Lecture 13: Lesson and Activity Packet

MATH 232: Introduction to Statistics

November 04, 2016

Homework and Announcements

- Homework 13 due in class today
- Exam 2 is on Wednesday next week
  - It will cover everything we do through the end of today
  - Monday will be a review day
  - Practice test will be on Canvas before Monday
- Book Review Project
  - Details to be given after the Election Project, **but** you should buy or procure (by legal means...) the book of your choice now, from among the following three choices:
    1. *Moneyball: The Art of Winning an Unfair Game*, by Michael Lewis
    2. *Dataclysm: Love, Sex, Race, and Identity—What Our Online Lives Tell Us about Our Offline Selves*, by Christan Rudder
    3. *How Not to Be Wrong*, by Jordan Ellenberg
  - You must make your choice before Monday, and **use groups on Canvas to sign up for your desired book**

Last time

- Conditional probability
- Independent events
- Multiplication rules

Recap of today

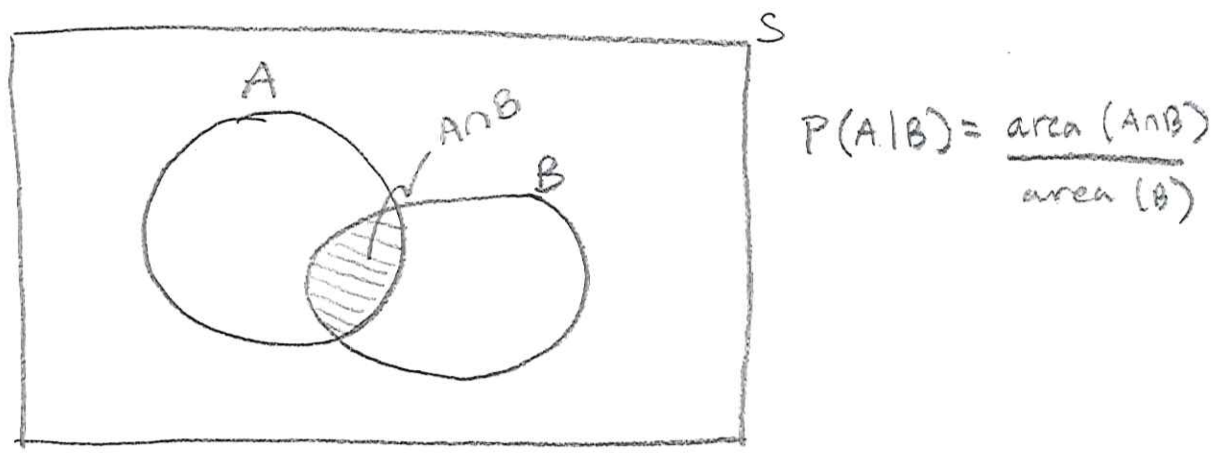
- Law of Total Probability
- Bayes' Theorem

We'll begin with a recap of conditional probability.

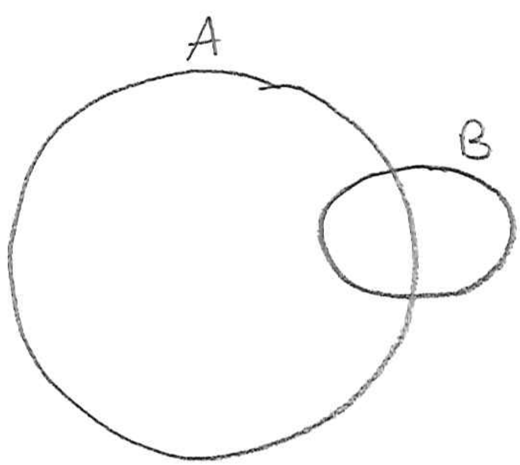
Remember that the conditional probability of A given B can be computed with the formula

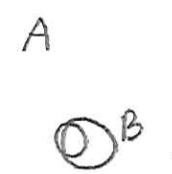
$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}.$$

What this says is that $P(A|B)$ is the **proportion** of the outcomes in B that are also in A. We can illustrate this using a Venn diagram:

