MODELING THE SPREAD OF DISEASE THROUGH A POPULATION

Megan Watson

Plan II Honors Program
Dean's Scholars Honors Program
University of Texas at Austin
Spring 2009

_____

Lauren Ancel Meyers, Ph.D.
College of Natural Sciences, Section of Integrative Biology, University of Texas at Austin
Principal Supervisor

_____

Sahotra Sarkar, Ph.D.
College of Liberal Arts, Department of Philosophy, University of Texas at Austin
Secondary Supervisor

ABSTRACT

Author: Megan Watson

Title: Modeling the Spread of Disease Through a Population

Principal Supervisor: Lauren Ancel Meyers, Ph.D.

As disease spreads through a population, scientists want to know who gets sick, how best to prevent a large outbreak, and if an epidemic may occur. For many years, mathematicians used models to approximate answers to these questions; however, these older models used simplifying assumptions about the host population that drastically reduced the accuracy of the model's predictions. Within the last ten years, however, graph theory was introduced to computational epidemiology so that we now have more realistic models of the contact patterns that facilitate the spread of disease. This thesis used these models in conjunction with probability generating functions to explore the effects of individuals changing contacts during the course of an epidemic on two different degree distributions, the Poisson and the power-law. We found that at all rates of swapping contacts, the total epidemic size for the Poisson distribution is larger than that of the power-law, but that the time to total epidemic size is lower for power-law than for Poisson.

TABLE OF CONTENTS

INTRODUCTION

Through globalization, social networks have inter-connected to the point that interactions between humans from around the world occur daily. Consequently, diseases will now affect larger portions of the human population so that the need for understanding the rapid spread of disease has grown. Knowing how illnesses (such as HIV/AIDS, Sexually Transmitted Infections, the flu, and even new diseases that jump from animals to humans) spread is increasingly vital to our own protection. How can we predict whether a disease will cause an epidemic, how many people it will infect, which people it will infect, and whether or not it is dangerous to society as a whole? Also, how can we determine which techniques to use in fighting an epidemic once it begins? One way to answer all of these questions is through mathematical modeling. Models help us gain a better understanding of the world around us, and as they become more sophisticated, they prove increasingly useful for answering questions such as those raised above.

In the past, mathematicians, like Bernoulli, have used basic differential equations to model the spread of such diseases as smallpox (Meyers 2007). During his lifetime, inoculation against smallpox was not commonly practiced, as there was a risk of death associated with the procedure. However, by using equations to model the benefits of smallpox variolation against the associated risk of death, Bernoulli concluded that inoculating all newborns in the country was still more advantageous to society as a whole than the individual risk of death associated with the inoculation itself. In part because of his work, the government decided to proceed with widespread variolation of all infants. Eventually, this and the mandatory vaccine, which was later discovered, led to the eradication of smallpox from England (Meyers 2007).

This example of Bernoulli's success demonstrates two aspects of models that are invaluable to science and to society—that models are ethical and that they are feasible. Since models don't require manipulation of actual people, experiments that would otherwise be considered "unfeasible or unethical" can be executed through simulations (Meyers 2007). For example, it would be wholly unethical to randomly assign real humans to become infected with the flu, just for the sake of science, or when quantifying the effects of large-scale quarantine for long periods of time, simulations allow us to test these otherwise impractical methods. In other words, models give us vital information for protecting the population—information that we would otherwise have to do without.

Compartmental or Mass-Action SIR Models

In the 1920's, compartmental models were developed by Reed and Frost, in which the host population is divided into subgroups that represent the disease status of its members. For example, the SIR compartmental model has three groups: susceptible (S), infected (I), and recovered (R). The susceptible group contains those who remain susceptible to the infection; the infected group consists of those who not only have the disease but are also in the contagious period of the disease; and the recovered group serves as a catchall state as it holds those who have recovered and retained immunity, who have been vaccinated, or who have died. This model assumes that all individuals are mixing in one large group and that all individuals have an equal likelihood of bumping into any other individual, no matter the individual's state. Now, infected individuals will spread the disease to other individuals in the population until they eventually move to the recovered class at an average rate of μ per unit time. So, $\frac{1}{\mu}$ actually

represents the average time interval during which an infected individual remains contagious (Meyers 2007).

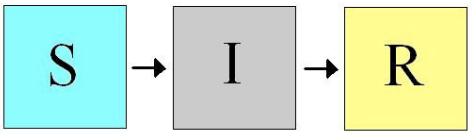


Figure 1.1 In the compartmental model, individuals move from the susceptible to infected to recovered states.

The model also assumes that members of the infected class will have infectious contacts with random individuals of all states at an average rate of β per unit time (Meyers 2007). However, disease transmission occurs as a result of this contact if and only if the infected individual comes into contact with a susceptible individual (Meyers 2007). This mass-action assumption states that "the number of new cases of disease in a time interval is proportional to the product of numbers infected and susceptible hosts in the previous time interval" (Newman 2002).

To calculate important statistics about the simulation, coupled nonlinear differential equations describe the flow of individuals from susceptible to infectious to recovered states (in the limit of large population, N) (Newman 2002). The equations $\frac{dS}{dt}$, $\frac{dI}{dt}$, and $\frac{dR}{dt}$ describe the changes in susceptible, infectious, and recovered subpopulations, respectively, and are found in Table 1.1.

$\frac{dS}{dt} = -\beta IS$
$\frac{dI}{dt} = \beta IS - \mu I$
$\frac{dR}{dt} = \mu I$

Table 1.1 The three equations that describe the compartmental model.

The equations S(t), I(t), and R(t) represent the proportions of susceptible, infected, and recovered subpopulations, respectively. Also, $1 = S(t) + I(t) + R(t)$ so that this model represents the entire population.

In order to analyze effectively the extent to which a disease will spread through a population, we use the basic reproductive rate, R_0, as defined in table 1.2. This value represents the "number of secondary infections produced by a single infected host in an entirely susceptible population" (Meyers 2007). If $R_0 > 1$, then each initially "infected host will transmit disease to at least one other host during its infectious period, and the model predicts that disease will spread through the population" (Meyers 2007). If $R_0 \leq 1$, then each initially infected host will, on average, transmit disease to less than one individual so that the model predicts that the disease will not cause a full-scale epidemic. In other words, $R_0 = 1$ is a critical value that determines whether the disease will cause an epidemic.

$$R_0 = \frac{\beta S}{\mu}$$

Table 1.2 R_0 is the critical value for determining whether an epidemic occurs in the compartmental model.

Although compartmental models were helpful in determining several characteristics of the disease spread, they also made several simplifying assumptions that made the model unrealistic and its predictions for the outbreak too high. For example, these models assume that all individuals have an equal likelihood of interacting with all other individuals and that each individual has the same number of contacts. So we can't determine anything about an individual or about the typical traits of those who become infected or those who remain susceptible. But

this information is important because knowing which members of the population could potentially infect a large number of individuals allows public health officials to know where vaccination or quarantine would best serve.

