DATA 8

Spring 2022

Lecture 37

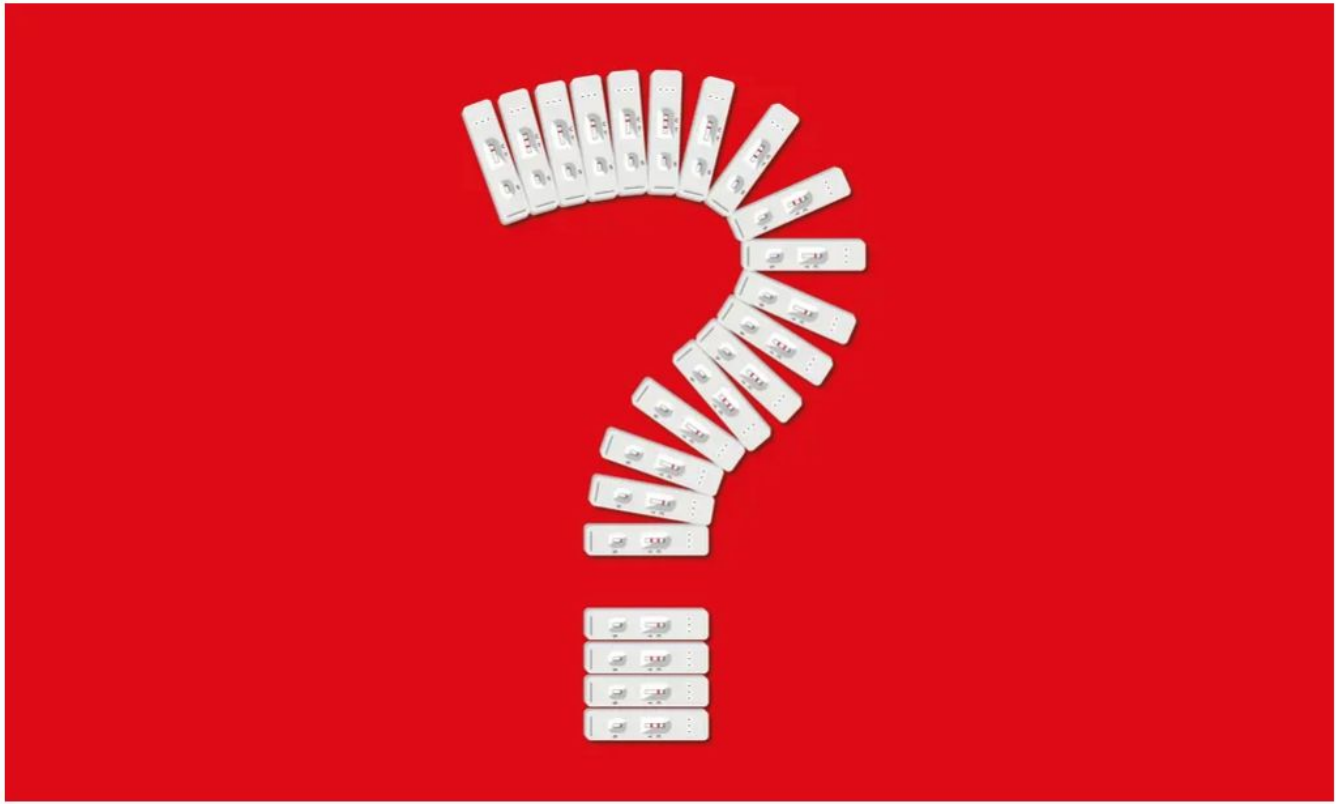
Updating Probabilities

Announcements

- **Lab 10** is optional
  - There will be a couple lab sessions this week
- **Project 3** is due Friday (12/2)
- GSIs will hold some **in person & online OH** this week
  - My OH are a bit longer today, 3:30-6:30pm @ FSM
- **Final Exam** is still happening in person as planned (12/15)
  - Review sessions next will be fully online

Updating Probabilities

The obscure maths theorem that governs the reliability of Covid testing



Source: Guardian

Updating Probabilities

Interpretation by Physicians of Clinical Laboratory Results (1978)

> "We asked 20 house officers, 20 fourth-year medical students and 20 attending physicians, selected in 67 consecutive hallway encounters at four Harvard Medical School teaching hospitals, the following question:
>
> "If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?"

Updating Probabilities

Interpretation by Physicians of Clinical Laboratory Results (1978)

> "Eleven of 60 participants, or **18%, gave the correct answer**. These participants included four of 20 fourth-year students, three of 20 residents in internal medicine and four of 20 attending physicians. The most common answer, given by 27, was that the chance that a person found to have a positive result actually has the disease was 95%.