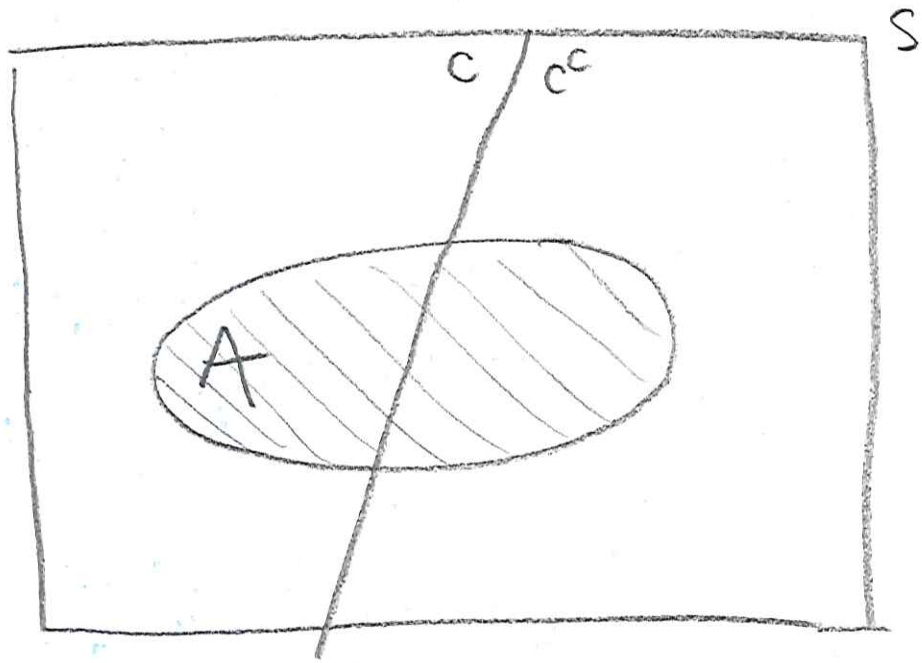


Notice that the size of A has little to do with $P(A|B)$. What's important is the **ratio** of the size of $A \cap B$ to the size of B itself. Two more examples that illustrate this:





We now introduce the **law of total probability** using Venn diagrams.



Theorem 1 (*Law of Total Probability*)

For any two events A and C,

$$P(A) = P(A|C)P(C) + P(A|C^C)P(C^C).$$

This law doesn't seem so useful now, but we will see its application in Bayes' Theorem before the end of this packet.

Although these examples show us that $P(A|B) \neq P(B|A)$, we might be interested in the question of whether there is some other formula that relates $P(A|B)$ and $P(B|A)$.
Recall the two versions of the multiplication rule:

$$P(E \text{ and } F) = P(F) \cdot P(E|F)$$

and

$$P(E \text{ and } F) = P(E) \cdot P(F|E).$$

Notice that these two version of the rule give two different ways of writing the same quantity $P(E \text{ and } F)$. So we can write:

$$P(F) \cdot P(E|F) = P(E) \cdot P(F|E),$$

which boils down to

$$P(E|F) = \frac{P(E) \cdot P(F|E)}{P(F)}.$$

Notice that this says that to go from $P(E)$ to $P(E|F)$, you multiply by the factor $\frac{P(F|E)}{P(F)}$.

Definition 1 (***Bayes' Factor***)

For any events E and F,

$$P(E|F) = \underbrace{\frac{P(F|E)}{P(F)}}_{\text{Bayes' factor}} P(E).$$

Recall that the law of total probability tells us the following about $P(F)$:

$$P(F) = P(F|E)P(E) + P(F|E^C)P(E^C).$$

Substituting this into the equation for $P(E|F)$, we obtain

$$P(E|F) = \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^C)P(E^C)},$$

which is exactly Bayes' Theorem.

Theorem 2 (***Bayes' Theorem***)

For any two events E and F,

$$P(E|F) = \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^C)P(E^C)}.$$

We will see an application of this to medical tests...

Suppose that the public health department tells us that 1 in 1,000 people has a certain disease; suppose further that researchers and clinicians tell us that a test for the disease yields a positive result 90% of the time when the disease is present, and 1% of the time when the disease is not present.

Question 1

Given that someone tests positive for the disease, what is the probability he or she actually has the disease?

We could use Bayes' Theorem to figure this out. But let's have an intuitive look ~~first~~ also, to see how Bayes' Theorem works...

First: Let D be the event that someone has the disease;

let T — " ——— tests positive.

Then we're given $P(D) = 0.001 = \frac{1}{1{,}000}$, $P(T \mid D) = 0.90$, and $P(T \mid D^c) = 0.01$.

We seek to compute $P(D \mid T)$.

