2016

Mathematical Modeling of Infectious Diseases with Latency: Homogeneous Mixing and Contact Network

Keith Carlson
University of Central Florida

MATHEMATICAL MODELING OF INFECTIOUS DISEASES WITH LATENCY: HOMOGENEOUS MIXING AND CONTACT NETWORK

by

KEITH A. CARLSON
M.S. University of Illinois at Urbana-Champaign, 1992
B.A. University of Missouri at Kansas City, 1983
B.A. University of Missouri at Kansas City, 1978

A thesis submitted in partial fulfilment of the requirements
for the degree of Master of Science
in the Department of Modeling and Simulation
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2016

ABSTRACT

In mathematical epidemiology, the standard compartmental models assume homogeneous mixing in the host population, in contrast to the disease spread process over a real host contact network. One approach to incorporating heterogeneous mixing is to consider the population to be a network of individuals whose contacts follow a given probability distribution. In this thesis we investigate in analogy both homogeneous mixing and contact network models for infectious diseases that admit latency periods, such as dengue fever, Ebola, and HIV. We consider the mathematics of the compartmental model as well as the network model, including the dynamics of their equations from the beginning of disease outbreak until the disease dies out. After considering the mathematical models we perform software simulations of the disease models. We consider epidemic simulations of the network model for three different values of $\mathcal{R}_0$ and compare the peak infection numbers and times as well as disease outbreak sizes and durations. We examine averages of these numbers for one thousand simulation runs for three values of $\mathcal{R}_0$. Finally we summarize results and consider avenues for further investigation.

ACKNOWLEDGMENTS

Thanks go to Dr. Zhisheng Shuai for his guidance and patience.

Thanks go to Dr. Ram Mohapatra for his advice and wisdom.

Thanks go to Mr. Donald Porchia for cajoling me to get into this.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The presence of communicable disease has always been an important part of life in human societies. Some diseases are endemic in many parts of the world, always present in the population, like typhus, malaria, or cholera. Other diseases such as SARS, influenza, and AIDS can spread from epidemics, operating on a relatively short time scale as they spread through the population. Mathematical epidemiology, the mathematical modeling of disease spread in a population, has a long history and many early developments in the mathematical modeling of diseases are due to public health physicians [1]. For example, Daniel Bernoulli, trained as a physician and a member of a famous family of mathematicians, attempted to model smallpox in 1760 while defending the practice of inoculation. Mathematical epidemiology can aid medical professionals who are trying to manage diseases. During a disease outbreak ethical considerations make it impossible for medical experiements to be performed on a population which would compare the effects of different disease management strategies. Therefore predictions based upon mathematical models may be essential in addressing the impact of communicable diseases and formulating strategies for fighting them.

Much of the basic theory of mathematical epidemiology was developed between 1900 and 1935 [1]. In the 1927 paper of Kermack and McKendrick [5], the authors partitioned a hypothetical population into sets (compartments) of people, including those in the population susceptible to disease infection (but not yet infected); those exposed and in a latent period who are infected but not yet infectious; those who are already infectious; and those who are removed from consideration (because they have recovered and are now immune, or because they have died) [4]. The numbers of people who are susceptible, exposed, infectious, or removed at time t are denoted by $S(t)$, $E(t)$, $I(t)$, and $R(t)$, respectively, and the total population $N = S + E + I + R$. This SEIR model assumes constant rates of mass-action incidence and disease recovery, and mean periods

of infectiousness and latency. There is also a constant $\mathcal{R}_0$, the basic reproduction number, whose value indicates whether the disease dies out ($\mathcal{R}_0 < 1$) or becomes an epidemic ($\mathcal{R}_0 > 1$). This Kermack-McKendrick epidemic model is a mass action (MA) model that assumes contacts are uniform between people in a population, similar to the mixing of chemicals. The model exhibits two flaws: it neglects social heterogeneity, that is, it neglects variation in contact rates between the susceptible and the infected; and it neglects partnership duration by assuming that all partnerships between people are only momentary in duration.

Since the year 2000 attention has turned to the usage of network models to allow for social heterogeneity, pioneered by Volz and Miller [8], [12], [16]. The objective in any of these so-called edge-based configuration models (EBCM) is, as always, to be able to calculate the numbers of susceptible, infectious, and recovered in the population. The EBCM approach begins with consideration of a randomly chosen test node u, and asks whether u is susceptible, infectious, or recovered. In the EBCM context, $S(t)$, $I(t)$, and $R(t)$ represent proportions of the population that are susceptible, infectious, or recovered, respectively, at time t (with $S(t) + I(t) + R(t) = 1$ since here they are proportions). The essential characteristic of the EBCM approach is the incorporation of the idea that test node u is susceptible as long as none of its neighbors (or partners) has ever transmitted infection to u [12]. This requires consideration of how many neighbors u might have as well as the possiblities that a randomly chosen neighbor v of u could itself be susceptible, infectious, or recovered.

Although the Miller-Volz model addresses social heterogeneity it does not consider exposed but non-infectious individuals. In some infectious diseases there is a period of time in which susceptible individuals have been exposed to disease but have not yet developed symptoms and are unable to transmit infection to anyone else [1]. We add to the model the proportion of exposed but not yet infectious individuals, $E(t)$. We then investigate this model. We derive a system of ordinary differential equations and determine and classify equilibria of the system. Finally we study its

long-term behaviour for various values of $\mathcal{R}_0$. We perform simulations and compare their results with the mathematical model.

In this thesis we investigate both homogeneous mixing and contact network models for disease epidemics with latency. In Chapter 2 the classic SEIR compartmental model with homogeneous mixing is revisited and the basic reproduction number $\mathcal{R}_0$ determines the final disease outbreak size. In Chapter 3 we construct and analyze an edge-based configuration model that includes an exposed class, and $\mathcal{R}_0$ is proven to serve as the threshold value for disease outbreak. In Chapter 4 we conduct numerical simulations to analyze and compare behavior of the system of equations from Chapter 2. In Chapter 5 we consider conclusions to be drawn from our investigations as well as ideas for future work.

CHAPTER 2: HOMOGENEOUS MIXING COMPARTMENTAL MODEL

2.1 Model Formulation

The disease model we are going to investigate is constructed within a static population, i.e., a population in which it is assumed that the population does not change from travel or births or deaths except possibly from deaths of people who are infected with the disease. In our model we are going to assume that the epidemic process is deterministic, that is, that population behavior is completely determined by its history and by the rules that describe the model [1]. A disease epidemic acts on a short time scale; an epidemic is a sudden outbreak of a disease that infects a substantial portion of a population before it disappears. Usually epidemics leave many members of the population untouched. Our population is divided into four sets (compartments) of people: those who are susceptible to disease infection but not yet infected; those who have been exposed, i.e., infected but not yet infectious; those who are infectious; and those who have been removed from being susceptible or infected by recovering and acquiring immunity to the disease. Therefore the population is made up of N people and $N = S + E + I + R$, where we define

$$S = S(t) = \text{number of individuals in the population susceptible at time } t,$$

$$E = E(t) = \text{number of individuals exposed (but not infectious) at time } t,$$

$$I = I(t) = \text{number of individuals infectious at time } t, \text{and}$$

$$R = R(t) = \text{number of individuals recovered at time } t.$$

The numbers in each compartment are integers, but if N is sufficiently large then S, E, I, and R can be treated in our model as continuous variables.

Denote the disease transmission rate (per infectious person) by β. This model assumes mass action

incidence, that an average member of the population makes contacts sufficient to transmit infection with βN others per unit time. For many infectious diseases there exists an exposed or latent period after transmission of infection from susceptibles to potentially infectious members of the population but before these potentially infectious individuals can transmit infection. Denote by ϵ the rate of transfer from exposure or latency to infectiousness; note that $1/\epsilon$ is then the mean period of exposure or latency. Denote by γ the recovery rate (so that $1/\gamma$ is the mean period of infectiousness).

If we assume that initially the entire population is susceptible, this is the same as writing $S(0) = N$. Given a susceptible individual, the probability that that person would have contact with an infectious individual is I/N since I/N is the fraction of the population that is infectious; so the rate of new infectious people per susceptible is $(\beta N)(I/N)$. Thus the rate of new infectious over the entire population is

$$(\beta N)\frac{I}{N}S = \beta IS. \tag{2.1}$$

The flowchart in Figure 2.1 shows the scheme of progression in the model from one compartment to the next.

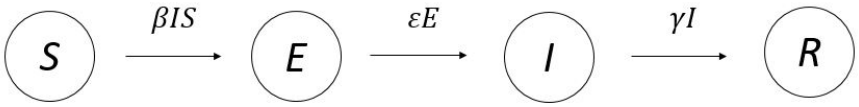


Figure 2.1: Flow chart for the SEIR model.

The model is determined by the system of ordinary differential equations (primes denote differen-

tiation with respect to time t):

$$S' = -\beta IS \tag{2.2}$$

$$E' = \beta IS - \epsilon E \tag{2.3}$$

$$I' = \epsilon E - \gamma I \tag{2.4}$$

$$R' = \gamma I. \tag{2.5}$$

The model is stated in terms of derivatives of the sizes of the different compartments, which assumes that the number of members of a compartment is a differentiable function of time. This is a reasonable assumption if there are many members of a compartment.

