Data Concepts – Regression on the the Calculator*

Kenny M. Felder

This work is produced by The Connexions Project and licensed under the
Creative Commons Attribution License †

Abstract

This module introduces the concept of modeling data using regression lines on a calculator.

1 Regression on the Calculator

What Kepler did is an example of "regression": finding an equation that models a particular set of data.

Kepler became famous because regression is hard. Who would have thought to look for $T^2 = ka^3$? Especially when you consider all the other equations that would still have a "this goes up, that goes up" relationship, such as $T = 2^a + 7$, or $T = \frac{1}{10-a}$, or maybe $T = 3\log 5a$?

Fortunately, we have a tool that Kepler did not have: the modern computer. Mathematical programs and graphing calculators can take a set of points, and find the line or curve of "best fit" to model the data.

As an example of this process, suppose that you have run an experiment and generated three data points: (2,1), (4,3), and (5,8). What function might model those data?

> NOTE: Note: the directions below are given for a TI-83 or compatible calculator. Many other calculators can perform the same functions, but the implementation details may look quite different.

1.1 Entering the data

1. Hit `STAT` to go into the Statistics menu.
2. Choose `Edit...` This brings you to a screen where you enter a bunch of L1 and L2 values.

*Version 1.1: Nov 15, 2008 11:41 am -0600
†http://creativecommons.org/licenses/by/2.0/

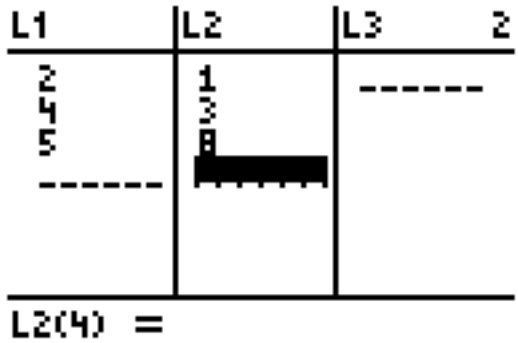
FIXME: A LIST CAN NOT BE A TABLE ENTRY. Enter the L1 and L2 values as follows: for each data point, the x-coordinate is in the L1 list, and the y-coordinate is in the L2 list.The screen to the right shows the points (2,1), (4,3), and (5,8).	

Table 1

1. Hit `2nd QUIT` return to the main screen.

1.2 Viewing the data, and guessing at the shape

Once you have entered your points into the L1 and L2 lists, your calculator can show you a "scatterplot"—which is a pointlessly fancy word for "a graph of a bunch of points," like you used to make when you were first learning what graphing was.

1. Hit `WINDOW` (near the upper-left-hand corner of the calculator).
2. Set the variables Xmin, Xmax, Ymin, and Ymax appropriately. For instance, for the three points shown above, the x-values are 2, 4, and 5, and the y-values are 1, 3, and 8. So it might make sense to set Xmin=0, XMax=10, Ymin=0, and Ymax=10—and that is how I did it in the drawing below. Of course, there are many other settings you could use. But if you go through the whole process and don't see any points, it's a reasonable guess that your window is not set properly.
3. Hit `Y=` (upper-left-hand corner of the calculator).
4. Then, hit the up-arrow key, so the focus moves to `Plot1` (which will start blinking).
5. Hit `ENTER`. It's actually impossible, at this point, to see that anything has happened. But if you down-arrow **away** from `Plot1`, you should see that it remains darkened (white letters on a black background, instead of the other way around). This indicates that it has been selected.

