Some Simple Epidemic Models

California State Polytechnic University, Pomona

and

Loyola Marymount University

Department of Mathematics Technical Report

Jill M. Anderson[*], Adrienne A. Byrne[†], Ruthie P. Fields[‡], Linda M. Segovia[§], Randall J. Swift [¶]

Applied Mathematical Sciences Summer Institute
Department of Mathematics & Statistics,
California State Polytechnic University Pomona
3801 W. Temple Ave.
Pomona, CA 91768

July 2006

[*]South Dakota School of Mines & Technology
[†]Buffalo State College
[‡]Boston College
[§]Honors College of Florida Atlantic University
[¶]California State Polytechnic University Pomona

Abstract

In this paper, a brief overview of the simple deterministic susceptible/infective (SI) epidemic model is detailed. The deterministic model is appropriate for large populations, where random interactions can be viewed as being averaged out. In smaller populations however, random interactions play a significant role. For example, N.T.J. Bailey (1963) considered a stochastic SI epidemic model. With the help of previous work by Gani and Swift (2006), Bailey's solution is modified to include a preventative quarantine of the susceptible population. From this new equation, the associated probability generating function, the transient probabilities, steady state probabilities, expected value, and expected epidemic duration can be found. Computer simulations support the results for this stochastic SI model.

1 Introduction

Epidemiology is the study of the distribution and determinants of chronic and infectious disease prevalence in humans [4]. In 1770, Daniel Bernoulli used mathematics to describe the spread of smallpox in a variolated population. Since then, epidemiological modeling has suffered dips in popularity and reprisals in interest. Today, the most interesting problems in disease tracking, control, and prevention are modeled mathematically using epidemiological processes. More recently, researchers are turning to computer simulations to verify analytic results. When analytic methods become too arduous or even impossible, these simulations can provide insight into complicated situations.

This report uses both analytic methods and computer simulation to describe three epidemic models: the simple deterministic epidemic model, the simple stochastic epidemic model and the simple epidemic model with preventative quarantine. The simple epidemic model with preventative quarantine utilizes recent work of Gani & Swift [5]. In Gani and Swift's work, the term *catastrophe* is equivalent to what we term a *preventative quarantine*: ending an epidemic with a random, instantaneous reduction the susceptible population to 0.

1.1 Basics of Epidemiology

Epidemiology is used to describe the distribution of disease (how the disease spreads), build and test theories (about an epidemic), plan, implement, and evaluate detection (of the disease), control and ultimately prevent disease. By creating mathematical models, diseases can be studied without endangering humans or acting inhumanely. Mathematicians can also calculate the most cost effective methods of handling a disease and predict how many people will be saved by implementing possible solutions.

As stated by Herbert Hethcote [4], "The art of epidemiological modeling is to make suitable choices in the model formulation, so that it is as simple as possible and yet is adequate for the question being considered." Some simple epidemic models take a few parameters into consideration. Although having fewer parameters allows for a simpler representation, this can prove to be a disadvantage. The fewer parameters a model has, the more naive and unrealistic the model becomes. In comparison, a general epidemic model is model which takes more parameters into consideration. This model is more representative of real-life, yet it increases the difficulty of obtaining parameter estimations and known behavior of all solutions.

Even though epidemic modeling began in the mid 1700's, it grew rapidly after World War II when research on the spread of infectious diseases became a high priority. Norman T.J. Bailey, one of the first well-known epidemic mathematical modelers, inspired mathematicians to model the spread of a disease. His results in epidemic modeling were important because they tested possible solutions for ending the epidemic with the least amount of infectives. With improved sanitation, antibiotics, and vaccination programs, the 1970's demanded modeling of newly emerging diseases such as Lyme disease and Legionnaire's disease. [3]. Since the study of disease is a never ending process, mathematical models are constantly being

updated and perfected to include more factors such as age, passive immunity and disease-acquired immunity [4]. To begin our study of disease, we will first consider the deterministic simple epidemic model.

2 The Deterministic Simple Epidemic Model

The simple epidemic model is one where the population consists only of susceptibles and infectives (SI). Once a susceptible becomes infected, he or she remains in that state forever. The deterministic SI model describes an infection with no latency. This infection occurs in a population with no immunity, where no one recovers or dies. For example, the beginning of a highly infectious, but not serious, upper respiratory infection might meet these criteria [1].

The deterministic SI model considers two disjoint populations: the susceptibles and the infectives. Let

$s(t)$ represent the number of susceptibles at time t,
$i(t)$ represent the number of infectives at time t.

For initial conditions, let

$$\begin{aligned} s(0) &= s_0, \\ i(0) &= i_0. \end{aligned}$$

Because the population remains constant, we also have

$$s(t) + i(t) = s_0 + i_0 \text{ for all } t.$$

In the work that follows, let

$$\begin{aligned} s_0 &= n, \\ i_0 &= 1, \end{aligned}$$

$$s(t) + i(t) = n + 1 \text{ for all } t. \tag{1}$$

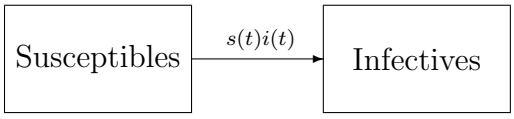

Figure 1: Deterministic SI Model

In other words, the epidemic begins with a single infective in a closed population of n susceptibles. Figure 1 is a pictorial representation of the simple epidemic model. In accordance with the *law of mass action*[2], susceptibles are becoming infectives at a rate proportional

to the product of the sizes of the two populations. So, if $\quad > 0$ is the *contact constant* then $s(t)i(t)$ is the amount of susceptibles become infective per unit of time.

This gives

$$\frac{ds}{dt} = - \ s(t)i(t), \tag{2}$$

$$\frac{di}{dt} = \ s(t)i(t), \tag{3}$$

where $s(t) \geq 0$, $i(t) \geq 0$ and $\quad > 0$ for all t. Looking at the qualitative behavior of these functions, we observe that $\frac{ds}{dt} \leq 0$, which implies $s(t)$ is a monotone non-increasing function. Similarly, $\frac{di}{dt} \geq 0$ which implies $i(t)$ is an monotone non-decreasing function.

Given $\quad > 0$ and $i(t) > 0$ for all t, the system (2)-(3) reaches the equilibrium if and only if $s(t) = 0$. Thus the epidemic ends when all the susceptibles have become infectives.

These same equations can be analyzed analytically. Substituting equation (1) into equation (2) we obtain

$$\frac{ds}{dt} = - \ s((n+1) - s).$$

Separation of variables and partial fraction decomposition gives

$$s(t) = \frac{(n+1)e^{-C \ t(n+1)}}{1 - e^{-C \ t(n+1)}}.$$

Using the initial conditions and solving for C gives

$$s(t) = \frac{n(1+n)}{n + e^{\ (n+1)t}}.$$

Similarly, we can substitute equation (1) into equation (3) to obtain

$$\frac{di}{dt} = \ i((n+1) - i)$$

this can be solved to give

$$i(t) = \frac{(n+1)e^{\ (n+1)t}}{n + e^{\ (n+1)t}}.$$

Figure 2 displays the behavior of the susceptibles and infectives with a population size of 10 and $\quad = 0.25$. This graph demonstrates the quantitative behavior we found with $s(t)$ being a logistically non-increasing function and $i(t)$ being a logistically non-decreasing function.

Another important graph is that of the *epidemic curve*. An example is figure 3. The y-values, $\frac{di}{dt}$, indicate how quickly the epidemic is spreading at time t. In particular, the maximum indicates when the infection is spreading fastest.

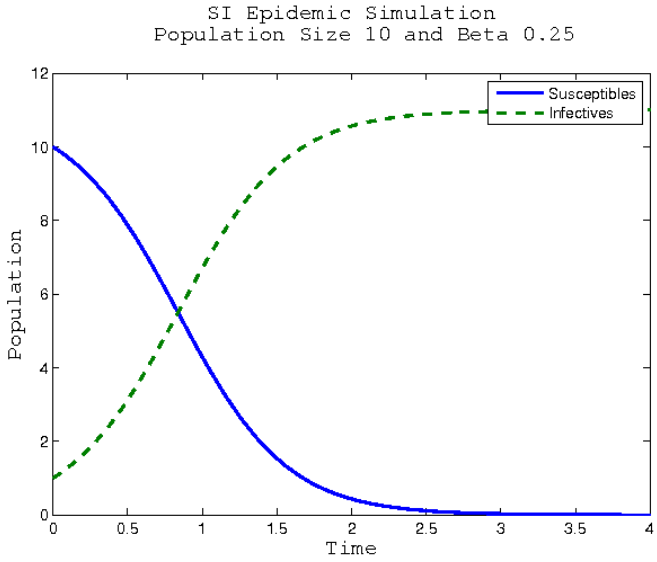

Figure 2: Graph of the SI Model

3 The Stochastic Simple Epidemic Model

3.1 Motivation

"While deterministic methods are useful to understand the spread of an epidemic in a large population these methods are not satisfactory for smaller populations." [2]. The deterministic model ignores the effect of an individual on the larger population. A stochastic model attempts to represent how an individual's behavior affects the spread of an epidemic. Furthermore, the average of the stochastic process describing the epidemic should be close to the values predicted by the deterministic model.

