Applying Bayes Theorem

In the last activity (see "Developing Bayes' Theorem"), you were introduced to Bayes' Theorem. In this activity, you will practice applying Bayes' Theorem to investigate questions in public health, psychology, and sports.

Manager:		Recorder:	
Presenter:		Reflector:	

Content Learning Objectives

After completing this activity, students should be able to:

- Apply Bayes' Theorem
- Construct probabilities from data to update beliefs

Process Skill Goals

During the activity, students should make progress toward:

- Identifying, planning, and executing a strategy that goes beyond routine action to find a solution to a situation or question (Problem Solving)

Model 1 Bayes Theorem Application

To dive into Bayes' Theorem, this activity will direct you to a website where you can work with an interactive web app to explore some of the probabilities that are important for using Bayes Theorem. The initial context is a medical one, where you will be asked to consider probabilities related to testing positive to a rare disease.

Medical Context for Bayes' Theorem from Seeing Theory

Click on the following link to go to Chapter 5 of "Seeing Theory: A visual introduction to probability and statistics" (created by Daniel Kunin while an undergraduate at Brown University). Seeing Theory: Chapter 5[4]

Read the first two paragraphs on the web-page and explore the population on the right, that is, try to make sense of the red and blue dots.

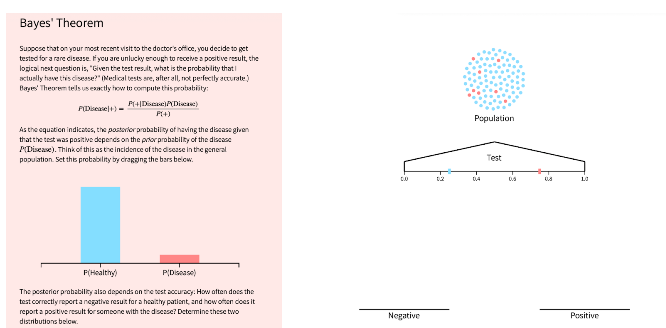

Figure 1: Seeing Theory App

Questions (15 min) Start time:

1. Let's explore the relationship between the bar plot and the cloud of dots.

a) What happens to the cloud of dots if you drag the blue and red bars to be equal to each other?

b) What happens to the cloud of dots when you drag the blue bar up to 1.00?

c) What happens to the cloud of dots when you drag the blue bar down to 0?

d) How does the bar plot on the left relate to the cloud of dots on the right?

2. Now take a look at the graph that has four bars (2 red and 2 blue). They are each labeled. Match the notation with its correct interpretation "in English."

	P(- | D)
	P(- | H)
	P(+ | D)
	P(+ | H)

English interpretations of conditional probabilities

A "Probability of negative test result given that the person is healthy"

B "Probability of positive test result given that the person is healthy"

C "Probability of negative test result given that the person has disease"

D "Probability of positive test result given that the person has disease"

Remember, three of the probabilities in the above discussion also have special names:

P(D) = probability of having disease (before taking a test) = prior
P(+ | D) = likelihood = sensitivity
P(- | H) = specificity

3. Refresh the web app so that we go back to 90% of the population being healthy and 10% having the disease (i.e. the prior is P(D) = 0.10). Then click "Test One patient"

a) Did that patient have the disease or were they healthy? How can you tell?
Note: your result might not match your neighbor's!

b) Did that patient test positive or negative?
Note: your result might not match your neighbor's!

4. Click "Test Remaining"

a) Look at the "Negative" pile - why are there red and blue dots?

b) Compare the number of dots in the two piles. Were there more "Negative" or "Positive" results?

c) Interpret - what do the blue dots in the "Negative" pile represent?

d) What do the red dots in the "Negative" pile represent?

e) Interpret - what do the blue dots in the "Positive" pile represent?

f) What do the red dots in the "Positive" pile represent?

g) Why are there more blue dots than red dots in the Negative and Positive piles?

5. Using the app, what is $P(D|+)$?

Confirm using Bayes Theorem. Show your work.

6. What happens to the posterior probability when prevalence changes?

7. What happens to the posterior probability when sensitivity changes?

8. What happens to the posterior probability when specificity changes?

Great job using Bayes' Theorem in the medical complex. In the next model we explore Bayes' Theorem in a different context.

Model 2 Updating Beliefs

The following example is adapted from the excellent 3Blue1Brown resource, by Grant Sanderson [5].