Contact Networks

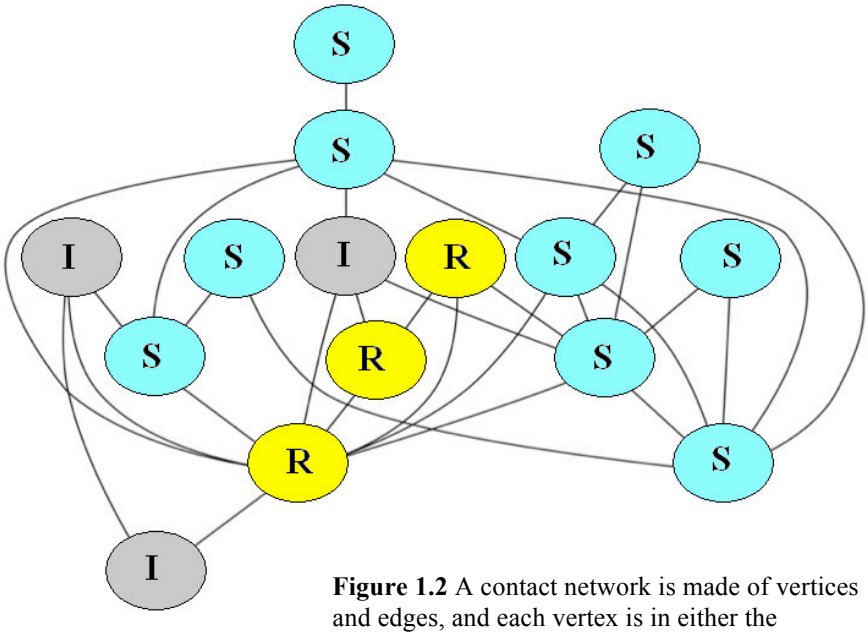


Figure 1.2 A contact network is made of vertices and edges, and each vertex is in either the susceptible, infected, or recovered state.

In order to address some of these concerns, the more recently developed "contact network" model changes the underlying structure of the population so that it can capture contact patterns among the hosts. A picture of these networks can be found in figure 1.2. This model was invented by Grassberger in 1986, but later improved upon by M.E.J. Newman in the early 2000's. Derived from graph theory, this more realistic model represents the population structure by a random graph with specified constraints on the number of vertices and edges. Each vertex represents an individual in the host population, and contacts between two individuals are

represented by an edge that connects the two. "An individual's connections are the set of people with whom the individual may have contact during the time he or she is infective—people that the individual lives with, works with, sits next to on the bus, and so forth" (Newman 2002). These connections do not guarantee the spread of disease from one individual to the next, but instead represent relationships between individuals that could potentially facilitate the spread of disease in the event of infection (Newman 2002). Let's compare the kind of edges we would need for modeling the flu versus HIV. When modeling the spread of influenza, we would need to include all contacts between individuals where the flu might spread through air, sharing food, or mutual contact surfaces. However, when modeling HIV, we would only include sexual contacts, the sharing of needles, blood transfusions, and various other contacts that allow the exchange of bodily fluids.

The probability of transmitting the disease from an infected to a susceptible individual along one of these edges or contacts is *T*, the average transmissibility (where $0 \le T \le 1$). Although the actual probability of transmission may vary among individuals, through some short calculations that can be found in Newman 2002, we can determine that the network as a whole will act as though each individual transmits disease with a probability, T, equal to the average of all probabilities of transmission for all individuals in the network.

In order for disease to begin spreading through a network, the disease must be introduced into the population, either through infecting a proportion of the population or through infecting one individual. As time moves forward, the disease will spread away from those initially infected, and two things may occur simultaneously at each time step. First, each infected individual will spread disease to each of its contacts with a probability *T*. Secondly, each infectious individual will recover at a rate, μ, at which point the individual will then no longer

infect any of its contacts. After the disease has run its course, we can determine how the disease affected the network by calculating various quantities that help us better understand the outbreak.

In order to understand an outbreak, we need to understand the shape of the network. For this, we must know how connected the network individuals are. For non-directed graphs, the number of edges emanating from a vertex is called the degree of a vertex, and the "distribution of the number of [edges] within a population is called the degree distribution" (Meyers 2007). The degree distribution tells us the shape of the overall network structure, and in turn, the paths that the disease could potentially follow as it moves through the network. In order to signify the degree distribution, we need all of the probabilities, p_k, that a randomly chosen vertex will be found to have degree *k*. These values can be determined from an actual population and entered individually into the model, or they can be determined by probabilistic distributions, such as the Poisson or power-law distributions. Probabilistic distributions simply use an underlying equation to generate a list of probabilities, as shown in Table 1.3, which was adapted from Volz and Meyers 2007. For pictures of these two types of networks, see figures 1.3 and 1.4.

Poisson	$g(x) = \exp\{z(x-1)\}$ where *z* is the average degree
power-law	$g(x) = \frac{\sum_{k} k^{-\alpha} x^{k}}{\sum_{i=1}^{\kappa} i^{-\alpha}}$ where κ = the maximum degree, *k* = the degree value, and α = average degree.

Table 1.3 The Poisson and power-law distributions are made by the respective underlying equations.

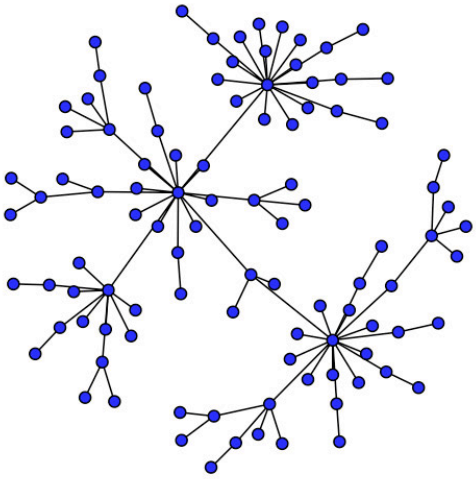


Figure 1.3 A power-law network, which tends to have hubs with high degrees and spokes with low degrees. This figure was taken from Lauren Ancel Meyers, Bio 337 class at the University of Texas at Austin.

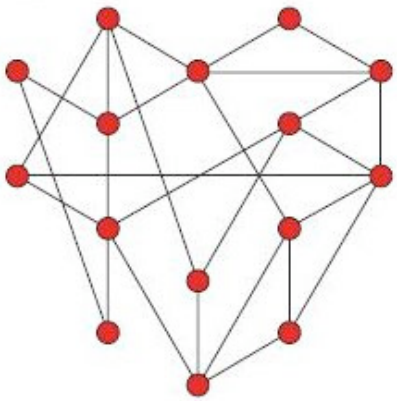


Figure 1.4 A Poisson network, where all the nodes tend to have similar, low degrees. This figure was taken from the website found at the following link: http://cns.bu.edu/%7Eheekko/CN710/Linked/barabasi2004_box2.JPG

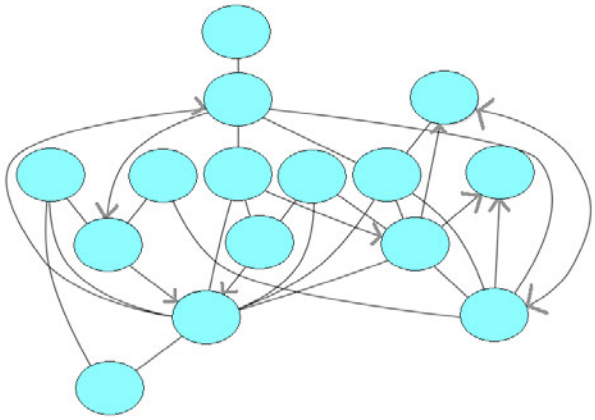


Figure 1.5 A semi-directed network where disease can only spread one way along an edge. These are useful for modeling the spread of disease from a patient to his or her doctor.