2.2 Model Analysis

We assume mass action incidence which implies that an average member of the population makes contacts sufficient to transmit infection with βN others per unit time, so that in the mean infectiousness period $1/\gamma$ a newly introduced infectious person could theoretically infect a total of $\beta N/\gamma$ individuals. Denote

$$\mathcal{R}_0 = \frac{\beta N}{\gamma} \tag{2.6}$$

the basic reproduction number. If we add together the equations (2.3) and (2.4) we get

$$\begin{aligned} E' + I' = (E+I)' &= \beta IS - \epsilon E + \epsilon E - \gamma I = \beta IS - \gamma I \\ &= (\beta S - \gamma) I. \end{aligned} \tag{2.7}$$

$E(t) + I(t)$ can increase only if $E' + I' > 0$. By (2.7) and because $I(t) > 0$, we conclude that $E' + I' > 0$ as long as

$$\beta S - \gamma > 0$$

or

$$\frac{\beta S}{\gamma} > 1. \tag{2.8}$$

With $S(0) = N$ this becomes

$$\frac{\beta N}{\gamma} > 1$$

or $\mathcal{R}_0 > 1$ by (2.6). Thus we have an epidemic if $\mathcal{R}_0 > 1$ and if $\mathcal{R}_0 < 1$ then $E(t) + I(t)$ decreases and the disease dies out.

It is desirable to be able to solve the system (2.2)–(2.5) and obtain an expression for $I(t)$. An expression for $I(t)$ was found in [5], but it did not model the number of infectious well when t was not close to zero [2]. An alternative is to add the equations for E' and I', then divide by the equation for S'.

$$\frac{d(E+I)}{dS} = \frac{\beta IS - \epsilon E + \epsilon E - \gamma I}{-\beta IS} = -1 + \frac{\gamma}{\beta S}$$

and multiplying through by dS gives

$$d(E+I) = -dS + \frac{\gamma dS}{\beta S}. \tag{2.9}$$

We next integrate both sides of this equation from 0 to t. Then we have

$$E(t) - E(0) + I(t) - I(0) = -(S(t) - S(0)) + \frac{\gamma}{\beta}(\ln S(t) - \ln S(0))$$

and since $E(0) = 0$ and $I(0) \approx 0$ this is equivalent to

$$E(t) + I(t) = N - S(t) + \frac{\gamma}{\beta} \ln \frac{S(t)}{N}. \tag{2.10}$$

Denote $\lim_{t\to\infty} S(t)$ by $S(\infty)$ and similarly for $\lim_{t\to\infty} E(t)$ and $\lim_{t\to\infty} I(t)$ by $E(\infty)$ and $I(\infty)$.

Lemma. The quantities $E(\infty) = 0$ and $I(\infty) = 0$.

Proof. We have already seen in (2.7) that $E + I$ is increasing when $S > \frac{\gamma}{\beta}$. By (2.7) when $S < \frac{\gamma}{\beta}$ then $E'+I' < 0$. Thus when $S = \frac{\gamma}{\beta}$, $E'+I' = 0$ and by the first derivative test $E+I$ has a maximum value there. $E + I$ is positive valued and thus is bounded below by zero, so $E(\infty) + I(\infty) = 0$ which implies $E(\infty) = 0$ and $I(\infty) = 0$.

Proposition. Let $\mathcal{R}_0$ be defined as in (2.6). Then $\ln \frac{S(\infty)}{N} = -\mathcal{R}_0(1 - \frac{S(\infty)}{N})$.

Proof. In equation (2.10) let t go to infinity:

$$E(\infty) + I(\infty) = N - S(\infty) + \frac{\gamma}{\beta} \ln \frac{S(\infty)}{N}$$

where the left side is zero by the lemma, so that

$$\begin{aligned} -\frac{\gamma}{\beta} \ln \frac{S(\infty)}{N} &= N - S(\infty) \\ &= N\Big(1 - \frac{S(\infty)}{N}\Big). \end{aligned}$$

Multiply both sides by $-\frac{\beta}{\gamma}$, then apply (2.6) to get

$$\begin{aligned} \ln \frac{S(\infty)}{N} &= -\frac{\beta}{\gamma} N\Big(1 - \frac{S(\infty)}{N}\Big) \\ &= -\mathcal{R}_0\Big(1 - \frac{S(\infty)}{N}\Big). \end{aligned} \tag{2.11}$$

CHAPTER 3: NETWORK EDGE-BASED COMPARTMENTAL MODEL

3.1 Model Formulation

The configuration model we are going to discuss here is a static network, i.e., a network where it is assumed that the population does not change except possibly from deaths of people who are infected with the disease. Thus we assume there are no changes in population as a result of births, non-disease deaths, or travel. In this model each individual in the population is represented by a node, and contact or association between nodes is modeled by edges. We assume that edges connecting nodes to other nodes are capable of transmitting disease between individuals. (We also assume no loops or multiple edges between nodes.) In the discussion to follow we often will consider a "test" node u and (synonymously) its neighbors or partners. In our model each node has degree $k \geq 0$ with probability $P(k)$, so any node u is assigned k_u stubs (half-edges) with probability $P(k_u)$. Denote by $\{P(k)\}_{k=0}^{\infty} = \{P(k)\}_{k=0}^{N-1}$ the probability distribution of node degrees in the network. Note that $P(k) \cdot N$ is the number of degree-k nodes in a population of size N.

The probability that a stub of node u connects to some stub of node v is proportional to k_v since the more stubs v has, the more stubs are available to connect to u. Thus

$$\begin{aligned} P_n(k) &= P(\text{a randomly selected neighbor } v \text{ of } u \text{ has degree } k) \\ &= (\text{total number of degree-}k \text{ network stubs}) \div (\text{total number of network stubs}) \\ &= \frac{kP(k)N}{\langle K \rangle N}, \text{ where } \langle K \rangle = \sum_k kP(k), \\ &= \frac{kP(k)}{\langle K \rangle}. \end{aligned} \tag{3.1}$$

Define

$$S = S(t) = \text{ proportion of population susceptible at time } t,$$

$$E = E(t) = \text{ proportion of population exposed (but not infectious) at time } t,$$

$$I = I(t) = \text{ proportion of population infectious at time } t, \text{ and}$$

$$R = R(t) = \text{ proportion of population recovered at time } t.$$

If we choose u randomly from our population then at a given time t, u will be susceptible, exposed, infectious, or recovered with probabilities according to the proportions $S(t)$, $E(t)$, $I(t)$, or $R(t)$, respectively.

There is an issue regarding u and its neighbors that needs to be addressed. When we assume a test node u is susceptible, exposed, infectious, or recovered with probability equal to the proportion of susceptible, exposed, infectious, or recovered in the population, respectively, we are implicitly assuming that the epidemic size grows deterministically. While we focus on the test node u, we make a simplifying assumption that the neighbors of u are acting independently [12]. Without making this assumption, we could have the situation that if neighbor w infects u, then u can infect another neighbor v, implying that the statuses of w and v are not independent. To see that this is not a concern we show now that ignoring transmissions from u to any neighbors will not affect the probability that u is in any state. Suppose we assume that u never transmits infection to any of its neighbors. This has no impact on the situation until after u is infected, so it does not affect the probability that u is susceptible or exposed. Once u is infectious, it may affect its neighbors. Whether it does or not does not affect the duration of infection of u; therefore it does not alter the probability that u is infected or recovered. Thus, preventing u from causing infection does not affect our calculations for S, E, I, or R based on the status of u.

Define $\theta = \theta(t)$ to be the probability that a randomly chosen neighbor v of u has not yet transmitted

infection to u. Initially, i.e., for $t = 0$, $\theta \approx 1$. For large networks, we have seen that we can assume neighbors of u are independent. If u has degree k then at time t, u is susceptible (but not yet infected) with probability $\theta(t)^k$. Also $S(t)$ can be thought of as the probability that none of u's neighbors have yet transmitted infection to u. Therefore, $S(t) = \sum_k P(k)\theta^k$. If we next define a probability generating function $\psi(x)$ whose coefficients are the discrete probability distribution of node degrees in the network, $\psi(x) = \sum_k P(k)x^k$, then we can write $S = \psi(\theta)$.

The probability θ can be expressed as the sum of mutually exclusive probabilities,

$$\theta = \phi_S + \phi_E + \phi_I + \phi_R, \tag{3.2}$$

where, assuming a chosen u with neighbor v,

$$\begin{aligned}
\phi_S &= P(v \text{ is susceptible but has not transmitted disease to } u),\\
\phi_E &= P(v \text{ is exposed but has not transmitted disease to } u),\\
\phi_I &= P(v \text{ is infectious but has not transmitted disease to } u), \text{and}\\
\phi_R &= P(v \text{ has recovered but did not transmit disease to } u).
\end{aligned}$$

We can construct a probability flux diagram that describes transitions between the four probabilities that comprise θ. For the disease in question we use the letters β, γ, and ϵ to denote respectively the rates of infection, recovery, and transition from the exposed/latent class to the infectious class. Using the probability generating function $\psi(\theta)$ we can represent ϕ_S in terms of θ. It will be useful to note that if $\psi(x) = \sum_k P(k)x^k$ then $\psi'(x) = \sum_k kP(k)x^{k-1}$ and $\psi'(1) = \langle K \rangle =$ the average degree of nodes in the network. Since $P_n(k) = \frac{kP(k)}{\langle K \rangle}$, we have

$$\phi_S = \sum_k P_n(k)\theta^{k-1} = \sum_k \frac{kP(k)}{\langle K \rangle}\theta^{k-1} = \frac{\psi'(\theta)}{\psi'(1)}. \tag{3.3}$$

We now have the probability flux diagram:

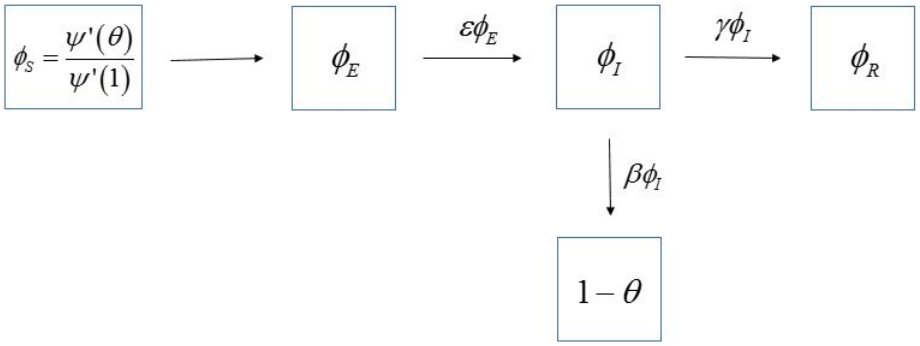


Figure 3.1: Flow/probability flux diagram for a static configuration model network.

