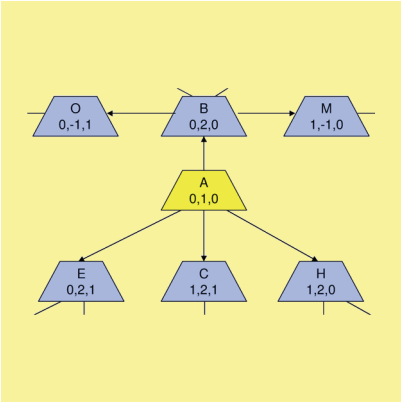
Chapter 19 / Case Study

Models of Infection: Person to Person



When faced with a spreading infection, public health workers would like to predict its path and severity in order to make decisions about vaccination strategies, quarantine policy, and use of public health resources. This is true whether dispersal of the pathogen is natural (as in SARS or the 1918 flu) or deliberate (for example, through terrorism). Effective mathematical models form a way to test the potential outcome of a public health policy and arrive at an effective response.

In this problem we focus on an overly simplified model of the spread of an infection and develop some tools that lend insight into its behavior.

To make our problem as easy as possible, we impose some rather artificial assumptions. Suppose that we have nm patients in a hospital ward and that their beds are arranged as n rows of m beds. Suppose that one of the patients, the one in the bed $\lceil m/2 \rceil$ in row $\lceil n/2 \rceil$, becomes infected and then can possibly infect patients in neighboring beds. How does this infection spread through the ward?

Insight through Monte Carlo Simulation

We'll need some parameters in our model. A patient, once infected, stays contagious for k days and then recovers, never to be infected again. During each day of infection, the probability that each susceptible neighbor (north, south, east or west) becomes infected is τ.

So there are three parts of the population to track. At day t, $I(t)$ is the proportion of the population that is infected and $S(t)$ is the proportion of the population that has never been infected. These quantities satisfy $0 \le I(t) \le 1$ and $0 \le S(t) \le 1$ for $t \ge 0$. The third part, $R(t)$, the proportion of the population that was once infected but has now recovered, can be derived from these two: $R(t) = 1 - I(t) - S(t)$.

We study this model by running a simulation. Each patient is in one of $k+2$ states: the patient has recovered from an infection, is susceptible to infection, or is in the ith day ($i = 1, \ldots, k$) of the k-day infection. It is convenient to use the integer values -1, 0, and

$1,\ldots,k$ to represent these different states. Each day, we update the status of each infected patient by incrementing the state of that patient (resetting it to -1 after k days), and for each susceptible neighbor, we generate a random number between 0 and 1; if that number is less than τ, then the neighbor's state is changed from 0 to 1, indicating infection. We continue this process until there are no infected patients; at that point, our model allows no possibility of any additional infections, so the epidemic ends.

Let's see how this model behaves.

CHALLENGE 19.1.

Run this model for $m = n = 10$, $k = 4$, $\tau = 0.2$ until there are no infected patients. Plot $I(t)$, $S(t)$, and $R(t)$ in a single graph.

If possible, display the epidemic as a movie. To do this, form a matrix of dimension $m \times n$, where the value of each entry corresponds to the state of the corresponding patient on a particular day. Using the `movie` command, we can display these matrices in sequence, day after day.

Since the model is stochastic, if we run it 10 times, we might get 10 different results, possibly ranging from no infections other than the original patient to infection of every patient, although these are both very low probability events. We need to investigate the variation in results, but first we'll add two complications.

The patients in our original model are immobile and only contact their four nearest neighbors. In most situations, population members move in more arbitrary ways. For example, epidemics jump from continent to continent by air or ship travel. In our hospital ward, let's assume that the nursing staff sometimes moves patients to other beds. For definiteness, we'll assume that each patient initiates a swap with probability δ. Then, at each time and for each patient we need to decide whether that patient initiates a swap. If so, we'll choose the bed indices for the second patient randomly, as $\lfloor r_2 n + 1 \rfloor, \lfloor r_2 m + 1 \rfloor$, where r_1 and r_2 are random samples from a uniform distribution on $[0, 1]$.

CHALLENGE 19.2.

Modify your model to include mobility and run it for $\delta = 0.01$ until there are no infected patients. Display the results as in Challenge 19.1.

There are two major tools used to slow the spread of epidemics: quarantine, to isolate infected individuals, and vaccination, to protect susceptible individuals. To reduce the infections in our hospital model, infected individuals should be moved to a corner of the ward, with recovered individuals separating them from susceptible ones whenever possible. You can experiment with this quarantine strategy, but in the next problem we turn our attention to the use of vaccinations. For convenience, you might use the value -2 to indicate a vaccinated patient.

