PROFESSOR: The law of total expectation will give us another important tool for reasoning about expectations. And it's basically a rule like the law of total probability, closely related to it really, for reasoning by cases about expectation. So it requires a definition of what's called conditional expectation.

So the expectation of a random variable R, given event A, is simply what you get by thinking of replacing the probability that R equals v by the probability that R equals v given A. So it's the sum over all the possible values that R might take of the probability that R takes that value, given A.

OK, with that definition, we can state the basic form of the law of total expectation, which says if you want to calculate the expectation of R, you can split it into cases, according to whether or not A occurs. It's simply the conditional expectation of R given A times the probability of A, plus the conditional expectation of R, given not A times the probability of not A. So it really looks [? as ?] the same format as the law of total probability.

Now, of course it generalizes to many cases. So the general form would say that I can calculate the expectation of R by breaking it up into the case that A 1 holds times the probability of A 1, the case that A 2 holds times the probability of A 2, through A n. And this could very well, and typically is, an infinite sum, where the [? A i's ?] of course, are a partition of the sample space-- so they're all the different cases, either A 1 or A 2 or A 3, they're disjoint. And altogether, they cover the entire set of possibilities.

Well, let's use this to get a nice different and simpler way-- more elementary way-- of calculating the expected number of heads and flips. So let's let of n be the expected number of heads and flips-- just shorthand, because the notational will be easier to work with than writing capital E brackets of H n. So what do we know about expectation of n?

Well, I can express it in terms of the expectation of the remaining flips. So if I have n flips to perform, they're independent. Then if I perform the first flip, something happens. And after that I'm going to do n more flips, and the expected number of flips is going to be the expected number on the remaining n minus 1 plus what happened now.

Well, if I flipped a head first, then I've got a 1 as adding to my total number of heads. And then I'm going to do n more flips, so the expected number of flips is going to be that 1 plus the expected number on the rest of them. If the first flip was not a head, it was a tail, then the total expected number of heads is simply the expected number of heads on the rest of the flips.

And these are two cases where I can apply total expectation. So by total expectation, the expected number in n flips is 1 plus e n minus 1 times the probability of a head, plus e n minus 1 times the probability of a tail. Well, now we could do a little algebra multiply through here by p-- that becomes a p, and this becomes a p times e n and minus 1. So I've got e n minus 1 times p, and e n minus 1 times q-- remembering that p plus q is 1, this simplifies to being simply e n minus 1 plus p.

Well, this is a very simple kind of recursive definition of e n, because you can see what's going to happen. Subtracting 1 from n adds a p. So if I subtract 2 from n, I add another p-- I get 2 p. And continuing all the way to the end, by the time I get to 0, I've gotten n times p.

And I've just figured out what I was familiar with already-- which we previously derived by differentiating the binomial theorem-- the expected number of heads in n flips is n times p. But this time I got it in a somewhat more elementary way, by appealing to total expectation.