Probability and Inductive Logic

Kenny Easwaran

March 23, 2018

1 Motivation for Inductive Logic

As discussed in the textbook, an argument is a set of sentences, one of which is designated as the conclusion, and the others are premises. The central point of logic is to assess arguments, and the way that the truth of the premises of an argument affects the truth of the conclusion. The strongest type of argument is a *valid* argument, where it is *impossible* for the premises to be true while the conclusion is false. Valid arguments are the subject matter for *deductive* logic.

The topic of this reading is *inductive* logic, which is the assessment of the strength of arguments that are *not* valid. Consider the following arguments.

First argument:

> Juan is either at home or at his second class today. Juan did not go to both his first two classes today. Therefore Juan is either at home, or did not go to his first class today.

If we let H stand for "Juan is at home", F stand for "Juan went to his first class today", and S stand for "Juan is at his second class today", then the following truth table shows that the argument is valid:

F	H	S	H	∨	S	¬	(F	∧	S)	H	∨	¬	F
T	T	T	T	T	T	F	T	T	T	T	T	F	T
T	T	F	T	T	F	T	T	F	F	T	T	F	T
T	F	T	F	T	T	F	T	T	T	F	F	F	T
T	F	F	F	F	F	T	T	F	F	F	F	F	T
F	T	T	T	T	T	T	F	F	T	T	T	T	F
F	T	F	T	T	F	T	F	F	F	T	T	T	F
F	F	T	F	T	T	T	F	F	T	F	T	T	F
F	F	F	F	F	F	T	F	F	F	F	T	T	F

As you can see, there is no row where both premises are true and the conclusion is false.

Second argument:

> It is either rainy or snowy today. It is not rainy and hot today. Therefore it is not hot today.

If we let R stand for "It is rainy today", S stand for "It is snowy today", and H stand for "It is hot today", then we get the following truth table:

H	R	S	R	∨	S	¬	(R	∧	H)	¬	H
T	T	T	T	T	T	F	T	T	T	F	T
T	T	F	T	T	F	F	T	T	T	F	T
T	F	T	F	T	T	T	F	F	T	F	T
T	F	F	F	F	F	T	F	F	T	F	T
F	T	T	T	T	T	T	T	F	F	T	F
F	T	F	T	T	F	T	T	F	F	T	F
F	F	T	F	T	T	T	F	F	F	T	F
F	F	F	F	F	F	T	F	F	F	T	F

As you can see, the third row is one where the premises are true and the conclusion is false. So the truth table doesn't, by itself, tell us that the argument is valid. However, in this third row, we see that S and H are both true.

However, once we bring back the meanings of the sentences, we see that this is impossible — it can't be both snowy and hot at the same time. Thus, there is no *possible* way for the premises to be true and the conclusion false. So the argument is valid, even though the truth table didn't show us by itself.

Third argument:

> It is either windy or rainy today. It is not windy and sunny today. Therefore it is not sunny today.

If we let W stand for "It is windy today", R stand for "It is rainy today", and S stand for "It is sunny today", then we get the following truth table:

R	S	W	W	∨	R	¬	(W	∧	S)	¬	S
T	T	T	T	T	T	F	T	T	T	F	T
T	T	F	F	T	T	T	F	F	T	F	T
T	F	T	T	T	T	T	T	F	F	T	F
T	F	F	F	T	T	T	F	F	F	T	F
F	T	T	T	T	F	F	T	T	T	F	T
F	T	F	F	F	F	T	F	F	T	F	T
F	F	T	T	T	F	T	T	F	F	T	F
F	F	F	F	F	F	T	F	F	F	T	F

Again, there is a row where the premises are both true and the conclusion is false (in this case, the second row). If we look at the meanings of the atomic sentences, we see that this situation would have to be one where it is rainy and sunny, but not windy.

Unlike with the previous argument, this scenario is not impossible. Thus, we have to say that this argument is not valid. However, this scenario is still pretty unlikely. If you learned these two premises, you would have a fairly good reason for believing the conclusion, even though they don't guarantee that the conclusion is true. Thus, this argument is fairly strong, even though our deductive techniques can't yet tell us that.

