An Introduction to Stochastic Epidemic Models

Linda J. S. Allen

Department of Mathematics and Statistics
Texas Tech University
Lubbock, Texas 79409-1042, U.S.A.
linda.j.allen@ttu.edu

1 Introduction

The goals of this chapter are to provide an introduction to three different methods for formulating stochastic epidemic models that relate directly to their deterministic counterparts, to illustrate some of the techniques for analyzing them, and to show the similarities between the three methods. Three types of stochastic modeling processes are described: 1) a discrete time Markov chain (DTMC) model, 2) a continuous time Markov chain (CTMC) model, and 3) a stochastic differential equation (SDE) model. These stochastic processes differ in the underlying assumptions regarding the time and the state variables. In a DTMC model, the time and the state variables are discrete. In a CTMC model, time is continuous, but the state variable is discrete. Finally, the SDE model is based on a diffusion process, where both the time and the state variables are continuous.

Stochastic models based on the well-known SIS and SIR epidemic models are formulated. For reference purposes, the dynamics of the SIS and SIR deterministic epidemic models are reviewed in the next section. Then the assumptions that lead to the three different stochastic models are described in Sects. 3, 4, and 5. The deterministic and stochastic model dynamics are illustrated through several numerical examples. Some of the MatLab programs used to compute numerical solutions are provided in the last section of this chapter.

One of the most important differences between the deterministic and stochastic epidemic models is their asymptotic dynamics. Eventually stochastic solutions (sample paths) converge to the disease-free state even though the corresponding deterministic solution converges to an endemic equilibrium. Other properties that are unique to the stochastic epidemic models include the probability of an outbreak, the quasistationary probability distribution, the final size distribution of an epidemic and the expected duration of an epidemic. These properties are discussed in Sect. 6. In Sect. 7, the SIS epi-

demic model with constant population size is extended to one with a variable population size and the corresponding SDE model is derived.

The chapter ends with a discussion of two well-known DTMC epidemic processes that are not directly related to any deterministic epidemic model. These two processes are chain binomial epidemic processes and branching epidemic processes.

2 Review of Deterministic SIS and SIR Epidemic Models

In SIS and SIR epidemic models, individuals in the population are classified according to disease status, either susceptible, infectious, or immune. The immune classification is also referred to as removed because individuals are no longer spreading the disease when they are removed or isolated from the infection process. These three classifications are denoted by the variables S, I, and R, respectively.

In an SIS epidemic model, a susceptible individual, after a successful contact with an infectious individual, becomes infected and infectious, but does not develop immunity to the disease. Hence, after recovery, infected individuals return to the susceptible class. The SIS epidemic model has been applied to sexually transmitted diseases. We make some additional simplifying assumptions. There is no vertical transmission of the disease (all individuals are born susceptible) and there are no disease-related deaths. A compartmental diagram in Fig. 1 illustrates the dynamics of the SIS epidemic model. Solid arrows denote infection or recovery. Dotted arrows denote births or deaths.

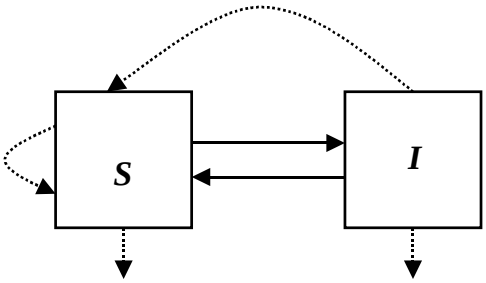

Fig. 1. SIS compartmental diagram.

Differential equations describing the dynamics of an SIS epidemic model based on the preceding assumptions have the following form:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{N}SI + (b+\gamma)I \\ \frac{dI}{dt} &= \frac{\beta}{N}SI - (b+\gamma)I, \end{aligned} \tag{1}$$

where $\beta > 0$ is the contact rate, $\gamma > 0$ is the recovery rate, $b \geq 0$ is the birth rate, and $N = S(t) + I(t)$ is the total population size. The initial conditions satisfy $S(0) > 0$, $I(0) > 0$, and $S(0) + I(0) = N$. We assume that the birth rate equals the death rate, so that the total population size is constant, $dN/dt = 0$. The dynamics of model (1) are well-known [25]. They are determined by the *basic reproduction number*. The basic reproduction number is the number of secondary infections caused by one infected individual in an entirely susceptible population [10, 26]. For model (1), the basic reproduction number is defined as follows:

$$\mathcal{R}_0 = \frac{\beta}{b + \gamma}. \tag{2}$$

The fraction $1/(b + \gamma)$ is the length of the infectious period, adjusted for deaths. The asymptotic dynamics of model (1) are summarized in the following theorem.

Theorem 1. *Let $S(t)$ and $I(t)$ be a solution to model (1).*

i) If $\mathcal{R}_0 \leq 1$, then $\lim_{t\to\infty}(S(t), I(t)) = (N, 0)$ (disease-free equilibrium).

ii) If $\mathcal{R}_0 > 1$, then $\lim_{t\to\infty}(S(t), I(t)) = \left(\frac{N}{\mathcal{R}_0}, N\left(1 - \frac{1}{\mathcal{R}_0}\right)\right)$ (endemic equilibrium).

In an SIR epidemic model, individuals become infected, but then develop immunity and enter the immune class R. The SIR epidemic model has been applied to childhood diseases such as chickenpox, measles, and mumps. A compartmental diagram in Fig. 2 illustrates the relationship between the three classes. Differential equations describing the dynamics of an SIR epidemic

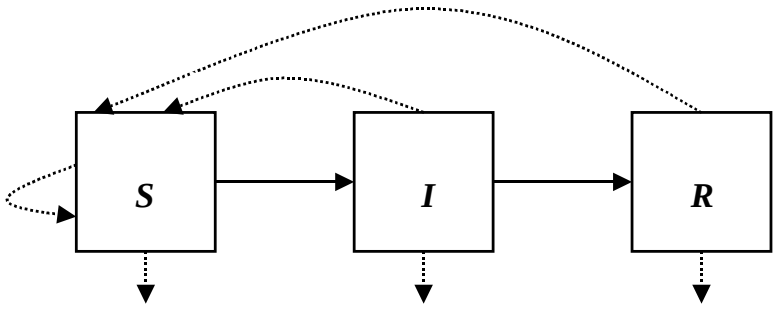

Fig. 2. SIR compartmental diagram.

model have the following form:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta}{N}SI + b(I+R)\\ \frac{dI}{dt} &= \frac{\beta}{N}SI - (b+\gamma)I\\ \frac{dR}{dt} &= \gamma I - bR,\end{aligned} \tag{3}$$

where $\beta > 0$, $\gamma > 0$, $b \geq 0$, and the total population size satisfies $N = S(t) + I(t) + R(t)$. The initial conditions satisfy $S(0) > 0$, $I(0) > 0$, $R(0) \geq 0$, and $S(0) + I(0) + R(0) = N$. We assume that the birth rate equals the death rate so that the total population size is constant, $dN/dt = 0$.

The basic reproduction number (2) and the birth rate b determine the dynamics of model (3). The dynamics are summarized in the following theorem.

Theorem 2. *Let $S(t)$, $I(t)$, and $R(t)$ be a solution to model (3).*

i) If $\mathcal{R}_0 \leq 1$, then $\lim_{t\to\infty} I(t) = 0$ (disease-free equilibrium).

ii) If $\mathcal{R}_0 > 1$, then

$$\lim_{t\to\infty}(S(t), I(t), R(t)) = \left(\frac{N}{\mathcal{R}_0}, \frac{bN}{b+\gamma}\left(1 - \frac{1}{\mathcal{R}_0}\right), \frac{\gamma N}{b+\gamma}\left(1 - \frac{1}{\mathcal{R}_0}\right)\right)$$

(endemic equilibrium).

iii) Assume $b = 0$. If $\mathcal{R}_0\dfrac{S(0)}{N} > 1$, then there is an initial increase in the number of infected cases $I(t)$ (epidemic), but if $\mathcal{R}_0\dfrac{S(0)}{N} \leq 1$, then $I(t)$ decreases monotonically to zero (disease-free equilibrium).

The quantity $\mathcal{R}_0 S(0)/N$ is referred to as the *initial replacement number*, the average number of secondary infections produced by an infected individual during the period of infectiousness at the outset of the epidemic [25, 26]. Since the infectious fraction changes during the course of the epidemic, the replacement number is generally defined as $\mathcal{R}_0 S(t)/N$ [25, 26]. In case iii) of Theorem 2, the disease eventually disappears from the population but if the initial replacement number is greater than one, the population experiences an outbreak.

3 Formulation of DTMC Epidemic Models

Let $\mathcal{S}(t)$, $\mathcal{I}(t)$, and $\mathcal{R}(t)$ denote discrete random variables for the number of susceptible, infected, and immune individuals at time t, respectively. (Calligraphic letters denote random variables.) In a DTMC epidemic model, $t \in \{0, \Delta t, 2\Delta t, \ldots\}$ and the discrete random variables satisfy

$$\mathcal{S}(t),\ \mathcal{I}(t),\ \mathcal{R}(t) \in \{0, 1, 2, \ldots, N\}.$$

The term "chain" (letter C) in DTMC means that the random variables are discrete. The term "Markov" (letter M) in DTMC is defined in the next section.

3.1 SIS Epidemic Model

In an SIS epidemic model, there is only one independent random variable, $\mathcal{I}(t)$, because $\mathcal{S}(t) = N - \mathcal{I}(t)$, where N is the constant total population size. The stochastic process $\{\mathcal{I}(t)\}_{t=0}^{\infty}$ has an associated probability function,

$$p_i(t) = \text{Prob}\{\mathcal{I}(t) = i\},$$

for $i = 0, 1, 2, \ldots, N$ and $t = 0, \Delta t, 2\Delta t, \ldots$,where

$$\sum_{i=0}^{N} p_i(t) = 1.$$

Let $p(t) = (p_0(t), p_1(t), \ldots, p_N(t))^T$ denote the probability vector associated with $\mathcal{I}(t)$. The stochastic process has the *Markov property* if

$$\text{Prob}\{\mathcal{I}(t + \Delta t)|\mathcal{I}(0), \mathcal{I}(\Delta t), \ldots, \mathcal{I}(t)\} = \text{Prob}\{\mathcal{I}(t + \Delta t)|\mathcal{I}(t)\}.$$

The Markov property means that the process at time $t + \Delta t$ only depends on the process at the previous time step t.

To complete the formulation for a DTMC SIS epidemic model, the relationship between the states $\mathcal{I}(t)$ and $\mathcal{I}(t + \Delta t)$ needs to be defined. This relationship is determined by the underlying assumptions in the SIS epidemic model and is defined by the transition matrix. The probability of a transition from state $\mathcal{I}(t) = i$ to state $\mathcal{I}(t + \Delta t) = j$, $i \to j$, in time Δt, is denoted as

$$p_{ji}(t + \Delta t, t) = \text{Prob}\{\mathcal{I}(t + \Delta t) = j|\mathcal{I}(t) = i\}.$$

When the transition probability $p_{ji}(t + \Delta t, t)$ does not depend on t, $p_{ji}(\Delta t)$, the process is said to be *time homogeneous*. For the stochastic SIS epidemic model, the process is time homogeneous because the deterministic model is autonomous.

To reduce the number of transitions in time Δt, we make one more assumption. The time step Δt is chosen sufficiently small such that the number of infected individuals changes by at most one during the time interval Δt, that is,

$$i \to i + 1, \quad i \to i - 1 \quad \text{or} \quad i \to i.$$

Either there is a new infection, a birth, a death, or a recovery during the time interval Δt. Of course, this latter assumption can be modified, if the time step cannot be chosen arbitrarily small. In this latter case, transition probabilities need to be defined for all possible transitions that may occur, $i \to i + 2$, $i \to i + 3$, etc. In the simplest case, with only three transitions possible, the transition probabilities are defined using the rates (multiplied by Δt) in the deterministic SIS epidemic model. This latter assumption makes the DTMC model a useful approximation to the CTMC model, described in Sect. 4. The transition probabilities for the DTMC epidemic model satisfy

$$p_{ji}(\Delta t) = \begin{cases} \dfrac{\beta i(N-i)}{N}\Delta t, & j = i+1 \\ (b+\gamma)i\Delta t, & j = i-1 \\ 1 - \left[\dfrac{\beta i(N-i)}{N} + (b+\gamma)i\right]\Delta t, & j = i \\ 0, & j \neq i+1, i, i-1. \end{cases}$$

The probability of a new infection, $i \to i+1$, is $\beta i(N-i)\Delta t/N$. The probability of a death or recovery, $i \to i-1$, is $(b+\gamma)i\Delta t$. Finally, the probability of no change in state, $i \to i$, is $1 - [\beta i(N-i)/N + (b+\gamma)i]\,\Delta t$. Since a birth of a susceptible must be accompanied by a death, to keep the population size constant, this probability is not needed in either the deterministic or stochastic formulations.

To simplify the notation and to relate the SIS epidemic process to a birth and death process, the transition probability for a new infection is denoted as $b(i)\Delta t$ and for a death or a recovery is denoted as $d(i)\Delta t$. Then

$$p_{ji}(\Delta t) = \begin{cases} b(i)\Delta t, & j = i+1 \\ d(i)\Delta t, & j = i-1 \\ 1 - [b(i)+d(i)]\Delta t, & j = i \\ 0, & j \neq i+1, i, i-1. \end{cases}$$

The sum of the three transitions equals one because these transitions represent all possible changes in the state i during the time interval Δt. To ensure that these transition probabilities lie in the interval $[0, 1]$, the time step Δt must be chosen sufficiently small such that

$$\max_{i\in\{1,\ldots,N\}} \{[b(i)+d(i)]\Delta t\} \leq 1.$$

Applying the Markov property and the preceding transition probabilities, the probabilities $p_i(t+\Delta t)$ can be expressed in terms of the probabilities at time t. At time $t+\Delta t$,

$$p_i(t+\Delta t) = p_{i-1}(t)b(i-1)\Delta t + p_{i+1}(t)d(i+1)\Delta t + p_i(t)(1-[b(i)+d(i)]\Delta t) \quad (4)$$

for $i = 1, 2, \ldots, N$, where $b(i) = \beta i(N-i)/N$ and $d(i) = (b+\gamma)i$.

A transition matrix is formed when the states are ordered from 0 to N. For example, the $(1,1)$ element in the transition matrix is the transition probability from state zero to state zero, $p_{00}(\Delta t) = 1$, and the $(N+1, N+1)$ element is the transition probability from state N to state N, $p_{NN}(\Delta t) = 1 - [b+\gamma]N\Delta t = 1 - d(N)\Delta t$. Denote the transition matrix as $P(\Delta t)$. Matrix $P(\Delta t)$ is a $(N+1) \times (N+1)$ tridiagonal matrix given by

$$\begin{pmatrix} 1 & d(1)\Delta t & 0 & \cdots & 0 & 0 \\ 0 & 1-(b+d)(1)\Delta t & d(2)\Delta t & \cdots & 0 & 0 \\ 0 & b(1)\Delta t & 1-(b+d)(2)\Delta t & \cdots & 0 & 0 \\ 0 & 0 & b(2)\Delta t & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & d(N-1)\Delta t & 0 \\ 0 & 0 & 0 & \cdots & 1-(b+d)(N-1)\Delta t & d(N)\Delta t \\ 0 & 0 & 0 & \cdots & b(N-1)\Delta t & 1-d(N)\Delta t \end{pmatrix},$$

where the notation $(b+d)(i) = [b(i)+d(i)]$ for $i = 1, 2, \ldots, N$. Matrix $P(\Delta t)$ is a *stochastic matrix*, i.e., the column sums equal one.

The DTMC SIS epidemic process $\{\mathcal{I}(t)\}_{t=0}^{\infty}$ is now completely formulated. Given an initial probability vector $p(0)$, it follows that $p(\Delta t) = P(\Delta t)p(0)$. The identity (4) expressed in matrix and vector notation is

$$p(t+\Delta t) = P(\Delta t)p(t) = P^{n+1}(\Delta t)p(0), \tag{5}$$

where $t = n\Delta t$.

