reason and argument

John Kulvicki

Probability

In valid deductive arguments it might be that your premises are outlandish, but you at least know that *if* they are true, the conclusion must follow. With inductive arguments, the best we can hope for is that the premises lend less-than-conclusive weight to the conclusion. What makes an inductive argument good, then, is the extent to which the premises support the conclusion.

Let's say that 'p' stands for the conjunction of an argument's premises, while 'c' stands for its conclusion. We use a vertical stroke '|' to indicate what is *given*, so the expression 'Pr(c|p)' reads as the probability that the conclusion is true, given that the premises are true. Terminology in hand, here are some ways to unpack what makes an inductive argument good. Which condition do you find most compelling? Why?

$Pr(c|p) > Pr(-c|p)$
$Pr(c|p) > Pr(c|-p)$
$Pr(c|p) > Pr(c)$
$Pr(c|p) > Pr(-c)$

None of these gives us as clear a sense of value as validity does in the deductive case, but which comes closest to what you want?

Probability helps us think about inductive arguments because they do not establish their conclusions in the way deductive arguments do. In addition, we can often turn an inductive argument into a deductive argument whose conclusion is a claim about the probability that something or other is true. So, it might not be that our evidence establishes a given conclusion, even though the evidence allows us to conclude that the probability of the conclusion is a certain value. For this reason, it's important to be able to think about probability and how the probabilities of complex statements depend on probabilities of their parts

Conjunction

Probabilities apply to statements, the kinds of things that can be true or false. What is the probability that it rains tomorrow? The statement, "It rains tomorrow", can be either true or false, and we are trying to figure out whether it's going to be pleasant walking to work. We can't be sure whether it will rain. The Eye on the Sky forecast says the probability is 40%. So Pr(it rains tomorrow) = 40% = 2/5.

Second, we often think of probability with respect so a set of things that remains the same, and this is typically implicit. What are the chances we can get this couch up the staircase (given we don't cut it in half, reduce it to ashes, etc.)? You will notice that typical examples of probability judgments happen against the background of a somewhat elaborate setup:

> Say you have a jar which contains beans that are either spherical or oblate, and either red or black (but not both!). Half of the black beans are spherical, the other half oblate, and half of the red beans are spherical, the other half oblate. 20% of the beans are red. Beans are chosen one at a time, at random, and then replaced. What is the probability that you draw a black bean, then an oblate bean, then a red, spherical bean?

Oh my. These setups establish some ground rules against which we can make judgments of probability. Unsaid, but understood, is that the beans do not change color or shape, that beans do not randomly appear and disappear, that your desires play no role in determining which bean you pull out of the urn, that the world does not end midway through your choosing, that you do not drop a bean and thereby fail to pull any bean out, and so on. The possibilities being considered here are the ones in which you pull out each one of the beans in the urn. They are the ones that count. So, what is the probability that the first bean is black? Well, the beans are either red or black and 20% of them are red, so 80% are black. Among the many possibilities here, it seems as though 4 out of 5 involve choosing a black bean. So,

> Pr(you choose a black bean on the first draw) = 4/5

Then you put that bean back and draw again. The second draw has nothing to do with the first draw, in that you are just picking a bean out of the same collection as in the first case. We know that half of the black beans are oblate and that half of the red beans are oblate and that all beans are either black or red, so we know that half of all the beans are oblate. So it would seem that in half of the possible choices you choose an oblate bean:

Pr(you choose an oblate bean on the second draw) = 1/2

Similar considerations apply to the third draw: it's independent of the first two. We know that 20% of the beans are red and half of those, or 10% of the total, are spherical, so:

Pr(you choose a red sphere on the third draw) = 1/10

Now we seem to have probabilities for the three parts of the question asked, viz., what is the probability of choosing a black bean, an oblate bean and a red, spherical bean, in that order? How should we put these partial answers together?

