Logic, First Course, Winter 2020. Week 9, Lecture 2. Back to course website

Probability

Deduction vs induction

Logic is the paradigmatic example of a deductive science, the other being mathematics. The common contrast is between deduction and induction. In deduction, the idea is that the truth of the premises guarantees the truth of the conclusion. In inductive settings, we replace this with the idea that the truth of the evidence makes more likely the truth of hypothesis that we are interested in. But how do we formalize the notion of "more likely"? One contemporary answer is: probability. Today we learn how to think about probability as a slight extension of truth-tables, and next time we learn how to make more precise the idea of "evidence making a hypothesis more likely".

Assigning probabilities to rows of the truth table

Suppose that we use the following key to do some simple translations, just like in Week 2:

> *f* = the coin flips heads on the *f*irst flip

> *s* = the coin flips heads on the *s*econd flip

Of course, we are thinking of the coin as a normal two-sided coin, with two sides: heads and tails. Hence, we could further translate any claim about landing tails by using a negation.
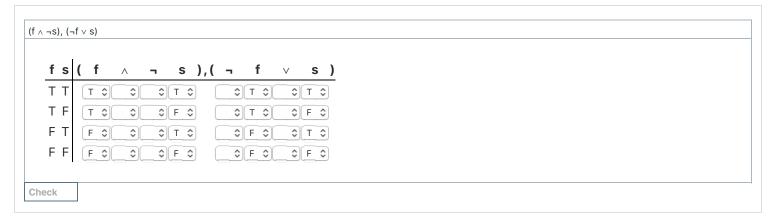

The coin flipped heads on the first slip and

The coin flipped heads on the first slip and tails on the second flip.

The coin flipped tails on the first flip or h

The coin flipped tails on the first flip or heads on the second flip.

And we could calculate the truth-values of these as normal (the commas separating the two formulas is just a way for us to write two truth-tables in one box):

f s	(	f	∧	¬	s	),(	¬	f	∨	s	)
T T		T ⬍	⬍	⬍	T ⬍		⬍	T ⬍	⬍	T ⬍	
T F		T ⬍	⬍	⬍	F ⬍		⬍	T ⬍	⬍	F ⬍	
F T		F ⬍	⬍	⬍	T ⬍		⬍	F ⬍	⬍	T ⬍	
F F		F ⬍	⬍	⬍	F ⬍		⬍	F ⬍	⬍	F ⬍	

Check

But recall what the rows of a truth-table represent: they represent possible outcomes or situations. Hence, a more descriptive presentation of the four rows would be:

TT = Heads on first flip, heads on second flip.
TF = Heads on first flip, tails on second flip.
FT = Tails on first flip, heads on second flip.
FF = Tails on first flip, tails on second flip.

It is natural to assign probabilities to each of these outcomes. If our coin was a fair coin, we would assign probabilities as follows:

Probability assignment for fair coin, for propositions *f* and *s*:

1/4 -- TT = Heads on first flip, heads on second flip.
1/4 -- TF = Heads on first flip, tails on second flip.
1/4 -- FT = Tails on first flip, heads on second flip.
1/4 -- FF = Tails on first flip, tails on second flip.

Since each of the rows is assigned equal probability, one can calculate the probability of a proposition simply by counting how many rows it is true at. For instance, $f \wedge \neg s$ was true at just one row, and so it has probability 1/4. But $\neg f \vee s$ was true at three rows, and so its probability would be 3/4.

Of course, in saying this we are reporting probabilities as numbers between 0 and 1. Instead of 1/4 we could also write 25% or .25. We will go back and forth between these different ways of representing numbers between 0 and 1 often in what follows.

Of course, our coin might have been an unfair coin, in which case the probabilities would be different. For instance, if it had a 1/3 chance of going heads on any given flip and 2/3 chance of going tails, then the probabilities would be as follows:

Probability assignment for biased coin, for propositions *f* and *s*:

1/9 -- TT = Heads on first flip and heads on second flip.
2/9 -- TF = Heads on first flip and tails on second flip.
2/9 -- FT = Tails on first flip and heads on second flip.
4/9 -- FF = Tails on first flip and tails on second flip.