Difference equations for the mean and the higher order moments of the epidemic process can be obtained directly from the difference equations in (4). For example, the expected value for $\mathcal{I}(t)$ is $E(\mathcal{I}(t)) = \sum_{i=0}^{N} ip_i(t)$. Multiplying equation (4) by i and summing on i leads to

$$\begin{aligned} E(\mathcal{I}(t+\Delta t)) &= \sum_{i=0}^{N} ip_i(t+\Delta t) \\ &= \sum_{i=1}^{N} ip_{i-1}(t)b(i-1)\Delta t + \sum_{i=0}^{N-1} ip_{i+1}(t)d(i+1)\Delta t \\ &\quad + \sum_{i=0}^{N} ip_i(t) - \sum_{i=0}^{N} ip_i(t)b(i)\Delta t - \sum_{i=0}^{N} ip_i(t)d(i)\Delta t. \end{aligned}$$

Simplifying and substituting the expressions $\beta i(N-i)/N$ and $(b+\gamma)i$ for $b(i)$ and $d(i)$, respectively, yields

$$\begin{aligned} E(\mathcal{I}(t+\Delta t)) &= E(\mathcal{I}(t)) + \sum_{i=1}^{N} p_{i-1}(t)\frac{\beta(i-1)(N-[i-1])}{N}\Delta t \\ &\quad - \sum_{i=0}^{N-1} p_{i+1}(t)(b+\gamma)(i+1)\Delta t \\ &= E(\mathcal{I}(t)) + [\beta-(b+\gamma)]\Delta t E(\mathcal{I}(t)) - \frac{\beta}{N}\Delta t E(\mathcal{I}^2(t)), \end{aligned}$$

where $E(\mathcal{I}^2(t)) = \sum_{i=0}^{N} i^2 p_i(t)$ (see e.g., [8]). The difference equation for the mean depends on the second moment. Difference equations for the second and the higher order moments depend on even higher order moments. Therefore,

these equations cannot be solved unless some additional assumptions are made regarding the higher order moments. However, because $E(\mathcal{I}^2(t)) \geq E^2(\mathcal{I}(t))$, the mean satisfies the following inequality:

$$\frac{E(\mathcal{I}(t+\Delta t)) - E(\mathcal{I}(t))}{\Delta t} \leq [\beta - (b+\gamma)]\, E(\mathcal{I}(t)) - \frac{\beta}{N} E^2(\mathcal{I}(t)). \tag{6}$$

As $\Delta t \to 0$,

$$\begin{aligned} \frac{dE(\mathcal{I}(t))}{dt} &\leq [\beta - (b+\gamma)]\, E(\mathcal{I}(t)) - \frac{\beta}{N} E^2(\mathcal{I}(t)) \\ &= \frac{\beta}{N}\,[N - E(\mathcal{I}(t))]\, E(\mathcal{I}(t)) - (b+\gamma) E(\mathcal{I}(t)) \end{aligned} \tag{7}$$

The right side of (7) is the same as the differential equation for $I(t)$ in (1), if, in equation (1), $I(t)$ and $S(t)$ are replaced by $E(\mathcal{I}(t))$ and $N - E(\mathcal{I}(t))$, respectively. The differential inequality implies that the mean of the random variable $\mathcal{I}(t)$ in the stochastic SIS epidemic process is less than the solution $I(t)$ of the deterministic differential equations in (1).

Some properties of the DTMC SIS epidemic model follow easily from Markov chain theory [6, 47]. States are classified according to their connectedness in a directed graph or digraph. The digraph of the SIS Markov chain model is illustrated in Fig. 3, where $i = 0, 1, \ldots, N$ are the infected states. The

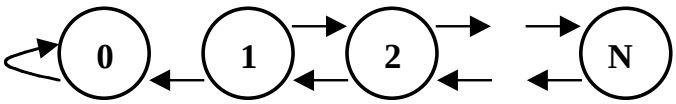

Fig. 3. Digraph of the stochastic SIS epidemic model.

states $\{0, 1, \ldots, N\}$ can be divided into two sets consisting of the recurrent state, $\{0\}$, and the transient states, $\{1, \ldots, N\}$. The zero state is an *absorbing state*. It is clear from the digraph that beginning from state 0 no other state can be reached; the set $\{0\}$ is closed. In addition, any state in the set $\{1, 2, \ldots, N\}$ can be reached from any other state in the set, but the set is not closed because $p_{01}(\Delta t) > 0$. For transient states it can be shown that elements of the transition matrix have the following property [6, 47]: Let $P^n = (p_{ij}^{(n)})$, where $p_{ij}^{(n)}$ is the (i, j) element of the nth power of the transition matrix, P^n, then

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0$$

for any state j and any transient state i. The limit of P^n as $n \to \infty$ is a stochastic matrix; all rows are zero except the first one which has all ones. From the relationship (5) and Markov chain theory, it follows that

$$\lim_{t\to\infty} p(t) = (1, 0, \ldots, 0)^T,$$

where $t = n\Delta t$. The preceding result implies, in the DTMC SIS epidemic model, the population approaches the disease-free equilibrium (probability of absorption is one), regardless of the magnitude of the basic reproduction number. Compare this stochastic result with the asymptotic dynamics of the deterministic SIS epidemic model (Theorem 1). Because this stochastic result is asymptotic, the rate of convergence to the disease-free equilibrium can be very slow. The mean time until the disease-free equilibrium is reached (absorption) depends the initial conditions and the parameter values, but can be extremely long (as shown in the numerical example in the next section). The expected duration of an epidemic (mean time until absorption) and the probability distribution conditioned on nonabsorption are discussed in Sect. 6.

3.2 Numerical Example

A *sample path* or *stochastic realization* of the stochastic process $\{\mathcal{I}(t)\}_{t=0}^{\infty}$ for $t \in \{0, \Delta t, 2\Delta t, \ldots\}$ is an assignment of a possible value to $\mathcal{I}(t)$ based on the probability vector $p(t)$. A sample path is a function of time, so that it can be plotted against the solution of the deterministic model. For illustrative purposes, we choose a population size of $N = 100$, $\Delta t = 0.01$, $\beta = 1$, $b = 0.25$, $\gamma = 0.25$ and an initial infected population size of $I(0) = 2$. In terms of the stochastic model,

$$\text{Prob}\{\mathcal{I}(0) = 2\} = 1.$$

In this example, the basic reproduction number is $\mathcal{R}_0 = 2$. The deterministic solution approaches an endemic equilibrium given by $\bar{I} = 50$.

Three sample paths of the stochastic model are compared to the deterministic solution in Fig. 4. One of the sample paths is absorbed before 200 time steps (the population following this path becomes disease-free) but two sample paths are not absorbed during 2000 time steps. These latter sample paths follow more closely the dynamics of the deterministic solution. The horizontal axis is the number of time steps Δt. For $\Delta t = 0.01$ and 2000 time steps, the solutions in Fig. 4 are graphed over the time interval $[0, 20]$. Each sample path is not continuous because at each time step, $t = \Delta t, 2\Delta t, \ldots$, the sample path either stays constant (no change in state with probability $1 - [b(i) + d(i)]\Delta t$), jumps down one integer value (with probability $d(i)\Delta t$), or jumps up one integer value (with probability $b(i)\Delta t$). For convenience, these jumps are connected with vertical line segments. Each sample path is continuous from the right but not from the left.

The entire probability distribution, $p(t)$, $t = 0, \Delta t, \ldots$, associated with this particular stochastic process can be obtained by applying (5). A MatLab program is provided in the last section that generates the probability distribution as a function of time (Fig. 5). Note that the probability distribution is bimodal, part of the distribution is at zero and the remainder of the distribution follows a path similar to the deterministic solution. Eventually, the probability distribution at zero approaches one. This bimodal distribution is important; the part of the distribution that does not approach zero (at

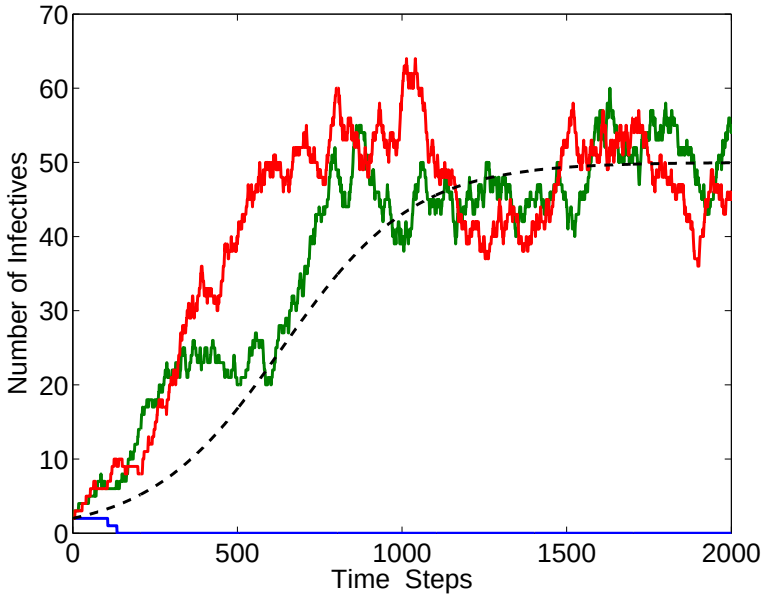

Fig. 4. Three sample paths of the DTMC SIS epidemic model are graphed with the deterministic solution (*dashed curve*). The parameter values are $\Delta t = 0.01$, $N = 100$, $\beta = 1$, $b = 0.25$, $\gamma = 0.25$, and $I(0) = 2$.

time step 2000) is known as the quasistationary probability distribution (see Sect. 6.2).

3.3 SIR Epidemic Model

Let $\mathcal{S}(t)$, $\mathcal{I}(t)$, and $\mathcal{R}(t)$ denote discrete random variables for the number of susceptible, infected, and immune individuals at time t, respectively. The DTMC SIR epidemic model is a bivariate process because there are two independent random variables, $\mathcal{S}(t)$ and $\mathcal{I}(t)$. The random variable $\mathcal{R}(t) = N - \mathcal{S}(t) - \mathcal{I}(t)$. The bivariate process $\{(\mathcal{S}(t), \mathcal{I}(t))\}_{t=0}^{\infty}$ has a joint probability function given by

$$p_{(s,i)}(t) = \text{Prob}\{\mathcal{S}(t) = s, \mathcal{I}(t) = i\}.$$

This bivariate process has the Markov property and is time-homogeneous.

Transition probabilities can be defined based on the assumptions in the SIR deterministic formulation. First, assume that Δt can be chosen sufficiently small such that at most one change in state occurs during the time interval Δt. In particular, there can be either a new infection, a birth, a death, or a recovery. The transition probabilities are denoted as follows:

$$p_{(s+k,i+j),(s,i)}(\Delta t) = \text{Prob}\{(\Delta\mathcal{S}, \Delta\mathcal{I}) = (k, j) | (\mathcal{S}(t), \mathcal{I}(t)) = (s, i)\},$$

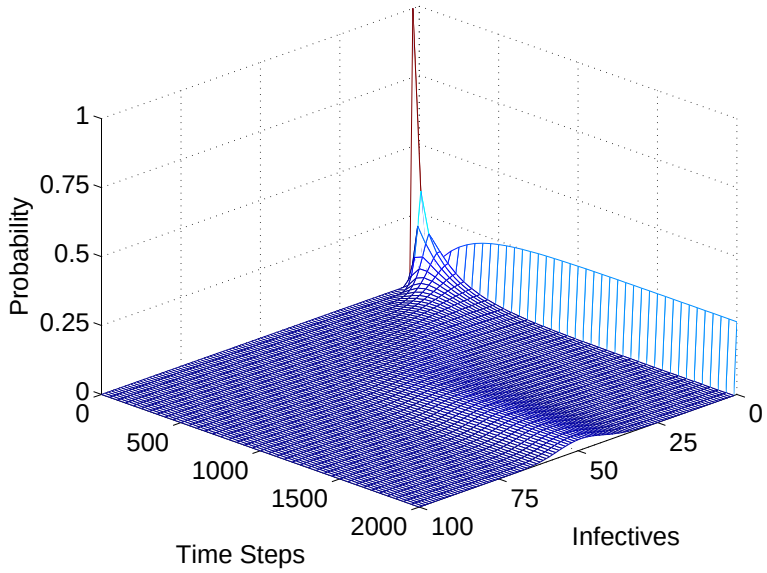

Fig. 5. Probability distribution of the DTMC SIS epidemic model. Parameter values are the same as in Fig. 4.

where $\Delta\mathcal{S} = \mathcal{S}(t+\Delta t) - \mathcal{S}(t)$. Hence,

$$p_{(s+k,i+j),(s,i)}(\Delta t) = \begin{cases} \beta is/N\Delta t, & (k,j) = (-1,1) \\ \gamma i\Delta t, & (k,j) = (0,-1) \\ bi\Delta t, & (k,j) = (1,-1) \\ b(N-s-i)\Delta t, & (k,j) = (1,0) \\ 1 - \beta is/N\Delta t & \\ \quad - [\gamma i + b(N-s)]\Delta t, & (k,j) = (0,0) \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

The time step Δt must be chosen sufficiently small such that each of the transition probabilities lie in the interval $[0,1]$. Because the states are now ordered pairs, the transition matrix is more complex than for the SIS epidemic model and its form depends on how the states (s,i) are ordered. However, applying the Markov property, the difference equation satisfied by the probability $p_{(s,i)}(t+\Delta t)$ can be expressed in terms of the transition probabilities,

$$\begin{aligned} p_{(s,i)}(t+\Delta t) = {} & p_{(s+1,i-1)}(t)\frac{\beta}{N}(i-1)(s+1)\Delta t + p_{(s,i+1)}(t)\gamma(i+1)\Delta t \\ & + p_{(s-1,i+1)}(t)b(i+1)\Delta t + p_{(s-1,i)}(t)b(N-s+1-i)\Delta t \\ & + p_{(s,i)}(t)\left(1 - \left[\frac{\beta}{N}is + \gamma i + b(N-s)\right]\Delta t\right). \end{aligned} \tag{9}$$

The digraph associated with the SIR Markov chain lies on a two-dimensional lattice. It is easy to show that the state $(N, 0)$ is absorbing ($p_{(N,0),(N,0)}(\Delta t) = 1$) and that all other states are transient. Thus, asymptotically, all sample paths eventually are absorbed into the disease-free state $(N, 0)$. Compare this result to the deterministic SIR epidemic model (Theorem 2).

Difference equations for the mean and higher order moments can be derived from (9) as was done for the SIS epidemic model, e.g., $E(\mathcal{S}(t)) = \sum_{s=0}^{N} s p_{(s,i)}(t)$ and $E(\mathcal{I}(t)) = \sum_{i=0}^{N} i p_{(s,i)}(t)$. However, these difference equations cannot be solved directly because they depend on higher order moments.

3.4 Numerical Example

Three sample paths of the DTMC SIR model are compared to the solution of the deterministic model in Fig. 6. In this example, $\Delta t = 0.01$, $N = 100$, $\beta = 1$, $b = 0$, $\gamma = 0.5$, and $(S(0), I(0)) = (98, 2)$. In the stochastic model,

$$\text{Prob}\{(\mathcal{S}(0), \mathcal{I}(0)) = (98, 2)\} = 1.$$

The basic reproduction number and the initial replacement number are both greater than one; $\mathcal{R}_0 = 2$ and $\mathcal{R}_0 S(0)/N = 1.96$. According to Theorem 2 part iii), there is an epidemic (an increase in the number of cases). The epidemic is easily seen in the behavior of the deterministic solution. Each of the three sample paths also illustrate an epidemic curve.

4 Formulation of CTMC Epidemic Models

The CTMC epidemic processes are defined on a continuous time scale, $t \in [0, \infty)$, but the states $\mathcal{S}(t)$, $\mathcal{I}(t)$, and $\mathcal{R}(t)$ are discrete random variables, i.e.,

$$\mathcal{S}(t),\ \mathcal{I}(t),\ \mathcal{R}(t) \in \{0, 1, 2, \ldots, N\}.$$

4.1 SIS Epidemic Model

In the CTMC SIS epidemic model, the stochastic process depends on the collection of discrete random variables $\{\mathcal{I}(t)\}$, $t \in [0, \infty)$ and their associated probability functions $p(t) = (p_0(t), \ldots, p_N(t))^T$, where

$$p_i(t) = \text{Prob}\{\mathcal{I}(t) = i\}.$$

The stochastic process has the *Markov property*, that is,

$$\text{Prob}\{\mathcal{I}(t_{n+1})|\mathcal{I}(t_0), \mathcal{I}(t_1), \ldots, \mathcal{I}(t_n)\} = \text{Prob}\{\mathcal{I}(t_{n+1})|\mathcal{I}(t_n)\}$$

for any sequence of real numbers satisfying $0 \le t_0 < t_1 < \cdots < t_n < t_{n+1}$. The transition probability at time t_{n+1} only depends on the most recent time t_n.