This brings us to the third point about probabilities: you can combine them, but you have to *be careful* about how you do it. And one way to be careful is to remember that probabilities have values between 0 and 1. If your method of combination is suggesting that the probability of some statement is 1.5 or -.25, you are doing something wrong. Notice too that combining probabilities is like combining statements. Probabilities apply to statements, and the ways we have of combining statements have been worked out, in part, by the sentential logic we learned a few weeks ago. We have three probabilities, and we are asking about the probability that the three statements are all true. We are asking about how to find the probability of a conjunction of statements given that we know the probabilities of each of the conjuncts.

This case is relatively easy, and I will show you others that get a bit more difficult. Think about it for a moment; 4/5 of the time you get the black bean, but on the second draw only half of those cases will yield an oblate bean. So, among the 4 out of 5 draws that yield a black bean, only half of those will yield an oblate bean, which means you divide the number of cases you started with by 2, or, more simply multiply it by one-half. That is, it seems like the probability of getting the first two options is

4/5 * 1/2 = 4/10 = 2/5

In two cases out of five, your first two draws will happen as we want. The third draw is unlikely: only one in ten cases will yield a red, spherical bean. This suggests we should divide the number we have so far by 10, or multiply it by 1/10:

4/5 * 1/2 * 1/10 = 4/100 = 1/25, or 4%

Notice that this reasoning suggests that the probability of the complex event is just the product of the probabilities of the component events. If we abbreviate each of the statements above with b, o, and r, respectively, we get:

Pr((b&o)&r) = (Pr(b)*Pr(o))*Pr(r)

These three events are *independent* of one another. How one of them works out has no bearing on how the other works out. That is to say, the probability that the second draw is an oblate bean given that the first draw was black is the same as the probability that the second draw is oblate given that the first was not black. We can write this as follows:

Pr(o|b) = Pr(o|-b) *The probability of o, given b, equals the probability of o, given not-b.*

We made sure this was true by insisting that you return the bean you choose to the bin before choosing another one. This formula *does not work* when the events in question are linked to one another. That is, the formula above fails when:

Pr(o|b) ≠ Pr(o|-b)

To see this point, imagine what happens if you get to keep the beans you draw. You draw a black bean, and set it aside. What is the probability that your next bean is oblate? You realize you are not sure because you are now choosing a bean from a smaller set of beans. How much smaller is it? The problem does not give you enough information to figure this out because it doesn't tell you how many beans are in the bin. If the bin has only 10 beans then keeping one off to the side greatly affects the odds for your next draw, while the effect is minimal if there are 1,000,000 beans in there. You also realize that your odds change depending on whether the drawn bean is oblate or spherical.

We know that the first draw was black, and let's say we are told there are 10 beans in all. Eight are black, but you have drawn one of those and put it to the side. Half of the time this happens it is an oblate bean while half the time is it spherical. So, there are seven black beans left (out of nine beans altogether) and either 3 or 4 of them are oblate, and two red beans, one of which is oblate. Thus, we know that there are either 4 or 5 oblate beans left out of nine. This is actually a nice result. When you think about it, after the first successful draw we have eighteen options: nine of them correspond to the nine beans left when you draw a black, oblate one and the other nine correspond to the nine beans left when you draw a black, spherical one. Of those eighteen, four in one group and five in another give you the result you want: an oblate bean. So, nine out of eighteen times you get the bean you want, or half the time. (This is the same result we got in the case where the draws were independent of one another, but things need not work out that way: check for yourself using different numbers of beans!)

> Pr(o|b) = ½ , *when there are 10 beans in the bin to start*

What about the final draw? For the final draw we want the red sphere. We already have taken out two beans, one black, the other oblate. The sphere has so far been untouched, so what is Pr(r|o&b)? There are eight beans left, and only one of them is the red sphere, so

> Pr(r|o&b) = 1/8, *when there are 10 beans in the bin to start*

We now have three components that can be combined to give us the answer we want:

> Pr(b&o&r) = Pr(b) * Pr(o|b) * Pr(r|o&b) = 4/5 * 1/2 * 1/8 = 1/20, or 5%.

