The Three Probability Axioms

PA1 **Nonnegativity**: For any statement A, $\Pr(A) \geq 0$.

PA2 **Certainty**: If you're certain that a statement A is true, then $\Pr(A) = 1$ (and if you're certain that A is false, then $\Pr(A) = 0$).

PA3 **Additivity** If A and B are never true at the same time, then $\Pr(A \vee B) = \Pr(A) + \Pr(B)$.

The following is a consequence of the probability axioms:

1. **Partitionality**: If $H_1, H_2, \ldots H_n$ are statements and you know that *exactly one* of them is true, then $\Pr(H_1) + \Pr(H_2) + \cdots + \Pr(H_n) = 1$.

The Ratio Formula and Theorems based on it

1. **The Ratio Formula for Conditional Probability**: If A and B are statements and $\Pr(B) \neq 0$, then $\Pr(A \mid B) = \frac{\Pr(A \& B)}{\Pr(B)}$.
2. **Bayes' Theorem**: If E and H are statements, then

$$\Pr(H \mid E) = \Pr(H)\frac{\Pr(E \mid H)}{\Pr(E)}.$$

3. **Total Probability**: If $H_1, H_2, \ldots H_n$ are statements and you know that *exactly one* of them is true, then for any statement E:

$$\Pr(E) = \Pr(E \mid H_1)\Pr(H_1) + \Pr(E \mid H_2)\Pr(H_2) \cdots + \Pr(E \mid H_n)\Pr(H_n).$$

If there exactly two hypotheses H and $\neg H$, then

$$\Pr(E) = \Pr(E \mid H)\Pr(H) + \Pr(E \mid \neg H)\Pr(\neg H).$$

Conditionalization

The Conditionalization Norm: If you have confidence $\Pr(H)$ in a hypothesis H, and then you learn E, then your new confidence in H upon learning E, $Pr_E(H)$, should equal your conditional probability in H given E:

$$Pr_E(H) \overset{\text{should}}{=} Pr(H \mid E).$$

Evidence For and Independence

1. **Definition**:
   (a) E is **evidence for** H if $\Pr(H \mid E) > \Pr(H)$.
   (b) E is **evidence against** H if $\Pr(H \mid E) < \Pr(H)$.
   (c) H is **independent of** E if $\Pr(H \mid E) = \Pr(H)$.
2. **The Evidence-For Lemma**: The following four are equivalent:

$$\Pr(H \mid E) > \Pr(H) \qquad \Pr(E \mid H) > \Pr(E)$$
$$\Pr(H \mid E) > \Pr(H \mid \neg E) \qquad \Pr(E \mid H) > \Pr(E \mid \neg H).$$

They are still equivalent if the ">" are replaced with "<" or "="

LOGIC, REASONING, AND PERSUASION 07; MIDTERM PRACTICE, WEDNESDAY, DECEMBER 3RD

- The midterm will have approximately this format.
- In the actual exam, please write or type your answers *in the bluebook*
- Show your calculations! Otherwise, wrong answers will get no partial credit.

1 | COURSE SURVEYS!

I'm an academic *very* early on in my career (hopefully). If you enjoyed this course or thought it was helpful for your education, I would be very grateful to you for filling out the survey. Surveys can be accessed at sirs.rutgers.edu/blue. This individual survey can be accessed at the QR code on the right.

2 | FORMAL PRACTICE

Suppose you have three hypotheses, H_1, H_2, H_3. You know *exactly one* of those hypotheses is true.

(1.1) *What is $Pr(H_1) + Pr(H_2) + Pr(H_3)$? Do not use any numbers from below.*

Right now, your estimations are that $Pr(H_1) = 0.3$, $Pr(H_2) = 0.3$, $Pr(H_3) = 0.4$. You're wondering whether some statement A is true. You know that

$$Pr(A \mid H_1) = 2/3, Pr(A \mid H_2) = 1/3, Pr(A \mid H_3) = 1/2.$$

By Partitionality (on the sheet), the probability is 1 (even if you know nothing else about the probabilities!)

