The Mathematics of Epidemics

Enduring Understandings:

- Epidemics follow a pattern of exponential growth and decay.
- Mathematical models can be used to represent complex and dynamic systems, such as the spread of an infectious disease.
- Using mathematical models of infection allows public health officials to predict the extent of an epidemic, as well as the effectiveness of alternative prevention strategies.

Essential Questions:

- What is an epidemic, and what factors affect it?
- How can mathematics be used to understand epidemics?
- What are the key factors in an AIDS epidemic?

Notes to the Teacher:

This lesson introduces a common model used in the study of epidemics—the SIR model. The lesson guides students through both a mathematical analysis of the model as well as a numerical analysis, conducted by running the model using a program such as Microsoft Excel.

You can find a related lesson, "Infectious Disease and Exponential Growth," in the resource list at the end of this lesson. It focuses on understanding exponential growth, using a simple model for the spread of rumors.

Duration of Lesson:

Approximately four 50-minute periods

Assessment:

Assessments are embedded in the lesson and may be given as either in-class or homework assignments.

Mathematics Standards
Indicators addressed by this lesson:

Standard 1: Understands and applies basic and advanced properties of functions and algebra

Level IV (Grades 9–12)

1. Uses a variety of strategies (e.g., identify a pattern, use equivalent representations) to understand new mathematical content and to develop more efficient solution methods or problem extensions
2. Constructs algorithms for multi-step and non-routine problems
7. Understands connections between equivalent representations and corresponding procedures of the same problem situation or mathematical concept (e.g., a zero of a function corresponds to an x-intercept of the graph of the function, the correspondence of binary multiplication to a series electrical circuit and the logical operation "and")
8. Understands the components of mathematical modeling (i.e., problem formulation, mathematical model, solution within the model, interpretation of solution within the model, validation in original real-world problem situation)

Standard 8: Uses a variety of strategies in the problem-solving process

Level IV (Grades 9–12)

7. Uses a variety of models (e.g., written statement, algebraic formula, table of input-output values, graph) to represent functions, patterns, and relationships
8. Understands the general properties and characteristics of many types of functions (e.g., direct and inverse variation, general polynomial, radical, step, exponential, logarithmic, sinusoidal)
9. Understands the effects of parameter changes on functions and their graphs

Materials Needed:

Scientific calculators

Computers with Microsoft Excel or a similar program

Activity: A simulated HIV/AIDS Epidemic requires several six-sided dice

Handout 1: Data Table for an SIR Model

Procedure:

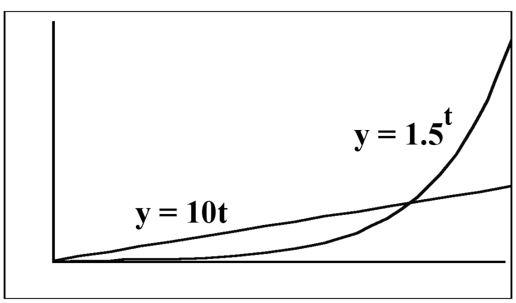
The SIR Model for the Spread of Infectious Disease

1. Introduce the lesson to students by giving them the following information:

   As the movie *Beat the Drum* makes clear, an HIV/AIDS epidemic is rampant in parts of Africa. Other epidemics, such as measles, bubonic plague, smallpox, malaria, diphtheria, scarlet fever, and influenza have killed millions of people around the world, and some diseases continue to do so. Between 1665 and 1666, London experienced a massive outbreak of bubonic plague that killed more than 75,000 Londoners (nearly 20 percent of the city's population). In 1918, in the United States alone, a flu epidemic resulted in 675,000 deaths. In 2002, worldwide, measles took 700,000 lives, malaria 1.1 million, tuberculosis 1.6 million, and AIDS 3.9 million.

   Explain to the class that the next several class periods will focus on understanding the mathematics of the spread of disease with its resulting epidemics. Students will use a simple, but very effective, mathematical model employed by epidemiologists to study ways to predict and influence the progress of epidemics. The lessons will use three different methods for understanding the underlying mathematics of disease: 1) in-class simulations with data collection and analysis, 2) algebraic analysis, and 3) computerized mathematical modeling.

2. Ask students what an epidemic is. Answers should emphasize the high rate of growth in new cases of the disease—much higher than expected, based on past experience. The word *epidemic* comes from the Greek *epi demos*, meaning "*upon people*."



Have a discussion about the rate of growth of a disease. How fast is fast? Have students draw a quick, rough graph that might approximate the number of new cases early in an epidemic. Ask what the abscissa and the ordinate should be. (x = Time, y = Number of new cases)

Some graphs may be linear and steep; others should be exponential. Illustrate both on the chalkboard or overhead.

