Fitting Functions to Data

Introduction

In a previous handout we discussed how easy it was to use Excel's extended graphical methods to fit a line to data, however, there may be circumstances under which we may wish to fit an arbitrary function to data. This document describes how we may used Excel's non-graphical function fitting routines to fit a line as well as some other common functions to data. Specifically, we will explain how to use the Excel spreadsheet "Function Fits" to fit four types of curves. To do this, we will need two columns of data: one will contain the values which would be plotted on the horizontal ("x") axis of a graph, and the other will contain the measured values which would be plotted on the vertical ("y") axis. Once you learn how to use it, the spreadsheet is quite handy, as it automates the fitting procedure. It is also an interesting example of how a spreadsheet may be set up to perform complicated tasks in a repeated way.

The handout is organized into two parts. The first explains the logic of the fitting procedure. The second explains how to use the spreadsheet.

Theory

When we fit a function to data, we're interested in finding the equation for some function that passes as closely as possible through the middle of our data. In addition to fitting as closely as possible, we would like this equation to be simple so we can use it summarize our data. First, we will discuss how to use the Excel function LINEST to fit a linear function to data. Next, we will discuss how to use what we have learned about fitting linear functions to data to fit three other functions.

Linear Function

As discussed in a previous handout, the formula for a straight line is a simple equation having two parameters, the slope and the *y*-intercept. Before we can fit a line we typically need two columns of data: one for the values on the horizontal ("*x*") axis and the other for the values on the vertical ("*y*") axis. Because we want a line that describes how the y-values change when we vary the *x*-values, we will need both those columns in estimating the best-fitting line through the data. Accordingly, the equation for a straight line contains both a "*y*" variable (which in our case represents the column of *y*-values) and an "*x*" variable (for the column of x-values): $y=m*x+b$.

In our terminology, *y* is the measured value for some dependent variable (e.g. perceived length, threshold intensity, reaction time, etc.) for a particular independent variable *x* (e.g. line length, masker intensity, number of items, etc.). The constant *m* is a number that determines the slope of the line that relates the *dependent variables* to the *independent variables.* The constant *b* is another number and determines where the line crosses the y-axis. What we want to do is use Excel to find the values for *m* and *b* that make the best fitting line. We often call *m* and *b* the parameters of the straight line that best fits the data in the "*y*" column.

To obtain the value of *m* of the line, type:

=LINEST(y-values,x-values)

You should replace **y-values** and **x-values** with the addresses of those two columns of data. For instance, if your column of *y*-values is found in cells B36 through B50, you would replace **y-values** with **B36:B50**. You could also enter the addresses by going to the function bar and positioning the cursor at the appropriate place in the function, then clicking on the first cell in the y-column and dragging down to the last cell. Do the same for the *x*-values. When you press the enter key, the value of the slope will appear in the cell.

To obtain the value of *b*, type:

=INDEX(LINEST(y-values,x-values),2)

To find the slope of a line that goes through zero on the *y*-axis (i.e., to force the *y*-intercept to be zero), type:

=LINEST(y-values,x-values,FALSE)

Fitting Other Functions to Data

The "Function Fits" spreadsheet fits three other two-parameter functions to data, using a procedure that's very similar conceptually to the one we use to find the best-fitting straight line.

For each type of equation, the spreadsheet again uses the Excel function LINEST to find the best-fitting *straight* line through the data; the important difference is that it first performs a logarithmic transform of the x-axis values, the y-axis values, or both. Fitting a straight line through transformed data is equivalent to fitting logarithmic, exponential, and power functions (which are typically curved) through the untransformed data. Thus, once we have made the appropriate transformations, we can express the equations for the different types of functions in the form of the equation for a straight line. This is a good thing to do because it's easy to fit a straight line through data in Excel using the LINEST function.

The subsections below review this theory in more detail, particularly for the power function.

