Logarithm Plots

Purpose

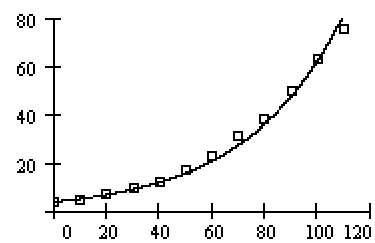
In this lab you will learn how to determine if data points can be fitted well by either a power function, $y = ct^k$, or by an exponential function, $y = ce^{kt}$, and you will learn how to choose the constants c and k.

Preview

The eighteenth and nineteenth century United States census data in the table below can be well represented by the function $P(t) = 4.19e^{0.027t}$, where t is the number of years since 1790.

Year	Population (in millions)
1790	3.9
1800	5.3
1810	7.2
1820	9.6
1830	12.9
1840	17.1
1850	23.2
1860	31.4
1870	39.8
1880	50.2
1890	62.9
1900	76.0

The plot below shows the function, $P(t) = 4.19e^{0.027t}$, superimposed on a plot of the data from the table.



You will learn how to test such data to determine if an exponential or power function fits the data well, and you will see how to determine the constants in the function.

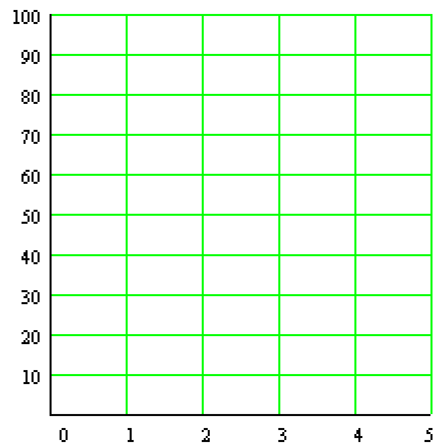
Overview

We will create some "data" points, and then we will pretend our data was gathered from an experiment so we can learn how to "discover" the original functional relationship between the variables. Once we have the technique and concept in hand, we can apply this knowledge to real data.

Part 1: Beginning of an Experiment

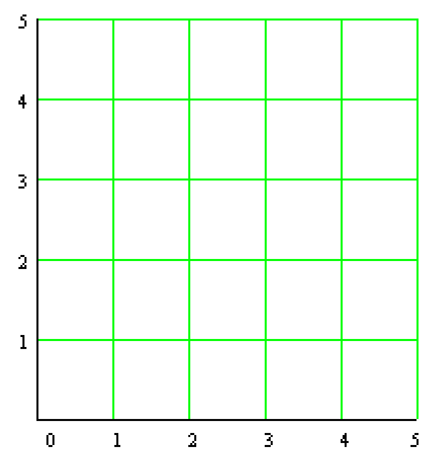
1. We will use the function, $y(t) = 3 \cdot 2^t$, to create some data points. Fill out the table below for the indicated values of t. After you've filled in the table, plot the points, (t, y), on the graph.

t	y
0	
1	
2	
3	
4	
5	



2. Use the values from the table in step 1 and your calculator to fill in the following table accurate to two decimal places. After you have filled in the table, plot the points, $(t, ln\, y)$, on the graph. Note the vertical scale, which is different from the scale used above.

t	$ln\, y$
0	
1	
2	
3	
4	
5	



An Explanation

Were you surprised by what you saw in the last graph? The type of graph that you just made is called a "semilog graph," because you used the logarithms of the second coordinates in your plot rather than the original y values. (The prefix "semi" refers to the fact that we did not take the logarithms of the first coordinate.) In this case we created our own "data" from a known function, so we could use that function to explain what happened in the second graph.

Here is the equation that expresses the relationship between y and t:

$$y = 3 \cdot 2^t$$

Take the (natural) logarithm of both sides of the equation to get a new equation:

$$ln\, y = ln(3 \cdot 2^t)$$

Applying some properties of logarithms, we see that

$$ln\, y = ln3 + t \cdot ln2.$$

The last equation tells us that $ln\, y$ is a linear function of t; i.e., a plot of $(t, ln\, y)$, must be linear!

Looking for Exponential Models

The explanation above gives us the information we need to test data points for an exponential fit. Suppose we have some data points, (t, y), and suppose also that the semilog plot of these data points is linear. This linearity of the semilog plot implies that

$$ln\, y = mt + b, \quad \text{for some constants } m \text{ and } b.$$

Making both sides of the equation an exponent of e, we get the equation

$$e^{lny} = e^{mt+b}.$$

Simplifying this equation gives us the following function:

$$\begin{aligned} y &= e^b e^{mt} \\ &= ce^{mt} \end{aligned}$$

(The two-symbol constant, e^b, was replaced with a single symbol for a constant.)

We have shown two important results:

- If the semilog plot is linear, then $y = ce^{mt}$, for some constants c and m.
- The slope of the semilog plot is the exponent m.

Part 2: Completion of the Experiment

1. We pretend here that we do not know the precise relationship between t and y for the "data" points in Part 1. We have only the numbers from the table in step 1, and we deduce from our semilog plot in step 2 of Part 1 that a function of the form $y = ce^{mt}$, for some constants m and c, should fit the data. We must now find the constants m and c.

With the information from the previous "Explanation" in mind, look back at the semilog plot and find the slope of that line. You now have the constant m. [If you used the points corresponding to $t = 1$ and $t = 5$ to compute m, you should have found that $m \approx .69$.]

2. Next we will find a value for c. You may have noted that c is the value of y at time $t = 0$, and we have $y(0) = 3$ listed in the table. In this case that choice for c is fine, because our semilog graph is a perfect line. However, with real data, the semilog plot may be close to being linear, but not perfectly linear.