This diagram can be used to derive a system of two coupled ordinary differential equations which can be expressed in terms of two functions, ϕ_I and θ.

First note from the vertical arrow in the probability flux diagram that

$$\theta' = -\beta\phi_I, \tag{3.4}$$

so $(1-\theta)' = \beta\phi_I$. Integrating with respect to time, from zero to t gives $(1-\theta(t)) - (1-\theta(0)) = \beta\int_0^t \phi_I d\tau$, or $1-\theta(t) = \beta\int_0^t \phi_I d\tau$. Here we have used information about initial conditions: $\theta \approx 1$ and $\phi_S = 1$, so $\phi_E + \phi_I + \phi_R \approx 0$ by equation (3.2) and thus $\phi_E = 0$, $\phi_I \approx 0$, and $\phi_R = 0$.

Again from Figure 3.1,

$$\phi_R' = \gamma\phi_I, \tag{3.5}$$

so again integrating with respect to time from 0 to t and using the initial condition $\phi_R(0) = 0$ we

get $\phi_R(t) = \gamma \int_0^t \phi_I d\tau$. Putting this together with the previous result we have

$$\frac{\phi_R(t)}{1-\theta(t)} = \frac{\gamma \int_0^t \phi_I d\tau}{\beta \int_0^t \phi_I d\tau} = \frac{\gamma}{\beta},$$

so that

$$\phi_R = \frac{\gamma}{\beta}(1-\theta). \tag{3.6}$$

Note that alternatively, we can derive ϕ_R as follows:

$$\frac{d\theta}{d\phi_R} = -\frac{\beta}{\gamma}\int d\theta = -\frac{\beta}{\gamma}\int d\phi_R.$$

Integrating from 0 to t, and using the initial conditions,

$$\theta(t) - \theta(0) = -\frac{\beta}{\gamma}\phi_R(t) + \frac{\beta}{\gamma}\phi_R(0),$$

or

$$\theta(t) - 1 = -\frac{\beta}{\gamma}\phi_R(t),$$

from which we again get equation (3.6).

From Diagram 3.1 we have

$$\phi_I' = \epsilon\phi_E - (\beta+\gamma)\phi_I.$$

We can rearrange (3.2) to get $\phi_E = \theta - \phi_S - \phi_I - \phi_R$. Substituting into this equation for ϕ_S and ϕ_R from (3.3) and (3.6) gives

$$\phi_E = \theta - \frac{\psi'(\theta)}{\psi'(1)} - \phi_I - \frac{\gamma}{\beta}(1-\theta). \tag{3.7}$$

By using this expression for ϕ_E in our equation for ϕ_I' we have

$$\phi_I' = \epsilon\Big[\theta - \frac{\psi'(\theta)}{\psi'(1)} - \phi_I - \frac{\gamma}{\beta}(1-\theta)\Big] - (\beta+\gamma)\phi_I,$$

which together with our equation for θ' gives us the system of two equations,

$$\phi_I' = \epsilon\Big[\theta - \frac{\psi'(\theta)}{\psi'(1)} - \frac{\gamma}{\beta}(1-\theta)\Big] - (\beta+\gamma+\epsilon)\phi_I \tag{3.8}$$

$$\theta' = -\beta\phi_I. \tag{3.9}$$

3.2 Model Analysis

In this section we investigate the dynamics of the system (3.8)–(3.9). Note that this system of two equations is entirely in terms of the functions $\phi_I(t)$ and $\theta(t)$, plus constants. We want to examine the behavior of this system by examining its equilibria. It will be convenient to refer to the right sides of the equations (3.8) and (3.9) as the functions $f(\phi_I, \theta)$ and $g(\phi_I, \theta)$, respectively. That is, we define $f(\phi_I, \theta) = \epsilon[\theta - \frac{\psi'(\theta)}{\psi'(1)} - \frac{\gamma}{\beta}(1-\theta)] - (\beta+\gamma+\epsilon)\phi_I$ and $g(\phi_I, \theta) = -\beta\phi_I$. Thus the system becomes

$$\phi_I' = f(\phi_I, \theta), \tag{3.10}$$

$$\theta' = g(\phi_I, \theta). \tag{3.11}$$

Before we set about finding equilibria for this system of ordinary differential equations (ODE) we first consider the quantity $\mathcal{R}_0$. In epidemic models $\mathcal{R}_0$ is usually defined as the number of new cases caused by a single randomly infected individual in a completely susceptible population. In this model we modify that definition to say that $\mathcal{R}_0$ is the number of new cases an average infected individual causes early in a disease outbreak [10]. When $\mathcal{R}_0 < 1$, no epidemic occurs in the

population; i.e., the disease dies out. If $\mathcal{R}_0 > 1$ an epidemic occurs. For an epidemic on a network, suppose u with degree k is chosen randomly from the population. If it is then infected by one of its neighbors, it could possibly infect $k-1$ other neighbors. Recall that $P($a randomly selected neighbor v of u has degree $k) = P_n(k)$, and

$$P_n(k) = \frac{kP(k)}{\langle K \rangle}.$$

To get an expression for $\mathcal{R}_0$ note also that since β is the rate of infection and γ is the rate of recovery, then the probability that an infected node will infect a neighbor before it itself recovers is $\frac{\beta}{\beta+\gamma}$. Recall that with $\psi(x) = \sum_k P(k)x^k$ we also have $\psi'(x) = \sum_k kP(k)x^{k-1}$ and $\psi''(x) = \sum_k k(k-1)P(k)x^{k-2}$; and when $x = 1$ we have $\psi''(1) = \sum_k k(k-1)P(k) = \langle K^2 - K \rangle$. We can write

$$\begin{aligned}
\mathcal{R}_0 &= \sum_k P_n(k)(k-1)\frac{\beta}{\beta+\gamma} \\
&= \frac{\beta}{\beta+\gamma}\sum_k \frac{kP(k)}{\langle K \rangle}(k-1) \\
&= \frac{\beta}{\beta+\gamma}\frac{\sum_k k(k-1)P(k)}{\langle K \rangle} \\
&= \frac{\beta}{\beta+\gamma}\frac{\psi''(1)}{\psi'(1)} \\
&= \frac{\beta}{\beta+\gamma}\frac{\langle K^2 - K \rangle}{\langle K \rangle}.
\end{aligned} \tag{3.12}$$

To find equilibria for the system of equations we wish to find values of ϕ_I and θ such that

$$\phi_I' = 0,$$
$$\theta' = 0.$$

By inspection, the second equation implies that $\phi_I = 0$ since β, the rate of infection, is assumed to be a positive quantity. If $\phi_I = 0$ then

$$f(0,\theta) = \epsilon\Big[\theta - \frac{\psi'(\theta)}{\psi'(1)} - \frac{\gamma}{\beta}(1-\theta)\Big].$$

Again by inspection, $f(0,\theta) = 0$ if $\theta = 1$. The next question is whether $(0,1)$ is the only equilibrium of our system, (3.8)–(3.9). We will see that the answer depends on the value of $\mathcal{R}_0$. To clarify our analysis, define

$$\overline{f}(\theta) = \theta - \frac{\psi'(\theta)}{\psi'(1)} - \frac{\gamma}{\beta}(1-\theta). \tag{3.13}$$

Then $\overline{f}(0) = -\frac{\gamma}{\beta}$ and $\overline{f}(1) = 0$. Note that θ is a probability function, so its values and thus the domain of the function $\overline{f}$ are the interval $[0,1]$. We can take derivatives with respect to θ to get

$$\overline{f}'(\theta) = 1 - \frac{1}{\psi'(1)}\frac{d}{d\theta}\psi'(\theta) + \frac{\gamma}{\beta} = \frac{\beta+\gamma}{\beta} - \frac{\sum_k k(k-1)P(k)\theta^{k-2}}{\langle K\rangle} \tag{3.14}$$

and

$$\overline{f}''(\theta) = -\frac{\sum_k k(k-1)(k-2)P(k)\theta^{k-3}}{\langle K\rangle}. \tag{3.15}$$

Since $\theta \in [0,1]$, the summands of the numerator summations for $\overline{f}'$ and $\overline{f}''$ as well as for $\langle K\rangle$ are all non-negative quantities. Therefore $\overline{f}'' < 0$ for all values of $\theta \in (0,1)$. We now discuss the equilibria of our system (3.10)–(3.11).

Theorem 1. If $\mathcal{R}_0 \leq 1$ then there exists a unique equilibrium $(\phi_I, \theta) = (0,1)$ for the system.

