Name: Math 130H Lab 4

Excel: Non-linear Curve Fitting and Transformations

In this lab, we continue to analyze paired data. The idea is that you take a sample, collect a pair of numbers from each member of your sample, make a scatter plot of the data, fit the data to a curve and use the equation of the curve to predict the value of *y* when you are given the value of *x*. In the last lab, the graph of the paired data looked like a line (or was modeled as a line) and so you used excel to find the best fitting line to the data. In this lab, you will fit the data to a curve that is NOT a line. There are 2 ways to do this. One way is to have excel fit the data to the desired curve directly, but this is too easy. So instead, in this lab we will dig deeper and use transformations to transform the given curve to a line, get excel to fit the transformed data to a line, then reverse the process to get the equation of the curve. Then you can finally use the equation of the curve to make a prediction.

Curve fitting (exponential curve)

Fitting paired data to a line means that you are trying to find the best value of the numbers *M*, and *B* that make the graph of the line $Y = MX + B$ as close as possible to the data points.

Fitting paired data to an exponential curve means that you are trying to find the best value of the numbers *a*, and *b* that make the graph of the line $y = a \cdot b^x$ as close as possible to the data points. Excel has uses an algorithm to fit data to straight lines, so if you want to fit data to a curve that is not a straight line, we need to transform the curve to a line. Let's go through the derivation of how to do this:

Let *x* and *y* stand for the *x*- and *y*-coordinated of the original paired data that you collected and let *X* and *Y* denote the coordinates of the data after we transform the data. Once the data is transformed, the graph will be a line and let *M* denote the slope of the best fitting line of the transformed data and let *B* denote the *y*-intercept of the best fitting line of the transformed data. In short, we will use lower case letters for the data that we start with and we will let capital letters denote the transformed data.

Our first goal is to derive the formulas for the best values of *a* and *b* once excel gives us the best values of *M* and *B*.

Your Turn 1: Take the natural log (ln) of both sides of the equation $y = a \cdot b^x$ and use properties of logs to expand the right side of the equation as much as possible.

Your Turn 2: Take the answer you got from part 1 and replace $\ln y$ with Y and x with X.

Your Turn 3: Now in terms of X and Y, the equation now looks like a line. That means that the number that is being multiplied by X is the slope. Set the term that is multiplied by X equal to M and solve the equation for b.

Your Turn 4: In your answer from part 2, the term on the right side of the equation that is not being multiplied by X is the y-intercept. Set this term equal to B and solve for a.

Example: Here is some data that we are going to fit to an exponential curve…

x	1.4	2.1	2.6	3.2	3.3	4.9	7.2	8.4
y	14.73	12.06	10.49	8.85	8.67	5.53	2.89	2.08

Type this data in excel exactly the way it appears above with the *x* in cell A1 and the 2.08 in cell I2. Graph the data and print it out and turn it in with this lab. Now type "X" in cell A3 and "Y" in cell A4. This will be our transformed data. The transformation we had above was $X = x$ and $Y = \ln(y)$.

To get the *X* row…
Since $X = x$, this means that we want our *X* row to be exactly the same as our *x* row. To accomplish this, click on cell B3 and type "=B1" then press enter (don't type the ""). This will copy the contents of cell B1 in cell B3. Then to do this for all other cells in the third row, click on cell B3, put your mouse over the bottom right corner of the cell until the mouse symbol turns into a black plus sign, then hold down the (left) mouse button and drag to the right until you get to cell I3, then release. This should have copied all of the numbers in the first row over again in the 3rd row.

To get the *Y* row…
Since $Y = \ln(y)$, this means that we want the *Y* row to be the natural log of the numbers in the *y* row. To accomplish this, click on cell B4 and type "=ln(B2)" then press enter (don't type the ""). This will put the natural log of the number in B2 in cell B4. Then to do this for all other cells in the 4th row, click on cell B4, put your mouse over the bottom right corner of the cell until the mouse symbol turns into a black plus sign, then hold down the (left) mouse button and drag to the right until you get to cell I4, then release. This should have put the natural log of the numbers in the 2nd row in the 4th row.

So now there are 2 sets of data, one is the original data (1st 2 rows) that we think can be modeled by an exponential equation, and the other is transformed data (last 2 rows) that should now form a near straight line.

