Solutions to Problems 1.2.24 and 1.2.27

Will Johnson

September 6, 2014

Since people asked me about these problems, I said I'd write up some solutions.

(27) A general power function looks like $f(x) = cx^b$. So for part (a), we want to find constants c and b such that $N \approx c \cdot A^b$ for each data point (A, N). Since we're not going to use a graphing calculator, a good idea is to just find a power law which goes through the first and last data points. So, we want

$$5 = c \cdot 4^b$$

$$75 = c \cdot 44218^b$$

Dividing these two equations, we get

$$\frac{75}{5} = \frac{c \cdot 44218^b}{c \cdot 4^b} = \frac{44218^b}{4^b} = \left(\frac{44218}{4}\right)^b$$

Simplifying the fractions, we want

$$15 = (11054.4)^b$$

Taking logs of both sides, we get

$$\ln 15 = \ln(11054.4^b) = b \ln 11054.4$$

So

$$b = \frac{\ln 15}{\ln 11054.4} = 0.29.$$

Finally we can solve for c:

$$5 = c \cdot 4^{0.29} = 1.49c,$$

so $c = 5/1.49 = 3.34$. So the final relation between A and N is

$$N = 3.34 \cdot A^{0.29}.$$

For part (b), we plug in 291 into the formula, getting

$$N = 3.34 \cdot (291)^{0.29} = 17.3,$$

So we expect there to be about 17 species.

(24) *Using an exponential function:*

A general exponential function looks like $f(x) = e^{ax+b}$. Since we aren't using a graphing calculator, we'll just find an exponential law which goes through the first and last data points. So, we want

$$30.4 = e^{1955a+b}$$

and

$$10.5 = e^{2000a+b}$$

to hold. Taking logarithms of both sides, we want

$$1955a + b = \ln 30.4 = 3.41$$

$$2000a + b = \ln 10.5 = 2.35$$

Subtracting these two equations, we get

$$45a = 2.35 - 3.41 = -1.06$$

$$a = -1.06/45 = -0.0236$$

Solving for b,

$$b = 2.35 + 2000 \cdot 0.0236 = 49.60$$

So we get the equation

$$y = e^{49.60-0.0236x}$$

Now, plugging in 1988, we get

$$e^{49.60-1988\cdot0.0236} = 13.9$$

So we predict that 13.9 percent of the population lived in rural areas in 1988.

Likewise, plugging in 2002, we get

$$e^{49.60-2002\cdot0.0236} = 10.0$$

So we predict that 10.0 percent of the population of Argentina lived in rural areas in 2002.

Using a linear function We want a line that goes through the points $(1955, 30.4)$ and $(2000, 10.5)$. The slope would be

$$m = \frac{10.5 - 30.4}{2000 - 1955} = -0.4422222$$

So the equation would be

$$y - 30.4 = -0.442222(x - 1955)$$

$$y = 30.4 - 0.442222x + 1955 \cdot 0.442 = 894.944 - 0.4422222x$$

Plugging in 1988 gives

$$894.944 - 0.442222 \cdot 1988 = 15.8,$$

so we'd predict that 15.8 percent of the population was rural in 1988.

Likewise, plugging in 2002 gives

$$894.944 - 0.442222 \cdot 2002 = 9.6,$$

so we'd predict that 9.6 percent of the population was rural in 2002.

