

## CHAPTER 1

**Introduction**

Mathematical sociology is not an oxymoron. There is a useful role for mathematics in the study of society and groups. In fact, that role is growing as social scientists and others develop better and better tools for the study of complex systems. A number of trends are converging to make the application of mathematics to society increasingly productive.

First, more and more human systems are complex, in a sense to be described soon. World economies are more and more interconnected. Our transportation and communication systems are increasingly worldwide. Social networks are less local and more global, making them more complex, producing new emergent communication patterns, a positive effect, but which also has made us increasingly vulnerable to pandemics, a negative effect. The Internet has connected us in ways that no one understands completely. Power grids are less and less local and are subject to more widespread failures than ever before. New species are increasingly introduced into local complex ecologies with unexpected effects. Our recent climate change has produced a situation in which it is more and more important to predict the future and the effects of human interventions in the complex system of the global weather. The mapping of the human genome makes available to biologists the possibility of studying the complex system of interactions between genes and proteins. All of these tendencies mean that scientists in a wide variety of areas—computer science, economics, ecology, genetics, climatology, epidemiology, and others—have developed mathematical tools to study complex systems, and these tools are available to us sociologists.

The second important trend is the growing power and ubiquity of the computer. Computer simulations and mathematics are complementary tools for the study of complex systems. They are two different ways of drawing implications for the future from what is known or assumed to be true. Mathematics can be used to draw far-reaching and sometimes unexpected conclusions using logic and mathematics. For example, many properties of networks have been proved to be true by mathematicians using traditional mathematical tools. Computer simulations use computer programs the coding of which embodies assumptions and whose conclusions are evident after the program has iterated. Computer simulations are useful in situations that are unsolved or intractable mathematically.

This text uses both mathematics and computer simulations. Sometimes the computer simulations demonstrate phenomena for which there is no exact mathematical solution. More frequently simulations are used to illustrate some model so that you, the reader of this book, will gain some understanding of how the model works and how it is affected by varying parameters even if a full mathematical treatment of the model is beyond the purpose of this book.

## EPIDEMICS

At the time we are writing this chapter there is an epidemic of concern over swine fever, a variant of influenza that seems to have captured the public's attention. Both the flu and fear of this flu spread through social networks, and we want now to illustrate some of the properties of epidemics through a very simple model. The model will be illustrated both with a little simple mathematics and with a computer simulation.

Suppose that a large population consists of  $N$  individuals. Suppose that each individual in the population has small probability  $p$  of being connected to each of the others in the population and that his connection to one individual has no bearing on his connection to any other person in the population. This creates a *random network* among the members of the population. Real social networks are not like random networks, but random networks are very tractable mathematically, and so they tend to be assumed by epidemiologists who study the spread of diseases. The powerful conclusions may be relatively unaffected by the unreality of the assumptions, much like the statistician may on occasion assume a normal distribution because the conclusions are not affected very much if the assumption of a normal distribution is not exactly true. We will examine how real social networks differ from random networks in much more detail in later chapters, but for the moment we will assume that they are useful descriptions of real networks.

Despite its unrealism, let's, for the sake of convenience, assume that the network is random. Suppose that initially just one person is sick with a contagious disease. If  $p$  is the probability this person is connected to anyone else, then we should expect this person to be connected to  $p \times N$  others, on the average. If  $p$  is small and  $N$  is large, then each of these  $pN$  individuals will be connected to  $pN$  others, and so the sick individual will be connected to  $(pN)^2$  individuals indirectly, through his contacts. If  $p$  is small and  $N$  is large we can ignore the unlikely event that some individuals will be connected to more than one of his direct connections. Similarly  $(pN)^3$  persons will be connected even less directly, through two intermediaries, and  $(pN)^{k+1}$  persons will be connected through  $k$  intermediaries. Let's look at this sequence:

$$1 + pN + (pN)^2 + (pN)^3 + (pN)^4 + \dots \quad (1.1)$$

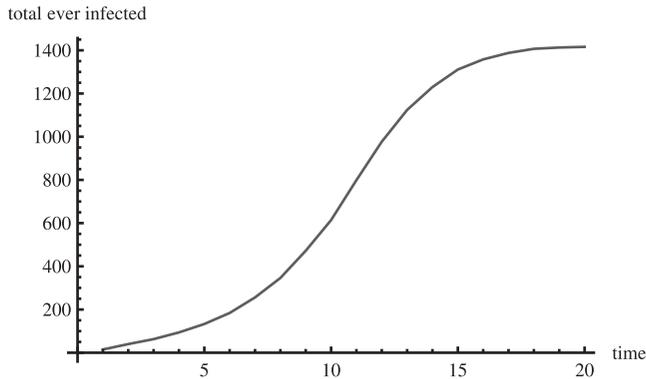


Figure 1.1. Total number of infected individuals when  $pN = 1.5$ .

This represents the expected or average number of people who get the contagious disease carried by one person. Of course, this number must in fact be limited because the population is of finite size  $N$ , but if  $pN \geq 1$ , this sum diverges: it just gets bigger and bigger, without limit. If  $pN < 1$ , however, the sum converges to a number, and this number may be quite small relative to  $N$ . The sequence in equation 1.1 converges to,

$$1 + pN + (pN)^2 + (pN)^3 + (pN)^4 + \dots = \frac{1}{1 - pN} \quad (1.2)$$

You can verify this for yourself by substituting a few values. When  $pN = .5$ , for example, the sums are 1, 1.5, 1.75, 1.875, and so on, getting closer and closer to 2, the limit. What this means is that if an infected person infects, on the average, less than one other person the disease will not become an epidemic affecting nearly everyone, but otherwise it will spread to the entire population.

The following figures were generated from a simulation based on a few simple assumptions. The network was of size 2,500. Ten individuals were initially infected. The probability that any two people were connected, also the probability that an infected person would infect a healthy person in any given time period, was set at .0006:  $p = .0006$ . The network was examined over 100 time periods. If infected in one time period, the person was assumed to be infectious at the next time period and immune thereafter. In this case  $p \times N = .0006 \times 2,500 = 1.5$ , and 1.5 is bigger than the critical value of 1.00 ties per person. We should expect the disease to spread.

Figure 1.1 shows the total number infected. After just 25 time periods most had been infected. The curve shows a familiar S-shaped figure, called the logistic curve. The disease spread slowly when few had it, then picked up speed, then slowed down as there were fewer and fewer who had not been infected and were not immune. Figure 1.2 shows the number of new cases in each time period, telling the same story as Figure 1.1 in a different way.

## 4 • Chapter 1

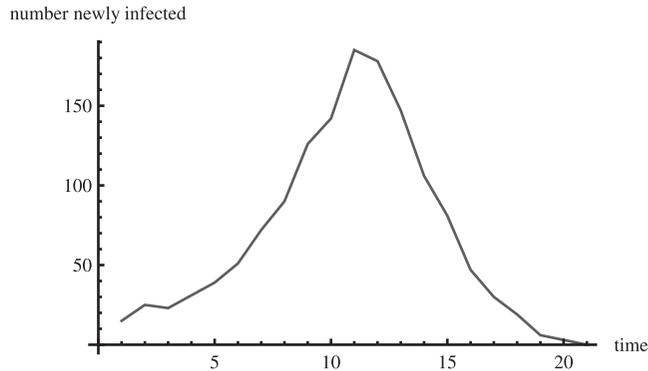


Figure 1.2. Number of newly infected individuals over time when  $pN = 1.5$ .

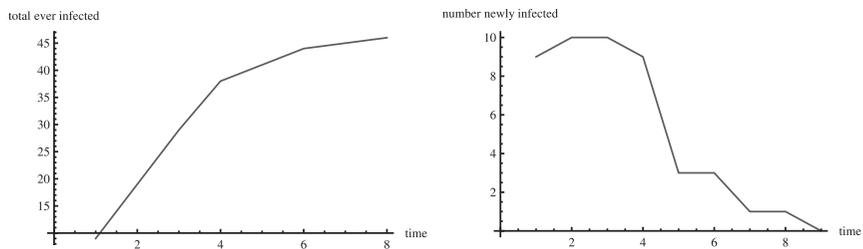


Figure 1.3. Number of total and newly infected individuals over time when  $pN = .75$ .