3. Provide the class the following information:

Those who study the spread of disease, known as *epidemiologists*, have developed mathematical models to help predict and control epidemics. The usefulness of any mathematical model ultimately depends on the extent to which it both matches reality and enhances our understanding of reality. To do this, the disease model must capture many of the essential features and dynamics of an actual epidemic. For our purposes, we will still need to make a number of simplifying assumptions so that we can successfully build a model. Since we are working with the film *Beat the Drum*, we will develop a model whose assumptions and characteristics are based on our understanding of HIV/AIDS.

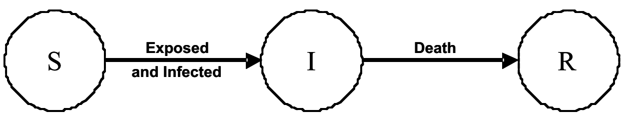
One model used for studying epidemics is called the SIR model. This model assumes that the population is divided into three distinct groups:

S group: ***Susceptible*** to catching the disease. For HIV/AIDS, this is everyone not already infected.

I group: ***Infected*** with the disease and contagious. For HIV/AIDS, once infected, people are contagious until they die.

R group: ***Removed*** from the population. For most diseases, this includes those who have had the disease and are now immune. For HIV/AIDS, no one recovers, so this group includes only those who have died from the disease.

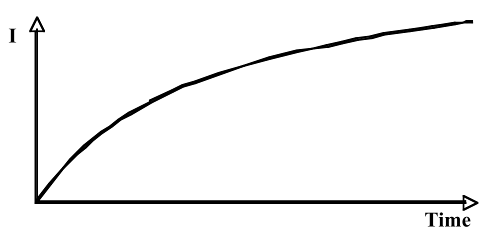
4. Ask students what moves a person between groups in the case of HIV/AIDS. Illustrate with the following diagram:



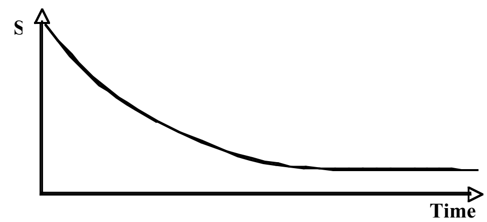
Discuss with students how the size of each group will change over time as an HIV/AIDS epidemic progresses, starting with the Infected group (I).

The Infected group probably starts with one infected individual and grows steadily. We know that for HIV/AIDS, not everyone in the population will become infected, but those who do will stay infected for life, and many will continue to infect others. In fact, on average, a person infected with HIV/AIDS will infect 3 to 5 other people during his or her lifetime.

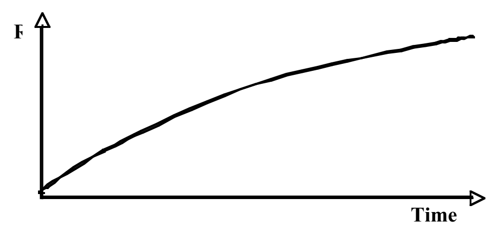
Ask students to sketch a graph for the size of the infected group over time. The graph should show exponential growth that tapers off over time, possibly with a maximum value:



What about the susceptible group (S)? As more and more susceptible people engage in unsafe practices with infected individuals, the size of this group declines over time. Ask students to sketch a graph for the size of the susceptible group over time. The graph should look like this:



For HIV/AIDS, the removed group (R) will grow as people with HIV contract AIDS and eventually die. Ask students to sketch a graph for group R over time.



5. Explain to the class that these graphs give us some idea of what an HIV/AIDS epidemic looks like, but they leave many questions unanswered:

- How many people will become infected?
- How fast will the epidemic progress?
- How many people will die?

A good model should help us answer some of these questions.

Suppose we know how big each group is at some point in time. Can we determine how big each group will be after some change in time?

6. In working with our mathematical epidemic model, we will assume that our population *N* stays constant at 6,000. This assumes that as people in group *I* die, new susceptibles move into the area to take their place. This means that $N = S + I$ remains valid.

Also, tell students to suppose we know from reports that for this population:

- There are currently 1,500 cases of HIV/AIDS.
- An average of 70 new cases have been reported each month.
- So far, 200 deaths from AIDS have been reported.
- An average of 19 new deaths from AIDS occur each month.

For our model, to keep things simple, we will not factor in normal births or normal deaths—only deaths from AIDS.

