Mathematics Coursework

Exponential and Logistic Growth

Deriving and expression for Logistic Growth and evaluating models of this principle

Written By: Abhay Dalmia

Introduction

The rate of occurrences is one of the most important practical aspects of the use of mathematics in the real world. Calculus defines this very idea, that the rate of change, whether increase or decrease of anything in this world can be modeled with the correct parameters.

I was first intrigued, when I learnt Calculus, by its ability to model and more importantly predict situations and forecast the progression of such situations. Exponential and Logistic Growth, in this context, is related to the change of rate of something occurring, normally analogous to the population growth model or the disease growth model. Logistic Growth can model the population growth of the world.

In this investigation, an exponential and logistic model for the spread of a disease or rumor mathematically will, be derived, and the more appropriate form will be chosen. And then, the reliability and the accuracy will be tested, by creating an actual model of spread of a rumor/disease by use of both, a computer simulation and real-life event. Hence, the principle of logistic growth will be evaluated in its usefulness for real-life situations.

Note: All graphs made and conclusions drawn are from the data gathered through the lava Program. The exact data can be obtained by inputting the conditions stated into the program reproduced in Appendix 1.

Exponential Model

The exponential model for growth/decay assumes that the rate of increase and decrease is constant as time progresses. This is represented through the equation -

$$
\frac{dy}{dt} = k \cdot y
$$

In this equation, the rate of change of a variable γ with respect to time is controlled by a constant k , which dictates the rate at which y increases or decreases with time. This equation is further integrated by separation of variables -

$$
\frac{1}{y}dy = k \cdot dt
$$

Both sides of this equation are then integrated, so that a general equation that models exponential growth can be obtained. By integrating both sides -

$$
\int \frac{1}{y} dy = \int k \cdot dt
$$

$$
\ln|y| = kt + C
$$

This equation, hence, models the variable y , which could be seen as population or the spread of a disease with time, and how this variable, increases in the case of population, with time. This equation is further simplified to $-$

$$
\log_e|y| = kt + C
$$

$$
y = e^{kt+C} = e^{kt} \cdot e^C
$$

$$
y(t) = Ce^{kt}
$$

The value of the constant C is found by using the initial condition that logically follows that at $t = 0$, when the rate of increase/decrease has not started, the value of y will result in the initial number of occurrences, in this case $-$

$$
y(0)=y_0
$$

Hence, by plugging in the value for time and the resulting value of y , the final equation that models this principle of exponential growth is $-$

$$
y(t) = y_0 \cdot e^{kt}
$$

Graphical Representation

This graph represents one example of the family of curves, of the final exponential model found. It shows one example of different values for the constants y_0 and k.

However, one problem with this model, for simulating the rate of growth of population/disease or the spread of a rumor is evident when compared to an actual graph of the growth of population reproduced below.

¹According to the rate expression $\frac{dy}{dt} = k \cdot y$ that was used to derive the exponential model for population growth, the rate of growth is constant. It may also be expected to increase. However, according to the graph above, provided by the U.S. Census Bureau, is that even though the population continues to grow, the rate of this growth has actually been decreasing in the last 50 years. The exponential model however, does not evaluate this decrease. This can be seen by the graphical representation of family of curves of the exponential model, where with time, the value of occurrences will continue increasing to infinity.

But actual data suggests that rate should decrease, eventually even reaching zero, causing the curve to attain a maximum value also known as the barrier value or carrying capacity.

¹ "Population Growth Rate." *U.S. Census Bureau*. US Government Organization, n.d. Web. 4 Dec. 2013. <http://b.static.trunity.net/files/118301_118400/118325/620px-Figure_1_longterm_population_growth.JPG>.

This is the main problem with the exponential model, which is corrected by the logistic growth model.