Since each row does not have equal probability, we then calculate as follows:

- The probability of $f \wedge \neg s$ is 2/9, since it is only true at row TF and this row has probability 2/9.
- The probability of $\neg f \vee s$ is 7/9 = 1/9+2/9+4/9, since it is true at rows TT and FT and FF, and the numbers in the sum are the probabilities of those rows.

In general, we can summarize as follows:

- A *probability assignment* assigns numbers between 0 and 1 to each row of the truth-table, so that the numbers all add up to 1.
- One then assigns probabilities $Pr(\phi)$ to propositions ϕ by adding up the probabilities of all those rows at which the proposition ϕ is true.

The only thing new here is the notation $Pr(\phi)$ for the probability which we assign to a proposition ϕ. Note that this is always relative to some assignment of probabilities to rows of a truth-table, which is usually fixed by context (i.e. by whether we are talking about a fair coin, a biased coin, etc.)

To illustrate, let's note that it is possible to assign a row of a truth-table probability zero. For instance, suppose we use the key

   p = it is sunny on Monday

   q = it is sunny on Tuesday

Then we might assign probabilities as follows, since we live in a part of the world where it is very unlikely to have two days in a row that are not sunny:

Probability assignment for an agent's degrees of belief about the weather, for propositions p and q:

3/5 -- TT = It is sunny on Monday and it is sunny on Tuesday.
1/5 -- TF = It is sunny on Monday and it is not sunny on Tuesday.
1/5 -- FT = It is not sunny on Monday and it is sunny on Tuesday.
0 -- FF = It is not sunny on Monday and it is not sunny on Tuesday.

Unlike in the previous cases about coins, here there are no intuitive rules that tell one how to get the precise numbers. Rather, we simply take this as given as part of the problem-- for instance, as the title suggests, we might take the numbers to be reflective of an agent's degrees of belief in the propositions. The view that probabilities are reflective of the degrees of beliefs of agents is called *Bayesianism* in philosophy. To give a feel for it, here is the opening pages of Jeffrey's *Subjective Probability* book:[1]

From this point of view probabilities are "in the mind"—the subject's, say, YOURS. If you say the probability of rain is 70% you are reporting that, all things considered, you would bet on rain at odds of 7:3, thinking of longer or shorter odds as giving an unmerited advantage to one side or the other.

For some more practice, let us fill out the following truth-table (the commas separating the two formulas is just a way for us to write two truth-tables in one box):

(p → q), (p ∧ ¬q)

p	q	(	p	→	q	),(	p	∧	¬	q	)
T	T		T		T		T			T	
T	F		T		F		T			F	
F	T		F		T		F			T	
F	F		F		F		F			F	

Check

What is Pr(p->q), relative to the agent's degrees of belief about the weather?

○ 1/5
○ 2/5
○ 3/5
○ 4/5

Check

The rules of probability

Suppose that we have a probability assignment which assigns probabilities to rows of a truth-table. And again, suppose we write $Pr(\phi)$ for the sum of the probabilities of the rows at which ϕ is true. There are a couple of helpful rules or axioms which one can then note:

- First rule: $0 \le Pr(\phi) \le 1$
- Second rule: if ϕ is a tautology, then $Pr(\phi) = 1$.
- Third rule: $Pr(\neg\phi) = 1 - Pr(\phi)$.
- Fourth rule: $Pr(\phi \vee \psi) = Pr(\phi) + Pr(\psi) - Pr(\phi \wedge \psi)$

We are going to show that these rules hold, given that we have defined the probability of a proposition as the sum of the probabilities of the rows where the proposition is true. In other settings where there are infinitely many propositions one is concerned with, one cannot define the probability of propositions in this way, and then these rules are laid down as axioms.