Your Turn 5: Make a scatter plot of the last 2 rows of data. Does the graph look like a straight line? Use the ideas from the lab 3 to fit the new data to a line and have excel give you the equation of the best fitting line to the new data. Write the equation below.

What we want is the equation of the best fitting exponential equation to the original data.

<u>Your Turn 6</u>: Take the slope of the line you obtained from step 5 and plug it into the equation you got for *b* back in step 3. Write the answer for *b* below.

<u>Your Turn 7</u>: Take the *y*-intercept of the line you obtained from step 5 and plug it into the equation you got for *a* back in step 4. Write the answer for *a* below.

<u>Your Turn 8</u>: Now take the answers you got for *a* and *b* from the last 2 steps and plug them into the formula $y = a \cdot b^x$ and write the answer below. This is the equation of the best fitting exponential curve to the original data.

<u>Your Turn 9</u>: Once you have the best fitting exponential curve to the original data you can use it to make predictions. Predict the value of *y* if $x = 6$.

Your Turn 10 (Sec. 4.5, Ex 4): The data in the table below represents the temperature of a pizza removed from an oven measured in 5-minute intervals. Newton's law of cooling models this type of situation with an exponential curve.

Time, x (minutes)	5	10	15	20	25	30	35
Temperature, y (Fahrenheit)	327	255	204	167	140	121	106

a) Use the same transformation of the data to get a new set of data and draw scatter plots for both sets of data.
b) Use excel to fit the transformed data to a line.
c) Use the formulas you derived earlier to find the equation of the best fitting exponential curve to your original data.
d) Use the exponential equation to determine the temperature of the pizza when it was removed from the oven.
e) Use the exponential equation to predict the temperature of the pizza 40 minutes after it was removed from the oven.

Curve fitting (power function)

The process we went through above is what you do when you think your paired data can best be modeled by an exponential equation $y = a \cdot b^x$. What if you think that the data needs to be modeled by some other curve? The idea is that you need to figure out a way to transform your data to a straight line. In this next part of the lab we are going to model our paired data with a power function $y = a \cdot x^r$ and get the values of a and r that make the power function's graph as close as possible to our data points.

To transform this curve to a straight line, we will use logs again. Let's have x and y stand for our original data (as before) and let X and Y denote the transformed data.

Derivation

Your Turn 11: Take the natural log (ln) of both sides of the equation $y = a \cdot x^r$ and use properties of logs to expand the right side of the equation as much as possible.

Your Turn 12: Take the answer you got from part 11 and replace $\ln y$ with Y and $\ln x$ with X.

Your Turn 13: Now in terms of X and Y, the equation now looks like a line. That means that the number being multiplied by X is the slope. Set the term in front of X equal to M and solve the equation for r. (If you think this step is too easy, it is!)

Your Turn 14: In your answer from part 12, the term on the right side of the equation that is not being multiplied by X is the y-intercept. Set this term equal to B and solve for a.

So whenever you have data that you feel can be modeled by a power function $y = a \cdot x^r$,

1) Transform the data to new data by taking the natural log (ln) of the original x-coordinates to get the new x-coordinates
2) Take the natural log (ln) of the original y-coordinates to get the new y-coordinates
3) Fit the transformed data to a line using excel
4) Take the slope and y-intercept you obtained from this line and plug them into your formulas in steps 13 and 14 to find a and r
5) Plug the values of a and r into $y = a \cdot x^r$ to get the best fitting power function to the original data
6) Use this equation to make predictions for y given the value of x

<u>Your Turn 15 (Sec. 4.5, Ex 5)</u>: Scott drops a ball from various heights and records the time, *x*, that it takes for the ball to hit the ground, using a motion detector. He obtains the data below. Newton's 2nd law in Physics models the relationship between the time *x* that it took the ball to reach the ground and the height y that the ball was dropped as a power function $y = a \cdot x^r$.

Time, x (seconds)	1.528	2.015	3.852	4.154	4.625
Distance, y (feet)	11.46	19.99	72.41	84.45	104.23

a) Transform the original data by taking the natural log (ln) of each row to obtain your transformed data, then make scatter plots of each data set.
b) Use excel to fit the transformed data to a line.
c) Use the formulas you derived earlier to find the equation of the best fitting power function to your original data.
d) Use the power function equation to predict the distance a ball will have to fall to take 4.2 seconds to hit the ground.