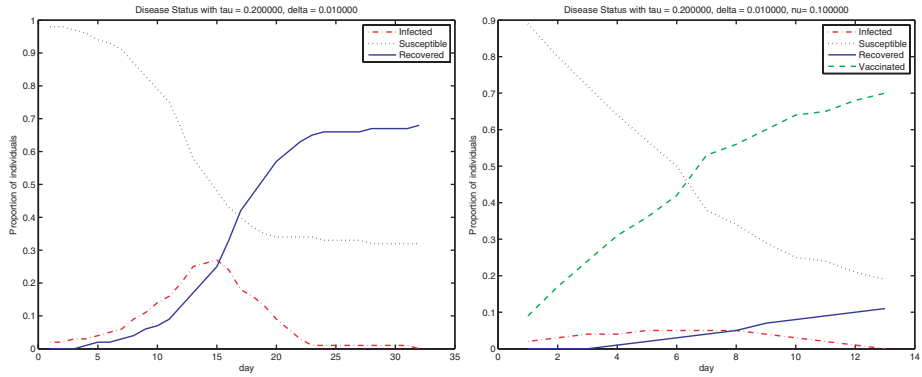


Figure 19.1. *Proportion of individuals (in one simulation) infected by day in a* 10×10 *grid of hospital beds, with infection rate* $\tau = 0.2$ *and mobility rate* $\delta = 0.01$ *and no vaccination (left) and with vaccination rate* $\nu = 0.1$ *(right).*

CHALLENGE 19.3.

Suppose that each day each susceptible individual has a probability ν of being vaccinated. Rerun your model with $\nu = 0.1$ until there are no infected patients. Display the results as in Challenge 19.1 and compare the results of the three models.

Some results from Challenges 19.2 and 19.3 are shown in Figure 19.1. Now we need to see how much variation is possible in the results if we run the model multiple times.

CHALLENGE 19.4.

Run the model of Challenge 19.3 1000 times, recording the number of individuals who become infected in each run. (Note that this is equal to the number of recovered individuals when the run is terminated.) Plot this data as a histogram (as in Figure 19.2), and compute the mean proportion of recovered individuals and the variance in this number. Try several different values of ν to see whether the variance changes.

From the results of Challenge 19.3, we can see that vaccinations can contain the spread of the epidemic. In Challenge 19.5, we take the role of a public health official trying to limit the spread of the epidemic.

CHALLENGE 19.5.

Develop a vaccination strategy that, on average, limits the epidemic to 20% of the population. Do this by using a nonlinear equation solver to solve the problem $R(\nu) - .2 = 0$, where $R(\nu)$ is the mean proportion of recovered individuals when we use a vaccination rate