Proof. We have just seen that by setting ϕ_I' and θ' equal to zero we find that $(\phi_I, \theta) = (0,1)$ is an equilibrium of our system. To see that it is unique we consider $\overline{f}'$. If $\mathcal{R}_0 \leq 1$ then by (3.12) we have

$$\mathcal{R}_0 = \frac{\beta}{\beta+\gamma}\frac{\langle K^2 - K\rangle}{\langle K\rangle} \leq 1 \tag{3.16}$$

so that

$$\frac{\langle K^2 - K\rangle}{\langle K\rangle} \leq \frac{\beta+\gamma}{\beta}. \tag{3.17}$$

In the expression (3.14) for $\overline{f}'$ the sum $\sum_k k(k-1)P(k)\theta^{k-2}$ is a monotonically increasing function of θ because each summation term $k(k-1)P(k)\theta^{k-2}$ is monotonically increasing from 0 to $k(k-1)P(k)$ as θ increases from 0 to 1. Therefore as θ increases from 0 to 1, the sum $\sum_k k(k-1)P(k)\theta^{k-2}$ increases monotonically from 0 to $\langle K^2 - K\rangle$ $(= \psi''(1))$, so by (3.14) $\overline{f}'(\theta) > 0$ for all $\theta \in (0,1)$. Therefore $\overline{f}$ increases monotonically from $\overline{f}(0) = -\frac{\gamma}{\beta}$ to $\overline{f}(1) = 0$. There is no number θ^* with $0 < \theta^* < 1$ such that $\overline{f}(\theta^*) = 0$, for if there were then by Rolle's Theorem a number θ^{**} between θ^* and 1 would exist such that $\overline{f}'(\theta^{**}) = 0$, contradicting $\overline{f}'(\theta) > 0$. Thus the equilibrium $(\phi_I, \theta) = (0,1)$ is the unique equilibrium of the system when $\mathcal{R}_0 \leq 1$.

Theorem 2. If $\mathcal{R}_0 > 1$ then there exist two equilibria for the system, $(\phi_I, \theta) = (0,1)$ and $(\phi_I, \theta) = (0, \theta^*)$, where $0 < \theta^* < 1$.

Proof. As in Theorem 1 there is one equilibrium for the system at $(0,1)$. If $\mathcal{R}_0 > 1$ then by (3.12) we have

$$\mathcal{R}_0 = \frac{\beta}{\beta+\gamma}\frac{\langle K^2 - K\rangle}{\langle K\rangle} > 1 \tag{3.18}$$

so that

$$\frac{\langle K^2 - K\rangle}{\langle K\rangle} > \frac{\beta+\gamma}{\beta}. \tag{3.19}$$

By equation (3.14) we have

$$\overline{f}'(\theta) = \frac{\beta+\gamma}{\beta} - \frac{\sum_k k(k-1)P(k)\theta^{k-2}}{\langle K\rangle}.$$

Since

$$\lim_{\theta\to 0^+} \overline{f}'(\theta) = \frac{\beta+\gamma}{\beta} > 0,$$

we know that near $\theta = 0$, $\overline{f}'(\theta) > 0$. And since by (3.19)

$$\lim_{\theta \to 1^-} \overline{f}'(\theta) = \frac{\beta + \gamma}{\beta} - \frac{\langle K^2 - K \rangle}{\langle K \rangle} < 0,$$

then we know that near $\theta = 1$, $\overline{f}'(\theta) < 0$. We know $\overline{f}'(\theta)$ is continuous on $(0, 1)$, therefore by the Intermediate Value Theorem there exists some number $c \in (0, 1)$ such that $\overline{f}'(c) = 0$. By the first derivative test $\overline{f}(c)$ is a maximum value of $\overline{f}(\theta)$ on $[0, 1]$. We claim that $\overline{f}(c) > \overline{f}(1) = 0$, for if $\overline{f}(c) \leq \overline{f}(1) = 0$ then $\overline{f}'' > 0$ for some $\theta \in (0, 1)$. But by equation (3.15) this is impossible. We thus have $\overline{f}(0) = -\frac{\gamma}{\beta} < 0$ and $\overline{f}(c) > 0$. By the Intermediate Value Theorem there exists some $\theta^* \in (0, c)$ such that $\overline{f}(\theta^*) = 0$. This number is unique because if there were more than one candidate for θ^* then we would not have that $\overline{f}''(\theta) < 0$ for all $\theta \in (0, 1)$ by (3.15). Thus we have two equilibria for our system when $\mathcal{R}_0 > 1$.

We now determine system dynamics near the equilibria we have found. In order to do this we will utilize the Jacobian of the functions $f(\phi_I, \theta)$ and $g(\phi_I, \theta)$. The Jacobian is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f}{\partial \phi_I} & \frac{\partial f}{\partial \theta} \\ \frac{\partial g}{\partial \phi_I} & \frac{\partial g}{\partial \theta} \end{bmatrix} = \begin{bmatrix} -(\beta + \gamma + \epsilon) & \epsilon\left[1 - \frac{\psi''(\theta)}{\psi'(1)} + \frac{\gamma}{\beta}\right] \\ -\beta & 0 \end{bmatrix} = \begin{bmatrix} -(\beta + \gamma + \epsilon) & \epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\psi''(\theta)}{\psi'(1)}\right] \\ -\beta & 0 \end{bmatrix}.$$

We are going to want to examine the eigenvalues for this matrix, so we want to find the determinant of the matrix $\lambda\mathbf{I} - \mathbf{J}$, where $\mathbf{I}$ is the 2-by-2 identity matrix:

$$\det\left(\lambda\mathbf{I} - \mathbf{J}\right) = \begin{vmatrix} \lambda + (\beta + \gamma + \epsilon) & \epsilon\left[\frac{\psi''(\theta)}{\psi'(1)} - \frac{\beta+\gamma}{\beta}\right] \\ \beta & \lambda \end{vmatrix} = \lambda^2 + (\beta + \gamma + \epsilon)\lambda + \beta\epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\psi''(\theta)}{\psi'(1)}\right].$$

This determinant is a quadratic polynomial in λ. Notice that the coefficient of λ is positive, and the last term could be positive or negative depending on the value of the bracketed expression. It will be useful to apply the quadratic formula to obtain an expression for λ:

$$\lambda = \frac{-(\beta+\gamma+\epsilon) \pm \sqrt{(\beta+\gamma+\epsilon)^2 - 4\beta\epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\psi''(\theta)}{\psi'(1)}\right]}}{2} \tag{3.20}$$

Theorem 3. If $\mathcal{R}_0 < 1$ then the equilibrium $(\phi_I, \theta) = (0, 1)$ is locally stable.

Proof. Here we have $\theta = 1$ so that in the discriminant of (3.20), the bracketed portion

$$\frac{\beta+\gamma}{\beta} - \frac{\psi''(\theta)}{\psi'(1)} = \frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}. \tag{3.21}$$

If $\mathcal{R}_0 < 1$ then by (3.12)

$$\frac{\beta}{\beta+\gamma}\frac{\langle K^2 - K\rangle}{\langle K\rangle} < 1 \tag{3.22}$$

so that

$$\frac{\langle K^2 - K\rangle}{\langle K\rangle} < \frac{\beta+\gamma}{\beta} \tag{3.23}$$

and thus the bracketed portion of the discriminant is positive. There are two possibilities, depending on the size of $4\beta\epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\langle K^2-K\rangle}{\langle K\rangle}\right]$. If the discriminant > 0 we have

$$\sqrt{(\beta+\gamma+\epsilon)^2 - 4\beta\epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}\right]} < \sqrt{(\beta+\gamma+\epsilon)^2} = \beta+\gamma+\epsilon \tag{3.24}$$

since β, γ, and ϵ are positive quantities, so that in the numerator of λ, both

$$-(\beta+\gamma+\epsilon) + \sqrt{(\beta+\gamma+\epsilon)^2 - 4\beta\epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}\right]} < 0 \tag{3.25}$$

and

$$-(\beta+\gamma+\epsilon) - \sqrt{(\beta+\gamma+\epsilon)^2 - 4\beta\epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}\right]} < 0. \tag{3.26}$$

Hence both eigenvalues λ are negative real numbers, implying that the system is locally stable at the equilibrium $(\phi_I, \theta) = (0, 1)$.

If the discriminant < 0 then the values of λ are complex conjugates, both with negative real parts $-\frac{\beta+\gamma+\epsilon}{2}$, again implying that the system is locally stable at the equilibrium $(\phi_I, \theta) = (0, 1)$. This proves the theorem.

Theorem 4. If $\mathcal{R}_0 > 1$ then the equilibrium $(\phi_I, \theta) = (0, 1)$ is unstable, and the equilibrium $(\phi_I, \theta) = (0, \theta^*)$ where $0 < \theta^* < 1$ is asymptotically stable.

Proof. First we have $\theta = 1$ so that in the discriminant of (3.20), the bracketed portion

$$\frac{\beta+\gamma}{\beta} - \frac{\psi''(\theta)}{\psi'(1)} = \frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}. \tag{3.27}$$

If $\mathcal{R}_0 > 1$ then by (3.12)

$$\frac{\beta}{\beta+\gamma}\frac{\langle K^2 - K\rangle}{\langle K\rangle} > 1 \tag{3.28}$$

so that

$$\frac{\langle K^2 - K\rangle}{\langle K\rangle} > \frac{\beta+\gamma}{\beta} \tag{3.29}$$

and thus the bracketed portion of the discriminant is negative. Thus the discriminant becomes the sum of two positive terms and we have

$$\sqrt{(\beta+\gamma+\epsilon)^2 - 4\beta\epsilon\Big[\frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}\Big]} > \sqrt{(\beta+\gamma+\epsilon)^2} = \beta+\gamma+\epsilon \tag{3.30}$$

since β, γ, and ϵ are positive quantities, so that in the numerator of λ,

$$-(\beta+\gamma+\epsilon) + \sqrt{(\beta+\gamma+\epsilon)^2 - 4\beta\epsilon\Big[\frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}\Big]} > 0 \tag{3.31}$$

and

$$-(\beta+\gamma+\epsilon) - \sqrt{(\beta+\gamma+\epsilon)^2 - 4\beta\epsilon\Big[\frac{\beta+\gamma}{\beta} - \frac{\langle K^2 - K\rangle}{\langle K\rangle}\Big]} < 0. \tag{3.32}$$

Thus both of the eigenvalues are real with one positive and the other negative, therefore the equi-

librium $(\phi_I, \theta) = (0, 1)$ is unstable.

