TANTON'S TAKE ON …

FITTING EQUATIONS TO DATA

CURRICULUM TIDBITS FOR THE MATHEMATICS CLASSROOM

MAY 2013

Standard algebra courses have students fit linear and exponential functions to two data points, and quadratic functions to three data points. Here are my thoughts and approaches to these matters.

THREE DATA POINTS:

This section is taken from a new online course currently under construction. For a sneak peak at ***"EVERYTHING ABOUT QUADRATICS"*** *have a look at:* ***www.gdaymath.org****. (All is there bar the videos to go with the text.)*

A typical, dry textbook question:

Find a quadratic function that fits the data:

x	2	5	7
y	7	10	3

That is, find a quadratic function that passes through the points $(2,7)$, $(5,10)$ *and* $(7,3)$.

The first, utterly appropriate response, should be: **WHY?**

Once (if?) sufficient context and impetus has been provided then my advice to students is:

JUST WRITE DOWN THE ANSWER!

(Don't bother with that unintuitive, unexplained `statlist/quadres` ugly calculator business.)

And here is the answer:

$$y = 7\cdot\frac{(x-5)(x-7)}{(-3)(-5)} + 10\cdot\frac{(x-2)(x-7)}{(3)\cdot(-2)} + 3\cdot\frac{(x-2)(x-5)}{(5)\cdot(2)}$$

Okay, I admit this looks scary.
But understanding and making sense of this crazy response is surprisingly straightforward.

First, if one <u>were to</u> expand this out we'd see that it is indeed a quadratic function of the form $y = ax^2 + bx + c$. (And since the author did not specify in which form she wanted the quadratic expression, there is no need to expand it! Why do work without a context or purpose to that work?)

Next… To understand this meaty formula, simply plug in some relevant x values.

Let's start with the first x value listed: $x = 2$. (Do it! Substitute it in.) Notice that the second and third terms are designed to vanish at $x = 2$ leaving us with only the first term to contend with:

$$7\cdot\frac{(x-5)(x-7)}{(-3)(-5)}$$

And with the value $x = 2$ inserted, the numerator and the denominator match (the denominator was designed to do this) giving us:

$$7\cdot 1$$

which is the value 7 as desired by the table provided.

In inserting $x = 5$ only the middle term

$$10 \cdot \frac{(x-2)(x-7)}{3 \cdot (-2)}$$

survives which, for $x = 5$, gives $10 \cdot \frac{3 \cdot (-2)}{3 \cdot (-2)} = 10$ as needed.

In the same way, for the value $x = 7$ only the third term is non-vanishing and has value

$$3 \frac{5 \cdot 2}{5 \cdot 2} = 3$$

Thus the quadratic

$$y = 7 \cdot \frac{(x-5)(x-7)}{(-3)(-5)} + 10 \cdot \frac{(x-2)(x-7)}{3 \cdot (-2)} + 3 \cdot \frac{(x-2)(x-5)}{5 \cdot 2}$$

does indeed produce the values 7, 10 and 3 for the inputs 2, 5, and 7!

A wee bit of easy simplification makes it a tad friendlier to read:

$$y = \frac{7}{15}(x-5)(x-7) - \frac{5}{3}(x-2)(x-7) + \frac{3}{10}(x-2)(x-5)$$

Despite the visual complication of the formula one can see that its construction is relatively straightforward:

i. *Write a series of numerators that vanish, in turn, at all but one of the desired inputs.*

ii. *Create denominators that cancel the numerators when a specific input is entered.*

iii. *Use the desired y-values as coefficients.*

COMMENT: It takes but a moment to get the hang of this. All my students thus far in my career, after a just a wee bit of practice, have mastered it with ease!

EXAMPLE: Find a quadratic that passes though the points $(3, 87)$, $(10, \pi)$ and $(35, \sqrt{2})$.