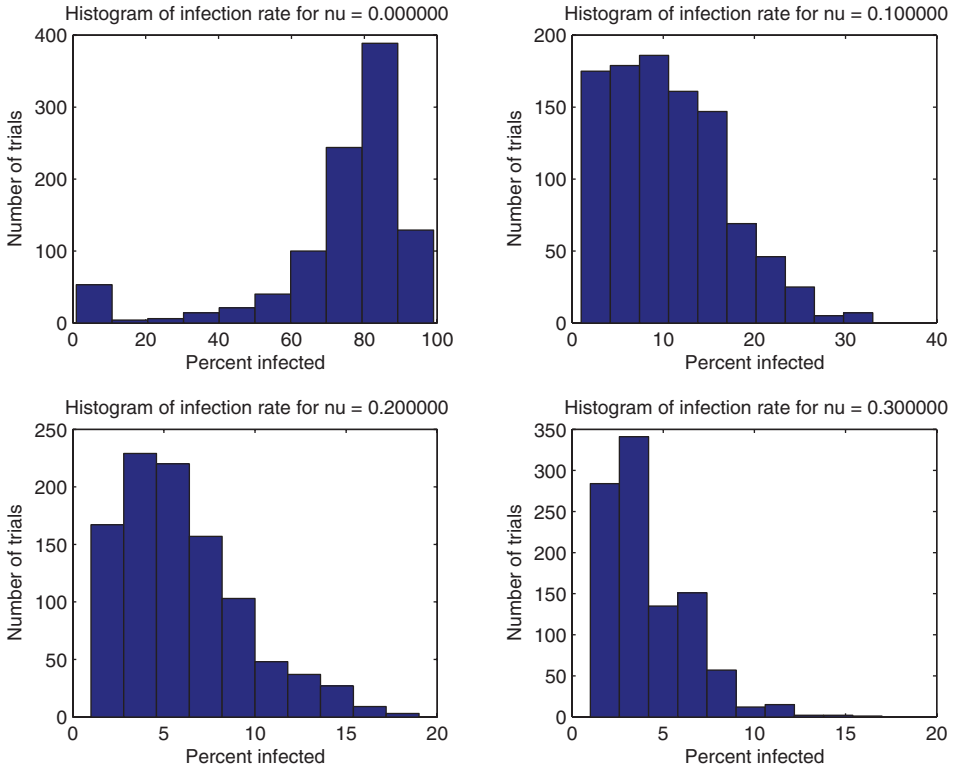


Figure 19.2. *Results of* 1000 *trials for a* 10×10 *grid of hospital beds, with infection rate* $\tau = 0.2$ *and vaccination rate* ν*, with* ν *varying.*

of ν. For each value of ν presented by the solver, you need to get a reliable estimate of R by running the model multiple times. Use the variance estimates from Challenge 19.4 to determine how many runs to use, and justify your choice.

A Markov Chain Model

The model we have developed has the **Markov property**: the status of each individual depends only on the status of the population on the previous day, and not on any older history. In fact, the system is a **Markov chain**. The **states** in the chain correspond to the possible statuses of the population; each state can be labeled $(d_1, \ldots, d_p)$, where there are p beds and d_i ranges from -2 to k, indicating that individual i $(i = 1, \ldots, p)$ is vaccinated, recovered, susceptible, or in day j $(1 \leq j \leq k)$ of the infection. There is an edge from one state to a second state if it is possible for the population to move from the first state to the second state the next day, and the weight on the edge is the probability of this happening. Figure 19.3 illustrates a Markov chain corresponding to three individuals in a single row of beds, with the middle one initially infected, a disease duration of $k = 2$ days, and no

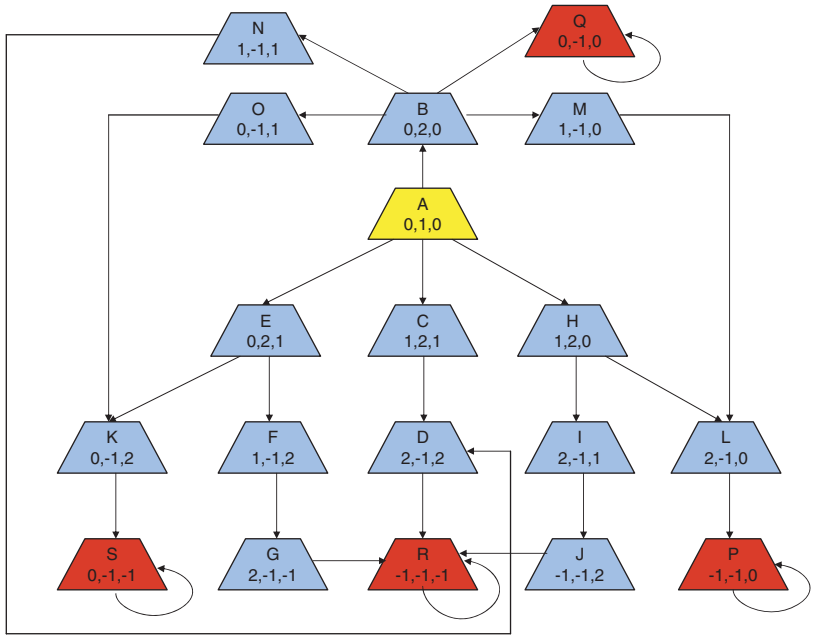


Figure 19.3. *A Markov chain that models three patients in a row of beds, with the middle patient infected and able to possibly infect two neighbors. The triple of numbers for each state gives the status of each of the three patients. The yellow state A is the state in which we start, and the red states are the possible outcomes when the infection runs its course, corresponding to* 1 *(state Q),* 2 *(states S and P), or* 3 *(state R) patients eventually infected.*

vaccination. (You will determine the edge weights in Challenge 19.6.) For this model, we are interested in three probabilities: the probability of terminating in

- state Q, corresponding to 33% of the population becoming infected,
- state R, corresponding to 100%,
- state P or state S, with the infection contained to 67% of the population.

CHALLENGE 19.6.

(a) Construct the **transition matrix** $\boldsymbol{A}$ corresponding to this Markov chain: element $a_{i,j}$ is the probability of transitioning to the ith state from the jth state and is therefore equal to the weight on that edge of the Markov chain in Figure 19.3.