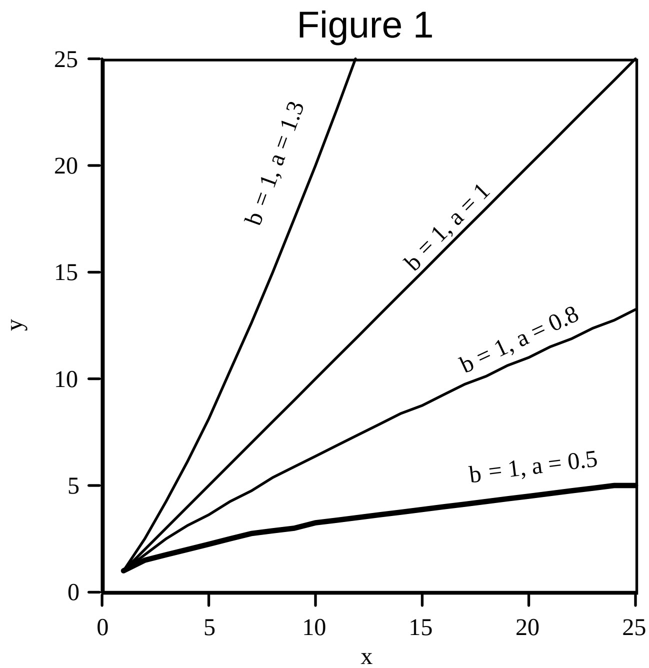
Power Function

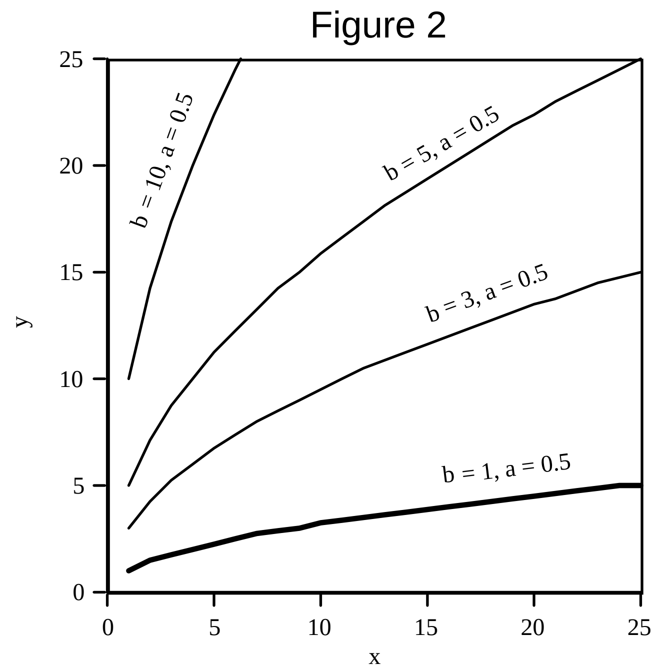
The formula for a power function is $y=b\,x^a$, where y denotes a measurement obtained for an independent variable x, and a and b are parameters that determine the particular form of the power function. Magnitude estimation data are often fit by a power function. In our Müller-Lyer experiment, for example, y would stand for the estimated length and **x** for the physical length of the lines.

Note that in the special case $a = b = 1$ the power function reduces to the equation $y = x$, which means that the subject's estimates are veridical.

Changing the parameters (a and b) affects the exact shape of the power function. Roughly speaking, the parameter b affects how steeply the function rises as it passes through its midpoint, while the parameter a affects where the midpoint occurs and whether the curve is concave or convex.

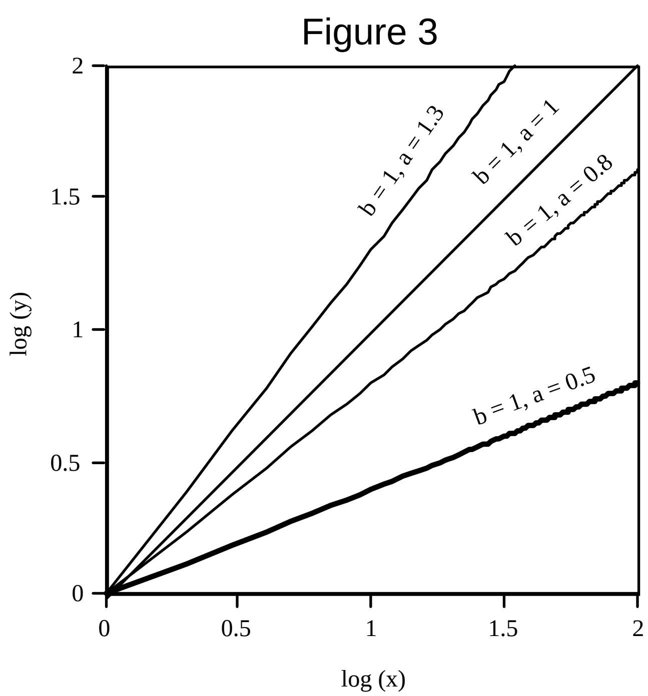
Figures 1 and 2 illustrate the effect of changing each of the parameters on the shape of the power function.