The stochastic SI model was first introduced by M.S. Bartlett (1947) and subsequently studied by Norman T.J. Bailey in the 1950's. In a series of papers (1953, 1955 and 1963), N.T.J. Bailey studied the stochastic SI model. The 1963 publication has a detailed elegant solution and we model our work after this solution.

Our work begins with a thorough comparison of the stochastic and deterministic SI models. The stochastic SI model is similar to the deterministic SI model in that it considers two disjoint populations and contact constant $\quad > 0$. $\mathcal{S}(t)$ represents the susceptibles and $\Phi(t)$ represents the infectives. The population is still closed with $\mathcal{S}(0) = \mathcal{S}_0$ and $\Phi(0) = \Phi_0$. However, for our work, we will take $\mathcal{S}_0 = n$ and $\Phi_0 = 1$. Therefore,

$$\mathcal{S}(t) + \Phi(t) = n + 1 \quad \text{for all} \quad t.$$

Unlike the deterministic model, the stochastic model does not produce an exact function for $\mathcal{S}(t)$. Let

$$P_j(t) = P[\,\mathcal{S}(t) = j \,|\, \mathcal{S}(0) = n]$$

be the probability that there are j susceptibles at time t given that there are initially n susceptibles. Since the population is closed, $P_j(t) = 0$ when $j > n + 1$. So, $\mathcal{S}(t)$ is a

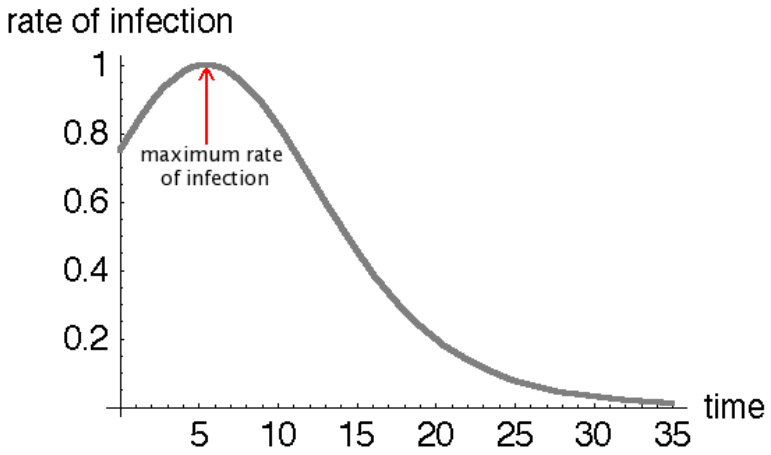

Figure 3: Epidemic Curve w/ $= 0.25$ The peak indicates the halfway point of the epidemic.

stochastic process with values of $0, 1, 2, \ldots, n$. At the initial time $t = 0$,

$$P_j(0) = P[\,\mathcal{S}(0) = j \,|\, \mathcal{S}(0) = n].$$

Thus $P_n(0) = 1$ and $P_j(0) = 0$ when $j \neq n$.

A stochastic version of the SI model can be obtained by assuming that the probability of exactly one contact between a susceptible and an infective in a small time interval is proportional to the product of the number of susceptibles and the number of infectives. That is,

$$\begin{aligned} P[\text{exactly 1 contact in}(t, t+\Delta t)] &= \mathcal{S}(t)\Phi(t)\Delta t \\ &= \mathcal{S}(t)[n+1-\mathcal{S}(t)]\Delta t. \end{aligned}$$

When a contact occurs, there is a probability that another susceptible moves into the infective category. The rate at which a person moves from being susceptible to infective is shown in Table 1 for each possible state size.

Transitions	Rates	
$j \to j-1$	$j(n+1-j)$	for $j = 0, \ldots, n$

Table 1: Transition Rate Table for Stochastic SI Model

We want to determine the probability that the population is in state j at time $t+\Delta t$. We assume Δt is so small that in any time period, only one of two events can occur - either there is a contact between a susceptible and an infective, i.e. one susceptible becomes infective, or there is no contact, i.e. no susceptible becomes infective. This means that there are only two ways the population could end up in state j at time $t + \Delta t$: either 1) the susceptible

population was at state $j+1$ at time t and a contact occurred or 2) the susceptible population was at state j at time t and no contact occurred. This transition is shown in Figure 4.

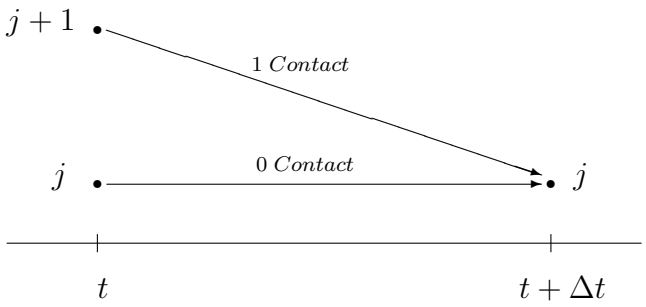

Figure 4: Time Dependent Transition Diagram for the Stochastic SI Model

The state diagram in Figure 5 indicates the probability of staying in state j as well as the probability of moving from state j to state $j-1$ for $j = 1, \ldots n$. The epidemic begins with n susceptibles; and it ends when the number of susceptibles equals 0. The law of mass action dictates that the transition rate between states is $\mathcal{S}(t)\Phi(t)$. The probability of having $n-1$ susceptibles at time t is the probability that

- there were n susceptibles at time $t - \Delta t$ and a contact occurred in $(t - \Delta t, t)$, so that susceptible population decreased by 1

or

- there was already $n - 1$ susceptibles and no contact occurred. The probability that this occurs is $1 - 2\ (n-1)\Delta t$ which is the complement of the probability that the population decreases from $n - 1$ to $n - 2$.

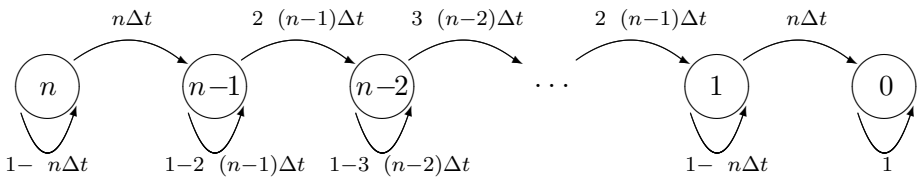

Figure 5: State Diagram for the Stochastic SI Model

In mathematical terms this diagram represents

$$\begin{aligned} P_j(t+\Delta t) &= P(0 \text{ contacts during } \Delta t \,|\, \mathcal{S}(t) = j) + P(1 \,\text{contact during } \Delta t \,|\, \mathcal{S}(t) = j+1) \\ &= (1 - \ j(n+1-j)\Delta t)P_j(t) + \ (j-1)(n-j)\Delta t\ P_{j+1}(t). \end{aligned}$$

Rearranging gives,

$$P_j(t+\Delta t) - P_j(t) \;=\; -\; j(n+1-j)\Delta t\; P_j(t) + \;\; (j-1)(n-j)\Delta t\; P_{j+1}(t),$$

so that

$$\frac{P_j(t+\Delta t) - P_j(t)}{\Delta t} \;=\; -\; j(n+1-j)P_j(t) + \;\; (j-1)(n-j)P_{j+1}(t).$$

Taking the limit of both sides as Δt goes to zero

$$\lim_{\Delta t \to 0} \frac{P_j(t+\Delta t) - P_j(t)}{\Delta t} \;=\; P_j'(t) = -\; j(n+1-j)P_j(t) + \;\; (j-1)(n-j)P_{j+1}(t)$$

gives the following equations which are known as the *forward Kolmogorov equations* [7] for the simple SI epidemic model:

$$P_j'(t) \;=\; -\; j(n+1-j)P_j(t) + \;\; (j-1)(n-j)P_{j+1}(t) \quad \text{for } j = 0, \ldots, n-1 \tag{4}$$

$$P_n'(t) \;=\; -\; nP_n(t). \tag{5}$$

The goal for these Kolmogorov equations is to solve for the probabilities $P_j(t)$. One method of solving these differential equations is to recursively solve the $P_j'(t)$. This method is extremely tedious and leads to an algebraic nightmare. An alternative approach is obtained by using probability generating functions (PGFs) and partial di erential equations (PDEs).