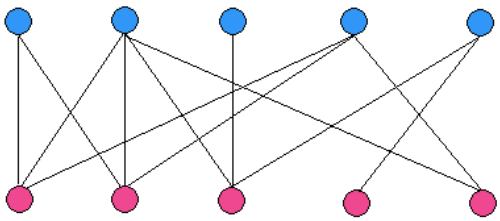


Figure 1.6 A bipartite network where the pink nodes only interact with blue nodes and, blue nodes only interact with pink nodes. These are useful for modeling the spread of disease in healthcare facilities.

Contact networks can be adapted to fit various problems. For example, creating bipartite networks (networks with two categories of vertices, and edges that only run between the two types, never among them) allows simplistic models of both sexually transmitted infections and the spread of disease in healthcare facilities (as seen in figure 1.6). In the first case, the two categories of vertices would represent males and females, and the edges would represent sexual encounters between the two groups; in the second case, the vertex types would represent nurses and patients, and the edges would represent routine interactions. Of course, a sexual bipartite network would ignore homosexuality, but could still represent a simple model. Other types of models require graphs with *directed* connections—edges that only allow infection to spread in one direction (as seen in figure 1.5). These one-way connections would accurately represent the relationship between patients and their doctors or nurses. If the patient falls ill, he or she will visit the doctor for treatment; however, if the doctor falls ill, the patient will have no reason to visit the doctor. Thus, the disease transmission is only in one direction, and it is under the condition that the patient be the one to spread disease.

By allowing a more descriptive population model, the contact networks let us create many parameters for the disease that other models would render impossible. They let us determine which individuals are more likely to fall ill and which individuals tend to spread disease more rapidly. Since the population can be divided into subgroups, we can determine which categories may be more susceptible to infection (eg: school children or the elderly). Other quantities of the outbreak that we would like to calculate include: outbreak size, presence of epidemic, epidemic size, and percentage of individuals infected in the outbreak that have n number of contacts (Newman 2002).

Volz Neighbor-Exchange Model

Although contact networks improve the sophistication of the model, in reality, social contact networks are often not as static as the model has been thus far. People often change their contacts several times over their lifetime. For example, when considering HIV, an individual may only have sexual contact with one person at any given time, but the identity of that person may change over the years. Out of this concept, Erik Volz and Lauren Ancel Meyers created the Neighbor-Exchange (NE) model in 2007. This model assumes that each individual's number of current contacts remains the same, but that the identity of those contacts will be changing at a steady rate, ρ. So each contact becomes "temporary" and when its time comes to an end, it will be replaced by a new contact (Volz and Meyers 2007). For example, an individual A has a contact with B while another individual C has a contact with D. According to this model, the contacts or edges will then switch from (A,B) and (C,D) to (A,C) and (D,B). This example is illustrated in figure 1.7.

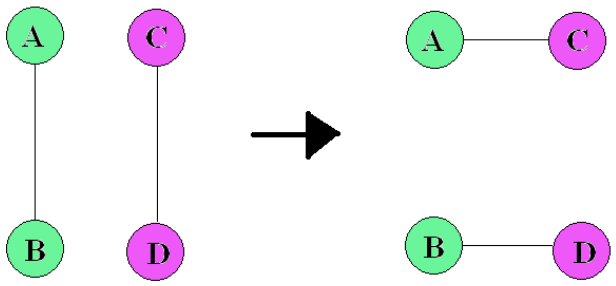


Figure 1.7 The Volz Neighbor-Exchange Model switches an individual's edges while keeping the degree constant for all nodes.

Since each contact is switching at a rate of ρ, each edge will be switching at a rate of $\frac{\rho}{2}$. Also, since a node's degree never changes during an edge swap, the degree distribution is

preserved, and we will continue to assume that an infected node spreads infection to its neighbors at a constant rate, *T*, and that an infected node recovers at a constant rate of μ. Actually as $\rho \to \infty$, the NE model approaches the mass-action model, or the SIR Compartmental model, because the rate of mixing is so high that all individuals are interacting with everyone else almost simultaneously. Actually, the probability of being connected to an infectious, susceptible, or recovered node becomes exactly proportional to the number of nodes in each state.

METHODS

M.E.J. Newman Model

In order to express the degree distribution, and other values of the network, in a functional way, Newman developed the idea of using Probability Generating Functions (PGF's) on networks. A PGF of a discrete random variable is a power series representation of the probability mass function of the random variable (Wikipedia). For example, $g(x) = p_o + p_1 x + p_2 x^2 + p_3 x^3 + ...$ generates the degree distribution for the network, where p_k represents the probability that a randomly chosen individual has degree *k*. Also, you could write it as $G_o(x) = \sum_{k=0}^{\infty} p_k x^k$. However, in general, *x* serves as a dummy variable, or place-holder, so that it is easy to find the probability associated with a particular value, *k*. PGF's have several nice qualities, including that the first derivative of the PGF, evaluated at $x = 1$, gives the average value of *k*. For example, the average degree of a vertex in the distribution would be found by $G_o'(1) = \langle k \rangle = \sum_k k p_k$. In general, PGF's are connected to the moments of the distribution. These functions prove beneficial not only for gathering information about properties of vertices but also for information about edges.

When we randomly choose an edge and follow it to a vertex, the probability is higher that we will reach a vertex of higher degree since vertices with higher degrees are attached to more edges (Newman 2002). However, when we approach a vertex from an edge, we are interested in how many possible paths we can follow away from this vertex. Since we can't go the way we already came, we are only interested in the $k-1$ other edges emanating from the vertex. So, we are really interested in the degree of the vertex minus one, also known as the *excess degree*

$= k-1$. The equation for the distribution of excess degree becomes $G_1(x) = \frac{G_o'(x)}{G_o'(1)}$. Again, taking the derivative of this function at the value of x=1 will give you the average excess degree for any vertex. In his 2002 paper, Newman goes on to derive many other PGF's and calculates many expected values, including the equation for the distribution of the number of occupied edges attached to a vertex, $G_0(x;T) = \sum_{m=0}^{\infty}\sum_{k=m}^{\infty} p_k \binom{k}{m} T^m (1-T)^{(k-m)} x^m = G_0(1+(x-1)T)$. Also, he derives the PGF that describes the distribution of occupied edges leaving a vertex, $G_1(x;t) = G_1(1+(x-1)T)$. Note that in his notation, $G_0(x;1) = G_0(x)$ and $G_0(1;T) = G_0(1)$. Using these derived PGF's recursively and in conjunction with each other, Newman can calculate the effects of disease transmission on the network. For a more in-depth derivation of this equation and more, see Newman 2002.

Volz Model

Instead of using these PGF's in their current form, Volz uses variable replacement to simplify Newman's equations and make them easier for creating dynamic modeling functions. Tables 2.1 and 2.2 are adapted from Volz 2007 and will be helpful for referencing throughout this section.

