Interdisciplinary Lively Application Project
Spread of Disease Activity

Title: Spread of Disease Activity

Authors: Bruce MacMillan
University of Colorado at Denver
Bruce.Macmillan@cudenver.edu

Lynn Bennethum
University of Colorado at Denver
Lynn.Bennethum@cudenver.edu

Date: January 20, 2005

Course Classification: Calculus 2 (MATH 2411 at CU-Denver),

Disciplinary Classifications: Mathematics, Biology

Prerequisite Skills:

1. Familiarity with Differential Equations
2. Scatterplot
3. Logistic Regression
4. Modeling

Mathematics Classifications: Modeling, Difference Equations Ordinary Differential Equations, Logistic Regression

Biological Classifications: Spread of Disease, Logistic Equation

Materials Required:

1. Dice (One die for each individual)
2. Graphics Calculator, preferably TI-89.

Interdisciplinary Lively Application Project:
Spread of Disease Activity

By Bruce MacMillan and Lynn Bennethum

Introduction Imagine that you are one of many people at a party and that, unknown to everyone else, one person arrives carrying an infectious disease. How quickly will the disease spread and what are the chances that you leave the party with the disease? In this project we explore these and several other questions related to the spread of disease and what is called the *logistic equation*.

During a later class we will conduct an activity that will model the spread of a disease in a closed environment. We will then analyze the data collected from the activity by using a discrete model (a difference equation) and a continuous model (a differential equation). First, however, we need to explore the type of data we will collect.

You will be asked to write up a series of miniature reports on this project. Though you should work in groups to discuss ideas and work through the student tasks, each person must turn in a report in his/her own words. Each report should include

1. A title page with the name of the ILAP, instructor, and course number-section.

2. A short introduction containing objectives of the activity.

3. A discussion which includes the task items. This should be written in proper English.

4. A conclusion.

The report may be handwritten or typed. It must be written grammatically correct. It is *not* enough to just turn in equations and graphs with no explanations (this sort of report will receive a failing grade). Graphs and equations should be motivated (why are we doing this?) and an explanation must be provided on how each new equation is derived.

I. Description of the Discrete Model

The goal is to create a mathematical model that describes the spread of disease at a party and also provides a crude description of real diseases. Imagine that there are N people at the party. Suppose the party is divided into M stages and let I_n represent the number of infected people at stage n, where $n = 0, 1, 2, ..., M$. The initial condition $I_0 = 1$ means that at the start of the party there is one infected person. The task is to devise a rule that tells us the number of infected people at stage $n+1$ if we know the number of infected people at stage n. Specifically, we look for a rule of the form

$$\underbrace{I_{n+1}}_{\substack{\text{Number of infected}\\ \text{at stage } n+1}} = \underbrace{I_n}_{\substack{\text{Number of infected}\\ \text{at stage } n}} + \underbrace{G_n}_{\substack{\text{Number of new}\\ \text{infected at stage } n}}, \tag{1}$$

which says that the number of infected people at the next stage is the number of infected people at the current stage plus the number of newly infected people at the next stage. The challenge is to determine G_n, the number (growth) of newly infected people.

Here is how we reason. New infections occur because of interactions between infected people and uninfected people. (We will neglect interactions between two infected people, which does not add to the number of newly infected people). The number of newly infected people is related to the number of interactions between the uninfected and infected. Consider Figure 1 where there are 7 people. If there is only 1 infected person, the number of interactions between infected and uninfected people is 6. If there are 2 infected people, the number of interactions for each infected person with uninfected people is 5. The number of infected people are 2, so the number of interactions between infected and uninfected is 2×5 or 10. If there are 3 infected people, then the number of interactions for each infected person is 4, so that the total number of interactions between infected and uninfected persons is 3×4 or 12.

