Probability Part 1: Rules and Patterns: Crash Course Statistics #13

Crash Course: Statistics
https://youtube.com/watch?v=OyddY7DlV58
https://nerdfighteria.info/v/OyddY7DlV58

===== (00:00) to (02:00) =====

Hi, I'm Adriene Hill and welcome back to Crash Course Statistics. If you've every seen a face in an onion or a grilled cheese or any other inanimate object, you've experienced pareidolia which is a product of our brains that causes us to see the pattern of a face in non-face objects. This happens because our brains are so good at seeing patterns that they sometimes see them when they're not really there, like a face in this bell pepper. And faces aren't the only patterns we see. Our brains recognize patterns everything especially in sequences of events like the kind we're going to be talking about today as we start talking about probability.

(Intro music)

Alright first let's just establish a more specific definition of what probability is, because the way we use the word in everyday life can be different from the way we use it in statistics. Statisticians talk about two types of probability: empirical and theoretical.

Empirical probability is something we observe in actual data, like the ratio of girls in each individual family. It has some uncertainty because, like the samples in experiments, it's just a small amount of the data that's available. Empirical probabilities like sample statistics give us a glimpse at the true theoretical probability, but they won't always be equal to it because of the uncertainty and randomness of any sample.

Theoretical probability, on the other hand, is more an idea or truth out there that we can't directly see. Just like we use samples of data to guess what the true mean or standard deviation of the population is, we can use a sample of data to guess what the true probability of an event is. Say you play a slot machine over and over and over and over and over and over. You'll be able to guess the probability of winning the jackpot by counting the number of times you win and dividing it by the number of times you played. If you play a hundred time a win six times, you can be pretty sure that the probability of getting a jackpot is around six out of a hundred or six percent.

===== (02:00) to (04:00) =====

Now, this isn't to say that you can rule out the true probability is 5% or even 10%, but you're relatively sure that it's close to 6% and not, say, 99% so the empirical probability can be a good estimation of the theoretical one, even if it's not exact.

So far, we've been talking about the probability of just one event, but often there may be two or more events that we want to consider. Like, what if you want to know the probability of picking a purple or a red Skittle from a bag? The proportion of each color in a bag of Skittles is roughly equal: 20% for each of the five colors. So let's say you randomly select a Skittle without looking. For this, we need the addition rule of probability. Since a Skittle can't be two different colors at once, the color possibilities are mutually exclusive. That means the probability of a Skittle being red and purple at the same time is 0, so we can use the simplified addition rule, which says that the probability of getting a red or purple Skittle is the sum of the probability of getting a red and the probability of getting a purple. Since we're gonna be talking a lot about probability in the next few episodes, I"m gonna introduce a little notation. Instead of writing out the probability of red, we can use the notation P(red). The probability of getting a red or a purple would be written p(red or purple). So far, we know what the probability of red or purple is: It's p(red) + p(purple), or .2+.2=.4, or 40%. I like all Skittles, so the probability that I will get a Skittle I like is .2 + .2 + .2 + .2 + .2, that's 100%. Good odds!

Red and purple Skittles are mutually exclusive, but not all the events we're interested in are. For example, if you roll a die and flip a coin, the probability of getting tails is not mutually exclusive of rolling a six. Since you can both roll a six and flip tails in the same turn. Since P(Tails), or six, is not equal to zero, these two events are not mutually exclusive and we'll need to adjust our addition rule accordingly. The full version of the addition rule states that P(tails or 6) = P(Tails) + P(6)-P(tails and 6).

===== (04:00) to (06:00) =====

When two things are mutually exclusive, the probability that they happen together is 0, so we ignored it, but now the probability of both these things happening is not 0, so we need to calculate it. You can see here there are 12 possible outcomes when flipping a coin and rolling a die. There are six outcomes with tails and two outcomes with a six. If we add all those together, we get eight, but by looking through the chart, we can tell there are only seven possible outcomes that have either tails or six. When we count tails and sixes independently, we double count the outcomes that have both. If we didn't subtract off the probabilities of tails and six, we would double count it.

Let's put these probabilities into a venn diagram, and we can see even more clearly why we need to subtract p(tails and 6). If this circle is all the times we flipped tails and the other circle is all the times we roll a 6, the overlapping area is counted twice if we simply added the two circles together. In this simple case, we can easily see what the probability of tails and six is, but sometimes it's not so easy to figure out. That's why we have the multiplication rule, which helps us figure out the probability of two or more things happening at the same time.

Let's say you just found out that actor Cole Sprouse goes to your local IHOP pretty often, and there's a 20% chance he'll be there for dinner any given night, and yeah, I know, that's not how people work, but we're gonna say that that's how Cole Sprouse works. Anyway, to top that off, your local IHOP has a promotion where they randomly select certain nights to be free ice cream nights, in the hopes that customers will keep coming back in case that night is the night. Each night there's a 10% chance that it will be free ice cream night. Now, you love ice cream and you like Cole Sprouse, as do we all, and your perfect night would include them both, so you try and calculate the probability that will happen on your visit tonight. Using the multiplication rule, multiply the probability that Cole Sprouse will be at the IHOP, .2, with the probability that it will be free ice cream night, .1 and you come to the sad realization there's only a 2% chance that you'll get to see Cole and get free dessert tonight.