Table 2.1 Definitions

r	Force of infection. The constant rate at which infectious nodes infect a neighbor. This value was previously referred to as T.
μ	Recovery rate. The constant rate at which infected nodes recover
p_k	The probability that a node will have degree k
$g(x)$	The probability generating function for the degree distribution p_k
S	The fraction of nodes susceptible at time t.
I	The fraction of nodes infectious at time t.
R	The fraction of nodes recovered at time t.
J	$J = I + R$. The cumulative epidemic incidence at time t.

Table 2.2 Definitions Continued	
ego	The first element of the ordered pair representing an edge
alter	The second element of the ordered pair representing an edge
d_v	Denotes the degree of a node, v
$\mathcal{A}_X$	Set of arcs (*ego, alter*) such that node $ego \in X$
$\mathcal{M}_X$	Fraction of arcs in set $\mathcal{A}_X$
$\mathcal{A}_{XY}$	Set of arcs (*ego, alter*) s.t. $ego \in X$ and $alter \in Y$
$\mathcal{M}_{XY}$	Fraction of arcs in set $\mathcal{A}_{XY}$

The following definitions and derivations are taken from Volz 2007. A network is a graph, $G = \{V, E\}$ where V is the set of all vertices in the graph and E is the set of all edges in the graph. Each edge is denoted as $\{a,b\}$ where $a,b \in V$. Now for two vertices, a and b, to have a connection upon which disease could possibly spread, there must exist an edge e, s.t. $e=\{a,b\}$ and $e \in E$. Two vertices that share an edge are called neighbors or connected. The degree of a node v is denoted as d_v. See Volz 2007 for a thorough explanation of how to construct this graph practically, and how to deal with the nuances of loops (an edge which connects a node to itself) and multiple edges (two nodes that have multiple edges between them). Now in order to keep track of the flow of infection, we need to specify a direction. So we will define each connection, $\{a,b\}$ as having two arcs—(a,b) and (b,a) (Volz 2007). We will call the first element of the ordered pair the *ego* and the second element the *alter*. The *ego* represents the transmitting node while the *alter* represents the receiving node.

Now let $\mathcal{A}$ denote the set of all arcs in the network. Then $\mathcal{A}_X$ denotes the set of arcs (*ego, alter*) such that node $ego \in X$. Also, $\mathcal{A}_{XY}$ denotes the set of arcs (*ego, alter*) such that $ego \in X$ and $alter \in Y$. In order to describe these subsets numerically, we need to know their

size in relation to the whole, $\mathcal{A}$. So, let $\mathcal{M}_{XY} = \frac{\#\{\mathcal{A}_{XY}\}}{\#\{\mathcal{A}\}}$, and this value denotes the fraction of arcs in the set $\mathcal{A}_{XY}$. For example, $\mathcal{M}_{SI}$ will be the fraction of arcs with a susceptible ego and an infectious alter, and $\mathcal{M}_S$ will be the fraction of arcs with a susceptible ego and an alter of any state. Let's think about a susceptible vertex *ego* at a time *t* with degree *k*. There are *k* arcs connected to *ego*, namely $\{(ego,alter_1),(ego,alter_2),...,(ego,alter_k)\}$. And for each of these arcs, there will be a uniform probability $p_I = \frac{\mathcal{M}_{SI}}{\mathcal{M}_S}$ that $alter_i$ is infectious.

Through some calculations which can be found in Volz 2007, we know that the probability that *ego* becomes infected at time *t* is $\lambda_k(t) = rkp_I(t)$. Now if we allow $u_k(t)$ to represent the probability that a node of degree *k* is susceptible at time *t*, we realize that the probability that a node of degree one is susceptible at time *t* is $u_1(t) = \exp\{-\int_{\tau=0}^{t} rp_I(\tau)d\tau\}$. Now let $\theta = u_1(t)$, and with a bit more manipulation, we get that $u_k = \theta^k$. For a more thorough derivation, see Volz 2007.

Table 2.3 Network-based Values	
θ	The fraction of degree one nodes that remain susceptible at time t
p_I	$p_I = \frac{\mathcal{M}_{SI}}{\mathcal{M}_S}$ The probability that an arc with a susceptible ego has an infectious alter
p_S	$p_S = \frac{\mathcal{M}_{SS}}{\mathcal{M}_S}$ The probability that an arc with a susceptible ego has a susceptible alter.
S	The fraction of nodes which remain susceptible at a time t.

Now, using θ, we can easily calculate the fraction of susceptible nodes at time *t*. $S = p_1 u_1 + p_2 u_2 + p_3 u_3 ... = p_1\theta + p_2\theta^2 + p_3\theta^3 + ... = g(\theta)$. Now we can clearly see that $g(\theta)$ takes advantage of the probability generating function for the degree distribution. In this case, $g'(1)$ the total number of arcs in the network. When determining the dynamics behind θ, $\dot{\theta} = -\theta r p_I$. However, since $\dot{\theta}$ depends on p_I, we must do further calculations on p_I in order for the dynamics of θ to be completely specified. The derivation of $\dot{p}_I$ is as follows:

$\dot{p}_I = \frac{d}{dt}\frac{\mathcal{M}_{SI}}{\mathcal{M}_S} = \frac{\dot{\mathcal{M}}_{SI}}{\mathcal{M}_S} - \frac{\dot{\mathcal{M}}_S \mathcal{M}_{SI}}{\mathcal{M}_S^2}$ and in order to remove these unwanted variables, we need to transform both $\mathcal{M}_S$ and $\mathcal{M}_{SI}$. So, $\mathcal{M}_S = \sum_k \frac{p_k k\theta^k}{g'(1)} = \frac{\theta g'(\theta)}{g'(1)}$ and

$\mathcal{M}_{SI} = \mathcal{M}_S \frac{\mathcal{M}_{SI}}{\mathcal{M}_S} = \mathcal{M}_S p_I = p_I \theta \frac{g'(\theta)}{g'(1)}$.

Using these new variables, we can calculate the number of new infectious nodes per time period, *dt*. We know that $-\dot{S}$ nodes become infected per time period and that $S = g(\theta)$ so that $\dot{S} = \frac{d}{dt}S = \frac{d}{dt}g(\theta) = \dot{\theta}g'(\theta) = -rp_I\theta g'(\theta)$. By manipulating $\mathcal{M}_{SS}$, $\mathcal{M}_{SI}$, and $\mathcal{M}_S$ you can get the dynamic equations for p_I, and p_S, and they are shown in the table below. For a more thorough derivation of these dynamic variables, consult Volz 2007. Tables 2.5 and 2.6 were adapted from Volz 2007 and Volz and Meyers 2007. With these equations, all epidemiological properties of the network can be found.