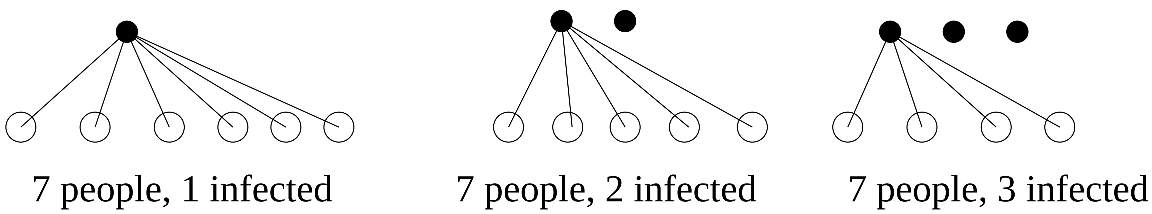

Figure 1: Interactions Between Infected and Uninfected

In general the number of interactions between infected people and uninfected people is proportional to the product of the number of infected and uninfected people, or $I_n \times (N - I_n)$. However this is the maximum number of interactions and not all interactions will result in a new infected person. Thus, we have

$$\underbrace{I_{n+1}}_{\substack{\text{Number of infected}\\ \text{at stage } n+1}} = \underbrace{I_n}_{\substack{\text{Number of infected}\\ \text{at stage } n}} + \underbrace{aI_n(N - I_n)}_{\substack{\text{Number of new}\\ \text{infected at stage } n}} \quad \text{for } n = 0, 1, 2, ..., M-1, \tag{2}$$

where $0 < I_0 \leq N$, and where a is a proportionality constant, $0 < a < 1$, which measures the fraction of interactions that are risky and lead to infections. This relationship that describes how the number of infected people evolves is called a *difference* or *discrete equation* because it involves no derivatives and determines the number of infected persons at discrete time step.

Student Task #1
We want to justify the form of G_n:

1. What happens to the difference equation when all the people are infected?
2. What happens when no one is infected ($I_n = 0$)?
3. For $G_n = aI_n(N - I_n)$, is it possible for I_n to decrease?
4. Using your answers from 1, 2, and 3 above, what assumptions do these properties imply about the disease?
5. Give an example of a disease which might be reasonably modeled by this difference equation, and a different disease which might not be reasonably modeled by this equation. Explain.

Student Task #2

In this example, let N (the number of people at the party) be 50, a (the proportionality constant) be 0.021, and let the initial condition be $I_0 = 1$.

1. Create a chart with headings n (Stage Number), and I_n (Number of Infected persons at stage n). The first row in the chart should have $n = 0$ and $I_0 = 1$. Using the difference equation (2) and the value of 0.021 for a, fill in the table to show the spread of the disease until the entire population has been infected. Note: While filling in the table, round to a whole number (on the TI calculator use 'round(expression,0)') since decimals do not make sense in this problem). Make a scatterplot of the values in your table and print the scatterplot using a computer.

2. The scatterplot shows a pattern that begins to grow "exponentially", but although it continues to increase, the concavity changes and the data points tend toward a horizontal asymptote. A function that fits data such as this is called a logistic growth function. To find such a function that fits your data, perform a "logistic regression" on your calulator/computer. Record the resulting regression equation and print the graph of the function on the scatterplot.

3. **Steady-state** is the term used to describe a solution which does not change in time (i.e. the rate of change of I, in this case, is zero). Using your scatter plot, what is the steady-state value of I as the number of stages increases to infinity? By using the logistic equation, does this make sense?

Student Task #3

1. To see the effect of the proportionality constant a, redo Student Task #2 with a value of $a = 0.015$ and then with $a = 0.026$.

2. What effect does the value of a have on the scatterplot and the logistic growth function?

IIA. Description of the Continuous Model

The previous model, which involved a difference equation, is called *discrete* because it gives the infected population at distinct (discrete) stages in time. The solution appears to lurch ahead in jumps at each stage. The result of the regression equation (using technology), is a more realistic model in which the infected population increases smoothly or continuously. To obtain this mathematical model, we can use a differential equation. An argument very similar to that used to obtain the discrete model can be used to derive an analogous differential equation that describes the growth of the infected population. We now let $I(t)$ denote the number of infected people at time $t \geq 0$. The corresponding differential equation is

$$I'(t) = aI(N - I) \tag{3}$$

where the initial condition is $I(0) = 1$, and the coefficient a is non-negative, $a \geq 0$. Note that the constant a in the continuous model is not the same constant as the proportionality constant in the discrete model.