Steve is very shy and withdrawn, invariably helpful but with very little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.

Which of the following do you find more likely?

- Steve is a librarian
- Steve is a farmer

If you think it's more likely that Steve is a librarian, you're not alone! But did you stop to consider whether there are more librarians or farmers in general?

This is a famous example from two cognitive psychologists and behavioral economists Daniel Kahneman & Amos Tverksy, who found that most people associate the descriptions "shy and withdrawn" and "meek and tidy soul" with librarians, but very few people take into account the base rate of how many people are farmers vs. librarians in their judgements [1]. This is known as the *base rate fallacy*, and is one of many types of "cognitive biases" Kahneman & Tversky's work have uncovered about the way often "intuitive" human judgements don't align with what laws of probability suggest as "rational decision-making". The two won a Nobel Prize for their work!

It turns out, Bayes' Theorem provides the key to reasoning about whether one statement is more likely to be true than another (e.g. Steve is a librarian vs. Steve is a farmer), and helps us overcome the base rate fallacy. Let's see how it works.

Questions (15 min) **Start time:** ☐

Let's suppose that there are actually 20 farmers for every 1 librarian. So for example, if there are 10 librarians, there would be 200 farmers.

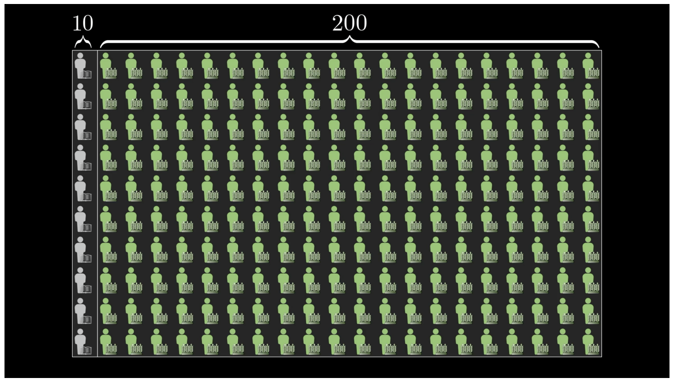

Figure 2: Base rate / prior / prevalence of librarians

9. Use Figure 1 to find the following probabilities:

a) P(Farmer) ☐

b) P(Librarian) ☐

There are surely some librarians AND some farmers who would fit Kahneman & Tyversky's description of a "meek and tidy soul". The graphic below highlights which people in both groups fit the description.

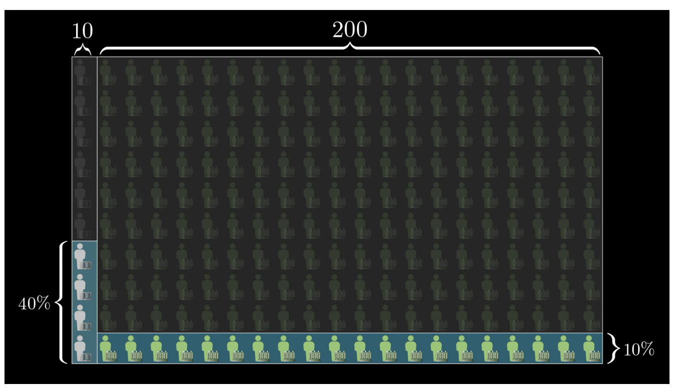

Figure 3: Likelihood of Fitting Description

10. Use Figure 2 to find the following probabilities:

a) P(description | librarian) ☐

b) P(description | farmer) ☐

11. Use Bayes' Theorem to find P(librarian | description) and P(farmer | description). Write out your work.

12. According to your results, is it more likely that Steve is a librarian or a farmer?

13. Let's recap with some Bayesian terminology.

a) Before knowing any information (description) about Steve, what was the prior probability that he was a librarian?

b) What was the likelihood that a librarian fits the description of a meek and tidy soul?

c) What was the likelihood that a farmer fits the description of a meek and tidy soul?

d) After incorporating the information that Steve is a meek and tidy soul, what was the posterior probability that he was a librarian?

$2/26 = 1/13 = 0.0769 = 7.69\% \ = 8\%$
In Q10 and Q13bc, you found that librarians were 4 times more likely than farmers to fit the description (40% vs. 10%). Even still, this was not enough to overcome the prior fact (base rate) that there are many more farmers than librarians.