Also notice that the order matters now in a way it didn't before. What is the probability of drawing the red sphere, followed by an oblate bean, followed by a black bean? I'll use the same letters, but remember that they stand for slightly different propositions here. For example, 'r' stands for 'the *first* draw is a red sphere' while earlier on it stood for the statement that the third draw was a red sphere, and so on for the others.

> Pr(r) = 1/10, *there is just one red sphere out of 10 beans*
> Pr(o|r) = 5/9, *there are nine beans left, five of which are oblate*
> Pr(b|o&r) = 15/16, *there are sixteen options, in eight of them, the oblate bean was red, leaving all the rest black, so all eight are good, if the oblate bean was black, there are 7 out of eight black beans, so 15 out of the total 16 cases work out here.*
> Pr(r&o&b) = 1/10 * 5/9 * 15/16 = 75/1440 or 5.21%

As in sentential logic, it is very important to keep track of what your sentence letters mean. Remember that in this latter case 'r' stands for the statement 'your first draw is a red sphere' while in the former case 'r' stood for 'your third draw is a red sphere'. I repeat this because I don't want you to think that the order of letters in the conjunction—'r&o&b'—is important. The order of sentence letters doesn't matter in a conjunction, which is what you learned when schematizing sentences. The order of draws is built into the statements themselves, not into the order in which letters appear in the conjunction. Just to beat a dead horse:

> Pr(r&o&b) = Pr(r&b&o) = Pr(o&r&b) = Pr(b&o&r) = Pr(b&r&o) = Pr(o&b&r)

That's just to say that the '&' in these expression works just like the '&' in sentential logic: order doesn't matter. In English, we sometimes use 'and' to indicate the order of events—"he got out of the car and fell into a puddle"—but we don't use it that way here.

Let's take a step back to note that in general,

> Pr(e&c) = Pr(e) * Pr (c|e)

In cases where e and c are independent of one another Pr(c|e) = Pr(c|-e) = Pr(c) and we have the original, simpler formula from above:

> Pr(e&c) = Pr(e) * Pr(c) , *when the events are independent*

Since the order of the conjuncts does not matter, we also know that

Pr(e&c) = Pr(e) * Pr (c|e) = Pr(c) * Pr (e|c)

We've been translating probabilities of conjunctions of statements into products of probabilities of statements. That's a useful tool. How might we translate probabilities of disjunctions into probabilities of individual statements? Earlier in the notes I mentioned that the order in which you try out independent events has no bearing on the probability that they happen. That's what being independent means. For example, I said the probability of getting black, then oblate, then red sphere are the same as getting red sphere, then black, then oblate, when the events are independent of one another. As we saw, this is *not* true if the events do depend on one another, but for now let's consider an analog of dependence that matters when thinking about disjunction: overlap.

Disjunction

What is the probability that you get either (black, oblate, red sphere) in that order, or (red, black, oblate) in that order? Assume we are replacing the beans we draw. Abbreviate the first statement with 'p' and the second with 'q'.

p: You draw black, then oblate, then a red sphere.
q: You draw a red sphere, then black, then oblate.

What is Pr(p v q)? These events are completely unconnected, and they are each possible outcomes of drawing three beans (and replacing them after each draw), and they are mutually exclusive, you get one sequence only if you don't get the other. The probability of the first is 5% and we know that the probability of the second is 5% as well. You can't get both of these results on three draws, so Pr(p & q) is zero. In this case it seems like the probability of getting one or the other is just the sum of the probabilities of getting each one:

Pr(p v q) = Pr(p) + Pr(q) = 8%

But what about cases in which the probability of getting both options is on the table? For example, what about the probability of getting a black bean (d) or getting a spherical bean (s) on one draw?

d: You draw a black bean.
s: You draw a spherical bean.

