LPS 31: Introduction to Inductive Logic
Lecture 1

Topics

Propositions
Logical connectives
Truth tables
Tautology, contradiction, and equivalence
Mutually exclusive propositions
Arguments
Validity and soundness

1 Propositions

A *proposition* (or a *statement*—for our purposes, these two terms are interchangeable) is a sentence that makes a definite claim or assertion. A propositions is the kind of thing that can be true or false of the world. To say that a give proposition is true or false is to assign a truth-value to that proposition. Many sorts of ordinary declarative sentences that we encounter in day-to-day life count as propositions.

- Today is Thursday.
- Socrates was mortal.
- All basketballs are orange with black stripes.
- Some pine trees grow in cold climates.

The words (or logical symbols) that appear in the expression of a proposition are different from the proposition being expressed. In other words, two sentences can correspond to the same proposition even if they use different words.

> "Snow is white" = "La neige est blanche" = "The form of precipitation common to New England winters and important for activities like skiing reflects light uniformly at all wavelengths."

Some declarative sentences are slightly more confusing because they express judgments. These, too, can be understood as propositions, though saying how to assign truth-values to these is sometimes be tricky because judgments can vary from person to person.

- Red is the prettiest color.
- Oranges taste good.
- The water is too cold for swimming.

For the purposes of this class, we are not interested in how truth values get assigned to propositions. So we will always understand these potentially problematic sentences as propositions, too. For the purposes of logical inference, they are just as good as any other kind of proposition.

Some sentences, meanwhile, are not propositions at all. These are sentences that do not make assertions about the world.

- Why is the sky blue?
- Eureka!
- Buy a gallon of milk when you go to the store.

In sentential (or propositional) logic, one often represents propositions using capitalized sentence letters. So, for instance, instead of saying "The snow is white," you might simply use the letter "S" to represent the proposition. This notation often makes logical manipulations more convenient.

2 Logical connectives

Propositions can be combined and modified to create new propositions using *logical connectives*. There are three logical connectives that we will use. These are: "not", "and", and "or". There are others that are sometimes used, and it is possible to define some of these in terms of the others. But we will not worry about such things in this course. The important thing is that you understand these four connectives. Informally, these work as follows. Note that each of the sentences in quotations marks counts as a proposition.

- "Today is not Thursday" is true if, and only if, "Today is Thursday" is false.
- "Today is Thursday and Socrates was mortal" is true if, and only if, "Today is Thursday" and "Socrates was mortal" are *both* true.
- "Today is Thursday or Socrates was mortal" is true if, and only if, "Today is Thursday" is true, "Socrates was mortal" is true, or both are true.

Note that "or" for logicians usually means "inclusive or". In ordinary usage, however, we usually mean "exclusive or". So if I say "I went to New York or Detroit" in an ordinary setting, it usually means that I either went to New York or I went to Detroit, but not both. In the context of logic, however, "I went to New York or Detroit" is true even if I went to New York *and* Detroit.

We will usually represent logical connectives symbolically. Here is a translation manual. (Assume A and B are propositions.)

Connective	Symbol(s)	Examples
not	$\sim$	$\sim A$
and	$\&, \wedge$	$A\&B$, $A \wedge B$
or	$\vee$	$A \vee B$

I cannot write the & symbol, so I will always use $\wedge$ for "and". I do not care which you use on homework or exams.

It is important to note that whenever you use logical connectives, it is important to use parenthesis. This is because expressions like $A \wedge B \vee C$ are ambiguous. So, whenever you combine two propositions with either $\wedge$ or $\vee$, be sure to add parentheses around the propositions you are combining. For example, if you have propositions $A \vee B$ and $\sim C$, and you want to express the proposition $A \vee B$ AND $\sim C$, you write $(A \vee B) \wedge (\sim C)$. Similarly, if you want to express NOT $A \vee B$, add parentheses around the original proposition, to get $\sim (A \vee B)$.

3 Truth tables

To understand complex propositions, i.e., propositions that are compiled out of sentences and logical connectives, it is often useful to use some formal method. We will focus on just one here: truth tables. Truth tables are a way of determining whether a given complex proposition is true or false when its simple constituent propositions are. Each of the logical connectives has an associated truth table.

A	B	$A \wedge B$
T	T	T
T	F	F
F	T	F
F	F	F

A	$\sim A$
T	F
F	T

A	B	$A \vee B$
T	T	T
T	F	T
F	T	T
F	F	F

These truth tables can be used to construct the truth tables for more complicated propositions. You do this by beginning with the propositions with the most parentheses around them and working your way out. So, for instance, the truth tree for the proposition $A \wedge (B \vee (\sim C))$ would be:

A	B	C	$\sim C$	$B \vee (\sim C)$	$A \wedge (B \vee (\sim C))$
T	T	T	F	T	T
T	T	F	T	T	T
T	F	T	F	F	F
T	F	F	T	T	T
F	T	T	F	T	F
F	T	F	T	T	F
F	F	T	F	F	F
F	F	F	T	T	F

Truth tables were studied extensively in LPS 29 and LPS 30, so I won't go into detail on how to construct them in general. We will not be making much use of them in this class, but they are a helpful tool for defining certain terms that we will need.

4 Tautology, contradiction, and equivalence

Some propositions are always true. These are called *tautologies*. Tautologies are propositions that only have T's marked in their column of a truth table. For example, $A \vee (\sim A)$ is a tautology. To see this, simply construct the relevant truth table:

A	$\sim A$	$A \vee (\sim A)$
T	F	T
F	T	T

As you can see, only T values appear in the column under the proposition $A \vee (\sim A)$. Conversely, some propositions are always false. These are called *contradictions*. They are propositions that only have F values appearing in their column of a truth table. An example is $A \wedge (\sim A)$.

A	$\sim A$	$A \wedge (\sim A)$
T	F	F
F	T	F

Tautologies an contradictions will be important later because they are propositions that get assigned special probabilities.

Sometimes two propositions can look different because they are expressed using different logical connectives, yet mean the same thing in the sense that they are true and false under the same circumstances. This means that whenever one of these propositions is true, the other is always true as well, and whenever one is false, the other is also false. When this happens, the two propositions are said to be *logically equivalent*. One way to test for logical equivalence again involves truth tables. Two propositions are logically equivalent if they have the same truth tables (neglecting any intermediate steps). Consider the propositions $A \wedge (B \vee (\sim C))$ and $(A \wedge B) \vee (A \wedge (\sim C))$. These two sentences look different from one another. But we can check whether they are logically equivalent nonetheless by looking at their truth tables. The truth table for the first is given above. The truth table for the second can be easily worked out. Ignoring the intermediate steps, we find,

A	B	C	$A \wedge (B \vee (\sim C))$
T	T	T	T
T	T	F	T
T	F	T	F
T	F	F	T
F	T	T	F
F	T	F	F
F	F	T	F
F	F	F	F

A	B	C	$(A \wedge B) \vee (A \wedge (\sim C))$
T	T	T	T
T	T	F	T
T	F	T	F
T	F	F	T
F	T	T	F
F	T	F	F
F	F	T	F
F	F	F	F

These are exactly the same. This means the two propositions $A \wedge (B \vee (\sim C))$ and $(A \wedge B) \vee (A \wedge (\sim C))$ are indeed logically equivalent. Determining whether two propositions are logically equivalent will be important later because if two propositions are logically equivalent, they should be assigned the same probability.

5 Mutually exclusive propositions

There is another relationship that can hold between two propositions. This topic is not usually discussed in LPS 29 or 30, but it is important in what will follow. I will explain this again later in the course, but it is appropriate to mention now. We will say that two propositions are *mutually exclusive* if they cannot both be true at the same time. In other words, two propositions are mutually exclusive if every time one of them is assigned the value "T" in a truth table, the other must be assigned the value "F".

Here are some examples of mutually exclusive propositions: A and $\sim A$; $A \wedge B$ and $A \wedge (\sim B)$; $A \vee B$ and $(\sim A) \vee (\sim B)$. How do we know these are mutually exclusive? One simple way is to check their truth tables. If every time on of the propositions is true, the other one is false (and vice versa), then they are mutually exclusive. So, for instance,

A	B	$A \vee B$	$(\sim A) \wedge (\sim B)$
T	T	T	F
T	F	T	F
F	T	T	F
F	F	F	T

You will notice that two propositions A and B are mutually exclusive whenever A is logically equivalent to $\sim B$ (all three of the examples above are of this form). But this is not the only situation in which two propositions can be mutually exclusive. For instance, $A \wedge (\sim A)$ and $B \vee C$ are also mutually exclusive. The first proposition is always false, which means that it is never the case that both propositions are true!

6 Arguments

An *argument*, in logic, is a list of propositions. One of these propositions (usually the last one) is the *conclusion* of the argument. The others are called the *premises* of the argument. The premises of an argument are propositions that are simply asserted. The conclusion, meanwhile, is expected to follow from the premises (more on this in a moment).

1	Socrates is a man	Premise
2	All men are mortal	Premise
3	Socrates is mortal	Conclusion

Sometimes arguments are presented in a more informal way. For instance,

> Socrates was certainly mortal. I know this because Socrates was a man, and everyone knows all men are mortal.

Even expressed this way, the sentences are still an argument. It is often helpful to re-express such informal arguments in logical form, by identifying the premises and conclusions and listing the premises first. The first example above shows how the informal argument just given might be put into logical form.

In some cases, a distinction is made between lists of propositions that, in some sense or another, are *intended* to demonstrate a conclusion. But saying what the intention of a list of propositions is can be difficult, and in any case, this kind of distinction does not matter here. (The distinction is important, if at all, only when one is worried about informal arguments, or putting informal arguments into logical form.) For our purposes, *any* list of propositions with one proposition marked as a conclusion and the rest as premises, counts as an argument. From this point of view, the only ways a list of sentences can fail to be an argument are if (1) none of the sentences is distinguished as a conclusion, or (2) one or more of the sentences in the list is not a proposition. So, the following is *not* an argument.

1	I need milk from the store.	Premise
2	John, get me my milk!	Premise
3	John will get me my milk.	Conclusion

7 Validity and soundness

An argument is *valid* if every time the conclusion is false, at least one of the premises is also false. (Equivalently, an argument if valid if, whenever all of the premises are true, the conclusion is also true.) Otherwise, it is *invalid.* Here's an example of a valid argument.

1	Malcolm is a banker or a soccer player	Premise
2	Malcolm is not a banker.	Premise
3	Malcolm is a soccer player	Conclusion

This argument could be expressed in symbols as follows.

1	$B \vee S$	Premise
2	$\sim B$	Premise
3	S	Conclusion

We know that this argument is valid because of the following truth table:

B	S	$B \vee S$	$\sim B$	S
T	T	T	F	T
T	F	T	F	F
F	T	T	T	T
F	F	F	T	F

There are two rows in which the conclusion (S) is false. In both of these lines, at least one premise is also false. Alternatively, one might just observed that there is one line in which both premises are true. In that line, the conclusion is also true. Thus the argument is valid.

A simple change would produce an invalid argument. Consider,

1	Malcolm is either a banker or a soccer player	Premise
2	Malcolm is a banker.	Premise
3	Malcolm is not a soccer player	Conclusion

This argument could be expressed in symbols as follows.

1	$B \vee S$	Premise
2	B	Premise
3	$\sim S$	Conclusion

Again, we can check that this is an invalid argument by checking the truth table.

B	S	$B \vee S$	B	$\sim S$
T	T	T	T	F
T	F	T	T	T
F	T	T	F	F
F	F	F	F	T

This argument is invalid because there is a line in the relevant truth table where the conclusion is assigned F, but both premises are assigned T.

An argument can be valid even if none of the premises happen to be true of the world. For instance, if one of the premises is a logical contradiction, the argument is automatically valid, since there is never a case in which the premises are true but the conclusion is false. Sometimes, however, we are interested in arguments that say something about the world. We call an argument *sound* if it is both valid, and all of the premises *are* true. Here is a sound argument.

1	Pine trees are either deciduous or coniferous.	Premise
2	Pine trees are not deciduous.	Premise
3	Pine trees are coniferous.	Conclusion

There are two ways in which this argument could be changed to make it unsound. For one, you could change the argument so that the premises are no longer true, but the argument is still valid. The following argument is not sound.

1	Pine trees are purple and either deciduous or coniferous.	Premise
2	Pine trees are not deciduous.	Premise
3	Pine trees are coniferous.	Conclusion

You can also change the argument to by keeping all of the premises true, but making the argument unsound. For instance, the following is not sound.

1	Pine trees are either deciduous or coniferous.	Premise
2	Pine trees are coniferous.	Premise
3	Pine trees are not deciduous.	Conclusion

Note that in this case, the premises really are true of the world, and the conclusion is *also* true. But the argument is not valid (for the reasons given above), and so it cannot be sound.

Checking whether an argument is sound can be a complicated empirical matter, since one needs a way of determining whether a given proposition is true of the world. Validity, however, only depends on the logical form of an argument. There are several tests for validity in propositional (and predicate) logic without worrying about what the world is like. These are studied in detail in LPS 29 and LPS 30. We will not discuss such topics further here. What is important for the current course is just that you understand what it means for an argument to be valid, and to be able to distinguish between a valid argument and a sound argument.

LPS 31: Introduction to Inductive Logic
Lecture 2

Topics
Deductive inference
Non-deductive inference and risk
Inductive inference
Deductive validity vs. inductive strength
Other non-deductive inferences

1 Deductive inference

In the last lecture, we reviewed various topics from LPS 29 and 30 involving propositions and arguments. In particular, we described what makes an argument *valid*. There are various argument patterns, or argument forms, that often appear in valid arguments. (Conversely, any argument constructed solely from these patterns is valid.) Some examples are:

- Disjunctive syllogism
- Modus ponens
- Modus tollens

Sometimes these argument forms are called *rules of inference*. *Inference* is the process of arriving at a conclusion from some set of premises; to *make an inference* is to draw a conclusion from some premises. The argument patterns just listed are rules that apply to deductive inference. An inference that follows these rules is a deductive inference, and any argument that one constructs from premises and a conclusion that is inferred by following these rules is a deductive argument. Some simple arguments make use of only one of these inferential rules. More complicated arguments, however, may rely on many such rules, applied one after another.

I will give some examples of these rules of inference (these aren't the only ones, but they give a good flavor). Strictly speaking, we are not working with the logical connective $\rightarrow$ (if ... then). This connective can be defined in terms of negation and disjunction ($A \rightarrow B$ is equivalent to $(\sim A) \vee B$). However, the most familiar of the deductive rules of inference are most naturally expressed in terms of material implication ($\rightarrow$). So I will (temporarily) use this connective to make the rules I describe look maximally familiar. I will also show how they work when one drops the conditional as a connective.

1.1 Disjunctive syllogism

Disjunctive syllogism is the rule that allows you to infer B from the premises $A \vee B$ and $\sim A$. An argument that uses this rule of inference would be written,

1	$A \vee B$	Premise
2	$\sim A$	Premise
3	B	Conclusion

Or, to use natural language,

1	You are either with me or against me.	Premise
2	You aren't with me.	Premise
3	You are against me.	Conclusion

1.2 Modus ponens

Modus ponens is the rule that allows you to infer B from the premises $A \to B$ and A. An argument that uses this rule of inference would be written,

1	$A \to B$	Premise
2	A	Premise
3	B	Conclusion

Or, to use natural language,

1	If it rains, I should have an umbrella.	Premise
2	It is raining.	Premise
3	I should have an umbrella.	Conclusion

Since we are not using the conditional, modus ponens is really an example of disjunctive syllogism. That is, since we are using $(\sim A) \vee B$ instead of $A \to B$, the argument above becomes,

1	$(\sim A) \vee B$	Premise
2	A	Premise
3	B	Conclusion

This is effectively disjunctive syllogism because $\sim(\sim A)$ is equivalent to A.

1.3 Modus tollens

Modus ponens is the rule that allows you to infer $\sim A$ from the premises $A \to B$ and $\sim B$. An argument that uses this rule of inference would be written,

1	$A \to B$	Premise
2	$\sim B$	Premise
3	$\sim A$	Conclusion

Or, to use natural language,

1	If it were raining, I would be all wet.	Premise
2	I am not all wet.	Premise
3	It isn't raining.	Conclusion

As with modus ponens, modus tollens is really an example of disjunctive syllogism. We can rewrite the above arguments as,

1	$(\sim A) \vee B$	Premise
2	$\sim B$	Premise
3	$\sim A$	Conclusion

2 Non-deductive inference and risk

Deductive inferential rules of the sort just described are used in constructing valid arguments. One can think of these inferential rules and the arguments constructed using them as "risk-free". This is because these rules always lead to valid arguments, and given any valid argument, the conclusion is always true whenever the premises are. You cannot go wrong if you accept the conclusion of a valid argument on the basis of its premises.

Unfortunately, the deductive rules of inference only apply in very special circumstances. If you happen to know that either A is true or B is true, and moreover you know that A is false, then you are in wonderful shape. But what happens if you are only pretty sure that A is false? Or if you know that A is often false—but right now, it might be true? This is the situation in which we typically find ourselves when trying to reason about the world. Deductive inference exists in a world of certainty; the real world is often uncertain. We are often forced to make use of arguments where the conclusion might be false even if the premises are all true. Real-world inferences often involve some amount of risk. You might be wrong. Here are some examples.

> A dedicated group of European ornithologists have conducted a multi-year study categorizing swan color. They have tracked thousands of swans from birth to death, across all of mainland Europe and the United Kingdom. This represents a substantial portion of the swan population in these regions. They have found that all of the swans they studied are white. Thus, the ornithologists conclude that all swans are white.
>
> John is sitting at a blackjack table in Las Vegas. He is showing a fifteen, while the dealer (who only reveals one of his or her two cards) is showing a five. John concludes the dealer will bust, so he does not take another card.
>
> When Mary arrived at school on Thursday, she saw that everyone was wearing a red t-shirt. Sometimes, Mary knows, clubs at the school organize fundraisers where they sell t-shirts and everyone wears them on the same day to support some cause. Mary concludes that there must have been such a fundraiser and that's why everyone's shirts are the same color.
>
> Cailin has only seen hairless cats. Therefore, she concludes, cats are always hairless.
>
> For a century and a half, biologists have accumulated evidence suggesting that species change over time to adapt to their environments. This evidence is well-explained by Darwin's view that evolution occurs as a result of natural selection. Thus, scientists conclude that Darwin's theory was correct.

These are not valid arguments. You will notice that the inferences used in constructing these arguments do not follow the rules of deductive inference. In some of these examples, the conclusions are known to be false (there are black swans and hairy cats), even though the arguments seem compelling or believable. In other cases, the conclusions may be supported by the premises in some way, but could still turn out to be false (Mary's schoolmates could all have worn red coincidentally, or to play a trick on Mary).

Nonetheless, these are often the best we have. This class will study how to understand at least some risky arguments. The goal will be to understand the rules of inference that are involved in these risky arguments, and to find a mathematical theory that allows us to judge how good an argument is.

3 Inductive inference

The variety of risky argument that we will focus on in this course is known as induction. Some of the examples already given (Cailin and the cats; the ornithologists; the blackjack player) are inductive arguments. For our purposes, inductive logic will be defined as the logic of uncertain evidential reasoning. This is a vague definition—I haven't said what evidence is. I will return to this definition later in the course. For now, we will focus on arguments involving propositions concerning populations and samples from those populations. A population is a collection of objects that are somehow related to one another (all of the orange in Florida; big cats; fishes; UCI undergraduates). Samples are sub-collections from a population (a particular box of oranges in Florida; a pride of lions; this fish; LPS 31). Sometimes a collection of objects can be a population *and* a sample of a still larger population (students in LPS 31 are a sample of students at UCI; students in the first row of this classroom are a sample of students in LPS 31). Moreover, a population can be something slightly more abstract, such as a long list of the results of successive coin flips (and a sample, then, could be a single coin flip, or a shorter list of coin flips). We will come back to this topic in coming weeks.