3.2 Solving PDEs and PGFs

A probability generating function (PGF) is a power series whose coefficients are probabilities. In the current context it takes the form

$$y(z,t) \;=\; \sum_{j=0}^{n} P_j(t) z^j = P_0(t) + P_1(t)z + \ldots + P_n(t)z^n.$$

Our goal is to find $P_j(t)$. But, since we know $P_j'(t)$, we take the partial derivative of the PGF with respect to t, so that

$$\frac{\partial y}{\partial t} \;=\; \sum_{j=0}^{n} P_j'(t) z^j. \tag{6}$$

Using the derivative in equation (6) and the Kolmogorov equations we substitute to obtain

$$\begin{aligned}
\frac{\partial y}{\partial t} \;&=\; \left[\sum_{j=0}^{n} \left[(j+1)(n+1-(j-1))P_{j+1}(t) - j(n+1-j)P_j(t)\right] z^j\right] \\
&=\; \left[\sum_{j=0}^{n} \left[-j(n+1)P_j(t)z^j + j^2 P_j(t) z^j + (j+1)(n+1)P_{j+1}(t)z^j - (j+1)^2 P_{j+1}(t) z^j\right]\right] \\
&=\; \left[-\sum_{j=0}^{n} j(n+1)P_j(t)z^j + \sum_{j=0}^{n} j^2 P_j(t) z^j + \sum_{j=0}^{n} (j+1)(n+I)P_{j+1}(t)z^j - \sum_{j=0}^{n} (j+1)^2 P_{j+1}(t) z^j\right]
\end{aligned}$$

Using a common substitution the resulting equation is

$$\begin{aligned}\frac{\partial y}{\partial t} &= \left[-(n+1)z\frac{\partial y}{\partial z}+z^2\frac{\partial^2 y}{\partial z^2}+z\frac{\partial y}{\partial z}+(n+1)\frac{\partial y}{\partial z}-z\frac{\partial^2 y}{\partial z^2}-\frac{\partial y}{\partial z}\right]\Bigg(\\ &= \left[\frac{\partial y}{\partial z}[-(n+1)(z-1)+(z-1)]+\frac{\partial^2 y}{\partial z^2}z(z-1)\right]\Bigg(\end{aligned}$$

After a little algebra, the resulting PDE is

$$\frac{\partial y}{\partial t} = \left[z(z-1)\frac{\partial^2 y}{\partial z^2}-(z-1)((n+1)-1)\frac{\partial y}{\partial z}\right]\Bigg(\tag{7}$$

N.T.J. Bailey solved equation 7 in 1963. The PDE in equation (7) can be solved using separation of variables. Let

$$y(z,t) = f(z)h(t),$$

so that the partial derivatives are

$$\frac{\partial y}{\partial t} = f(z)h'(t) \qquad \text{and} \qquad \frac{\partial y}{\partial z} = f'(z)h(t)$$
$$\frac{\partial^2 y}{\partial z^2} = f''(z)h(t).$$

These new equations can be substituted into equation (7) and then simplified, which yields

$$f(z)h'(t) = [z(z-1)f''(z)h(t)-(z-1)((n+1)-1)f'(z)h(t)].$$

Factoring out $h(t)$, we obtain

$$f(z)h'(t) = h(t)[z(z-1)f''(z)-(z-1)((n+1)-1)f'(z)]$$

which is

$$\frac{h'(t)}{h(t)} = \frac{}{f(z)}[z(z-1)f''(z)-(z-1)((n+1)-1)f(z)].$$

Since the two sides of this equation are equal, then

$$-c = \frac{h'(t)}{h(t)} \tag{8}$$

and

$$-c = \frac{}{f(z)}[z(z-1)f''(z)-(z-1)((n+1)-1)f(t) \tag{9}$$

where c is a constant to be determined.

Solving equation (8) results in

$$h(t) \;=\; ke^{-ct}.$$

To solve for $f(z)$ in equation (9), we simplify the equation

$$-c \;=\; \frac{}{f}[z(z-1)f'' - (z-1)((n+1)-1)f']$$

so that

$$0 \;=\; z(1-z)f'' - n(1-z)f' - \frac{cf}{}. \tag{10}$$

3.3 Transient Probability

The PDE achieved in the previous section is extremely difficult to sovle because of repeated eigenvalues. To get around this, we follow Bailey and employ perturbation: let $N = n + \epsilon$ where $\epsilon > 0$. This change makes (10) solvable and with *Mathematica* we achieve

$$y(z,t) = \sum_{j=0}^{n} \left(d_j e^{-j(N+1-j)\ t}\, {}_2F_1(-j, j-N-1, -N, z) \right. \tag{11}$$

where ${}_2F_1(-j, j-N-1, -N, z)$ is a hypergeometric function, and

$$d_j \;=\; \frac{(-1)^j n!(N-2j+1)N!}{j!(n-j)!(N-n)! \prod_{r=0}^{n}(N-j-r+1)}\,. \tag{12}$$

But what is ${}_2F_1[-j, j-N-1, -N, z]$? The standard hypergeometric function ${}_2F_1$ is

$${}_2F_1[a,b,c,z] \;=\; \sum_{\ell=0}^{\infty} \left(\frac{(a)_\ell\,(b)_k}{(c)_\ell}\,\frac{z^\ell}{\ell!} \right. \tag{13}$$

where $(a)_\ell \;=\; a(a+1)(a+2)\ldots(a+\ell-1)$ with $(a)_0 = 1$ is the rising factorial notation or Pochhammer symbol.

Recall from 3.2,

$$\begin{aligned} y(z,t) \;&=\; \sum_{k=0}^{n} \left(P_k(t)\, z^k \right. \\ &=\; P_0(t) + P_1(t)z + P_2(t)z^2 + \ldots + P_n(t)z^n, \end{aligned}$$

where the transient probabilities for this process are simply the coefficients of this polynomial.

Consequently setting $z = 0$ gives,

$$y(0,t) \;=\; P_0(t)$$

where

$$P_0(t) \;=\; \sum_{j=0}^{n} d_j\, e^{-j(N-j+1)\;t}\, {}_2F_1(-j, j-N-1, -N, 0)\left(\frac{1}{0!}\right)$$

is our PGF evaluated at $z = 0$.

Expanding ${}_2F_1$ gives

$$\begin{aligned} {}_2F_1(-j, j-N-1, -N, z) &= \sum_{\ell=0}^{\infty} \frac{(-j)_\ell (j-N-1)_\ell}{(-N)_\ell} \frac{z^\ell}{\ell!} \\ &= 1 + \frac{(-j)_1 (j-N-1)_1}{(-N)_1} \frac{z}{1!} + \frac{(-j)_2 (j-N-1)_2}{(-N)_2} \frac{z^2}{2!} + \cdots . \end{aligned}$$

From this expansion, one can see that by substituting $z = 0$ into ${}_2F_1$, gives

$${}_2F_1(-j, j-N-1, -N, 0) \;=\; 1.$$

Therefore

$$P_0(t) \;=\; \sum_{j=0}^{n} d_j\, e^{-j(N-j+1)\;t}.$$

Taking derivatives of our general PGF and evaluating them at $z = 0$ yields

$$P_1(t) \;=\; \left(\frac{1}{1!}\right) \frac{\partial y}{\partial z}_{\;z=0}, \qquad P_2(t) \;=\; \left(\frac{1}{2!}\right) \frac{\partial^2 y}{\partial z^2}_{\;z=0}, \;\ldots, \quad P_k(t) \;=\; \left(\frac{1}{k!}\right) \frac{\partial^k y}{\partial z^k}_{\;z=0}.$$

Since ${}_2F_1(-j, j-N-1, -N, z)$ is the only part of $y(z,t)$ that contains z, then the rest of the equation will be treated as a constant when evaluating the partial derivatives of $y(z,t)$ The expansion

$${}_2F_1(-j, j-N-1, -N, z) \;=\; 1 + \frac{(-j)_1 (j-N-1)_1}{(-N)_1} \frac{z}{1!} + \frac{(-j)_2 (j-N-1)_2}{(-N)_2} \frac{z^2}{2!} + \ldots \tag{14}$$

(15)

has the derivative of ${}_2F_1$ as

$$\begin{aligned} \frac{\partial}{\partial z}\, {}_2F_1 &= 0 + \frac{(-j)_1 (j-N-1)_1}{(-N)_1}\, 1 + \frac{(-j)_2 (j-N-1)_2}{(-N)_2}\, z \\ &\quad + \frac{(-j)_3 (j-N-1)_3}{(-N)_3}\, \frac{z}{2} + \cdots . \end{aligned}$$

When this derivative is evaluated at $z = 0$, we have

$$\frac{\partial}{\partial z}\,{}_2F_1\Big|_{z=0} = \frac{(-j)_1(j-N-1)_1}{(-N)_1},$$

which gives the probability for $P_1(t)$ as

$$P_1(t) = \sum_{j=0}^{n} d_j\, e^{-j(N-j+1)\ t}\, \frac{(-j)_1(j-N-1)_1}{(-N)_1}.$$

In a similar manner, $P_k(t)$ is obtained as

$$P_k(t) = \sum_{j=0}^{n} d_j\, e^{-j(N-j+1)\ t}\, \frac{(-j)_k(j-N-1)_k}{(-N)_k}\, \frac{1}{k!} \quad \text{for } k = 0, ..., n.$$