Next consider the equilibrium $(\phi_I, \theta) = (0, \theta^*)$ where $0 < \theta^* < 1$. For this equilibrium we return to the Jacobian of the system,

$$\mathbf{J} = \begin{bmatrix} -(\beta + \gamma + \epsilon) & \epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\psi''(\theta)}{\psi'(1)}\right] \\ -\beta & 0 \end{bmatrix}.$$

If we can show the trace of $\mathbf{J}$ is negative and the determinant is positive, we have shown the equilibrium is locally stable. Because β, γ, and ϵ are all positive quantities, the trace $-(\beta+\gamma+\epsilon) < 0$. The determinant $\beta\epsilon\left[\frac{\beta+\gamma}{\beta} - \frac{\psi''(\theta^*)}{\psi(1)}\right] > 0$ since the bracketed portion is positive. Thus the system is locally stable at $(\phi_I, \theta) = (0, \theta^*)$.

Following [12], put $\theta' = 0$. Define $\theta(\infty) = \lim_{t\to\infty} \theta(t)$. When the epidemic has ended we assume all derivatives $= 0$. Then equation (3.8) becomes

$$\theta(t) = \frac{\gamma}{\beta+\gamma} + \frac{\beta}{\beta+\gamma}\frac{\psi'(\theta(t))}{\psi'(1)}$$

and letting t go to infinity we get

$$\theta(\infty) = \frac{\gamma}{\beta+\gamma} + \frac{\beta}{\beta+\gamma}\frac{\psi'(\theta(\infty))}{\psi'(1)}.$$

When $\mathcal{R}_0 > 1$ this has two solutions for $\theta(\infty)$, one of which is $\theta = 1$ (the pre-disease equilibrium). We can numerically solve for the smaller $\theta(\infty)$, and use it to compute $R(\infty)$. Therefore when an epidemic occurs the total fraction of the population infected is

$$R(\infty) = 1 - \psi(\theta(\infty)).$$

CHAPTER 4: NUMERICAL SIMULATIONS

4.1 Random Network Compartmental Model Simulation

In the following simulations we are using NetworkX to simulate disease epidemics in a population of one thousand people. NetworkX is an open-source graph and network software package programmed in Python and developed at the Los Alamos National Laboratory. For our simulations NetworkX first constructs an Erdos-Renyi random graph $(N = 1000, p = 0.05)$, in which the probability of an edge connecting two nodes = 0.05. The value of $\mathcal{R}_0$ is assigned to be 0.8, 1.5, 2, or 5. One randomly picked node is chosen to be infectious, while all the other nodes are susceptible. Then a loop begins. Each susceptible neighbor of the infectious node may become exposed with a pre-determined probability and move out of the susceptible class, or else remain susceptible. Then the infectious node becomes recovered. Next, the algorithm looks at exposed nodes and decides (using a pre-determined probability) whether they become infectious or just stay exposed. As the process continues the number of infectious and exposed nodes will grow if an epidemic breaks out. The process repeats by going back to all susceptible neighbors of each infectious node. The loop continues until the number of exposed nodes plus the number of infectious nodes is zero.

In the first two graphs and the first three sets of six graphs, the simulation is run once and the numbers of susceptible, exposed, infectious, and removed nodes through the life of the epidemic are plotted. The fourth set of six graphs has plots of average numbers of infectious nodes over 1000 runs of the simulation. Three of the simulation runs average 1000 disease runs on a single graph; the other three runs average 1000 disease runs where one disease epidemic is run for each of 1000 graphs.

We first run two simulations for which $\mathcal{R}_0 = 0.8$. As $\mathcal{R}_0 < 1$ in this case, we do not expect a disease epidemic in the population. Results of the two simulations are shown in Table 4.1 and Figure 4.1. In each case there are multiple generations in which the maximum number of infectious individuals occurs. Although the disease is active beyond 20 generations in each case, the disease does not break into an epidemic.

Table 4.1: Two sample runs, $\mathcal{R}_0 = 0.8$

Sample	Infection Peak $I_{\max}$	Peak Moment(s) $I_{\max}$ Generation(s)	Final Infection Size $R_{\text{terminal}} = R(\infty)$	Disease Duration End Generation
(a)	1	0,3,4,7,9,10,11,15,16,18,19	11	21
(b)	3	9,13,21,37,42	46	49

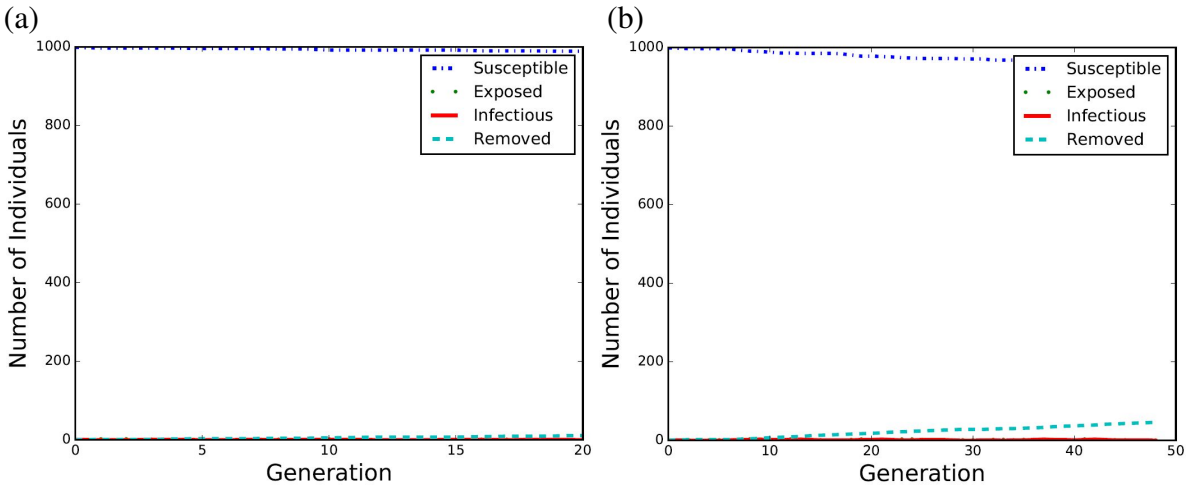


Figure 4.1: Sample simulations for the spread of infectious diseases over a random social network with $\mathcal{R}_0 = 0.8$.

In Table 4.2 and Figure 4.2 we set $\mathcal{R}_0 = 1.5$. The graph of Figure 4.2(a) shows that after one of the susceptible individuals becomes infectious in generation 0 the disease dies out immediately in generation 1 with no outbreak. In this instance the simulation is as brief as can happen, where the population goes immediately from one infectious individual to none just one generation after the simulation begins. In Figure 4.2(b) the disease maintains a low level, with never more than one infectious individual in the population during any generation. After generation 0 the peak number of infectious individuals occurs six more times until the outbreak ends at generation 14 with seven individuals having been infectious and then recovered from the disease and removed from susceptibility. In Figure 4.2(c) disease outbreak occurs. The number of infectious individuals peaks at $I_{\max} = 41$ in generation 27 and the oubreak ends at generation 76 with a total of $R(\infty) = 612$ individuals removed from the susceptible portion of the population. In the disease outbreaks of Figures 4.2(d), (e), and (f) the infectious peaks $I_{\max}$ occur in successively later generations. In Figure 4.2(e) $I_{\max}$ occurs on four occasions, demonstrating that it is possible to have $I_{\max}$ occur multiple times even if not at the beginning of a disease outbreak. Final outbreak sizes $R(\infty)$ are 668, 544, and 575 at generations 68, 112, and 99, respectively.

Table 4.2: Six sample runs, $\mathcal{R}_0 = 1.5$

Sample	Outbreak Peak $I_{\max}$	Peak Moment(s) $I_{\max}$ Generation(s)	Final Outbreak Size $R_{\text{terminal}} = R(\infty)$	Outbreak Duration End Generation
(a)	1	0	1	1
(b)	1	0,1,2,3,9,12,13	7	14
(c)	41	27	612	76
(d)	53	39	668	68
(e)	20	54,55,57,59	544	112
(f)	32	60	575	99