Fourth argument:

> It is either windy or rainy today. It is either windy or sunny today. Therefore it is sunny today.

Using the same abbreviations as above, we get:

R	S	W	W	∨	R	W	∨	S	S
T	T	T	T	T	T	T	T	T	T
T	T	F	F	T	T	F	T	T	T
T	F	T	T	T	T	T	T	F	F
T	F	F	F	T	T	F	F	F	F
F	T	T	T	T	F	T	T	T	T
F	T	F	F	F	F	F	T	T	T
F	F	T	T	T	F	T	T	F	F
F	F	F	F	F	F	F	F	F	F

In this case, there are five rows where the premises are both true, and three of them make the conclusion true, while two make the conclusion false. Furthermore, the two cases where the premises are true and the conclusion is false are ones that are perfectly reasonable — rainy and windy without being sunny, or windy while neither rainy nor sunny. Thus, the premises of this argument don't give you any particularly good reason to believe the conclusion. It's a weak argument.

The goal of this reading is to develop some basic techniques for quantifying how strong or weak an argument is. The central idea is that we want to see how "close" the argument is to being valid. In a valid argument, once you assume the premises are true, you can *guarantee* that the conclusion is true. In a strong argument, once you assume the premises are true, the conclusion is *probably* true. In a weak argument, assuming the premises doesn't particularly make the conclusion more probable at all. The tool we will use to measure this sort of "closeness" is *probability* — we assign each possibility a number on a scale from 0 to 1 (which you might sometimes think of as a scale from 0% to 100%) of how likely it is, and see whether and how much the numbers for the conclusion change when you take the premises into account.

2 Probability

There are several ways to think about what probability is. But the two main ways to think of it are in terms of what fraction of possibilities come out a certain way, or how much you should be willing to pay for a bet where you win $1 if the truth comes out a certain way. For instance, you might have probabilities like the following:

Probability	Rainy	Sunny
.05	T	T
.25	T	F
.3	F	T
.4	F	F

One way to think about what this means is that out of the past 100 days, 5 of them were both rainy and sunny, 25 of them were rainy and not sunny, 30 of them were sunny and not rainy, and 40 of them were neither rainy nor sunny.

The probability of any sentence ends up being the sum of the probabilities of the individual rows in which it is true. $5+25=30$ of the days were rainy and $30+40=70$ weren't, and 35 of the days were sunny but 65 weren't. If you made a bet on rain every day, you would win $30, so you would have made money if the price of each bet was less than 30 cents, and you would have lost money if the price was more than 30 cents, and you would break even if the price was exactly 30 cents. So if we think of each day as a possibility, then there is a connection between this "fraction of possibilities" idea and the "betting price" idea.

Of course, we also know that the weather changes from day to day, and some days have a high chance of rain and others have a low chance of rain. So looking at the last 100 days isn't necessarily what we mean when we talk about probability. An equally good way to think about it is that there are 100 different ways that tomorrow's weather might go, and out of those 100 different ways, 5 are rainy and sunny, 25 are rainy and not sunny, 30 are sunny and not rainy, and 40 are neither rainy nor sunny. In this case these possibilities aren't actual days, but are either parallel universes that all branch off from the present, or are just scenarios in your mind that you are imagining. The connection to the betting price is not quite as close, but if you imagine you and your counterparts in these parallel scenarios all betting as a team, again the probability corresponds to the price where your team breaks even.

This is just a very brief introduction to the meaning of the concept of probability, and there are a lot of important issues that go much further. If you want to read more, check out the Stanford Encyclopedia of Philosophy entry on "Interpretations of Probability": `https://plato.stanford.edu/entries/probability-interpret/`

The important points about probability are the following. Each possibility has some number, representing something like the fraction of ways it could be true, or the fraction of the winnings that would be fair to pay for a bet. If you want to know the probability of any sentence, you just add up the numbers on all the rows of the truth table where that sentence comes out true. Because the numbers represent fractions, if you add them all together, you should get 1 (or 100%).

3 Conditional probability

The thing that makes probability most interesting, and most relevant for logic, is the way that probabilities change when more information is given. The basic assumption in most work on probabilistic learning is that information only ever works by *eliminating* possibilities and leaving others behind. As Sherlock Holmes says, "Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth." (Holmes calls his method "deduction", but in the terminology favored by most logicians, most of what he does is what we would call "induction".)

