00-welcome-to-stat218

Wednesday, September 20, 2023 9:08 PM

Chapter 0: Welcome to Stat 218 - Applied Statistics for Life Science

Hello, I'm Dr. Robinson ✋ Welcome to Stat 218!

Data is a set of *variables* that capture various aspects of the world. It also contains *observations* over those variables.

Sometimes data is collected intentionally

- The Census
- Pew Research surveys
- Science!
- …and of course, many more

Sometimes data is collected for one reason and used for another

- Health information about you at the doctor
- Emails
- Location information from social media posts
- Paying with a credit card at a gas station
- …and more!

1. Make a list of all the places you left "data exhaust" this week (MentiMeter Poll).

Introduce Yourself!

1. What are you currently obsessed with? For me, it's pickleball and nectarines!

What to Expect

… let's look at the syllabus together!

Sometimes data is collected for one reason and used for another

- Health information about you at the doctor
- Emails
- Location information from social media posts
- Paying with a credit card at a gas station
- ...and more!

1. Make a list of all the places you left "data exhaust" this week (MentiMeter Poll).

Introduce Yourself!

1. What are you currently obsessed with? For me, it's pickleball and nectarines!

What to Expect

... let's look at the syllabus together!

Manages communications	ideas. In lab assignments, responsible for collecting, organizing, and recording answers to the assignment during the discussions, compiling the summary of the answers discussed, soliciting feedback on summaries from other group members, and submitting the final version of the lab assignment.
Resource Manager *Manages team resources*	Responsible for pulling up course notes and other useful resources (e.g., R note cards, applets, homework assignments, etc.). Contacts professor/TA for *group questions only*. Ensures everyone understands team questions. Collects team supplies and organizes clean-up.
Team Captain *Manages team participation*	Responsible for making sure everyone has a chance to contribute, asking quiet team members to speak up, asking loud team members to listen to others, and bringing the conversation back to the lab assignment if it deviates.

If you are in a group of *three*, the Facilitator also acts as the Team Captain*.

1. Please find your team members and introduce yourself by sharing something about your name.

Suggested Class Norms

If you are in a group of *three*, the Facilitator also acts as the Team Captain*.

1. Please find your team members and introduce yourself by sharing something about your name.

Suggested Class Norms

- Share talk time
- Listen to understand
- Everyone has expertise
- Be present*
- Critique ideas, not people
- Embrace discomfort
- Treat everyone with respect
- Normalize time to think / process
- (Others? Changes?)

Example 0.1: Question the Arson

I am a detective in San Luis Obispo, California, and everyone in this room is a suspect in the three fires that have been set in the last 6 weeks. I have called all of you in for questioning. Given are 16 questions I want you to answer about these fires. Please answer all 16 questions.

THE TRIBUNE

SLO police arrest resident suspected of lighting 3 fires at Madonna Plaza

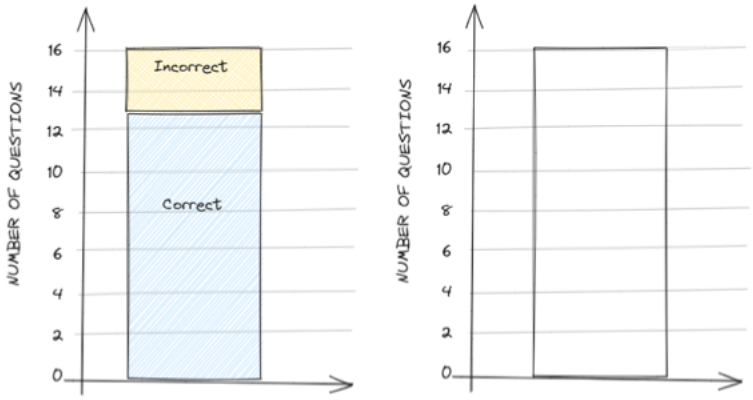
• One of the fires was started using gasoline.	T	F
• Matches were used to set the fires.	T	F
• The trash was on the curb in front of the first house that caught on fire.	T	F
• There was a doghouse in the backyard at the scene of the second fire.	T	F
• The residence of the third fire had off-street parking.	T	F
• The second residence had a screened porch.	T	F
• The suspected entered the first residence through the kitchen window.	T	F
• The suspect left a flashlight behind at the scene of the second fire.	T	F
• All three fires took place on Tuesdays.	T	F
• The porch light was on at the residence where the first fire took place.	T	F
• The third fire was started in the garage.	T	F
• There was a dog barking at the house next door to the third fire.	T	F
• The suspect wore work boots at the scene of the first fire.	T	F
• The home security system was triggered when the suspect fled the scene of the third fire.	T	F
• The lock on the back door at the first residence was broken.	T	F
• Multiple fires were set at the second residence.	T	F