Answer:

$$y = 87 \frac{(x-10)(x-35)}{(-7)(-32)} + \pi \frac{(x-3)(x-35)}{(7)(-28)} + \sqrt{2} \frac{(x-3)(x-10)}{32 \cdot 25}$$

CHECK: Put in $x = 3$. Do you get the output 87? Also, put in $x = 10$ and then $x = 35$.

EXAMPLE: Find a quadratic that fits the data

x	1	2	3
y	a	b	c

Answer:

$$y = a \frac{(x-2)(x-3)}{(-1)(-2)} + b \frac{(x-1)(x-3)}{1 \cdot (-1)} + c \frac{(x-1)(x-2)}{2 \cdot 1}$$

$$= \frac{a}{2}(x-2)(x-3) - b(x-1)(x-3) + \frac{c}{2}(x-1)(x-2)$$

PRACTICE: Find a quadratic that fits the data

x	0	4	5
y	3	3	18

Do a tiny bit of simplification to your answer.

IF YOU NEED TO SIMPLIFY ALL THE WAY… It can be done and actually isn't as bad as you might first think.

For example, consider the data:

x	2	-1	3
y	3	6	10

A quadratic that fits this data is:

$$y = 3 \frac{(x+1)(x-3)}{(3)(-1)} + 6 \frac{(x-2)(x-3)}{(-3)(-4)} + 10 \frac{(x-2)(x+1)}{(1)(4)}$$

Side Question: Is it: "**A** quadratic that fits the data is…" or "**The** quadratic that fits the data is …"?

A tiny bit of simplifying gives:

$$y=-(x+1)(x-3)+\frac{1}{2}(x-2)(x-3)+\frac{5}{2}(x-2)(x+1)$$

To handle the fractions, it might be easiest to put terms over a common denominator:

$$y=\frac{-2(x+1)(x-3)+(x-2)(x-3)+5(x-2)(x+1)}{2}$$

Expanding each product gives:

$$y=\frac{-2(x^2-2x-3)+(x^2-5x+6)+5(x^2-x-2)}{2}$$

We now see:

The x^2 *terms are:* $\frac{-2+1+5}{2}x^2=2x^2$

The x *terms are:* $\frac{4-5-5}{2}x=-3x$

The constant term is: $\frac{6+6-10}{2}=1$

Thus the quadratic that fits the data is:

$$y=2x^2-3x+1.$$

PRACTICE: Find a quadratic that goes through the points $(-3,-14)$, $(2,1)$ and $(3,-2)$. For fun, simplify your answer all the way!

PRACTICE: Something interesting happens if one tries to find a quadratic that fits the points $(2,7)$, $(3,9)$ and $(6,15)$.

a) Write down a quadratic that seems to fit these data points and simplify your answer.

b) What happened and why?

PRACTICE: Something goes wrong if one tries to find a quadratic that fits the data $(1,0)$, $(1,-2)$ and $(-1,-1)$.

a) Try to write a quadratic that fits this data.
b) What goes wrong and why?
c) Find an equation of the form $x=ay^2+by+c$ that fits this data!

PRACTICE: Why stop at quadratics?

Here is some data and here is a polynomial formula that fits it:

x	3	6	8	9
y	4	8	-3	12

$$y=4\frac{(x-6)(x-8)(x-9)}{(-3)(-5)(-6)}+8\frac{(x-3)(x-8)(x-9)}{(3)(-2)(-3)}-3\frac{(x-3)(x-6)(x-9)}{(5)(2)(-1)}+12\frac{(x-3)(x-6)(x-8)}{(6)(3)(1)}$$

a) *Write down a formula that fits the data:*

x	-3	5	6	8	10
y	2	8	16	13	5

b) Here is a table of data that spells my name. (Do you understand the connection?)

x	0	1	2	3	4
y	10	1	13	5	19
	J	A	M	E	S

I wrote a formula that fits the data perfectly and then had a computer algebra system do the simplifying for me. Here's the polynomial that spells my name:

$$p(x)=\frac{83}{24}x^4-\frac{331}{12}x^3+\frac{1657}{24}x^2-\frac{647}{12}x+10$$

[So putting in $x=0$ gives $p(0)=10$, and putting in $x=1$ gives $p(1)=1$, and so on.]