Now suppose we make the parameter  $p$  half as big,  $p = .0003$ , so that each individual averages .75 contacts instead of 1.50. This is below the critical value of one contact per person on average. We would expect the outbreak not to become an epidemic and to peter out after the initial set of infected individuals fail to reproduce themselves. Figure 1.3 shows that the total number never passes 50 and illustrates the declining number of new cases.

In later chapters we'll see how these inferences are a consequence of deep and nonobvious properties of random networks. We will also see how the inferences must be modified for other classes of nonrandom networks. Even with these qualifications, the results are interesting. First, they apply to phenomena other than the spread of disease. Information and rumor transition can also be modeled. Coleman et al. (1966) used this model to account for diffusion in the use of a new antibiotic among doctors in Midwestern communities. In this case, what was being spread was not disease but information and influence. The existence of a critical point has policy implications in epidemiology. It means that not everyone need be inoculated for an inoculation campaign to be effective. It also helps explain why Apple computers are less subject to viruses than are PCs. Since computer viruses are targeted for one operating system, any Mac

virus will spread only from Mac to Mac, while almost all the computers a Mac is connected to will be PCs. The Mac to Mac network will be very sparse while the PC to PC network will have a lot higher *density* (a higher value of  $p$ ). Thus, the lower frequency of Mac infections need not be due to any superiority of the Mac operating system but simply due to the fact that very few people use Macs.

The spread of a disease or information depends not only on the density of the network but also on the presence or absence of long-range connections in the network, and this topic can be examined both mathematically and through the use of computer simulations. When in history almost all ties between people were strictly local, epidemic diseases spread much more slowly. The “Black Death,” a plague that decimated Europe in the Middle Ages, was carried by sailors from Asia to Italian port cities in 1347, but it did not reach England until 1349 and Scandinavia until 1350. This slow spread occurred because long-distance movements hardly existed. Most people never saw anyone outside their own small village. Nowadays worldwide influenza pandemics occur every year.

## Simulated Epidemics

Let’s explore this difference using the demonstration *Simulated Epidemics*. This demonstration offers the possibility of examining contagion in two different types of networks, random networks and grids. In grid networks all ties are local, like in a farming community where farmers have relations only to those in neighboring farms; there are no long-distance ties. In a random network there are no constraints on the ties at all. In the demonstration, the number of individuals and connections are the same in a grid and random network: there are 400 individuals and the average number of connections is four. Play with the demonstration for a while and you see that there are two major easily observable differences. First, the diseases spread much more rapidly in the random network. Second, the shape of the curves for new cases is quite different. In the grid the number of new cases goes up in a straight line until the edges of the grid are reached. In the random network the number of cases seems to go up exponentially at the beginning.

Why are there these differences? In a grid the disease can expand only at the circumference of the infected area. It is only individuals at the border of the infected area who come in contact with uninfected individuals. On the other hand, in a random network each newly individual in the early stages is coming in contact with four uninfected individuals. In a random network we will have 1 infected individual, then 4, then each of these 4 will infect 4 others so that there will be 16 new cases, then each of these 16 cases will infect 4 more for 64 more cases, and so on. The number of new cases will not be proportional to time,  $t$ , but  $4^t$ .

## RESIDENTIAL SEGREGATION

Residential segregation by race and class has many causes. In the past some of it was legal: residential covenants prevented whites from selling their homes to nonwhites. Patterns established in this way persist. Some segregation is economic because race and income are correlated and neighborhoods with different priced homes may become segregated primarily by income but also incidentally by race. However, some segregation patterns result from individual choices: individuals may wish to avoid being a minority in their own community, or at least being a small minority. Such voluntary segregation based on the desire to be in proximity to similar others can produce segregation by sex or social class as well as race. In a lunchroom of young children or at a party among middle-aged adults, there may be segregation by gender. People of different social classes don't meet each other in unstructured situations, but our experience in public schools was that school social events and cafeteria lunch groups were segregated by class, sex, and race.