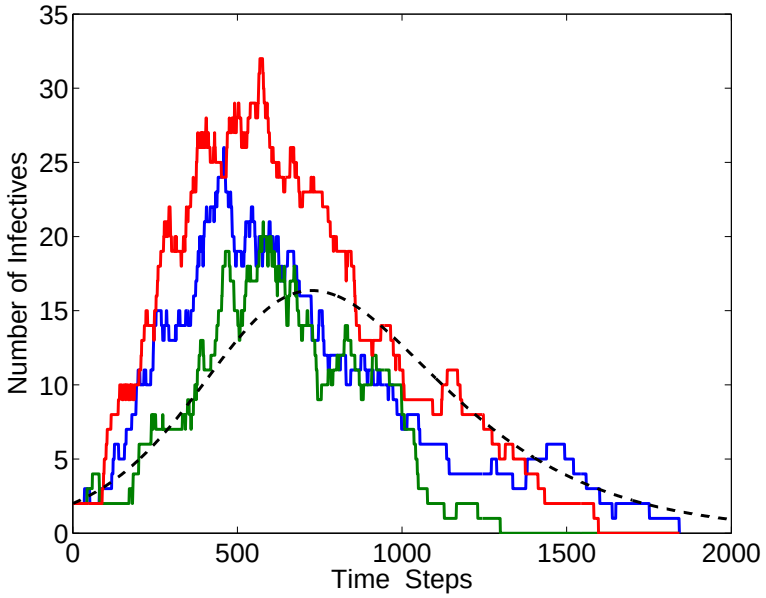

Fig. 6. Three sample paths of the DTMC SIR epidemic model are graphed with the deterministic solution (*dashed curve*). The parameter values are $\Delta t = 0.01$, $N = 100$, $\beta = 1$, $b = 0$, $\gamma = 0.5$, $S(0) = 98$, and $I(0) = 2$.

The transition probabilities are defined for a small time interval Δt. But in a CTMC model, the transition probabilities are referred to as *infinitesimal transition probabilities* because they are valid for sufficiently small Δt. Therefore, the term $o(\Delta t)$ is included in the definition $[\lim_{t\to\infty}(o(\Delta t)/\Delta t) = 0]$. The infinitesimal transition probabilities are defined as follows:

$$p_{ji}(\Delta t) = \begin{cases} \dfrac{\beta}{N} i(N-i)\Delta t + o(\Delta t), & j = i+1 \\ (b+\gamma)i\Delta t + o(\Delta t), & j = i-1 \\ 1 - \left[\dfrac{\beta}{N} i(N-i) + (b+\gamma)i\right]\Delta t + o(\Delta t), & j = i \\ o(\Delta t), & \text{otherwise,} \end{cases}$$

Because Δt is sufficiently small, there are only three possible changes in states:

$$i \to i+1, \quad i \to i-1, \text{ or } \quad i \to i.$$

Using the same notation as for the DTMC model, let $b(i)$ denote a birth (new infection) and $d(i)$ denote a death or recovery. Then

$$p_{ji}(\Delta t) = \begin{cases} b(i)\Delta t + o(\Delta t), & j = i+1 \\ d(i)\Delta t + o(\Delta t), & j = i-1 \\ 1 - [b(i) + d(i)]\Delta t + o(\Delta t), & j = i \\ o(\Delta t), & \text{otherwise.} \end{cases}$$

Applying the Markov property and the infinitesimal transitional probabilities, a continuous time analogue of the transition matrix can be defined. Instead of a system of difference equations, a system of differential equations is obtained. Assume $\text{Prob}\{\mathcal{I}(0) = i_0\} = 1$. Then $p_{i,i_0}(\Delta t) = p_i(\Delta t)$ and

$$\begin{aligned} p_i(t+\Delta t) = p_{i-1}(t)b(i-1)\Delta t + p_{i+1}(t)d(i+1)\Delta t \\ + p_i(t)(1-[b(i)+d(i)]\Delta t) + o(\Delta t). \end{aligned}$$

These equations are the same as the DTMC equations (4), except $o(\Delta t)$ is added to the right side. Subtracting $p_i(t)$, dividing by Δt, and letting $\Delta t \to 0$, leads to

$$\frac{dp_i}{dt} = p_{i-1}b(i-1) + p_{i+1}d(i+1) - p_i[b(i)+d(i)] \tag{10}$$

for $i = 1, 2, \ldots, N$ and $dp_0/dt = p_1 d(1)$. These latter equations are known as the *forward Kolmogorov differential equations* [47]. In matrix notation, they can be expressed as

$$\frac{dp}{dt} = Qp, \tag{11}$$

where $p(t) = (p_0(t), \ldots, p_N(t))^T$ and matrix Q is defined as follows:

$$Q = \begin{pmatrix} 0 & d(1) & 0 & \cdots & 0 \\ 0 & -[b(1)+d(1)] & d(2) & \cdots & 0 \\ 0 & b(1) & -[b(2)+d(2)] & \cdots & 0 \\ 0 & 0 & b(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & d(N) \\ 0 & 0 & 0 & \cdots & -d(N) \end{pmatrix},$$

$b(i) = \beta i(N-i)/N$ and $d(i) = (b+\gamma)i$. Matrix Q is referred to as the *infinitesimal generator matrix* or simply the *generator matrix* [6, 47], More generally, the differential equations $dP/dt = QP$ are known as the forward Kolmogorov differential equations, where $P \equiv (p_{ji}(t))$ is the matrix of infinitesimal transition probabilities. It is interesting to note that the transition matrix $P(\Delta t)$ of the DTMC model and the generator matrix Q are related as follows:

$$Q = \lim_{\Delta t \to 0} \frac{P(\Delta t) - I}{\Delta t}.$$

The generator matrix Q has a zero eigenvalue with corresponding eigenvector $(1, 0, \ldots, 0)^T$. The remaining eigenvalues are negative or have negative real part. This can be seen by applying Gershgorin's circle theorem and the fact that the submatrix $\tilde{Q}$ of Q, where the first row and the first column are deleted, is nonsingular [43]. Therefore, $\lim_{t\to\infty} p(t) = (1, 0, 0, \ldots, 0)^T$. Eventual absorption occurs in the CTMC SIS epidemic model. Compare this stochastic result with Theorem 1.

Differential equations for the mean and higher order moments of $\mathcal{I}(t)$ can be derived from the differential equations (11). As was shown for the DTMC epidemic model, the differential equations (10) can be multiplied by i, then summed over i. However, we present an alternate method for obtaining the differential equations for the mean and higher order moments using generating functions. Either the probability generating function (pgf) or the moment generating function (mgf) can be used to derive the equations. The pgf for $\mathcal{I}(t)$ is defined as

$$\mathcal{P}(\theta, t) = E(\theta^{\mathcal{I}(t)}) = \sum_{i=0}^{N} p_i(t)\theta^i$$

and the mgf as

$$M(\theta, t) = E(e^{\theta\mathcal{I}(t)}) = \sum_{i=0}^{N} p_i(t)e^{i\theta}.$$

We use the mgf to derive the equations because the method of derivation is simpler than with the pgf. In addition, the moments of the distribution corresponding to $\mathcal{I}(t)$ can be easily calculated from the mgf,

$$\left.\frac{\partial^k M}{\partial\theta^k}\right|_{\theta=0} = E(\mathcal{I}^k(t))$$

for $k = 1, \ldots, n$.

First, we derive a differential equation satisfied by the mgf. Multiplying the equations in (10) by $e^{i\theta}$ and summing on i, leads to

$$\begin{aligned}\frac{\partial M}{\partial t} &= \sum_{i=0}^{N} \frac{dp_i}{dt} e^{i\theta} \\ &= e^{\theta} \sum_{i=1}^{N} p_{i-1} e^{(i-1)\theta} b(i-1) + e^{-\theta} \sum_{i=0}^{N-1} p_{i+1} e^{(i+1)\theta} d(i+1) \\ &\quad - \sum_{i=0}^{N} p_i e^{i\theta}[b(i) + d(i)].\end{aligned}$$

Simplifying and substituting $\beta i(N-i)/N$ and $(b+\gamma)i$ for $b(i)$ and $d(i)$, respectively, yields

$$\begin{aligned}\frac{\partial M}{\partial t} &= \beta(e^{\theta} - 1) \sum_{i=1}^{N} i p_i e^{i\theta} + (b+\gamma)(e^{-\theta} - 1) \sum_{i=1}^{N} i p_i e^{i\theta} \\ &\quad - \frac{\beta}{N}(e^{\theta} - 1) \sum_{i=1}^{N} i^2 p_i e^{i\theta}.\end{aligned}$$

The summations in the previous expression can be replaced with $\partial M/\partial\theta$ or $\partial^2 M/\partial\theta^2$ so that the following second order partial differential equation is obtained for the mgf:

$$\frac{\partial M}{\partial t} = [\beta(e^{\theta} - 1) + (b + \gamma)(e^{-\theta} - 1)]\frac{\partial M}{\partial \theta} - \frac{\beta}{N}(e^{\theta} - 1)\frac{\partial^2 M}{\partial \theta^2}. \quad (12)$$

Bailey [13] derives a general expression for the partial differential equation satisfied by the mgf (and also the pgf) based on the infinitesimal transition probabilities for the process.

The partial differential equation for the mgf, equation (12), is used to obtain an ordinary differential equation satisfied by the mean of $\mathcal{I}(t)$. Differentiating the equation (12) with respect to θ and evaluating at $\theta = 0$ yields an ordinary differential equation satisfied by the mean $E(\mathcal{I}(t))$,

$$\frac{dE(\mathcal{I}(t))}{dt} = [\beta - (b + \gamma)]E(\mathcal{I}(t)) - \frac{\beta}{N}E(\mathcal{I}^2(t)).$$

Because the differential equation for the mean depends on the second moment, it cannot be solved directly, but as was shown for the DTMC SIS epidemic model in (7), the mean of the stochastic SIS epidemic model is less than the deterministic solution. The differential equations for the second moment and for the variance depend on higher order moments. These higher order moments are often approximated by lower order moments by making some assumptions regarding their distributions (e.g., normality or lognormality), referred to as moment closure techniques (see e.g., [27, 34]). Then these differential equations can be solved to give approximations for the moments.

4.2 Numerical Example

To numerically compute a sample path of a CTMC model, we need to use the fact that the interevent time has an exponential distribution. This follows from the Markov property. The exponential distribution has what has been called the "memoryless property".

Assume $\mathcal{I}(t) = i$. Let T_i denote the interevent time, a continuous random variable for the time to the next event given the process is in state i. Let $H_i(t)$ denote the probability the process remains in state i for a period of time t. Then $H_i(t) = \text{Prob}\{T_i > t\}$. It follows that

$$H_i(t + \Delta t) = H_i(t)p_{ii}(\Delta t) = H_i(t)(1 - [b(i) + d(i)]\Delta t) + o(\Delta t).$$

Subtracting $H_i(t)$ and dividing by Δt, the following differential equation is obtained:

$$\frac{dH_i}{dt} = -[b(i) + d(i)]H_i.$$

Since $H_i(0) = 1$, the solution to the differential equation is $H_i(t) = \exp(-[b(i)+ d(i)]t)$. Therefore, the interevent time T_i is an exponential random variable with parameter $b(i) + d(i)$. The cumulative distribution of T_i is

$$F_i(t) = \text{Prob}\{T_i \leq t\} = 1 - \exp(-[b(i) + d(i)]t)$$

[6, 47].

The uniform random variable on $[0, 1]$ can be applied for numerical computation of the interevent time. Let U be a uniform random variable on $[0, 1]$. Then

$$\begin{aligned}\text{Prob}\{F_i^{-1}(U) \le t\} &= \text{Prob}\{F_i(F_i^{-1}(U)) \le F_i(t)\} \\ &= \text{Prob}\{U \le F_i(t)\} = F_i(t)\end{aligned}$$

The interevent time T_i, given $\mathcal{I}(t) = i$, satisfies

$$T_i = F_i^{-1}(U) = -\frac{\ln(1-U)}{b(i)+d(i)} = -\frac{\ln(U)}{b(i)+d(i)}.$$

In Fig. 7, three sample paths for the CTMC SIS epidemic model are compared to the deterministic solution. Parameter values are $b = 0.25$, $\gamma = 0.25$, $\beta = 1$, $N = 100$, and $I(0) = 2$. For the stochastic model,

$$\text{Prob}\{\mathcal{I}(0) = 2\} = 1.$$

The basic reproduction number is $\mathcal{R}_0 = 2$. One sample path in Fig. 7 is absorbed rapidly (the population following this path becomes disease-free). The sample paths for the CTMC model are not continuous for the same reasons given for the DTMC model. With each change, the process either jumps up one integer value (with probability $b(i)/[b(i) + d(i)]$) or jumps down one integer value (with probability $d(i)/[b(i) + d(i)]$). Sample paths are continuous from the right but not from the left. Compare the sample paths in Fig. 7 with the three sample paths in the DTMC SIS epidemic model in Fig. 4.

4.3 SIR Epidemic Model

A derivation similar to the SIS epidemic model can be applied to the SIR epidemic model. The difference, of course, is that the SIR epidemic process is bivariate, $\{(\mathcal{S}(t), \mathcal{I}(t))\}$, where $\mathcal{R}(t) = N - \mathcal{S}(t) - \mathcal{I}(t)$. Assumptions similar to those for the DTMC SIR epidemic model (8) apply to the CTMC SIR epidemic model, except that $o(\Delta t)$ is added to each of the infinitesimal transition probabilities.

For the bivariate process, a joint probability function is associated with each pair of random variables $(\mathcal{S}(t), \mathcal{I}(t))$, $p_{(s,i)}(t) = \text{Prob}\{(\mathcal{S}(t), \mathcal{I}(t)) = (s, i)\}$. A system of forward Kolmogorov differential equations can be derived,

$$\begin{aligned}\frac{dp_{(s,i)}}{dt} &= p_{(s+1,i-1)}\frac{\beta}{N}(i-1)(s+1) + p_{(s,i+1)}\gamma(i+1) \\ &\quad + p_{(s-1,i+1)}b(i+1) + p_{(s-1,i)}b(N-s+1-i) \\ &\quad - p_{(s,i)}\left[\frac{\beta}{N}is + \gamma i + b(N-s)\right].\end{aligned}$$

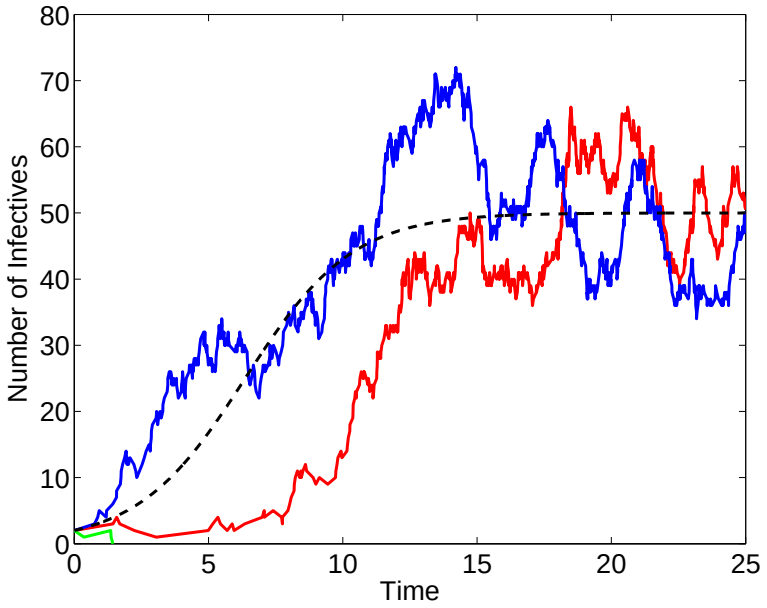

Fig. 7. Three samples paths of the CTMC SIS epidemic model are graphed with the deterministic solution (*dashed curve*). The parameter values are $b = 0.25$, $\gamma = 0.25$, $\beta = 1$, $N = 100$, and $I(0) = 2$. Compare with Fig. 4.

These differential equations are the limiting equations (as $\Delta t \to 0$) of the difference equations in (9). Differential equations for the mean and higher order moments can be derived. However, as was true for the other epidemic processes, they do not form a closed system, i.e., each successive moment depends on higher order moments. Moment closure techniques can be applied to approximate the solutions to these moment equations [27, 34].

The SIR epidemic process is Markovian and time homogeneous. In addition, the disease-free state is an absorbing state. In Sect. 6.3, we discuss the final size of the epidemic, which is applicable to the deterministic and stochastic SIR epidemic model in the case $\mathcal{R}_0 > 1$ and $b = 0$ (Theorem 2, part iii)).