Broadly speaking, there are three kinds of inductive inference:

- Inference from propositions about a population to a proposition about a sample from that population.
- Inference from propositions about a sample of a population to a propositions about the whole population.
- Inference from propositions about one sample of a population to a proposition about another sample of the same population.

Here are some examples of each.

3.1 Population ⇒ Sample

Consider the following.

> Suppose you decide to buy a new umbrella. You happen to know that most Leaky brand umbrellas are reliable. You thus conclude that if you were to go to a store and buy a Leaky umbrella, that the particular umbrella you buy will also be reliable.

In this example, you happen to know something about a large group of objects—a *population*. This knowledge can be expressed by the following proposition.

Most Leaky umbrellas are reliable.

This serves as a premise for your argument. The conclusion is that:

The Leaky umbrella I buy will be reliable.

This is a risky argument. You know that most, or perhaps almost all, Leaky umbrellas are reliable. But the manufacturing process may be imperfect, or the box with the particular umbrella you are destined to buy may have fallen off a truck on its way to the store. Or perhaps someone else already purchased the umbrella and punctured it on a tree branch, and then returned it. You do not have the right sort of information to make a deductive inference. (To make a deductive inference, you would need a strong premise—something like, "All Leaky umbrellas are reliable.")

The inference being used in both examples can be expressed as follows:

1	Most/many/some/no x's are P.	Premise
3	This x is probably/maybe/probably not P.	Conclusion

Note that not all inferences of this form are going to be strong. This is a way in which inductive inference differs from deductive inference: the form of the argument is not enough to determine whether the argument is any good. Take the following example.

A lot of oranges from Florida are juicy. I just ordered a case of oranges from Florida. They will be juicy.

This is an argument that reasons from a proposition concerning a population (Florida oranges) to a proposition concerning a sample (the particular box I ordered). But there are all sorts of reasons to think that even if *many* Florida orange are juicy, many more, or even most, Florida oranges are not juicy. This kind of argument is very difficult to evaluate in its present form. This is why we are going to use probability theory to analyze inductive inferences.

3.2 Sample ⇒ Population

Take the examples above of Cailin and her cats, or of the ornithologists and the swans. In both of these examples, one happens to know something about several members of a population—i.e., a sample—and one uses that information to infer something about the whole population. So, for instance, in the swan example, the premise is,

All of the swans we have ever seen, representing a substantial portion of all swans, are white.

The conclusion is that this feature of the sample applies to the whole population, i.e.,

All swans are white.

This variety of inference need not be categorical.

My shipment of oranges from Florida just arrived. I have tried to eat five of them, and only two of them were juicy. I conclude that less than half the oranges I bought are juicy. Indeed, less than half the oranges in Florida are juicy.

In this case, one is again inferring something about a whole population (indeed, two populations: the oranges in my box, and all of the oranges in Florida), based on some sample. But I am not inferring something about *all* Florida oranges—rather, I am inferring a different kind of feature of the population. It is a feature that concerns the distribution of properties within the population.

This form of inference can be characterized as follows:

1	These particular x's are never/sometimes/often/always P.	Premise
3	All x's are never/sometimes/often/always P.	Conclusion

The last example, with the oranges, shows how this form of inference, too, can go wrong. Intuitively, it seems absurd to infer something about all of the oranges in Florida from a sample of five oranges. It is less absurd, perhaps, to infer something about the box of oranges that the five oranges came in. But the same form of inference is being used in both cases. Once again, it is difficult to evaluate such arguments from the form of their inference alone.

3.3 Sample ⇒ Sample

The final variety of inductive inference involves reasoning from propositions concerning one sample to propositions concerning another. Suppose we modified the orange example above slightly.

> My shipment of oranges from Florida just arrived. I have tried to eat five of them, and only two of them were juicy. I conclude that of the next five I try to eat, two will be juicy.

Here we are inferring something about one sample (the second batch of five oranges) from something about another sample (the first batch of five oranges). Once again, this sort of inference could go wrong: perhaps I lucked out and ate both juicy oranges in the whole box with the first five I picked out. Then there is no chance that the next five oranges will include two juicy oranges.

The form of inference can be characterized as follows:

1	These particular x's are never/sometimes/often/always P.	Premise
3	Those particular x's are never/sometimes/often/always P.	Conclusion

The example above, concerning the blackjack player, is another example of a sample to sample inference, though it is somewhat different because it relies on information concerning the whole population as well (specifically, how likely it is that various cards will turn up). It is best understood as a combination of several simpler inductive inferences.

4 Deductive validity vs. inductive strength

I have already alluded to this, but it is important so I will say it again. Deductive validity is a yes-or-no affair. Either an argument is deductively valid or it is not. And any argument constructed exclusively from the rules of deductive inference is valid. There is no gray area. There is much more to be said, meanwhile, of an argument that makes use of inductive inferences. There is a lot of gray area. Ultimately, we will be interested in evaluating the inductive *strength* of an argument. This will be a measure of how good an argument using inductive inference is. One way of thinking about deductive arguments is that if the premises are true, then the conclusion is certain to be true. With inductive arguments, if the premises are true, then the conclusion may be true or false. We want to distinguish between arguments such that, if the premises are true, the conclusion is *very likely* to be true, and argument where if the premises are true, the conclusion is only *somewhat likely* or even *unlikely* to be true. Arguments of the first sort will be called *inductively strong* arguments. If the conclusion is only somewhat likely given the premises, then the argument is not inductively strong. Again, there will be more to say on this topic.

5 Other non-deductive inferences

Induction is not the only kind of non-deductive reasoning that comes up in ordinary discourse. (Indeed, sometimes other forms of non-deductive inference are important for science.) To give an example, another variety of non-deductive reasoning is abductive inference. Hacking discusses abductive inference briefly, but I do not think he adequately explains how it differs from inductive inference.

5.1 Abductive inference

Abductive inference is sometimes called "inference to a plausible explanation" or "inference to the best explanation." It is inference from some set of premises to a conclusion that *explains* those premises. What constitutes an explanation is a famously difficult question. It may be that there is no way to characterize a "good" explanation in general. But for the purposes of identifying abductive inference, knowing whether an explanation is good is not relevant to determining whether the inferential rule used in an argument is abductive. What matters is that the conclusion is drawn because it would account for the truth of the premises.

One simple variety of abductive inference that you have seen before is the fallacy of affirming the consequent. Consider the following argument.

1	If it were raining, I would be all wet.	Premise
2	I am all wet.	Premise
3	It must be raining.	Conclusion

This is not a valid argument. It may be that I am wet because someone threw a water balloon at me, or I just took a shower. However, since if it *were* raining, that would mean I was wet, to postulate that is *is* raining is a plausible explanation of the premise that I happen to be wet. Note that this kind of inference is not one of the three inductive inferences I described above. There are no samples or populations anywhere in sight. I am not inferring from particulars to general features of the world or vice versa. I am trying to explain a premise by postulating another proposition. Another example of an abductive inference is the one given above, concerning Mary and her schoolmates' t-shirts. Mary is explaining an observation by postulating a proposition that, if true, would imply the observed fact.

Abductive inference is not the focus of this class. It is not clear that a satisfactory account of when an abductive inference is good, and of how good a given inference is, can be given in general. But in at least some cases, it is possible to spell an abductive inference out in terms of inductive inferences and probability theory. I will say more about this once we have some probability theory under our belts.

LPS 31: Introduction to Inductive Logic
Lecture 3

Topics
Probability and inductive inference
Test case: roulette
Bias
Independence
Three inferences in roulette

1 Probability and inductive inference

On Tuesday, we discussed three varieties of inference that I called inductive inferences. These were inferences from propositions concerning a sample to a proposition concerning a population, inferences from propositions concerning a population to a proposition concerning a sample, and inferences from propositions concerning one sample to a proposition concerning another sample. Our principle tool for studying arguments involving such inferences will be probability theory.

Over the next several weeks, we will develop the theory of probability. Once we have this set of tools under our belts, we will return to the question of how probability theory is useful in understanding inductive inference and the inductive strength of arguments. Today is a warm up to probability theory—a transition, if you will, from thinking about arguments to thinking about probability. The goal is just to get some of the basic ideas on the table and begin thinking about how probabilistic reasoning does (and doesn't) work. Next week I will talk about some of the rules for thinking about and working with probability theory. That will be the end of the first module of the class. The first midterm will be a week from Tuesday. After that, we will dive into probability theory more seriously. After the second midterm, we will return to the connection between probability and induction.

One of the principle questions we will address in the third module of the course will concern what probabilities are supposed to represent. For now, though, we can thinking of them as follows. To assign a probability to a proposition is to make some assertion about the *likelihood* that the proposition is true. Likelihood is not a perfectly clear notion either, as we will discuss later in the course. But in the meantime, we have an intuitive sense of what likelihood means. Most importantly for current purposes, we all have a sense of what it means to say that one proposition is *more* likely to be true (has greater likelihood) than another. It is very likely that most of us will leave this room before 9:30, but after 9:00. It is much less likely that we will leave before 9:00, although it is within the realm of possibility since, for instance, a fire alarm could go off. It is much less likely still that one of us will turn into a dog before the class is through, or that it will start snowing in the classroom, or that a desk will spontaneously combust.

I don't want to try to say any more about what this intuitive notion of "likelihood" means, at least for now. But what probability theory does is attempt to assign numbers to propositions like "Most of us will leave this room before 9:30 and after 9:00" or "Most of us will leave this room before 9:00" that capture the fact that one of these propositions is more likely to be true than the other. The way the theory does this is by assigning a larger number to the more likely proposition. The idea is then to find a way of assigning these numbers and of manipulating them to respect the intuitive notion of likelihood that we were trying to capture to begin with.

This kind of assignment of numbers has worked its way into ordinary usage. We often say things like, "there's a fifty-fifty chance that the Dodgers will make it to the play-offs this year" or "there's a one in five chance that it will rain tomorrow". When we do this, we are assigning numbers to the propositions "the Dodgers will make it to the play-offs this year" and "it will rain tomorrow" that reflect how likely we take those propositions to be. Probability theory is just an extension of that kind of usage, to make the assignments of numbers consistent and mathematically rigorous.

To give a sense of where all of this is going, let me try to say up front what connection this language is going to bear to inductive logic. Last week, we discussed the notion of inductive strength. We said that a valid argument is one in which, if the premises are all true, then the conclusion is true. Similarly, an inductively strong argument is going to be one in which, if the premises are all true, then the conclusion is very likely to be true, given the premises. (This last caveat is something we will return to later in the course—it is added to avoid certain undesirably problem cases.) Probability theory is going to be what lets us make sense of this criterion that a proposition is likely, or that it is more likely given that certain other propositions are true. It will turn out, however, that we can do somewhat more with probability theory than just say how inductively strong certain arguments are. Probability theory will give us a way of navigating, acting within, and learning about an uncertain world in a precise way. In LPS 29 and 30, much emphasis was put on identifying valid arguments. In this course, we will discuss inductive strength, the corollary to deductive validity. But more emphasis is going to be placed on inductive inference and its roles in decision making and learning.

Before proceeding, let me say one more thing. I will try to consistently speak of probabilities as numbers assigned to propositions representing the likelihood of their truth. But sometimes, it is equally natural (or perhaps, more natural) to speak of probabilities as numbers associated with future events. So, one might say the probability of 1/2 that a coin flip will come up heads. The event, here, is the landing of the coin with its head facing up. But you can always translate this event language into proposition language. One way is to simply say "The coin will land facing up." This is a declarative sentence, the sort of thing that might be true or false. (One might wonder whether it *is* true or false before the coin actually lands, but this is a question beyond the scope of this course.) It is more difficult to see how the proposition language can be translated into event language in all cases. For instance, I might want to assign a probability to the proposition "All swans are white." This is certainly an example of the kind of proposition we might want to make inductive inferences about. But what is the corresponding event? I don't mind if you use the event language. Sometimes I will, too. But for the purposes of the class, I think the proposition language will tend to be more useful.

2 Test case: Roulette

To get started in thinking about probability theory, let's first think about the game of roulette. Roulette, at least in the U.S., is a casino game based on a wheel with 38 numbered pockets uniformly distributed around the circumference of the wheel. The numbers run from 1-36, plus 0 and 00. 18 of these numbers are colored red on the wheel; another 18 are colored black. The 0 and 00 are both colored green. The game involves a casino worker, called the croupier, who spins the wheel and drops a ball into the middle of the wheel, which curves downward towards the pockets. The ball bounces around while the wheel spins, before finally settling down into one of the pockets as the wheel slows.

There are several ways to play roulette. They involve betting on different possible outcomes of a spin of the wheel. The ball can land in exactly one of the 38 pockets. One way to play is to bet on a particular pocket. If you do this, and you win, you get back 37 times your bet (i.e., the bet itself, plus 36 times the bet—this is often expressed as a payout of 36:1). But there are other ways to play. You can bet on even or odd numbers (where neither 0 nor 00 counts as even). Or you can bet on color. In either of these latter cases, the payout is 1:1, i.e., you receive double your money if you win. Another way to play is to be on the first dozen, the second dozen, or the third dozen. These payout 2:1, i.e., you receive three times your money if you win.

We are going to use roulette as a test case to explore some of the basic ideas of probability theory.

3 Bias

Roulette is a great place to begin thinking about probability theory because a roulette wheel is (usually) engineered so that the wheel is perfectly symmetric about its axis, and so that each of the 38 pockets is exactly the same size and exactly evenly spaced around the wheel. The wheel is then carefully calibrated so that it rotates perfectly, without any tilt. The purpose of all of this care is to make each of the 38 pockets equally likely. Using the language of probability theory, this means that each of the pockets should have equal probability associated with it. (Actually, in the proposition language, we should say that, for any given spin of the wheel, the probabilities associated with the 38 propositions "The ball will land in pocket x" are equal.) As far as probabilistic reasoning goes, this is as simple a test case as you can hope for.

In the context of probability theory, the fact all 38 possible pockets are equally likely is expressed by saying that roulette wheels are *unbiased*. If a roulette wheel is *biased*, meanwhile, it will tend to yield some numbers more often than others, perhaps because some of the pockets are larger than others, or because the wheel is tilted on its axis in a way that favors some numbers over others.

To think about bias precisely, it's going to be useful to introduce some terminology. Roulette is an example of what we will call a *chance setup*. Some other examples of chance setups are things like coin flips, dice rolls, drawing cards, pulling colored balls out of urns, etc. These chance setups are simple, sometimes idealized scenarios that are useful in probabilistic reasoning (and in applying probability theory to the world, at least in a simple way). Chance setups are characterized by their *possible outcomes*. The possible outcomes of roulette are that the ball will land in one of the 38 pockets. These outcomes are represented by the numbers associated with the pockets, so 00, 0, 1, 2, ... The possible outcomes of a coin flip are even simpler—just heads or tails. These can be represented by 0 and 1, or by H and T. The possible outcomes of a die roll are any of the six faces of the die, i.e., the numbers 1, 2, 3, 4, 5, and 6. If you play Dungeons and Dragons, the possible outcomes are any of the 20 numbers associated with a 20-sided die. The possible outcomes of pulling balls out of an urn will depend on what balls are put in the urn in the first place.

There is some idealization going on here. There are possible outcomes of spinning a roulette wheel that don't involve the ball falling in one of the pockets. The ball could fly off the wheel. The wheel could fly off its axis. The casino could be hit by a meteorite before the ball stops, obliterating the ball, the wheel, and the bettors. When considering chance setups, though, we ignore these kinds of possibilities. We assume the setup is such that the *only* possible outcomes are the ones described by the setup. Usually this is a harmless assumption—the likelihood that the ball will land in a pocket is much higher than that it will land in someone's martini. But it is an assumption, and one that we will consistently make.

The final concept we need is that of a *trial*. A trial is an event in which on of the possible outcomes is achieved. For roulette, a trial is one spin of the wheel. A trial for coin flipping is just a single flip of the coin; for dice, it's rolling a die. Each trial leads to a *result*, which is one of the possible outcomes. The possible outcomes are the things that might happen in a trial; the result is the thing that actually happens.

Given this vocabulary, we can define an *unbiased chance setup* as a chance setup such that, given any trial, the likelihoods that each of the possible outcomes will be the result are equal. In other words, the probabilities associated with the propositions "The result of the trial is x" where x is any of the possible outcomes are all equal. A *biased chance setup* is a chance setup such that, given any trial, the possible outcomes are not all equally likely to be the result. Because roulette wheels are usually manufactured to such high precision, it tends to be safe to assume that a game of roulette is an unbiased chance setup.

4 Independence

Often, we want to think about several trials of a chance setup in a row. This is certainly true in casinos—one rarely plays a single round of roulette, and even a less a single hand of poker. A second reason that roulette is a particularly simply game to think about probabilistically is that a given trial does not depend on the results of previous trials. This is again because of the careful manufacture of roulette wheels and balls: each time you spin the wheel, the wheel and the ball are virtually identical to the time before and the time after. The pockets are all the same size before a trial as after; the tilt of the wheel is still the same.

If the results of successive trials of a chance set up do not influence the likelihood of future results, the results are called *independent*. Roulette results are usually assumed to be independent, which is once again safe because of the high standards for roulette wheel manufacturing. If a chance setup is unbiased and its trials and independent, then it is unbiased for each successive trial. The likelihoods of the possible outcomes do not change as you run successive trials.

Independence is a notorious counterintuitive concept. We will do an experiment in class that, I think, will show this.

It is possible to have chance setups that are biased but whose trials are independent; unbiased but not independent; and biased and not independent.

4.1 Biased but independent

Some dice are heavier on one side than the others. They are called loaded dice. The unequal distribution of weight makes it more likely for the heavy side to land facing down than the lighter sides, which in turn makes the number opposite the heavy side more likely to come up than the others. But successive rolls of loaded dice are still independent of one another: the weight distribution does not change from roll to roll.

4.2 Unbiased but not independent

Suppose you take a standard deck of cards. The chance setup will be drawing the cards, without replacing them. The possible outcomes of the first trial are any of the fifty two cards in the deck. Any of these are equally likely, which means that the chance setup is unbiased. The possible outcomes of the second trial, meanwhile, are any of the fifty one cards remaining in the deck. Once again, these are equally likely, so the setup remains unbiased. But the two trials are not independent: you cannot draw the same card twice. The result of the first trial influences the likelihood of the results of the second trial.

4.3 Biased and not independent

Imagine a large urn filled with balls that are two different colors. 3 of the balls are blue, and 6 of them are red. The chance setup involves sticking your hand in any pulling out a ball without looking, and then not replacing the ball. The possible outcomes are that you draw a blue ball or that you draw a red ball. The likelihood of drawing a blue ball is much lower than of drawing a red ball, so the setup is not biased. But successive trials are not independent either. If you draw a blue ball on the first trial, the likelihood of drawing a red ball on the second trial are even higher than before (because now there are only 2 blue balls, and 6 red ones). Meanwhile, if you draw a red ball on the first draw, the likelihood of drawing a second red ball is still higher than that of drawing a blue ball, but not as high as it was in the first trial.

