Line 3 "flipped" the coin 8 times.

4) 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?

   My result: 5
   This is essentially each of us flipping a coin 8 times. This is a random process, so difference people get different results.

5) 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

   Each person did this individually and should get slightly different results.

6) 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?

   Several of us saw results of 5 correct guesses out of 8, so this would not convince us she had an ability to tell the difference.

7) 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?

   For those who saw 7 out of 8 in their simulation, this result may not convince them that she can tell the difference. For someone who didn't have this happen, they may find this to be convincing.

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

   She got all 8!

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

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

   With only 10 simulations, we don't get a very full sense of what happens when people guess in a situation like this. With 1000 repetitions, we have a better sense of how often she might guess 7 or 8 cups correct by chance alone.

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

   This graph we generated in our simulation shows that it's actually incredibly rare to get 8 cups right if you're guessing. This happened only 5 out of 1000 times in the simulation. This suggests that despite only having looked at 8 cups of tea, we would probably still be comfortable concluding Dr. Bristol can tell whether milk is added first or last to tea. This is because her result would be really unusual or surprising if she couldn't tell the difference.

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))
```