5 Formulation of SDE Epidemic Models

Assume the time variable is continuous, $t \in [0, \infty)$ and the states $\mathcal{S}(t)$, $\mathcal{I}(t)$, and $\mathcal{R}(t)$ are continuous random variables, that is,

$$\mathcal{S}(t), \mathcal{I}(t), \mathcal{R}(t) \in [0, N].$$

5.1 SIS Epidemic Model

The stochastic SIS epidemic model depends on the number of infectives, $\{\mathcal{I}(t)\}$, $t \in [0, \infty)$, where $\mathcal{I}(t)$ has an associated probability density function (pdf), $p(x,t)$,

$$\text{Prob}\{a \le \mathcal{I}(t) \le b\} = \int_a^b p(x,t)dx.$$

The stochastic SIS epidemic model has the *Markov property*, i.e.,

$$\text{Prob}\{\mathcal{I}(t_n) \le y | \mathcal{I}(t_0), \mathcal{I}(t_1), \ldots, \mathcal{I}(t_{n-1})\} = \text{Prob}\{\mathcal{I}(t_n) \le y | \mathcal{I}(t_{n-1})\}$$

for any sequence of real numbers $0 \le t_0 < t_1 < \cdots < t_{n-1} < t_n$. Denote the transition pdf for the stochastic process as

$$p(y, t + \Delta t; x, t),$$

where at time t, $\mathcal{I}(t) = x$, and at time $t + \Delta t$, $\mathcal{I}(t + \Delta t) = y$. The process is *time homogeneous*; the transition pdf does not depend on t but does depend on the length of time, Δt. The stochastic process is referred to as a *diffusion process* if it is a Markov process in which the infinitesimal mean and variance exist. The stochastic SIS epidemic model is a time homogeneous, diffusion process. The infinitesimal mean and variance are defined next.

For the stochastic SIS epidemic model, it can be shown that the pdf satisfies a forward Kolmogorov differential equation. This equation is a second order partial differential equation [6, 21], a continuous analogue of the forward Kolmogorov differential equations for the CTMC model in (10). Assume $\text{Prob}\{\mathcal{I}(0) = i_0\} = 1$ and let $p(i, i_0; t) = p(i,t) = p_i(t)$. Then the system of differential equations in (10) can be expressed as a finite difference scheme in the variable i with $\Delta i = 1$,

$$\begin{aligned}\frac{dp_i}{dt} &= p_{i-1}b(i-1) + p_{i+1}d(i+1) - p_i[b(i) + d(i)] \\ &= -\frac{\{p_{i+1}[b(i+1) - d(i+1)] - p_{i-1}[b(i-1) - d(i-1)]\}}{2\Delta i} \\ &\quad + \frac{1}{2}\frac{\{p_{i+1}[b(i+1) + d(i+1)] - 2p_i[b(i) + d(i)] + p_{i-1}[b(i-1) + d(i-1)]\}}{(\Delta i)^2}.\end{aligned}$$

Let $i = x$, $\Delta i = \Delta x$ and $p_i(t) = p(x,t)$. The limiting form of the preceding equation (as $\Delta x \to 0$) is the forward Kolmogorov differential equation for $p(x,t)$:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}\{[b(x) - d(x)]p(x,t)\} + \frac{1}{2}\frac{\partial^2}{\partial x^2}\{[b(x) + d(x)]\, p(x,t)\}.$$

Substituting $b(x) = \beta x(N - x)/N$ and $d(x) = (b + \gamma)x$ yields

$$\frac{\partial p(x,t)}{\partial t} = \frac{\partial}{\partial x}\left\{\left[\frac{\beta}{N}x(N-x) - (b+\gamma)x\right] p(x,t)\right\}$$

$$+\frac{1}{2}\frac{\partial^2}{\partial x^2}\left\{\left[\frac{\beta}{N}x(N-x)+(b+\gamma)x\right]p(x,t)\right\}.$$

The coefficient in the first term on the right side of the preceding equation, $[\beta x(N-x)/N-(b+\gamma)x]$, is the infinitesimal mean and the coefficient in the second term, $[\beta x(N-x)/N+(b+\gamma)x]$, is the infinitesimal variance. More generally, the forward Kolmogorov differential equations are expressed in terms of the transition probabilities, $p(y,s;x,t)$. To solve the differential equation requires boundary conditions for $x=0,N$ and initial conditions for $t=0$. An explicit solution is not possible because of the nonlinearities. We derive a SDE that is much simpler to solve numerically and whose solution is a sample path of the stochastic process.

A SDE for the SIS epidemic model can be derived from the CTMC SIS epidemic model [5]. The assumptions in the CTMC SIS epidemic model are restated in terms of $\Delta\mathcal{I}=\mathcal{I}(t+\Delta t)-\mathcal{I}(t)$. Assume

$$\text{Prob}\{\Delta\mathcal{I}=j|\mathcal{I}(t)=i\}=\begin{cases}b(i)\Delta t+o(\Delta t), & j=i+1\\ d(i)\Delta t+o(\Delta t), & j=i-1\\ 1-[b(i)+d(i)]\Delta t+o(\Delta t), & j=i\\ o(\Delta t), & j\neq i+1,i-1,i\end{cases}$$

In addition, assume that $\Delta\mathcal{I}$ has an approximate normal distribution for small Δt. The expectation and the variance of $\Delta\mathcal{I}$ are computed.

$$\begin{aligned}E(\Delta\mathcal{I})&=b(\mathcal{I})\Delta t-d(\mathcal{I})\Delta t+o(\Delta t)\\&=[b(\mathcal{I})-d(\mathcal{I})]\Delta t+o(\Delta t)=\mu(\mathcal{I})\Delta t+o(\Delta t).\end{aligned}$$

$$\begin{aligned}Var(\Delta\mathcal{I})&=E(\Delta\mathcal{I})^2-[E(\Delta\mathcal{I})]^2\\&=b(\mathcal{I})\Delta t+d(\mathcal{I})\Delta t+o(\Delta t)\\&=[b(\mathcal{I})+d(\mathcal{I})]\Delta t+o(\Delta t)=\sigma^2(\mathcal{I})\Delta t+o(\Delta t),\end{aligned}$$

where the notation means $b(\mathcal{I})=\beta i(N-i)/N$ and $d(\mathcal{I})=(b+\gamma)i$ given that $\mathcal{I}(t)=i$. Because the random variable $\Delta\mathcal{I}$ is approximately normally distributed, $\Delta\mathcal{I}(t)\sim\mathbf{N}(\mu(\mathcal{I})\Delta t,\sigma^2(\mathcal{I})\Delta t)$,

$$\begin{aligned}\mathcal{I}(t+\Delta t)&=\mathcal{I}(t)+\Delta\mathcal{I}(t)\\&\approx\mathcal{I}(t)+\mu(\mathcal{I})\Delta t+\sigma(\mathcal{I})\sqrt{\Delta t}\,\eta,\end{aligned}$$

where $\eta\sim\mathbf{N}(0,1)$.

The difference equation $\mathcal{I}(t+\Delta t)=\mathcal{I}(t)+\mu(\mathcal{I})\Delta t+\sigma(\mathcal{I})\sqrt{\Delta t}\,\eta$ is Euler's method applied to the following Itô SDE:

$$\frac{d\mathcal{I}}{dt}=\mu(\mathcal{I})+\sigma(\mathcal{I})\frac{dW}{dt},$$

where W is the Wiener process, $W(t+\Delta t)-W(t)\sim\mathbf{N}(0,\Delta t)$ [21, 31, 32]. Euler's method converges to the Itô SDE provided the coefficients, $\mu(\mathcal{I})$ and

$\sigma(\mathcal{I})$, satisfy certain smoothness and growth conditions [31, 32]. The coefficients for the stochastic SIS epidemic model are $\mu(\mathcal{I}) = b(\mathcal{I}) - d(\mathcal{I})$ and $\sigma(\mathcal{I}) = \sqrt{b(\mathcal{I}) + d(\mathcal{I})}$, where

$$b(\mathcal{I}) = \frac{\beta}{N}\mathcal{I}(N - \mathcal{I}) \quad \text{and} \quad d(\mathcal{I}) = (b + \gamma)\mathcal{I}.$$

Substituting these values into the Itô SDE gives the SDE SIS epidemic model,

$$\frac{d\mathcal{I}}{dt} = \frac{\beta}{N}\mathcal{I}(N - \mathcal{I}) - (b + \gamma)\mathcal{I} + \sqrt{\frac{\beta}{N}\mathcal{I}(N - \mathcal{I}) + (b + \gamma)\mathcal{I}}\frac{dW}{dt}. \tag{13}$$

From the Itô SDE, it can be seen that when $\mathcal{I}(t) = 0$, $d\mathcal{I}/dt = 0$. The disease-free equilibrium is an absorbing state for the Itô SDE.

We digress briefly to discuss the Wiener process $\{W(t)\}$, $t \in [0, \infty)$. The Wiener process depends continuously on t, $W(t) \in (-\infty, \infty)$. It is a diffusion process, but has some additional nice properties. The Wiener process has stationary, independent increments, that is, the increments ΔW depend only on Δt. They are independent of t and the value of $W(t)$:

$$\Delta W = W(t + \Delta t) - W(t) \sim \mathbf{N}(0, \Delta t).$$

Two sample paths of a Wiener process are graphed in Fig. 8.

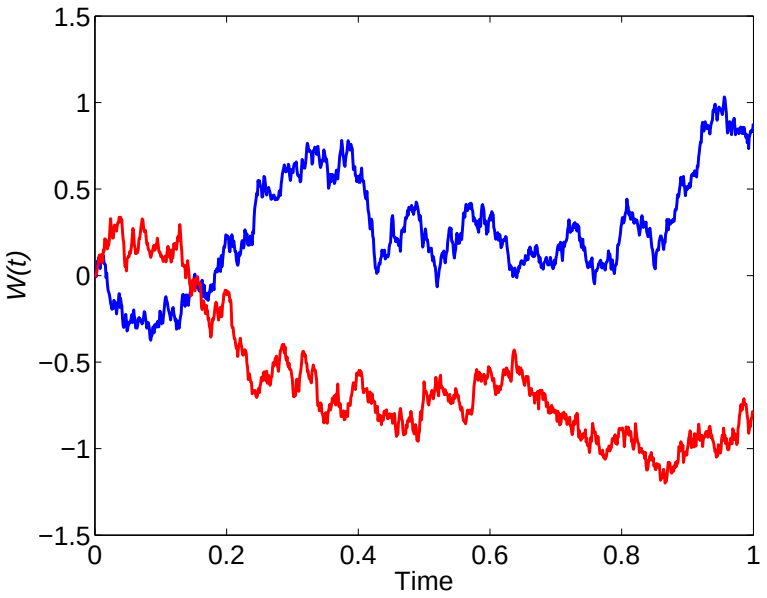

Fig. 8. Two sample paths of a Wiener process.

The notation $dW(t)/dt$ is only for convenience because sample paths of $W(t)$ are continuous but nowhere differentiable [12, 21]. The Itô SDE (13) should be expressed as a stochastic integral equation but the SDE notation is standard.

5.2 Numerical Example

Three sample paths of the SDE SIS epidemic model are graphed in Fig. 9. The parameter values are $b = 0.25$, $\gamma = 0.25$, $\beta = 1$, and $N = 100$. The initial condition is $I(0) = 2$. For the stochastic model the pdf for the initial condition is $p(x, 0) = 2\delta(x - 2)$, where $\delta(x)$ is the Dirac delta function. The basic reproduction number is $\mathcal{R}_0 = 2$, so that the deterministic solution approaches the endemic equilibrium $\bar{I} = 50$. The MatLab program which generated these sample paths is given in the last section. Compare the sample paths of the Itô SDE in Fig. 9 with those for the DTMC and the CTMC models in Figs. 4 and 7. The sample paths for the Itô SDE are continuous, whereas the sample paths of the DTMC and the CTMC models are discontinuous.

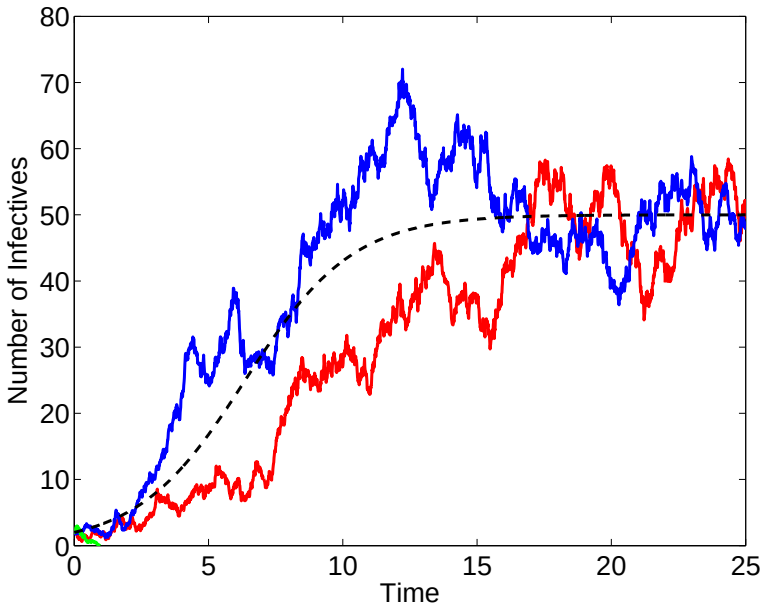

Fig. 9. Three sample paths of the SDE SIS epidemic model are graphed with the deterministic solution (*dashed curve*). The parameter values are $b = 0.25$, $\gamma = 0.25$, $\beta = 1$, $N = 100$, $I(0) = 2$. Compare with Figs. 4 and 7.

5.3 SIR Epidemic Model

A derivation similar to the Itô SDE for the SIS epidemic model can be applied to the bivariate process $\{(\mathcal{S}(t), \mathcal{I}(t))\}$ [5, 6]. Similar assumptions are made regarding the change in the random variables, $\Delta\mathcal{S}$ and $\Delta\mathcal{I}$, as in the transition probabilities for the DTMC and CTMC models. In addition, we assume that the change in these random variables is approximately normally distributed. To simplify the derivation, we assume there are no births, $b = 0$, in the SIR epidemic model.

Let $\Delta X(t) = (\Delta\mathcal{S}, \Delta\mathcal{I})^T$. Then the expectation of $\Delta X(t)$ to order Δt is

$$E(\Delta X(t)) = \begin{pmatrix} -\dfrac{\beta}{N}\mathcal{S}\mathcal{I} \\ \dfrac{\beta}{N}\mathcal{S}\mathcal{I} - \gamma\mathcal{I} \end{pmatrix} \Delta t.$$

The covariance matrix of $\Delta X(t)$ is $V(\Delta X(t)) = E(\Delta X(t)[\Delta X(t)]^T) - E(\Delta X(t))E(\Delta X(t))^T \approx E(\Delta X(t)[\Delta X(t)]^T)$ because the elements in the second term are $o([\Delta t]^2)$. Then the covariance matrix of $\Delta X(t)$ to order Δt is

$$V(\Delta X(t)) = \begin{pmatrix} \dfrac{\beta}{N}\mathcal{S}\mathcal{I} & -\dfrac{\beta}{N}\mathcal{S}\mathcal{I} \\ -\dfrac{\beta}{N}\mathcal{S}\mathcal{I} & \dfrac{\beta}{N}\mathcal{S}\mathcal{I} + \gamma\mathcal{I} \end{pmatrix} \Delta t$$

[5, 6]. The random vector $X(t + \Delta t)$ can be approximated as follows:

$$X(t + \Delta t) = X(t) + \Delta X(t) \approx X(t) + E(\Delta X(t)) + \sqrt{V(\Delta X(t))}. \qquad (14)$$

Because the covariance matrix is symmetric and positive definite, it has a unique square root $B\sqrt{\Delta t} = \sqrt{V}$ [43]. The system of equations (14) are an Euler approximation to a system of Itô SDEs. For sufficiently smooth coefficients, the solution $X(t)$ of (14) converges to the solution of the following system of Itô SDEs:

$$\begin{aligned} \frac{d\mathcal{S}}{dt} &= -\frac{\beta}{N}\mathcal{S}\mathcal{I} + B_{11}\frac{dW_1}{dt} + B_{12}\frac{dW_2}{dt} \\ \frac{d\mathcal{I}}{dt} &= \frac{\beta}{N}\mathcal{S}\mathcal{I} - \gamma\mathcal{I} + B_{21}\frac{dW_1}{dt} + B_{22}\frac{dW_2}{dt} \end{aligned}$$

where W_1 and W_2 are two independent Wiener processes and $B = (B_{ij})$ [31, 32].

5.4 Numerical Example

Three sample paths of the SDE SIR epidemic model are graphed with the deterministic solution in Fig. 10. The parameter values are $\Delta t = 0.01$, $\beta = 1$, $b = 0$, $\gamma = 0.5$, and $N = 100$ with initial condition $I(0) = 2$. The basic reproduction number and initial replacement number are $\mathcal{R}_0 = 2$ and $\mathcal{R}_0 S(0)/N = 1.96$, respectively. Compare the sample paths in Fig. 10 with the sample paths for the DTMC SIR epidemic model in Fig. 6.