5 Three inferences in roulette

Using the ideas of bias and independence, let's think about the following three roulette related arguments, all of which use inductive (sample to sample) inferences. Imagine you are in a casino playing roulette. You are betting on red and black, depending on what feels right for each spin. You notice, however, that red has come up ten times in a row. That is, the results of ten successive trials have been red numbers. What should you think? You should take the rules of the chance setup for granted as additional premises, if you like.

5.1 *Argument 1*

P	The results of the last ten trials have been "red".
C	The likelihood is higher than normal that the next trial will yield "black".

5.2 *Argument 2*

P	The results of the last ten trials have been "red".
C	The likelihood that the next trial will yield "black" is unchanged.

5.3 *Argument 3*

P	The results of the last ten trials have been "red".
C	The likelihood is lower than normal that the next trial will yield "black".

What reasons might be given in support of each of these? Well, in the first case, you might argue, we know that the game is unbiased. Hence the likelihood of getting black and getting red are the same. But likelihood reflects something about how often the possible outcomes will obtain. If the likelihood of "red" and "black" are equal, then each outcome must obtain equally often. This means that if you get ten reds in a row, a black must be "due" sooner rather than later, because otherwise you will have more reds than blacks.

This reasoning is often called the "gambler's fallacy". Where is the fallacy? It comes in the fact that we have assumed that roulette is both unbiased and that the successive trials are independent. This means that a string of "red" results, no matter how long, does not affect the likelihood that the next result will be "black". Independence can be thought of as a kind of lack of memory. The anxious roulette player might be carefully watching the results of the roulette wheel, but the roulette wheel has no idea what has happened before. As we have seen already, independence can be a counter-intuitive notion. But if roulette really is both unbiased and independent, then argument 1 is incorrect (indeed, inconsistent—hence the term "fallacy"), and argument 2 is correct.

But what about argument 3? You might think that argument 3 goes afoul for the same reason as argument 1. In a sense it does. But argument 3 is more subtle than argument 1. When you use probability theory to analyze something like roulette, you are making assumptions and idealizations to simplify the situation. These assumptions are called a *probabilistic model.* A probabilistic model consists of stating the setup, the possible outcomes, and various assumptions about the setup, such as that it is unbiased or independent. Models make probability a useful tool for studying the real world. But sometimes we are in a position to question the assumptions of a model. If ten trials in a row come up "red," you might begin to wonder if the game really is unbiased. If a hundred trials come up "red", you would have even more cause to worry. The game might not be unbiased after all.

Fortunately, probability theory gives us tools that will help us understand this case as well. Using Bayes' theorem, we will be able to study the circumstances under which we might want to change our assumptions about the roulette wheel. The important point here, though, is that argument 3 is different than argument 1. Argument 1 is based on the idea that the model holds, but with a misunderstanding about how bias is related to independence. Argument 3 is based on the idea that the model you are working with is inadequate. Argument 1 is never a good one; argument 3 sometimes can be.

LPS 31: Introduction to Inductive Logic
Lecture 4

Topics
Tautologies and logical contradictions
Mutually exclusive propositions
Exhaustive propositions
Independent propositions

1 Some conventions of probability theory

Last Thursday, we began talking about some of the basic ideas underlying probability theory. The two ideas we focused on were bias and independence—but we also discussed what a chance setup is and how chance setups can serve as probability models when we make some simple idealizing assumptions about the world. I also talked about some fallacies that can come up when we try to reason probabilistically and which help show what assumptions about bias and independence mean. Today we are going to push forward with our study of probability theory. At this stage, the discussion is still informal, although today we will begin to actually assign probabilities to proposition. The first midterm is a week from today. After that I will return to the topics discussed today and Thursday with a little more precision.

As I said last week, the goal of probability theory (as we are studying it) is to find a mathematical way of capturing the intuition that some propositions are more likely to be true than others. Probability theory does this by assigning numbers of propositions, where, roughly speaking, the larger the number is, the more likely the proposition is to be true. We can represent this symbolically as follows. If we start with some proposition A, we will represent the probability that A is true by

$$\Pr(A).$$

So what number should we assign to A? As a first pass, we should expect that $\Pr(A)$ will be some number such that if B is any other proposition, and if B is more likely to be true than A, then $\Pr(B)$ will be greater than $\Pr(A)$. Conversely, if B is less likely than A, we should require that $\Pr(B) < \Pr(A)$. But one might ask if, given any proposition, there is always some other proposition that is still more likely to be true. Another way to ask this question would be, are there any propositions that are *so* likely to be true that there are no other propositions that are more likely to be true? Are some propositions maximally likely, or certain?

In a sense, this question has come up before, in deductive logic. In our brief review a few days ago, we discussed tautologies. Tautologies are propositions that are *always* true—i.e., there is no likelihood at all that they are false. So whatever else is the case, you should expect that tautologies are at least as likely to be true as any other proposition. But this thought suggests an answer to the question I asked above. There are some propositions that are maximally likely, namely the tautologies.

In terms of probability theory, the fact that tautologies are maximally likely might be expressed as follows: for any propositions A and B, if B is a tautology, then $\Pr(A) \leq \Pr(B)$.

In other words, there should be a largest number that we assign to propositions, corresponding to the probability we associate with tautologies. One of the principal conventions of probability theory is that this largest number is 1. It turns out that there are several reasons why 1 is a very natural choice. These reasons are due to how probabilities relate to one another. This will be more obvious later in the course.

This whole argument can be flipped on its head, too. We can ask if there are any propositions that are maximally unlikely, or in other words, that are at least as likely to be false as any other proposition. If tautologies are maximally likely to be true, we might take logical contradictions as an example of propositions that are *never* true. By the same reasoning that led us to assign a particular highest number to tautologies (the choice of which is a convention), we might be inclined to assign a particular *lowest* number to logical contradictions. Once again, it turns out that there is a particularly convenient choice: a second principal convention of probability theory is that the smallest value assigned to any proposition is 0.

Any other proposition is at least as likely as a logical contradiction (which is never true), and at most as likely as a tautology (which is always true). This means that the probabilities we will associate with any proposition will always be some number that falls between 0 and 1. To recap:

- The probability assigned to any proposition is some number that falls between 0 and 1.
- The probability assigned to tautologies is always 1.
- The probability assigned to logical contradictions is always 0.

2 Mutually exclusive propositions

We now have a way of assigning probabilities to some very special varieties of propositions—namely, tautologies and logical contradictions. We might ask, though, how to come up with probabilities for other propositions, i.e., for propositions that are not certainly true or certainly false. In general, this is very difficult to do. This course will focus on assigning probabilities to chance setups, which as we have discussed can provide useful probabilistic models. In the special case of a chance setup, it is generally possible to assign probabilities. But before we can say how that is done, we need to say a little bit about how probabilities need to relate to one another. Once we have done that, we can think about how to assign probabilities to other propositions.

During the overview of deductive logic a few weeks ago, I introduced the notion of mutually exclusive propositions. I said that two propositions were mutually exclusive if there were no conditions under which both were true. You could always check to see if two complex propositions were mutually exclusive by checking their truth tables. But sometimes one wants to use this concept to describe simpler propositions, or to describe propositions expressed with ordinary language. So, to take an example, if I roll a die, the propositions "the die will show a 1" and "the die will show a 5" are mutually exclusive propositions, even though there is no way to reveal this using truth tables. Instead, the sense in which these are mutually exclusive is given by the structure of the chance setup. Given any chance setup, the distinct possible outcomes are typically understood to be mutually exclusive. This is because, on any given trial, the result can only be *one* of the several possible outcomes. Thus a proposition asserting that the result of a trial is a particular outcome cannot be true if another proposition asserting that the result of the trial is a *different* particular outcome is true. A roulette ball cannot land in the 7 pocket and in the 5 pocket. A single draw from a deck of cards cannot be both the seven of hearts and the ace of spades.

Mutually exclusive propositions have a special property in probability theory. If two propositions are mutually exclusive, then the probability that one of them is true or the other of them is true is equal to the sum of the probabilities of each of them individually. Using logical connectives, this assertion can be expressed as follows. If A and B are mutually exclusive propositions, then the probability of the proposition $A \vee B$ is

$$\Pr(A \vee B) = \Pr(A) + \Pr(B).$$

In other words, the probability that A or B is true is given by the probability that A is true plus the probability that B is true.

Why should we expect probabilities to behave in this way? Here is one way of thinking about it. Imagine you are rolling a die over and over again. Some of the time, a 1 will land facing up. Other times, a 4 will land facing up. These are mutually exclusive, since only one of them can land facing up for any particular trial. If you interpret the likelihood of getting a 1 as having something to do with how often the die actually lands with 1 facing up, likewise if the likelihood of getting a 4 has something to do with how often the die actually lands with a 4 facing up, then you should expect the likelihood of getting *either* 1 or 4 to having something to do with how many times either 1 comes up or 4 comes up. So, for instance, if you rolled 100 times, and 1 can up 13 times, and 4 came up 20 times, then either 1 or 4 came up 33 times in total. This suggests that probabilities should behave the same way.

3 Exhaustive propositions

Think about flipping a coin. There are two possible outcomes, at least as we have defined the chance setup: the coin can land heads or tails. These are mutually exclusive, in the sense that the propositions "the coin landed heads up" and "the coin landed tails up" in relation to a particular trial cannot both be true. But they have another property, too: they are the only possible propositions. In the terms of probability theory, these are *exhaustive propositions.*

In real life, the proposition "the coin will come up heads or tails" is extremely likely, but it is not quite a tautology in the technical sense. Within the context of the chance setup, though, it is absolutely certain that exactly one of the possible outcomes will be the result of any given trial. In other words, it "the coin will come up heads or tails" is always true within the setup by definition. So, at least within the chance setup, "the coin will come up heads or tails" has the same likelihood as a tautology, since it is certain to be true. For this reason, by the reasoning given above, we should assign the proposition "the coin will come up heads or tails" probability 1. Symbolically, if we say H is the proposition "the coin will come up heads" and T is the proposition "the coin will come up tails," then,

$$\Pr(H \vee T) = 1.$$

This property of exhaustive propositions always holds. Given any chance setup whose possible outcomes are represented by the mutually exclusive, exhaustive propositions $A_1, A_2, \ldots, A_n$, the probability that A_1 or A_2 or ... or A_n will occur is given by $\Pr(()A_1 \vee A_2 \vee \cdots \vee A_n) = 1$.

Let's take another example. For a standard die, there are six possible outcomes. These are that any of the six faces will end up facing upwards. These correspond to propositions like "the die will land so that 1 faces up" or "the die will land so that 5 faces up." Then, the probability that the proposition "the die will land so that 1, 2, 3, 4, 5, or, 6 faces up" is 1. This is the same as saying $\Pr(A_1 \vee A_2 \vee A_3 \vee A_4 \vee A_5 \vee A_6) = 1$, where $A_n =$"the die will land so that n faces up."

Combining the properties of exhaustive propositions and mutually exclusive propositions gives us a way for calculating the probabilities for individual propositions, at least for unbiased chance setups. Consider the following. We know that if some propositions are mutually exclusive, then the probability that one of them is true is given by the sum of their probabilities. Meanwhile, if the propositions are exhaustive, then the probability that one of them is true is 1. Finally, if a chance setup is unbiased, then the probabilities associated with each of its possible outcomes are equal (this is from last week).

These three things let us make the following calculations. First, let's consider an unbiased coin toss. There are two possibilities, corresponding to the propositions H and T. Since these are the only possibilities, they are exhaustive; moreover, since they are two different possible outcomes, they are mutually exclusive. Thus, we can say the following

$$\Pr(H \vee T) = 1$$
$$\Pr(H \vee T) = \Pr(H) + \Pr(T)$$
$$\Pr(H) = \Pr(T)$$

Thus,

$$1 = \Pr(H \vee T) = \Pr(H) + \Pr(T) = \Pr(H) + \Pr(H) = 2\Pr(H)$$

and so $\Pr(H) = \Pr(T) = 1/2$. In other words, our other assumptions force us to conclude that the probabilities of getting heads is $1/2$. Similar calculations allows us to say that the probability of getting any particular number between 1 and 6 if you roll an unbiased die is $1/6$. The probability of getting any of the 38 possible roulette outcomes (for an unbiased roulette wheel) is $1/38$. Etc.

This suggests a general way of coming up with probabilities, at least for unbiased chance setups. You begin by counting up all of the possible outcomes. Suppose there are N distinct possible outcomes. If the game is unbiased, then each of these outcomes is equally likely, and so the probability of any one of them must be $1/N$.

What about for biased chance setups? There isn't always a rule — for biased coins, for instance, one has no way to calculate the probabilities from scratch. But in some cases, it is possible to re-imagine a biased chance setup as an unbiased one in which one forgets that some of the possible outcomes are different. Consider an urn with 4 black balls and 5 blue balls. In one sense, if you think of there being two possible outcomes (black or blue), then this setup is biased, since the likelihood of getting a blue ball is higher than that of getting a black ball if you reach in and pull a ball out. But if you imagine for a moment that each of these balls also has a unique number written on it, then you can immediately see that there are actually 9 distinct possible outcomes for this chance setup, rather than just two. And now the setup is unbiased, since each of the 9 balls is equally likely to be pulled from the urn. This means that the probability of getting any of the 9 balls is 1/9, by the reasoning given just above.

But we know that there are actually 4 black balls and 5 blue balls. Suppose that the four black balls happen to be the ones numbered 1-4, and the 5 black balls are numbered 5-9. Then you can say that the probability of getting a black ball is given by the probability of getting ball 1, ball 2, ball 3, or ball 4. Since these are mutually exclusive, we know how to calculate this probability: it is given by $\Pr(\text{“black ball”}) = \Pr(B_1 \vee B_2 \vee B_3 \vee B_4) = \Pr(B_1) + \Pr(B_2) + \Pr(B_3) + \Pr(B_4) = 4/9$. The probability of getting a blue ball, likewise, is given by 5/9.

This suggests a new, more general rule. Given a biased chance setup, try to identify a way to understand it as an unbiased chance setup, that is, find a way of individuating the possible outcomes in a way that makes each of them equally likely. In the urn case just described, this meant individuating the possible outcomes by counting balls, rather than by counting colors. A given color, then, can be understood as a property shared by some, but not all, of the possible outcomes. Then, the probability that a given trial will yield an outcome with a particular property (for instance, blue) is given by

$$\frac{\text{the number of possible outcomes with the given property}}{\text{the number of possible outcomes}}.$$

In the above example, there were 9 possible outcomes (each of the balls could be drawn from the urn). 4 of those outcomes had the property “black”; the other 5 had the property “blue”. Thus the probability of getting a black ball was 4/9 and a blue ball was 5/9. This kind of reasoning can also be applied when one asks about properties of outcomes of chance setups that are already unbiased. So, I might ask, what is the probability of rolling an even number on an unbiased standard die? (It’s 3/6=1/2 because there are three even numbers on the die.) Or, what is the probability of drawing a queen from a standard deck of cards? (It’s 4/52=1/12 because there are four queens in a fifty two card deck.)

4 Independent propositions

Last week, we discussed a property that a chance setup could have called “independence”. The idea was that the results of a series of trial of a chance setup were independent if the probabilities of the outcomes of one trial did not depend on the outcomes of previous trials. This notion can be generalized slightly. We will say that two propositions are *independent* if the truth of one does not influence the truth of the other. Note that this subsumes the previous definition, since if propositions concerning the results of different trials of a chance setup are independent in the new sense, then they are also independent in the old sense.

If two propositions are independent, then the probability that they are both true is given by the product of the probabilities that each one is true. So, if the trials of an unbiased coin flipping setup are independent, then the probability that I will get heads on the first flip is 1/2. The probability that I will get heads on the second flip is also 1/2. Likewise for the third flip and so on. But since these are independent, then the probability that I will get heads on *both* the first flip and the second flip is given by $\Pr(H_1 \wedge H_2) = \Pr(H_1) \times \Pr(H_2) = 1/4$.

Why should we expect this to be the case? Think about the coin flipping again. If the trials are independent, then we can understand flipping a coin twice as a new chance setup, with four possible outcomes. These are getting heads twice, getting heads on the first flip and tails on the second, and getting tails twice. These could be represented by HH, HT, TH, and TT. Thus there is one way, out of four possible outcomes, to get heads twice. This gives a probability of $\Pr(HH) = \Pr(H_1 \wedge H_2) = 1/4$, which is consistent with the multiplication rule.

More generally, one can think of the multiplication rule as follows. (Disclaimer: for some people, this explanation will do more harm than good. If it seems confusing, please just ignore it.) Suppose you have two propositions, A and B. Under some conditions, A will be true; under other conditions, it will be false. You can think of these conditions as possible worlds (this shouldn't be taken seriously – it is just to give an intuitive sense of what is going on here). So, suppose that in some possible worlds A is true, and in some A is false. Suppose the possible worlds in which A is true are called A worlds. But what about B? Once again, B is true in some possible worlds and false in others. But since A and B are independent, one should expect that whether A is true or not should not affect whether B is true. This means that the fraction of A worlds in which B is true should be the same as the fraction of *all* worlds in which B is true. So if B is true in 1 out of every 3 possible worlds, you should expect it to also be true in 1 out of every 3 A worlds. This, after all, is what independence means. The A worlds also make up some fraction of all possible worlds. If the A worlds are some fraction of all possible worlds, and the A worlds in which B is also true is some (known) fraction of the A worlds, you would expect the fraction of the worlds in which A is true and B is true to be the fraction of worlds in which A is true times the fraction of worlds in which B is true. This is what yields the multiplication rule.

LPS 31: Introduction to Inductive Logic
Lecture 5

Topics
Compounding probabilities
Categorical vs. conditional propositions
Conditional probability

1 Compounding probabilities

At the end of lecture last week, I described a rule for how to relate the probability that two independent propositions were both true to the probabilities that each of them was true on their own. That rule was that, if A and B are independent, then

$$\Pr(A \wedge B) = \Pr(A) \times \Pr(B).$$

This rule is especially useful for calculating probabilities for what might be called compound events (or, to be more precise, propositions concerning compound events).

What do we mean by compound events? Often, one is faced with questions concerning more than one chance setup (or more than one trial of a chance setup). We have already seen how to calculate probabilities concerning a single die—but what happens if you roll two dice, or three? How are you to calculate the probabilities for various possible outcomes then? To take another example, one might consider a compound chance setup in which you began with two urns, each filled with different numbers of colored balls. You flip a coin to decide which urn to draw from. What is the probability of getting a ball of a particular color?

These examples are cases in which one has more than one chance setup, and you are interested in the likelihood that propositions that concern trials of both setups are true. To calculate these probabilities, one begins with the assumption that trials to two different chance setups (or, two trials of a single chance setup) are independent. Then, you use the rule just stated.

To take the first example just described, suppose you are rolling two standard, unbiased dice, instead of one. For simplicity, let's assume one of the die is red and the other is white. Each of these dice can be thought of as a chance setup in its own right, and so we can take for granted that the results of each die will be independent. So suppose we want to ask: what is the probability that the red die will yield a 4 and the white die will yield a 2? If we label propositions concerning the possible outcomes of the red die's trial by $R_1, R_2, \ldots, R_6$ and the ones concerning the outcomes of the white die's trial by $W_1, W_2, \ldots, W_6$, then we are looking for $\Pr(R_4 \wedge W_2)$, which is given by

$$\Pr(R_4 \wedge W_2) = \Pr(R_4) \times \Pr(W_2) = (1/6) \times (1/6) = 1/36.$$

This is the probability of a proposition concerning a compound event.