(1.2) *What is $Pr(A)$? Use the law of total probability.*

By total probability,

$$\begin{aligned} Pr(A) &= Pr(A \mid H_1)Pr(H_1) + Pr(A \mid H_2)Pr(H_2) + Pr(A \mid H_3)Pr(H_3) \\ &= 2/3 \cdot 0.3 + 1/3 \cdot 0.3 + 1/2 \cdot 0.4 = 0.2 + 0.1 + 0.2 = 0.5. \end{aligned}$$

(1.3) Suppose you learn that A is in fact true. If you used conditionalization to update your opinions, what would your new estimates about the hypotheses be? Use Bayes' Theorem.

(a) $Pr_A(H_1) =$

$$\begin{aligned} Pr_A(H_1) &= Pr(H_1 \mid A) && \text{(By Conditionalization)} \\ &= Pr(H_1)\frac{Pr(A \mid H_1)}{Pr(A)} && \text{(By Bayes' Theorem)} \\ &= 0.3 \cdot \frac{2/3}{1/2} = 0.4. \end{aligned}$$

(b) $Pr_A(H_2) =$

$$\begin{aligned} Pr_A(H_2) &= Pr(H_2 \mid A) && \text{(By Conditionalization)} \\ &= Pr(H_2)\frac{Pr(A \mid H_2)}{Pr(A)} && \text{(By Bayes' Theorem)} \\ &= 0.3 \cdot \frac{1/3}{1/2} = 0.2. \end{aligned}$$

(c) $Pr_A(H_3) =$

$$\begin{aligned} Pr_A(H_3) &= Pr(H_3 \mid A) && \text{(By Conditionalization)} \\ &= Pr(H_1)\frac{Pr(A \mid H_3)}{Pr(A)} && \text{(By Bayes' Theorem)} \\ &= 0.4 \cdot \frac{1/2}{1/2} = 0.4. \end{aligned}$$

(1.4) Evidence:

(a) Is A evidence for H_1? By the evidence lemma, A is evidence for H_1 if $Pr(H_1 \mid A) > Pr(H_1)$. Since $Pr(H_1 \mid A) = 0.4$ and $Pr(H_1) = 0.3$, A is evidence for H_1.

(b) Is A evidence for H_2? Since $Pr(H_2 \mid A) = 0.2$ and $Pr(H_2) = 0.3$, A is evidence *against* H_2.

(c) Is A evidence for H_3? Since $Pr(H_3 \mid A) = 0.4$ and $Pr(H_1) = 0.4$, A is independent of H_3.

(d) Is H_1 evidence for A?

(e) Is H_2 evidence for A?

(f) Is H_3 evidence for A?

Notice that by the evidence Lemma, If A is evidence for H then H is evidence for A. So we don't have to do the calculations again. H_1 is evidence for A, H_2 is evidence against A, and H_3 is independent of A.

3 | Nasim's Studying

The probability that Nasim will study for her test is 4/10. The probability that she will pass, given that she studies, is 9/10. The probability that she passes, given that she does not study, is 3/10. Nasim passes the test. What is the probability that she has studied? (Modified slightly from Weisberg, *Odds and Ends*)

Let P be the statement that Nasim passes the test. Let S be the statement that Nasim studies. Then

1. $Pr(S) = 0.4$
2. $Pr(P \mid S) = 0.9$
3. $Pr(P \mid \neg S) = 0.3$

We learn that Nasim has passed the test. So we want to know $Pr_{new}(S) = Pr_P(S)$. If we use conditionalization, then $Pr_P(S) = Pr(S \mid P)$. Then, by Bayes' Theorem,

$$Pr_P(S) = Pr(S \mid P) = Pr(S)\frac{Pr(P \mid S)}{Pr(P)} = 0.4\frac{0.9}{???}$$

The issue is that we don't know the *prior* probability that Nasim would pass, $Pr(P)$. To calculate this, we need total probability:

$$\begin{aligned} Pr(P) &= Pr(P \mid S)Pr(S) + Pr(P \mid \neg S)Pr(\neg S) && \text{(total probability)} \\ &= 0.9 \cdot 0.4 + 0.3 \cdot (1 - 0.4) = 0.54. \end{aligned}$$

So we can plug this value back into Bayes' Theorem:

$$Pr_P(S) = Pr(S \mid P) = Pr(S)\frac{Pr(P \mid S)}{Pr(P)} = 0.4\frac{0.9}{0.54} = 2/3.$$

4 | Brain Scans

Things to Use: *Total Probability, Bayes' Theorem, Conditionalization, Evidence Lemma.* You're a doctor looking at a patient's symptoms. Based on their symptoms, you think there is a 1/3 chance that they have brain cancer. Right now, that's the hypothesis that you think is most likely.