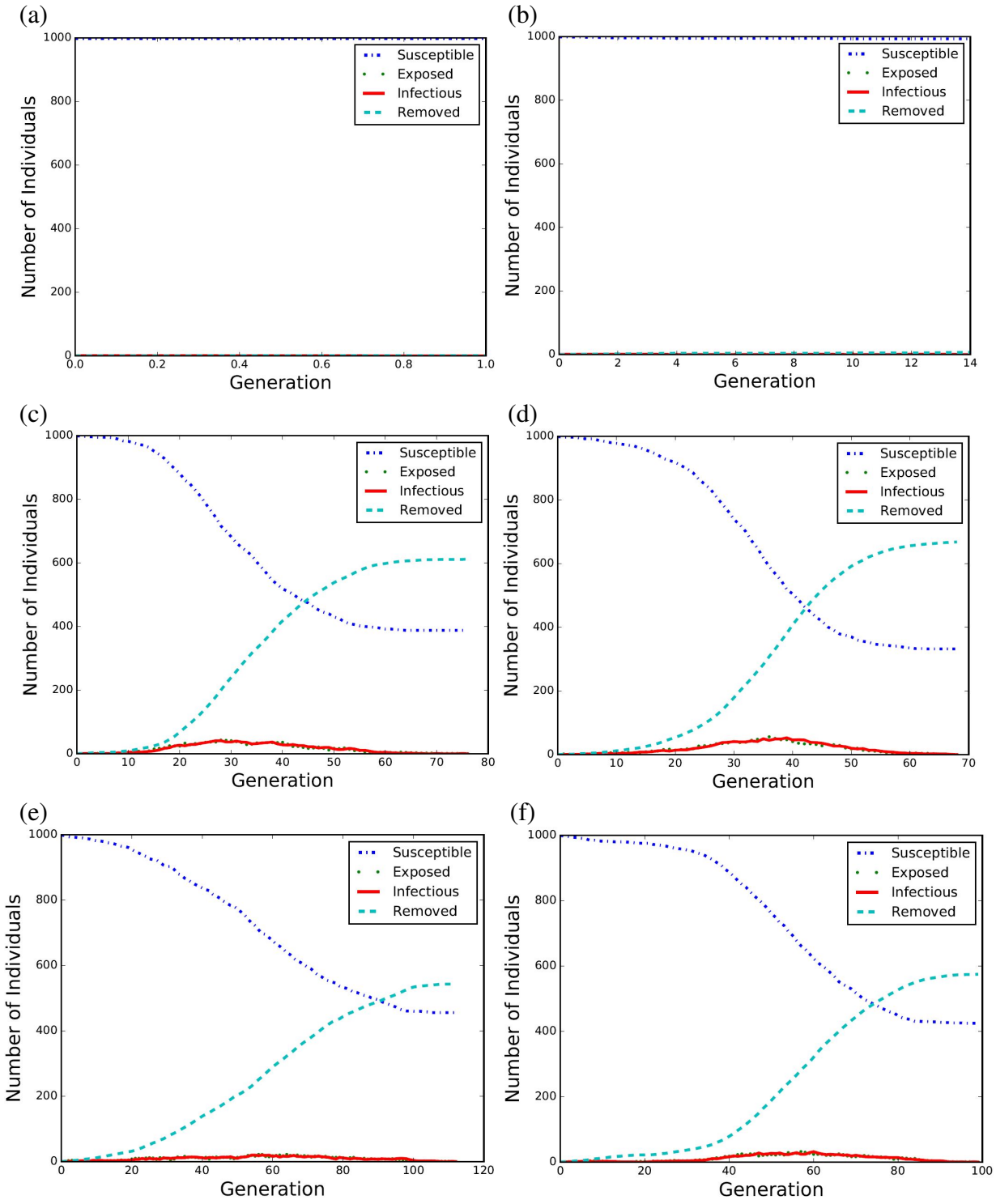


Figure 4.2: Sample simulations for the spread of infectious diseases over a random social network with $\mathcal{R}_0 = 1.5$.

In Table 4.3 and Figure 4.3 we set $\mathcal{R}_0 = 2$, meaning an increase in the likelihood that an infected individual in the model will infect other nodes. As can be seen from Figures 4.3(a) and (b), it is still possible with $\mathcal{R}_0 = 2$ for a disease outbreak not to occur, with the disease maintaining a low level until dying out at generations 8 and 17 respectively. The simulations (c), (d), (e), and (f) are disease outbreaks with peak numbers of infections $I_{\max}$ occurring in successively later generations, between 21 and 40. Compared to the sample simulation runs for outbreaks with $\mathcal{R}_0 = 1.5$, the sample runs for $\mathcal{R}_0 = 2$ outbreaks have greater output peaks with $I_{\max}$ between 69 and 87 versus 20 to 53, and the peak moments for $\mathcal{R}_0 = 2$ occur overall somewhat earlier, between generations 21 and 40 as opposed to generations 27 to 60. The final outbreak sizes for $\mathcal{R}_0 = 2$ range between 767 and 818 as opposed to between 544 and 668 for $\mathcal{R}_0 = 1.5$, and the outbreak end generation ranges between 58 and 68 for $\mathcal{R}_0 = 2$ as opposed to 68 to 112 for the $\mathcal{R}_0 = 1.5$ sample runs. Generalizing from these simulation samples for the two values of $\mathcal{R}_0$, one can say that when $\mathcal{R}_0$ is increased from 1.5 to 2 and one compares disease outbreaks, the outbreak peaks increased with peak moments generally in earlier generations, and final outbreak sizes $R(\infty)$ increased with a decrease in the overall disease outbreak durations.

Table 4.3: Six sample runs, $\mathcal{R}_0 = 2$

Sample	Outbreak Peak $I_{\max}$	Peak Moment(s) $I_{\max}$ Generation(s)	Final Outbreak Size $R_{\text{terminal}} = R(\infty)$	Outbreak Duration End Generation
(a)	1	0,7	2	8
(b)	3	6	14	17
(c)	87	21	818	63
(d)	73	24,25	801	58
(e)	72	30	817	62
(f)	69	40	767	68

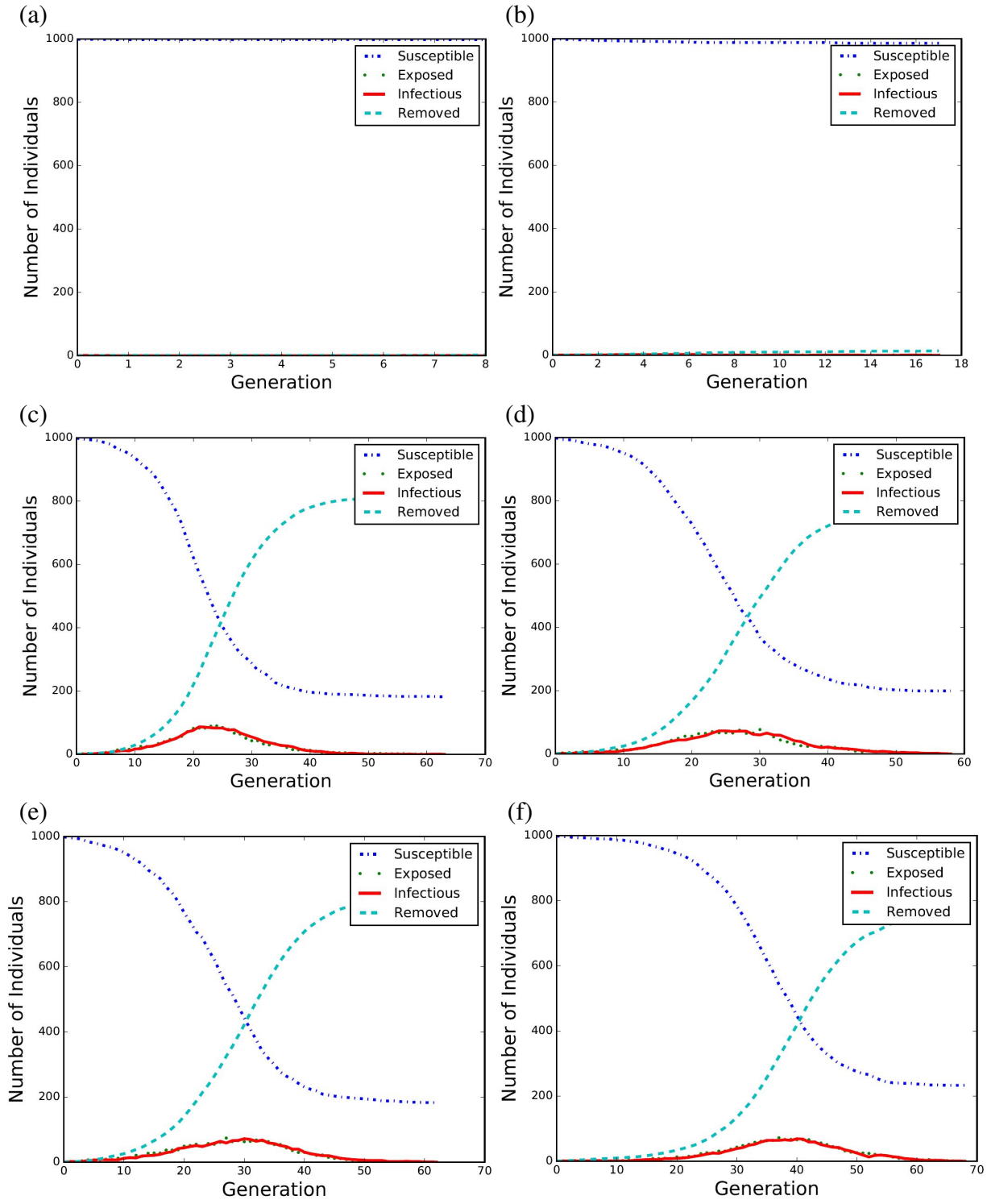


Figure 4.3: Sample simulations for the spread of infectious diseases over a random social network with $\mathcal{R}_0 = 2$.

Table 4.4 and Figure 4.4 give details of six sample simulation runs with $\mathcal{R}_0$ set to 5. The results of the six sample runs show greater uniformity for $\mathcal{R}_0 = 5$ than for either $\mathcal{R}_0 = 1.5$ or 2. All six of the samples had disease outbreaks. Outbreak peak values $I_{\max}$ in these six sample range between 227 and 240, higher values and a narrower range of values than with $\mathcal{R}_0 = 1.5$ or 2 (20 to 53 and 69 to 87, respectively). The generations in which $I_{\max}$ occurred for $\mathcal{R}_0 = 5$ are earlier as well as in a shorter time interval, 10 to 15 as opposed to 27 to 60 and 21 to 40. Final outbreak sizes $R(\infty)$ for $\mathcal{R}_0 = 5$ are greater and confined to a shorter interval: 988 to 996 versus 544 to 668 and 767 to 818. The outbreak durations for $\mathcal{R}_0 = 5$ are a set of generations that is the earliest as well as in the shortest time interval: generations 27 to 32 versus 68 to 112 and 58 to 68.

As can be seen from the outbreak information of the sample runs for the three chosen values of $\mathcal{R}_0$, as $\mathcal{R}_0$ increases the numbers of infectious individuals increase, both at the time of peak outbreak as well as over the entire life span of the epidemic. With increasing $\mathcal{R}_0$ the growth of infectiousness occurs at a faster rate, resulting in earlier peak generations of infectiousness as well as a disease epidemic over a smaller number of generations. The greater the value of $\mathcal{R}_0$ the faster and more aggressive the progression of the disease in the host population.