Table 2.5 Dynamic Equations for Volz Model
$\dot{\theta} = -rp_I\theta$ $\dot{p}_I = rp_S p_I \theta \frac{g''(\theta)}{g'(\theta)} - rp_I(1-p_I) - p_I\mu$ $\dot{p}_S = rp_S p_I (1 - \theta \frac{g''(\theta)}{g'(\theta)})$
$\dot{S} = -rp_I\theta g'(\theta)$ $\dot{\mathcal{M}}_I = -\mu\mathcal{M}_I + rp_I(\frac{\theta^2 g''(\theta) + \theta g'(\theta)}{g'(1)})$ $\dot{I} = rp_I\theta g'(\theta) - \mu I$

Table 2.6 PGF's for Volz Model
$g(\theta) = \sum_k \mathrm{p}_k \theta^{\mathrm{k}}$
$g'(\theta) = \sum_k \mathrm{k}\mathrm{p}_k \theta^{\mathrm{k-1}}$
$g''(\theta) = \sum_k \mathrm{k(k-1)}\mathrm{p}_k \theta^{\mathrm{k-2}}$

Now θ represents the number of degree one nodes that are still susceptible at time t, and in the equation of $\dot{\theta}$, the term, $-rp_I\theta$, accounts for those nodes no longer susceptible due to disease spreading to them from their infected contact. For the equation, $\dot{p}_I$, there are three main terms. Remember that p_I represents the probability that an arc with a susceptible ego has an infectious alter. Thus the term $rp_S p_I \theta \frac{g''(\theta)}{g'(\theta)}$ accounts for the fraction of arcs gained, in one time step, from nodes who became infected but that were attached to some susceptible nodes (so that Susceptible-Susceptible edges became Infected-Susceptible edges); and the term $-rp_I(1-p_I)$ represents the arcs lost when the susceptible node turned infected (so that S-I edges became I-I edges); and the term $-p_I\mu$ accounts for those nodes who were infectious and connected to susceptibles but that recovered (so that the I-S edges became R-S edges). A similar argument applies to $\dot{p}_S$, remembering that p_S represents the probability that an arc with a susceptible ego

has a susceptible alter. The equation only has one main term, $rp_S p_I(1-\theta\frac{g''(\theta)}{g'(\theta)})$, which accounts for those arcs lost when a susceptible node becomes infectious (when S-S edges become S-I edges). For the change in number of susceptible individuals, S, we only must account for those who were susceptible that became infectious. For the change in number of infectious individuals, we must only account for those infectious individuals who recover and those susceptible individuals who become infected.

Volz-Meyers Neighbor-Exchange Additions to Volz Model

Now when we use the Volz-Meyers model but want to add mixing, we just need to change these equations to account for the mixing. However the only two equations which will change due to mixing are $\dot{p}_I$ and $\dot{p}_S$. Volz and Meyers describe the effect of mixing as changing $\dot{p}_I$ and $\dot{p}_S$ each by a term that describes the effect of the NE model. The equation $\dot{p}_I$ becomes $\dot{p}_I = rp_S p_I\theta\frac{g''(\theta)}{g'(\theta)} - rp_I(1-p_I) - p_I\mu + f_I(p_I,\mathcal{M}_I)$ where $f_I(p_I,\mathcal{M}_I)$ represents the effect of the NE model on $\dot{p}_I$. First, they examine the decrease of $\dot{p}_I$ due to the NE model:

- At a rate ρ, a particular arc (*ego, alter*) will transform to (*ego, alter'*)
- Given that *ego* is susceptible, *alter* is infected with probability p_I
- With probability $1-\mathcal{M}_I$, *alter'* is not infected

Taking all of these factors into consideration, the NE model decreases $\dot{p}_I$ by $\rho p_I(1-\mathcal{M}_I)$. Now, the increase of $\dot{p}_I$ due to NE dynamics is:

- At a rate ρ, a particular arc (*ego, alter*) will transform to (*ego, alter'*)
- Given that *ego* is susceptible, *alter* is not infected with probability $1-p_I$

- With probability $\mathcal{M}_I$, *alter'* is infected

So, the NE model increases by $\rho(1-p_I)\mathcal{M}_I$. Adding the increase and decrease together, Volz and Meyers show that $f_I(p_I,\mathcal{M}_I)=\rho(\mathcal{M}_I-p_I)$.

By doing the same thing for $\dot{p}_S$, we get that $f_S(p_S,\mathcal{M}_S)=\rho(\mathcal{M}_S-p_S)$. Thus, the entire equation set can be found in Table 2.7.

Table 2.7 Dynamic Equations for the NE model

$\dot{\theta}=-rp_I\theta$
$\dot{p}_I=rp_Sp_I\theta\frac{g''(\theta)}{g'(\theta)}-rp_I(1-p_I)-p_I\mu+\rho(\mathcal{M}_I-p_I)$
$\dot{p}_S=rp_Sp_I(1-\theta\frac{g''(\theta)}{g'(\theta)})+\rho(\frac{\theta g'(\theta)}{g'(1)}-p_S)$
$\dot{S}=-rp_I\theta g'(\theta)$
$\dot{\mathcal{M}}_I=-\mu\mathcal{M}_I+rp_I(\frac{\theta^2g''(\theta)+\theta g'(\theta)}{g'(1)})$
$\dot{I}=rp_I\theta g'(\theta)-\mu I$

Distributions and Parameters

In order to run the model, we need initial conditions for $t=0$. These initial conditions are found in Volz 2007. If we select ε random nodes to initially infect, the fraction of arcs with infectious egos will also be $\mathcal{M}_I=\varepsilon$. Since ε is an extremely tiny value, it is unlikely that two initially infected nodes will be connected so that $\mathcal{M}_{SI}\approx\mathcal{M}_I=\varepsilon$. Since θ represents the number of degree 1 nodes that are susceptible, $\theta=1-\varepsilon$. Also, $\mathcal{M}_S=1-\mathcal{M}_{SI}=1-\varepsilon$ and $\mathcal{M}_{SS}=\mathcal{M}_S-\mathcal{M}_{SI}=1-2\varepsilon$. Also, since $S=g_0(\theta)$, we know that $S=g_0(1-\varepsilon)$ and we know that $I=1-S=1-g_0(1-\varepsilon)$. Thus the initial conditions can be summarized in table 2.8.

Table 2.8 Initial Conditions
$\theta(t=0)=1-\varepsilon$
$p_I(t=0)=\frac{\mathcal{M}_{SI}}{\mathcal{M}_S}=\frac{\varepsilon}{1-\varepsilon}$
$p_S(t=0)=\frac{\mathcal{M}_{SS}}{\mathcal{M}_S}=\frac{1-2\varepsilon}{1-\varepsilon}$
$S(t=0)=g(1-\varepsilon)$
$\mathcal{M}_I(t=0)=\varepsilon$
$I(t=0)=1-g(1-\varepsilon)$

Also, we need to give values to our parameters. The population size, $N=10{,}000$ was arbitrarily chosen by Volz and Meyers, and $\varepsilon=.0001$ was the corresponding proportion of the population chosen (the population is 10^4 and ε is 10^{-4}). The force of infection was chosen to be $r=.2$ because it seems to be a highly infectious value, yet low enough that there would be room for the epidemic size to grow as expected. And the value for μ was chosen as .2 in order to match *r* so that the rate of infection would be equal to the rate of recovery. The total time for running the model was $t=200$ just so that the epidemic would have plenty of time to run its course. Correspondingly, the time steps were chosen as $\tau=.1$ in order to get clearer, more accurate resolution for the graphs. A summary of the parameter values can be found in table 2.9.

