The Basics of Probability

What is Probability?

Probability is the mathematical representation of the likelihood of an event under a given set of circumstances (conditions) in a given period of time. We will say, for example, that the probability of winning the jackpot in the lottery from buying one ticket this week is some (very tiny) number like .00000003.

We also sometimes use probability language to refer to states of affairs—we might say, for example, that given what I know about the number of buses in town and the color of the bus that hit the plaintiff, the probability that the bus that hit the plaintiff belonged to the Acme Bus Company is .75. (This comes from a very famous thought experiment, described in Charles Nesson, The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts, 98 Harv. L. Rev. 1357 (1985)) Usually, you won't need to think too hard about the difference between these two ideas.

We represent probability as a fraction between 0 and 1, where 0 means no chance that an event will happen, and 1 means that an event will definitely happen. If someone starts talking about a probability higher than 1 or lower than 0, they're either clueless or lying.

Normally, we express probability in decimals like .75 or percentages like 75%. We can also express them in fractions like 3/4, but it's a little easy to confuse those with *odds*, which is a different representation of probability, so we'll (mostly) be avoiding fractional probability.

There are two big theoretical interpretations of probability, the frequentist and the Bayesian.

Very broadly speaking, the frequentist interpretation of probability says that we are, as the name suggests, talking about frequencies. To say that you have a .75 chance of winning the lottery, to the frequentist, says that if if we picked a large number of groups of 100 people just like you, then we would expect that on average 75 people in each of those groups would win the lottery.

By contrast, the Bayesian interpretation is about subjective beliefs. On the Bayesian interpretation, when I say that I have a subjective probability of .75 in winning the lottery, what I mean is that if I make a large number of predictions with a .75 probability, I expect them to be true about 75% of the time. That is, a Bayesian probability represents an attempt to quantify my level of uncertainty.

For the most part, you don't need to care about these different interpretations. Statisticians and philosophers care about them a lot, but the practical implications of the two worldviews won't come up in much that lawyers need to deal with. Importantly, in a bit I'll introduce an idea called *Bayes Rule*, but this is *not* a concept that is limited to Bayesians, frequentists recognize it too—it's a

basic mathematical fact.

Probability math

A single event

For the purposes of having an intuition for this kind of thing, you can think of the probability of a single event as just the proportional frequency of an event in the world of possible events under similar circumstances. If 5 of every 100 people who file suits for negligence win their cases at trial, then the probability of winning a negligence case at trial, assuming no other information, is .05.

Usually we express the probability of some event with a capital P, like P(Win your negligence suit).

Then it's real easy to calculate the probability of a single event:

$$P(event) = \frac{ways - that - event - can - happen}{ways - anything - can - happen}$$

For example, suppose I have a jar of 30 marbles, and 3 of them are red. And suppose I draw exactly 1 marble. There are three ways I can pick a red marble, and 30 total ways I can pick any marble (if we interpret "grabbing a marble" as a "way"). Therefore, by the math above, the probability of getting a red marble is .1. It's really that simple.

Multiple Events

The real interesting kind of probability math is what happens when we consider multiple events. But before we dig into that math, we should first define some terms and simple rules.

Basic ideas and rules

First, two (or more) events are **independent** if they don't affect one another, or, put differently, if knowing about one of them doesn't affect our judgment about how probable the other is. For example, who I call on in Torts class and the winner of the Kentucky Derby are independent: horse races don't affect who I call on, and vice versa. (If you're a bit of a jerk, you might imagine degenerate cases where this might not be true. For example, suppose the owner of the winning horse is in my class, but quits law school after the race makes him/her rich? Or suppose my questions are so unkind that I demoralize one of the students, who quits law school to become a horse trainer? Fine, *in the ordinary course of events* those two are independent.)

Two (or more) events are **mutually exclusive** if they cannot occur at the same time. For example, my becoming Pope and my being a Protestant are mutually exclusive.

A set of events is **exhaustive** if they describe all possible events under the circumstances. For example, here is an exhaustive set of the relationships I might have with the University of Iowa College of Law:

- I'm a faculty member.
- I'm a staff member.
- I'm an administrator.
- I'm a student.
- I'm not affiliated with the law school.