You have to be very careful, though, about just what is being asked for. Suppose that instead of asking for the probability that the red die will yield a 4 and the white die will yield a 2, I instead asked for the probability that at least one of them would yield a 4. This is a question concerning a compound event, but it is not a question that uses the multiplication rule. Why? Because there are two ways in which (at least) one of the die could yield a 4: either the red one yields and 4, or the white one yields a 4. In logical terms, the complex proposition we are interested in is $W_4 \vee R_4$. How do we handle this? Actually, we don't yet have a rule that will tell us what to do here, because W_4 and R_4 are *not mutually exclusive*. It is possible for *both* W_4 and R_4 to be true at the same time, which means that the addition rule we used last class does not apply. Next week we will learn the rule for handling this kind of case, when we develop the probability calculus more formally.

Here's another kind of question that one might ask concerning compound probabilities. Suppose that instead of asking about particular possible outcomes for each die, I instead asked a question concerning the probability that the *sum* of the two results took a particular value? That is, suppose I asked for the probability that I would roll a 7, calculated by adding the two die's individual results? Here we want to think carefully, because so far we have only talked about the possible outcomes of the two chance setups individually. Now we want to think about how these might relate to one another.

The first step to calculating this compound probability is to try to think of what is being asked for, in terms of quantities we already understand. There are several ways in which the two dice might add to 7. You could have: a 6 on the red die and a 1 on the white one; a 5 on the red die and a 2 on the white one; a 4 on the red die and a 3 on the white one; a 3 on the red die and a 4 on the white one; a 2 on the red die and a 5 on the white one; or a 1 on the red die and a 5 on the white one. These are all the ways in which you can get the two die to sum to 7. The second step is to express these cases logically. The proposition whose probability we are interested in can be written as: $(R_6 \wedge W_1) \vee (R_5 \wedge W_2) \vee (R_4 \wedge W_3) \vee (R_3 \wedge W_4) \vee (R_2 \wedge W_5) \vee (R_1 \wedge W_5)$. What can we do with this? Well, first note that each term of the form $(R_i \wedge W_j)$ is mutually exclusive with each of the other terms of that form. This is because the red die can only take one value, and the white die can only take one value, so it cannot be true that the red die yields 5 and the white die yields 1 and *also* be true that the red die yields 4 and the white die yields 1. Meanwhile, we already know that propositions concerning one die are independent of propositions concerning the other. This allows us to calculate the following:

$$\begin{aligned}
\Pr(\text{Rolling a 7}) &= \Pr((R_6 \wedge W_1) \vee (R_5 \wedge W_2) \vee (R_4 \wedge W_3) \vee (R_3 \wedge W_4) \vee (R_2 \wedge W_5) \vee (R_1 \wedge W_5)) \\
&= \Pr(R_6 \wedge W_1) + \Pr(R_5 \wedge W_2) + \Pr(R_4 \wedge W_3) + \Pr(R_3 \wedge W_4) \\
&\quad + \Pr(R_2 \wedge W_5) + \Pr(R_1 \wedge W_5) \\
&= \Pr(R_6) \times \Pr(W_1) + \Pr(R_5) \times \Pr(W_2) + \Pr(R_4) \times \Pr(W_3) + \Pr(R_3) \times \Pr(W_4) \\
&\quad + \Pr(R_2) \times \Pr(W_5) + \Pr(R_1) \times \Pr(W_5) \\
&= 6 \times (1/36) = 1/6
\end{aligned}$$

There is another way of thinking about this problem that gives the same answer. Instead of thinking of the compound setup as two separate chance setups, you can think of it as a single (unbiased) setup with 36 outcomes. These outcomes would be

$$\begin{array}{cccccc}
R_1, W_1 & R_1, W_2 & R_1, W_3 & R_1, W_4 & R_1, W_5 & R_1, W_6 \\
R_2, W_1 & R_2, W_2 & R_2, W_3 & R_2, W_4 & R_2, W_5 & R_2, W_6 \\
R_3, W_1 & R_3, W_2 & R_3, W_3 & R_3, W_4 & R_3, W_5 & R_3, W_6 \\
R_4, W_1 & R_4, W_2 & R_4, W_3 & R_4, W_4 & R_4, W_5 & R_4, W_6 \\
R_5, W_1 & R_5, W_2 & R_5, W_3 & R_5, W_4 & R_5, W_5 & R_5, W_6 \\
R_6, W_1 & R_6, W_2 & R_6, W_3 & R_6, W_4 & R_6, W_5 & R_6, W_6
\end{array}$$

Each of these is equally likely, and so for any one of them, $\Pr(R_i, W_j) = 1/36$. You can then ask what the probability of getting a seven would be, and use the rule described on Tuesday for cases in which there is more than one way to get a particular result:

$$\Pr(P) = \frac{\text{the \# of possible outcomes with } P}{\text{the total \# of possible outcomes}}.$$

There are 6 possible outcomes with the property that the two dice add to 7, and so one would find $\Pr(\text{Rolling a } 7) = 6/36 = 1/6$. Thinking of the situation in this way should give some additional motivation for the multiplication rule from last week: after all, the outcome R_i, W_j understood in the new setup is equivalent to the outcome $R_i \wedge W_i$ understood as a complex propositions about two distinct setups.

2 Categorical vs. conditional likelihood

Thus far in our discussion of probability theory, we have focused on what might be called the *categorical* probability of various propositions. These are probabilities that capture the likelihood that assertions about the world (or about a particular chance setup) are true, without qualification. But often in probability theory, we are interested in something slightly different, which might be called a *conditional* probability of a proposition. Conditional probabilities represent the likelihood that one proposition is true, supposing that another one is true. It is very often the case that the likelihood of different propositions depend on one another. If I asked you what the likelihood was that you would be in class next Tuesday, you might say it was pretty high, given that we have a midterm. But if I then told you that I happened to know for a fact that a large meteorite would hit southern California over the weekend, and then asked what the likelihood was that you would be in class next Tuesday given this new piece of information, you would probably revise your estimate. If a meteorite hit southern California, it would be much less likely that class would happen at all, midterm or no midterm, than if no meteorite hit.

Here are some more examples of informal statements that make use of categorical and conditional probabilities:

- **Categorical** The Yankees have a fifty percent chance of making it to the playoffs this year.
- **Conditional** If the Red Sox bullpen does not improve, the Yankees have a fifty percent chance of making it to the playoffs this year.

- **Categorical** You stand a good chance of doing well on your midterm.
- **Conditional** Given how hard you've studied, you stand to a good chance of doing well on your midterm.
- **Categorical** The odds of drawing two aces in a row from a shuffled deck of cards are 1/221.
- **Conditional** The odds of drawing two aces in a row from a shuffled deck of cards, given that you've already drawn one ace are 3/51.
- **Categorical** The odds that you will get a five if you roll two dice is 1/9.
- **Conditional** The odds that you will get a give if you roll two dice, given that one of the dice came up 4 is 1/6.

3 Conditional probability

Conditional probabilities might seem like a complicated matter, but it turns out that there is a simple rule for calculating a conditional probability using the tools we have already developed. First, though, here is some notation. The probability that A is true given that B is true (or, equivalently, assuming B, or conditional on B) is expressed by: $\Pr(A|B)$. (Hacking writes $\Pr(A/B)$.) This number is calculated using the following formula. As long as $\Pr(B) \neq 0$, the probability of A given B is

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

An intuitive reason why we should expect this formula to be true goes something like as follows: if you imagine all of the possible worlds in which A is true, and the ones in which B is true, to ask what the probability that A is true given B, you want to first focus in on the worlds in which B is true and then ask what fraction of *those* worlds also have A true. But this is just another way of saying that you want to know the fraction of B worlds in which A and B are both true, which is what is represented by the formula (again, intuitively speaking).

Note that we do not require that A and B be independent here—indeed, to do so would defeat the purpose, because we are interested in just those cases where $\Pr(A)$ changes if we assume that B is true. But this means that we can extract some interesting information to generalize our earlier formula for $\Pr(A \wedge B)$. First, note that if A and B *are* independent, then we can write $\Pr(A \wedge B) = \Pr(A) \times \Pr(B)$, and so $\Pr(A|B) = \frac{\Pr(A \wedge B)}{\Pr(B)} = \frac{\Pr(A) \times \Pr(B)}{\Pr(B)} = \Pr(A)$. In other words, if A and B are independent, then the probability of A given B is just the same as the probability of A (i.e., the truth of B does not affect the probability of A, just as we should expect). Meanwhile, even if A and B are not independent, we can rewrite the formula above to find

$$\Pr(A \wedge B) = \Pr(A|B) \times \Pr(B).$$

This gives us a way a treating conjunctions of propositions that are not independent. Note that this means that you can use the conditional probability of one proposition given another to calculate the probability that both (not necessarily independent) propositions are true, or vice versa. But you need at least one of these pieces of information to figure out the other. These calculations can go in either direction, depending on the context.

First, let's consider the example treated above. You have two dice, one red and one white. Here are some conditional probabilities that you might calculate:

- What is the probability of rolling a 8, given that the red die rolled a 4?
- What is the probability of rolling an even number (total), given that the red die rolled a 4?
- What is the probability of rolling an 8, given that the red die rolled an even number?
- What is the probability of rolling an 8, given that one of the two dice rolled an odd number?

I will solve the first of these explicitly in the notes; in class, I will work through the others depending on time.

To answer this first question, we need to first figure out (as always), what is being asked. In symbolic terms, we want $\Pr(\text{Rolling an } 8|R_4)$. This isn't quite good enough, though, because we do not yet know what "Rolling an 8" means. So we need to unpack that in term of propositions whose probabilities we know. There are several ways to get your two dice to add up to 8. You could have: red land on 2 and white land on 6; red land on 3 and white land on 5; red land on 4 and white land on 4; red land on 5 and white land on 3; or red land on 6 and white land on 2. Now we want to translate these possibilities into logical expressions. In other words, "Rolling and 8" is equivalent to the proposition $P_8 = (R_2 \wedge W_6) \vee (R_3 \wedge W_5) \vee (R_4 \wedge W_4) \vee (R_5 \wedge W_3) \vee (R_6 \wedge W_2)$.

From here, we can proceed in one of two ways. The first is essentially brute force: we simply apply our formula. We know that

$$\Pr(P_8|R_4) = \frac{\Pr(P_8 \wedge R_4)}{\Pr(R_4)} = \frac{\Pr(P_8 \wedge R_4)}{1/6}$$

Meanwhile $P_8 \wedge R_4$ can be expanded (using the distributive and associative rules) to find,

$$\begin{aligned} P_8 \wedge R_4 &= ((R_2 \wedge W_6) \vee (R_3 \wedge W_5) \vee (R_4 \wedge W_4) \vee (R_5 \wedge W_3) \vee (R_6 \wedge W_2)) \wedge R_4 \\ &= ((R_2 \wedge W_6) \wedge R_4) \vee ((R_3 \wedge W_5) \wedge R_4) \vee ((R_4 \wedge W_4) \wedge R_4) \vee ((R_5 \wedge W_3) \wedge R_4) \\ &\quad \vee ((R_6 \wedge W_2) \wedge R_4) \\ &= ((R_2 \wedge R_4) \wedge W_6) \vee ((R_3 \wedge R_4) \wedge W_5) \vee ((R_4 \wedge R_4) \wedge W_4) \vee ((R_5 \wedge R_4) \wedge W_3) \\ &\quad \vee ((R_6 \wedge R_4) \wedge W_2) \end{aligned}$$

First note, that this is a disjunction of several different propositions. But it is easy to see that these are mutually exclusive, since, for instance, they each assert different things about the result of the next roll of the white die (which can only have one result per trial). But we might also note that some of the conjuncts have a striking property. Take, for instance, $R_2 \wedge R_4$. This proposition represent the assertion that on the next trial, the red die will come up both 2 and 4! But this is impossible, given the setup. So when we express the probability $\Pr(P_8 \wedge R_4)$, we find terms dropping out rapidly:

$$\begin{aligned}
\Pr(P_8|R_4) &= 6 \times \Pr(P_8 \wedge R_4) \\
&= 6 \times \Pr(((R_2 \wedge R_4) \wedge W_6) \vee ((R_3 \wedge R_4) \wedge W_5) \vee ((R_4 \wedge R_4) \wedge W_4) \vee ((R_5 \wedge R_4) \wedge W_3) \\
&\quad \vee ((R_6 \wedge R_4) \wedge W_2)) \\
&= 6 \times (\Pr((R_2 \wedge R_4) \wedge W_6) + \Pr((R_3 \wedge R_4) \wedge W_5) + \Pr((R_4 \wedge R_4) \wedge W_4) \\
&\quad + \Pr((R_5 \wedge R_4) \wedge W_3) + \Pr((R_6 \wedge R_4) \wedge W_2)) \\
&= 6 \times (\Pr(R_2 \wedge R_4) \times \Pr(W_6) + \Pr(R_3 \wedge R_4) \times \Pr(W_5) + \Pr(R_4 \wedge R_4) \times \Pr(W_4) \\
&\quad + \Pr(R_5 \wedge R_4) \times \Pr(W_3) + \Pr(R_6 \wedge R_4) \times \Pr(W_2)) \\
&= 6 \times (0 \times \Pr(W_6) + 0 \times \Pr(W_5) + \Pr(R_4) \times \Pr(W_4) + 0 \times \Pr(W_3) + 0 \times \Pr(W_2)) \\
&= 6 \times (\frac{1}{6} \times \frac{1}{6}) = \frac{1}{6}
\end{aligned}$$

So the probability of rolling an 8, given that you have rolled a 6 with one die, is 1/6.

This approach has the benefit that it is perfectly clear how we are using the rules all the way through. But it is quite complicated. There is another way of thinking about this problem, though, that is much simpler—and which shows how one can often calculate conditional probabilities directly and then use them to calculate the probabilities of conjunctions of non-independent propositions. This corresponds to the second way of calculating the conjunctive probability above. To begin, instead of writing down all of the possible outcomes in which we get an 8, we want to write down all of the possible outcomes for the two dice in which the red die shows a 4. These are: the red die shows a 4 and the white die shows a 1; the red die shows a 4 and the white die shows a 2; the red die shows a 4 and the white die shows a 3; the red die shows a 4 and the white die shows a 4; the red die shows a 4 and the white die shows a 5; or the red die shows a 4 and the white die shows a 6. In other words, given that the red die shows a 4, there are only six possible outcomes. Of these, exactly one has both dice adding up to 8. This means we can use the rule for properties to conclude that

$$\Pr(\text{Roll an } 8|R_4) = \frac{\#\text{ of outcomes that add up to 8, where red shows 4}}{\#\text{ of outcomes in which red shows 4}} = \frac{1}{6}.$$

This agrees with our previous answer.

LPS 31: Introduction to Inductive Logic
Lecture 6

Topics
Probability Calculus
Total Probability Theorem

1 Probability Calculus

Thus far, we have treated probability relatively informally. Today we are going to present the (more) formal theory of probability, known as the probability calculus. Some of this is material that we have already covered. Consider that material review. Some of today's material shows connections between the different rules that we have only hinted at previously. In some cases, we will give the final, formal definitions of concepts that have been introduced on informally before. The rules (including the rule numbers) listed here are the ones we will use for the purposes of this class. I make no effort to make these rules independent of one another.

Rule 1 (Normalization)**.** *For any proposition A, $0 \leq Pr(A) \leq 1$.*

Rule 2 (Tautology)**.** *For any proposition A, $Pr(A) = 1$ if and only if A is necessarily true.*

So far we have encountered two cases of necessary truth. A is necessarily true if it is either (a) a tautology or (b) if it is sure to occur in a particular trial of a chance setup, given the definition of the setup.

Rule 3 (Contradiction)**.** *For any proposition A, $Pr(A) = 0$ if and only if $\sim A$ is necessarily true.*

Again, we have encountered two such cases. $\sim A$ is necessarily true if either (a) A is a logical contradiction, or (b) A is impossible in the context of a particular chance setup.

Rule 4 (Equivalence)**.** *For any two propositions A and B, if A and B are logically equivalent, then $Pr(A) = Pr(B)$.*

We have said this before, but now it is official.

Rule 5 (Special Disjunction)**.** *A and B are mutually exclusive propositions if and only if $Pr(A \vee B) = Pr(A) + Pr(B)$.*

(This can be taken as a definition of mutually exclusive. Note that, in conjunction with rule 8 below, this rule implies that if A and B are mutually exclusive, then $\Pr(A \wedge B) = 0$. This, too, could be taken as a definition of mutually exclusive propositions.)

Rules 2 and 5 together allow us to deduce Rule 6. Let A be some proposition. Then $A \vee (\sim A)$ is a tautology. But A and $\sim A$ are mutually exclusive. Thus, $1 = \Pr(A \vee (\sim A)) = \Pr(A) + \Pr(\sim A)$ and so, subtracting $\Pr(\sim A)$ from both sides, we find $\Pr(A) = 1 - \Pr(\sim A)$.

Rule 6 (Special Exhaustion)**.** *For any proposition A, $Pr(A) = 1 - Pr(\sim A)$.*

Rule 6 can also be understood as a special case of the following rule, which makes our notion of "exhaustion" more precise (this could be taken as a definition of "exhaustive propositions").

Rule 7 (General Exhaustion)**.** *A list of propositions $A_1, \ldots, A_n$ is exhaustive if and only if $Pr(A_1 \vee \cdots \vee A_n) = 1$.*

In general, we will limit attention to exhaustive lists of pairwise mutually exclusive propositions.

Rules 4 and 5 also allow us to derive a more general rule for disjunctions. Let A and B be any two propositions. The consider the following three propositions: $A \wedge B$, $(\sim A) \wedge B$, and $A \wedge (\sim B)$. These are all clearly mutually exclusive, though one might check using truth tables. Since they are mutually exclusive, we can use Rule 5 to show that:

$$\Pr((A \wedge B) \vee (A \wedge (\sim B))) = \Pr(A \wedge B) + \Pr(A \wedge (\sim B)) \tag{1}$$
$$\Pr((A \wedge B) \vee ((\sim A) \wedge B)) = \Pr(A \wedge B) + \Pr((\sim A) \wedge B) \tag{2}$$
$$\Pr((A \wedge B) \vee ((A \wedge (\sim B)) \vee ((\sim A) \wedge B))) = \Pr(A \wedge B) + \Pr(A \wedge (\sim B)) + \Pr((\sim A) \wedge B) \tag{3}$$

However, it is easy to show that: $(A \wedge B) \vee (A \wedge (\sim B))$ is logically equivalent to A; $(A \wedge B) \vee ((\sim A) \wedge B)$ is logically equivalent to B; and $(A \wedge B) \vee ((A \wedge (\sim B)) \vee ((\sim A) \wedge B))$ is logically equivalent to $p \vee q$. This means we can rewrite Eqs. (1) and (2) as

$$\Pr(A) = \Pr(A \wedge B) + \Pr(A \wedge (\sim B)) \tag{4}$$
$$\Pr(B) = \Pr(A \wedge B) + \Pr((\sim A) \wedge B) \tag{5}$$

Adding both sides, we can conclude that

$$\Pr(A) + \Pr(B) = 2\Pr(A \wedge B) + \Pr(A \wedge (\sim B)) + \Pr((\sim A) \wedge B),$$

which yields (by subtracting $\Pr(A \wedge B)$ from both sides),

$$\Pr(A) + \Pr(B) - \Pr(A \wedge B) = \Pr(A \wedge B) + \Pr(A \wedge (\sim B)) + \Pr((\sim A) \wedge B). \tag{6}$$

Now consider Eq. (3). Using the logical equivalence above, we can write this as

$$\Pr(A \vee B) = \Pr(A \wedge B) + \Pr(A \wedge (\sim B)) + \Pr((\sim A) \wedge B) \tag{7}$$

Now, though, if we examine Eqs. (6) and (7), we see that their right hands sides are identical. This means that the left hands sides must be equal, too. We conclude that,

Rule 8 (General Disjunction)**.** *For any propositions A and B, $Pr(A \vee B) = Pr(A) + Pr(B) - Pr(A \wedge B)$.*

That does it for disjunction. Let's now move to conditional probability and conjunction. We have already defined conditional probability, but now let's enshrine it as a rule (really, it's a definition, but who's counting).