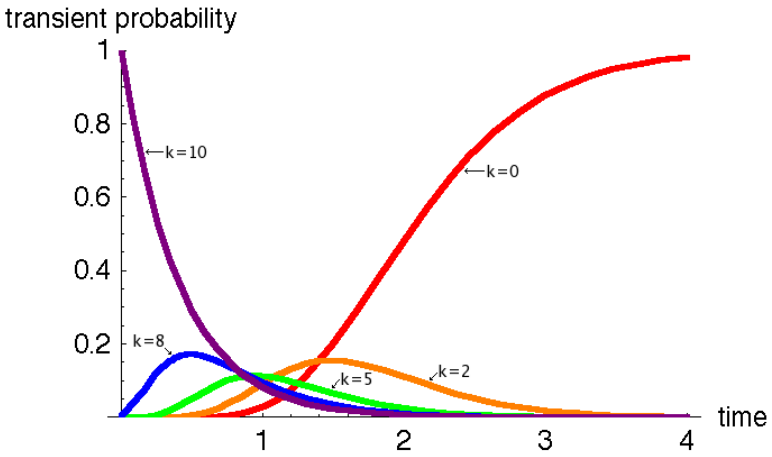

Figure 6: Transient Probabilities for stochastic SI Model with = 0.25 and n=10.

Figure 6 is a graph of the transient probabilities for various $P_k(t)'s$ when = 0.25 and the total population size is $n + 1 = 11$. We can see, for example, that as time increases, the probability that all ten susceptibles remain uninfected, $P_{10}(t)$, decreases exponentially to zero. The probability that all ten of the susceptibles become infected, $P_0(t)$, increases to 1 as time increases. This steady state probability will be shown in the following section.

3.4 Steady State Probabilities

The steady state probability for any given PGF is the limit of the transient probability as time, t, goes to infinity. Since

$$\lim_{t\to\infty} P_k(t) = \lim_{t\to\infty} \sum_{j=0}^{n} d_j\, e^{-j(N-j+1)t}\, \frac{(-j)_k(j-N-1)_k}{(-N)_k}\, \frac{1}{k!} = \begin{cases} 1 & \text{for }, \ k = 0 \\ 0 & \text{for }, k > 0 \end{cases}$$

This is because

$$\lim_{t\to\infty} P_0(t) \;=\; \lim_{t\to\infty} \sum_{j=0}^{n} d_j\, e^{-j(N-j+1)t}$$

. So for $j = 0$ then

$$\lim_{t\to\infty} P_0(t) \;=\; \lim_{t\to\infty} d_0\, e^0 = 1.$$

So as time goes on, there is a 100% chance that all the susceptibles will get infected

And

$$\lim_{t\to\infty} e^{-j(N-j+1)t} = e^{-\infty} = 0$$

3.5 Moments In Time

The PGF $y(z,t)$ can also be used to find the *factorial moment* of the process $\mathcal{S}(t)$. Repeated differentiation and evaluation $z = 1$ gives

$$E[(\mathcal{S}(t)(\mathcal{S}(t)-1)(\mathcal{S}(t)-2)\ldots(\mathcal{S}(t)-k)] \;=\; \frac{\partial^k y}{\partial z^k}\Big|_{z=1}.$$

The expected value is

$$E[\mathcal{S}(t)] \;=\; \frac{\partial y}{\partial z}\Big|_{z=1}.$$

Remember that the hypergeometric function ${}_2F_1$ is

$${}_2F_1(-j, j-N-1, -N, z) \;=\; \sum_{k=0}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{z^k}{k!}$$

and that expanded derivative for it is

$$\begin{aligned}\frac{d}{dz}\,{}_2F_1(-j, j-N-1, -N, z) \;=\; & 0 + \frac{(-j)_1(j-N-1)_1}{(-N)_1} + \frac{(-j)_2(j-N-1)_2}{(-N)_2} \\ & + \frac{(-j)_3(j-N-1)_3}{(-N)_3}\frac{1}{2} + \ldots\end{aligned}$$

When this derivative is evaluated at $z = 1$, the it becomes

$$\frac{d}{dz}\,{}_2F_1(-j, j-N-1, -N, 1) \;=\; \sum_{k=1}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-1)!}\,.$$

Therefore,

$$E[\mathcal{S}(t)] \quad = \quad \sum_{j=0}^{n} d_j e^{-j(N-j+1)\ t} \sum_{k=1}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-1)!}.$$

Similarly,

$$\frac{d^i}{dz^i}\ {}_2F_1(-j, j-N-1, -N, 1) \quad = \quad \sum_{k=i}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-i)!} \quad \text{for } i \ = \ 0 \ldots n.$$

Therefore,

$$E[(\mathcal{S}(t)(\mathcal{S}(t)-1)(\mathcal{S}(t)-2)\ldots(\mathcal{S}(t)-k)] \quad = \quad \sum_{j=0}^{n} d_j e^{-j(N-j+1)\ t} \sum_{k=i}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-i)!}.$$

Knowing the general form for the *factorial moment*, we can easily find the steady state as time goes to infinity. Since

$$\lim_{t\to\infty} e^{-j(N-j+1)\ t} \quad = \quad 0$$

Then

$$\lim_{t\to\infty} E[S(S-1)(S-2)\ldots(S-k)] \quad = \quad 0$$

So the expected number of susceptibles that do not get infected is zero.

3.6 Duration of Epidemic

The duration of the stochastic epidemic tells us how long the epidemic lasts . Keeping that in mind allows us to find an analytic formula for it. Since $\mathcal{S}(t)$ is the number of susceptibles at time t and if $\mathcal{S}(0) = n$, then $\mathcal{S}(t)$ is a death process that evolves by unit decrements at the epochs $t_n, t_{n-1}, \ldots, t_1$. This behavior is displayed in Figure 7.

Let T_j be the length of the time interval for which there are exactly j susceptibles if there is initially 1 infective.The T_j's are independent exponentially distributed random variables with expected values.

$$E[T_j] \quad = \quad \frac{1}{j(n+1-j)} \quad 1 \le j \le n$$

Thus, letting $\tau = \inf\{t : s(t) = 0\}$ so that τ is the duration time of the epidemic, we have

$$\tau = \sum_{j=1}^{n} T_j.$$

Therefore,

$$E[\tau] = E\left[\sum_{j=1}^{n} T_j\right] = \sum_{j=1}^{n} E[T_j] = \sum_{j=1}^{n} \frac{1}{j(n+1-j)}$$

is the expected duration.

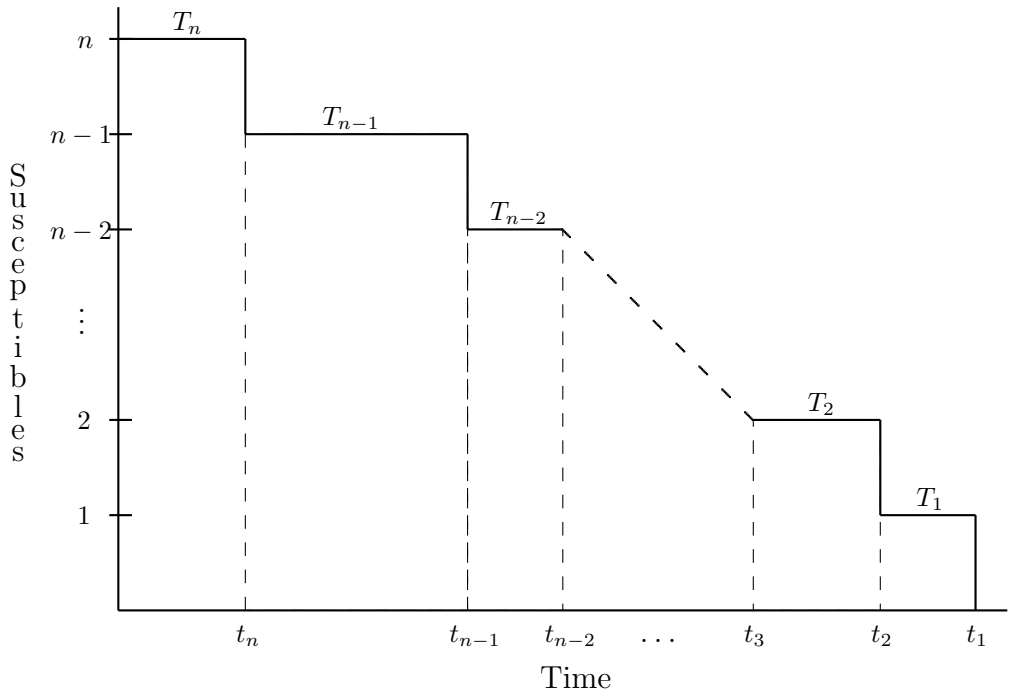

Figure 7: Stochastic SI Expected Epidemic Duration Graph

3.7 Simulation Results

Computer simulations are an alternative method of determining the expected duration of the epidemic. To support our formulas, wrote *Matlab* simulations to model the duration of the epidemic. Figure 8 compares three different values. Table 2 shows that simulation averages support expected duration formula. Notice that for increasing the duration is lower because the infection spreads more quickly.

