STUDENT VERSION

Going Viral

Bill Skerbitz
Wayzata High School
Plymouth MN USA

STATEMENT

We wish to model the spread of a non-fatal but incurable virus throughout a healthy (i.e. non-infected) population. We will then analyze the model by considering two viewpoints. First, we will simulate the behavior of the model in order to obtain a plot of the number of people infected versus time and to get an intuitive "feel" for how the spread of the virus progresses. Then, we will analyze the expected behavior of the spread of the virus in order to develop a differential equation whose behavior matches, at least qualitatively, the behavior of the data obtained via the simulation. Finally, we will test and compare our two approaches by solving the differential equation analytically and observing how well it fits the simulated data. Our tasks will include 1) Consider Assumptions, 2) Devise and Perform a Simulation, 3) Devise a Differential Equation, 4) Solve the Differential Equation, 5) Compare Plots of the Results, and 6) Summarize and Comment.

Assumptions

To simulate the spread of the virus and create a viable model, we must make some **assumptions**:

a. The spread of the virus begins via the infection of a single member of the population.
b. Each carrier of the virus can potentially (but might not) pass the virus to at most one other individual on a given day.
c. Although an individual who has contracted the virus can transmit it to at most one other individual, they are themselves no longer susceptible to reinfection because once infected, an individual remains infected for all time.
d. Infection is never fatal, i.e. the population never decreases.

Students will number themselves from 1 to *n*. For ease of discussion, we will assume there are $n = 32$ students for the remainder of these instructions. To start, all students stand next to their chairs. Here are the rules for "going viral" that attempt to satisfy our assumptions.

1) Pick one student at random who will serve as the original carrier of the virus, and have that student sit. This student represents the initial condition $P(0)=1$, where $P(t)$ represents the number of individuals (students) in the population who have been infected with the virus at time *t* in days (label them as "infecteds"). As we progress through the simulation, the students who do NOT have the virus remain standing while the infecteds will sit.

2) Each iteration will represent one day in the life of this viral spread. At each iteration select *j* random integers from the TOTAL set of $n = 32$ numbers, with replacement, where *j* is the number of infecteds at the start of the iteration. For each number selected in this manner, if the number selected is already associated with an infected student, do nothing. Otherwise, move the student associated with this number to the infected state and have the student sit. For each day in the simulation, note the number of newly infected students on this day as well as the total number of students infected at the end of this day.

Sample Simulation Data

We proceed with our simulation, using the TI-83/84 calculator, in infecting our class with the virus. In Table 1, we show the results of a sample simulation using the randInt(a, b, c) command found on a TI-83/84.

Day	**Total number of students infected**	**Specific students infected (i.e. student #5 was infected first)**
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
…	…	…

(t, $P(t)$) = (day, total # of students infected at the end of day t).

a) Before you plot the points resulting from the simulation, consider:
   i) How many people are infected with the virus on day 0, i.e. what is $P(0)$?
   ii) What is the largest number of people that could ever have the virus (i.e. what is the maximum value of P)?
   iii) Describe the size of the infected population P when:
      a. the infected population is growing the slowest (# of pop units per day).
      b. the infected population is growing the fastest (# of pop units per day).

b) Enter the "Day" and "Number Infected" lists into the data editor on your calculator and do a scatterplot of "Number Infected" vs. "Day." You should recognize an S-shaped curve called the "logistic curve" (more generally, a "sigmoid curve"). If using the TI-83/84 calculator one may adjust the window accordingly … ZOOM:Stat usually works well. If using software instead of the calculator, use an appropriate plotting tool such as Geogebra or desmos.com to plot the ordered pairs.

Development of the differential equation

In light of the discussion regarding the rate of growth above, we can now discuss the development of a differential equation that could govern our viral spread.