Table 2.9 Parameter Values

Poisson		power-Law	
λ	2.5	α	2.218
		max degree	75
N	10,000	N	10,000
r	0.2	r	0.2
μ	0.2	μ	0.2
ε	0.0001	ε	0.0001
total time	200	total time	200
step	0.1	step	0.1

To explore the effects of mixing on networks, comparing different degree distributions adds to our understanding. Using Poisson and power-law distributions for comparison seems the most reasonable as each distribution has desirable characteristics for drawing conclusions about the effects of mixing. Both distributions are well studied. The power-law distribution is heterogeneous as the majority of individuals have low degrees while a few individuals have high degrees. This makes for a hub-and-spoke network, and actually the power-law distribution is often used to estimate biological and sexual networks (Bansal, Grenfell, and Meyers 2007). A picture of a power-law network can be found in figure 1.3. The Poisson distribution, on the other hand, is homogenous so that all individuals in the network have almost the same degree distribution, and it makes a nice contrast to the power-law distribution (Bansal, Grenfell, and Meyers 2007). For a picture of a Poisson network, see figure 1.4.

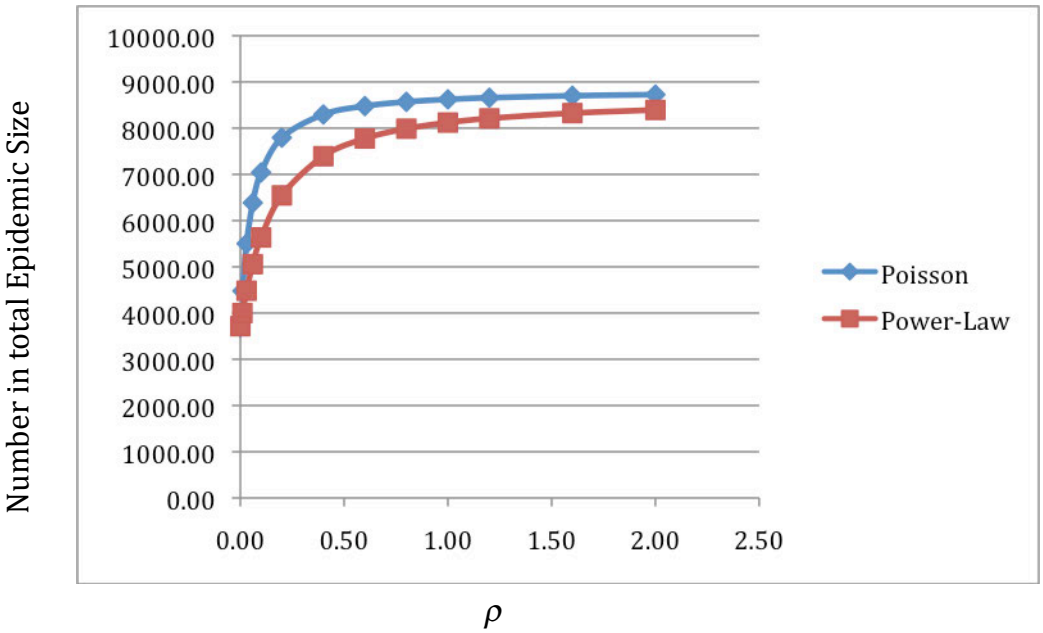
In order to make a fair comparison between the two networks, we must fix all the parameters while still making the total epidemic size (when $\rho = 0$) the same for both networks. This way we can guarantee that any differences in the outbreaks arises from the degree distributions and not from a different factor. Beginning with the Poisson distribution, an average

degree of 2.5 was used, and with the parameters as stated above, the total epidemic size was 3716.18 people. To get an epidemic size that closely matches that of the Poisson distribution, we set the maximum degree equal to 75 and then tinkered with alpha. Finally, we set $\alpha = 2.218$, and with this value were able to get a total epidemic size of 3716.21 people. Thus when $\rho = 0$, the two distributions start at an equivalent baseline.

RESULTS

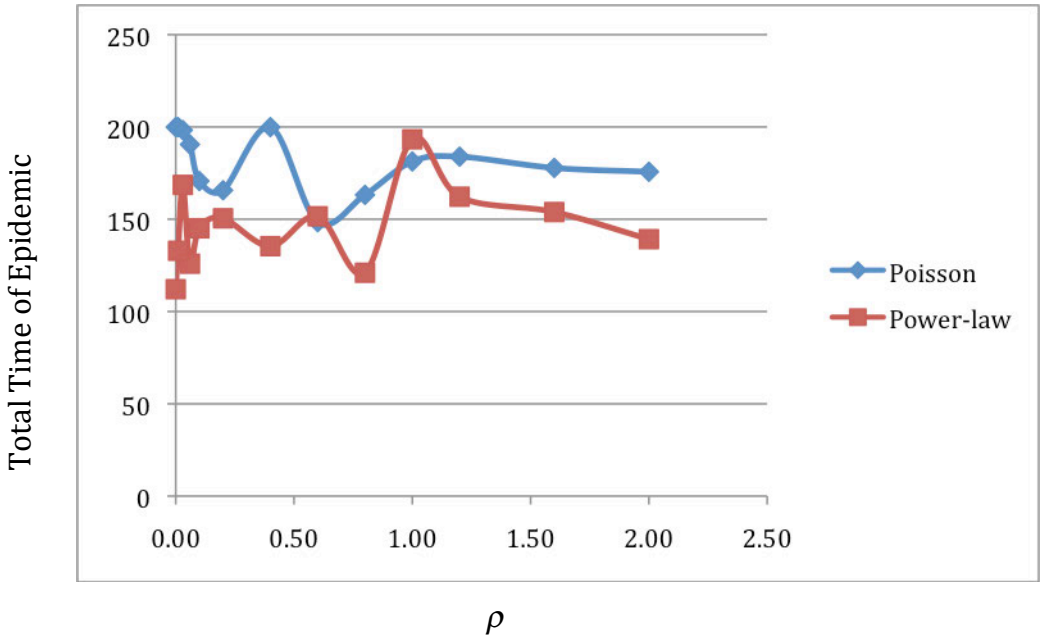
The equations were integrated using the programming language, Python, which has a function derv that will integrate multiple differential equations over a specified interval. Comparing the two distributions, Poisson and power-law, can shed light on how the mixing parameter will affect a model. First we will look at the total epidemic size ($J = I + R$) plotted against ρ in figure 3.1. For both distributions, the total epidemic size increases as $\rho \to \infty$, as was exptected. Actually, the total epidemic size for both is approaching the epidemic size produced by the mass-action model because as $\rho \to \infty$, the NE model approaches the compartmental model. Now, at each value of ρ, the Poisson distribution has a higher epidemic size than the power-law distribution; however, as ρ increases, the difference becomes smaller. The epidemic size for $\rho = 1000\mu$ was left off for reasons of scale, but the corresponding peak epidemic size is 8818.07 for Poisson and 8663.2 for power-Law so that the total epidemic sizes continue to converge.

Total Epidemic Size versus ρ (figure 3.1)



In order to better understand the speed of the spread of disease on both networks, we need to examine the total time of the epidemic as seen in figure 3.2. At low and high values of ρ, the power-law distribution has a shorter time of total epidemic than does the Poisson distribution. For medium values of ρ such as $\rho = 5\mu$ and $\rho = 3\mu$, the time fluctuates more, so that sometimes Poisson has the shortest time while other times power-law has the shortest time. However, as ρ increases, the time decreases overall for both distributions. The value of $\rho = 1000\mu$ was omitted from this graph as well for reasons of scale, but the time to the total epidemic was 199.9 for Poisson and 180 for power-Law. So although the Poisson distribution generally requires more time to reach the end of the epidemic, the total epidemic size is actually larger than that of the power-law distribution.