Student Task #4
We want to justify the form of the continuous equation (3).

1. Using the limit definition of derivative, can you make an argument as to why the continuous equation and the difference equation are roughly equivalent?

2. The dimensions of a variable are the generic units a variable might have. For example, velocity has units of distance/time. Acceleration has units of distance/time^2 . Each term in an equation must have the same units. So in the continuous equation, I has units of number (e.g number of infected people), and $I'(t)$ has units of number/time. What are the units of a?

3. For what values of I is the solution at steady-state? (Hint: when is $I'(t) = 0$? Look at equation (3)). Explain why this makes sense.

4. Using the continuous equation (3), determine the values of I which make $I'(t)$ positive. Explain why this makes sense.

Student Task #5

We want to see the effect the constant a has on the differential equation.

1. Draw the slope field for equation (3) with $a = 0.04$, 0.01, 0.03, and 0.07. Let $N = 50$ and be sure the "window" of your calculator/computer shows the different resulting slope fields. On each slope field sketch the particular solution when $I(0) = 1$. Print the slope fields using a computer and give the window values which you used to show that for each value of a you still have logistic growth.

2. Give a physical interpretation of the constant a. Compare your interpretation with the units of a of Student Task #4. Are your answers consistent? Explain.

Student Task #6

1. By hand (without technology), solve equation (3). Show that the solution can be written as

$$I(t) = \frac{N}{1 + Be^{-aNt}} \quad . \tag{4}$$

   Be sure to show all steps neatly in your analysis.

2. Using the conditions $N = 50$, $a = 0.03$, and $I(0) = 1$, find the particular solution to the differential equation and compare it with the appropriate slope field.

IIB. Data Collection Activity

Data Collection: The procedure for the collection of the data will be described in class. Each student needs a standard die and the Data Collection Sheet attached.

Student Task #7

Make a scatterplot of the data collected from the activity in class. To find the particular function that will fit your data (and is also of the form of your solution (4)), we need to find values for the three constants in your solution. What is the value of N in your solution? To find the other two constants, we need two points from the scatterplot. Write down the initial point ($t = 0$) and use it to find the second constant in your equation. By tracing on the scatterplot, find another good point that represents the data. Use it to solve for the third constant. Show the algebra used in finding these constants. Write down your final equation, graph it on the scatterplot, and see how it fits. (If it is not a good fit, try another point to find the third constant, and regraph.)

Data Collection Sheet
Your ID Number: ________________

Stage #1	Stage #2	Stage #3	Stage #4	Stage #5	Stage #6

Stage Number	Number of Newly Infected Individuals	Number of Total Infected Individuals
0		
1		
2		
3		
4		
5		
6		

III. Population Model

The Logistic equation can also be used to model population growth. We change variables to denote the different problem we are modeling. If N is the population then the simplest population growth model is

$$N'(t) = rN, \tag{5}$$

where $N(t)$ is the population at time t and r is the (constant) growth rate.

Student Task #8
As you may know the world population reached 5 billion in 1988 and was growing at a rate of 1.6% per year. Assuming the growth rate remains constant and that the world population reached 5 billion at the beginning of 1988, use equation (5) to determine the function that describes the world population growth for time after 1988. According to this function:

1. When will the population reach 6 billion?
2. How many people will be added (net gain) in 1999?
3. How many people will be added (net gain) in 2000?
4. *If* you assume that the human race began with two individuals, what date does this model give for the "creation"? Is the model accurate? Why or why not?

Student Task #9
Now modify the simplest population model to a more realistic model - the logistic equation

$$N'(t) = rN(1 - \frac{N}{k}).$$

Explain how this model is an improvement over equation (5) and what the meaning of the coefficient k is.