Put the following table on the board or overhead. Ask students how it can be filled in.

Months from Now (t)	S	I	New Infections (newI)	New Deaths (newR)	R
0					
1					
2					
3					

Note that each time a new person is infected, S decreases and I increases. Each time an infected person dies, I decreases, and R increases. This reasoning leads to the following set of equations:

$$S(t+1) = S(t) - newI + newR$$
$$I(t+1) = I(t) + newI - newR$$
$$R(t+1) = R(t) + newR$$

where, newI is the *fixed* number of new infections each month, and newR is the *fixed* number of new deaths each month. Point out that newR also represents the number of new S's moving into the area each month to replace those who died from AIDS.

With students, look at the first iteration from $t = 0$ to $t = 1$ (the next month):

$$S(1) = S(0) - newI + newR$$
$$= 4{,}500 - 70 + 19$$
$$= 4{,}449$$
$$I(1) = I(0) + newI - newR$$
$$= 1{,}500 + 70 - 19$$
$$= 1{,}551$$
$$R(1) = R(0) + newR$$
$$= 200 + 19$$
$$= 219$$

Have students work with you to fill in the table using the above equations. Since *I* = 1,500 currently, *S* = 6,000 – 1,500 = 4,500.

Months from Now (t)	S	I	New Infections (newI)	New Deaths (newR)	R
0	4,500	1,500	70	19	200
1	4,449	1,551	70	19	219
2	4,398	1,602	70	19	238
3	4,347	1,653	70	19	257

7. Explain to students that in reality, *newI* and *newR* are not fixed values. The values change over time. Ask students what they think will affect each of these rates of change. Answers should include:

newI: Depends on the number of *S*'s engaging in unsafe practices with *I*'s

Depends on the probability that the unsafe practice will result in a new infection

newR: Depends on the current number of *I*'s
Depends on the average death rate from AIDS

Define a few new variables:

C: The average proportion of *S*'s that a *single infected individual* "contacts" each month (i.e., engages in unsafe practices with).

In one month, an infected person will contact only a small fraction of the total *S* population. For example, if S = 5,000, and they contact an average of only 10 each month, then C = 10÷5,000 = 0.002.

F: The average probability of infection during an unsafe practice.

Not everyone who engages in an unsafe practice with an infected individual becomes infected.

We can combine these two variables to get one variable that gives us the proportion of *S*'s infected each month by a *single infected individual*:

ß: The proportion of *S*'s infected each month by *one I*.

So, $ß = C \times F$

Let,
μ: The average death rate from AIDS.

Ask students to suggest formulas for *newI* and *newR* that use *ß* and *μ*.

Answers:

$newI = ßS(t)I(t)$ — $ßS(t)$ gives the number of *S's* infected by *each I*.

We take that value and multiply by the number of *I*'s.

$newR = \mu I(t)$ — The number of deaths depends only on the number of *I's*.

8. Explain to students that *newI* and *newR* in the original model can now be replaced with the new variables. This gives the following model:

$$\begin{aligned} S(t+1) &= S(t) - newI + newR \\ I(t+1) &= I(t) + newI - newR \\ R(t+1) &= R(t) + newR \end{aligned} \quad \rightarrow \quad \begin{aligned} S(t+1) &= S(t) - \beta S(t)I(t) + \mu I(t) \\ I(t+1) &= I(t) + \beta S(t)I(t) - \mu I(t) \\ R(t+1) &= R(t) + \mu I(t) \end{aligned}$$

Review the assumptions made earlier:

The population was fixed at 6,000. We knew from reports that for this population:

- There are currently 1,500 cases of HIV/AIDS.
- An average of 70 new cases have been reported each month.
- So far, only 200 deaths from AIDS have been reported.
- An average of 19 new deaths from AIDS occur each month.

Since we know that infections depend on the sizes of *I* and *S*, and that AIDS deaths depend on the size of *I*, we can estimate *ß* and *μ* from this information.

$I = 1{,}500 \rightarrow S = 4{,}500$. Since we know that initially $newI = 70$, and from our model we know that $newI = \beta S(t)I(t)$, we have:

$$70 = \beta S(t)I(t) = \beta(4500)(1500), \text{ so}$$

$$\beta = \frac{70}{4500 \times 1500} \cong 0.00001037$$

We also know that $newR = 19$, and from our model, we know that $newR = \mu I(t)$. So we have:

$$19 = \mu I(t) = \mu(1500) \rightarrow \mu = \frac{19}{1500} \cong 0.01267$$

9. Ask students to fill in the table in **Handout 1: Data Table for an SIR Model** using our current model and the values for *ß* and *μ* that we've just computed. Students should round each answer to a whole number, representing the number of people affected. You may want to allow students to use a program like Microsoft Excel to do this assignment.