What is your personalized polynomial?

CHALLENGE: Prove that your polynomial has the property that, despite all the fractions, it is sure to give an integer output for each and every integer input! (Whoa!)

PRACTICE: Explain why $\sqrt{17}$ is a valid answer to this intelligence test question:

What is the next number in the sequence?

2 4 6 8 __

TWO DATA POINTS:

The standard curriculum wants students to do two things with two data points. Fit a straight line through them and/or fit an exponential function to them.

Two Data Points: Straight Line Fits

Let's illustrate two techniques through a specific example.

Find a linear function that fits this data:

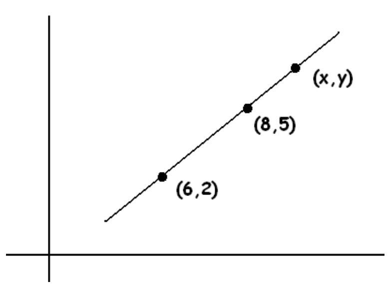
x	6	8
y	2	5

Approach 1: Just follow the previous technique and write down the answer!

$$y = 2 \cdot \frac{(x-8)}{(-2)} + 5 \cdot \frac{(x-6)}{(2)}$$

That is, $y = -(x-8) + \frac{5}{2}(x-6)$.

Approach 2: We like to believe that straight lines have the same value for slope no matter how we compute it. We have two points, $(6,2)$ and $(8,5)$, on the line we want and we want a formula that must be true for the coordinates of any other general point (x, y) on the line.



I can see three ways to compute the slope:

$$\frac{5-2}{8-6} = \frac{3}{2} \quad \text{or} \quad \frac{y-2}{x-6} \quad \text{or} \quad \frac{y-5}{x-8}$$

and they must all be equal. Equating any two of them gives an equation for the line!

For instance, $\frac{y-2}{x-6} = \frac{3}{2}$ is one possible equation. $\frac{y-2}{x-6} = \frac{y-8}{x-5}$ is another valid* equation.

COMMENT* Well .. These equations are "valid" at all locations where the denominators involved are not zero. Most everyone likes to perform a little algebra on these equations to avoid possible worries of dividing by zero. For example, $\frac{y-2}{x-6} = \frac{3}{2}$ can be presented as $2(y-2) = 3(x-6)$ - and this holds even for $x = 6, y = 2$ - and $\frac{y-2}{x-6} = \frac{y-8}{x-5}$ as can be given as: $(x-5)(y-6) = (x-6)(y-8)$, which also holds valid at the previous trouble spots.

Two Data Points: Exponential Fits

An exponential function is a function of the form $f(x) = a \cdot b^x$. (Note $f(0) = a$ is the function's *initial value*.) Since exponential functions arise in models of population growth, one might be interested in fitting exponential curves to data.

EXAMPLE: *A scientist grows a yeast culture in a Petri dish. After 3 days the mass of the yeast was 8 grams. After 7 days, 13 grams. Assuming that the yeast population grows exponentially, find an exponential function that fits this data.*

Answer: We have the table:

t	P(t)
3	8
7	13

where *t* is time (in days) and $P(t)$ is the population of yeast measured in terms of weight (grams). We seek a formula:

$$P(t) = a \cdot b^t$$

that fits this data.

There are two approaches.

