Home Page

Title Page

◀◀   ▶▶

◀   ▶

Go Back

Full Screen

Close

Quit

Supplementary Notes on Bayes' Theorem

1. Bayes Theorem

Estimate the probabilities in the following two examples.

Example One: Exactly two cab companies operate in Belleville. The Blue Company has blue cabs, and the Green Company has Green Cabs. Exactly 85% of the cabs are blue and the other 15% are green. A cab was involved in a hit-and-run accident at night. A witness, Wilbur, identified the cab as a Green cab. Careful tests were done to ascertain peoples' ability to distinguish between blue and green cabs at night. The tests showed that people were able to identify the color correctly 80% of the time, but they were wrong 20% of the time. What is the probability that the cab involved in the accident was indeed a green cab, as Wilbur says?

Example Two: Suppose that we have a test for the HIV virus. The probability that a person who really has the virus will test positive is .90 while the probability that a person who does *not* have it will test positive is .20. Finally, suppose that the probability that a person in the general population has the virus (the base rate of the

virus) is .01. Smith's test came back positive. What is the probability that he really is infected?

Problems like these require us to *integrate* knowledge about the present case (e.g., what the eyewitness says, how Smith's medical test came out) with prior information about base rates (what proportion of the cabs are green, what proportion of people have the HIV virus). In many cases we focus too much on the present information and ignore information about base rates. The correct way to give both pieces of information their proper due is to use Bayes' Rule.

1.1. Bayes' Rule

Bayes' Rule (named after the Reverend Thomas Bayes, who discovered it in the eighteenth century) is really just another rule for calculating probabilities. It tells us how we should modify or update probabilities when we acquire new information. It gives the *posterior probability* of a hypothesis A in light of a piece of new evidence B as

$$(1) \qquad \Pr(A \,|\, B) = \frac{\Pr(A) \times \Pr(B \,|\, A)}{\Pr(B)}$$

We say that $\Pr(A)$ is the *prior probability* of the hypothesis, $\Pr(B \,|\, A)$ the *likelihood* of B given A, and $\Pr(B)$ the *prior probability* of evidence B. For example, A might be the hypothesis that Smith is infected by the HIV virus and B might be the datum (piece of evidence) that his test came back positive.

By way of illustration let's work through example one above (we will work through example two below). This example involved the Blue Company, which has

all blue cabs, and the Green Company, which has all green cabs. Exactly 85% of the cabs are blue and the other 15% are green. A cab was involved in a hit-and-run accident at night. An honest eyewitness, Wilbur, identified the cab as a green cab. Careful tests were done to ascertain witness' ability to distinguish between blue and green cabs at night; these showed that people were able to identify the color correctly 80% of the time, but they were wrong 20% of the time.

What is the probability that the cab involved in the accident was indeed a green cab? The fact that witnesses are pretty reliable leads many people to suppose that Wilbur probably right, even that the probability that he is right is .8. But is this correct?

A large part of the battle is often setting up clear notation, so let's begin with that:

- $\Pr(G) = .15$ This is the *base rate* of green cabs in the city. It gives the *prior probability* that the cab in the accident is green. Similarly, $\Pr(B) = .85$.

- $\Pr(SG \mid G) = .80$ This is the probability the witness will be correct in *s*aying *g*reen when the cab in fact is *g*reen, i.e., given that the cab really was *g*reen. Similarly, $\Pr(SB \mid B) = .80$.

    These are the probabilities that witnesses are correct, so by the negation rule, the probabilities of misidentifications are:

- $\Pr(SG \mid B) = .2$ and $\Pr(SB \mid G) = .2$

What we want to know is the probability that the cab really was *g*reen, given that Wilbur *s*aid it was *g*reen, i.e., we want to know $\Pr(G \mid SG)$.

According to Bayes' Rule, this probability is given by:

$$\Pr(G\,|\,SG) = \frac{\Pr(G) \times \Pr(SG\,|\,G)}{\Pr(SG)}$$

We have the values for the two expressions in the numerator—$\Pr(G) = .15$ and $\Pr(SG\,|\,G) = .8$—but we have to do a little work to determine the value for the expression—$\Pr(SG)$—in the denominator.