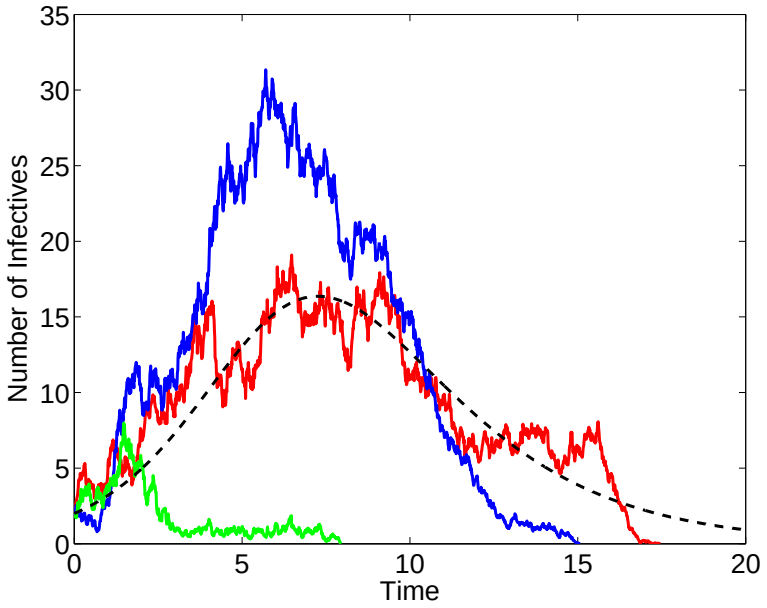

Fig. 10. Three sample paths of the SDE SIR epidemic model are graphed with the deterministic solution (*dashed curve*). The parameter values are $\Delta t = 0.01$, $\beta = 1$, $b = 0$, $\gamma = 0.5$, $N = 100$, and $I(0) = 2$. Compare with Fig. 6.

6 Properties of Stochastic SIS and SIR Epidemic Models

In the next subsections, we concentrate on some of the properties of these well-known stochastic epidemic models that distinguish them from their deterministic counterparts. Four important properties of stochastic epidemic model include the following: probability of an outbreak, quasistationary probability distribution, final size distribution of an epidemic and expected duration of an epidemic. Each of these properties depend on the stochastic nature of the process.

6.1 Probability of an Outbreak

An outbreak occurs when the number of cases escalates. A simple random walk model (DTMC) or a linear birth and death process (CTMC) on the set $\{0, 1, 2, \ldots\}$ can be used to estimate the probability of an outbreak. For example, let $X(t)$ be the random variable for the position at time t on the set $\{0, 1, 2, \ldots\}$ in a random walk model. State 0 is absorbing and the remaining states are transient. If $X(t) = x$, then in the next time interval, there is either a move to the right $x \to x+1$ with probability p or a move to the left, $x \to x-1$ with probability q, with the exception of state 0, where there is no movement

$(p+q=1)$. In the random walk model, either the process approaches state 0 or approaches infinity. The probability of absorption into state 0 depends on p, q, and the initial position. Let $X(t)=x_0>0$, then it can be shown that

$$\lim_{t\to\infty}\text{Prob}\{X(t)=0\}=\begin{cases}1, & \text{if } p\le q\\ \left(\frac{q}{p}\right)^{x_0}, & \text{if } p>q\end{cases} \tag{15}$$

(e.g., [6, 13, 45]).

The identity (15) is also valid for a linear birth and death process in a DTMC or CTMC model, where b and d are replaced by λi and μi, where i is the position. In the linear birth an death process, the infinitesimal transition probabilities satisfy

$$p_{i+j,i}(\Delta t)=\begin{cases}\lambda i\Delta t+o(\Delta t), & j=1\\ \mu i\Delta t+o(\Delta t), & j=-1\\ 1-(\lambda+\mu)i\Delta t+o(\Delta t), & j=0.\end{cases}$$

The identity (15) holds with λ replacing p and μ replacing q. The probability of absorption is one if $\lambda\le\mu$. But if $\lambda>\mu$ the probability of absorption decreases to $(\mu/\lambda)^{x_0}$. In this latter case, the probability of population persistence is $1-(\mu/\lambda)^{x_0}$. This identity can be used to approximate the probability of an outbreak in the DTMC and CTMC SIS and SIR epidemic models, where population persistence can be interpreted as an outbreak. The approximation improves the larger the population size N and the smaller the initial number of infected individuals.

Suppose the initial number of infected individuals i_0 is small and the population size N is large. Then the 'birth' and 'death' functions in an SIS epidemic model are given by

$$\text{Birth}=b(i)=\frac{\beta}{N}i(N-i)\approx\beta i$$

and

$$\text{Death}=d(i)=(b+\gamma)i.$$

Applying the identity (15) and the preceding approximations for the birth and death functions leads to the approximation $\mu/\lambda=(b+\gamma)/\beta=1/\mathcal{R}_0$, that is,

$$\text{Prob}\{\mathcal{I}(t)=0\}\approx\begin{cases}1, & \text{if } \mathcal{R}_0\le 1\\ \left(\frac{1}{\mathcal{R}_0}\right)^{i_0}, & \text{if } \mathcal{R}_0>1\end{cases}.$$

Therefore, the probability of an outbreak is

$$\text{Probability of an Outbreak}\approx\begin{cases}0, & \text{if } \mathcal{R}_0\le 1\\ 1-\left(\frac{1}{\mathcal{R}_0}\right)^{i_0}, & \text{if } \mathcal{R}_0>1\end{cases}. \tag{16}$$

The estimates in (16) apply to the stochastic SIS and SIR epidemic models only for a range of times, $t \in [T_1, T_2]$. In the stochastic epidemic models, eventually $\lim_{t\to\infty} \text{Prob}\{\mathcal{I}(t) = 0\} = 1$ because zero is an absorbing state. The range of times for which the estimate (16) holds can be quite long when N is large and i_0 is small (see Fig. 5). In Fig. 5, $N = 100$, $\mathcal{R}_0 = 2$, and $i_0 = 2$, so that applying (16) leads to the estimate for the probability of no outbreak as $(1/2)^2 = 1/4$. The value $1/4$ is very close to the mass of the distribution concentrated at zero, $\text{Prob}\{\mathcal{I}(t) = 0\}$. In Fig. 11, $\text{Prob}\{\mathcal{I}(t) = 0\}$ for the DTMC SIS epidemic model is graphed for different values of $\mathcal{R}_0$. There is close agreement between the numerical values and the estimate $(1/\mathcal{R}_0)^{i_0}$ when $i_0 = 1, 2, 3$ $[(1/\mathcal{R}_0)^{i_0} = 0.5, 0.25, 0.125]$.

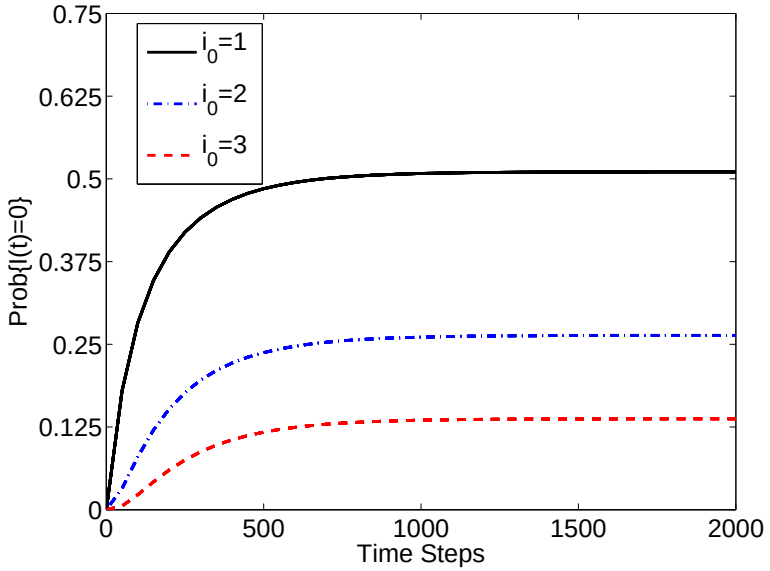

Fig. 11. Graphs of $\text{Prob}\{\mathcal{I}(t) = 0\}$ for $\mathcal{R}_0 = 2$, $N = 100$, and $\text{Prob}\{\mathcal{I}(0) = i_0\} = 1$, $i_0 = 1, 2, 3$.

6.2 Quasistationary Probability Distribution

Because the zero state in the stochastic SIS epidemic models is absorbing, the unique stationary distribution approached asymptotically by the stochastic process is the disease-free equilibrium. However, as seen in the previous section and in Fig. 5, prior to absorption, the process approaches what appears to be a stationary distribution that is different from the disease-free equilibrium. This distribution is known as the quasistationary probability distribution (first investigated in the 1960s [18]). The quasistationary probability distribution

can be obtained from the distribution conditioned on nonextinction (i.e., conditional on the disease-free equilibrium not being reached).

Let the distribution conditioned on nonextinction for the CTMC SIS epidemic model be denoted as $q(t) = (q_1(t), \ldots, q_N(t))^T$. Then $q_i(t)$ is the probability $\mathcal{I}(t) = i$ given that $\mathcal{I}(s) > 0$ for $t > s$ (the disease-free equilibrium has not been reached by time t), i.e.,

$$q_i(t) = \text{Prob}\{\mathcal{I}(t) = i | \mathcal{I}(s) > 0,\ t > s\},$$

$i = 1, 2, \ldots, N$. Because the zero state is absorbing, the probability $\text{Prob}\{\mathcal{I}(s) > 0,\ t > s\} = 1 - p_0(t)$. Therefore,

$$q_i(t) = \frac{p_i(t)}{1 - p_0(t)}, \quad i = 1, 2, \ldots, N. \tag{17}$$

The forward Kolmogorov differential equations for p_i given in (10) can be used to derive a system of differential equations for the q_i.

Differentiating the expression for q_i in (17) with respect to t and applying the identity for dp_i/dt in (10) leads to

$$\frac{dq_i}{dt} = \frac{dp_i/dt}{1 - p_0} + (b + \gamma) q_1 \frac{p_i}{1 - p_0}$$

for $i = 1, 2, \ldots, N$. In matrix notation, the system of differential equations for $q = (q_1, \ldots, q_N)^T$ are similar to the forward Kolmogorov differential equations,

$$\frac{dq}{dt} = \tilde{Q} q + (b + \gamma) q_1 q,$$

where matrix $\tilde{Q}$ is the same as matrix Q in (11) with the exception that the first row and column deleted. Matrix $\tilde{Q}$ is

$$\begin{pmatrix} -[b(1) + d(1)] & d(2) & \cdots & 0 \\ b(1) & -[b(2) + d(2)] & \cdots & 0 \\ 0 & b(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d(N) \\ 0 & 0 & \cdots & -d(N) \end{pmatrix},$$

where $b(i) = \beta i(N - i)/N$ and $d(i) = (b + \gamma)i$.

Now, the quasistationary probability distribution can be defined. The *quasistationary probability distribution* is the stationary distribution (time-independent solution) $q^* = (q_1^*, \ldots, q_N^*)^T$ satisfying

$$\tilde{Q} q^* = -(b + \gamma) q_1^* q^*. \tag{18}$$

Although q^* cannot be solved directly from the system of equations (18), it can be solved indirectly via an iterative scheme (see e.g., [38, 39]).

The quasistationary distribution is related to the eigenvalues of the original matrix Q, where $dp/dt = Qp$. The solution to the forward Kolmogorov differential equations (11) satisfy

$$p(t) = v_0 + v_1 e^{r_1 t} + \cdots + v_N e^{r_N t},$$

where $v_0 = (1, 0, 0, \ldots, 0)^T$ [28, 38, 39]. Since matrix Q is the same as $\tilde{Q}$, with the first row and column deleted, the vector $v_1 = (-1, q_1^*, q_2^*, \ldots, q_N^*)^T$ is an eigenvector of Q corresponding to the eigenvalue $r_1 = -(b+\gamma)q_1^*$, that is,

$$Qv_1 = r_1 v_1$$

so that

$$p(t) = (1, 0, 0, \ldots, 0)^T + (-1, q_1^*, q_2^*, \ldots, q_N^*)^T e^{r_1 t} + \cdots + v_N e^{r_N t}.$$

Nåssell discusses two approximations to the quasistationary probability distribution [38, 39, 40]. One approximation assumes $d(1) = 0$. For this approximation, the system of differential equations for q simplify to

$$\frac{dq}{dt} = Q_1 q, \tag{19}$$

where

$$Q_1 = \begin{pmatrix} -b(1) & d(2) & \cdots & 0 \\ b(1) & -[b(2)+d(2)] & \cdots & 0 \\ 0 & b(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d(N) \\ 0 & 0 & \cdots & -d(N) \end{pmatrix}.$$

System (19) has a unique stable stationary distribution, $p^1 = (p_1^1, \ldots, p_N^1)^T$, where $Q_1 p^1 = \mathbf{0}$. Because matrix Q_1 is tridiagonal, p^1 has an explicit solution given by

$$p_i^1 = p_1^1 \frac{(N-1)!}{i(N-i)!} \left(\frac{\mathcal{R}_0}{N}\right)^{i-1}, \; i = 2, \ldots, N,$$
$$p_1^1 = \left[\sum_{k=1}^{N} \frac{(N-1)!}{k(N-k)!} \left(\frac{\mathcal{R}_0}{N}\right)^{k-1}\right]^{-1}.$$

[8, 38, 39, 40] A simple recursion formula can be easily applied to find this approximation:

$$p_{i+1}^1 = \frac{b(i)}{d(i+1)} p_i^1$$

with the property that $\sum_{i=1}^{N} p_i^1 = 1$. The exact quasistationary distribution and the first approximation (for the DTMC and the CTMC epidemic models)

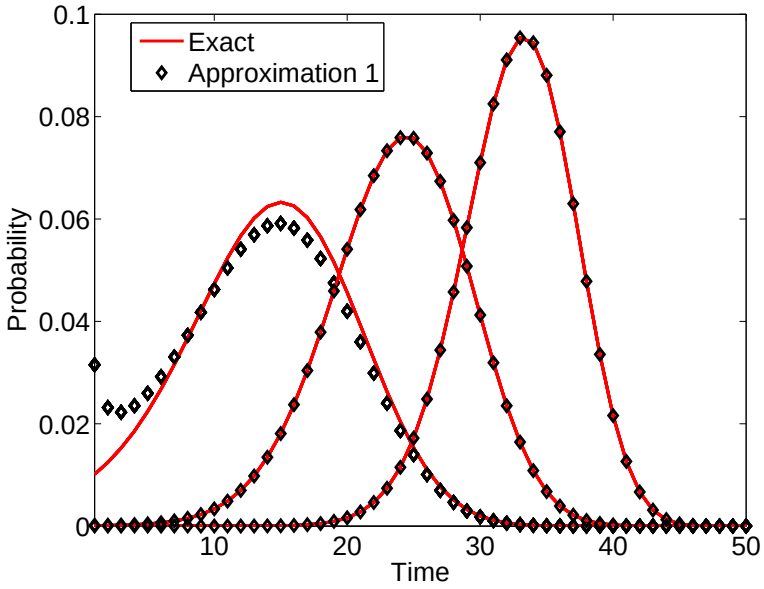

Fig. 12. Exact quasistationary distribution and the first approximation to the quasistationary distribution for $\mathcal{R}_0 = 1.5, 2$, and 3 when $N = 50$.

are graphed for different values of $\mathcal{R}_0$ in Fig. 12. Note that the agreement between the exact quasistationary distribution and the approximation improves as $\mathcal{R}_0$ increases. In addition, note that the mean values are close to the stable endemic equilibrium of the deterministic SIS epidemic model.

The second approximation to the quasistationary probability distribution replaces $d(i)$ by $d(i-1)$. Then the differential equations for q simplify to

$$\frac{dq}{dt} = Q_2 q,$$

where

$$Q_2 = \begin{pmatrix} -b(1) & d(1) & \cdots & 0 \\ b(1) & -[b(2)+d(1)] & \cdots & 0 \\ 0 & b(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d(N-1) \\ 0 & 0 & \cdots & -d(N) \end{pmatrix}.$$

The stable stationary solution is the unique solution p^2 to $Q_1 p^2 = \mathbf{0}$. An explicit solution for p^2 is given by

$$p_i^2 = p_1^2 \frac{(N-1)!}{(N-i)!} \left(\frac{\mathcal{R}_0}{N}\right)^{i-1}, \ i = 2, \ldots, N,$$
$$p_1^2 = \left[\sum_{k=1}^{N} \frac{(N-1)!}{(N-k)!} \left(\frac{\mathcal{R}_0}{N}\right)^{k-1}\right]^{-1}$$

(see [8, 38, 39, 40]).