- One of the fires was started using gasoline. T F
- Matches were used to set the fires. T F
- The trash was on the curb in front of the first house that caught on fire. T F
- There was a doghouse in the backyard at the scene of the second fire. T F
- The residence of the third fire had off-street parking. T F
- The second residence had a screened porch. T F
- The suspected entered the first residence through the kitchen window. T F
- The suspect left a flashlight behind at the scene of the second fire. T F
- All three fires took place on Tuesdays. T F
- The porch light was on at the residence where the first fire took place. T F
- The third fire was started in the garage. T F
- There was a dog barking at the house next door to the third fire. T F
- The suspect wore work boots at the scene of the first fire. T F
- The home security system was triggered when the suspect fled the scene of the third fire. T F
- The lock on the back door at the first residence was broken. T F
- Multiple fires were set at the second residence. T F

8 since that is half of 16 "guessing"

more than half, enough to show you have more inform..

Only one person in class had 13 or more correct... maybe they were unlucky



3. How do your results compare to the convicted suspect's results? How about the other suspects (i.e., your neighbor's) results?

More convinced, stronger evidence

4. Are you convinced the arson suspect was rightfully convicted?

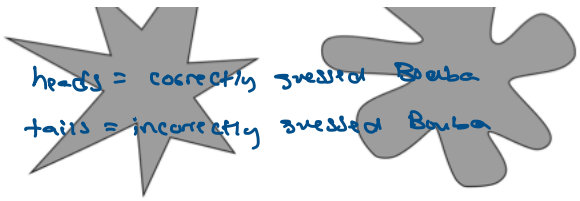
less convinced, more likely to be guessing

5. How would your answer change if the convicted arson suspect answered...

suspicious 32

31

16 - 50/50 chance of correct



4. Are you convinced the arson suspect was rightfully convicted?

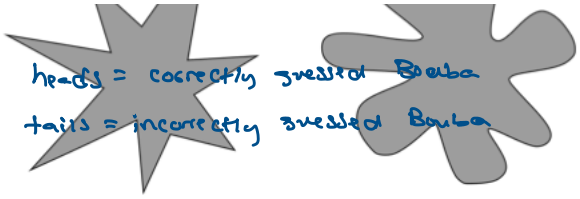
less convinced, more likely to be guessing correct

5. How would your answer change if the convicted arson suspect answered...

suspicious 32

31

1/2 - 50/50 chance of correct



1. Were you correct or incorrect in identifying Bouba?

2. How many students are in class today?

3. How many students in the entire class were correct in identifying Bouba? (MentiMeter Poll 1)

4. If we really don't know Martian and are just guessing which is Bouba, how many students would you *expect* to choose Bouba correctly? Explain your reasoning.

everyone flip a coin, aggregate results.

5. How could we use a coin to simulate each student "just guessing" which Martian letter is Bouba?

incorrect

21/32 were correct

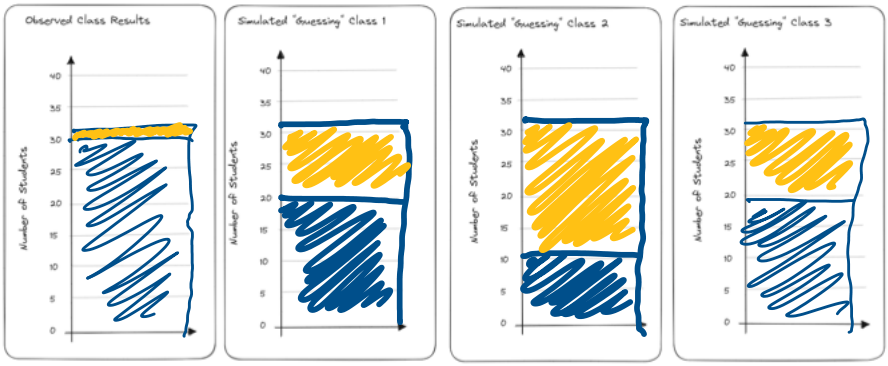
More extreme than the simulated results

12/32 correct

9. Let's flip our coins one more time. How many students "guessed" correctly in the third class simulation? (Sketch your results in Q10)