Total time of Epidemic versus ρ (figure 3.2)



In Table 3.1, the epidemic sizes are shown for both distributions, as well as the value of ρ and the difference between the two distributions. From the chart, we can see that although the distributions start off at the same total epidemic size, they quickly diverge, with Poisson becoming larger. However, as ρ continues to increase, the divergence decreases and the difference seems to approach zero again.

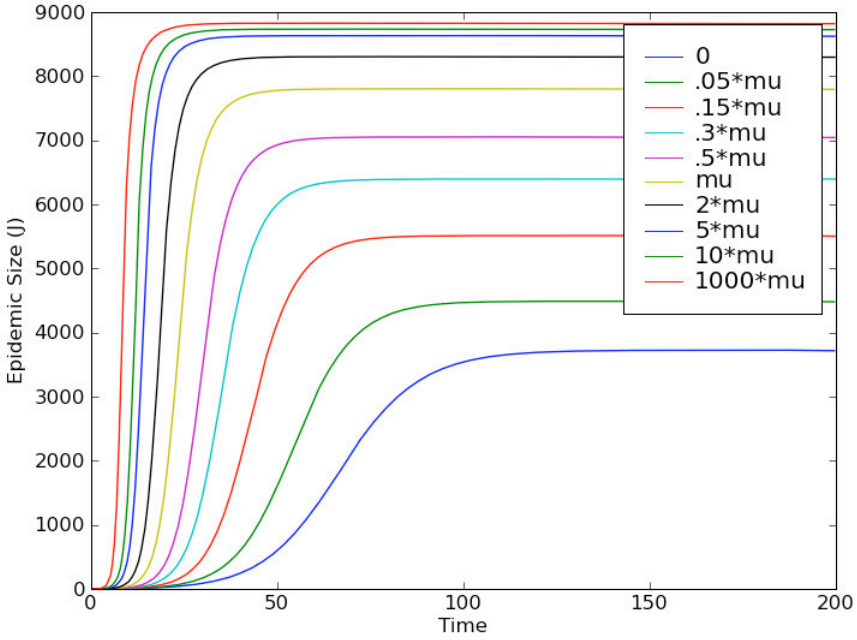
Epidemic Size (N=10,000) (**Table 3.1**)

Ro(mu)	ρ	Poisson Size	power-law Size	Size Difference (Poisson-power)
0	0.00	3716.18	3716.21	-0.03
$.05\mu$	0.01	4479.35	3999.23	480.13
$.1\mu$	0.03	5503.06	4482.46	1020.61
$.3\mu$	0.06	6384.91	5057.43	1327.48
$.5\mu$	0.10	7042.44	5635.21	1407.23
μ	0.20	7794.36	6545.68	1248.68
2μ	0.40	8295.29	7392.72	902.57
3μ	0.60	8477.68	7775.96	701.72
4μ	0.80	8568.84	7988.07	580.77
5μ	1.00	8622.62	8120.83	501.80
6μ	1.20	8657.80	8211.07	446.73
8μ	1.60	8700.71	8325.14	375.57
10μ	2.00	8725.75	8393.89	331.86
1000μ	200.00	8818.07	8663.2	154.87

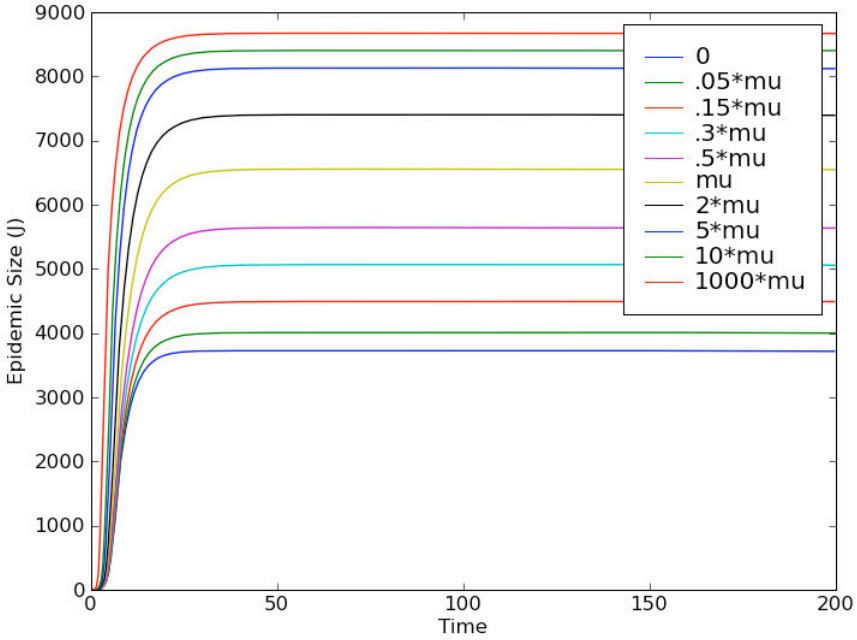
To better understand the total epidemic size and the time to the total epidemic, we can look at each distribution individually over all values of ρ. The Poisson distribution can be seen in figure 3.3 and the power-law distribution can be seen in figure 3.4. In figure 3.3, we can see that the total epidemic size for the Poisson distribution systematically grows as $\rho \to \infty$, but that it grows faster for smaller values of ρ and much slower for larger values of ρ. Also, we can see that the length of the duration of the epidemic gets shorter and shorter because as ρ gets larger, the total size occurs more quickly. For the power-law distribution in figure 3.4, we see that the

total epidemic size grows similarly to the way the Poisson distribution grows, but that we don't see the drastic change in time that we see in the graph of the Poisson distribution. In figure 3.4, the time for the power-law distribution seems to remain quite stable, despite the value of ρ.

Poisson: Total Epidemic Size versus Time (figure 3.3)