Since probabilities represent the fraction of ways in which a possibility is true, we said they always have to add up to 1. Thus, the way we update the probabilities in a table is by first eliminating the rows in which the information we gained is false, and then figuring out what fraction of what is left is represented by each row. That is, we divide the number in each row into the sum of all remaining rows, and the result is our new probabilities.

Let's look at an example of what happens in one case. Imagine that we start with this set of probabilities (for now, we will always just take the starting probabilities in the rows as given, which we call the "prior probabilities" — later we will discuss some ways of calculating them in particular circumstances).

Probability	A	B	C
.15	T	T	T
.25	T	T	F
.1	T	F	T
.05	T	F	F
.2	F	T	T
.05	F	T	F
.15	F	F	T
.05	F	F	F

Now imagine that we learn that (A∧(B∨C)) is true:

Probability	A	B	C	A	∧ (	B	∨	C)
.15	T	T	T	T	T	T	T	T
.25	T	T	F	T	T	T	T	F
.1	T	F	T	T	T	F	T	T
.05	T	F	F	T	F	F	F	F
.2	F	T	T	F	F	T	T	T
.05	F	T	F	F	F	T	T	F
.15	F	F	T	F	F	F	T	T
.05	F	F	F	F	F	F	F	F

Since the information we gained is only true in the first three rows, we eliminate all the rest:

Probability	A	B	C	A	∧ (	B	∨	C)
.15	T	T	T	T	T	T	T	T
.25	T	T	F	T	T	T	T	F
.1	T	F	T	T	T	F	T	T

These probabilities don't add up to 1, so we have to look at the new total ($.15 + .25 + .1 = .5$) and divide by that to get new values, called the "posterior probabilities". $.15/.5 = .3$, $.25/.5 = .5$, and $.1/.5 = .2$. Thus, we end up with:

Probability	A	B	C	A	∧	(B	∨	C)
.3	T	T	T	T	T	T	T	T
.5	T	T	F	T	T	T	T	F
.2	T	F	T	T	T	F	T	T

Thus, if we are given as information the sentence (A∧(B∨C)) with the prior probabilities above, then under the posterior probabilities, the probability of C is $.3 + .2 = .5$, the probability of B is $.3 + .5 = .8$, and the probability of A is $.3 + .5 + .2 = 1$.

If we want to know how strong various arguments are, then we need to compare these posterior probabilities to the prior probabilities, to see whether the information that was given is actually responsible for the probability being high or not. Looking at the original chart, we see that the prior probability of A was $.15+.25+.1+.05 = .55$, the prior probability of B was $.15+.25+.2+.05 = .65$, and the prior probability of C was $.15 + .1 + .2 + .15 = .6$. Thus, the information that (A∧(B∨C)) gave us a valid argument for A (since the probability went from .55 all the way up to 1), a moderately strong argument in favor of B (since the probability went from .65 up to .8), and a weak argument *against* C (since the probability went from .6 *down* to .5).

At this point, we still don't have any method for finding prior probabilities (and in any case, the existence of objective methods for finding prior probabilities is the biggest controversy in statistics these days), but if we fill in some plausible numbers in our weather arguments from the first section, we can validate the intuitions described there.

First weather argument (second argument):

> It is either rainy or snowy today. It is not rainy and hot today. Therefore it is not hot today.

Here is a set of plausible prior probabilities:

Probability	H	R	S	R	∨	S	¬	(R	∧	H)	¬	H
0	T	T	T	T	T	T	F	T	T	T	F	T
.1	T	T	F	T	T	F	F	T	T	T	F	T
0	T	F	T	F	T	T	T	F	F	T	F	T
.3	T	F	F	F	F	F	T	F	F	T	F	T
.1	F	T	T	T	T	T	T	T	F	F	T	F
.2	F	T	F	T	T	F	T	T	F	F	T	F
.1	F	F	T	F	T	T	T	F	F	F	T	F
.2	F	F	F	F	F	F	T	F	F	F	T	F

The important point here is that the prior probability of it being both hot and snowy was 0. (The rest of the chart was just filled out by imagining a day in Missouri in late March, when temperatures could get up into the 80's, but there's also a chance of an Arctic blast bringing snow.)