If a set of events is mutually exclusive and exhaustive, then their probabilities always sum to 1. Think about our example: it's definitely true that I'm either faculty, staff, administrator, student, or not affiliated.

An event and its **complement** are always mutually exclusive and exhaustive. "Complement" is just a fancy set theory way of saying "add the word 'not' to the front of it." For example, the following two events are complements:

- I'm king of the world.
- I'm not king of the world.

It follows that P(Gowder is king of the world) + P(Gowder is not king of the world) = 1, because, being complements, they're mutually exclusive and exhaustive.

By some basic algebra, we also get **the subtraction rule**: the probability of any event is just the probability of its complement, subtracted from 1. For example, if I have a .8 chance of being king of the world, then it follows that I have a 1 - .8 = .2 chance of not being king of the world.

Alternative probabilities

For mutually exclusive events, we have **the addition rule**: if A and B are mutually exclusive events, then the probability of one or the other being true is just equal to the sum of their individual probabilities.

Here's a concrete example: what's the probability of me being either Pope or a Protestant? Well, let's suppose the probability of my being Pope is .00001, and the probability of my being Protestant is .05. Then the probability of my being either Pope or Protestant is .05001, the sum P(Pope) + P(Protestant).

In set-theory language, this "or" relationship is called the **union** of the two events; in other words, the probability that either A or B occurs is called the probability of the union of A and B is denoted:

$$P(A \cup B)$$

The addition rule can be extended to the case where A and B are not mutually exclusive. Under those circumstances, the probability that either A or B occurs is just the sum of the probability of A plus the probability of B, *minus* the probability of both A and B happening.

The probability of both A and B happening also has a set theory term and notation. It's called the **intersection** of A and B, and is denoted:

$$P(A \cap B)$$

Our expanded addition rule can then be expressed in mathematical notation as following:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Note that we can derive the shorter version of the addition rule that we started with from the longer version I just gave you. If two events are mutually exclusive, then the probability of their union is 0, so the third term in that longer equation just drops away, leaving us with the shorter version.

Joint Probabilities

But how do we calculate the intersection of probabilities? The intersection is also known as the joint probability of events A and B. To remind you, that just means the probability that both are true at the same time.

If A and B are independent, the intersection is just the product of the probabilities of each. This is the **multiplication rule**:

$$P(A \cap B) = P(A)P(B)$$

Here's an example. Suppose I flip a coin twice. Each coin flip is an independent event. So what's the probability of my getting heads twice? Let's reason through it in words. Half of the time, I get heads on the first flip. Then, *half of that first half* I get heads on the second flip. Put differently, I flip two coins every day for 100 days, I expect that on roughly 50 of those days I'll get heads the first time. And then on roughly half of those 50 days, or 25 days, I'll get a second head too.

An immediate application

The notion of joint probability reveals a potential puzzle in the law. Suppose a trial finishes in a case that hinges on two independent factual issues. For example, suppose we have a toxic tort case. Defendant denies that the chemical produced by its plant was in the water supply, but also denies that the chemical could, as a matter of biology, have caused the ill effects plaintiff suffered. Plaintiff needs to prove both of those things in order to win, but the probability of one being true is not affected by the probability of the other being true (whether the chemical is poisonous doesn't affect whether it was in the water supply, and vice versa).

Now suppose a juror, based on the evidence presented by plaintiff, assigns a subjective probability of .7 to the chemical being in the water supply, and also assigns a subjective probability of .7 to the chemical's being poisonous. Should she vote for a verdict for the plaintiff?

Well, it's not clear. If the rule is "plaintiff must prove every individual fact at issue by a preponderance of the evidence," then it looks like plaintiff wins: it's more likely than not that the chemical was in the water supply, and it's more likely than not that it's poisonous. On the other hand, if the rule is "plaintiff must prove all the facts at issue by a preponderance of the evidence," then it sounds like we're concerned with the *joint* probability of the two claims, i.e. .7 * .7 = .49. Whoops, that's less than half! Plaintiff loses.