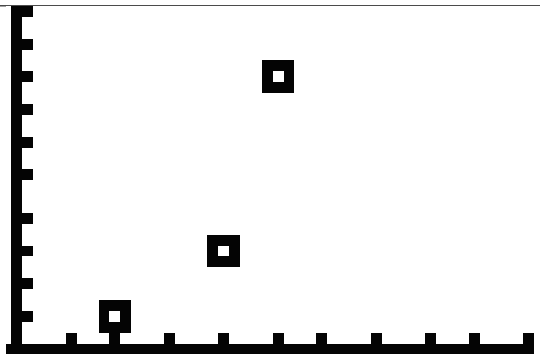
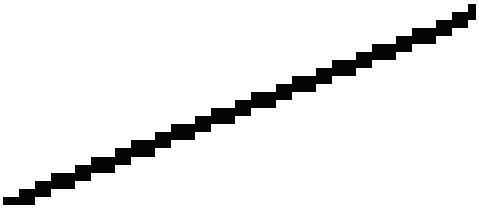
FIXME: A LIST CAN NOT BE A TABLE ENTRY. Hit `GRAPH` (upper-right-hand corner of the calculator). The calculator now displays the points. From the image, you can see that a quadratic (parabolic) or exponential function might be a reasonable guess, whereas a line or logarithmic function would be unlikely to fit.	

Table 2

At this point, looking at the data, it is often useful to categorize it in two ways.

First: is it **increasing or decreasing?** In our example, of course, the points are increasing. (Some data, of course, may be doing both at different times: consider, for instance, a parabola.)

Now, in the case of an increasing function, you can categorize it as one of the following.

	If it is increasing steadily, that suggests a line. (Remember that what makes a linear function linear is that it always goes up at the same rate, or slope!)
	continued on next page

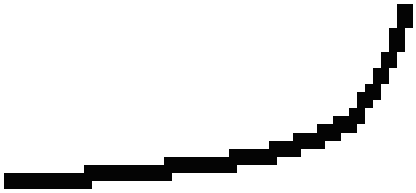
	If it is increasing more and more slowly, that suggests a logarithmic function. (A square root would also have this basic shape, but you cannot do a square-root-regression.)
	If it is increasing more and more quickly, that suggests an exponential function, or possibly the right side of a quadratic function (a parabola).

Table 3

Decreasing functions can be categorized similarly, of course. If a function decreases and **then** increases, a parabola is probably the best fit. Functions that go up, then down, then up again, are most likely to be higher-order polynomials.

1.3 Finding the formula

Once you have decided on the right shape, the hard work is done: the calculator takes care of the rest.

1. Hit `STAT` to return again to the Statistics menu.
2. Hit the right-arrow key to go to `CALC`.
3. At this point, you have several choices. `LinReg` will give you a line that best fits your points. `QuadReg` will give you a quadratic function, **aka** a second-order polynomial. There are also options for "cubic" (third-order polynomial), "quartic" (fourth-order polynomial), logarithmic, or exponential curves. Choose the one you want and hit `ENTER`.

The calculator does **not** graph your curve for you, but it does tell you what the curve is. For instance, if I run a `QuadReg` on the data above, the calculator gives me:

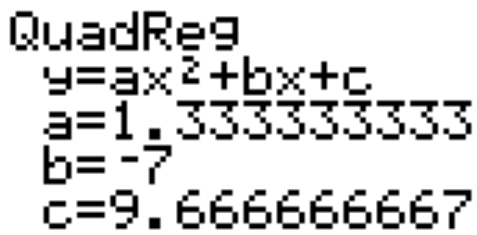


Figure 1

This tells me that the best quadratic fit for my data is the curve $y = 1.33x^2 - 7x + 9.67$. One way to double-check this, of course, is to enter $Y1 = 1.33x^2 - 7x + 9.67$ and then graph it, and see how closely it approximates the points!

NOTE: Remember that whatever option you choose, it will operate on the points you have entered in the L1 and L2 lists, so make sure your data is correctly entered there!

1.3.1 Starting over

There's just one more thing you have to know: once you've done this once, how do you clear out the lists to enter new ones? Here is one way to do it.

1. Hit `MEM` (you do this by hitting the yellow `2nd` key, and then hitting `+`).
2. This brings up a menu. Choose `ClrAllLists`.
3. Then—after you return to the main screen—hit `ENTER` and the lists are emptied out.