	Formula	*Simulation*
0.25	2.01	2.13
0.50	1.07	1.01
0.75	0.67	0.71

Table 2: A comparison of expected duration formulas with the average duration for 100,000 trials with $n = 10$.

4 Stochastic SI Model with Preventative Quarantine

Having considered the simple stochastic process, we would like now to add to this process a *preventative quarantine*. The idea of preventative quarantine is similar to that of a process with catastrophes. "The simple death process, the survival of susceptibles in a carrier-borne epidemic, the birth-death and immigration process, the unbiased random walk and the barber shop queue, are all subject to random catastrophes occurring as a Poisson process" [6].

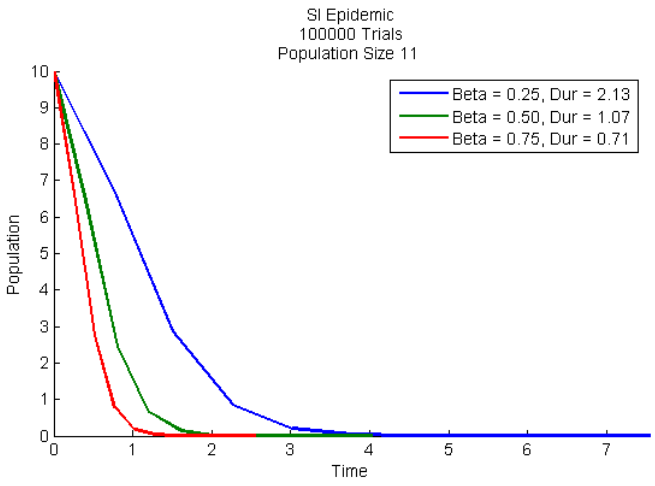

Figure 8: Three sample epidemic runs with varying s

The epidemic ends when there are no more susceptibles. This can happen two ways; either the susceptibles become quarantined at a random point in time or the susceptibles all become infective. Remember we are modelling the susceptibles and not the infectives since N.T.J. Bailey solved the stochastic SI model for the susceptibles. Therefore in this model, we quarantine the susceptibles. We named the process of removing the susceptibles before they become infective, *preventative quarantine* so that it is clear we are saving the susceptibles from becoming sick. Therefore, our type of quarantine is one of prevention.

We assume that there is some way of detecting who is infected and who is not. Then upon detection we let $\delta > 0$ be the rate of how quickly we intervene and quarantine the susceptibles. The rest of the assumptions are the same as 3. Let $\mathcal{S}(t)$ be the number of susceptibles in the population at time t.

Let $\Phi(t)$ be the number of infectives at time t. Suppose at t_0 we have only one infective then $\Phi_0 = 1$ and our total population is $n + 1$. Suppose that there are $\mathcal{S}_0$ susceptibles initially, then $\mathcal{S}(0) = S_0$, and $\mathcal{S}(t) + \Phi(t) = n + 1$. Let > 0 be the contact constant. Then the rate at which the susceptibles are becoming infected is $\mathcal{S}(t)\Phi(t)$. A pictorial representation shown in Figure 9 is helpful to clarify the process. Similar to the SI model

without quarantine, when a contact occurs, another susceptible moves into the infective category. With quarantine, however, the epidemic can end at any point in time. The rate at which a person might move from being susceptible to infective is shown in Figure 10 for each possible state size. A pictorial representation of this is shown in Figure 4.

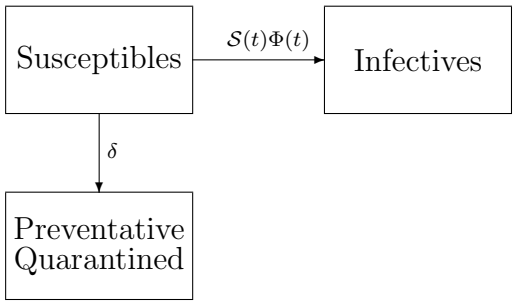

Figure 9: Stochastic SI Model with Preventative Quarantine

Transitions	Rates
$j \to j-1$	$j(n+1-j)+\delta$
$j \to 0$	δ

for j = n, n-1, . . . , 2, 1

Figure 10: Transition Diagram for Stochastic SI Model

Our goal is to determine

$$P_j(t) = P[S(t) = j | S(0) = S_0].$$

As with the SI model, we can create the Kolmogorov equations to get the $P_j(t)$'s. It is helpful to see the state diagram when creating the Kolmogorov equations. Observing Figure 4 we

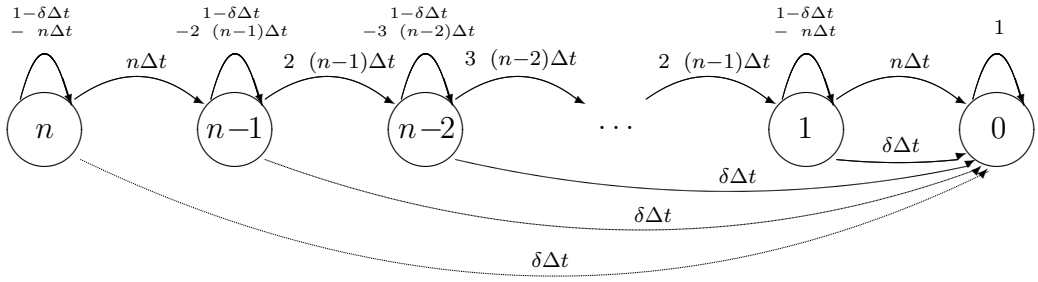

Figure 11: State Diagram for Stochastic SI Model with Preventative Quarantine

can see that

$$P[\text{staying at state } n] \;=\; 1 - \delta\Delta t - \;\; n\Delta t$$

and

$$P[\text{moving from state } n \text{ to } n-1] \;=\; \;\; n\Delta t + \delta\Delta t.$$

This produces the forward Kolmogorov equations

$$P_j'(t) \;=\; \;\; (j+1)(n-j)P_{j+1}(t) - (\;\; j(n+1-j)+\delta)P_j(t) \quad \text{for } 1 \le j < n$$

$$P_n'(t) \;=\; -(\;n+\delta)\,P_n(t).$$

We could derive the PGF for this distribution by first principles as was done for the stochastic epidemic model with out preventative quarantine, however, this process would be quite lengthy. An alternative approach is to follow a recent method of J. Gani and R.J. Swift (2006) for modifying the PGF of a random process to include a random catastrophe. Given a PGF, $y(z,t)$, then $G(z,t)$ is the PGF for the same process with a catastrophe constant δ:

$$G(z,t) \;=\; e^{-\delta t}y(z,t) + \int_0^t \delta e^{-\delta v}y(z,v)dv, \tag{16}$$

thus, this PGF is the general form for any process with a catastrophe. Substituting the PGF we developed for the simple stochastic model without quarantine,

$$y(z,t) \;=\; \sum_{j=0}^{n}\left(d_j\, e^{-j(N-j+1)\;t}\;{}_2F_1(-j,j-N-1,-N,z), \right. \tag{17}$$

where

$$d_j \;=\; \frac{(-1)^j\, n!\,(N-2j+1)N!}{j!\,(n-j)!(N-n)!\;\prod_{k=0}^{n}\,(N-j-k+1)},$$

and substituting this into equation (16) we are able to create a PGF for the stochastic simple epidemic model with preventive quarantine.