Given this set of priors, the posteriors are given by taking just the rows where both premises are true, and dividing the probabilities for those rows by the total that remains:

Probability	H	R	S	R	∨	S	¬	(	R	∧	H	)	¬	H
0	T	F	T	F	T	T	T		F	F	T		F	T
.1/.4=.25	F	T	T	T	T	T	T		T	F	F		T	F
.2/.4=.5	F	T	F	T	T	F	T		T	F	F		T	F
.1/.4=.25	F	F	T	F	T	T	T		F	F	F		T	F

In this case, the prior probability of the conclusion was $.1 + .2 + .1 + .2 = .6$, and the posterior probability is $.25 + .5 + .25 = 1$, so as we said before, the argument is valid.

Next argument:

> It is either windy or rainy today. It is not windy and sunny today. Therefore it is not sunny today.

Here is a plausible set of prior probabilities:

Probability	R	S	W	W	∨	R	¬	(	W	∧	S	)	¬	S
.05	T	T	T	T	T	T	F		T	T	T		F	T
.05	T	T	F	F	T	T	T		F	F	T		F	T
.15	T	F	T	T	T	T	T		T	F	F		T	F
.1	T	F	F	F	T	T	T		F	F	F		T	F
.15	F	T	T	T	T	F	F		T	T	T		F	T
.15	F	T	F	F	F	F	T		F	F	T		F	T
.2	F	F	T	T	T	F	T		T	F	F		T	F
.15	F	F	F	F	F	F	T		F	F	F		T	F

The important point here is that it is unlikely to be both rainy and sunny together.

The posterior probabilities result from deleting the rows where not all the premises are true, and dividing through by the total remaining probability:

Probability	R	S	W	W	∨	R	¬	(	W	∧	S	)	¬	S
.05/.5=.1	T	T	F	F	T	T	T		F	F	T		F	T
.15/.5=.3	T	F	T	T	T	T	T		T	F	F		T	F
.1/.5=.2	T	F	F	F	T	T	T		F	F	F		T	F
.2/.5=.4	F	F	T	T	T	F	T		T	F	F		T	F

To see the strength of this argument, we compare the prior of the conclusion, which is $.15 + .1 + .2 + .15 = .6$ to the posterior, which is $.3 + .2 + .4 = .9$. Thus, we see that this is a moderately strong argument, as we had stated.

For the last argument:

> It is either windy or rainy today. It is either windy or sunny today. Therefore it is sunny today.

we can use the same prior probabilities:

Probability	R	S	W	W	∨	R	W	∨	S	S
.05	T	T	T	T	T	T	T	T	T	T
.05	T	T	F	F	T	T	F	T	T	T
.15	T	F	T	T	T	T	T	T	F	F
.1	T	F	F	F	T	T	F	F	F	F
.15	F	T	T	T	T	F	T	T	T	T
.15	F	T	F	F	F	F	F	T	T	T
.2	F	F	T	T	T	F	T	T	F	F
.15	F	F	F	F	F	F	F	F	F	F

The posteriors result from the following table:

Probability	R	S	W	W	∨	R	W	∨	S	S
.05/.6=1/12	T	T	T	T	T	T	T	T	T	T
.05/.6=1/12	T	T	F	F	T	T	F	T	T	T
.15/.6=1/4	T	F	T	T	T	T	T	T	F	F
.15/.6=1/4	F	T	T	T	T	F	T	T	T	T
.2/.6=1/3	F	F	T	T	T	F	T	T	F	F

In this case, the prior of the conclusion is $.05 + .05 + .15 + .15 = .4$, and the posterior is $1/12 + 1/12 + 1/4 = 5/12 \approx .4166$. Thus, this argument is very weak.