Logistic Growth Model

This model is similar to the model of exponential growth, however, takes into account the maximum value that occurrences can attain, for example population, because it is only logical for example, that if a disease starts to spread, it cannot spread to an infinite value. It is limited by a certain number, for example, in this case the number of humans. The number of humans infected by the disease cannot be greater than the total population itself (something that would be possible under the exponential model)

Hence while the rate of growth of the exponential model was

$$
\frac{dy}{dt}_{\exp.} = k \cdot y
$$

The equation can be limited to a maximum value of M by multiplying the rate of exponential growth, by the fraction of the variable ν that has not been affected. For example, following the analogy above, the fraction of the people not yet affected by the disease. This is logical, because it would show that at the end, when carrying capacity is reached, the rate of growth would tend towards zero. (*Further elaborated with graphical representation of this model*). Hence, the rate of growth of a disease, by this model would be $-$

$$
\frac{dy}{dt_{\text{log.}}} = \frac{dy}{dt_{\text{exp.}}} \cdot \frac{M - y(t)}{M}
$$

$$
\frac{dy}{dt_{\text{log.}}} = k \cdot y \left(\frac{M - y(t)}{M}\right)
$$

Again this is only the rate if growth. Now, the equation in terms of a variable y and time t has to be derived in order to model the growth. Hence, by separating variables and integrating -

$$
\int \frac{1}{y \cdot \left(1 - \frac{y(t)}{M}\right)} \, dy = \int k \, dt
$$

The term on the left hand side is a rational polynomial function that cannot be further simplified. Hence, in its current form, it is very difficult to integrate. Hence, it has to be simplified and then integrated by use of the principle of partial fractions.

Simplification of the Complex Term

The complex denominator has to be separated to two or more terms that can be individually integrated. This approach is called integration by partial fractions. Hence, to divide the denominator into more terms -

$$
\frac{1}{y \cdot \left(1 - \frac{y(t)}{M}\right)} = \frac{A}{y} + \frac{B}{\left(1 - \frac{y(t)}{M}\right)}
$$

This arrangement allows the values of A and B to be found, since the denominators and numerators can be equated. The expression on the right hand side, if an attempt to combine terms is made, will result in the same denominator as in the left hand side. Hence, the numerator created in this process should also equate to the numerator on the left hand side -

$$
\frac{1}{y \cdot \left(1 - \frac{y(t)}{M}\right)} = \frac{A\left(1 - \frac{y(t)}{M}\right) + By}{y \cdot \left(1 - \frac{y(t)}{M}\right)}
$$

Therefore, an attempt to find the values of A and B is made by equating the two numerators.

$$
1 = A\left(1 - \frac{y(t)}{M}\right) + By = A - \left(\frac{A}{M} + B\right)y
$$

Hence, the coefficient of y should be 0, since there is no term in y on the left hand side, and the term without y should be equal to 1. Therefore, the two equations created by this is -

(1)
$$
A = 1
$$

(2) $-\frac{A}{M} + B = 0$

By solving the equations simultaneously (although one equation directly lead to an answer in this case), the value of A and B is found to be –

$$
A=1, B=\frac{1}{M}
$$

Hence, these values are finally replaced back into the equation with partial fractions at the start of this page $-$

$$
\frac{1}{y \cdot \left(1 - \frac{y(t)}{M}\right)} = \frac{1}{y} + \frac{\frac{1}{M}}{\left(1 - \frac{y(t)}{M}\right)}
$$

By this method, the complex term has been split into two other terms, which can be individually integrated. This final simplification, is finally inserted back into the integration previously $-$

$$
\int \frac{1}{y \cdot \left(1 - \frac{y(t)}{M}\right)} dy = \int k dt
$$

$$
\int \frac{1}{y} + \frac{\frac{1}{M}}{\left(1 - \frac{y(t)}{M}\right)} dy = \int k dt
$$

$$
\ln|y| - \ln\left|1 - \frac{y}{M}\right| = kt + C
$$

As before, this I further simplified to produce an equation in y and t where the constant value is found -

$$
\ln \left| \frac{y}{1 - \frac{y}{M}} \right| = kt + C
$$

$$
\frac{y}{1 - \frac{y}{M}} = Ce^{kt}
$$

Finally analogous to the exponential model, at $t = 0$, the initial value of $y = y_0$. By inserting these values into the equation, the constant is found and the $\frac{1}{2}$ variables are separated to derive the world-famous equation for logistic growth -

$$
y = \frac{M \cdot y_0}{y_0 + (M - y_0)e^{-kt}}
$$

Graphical Representation

In the graphs above, the maximum capacity of the occurrences M , assumed as 100 is drawn with different values of the constants y_0 and k.