To see what this value—the probability that a witness will say that the cab was green—must be, note that there are two conditions under which Wilbur could say that the cab in the accident was green. He might say this when the cab was green *or* when it was blue. This is a disjunction, so we add these two probabilities. In the first disjunct Wilbur says green *and* the cab is green; in the second disjunct Wilbur says green *and* the cab is blue. Putting this together we get:

$$\Pr(SG) = \Pr(G \wedge SG) + \Pr(B \wedge SG)$$

Now our rule for conjunctions tells us that $\Pr(G \wedge SG) = \Pr(G) \times \Pr(SG\,|\,G)$ and $\Pr(B \wedge SG) = \Pr(B) \times \Pr(SG\,|\,B)$. So

$$
\begin{aligned}
\Pr(SG) &= \Pr(G) \times \Pr(SG\,|\,G) + \Pr(B) \times \Pr(SG\,|\,B) \\
&= (.15 \times .80) + (.85 \times .20) \\
&= .12 + .17 \\
&= .29
\end{aligned}
$$

Home Page

Title Page

◀◀   ▶▶

◀   ▶

Go Back

Full Screen

Close

Quit

Finally, we substitute this number, .29, into the denominator of Bayes' Rule:

$$
\begin{aligned}
\Pr(G\,|\,SG) &= \frac{\Pr(G) \times \Pr(SG\,|\,G)}{\Pr(SG} \\
&= \frac{.15 \times .80}{.29} \\
&= .414
\end{aligned}
$$

So the probability that the witness was correct in saying the cab was green is just a bit above .4—definitely less than fifty/fifty—and (by the negation rule) the probability that he is wrong is nearly .6. This is so even though witnesses are pretty reliable. How can this be? The answer is that the high *base rate* of blue cabs and the low *base rate* of green cabs makes it somewhat likely that the witness was wrong in this case.

1.2. Rule for Total Probabilities

In calculating the value of the denominator in the above problem we made use of the *Rule for Total Probabilities*. We will use a fairly simple, but widely applicable version of this rule here. If a sentence B is true, then either B and A are true or else B and $\neg A$ are true (since any sentence A is either true or false). This allows us to express the probability of B—$\Pr(B)$—in a more complicated way. This may seem like a strange thing to do, but it turns out that it is often possible to calculate the probability of the more complicated expression when it isn't possible to obtain the

probability of B more directly.

$$
\begin{aligned}
\Pr(B) &= \Pr[(B \wedge A) \vee (B \wedge \neg A)] \\
&= \Pr(B \wedge A) + \Pr(B \wedge \neg A) \\
&= [\Pr(A) \times \Pr(B|A)] + [\Pr(\neg A) \times \Pr(B|\neg A)]
\end{aligned}
$$

In short

$$
\Pr(B) = [\Pr(A) \times \Pr(B|A)] + [\Pr(\neg A) \times \Pr(B|\neg A)].
$$

This rule can be useful in cases that do not involve Bayes' Rule, as well as in many cases that do. It is particularly useful when we deal with an outcome that can occur in either of two ways.

Example: A factory has two machines that make widgets. Machine A makes 800 per day and 1% of them are defective. The other machine, call it $\neg A$, makes 200 a day and 2% are defective. What is the probability that a widget produced by the factory will be defective (D)? We know the following:

- $\Pr(A) = .8$ – the probability that a widget is made by machine A is .8 (since this machine makes 800 out of the 1000 produced everyday)
- $\Pr(\neg A) = .2$ – the probability that a widget is made by the other machine $\neg A$ is .2
- $\Pr(D|A) = .01$ – the probability that a widget is defective given that it was made by machine A, which turns out 1% defective widgets, is .01p
- $\Pr(D|\neg A) = .02$ – the probability that a widget is defective given that it was made by machine $\neg A$, which turns out 2% defective widgets, is .02p

Home Page

Title Page

◀◀   ▶▶

◀   ▶

Go Back

Full Screen

Close

Quit

Plugging these numbers into the theorem for total probability we have:

$$
\begin{aligned}
\Pr(D) &= [\Pr(A) \times \Pr(D\,|\,A)] + [\Pr(\neg A) \times \Pr(D\,|\,\neg A)] \\
&= [.8 \times .01] + [.2 \times .02] \\
&= 0.012
\end{aligned}
$$

2. Odds Version of Bayes' Rule

Particularly when we are concerned with mutually exclusive and disjoint hypothesis (cancer vs. no cancer), the odds version of Bayes' rule is often the most useful version. It also allows avoid the need for a value for $\Pr(e)$, which is often difficult to come by. Let H be some hypothesis under consideration and e be some newly acquired piece of evidence (like the result of a medical test) that bears on it. Then

$$
(2) \qquad \frac{\Pr(H\,|\,e)}{\Pr(\neg H\,|\,e)} = \frac{\Pr(H)}{\Pr(\neg H)} \frac{\Pr(e\,|\,H)}{\Pr(e\,|\,\neg H)}
$$