Table 4.4: Six sample runs, $\mathcal{R}_0 = 5$

Sample	Outbreak Peak $I_{\max}$	Peak Moment $I_{\max}$ Generation	Final Outbreak Size $R_{\text{terminal}} = R(\infty)$	Outbreak Duration End Generation
(a)	228	10	988	27
(b)	230	11	990	30
(c)	239	11	993	28
(d)	232	12	992	30
(e)	227	13	993	30
(f)	240	15	996	32

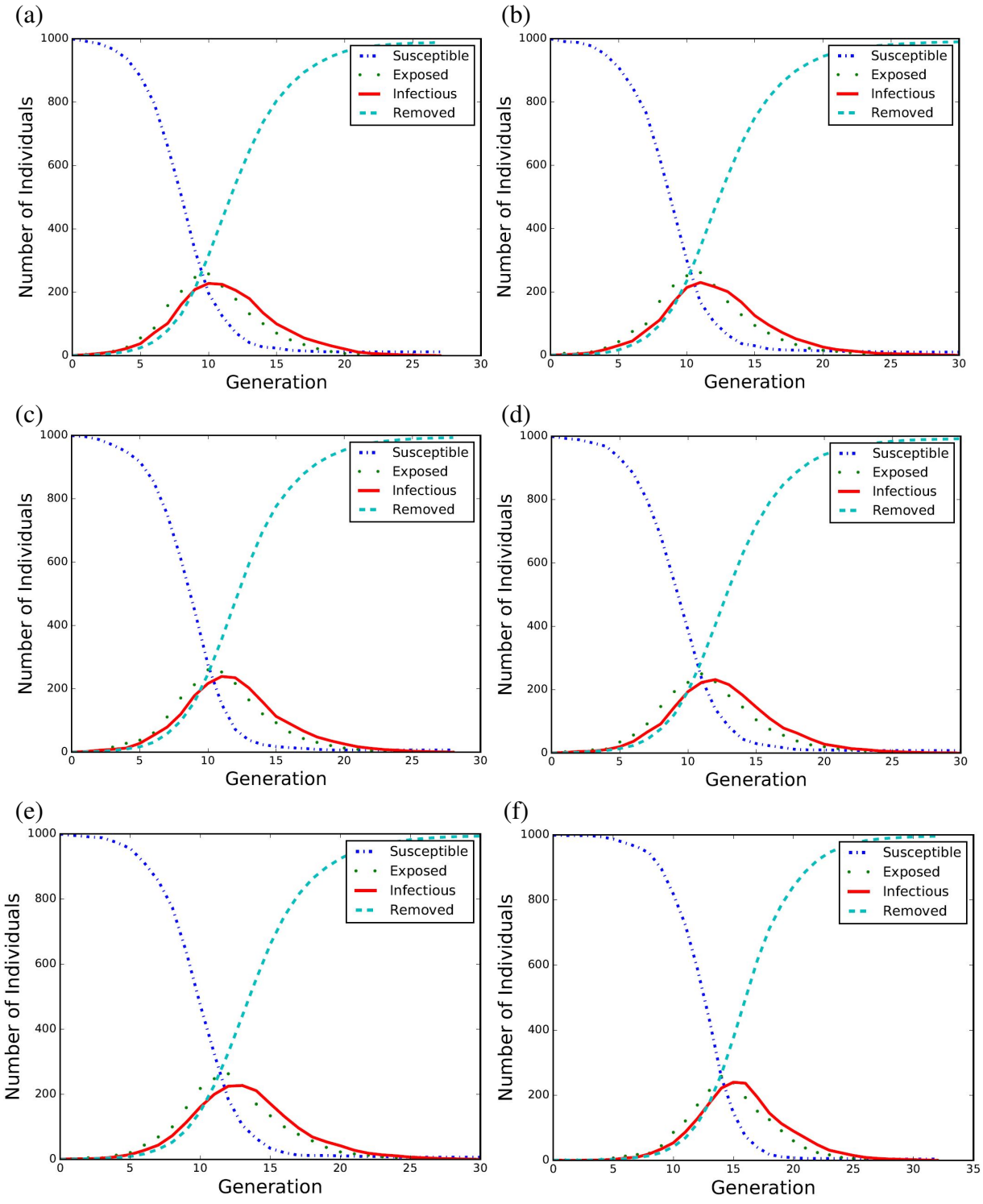


Figure 4.4: Sample simulations for the spread of infectious diseases over a random social network with $\mathcal{R}_0 = 5$.

In Table 4.5 and Figure 4.5 we display the results of six averages of one thousand sample runs, two averages each for all three values of $\mathcal{R}_0$. For each value of $\mathcal{R}_0$ the first sample ("fixed graph") is an average of one thousand runs of randomly generated disease epidemics over the same randomly generated graph of one thousand nodes. The second sample ("random graph") for each value of $\mathcal{R}_0$ is an average of one thousand runs in which one disease epidemic is simulated on each of one thousand different randomly generated graphs of one thousand nodes. In Table 4.5 the results are paired in adjacent rows by $\mathcal{R}_0$ value.

As can be seen from Table 4.5, averages for fixed and random samples for each value of $\mathcal{R}_0$ are very similar. For sample averages (1.5) and (1.5′), the outbreak peak value $I_{\max}$ for (1.5′) is only about 0.9% greater than that for (1.5). The peak generation for (1.5) is 41, for (1.5′) is 40. Final outbreak size $R(\infty)$ for (1.5) is about 1.6% more than that for (1.5′), and the outbreak duration for (1.5) is about 2.7% greater than that for (1.5′). When $\mathcal{R}_0 = 2$, $I_{\max}$ for (2) is about 2.2% greater than that of (2′), with both outbreak peaks occurring in generation 28. For sample average (2) $R(\infty)$ is about 1.3% greater than $R(\infty)$ for (2′), and the outbreak duration for (2) is about 0.9% greater than that for (2′). For $\mathcal{R}_0 = 5$, $I_{\max}$ for (5′) is about 0.4% greater than that for (5), both peaks occurring in generation 12. For (5′) $R(\infty)$ is about 0.09% greater than that for (5) and the outbreak duration for (5) is about 0.1% greater than that for (5′).

Table 4.5: Averages for 1000 fixed graph runs and 1000 random graph runs, $\mathcal{R}_0 = 1.5, 2,$ and 5

	Outbreak Peak $I_{\max}$	Peak Moment $I_{\max}$ Generation	Final Outbreak Size $R_{\text{terminal}} = R(\infty)$	Outbreak Duration End Generation
(1.5) Fixed	15.841	41	342.946	54.085
(1.5′) Random	15.981	40	337.641	52.684
(2) Fixed	54.045	28	640.56	49.331
(2′) Random	52.872	28	632.427	48.88
(5) Fixed	214.201	12	984.191	30.261
(5′) Random	215.126	12	985.104	30.22

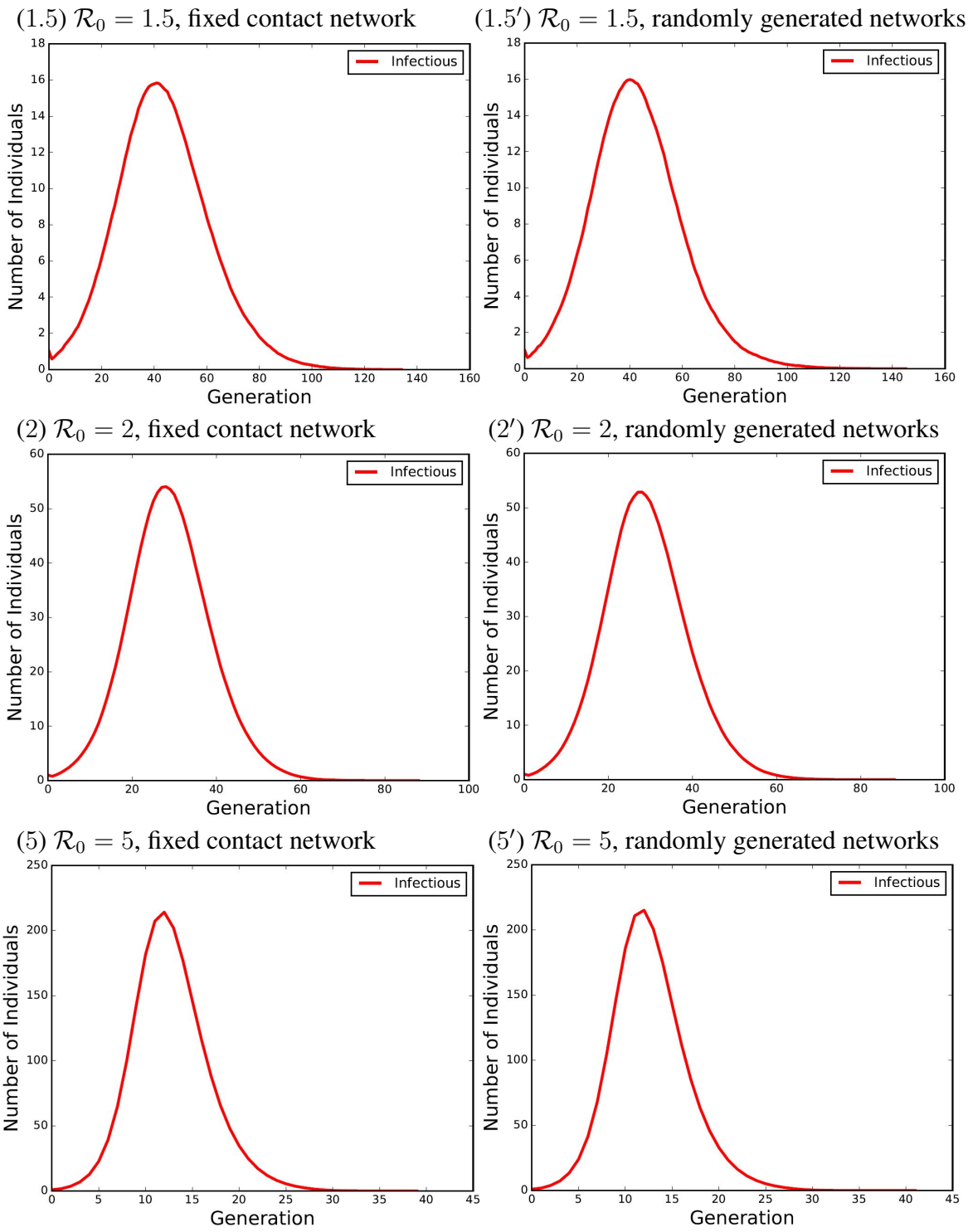


Figure 4.5: Averages of 1000 sample simulations for the spread of infectious diseases over a random social network with $\mathcal{R}_0 = 1.5$, 2, and 5. Left column: fixed network. Right column: average over 1000 random networks.