DIFFICULT APPROACH

Put $t=3$ and $t=7$ into this formula to gain two equations:

$$a \cdot b^3 = 8$$

$$a \cdot b^7 = 13$$

Solve for *a* in one equation and substitute into the second … OR … divide the two equations to obtain:

$$b^4 = \frac{13}{8}.$$

This gives:

$$b = \left(\frac{13}{8}\right)^{\frac{1}{4}}.$$

Now substitute this value of *b* into the first equation, say, to obtain:

$$a\left(\frac{13}{8}\right)^{\frac{3}{4}} = 8$$

giving

$$a = 8 \cdot 8^{\frac{3}{4}} \cdot 13^{-\frac{3}{4}} = 8^{\frac{7}{4}} \cdot 13^{-\frac{3}{4}}.$$

Thus we have:

$$P(t) = a \cdot b^t = 8^{\frac{7}{4}} \cdot 13^{-\frac{3}{4}} \cdot \left(\frac{13}{8}\right)^{\frac{t}{4}}.$$

SIMPLER APPROACH:
JUST WRITE DOWN THE ANSWER!

One has to be savvy, but one can do it.

First note that if the data were simpler, then it would be easier to read off a formula that fits. For example, for the table:

t	P(t)
0	8
1	13

we see that the initial value is 8 (so $a=8$) and after 1 day, the data has grown by a factor of $\frac{13}{8}$. Thus the function

$$8 \cdot \left(\frac{13}{8}\right)^t$$

fits.

CHECK: Put in $t=0$ and $t=1$ to see that the values 8 and 13 do indeed appear.

Now suppose the data were "slowed down" by a factor of 4:

t	P(t)
0	8
4	13

That is, we are running though the time values four times as slowly:

Since "time has changed" by a factor of four, the formula fitting the data changes to:

$$8 \cdot \left(\frac{13}{8}\right)^{\frac{t}{4}}$$

CHECK: Put in $t=0$ and $t=4$ to see that the values 8 and 13 do indeed appear.

But our original data is not only "slow" by a factor of four, but also starts at $t=3$ rather than $t=0$.

t	P(t)
3	8
7	13

This means that $t=3$ is "behaving like zero" for the time values. We have that:

$$P(t) = 8 \cdot \left(\frac{13}{8}\right)^{\frac{t-3}{4}}$$

fits the data perfectly!

CHECK: Put in $t=3$ into this formula to see what I mean by "3 behaves like zero" for the time values. See we get $P(3)=8$. Also put $t=7$ to check that the value 13 does indeed appear.

EXERCISE: Show that $P(t) = 8 \cdot \left(\frac{13}{8}\right)^{\frac{t-3}{4}}$ agrees with our previous answer of $P(t) = 8^{\frac{7}{4}} \cdot 13^{-\frac{3}{4}} \cdot \left(\frac{13}{8}\right)^{\frac{t}{4}}$.

EXAMPLE: Find an exponential function that fits the data:

x	f(x)
5	30
12	40

Answer: This function has "initial value" 30, but with $x = 5$ behaving as zero. The data grows by a factor of $\frac{40}{30} = \frac{4}{3}$, but over a period of seven x-values. (It is seven times as slow as "normal"). We must have:

$$f(x) = 30\left(\frac{4}{3}\right)^{\frac{x-5}{7}}$$

CHECK: Put in $x = 5$ and $x = 12$ to see that we are correct!

EXAMPLE: Find an exponential function that fits the decreasing data:

x	f(x)
9	18
20	5

Answer: This function has "initial value" 18, but with $x = 9$ as the new zero. It "grows" by a factor of $\frac{5}{18}$, but over a period of eleven x-values. We must have:

$$f(x) = 18\left(\frac{5}{18}\right)^{\frac{x-9}{11}}$$

CHECK: Put in $x = 9$ and $x = 20$ to see we are right!

PRACTICE:
A silly pretend real-world question

A scientist is studying the growth rate of yerbits. She has two data points for a population model. (Why bother to have more than two data points?)

t	P(t)
2	5
6	11

Here time t is in weeks, and population is measured in counts of a thousand.