It would be a **mistake** to reason as follows:

Pr(d v s) = Pr(d) + Pr(s) = 4/5 + 5/10 = 13/10 or 130%
*******!!!NO!!!********

What went wrong? Well, we easily forget that some things are both black and spherical. We are counting those things twice, which can easily put your probability estimates into uninterpretable territory. To avoid counting twice, you always need to subtract the probability that something is both spherical and black (or more generally the probability that both of the disjuncts are true), as follows:

Pr(d v s) = Pr(d) + Pr(s) – Pr(d & s) = 8/10 + 5/10 – 4/10 = 9/10 or 90%

And when you think about it, eight beans are black, and one of the red ones is spherical, so 9 out of ten are black or spherical, or both (remember the 'v' sign is an inclusive 'or'). We subtracted 4/10 because 4 spheres out of 10 are both black and spherical: don't want to count them twice! It's very important to note that we are not *ignoring* the cases in which the beans are both black and spherical. We count those cases, but because beans can be black and spherical at once we count them *twice* if we just consider Pr(d) and Pr(s), once for the beans that are spherical and once for the ones that are black. We subtract the probability that they both occur in order to avoid double counting, not in order to avoid cases in which you get both. It's a common mistake to think that because we subtract the overlap we are actually calculating the probability for an exclusive 'or'. We are not. The 'v' symbol used here is the same as the one used earlier in the sentential logic notes.

So far, then, we know that

Pr(e&c) = Pr(e) * Pr(c|e) and
Pr(e v c) = Pr(e) + Pr(c) – Pr(e & c)

In some special cases Pr(c|e) = Pr(c) and in some special cases, where there is no overlap, Pr(e & c) = 0. Now that we have those connectives all sorted out we are just left with negation. The nice thing is that negation is easier than conjunction and disjunction.

Negation, etc.

If we know Pr(e) then what is Pr(-e)? Remember: probabilities fall between 0 and 1. Also remember that we take it for granted that any statement—the kind of thing that can be true or false—is either true or false. So,

Pr(-e) = 1 – Pr(e)

The truth-functional connectives—'-', 'v', and '&'—are important for understanding relations between parts of complex statements, and these exercises have tried to show how to derive probabilities for complex statements based on probabilities of their parts. This is a subtle affair, because sometimes the probability of one event depends in no small part on the probability of another event. So, we need to keep our eyes on whether events are independent or not. We also need to keep track of what statements we are calculating probabilities for. It's easy to forget.

By the way, what about the '→'? Well, we know that the arrow, understood truth-functionally, is equivalent to a disjunction, so

Pr(p → q) = Pr(-p v q) = Pr(-p) + Pr(q) – Pr(-p & q)

Here's a cautionary tale about applying the above formulae.

Let's say you flip a coin, over and over. You want to know the probability of getting heads at least once in n flips. The probability of getting heads on the first try is ½, of course, but how do you figure out the probability for getting heads at some point in a string of, say, five tosses? It is *very* easy to make a mistake here. This is a question about a disjunction: what is the probability that the first toss is heads, or the second toss is heads, or the third toss is heads, and so on. Applying the formula for a disjunction to this case, however, will likely end in heartbreak. Let's say h1 is "The coin is heads on the first flip", h2 is "The coin is heads on the second flip", and so on. Then you might think:

Pr(h1 v h2 v h3 v h4 v h5) = Pr(h1) + Pr(h2) + Pr(h3) + Pr(h4) + Pr(h5) = 2.5, or 250%   *************!!!NO!!!*************

Where did we go wrong? Aren't the tosses independent of one another? They are. Do they overlap? In a sense, they do, though this might not be obvious to you. We are clearly counting things twice in the above situation, but what are we counting twice? It turns out we are counting *a lot* of possibilities twice. For example, imagine the first toss is heads. Then we know that the disjunction—"h1 v h2 v h3 v h4 v h5"—is true. It doesn't matter whether the second toss is heads or tails. So we need to subtract from the total above the probability that *both* the first and second tosses are heads: Pr(h1 & h2). Once we see this, it becomes apparent that we need to subtract the probability that the first three are all heads, the first four are all heads, and the first five are all heads. We also need to subtract the probability that the first, third, and fifth are all heads, and so on and so on. A simple calculation has quickly come to seem a royal pain. Luckily, there is a way out of this mess, though a simple application of the formulas we have already derived.