The first rule is true simply because we assigned numbers between 0 and 1 to rows with a sum totaling to 1, and hence when we add up the rows where a proposition is true, we get a number between 0 and 1. The second rule is true because if ϕ is a tautology then ϕ is true at all rows of the truth-table, and since the numbers we assigned to rows adds up to 1, we have that $Pr(\phi) = 1$. The third rule is true because every row is such that either ϕ is true or $\neg\phi$ is true at the row but not both, and so $Pr(\neg\phi) + Pr(\phi) = 1$. For the fourth rule, let us introduce the following abbreviations:

- $A = \phi \wedge \psi$
- $B = \phi \wedge \neg\psi$
- $C = \neg\phi \wedge \psi$

Note that the rows where $\phi \vee \psi$ is true are exactly the rows where A is true, B is true, or C is true; and no row can be such that more than one of A, B, C is true at it. And the rows where ϕ is true are exactly the rows where A is true or B is true. Likewise the rows where ψ is true are exactly the rows where A is true or C is true. Since the probabilities of propositions are obtained by counting up probabilities of rows where propositions are true, we then get:

- $Pr(\phi \vee \psi) = Pr(A) + Pr(B) + Pr(C)$
- $Pr(\phi) = Pr(A) + Pr(B)$
- $Pr(\psi) = Pr(A) + Pr(C)$.

Putting these together, we can then verify the fourth rule:

- $Pr(\phi \vee \psi) = Pr(A) + Pr(B) + Pr(C) = Pr(\phi) + Pr(\psi) - Pr(A) = Pr(\phi) + Pr(\psi) - Pr(\phi \wedge \psi)$

These rules save us from essentially redoing several calculations with truth-tables over and over again.

Example 1. Suppose that $Pr(p) = .4$ and $Pr(q) = .5$ and $Pr(p \wedge q) = .2$. What is $Pr(p \vee q)$?

Example 2. Suppose that $Pr(p) = .3$ and $Pr(q) = .2$ and $Pr(p \lor q) = .45$. What is $Pr(p \land q)$?

Example 3. Suppose that $Pr(p) = .3$ and $Pr(q) = .2$ and $Pr(\neg p \land \neg q) = .5$. What is $Pr(\neg p \lor \neg q)$?

Connections between probability and logic

There are some simple connections between probability and logic which are useful to draw explicit attention to:
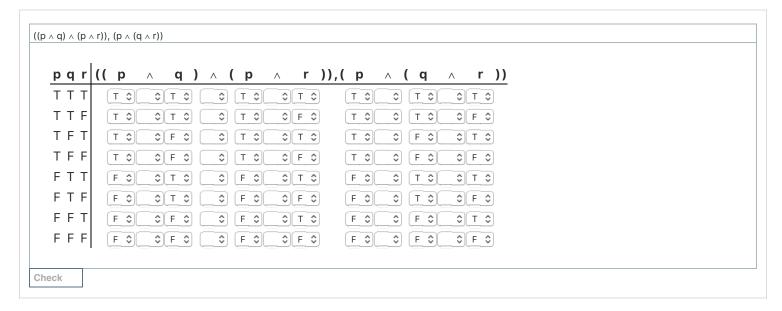
- If $\phi \rightarrow \psi$ is a tautology, then $Pr(\phi) \leq Pr(\psi)$.
- If ϕ and ψ are equivalent, then $Pr(\phi) = Pr(\psi)$.

The second just follows from the first since equivalences are biconditionals, which break down into a conjunction of conditionals. For the first, simply note that if $\phi \rightarrow \psi$ is a tautology, then any row where ϕ is true is also a row where ψ is true. Hence, every row which counts towards $Pr(\phi)$ will also count towards $Pr(\psi)$.

As a simple application of this, consider the following:

Example 4. Show that $Pr(\phi \wedge (\psi \vee \xi)) = Pr(\phi \wedge \psi) + Pr(\phi \wedge \xi) - Pr(\phi \wedge (\psi \wedge \xi))$.