Rule 9 (Conditional Probability)**.** *Let A and B be propositions, with $Pr(B) \neq 0$. Then the conditional probability of A assuming B is $Pr(A|B) = Pr(A \wedge B)/Pr(B)$.*

Rule 9 suggests a way of understanding conjunction. (Actually, it is not a strict derivation—Rule 10 is more general.)

Rule 10 (General Conjunction)**.** *For any propositions A and B, if either $Pr(B) = 0$ or $Pr(A) = 0$, then $Pr(A \wedge B) = 0$. Otherwise, $Pr(A \wedge B) = Pr(A|B) \times Pr(B)$.*

Rule 9 allows us to make our earlier notion of independence more precise. (This also could count as a definition.)

Rule 11 (Independence)**.** *Two propositions A and B are independent if and only if either $Pr(B) = 0$ or $Pr(A|B) = Pr(A)$.*

You might worry that there is an asymmetry here. But it doesn't matter. First, suppose $\Pr(A) = 0$. Then $\Pr(A|B) = \Pr(A \wedge B)/\Pr(B) = 0 = \Pr(A)$, so A and B are automatically independent. Meanwhile, if $\Pr(A) \neq 0$ and $\Pr(A|B) = \Pr(A)$, then by rule 9, $\Pr(A \wedge B)/\Pr(B) = \Pr(A)$, which implies $\Pr(A \wedge B) = \Pr(A) \times \Pr(B)$. But by rule 4, since $A \wedge B$ is equivalent to $B \wedge A$, $\Pr(A \wedge B) = \Pr(B \wedge A)$, and so $\Pr(B|A) = \Pr(B \wedge A)/\Pr(A) = \Pr(A) \times \Pr(B)/\Pr(A) = \Pr(B)$. Thus $\Pr(A|B) = \Pr(A)$ implies $\Pr(B|A) = \Pr(B)$, and so no asymmetry actually arises. Incidentally, we have proved another rule.

Rule 12 (Special Disjunction)**.** *Let A and B be independent propositions. Then $Pr(A \wedge B) = Pr(A) \times Pr(B)$.*

These are almost all of the rules of probability. However, there are two more very important rules (they might be called theorems, because they are important derived consequences of the other rules) that we will discuss.

2 Total probability theorem

Suppose that $A_1, \ldots, A_n$ are mutually exclusive, exhaustive propositions, and suppose that B is any other proposition. First, I claim that the proposition $(B \wedge A_1) \vee (B \wedge A_2) \vee \cdots \vee (B \wedge A_n)$ is logically equivalent to B. There are several ways to see why this is. The simplest one is to note that, since $A_1, \ldots, A_n$ are mutually exclusive and exhaustive, exactly one of them is always true. Suppose A_i is the true one. This means that the whole proposition is true just in case the conjunction $B \wedge A_i$ is true, since all of the other conjunctions $B \wedge A_j$ for $i \neq j$ are false by construction. But since A_i is true, $B \wedge A_i$ is true just in case B is true. Thus the whole thing is true if and only if B is true, which means the proposition is logically equivalent to B.

This means, by Rule 4, that $\Pr(B) = \Pr((B \wedge A_1) \vee (B \wedge A_2) \vee \cdots \vee (B \wedge A_n))$. Moreover, each of the terms $(B \wedge A_i)$ is mutually exclusive of each of the others, simply because each of the A_i propositions is mutually exclusive of each of the others. Thus by rule 5 we can write the right hand side of this equation as $\Pr((B \wedge A_1) \vee (B \wedge A_2) \vee \cdots \vee (B \wedge A_n)) = \Pr(B \wedge A_1) + \Pr(B \wedge A_2) + \ldots + \Pr(B \wedge A_n)$. But then, by Rule 10, $\Pr(B \wedge A_1) + \Pr(B \wedge A_2) + \ldots + \Pr(B \wedge A_n) = \Pr(B|A_1) \times \Pr(A_1) + \Pr(B|A_2) \times \Pr(A_2) + \ldots + \Pr(B|A_n) \times \Pr(A_n)$. This allows us to conclude

Rule 13 (Total Probability Theorem)**.** *Given any list of mutually exclusive, exhaustive propositions $A_1, A_2, \ldots, A_n$ and any proposition B,*

$$Pr(B) = Pr(B|A_1) \times Pr(A_1) + Pr(B|A_2) \times Pr(A_2) + \ldots + Pr(B|A_n) \times Pr(A_n).$$

This is an important result. Intuitively, it means that if you want to know whether a given proposition is true, you can ask whether it is true under various mutually exclusive, exhaustive circumstances, and then derive the general probability from that. Here's an example of how the theorem can be useful. Suppose you have two urns. Urn 1 has 3 black balls and 5 green balls. Urn 2 has 4 black balls and 1 green ball. Suppose I flip a coin to decide which urn to draw from. If the coin lands heads, then I draw from urn 1. If it lands tails, I draw from urn 2. What is the probability of drawing a green ball? Rule 13 gives us an easy way to think about this. (There's a harder way, too, using things we've already learned.) It tells us that $\Pr(G) = \Pr(G|H) \times \Pr(H) + \Pr(G|T) \times \Pr(T)$ But $\Pr(G|H)$ is just the probability of drawing a green ball from urn 1, which is 5/8. And $\Pr(G|T)$ is the probability of drawing a green ball from urn 2, which is 1/5. So $\Pr(G) = (5/8) \times (1/2) + (1/5) \times (1/2) = 5/16 + 1/10 = 33/80$.

You could have done this calculation without using Rule 13, because flipping the coin is independent of drawing the balls. But in cases where the relevant propositions are not independent, the theorem proves essential. Suppose, now, that we change the setup above as follows. First, we flip the coin to choose an urn. Then we draw a ball. Then we flip the coin again to determine where to put the ball back in (heads, again, is urn 1, and tails is urn 2). Now we can ask various questions. For instance, what is the probability of drawing a green ball from urn 1 on this second round? This is a much harder calculation, since the probability depends on the particular chain of events. This means we can't use the rule of special conjunction. So how should we think about it?

The way to do this sort of problem is to first break it up into cases that you *can* assign probabilities to. In case 1, the coin comes up heads, you draw a green ball, and then the coin comes up heads again and the green ball goes back in the first urn. We can write this as $C_1 = H_1 \wedge G_1 \wedge H_2$. (The subscript on the H corresponds to the time you've flipped the coin; the subscript on the G corresponds to the time that you've pulled the ball out.) Case 2, then, will be that we get heads, draw a green ball, and then get tails, and the green ball does into the second urn. We can write this as $C_2 = H_1 \wedge G_1 \wedge T_2$. And so on. We get 8 cases: $C_3 = H_1 \wedge B_1 \wedge H_2$, $C_4 = H_1 \wedge B_1 \wedge T_2$, $C_5 = T_1 \wedge G_1 \wedge H_2$, $C_6 = T_1 \wedge G_1 \wedge T_2$, $C_7 = T_1 \wedge B_1 \wedge H_2$, and $C_8 = T_1 \wedge B_1 \wedge T_2$. Given the chance setup, these are mutually exclusive and they are exhaustive. This means we can apply Rule 13. This gives

$$\begin{aligned}\Pr(G_2) = {} & \Pr(G_2|C_1) \times \Pr(C_1) + \Pr(G_2|C_2) \times \Pr(C_2) + \Pr(G_2|C_3) \times \Pr(C_3) \\ & + \Pr(G_2|C_4) \times \Pr(C_4) + \Pr(G_2|C_5) \times \Pr(C_5) + \Pr(G_2|C_6) \times \Pr(C_6) \\ & + \Pr(G_2|C_7) \times \Pr(C_7) + \Pr(G_2|C_8) \times \Pr(C_8).\end{aligned}$$

We have the apparatus, though, to calculate each of these terms. Consider case 1. We know $C_1 = H_1 \wedge G_1 \wedge H_2$. But the second coin toss is independent of everything that came before, and so $\Pr(H_1 \wedge G_1 \wedge H_2) = \Pr(H_1 \wedge G_1) \times H_2$. And $\Pr(H_1 \wedge G_1) = \Pr(G_1|H_1) \times \Pr(H_1)$, where $\Pr(G_1|H_1)$ is just the initial probability of drawing a green ball from urn 1. So

$\Pr(C_1) = \Pr(H_1) \times \Pr(H_2) \times \Pr(\text{Green from urn 1.}) = (1/2) \times (1/2) \times (5/8) = 5/32$. Similar calculations give $\Pr(C_2) = \Pr(H_1) \times \Pr(T_2) \times \Pr(\text{Green from urn 1.}) = 5/32$. Likewise, $\Pr(C_3) = \Pr(C_4) = (1/4) \times (3/8) = 3/32$; $\Pr(C_5) = \Pr(C_6) = (1/4) \times (1/5) = 1/20$; and $\Pr(C_7) = \Pr(C_8) = (1/4) \times (4/5) = 1/5$. But what about $\Pr(G_2|C_1)$, and the other terms like that? This requires a little thought. This is the probability of drawing a green ball from urn 1, assuming we took a green ball out of urn 1 and put it back into urn 1. So, this is the same as the probability of drawing a green ball initially (because the distributions haven't changed). Thus $\Pr(G_2|C_1) = 5/8$. In the second case, though, we have moved a green ball from urn 1 to urn 2. This means that now urn 1 has 4 green balls and 3 black ones, whereas urn 2 has 4 black balls and 2 green ones. So $\Pr(G_2|C_2) = 4/7$. Similarly for the others, $\Pr(G_2|C_3) = 5/8$ (because the black ball was removed from urn 1, and then replaced in urn 1); $\Pr(G_2|C_4) = 5/7$ (because a black ball was moved from urn 1 to urn 2); $\Pr(G_2|C_5) = 6/9$ (because a green ball has been moved from urn 2 to urn 1); $\Pr(G_2|C_6) = 5/8$ (because a green ball is removed from urn 2 and put back in urn 2, leaving urn 1 intact); $\Pr(G_2|C_7) = 5/9$ (because a black ball is moved from urn 2 to urn 1); and $\Pr(G_2|C_8) = 5/8$ (because urn 1 remains intact).

Putting all of this together, we find:

$$\begin{aligned}
\Pr(G_2) &= (5/8) \times (5/32) + (4/7) \times (5/32) + (5/8) \times (3/32) + (5/7) \times (3/32) \\
&\quad + (6/9) \times (1/20) + (5/8) \times (1/20) + (5/9) \times (1/5) + (5/8) \times (1/5) \\
&= 25/256 + 5/56 + 15/256 + 5/224 + 1/30 + 1/32 + 1/9 + 1/8 \\
&\approx .57
\end{aligned}$$

The probability of drawing a green ball from urn 1 with the original distributions, meanwhile, was 5/8=.625. So the probabilities have changed, given the (potential) redistribution.

LPS 31: Introduction to Inductive Logic
Lecture 7

Topics
Bayes' Theorem

1 Bayes' Theorem

Thus far, we have used conditional probability to ask questions of the following sort: What is the probability that a particular proposition is true, given that some other proposition is true? For instance,

- What is the probability that I will bring an umbrella to work today, given that it is raining?
- What is the probability that you will do well on the next midterm, given that you studied?
- If you draw two cards from a standard deck, what is the probability that the second one will be an ace given that the first one was an ace?

For each of these cases, however, we can also ask what you might take to be the opposite question.

- What is the probability that it is raining, given that I have an umbrella today?
- What is the probability that you studied, given that you did well on the midterm?
- If you draw two cards from a standard deck, what is the probability that the first one was an ace, given that the second one was an ace.

The first thing to emphasize is that these questions really are different from one another. We are not just asking the same thing in different terms. If the first questions are of the form $\Pr(A|B)$, then the second questions are $\Pr(B|A)$. In the first case, we assume one thing is true, and ask how that affects the probability of the other; in the second questions, we reverse the relationship. We assume we have an umbrella, and ask how it affects the chances of rain.

This second kind of question can seem a little weird at first—less intuitive, in a way. It is easy to see how, intuitively speaking anyway, the fact that it is raining can *make* you bring an umbrella. In probabilistic terms, we can see how the fact that it is raining can make it more likely that someone is carrying an umbrella. (Similarly, we can make sense of the idea that the fact that it is *not* raining can make it less likely that someone is carrying an umbrella.) By why should the fact that I have an umbrella make any difference to the weather? I can't make it rain by carrying an umbrella (or keep it sunny by leaving my umbrella at home).

This apparent difference arises because we tend to think about the relationship between two (dependent) propositions in terms of causation. Rain can cause us to carry an umbrella. Studying can cause you to do well on an exam. But causal arrows tend to go in one direction:

it seems wrong to say that carrying an umbrella can cause it to rain. Why does this seem wrong? One problem concerns what kinds of things can have influence over other things. We do not usually think we can affect the weather—least of all by carrying an umbrella. Moreover, we usually assume that a cause precedes its effects. So, if you do well on a midterm, that cannot change whether you studied in advance. Studying can cause you to do well, but not vice versa.

Unfortunately, causation has proved to be a very difficult concept to make precise. We aren't going to try to say what it really means for one thing to cause another. In fact, in the present context, when trying to understand conditional probability, it is most useful to try to ignore your instincts about what causes what. From the point of view of probability theory, the two kinds of question noted above are treated in exactly the same way. When we write $\Pr(A|B)$, there is nothing distinctive or special about A and B—we do not require, somehow, that B causes A. We do not require anything of A and B at all. Indeed, we have a precise definition of $\Pr(A|B)$, using rule 9.

Rule 9 (Conditional Probability)**.** *Let A and B be propositions, with $Pr(B) \neq 0$. Then the conditional probability of A assuming B is $Pr(A|B) = Pr(A \wedge B)/Pr(B)$.*

In this rule, A and B are arbitrary propositions. This means that, for fixed propositions A and B, we can use the rule to calculate

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

Thus, although questions like "What is the probability that it is raining, given that I have an umbrella" might seem confusing, there is nothing mysterious about the relevant calculations.

$$\Pr(\text{umbrella}|\text{rain}) = \frac{\Pr(\text{umbrella and rain})}{\text{rain}}$$
$$\Pr(\text{rain}|\text{umbrella}) = \frac{\Pr(\text{rain and umbrella})}{\text{umbrella}}$$

The important thing isn't what *makes* something true, it's whether one thing is true more often than usual when a second thing is true.

One of the most important results in basic probability theory provides a way of relating the probabilities $\Pr(A|B)$ and $\Pr(B|A)$ for any propositions A and B (as long as $\Pr(A) \neq 0$ and $\Pr(B) \neq 0$). This result is our 14th and final rule of probability theory. It is known as Bayes' theorem. It's a simple consequence of what we already know. First, consider the formula from rule 9, $\Pr(A|B) = \Pr(A \wedge B)/\Pr(B)$. This can be written as in rule 10, so that $\Pr(A \wedge B) = \Pr(A|B) \times \Pr(B)$. An identical calculation can be done for $\Pr(B|A)$, to yield $\Pr(B \wedge A) = \Pr(B|A) \times \Pr(A)$. Finally, using rule 4, we know that $\Pr(A \wedge B) = \Pr(B \wedge A)$ because $A \wedge B$ and $B \wedge A$ are logically equivalent. This allows us to conclude that $\Pr(A|B) \times \Pr(B) = \Pr(B|A) \times \Pr(A)$. This, essentially, is Bayes' theorem, though it is usually written in a slightly different form, as in the following rule.

Rule 14 (Bayes' Theorem)**.** *Let A and B be propositions, with $Pr(A) \neq 0$ and $Pr(B) \neq 0$. Then the conditional probability of A assuming B is*

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

It is often useful to write the denominator in Bayes' theorem using the total probability theorem (rule 13). This gives the equivalent result,

$$\Pr(A|B) = \frac{\Pr(B|A) \times \Pr(A)}{\Pr(B|A) \times \Pr(A) + \Pr(B| \sim A) \times \Pr(\sim A)}.$$

Bayes' theorem will play an essential role in what follows in this course. It is most useful when trying to evaluate new evidence. Consider the following example. Suppose that you are a doctor and a patient comes in showing the symptoms of a rare and deadly disease. Suppose that this disease afflicts about 1 out of every 10000 people in the world. Fortunately, however, there is a test that can be performed to check whether the person has the disease. The test is very good: 99% of the people who have the disease and are given the test receive positive results (this is called the sensitivity of the test). Sometimes there are false positives—people who don't have the disease getting positive results—though these are also rare. 98% of the people who *don't* have the disease but who take the test get negative results (this is called the specificity of the test). You give the patient this test, and the results come back positive. What is the probability that the patient has the disease?

Most people have the instinct that with such a good test, the likelihood that someone who has gotten a positive result has the disease is high. But let's see. This is just the kind of case where Bayes' theorem is useful. Let D represent the proposition "the patient has the disease". Let P be the proposition "the test results were positive." Then, we have the following probabilities given to us already. $\Pr(D)$ is just the probability that a person has the disease, with no other information. In this case, $\Pr(D) = 1/10000$. This means that $\Pr(\sim D) = 1 - 1/10000 = 9999/10000$. $\Pr(P|D)$ is the probability of getting a positive result, given that the patient has the disease. This is the sensitivity of the test. We have $\Pr(P|D) = .99$. We also know the probability of getting a positive result given that you don't have the disease, which is $\Pr(P| \sim D) = 1 - \Pr(\sim P| \sim D) = 1 - .98 = .02$.

We can now plug this into Bayes' theorem. We want $\Pr(D|P)$, which is,

$$\begin{aligned}\Pr(D|P) &= \frac{\Pr(P|D) \times \Pr(D)}{\Pr(P|D) \times \Pr(D) + \Pr(P| \sim D) \times \Pr(\sim D)} \\ &= \frac{(.99) \times (.0001)}{(.99) \times (.0001) + (.02) \times (.9999)} \approx .005\end{aligned}$$

This should lead to a big sigh of relief. It means the probability that the patient has the disease, even after a positive result, is only twice as likely as it was to begin with. Why would this be? It's because false positives are quite common relative to how frequently the disease actually appears.

LPS 31: Introduction to Inductive Logic
Lecture 8

Topics
Acts and Consequences
Utility
Expected Value
Fair Price

1 Acts and Consequences

After the midterm next week, we are going to return to Bayes' theorem and to questions concerning evidence and learning of the sort that originally motivated our investigation into probability theory. We will also spend some time talking about the philosophical foundations of probability. Today and in the next class, though, we are going to take a brief detour and discuss an application of probability theory called decision theory.

Decision theory is a way of analyzing choices that a person can make. The basic idea, at least at the level at which we are going to engage with the subject, is that before making a choice, we want to understand its consequences. Often, however, the consequences of a particular choice are uncertain. If you decide to drive a car, for instance, you take certain risks. Most of the time, you will safely arrive at your destination. But some people have car accidents, and so there are possible downsides to the decision to drive a car. We want to know how to weigh these positive and negative consequences so that we can make an informed and rational choice about how to act. The idea is that probability gives us a way to deal with the uncertainty associated with this kind of decision.

