Robert_S2_L14

Wed, 1/12 1:33PM 8:58

SUMMARY KEYWORDS

covariance, correlation, standard deviation, equal, variance, denominator, multiplied, square root, numerator, squared, divided, deviations, calculate, algebraically, explained, series, rewrite, formula, violate, give

SPEAKERS

Robert McKeown

R Robert McKeown 00:04

Hello and welcome. In a previous video, we explained or at least show to you the correlation equation, which is this what you see on the slide in front of you, we explained how the numerator is equal to something called the covariance. And it was a measure of how the series of values x, and the series of values y moved together. Now, let's turn our attention to the denominator, the denominator might be intimidating to you. But if you look closely, I hope that you can recognize a familiar formula, namely that this right here is actually the variance of y. And this right here, well, that is the variance of x. The denominator is the square root of the variance of y, multiplied by the variance of x, or the denominator is the square root of the product of the variance of X and the variance of y. So this can be rewritten as the correlation between x and y is equal to the covariance between X and Y divided by the variance of x times the variance of y with the square root. And, of course, the square root of a square is just itself. And if you look here, this is exactly what I showed for you. Remember that, n over n then we had before that was just equal to one. So there was nothing algebraically wrong with what we were doing didn't violate any of the rules of algebra, the correlation was still the same. But we can calculate the correlation just by taking the covariance and dividing it by the standard deviation of x and the standard deviation of y. And that's what we have written right here in this note. Here's an example. Where we've got two series X and Y, you're being asked to solve for the means the standard deviations, the covariance, and the correlation between X and Y, or of x and y. The mean of x is equal to four and the mean of y is equal to one plus seven divided by three, which is equal to five. What about the standard deviation of x and y? Well, we have well, I may title this, we have three minus four, four minus four, and five, minus four, which give us negative one, zero and one respectively. They're the deviations from the mean, if we want to square each of these that's going to give us plus one zero and one, we sum them up. It gives us two. And if we divide by three, we get two thirds standard deviation of what about the deviation from the mean and Y. We've got one minus five squared, seven minus five squared, seven minus five squared. That's going to give us 416. Negative four squared is 16. It's going to be as four and that's going to give us four. We sum them all up, we get 24. And 24 divided by three is going to be equal to eight. What are we calculated here? Well, we calculated the variances So we have the variance of X and the variance of y. If we want to know, what is the standard deviation, we can take the square root. And there we go. What about the covariance?

So the covariance, remember, that's in our numerator here. If we want the covariance, we're going to take the product of the deviations from the mean and x multiplied by the product, or multiplied to say the product twice, multiplied by the deviations of the mean from the mean and y. So we're going to have three minus four times one minus five, four minus four, times seven minus five, and we're going to have five minus four times seven minus five again, this is going to give us negative one times negative four, we get four there. Okay, zero here, zero times any number is zero, and we've got two here. And so the covariance is equal to six. But of course, we have to do also divide by n. So six divided by three is going to be equal to two, and there is our covariance. Now if we want to calculate the correlation, it's going to be equal to covariance divided by the standard deviation of y multiplied by the standard deviation of x. And the numerator is just going to be two. So there's our covariance. Let me rewrite the formula that I'm using. And so what's the correlation here? I find that the correlation is equal to 0.866. Now, I wanted to introduce to you how to calculate the correlation covariance. Specifically, now we want to talk about how do we interpret this correlation? What does a correlation of 0.866 mean? And it turns out that there are boundaries on the value of our correlation, and we're going to look at that next.