The expression on the left of (2) gives the posterior odds of the two hypotheses, $\Pr(H)/\Pr(\neg H)$ gives their prior odds, and $\Pr(e\,|\,H)/\Pr(e\,|\,\neg H)$ is the likelihood (or diagnostic) ratio. In integrating information our background knowledge of base rates is reflected in the prior odds, and our knowledge about the specific case is represented by the likelihood ratio. For example, if e represents a positive result in a test for the presence of the HIV virus, then $\Pr(e\,|\,H)$ gives the *hit rate* (*sensitivity*) of the test and $\Pr(e\,|\,\neg H)$ gives the *false alarm rate*. As the multiplicative relation in (2) indicates,

Home Page

Title Page

◀◀ ▶▶

◀ ▶

Go Back

Full Screen

Close

Quit

when the prior odds are quite low even a relatively high likelihood ratio won't raise the posterior odds dramatically.

The importance of integrating base rates into our inferences is vividly illustrated by example two above. Suppose that we have a test for the HIV virus. The probability that a person who really has the virus will test positive is .90 while the probability that a person who does *not* have it will test positive is .20. Finally, suppose that the probability that a person in the general population has the virus (the base rate of the virus) is .01. How likely is Smith, whose test came back positive, to be infected?

Because the test is tolerably accurate, many people suppose that the chances are quite high that Smith is infected, perhaps even as high as 90%. Indeed, various studies have shown that many physicians even suppose this. It has been found that people tend to confuse probabilities and their converses. The probability that we'll get the positive test, e, if the person has the virus is .9, i.e., $\Pr(e\,|\,H) = .9$. But we want to know the probability of the converse, the probability that the person has the virus given that they tested positive, i.e., $\Pr(H\,|\,e)$. These need not be the same, or even close to the same. In fact, when we plug the numbers into Bayes' Rule, we find that the probability that someone in this situation is infected is actually quite low.

To see this we simply plug our numbers into the odds version of Bayes' rule. We have been told that

- $\Pr(e\,|\,H) = .9$
- $\Pr(e\,|\,\neg H) = .2$
- $\Pr(H) = .01$
- $\Pr(\neg H) = .99$.

Hence

$$\frac{\Pr(H\,|\,e)}{\Pr(\neg H\,|\,e)} = \frac{\Pr(H)}{\Pr\neg H} \times \frac{\Pr(e\,|\,H)}{\Pr(e\,|\,\neg H)} = \frac{.9}{.2} \times \frac{.01}{.99} = \frac{1}{22}$$

This gives the odds: 22 to 1 against Smith's being infected. So $\Pr(H\,|\,e) = \frac{1}{23}$. Although the test is fairly accurate, the low base rate means that there is only 1 chance in 23 that Smith has the virus.

Similar points apply to other medical tests and to drug tests, as long as the base rate of the condition being tested for is low. It is not difficult to see how policy makers, or the public to which they respond, can make very inaccurate assessments of probabilities, and hence poor decisions about risks and remedies, if they overlook the importance of base rates.

As we noted in Chapter 16, it will help you to think about these matters intuitively if you try to rephrase probability problems in terms of *frequencies* or *proportions* whenever possible.

2.1. Further Matters

Bayes' Rule makes it clear when a conditional probability will be equal to it's converse. When we look at

$$\Pr(A\,|\,B) = \frac{\Pr(A) \times \Pr(B\,|\,A)}{\Pr(B)}$$

We see that $\Pr(A\,|\,B) = \Pr(B\,|\,A)$ exactly when $\Pr(A) = \Pr(B)$, so that they cancel out (assuming, as always, that we do not divide by zero). Dividing both sides of Bayes' Rule by $\Pr(B\,|\,A)$ gives us the following ratio

$$\frac{\Pr(A\,|\,B)}{\Pr(B\,|\,A)} = \frac{\Pr(A)}{\Pr(B)}$$

which is often useful.

2.1.1. Updating via Conditionalization

One way to update or revise our beliefs in light of the new evidence *e* is to set our new probability (once we learn about the evidence) as:

$$New \Pr(H) = Old \Pr(H\,|\,e),$$

with *Old* $\Pr(H\,|\,e)$ determined by Bayes' Rule. Such updating is said to involve *Bayesian conditionalization*.

3. Reference Classes

There are genuine difficulties in selecting the right group to consider when we use base rates. This is known as the *problem of the reference class*. Suppose we are considering the effects of smoking on lung cancer. *Which* reference class should we consider: all people, just smokers, just heavy smokers?