Our focus in these two lectures is going to be on *acts* and their possible *consequences*. An act is a thing that you might choose to do. For instance, you might choose to drive to L.A. Driving to L.A. is an act. You might also choose to bet a dollar that a coin will come up heads. This is another act. A third example of an act might be buying a lottery ticket. Each of these acts corresponds to a decision that you might make, or that you might decide not to make.

Each of these acts also has possible consequences. It is difficult to list all of the possible consequences of driving to L.A.—you might arrive safely at your destination in time for your appointment; you might get lost or hit traffic and miss your appointment; you might get into an accident on the 405 freeway and damage your car; you could get into an accident and injure yourself or even die. There are likely other possible consequences of the act of driving to L.A. If you make a bet on a flip of a coin, the consequences are a little easier to state. One possible consequence is that the coin will come up heads and you will win \$1. The other possible consequence is that the coin will come up tails and you will lose \$1. (Note the connection between possible consequences and possible outcomes for the corresponding chance setup—each of the possible consequences corresponds to a possible outcome, but with additional information regarding how you fared in your gamble.) The possible outcomes of buying a lottery ticket are that you either win the lottery and make millions of dollars, or you do not win the lottery and you lose the cost of the ticket.

In each of these cases, the consequences of your action could be positive or negative. Decision theory attempts to answer the following question. Given that these actions have the capacity to have both positive and negative consequences, should you perform the action? In other words, under what circumstances does the good outweigh the bad?

2 Utility

The first step in answering the question of when the good consequences of an act outweigh the bad is to figure out a way to quantify what we mean by "good" and "bad". This is done by assigning a number, called *utility*, to each of the consequences. Utility is simply a measure of "goodness". A positive number for utility indicates that, on a whole, a particular consequence is good for you. A negative number indicates that a particular consequence is bad for you. The larger the number, the better the consequence. Utility is often measured in units called "utiles." Utiles don't really correspond to anything in the world, they're just a convenient way of quantifying how good something is for you. Sometimes it is more convenient to measure utility in terms of money.

In some cases, particularly in cases where a bet is being made, the most natural way of assigning utilities is in terms of the financial ramifications of each of the possible outcomes. So, in the example given above of gambling on the flip of a coin, you might say that the utility of guessing correctly and winning a dollar is 1 utile. (Alternatively, you might just say that the utility of winning the coin flip is 1$.) Conversely, the utility of *losing* the coin flip would be -1 utile (or -1$). These numbers mean that winning the coin flip is good for you, whereas losing is bad. Moreover, the quantity of goodness that is associated with winning the coin flip is the same as the quantity of badness associated with losing.

As another example, suppose that the lottery described above works as follows. Each ticket costs 1$. If you win the lottery, you receive 1,000,000$ (a million dollars). If you lose, you don't get anything. In this case, one natural way to assign utilities would be to say that the utility associated with winning the lottery is 1,000,000 utiles (or 1,000,000$). The utility of losing the lottery is -1 utile (or -1$). This means that the quantity of goodness that is associated with winning the lottery is a million times greater than the quantity of badness associated with losing.

There are no strict rules for how you assign utilities. In particular, the utilities of various outcomes can vary from person to person. For instance, suppose that you very much want to buy a yacht that costs 1,000,000$. However, you don't have that kind of money sitting in your bank account, and so you would get a lot of benefit from suddenly having a million dollars to spend. A dollar, however, doesn't mean much to you—it's just a dollar, and losing a dollar doesn't have much effect on your long-term yacht aspirations. Under these circumstances, you might say that the amount of goodness associated with winning the lottery is more than 1,000,000 utiles, because that million dollars could be put to immediate use to purchase something worth far more to you personally. Under these circumstances, you might assign a higher number to the utility of winning the lottery, a number that more accurately captures how badly you would like to own a yacht. You might assign a number like 5 million, or 10 million, or a billion utiles (depending on how badly you want to own the yacht). But since a dollar is still a dollar, you might continue to assign a utility of -1 utiles to the case in which you lose the lottery.

The same kind of reasoning can go the other way as well. Suppose that you are very poor and a dollar could be the difference between having dinner or not having dinner. In this case, winning a million dollars would be great. But losing your last dollar would be disastrous because it could mean you starve. Under these circumstances, you might assign a larger negative number to the outcome where you lose the dollar, say 5 utiles, or 10 utiles, or maybe a million utiles if you are so hungry that you might die if you don't eat a meal. Winning the lottery, meanwhile, would be nice, but not essential, and so you might continue to assign that possible consequence a value of 1,000,000 utiles.

The point of these examples is to say that utility can be a very subjective matter. How much utility you assign to a possible consequence of an act depends entirely on your particular situation and it can vary from person to person. Decision theory doesn't tell you how to assign utilities, only what to do to evaluate possible acts once you have established what the utilities of their consequences would be. For this reason, decision theory doesn't tell you how to act. It gives you a way to think about possible actions, but only after you have figured out how good the consequences of those outcomes would be for you.

This point is especially clear in the context of the car trip described above. In general, it will be difficult to assign financial values to things like arriving late on account of traffic, or even worse, to sustaining an injury in a car accident. This means that utilities can only be assigned to these consequences on the basis of subjective judgments of how good or bad these would be for you.

3 Expected Value

We have now seen some examples of ways in which you might assign utilities to the consequences of an act. There is no unique way to do this—it depends on the circumstances of the person assigning the utilities. However, once you *have* assigned utilities to the possible consequences of an act, probability theory gives you a way of deciding whether the act is, on the whole, good to perform. It also gives you a way of saying *how* good the act is. Note that we don't mean good or bad in an ethical sense. This isn't a theory of whether an act is morally right or morally wrong. Instead, it is a theory of how good an act can be expected to be for you, in the same sense of good and bad that is captured by utility.

The tool for establishing how good an act with uncertain consequences is for you, on the whole, is called the *expected value* of that act. We will represent acts by bold-faced letters, like **A**. We will represent the possible consequences of an act by non-bold-faced letters, in the same way as propositions, so the n possible consequences of act **A** might be written $C_1, C_2, \ldots, C_n$. In fact, each of these can be understood as a proposition of the form "Such and such consequence occurs." Finally, we can represent the utility of a given consequence C by $\mathrm{U}(C)$. Then the expected value of an act **A**, written $\mathrm{EV}(\mathbf{A})$, is given by the formula

$$\mathrm{EV}(\mathbf{A}) = \mathrm{U}(C_1) \times \Pr(C_1) + \mathrm{U}(C_2) \times \Pr(C_2) + \cdots + \mathrm{U}(C_n) \times \Pr(C_n).$$

What does this formula mean? It means that the expected value of an act is given by the utility of its first possible consequence times the probability of that consequence being realized, plus the utility of the second possible consequence times the probability of that consequence being realized, and so on for all of the possible consequences of the act. The expected value of an act is also measured in units of utiles.

Let's see how this works in practice. Consider the coin flipping example above. Call the act of betting on the coin **B**. Then call the consequence where you get heads and earn 1$ H and the consequence where you get tails and lose 1$ T. Then the formula above tells us that,

$$\mathrm{EV}(\mathbf{B}) = \mathrm{U}(H) \times \Pr(H) + \mathrm{U}(T) \times \Pr(T) = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0.$$

This means that the expected value of the coin flipping gamble is zero, which in turn means that taken as a whole, the act is neither good or bad for you. The good cancels the bad.

Suppose we changed the example just a little bit, and instead of winning 1$ if you get heads, instead you win 2$. If you get tails, meanwhile, you still lose 1$. What is the expected value then? The calculation only changes a little. We get,

$$\mathrm{EV}(\mathbf{B}) = \mathrm{U}(H) \times \Pr(H) + \mathrm{U}(T) \times \Pr(T) = 2 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 1.$$

This means that the modified bet has a *positive* expectation value. Taken as a whole, the good consequences of the act outweigh the bad. From the point of view of simple decision theory, you should choose to perform this act.

Let's take another example. Suppose that someone offered to make the following bet, based on rolling an unbiased, standard 6-sided die. The bet works as follows. You roll the die. If an odd number comes up, your opponent agrees to pay you the face value of the die in dollars. If an even number comes up, you have to pay your opponent the face value on the die, also in dollars. If we assume that one dollar corresponds to 1 utile, then we have 6 possible consequences. If the die lands with a 1 facing up, then you make a dollar. This consequence has positive utility $\mathrm{U}(D_1) = 1$. If the die lands with a 2 facing up, then you have to pay two dollars. This consequence has negative utility for you, with $\mathrm{U}(D_2) = -2$. For the other consequences, we have $\mathrm{U}(D_3) = 3$, $\mathrm{U}(D_4) = -4$, $\mathrm{U}(D_5) = 5$ and $\mathrm{U}(D_6) = -6$. Call the act of the bet **D**. Then the expected value is given by,

$$\begin{aligned}\mathrm{EV}(\mathbf{D}) &= \mathrm{U}(D_1) \times \Pr(D_1) + \mathrm{U}(D_2) \times \Pr(D_2) + utD_3 \times \Pr(D_3) + \mathrm{U}(D_4) \times \Pr(D_4) \\ &\quad + \mathrm{U}(D_5) \times \Pr(D_5) + \mathrm{U}(D_6) \times \Pr(D_6) \\ &= 1 \times (1/6) - 2 \times (1/6) + 3 \times (1/6) - 4 \times (1/6) + 5 \times (1/6) - 6 \times (1/6) \\ &= -3 \times (1/6) = -1/2\end{aligned}$$

Should you take this bet? No. Taken as a whole, the bad consequences outweigh the good. You should not choose to perform the act.

4 Fair Bet

If the expected value of an act is zero, then we call that act a *fair bet.* (Sometimes, if the act isn't a gamble, we say that it has a *fair price.*) If the expected value of an act is positive, then that act is *favorable* or a *favorable bet.* Conversely, if the expected value of an act is negative then we say that that act is *unfavorable.* Betting a dollar on an unbiased coin is a fair bet. You and your opponent are equally likely to walk away happy—no one has an advantage over the other. Betting a dollar on an unbiased coin, but where you get two dollars back if you win is a favorable bet. The odds are in favor of you making money on that action. Conversely, the die game described above is an unfavorable bet. Odds are, you will lose money. (Note that if a bet is unfavorable for you, then it is favorable for your opponent).

Sometimes it is useful to work backwards and figure out what you would have to win in order to make a bet fair. (Alternatively, if you stipulate how much you win in a particular consequence, you might ask how much you would have to pay, i.e., how much you would lose if you lost the bet, to make the bet fair.) So suppose you are playing roulette in a casino. The probability of winning is 1/38, but if you win you only get 35$. (You also get the dollar you bet back, but that doesn't count as part of the utility of winning. After all, you had the dollar to begin with!) So what is the expected value of betting 1$ on a single number in roulette? The utility of winning is $\mathrm{U}(W) = 35$, and the probability of winning is $\Pr(W) = 1/38$. The utility of losing is $\mathrm{U}(L) = -1$, with probability $\Pr(L) = 37/38$. This means that the expected value is

$$\mathrm{EV}(\mathbf{R}) = \mathrm{U}(W) \times \Pr(W) + \mathrm{U}(L) \times \Pr(L) = 35 \times (1/38) - 1 \times (37/38) = -2/38 = -1/19.$$

This means that roulette is an unfavorable bet. (Indeed, every game you can play in a casino is an unfavorable bet!)

But what would the payout have to be in order to make it a fair bet? That is, how much would the casino need to give you to make wagering a dollar on roulette fair? To answer this question we want to find the value of $\mathrm{U}(W)$ such that $\mathrm{EV}(\mathbf{R}) = 0$, because a fair bet is one with an expected value of 0. This means we want $\mathrm{EV}(\mathbf{R}) = \mathrm{U}(W) \times (1/38) - 1 \times (37/38) = 0$. The solution is given by $\mathrm{U}(W) = (37/38) \times (38) = 37$. In other words, roulette would need to pay you 37$ every time you won to cancel the possibility of losing a dollar every time you lost.

LPS 31: Introduction to Inductive Logic
Lecture 9

Topics

More expected value calculations
Expected value rule
Risk aversion
Generalized expected value rule

1 More expected value calculations

On Tuesday, we discussed a notion of the expected value of an act. Today I am going to start with an additional expected value calculations. Imagine you are buying a new computer. The new computer costs \$1,000. At the store, as you prepare to check out, the salesman offers to sell you an in-store warranty. The warranty is a kind of insurance policy. For just \$200, the salesman says, the store will replace the computer with a new computer of comparable value if there is a hardware malfunction any time during the first two years that you own the machine.

First, suppose that the store thinks taking the insurance is a fair bet. What probability does the store assign the possibility that the computer will fail during its first two years in service? This can be calculated by going backwards from the assumptions given, if we assume that the store is solely interested in profit. There are two possible cases: either the computer breaks during the first two years, in which case the store has to pay $\$1,000 - \$200 = \$800$; or else, the computer does not break, in which case the store gains \$200. Since we are assuming the bet is fair, we know that the expected value of selling the insurance equals 0 and so,

$$\begin{aligned} \mathrm{EV}(\mathbf{S}) &= \mathrm{U}(C_1) \times \Pr(C_1) + \mathrm{U}(C_2) \times \Pr(C_2) \\ &= (-800)\Pr(C_1) + 200(1 - \Pr(C_1)) = 0 \end{aligned}$$

This means that $-1000\Pr(C_1) = -200$, and so $\Pr(C_1) = 2/10 = .2$.

Suppose you happen to know, based on some independent research, that the computer you are buying has a 15% failure rate in the first two years. What is the expected value, for you, of taking the insurance, assuming you are only interested in the money? First, note that the utilities of the consequences of the act of buying the insurance are just the reverse of those for the store selling the computer. So, if the computer breaks, nothing will happen, you will get a new computer, basically for the \$200 you already spent. And if the computer doesn't break, you will be out \$200. So, no matter what

$$\begin{aligned} \mathrm{EV}(\mathbf{B}) &= \mathrm{U}(C_1) \times \Pr(C_1) + \mathrm{U}(C_2) \times \Pr(C_2) \\ &= -200 \end{aligned}$$

This means that your expected value is -\$200.

2 Expected value rule

If you think about the case above, it begins to look like a raw deal. Buying insurance would be a fair bet for you if the probability of your computer breaking in two years were 30%. But you happen to know that the probability is only 15%, which means that the insurance is actually an unfavorable bet for you. On the whole, the store should expect to make a profit on selling insurance at these rates (especially if they are selling this kind of insurance to many people).

So, should you buy the insurance? Before trying to answer this, let's switch to another example. Suppose someone stops you in the middle of the street and makes the following proposal. Either you have to bet him $20 dollars that a die will come up 6 (if you win, you get $20, otherwise you have to give him $20), or else he's going to steal your wallet. You happen to have $20 in your wallet. What should you do? You have two choices: either you make the bet or you don't. First, let's calculate the expected value of the bet. There's a 1/6 chance that the die will come up 6. This means your expected value is

$$\begin{aligned}\mathrm{EV}(\mathbf{B}) &= \mathrm{U}(C_1) \times \Pr(C_1) + \mathrm{U}(C_2) \times \Pr(C_2) \\ &= (20)(1/6) - 20(5/6) = -80/6 = -13.33\end{aligned}$$

This is a negative expectation. Chances are that you are going to lose money.

But what if you don't agree to the bet? This, too, has ramifications, since you know he is going to steal your wallet. If that happens, you're guaranteed to lose $20. In other words, the expectation value of *not* taking his bet is $\mathrm{EV}(\sim \mathbf{B}) = -20$. This is also negative, but it's *more* negative. In both cases you should plan to lose money, but one of the two acts has a higher expectation value than the other. This suggests a rule.

Rule 1. *If presented with a choice between several acts, choose the act with the largest expected value.*

3 Risk aversion

Now let's return to the insurance question. You might think that the answer is obvious: no, you shouldn't, because the expected value is negative. But be careful. In the situation just described, just as with the example just given, you actually have a choice between two different acts. One of those acts is buying the insurance. The other act is *not* buying insurance. Not buying insurance might seem different than an act—somehow, it is the absence of an act. But the important thing, from a decision theoretic standpoint, is that *not* buying insurance also has possible consequences. In particular, your computer could still break. In deciding between several acts, you cannot look at the expected values of any one act in isolation. You need to compare the expected values of both acts.

So what is the expected value of not buying insurance? Well, there are still two cases to consider. One is that the computer breaks in the first two years. If that happens, you need to shell out $1,000 to buy a new computer (we're assuming that you just cannot live without a computer). The other possibility is that the computer doesn't break—in which case, you don't have to pay anything. You already know the probabilities, and so it is just a matter of plugging everything into the formula.

$$\begin{aligned}\text{EV}(\sim \mathbf{B}) &= \text{U}(C_1) \times \text{Pr}(C_1) + \text{U}(C_2) \times \text{Pr}(C_2) \\ &= (-1000)(3/20) - 0(17/20) = -3000/20 = -150\end{aligned}$$

What does this mean? Well, for one it means that *not* buying the insurance also has a negative expectation value.

But what should you do? If you follow the rule above, then you should not buy the insurance because it has a lower expectation value. But their is a big difference between the two acts that isn't fully captured by the expectation value. In one case, you have to pay $200 no matter what. That's not an insignificant sum, but it does guarantee that if your computer breaks, you will get a new one without having to pay any more money. In the other case you don't have to pay anything up front. But there's a chance that you will have to pay $1,000, the price of a new computer and substantially more than the insurance premium.

Different people have different instincts about which option is more desirable. Some people are inherently *risk-averse*, that is, if given a choice between two options of the same expectation value, they will tend to choose the act that is less risky. In fact, most people are *so* risk-averse that they would be willing to pay *extra* in order to avoid risk. This extra cost is called a *risk premium.* In this case, if you take the $200 insurance policy, then you are paying a risk premium of $50. This $50 is buying you peace of mind. The alternative to risk-averse is usually called *risk-loving.* Some people, for whatever reason, seem to enjoy risk. These people don't mind the possibility of paying $1,000 sometimes, because they are willing to take chances. In most situations where the expected values of two or more acts are similar, there are good reasons to prefer being risk-averse to risk-loving.

LPS 31: Introduction to Inductive Logic
Lecture 10

Topics
Interpretations of Probability
Relative Frequency and Chance

1 Interpretations of Probability

Thus far in the course, we have spent a lot of time thinking about how to assign probabilities to various propositions (and in some cases events or consequences, though we have attempted to use the proposition language whenever possible). All of this work has been predicated on an intuitive notion of "likelihood." But what is "likelihood" supposed to mean?

To see what the question is, consider flipping an unbiased coin. What do we mean when we say that the probability of it landing heads is 1/2? One instinct that many people have is to say that to assign probability 1/2 in this case indicates something about how often the coin is going to land heads if you flip the coin over and over again. Probability, you might think, is saying something like the following: if you flip the coin many times, and the probability of it landing heads is 1/2, then the coin should land heads about 1/2 of the time. In other words, probability is saying something about how *frequently* an outcome will occur, relative to the other possible outcomes. This is known as a *relative frequency* interpretation of probability.