These graphs, in my opinion, greatly follow logic. Assuming that this is a model for the spread of a disease, in the starting, the rate increases from 0 to a maximum value. This follows logic, as in the starting, if only 1 person is infected, the rate of spread will be relatively slow, as only the people who meet the infected will also be infected. However, as more people are infected, the rate increases as the infected people meet more non-infected people. From this information alone, we could come to the conclusion since more people are being infected per unit time, the rate of spread of the disease will continue to increase (like the exponential model).

However, there are a maximum number of people. The rate decreases to 0 as it reaches this point, because as people meet, there is a lower probability that either one is infected. Hence, as a greater proportion of society is infected, the probability that an infected and non-infected meeting occurs decreases in probability. This decrease is shown by the rate decreasing eventually to reach 0, when all people have been infected.

This logic seems to lend reliability to this theory.

Real-Life Modeling

At first, I couldn't exactly understand from the equation of the sigmoid itself. under what conditions of population growth or spread of a disease was the equation derived. However, after reading the original manuscripts of Malthus (1789) who invented the exponential model which was further modified by his student, Pierre-François Verhulst (1804{1849), who added the further principle of carrying capacity and resistance to further growth.

Verhulst employed the real values of Belgium, and created the sigmoid curve. This is further applied to the spread of a disease or rumour, particularly by the research paper by Dr. D. Batic and Mr. D. Dunn, which will be explored by both the computer simulation and the real-life experiment. This equation was created, by considering that meetings of two people take place, and if anyone has the disease or knows the rumour, it is passed to the other person. This is how the spread takes place. Hence, this exact condition will first be explored and then different conditions, and the affect of different conditions on the sigmoid curve of occurrence vs. time.

Computer Simulation

The computer simulation is a Java Program made on BlueJ (**Appendix 1**). I made this program myself as a method of running a model simulation that generates the required data.

This Java program is imperative. Basically, this program creates an array, like a list, of 0 's the size of the carrying capacity. Hence, if the carrying capacity is 100 people, then this list will have 100 zeros. Then, some of the 0's are converted to 1's randomly, based on the initial value input by the user. This initial value is the initial number of people out of the 100 that have the disease. Once these two values are input, there is a list of 1's (those who have he disease) and 0 's (those who do not have the disease.

After this, two elements are compared randomly, and if any of the two are diseased (value of 1), then both will be given a value of 1. Then, every number of particular increments, also input by the user (example $-$ 30), the number of 1's will be output again (the current number of people who have the disease). This will be run a number of times input by the user (example 600). A sample output is given below.

Hence, every 30 increments, the number of diseased people is output by the program for up to 600 runs, according to the data entered into the program.

A graph was drawn for the following parameters using the program. The Java program simulated a spread of disease amongst people, and output values of people infected after each increment. Sample of data output by the program is given in Appendix 2.

This graph was created with 63,765 data points. Each red point on the **graph stands for 1000 actual data points' output by the program.** The program could output an even greater number of data points, if a smaller increment was taken. Hence, the graph can be made as accurate as needed/possible.

To prove whether the curve follows the equation for logistic growth above, the equation from the graph is compared to what the equation should be according to the equation of logistic growth.

From Graph -

$$
y = \frac{150000 \cdot 5}{5 + 149999 \cdot e^{-0.00032x}}
$$

Logistic Growth Equation theoretically -

$$
y = \frac{150000 \cdot 5}{5 + 149999 \cdot e^{-kx}}
$$

As seen by the similarity in the two equations, the theoretical equation is almost congruent to the equation derived from the computer simulation. There are three conclusions that can be drawn from this $-$

- 1. The equation for logistic growth is in fact derived correctly, as a computer simulation gives exactly the same result. This is astounding, because mathematicians created this equation in the 1800's when computer simulation was not available to form or prove relationships. Now, with the use of computer simulation, we can see just how mathematically correct the equation is.
- 2. The equation of k found for this curve is found by comparing the two equations above (the theoretical and actual equations).

 $k = 0.00032$ when $M = 150,000$ and $y_0 = 5$

However, this value of k is only appropriate when the carrying/maximum capacity is 150,000 and the initial value of number of people diseased is 5.