4 How to work out answers

If you are given a set of prior probabilities (that is, a truth table with a number attached to each row, where the numbers add up to 1), then you can figure out how strong an argument is by going through the following process:

1. Draw up a truth table to figure out whether the premises and conclusion are true or false in each row.

2. Calculate the prior probability of the conclusion by adding up the numbers on the rows where the conclusion is true.

3. Eliminate all the rows except the ones where all the premises are true.

4. Calculate the posterior probabilities of the rows by dividing the number on each row by the total of the numbers that are left.

5. Calculate the posterior probability of the conclusion by adding up the posterior probabilities of the rows where it is true.

6. Figure out how strong the argument is by seeing how much the prior probability moved towards either 1 or 0 when you got to the posterior probability.

You will express the strength of the argument in ordinary language using vague and imprecise terms like "weak" or "moderately strong" or "strong". There is no hard and fast rule here, but you might say that an argument is very weak if it moves a probability of .2 to .25 or .9 to .91, and it's moderately strong if it moves .3 to .7 or .85 to .95, and it's very strong if it moves .1 to .9 or .9 to .99. Of course, if it moves the probability *down*, then the premises form an argument *against* the conclusion (which could itself be weak or strong), and if the probability goes all the way to 1 then the argument is *valid*.

5 Calculating priors when given conditional probabilities

So far, we have discussed posterior probabilities as the probabilities that you have after learning some evidence. However, before you have learned the evidence, it is standard to call these *conditional* probabilities, because they are the probabilities that you would have *if* you were to learn some particular piece of evidence.

If we look through the procedure from before, we can see that the conditional probability of one sentence given another can be calculated by first calculating the probability that both sentences are true, and then dividing it by the one that is the condition. That is, the probability of A given B is the probability of $A \wedge B$ divided by the probability of B. It is common to write conditional probability with a vertical bar separating the claim whose probability you are interested in on the left, from the claim that might be given as evidence on the right, so that the probability of A given B is written as $P(A|B)$. (Note that this is the opposite order from how conditionals are written within the logical language, where "if B, then A" is written as $B \to A$.)

With all this in mind, we thus have the equation:

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

which summarizes much of what was said earlier. The probability of B is the result of summing all the rows where the evidence B is true, and the probability of $A \wedge B$ is the result of summing the rows among those where A is also true.

This equation can be rewritten by a little bit of algebra, moving $P(B)$ from the denominator of one side over to a multiplication on the other side:

$$P(A \wedge B) = P(A|B)P(B).$$

Thus, if you are told the probability of A given B, and are told the probability of B, then you can find the probability of $A \wedge B$.

For instance, suppose we know there is a football game tomorrow between the Thunder Bay Stormers and the Sunshine Coast Rays. We are told that there is a 40% chance of rain, and that the Stormers win 3/4 of their games against the Rays that are played on rainy days. Given this information, we want to find the probability that there is rain and the Stormers win.

Probability	Rain	TB wins
$.4 \cdot 3/4 = .3$	T	T
	T	F
	F	T
	F	F

Another way to think about this is that because the probability of rain is 40%, we know that the first two rows together have to add up to .4. Then, we know that out of that .4, 3/4 of the time is a win by Thunder Bay, and 1/4 is a win by Sunshine Coast.

Probability		Rain	TB wins
.4 {	.3	T	T
	.1	T	F
		F	T
		F	F

Since we know that the first two rows add up to .4, and the whole set of rows must add up to 1, we know that the bottom two rows must add up to .6. However, if we want to find the precise numbers there, we will need to know another conditional probability, which tells us how likely Thunder Bay or Sunshine Coast are to win given that it *doesn't* rain. If we know that Sunshine Coast has a 2/3 chance of winning when it doesn't rain, then we can fill this in as follows:

Probability		Rain	TB wins
.4 {	.3	T	T
	.1	T	F
.6 {	.2	F	T
	.4	F	F

Thus, we can set up the following procedure for calculating all the probabilities in a case where we are considering two atomic sentences A and B, and are told the probability of A, and the probability of B given A, and the probability of B given $\neg A$.

1. Create the chart and put in the known probabilities of A and $\neg A$.

Probability		A	B
$P(A)$ {	?	T	T
	?	T	F
$1 - P(A)$ {	?	F	T
	?	F	F

2. Use the conditional probabilities of B given A and given $\neg A$ to figure out how much of each probability goes on the first and third line respectively.

