So for practice with expectation, let's calculate the expected number of heads in n coin flips, and just working directly from the definition, because we have tools to do that.

So we're imagining n independent flips of a coin with bias p. So the coins might not be fair. The probability of heads is p. It would be biased in favor of heads if p is greater than 1/2 and biased against heads if p is less than 1/2. And we want to know how many heads are expected. This is a basic question that will come up again and again when we look at random variables and probability theory.

So what's the expected number of heads? Well, we already know-- we've examined the binomial distribution B n,p. B n,p is telling us how many heads there are in n independent flips. So we're asking about the expectation of the binomial variable B n,p.

Well, let's look at the definition. The definition of B n,p is it's the sum over all the possible values of B, namely all the numbers from 0 to n-- that's k-- of the probability of getting k heads. And this formula here is the probability of getting k heads, which we've worked out previously. n choose k times p to the k, 1 minus p to the n minus k.

Well, let's introduce an abbreviation, a standard abbreviation. Let's replace 1 minus p by q, where-- so p plus q equals 1, and they're both not negative and between 0 and 1. And when I express the expectation this way, it starts to look like something a little bit familiar. And our strategy is going to be to use the binomial theorem, and then the trick of differentiating it is going to wind up giving us a closed formula for this expression for the expectation of the binomial random variable.

So let's remember the binomial theorem says that the nth power of x plus y is the sum all from k equals 0 to n of n choose k, x to the k, y to the n minus k. And if I differentiate this, what happens is that on the left hand side, if I differentiate with respect to x, I get x plus y to the n minus 1 times n.

And if I differentiate the right hand side-- let's differentiate it term by term. And differentiating with respect to x is going to turn this n choose k x to the k, y to the n minus k into an x to the k minus 1 times k term. But I'd like to keep the n-- the k here and the k there matching. So that after differentiating, that becomes an x to the k minus 1. Let's multiply it by x to make it x to the k. And of course, I have to undo that multiplication by dividing the whole thing by 1/x.

So by differentiating the binomial formula, we get the following formula for this sum that is starting to look just like the expectation of B n,p, 1/x times the sum from k equals 0 to 1 of k times n choose k, x to the k, y to the n minus k.

Well, let's compare the two terms. So here's this term and there's this one. I'm going to replace this line by the formula for expectation of the binomial random variable. So this is what we're trying to evaluate, and I have this great theorem. You can see how they match up.

So what I'm going to do is replace p and q-- replace x and y in this general formula that I got by differentiating the binomial theorem with p and q. And what happens? So I just plug in the p and q. Now, the left hand side. p plus q is 1. So the left hand side is going to become n. And this right hand side now is exactly the expectation of B n,p-- this part of it, anyway. So what I'm going to wind up with is that n is equal to 1/p times the expectation of B n,p.

In other words, the expectation of B n,p is n times p, and that is the basic formula that we were deriving by first principles without using any general properties of expectation, just the definition of expectation and the stuff that we had already worked out in terms of the binomial theorem.