Similar problems arise when we think about regression to the mean. We know that extreme performances and outcomes tend to be followed by ones that are more average (that regress to the mean). If Wilma hits 86% of her shots in a basketball

game, she is likely to hit a lower percentage the next time out. People do not seem to have an intuitive understanding of this phenomenon, and the difficulty is compounded by the fact that, as with base rates, there is a problem about reference class. *Which* average will scores regress to: Wilma's average, her recent average, her team's average?

Up to a point smaller reference classes are likely to underwrite better predictions, and we will also be interested in reference classes that seem causally relevant to the characteristic we care about. But if a reference class becomes too small, it won't generate stable frequencies, and it will often be harder to find data on smaller classes that is tailored to our particular concerns.

There is no one right answer about which reference class is the correct one. It often requires a judicious balancing of tradeoffs. With base rates it may be a matter of weighing the costs of a wrong prediction against the payoffs of extreme accuracy, or the value of more precise information against the costs of acquiring it or (if you are a policy maker) the fact that the number of reference classes grows exponentially with each new predictor variable. But while there is no uniquely correct way to use base rates, it doesn't follow that it is fine to ignore them; failing to take base-rate information into account when we have it will often lead to bad policy.

4. Exercises

1. Above (6) we encountered the factory with two machines that make widgets. Machine *A* makes 800 per day and 1% of them are defective. The other machine, call it ¬*A*, makes 200 a day and 2% are defective. What is the probability that a widget

is produced by machine A, given that it is defective? We know the probability that it is defective if it is produced by A—$\Pr(D|A) = .01$—but this asks for the opposite or *converse probability*; what is $\Pr(A|D)$? Use Bayes' Rule to calculate the answer.

2. Officials at the suicide prevention center know that 2% of all people who phone their hotline actually attempt suicide. A psychologist has devised a quick and simple verbal test to help identify those callers who will actually attempt suicide. She found that

1. 80% of the people who will attempt suicide have a positive score on this test.
2. but only 5% of those who will not attempt suicide have a positive score on this test.

If you get a positive identification from a caller on this test, what is the probability that he would actually attempt suicide?

3. A clinical test, designed to diagnose a specific illness, comes out positive for a certain patient. We are told that:

1. The test is 79% accurate: the chances that you have the illness if it says you do is 79% and the chances that you do not have the illness if it says you don't is also 79%.
2. This illness affects 1% of the population in the same age group as the patient.

Taking these two facts into account, and assuming that you know nothing about the patient's symptoms or signs, what is the probability that this particular patient actually has the illness?

Home Page

Title Page

◀◀ ▶▶

◀ ▶

Go Back

Full Screen

Close

Quit

4. **Conservatism** Suppose that you handed two bags of poker chips, but from the outside you can't tell which is which.

1. **Bag 1**: 70 red chips and 30 blue chips
2. **Bag 2**: 30 red chips and 70 blue chips.

You pick one of the two bags at random and draw a chip from it. The chip is red. You replace it, draw again, and get another chip, and so on through twelve trials (i.e., twelve draws). In the twelve trails you get 8 red chips and 4 blue chips. What is the probability that you have been drawing chips from **Bag One** (with 70 red and 30 blue) rather than from **Bag Two** (with 30 red and 70 blue)? Most people answer that the probability that you have been drawing from **Bag One** is around .75. In fact, as Bayes' Rule will show, it is .97. But people often revise their probabilities *less* that Bayes' Rule says they should. People are **conservative** when it comes to updating their probabilities in the light of new evidence. Use Bayes' Rule to show that .97 is the correct value.

5. **Monte Hall Problem** In an earlier chapter we approached the Monty Hall Problem in an intuitive way. Now we will now verify our earlier answers using Bayes' Rule. Recall that in this problem you imagine that you are a contestant in a game show and there are three doors in front of you. There is nothing worth having behind two of them, but there is $50,000 behind the third. If you pick the correct door, the money is yours.

You choose door A. But before the host, Monte Hall, shows you what is behind that door, he opens one of the other two doors, picking one he knows has nothing

behind it. Suppose he opens door B. This takes B out of the running, so the only question now is about door A vs. door C.

Monte ('M', for short) now allows you to reconsider your earlier choice: you can either stick with door A or switch to door C. Should you switch?

