We end this lecture sequence with the most important property of expectations, namely linearity. The idea is pretty simple. Suppose that our random variable, X, is the salary of a random person out of some population. So that we can think of the expected value of X as the average salary within that population.

And now suppose that everyone gets a raise, and Y is the new salary. And generously, the new salary is twice the old salary plus a bonus of $100. What happens to the expected value of the salary, or the average salary?

Well the new average salary, which is the expected value of 2X plus 100, is twice the old average plus 100. So doubling everyone's salary and giving to everyone an additional $100, what it does to the average is that it doubles the average and adds 100 to it.

This is the linearity property of expectation in one particular example. It's a most intuitive property, but it's worth also deriving it in a formal way. And the derivation proceeds through the expected value rule.

We're dealing here with a particular function, g, which is a linear function. So we're dealing with a linear function, ax plus b. And we're dealing with a random variable, Y, which is g applied to an original random variable, X.

So the expected value of Y can be calculated according to the expected value rule. It's the sum over all x's of g of x times the probability of that particular x.

And we plug-in the specific form of the function, g, which is ax plus b. And then we separate the sum into two sums. The first sum, after pulling out a constant of a, takes this form. And the second sum, after pulling out the constant, b, takes this form.

Now, the first sum is a times the expected value of X. This is just the definition of the expected value. As, for the second sum, we realize that this quantity is equal to 1 because it is the sum of the probabilities of all the different values of X. And this concludes the proof of the linearity of expected values.

Notice that for expected values, what we have is that the expected value of Y, which is expected value of g of X, is this same as g of the expected value of X. The expected value of a linear function is the same linear function applied to the expected value. But this is an exceptional case. This does not happen in general. It's an exceptional function g that makes this happen. This property is true for linear functions. But for non-linear functions, it is generally false.