That is,

$$\begin{aligned}
G(z,t) \;&=\; e^{-\delta t}\left[\sum_{j=0}^{n}\left(d_j e^{-j(N-j+1)\;t}\;{}_2F_1(-j,j-N-1,-N,z)\right]\right(\\
&\quad + \int_0^t \delta e^{-\delta v} d_j e^{-j(N-j+1)\;v}\;{}_2F_1(-j,j-N-1,-N,z)dv. \\
&=\; e^{-\delta t}\sum_{j=0}^{n}\left(d_j\, e^{-j(N-j+1)\;t}\;{}_2F_1(-j,j-N-1,-N,z)\right. \\
&\quad + \delta\, d_j\;{}_2F_1(-j,j-N-1,-N,z)\int_0^t e^{-\delta v}\, e^{-j(N-j+1)\;v}dv \\
&=\; \sum_{j=0}^{n}\left(d_j\, e^{t(j^2\;-j(1+N)\;-\delta)}\;{}_2F_1(-j,j-N-1,-N,z)\right. \\
&\quad + \delta\, d_j\;{}_2F_1(-j,j-N-1,-N,z)\,\frac{e^{t(j^2\;-j(1+N)\;-\delta)}-1}{j^2\;\;-j(1+N)\;\;-\delta} \\
&=\; \sum_{j=0}^{n}\left(d_j\;{}_2F_1(-j,j-N-1,-N,z)\quad e^{t(j^2\;-j(1+N)\;-\delta)}+\frac{(e^{t(j^2\;-j(1+N)\;-\delta)}-1)}{j^2\;\;-\,j(1+N)\;\;-\delta}\right).
\end{aligned}$$

If we let

$$A_j(t) \;=\; e^{\;j(j-N-1)t-t\delta} \;+\; \frac{(e^{t(j^2\;-j(1+N)\;-\delta)}\;-\;1)}{j^2\;\;-\;j(1+N)\;\;-\delta},$$

then we can write this PGF as

$$G(z,t) \;=\; \sum_{j=0}^{n} d_j\, A_j(t)\, {}_2F_1(-j, j-N-1, -N, z).$$

4.1 Transient Probabilities

The transient probabilities for this model can be found using the same process as before, since

$$\begin{aligned} G(z,t) &= \sum_{k=0}^{\infty} P_k(t)\, z^k \\ &= P_0(t) + P_1(t)z + P_2(t)z^2 + \ldots \end{aligned}$$

then $G(0,t) = P_0(t)$ so

$$P_0(t) \;=\; \sum_{j=0}^{n} d_j\, A_j\, {}_2F_1(-j, j-N-1, -N, 0)\left(\frac{1}{0!}\right).$$

Now recalling equation (15)

$$\begin{aligned} {}_2F_1(-j, j-N-1, -N, z) &= \sum_{k=0}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k}\frac{z^k}{k!} \\ &= 1 + \frac{(-j)_1 (j-N-1)_1}{(-N)_1}\frac{z}{1!} + \frac{(-j)_2 (j-N-1)_2}{(-N)_2}\frac{z^2}{2!} + \ldots \end{aligned}$$

we have

$${}_2F_1(-j, j-N-1, -N, 0) \;=\; 1$$

. We see that

$$P_0(t) = \sum_{j=0}^{n} d_j\, A_j(t).$$

As with the simple stochastic model, we can find the transient probabilities, $P_1(t), P_2(t), \ldots, P_i(t)$:

$$P_1(t) \;=\; \left(\frac{1}{1!}\right)\frac{\partial G}{\partial z}\,|_{z=0}\;, \quad P_2(t) \;=\; \left(\frac{1}{2!}\right)\frac{\partial^2 G}{\partial z^2}\,|_{z=0}\;, \quad \ldots, \quad P_k(t) \;=\; \left(\frac{1}{k!}\right)\frac{\partial^k G}{\partial z^k}\,|_{z=0}\;.$$

It follows that

$$P_k(t) \;=\; \sum_{j=0}^{n} d_j\, A_j(t)\, \frac{(-j)_k (j-N-1)_k}{(-N)_k}\, \frac{1}{k!}.$$

A sample graph is show in Figure 12. This graph shows the transient probabilities for various states of j when $\quad = 0.25$ and $\delta = 1$ for a total population of $n + 1 = 11$. We again see that as time increases, the probability of all ten susceptibles remain uninfected decreases. And the probability that all ten of the susceptibles get infected, $P_0(t)$, increases. However, if $\delta > 0$, the steady state probability for $j = 0$, (all the susceptibles getting infected), will never reach 1. And the probability of reaching any other state will never reach 0, not everyone will get infected. Note that the graph in figure (12) shows that the SIQ process has steady-state probabilities. We find the

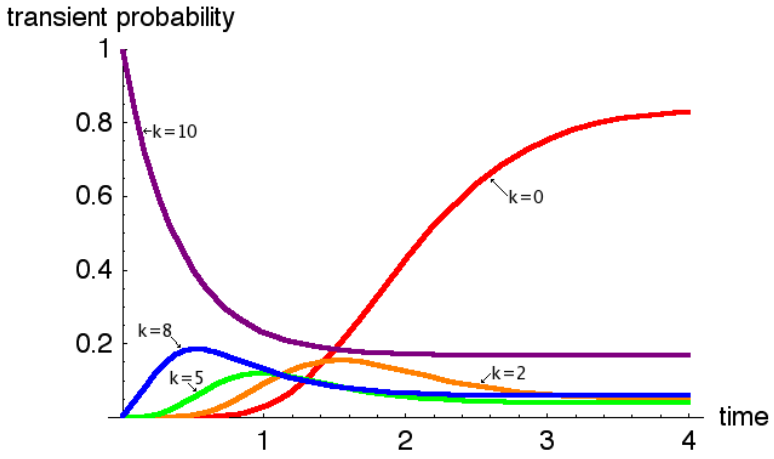

Figure 12: Transient Probabilities for stochastic SI Model with Preventative Quarantine with $\quad = 0.25, \delta = 0 and n = 10$.

[explain]

4.2 Steady State Probabilities

Now that we have our transient probabilities, we can readily find the steady state probabilities. Since $A_j(t)$ is the only part that contains t, it follows that

$$\lim_{t\to\infty} e^{t(j^2 \quad -j(1-N) \quad -\delta)} + \frac{\delta(e^{t(j^2 \quad -j(1+N) \quad -\delta)} - 1)}{j^2 \quad -j(1+N) \quad -\delta} = \frac{-\delta}{j^2 \quad -j(1+N) \quad -\delta}$$

because $j^2 \;\; - j \;\; - j\,N \;\; - \delta$ will always be negative as $j\,N > j^2$. Therefore

$$\begin{aligned}
P_0 &= \sum_{j=0}^{n} d_j \, \frac{-\delta}{j^2 \;\; - j(1+N) \;\; - \delta} \\
P_1 &= \sum_{j=0}^{n} d_j \, \frac{-\delta}{j^2 \;\; - j(1+N) \;\; - \delta} \, \frac{j(j-N-1)}{-N} \\
&\vdots \\
P_k &= \sum_{j=0}^{n} d_j \, \frac{-\delta}{j^2 \;\; - j(1+N) \;\; - \delta} \, \frac{(j)_i (j-N-1)_i}{(-N)_i} \quad \text{for } k = 0, \ldots, n.
\end{aligned}$$

We next consider the moments of the process.

4.3 Moments In Time

Recall the PGF can be used to find the factorial moment of the process $\mathcal{S}(t)$ where

$$E[(\mathcal{S}(t)(\mathcal{S}(t)-1)(\mathcal{S}(t)-2)\ldots(\mathcal{S}(t)-k)] = \frac{\partial^k G}{\partial z^k}_{\;\; z=1}$$

and

$$E[\mathcal{S}(t)] = \frac{\partial G}{\partial z}_{\;\; z=1}.$$

Using methods discussed earlier we find

$$E[\mathcal{S}(t)] \;\; = \;\; \sum_{j=0}^{n} d_j A_j(t) \sum_{k=1}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{z^k}{(k-1)!},$$

and

$$E[(\mathcal{S}(t)(\mathcal{S}(t)-1)(\mathcal{S}(t)-2)\ldots(\mathcal{S}(t)-k)] \;\; = \;\; \sum_{j=0}^{n} d_j \, A_j(t) \sum_{k=i}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-i)!} \, .$$

Knowing this, we can easily find the steady state factorial moments

$$E[\mathcal{S}(\mathcal{S}-1)(\mathcal{S}-2)\ldots(\mathcal{S}-k)] \;\; = \;\; \sum_{j=0}^{n} d_j \, \frac{-\delta}{j^2 - j(1+N) - \delta} \sum_{k=i}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-i)!} \, .$$

From the factorial moments we are able to produce the variance for our distribution. The variance is

$$V[\mathcal{S}] \; = \; E[\mathcal{S}(\mathcal{S}-1)] + E[\mathcal{S}] - (E[\mathcal{S}])^2.$$

So that

$$V[\mathcal{S}] = \sum_{j=0}^{n} d_j \frac{-\delta}{j^2 - j(1+N-)\delta} \sum_{k=1}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-2)!} \Bigg) + \sum_{j=0}^{n} d_j \frac{-\delta}{j^2 - j(1+N-)\delta} \sum_{k=1}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-1)!} \Bigg) - \sum_{j=0}^{n} d_j \frac{-\delta}{j^2 - j(1+N) - \delta} \sum_{k=1}^{\infty} \frac{(-j)_k (j-N-1)_k}{(-N)_k} \frac{1}{(k-1)!} \Bigg)^2 .$$

While this equation is impractical for computation purposes, we return to the simulations used earlier to give numerical approximations.

4.4 Expected Duration

Recall the duration of an epidemic without quarantine is the time when all the susceptibles become infected. In the SIQ model, the epidemic can end at any point in time. This model is primarily dependent on δ, the rate at which preventive quarantine in introduced. The $T_j's$ are still independently distributed random variables but the expected values are now:

$$E[T_j] = \frac{1}{j(n+1-j)+\delta} \quad 1 \leq j \leq n$$

By letting $\tau = \inf\{t : s(t) = 0\}$ where τ is the duration time of the epidemic, we have

$$\sum_{j=1}^{n} T_j \Bigg(\tag{18}$$

so that

$$E[\tau] = E\left[\sum_{j=1}^{n} T_j\right] \Bigg(= \sum_{j=1}^{n} E[T_j] = \sum_{j=1}^{n} \Bigg(\frac{1}{j(n+1-j)+\delta}$$

is the expected duration.