20/32 correct

10. Sketch the observed class results and the three simulated "guessing" class results.



21 / 32 were correct

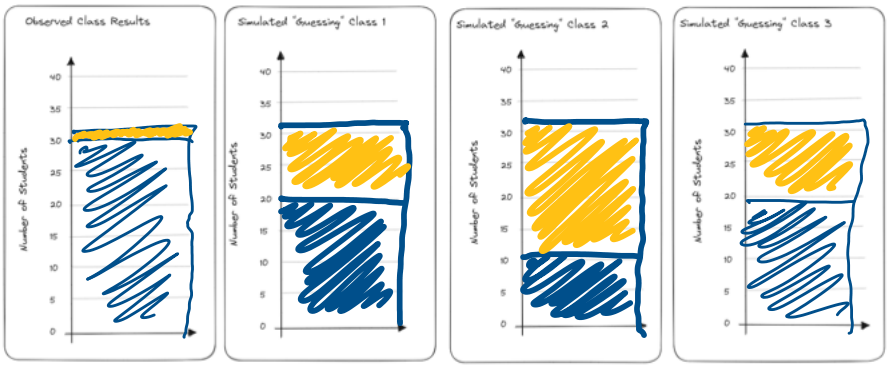
More extreme than the simulated results

12/32 correct

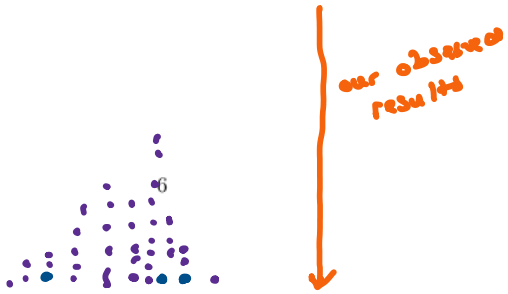
9. Let's flip our coins one more time. How many students "guessed" correctly in the third class simulation? (Sketch your results in Q10)

20/32 correct

10. Sketch the observed class results and the three simulated "guessing" class results.



11. How do the observed class results compare to the simulated guessing classes results?



We still only have a few simulations to compare our class data to. It would be much better to be able to see how our class compared to hundreds or thousands of "just-guessing" classes. Since we don't want to flip coins all class period, we will use a computer simulation (`Canvas > Online Simulation Applets > One proportion inference`) to get 100 "random guessing" statistics.

13. Carry out the applet simulation. Note that you should consider the following questions when designing your simulation study:

- What are the two possible outcomes on each of the trials?
- What is the probability that a Bouba is accurately identified under the assumption that we are "just-guessing"? Change your `Probability of heads` accordingly.

  Probability of heads: 0.5

- How many students were there in this study? Keep this value in mind when setting the `Number of tosses` value.

  Number of tosses: 32

Carry out the simulation study 100 times overall, keeping track of the probability of employees chosen for management that were female on each of the simulated experiments. Sketch the 100 simulated guessing class results on the dotplot in Q12.

Number of repetitions: 100

yes, all of the guessing classes had results for # of students correct less than our observed results

yes, if we continued to simulate, maybe we would see 31.

- What are the two possible outcomes on each of the trials?
- What is the probability that a Bouba is accurately identified under the assumption that we are "just-guessing"? Change your `Probability of heads` accordingly.

  Probability of heads: 0.5

- How many students were there in this study? Keep this value in mind when setting the `Number of tosses` value.

  Number of tosses: 32

Carry out the simulation study 100 times overall, keeping track of the probability of employees chosen for management that were female on each of the simulated experiments. Sketch the 100 simulated "guessing" class results on the dotplot in Q12.

Number of repetitions: 100

yes, all of the guessing classes had results for fewer students correct test than our observed results

yes, if we continued to simulate, maybe we would see 31.

No, none in our 100 simulations were as extreme

17. Does this activity provide evidence that students were not just guessing at random? If so, what do you think is going on here? Can we as class read Martian?

Looks like there might be evidence.

TED Talk: 3 Clues to Understanding Your Brain by Vilayanur Ramachandran (2007)

The synesthesia part begins at roughly 17:30 minutes: https://www.ted.com/talks/vs_ramachandran_3_clues_to_understanding_your_brain