- 3. From the equation, it is noted that logistic growth is only affected by the maximum capacity and initial number of occurrences. These two factors will transform the curve and change the value of k . k Itself is a constant and hence logically will not affect the curve, but instead is affected by transformations in the curve
- 4. Since the theoretical equation of logistic curve is proved perfectly correct and accurate through the computer simulation, it can be inferred that the rate of change equation used to derive this expression is also correct. Hence, the rate of growth/decrease at any point of the process can be pinpointed by inserting the value of the variable ν at that instance

$$
\frac{dy}{dt_{\text{log.}}} = k \cdot y \left(1 - \frac{y(t)}{M} \right) = 0.00032 \cdot y \left(1 - \frac{y}{150000} \right)
$$

5. As time increases to an indefinite point, the value of y and hence occurrences reaches the maximum/carrying capacity.

$$
\lim_{t\to\infty}y=M
$$

6. The point of inflection, on differentiating the equation and equating it to 0 , is exactly half of the carrying capacity. This is very interesting, as this means that the maximum rate of spread of a disease will occur when exactly half the population has it. After that it will start decreasing again.

$$
\max_{0 \le t < \infty} \frac{dy}{dx} = \frac{M}{2}
$$

Transformations of the Logistic Curve

According to the 3^{rd} conclusion drawn above, the affect of the two factors that affect the logistic curve is explored, for complete understanding of logistic curves and how they are transformed. Hence, the factors that are evaluated are -

1. Initial Number of People

For all Graphs Below (conditions to obtain exact same data) -

```
Maximum Capacity M = 10,000Increments = 100Number of Times Run = 1,000,000
```


This graph lead to another surprising conclusion, in that as the initial number of people is increased from 1 to 5 and finally 10, the graph is simply shifted to the left. Corresponding points on each curve have exactly the same rate of growth (gradient of this graph). There is no change in gradient, but the graphs are simply translated horizontally.

This also makes logical sense, since if the initial value is 5 instead of 1, then from the starting itself, there is a greater probability and hence rate of spread of the disease. At some point, when the initial condition is 1, it also becomes 5 and from there it follows the exact same path as when the initial condition is 5.

Also, from the graph, it is concluded that a change in the initial number of people will lead to a constant transformation. For example, the transformation between

1 and 5 will be the same as 5 to 9 and so on. Hence, for every point of initial value, the horizontal translation can be quantified.

2. Maximum/Carrying Capacity

For all Graphs Below (conditions to obtain exact same data) -

Initial Number of People $y_0 = 5$ Increments $= 50$ Number of Times $Run = 1,000,000$

This graph was very interesting. Not only was it concluded from the previous factor that changes in initial number of occurrences will simply translate the curve, but essentially, the same path in terms of growth rate will be taken.

In this case too, it can be seen that very similar paths are taken between curves. When the carrying capacity is increased from 5000 to 10000, the relative corresponding rate of logistic growth stays the same. In other words, the rate halfway through the blue curve is equal to that of the red curve.

Hence, the graph seems to be only stretched upwards and sideways, between all three curves. This is the effect of change in carrying capacity with time.

Conclusion

This was an extremely interesting investigation into sigmoid curves and logistic growth. This mathematical concept is important due to its far-reaching consequences, in predicting the frightening future of the world, where population growth is concerned. A human link to society, this logistic curve represents that the population will stagnate in the future and might also start to decrease. This is frightening because that means that some time in history, the birth and death rate will be equal. The logistic growth, represented by a sigmoid curve, also models the spread of a disease.

This is perhaps the most important application, as the rate at which an outbreak of a disease in the future can be mathematically calculated. Even in this respect, the sigmoid curve paints a frightening picture, showing how quickly, as more people are affected, the rate at which the disease will spread will increase. But it also shows that towards the end, after half way, the rate at which the spread of disease takes place will actually begin to slow. This phenomenon is ground breaking and definitely extremely interesting.

From this investigation, the separate rates and equations for both exponential and logistic growth were derived. The preference of logistic growth over exponential was explained, as logistic growth takes into account a limit that occurrences can attain (since infinite number of occurrences in for example diseases is not only not possible but also useless). Then, by use of computer simulation, the theoretical derivation of logistic growth was proved. Hence, the aim was realized. A further investigation was carried out by use of the computer simulation, as to the effects of the maximum capacity and initial occurrences on the curve.