We will say more about the details of this proposal in a moment. But for now, consider a second question. Suppose you buy a raffle ticket. You know that exactly one person is going to win the raffle, and that the raffle is happening just one time with these tickets. What does it mean to say that the probability of *you* winning the raffle is 1/100 (say)? It seems strange, certainly, to thinking about how often you would win it you repeated the same raffle over and over again—after all, each time would be a different raffle. To see the point more clearly, think about a school of fish being chased by a shark. The shark is going to eat just one of the 100 fish. You might reason that the probability that the shark will eat a particular fish is 1/100. But at the end of the day, the fish is either eaten or not. You cannot keep running the chance setup. In these cases, there probability has something to do with "the luck of the draw." One person will win the raffle; one fish will get eaten. It is hard to think about running these scenarios over and over again to extract some notion of relative frequency of each of the outcomes. Instead, it seems like the probability is measuring something more basic, concerning the *chances* of some single thing happening.

Here's a third kind of question. Suppose I say that there's a 90% chance that a certain kind of dinosaur spent most of its time in the water. I might give some evidence for this: I could say, for instance, that a lot of this kind of dinosaur's bones have been found near shallow lakes; I could say that the dinosaur was apparently so large that it is difficult to see how it could have supported its own weight on land. Is this a statement about relative frequency? It would seem not—I am not saying that if this kind of dinosaur were to evolve over and over again, that about 9 out of every 10 times, it would spend most of its time in the water. (It's hard to even figure out how to express this idea.) But I am also not saying that it's the luck of the draw, or that it has anything to do with the chances of a dinosaur evolving in a particular way. Instead, I am expressing something about my *degree of belief* that the dinosaur spent most of its time in the water. This degree of belief may be based in part on some evidence that I can present regarding the dinosaur's habits, but it is ultimately saying something about my own confidence level.

Each of these three options are reasonable interpretations of (some) probability statements. In this class, I am going to take the view that there isn't a single, unique interpretation of probability that applies in all situations. Instead, probability can be used in multiple ways. Indeed, in many cases, a given probability statement can have more than one interpretation. For instance, if I say that the probability that a coin will land heads is 1/2, I might be interpreted as saying something about the relative frequency of heads, as opposed to tails. But I might also be saying something about how *this* trial is going to go—in which case, I might be interpreted as saying something about chance, in the sense of luck of the draw. I could also be expressing my personal degree of belief that the coin is going to come up heads— a degree of belief based, for instance, on a long experience with coins.

Over the next few weeks, we are going to focus on the degree of belief interpretation. But the other interpretations are also important, so I am going to say a little more about them here.

2 Relative Frequency and Chance

There is a sense in which the relative frequency interpretation of probability statements is both the most obvious, and the thorniest. On the one hand, many people *do*, instinctively, interpret most probability statements in this way. And it's reasonable—this is the kind of probability most often associated with statistics, and so we often encounter statements that are naturally interpreted as relative frequency assertions in fields like politics or medicine. If I conduct a poll, for instance, and report the percentage of people who are in favor of candidate 1, and the percentage of people in favor of candidate 2, I am simply reporting frequencies, i.e., how often the people who were polled reported a preference for one candidate or the other. If I take these numbers to reflect the probability that a person picked randomly off the street would support one candidate or the other, then that kind of probability would most naturally be interpreted as a relative frequency assertion. The poll has given me some indication of what would happen if I stopped a lot of people randomly on the street and asked them about their political beliefs.

Similarly, if I say that there's a ten percent chance that a person is left handed, that usually means that one in every ten people is left handed. But it *doesn't* mean that if I went out an collected ten people, that one of them would *have* to be left handed. And it doesn't mean that if I talked to nine right handed people, that the next person I talked to would have to be left handed. (This would be gambler's fallacy style reasoning.) This is what makes relative frequency interpretations a little bit difficult to make sense of. They say something about how often a thing occurs, but just what that is is subtle.

The point is clearest with the coin flipping example. Suppose I flip a coin over and over again. Am I ever guaranteed to get exactly the 50-50 split suggested by the probability? No, at least not if the trials are independent. (This is especially obvious is you consider an odd number of flips—there's no way that these could break down into a 50-50 distribution.)

But you can say something else: you can say that on *average* the distribution of coin flips will tend towards the frequency predicted by the (correct) probability assignment. This can be made a little more precise as follows: on the relative frequency interpretation, to say that the coin has a 1/2 probability of coming up heads is to say that if you flip the coin an infinite number of times, then the probability of *actually* getting heads half the time will get closer and closer to 1. This is called the (weak) law of large numbers. It can be illustrated with a picture that I will show in class. The important thing is that there's no guarantee that after any particular number of flips, you will have anything like a 50-50 split.

The law of large numbers can also be a way of understanding how chance and relative frequency relate, in cases like the flip of a coin, where it can be difficult to say which is the correct interpretation. One thing you might say is that a coin has (chance) probability 1/2 of landing heads if the (chance) probability that the actual frequency of heads will be 1/2 gets closer and closer to 1 the more times you flip the coin.

LPS 31: Introduction to Inductive Logic
Lecture 11

Topics

Probability as degree of belief
Identifying degree of belief: first method
Betting rates and payoffs
Identifying degree of belief: second method
Conditional belief

1 Probability as degree of belief

Last week, we introduced three interpretations of probability. Two of them were based on the idea that probabilities said something about how the world is. These were the relative frequency and objective chance interpretations. The third one was based on the idea that probabilities say something about us. This is the degree of belief interpretation. We spent some time then talking about how to make sense of these interpretations, and how to recognize statements that seem to fit naturally with one or another of these interpretations. The focus for the remainder of the course is going to be on the degree of belief interpretation, and particularly on how to think about the relationship between belief, evidence, and learning in terms of probability theory. The philosophical term for what we are going to spend the next two weeks talking about is "Bayesian epistemology." "Epistemology" is the philosopher's word for "theory of knowledge," from the Greek "episteme" for "knowledge" and "logos" for "study of". And "Bayesian" because, as we will see next week, Bayes' theorem plays an essential role in how evidence affects belief.

The next two weeks can be thought of as a sustained argument in favor of a particular way of thinking about belief and evidence. We have already seen that there are plenty of examples of case where we say things about probability that are most naturally interpreted as statements about our beliefs (say, that there is a 90% chance that sauropods spent most of their time in the water). But it is a little more difficult to understand how to take our beliefs more generally and assign numbers to them. As a starting point for this discussion, we are going to spend today, and possible the beginning of class on Thursday, thinking about how we might go about doing just this. Then, we will talk about why these numbers should be expected to satisfy the rules of probability theory (which is not necessarily obvious). And then, we will make the argument that if our beliefs can be represented as numbers, and moreover that these numbers must satisfy the rules of probability theory, then Bayes' theorem provides a powerful tool for understanding how to revise our degrees of belief in light of new information about the world.

2 Identifying degree of belief: first method

Beliefs come in degrees. Sometimes we are absolutely certain that something is true. Sometimes we are certain it is false. But usually, we are somewhere in between: pretty sure, reasonably confident, not so certain. For instance, think about your answer to the following question. Do you believe that it is going to rain tomorrow? Probably not—it rains infrequently in Irvine. But how strongly do you believe that it won't rain tomorrow? Are you certain that it won't rain? Or just pretty sure? The answer to these questions, at least for me on any given day in Irvine, is that I am pretty sure, quite confident even, that it won't rain tomorrow. But I am not absolutely certain.

Suppose you were having a conversation like this with a friend. Suppose you said you were really certain it wouldn't rain tomorrow. What if your friend asked something like, would you bet a million dollars that it isn't going rain tomorrow? This kind of question feels like a figure of speech. But it suggests something useful. We already know, from our foray into basic decision theory before the midterm, that probability can be a useful tool for thinking about rational action and decision making. Now we are going to turn that on its head. The idea is to use thought experiments of the form "would you bet... that" to figure out how strongly we believe things.

As a starting point, we can think about very simple gambles. Here are two options I could offer you. (A) I give you ten dollars if it rains tomorrow, and nothing otherwise. (B) I give you ten dollars if it does not rain tomorrow, and nothing otherwise. In both cases, there's no cost to you—it's all upside, but only if you choose wisely. Which would you pick? There isn't a right answer here—we are trying to probe your personal degrees of belief. For me, I would pick option (B). But this is indicative of something: if we assume I am acting to maximize my expected value, then I must think it is more likely that it won't rain tomorrow than that it will rain.

This is progress. I now have a way of determining an order of my degrees of belief: if I would prefer to tie getting something I want to one belief than to another, then it follows that I must be more confident of the first thing than I am of the second thing. But really, we would like to figure out how to assign numbers to my beliefs, to represent as precisely as I'd like how certain I am that various things are true.

This suggests a modification of the last procedure. That gave me a way of determining relative degrees of belief. But what if instead I compared my certainty that it is going to rain tomorrow to my certainty that something whose probability I already know will happen? Consider the following choice. (A) I will give you ten dollars if this coin lands heads, and nothing otherwise. (B) I will give you ten dollars if it does not rain tomorrow, and nothing otherwise. I would pick (B) again. But this is more telling: now, I know that the probability I would assign to the likelihood that it won't rain tomorrow is higher than 1/2. (Actually, I haven't really learned anything that I didn't already know from the last choice. Why not?)

This method can be expanded as follows. Imagine a series of such choices, where in each case (A) was based on a more likely event. So, (A) I will give you ten dollars if this unbiased die landed on a number other than 1, and nothing otherwise. Or (A) I will give you ten dollars if this roulette wheel does not land on 00 and nothing otherwise. If you would still choose (B) in these cases, it would suggest that your degree of belief that it will rain is quite high indeed—in the first case, it would mean that you think the chance that it will rain is higher than 5/6, and in the second case it would mean that your degree of belief that it will rain is higher than 37/38.

You can also do it by taking the other possibility, i.e., by making comparisons between the proposition that it will rain tomorrow and increasingly *unlikely* events. So, I might ask you to choose between: (A) I will give you ten dollars if this coin comes up heads twice in a row (or three times, or four times...); and (B) I will give you ten dollars if it rains tomorrow. If you pick (A), then you must think that the likelihood that it will rain is less than 1/4 (or 1/8, or 1/16). If you find a point where you stop, that is, if you would pick flipping a coin three times, but would pick rain instead of picking flipping a coin 4 times, that gives you a range of where your belief must lie. (Specifically, you would say that you think the odds of it raining tomorrow are somewhere between 1/8 and 1/16.)

This new method gives a way of coming up with ranges of numbers to assign to beliefs. But we'd like to do better, still. We'd like to figure out a way to assign numbers more directly.

3 Betting rates and payoffs

The idea that using questions (or, to use Hacking's expression, thought experiments) concerning potential gambles to identify your degrees of belief is a good one. In fact, we can take it a step further. But first, we need to develop a little bit of the language of gambling.

Suppose two people, Jack and Jill, are making a bet. Jack says he is willing to bet $\$X$ that something will happen (say, the Yankees will win the world series). Jill, meanwhile, says that she will bet $\$Y$ that the thing won't happen. What does this mean? Well, if the thing happens, if the Yankees win the world series, then Jill has to pay $\$Y$ to Jack. If the thing *doesn't* happen, then Jack needs to pay $\$X$ to Jill. (Note that $\$X$ and $\$Y$ need not be equal to one another.) In a moment we will worry about what it means to say that Jack and Jill accept this bet. First, though, I just want to develop a way of talking about what's going on. First off, we will call $\$X$ Jack's *bet*. $\$Y$, then, is Jill's bet. A person's bet is how much money they stand to lose on a given gamble.

Next we can define the *stake*. We will usually call the stake S. This is the total amount of money that is being considered. It is the sum of the amount that Jack has bet and the amount that Jill has bet, so $(\$X + \$Y)$. The final thing to define is the *betting rate*. We will usually call the betting rate p. The better rate says what portion of the stake you are responsible for. For Jack, the betting rate is his bet divided by the state, so

$$p_{\text{Jack}} = \frac{X}{X+Y}.$$

Jill's betting rate is

$$p_{\text{Jill}} = \frac{Y}{X+Y} = (1 - p_{\text{Jack}}).$$

A person's betting rate connects up his or her bet with the stakes of the gamble. So, $X = p_{\text{Jack}}S$ and $Y = p_{\text{Jill}}S$. Note that it is always the case that if you add up the betting rates of all of the people involved in a gamble, you will always get 1.

A good way to represent a bet of this sort, which will come in handy later on, is as a payoff matrix. A payoff matrix looks something like this.

	Jack	Jill
Yankees win	$\$Y$	$-\$Y$
Yankees don't win	$-\$X$	$\$X$

In general, a payoff matrix for a general proposition A, with betting rate p for the person betting on A and stake S, can always be written,

	Bet that A	Bet that $\sim A$
A	$(1-p)S$	$-(1-p)S$
$\sim A$	$-pS$	pS

4 Identifying degree of belief: second method

We have already discussed the notion of a fair bet, and conversely, of advantageous and disadvantageous bets. One way to determine a person's degree of belief is to try to determine whether they would think that various bets are advantageous or disadvantageous. This is essentially what we were doing before, although the bets we were considering were extremely simple. But now we can make the idea more precise. Here are two ways to think about it.

First, suppose I offered you the following bet. I told you I would bet you \$10 that it won't rain tomorrow. How much would you have to bet in return in order for you to think the bet was fair (or even more, advantageous)? Recall that this is the same as asking what the cost of losing the bet would have to be in order for you to consider the expected value of the bet to be 0. If we knew the probabilities involved, then we would be able to plug the numbers directly into the expected value formula. But now we are trying to go backwards. We are trying to figure out your degrees of belief, i.e., trying to figure out what probabilities you assign. So what we want is come up with a number that you would intuitively think of as fair for a bet like this, and then use the expected value formula to work out what probabilities you must assign. Suppose I said I would bet you \$10 that it won't rain tomorrow. You think about it about you decide that, if I am betting \$10 that it won't rain, then you would be willing to bet \$1 it *will* rain tomorrow. If you think this is a fair bet, then we can work backwards to find the probabilities you assign to the proposition that it is going to rain tomorrow, using

$$0 = \text{EV}(R) = 10\Pr(R) - 1\Pr(\sim R) = 10\Pr(R) - 1(1-\Pr(R)) = 11\Pr(R) - 1.$$

Solving for $\Pr(R)$, we conclude that you must think the probability that it will rain tomorrow is $1/11$.

Here is another way to do it. Suppose that you are designing a bet, by stipulating the betting rate. (Imagine that the stakes are fixed, let's say at \$100.) The negotiations about the bet proceed in two steps. First, you say what you think the betting rate should be. Then your friend decides which side of the bet to take. You have to take the other side, no matter what the friend picks. This means that you want to pick the betting rate so that you think the bet is fair—otherwise, you risk giving your friend and advantage over you. How would you set the betting rate?

To be concrete, think about the question in terms of a payoff matrix, with a \$100 stake, that the Yankees are going to win the world series.

	Bet that Yankees win	Bet that Yankees don't win
The Yankees win	$(1-p)100$	$-(1-p)100$
The Yankees don't win	$-p100$	$p100$

You get to pick p, and then your friend picks whether to bet that the Yankees will win or that the Yankees will not win. Your goal is to pick p so that you simply do not care which side your friend takes. If you do this right, then p is a good reflection of your degree of belief that the Yankees are going to win the world series. This is a general method for associating numbers with your beliefs—setting up the betting rates at which you wouldn't care which side of the bet you were on.

5 Conditional belief

So far, we have only thought about cases in which we are betting on propositions directly. But often our beliefs depend on one another. For instance, if I wake up tomorrow and it is cloudy, I will be less confident that it is not going to rain than if I wake up in the morning it is is sunny. We have a notion of conditional probability already, but we now want to apply this to beliefs. We want to figure out how to assign numbers to our conditional beliefs, i.e., our beliefs that something is (or will be) true, assuming that some other things that we aren't perfectly certain of are also true. So, for instance, your degree of belief that the Yankees will win the world series likely would go up if you learned that they have made it to the playoffs.

It turns out that conditional beliefs can be given numbers in much the same way as normal beliefs, using the same sort of betting rules. The difference is that we need to consider a third possibility, where the thing that we are conditionalizing on turns out to be false. We will always assume that if this happens, then no money changes hands. So for instance, if I bet you that the Yankees will win the world series, assuming they make it to the playoffs, and the Yankees don't make it to the playoffs, then no money changes hands.

We can represent this sort of scenario using a payoff matrix again.

	Bet that Yankees win, given that they make i
The Yankees make it to the playoffs and win	$(1-p)100$
The Yankees make it to the playoffs and don't win	$-p100$
The Yankees don't make it to the playoffs	0

We can use the same procedure to come up with a number p here: just assume you are trying to assign a number so that you don't care which side of the bet you end up taking. This number will once again reflect your degree of belief—except now, instead of representing your degree of belief that the Yankees will win, it represents your degree of belief that the Yankees will win, given that they make it to the playoffs. In probabilistic terms, the number is $\Pr(Y|P)$, as opposed to $\Pr(Y)$.

LPS 31: Introduction to Inductive Logic
Lecture 12

Topics
Dutch Books
Coherence

1 Dutch books

Last class, we introduced a way of associating a number with your degree of belief in a proposition. It was by way of the betting rate that you would assign to a proposition if you didn't know which side of the bet you would be on. This argument had two parts. First, we argued that if you didn't know what side of a bet you would be on, you would try to set the betting rate in such a way that the bet would be fair—in other words, you would want to set the betting rate in such a way that neither side would have negative expected value. The second half of the argument was to then remark that a bet would seem fair to you if you thought that the probability that thing you were betting on was equal to the betting rate you assigned it. The meant that we could identify betting rates with degree of belief probabilities. If you think a bet is fair, then you can read your degree of belief off of the betting rate for the bet.

I said then that the next thing we would do would be to argue that the numbers we associate with degrees of belief ought to satisfy the rules of probability. The argument for this is going to be indirect, based on betting rates. In particular, we will show how you can run into problems if you set your betting rates in a way that doesn't satisfy the rules of probability. Since we already know that there is a close connection between degrees of belief and betting rates, it follows that our degrees of belief should satisfy the rules of probability. Now, the sense of "should" here is subtle. The claim is going to be that it is *irrational* to have degrees of belief that do not satisfy the rules of probability. Of course, people are often irrational, and so it should come as no surprise that in many cases, our beliefs do not satisfy the rules of probability. But as we shall see, if your beliefs do not satisfy the rules of probability, then you can get into trouble while gambling.

The argument involves series of bets called Dutch books. A Dutch book is a set of gambles such that you are guaranteed to lose, no matter what happens. (Sometimes they are called sure-loss contracts, out of sensitivity to the Dutch.) The claim is going to be that if your degrees of belief do not satisfy the rules of probability, then you will be open to Dutch books. We should be careful here to say just what it means to say that you are "open" to such sets of gambles. The idea is that if your beliefs are incoherent, then you will willing accept a series of bets, thinking that each of them is fair (or even advantageous), even though you are guaranteed to lose money on the whole set.

How does this work? Let's consider a very simple example. Suppose that you set your betting rates as follows. You set your rate on the proposition "The critics are going to like the last Harry Potter movie" at 3/4. And then you set your betting rate on the proposition "The critics are not going to like the last Harry Potter movie" at 1/2. First, note that these do not satisfy the rules of probability. The first one would correspond to assigning probability 3/4 to a proposition, L. The second one corresponds to assigning probability 1/2 to the negation of that proposition, $\sim L$. But the rules of probability say that $\Pr(L) + \Pr(\sim L) = 1$, and your subjective degrees of belief do not satisfy this rule.