Thomas Schelling (1969, 1972), an economist, was the first to explore this phenomenon in simulations. Schelling devised what would now be called an agent-based model. In this simulation individual actors were placed in a two-dimensional grid, an eight by eight checkerboard. Actors were of two different types: "X" and "O" actors. Each actor could change his position, following certain simple rules. Each actor wanted a minimum percentage of her neighbors to be of the same type. On this grid each actor had up to eight neighbors: to the right, left, above, below, upper right, upper left, lower right, and lower left. Actors at the corners of the grid had only three neighbors, while actors on an edge but not a corner had five neighbors. Some squares of the grid would be empty, to permit movement.

Each actor followed a simple rule: he required that a certain minimum proportion of his neighbors be of his type. Suppose, for example, that desired minimum proportion were greater than one-third. Then if a person had just one neighbor, that actor would have to be of the same type; if two neighbors, at least one must be of the same type; if three, four, or five neighbors, at least two; and if six, seven, or eight neighbors, at least three. If an actor were dissatisfied she would move to the nearest position on the grid that was satisfactory in terms of the composition of the neighborhood.

The interesting thing about such situations is that they can lead to cascades of movement. If one actor moves he shifts the character of his previous neighborhood, increasing the prominence of the other type, so that others who were satisfied before his move may, as a consequence of his move, become dissatisfied with the composition of their neighborhood. These cascades are *emergent properties* of the simulation: there is usually no simple way of predicting the outcome. Schelling showed that even such simple rules could have unexpected outcomes.

Schelling presented this model in 1972, when computers were not as prominent in research as they are today. He used a small eight-by-eight grid

○	×	○	○	○		×	
○	○		×	○			×
×	○	×				×	
×	○	○	×	○	○	○	○
×	×	○	×		×		
○		○	×			×	○
	○		×	×		○	
×			○	×		×	○

Figure 1.4. Initial locations of actors in the Schelling simulation.

that he could manipulate by eye, instead of the much larger grid he could have used even with the computers of his day. However, there is another effect of his choice not to use computers. Using a checkerboard-sized grid and moving the pieces by eye emphasized his larger point—that simple assumptions, easily implemented, can produce unexpected outcomes.

Of course, there are outcomes that are not surprising. If neither type is willing to be a minority in its neighborhood, the result will be complete segregation of the pieces, each type occupying its own separate areas of the grid. But suppose that each actor wants her own type to constitute more than a third of her neighbors. Figure 1.4 shows an initial distribution with about one-third X type, one third O type, and one third empty squares.

O	X	O	O	O		X	
O	O	X	X	O			X
	O	X				X	
X	O	O	X	O	O	O	O
X	X	O	X		X		
O		O	X			X	O
	O		X	X		O	
X			O	X		X	O

Figure 1.5. Schelling grid after one move.

Squares in which one-third or less of the neighbors are of the same type (the dissatisfied actors) are shown in a smaller font.

In all, 27 actors are satisfied with their neighborhoods and 15 are not. On the average, 45% of anyone's neighbors are of the same color. Suppose that one of the unhappy actors moves. For example, the X actor in the third row and first column has five neighbors, only one of whom is a fellow X. A move to the square in the second row and third column will give him three X and three O neighbors, 50% with his type. Figure 1.5 shows the grid after his move.

As a result of this move, not only is he (the actor initially on (3,1) who moved to (2,3)) happier, but the actor in the third row and third column now also has more than one-third of his neighbors of his type.

○		○	○	○		×	
○	○	×	×	○		×	×
	○	×	×		○		
	○	○	×	×	○	○	○
		○	×	×	×		
○	○	○	×		×		○
	○	×	×	×		○	
	×			×	×		○

Figure 1.6. Schelling grid after sixteen moves.