Could plug into Bayes' Thm: $$P(D \mid T) = \frac{P(T \mid D)P(D)}{P(T \mid D)P(D) + P(T \mid D^c)P(D^c)}$$

$$= \frac{P(T \mid D)P(D)}{P(T \mid D)P(D) + P(T \mid D^c)(1 - P(D))}$$

$$= \frac{0.90 \cdot 0.001}{0.90 \cdot 0.001 + 0.01 \cdot (1 - 0.001)} \approx 0.0826, \text{ so } 8.3\%.$$

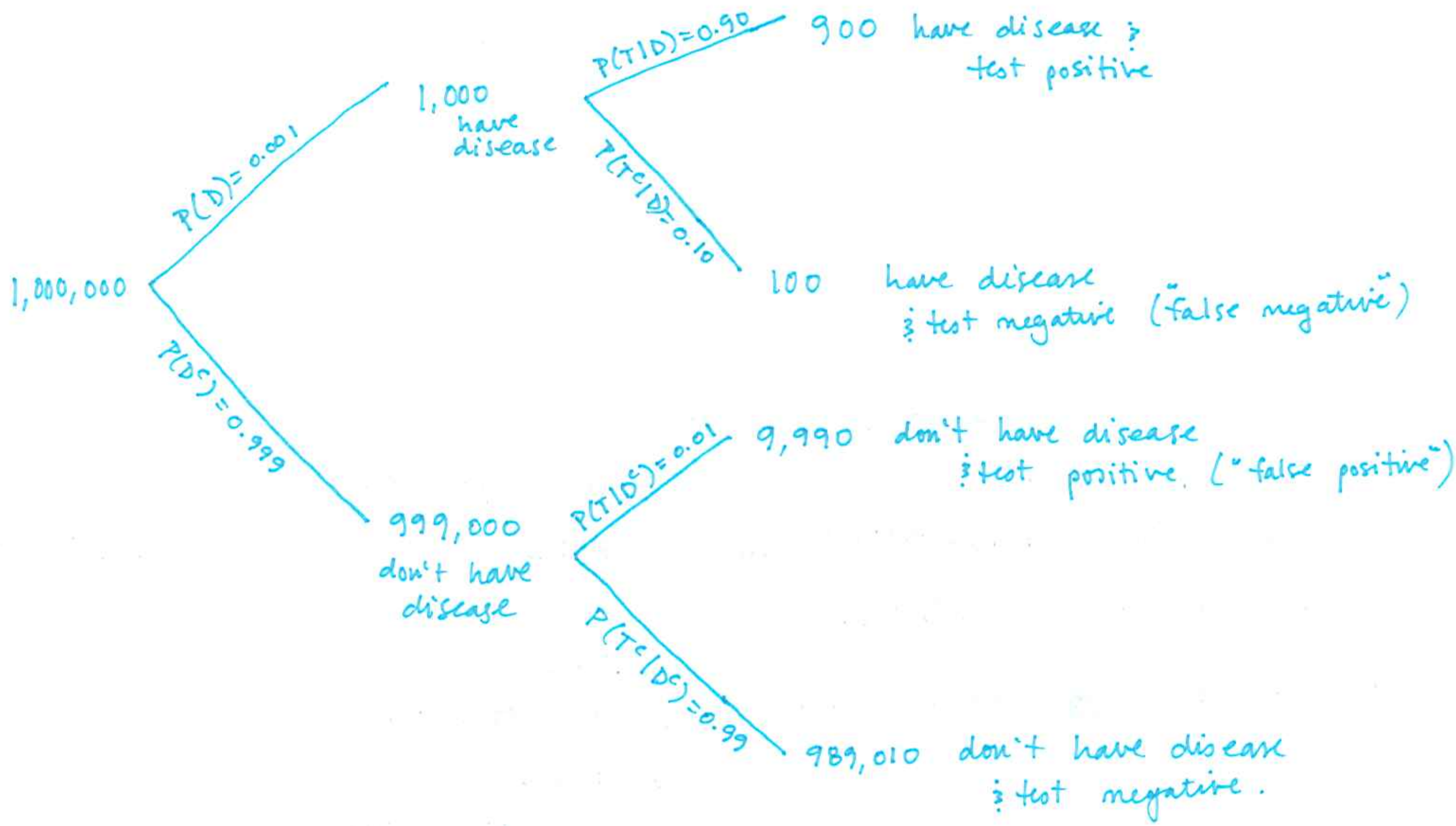
Could see where this comes from by assuming a population of 1,000,000 people (or any other number that makes the arithmetic work nicely), and computing how many people are in each category of the table:

	T Test pos.	T^c Test neg.	
D disease			
D^c no disease			

We do this using a tree diagram (see next pg.) ⟶

Given: $P(D) = \frac{1}{1{,}000} = 0.001$, $P(T \mid D) = 0.90$, $P(T \mid D^c) = 0.01$,

and a population of 1,000,000 people:



So the table looks like:

	T	T^c	TOTAL
D	900	100	1,000
D^c	9,990	989,010	999,000
TOTAL	10,890	989,110	1,000,000

Therefore, $P(D \mid T) = \frac{P(T \cap D)}{P(T)} = \frac{900}{10{,}890} \approx 0.0826$, or 8.3%.

Please see the Science News article given in class on Friday!