CLARK UNIVERSITY 1887

Conditional Probability
Math 217 Probability and Statistics

Prof. D. Joyce, Fall 2014

The rare disease Xox. This example also has to do with the general question of evidence.

Suppose you have a medical exam and part of that exam includes a blood test. Your blood sample is sent to a lab that checks for a whole slew of things. The lab finds your blood sample tests positive for a rare disease Xox. Your doctor tells you the results and tries to assure you not to be too worried about it until the disease is confirmed. But you find out that (1) this is a pretty rare disease in that only 1 in 1000 people have Xox; and (2) the lab is pretty accurate—it's only wrong in 1 of 1000 lab tests for Xox, that is to say, if a person doesn't have Xox, the lab test will be negative 999 times out of 1000, and if a person does have Xox, the lab test will be positive 999 times out of 1000. Are you worried that you have Xox?

We'll discuss this in class, and if we don't come to a conclusion, we'll revisit it later.

Conditional probability. The idea. We'll look at a couple of examples. We'll try to figure out what "the probability that event F occurs given that event E has occurred." We'll pronounce that more briefly as "the probability of F given E," and write it $P(F|E)$.

A uniform discrete example. Let's start with a uniform discrete probability distribution to see what's going on. Let's toss two fair dice, a red one and a green one. For this situation, the sample space Ω has 36 outcomes, each outcome being an ordered pair of numbers, the first being what the red die shows, the second what the green die shows.

Suppose someone tosses the dice and tells you that their sum is 6. We can ask, now, what the probability is that one of the dice came up 4.

Here, E is the event that the sum is 6, and F is the event that one of the dice came a 4. Without knowing that E has occurred, we would say the probability that one die shows 4 is $\frac{11}{36}$, since there are 11 of the 36 outcomes with at least one of the dice showing 4. Thus, $P(F) = \frac{11}{36}$. But that's not the question here. We need to figure out $P(F|E)$. Since E has occurred, our new sample space isn't the old sample space Ω that has 36 elements. Instead, our new sample space is the subspace E which has only 5 outcomes in it:

$$E = \{(5,1),(4,2),(3,3),(2,4),(1,5)\}.$$

Now, inside this sample space, we want to know how many outcomes belong to F. Those will be the outcomes in the intersection $E \cap F$. There are two such outcomes

$$E \cap F = \{(4,2),(2,4)\}.$$

Since the original 36 outcomes were equally probable, now that we've restricted our attention to E, the 5 outcomes in E are equally probable. Since two of these are in $E \cap F$, we conclude that $P(F|E)$ equals $\frac{2}{5}$.

Although we computed this conditional probability by taking the ratio of outcomes in $E \cap F$ to outcomes in E, that is, $\dfrac{|E \cap F|}{|E|}$, we could have taken the ratio of their probabilities $\dfrac{P(E \cap F)}{P(E)}$ and gotten the same number. It's this last formulation

$$P(F|E) = \frac{P(E \cap F)}{P(E)}$$

that generalizes to situations besides uniform discrete probabilities.

A nonuniform discrete example. A Bernoulli trial has one of two outcomes, one we'll call success, and the other failure. The probabilities of success and failure are assumed to be 1/2 but could be any fixed probability. We'll use p to denote the probability of success and $q = 1 - p$ to be the probability of failure.

For this example, let's take a sequence of 6 Bernoulli trials, each with probability p of success. The sample space Ω has $2^6 = 64$ outcomes, but they're not all equally probable. The probability of a particular outcome that contains k successes and $6 - k$ failures is $p^k q^{6-k}$.

Suppose that the first trial is a success. We can ask, then, what the probability that there are 4 successes in all.

Here, the event E is that the first outcome is success. The event F is that exactly 4 of the 6 trials are successes. We want to know $P(F|E)$, the probability that exactly 4 of the 6 trials are successes given that the first is.

To answer this question, we change from the original sample space Ω to the subspace E. Since this involves nonuniform probabilities, we can't simply count outcomes like we did in the last example. Before we go on, we have to decide how to assign probabilities to the outcomes in E when E is the whole sample space, rather than just an event in Ω. The problem is, that as an event in Ω, E only has a probability of p (since the probability that the first trial is a success is p). But when E is the entire probability space we need that probability to change to 1. In other words, we need to scale up the probabilities. We can do that by multiplying them by $1/p$, that is, dividing them by p. Then the new probability of E will be $p/p = 1$. We'll scale up the probabilities of all the outcomes in E as well.

So, what does that do in this example. We want to know the probability of F given E, and that should be the scaled up probability for the event $E \cap F$ as an event in the new sample space E. The probability of $E \cap F$ as an event in Ω was $P(E \cap F)$, but we need to divide that by p to scale up its probability as an event of E. Therefore $P(F|E)$ equals $P(E \cap F)/p$. Let's see what that value is. The event $E \cap F$ is when the first trial is a success and there are exactly 4 trials among the 6 trials. That's the same as saying the first trial is a success and there are exactly 3 trials among the remaining 5 trials. Therefore,

$$P(E \cap F) = p \binom{5}{3} p^3 q^2.$$

Thus, the conditional probability is

$$P(F|E) = \frac{P(E \cap F)}{p} = \binom{5}{3} p^3 q^2.$$

This conclusion makes sense because, of course, the probability that exactly 4 of 6 are successes given the first was a success should be the same as the probability that exactly 3 of 5 are successes.

Conditional probability. The formal definition. So long as $P(E)$ is not 0, we define the conditional probability

$$\boxed{P(F|E) = \frac{P(F \cap E)}{P(E)}}$$

If $P(E)$ does equal 0, then we won't define the conditional probability $P(F|E)$. (There are some circumstances with continuous probabilities where $P(F|E)$ can be defined somehow even when $P(E) = 0$, but it isn't done through this definition.)

The multiplication rule. In many situations in order to find the probability of an intersection of events, it's easier to break it down into steps and find the conditional probabilities for each step. The definition of conditional probability yields this identity for the intersection of two events

$$P(E \cap F) = P(F|E)\, P(E)$$

which generalizes to the intersection of three events or more

$$P(E \cap F \cap G) = P(G|E \cap F)\, P(F|E)\, P(E).$$

Math 217 Home Page
at `http://math.clarku.edu/~djoyce/ma217/`