This has actually been the object of some discussion in academic journals. *See, e.g.*, Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 YALE L. J. 1254 (2013) (calling this the "conjunction paradox" and proposing one possible solution). *See also* Ronald J. Allen & Alex Stein, *Evidence, Probability and the Burden of Proof*, 55 ARIZ. L. REV. 557 (2013) (excellent discussion of related problems and literature on same; claiming that practical adjudication rarely involves making probabilistic calculations).

This provides all kind of opportunities for the skillful lawyer who isn't afraid of math. Suppose you're defense counsel, and plaintiff's experts have testified that the probability of those factual issues is each .7. The moment plaintiff rests, if you understand probability, you should instantly move for a directed verdict and argue for the judge to apply the all-the-facts rule. You might not win, especially if those scholars who argue that the preponderance of the evidence standard isn't probabilistic are right, but it's a sufficiently compelling argument to be well worth trying. This is also something that you could argue to the jury, or try to get in a jury instruction.

But what if events aren't independent?

Conditional Probabilities

One of the most important probability topics to understand, and one that

people frequently get wrong—sometimes to dire consequences—is conditional probability. Conditional probability is what we need when we realize that some events are not independent.

The conditional probability of some event A, given some other event B, is just the probability of A *taking into account the fact that non-independent event B happened*. It expresses the effect that one event happens on the probability of another.

For example, the probability that I'll get a concussion today is pretty low, but the probability that I'll get a concussion *given* that I play for the Hawkeyes and there's a football game today is much higher. If we're trying to make judgments about my likelihood of getting a head injury, it would be pretty stupid to fail to take into account that I'm involved in an activity that primarily involves gigantic people bashing my head.

The mathematical notation for conditional probability is:

$$P(B|A)$$

The vertical line signifies the conditional nature of the relationship, i.e., the probability of B conditional on A (imagine B as the concussion, and A as my playing in a football game).

There's a simple formula for calculating conditional probabilities. I'll give you the math first, and then we'll talk through it:

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

This formula is known as **Bayes Rule** (sometimes Bayes Theorem or Bayes law, and sometimes we stick an apostrophe in there). It's super-important.

In words: the probability of event B conditional on event A is the product of the probability of event A conditional on event B and the probability of event B, all divided by the probability of event A.

Here's a concrete example. Suppose you want to know the probability that I'm a football player, given that I have a concussion. Here's how you'd calculate it:

$$P(Football|Concussion) = \frac{P(Concussion|Football)P(Football)}{P(Concussion)}$$

So let's get some numbers up in this exploitative college sport/financial scam. Assume that:

The probability of my having a concussion, given that I'm a football player, is .5:

$$P(Concussion|Football) = .5$$

The probability of my being a football player without knowing anything else about me (the "base rate" of football players in the population, or, to Bayesians, the "prior" that you have in my being a football player) is, blessedly, low, say .01:

$$P(Football) = .01$$

The probability of my having a concussion without knowing anything else about me (the base rate of concussions) is also pretty low, let's call that .03:

$$P(Concussion) = .03$$

So then we can just plug the numbers into our Bayes' Rule equation and come up with our answer:

$$P(Football|Concussion) = \frac{(0.5)(0.01)}{(0.03)} = 0.17$$

Conceptually, that equation represents the additional information you get about whether I'm likely to be a football player, as a result of learning that I have a concussion. If you're a Bayesian, we'd say that you now have .17 subjective probability in my being a football player, as opposed to merely .01 before (your "posterior probability" as opposed to your "prior probability"); if you're a frequentist you'd say that you expect about 17% of the people you meet who have concussions to be football players, as opposed to 1% of people in the general population.

Independence and Joint Probability Revisited.

Now that we have the notion of conditional probability, we can go back and state a formal definition of independence. Events A and B are independent if:

$$P(A) = P(A|B)$$

That is, they're independent if whether or not B happens doesn't affect A.

If you want to practice your algebra, you can prove that if the condition above is true, it's also true that:

$$P(B) = P(B|A)$$

7

Hint: try plugging stuff into Bayes Rule and then factoring stuff out.

Another (equivalent) definition of independence is that A and B are independent if:

$$P(A \cap B) = P(A)P(B)$$

That's just the multiplication rule flipped around, if you understand our coin flipping example you understand why that's true.