Go over student answers. Here is the correct table: (Also in the Answers to **Handout 1: Data Table for an SIR Model** if you wish to give students a copy):

			0.00001037	0.01267	
Months from Now (*t*)	***S***	***I***	$newI = \beta SI$	$newR = \mu I$	***R***
0	4,500	1,500	70	19	200
1	4,449	1,551	72	20	219
2	4,397	1,603	73	20	239
3	4,344	1,656	75	21	259
4	4,290	1,710	76	22	280
5	4,236	1,764	77	22	302
6	4,181	1,819	79	23	324
7	4,125	1,875	80	24	347
8	4,069	1,931	81	24	371
9	4,012	1,988	83	25	395
10	3,954	2,046	84	26	420

10. Ask students to look at what can be learned from our model. In particular, does the model allow us to predict how bad the epidemic will be?

From our model, we have:

$$I(t+1) = I(t) + \beta S(t)I(t) - \mu I(t)$$

If we define ΔI as the change in the size of group *I* from time t to time *t*+1, then we have:

$$\Delta I = I(t+1) - I(t) = \beta S(t)I(t) - \mu I(t)$$

We simplify this notation by writing *S* for *S(t)* and *I* for *I(t)*:

$$\Delta I = \beta SI - \mu I$$

Now, one of our model's assumptions was that the total population *N* stays constant:

$$N = S + I$$

So, we can replace *S* with the expression *(N-I)* in our equation for ΔI to get:

$$\Delta I = \beta(N - I)I - \mu I$$

Ask students to simplify this equation to get:

$$\Delta I = \beta I\left(N - I - \frac{\mu}{\beta}\right)$$

If we define a new constant, $K = N - \frac{\mu}{\beta}$, then we have:

$$\Delta I = \beta I(K - I)$$

In an AIDS epidemic, it is hoped that the number of infected individuals will eventually stop growing, i.e., I(t) will reach a maximum value.

11. Ask student to look at the equation for ΔI. Ask if *I*(t) is to reach a peak, what does the value of ΔI need to reach? (Answer: *I*(t) will reach a maximum value when $\Delta I = 0$.)

Ask students, given , > 0, what values for *I* and *K* will result in $\Delta I = 0$? (Answer: $I = 0$, or $I = K$.)

$I = 0$ means there are no more infected individuals, so the epidemic is over.

$I = K$ means that the number of infected individuals remains constant at level *K*.

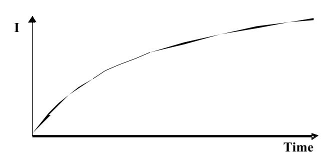
Therefore, K tells us what the maximum level of infection will be.

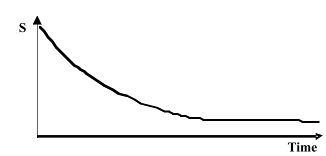
In our model, we have:

$$K = N - \frac{\mu}{\beta} = 6000 - \frac{0.01267}{0.00001037} = 4{,}778$$

This means our model has predicted that if nothing is done to stem the epidemic, approximately 4,778 individuals out of a population of 6,000 (that's 80%) will eventually become infected with HIV/AIDS!

12. Early in our discussion of the SIR model, students sketched graphs for several of the groups in the model:





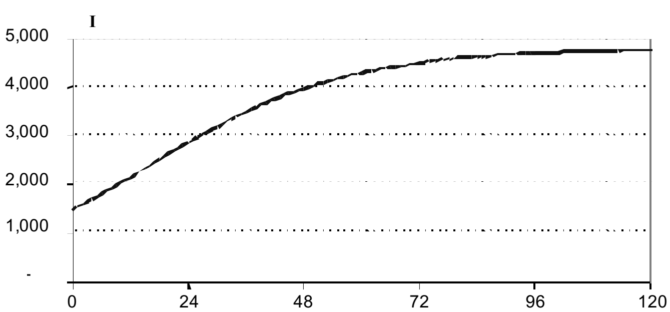
Ask students to look at those graphs to see how they compare with the model.

13. Have students use a program such as Microsoft Excel to extend their model data from Assignment 9 from 10 months out to 142 months. Have them generate separate graphs for *I, S, R,* and newI as functions of *t* (months).