Consider the following qualitative behavior of the viral growth:

c) What would the growth of the virus be if $P=0$ or $P=M$ (where M is the maximum number infected)?

d) For what value of P would the number of infecteds be growing the fastest? Why is this a reasonable conclusion?

e) The growth starts out looking like the exponential function $P(t)=2^t$. Superimpose a plot of $y=2^t$ along with the (t, P) data on the calculator to see how well it fits.

f) Assuming the derivative to be a continuously varying quantity, discuss and devise a smooth function $f\left(P\right)$ that has zeros at $P=0$ and $P=M$ and a maximum at $P=\frac{M}{2}$. Hint: See Figure 1.

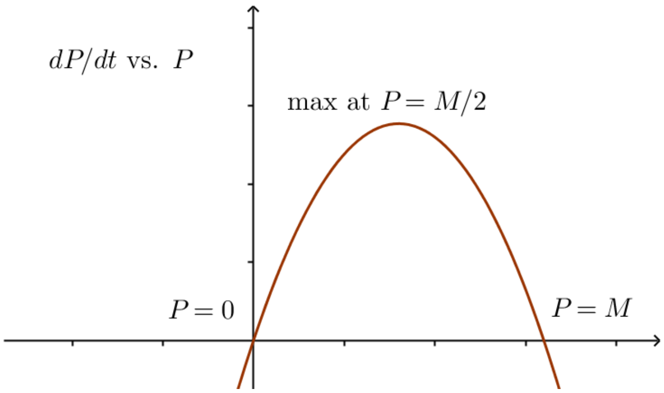

Figure 1. Sketch of smooth function in response to (f).

g) We know that the model behaves like $P(t)=2^t$ when $P\approx 0$. We have seen previously that a differential equation having the function $P(t)=2^t$ as a solution is given by $\frac{dP}{dt} = (\ln 2)P(t)$. Furthermore, when $P \approx 0$, the equation $\frac{dP}{dt} = K \cdot P(t) \cdot (M - P(t))$ will behave in a manner similar to $\frac{dP}{dt} = K \cdot P(t) \cdot M$. Now, in our example case, M = 32. Therefore, we would conclude that $KM = 32K \approx \ln 2$, or $K \approx \frac{\ln 2}{32}$ (and in general, under the assumption that the model starts out closely modeling $P(t) = 2^t$, K will work out to be $K = \frac{\ln 2}{M}$).

h) Our differential equation is therefore of the form $\frac{dP}{dt} = \frac{\ln 2}{32} P \cdot (32 - P)$. A common alternative form is given by $\frac{dP}{dt} = (\ln 2) P \cdot \left(1 - \frac{P}{32}\right)$. Each form has its advantages.

i) Plot a slope field with a CAS such as Mathematica (or Wolfram Alpha), Geogebra, etc., or on your graphing calculator. How does it compare to Figure 2?

j) Use our methods for separating variables and partial fraction decomposition to solve the differential equation (either form) in (h) (see discussion below on solving the logistic differential equation).

k) Now, plot the solution function from (j) along with the original data. What do you see? Offer comments and summarize our activity.

l) Plot the data, the slope field, and the solution on the same axes and comment.

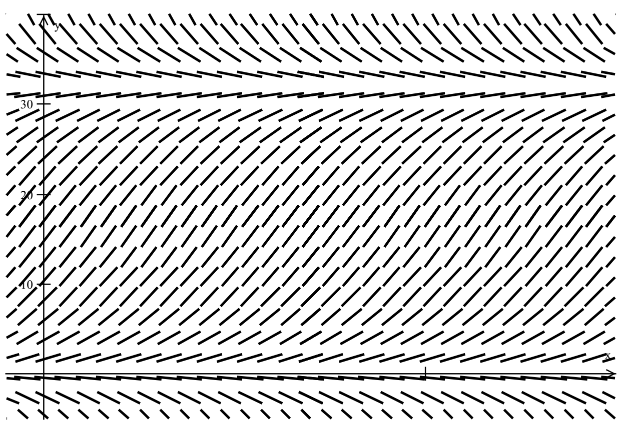

Figure 2. Sketch of slope field in response to (i).