We can also conclude that the conditional probability of an event given its complement is 0. There are lots of ways to write this (complement notation doesn't seem to be particularly standardized), but here's one way:

$$P(A|\overline{A}) = 0$$

Why? Well, because we know that if not-X happened, then X definitely didn't happen. The conditional probability of my being king of the world, given that I'm not king of the world, has to be zero.

Now that we have the notion of conditional probability, we can also get a formula for calculating the joint probability of dependent events. Where A and B are dependent (that is, *not independent*):

$$P(A \cap B) = P(A)P(B|A)$$

This is just our earlier multiplication rule for joint probability again: we multiply the probability of A by the probability of A given B to get the joint probability of A and B.

Note how we can derive our original version of the multiplication rule (for independent events) from this expanded version. If A and B are independent, then:

$$P(B) = P(B|A)$$

hence we can substitute that into the above equation, and get our original multiplication rule:

$$P(A \cap B) = P(A)P(B)$$

The classic example to understand the reasoning behind the expanded multiplication rule is drawing cards. If you draw cards from a normal deck and don't put them back, then the probability of the first card being a heart is just 13/52 = 1/4. What's the probability of drawing a second heart?

Well, imagine that instead of drawing from the same deck, someone just gave you a *second* deck of cards with 51 cards in it, 12 of which were hearts. We know the probability of drawing a heart in that second deck, right? 12/51. That world is just the world described by the conditional probability of drawing a second heart given that you've already drawn a first. So what's the probability of drawing a heart out of the normal deck and a heart out of the weird second deck? Easy: 1/4 * 12/51.

We can also chain this rule, i.e.:

$$P(A \cap B \cap C) = P(A)P(B|A)P(C|(A \cap B))$$

Think about the probability of drawing a third heart in a row. And then a fourth, and so on.

Conditional independence

In the real world, full-fledged independence is pretty rare. But conditional independence—two events being independent conditional on a third variable—is much more common and just as useful. Here's an example of conditional independence: having multiple concussions and weighing more than 300 pounts might be independent conditional on being a football player. That is, given I know someone is a football player, I don't learn anything more about whether they're likely to have multiple concussions from knowing they're really big, or vice versa. Formally speaking,

$$P(concussions|football, gigantic) = P(concussions|football)$$

and

$$P(gigantic|football, concussions) = P(gigantic|football)$$

The trick is that two variables can be conditionally independent on some third variable without being independent. If I see someone really huge walking around campus, I do learn something about the likelihood of them having had multiple concussions—because really big people are more likely to be football players, and football players are more likely, in turn, to suffer multiple concussions.

The nice thing about conditional independence is the multiplication rule still works. That is:

$$P((gigantic \cap concussions)|football) = P(gigantic|football)P(concussions|football)$$

A Collection of Common Mistakes

There are a bunch of frequent errors that people make with probability. Don't make them.

Ignoring Base Rates

The first mistake is just not getting the whole Bayes thing, like, at all. Consider this classic example:

Suppose there's a medical test for some disease, and let's say that it has a 95% accuracy rate. To be more precise, let's say that 95% of the time it says what it should say, given the state of the world: if a patient has the disease, 95% of the time it returns a positive result; if the patient does not have the disease, 95% of the time it returns a negative result.

(Note on statistical lingo you might see floating around: a Type I error is a "false positive," i.e., if the machine sees someone without the disease and returns a positive result; a Type II error is a "false negative," i.e., if the machine sees someone with the disease and returns a negative result. For simplicity, we're assuming here that the test has identical Type I and Type II error rates.)

Imagine you're a doctor, and you administer that test to a patient. It reads positive. *Are you justified in concluding that there's a 95% chance they have the disease?* The answer is "no, not without first taking into account what proportion of people have the disease ('the base rate')."

Let's work through this problem. We can do it in two ways. Both ways require making an assumption about the base rate of the disease in the population. So let's suppose that 2% of the people in the population have the disease.

First we'll talk through it using our intuitions. Suppose we test 1000 people. 20 of them will have the disease. Of those, the machine will correctly identify 19, so it'll return 19 positive results and 1 negative result. 980 of our 1000 people will *not* have the disease, but our 95% accuracy machine will only identify 931 of those correctly: it'll return 931 negative results and 49 positive results.