Having never studied yerbits before, the scientist is not sure if population is growing linearly or exponentially.

a) Fit a linear function to this data. What does this model predict for the population of yerbits in week 10?

b) Fit an exponential function to this data. What does this model predict for the population of yerbits in week 10?

c) The scientist waits until week 10 and counts 22.3 thousand yerbits. Which model seems most reasonable?

USING LOGARITHMS FOR EXPONENTIAL FITS

If one suspects data fits an equation $y = a \cdot b^x$, then one would expect $\log y = x \log b + \log a$ to hold, that is, for x and $\log y$ to have a linear relationship with slope $\log b$ and intercept $\log a$. We could work with x and $\log y$ directly and look a linear fit.

[Did you ever plot data on log-graph paper and look for linear patterns?]

EXAMPLE: Find an exponential fit to the data using a linear fit on logarithmic values.

x	y
2	5.88
5	16.13

Answer: Let's add of column of $\log y$ values.

x	y	log y
2	5.88	0.769
5	16.13	1.207

A linear equation that fits x and $\log y$ is:

$$\frac{\log y - 0.769}{x-2} = \frac{1.207-0.769}{5-2}$$

That is,

$$\log y = 0.146x + 0.477$$

giving:

$$y = 10^{0.477} \cdot \left(10^{0.146}\right)^x \approx 3 \cdot (1.4)^x .$$

[How does this answer to compare to the answer $5.88\left(\frac{16.13}{5.88}\right)^{\frac{x-2}{3}}$?]

CHALLENGE EXERCISE: Find a function of the form $y = ax^b$ that fits the data

x	y
2	1.6
4	51.2

SOME MORE PRACTICE PROBLEMS:

All this latter material comes from *THINKING MATHEMATICS! Vol 4: Functions and their Graphs* available at www.lulu.com.

Question 1: For each of these two data sets:

x	14	67
y	772	62

x	88772	7777263
y	43	43

find the equation of a straight line $y = mx + b$ that passes through the two data points and find the equation of an exponential function $y = ab^x$ that passes through the two data points.
(The second data set is curious!)

Question 2: Find an exponential function that fits this data. (??)

x	f(x)
1	23
9	0

Question 3: Consider the data:

x	y
1	4.2
6	1.8

a) Write down an exponential function that fits this data.

b) Draw a table of x-values and of $\log y$ values and find a linear function that fits this new table. Use the linear formula to again find an exponential function that fits the original data. Do you have the same function as the one you obtained in a) ?

c) If possible, find a function of the form $y = ax^b$ that fits this data.

d) If possible, find a function of the form $y = x^a + b$ that fits this data

Question 4:

a) Find a quadratic that passes through the points $(4,7)$, $(5,5)$ and $(10,-5)$ and simplify your answer. What do you notice? Explain.

b) Find a quadratic that passes through the points $(5,6)$, $(5,8)$ and $(7,5)$.
What do you notice?

c) Find a quadratic of the form $x = ay^2 + by + c$ that passes through the points $(5,6)$, $(5,8)$ and $(7,5)$. Sketch it!

MORE DATA FITTING:

There are two more appearances of data-fitting in the standard curriculum. Each of these will be discussed in later curriculum letters. But for now, here are some places to learn about them.

1. Periodic Data
If you suspect your data is cyclic - following some smooth "up and down" pattern over a fixed repeating period - then, typically, in a pre-calculus course, students will be asked to find a trigonometric function that fits the data. This work is also discussed in *THINKING MATHEMATICS! Vol 4* available at www.lulu.com.

2. Scatter Plots and Linear Regression
"Lines of best fit" and correlation coefficients measuring the "goodness" of fit, are discussed in statistics courses. These ideas are developed and explained in *THINKING MATHEMATICS! Vol 8: Beginning Probability and Statistics*. The short video www.jamestanton.com/?p=1175 also delivers the basics on this.

tanton.math@gmail.com