(b) Let $\boldsymbol{e}_1$ be the first column of the identity matrix. If we begin in day 1 in the first state, then the vector $\boldsymbol{A}\boldsymbol{e}_1$ tells us the probabilities of being in each of the states on day 2. Prove this.

(c) Similarly, A^2e_1 gives the probabilities for day 3. For efficiency, this should be computed as $A(Ae_1)$ rather than as $(A^2)e_1$. Explain why, by doing the operations counts.

(d) If we compute $z = A^j e_1$ for a large enough j, we have the (exact) probabilities of being in each state after the epidemic passes. Use this fact to compute the probabilities of having 1, 2, or 3 infected individuals, and compare these probabilities with the results of a Monte Carlo experiment performed as in Challenge 19.4 but using 3 individuals. How many Monte Carlo simulations does it take to get 2 digits of accuracy in the probabilities?

(e) In this simple problem, you can determine the three probabilities directly from Figure 19.3, by determining the probability of a transition from state A to states P, Q, R, and S. Show how these probabilities can be derived, giving the same answer as obtained by the Markov chain computation above.

In the preceding challenge, we estimated the probabilities of being in each state of a Markov chain at **steady state**, after a long time has passed. These estimates were formed by computing $A^j e_1$, where A is the transition matrix. The matrix A has several interesting properties, and it is worthwhile to take the time to verify them to see the relation between the steady state probabilities and the eigenproblem for A.

CHALLENGE 19.7.

Verify that the matrix A has the following properties:

- Its elements are all nonnegative, since they represent probabilities.
- The sum of the elements in every column is 1, since the transition probabilities out of a given state must sum to 1. We can write this as $e^T A = e^T$, where e is the vector with each component equal to 1.
- Therefore, e is a left eigenvector of the matrix A corresponding to eigenvalue $\lambda = 1$. So we can find the vector of steady state probabilities for a Markov chain by using any algorithm for solving the matrix eigenvalue problem. Any right-eigenvector z corresponding to $\lambda = 1$ is a steady state vector, since it satisfies $Az = z$.
- We can use the Gerschgorin theorem (See Pointer 5.4) to verify that no eigenvalue of A can be outside the unit circle.
- (More difficult) It turns out that for this problem, A has four eigenvalues equal to 1, and no other eigenvalue of A has magnitude 1. Using this fact, we can verify that the limit of the sequence $A^k e_1$ as $k \to \infty$ is a right eigenvector corresponding to the eigenvalue $\lambda = 1$. (You may assume that the matrix has a complete set of right eigenvectors $u_1, \ldots, u_n$ that form a basis for the space of $n \times 1$ vectors. Express e_1 as $\alpha_1 u_1 + \cdots + \alpha_n u_n$, where $\alpha_1, \ldots, \alpha_n$ are some coefficients, and then compute $A^k e_1$.)

Our Markov chain has an enormous number of states ($(k+3)^p$ for p patients), but many of these states provide more detail than we might need. For example, in Figure 19.3,

POINTER 19.1. Further Reading.

The 1918 flu epidemic killed more than 20 million people. Some investigators believe that it may have first taken hold in a US army base, but the disaster was worldwide, with millions killed in India alone. Travel of soldiers in World War I aided the spread of the infection. A book by Gina Kolata chronicles these events [94].

When doing Monte Carlo experiments, it is wise to use a high-quality pseudorandom number generator in order to get valid results. The classic reference for understanding such programs is Donald Knuth's book [92].

To go beyond the simple models investigated in this chapter, see, for example, the books by Britton [21] and by Hoppensteadt and Peskin [80] or the article by Zhuge [155].

Learn about Markov chain models and computing in the book by Stewart [142].

There are many approaches to aggregation of Markov chains. One starting point is an article by Marek [105].

---

if we are in state C, we always make the transition to state D, so these two states can be combined or **aggregated** without loss of information about infection totals. More importantly, states P and S represent different outcomes, but they are equivalent to us, since in each case 67% of the population becomes infected.

By aggregating states, we can reduce the size of the problem. Sometimes this can be done analytically, but when the model is too complicated (for instance, once mobility is added), we can do it by simulation, gathering data to determine the probability of transitions between aggregated states.

These simple Markov models can yield some insight into epidemics, but we have seen that the work of the Monte Carlo experiment, or the number of states in the original Markov chain, quickly grows with the size of the population. In the case study of Chapter 21 we investigate an alternative set of models.

Meanwhile, you might want to modify the models to explore more realistic variations. You also might consider how to model related systems, such as spread of a fungus in a tree farm, spread of contamination in a set of chicken coops, or spread of disease in a dormitory when the students also interact at school.