6.3 Final Size of an Epidemic

In the SIR epidemic model, eventually the epidemic ends. Of interest is the total number of cases during the course of the epidemic, i.e., the final size of the epidemic. If the epidemic is short term and involves a relatively small population, it is reasonable to assume that there are no births nor deaths. In addition, at the beginning of the epidemic, we assume all individuals are either susceptible or infected, $R(0) = 0$. The initial population size is $N = S(0) + I(0)$. Then the final size of the epidemic is the number of susceptible individuals that become infected during the epidemic plus the initial number infected.

In the deterministic model, the final size of the epidemic can be computed directly from the differential equations (3) (see introductory chapter by F. Brauer). Integrating the differential equation $dI/dS = -1 + N\gamma/\beta S$, leads to

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

Letting $t \to \infty$,

$$S(\infty) = I(0) + S(0) + \frac{N\gamma}{\beta} \ln \frac{S(\infty)}{S(0)}.$$

The final size of the epidemic is

$$R(\infty) = N - S(\infty).$$

The final sizes in the deterministic SIR epidemic model are summarized in Table 1 when $I(0) = 1$ and $\gamma = 1$ for various values of $\mathcal{R}_0$ and N.

In the stochastic SIR epidemic model there is a distribution associated with the final size of the epidemic. Let (s, i) denote the ordered pairs of values for the susceptible and infected individuals in the CTMC model. The epidemic ends when $\mathcal{I}(t)$ reaches zero. When the epidemic ends, the random variable for the number of susceptible individuals ranges from 0 to $N - \mathcal{I}(0) = N - i_0$. In particular, the set $\{(s, 0)\}_{s=0}^{N-i_0}$ is absorbing,

$$\lim_{t\to\infty} \sum_{s=0}^{N-i_0} p_{(s,0)}(t) = 1.$$

Table 1. Final size of an epidemic when $\gamma = 1$ and $I(0) = 1$ for the deterministic SIR epidemic model.

$\mathcal{R}_0$	N		
	20	100	1000
0.5	1.87	1.97	2.00
1	5.74	13.52	44.07
2	16.26	80.02	797.15
5	19.87	99.31	993.03
10	20.00	100.00	999.95

Daley and Gani [17] discuss two different methods to compute the probability distribution associated with the final size. The simpler method, originally developed by Foster [20], depends on the embedded Markov chain, that is, the DTMC model associated with the CTMC model. To apply this method, the transition matrix for the embedded Markov chain needs to be computed. This requires computing the probability of a transition between the states (s, i), where the states lie in the set $\{(s, i) : s = 0, 1, \ldots, N; i = 0, 1, \ldots, N - s\}$. In the embedded Markov chain for the final size, the times between transitions are not important, only the probabilities.

For example, suppose $N = 3$, then the states in the transition matrix are

$$(s, i) \in \{(3, 0), (2, 0), (1, 0), (0, 0), (2, 1), (1, 1), (0, 1), (1, 2), (0, 2), (0, 3)\}, \quad (20)$$

i.e., there are 10 ordered pairs of states. There are only two types of transitions, either an infected individual recovers, $(s, i) \to (s, i - 1)$ or a susceptible individual becomes infected, $(s, i) \to (s-1, i+1)$. In the first type of transition, an infected individual recovers with probability

$$p_s = \frac{\gamma i}{\gamma i + (\beta/N) is} = \frac{\gamma}{\gamma + (\beta/N)s}, \quad s = 0, 1, 2.$$

In the second type of transition, a susceptible individual becomes infected with probability $1 - p_s$. If the 10 states are ordered as in (20), then the transition matrix for the embedded Markov chain is a 10×10 matrix with the following form:

$$T = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & p_2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & p_1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & p_0 & 0 & 0 & 0 \\
- & - & - & - & - & - & - & - & - & - \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & p_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_0 & 0 \\
- & - & - & - & - & - & - & - & - & - \\
0 & 0 & 0 & 0 & 1-p_2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1-p_1 & 0 & 0 & 0 & p_0 \\
- & - & - & - & - & - & - & - & - & - \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p_1 & 0 & 0
\end{pmatrix}$$

The upper left 4×4 corner of matrix T is the identity matrix because these are the four absorbing states. The first four rows are the transitions into these four absorbing states. Matrix T is a stochastic matrix, whose column sums equal one (note that $p_0 = 1$). Given the initial distribution for the states $p(0)$, then the distribution for the final size can be found from the first four entries of $\lim_{t\to\infty} T^t p(0)$ (the remaining entries are zero). However, it is not necessary to compute the limit as $t \to \infty$, since the limit converges by time $t = 2N - 1$. For this example, it is straightforward to compute the final size distribution. The final size is either 1,2, or 3 with corresponding probabilities p_2, $p_1^2(1-p_2)$ and $(1-p_1^2)(1-p_2)$, respectively. In Fig. 13, there are graphs of three final size distributions for different values of $\mathcal{R}_0$ when $\gamma = 1$, $\text{Prob}\{\mathcal{I}(0) = 1\} = 1$, and $N = 20$.

When $\mathcal{R}_0$ is less than one or very close to one, then the final size distribution is skewed to the right, but if $\mathcal{R}_0$ is much greater than one, then the distribution is skewed to the left. The average final sizes for the stochastic SIR when $N = 20$ and $N = 100$ are listed in Table 2. Compare the values in Table 2 to those in Table 1. For values of $\mathcal{R}_0$ less than one or much greater than one, the average final sizes for the stochastic SIR epidemic model are closer to the values of the final sizes for the deterministic model.

Table 2. Average final size of an epidemic when $\gamma = 1$, $b = 0$, and $\text{Prob}\{\mathcal{I}(0) = 1\} = 1$ for the stochastic SIR epidemic model.

$\mathcal{R}_0$	N	
	20	100
0.5	1.76	1.93
1	3.34	6.10
2	8.12	38.34
5	15.66	79.28
10	17.98	89.98

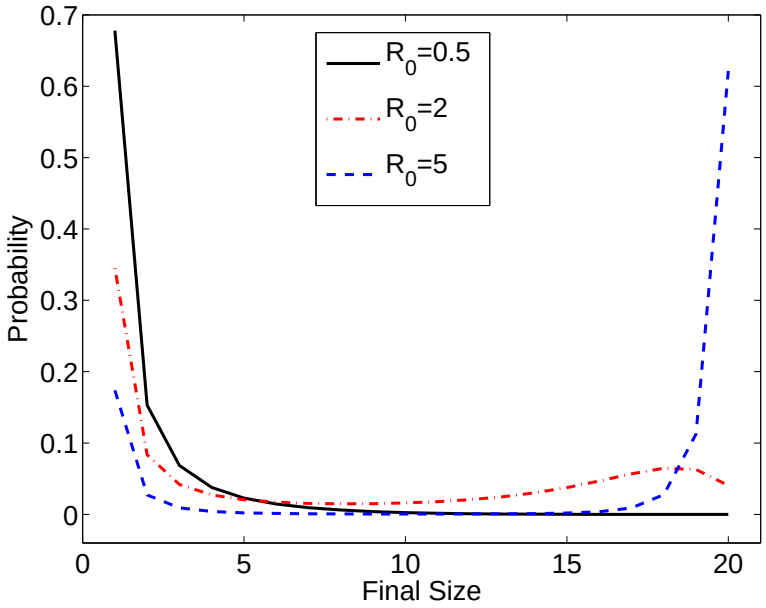

Fig. 13. Distribution for the final size of an epidemic for three different values of $\mathcal{R}_0$ when $\gamma = 1$, $N = 20$, and Prob$\{\mathcal{I}(0) = 1\} = 1$.

6.4 Expected Duration of an Epidemic

The duration of an epidemic corresponds to the time until absorption, i.e., the time T until $\mathcal{I}(T) = 0$. For the stochastic SIS epidemic model, the probability of absorption is one, regardless of the value of $\mathcal{R}_0$. However, depending on the initial number infected, i, the population size N, and the value of $\mathcal{R}_0$, the time until absorption can be very short or very long. Here, we derive a system of equations that can be solved to find the expected time until absorption for a stochastic SIS epidemic model.

Let T_i denote the random variable for the time until absorption and let

$$\tau_i = E(T_i)$$

denote the expected time until absorption beginning from an initial infected population size of i, $i = 0, 1, \ldots, N$. Let the higher order moments for the time until absorption be denoted as

$$\tau_i^r = E(T_i^r),$$

$i = 0, 1, \ldots, N$. Note that $\tau_0 = 0 = \tau_0^r$. Then, considered as a birth and death process, the mean time until absorption in the DTMC SIS epidemic model satisfies the following difference equation:

$$\tau_i = b(i)\Delta t(\tau_{i+1} + \Delta t) + d(i)\Delta t(\tau_{i-1} + \Delta t) + (1 - [b(i) + d(i)]\Delta t)(\tau_i + \Delta t), \quad i = 1, \ldots, N \tag{21}$$

The CTMC SIS epidemic model satisfies the same relationship as equations (21), except that a term $o(\Delta t)$ is added to the right side of each equation. Simplifying the equations in (21) leads to a system of difference equations for the expected duration of an epidemic (for both the CTMC and the DTMC models)

$$d(i)\tau_{i-1} - [b(i) + d(i)]\tau_i + b(i)\tau_{i+1} = -1$$

where $b(i) = i(N - i)(\beta i/N)$ and $d(i) = (b + \gamma)i$ [7, 33]. Similar difference equations apply to the higher order moment τ_i^r in the CTMC SIS epidemic model:

$$d(i)\tau_{i-1}^r - [b(i) + d(i)]\tau_i^r + b(i)\tau_{i+1}^r = -r\tau_i^{r-1}$$

[7, 22, 41, 42].

The mean and higher order moments can be expressed in matrix form. Let $\tau = (\tau_1, \tau_2, \ldots, \tau_N)^T$, $\tau^r = (\tau_1^r, \tau_2^r, \ldots, \tau_N^r)^T$ and $\tau^1 = \tau$. Then

$$D\tau = -\mathbf{1} \quad \text{and} \quad D\tau^r = -r\tau^{r-1}.$$

where $\mathbf{1} = (1, \ldots, 1)^T$ and

$$D = \begin{pmatrix} -[b(1) + d(1)] & b(1) & 0 & \cdots & 0 & 0 \\ d(2) & -[b(2) + d(2)] & b(2) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & d(N) & -d(N) \end{pmatrix}.$$

Matrix D is nonsingular because it is irreducibly diagonally dominant [43]. Hence, the solutions τ and τ^r are unique.

A solution for the expected time until absorption, based on a system of SDEs, can be derived also [7]. The relationship satisfied by τ follows from the backward Kolmogorov differential equations. Let $\tau(y)$ denote the expected time until absorption beginning from an infected population size of $y \in (0, N)$. Then it can be shown that $\tau(y)$ is the solution to the following boundary value problem:

$$[b(y) - d(y)]\frac{d\tau(y)}{dy} + \frac{[b(y) + d(y)]}{2}\frac{d^2\tau(y)}{dy^2} = -1, \tag{22}$$

where

$$\tau(0) = 0 \quad \text{and} \quad \left.\frac{d\tau(y)}{dy}\right|_{y=N} = 0,$$

$b(y) = (N - y)(\beta y/N)$ and $d(y) = (b + \gamma)y$ in the SDE SIS epidemic model [7].

It is interesting to note that if the derivatives in the boundary value problem for $\tau(y)$ in (22) are approximated by finite difference formulas, then the

difference equations for τ_i, given in (21), for the CTMC and DTMC epidemic models are obtained [7]. For $y \in [i, i+1]$, let

$$\frac{d\tau(y)}{dy} \approx \frac{\tau_{i+1} - \tau_{i-1}}{2},$$

where $\tau_i = \tau(i)$ and $\tau_{i+1} = \tau(i+1)$. In addition, let

$$\frac{d^2\tau(y)}{dy^2} \approx \tau_{i+1} - 2\tau_i + \tau_{i-1}.$$

With these approximations, the boundary value problem for $\tau(y)$ in (22) is approximated by the difference equations for τ_i in (21).

The expected duration of an SIS epidemic can be calculated from the solution to the equations (21) or (22). Allen and Allen [7] compare the mean and the variance for the time until population extinction for the three different types of stochastic formulations considered here. However, their population model is logistic growth (similar to the SIS epidemic model).

As an example, consider the expected duration for an SIS epidemic, based on the DTMC or CTMC model. Because the DTMC and CTMC models satisfy the same set of equations for the expected duration, these results apply to both models. With a population size of $N = 25$ and either $\mathcal{R}_0 = 2$ or $\mathcal{R}_0 = 1.5$. The solution $\tau = -D^{-1}\mathbf{1}$ is graphed in Fig. 14. If the population size is increased to $N = 50$ or $N = 100$ with the same value for $\mathcal{R}_0 = 1.5$, the expected duration for large i increases to $\tau_i \approx 160$ and $\tau_i \approx 3,500$, respectively. At population sizes of $N = 50$ and $N = 100$ but a basic reproduction of $\mathcal{R}_0 = 2$, the expected duration for large i is much larger, $\tau_i \approx 25,000$ and $\tau_i \approx 2.6 \times 10^8$, respectively. Of course, the expected duration depends on the particular time units of the model. For example, if the time units are days, then $\tau_i \approx 160 \approx 5.3$ months and $\tau_i \approx 25,000 \approx 68.5$ years. This latter estimate is much longer than a reasonable epidemiological time frame, implying that the disease does not die out but persists. Hence, for these examples, when $N \geq 100$ and $\mathcal{R}_0 \geq 2$, if the outbreak begins with a sufficient number of infected individuals, then the results for the stochastic SIS epidemic are in close agreement with the predictions of the deterministic SIS epidemic model: the disease becomes endemic.

7 Epidemic Models with Variable Population Size

Suppose the population size N is not constant but varies according to some population growth law. To formulate an epidemic model, an assumption must be made concerning the population birth and death rates which depend on the population size N. Here, we assume, for simplicity, that the birth rate and death rates have a logistic form,

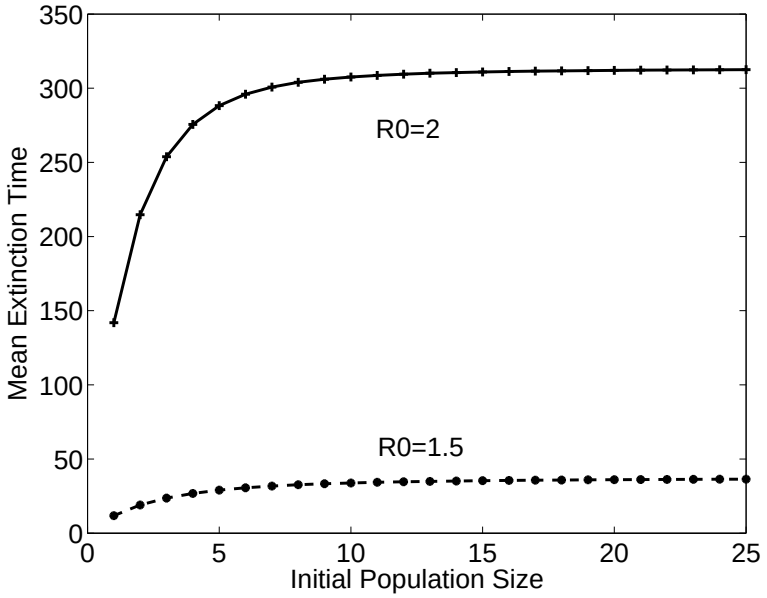

Fig. 14. Expected duration of an SIS epidemic with a population size of $N = 25$; $\mathcal{R}_0 = 1.5$ ($b = 1/3$, $\gamma = 1/3$ and $\beta = 1$) and $\mathcal{R}_0 = 2$ ($b = 1/4$, $\gamma = 1/4$ and $\beta = 1$).

$$\lambda(N) = bN \quad \text{and} \quad \mu(N) = b\frac{N^2}{K},$$

respectively. Then the total population size satisfies the logistic differential equation

$$\frac{dN}{dt} = \lambda(N) - \mu(N) = bN\left(1 - \frac{N}{K}\right),$$

where $K > 0$ is the carrying capacity. There are many functional forms that can be chosen for the birth and death rates [7]. Their choice should depend on the dynamics of the particular population being modeled. For example, in animal diseases (e.g., rabies in canine populations [37, 46] and hantavirus in rodent populations [2, 3, 9, 44]), logistic growth is assumed, then the choice of $\lambda(N)$ and $\mu(N)$ depends on whether the births and deaths are density-dependent. For human diseases, a logistic growth assumption may not be very realistic.