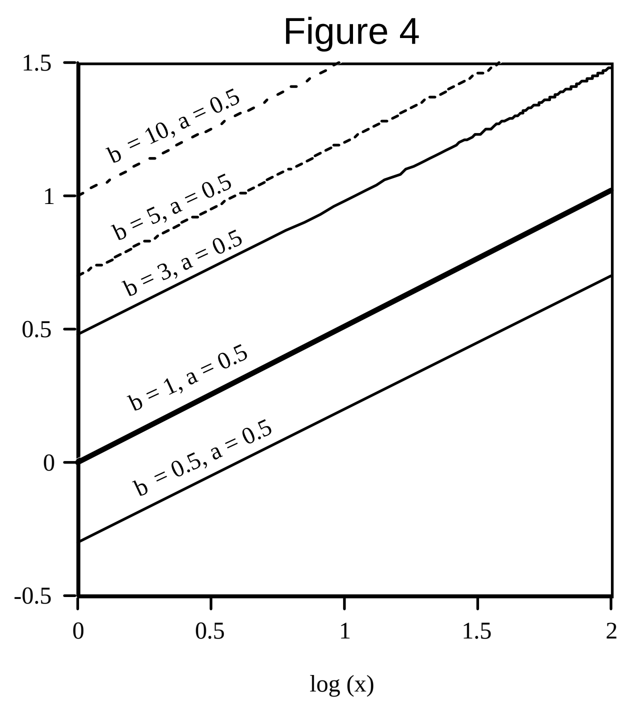




Converting to log coordinates. As you can see from Figures 1 and 2, a power function may have many different shapes. It is difficult to decide just by staring at a graph whether the data follow a power function or not. But if we replot the data in log coordinates, things get simpler.

Plotting in log coordinates means that we plot the logarithm of *y* versus the logarithm of *x*. We will use log base 10 in this course, so we plot log(*y*) versus log(*x*).





In Figures 3 and 4, the data from Figures 1 and 2 are replotted in log coordinates. Notice that now all of the different power functions fall along a straight line. This is not a coincidence. If

$$y = bx^a$$

Then

$$\begin{aligned}\log(y) &= \log(bx^a) \\ &= \log(x^a) + \log(b) \\ &= a\log(x) + \log(b).\end{aligned}$$

The last form of the equation is the equation of a straight line, with slope *a* and intercept log(*b*). This derivation depends on standard facts about logarithms which you probably already know. But they are listed in a separate handout.

To decide whether a measured function follows a power function, you can thus plot the data in log coordinates and see if it lies along a straight line. If it does, the slope and intercept of the line may be used to determine the parameters of the power function. The *slope* of the straight line through the log data is the exponent *a* in the best-fitting power function curve through the original (untransformed) data; to find the value of the other constant in the power function, *b*, we raise 10 to the value of the *intercept* of the straight line.

Logarithmic and Exponential functions

Two other simple function forms are the logarithmic and exponential functions. They are both two-parameter functions, and like the power function, they are typically curved. In some cases, either one of these functions might provide a better fit to data than either a straight line or a power function would. Another property they share with the power function is that they can be expressed in the form of a linear equation if the axes are appropriately transformed.

To obtain the formula for a *logarithmic function*, all we have to do is convert the *x*-values to log(*x*). What we get is an equation with slope *a* and intercept *b*:

$$y = a\log(x) + b.$$

The formula for an *exponential function* is

$$y = b * 10^{ax}.$$

Taking the logarithm of both sides, we get

$$\begin{aligned}\log(y) &= \log(b * 10^{ax}) \\ &= \log(10^{ax}) + \log(b) \\ &= ax + \log(b).\end{aligned}$$

This leaves us with slope *a* and intercept log(*b*).

Summary

Power, logarithmic, and exponential functions, then, will all plot as straight lines when we transform the x-axis values, the y-axis values, or both, to logarithms. This is a consequence of the fact that when we perform those transformations, we can express the different types of functions as formulas for straight lines.

And because this is true, we can find the best-fitting power, logarithmic and exponential functions by transforming the data appropriately and finding the best-fitting straight line through the transformed data.

Using the Spreadsheet

Once you have pasted your data into the X and Y columns, there are several things you'll need to alter in the spreadsheet to make it work properly. First, make sure that the last row of your data is the last line on the sheet—in most cases it will NOT be. You will probably have to delete or add rows at the bottom.

Deleting Rows. If your data set is shorter than the existing data in the "Function Fits" spreadsheet, you will need to delete all the rows below the last line of your data. First, select the rows, then choose Clear... under the Edit menu. Be sure to include the all the cells in the rows you're clearing (that is, clear columns A to N, starting at the row just under the last line of your data).

Adding Rows. If your data set is longer than the existing data, add more rows by highlight an area that includes: the last row of the spreadsheet containing data in columns C through N, *and* the empty area adjacent to your own data. Now choose Fill-Fill Down from the Edit menu.