Notice that

$$\frac{1}{j(n+1-j)+\delta} \leq \frac{1}{j(n+1-j)} \quad \text{for} j = 1, \ldots, n.$$

Therefore the expected epidemic duration with preventative quarantine is less than the expected epidemic duration without quarantine which is represented in Figure 13.

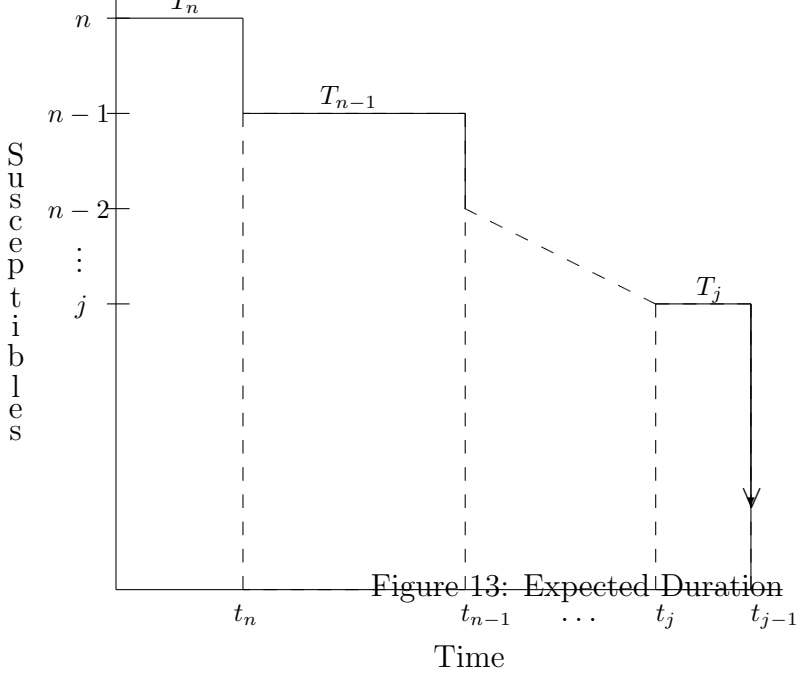

Figure 13: Expected Duration

4.5 Simulation Results

To determine if the derived formulas were correct, we wrote Matlab simulations to model the duration of the epidemic. Figure 14 compares three different values and Figure 15

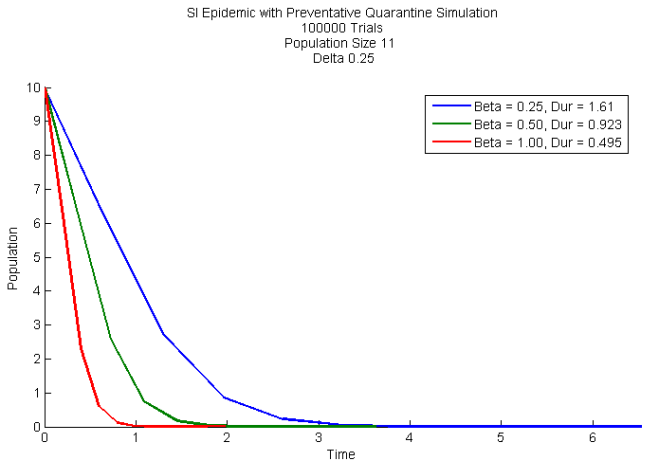

Figure 14: Three sample epidemic runs with varying s and fixed δs

5 Conclusion

We approached our model by first looking at the deterministic SI model and the stochastic SI model where the epidemic ends when all susceptibles become infected. As a way of shortening the duration of the epidemic we added a preventative quarantine which allowed for a random removal of the susceptibles from the population resulting in less infecteds. Once we formulated our model, we calculated the average duration of the epidemic analytically

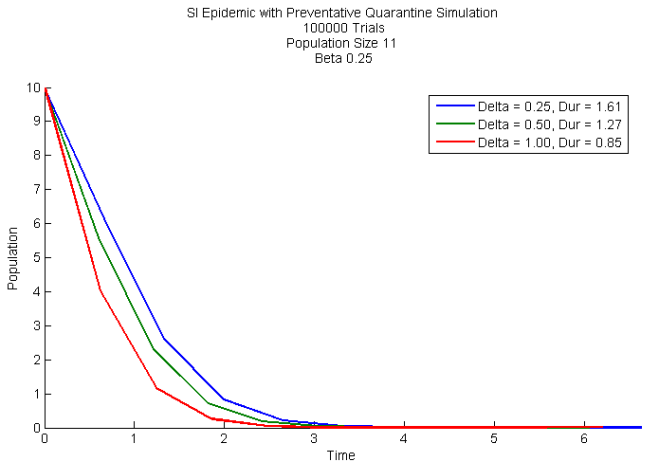

Figure 15: Three sample epidemic runs with varying s and fixed δs

and by running computer simulations. In the future we would like to apply preventative quarantine to the general stochastic SIR model.

6 Acknowledgements

This research was conducted at the Applied Mathematical Sciences Summer Institute (AMSSI) and has been partially supported by grants given by the Department of Defense (through its ASSURE program), the National Science Foundation (DMS-0453602), and National Security Agency (MSPF-06IC-022). Substantial financial and moral support was also provided by Don Straney, Dean of the College of Science at California State Polytechnic University, Pomona. Additional financial and moral support was provided by the Department of Mathematics at Loyola Marymount University and the Department of Mathematics & Statistics at California State Polytechnic University, Pomona. Thanks to Michelle Kradic, Laura Smith, and the rest of the AMSSI faculty for their support and advice. A special thanks to Dr. Joe Gani who willingly gave constant encouragement and advice from across the pacific. Let it be known this research would not have been possible without the Lady's Man, our advisor, Dr. Randall Swift who practically put his life on hold to ensure our success. The authors are solely responsible for the views and opinions expressed in this research; it does not necessarily reflect the ideas and/or opinions of the funding agencies and/or Loyola Marymount University or California State Polytechnic University, Pomona.

References

[1] Norman T.J. Bailey. The simple stochastic epidemic: A complete solution in terms of known functions. *Biometrika*, 50:235–240, 1963.

[2] D. J. Daley and J. Gani. *Epidemic Modelling: An Introduction*. Cambridge University Press, 1999.

[3] Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM (Society for Industrial and Applied Mathematics) Review*, 42(4):599–653, 2000.

[4] Herbert W. Hethcote. The basic epidemiology models i and ii, expressions for r_0, parameter estimation, and applications. 2005.

[5] R. J. Swift J. Gani. A simple approach to birth processes with random catastrophes. 2006.

[6] R.J. Swift J. Gani. Death and birth-death and immigration processes with catastrophes. 2006.

[7] Steven H. Strogatz. *Nonlinear Dynamics and Chaos*. Westview Press, 1994.

