Robert_S2_L07

Wed, 1/12 12:25PM 12:46

SUMMARY KEYWORDS

variability, range, equal, series, deviations, values, salaries, measure, calculate, central tendency, variance, social sciences, quartile range, highest, toronto maple leafs, standard deviation, spread, important, lowest, numbers

SPEAKERS

Robert McKeown

R Robert McKeown 00:05

Hello, and welcome to our module on an introduction to variability. Variability is very important in the social sciences, it's really important in business because, of course, business is risky, risky business, anything that's risky, is going to have variability. And hopefully when I make that connection to you, you can see how it's going to have applications to the social sciences. Now, variability, as it says in this slide in front of you reflects the difference, and values or how different is one value from another. And you can see that based in the name and the alternative names for variability, sometimes it's called dispersion from the word dispersed, which means that things are far apart. And quite literally, the spread, how, what's the what's the distance, how spread out are the values from each other. And so previously, we looked at finding a single measure for central tendency. Now, variability is going to try to do the same thing, but with measuring how far apart the values are. And first, I'm going to introduce you to this idea of variability. I'll show you a series that has a higher variability one that has a lower, and then I'll introduce to you a number of concepts, the range, the inter quartile range, the variance and standard deviation, which are all attempts to measure variability. Let's take a look at an example. We've got a series of numbers we've got X. And it's a series and we've got a Y series of values. And we're being asked to compute the mean of X and why. And then we're being asked to compute the deviations from the mean, and x and y. And if we want to calculate the mean and X, we add up all the values in the series and divide by the number of values in the series. So we're going to end up with 14 divided by four, which is equal to 3.5. And let's calculate the mean of Y, we've got zero plus one plus six plus seven, two also divided by four. And that's going to give us 14 over four, which equals 3.5. So according to the average measure of central tendency, these two series appear identical, they're both have a mean equal to 3.5. But you can see with your eyes that they're not exactly the same, their values are different access just threes and fours, while y has 016 and seven. Now, we might be interested in understanding which of these is more variable, which one has a higher spread than the other. And maybe you can see it clearly just by looking at the two series. But let's go ahead and compute the deviations from the mean for X and Y. So if we're considering the X series, we're going to have three minus 3.5 which equals negative .5, got another value that's identical. .5 and four minus 3.5, which equals .5. And of course, they sum up to zero because the mean when we take deviations from the mean, they're always going to sum up to zero. Now let's

compare the X series to the Y series we've got zero minus 3.5 equals negative 3.5. We have one minus 3.5, negative 2.5. And we have six minus 3.5 2.5 and seven minus 3.5. is equal to 3.5. And so although all the deviations add up to zero, you can see that each individual variation is larger than its similarly ordered counterpart in the X series. And so I haven't given you we're gonna do something along these lines actually measure the variance, or kind of gave it away the variability and why the variability acts. But for now, you can see that the deviations from the mean and y are an absolute terms larger than the deviations from the mean and X, Y varies more than X. Just looking at the numbers, we can kind of see that three and four are closer together than zero and seven, or one and six. And when we calculate deviations of the mean, we can see it again. Now, suppose we add in a new series, an additional series, Z, or is that it's got four numbers for values. And each value is equal to 3.5. And we're being asked the question, do any of the values in Z deviate from the mean, at all. And so it should be pretty clear, just by looking at the mean, Z is equal to 3.5. And all its deviations from the mean, are just going to be equal to zero. And so we can see here that Z has no variability. It has the same mean as X and Y, but it has no variability. It's one might even say it's constant. If we think back to some of the things you may have learned in precalculus, looks constant, it's always equal to 3.5. Now, in this part of our series, looking at various descriptive statistics, I'm going to introduce three actually four to you that were to look at the range, we're also going to look at something called IQR, the inter quartile range, which is related to the range, we're going to look at the variance and the standard deviation. And of these four, the most important is the standard deviation. That's the most intuitive one, one that you really want to walk away from understanding, if you never take another course in statistics, it's really valuable. And it's got an intuition behind it in terms of interpretation. Variance is very important. If you're when you're in finance, it's very important. And in economics, it's very important and other stuff with other statistical applications. But it's kind of hard to interpret on its own. And you'll see that mathematically, it's also very similar to standard deviation. The range is a simple measure of the variability in a series. So we're going to look at that next. And you might wonder, Well, why do we have dice over here as a picture? Well, because in the social sciences, variability, we want to be able to measure variability. So it could be a stand. And for our measures of risk, it's very useful in business and social sciences to have a measure for risk that comes up in statistics also comes up in Portfolio Management, and understanding your personal finances. Now, the rest of this video, we're going to look at the range, the simplest, simplest measure of variability. The range measures how far apart the while the endpoints are really, we're going to be looking at the M endpoints, because the range is measured as the highest value in a series, less the lowest value in a series. So if we want to find the range, X, it's going to be equal to the highest value less the lowest, so four minus three is equal to one, and it has a range of one. The range of Y. Well, it's just going to be hate and zero, so it's eight minus zero, the range of Y is equal to eight. And the range of Z is just equal to zero and has no range has no variability. And we can actually see that when we measure the range So if we were just if all we had was information about the range. For these three series, we could say that y has the highest variability followed by X, and followed by Z, which has no variability at all. We'd look back to some series we've seen before we've got stem and leaf tables are plots for the Toronto Maple Leaf salary and the Toronto Raptors salaries. If we want to calculate the range for the Toronto Maple Leafs, we've got $11.6 million minus $700,000. And the range for the Toronto Maple Maple Leafs is $10.9 million and so we just took the highest value, and subtracted the lowest value and that gave us the range. Looking at the the rapture series, this is 10s of millions and this over here is hundreds of this line at the top is hundreds of 1000s. So write it out in full, we've got $33 million, there's spicy P Pascal Siakam and the lowest paid raptor we have here earning $600,000 And so we're going to end up with 32,400,000. So the range is pretty simple, simple to use NBA, or at least the Toronto Raptors in the NBA. Their salaries are have a higher variability, a bigger spread, or more dispersed than NHL hockey players, or at least those

players who play for the Toronto Maple Leafs. And we can see that because the range of Maple Leaf salaries is 10 point 9 million. While the range Toronto raptor player salaries is 32 point 4 million