One can check the equivalence mentioned at the end of this video by doing a truth-table like follows:

$((p \wedge q) \wedge (p \wedge r)), (p \wedge (q \wedge r))$

p q r	((p	∧	q)	∧	(p	∧	r)),(p	∧	(q	∧	r))
T T T	T		T		T		T	T		T		T
T T F	T		T		T		F	T		T		F
T F T	T		F		T		T	T		F		T
T F F	T		F		T		F	T		F		F
F T T	F		T		F		T	F		T		T
F T F	F		T		F		F	F		T		F
F F T	F		F		F		T	F		F		T
F F F	F		F		F		F	F		F		F

Check
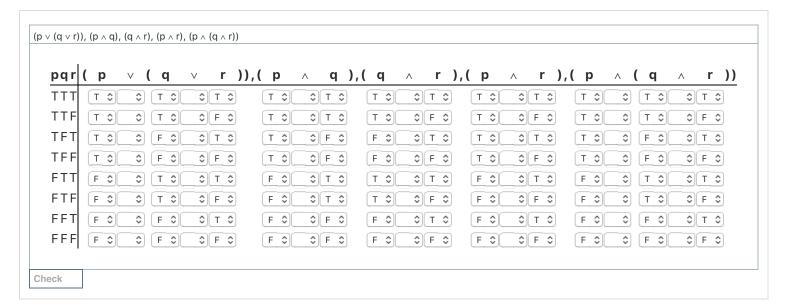
One can calso check the equivalence by doing two simple proofs:

((p ∧ q) ∧ (p ∧ r)) ⊢ (p ∧ (q ∧ r))

    1. (p∧q)∧(p∧r) :assumption

(p ∧ (q ∧ r)) ⊢ ((p ∧ q) ∧ (p ∧ r))

    1. p∧(q∧r) :assumption

Example 5. Show that $Pr(\phi \lor (\psi \lor \xi)) = Pr(\phi) + Pr(\psi) + Pr(\xi) - Pr(\phi \land \psi) - Pr(\psi \land \xi) - Pr(\phi \land \xi) + Pr(\phi \land (\psi \land \xi))$.

This is a special case of the so-called 'Inclusion-Exclusion' formula.

Example 6. Let us verify this formula, in a specific case. Consider the following truth-table, which we complete just like in Week 1 (the commas separating the formulas is just a way for us to write many truth-tables in one box):

(p ∨ (q ∨ r)), (p ∧ q), (q ∧ r), (p ∧ r), (p ∧ (q ∧ r))

p q r	(p	∨	(q	∨	r)),	(p	∧	q),	(q	∧	r),	(p	∧	r),	(p	∧	(q	∧	r))
T T T	T		T		T	T		T	T		T	T		T	T		T		T
T T F	T		T		F	T		T	T		F	T		F	T		T		F
T F T	T		F		T	T		F	F		T	T		T	T		F		T
T F F	T		F		F	T		F	F		F	T		F	T		F		F
F T T	F		T		T	F		T	T		T	F		T	F		T		T
F T F	F		T		F	F		T	T		F	F		F	F		T		F
F F T	F		F		T	F		F	F		T	F		T	F		F		T
F F F	F		F		F	F		F	F		F	F		F	F		F		F

Check

Suppose that each of the rows is weighted equally, that is each has equal chance of happening, namely a 1/8 chance of happening.

Then one can directly calculate

- Pr(p∨(q∨r))=7/8
- Pr(p)=4/8
- Pr(q)= 4/8
- Pr(r)= 4/8
- Pr(p∧q) = 2/8
- Pr(q∧r) = 2/8
- Pr(p∧r) = 2/8
- Pr(p/(q∧r)) = 1/8

To verify the formula from Example 5, we just need to check that

- 7/8 = 4/8+4/8+4/8-2/8-2/8-2/8+1/8

and as one quickly sees, this identity indeed holds.

These are lecture notes written for this course.[2]

---

1. Jeffrey, Richard. 2004. Subjective Probability: The Real Thing. Cambridge University Press.↩

2. It is run on the Carnap software, which is↩