(1.1) *How could you think brain cancer is the most likely hypothesis if you think it is less than 50% likely?*

Here's one way: you think it's 1/3 likely to be brain cancer, 1/6 likely to be a a benign brain tumor, and 1/6 likely to be something else (that you don't know). So none of the hypothese are more than 50% likely, but brain cancer is the most likely one.

The standard thing to do to learn more is to take a brain scan to check for tumors. There is a particular pattern in scans that occurs in 20% of scans of patients who don't have brain cancer and 60% of scans of patients with brain cancer. So the patient is 20% likely to have this pattern if they don't have brain cancer, and 60% likely if they have cancer.

(1.2) *Let C be the statement "the patient has brain cancer". Let P be the statement "the patient's scan has the pattern". Write out all the probabilities you already know based on the problem statement. There should be three.*

1. $Pr(C) = 1/3$
2. $Pr(P \mid C) = 0.6$
3. $Pr(P \mid \neg C) = 0.2$.

(1.3) *You've ordered the test, but haven't seen the results yet. How likely do you think it is that the patient's scan has this pattern?*

$$\begin{aligned} Pr(P) &= Pr(P \mid C)Pr(C) + Pr(P \mid \neg C)Pr(\neg C) && \text{(Total Probability)} \\ &= 0.6 * 1/3 + 0.2 * 2/3 = 1/3. \end{aligned}$$

(1.4) *How likely do you think it is that the patient has brain cancer, supposing that their scan has the pattern?*

$$Pr(C \mid P) = Pr(C)\frac{Pr(P \mid C)}{Pr(P)} = 1/3\frac{0.6}{1/3} = 0.6. \qquad \text{(Bayes)}$$

(1.5) *How likely do you think it is that the patient has brain cancer, supposing that their scan does* ***not*** *have the pattern?*

$$Pr(C \mid \neg P) = Pr(C)\frac{Pr(\neg P \mid C)}{Pr(\neg P)} = 1/3\frac{0.4}{2/3} = 0.2. \qquad \text{(Bayes)}$$

The base rate of brain cancer is around 0.5% (1 in 200).

(1.6) *What is the probability that a random person in the population, if you scanned their brain, would show the pattern?*

Let C_R be the probability that some random person has cancer.

$$\begin{aligned} Pr(P) &= Pr(P \mid C_R)Pr(C) + Pr(P \mid \neg C_R)Pr(\neg C_R) \qquad \text{(Total Probability)} \\ &= 0.6 * 0.005 + 0.2 * 0.995 = 0.202 \approx 20\%. \end{aligned}$$

(1.7) *Medical statistics show that around 35% of brain scans run show this pattern. If you did the math right, your answer to (1.6) is lower than this. How could this be?*

About 35% of people scanned show the pattern. 20.2% of people in the population *would* should the pattern if they got scanned. This must mean that the people who are actually scanned are more likely to show the pattern than a random sample of the population. And this makes sense: people who get these sorts of brain scans to check for brain cancer are typically those whom the doctors suspect *might* have brain cancer. In fact, I calculated the 35% by supposing that, on average, patients whom doctors ordered tests were 40% likely to have brain cancer. Let C_S be the probability that some person who got a brain scan has brain cancer. Then, if we use total probability,

$$\begin{aligned} Pr(P) &= Pr(P \mid C_S)Pr(C_S) + Pr(P \mid \neg C_S)Pr(\neg C_S) \qquad \text{(Total Probability)} \\ &= 0.6 * 0.4 + 0.2 * 0.6 = 0.36 \approx 35\%. \end{aligned}$$

The scan results come back. You don't see a pattern.