Editing Array Functions. Starting in cell I20 you will see five columns of data used by the spreadsheet to compute the sum of squared error (SSE), which we can use as a relative measure of how well each type of line fits the data.

These are the parameters from the four fits along with the SSE for each.
Divide the SSE by the number of rows to get the MSE if you prefer.

A	B	a	b	SSE
-13.797041	3166.56999	-13.797041	3166.56999	552911.758
-1916.9027	5495.45743	-1916.9027	5495.45743	252933.563
-0.0022581	3.4769269	-0.0022581	2998.65777	536420.865
-0.2768218	3.79451654	-0.2768218	6230.40876	174911.013

If you click on the first cell under the column labeled "A" here, you will see a function appear in the function bar that looks something like this: {=LINEST(B30:B169,A30:A169)}. If you deleted OR added rows, you'll need to change the references to the last row of your data. These references appear after the colons in this function; in this example, you'll need to change B169 and A169 to whatever the addresses are that contain the last row of your data.

Starting at cell I20, we'll work our way down to cell I23. First, look at the last row of your data, and make a note of the corresponding row number (at the extreme left-hand side of the Excel window). Now, find cell I20 and click in it. You'll notice the curly brackets **{}** that surround the formula in the formula bar disappear. Change the *number part only* of the cell addresses that appear after the colons. So, if B169 and A169 come after the colon, change the 169's to whatever the row number was of the last row of your data.

Now, instead of hitting the enter button as we usually would to complete the function, *hold down the Control key and the Shift key while you hit the enter button*. If you don't, you'll see an error message that says: "Cannot change part of an array". If you see this message, just change the numbers again and be sure to hold down the Control and Shift keys as you hit the enter button. Repeat this procedure for cells I21 to I23. You'll note that the letter part of the cell addresses that come after the colons will change, but as long as you only change the number part, you'll be OK.

Finally: Use a similar procedure to change the range of rows used in the SSE column; again, change the number part of the cell address that comes after the colon. This time, you can simply hit the enter button without holding down the Control and Shift key.

After making these changes, the spreadsheet will calculate the best fitting lines through your data and the sum of squared error for each type of line.

A More In-Depth Look at the "Function Fits" Spreadsheet

Starting in cell A20, you'll see some text that outlines the equations used to fit lines to the data contained in columns X and Y:

Type	Raw Equation	Effective Equation	What are a and b
Linear	Y = A*X + B	Y = a*X + b	(a = A, b = B)
Logarithmic	Y = A*LOG(X) + B	Y = a*LOG(X) + b	(a = A, b = B)
Exponential	LOG(Y) = A*X + B	Y = b*10^(a*X)	(a = A, b = 10^B)
Power	LOG(Y) = A*LOG(X) + B	Y = b*X^a	(a = A, b = 10^B)

First notice the Raw Equation for a linear (straight line) function:

$$Y = A * X + B.$$

Y is the data you measured in your experiment, and *X* is the physical variable you manipulated. The slope here is *A*, and the *y*-intercept is *B*.

If you look at the Raw Equation for a power function in the spreadsheet, you can probably see that the formula still looks a lot like the formula for a straight line:

$$Y = A * LOG(X) + B.$$

We have simply replaced *X* with the logarithm of *X*, LOG(*X*). Otherwise, there is still a slope, *A*, and a *y*-intercept, *B*. Similarly, the other Raw Equations replace either *X* or *Y*, or both, with their logarithms, without otherwise changing the form of the equation.

The Effective Equations show another way of expressing the Raw Equations, such that *Y* is always expressed in untransformed (i.e., linear) terms. To express the equations this way, we need to transform *A* or *B* in some of the equations; we denote this by converting *A* to *a* and *B* to *b*. The column labeled "What are a and b" shows us how to transform *A* and *B*, if we need to, so that the equations work out right.

Estimating the Goodness of Fit

There are two ways we can get a quick idea of how the different types of functions stack up in terms of how well they fit the data:

We have already hinted at the first way—simply look at a plot of your data along with the calculated best-fitting curve for each type of function (see below). The degree to which each curve passes through the middle of the data is an indication of how well that type of function fits the data. If you're lucky, you might be able to tell immediately that one type of function passes more nearly through the center of the data than some or all of the other functions do.

Another, more numerical way of comparing the relative goodness of fits is by comparing the sum-of-squared-errors (SSE) obtained for each type of function (look in cells M20 to M23 in the Function Fits spreadsheet). This number provides a measure of how far, on average, the y-value we predict by plugging the *x*-values, *a* and *b* into the formula for that type of function deviates from the actual y-value. There are some other things that go into the calculation of this statistic which need not concern us right now, but the important thing is that the lower the SSE, the better the fit.