Conditional Probability

Scenario 1

- Scenario:
  - Class consists of second years (60%) and third years (40%)
  - 50% of the second years have declared their major
  - 80% of the third years have declared their major
- **I pick one student at random.**
- Which is more likely: Second year or Third year?
  - Second year, because they are 60% of the class

Scenario 2

- Slightly different scenario:
  - Class consists of second years (60%) and third years (40%)
  - 50% of the second years have declared their major
  - 80% of the third years have declared their major
- **I pick one student at random...**
  That student has declared a major!
  (Demo)
- Which is more likely: Second Year or Third Year?

Bayes' Rule

Purpose of Bayes' Rule

- Update your prediction based on new information
- In a multi-stage experiment,
  find the chance of an event at an earlier stage,
  given the result of a later stage

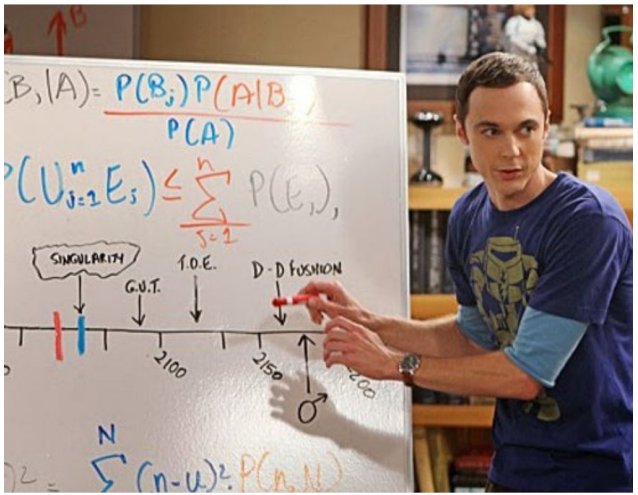
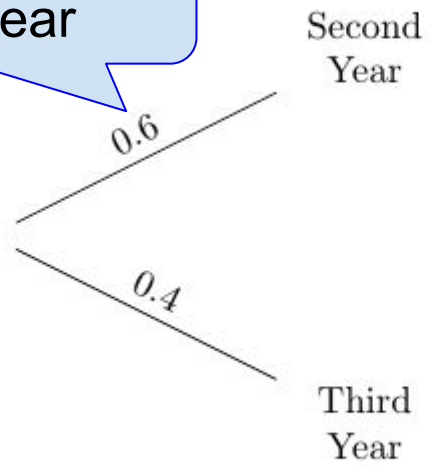


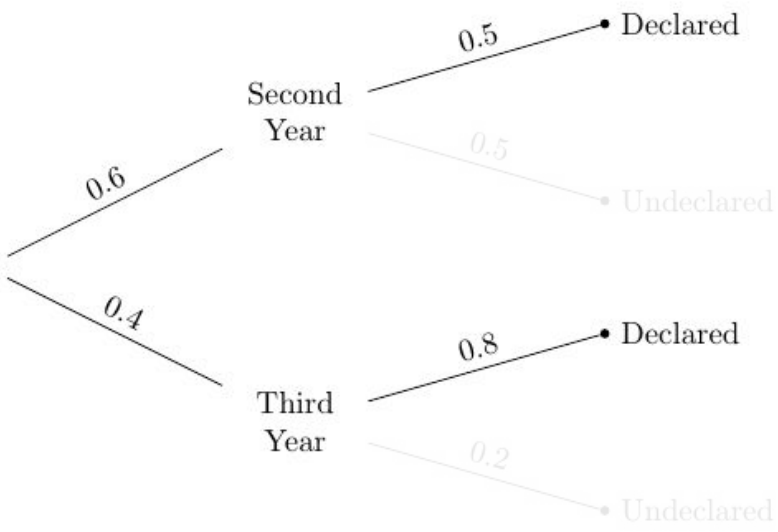
Diagram and Terminology

Prior probability of being a 2nd year



Likelihood of declared, given 2nd year

Data & Calculation



Pick a student at random.