How can this cause you problems? The idea is that when you set your betting rates so that $p(L) = 3/4$ and $p(\sim L) = 1/2$, you are claiming that you are happy to take either side of a bet at those rates. Now suppose a cunning gambler comes along and offers you the following bets. First he say he wants to accept your bet on L, for \$20 stakes. This means that if the critics do not like the last Harry Potter movie, you will owe the cunning gambler $p(L)S = \$15$, and if the critics do like the last movie, he will owe you \$5. But then he says that he also wants to accept your bet on $\sim L$, again for \$20 stakes. This means that if the critics do not like the last Harry Potter movie, he will pay you \$10. And if they do like it, you will pay him \$10. (Note that in this second case, you are betting that $\sim L$ will be true, which means you will owe $p(\sim L)S$ if $\sim L$ is false.) The important point is that, given your betting rates (and hence your beliefs), you will accept both of these bets. Indeed, you will think both are fair.

But what happens? There are two possibilities. Either the critics like the movie or they don't. If they like the movie, then you win your first bet, which gets you \$5. But you lose your second bet, which means you have to pay \$10. On a whole, you lose \$5, because you gain \$5 and lose \$10. Fair enough—you win some, you lose some. But now think about what happens if the critics *don't* like the movie. In this case you lose your first bet, which means that you have to pay \$15. On the other hand, you win your second bet—but you only get \$10, and so once again, you find yourself losing \$5 on the whole. This means that no matter what happens, you lose \$5.

This Dutch book can be expressed in terms of the following payoff matrices. For your first bet, with betting rate 3/4 for L,

	Bet for L	Bet against L
L	$(1-\frac{3}{4})20 = 5$	$-(1-\frac{3}{4})20 = -5$
sim L	$-(\frac{3}{4})20 = -15$	$(\frac{3}{4})20 = 15$

For your second bet, with betting rate 1/2 for $\sim L$,

	Bet for $\sim L$	Bet against $\sim L$
L	$(1-\frac{1}{2})20 = 10$	$-(1-\frac{1}{2})10 = -10$
sim L	$-(\frac{1}{2})20 = -10$	$(\frac{1}{2})20 = 10$

Combining these two bets (with bet 1 being your bet for L, and bet 2 being your bet for $\sim L$),

	Bet for L	Bet for $\sim L$	Payoff
L	5	-10	-5
$\sim L$	-15	10	-5

What has gone wrong here? Evidently, something was amiss in the way you have set your betting rates (and by extension your probabilities), because you have accepted as fair a series of bets that have guaranteed that you will lose money. Clearly something has gone wrong here.

2 Coherence

Let's define a new term. Let's say your beliefs are *coherent* if it is not possible to construct a Dutch book that you would be willing to accept. (We will use "coherent" to describe a set of beliefs or a set of betting rates you believe to be fair, interchangeably.) Incoherent beliefs will lead you to make choices that are always against your best interests, and so it seems that you want to avoid incoherent beliefs whenever possible. How can you do this? **Claim:** A set of betting rates (and by extension beliefs) is coherent if and only if they satisfy the rules of probability.

What does this mean? It means, for instance, that every time you have already set betting rates for propositions A and B, for instance, that you have to set your betting rate for $A \vee B$ to $p(A) + p(B)$. Likewise, if A and B are mutually exclusive and exhaustive, then you require that your betting rates add to 1, so $p(A) + p(B) = 1$. We have just seen that if your betting rates for L and $\sim L$ do not add to 1, that a Dutch book can be constructed against you. Conversely, if they *do* add to 1, then no Dutch book can be constructed. To see this point, modify the last example as follows. Instead of taking $p(\sim L) = 1/2$, take $p(\sim L) = 1 - p(L) = 1/4$. Then, no matter what bets get made, there is a situation in which you will win. Consider,

	Bet for L	Bet for $\sim L$	Payoff
L	$(1 - p(L))S_1$	$-p(\sim L)S_2 = -(1 - p(L))S_2$	$p(\sim L)(S_1 - S_2)$
$\sim L$	$-p(L)S_1$	$(1 - p(\sim L))S_2 = p(L)S_2$	$p(L)(S_2 - S_1)$

If $S_1 > S_2$, that is, if the stake of your bet on L is larger than the stake of your bet on $\sim L$, then you will make money if L is true; if the stake of your bet on $\sim L$ is larger than the stake of your bet on L, then you will make money if L is false. But either way, there is some set of circumstances were you make money. This is not a Dutch book, no matter what values the stakes of your bets are.

There's an important point to emphasize, here. The bets we have just described may not be fair—you could lose far more often than you win. But they are not a Dutch book. A Dutch book is a set of bets that you are *guaranteed* to lose. If your betting rates are coherent but do not line up with the real probabilities (in cases where it makes sense to talk about real probabilities), you might take a bet to be fair that really was terribly disadvantageous. You might lose far more often than you win. But there are nonetheless situations in which you do win. If your bets are incoherent, then there are sets of bets that you will lose no matter what. These are Dutch books.

I am not going to prove the claim above in any kind of formal way. But I will show some examples of how, if you break various basic rules of probability with your betting rates (beliefs), Dutch books can be constructed. I won't go through the argument in the other direction, but the example above should give you a sense of how it works. What is going to be important for our purposes is that you understand how the argument is supposed to go: that the reason why betting rates have to satisfy the rules of probability is that if they *don't*, then you will accept bets that are guaranteed to lose you money.

We have already seen one example of how a broken rule (betting rates for mutually exclusive, exhaustive propositions do not add to 1) can lead to a Dutch book. Let's now consider another rule—a basic one, that a tautology (or otherwise certain proposition) should be assigned probability 1. What happens if you don't assign a betting rate of 1 to a tautology (and likewise, 0 to a contradiction)? Suppose you assign a betting rate to $L \vee \sim L$ of .9. Remember that we are considering bets that you consider fair, that is, where you are happy to be on either side of the bet. So in this case, the cunning gambler simply has to take the positive side of the bet. That is, the cunning gambler tells you that he wants to bet on $L \vee \sim L$, and have you take the position against $L \vee \sim L$ at your rate. Then, if the stakes are \$10, you will always lose $\$(1 - .9)10 = \$1$, because the thing you are betting against is always true. No matter what, you lose money.

Let's take another rule: additivity. Consider the following set of incoherent betting rates. We are rolling an unbiased die. I say that the betting rate for a 1 will be 1/6, and the betting rate for a 2 will be 1/6. These are mutually exclusive, and so in order to make my betting rates coherent, I would have to set the betting rate for $D_1 \vee D_2$ to 1/3. Suppose that instead, I set it to 1/2.

Now the cunning gambler comes along and offers me the following three bets, all with \$60 stakes. We says he wants to bet on D_1 and D_2, so you have to take the position against D_1 and D_2. This means that if any number other than a 1 gets rolled, he has to pay you $(1/6)60 = \$10$. But if a 1 does get rolled, you pay him $(5/6)60 = \$50$. Likewise, if anything other than a 2 gets rolled, he pays you $(1/6)60 = \$10$, and if a 2 gets rolled, you pay him $(5/6)60 = \$50$. Then he says, he wants to bet *against* $D_1 \vee D_2$. This means you have to take the position in favor of $D_1 \vee D_2$. The betting rate on $D_1 \vee D_2$ was 1/2, so if anything other than a 1 or a 2 gets rolled, you have to pay $(1/2)(60) = \$30$. Meanwhile, if a 1 or a 2 is rolled, you get payed $(1/2)(60) = \$30$.

What happens? First, if a 1 gets rolled, you lose the first bet and win the second and third bets. This means you have to pay out \$50, but you make back $\$10 + \$30 = \$40$, for net loss of \$10. The same thing happens if a 2 gets rolled (now you lose the second bet, but win the first and third). Meanwhile, if anything else gets rolled, you win the first and second bets, but lose the last one. This means you win $\$10 + \$10$, but you have to pay \$30, and so you net a loss of \$10. In other words, you lose \$10 no matter what.

This can be summarized in the following payoff matrix.

	Bet against D_1	Bet against D_2	Bet for $D_1 \vee D_2$	Payoff
D_1	$-\$50$	\$10	\$30	$-\$10$
D_2	\$10	$-\$50$	\$30	$-\$10$
$\sim (D_1 \vee D_2)$	\$10	\$10	-\$30	$-\$10$

Similar arguments hold for the other basic rules of probability. If your betting rates violate any of them, then it is always possible to find a Dutch book that you will be willing to accept.

LPS 31: Introduction to Inductive Logic
Lecture 13

Topics
Conditional Coherence
Bayesianism
Hypotheses and Evidence
Prior and Posterior Probabilities

1 Conditional Coherence

So far, we have seen how if your degrees of belief do not satisfy the basic rules of probability, then it is possible to construct a system of bets such that you will accept each bet as fair, but where you are guaranteed to lose no matter what happens. This gives a standard by which it is irrational to have degrees of belief that do not satisfy the rules of probability. The last thing to check is whether our degrees of belief also need to satisfy the rules of conditional probability. The claim is that they do, and the argument has a now familiar form. We will show that if your degrees of belief do not satisfy the rules of conditional probability, then it is possible to construct a Dutch book against you.

What would it look like for betting rates to satisfy the rules of conditional probability? In particular, we would require that if you made a bet on A, conditional on B (where the betting rate on B is non-zero)—something like a bet that the Yankees will win the World Series, conditional on them getting to the playoffs—that the betting rate for the conditional bet would have to equal your betting rate for $A \wedge B$, divided by your betting rate for B.

Let's take an example. Suppose that you would accept as fair a bet that the Yankees are going to win the World Series at a betting rate of 1/4 (note that they can't win the World Series without making it to the playoffs first, so this is the same as a bet that they will both make it to the playoffs and win the World Series). Now, you also would accept a bet that the Yankees are going to make it to the playoffs at 1/2. But you think that if the Yankees *do* make the playoffs, then they are very likely to win the World Series. So likely, in fact, that you would accept a betting rate of 3/4.

Now consider the following gambles, all at the betting rates just given. Unlike the previous Dutch books, these gambles will involve different stakes. The first one is a gamble with stakes of \$100, that the Yankees are going to win the World Series. This has a payoff matrix of

	Bet for W	Bet against W
$W \wedge P$	\$75	−\$75
$\sim (W \wedge P)$	−\$25	\$25

The second one is a gamble that the Yankees will make it to the playoffs, with stakes of \$60. This has a payoff matrix of

	Bet for P	Bet against P
P	\$30	−\$30
$\sim P$	−\$30	\$30

The last one is a conditional bet, that the Yankees will win the World Series if they make it to the playoffs, with stakes of \$100. This one has a payoff matrix of,

	Bet for $W\|P$	Bet against $W\|P$
$W \wedge P$	\$25	−\$25
$\sim W \wedge P$	−\$75	\$75
$\sim P$	0	0

Given your betting rates, you would be happy to accept either side of any of these gambles. And so, if the cunning gambler comes along and asks you to bet *against* the Yankees winning the World Series, at \$100 stakes, and then bet *for* the Yankees making it to the playoffs with \$50 stakes, and for the Yankees winning the World Series, given that they make it to the playoffs, again at \$100 stakes, you would think this was a fair bet. But now consider the consequences.

	Bet against W	Bet for P	Bet for $W\|P$	Payoff
$W \wedge P$	−\$75	\$30	\$25	−\$20
$\sim W \wedge P$	\$25	\$30	−\$75	−\$20
$\sim P$	\$25	−\$30	0	$\$ - 5$

This means that if your betting rates for conditional bets don't satisfy the rule of conditional probability then you are once again susceptible to a Dutch book.

2 Bayesianism

This last argument, that in addition to satisfying the basic rules of probability, your degrees of belief (represented by your fair betting rates) need to satisfy the rules of conditional probability, may seem like a boring extension of the Dutch book argument we have already seen. But it actually has some significant consequences. If our *conditional* degrees of belief need to satisfy the rules of probability theory, then in particular it is easy to see that our degree of belief need to satisfy Bayes' theorem, since that is a simple consequence of the rules of conditional probability.

On one way of looking at things, that our degrees of belief must satisfy Bayes' theorem shouldn't be a surprising, or remarkable, discovery. But for a certain kind of problem, a kind of problem that philosophers and scientists have found vexing for hundreds of years, Bayes' theorem provides a powerful tool. The kind of problem is the one we have been working our way towards solving for this whole quarter. It concerns how to understand the relationship between beliefs about the world and evidence for those beliefs.

To make some contact with ideas discussed at the very beginning of the quarter, think of the problem in this way. When we construct a *sound* argument for some conclusion, then we know with absolute certainty that the conclusion is true. But what about invalid arguments (with true premises)? Sometimes we want to reject such argument altogether. But in at least some cases, the arguments can be understood to rely on some variety of what I have called inductive inference. In these cases, we have an intuitive sense that the premises of the argument should be understood to provide a kind of evidence for the conclusion. But what should this kind of inductive evidence tell us? We know that inductive evidence cannot guarantee that a conclusion is true. If I tell you that every swan I have ever seen is white, that doesn't make it certain that there are no black swans. But it does *something*. In particular, it should make you more confident that there are no black swans.

We now have a way of making this idea of "more confident" precise. The claim is that evidence in favor of a proposition should *increase* your degree of belief that the proposition is true. Evidence should make the betting rate you would accept for a bet in favor of the proposition higher. How should it do that? Well, we now know that our degrees of belief have to satisfy the rules of probability, including conditional probability, or else we will be susceptible to Dutch books. This gives us a guide to how to change our beliefs in light of evidence, in order to ensure that our beliefs remain coherent. The basic idea underlying Bayesianism is that our degree of belief that a proposition is true in light of new evidence should increase (or decrease) to reflect the evidence, via Bayes theorem. Let's see how this works in detail.

3 Hypotheses and Evidence

As a starting point for thinking about how to update our beliefs in light of evidence, let's first distinguish between two types of propositions. A *hypothesis* is a belief about the world. It is something about which you are uncertain, the sort of thing that you might be looking to confirm (or disconfirm). In general, virtually any proposition can be a hypothesis. For instance, you might have a hypothesis that all swans are white. Or you might have a hypothesis that a particular patient has strep throat. Or you could have a hypothesis that a coin is fair.

The second kind of proposition that we are going to be interested in is a piece of *evidence*. Evidence is something that you learn about the world (usually with non-Bayesian methods) that might affect your degree of belief. What counts as evidence? As far as the Bayeian is concerned, any proposition that is not independent of your hypotheses can count as evidence for or against your hypotheses. For instance, if you have a hypothesis that a patient has strep throat, then a positive result on a strep test counts as evidence for your hypothesis. A negative result counts as evidence against your hypothesis. If you think that a coin is fair, then a run of 100 heads could count as evidence that it isn't fair after all.

To reason about a give hypothesis, you need to make sure you understand the alternatives. If not all swans are white, then some swans must be a different color. An alternative hypothesis might be that a small number of swans are black. Perhaps your patient doesn't have strep throat, but instead has a cold. This is another alternative hypothesis. When trying to understand how evidence should affect your confidence in a hypothesis, you should start by coming up with a list of alternative hypotheses. The goal is to come up with a list of mutually exclusive, and exhaustive hypotheses. There is a big problem here, which is that in many cases you can't say in advance just what the alternative hypotheses are! For this reason, we usually include a final hypothesis of "None of these!", called the catch-all hypothesis. Sometimes you can't do any better than simply taking the catch-all as your alternative, and sometimes this isn't a bad thing (for instance, if your hypothesis is that a coin is fair, it is okay to take "the coin isn't fair" as your only alternative hypothesis).

We will usually use H to represent hypotheses. So if we have a list of exhaustive, mutually exclusive hypotheses, we would write them as $H_1, \ldots, H_n$. Evidence, meanwhile, we will represent as either E or, in some cases, by a letter representing the particular bit of evidence.

4 Prior and Posterior Probabilities

The Bayesian picture looks something like as follows. It involves having degrees of belief that evolve over time. One begins with a bunch of beliefs about the world. In fact, we assume that we have beliefs about every proposition there is. These are expressed as probabilities that various propositions, including a number of hypotheses, are true. Our initial degrees of belief are called *prior probabilities*, or sometimes just *priors*. At first, at least in general, we don't have any principled way to assign numbers to our priors (some versions of Bayesianism maintain that there are objective ways to assign these probabilities — for our purposes, we will assume that we just assign them based on our guts, though with some education about how, for instance, chance setups usually work if they are unbiased, etc.). We just come up with numbers, using hypothetical gambling situations to probe our degrees of belief. The most important thing is that our beliefs be coherent.

The only other constraint on our prior probabilities is that we should never be *dogmatic*, that is, we should only assign probability 1 or 0 to tautologies or logical contradictions, no matter how certain we feel about something. You might be absolutely sure that something is true. It so, then you might assign a prior probability of .9999999 to it, or even higher. But never assign a probability of 1, because that means that no matter what happens, you will not be able to change your belief.

Now, since we assume that we have prior probabilities for *all* of our beliefs, in particular we have prior probabilities for various hypotheses about the world. For instance, a doctor might have prior probabilities about whether or not a particular patient has strep throat, whether she has a cold, whether she has tonsilitis, etc. But we also have prior conditional probabilities. For instance, we have prior probabilities concerning how likely various consequences of our hypotheses would be, supposing the hypotheses are true. For instance, we have an idea of how likely it is for a patient to get a positive result on a strep test if she has strep, how likely it is for a patient to get a positive result if she has a cold, how likely it is for a patient to get a positive result if she has tonsilitis, etc.

We have already done a number of problems of the following sort. We imagine a doctor has a patient come in, and the doctor gives the patient a strep test, and we try to figure out what the probability that the patient has strep is, given that the patient gets a positive result on the strep test. This kind of problem is a classic Bayes' theorem problem. Recast in our current terms, though, its significance becomes clearer. Here's how we write Bayes' theorem in this case:

$$\Pr(D|+) = \frac{\Pr(+|D)\Pr(D)}{\Pr(+|D)\Pr(D) + \Pr(+|\sim D)\Pr(\sim D)}$$

As we've been discussing the terms, the hypothesis here is that the patient has strep. Let's call this H. We can take the alternatives as a catch-all hypothesis, $\sim H$. The evidence in this case is the positive result on the test. To begin with, we have prior probabilities for H, and we have prior probabilities that the patient will get a positive result on the test, given that she has the disease, and given that she doesn't. All of these go on the right hand side of Bayes' theorem.

The important thing is how to interpret the left hand side of the Bayes' theorem calculation. This tells us what probability to assign to our hypothesis, given the evidence. In other words, this tell us how to *update* our degree of belief in the hypothesis, that the patient has strep throat, *in light of the evidence.* This is called the *posterior probability* that the hypothesis is true. It is the probability of that hypothesis is true, given the evidence we have just gained.

The Bayesian approach to understanding beliefs and evidence is a kind of iterative process. We start with some prior probabilities. Then we get evidence. We use Bayes' theorem to calculate posterior probabilities. *And then we accept the posterior probabilities as our new degrees of belief.* In other words, in answer to the question "how should evidence affect our beliefs?" the Bayesian will answer: in light of evidence, your new beliefs should be represented by your posterior probabilities. In other words, at the end of the process, your posterior probabilities, as calculated using Bayes' theorem, should now be your new prior probabilities, which you plug into Bayes' theorem the next time a piece of evidence comes along. This means that each time you gain evidence about the world, you change your beliefs in such a way as to make sure they remain coherent, but applying Bayes' theorem. Your beliefs constantly change in light of new evidence, in a precise and systematic way.

In class, we will do a number of problems that show how this works in detail.