12 FOM Regression with a Logarithmic Function

From: https://courses.lumenlearning.com/ivytech-collegealgebra/chapter/build-a-logarithmic-model-from-data/

A logarithmic function has the general form $y = a + b\log_c x$, where c can be any base.

However, as it is possible to change the base for a logarithm, we tend to use one of two bases: 10 or e

$y = a + b\log x$ is the notation for a logarithmic function using base 10.

$y = a + b\ln x$ is the notation for a logarithmic function using the base e, where $e = 2.718\ldots$

The GeoGebra command to calculate a logarithmic **model** for a data set is 'fitlog'.

Build a logarithmic model from data

Just as with exponential functions, there are many real-world applications for logarithmic functions: intensity of sound, pH levels of solutions, yields of chemical reactions, production of goods, and growth of infants. As with exponential models, data modeled by logarithmic functions are either always increasing or always decreasing as time moves forward. Again, it is the *way* they increase or decrease that helps us determine whether a **logarithmic model** is best.

Recall that logarithmic functions increase or decrease rapidly at first, but then steadily slow as time moves on. By reflecting on the characteristics we've already learned about this function, we can better analyze real world situations that reflect this type of growth or decay. When performing logarithmic **regression analysis**, we use the form of the logarithmic function most commonly used on graphing utilities, $y = a + b\ln(x)$. For this function

- All input values, *x*, must be greater than zero.
- The point (1, *a*) is on the graph of the model.
- If *b* > 0, the model is increasing. Growth increases rapidly at first and then steadily slows over time.
- If *b* < 0, the model is decreasing. Decay occurs rapidly at first and then steadily slows over time.

A GENERAL NOTE: LOGARITHMIC REGRESSION

Logarithmic regression is used to model situations where growth or decay accelerates rapidly at first and then slows over time. We use the command "LnReg" on a graphing utility to fit a logarithmic function to a set of data points. This returns an equation of the form,

$$y = a + b\ln(x)$$

Note that

- all input values, *x*, must be non-negative.
- when *b* > 0, the model is increasing.
- when *b* < 0, the model is decreasing.

EXAMPLE 2: USING LOGARITHMIC REGRESSION TO FIT A MODEL TO DATA

Due to advances in medicine and higher standards of living, life expectancy has been increasing in most developed countries since the beginning of the 20th century.

The table below shows the average life expectancies, in years, of Americans from 1900–2010.[1]

Year	1900	1910	1920	1930	1940	1950
Life Expectancy(Years)	47.3	50.0	54.1	59.7	62.9	68.2
Year	1960	1970	1980	1990	2000	2010
Life Expectancy(Years)	69.7	70.8	73.7	75.4	76.8	78.7

1. Let x represent time in decades starting with $x = 1$ for the year 1900, $x = 2$ for the year 1910, and so on. Let y represent the corresponding life expectancy. Use logarithmic regression to fit a model to these data.
2. Use the model to predict the average American life expectancy for the year 2030.

TRY IT 2

Sales of a video game released in the year 2000 took off at first, but then steadily slowed as time moved on. The table below shows the number of games sold, in thousands, from the years 2000–2010.

Year	2000	2001	2002	2003	2004	2005
Number Sold (thousands)	142	149	154	155	159	161
Year	2006	2007	2008	2009	2010	–
Number Sold (thousands)	163	164	164	166	167	–

a. Let x represent time in years starting with $x = 1$ for the year 2000. Let y represent the number of games sold in thousands. Use logarithmic regression to fit a model to these data.
b. If games continue to sell at this rate, how many games will sell in 2015? Round to the nearest thousand.

Solution

Model: a function will rarely fit a data set perfectly. A real life data set may have a logarithmic trend, or linear/quadratic/sinusoidal etc but due to the many contributing variables will never be perfectly logarithmic/linear/quadratic/sinusoidal etc.

We use the word 'model' when calculating a function that fits the data 'well'. It is a word for applied math. We also have formulae for measuring how 'well' the function fits as you will have noticed that some functions almost seem to go right through all the points, and others go through none of the points. The regression process is to find the curve that minimizes the sum of distance between the curve and the points.