A deterministic SIS epidemic model is formulated for a population satisfying the logistic differential equation. Again, for simplicity, we assume there are no disease-related deaths and no vertical transmission of the disease; all newborns are born susceptible. Then the deterministic SIS epidemic model has the form:

$$\begin{aligned}\frac{dS}{dt} &= \frac{S}{N}(\lambda(N) - \mu(N)) - \frac{\beta}{N}SI + (b+\gamma)I \\ \frac{dI}{dt} &= -\frac{I}{N}\mu(N) + \frac{\beta}{N}SI - \gamma I,\end{aligned} \tag{23}$$

where $S(0) > 0$ and $I(0) > 0$. It is straightforward to show that the solution to this system of differential equations depends on the basic reproduction number $\mathcal{R}_0 = \beta/(b+\gamma)$.

Theorem 3. *Let $S(t)$ and $I(t)$ be a solution to model (23).*

i) If $\mathcal{R}_0 \leq 1$, then $\lim_{t\to\infty}(S(t), I(t)) = (K, 0)$.
ii) If $\mathcal{R}_0 > 1$, then $\lim_{t\to\infty}(S(t), I(t)) = (K/\mathcal{R}_0, K(1 - 1/\mathcal{R}_0))$.

Stochastic epidemic models for each of the three types (CTMC, DTMC, and SDE models) can be formulated. Because $S(t) + I(t) = N(t)$, the process is bivariate. We derive a SDE model and compare the graph of a sample path for the stochastic model to the solution of the deterministic model.

Let $\mathcal{S}(t)$ and $\mathcal{I}(t)$ be continuous random variables for the number of susceptible and infected individuals at time t,

$$S(t),\ \mathcal{I}(t) \in [0, \infty).$$

Then, applying the same methods as for the SDE SIS and SIR epidemic models [5, 6],

$$\begin{aligned}\frac{d\mathcal{S}}{dt} &= \frac{\mathcal{S}}{\mathcal{N}}(\lambda(\mathcal{N}) - \mu(\mathcal{N})) - \frac{\beta}{\mathcal{N}}\mathcal{S}\mathcal{I} + (b+\gamma)\mathcal{I} + B_{11}\frac{dW_1}{dt} + B_{12}\frac{dW_2}{dt} \\ \frac{d\mathcal{I}}{dt} &= -\frac{\mathcal{I}}{\mathcal{N}}\mu(\mathcal{N}) + \frac{\beta}{\mathcal{N}}\mathcal{S}\mathcal{I} - \gamma\mathcal{I} + B_{21}\frac{dW_1}{dt} + B_{22}\frac{dW_2}{dt},\end{aligned}$$

where $B = (B_{ij})$ is the square root of the following covariance matrix:

$$\begin{pmatrix} \frac{\mathcal{S}}{\mathcal{N}}(\lambda(\mathcal{N}) + \mu(\mathcal{N})) + \frac{\beta}{\mathcal{N}}\mathcal{S}\mathcal{I} + (b+\gamma)\mathcal{I} & -\frac{\beta}{\mathcal{N}}\mathcal{S}\mathcal{I} - \gamma\mathcal{I} \\ -\frac{\beta}{\mathcal{N}}\mathcal{S}\mathcal{I} - \gamma\mathcal{I} & \frac{\mathcal{I}}{\mathcal{N}}\mu(\mathcal{N}) + \frac{\beta}{\mathcal{N}}\mathcal{S}\mathcal{I} + \gamma\mathcal{I} \end{pmatrix}.$$

The variables W_1 and W_2 are two independent Wiener processes. The absorbing state for the bivariate process is total population extinction, $\mathcal{N} = 0$.

7.1 Numerical Example

As might be anticipated, the variability in the population size results in an increase in the variability in the number of infected individuals. As an example, let $\beta = 1$, $\gamma = 0.25 = b$, and $K = 100$. Then the basic reproduction number is $\mathcal{R}_0 = 2$. The SDE SIS epidemic model with constant population

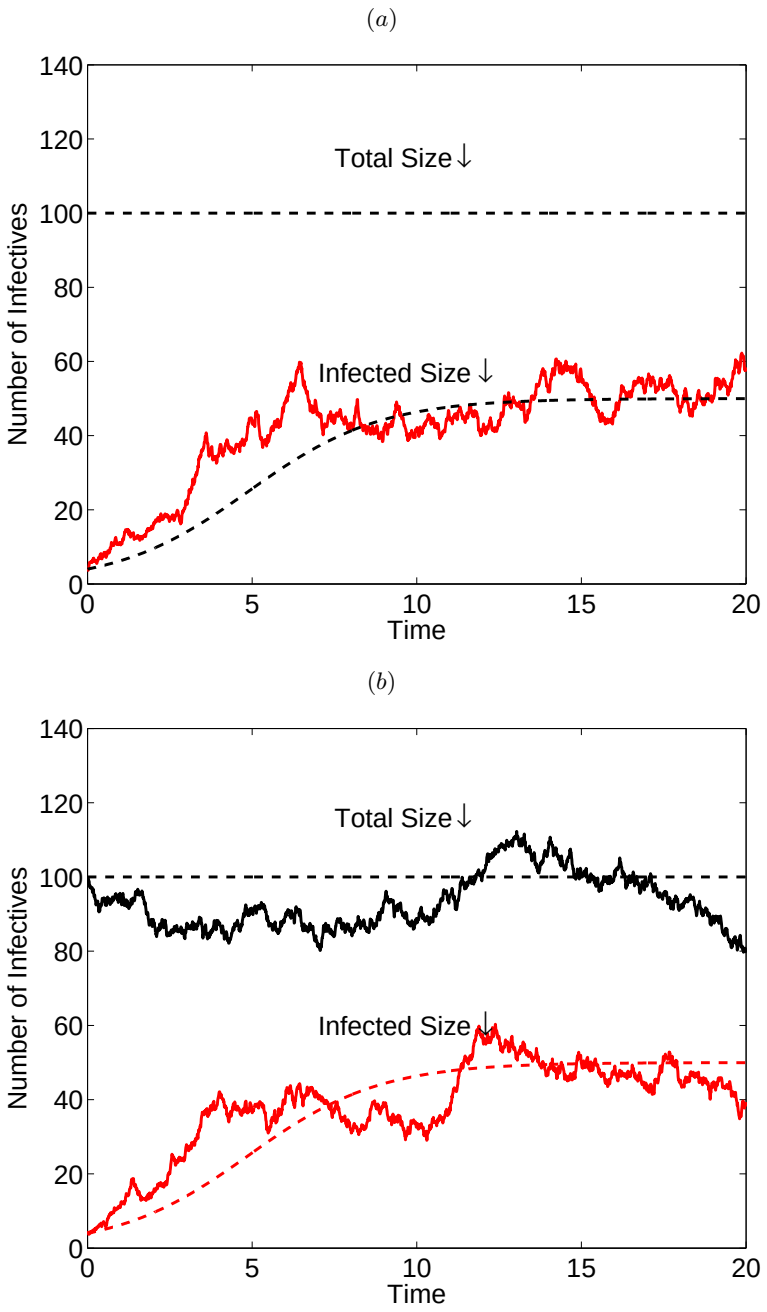

Fig. 15. The SDE SIS epidemic model (**a**) with constant population size, $N = 100$ and (**b**) with variable population size, $\mathcal{N}(t)$. The parameter values are $\beta = 1$, $\gamma = 0.25 = b$, $K = 100$, and $\mathcal{R}_0 = 2$.

size, $N = 100$, is compared to the SDE SIS epidemic model with variable population size, $\mathcal{N}(t)$, in Fig. 15. One sample path of the SDE epidemic model is graphed against the deterministic solution.

More realistic stochastic epidemic models can be derived based on their deterministic formulations. Excellent references for a variety of recent deterministic epidemic models include the books by Anderson and May [10], Brauer and Castillo-Chavez [15], Diekmann and Heesterbeek [19], and Thieme [48] and the review articles by Hethcote [26] and Brauer and van den Driessche [16].

In this chapter, the simplest types of epidemic models were chosen as an introduction to the methods of derivation for various types of stochastic models (DTMC, CTMC, and SDE models). In many cases these three stochastic formulations produce similar results, if the time step Δt is small [7]. There are advantages numerically in applying the discrete time approximations (DTMC model and the Euler approximation to the SDE model) in that the discrete simulations generally have a shorter computational time than the CTMC model. Mode and Sleeman [36] discuss some computational methods in stochastic processes in epidemiology. The most important consideration in modeling, however, is to choose a model that best represents the demographics and epidemiology of the population being modeled.

We conclude this chapter with a discussion of some well-known stochastic epidemic models that are not based on any deterministic epidemic model.

8 Other Types of DTMC Epidemic Models

Two other types of DTMC epidemic models are discussed briefly that are not directly related to any deterministic epidemic model. These models are chain binomial epidemic models and epidemic branching processes.

8.1 Chain Binomial Epidemic Models

Two well-known DTMC models are the Greenwood and the Reed-Frost models. These models were developed to help understand the spread of disease within a small population such as a household. They are referred to as chain binomial epidemic models because a binomial distribution is used to determine the number of new infectious individuals. The Greenwood model developed in 1931, was named after its developer [23]. The Reed-Frost model, developed in 1928, was named for two medical researchers, who developed the model for teaching purposes at John's Hopkins University. It wasn't until 1952 that the Reed-Frost model was published [1, 17].

Let $\mathcal{S}_t$ and $\mathcal{I}_t$ be discrete random variables for the number of susceptible and infected individuals in the household at time t. Initially, the models assume that there are $\mathcal{I}_0 = i_0 \geq 1$ infected individuals and $\mathcal{S}_0 = s_0$ susceptible individuals. The progression of the disease is followed by keeping track of the

number of susceptible individuals over time. At time t, infected individuals are in contact with all the susceptible members of the household to whom they may spread the disease. However, it is not until time $t+1$ that susceptible individuals who have contracted the disease are infectious. The period of time from t to $t+1$ is the latent period and the infectious period is contracted to a point. Only at time t can the infectious individuals $\mathcal{I}_t$ infect susceptible members $\mathcal{S}_t$. After that time, they are no longer infectious. It follows that the newly infectious individuals at time $t+1$ satisfy

$$\mathcal{S}_{t+1} + \mathcal{I}_{t+1} = \mathcal{S}_t.$$

These models are bivariate Markov chain models that depend on the two random variables, $\mathcal{S}_t$ and $\mathcal{I}_t$, $\{(\mathcal{S}_t, \mathcal{I}_t)\}$.

The models of Greenwood and Reed-Frost differ in the assumption regarding the probability of infection. Suppose there are a total of $\mathcal{I}_t = i$ infected individuals at time t. Let p_i be the probability that a susceptible individual does not become infected at time t. The Greenwood model assumes that $p_i = p$ is a constant and the Reed-Frost model assumes that $p_i = p^i$. For each model, the transition probability from state (s_t, i_t) to (s_{t+1}, i_{t+1}) is assumed to have a binomial distribution. Sample paths are denoted as $\{s_0, s_1, \ldots, s_{t-1}, s_t\}$. The epidemic stops at time t when $s_{t-1} = s_t$ because there are no more infectious individuals to spread the disease, $i_t = s_{t-1} - s_t = 0$.

Greenwood Model

In the Greenwood model, the random variable $\mathcal{S}_{t+1}$ is a binomial random variable that depends on $\mathcal{S}_t$ and p, $\mathcal{S}_{t+1} \sim b(\mathcal{S}_t, p)$. The probability of a transition from (s_t, i_t) to (s_{t+1}, i_{t+1}) depends only on s_t, s_{t+1}, and p. It is defined as follows:

$$p_{s_{t+1}, s_t} = \binom{s_t}{s_{t+1}} p^{s_{t+1}} (1-p)^{s_t - s_{t+1}}.$$

The conditional mean and variance of $\mathcal{S}_{t+1}$ and $\mathcal{I}_{t+1}$ are given by

$$E(\mathcal{S}_{t+1}|\mathcal{S}_t) = p\mathcal{S}_t, \quad E(\mathcal{I}_{t+1}|\mathcal{S}_t) = (1-p)\mathcal{S}_t$$

and

$$\mathrm{Var}(\mathcal{S}_{t+1}|\mathcal{S}_t) = p(1-p)\mathcal{S}_t = \mathrm{Var}(\mathcal{I}_{t+1}|\mathcal{S}_t).$$

Four sample paths of the Greenwood model when $s_0 = 6$ and $i_0 = 1$ are illustrated in Fig. 16. Applying the preceding transition probabilities, it is clear that the sample path $\{6, 6\}$ occurs with probability $p_{6,6} = p^6$ and the sample path $\{6, 5, 5\}$ occurs with probability $p_{6,5}p_{5,5} = 6p^{10}(1-p)$. The probability distributions associated with the size and the duration of epidemics in the chain binomial models can be easily defined, once the probability distribution

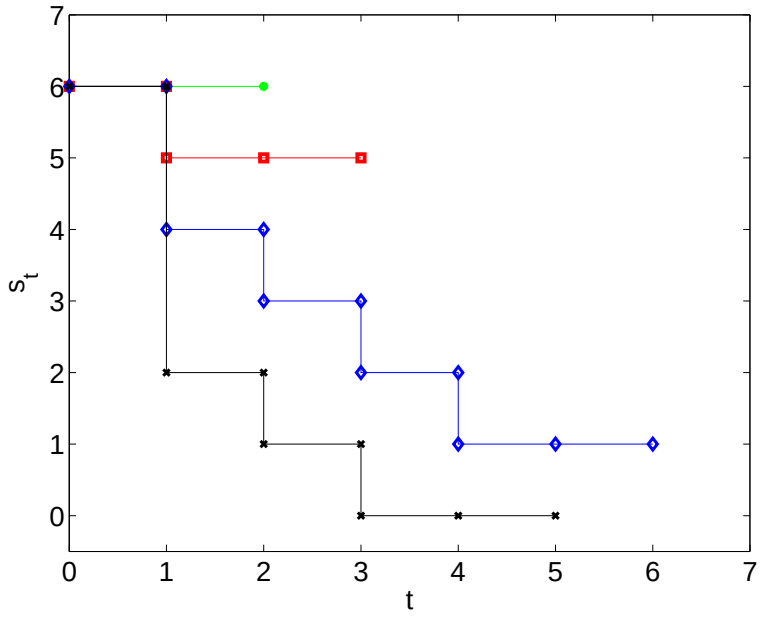

Fig. 16. Four sample paths for the Greenwood chain binomial model when $s_0 = 6$ and $i_0 = 1 : \{6, 6\}$, $\{6, 5, 5\}$, $\{6, 4, 3, 2, 1, 1\}$, and $\{6, 2, 1, 0, 0\}$.

associated with each sample path are determined. The discrete random variable $W = \mathcal{S}_0 - \mathcal{S}_t$ is the size of the epidemic and the discrete random variable T is the length of the path, e.g., if $\{s_0, s_1, \ldots, s_{t-1}, s_t\}$, then $T = t$.

Table 3 summarizes the probabilities associated with the Greenwood and Reed-Frost epidemic models when $s_0 = 3$ and $i_0 = 1$ (see [17]).

Table 3. Sample paths, size T, and duration W for the Greenwood and Reed-Frost models when $s_0 = 3$ and $i_0 = 1$.