(1.8) *Are the scans evidence for or against the hypothesis that the patient has cancer?*

The scans, because they've come back without the pattern are evidence *against* the hypothesis that the patient has cancer. The probability that the patient has cancer, *given* that they scan does not have a pattern, is lower than the prior probability that the patient has cancer.

$$Pr(C \mid \neg P) = 0.2 < 0.333 \approx Pr(C).$$

(1.9) *What should your new estimate be for how likely the patient is to have brain cancer?*
$Pr_{\neg P}(C) = Pr(C \mid \neg P) = 0.2 < 1/3.$

(1.10) *Should you rule out the hypothesis that your patient has brain cancer?*
No. 20% is still far greater than zero.

One of your interns has a great idea. Since getting brain scans without the pattern *lowers* your confidence that the patient has brain cancer, you should just scan the patient again, and again, and again, until your confidence that the patient has brain cancer is very low. Then you can be pretty sure the patient does not have brain cancer.

(1.11) *Is this intern correct? Why or why not? Assume that if you put a patient through a scanner any number of times, the scan will come out the same every time.*

The intern is not correct, because the scans are not independent. After you learn that the patient's first scan did not have the pattern, then your opinions about how likely you are to see a pattern if the patient does / does not have cancer are as follows: $Pr(P \mid C) = Pr(P \mid \neg C) = 0$. That is, no matter whether the patient has cancer or not, you are sure there will be no pattern. Why? Because you already did a scan and saw no pattern, and you know that if you do a scan again, you'll see the same thing.

5 | Bookbags

This is a variant of the problem from Monday used to illustrate Hindsight Bias:

> There are two bookbags, one containing 700 red and 300 blue chips, the other containing 300 red and 700 blue. Take one of the bags. Now, you sample, randomly, with replacement after each chip. You get this sequence:
>
> blue blue red blue red
>
> What is the probability that you chose the bookbag with mostly blue chips (700 blue, 300 red)?

Solve this rigorously. Assume that the samplings are *independent*: each sample does not affect the probabilities of the other samples. This means that you can multiply the probabilities of two sample outcomes to get the probability that both occurred:

> If A and B are independent (remember: this means that $Pr(A \mid B) = Pr(A)$), then $Pr(A\&B) = \Pr(A) \cdot \Pr(B)$.

For example,

$$Pr(\text{1st sample is red AND 2nd sample is blue})$$
$$=$$
$$Pr(\text{1st sample is red}) \cdot Pr(\text{2nd sample is blue}).$$

Solution: (It may be worth it to read the midterm review advice before reading this!) Let S be the statement "you get the sequence blue, blue, red, blue, red", B be the statement "you chose the bag with mostly blue chips," and R be the statement "you chose the bag with mostly red chips".

What is the problem asking? The problem statement says we've sampled and gotten a sequence of five (S), and is asking how likely it is that we chose the bookbag with the mostly blue chips (B). That is, we learn that S is true. So in the setup we learn something, like in this diagram:

$$Pr(B) \longrightarrow \text{learn } S \longrightarrow Pr_S(B).$$

The problem is asking us what $Pr_S(B)$ is. We'll assume that we do conditionalization, so that $Pr_S(B) = Pr(B \mid S)$. So the problem is really asking us what $Pr(B \mid S)$ is.

1. **Setting Up**. Given the setup of the case, we already know the following probabilities:

$$Pr(B) = 0.5 \qquad\qquad Pr(R) = 0.5.$$

We can also solve for the probability that we would get the sequnce S from the mostly-blue bag or the mostly-red bag. Let a number followed by a colon and a color (like 1:blue) mean that that number draw was that color. Since the samples are independent, we have

$$\begin{aligned} Pr(S \mid B) &= Pr(1{:}\text{blue}|B)Pr(2{:}\text{blue}|B)Pr(3{:}\text{red}|B)Pr(4{:}\text{blue}|B)Pr(5{:}\text{red}|B) \\ &= 0.7 \cdot 0.7 \cdot 0.3 \cdot 0.7 \cdot 0.3 = 0.03087 \\ Pr(S \mid R) &= Pr(1{:}\text{blue}|R)Pr(2{:}\text{blue}|R)Pr(3{:}\text{red}|R)Pr(4{:}\text{blue}|R)Pr(5{:}\text{red}|R) \\ &= 0.3 \cdot 0.3 \cdot 0.7 \cdot 0.3 \cdot 0.7 = 0.01323. \end{aligned}$$

So we have

$$\begin{aligned} Pr(B) &= 0.5 & Pr(R) &= 0.5. \\ Pr(S \mid B) &= 0.03087 & Pr(S \mid R) &= 0.01323. \end{aligned}$$

2. **Solving for** $Pr(B \mid S)$. Should we use Bayes' Theorem or Total Probability to figure out $Pr(B \mid S)$? Because this is a *conditional* probability, with a bar in the middle, we should use **Bayes' Theorem**.