So out of our 1000 people, we'll get a total of 68 positive results. Of those results, only 19 of the people will have the disease. That's only about 28% (after some rounding)! So you can only conclude that an individual person has a 28% chance of having the disease, not a 95% chance.

Ok, let's do this a bit more formally using Bayes' Rule. This will require some thinking. We want to know the probability of the patient having the disease, given that s/he tested positive. We know the probability of the patient testing positive, given that s/he has the disease (.95), and we know the base rate probability of the patient having the disease (.02), but we seem to be missing

a term in our equation: we don't know the base rate probability of getting a positive result on the test. Or do we?

Well, actually, we can calculate it simply by exhaustively describing the possible states of the world. After all, we know that a patient can only have or not-have the disease. Therefore, we know (by the addition rule) that the probability of the test being positive is just the sum of the probability of the two mutually exclusive states of affairs in which it can be positive:

$$P((Positive|Disease) \cap Disease) + P((Positive|Healthy) \cap Healthy)$$

And we know how to calculate these unions. They're all going to be independent events, so we just apply the multiplication rule.

To make this concrete, let's imagine a patient. 98% of the time, the patient does not have the disease, and 5% of that 98% time, that patient will get a positive result. 2% of the time, the patient will have the disease, and 95% of that 2% time, that patient will get a positive result. So then we can calculate the base rate of the machine returning positive as:

$$P(Positive) = (.98 * .05) + (.02 * .95) = .068$$

Now we can plug this all right into Bayes' Rule:

$$P(Disease|Positive) = \frac{P(Positive|Disease)P(Disease)}{P(Positive)} = \frac{(.95)(.02)}{(.068)} \ .28$$

Unsurprisingly, this is the same result we got before.

Now imagine this reasoning applied in common legal contexts. Do defense lawyers know the relevant error rates of breathalyzer machines, and the base rate of drunk driving in the population? Let's hope so!

The Conjunction Fallacy.

Here's another classic probability oopsie:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.
>
> a) Linda is a teacher in elementary school.
>
> b) Linda works in a bookstore and takes Yoga classes.

c) Linda is active in the feminist movement.

d) Linda is a psychiatric social worker.

e) Linda is a member of the League of Women Voters.

f) Linda is a bank teller.

g) Linda is an insurance salesperson.

h) Linda is a bank teller and is active in the feminist movement.

Asked to rank these statements in order of probability, the "great majority" of people said that (h) is more probable than (f), according to Amos Tversky & Daniel Kahneman, *Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment,* 90 PSYCHOLOGICAL REV. 293 (1983). (I've slightly modified the example given in that paper to more clearly enumerate the options.)

Here's the problem with that judgment: it's *impossible.* Remember that probabilities lie on the range 0-1 inclusive, and that the joint probability of two independent events is the product of their individual probabilities. It follows that the *maximum* probability of Linda being active in the feminist movement is 1, in which case $P(h) = P(f)$; in all other possible cases, $P(h) < P(f)$.

(Note that there's been a ton of debate on the original study, mainly focused on whether experimental subjects really understood (f) to be talking about P(Bank teller) alone and whether they understood (h) to be asking about the conjunction independent events P(Bank teller) and P(Feminist). There have been lots of replication attempts for this experiment in slightly different forms in order to nail it down. If you're worried about them, go read the literature. For present purposes, I just want you to see the error: joint probabilities of independent events can never be more than the individual probability of the least probable individual event.)

Monty Hall

There's a famous puzzle called the "Monty Hall Problem." It goes like this: imagine you're in a game show, and the host invites you to pick from three closed doors. One of the doors holds a fancy new sports car, and the other two hold goats. (Let us suppose, perhaps counterfactually, that you prefer a sports car to a goat.) After you choose, the host opens one of the *other* doors, always to reveal a goat.

So now you've picked a door, and the host shows you a goat behind one of the other doors. I want to emphasize this, because it's very important: *the host knows what door has the car, and s/he always opens a door you didn't choose, and it's always a goat door.* Then the host asks if you want to stick with your

original choice, or switch to the remaining closed door. If you want to get the sports car, should you stay or switch?