Sample Paths $\{s_0, \ldots, s_{t-1}, s_t\}$	Duration T	Size W	Greenwood Model	Reed-Frost Model
3 3	1	0	p^3	p^3
3 2 2	2	1	$3(1-p)p^4$	$3(1-p)p^4$
3 2 1 1	3	2	$6(1-p)^2p^4$	$6(1-p)^2p^4$
3 1 1	2	2	$3(1-p)^2p^2$	$3(1-p)^2p^3$
3 2 1 0 0	4	3	$6(1-p)^3p^3$	$6(1-p)^3p^3$
3 2 0 0	3	3	$3(1-p)^3p^2$	$3(1-p)^3p^2$
3 1 0 0	3	3	$3(1-p)^3p$	$3(1-p)^3p(1+p)$
3 0 0	2	3	$(1-p)^3$	$(1-p)^3$
Total			1	1

Reed-Frost Model

In the Reed-Frost model, the random variable $\mathcal{S}_{t+1}$ is binomially distributed and satisfies $\mathcal{S}_{t+1} \sim b(\mathcal{S}_t, p^{\mathcal{I}_t})$. The probability of a transition from (s_t, i_t) to (s_{t+1}, i_{t+1}) is defined as follows:

$$p_{(s,i)_{t+1},(s,i)_t} = \binom{s_t}{s_{t+1}} (p^{i_t})^{s_{t+1}} (1-p^{i_t})^{s_t - s_{t+1}}.$$

The conditional mean and and variance associated with $\mathcal{S}_{t+1}$ are

$$E(\mathcal{S}_{t+1}|(\mathcal{S}_t, \mathcal{I}_t)) = \mathcal{S}_t p^{\mathcal{I}_t}, \quad E(\mathcal{I}_{t+1}|(\mathcal{S}_t, \mathcal{I}_t)) = \mathcal{S}_t(1-p^{I_t})$$

and

$$\mathrm{Var}(\mathcal{S}_{t+1}|(\mathcal{S}_t, \mathcal{I}_t)) = \mathcal{S}_t(1-p^{\mathcal{I}_t})p^{\mathcal{I}_t} = \mathrm{Var}(\mathcal{I}_{t+1}|(\mathcal{S}_t, \mathcal{I}_t)).$$

The Greenwood and Reed-Frost models differ when $\mathcal{I}_t > 1$ for $t > 0$ (see Table 3). For additional information on the Greenwood and Reed-Frost models, and epidemics among households consult Ackerman et al. [4], Ball and Lyne [14], and Daley and Gani [17].

8.2 Epidemic Branching Processes

Branching processes can be applied to epidemics. We illustrate with a simple example of a Galton-Watson branching processes. Let $\mathcal{I}_t$ be the number of new cases at time t. We assume during the time interval t to $t+1$ that new infectious individuals are generated by contacts between the new cases at time t and the susceptible population. Suppose each infected individual infects on the average $\mathcal{R}_0$ susceptible individuals. In a Galton-Watson process, the simplifying assumption is that each infected individual is independent from all other infected individuals.

Let $\{p_k\}_{k=0}^{\infty}$ be the probabilities associated with the number of new infections per infected individual. Then the probability generating function (pgf) for the the number of new infections is

$$f(t) = \sum_{k=0}^{\infty} p_k t^k$$

with mean $f'(1) = \mathcal{R}_0$.

An important result from the theory of branching processes states that the probability of extinction (probability the epidemic eventually ends), $\lim_{t\to\infty} \mathrm{Prob}\{\mathcal{I}_t = 0\}$, depends on the pgf $f(t)$. If $0 \le p_0 + p_1 < 1$ and $\mathcal{R}_0 > 1$, then there exists a unique fixed point $q \in [0,1)$ such that $f(q) = q$. The assumption $0 \le p_0 + p_1 < 1$ guarantees that there is a positive probability of infecting more than one individual. It is the value of q and the initial number of infected individuals in the population that determine the probability of extinction. The next theorem summarizes the main result concerning the probability of extinction. For a proof of this result and extensions, please consult the references [6, 24, 29, 30, 35, 45].

Theorem 4. *Suppose the pgf* $f(t)$ *satisfies* $0 \leq f(0) + f'(0) < 1$ *and* $\text{Prob}\{\mathcal{I}_0 = i_0\} = 1$, *where* $i_0 > 0$.

i) If $\mathcal{R}_0 \leq 1$, *then* $\lim_{t\to\infty} \text{Prob}\{\mathcal{I}_t = 0\} = 1$.

ii) If $\mathcal{R}_0 > 1$, *then* $\lim_{t\to\infty} \text{Prob}\{\mathcal{I}_t = 0\} = q^{i_0}$, *where* q *is the unique fixed point in* $[0, 1)$ *such that* $f(q) = q$.

As a consequence of this theorem, the probability the epidemic persists in the population (the disease becomes endemic) is $1 - q^{i_0}$, provided $\mathcal{R}_0 > 1$.

Antia et al. [11] assume that the number of cases $\mathcal{I}_t$ follows a Poisson distribution with mean $\mathcal{R}_0$. The pgf of a Poisson probability distribution satisfies

$$f(t) = \sum_{k=0}^{\infty} \exp(-\mathcal{R}_0)\frac{\mathcal{R}_0^k}{k!}t^k = \exp(-\mathcal{R}_0(1-t)).$$

Applying Theorem 4, we can estimate the probability the disease becomes endemic. If $\mathcal{R}_0 > 1$, the fixed point of f satisfies

$$q = \exp(-\mathcal{R}_0(1-q)).$$

For example, if $\mathcal{R}_0 = 1.5$ and $\text{Prob}\{\mathcal{I}_0 = 1\} = 1$, then $1 - q = 0.583$, but if $\text{Prob}\{\mathcal{I}_0 = 2\} = 1$, then $1 - q^2 = 0.826$. If $\mathcal{R}_0 = 2$ and $\text{Prob}\{\mathcal{I}_0 = 2\} = 1$, then $1 - q^2 = 0.959$.

9 MatLab Programs

The following three MatLab programs were used to generate sample paths and the probability distribution associated with the stochastic SIS epidemic model. MatLab Program # 1 computes the probability distribution for the DTMC SIS epidemic model. MatLab Programs # 2 and # 3 compute sample paths associated with CTMC and SDE SIS epidemic models, respectively.

```
% MatLab Program # 1
% Discrete Time Markov Chain
% SIS Epidemic Model
% Transition Matrix and Graph of Probability Distribution
clear all
set(gca,'FontSize',18);
set(0,'DefaultAxesFontSize',18);
time=2000;
dtt=0.01; % Time step
beta=1*dtt;
b=0.25*dtt;
gama=0.25*dtt;
```

```
N=100; % Total population size
en=50; % plot every enth time interval
T=zeros(N+1,N+1); % T is the transition matrix, defined below
v=linspace(0,N,N+1);
p=zeros(time+1,N+1);
p(1,3)=1; % Two individuals initially infected.
bt=beta*v.*(N-v)/N;
dt=(b+gama)*v;
for i=2:N % Define the transition matrix
    T(i,i)=1-bt(i)-dt(i); % diagonal entries
    T(i,i+1)=dt(i+1); % superdiagonal entries
    T(i+1,i)=bt(i); % subdiagonal entries
end
T(1,1)=1;
T(1,2)=dt(2);
T(N+1,N+1)=1-dt(N+1);
for t=1:time
    y=T*p(t,:)';
    p(t+1,:)=y';
end
pm(1,:)=p(1,:);
for t=1:time/en;
    pm(t+1,:)=p(en*t,:);
end
ti=linspace(0,time,time/en+1);
st=linspace(0,N,N+1);
mesh(st,ti,pm);
xlabel('Number of Infectives');
ylabel('Time Steps');
zlabel('Probability');
view(140,30);
axis([0,N,0,time,0,1]);
```

% Matlab Program # 2

```
% Continuous Time Markov Chain
% SIS Epidemic Model
% Three Sample Paths and the Deterministic Solution
clear
set(0,'DefaultAxesFontSize', 18);
set(gca,'fontsize',18);
beta=1;
b=0.25;
gam=0.25;
N=100;
```

```
init=2;
time=25;
sim=3;
for k=1:sim
    clear t s i
    t(1)=0;
    i(1)=init;
    s(1)=N-init;
    j=1;
    while i(j)>0 & t(j)<time
       u1=rand; % uniform random number
       u2=rand; % uniform random number
       a=(beta/N)*i(j)*s(j)+(b+gam)*i(j);
       probi=(beta*s(j)/N)/(beta*s(j)/N+b+gam);
       t(j+1)=t(j)-log(u1)/a;
       if u2 <= probi
          i(j+1)=i(j)+1;
          s(j+1)=s(j)-1;
       else
          i(j+1)=i(j)-1;
          s(j+1)=s(j)+1;
       end
       j=j+1;
    end
    plot(t,i,'r-','LineWidth',2)
    hold on
end
```

% Matlab Program # 3

```
% Stochastic Differential Equation
% SIS Epidemic Model
% Three Sample Paths and the Deterministic Solution
clear
beta=1;
b=0.25;
gam=0.25;
N=100;
init=2;
dt=0.01;
time=25;
sim=3;
for k=1:sim
    clear i, t
    j=1;
```

```
    i(j)=init;
    t(j)=dt;
    while i(j)>0 & t(j)<25
       mu=beta*i(j)*(N-i(j))/N-(b+gam)*i(j);
       sigma=sqrt(beta*i(j)*(N-i(j))/N+(b+gam)*i(j));
       rn=randn; % standard normal random number
       i(j+1)=i(j)+mu*dt+sigma*sqrt(dt)*rn;
       t(j+1)=t(j)+dt;
       j=j+1;
    end
    plot(t,i,'r-','Linewidth',2);
    hold on
end
% Euler's method applied to the deterministic SIS model.
y(1)=init;
for k=1:time/dt
    y(k+1)=y(k)+dt*(beta*(N-y(k))*y(k)/N-(b+gam)*y(k));
end
plot([0:dt:time],y,'k--','LineWidth',2);
axis([0,time,0,80]);
xlabel('Time');
ylabel('Number of Infectives');
hold off
```

References

1. Abbey, H.: An examination of the Reed-Frost theory of epidemics. Hum. Biology, **24**, 201–233 (1952).
2. Abramson, G., Kenkre, V. M.: Spatiotemporal patterns in hantavirus infection. Physical Review E, **66**, 011912-1–5 (2002)
3. Abramson, G., Kenkre, V. M., Yates, T. L., Parmenter, R. R.: Traveling waves of infection in the hantavirus epidemics. Bull. Math. Biol., **65**, 519–534 (2003)
4. Ackerman, E., Elveback, L. R., and J. P. Fox, J. P.: Simulation of Infectious Disease Epidemics. Charles C. Thomas Pub., Springfield, Illinois (1984)
5. Allen, E. J.: Stochastic differential equations and persistence time for two interacting populations. Dyn. Cont., Discrete and Impulsive Systems, **5**, 271–281 (1999)
6. Allen, L. J. S.: An Introduction to Stochastic Processes with Applications to Biology. Prentice Hall, Upper Saddle River, N.J. (2003)
7. Allen, L. J. S., Allen, E. J.: A comparison of three different stochastic population models with regard to persistence time. Theor. Pop. Biol., **64**, 439–449 (2003)
8. Allen, L. J. S., Burgin, A. M.: Comparison of deterministic and stochastic SIS and SIR models in discrete time. Math. Biosci., **163**, 1–33 (2000)
9. Allen, L. J. S., Langlais, M., Phillips, C.: The dynamics of two viral infections in a single host population with applications to hantavirus. Math. Biosci. **186**, 191-217 (2003)

10. Anderson, R. M., May, R. M.: Infectious Diseases of Humans. Oxford Univ. Press, Oxford (1992)
11. Antia, R., Regoes, R. R, Koella, J. C., Bergstrom, C. T.: The role of evolution in the emergence of infectious diseases. Nature, **426**, 658–661 (2003)
12. Arnold, L.: Stochastic Differential equations: Theory and Applications. John Wiley & Sons, New York (1974)
13. Bailey, N. T. J.: The Elements of Stochastic Processes with Applications to the Natural Sciences. John Wiley & Sons, New York (1990)
14. Ball, F. G., Lyne, O. D.: Epidemics among a population of households. In: Castillo-Chavez, C., Blower, S., van den Driessche, P., D. Kirschner, D., Yakubu, A. -A. (eds) Mathematical Approaches for Emerging and Reemerging Infectious Diseases An Introduction. Springer-Verlag, New York, pp. 115–142 (2002)
15. Brauer, F., Castillo-Chávez, C.: Mathematical Models in Population Biology and Epidemiology. Springer, New York, Berlin, and Heidelberg (2001)
16. Brauer, F., van den Driessche, P.: Some directions for mathematical epidemiology. In: S. Ruan, S., Wolkowicz, G. S. K., Wu, J. (eds) Dynamical Systems and Their Applications to Biology. Fields Institute Communications 36, AMS, Providence, RI, pp. 95–112 (2003)
17. Daley, D. J., Gani, J.: Epidemic Modelling An Introduction. Cambridge Studies in Mathematical Biology, Vol. 15. Cambridge Univ. Press, Cambridge (1999)
18. Darroch, J. N., Seneta, E: On quasi-stationary distributions in absorbing continuous-time finite Markov chains. J. Appl. Prob., **4**, 192–196 (1967)
19. Diekmann, O., Heesterbeek, J. A. P.: Mathematical Epidemiology of Infectious Diseases: Model building, Analysis and Interpretation. John Wiley & Sons (2000)
20. Foster, F. G.: A note on Bailey's and Whittle's treatment of a general stochastic epidemic. Biometrika, **42**, 123–125 (1955)
21. Gard, T. C.: Introduction to Stochastic Differential Equations. Marcel Dekker, Inc., New York and Basel (1988)
22. Goel, N. S., Richter-Dyn, N.: Stochastic Models in Biology. Academic Press, New York (1974).
23. Greenwood, M.: On the statistical measure of infectiousness. J. Hyg. Cambridge **31**, 336–351 (1931)
24. Harris, T. E.: The Theory of Branching Processes. Springer-Verlag, Berlin (1963)
25. Hethcote, H. W.: Qualitative analyses of communicable disease models. Math. Biosci. **28**, 335–356 (1976)
26. Hethcote, H. W.: The mathematics of infectious diseases. SIAM Review, **42**, 599–653 (2000)
27. Isham, V.: Assessing the variability of stochastic epidemics. Math. Biosci., **107**, 209–224 (1991)
28. Jacquez, J. A., Simon,C. P.: The stochastic SI epidemic model with recruitment and deaths I. Comparison with the closed SIS model. Math. Biosci., **117**, 77–125 (1993)
29. Jagers, P.: Branching Processes with Biological Applications. Wiley, London (1975)
30. Kimmel, M., Axelrod, D. E.: Branching Processes in Biology. Springer, New York, Berlin, Heidelberg (2002)
31. Kloeden, P. E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer-Verlag, New York (1992)

32. Kloeden, P. E., Platen, E., Schurz, H.: Numerical Solution of SDE through Computer Experiments. Springer-Verlag, Berlin (1997)
33. Leigh, E. G., The average lifetime of a population in a varying environment. J. Theor. Biol., **90**, 213–219 (1981)
34. Lloyd, A. L.: Estimating variability in models for recurrent epidemics: assessing the use of moment closure techniques. Theor. Pop. Biol., **65**, 49–65 (2004)
35. Mode, C. J.: Multitype Branching Processes. Elsevier, New York (1971)
36. Mode, C. J. Sleeman, C. K.:. Stochastic Processes in Epidemiology. HIV/AIDS, Other Infectious Diseases and Computers. World Scientific, Singapore, New Jersey (2000)
37. Murray, J. D., Stanley, E. A., Brown, D. L.: On the spatial spread of rabies among foxes. Proc. Roy. Soc. Lond. B, **229**, 111-150 (1986)
38. Nåsell, I.: The quasi-stationary distribution of the closed endemic SIS model. Adv. Appl. Prob., **28**, 895–932 (1996)
39. Nåsell, I.: On the quasi-stationary distribution of the stochastic logistic epidemic. Math. Biosci., **156**, 21–40 (1999)
40. Nåsell, I.: Endemicity, persistence, and quasi-stationarity. In: Castillo-Chavez, C., Blower, S., van den Driessche, P., D. Kirschner, D., Yakubu, A. -A. (eds) Mathematical Approaches for Emerging and Reemerging Infectious Diseases An Introduction. Springer-Verlag, New York, pp. 199–227 (2002)
41. Nisbet, R. M., Gurney, W. S. C.: Modelling Fluctuating Populations. Wiley, Chichester, New York, Brisbane, Toronto, and Singapore (1982)
42. Norden, R. H.: On the distribution of the time to extinction in the stochastic logistic population model. Adv. Appl. Prob., **14**, 687–708 (1982)
43. Ortega, J. M.: Matrix Theory a Second Course. Plenum Press, New York (1987)
44. Sauvage, F., Langlais, M., Yoccoz, N. G.,. Pontier, D.: Modelling hantavirus in fluctuating populations of bank voles: the role of indirect transmission on virus persistence. J. Anim. Ecology, **72**, 1–13 (2003)
45. Schinazi, R. B.: Classical and Spatial Stochastic Processes. Birkhäuser, Boston (1999)
46. Suppo, Ch., Naulin, J. M., Langlais, M., Artois, M.: A modelling approach to vaccination and contraception programmes for rabies control in fox populations. Proc. Roy. Soc. Lond. B, **267**, 1575-1582 (2000)
47. Taylor, H. M., Karlin, S.: An Introduction to Stochastic Modeling. 3rd Ed. Academic Press, San Diego, London, Boston, and New York (1998)
48. Thieme, H. R.: Mathematics in Population Biology. Princeton Univ. Press, Princeton and Oxford (2003)

Index