A Sample Code

A.1 The SI Simulation

```
% Simple Stochastic Epidemic Model with no Catastrophe
% AMSSI 2006 - Randy's Angels
% Public Domain.  Mess with this as much as you want.
%
% contact = Contact Constant
% initPop = Initial population

function simpleSI(c,initPop,trials) global contact number contact =
c; number = initPop;

numPaths = 3;               % Number of sample paths to plot with average

% First we generate the population growths and times in their own cell
fprintf('Generating %d trials\n', trials);
tic;                              % Start timer
A=cell(1,trials);                 % Cell array to avoid a sparse matrix
for i = 1:trials
    popSusc = zeros(1,initPop);   % Initialize population and time vectors
    t = zeros(1,initPop);
    popSusc(1) = initPop - 1;        % Fill in initial population and time
    t(1) = 0;
    j = 1;                        % j = Step counter for current simulation

    % This while loop will run will there is still individuals left
    % This time, maxStep is not critical, because the susceptible
    % population will go to zero in exactly initPop - 1 steps anyway
    while (popSusc(j) > 0)
        % Evaluates the beta constant
        total = contact*(initPop - popSusc(j))*popSusc(j);

        % Someone becomes infected
        popSusc(j+1) = popSusc(j) - 1;

        % Generate new time step and add to current time
        t(j+1) = -log(rand)/total + t(j);%indicates time of next event
        j = j+1;%  reassigns the j value
    end

    A{i}=[popSusc(1:j) ; t(1:j)];     % Save pop and time vectors in a cell
    if (mod(i,fix(trials/10))==0)
```

```
        fprintf('.');               % Print out progress dots
    end
end

% Keep track of the length of runs and number of steps
timeLen = zeros(1,trials);          % Make a time length vector

for i=1:trials
    timeLen(i) = A{i}(2,end);%indicates duration of the epidemic
end

% Stats in case anyone is interested
avgDuration = mean(timeLen); maxTime = max(timeLen); medianTime =
median(timeLen);


% Now we quantize the data into a discrete time interval
plotTime = linspace(0,maxTime,initPop);      % Time vector to plot
plotPop  = zeros(trials,initPop);            % Pop vector to plot
plotPop(:,1) = initPop-1;                    % Initialize Pop vector

fprintf('\nQuantizing\n');
for i = 1:trials                    % i = Current trial
   j = 2;                           % j = Current quantized step
   k = 1;                           % k = Current time step of actual data
   while (k < length(A{i}) && j <= initPop)
        %If next actual time > quantized step
       if (A{i}(2,k+1) >= plotTime(j))
           plotPop(i,j) = A{i}(1,k);      % Record population as current
           j = j+1;
       else
           k = k+1;
       end
   end % End trial

   if (mod(i, fix(trials/10))==0)
       fprintf('.');             % Print out progress dot
   end
end % End quantizing

% We take non-zero population entries and find average and variance
fprintf('\nFinding Mean and Variance\n');
varPop=zeros(1,initPop);         % Initialize variance and average vectors
avgPop=zeros(1,initPop); for i=1:initPop
   w = plotPop(:,i);             % w = Use every population run
```

```
   avgPop(i)=mean(w);            % Average non-zero values
   varPop(i)=var(w);             % Variance of non-zero values
   if (mod(i,fix(initPop/10))==0)
       fprintf('.');             % Print out progress dot
   end
end

% Plot the graphs and label 'em proper
fprintf('\nPlotting.\n');
idxPaths  = 1:ceil(trials/numPaths):trials;    % Indices of paths to show
plotPaths = plotPop(idxPaths,:);               % Sample paths to plot
maxPop    = max([max(plotPaths) avgPop]);      % Max population for y axis
maxTime   = max(plotTime); figure('Position',[100 100 500 500]);

subplot(2,1,1); hold off
plot(plotTime,avgPop,'k','LineWidth',2);       % Plot the average
hold on
[t,Y]=ode45(@diffpop,plotTime,initPop-1);        % Plot deterministic
plot(t,Y,'--r','LineWidth',2);
stairs(plotTime,plotPaths',':');               % Plot the sample paths
axis([0 maxTime 0 maxPop]);                    % Set axes and labels
set(gca,'FontName','Courier New'); xlabel('Time','FontName','Courier
New'); ylabel('Population','FontName','Courier New');
legend('Average Stochastic','Deterministic','Sample Paths',...
      'Location','Northeast');
title(sprintf('SI Epidemic Simulation\n %d Trials with Population Size %d,
       and Beta %g\n Average Epidemic Duration %g, \n Maximum Epidemic
       Duration %g, Median Epidemic Duration %g' ,...
       trials,initPop,contact,avgDuration,maxTime,medianTime));

subplot(2,1,2);
plot(plotTime,varPop,'k','LineWidth',2);   % Plot variance
axis([0 maxTime 0 max(varPop)]); set(gca,'FontName','Courier New');
xlabel('Time','FontName','Courier New');
ylabel('Variance','FontName','Courier New');

hold off
fprintf('Took %.2f secs\n',toc);           % Log time

function dxdt = diffpop(t,x) global contact number dxdt =
-contact*x*number+contact*x^2;
```

A.2 The SIQ Simulation

```
% Simple Stochastic Epidemic Model with Catastrophe
```

```
% AMSSI 2006 - Randy's Angels
% Public Domain.  Mess with this as much as you want.
%
% beta = Contact Constant
% delta = Preventative Quarentine Constant
% initPop = Total Initial population

function SIQ(b,d,initPop,trials) global beta number beta = b; delta
= d; number = initPop; numCats = 0; quarPop = zeros(1,1);

numSteps = 40;             % Number of graph points to quantize to
numPaths = 3;              % Number of sample paths to plot with average

% First we generate the population growths and times in their own cell
fprintf('Generating %d trials\n', trials);
tic;                               % Start timer
A=cell(1,trials);                  % Cell array to avoid a sparse matrix
for i = 1:trials
    popSusc = zeros(1,initPop);    % Initialize population and time vectors
    t = zeros(1,initPop);
    popSusc(1) = initPop - 1;      % Fill in initial population and time
    t(1) = 0;
    j = 1;                         % j = Step counter for current simulation

    % This while loop will run will there is still individuals left
    % This time, maxStep is not critical, because the susceptible
    % population will go to zero in exactly initPop - 1 steps anyway
    while (popSusc(j) > 0)

        % Evaluates the beta constant
        total = beta*(initPop - popSusc(j))*popSusc(j) + delta;

        % Generate new time step and add to current time
        t(j+1) = -log(rand)/total + t(j);  % indicates time of next event


        if rand < (beta*(initPop - popSusc(j))*popSusc(j))/total
            % Someone becomes infective
            popSusc(j+1) = popSusc(j) - 1;
        else
            % A quarantine occurs
            popSusc(j+1) = 0;

            % counts the number of quarantines
            numCats = numCats + 1;
```

```
            quarPop(numCats) = popSusc(j);
            % used to keep track of the number saved
        end

        j = j+1;%  reassigns the j value
    end

    A{i}=[popSusc(1:j) ; t(1:j)];   % Save pop and time vectors in a cell
    if (mod(i,fix(trials/10))==0)
        fprintf('.');               % Print out progress dot
    end
end

% Keep track of the length of runs and number of steps
timeLen = zeros(1,trials);          % Make a time length vector

for i=1:trials
    timeLen(i) = A{i}(2,end);%indicates duration of the epidemic
end

% stats someone might be interested in
avgDuration = mean(timeLen); avgSaved = mean(quarPop); totalAvgSaved
= sum(quarPop)/trials; maxTime=max(timeLen); medianTime =
median(timeLen);

% Now we quantize the data into a discrete time interval
plotTime = linspace(0,maxTime,initPop);      % Time vector to plot
plotPop  = zeros(trials,initPop);            % Pop vector to plot
plotPop(:,1) = initPop-1;                    % Initialize Pop vector

fprintf('\nQuantizing\n');
for i = 1:trials                  % i = Current trial
   j = 2;                         % j = Current quantized step
   k = 1;                         % k = Current time step of actual data
   while (k < length(A{i}) && j <= initPop)
       %If next actual time > quantized step
       if (A{i}(2,k+1) >= plotTime(j))
           plotPop(i,j) = A{i}(1,k);      % Record population as current
           j = j+1;
       else
           k = k+1;
       end
   end % End trial

   if (mod(i, fix(trials/10))==0)
```

```
        fprintf('.');             % Print out progress dot
    end
end % End quantizing

% We take non-zero population entries and find average and variance
fprintf('\nFinding Mean and Variance');
varPop=zeros(1,initPop);          % Initialize variance and average vectors
avgPop=zeros(1,initPop); for i=1:initPop
    w = plotPop(:,i);             % w = Use every population run
    avgPop(i)=mean(w);            % Average non-zero values
    varPop(i)=var(w);             % Variance of non-zero values
    if (mod(i,fix(initPop/10))==0)
        fprintf('.');             % Print out progress dot
    end
end

% Plot the graphs and label 'em proper
fprintf('\nPlotting.\n');
idxPaths  = 1:ceil(trials/numPaths):trials;    % Indices of paths to show
plotPaths = plotPop(idxPaths,:);               % Sample paths to plot
maxPop    = max([max(plotPaths) avgPop]);      % Max population for y axis
maxTime   = max(plotTime); figure('Position',[100 100 500 500]);

subplot(2,1,1); hold off
plot(plotTime,avgPop,'k','LineWidth',2);       % Plot the average
hold on
stairs(plotTime,plotPaths',':');               % Plot the sample paths
axis([0 maxTime 0 maxPop]);                    % Set axes and labels
set(gca,'FontName','Sans serif','FontSize',13);
xlabel('Time','FontName','sans serif','FontSize',13);
ylabel('Population','FontName','sans serif','FontSize',13);
legend('Average Stochastic','Sample Paths',...
      'Location','Northeast');
title(sprintf('SI Epidemic with Preventative Quarantine Simulation\n %d
       Trials with Population Size %d, Beta %g and Delta %g\n Average
       Epidemic Duration %g \n Cases Ending in Quarantine %d,  Expected
       Number Saved %g' ,...
       trials,initPop,beta,delta,avgDuration,numCats,totalAvgSaved));

subplot(2,1,2);
 plot(plotTime,varPop,'k','LineWidth',2);   % Plot variance
 axis([0 maxTime 0 max(varPop)]);
 set(gca,'FontName','sans serif','FontSize',13);
 xlabel('Time','FontName','sans serif','FontSize',13);
 ylabel('Variance','FontName','sans serif','FontSize',13);
```

```
hold off
fprintf('Took %.2f secs\n',toc);           % Log time
```