Back in 1990, Marilyn vos Savant, who apparently occupies the spot in the Guinness World Record book for "highest IQ," published this problem in a magazine, and declared that the correct answer was to always switch. In response, many, many people, including a number of mathematicians, blew their stacks and wrote angry letters telling her how wrong she was. (To be fair, her original description of the problem was a bit ambiguous about what the host knows and whether the host always revealed a goat.)

Here's how you might (incorrectly) reason to conclude that she was wrong. At the start of the game, each door had a .33[3333...] chance of having the sports car behind it. Initially, you picked a door at random, so you have a .33 chance of being right to stick. The remaining door also had a .33 chance of being the one with a sports car behind it. Nothing about what the host did changed those probabilities, so you have an equal chance of getting the sports car whether you stick with your original choice or switch.

By contrast, here's the correct way to reason. The host actually gave you additional information when s/he opened the door. Originally, your choice (call it the A-door) had a .33 chance of being correct, and hence a .66[6666666...] chance of the sports car being in one of the two remaining doors (call them the B-doors). So now the host has ruled out one of the B-doors for you, you still know that there's a .33 chance of the car being behind the A-door and a .66 chance of the car being behind the B-doors... but the number of B-doors has been reduced from 2 to 1! So there's a .66 chance of the car behind behind the one door left. You should switch. (Exercise for the reader: use Bayes' rule to prove this mathematically.)

This is a famously difficult problem, and the vast majority of people initially tend to pick the wrong answer. It tends to break our ordinary intuitions, and as a result serves as an excellent demonstration of the fact that you really can't trust your intuitive judgments when it comes to probability.

In case you still don't believe me (and I understand if you don't), you can see this explanation paired with a computer simulation of switching and sticking strategies at http://paul-gowder.com/montyhall/ — which even includes a link to a mathematical explanation (but really, try to do that part yourself before following the link).

Here's another way to think of the problem (with thanks to my friend Corey Yanofsky):

1. Host never offers you the chance to switch; you win 1/3 of the time.

2. Host offers, you're stubborn and never switch: always produces the same outcome as if the host never gave you a chance to switch. Hence 1/3 of the time.

3. Always switch: wins whenever a stubborn player would lose and loses whenever the stubborn player would win; therefore the switcher wins 2/3 of the time.

Here's still another way to think of the problem: suppose instead of 3 doors, there are 100 doors, and the host opens 98 of them. Does switching seem more plausible now?

Application: Expected Value

Economists define the expected value of a risk as the sum of the probabilities of every possible result of that risk, multiplied by their value. For example, suppose you buy a lottery ticket with a 1% chance of a 10,000 dollar payoff and a 10% chance of a 100 dollar payoff, and otherwise you get nothing. Then the expected value (in dollars) is:

$$.01 * 10000 + .1 * 100 = 110$$

For purposes of valuing options, outcomes with identical expected values are usually treated as equal. For example, it would be roughly equivalent to pay $110 (give up a 100% chance at $110) for that lottery ticket, because that trade is equal expected value. (This is not to say that any actual person would be willing to make that trade. Some people are "risk averse" – they might only be willing to pay, say, $100 for that lottery ticket, because they don't want to take the risk of ending up with nothing. Others, like gamblers, might be risk-seeking: people in Vegas regularly pay more than the expected value for the lotteries they play.)

Loss of a Chance

Now let's work through a common torts problem. Suppose that a plaintiff experiences a loss in probability of something as a result of the defendant's negligence.

For example, suppose that the plaintiff went into the defendant doctor's clinic complaining of leg pain. The patient actually has an infection, but the doctor negligently misdiagnoses it as early-stage arthritis and tells the patient to just take some aspirin. A week later, the patient comes back complaining of much worse pain. This time, the doctor can't miss the infection, and rushes the patient into the ER for a massive dose of antibiotics. Unfortunately, it's too late, and the patient loses the leg.

You're a juror at trial, and you hear credible expert testimony suggesting that the plaintiff had a 60% chance of keeping the leg if the doctor had administered appropriate treatment at the first visit, but that went down to 10% by the time

of the second visit. You also hear credible expert testimony placing the economic, emotional, etc. value of a lost leg at $500,000. Can you use the idea of expected value to calculate plaintiff's damages?