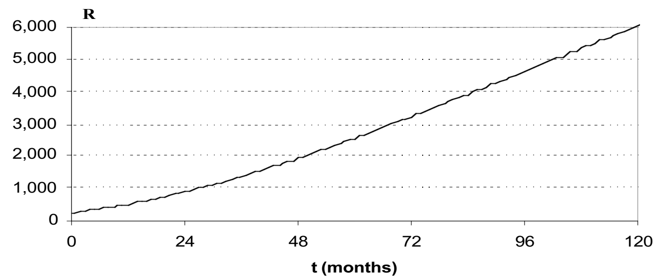
For best results have them set the scales for the y-axes as follows:

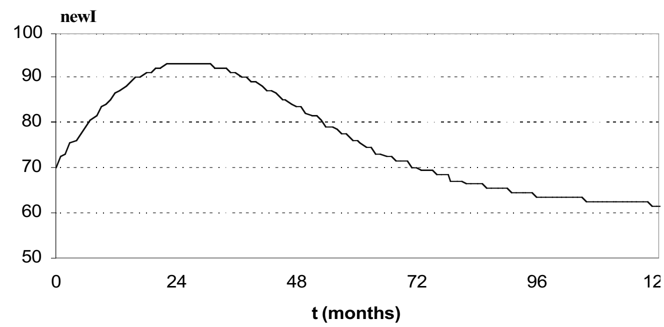
Graph	Min	Max	Major Unit
I	0	5,000	1,000
R	0	5,000	1,000
S	0	6,000	1,000
newI	50	100	10

Graphs should look like this:



Remind students that the model predicted a maximum value for *I*(t) of approximately 4,778. Our model reached a maximum of 4,776!





Have students describe the meaning of the graph for *newI*.

Answer: The number of new HIV/AIDS cases continues to climb for a little more than two years, levels off, and then declines exponentially to a steady state after 10 months of approximately 60 new cases per month. This figure exactly matches the eventual steady-state for the number of AIDS deaths per month. Hence, the total number of infected individuals remains constant from this point on.

Our original predicted graph shapes do match the model's computer graphs fairly well.

14. Summarize the lesson with the following key points:

- It is possible to build a fairly simple mathematical model for an epidemic.

Mathematical analysis of our SIR model shows that the maximum extent of the modeled epidemic (i.e., the maximum number of infected individuals) is given by a constant *K*, which depends on only three numbers: population size (*N*), infection rate (*ß*), and mortality rate (μ).

- Numerical analysis of our model shows that it produces intuitively correct data (i.e., graph shapes coincide with reality).

Additional Resources

Internet resources

Callahan, Jim. *The Spread of Disease*
http://www.math.smith.edu/Local/cicchap1/node2.html

The HIV Epidemic in Vancouver's Lower East Side
http://www.ugrad.math.ubc.ca/coursedoc/math100/notes/mordifeqs/hiv.html

Harrison, Brink. *What Does Math Have to Do With Getting Sick?*
http://pulse.pharmacy.arizona.edu/math/what_does.html

Keeling, Matt. *The Mathematics of Diseases*
http://plus.maths.org/issue14/features/diseases/

Link, Don. *Infectious Disease and Exponential Growth*
http://teacherweb.com/MD/IndianCreekSchool/LinkD/WQR2.stm

Handout 1

Data Table for an SIR Model

Our Current Model:

$$S(t+1) = S(t) - \beta S(t)I(t) + \mu I(t)$$
$$I(t+1) = I(t) + \beta S(t)I(t) - \mu I(t)$$
$$R(t+1) = R(t) + \mu I(t)$$

			0.00001037	0.01267	
Months from Now (t)	*S*	*I*	*newI ? ? SI*	*newR ? ? I*	*R*
0	**4,500**	**1,500**	**70**	**19**	**200**
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Handout 1 ▸ Answers

Our Current Model:

$$S(t+1) = S(t) - \beta S(t)I(t) + \mu I(t)$$
$$I(t+1) = I(t) + \beta S(t)I(t) - \mu I(t)$$
$$R(t+1) = R(t) + \mu I(t)$$

			0.00001037	0.01267	
Months from Now (t)	*S*	*I*	*newI ? ? SI*	*newR ? ? I*	*R*
0	4,500	1,500	70	19	200
1	4,449	1,551	72	20	219
2	4,397	1,603	73	20	239
3	4,344	1,656	75	21	259
4	4,290	1,710	76	22	280
5	4,236	1,764	77	22	302
6	4,181	1,819	79	23	324
7	4,125	1,875	80	24	347
8	4,069	1,931	81	24	371
9	4,012	1,988	83	25	395
10	3,954	2,046	84	26	420