We write out Bayes' Theorem with the right statement letters:

$$Pr(B \mid S) = Pr(B)\frac{Pr(S \mid B)}{Pr(S)}.$$

Now, we already know (from our list above) that $Pr(B) = 0.5$ and $Pr(S \mid B) = 0.01323$. So we can substitute these in:

$$Pr(B \mid S) = 0.5\frac{0.01323}{Pr(S)}.$$

But we don't know what $Pr(S)$ is, so we'll have to calculate it.

3. **Finding** $Pr(B \mid S)$. Should we use Bayes' Theorem or Total Probability to figure out $Pr(B \mid S)$? Because this is a *unconditional* probability, with *no* bar in the middle, we should use **Total Probability**.

What are the two ways in which S could be true? Well, we could either have gotten the sequence from the mostly blue bag or from the mostly red bag. So the two hypotheses, exactly one of which is true, are B and R. Let's write out total probability with the right statement letters:

$$Pr(S) = Pr(S \mid B)Pr(B) + Pr(S \mid R)Pr(R).$$

And we see from above that we know all these values already, so let's substitute them in:

$$Pr(S) = 0.03087 \cdot 0.5 + 0.01323 \cdot 0.5 = 0.02205.$$

4. **Finding $Pr(B \mid S)$, resumed**: Now that we have $Pr(S) = 0.02205$ we can substitute it back into Bayes' Theorem:

$$Pr(B \mid S) = 0.5\frac{0.03087}{0.02205} = 0.7.$$

So our final answer is $Pr(B \mid S) = 0.7$.

6 | Birthdays

[I won't have anything like Birthdays on the exam!]

7 | Chocolates

Willy Wonka Chocolates Inc. makes two kinds of boxes of chocolates. The "wonk box" has four caramel chocolates and six regular chocolates. The "zonk box" has six caramel chocolates, two regular chocolates, and two mint chocolates. A third of their boxes are wonk boxes, the rest are zonk boxes.

They don't mark the boxes. The only way to tell what kind of box you've bought is by trying the chocolates inside. In fact, all the chocolates look the same; you can only tell the difference by tasting them.

If you buy a random box, try a chocolate at random, and find that it's caramel, what is the probability you've bought a wonk box? (Weisberg, *Odds and Ends*)

Solution: see Midterm Review Advice!

8 | For Fun: Newcomb's Problem

[I won't have anything like Newcomb on the exam, but I may discuss on the last day!]

Here is a game. You walk into a room, and there is a clear box and an opaque box. You see that the clear box has $100. You can't see into the opaque box. Your choices are

1. Take *only* the opaque box.
2. Take *both* the opaque box and the clear box.

A predicting machine has predicted whether you will take the clear box (option 2).

1. If the predictor predicted you would **not** take the clear box (option 1), then $10,000 was placed in the opaque box.
2. Otherwise, nothing was placed in the opaque box.

We don't know how, but the predictor is 99% accurate at predicting. This game has been run a bunch of times on a bunch of people, and when people took both boxes, the predictor predicted they would 99% of the time, and when people took only the opaque box, the predictor *also* predicted they would 99% of the time.

This means that if you choose option 1, there's a 99% chance that the predictor predicted you would, and there's $10,000 in the opaque box. And if you choose option 2, there's a 99% chance the predictor predicted you would, and there's nothing in the opaque box.

What should you do? Note: *most* philosophers agree on an answer to this question, but it is by no means controversial. This is not a problem where you can solve it just by rote application of the methods we've learned in this course. But you certainly have the resources to ponder it and decide what you would do!