Here's one way you might want to do it. At the time of the first visit, absent negligence, the expected value the plaintiff had in the possible states of affairs relating to that leg was .6 * 500000 + .4 * 0 = $300,000. At the time of the second visit, the expected value the plaintiff had in the leg was .1 * 500000 + .9 * 0 = 50,000. Therefore, the plaintiff suffered a net loss in expected value terms of $250,000.

Interestingly, it turns out that lots of cases don't go down this way. First of all, many courts would consider the plaintiff's actual injury to be a reduction to 0% rather than 10% of the value of the leg, based on the principle that the defendant is responsible for the actual foreseeable injury that plaintiff suffers as a result of his/her negligent conduct, no matter how likely or unlikely. But second, while some courts award proportional damages that aim at something like the expected value of the lost probability of plaintiff's avoiding damages, there are also cases awarding full damages in this situation, and other cases awarding no damages. Sometimes this has depended on whether the original chance was greater or less than 50%. Some courts engage in some loose talk about whether the negligence was a "substantial factor" in the ultimate harm, which basically means "punt and let the jury figure it out." Courts do all kinds of weird math. The doctrine is a mess. (An extended summary can be found in Tory A. Weigand, *Loss of Chance in Medical Malpractice: The Need for Caution*, 87 Massachusetts Law Review n. 1 (2002), online here.) Should you find yourself with such a case, it's critical to know your jurisdiction's law.

But maybe if more lawyers actually understand how probability work, we can get the doctrine sorted out one day. Such, at least, is the dream.

Some more examples of probability in the law

The law is shot through with ideas about probability, uncertainty, risk, and chance. Consider the following examples, all from core law school courses, and most from the first year:

- Warrants and arrests must be supported by "probable cause."

- A civil plaintiff must ordinarily prove her case by a "preponderance" of the evidence, which is usually understood to mean that the elements of her case are more likely than not to be true.

- FRE 401(a) provides that the first element of the test for relevant evidence is that "it has any tendency to make a fact more or less probable than it would be without the evidence."

- In criminal law as well as in torts, we say that one the rationales for punishment is deterrence, that is, to make doing harm to others too costly to the defendant to be worth doing. The usual way we calculate the deterrent value of a punishment is probabilistically: we say (roughly) that ideal deterrence should set the product of the cost of the punishment and the probability of having that punishment inflicted to be equal to the harm done to others by the punishable act.

- In torts, the Hand formula for breach of the duty of care in negligence cases, BPL, is meant to capture the burden of taking a precaution, balanced against the probability that the loss will occur absent that precaution as well as the cost of the loss.

A Footnote on Odds

Odds is the gambling-style representation of probability, like "4 to 1" and so forth. "4 to 1 odds that e" means "the event E is 4 times as likely as Not-E," which translates to ordinary probability language as an 0.8 probability of E. It's easy to convert from odds back to decimal probability. It's easy to convert back and forth from odds to probability. In order to do so, let's represent odds with a doublecolon (sometimes it gets represented as a fraction, but that just makes the notation super confusing and I refuse to do it) such that A to B odds is A::B.

To go from odds to probability:

$$A :: B = \frac{A}{A + B}$$

So, in our example, "4 to 1 odds" converts to probability as $4 :: 1 = \frac{4}{4+1} = \frac{4}{5} = .8$

Actually, odds isn't exactly the same thing as a representation of probability, it's technically defined as a different concept derived from probability. So we can represent our 4:1 odds as the fraction $\frac{4}{1}$ and actually treat it like a fraction, for example, and simplify it to 4. Then we could say "odds of 4" which is equivalent to saying "probability of .8." But I'll just use the double colon notation to keep it clear what's happening.

What odds really are is the probability of the event divided by the probability of not the event. Which means that conversion from odds to probability is as simply as computing that. Assign the variable P to the probability of our event, and then odds are just

$$\frac{P}{1 - P}$$

For the most part, people in the sciences talk about probability rather than odds. People in gambling talk about odds, and some Bayesian statisticians do as well,

mainly because it makes a few computations easier. In this course, we'll almost entirely talk about probability, not odds (with one brief exception toward the end).