CS130/230 Lecture 9
Nonlinear Regression

Tuesday, March 2, 2004

Nonlinear Regression

Often, relationships are nonlinear and we need a different type of graph to fit the data. Excel provides us with different types of nonlinear functions that we can use to fit data. These functions include polynomial, exponential, logarithmic and power.

Problem 1

Let us consider the following data which represents the number of deaths, N, from AIDS in the United States from 1981 to 1996, where t denotes the number of years after 1980.

t	N
1	159
2	622
3	2130
4	5635
5	12607
6	24717
7	41129
8	62248
9	90039
10	121577
11	158193
12	199287
13	243923
14	292586
15	340957
16	375904

After typing the data into an Excel spreadsheet, try to fit different types of nonlinear functions to the data. Which works the best? How do we know?

Can we predict what the number of deaths from AIDS would be in 1997?

Solving Exponential and Logarithmic Equations

Recall that to solve an equation of the form $y = ae^{bx}$ for x (where a and b are just constants), you first divide by a to obtain $y/a = e^{bx}$. Now, you must take the natural logarithm of each side to obtain $\ln(y/a) = bx$. Dividing by b yields $x = (1/b)\ln(y/a)$.

Recall that to solve an equation of the form y = a ln(bx) for x (where a and b are just constants), you again divide by a to obtain y/a = ln(bx). Now, you must exponentiate each side to obtain e y/a = bx. Dividing by b yields x = (1/b)e y/a .

Problem 2

The following data is from an actual study that considered how memory decreases with time. The subjects each read a list of 20 words slowly aloud, and later, at different time intervals, were shown a list of 40 words containing the 20 words that he or she had read. The percentage, P, of words recognized was recorded as a function of the time t elapsed in minutes. The table below shows the averages for 5 different subjects. Your task is to; using the regression capabilities of EXCEL, determine the best-fit model for this data AND to use this model to predict the time at which the subjects will recognize 15% of the words.

T,min	5	15	30	60	120	240	480	720	2880	5760
P%	73.0	61.7	58.3	55.7	50.3	46.7	38.3	29.0	24.0	18.7

Midterm

On Thursday we will have our first midterm, which will include all the material that we have covered in class up to this point. Part of the exam is written and part will be hands on. The first part of the exam is the written part. Once you have handed in the written portion, I will give you the hands-on portion that will be solved using Excel and placed in the CS130 Drop Box. Remember, I will not answer any computer questions, so make sure you can do everything we have discussed in class. I will only clarify questions on the exam if you do not understand something.

The written portion of the exam can consist of multiple choice, true/false, and short answer questions. The hands on portion will be two or three problems of the type we have done in class.

Review Questions

Problem 1

Molly begins her IRA at age 18 and puts $1,000 per year into her IRA at the beginning of each year for eight years. At the end of eight years, she just lets her money accumulate interest until age 65 without putting any more of her money in the IRA. Polly on the other hand begins her IRA at age 31 and puts $1,000 per year into the IRA at the beginning of each year until age 65. If both IRAs pay 10% interest compounded yearly, how much has each person accumulated by age 65 (include age 65 in your accumulation)? Work this out using the repetition method (table method) only. Make this solution so that the user can input the amount invested per year and the interest rate for each person. Place a border around the cell containing Molly's final total and place a border around the cell containing Polly's final total. No other borders are to be in this worksheet. Design this

worksheet such that the last row for Molly is age 65 and the last row for Polly is also age 65. That way I can directly compare the ending values. Further, split your screen so that I can input information in the top half of the screen and look at the age 65 results in the bottom half of the screen. Save your file this way.

Problem 2

You have been hired to work for the month of January and have been given the choice of one of two ways to be paid. Method 1: You can receive $500 per day for 20 days. Method 2: You can receive 1 penny the first day, 2 pennies the second, 4 pennies the third, 8 pennies the fourth, and so on. In a worksheet, show side-by-side comparisons of the money made each day for each of these methods. As a final figure at the bottom of each column, show how many dollars were made for each method. Create a graph (properly labelled) that shows the money made each day of both methods.

Problem 3

Using the file candy.xls, calculate the average calories for M&M/Mars candy bars using IF or SUMIF. Remember, you first need to calculate the sum of all calories of M&M/Mars candy bars and then calculate the number of M&M/Mars candy bars. To get full credit for the problem, you must calculate the number of M&M/Mars candy bars (you can do this with IF or COUNTIF). Highlight the average value you calculated in your spreadsheet. (Make sure you put it in a spot that's easily seen on your spreadsheet!)