Figure 1.6 shows the results after 16 moves. Note that all actors live in acceptable neighborhoods: no letters are in small fonts. The degree of segregation into primarily X and O regions is quite evident. One possible measure of segregation is the average proportion of neighbors who are of the same type. Figure 1.7 shows that this measure climbed regularly. By the sixteenth move the degree of segregation is almost three-quarters, even though the actors are programmed to require that more than one third of their neighbors be of their own color. This is one of the surprising aspects of this simulation—the high degree to which the actual segregation exceeds the minimum desired degree. This high proportion is clearly an emergent property of the simulation. As Schelling points out, no one may want to live in such a highly segregated society. Everyone may actually prefer more

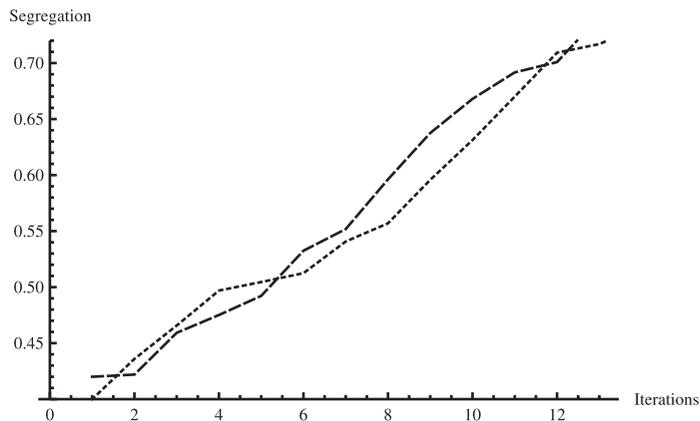


Figure 1.7. Segregation after sixteen moves.

integration, but a highly segregated society can emerge from independent actions by unwitting actors who do take the larger context into account.

### Schelling's Model of Residential Segregation

This demonstration enables you to examine Schelling's model and to vary its parameters. You can change the relative sizes of the majority and minority (represented by blue and green squares), the proportion of empty squares, and the preferences of the majority and minority for having neighbors of their own color. Some homework problems also use this demonstration. In this chapter we have tried to give you, the reader and teacher, a sense of what this book is like. No mathematical background beyond simple algebra is assumed. All additional mathematics will be introduced as needed. The tools are simple mathematical models and computer simulations of social processes. Computer simulations are used when the phenomenon itself is intractable mathematically and to give you the reader a sense of how a mathematical model develops and is sensitive to changes in its parameters. The emphasis is on conclusions that would not be obvious without the use of mathematics or simulations. These are the unexpected emergent properties of social systems of complexly interconnected individuals, and it is the study of these emergent properties that constitutes the bailiwick of sociology.

### Chapter Demonstrations

- *Schelling's Model of Residential Segregation* simulates the Schelling model of residential segregation.

- *Simulated Epidemics* simulates the spread of an epidemic in a random grid network. The density of contagious connections in the network can be varied.

## EXERCISES

1. The simulation *Simulated Epidemics* allows for variation in the size of a population living on a square grid (the size of the population is the squared length of a side of the grid) and the probability of a random connection between any two nodes, labeled “Density of Network.” Below the critical value of the density, the disease dies out and above that it becomes an epidemic. What are the corresponding critical values of the density for each grid? Determine these results theoretically from what the chapter says about critical values.
2. Verify these answers by showing, with *Simulated Epidemics*, that values 50% larger than the critical value of the density produce epidemics that spread to most of the population.
3. For a 50 by 50 grid, what would you expect to be the effect of network density on the time it takes for the epidemic to reach 80% of the population? Would increasing the density increase or decrease the time it takes for the epidemic to infect 80% of the population? Provide evidence for your conjecture by running the simulation with different densities.
4. Using the *Schelling* demonstration, create a community in which 50% of the squares are green, 25% are blue, and both types prefer that at least one-third of their neighbors are the same type. Move unhappy squares as long as you can (the simulation will stop when no square is dissatisfied with its neighborhood). Do this a few times. The segregation index is simply the average proportion of same-color neighbors. Although people are willing to live in neighborhoods with one-third of the other color, what is the average value of the index?
5. Now change the last simulation so that both types of squares want to live in neighborhoods in which their color is in the majority. What happens to the segregation index?