1. What is the probability that the money is behind door A?
2. What is the probability that the money is behind door C?

Since we are asking these questions once Monte has opened door B, they are equivalent to asking about:

1′ Pr($ behind A | M opened B)
2′ Pr($ behind C | M opened B)

Bayes' Rule tells us that Pr(Money behind A | M opened B) =

$$\frac{\text{Pr(\$ behind A)} \times \text{Pr(M opened B | \$ behind A)}}{\text{Pr(M opens B)}}$$

We know the values for the two items in the numerator:

1. Pr($ behind A): the prior probability that the money is behind door A is $\frac{1}{3}$ (going in, it could equally well be behind any of the three doors).
2. Pr(M opened B | $ behind A) is $\frac{1}{2}$ (there is a fifty/fifty chance that he would open either door B *or* door C, when the money is behind door A).
3. But Pr(M opens B), the number in the denominator, requires some work.

To see the value of the denominator—Pr(M opens B)—note that Monte will *never* open door B if the money is behind B. Hence, he could open B under exactly two conditions. He could open B when the money is behind A or he could open B when the money is behind C. So

$$
\begin{aligned}
\text{Pr(M opens B)} &= [\text{Pr(\$ behind A)} \times \text{Pr(M opens B} \,|\, \text{\$ behind A)}] \\
&+ [\text{Pr(\$ behind C)} \times \text{Pr(M opens B} \,|\, \text{\$ behind C)}] \\
&= [1/3 \times 1/2] + [1/3 \times 1] \\
&= [1/6] + [1/3] \\
&= 3/6 \\
&= 1/2
\end{aligned}
$$

Plugging this into the denominator in Bayes' Rule, we get:

$$
\begin{aligned}
\text{Pr(\$ behind A} \,|\, \text{M opened B)} &= \frac{\text{Pr(\$ behind A)} \times \text{Pr(M opened B} \,|\, \text{\$ behind A)}}{\text{Pr(M opens B)}} \\
&= \frac{1/3 \times 1/2}{1/2} \\
&= \frac{1/6}{1/2} \\
&= 1/3
\end{aligned}
$$

So the probability that the money is behind your door, door A, given that Monte has opened door B, is $\frac{1}{3}$. Since the only other door left is door C, the probability that the

money is there, given that Monte opened door B, should be $\frac{2}{3}$. Use Bayes' Rule to prove that this is correct.

4.0.2. Answers to Selected Exercises

1. We are given all of the relevant numbers except that for $\Pr(D)$, but we calculated this in the section on total probabilities. By Bayes' Rule:

$$
\begin{aligned}
\Pr(A\,|\,D) &= \frac{\Pr(A) \times \Pr(D\,|\,A)}{\Pr(D)} \\
&= \frac{.8 \times .01}{.012} \\
&= \frac{2}{3}
\end{aligned}
$$

4.1. Derivation of Bayes' Rule

It is straightforward to derive Bayes' Rule from our rules for calculating probabilities. You don't need to worry about the derivation, but here it is for anyone who likes such things.

$$
\begin{aligned}
\Pr(A\,|\,B) &= \frac{\Pr(A \wedge B)}{\Pr(B)} \\
&= \frac{\Pr(A) \times \Pr(B\,|\,A)}{\Pr(B)}
\end{aligned}
$$

In actual applications we often don't have direct knowledge of the value of the denominator, $\Pr(B)$, but in many cases we do have enough information to calculate it using the Rule for Total Probability. This tells us that:

$$
\begin{aligned}
\Pr(B) &= \Pr(B \wedge A) + \Pr(B \wedge \neg A) \\
&= [\Pr(A) \times \Pr(B \mid A)] + [\Pr(\neg A) \times \Pr(B \mid \neg A)]
\end{aligned}
$$

So the most useful version of Bayes' Rule is often

$$
\Pr(A \mid B) = \frac{\Pr(A) \times \Pr(B \mid A)}{[\Pr(A) \times \Pr(B \mid A)] + [\Pr(\neg A) \times \Pr(B \mid \neg A)]}
$$

Bayes' Rule can take more complex forms, but this is as far as we will take it in this book.[1]

---

[1]Many of the ideas and experimental results described here come from an extremely important series of papers by Amos Tversky and Daniel Kahneman. Many of these are collected in Kahneman, Slovic and Amos Tversky (1982); D. M. Eddy, "Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities," in the book edited by Kahneman, Slovic & Tversky documents our strong tendency to confuse conditional probabilities with their converses (also called their *inverses*). Good additional discussion may be found in Nisbett and Ross (1980). More recent treatments will be found in [to be supplied].

Home Page

Title Page

◀◀   ▶▶

◀   ▶

Go Back

Full Screen

Close

Quit

References

Kahneman, Daniel, Slovic, P., and Tversky, Amos. 1982. *Judgments Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Nisbett, Richard and Ross, Lee. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, N.J.: Prentice-Hall.