10-2, 10-3 Linear Correlation and Regression

Screen 1
In this lecture we’re going to look at section 9-2, “Correlation”, and section 9-3, “Regression”.

First we’re going to setup the calculator. To do this, press 2nd-0, which brings up the catalog function, and then scroll quite far down to get to the option ‘DiagnosticsOn’, and just select it and press ‘Enter’. The calculator will say ‘done’, but it doesn’t look like like anything has happened.

You’ll also be using table A-6, so have that handy. You can find it in the back of the book. It’s also in the formulas and tables handout.

So what we’re going to do is number-12 in the both sections 9.2 and 9.3. These two problems use the same data-lists. So go ahead and take a moment – and put the x-values, the temperatures, into L1, and the y-values which are the times into L2.

You may notice your calculator rounds in the display for the times, but that’s ok. So in 9-2, #12 we’re asked: Is there a significant linear correlation? Our process for answering this will be to find ‘r’, the correlation coefficient, and compare this to the critical value that we’ll find from table A-6.

While the book has you do a hypothesis test to find ‘r’, there’s another way that I’ll suggest here. To find ‘r’, you can press ‘STAT’, and then choose the ‘CALC’ menu option under the ‘STAT’ menu, and then ‘Linear Regression’, (ax + b), ‘LinReg (ax + b)’, then tell the calculator where the data is, L1, and L2, with a comma in between – and then press ‘ENTER’.

I get ‘r’ = 0.183 rounded. For the critical value we go to table A-6, and then (notice that we have 8 data pairs in our table – 1-2-3-4-5-6-7-8 pairs, and then alpha is equal to 0.05 unless we’re given some other value in the problem.) So from table A-6, n = 9, alpha = 0.05, I get a critical value of 0.707.

Now we can make our comparison compare our ‘r’ value to the critical value. We make the comparison, and since the absolute value of our critical value, 0.183, is less than 0.707, there is not a significant linear correlation.

Screen 2
Ok! Now we’re going to 9-3 # 12, which asks us a couple of things. We’re going to find the equation first of all to start off. That’s in the blue text up above the problems, and then we’re going to make a prediction – and we’re possibly going to use our equation – possibly not. We’ll talk about that when we get there.

So first of all we’re going to find the regression equation. To find the equation we do the same calculations as we did in 9-2. We’ll go ‘STAT’ - ‘CALC’ – ‘LinReg(ax + b)’, and then tell the calculator that the data in L1 and L2 in particular, and then press ‘ENTER’.

The parts of the display that I need are (First of all, y=ax + b reminds me of a format of the line I’m going to write down.), and then I just need the ‘a’ value. 0.032, I’ve rounded a few places – and the ‘b’ value, 145.186, again rounded to 3 places. We’re just gonna drop those in for ‘a’, and ‘b’ in the equation.

So notice I write my equation as y-hat because this is now based on data other than any sort of real world correlations – this is just our best approximation based on our data. So ‘y-hat’ becomes 0.032x (plain old ‘x’) + 145.186. That is my equation. It’s that part of the problem.
Ok. Now we’re going to move to the prediction part of the problem. Given that ‘x’ is now going to be equal to 73, we want to find the best predicted value of ‘y’. To be honest there may be a mistake here since 73 is actually one of our data values – I don’t know why we don’t just use the ‘y-value’ from the table, but in any case, we’re going to follow the rules and make a prediction for ‘x=73’.

So first of all we have to make a choice, and this is the key part of the problem. We may use $y\hat{\text{=}} 0.032x + 145.186$. We just found this equation, so why not just use it – but there’s another option. We may need to use $y\bar{\text{=}}$, that we can find by using one-variable stats on it.

So using ‘1-var’ stats on ‘L2’ where our y-values are, I find that $y\bar{\text{=}}$ is equal to 147.077. So I have to choose between one of these two things, and our critical value linear-correlation comparison from the last section will tell us. So we saw in 9-2 that we got a correlation coefficient of 0.183, and a critical value of 0.707.

Since our correlation coefficient of 0.183 is less than the critical value 0.707, we cannot use the equation to make predictions. Instead we choose to use our average ‘y-value’, $y\bar{\text{=}}$ equal to 147.077 – and so our best predicted value for ‘x=73’ is 147.077.

And we’re done!