===== (06:00) to (08:00) =====

When we wanna know the probabiity of two things happening at the same time, we first need to look at only the times when one thing, Cole Sprouse at IHOP is true, which is 20% of the time. Now that we reduced our options to just Cole nights, out of all of those Cole times, how often is it free ice cream time? Only 10% of Cole nights. 10% of the original 20% leaves only a 2% chance that both will happen at the same time, but you can always change your expectations and calculate the probabilty of getting either by using the addition rule. Cole or free ice cream, which is calculated by adding the probability of Cole to the probability of free ice cream, minus the probability of both, so we don't double count anything. You realize there's a 28% chance that something good will happen tonight, so you decide to go. No matter what, you're gonna get french toast.

Cole Sprouse and free ice cream night are independent. Cole doesn't have any secret knowledge about when is free ice cream night, so it has never affected his decision to come. Two events are considered independent if the probability of one event occuring is

not changed by whether or not the second event occurred. In more concrete terms, if Cole's decision to go to IHOP is independent of IHOP's decision to give out free ice cream, then the probability of Cole showing up should be the same on both ice cream and non-ice cream nights, since he's just choosing randomly.

We write conditional probabilities as P(Event 1) given (Event 2). Conditional probabiliites tell us the probability of Event 1 given that Event 2 has already happened. If two events are independent, like Cole and ice cream night, then we expect P(Cole) given ice cream night to be the same as just plain old P(Cole) since the two things are unrelated. If P(Cole) given ice cream wasn't the same as plain old P(Cole), that means that ice cream night somehow affect Cole's decision to show up at IHOP. We calculate conditional probability, P(Event 2) given Event 1 by dividing the probabilty of Event 1 and Event 2 by the probability of Event 1.

===== (08:00) to (10:00) =====

The role of conditional probabilities are particularly important when we consider medical screenings. For example, when screening for cervical cancer, it used to be recommended that all adult women get screened once a year, but sometimes the results of the screenings are wrong. Either they can say there's something abnormal when there isn't, that's called a false positive, or that everything is all clear when it's really not, called a false negative. This is exactly the kind of scenario where knowing the likelihood that something is actually abnormal, in this case cervical cancer, given that you've gotten positive test results, would be useful. That is, P(Cancer) given positive test. When looking at the data of people who don't have cancer, three percent will get a false positive, and people who do have cancer will get false negatives 46% of the time. This means we miss a lot and maybe freak some people out who don't need to be freaked out. The logic of conditional probabilities can help us make sense of why doctors have recently recommended that these tests be done less frequently in some cases.

In the United States, the rate of cervical cancer is about .0081%, so only about eight in 100,000 women get cervical cancer. Using our rates of false negatives and positives, we can see that for every 100,000 women in the US, only about four, and we are rounding here, of the about 3,004 people with positive tests actually had abnormal growth. That means the conditional probability of having cancer, given that you've got a positive test, is only .1%, give or take, we're rounding, and these positive tests require expensive and invasive follow up tests, and I wanna point out here that conditional probabilities aren't reciprocal. That is to say, P(Cancer) given positive test isn't the same as P(Positive test) given cancer. That's about 50%. In real life, you're not always gonna know the probability of Cole Sprouse showing up at the IHOP.

===== (10:00) to (12:00) =====

He's more unpredictable than I gave him credit for. Unpredictable like pretty much the rest of life. It can be really difficult to put a specific probability on a lot of everyday situations, like how likely it is that your teacher will call in sick today, like whether or not you're gonna catch all the red lights on your way to school. Probabilities can, as we've seen, require a lot of calculations and there's not always time for that, but that doesn't mean they belong only on the school side of your brain.

Say you wanna go out on Friday night with friends. More than anything, you don't want it to suck. Last week, you wound up on the couch watching Sandy Wexler again. You know it'll be hard to get tickets to see Black Panther, so you make a backup plan just in case. You can always stream Get Out without stealing it of course, but if you're determined to see Black Panther in the theater, probability will help you set your expectations. If you'll only settle for center row tickets, you're more likely to be disappointed. Your chance of seeing Black Panther is gonna be greater if you're willing to settle for whatever tickets you can get.

Probabilities help us understand why it makes sense to apply to more than one college, why you shouldn't expect that the first short story you write will get an A and be published in The New Yorker, and how likely it is that you'll get mono, given your significant other has mono. Probability can help you figure out that, too.

Thanks for watching, we'll see you next time.

CrashCourse Statistics is filmed in the Chad and Stacey Emigholz Studio in Indianapolis, Indiana, and it's made with the help of all these nice people. Our animation team is Thought Cafe. If you'd like to keep CrashCourse free for everyone forever, you can support the series at Patreon, a crowdfunding platform that allows you to support the content you love. Thank you to all our Patrons for your continued support. CrashCourse is a production of Complexly. If you like content designed to get you thinking, check out some of our other channels at complexly.com. Thanks for watching.

===== (12:00) to (12:01) =====