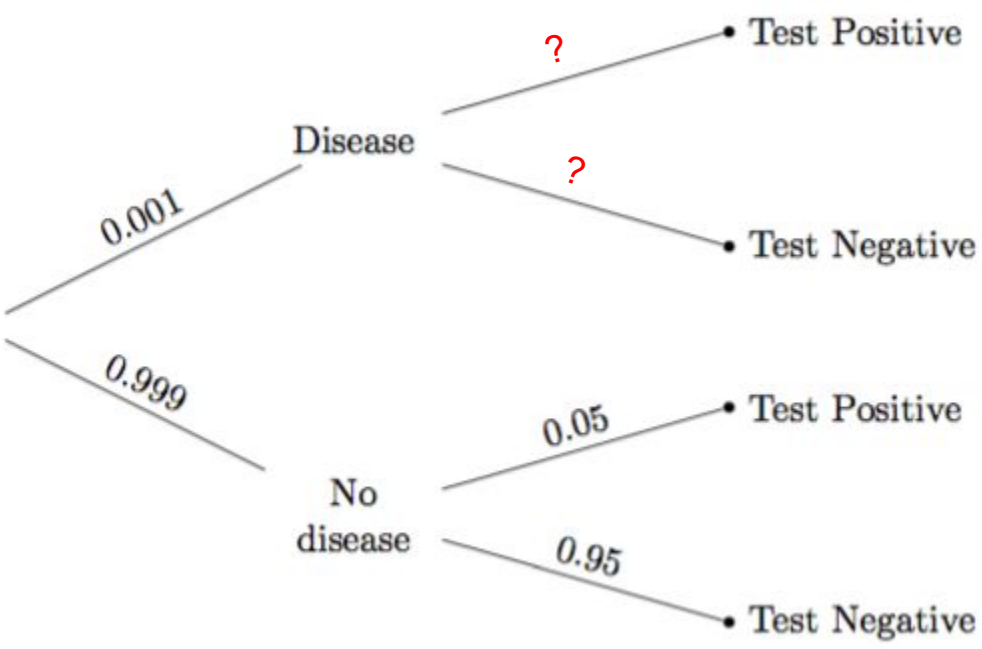
Posterior probability:

P(Second Year | Declared)

$$= \frac{0.6 \times 0.5}{(0.6 \times 0.5) + (0.4 \times 0.8)}$$

= 0.4839…

Example: Doctors & Clinical Tests

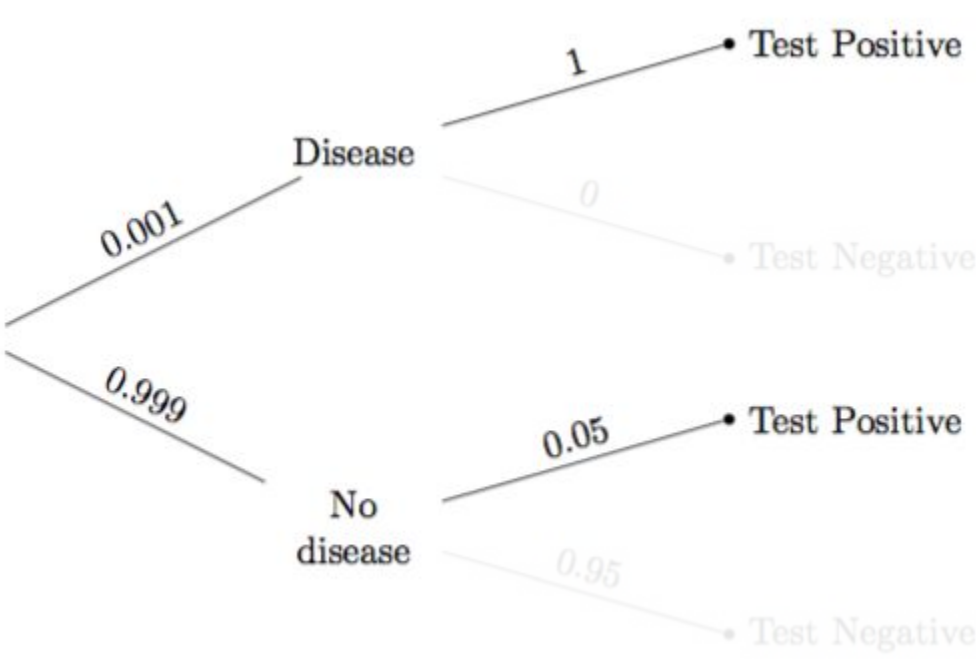


Problem did not give the *true positive* rate.

That's the chance the test says "positive" if the person has the disease.

It was assumed to be 100%.