	Probability	A	B
$P(A)$	$P(A) \cdot P(B\|A)$	T	T
	?	T	F
$1 - P(A)$	$(1 - P(A) \cdot P(B\|\neg A)$	F	T
	?	F	F

3. Fill in the remaining rows with whatever is left from the probabilities that you know from step 1.

	Probability	A	B
$P(A)$	$P(A) \cdot P(B\|A)$	T	T
	$P(A) - (P(A) \cdot P(B\|A))$	T	F
$1 - P(A)$	$(1 - P(A) \cdot P(B\|\neg A)$	F	T
	$(1 - P(A)) - ((1 - P(A)) \cdot P(B\|\neg A))$	F	F

Another way to think of this directly is:

Probability	A	B
$P(A) \cdot P(B\|A)$	T	T
$P(A) \cdot P(\neg B\|A)$	T	F
$P(\neg A) \cdot P(B\|\neg A)$	F	T
$P(\neg A) \cdot P(\neg B\|\neg A)$	F	F

6 Bayes's Theorem

Once you have calculated all this information, you can feed it back into the procedure from earlier if you need to determine a different conditional probability. Thus, if you know $P(A)$, $P(B|A)$ and $P(B|\neg A)$, you can use these to calculate $P(A|B)$. This is often useful if you know the outcome of some process, and want to know how likely it was that some earlier step in the process actually occurred.

Returning to the football example, recall that we knew that the chance of rain was .4, and that Thunder Bay wins 3/4 of games when it rains, while Sunshine Coast wins 2/3 of the games when it doesn't rain. Imagine now that we are told that Thunder Bay won, and want to know whether this is good evidence that it rained.

First, we can fill out the chart using the methods given in the previous section:

Probability	Rain	TB wins
.3	T	T
.1	T	F
.2	F	T
.4	F	F

Then we can use the methods of the earlier sections to figure out how this changes if we know that Thunder Bay won.

Probability	Rain	TB wins
$.3/.5 = .6$	T	T
$.2/.5 = .4$	F	T

Thus, we see that the probability of rain went from .4 to .6, so learning that Thunder Bay won gives some moderate support to the conclusion that it must have rained.

If you just want to calculate answers, this is enough. But if you want to understand the general mathematics behind this, you can read on.

Implicit in the way we are doing this whole calculation is the fact that the same probability can be calculated in two different ways. We know that $P(A \wedge B) = P(A) \cdot P(B|A)$. But we also know that $P(B \wedge A) = P(B) \cdot P(A|B)$. Since $A \wedge B$ and $B \wedge A$ are the same statement, these must be equal:

$$P(A) \cdot P(B|A) = P(B) \cdot P(A|B).$$

If we are trying to calculate one conditional probability when we are told information about the other, this is often useful in the slightly rearranged form:

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

This equation is known as "Bayes's Theorem", because a version of it was first proved in an essay by the Rev. Thomas Bayes in 1763.

Usually we are not given the probability of A directly, so it is often useful to consider that $P(A) = P(A \wedge B) + P(A \wedge \neg B)$, as well as the fact that $P(A \wedge B) = P(B) \cdot P(A|B)$ and $P(A \wedge \neg B) = P(\neg B) \cdot P(A|\neg B)$ to reformulate it as follows:

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(B) \cdot P(A|B) + P(\neg B) \cdot P(A|\neg B)}.$$

In this formulation, $P(A|B)$ and $P(A|\neg B)$ are known as the "likelihoods" that tell you how likely A is to result when B is true or false, while $P(B)$ and $P(\neg B)$ are the "priors", and $P(B|A)$ is the "posterior" probability of B once you've learned A. What this result shows is that if you want to calculate how strong some event A is as evidence for or against some condition B, you can figure this out if you know the prior probability of B, and the likelihoods of A given B and given $\neg B$. This sort of situation arises all the time, when B is some underlying condition that can't directly be observed (whether it's the presence of a disease, or the bias of a coin, or yesterday's weather) and A is some test that depends on that condition (whether it's a medical test, or the result of a coin flip, or the outcome of a football game). Thus, this set of techniques is often extremely useful in dealing with probabilistic reasoning.