4.2 ODE Compartmental Model Simulation

In the next table and figures we display behaviour of the system of ODEs from Chapter 2, equations (2.2)–(2.5). The Runge-Kutta (4,5) algorithm is used by MATLAB/GNU Octave code to solve the system of equations once the value of $\mathcal{R}_0$ is set. Because the system is deterministic with initial conditions, only one run for each $\mathcal{R}_0$ value is required.

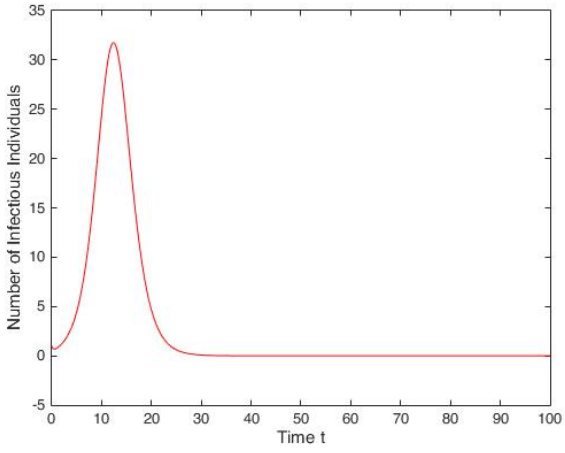
In comparing results of this deterministic model with the previous contact network model it is easiest to compare results of Table 4.5, averages over 1000 runs, with Table 4.6. The results of the Chapter 2 SEIR model can be generalized to say that the assumption of homogeneous mixing increases numbers of infectious individuals and accelerates the spread of infectiouness in the population, compared to the contact network model. Comparison of homogeneous mixing versus contact network for $\mathcal{R}_0 = 1.5$ shows that $I_{\max}$ is 31.76 versus about 15.9, occurring around generation 12 versus generation 40 in the contact network model. When $\mathcal{R}_0 = 2$, $I_{\max}$ is 76.26 occurring around generation 8 versus about 53 at generation 28 in the contact network model. When $\mathcal{R}_0 = 5$, $I_{\max}$ is 226.96 occurring around generation 3 versus about 214 at generation 12 in the contact network model.

Similarly for comparing final outbreak sizes $R(\infty)$ and when they occur, for all values of $\mathcal{R}_0$ the $R(\infty)$ values are greater and epidemic duration times are shorter in the deterministic model.

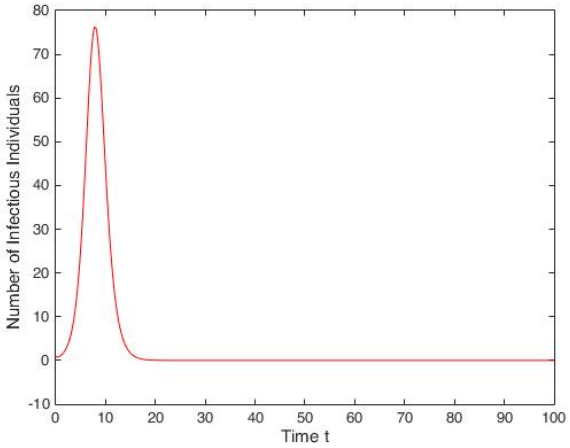
Table 4.6: Infectious in Chapter 2 SEIR model, $\mathcal{R}_0 = 1.5, 2$, and 5

$\mathcal{R}_0$	Outbreak Peak $I_{\max}$	Peak Moment $I_{\max}$ Generation	Final Outbreak Size $R_{\text{terminal}} = R(\infty)$	Outbreak Duration End Generation
1.5	31.76	12.34	583.92	51.11
2	76.26	7.89	797.15	33.32
5	226.96	3.21	993.05	14.63

(a)



(b)



(c)

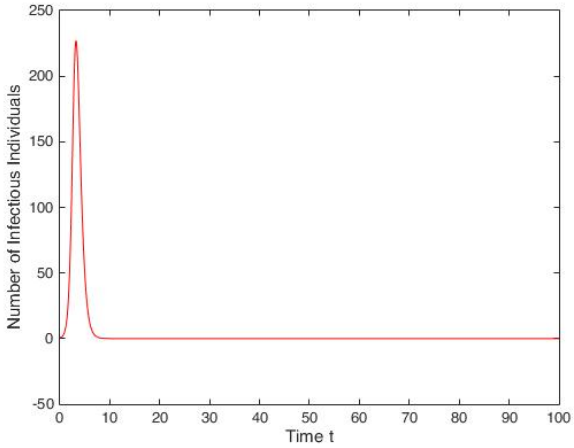


Figure 4.6: Infectious numbers in SEIR model for infectious disease spread in population with homogeneous mixing: (a) $\mathcal{R}_0 = 1.5$, (b) $\mathcal{R}_0 = 2$, (c) $\mathcal{R}_0 = 5$.

CHAPTER 5: CONCLUSION AND FUTURE STUDY

In the preceding chapters we have considered the SEIR model, a variation of the SIR infectious disease model that admits a latency period for the disease. We first examined the compartmental model that assumes homogeneous mixing in the host population. We discussed the mathematics of the model as well as its numerical simulation. Next we examined a network model that allows heterogeneous mixing beause connections between individual nodes are defined in terms of a probability distribution. We employed numerical simulation for this model as well. We have extensively explored numerical simulations of disease spread on randomly generated networks of one thousand nodes each. Examples of results have shown that as $\mathcal{R}_0$ increases in value, epidemics occur in less time and result in greater numbers of infected individuals. As tables of figures show, as $\mathcal{R}_0$ increases final outbreak size increases while outbreak duration decreases.

It was found that in the compartmental model that assumes mass action incidence and homogeneous mixing, there are greater numbers of infectious individuals and the disease spread during an epidemic occurs more rapidly. It appears that disease spread is less acute when a more realistic assumption is made of heterogeneous mixing of individuals within the population.

There are several possible avenues of future study. One focuses on developing and analyzing variations of the SEIR model. For example, there could be more than one infectious stage in the model, so the model could be SEIIR or SEIIIR or something similar. With more stages the challenge increases as to how one could define the model so that meaningful and useful mathematical analysis can be employed without paying too high a penalty in complexity. An alternative avenue of study would be to take the SEIR model discussed here and try applying it to a particular disease whose behavior seems similar. Ideally the model could shed light on the disease behavior in a way that would assist health care workers in the containment or treatment of the disease. It is possible that

a variation of the SEIR model would be most helpful in this regard. Computer simulations could help determine the possible effectiveness of various treatment or containment options.

LIST OF REFERENCES

[1] F. Brauer, C. Castillo-Chavez, *Mathematical Models for Communicable Diseases*, SIAM, 2013

[2] F. Brauer, P. van den Driessche, J. Wu (Eds.), *Mathematical Epidemiology*, Springer-Verlag, 2008

[3] S.L. Feld, Why your friends have more friends than you do, *American Journal of Sociology*, **96** (1991) 1464–1477

[4] H.W. Hethcote, The mathematics of infectious diseases, *SIAM Review* **42** (2000) 599–653.

[5] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Royal Society of London Proceedings Series A* **115** (1927) 700–721

[6] O. Krylova, D.J.D. Earn, Effects of the infectious period distribution on predicted transitions in childhood disease dynamics, *Journal of the Royal Society Interface* **10** (2013)

[7] L.A. Meyers, B. Pourbohloul, M.E.J. Newman, D.M. Skowronski, R.C. Brunham, Network theory and SARS: predicting outbreak diversity, *Journal of Theoretical Biology*, **232** (2005) 71–81

[8] J.C. Miller, A note on a paper by Erik Volz: SIR dynamics in random networks, *Journal of Mathematical Biology*, **62** (2011) 349–358

[9] J.C. Miller, I.Z. Kiss, Epidemic spread in networks: existing methods and current challenges, *Mathematical Modeling of Natural Phenomena*, **9** (2014) 4–42

[10] J.C. Miller, E.M. Volz, Incorporating disease and population structure into models of SIR disease in contact networks, *PLOS ONE*, **8** (2013) 1–14

[11] J.C. Miller, E.M. Volz, Model hierarchies in edge-based compartmental modeling for infectious disease spread, *Journal of Mathematical Biology* **67** (2013) 869–899

[12] J.C. Miller, A.C. Slim, E.M. Volz, Edge-based compartmental modeling for infectious disease spread, *Journal of the Royal Society Interface* **9** (2012) 890–906

[13] M.E.J. Newman, The spread of epidemic disease on networks, *Physical Review E*, **66** (2002)

[14] M.E.J. Newman, The structure and function of complex networks, *SIAM Review*, **45** (2003) 167–256

[15] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, *Physical Review Letters*, **86** (2001) 3200–3203

[16] E. Volz, SIR dynamics in random networks with heterogeneous connectivity, *Journal of Mathematical Biology*, **56** (2008) 293–310