Plotting a Line Using the Estimated Parameters

Just as we used the parameters of the best fitting straight line to calculate coordinates of points along the line, we can use a similar procedure to plot the coordinates of power, logarithmic and exponential functions through data.

You should first decide whether you want to plot your data in linear or log coordinates. As I stated before, the extent to which your data fall along a straight line when the axes are transformed appropriately is an indication of how well the function in question fits your data. This is an easy visual rule of thumb that you may want to take advantage of when making your figures. Another advantage of plotting your data in the transformed axes is that the slope and intercept parameters of the straight line may be easier to conceptualize and interpret than are the exponential parameters in the effective equations.

Alternatively, it is fairly easy to plot each type of function in linear coordinates; this makes it convenient to visually compare different fits by overlaying several different functions in the same graph.

Plotting the Data in Log Coordinates

First, you may want to open a new worksheet and copy column *X* into it. Now, look at the raw equations starting in cell B20 of the Function Fits spreadsheet. Find the raw equation that corresponds to the function you want to graph. If log(*X*) appears in the equation, you'll need to make a new column to contain the transformed values. To transform the data in column *X* to logarithms, click on the first cell of the column you want the logs to go into, and type: =log(*a13*) (replace *a13* with the address of the first cell of your column *X* data.) Now use Fill Down to convert the rest of the column to logs.

Next we want to calculate the y-coordinate of a point on the fitted line that results when we plug a particular value from the *X* or log(*X*) column into the equation for the appropriate type of function. To do this, we start with the raw equation of the function in question, and replace the constants in that expression with the parameters *A* and *B* calculated by the Function Fits spreadsheet.

For example: suppose you have a column of 40 values in the *X* column, smoothly increasing from 0 to 500. These are the values that will be plotted on the x-axis of a graph. To the right of that column, you have an empty column which will contain the y-values calculated from these *x*-axis values. Suppose that previously you had fit a straight line in log-log coordinates to your data, using the Function Fits spreadsheet, and calculated a slope *A* of .28 and a *y*-intercept log(*B*) of 3.79. (These would be found in cells I23 and J23.)

The raw equation for the power function fit is

$$\log(Y) = A\log(X) + B.$$

We'll want to replace *A* and *B* in this equation with the values returned by the Function Fits spreadsheet:

$$\log(Y) = (.28)\log(X) + 3.79.$$

In terms of our new Excel spreadsheet, log(*X*) will of course be the value in the log(*X*) column, and log(*Y*) will be the *y*-axis value we calculate based on that *x*-axis value. So, if the first log(*X*) value is in cell *b3*, to calculate the associated *y*-axis value, type the following formula in cell *c3*, type

=.28*b3 + 3.79

Then use Fill Down to calculate y-axis values for the rest of the log(*X*) column. Now, to plot the line, copy this column and the log(*X*) values into another spreadsheet and make a new line graph. Name the two columns "log(X)" and "log(Y)". When you choose log(X) as the x-axis variable and log(Y) as the y-axis variable, the line plotted should be straight.

Similarly, to plot any of the other functions, write a formula for the raw equation of the function you want to plot. Plug in the values of A and B calculated by the Function Fits spreadsheet (in columns I and J). Replace X or log(X) in the raw equation with the address of the first cell of the appropriate column in your data sheet. Then use Fill Down to calculate the rest of the values in the column, and copy this column and either column X or log(X) (whichever is appropriate) into another spreadsheet to graph.

Plotting the Data in Linear Coordinates

To plot a line in linear coordinates, we will refer to the Effective Equations instead of the Raw Equations from the Function Fits spreadsheet. The Effective Equation for the power function is:

$$Y = b * X^\wedge a.$$

Again, we'll want to replace a and b with the values returned by the Function Fits spreadsheet. This time instead of looking in columns I and J, look in columns K and L for the parameter values to plug into the effective equation. We might find, for example, that a and b for the power function are -.27 and 6230.41, respectively. So, plugging these values into the Effective Equation, we get

$$Y = 6230.41 * X^\wedge.27.$$

Translating this into an Excel formula, we would type

=6230.41**c3*^.27 (Replace *c3* with the address of the first cell of column *X*.)

Use Fill Down to calculate the rest of the values in the column, then copy this column and column X into another spreadsheet to graph it.

Use a similar method to plot any of the other functions: find the values of a and b for the function you want to plot (look in columns J and K). Write a formula for the effective equation for that function, plugging in a and b, and replace X with the address of the first cell of column X. Then Fill Down, and copy the column into another spreadsheet to graph.