This is the key to Bayesian reasoning: your beliefs and decisions (e.g. whether its more likely that Steve is a librarian or farmer) should not be entirely based on new evidence in a vacuum (e.g. Steve is a meek and tidy soul), but rather new evidence should update your prior beliefs and knowledge about the world (e.g. that there are many more farmers than librarians)! In the same way, you shouldn't base your beliefs entirely on prior information, ignoring all new evidence. **Your posterior beliefs should be a combination informed by both prior information and new evidence.**

Bayesian statistics provides a mathematical framework for the process of updating beliefs by incorporating new evidence. Stated differently, Bayesian thinking allows us to evaluate a hypothesis, in light of evidence.

$$P(Hypothesis|Evidence) = \frac{P(Evidence|Hypothesis)P(Hypothesis)}{P(Evidence)}$$

$$posterior = \frac{likelihood \times prior}{P(Evidence)}$$

14. What are one or two additional scenarios you can think of when Bayesian thinking might be useful? This could be a scientific decision-making situation (e.g. P(Flu | +)) or an informal decision-making situation in your own life (e.g. updating a belief based on new information).

Model 3 Sports Predictions

Probability and statistical models are often used in sports to predict the outcome of a game, based on data about a team's past performance. The following data provides an example for soccer. Suppose you want to know the probability that your team will win, given that they are currently down by 1 goal against their opponent. The following information provides data for how your team performed in its 34 games last season, broken down by whether they faced a 1-goal deficit (down by 1 goal) during those games or not.

- They won 25 of their 34 games.
- Altogether, they played in 13 games where they faced a 1-goal deficit at some point during the game. Of these, they won 5, lost 7, and tied 1.

Questions (15 min) Start time:

15. Before the game begins, what's the probability our team will win?

16. Assuming your team is down by 1, what is the (posterior) probability they will come from behind and win the game? What's the probability they will lose? Tie? Use Bayes' Theorem and show your work.

Congrats, you've just used Bayes' Theorem as a predictive model to predict the outcome of a sporting event! In today's world, most professional sports teams have a team of data scientists working for them to run lots of predictive models about player performance and game outcomes. In our simple model above we used a single categorical variable - team record (win/loss/tie) - to inform our prior probabilities and a single categorical variable - whether the team faced a one-goal deficit (yes/no) - as "evidence" to update the probabilities. It turns out, Bayesian statistics is not just limited to this scenario of utilizing two categorical variables; it can be extended to take lots of complex data into account. In real life, data scientists build sports analytics models that take in lots of data to inform their prior win probabilities (not just team record) and complex real-time "evidence" to update their probabilities (e.g., all the factors of the game and the opponent up to that point in the game). Yet the same Bayesian reasoning applies (e.g. priors, likelihoods, posteriors). In the next activity, you will be introduced to a new powerful Bayesian model - the Beta-Binomial - that allows for a lot of flexibility in determining a prior model.

If your curiosity is peaked, consider reading more about Bayesian reasoning through these freely accessible online textbooks.

- Bayes Rules! An Introduction to Applied Bayesian Modeling (by Alicia Johnson, Miles Ott, and Mine Dogucu). Bayes Rules[3]

- Probability and Bayesian Modeling (by Jim Albert and Jingchen Hu). Probability and Bayesian Modeling[2]

References

[1] Daniel Kahneman. ***Thinking, Fast and Slow***. Farrar, Straus and Giroux, Oct. 2011. ISBN: 978-1-4299-6935-2.

[2] Jim Albert and Jingchen Hu. ***Probability and Bayesian Modeling***. 2020. URL: https://bayesball.github.io/BOOK/probability-a-measurement-of-uncertainty.html.

[3] Alicia Johnson, Miles Ott, and Mine Dogucu. ***Bayes Rules! An Introduction to Applied Bayesian Modeling***. 2021. URL: https://www.bayesrulesbook.com/.

[4] Daniel Kunin. ***Seeing Theory-A visual introduction to probability and statistics.*** URL: http://seeingtheory.io. (accessed: 05.29.2024).

[5] Grant Sanderson and Josh Pullen. ***3Blue1Brown - Bayes' Theorem***. URL: https://www.3blue1brown.com/lessons/3blue1brown.com. (accessed: 05.29.2024).