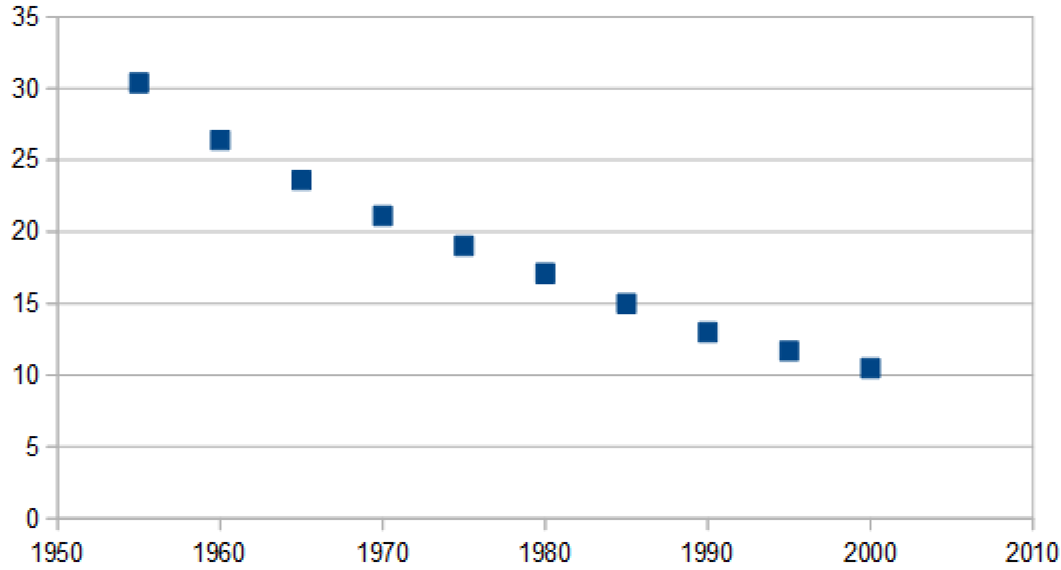
1 Comments on these problems

This kind of problem will not occur on the exams. As a general rule of thumb, any problem requiring a calculator will not appear on the exams. A problem similar to 24 will appear on the September 11th homework, but other than that, this kind of "regression" problem won't make any further appearances in this class.

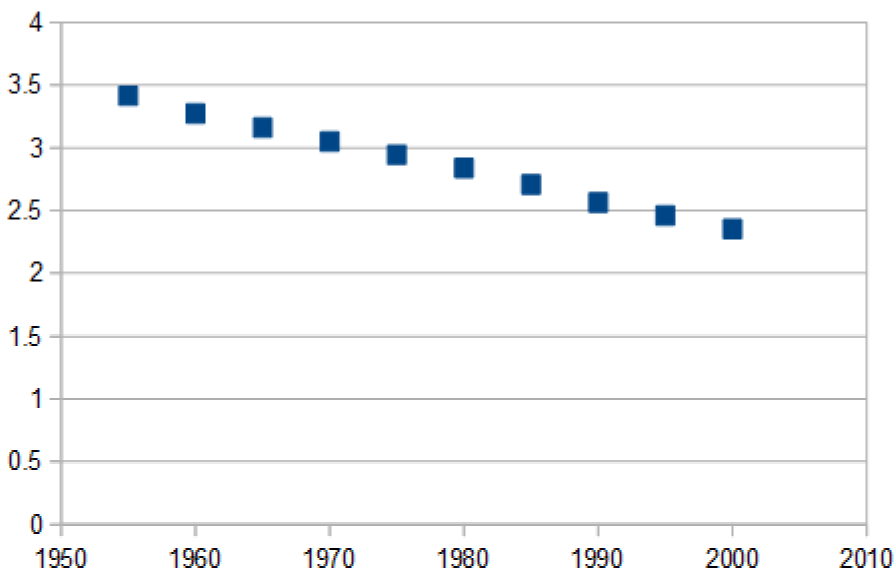
How do we know what kind of model to use for problem 24? An exponential model is probably better than a linear model, but how do you see this? One very crude reason is that this is a problem about population, and things involving population often tend out to be exponential rather than linear. The linear model also predicted that 15.8 percent of the population in 1988 was rural, which is a little suspicious: from the table we know that the rural percentage was decreasing steadily, and in 1985 it was 15.0, while in 1990 it was 13.0. If it was decreasing the whole time, the rural percentage in 1988 ought to be between 15.0 and 13.0. But our prediction of 15.8 is not in this range.

Also, it was strongly implied on the in-class worksheet that the dependent and independent variables in problem 24 were not linearly related, but that the *natural logarithm* of the dependent variable was linearly related to the independent variable.

Letting y denote the rural percentage, and x denote the year, here is a graph of y vs x:



. . . and here is a graph of $\ln(y)$ vs x:



The first graph appears to be slightly curved, while the second graph looks a little more linear. On the worksheet from lecture, the business of taking successive differences was intended to convey this same idea, without drawing an actual graph.

So it seems like rather than a model like

$$y = mx + b$$

we instead want

$$\ln(y) = mx + b,$$

or equivalently,

$$y = e^{mx+b}.$$

But this is just the same as an exponential model.