Think about the problem this way. We want to know the probability that in five tosses we get at least one head. Why not figure out the probability that in five tosses we get *no* heads? The probability that we get at least one toss heads will just be 1 minus the probability that we get no heads. So,

Pr(h1 v h2 v h3 v h4 v h5) = 1 – Pr ~(h1 v h2 v h3 v h4 v h5) *this is the negation formula*

Why might this make our lives easier? The right hand side of the equation is looking for the probability, not of a disjunction, but of the negation of a disjunction, which, if you remember your sentential logic, is just a conjunction:

-(h1 v h2 v h3 v h4 v h5) *is logically equivalent to*
-(-(-h1 & -h2 & -h3 & -h4 & h5))*which is logically equivalent to*
-h1 & -h2 & -h3 & -h4 & -h5 *which we can now substitute back:*
Pr(h1vh2vh3vh4vh5) = 1 – Pr(-h1 & -h2 & -h3 & -h4 & -h5)

That's progress! We know that the coin tosses are independent of one another, and we have an easy formula for transforming a probability of a conjunction of independent events into a product of probabilities for those events:

Pr(-h1 & -h2 & -h3 & -h4 & -h5) =
Pr(-h1)*Pr(-h2)*Pr(-h3)*Pr(-h4)*Pr(-h5)

Substituting back into the original formula yields:

Pr(h1 v h2 v h3 v h4 v h5) =
1 – Pr(-h1)*Pr(-h2)*Pr(-h3)*Pr(-h4)*Pr(-h5)

And we know that the probability for each of the statements in the right hand side is ½.

1 – Pr(-h1)*Pr(-h2)*Pr(-h3)*Pr(-h4)*Pr(-h5) =
1 – ½*½*½*½*½ = 1 – $(½)^5$ = 1 – 1/32 = 31/32

So,

Pr(h1 v h2 v h3 v h4 v h5) = 31/32

This approach applies quite generally. As long as h1 – h5 refer to independent events, and you want to know the probability that at least one of these events happen, this approach will help you. A certain kind of specific case comes up often enough to make it worth its own formula. Often, you want to know the probability that a certain thing will happen in a string of identical, independent events, like coin tosses, rolls of the dice, spins of the roulette wheel, and so on. In the above example, we wanted to know the probability that a coin toss was heads. Rather than come up with a different statement for each coin toss—h1, h2, h3, etc.—let's call h the statement that "The coin comes up heads". In cases like this:

Pr(the coin comes up heads at least once in n tries) =
1 – $[Pr(-h)]^n$

Try out these ideas yourself. For example, what is the probability that you roll a 3, if you roll one six-sided die four times? Or, what is the probability that you roll a 7 if you roll two six sided dice three times?

The foregoing example serves to make a very important point. In many cases it's easy to see that events are independent, and thus when you can apply the formula that makes a probability of a conjunction into a product of probabilities. It is comparatively harder to notice situations when there is overlap between events that leads to problems for the formula for disjunctions. Independence and overlap are not the same thing! Be careful when applying these formulas. But remember that there is typically a nice way to check whether you are going down the wrong path. If you get answers for probabilities that are either above 1 or below zero you have made a mistake somewhere. Making a mistake is not so bad if you have the tools for noticing that you have done so and for fixing it once it has been noticed.

The logical lesson in the foregoing is that because the probability of p is 1 minus the probability of not-p, we have a simple way to turn a calculation about a disjunction into a calculation about a conjunction. That's so because negating a disjunction gives a conjunction, and vice-versa, which is something you learned in the earlier notes. For this reason, among others, it's helpful to have sentential logic in hand when you are starting to study probability.

A particularly recalcitrant problem, in which overlap is difficult to discern, is found in the so-called Monty Hall cases, which we will encounter soon. The next two sets of notes look at the importance of what we call inverse probabilities.

*