Data and Calculation



P(Disease | Test +)

=

$$\frac{0.001 * 1}{(0.001 * 1) + (0.999 * 0.05)}$$

= 0.0196270…

(Demo)

Visual Explanation

Subjective Probabilities

Subjective Probabilities

A probability of an outcome is…

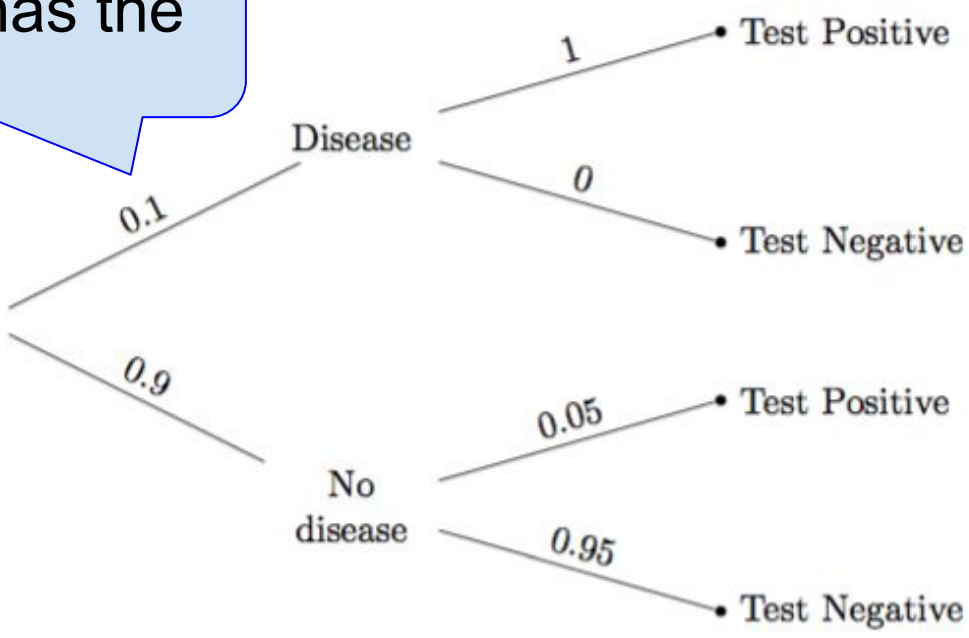
- The frequency with which it will occur in repeated trials, *or*
- The subjective degree of belief that it will (or has) occurred

Why use subjective priors?

- In order to quantify a belief that is relevant to a decision
- If the subject of your prediction was not selected randomly from the population

A Subjective Opinion

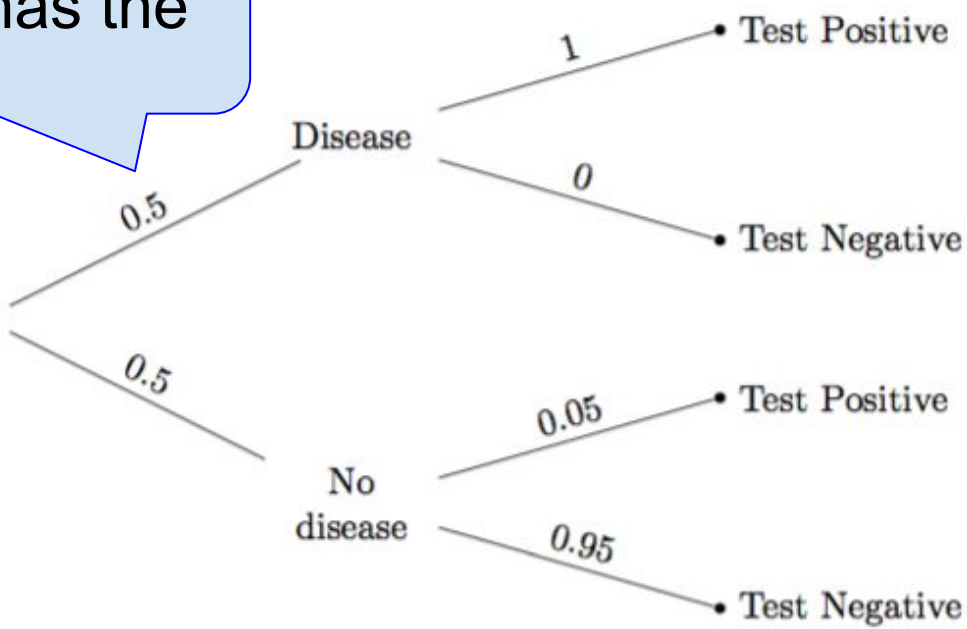
prior probability that the person has the disease



(Demo)

A Different Subjective Opinion

prior probability that the person has the disease



(Demo)