If you draw by hand a line to fit an approximately linear plot, then the line you think best fits the set of all points may not pass through the individual data point at time $t = 0$. In fact you may not have a data point corresponding to $t = 0$. In such cases you should use the fact the intercept of your fitted line on the $ln\, y$ axis is b; c is then given by the formula $c = e^b$. (Can you explain why?)

3. We have outlined the procedure for testing data to determine if it can be fitted with an exponential function, and we have seen how to find such a function. If you've followed along with the computations, then you probably found a function similar to this one:

$$\begin{aligned} y &= 3e^{0.69t} \\ &= 3(1.9937)^t \end{aligned}$$

But you know in this case that we used the function $y = 3 \cdot 2^t$ to create the data, which we then used to "rediscover" the function. Explain why the function we fitted to our data is a slightly different function from the original one.

Exercises

1. Refer to the US population data shown in the preview. Check to see that the semilog plot is approximately linear, and construct your own exponential function to fit the data.

2. Test the data in the following tables to determine if an exponential function could provide a good fit. If an exponential function is appropriate, then find it and make a graph of the exponential function superimposed on a scatter plot of the data. If using an exponential function is not appropriate, then explain carefully what conclusions you can draw from the semilog plot.

(a)

t	1	2	3	4	5	6	7
y	3.65	3.18	2.77	2.40	2.09	1.82	1.58

(b)

t	1	2	3	4	5	6	7	8	9
y	.111	.125	.143	.167	.200	.250	.333	.500	1.00

Power Function Fits

Another type of function that we often use to fit to data is the family of power functions, $y = c\,t^k$. It is as easy to test for this type of fit as it was for the exponential fit; indeed, we use the same method to discover what kind of plot to use: take logarithms of the equation for a power function, and use log properties to simplify the expression.

Assume $y = ct^k$.

It follows that

$$\begin{aligned} ln\,y &= ln(ct^k) \\ &= ln(c) + ln(t^k) \\ &= ln(c) + k\,ln(t) \end{aligned}$$

Because c and k are constants, this equation implies $ln(y)$ is a linear function of $ln(t)$. The implication here is that if $y = ct^k$, then a plot of the points $(\,ln(t), ln(y)\,)$ will be a line and the slope of this line is the degree of the power function. Plots of the points $(\,ln(t), ln(y)\,)$ are called “log-log plots.”

We can summarize these results as follows:

- If the log-log plot is linear, then $y = ct^k$, for some constants c and k.
- The slope of the log-log plot is the exponent k.

Exercise

Show that the log-log plot of the following data is linear. Find a function that fits the data well, and superimpose the function over a scatter plot of the data. Make a semilog plot of the same data, and compare that plot to the log-log plot. Explain why the semilog plot has the shape that it does.

t	.09	.24	.39	.54	.69	.84	.99
y	1.64	2.09	2.37	2.57	2.73	2.87	2.99

Investigations

1. When he was trying to determine the laws of motion, Galileo Galilei (1564-1642) at first tried the hypothesis that the velocity of a free-falling object would be proportional to the distance fallen by the object. Use the following data to determine if Galileo's first conjecture could be true. If it cannot be true, then determine what the correct relationship between distance and velocity is. (You should find an expression that gives the distance fallen as a function of velocity, and you should test your function by graphing it over a plot of the data points listed in the table.)

Distance fallen in meters	7.056	56.64	82.369	117.649	137.64
Velocity in meters per second	11.76	33.32	40.18	48.02	51.94

2. While at a clinic in Roxboro, North Carolina, a patient was given a dose of Xylopain, after which he lost consciousness. He was rushed to Duke Hospital, where the level of the drug was checked hourly. The readings are recorded in the following table:

Hours elapsed since dose of Xylopain	2	3	4	5	6	7	8
Concentration of Xylopain in mg/ml	6	4.25	3	2.12	1.5	1.06	.75

(a) Find a function, $C(t)$, that approximates the level of Xylopain in the patient's blood t hours after the drug was given to the patient. To test your function, $C(t)$, show a graph of $C(t)$ superimposed on a plot of the data points.

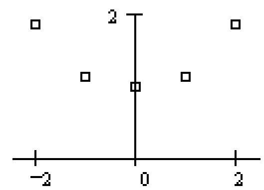
(b) The patient, who recovered from the incident, has since discovered that the maximum allowable concentration of Xylopain is 10 *milligrams per milliliter*. He believes that the clinic exceeded this maximum, and he plans to sue the clinic for negligence. Making use of the function you found in part (a), write a brief outline of the argument that the patient's attorney might make in court.

(c) Write a brief outline of a response that the defense attorney could make on behalf of the clinic.

Concluding Cautions and Comments

In an earlier lab you learned how to find a linear function to fit data which appears to be approximately linear. In this lab you learned how to fit exponential and power functions to data. But you should note that the only type of function we can identify simply by looking at its graph is a linear function. Indeed, we found the exponential and power functions by recognizing a line on a semilog plot. To illustrate the general difficulty of visually identifying nonlinear relationships, consider the data shown in the table and scatter plot below.

t	-2	-1	0	1	2
y	1.85	1.14	1.00	1.14	1.85



If you tried to guess the curve that fits best, you might think of functions of the form $y = ct^k + a$ because these are common, simple functions that can resemble the scatter plot above. Actually, we used the function $y = sec(\frac{t}{2})$ to create these "data" points.

Report

Your report should include the following:

1. A description of your work and conclusions from the "Investigation" of Galileo's first hypothesis.

2. Your responses to the three parts of the second "Investigation" concerning the concentration of a drug in a patient.