Hence, overall, through this investigation. I learnt exactly how to construct and model situations to a logistic curve, in order to make predictions based on its future rates and values. It was an exciting journey to drill down to how these curves are created, modified and interpreted.

The computer program was prefect in aiding this investigation into mathematics and the true power of computation, in its ability to carry out a process millions of times in matter of seconds, was realized. Computational processing power hugely complemented mathematics, and allowed to also derive the equation of the logistic curve without using calculus and logic as used primarily.

I am left with a desire to further perfect this logistic model by taking into account the many other factors that affect the variable y with time, not just the maximum capacity and initial occurrence value. A further investigation could be carried out to include factors that are included in F.J.Richards and J.A.Nelder form of logistic growth. They modified and adapted the curve and added further factors to it, in order to be able to map biological systems mathematically.

Appendix 1

```
import java.util.*;
public class log1
\mathbf{f}public static void main(String □args)
    Ł
        int number;
        Scanner scan=new Scanner(System.in);
        System.out.println("Maximum Number of People?");
        int max=scan.nextInt();
        System.out.println("Increments?");
        int inc=scan.nextInt();
        System.out.println("Inital Number of People?");
        int init=scan.nextInt();
        System.out.println("Number of Times to Run?");
        int run=scan.nextInt();
        int a□=new int[max];
        for (int i=0; i=max; i++)\mathfrak{c}[a[i]=0;3
        for (int i=0; i<init; i+1)
         ſ
             int randomNumber = (int ) Math.random() * max );
             if (a[randomNumber]==0)
             \mathbf{f}a[randomNumber]=1;
             Þ
             else
             \mathbf{f}\left| i=i-1; \right.<sub>B</sub>
        l3
         for (int i=1; i <= run; i++)
         Ł
             int randomNumber1 = (int ) Math.random() * max );
             int randomNumber2 = (int )( Math.random() * max );
             if (a[randomNumber1]==1 | | a[randomNumber1]==1)
             \mathbf{f}a[randomNumber1]=1;
                 a[randomNumber2]=1;
             <sub>3</sub>
             if (i%inc==0)
             \mathbf{f}int j;
                  number=0:
                  for (j=0; j<sub>max</sub>; j++)ł
                      if (a[j]=1)Ł
                           number=number+1;
                      l?
                  System.out.println(+i+"
                                                "+number);
             \mathbf{B}|1|3
B
```
Appendix 2

The tables above are a sample of the data collected. The program would output as much data as the increments allowed. The lower the increment values, the more precise and accurate data was generated. The most accurate graph created above is the first graph featuring almost $65,000$ data points. This translates to 65,000 rows of data on excel.

Hence, the methods used in this investigative task lead to accurate and precise conclusions and hence, computer simulation and programming greatly benefitted this endeavor.

Abhay Dalmia

Bibliography

- 1. "Logistic Equation." *-- from Wolfram MathWorld*. N.p., n.d. Web. 04 Apr. 2014. <http://mathworld.wolfram.com/LogisticEquation.html>.
- 2. "Logistic Function." *Wikipedia*. Wikimedia Foundation, 04 Mar. 2014. Web. 04 Apr. 2014. <http://en.wikipedia.org/wiki/Logistic_function>.
- 3. "Mathwords: Logistic Growth." *Mathwords: Logistic Growth*. N.p., n.d. Web. 04 Apr. 2014. <http://www.mathwords.com/l/logistic_growth.htm>.
- 4. *Nature.com*. Nature Publishing Group, n.d. Web. 04 Apr. 2014. <http://www.nature.com/scitable/knowledge/library/how-populations-growthe-exponential-and-logistic-13240157>.
- 5. "Population Growth Rate." *U.S. Census Bureau*. US Government Organization, n.d. Web. 4 Dec. 2013.

<http://b.static.trunity.net/files/118301_118400/118325/620px-

Figure 1 long-term population growth.JPG>.

6. Thompson, John N. "Logistic Population Growth." *Encyclopedia Britannica Online*. Encyclopedia Britannica, n.d. Web. 04 Apr. 2014. <http://www.britannica.com/EBchecked/topic/470416/populationecology/70579/Logistic-population-growth>.