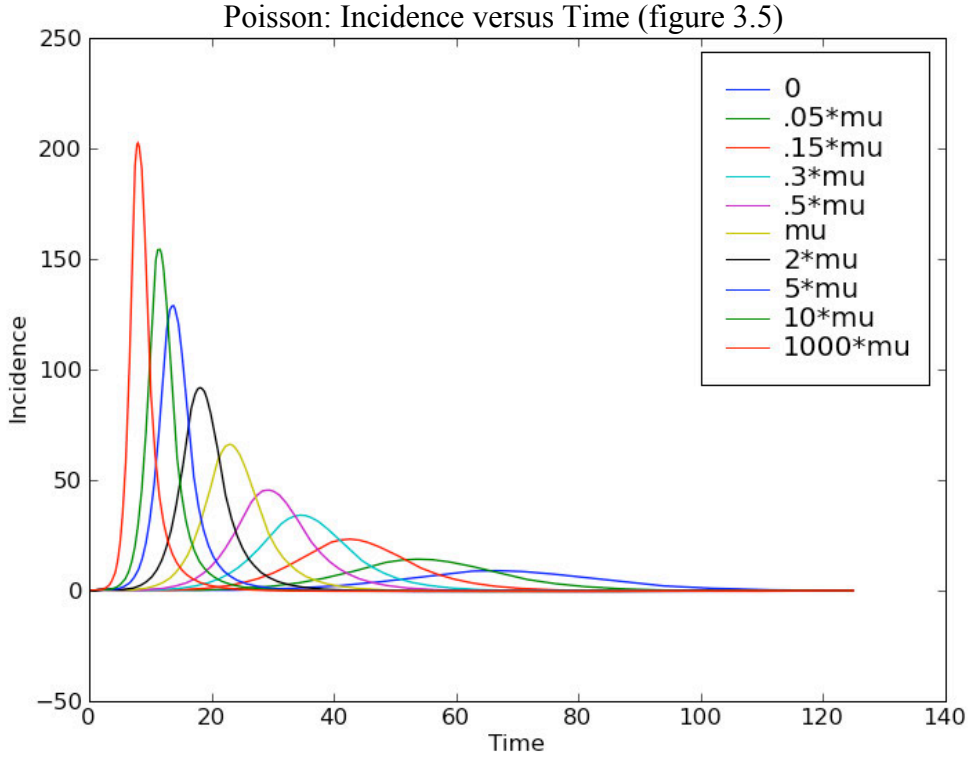


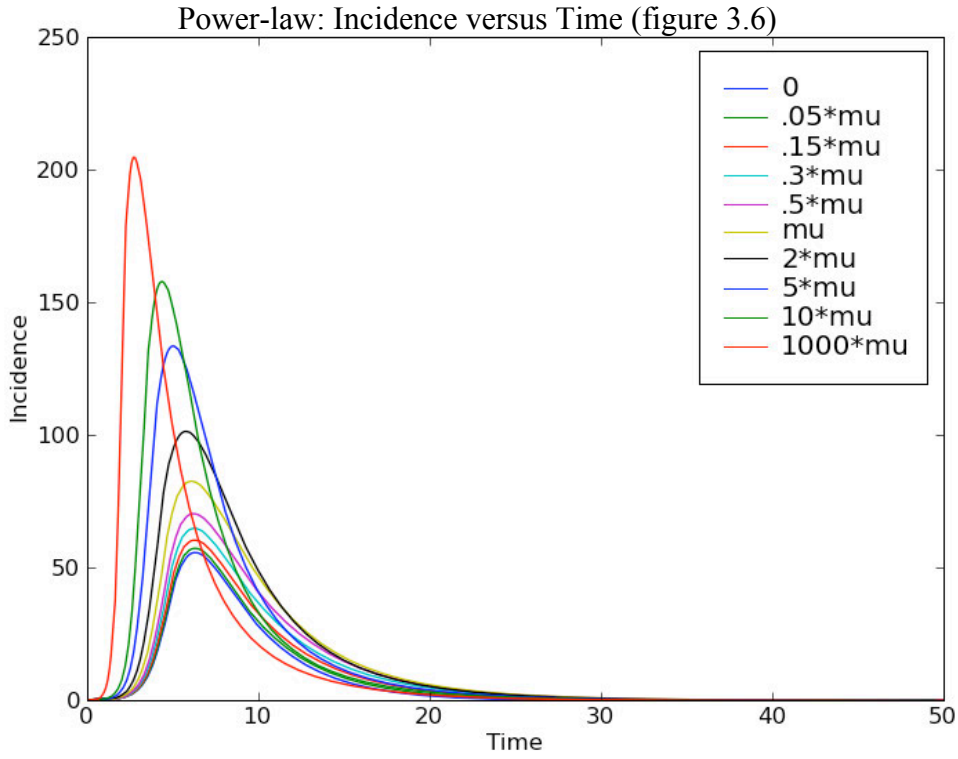
Power-law: Total Epidemic Size versus Time (figure 3.4)



Next, we can look at Incidence, which is defined as the number of new cases of infected individuals per unit time. Figures 3.5 and 3.6 were scaled differently on the time axis because for low values of ρ, the Poisson total epidemic time lasts longer than does the power-law distribution. So, figure 3.5 for Poisson reaches 140 time units while figure 3.6 for power-law only reaches 50 time units. For both figures 3.5 and 3.6, we see an increase in incidence as more individuals are infected per time period with an increase in ρ. In figure 3.5 for the Poisson distribution, again we see a difference in time as well. As ρ increases, the time to reach peak incidence decreases. But in figure 3.6 for the power-law distribution, as ρ increases, time to peak incidence remains basically stable despite the fact that the incidence itself is growing.

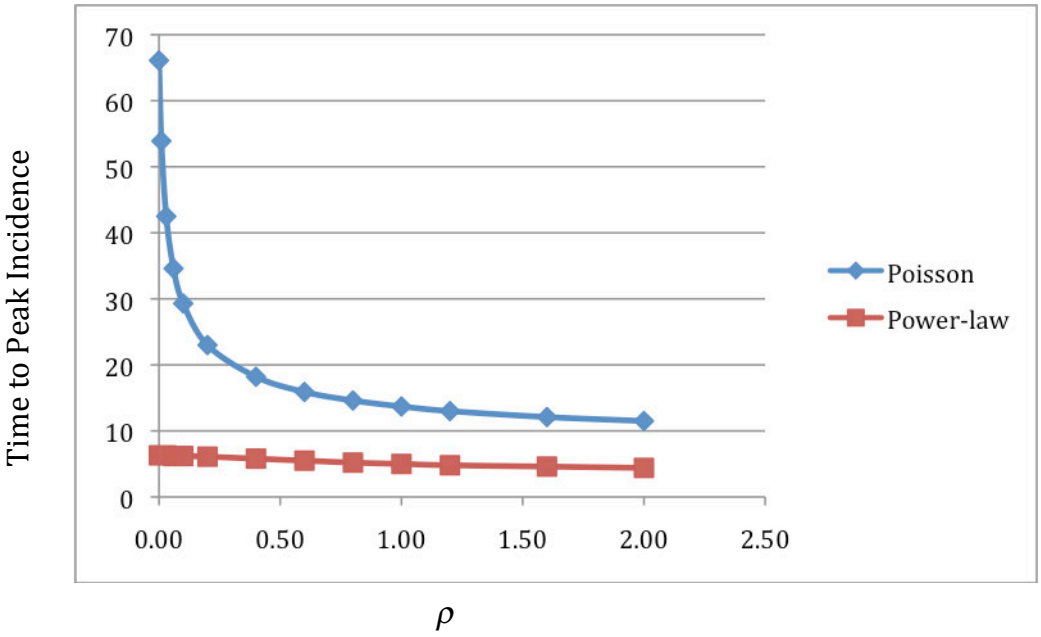
Poisson: Incidence versus Time (figure 3.5)





In figure 3.7, we can see this time difference more clearly as time to peak incidence is plotted against ρ. For the power-law distribution, at each value of ρ, the time does not change drastically and is always somewhere between 0 and 10 time units. (Although, the time does decrease a bit as ρ increases). But for the Poisson distribution, the time to peak incidence decreases rapidly for low values of ρ and decreases less rapidly for higher values of ρ, and the time drops from about 65 time units to nearly 10 time units. It appears that the time to peak incidence for Poisson is approaching that of the power-law.

Time to Peak Incidence versus ρ (figure 3.7)



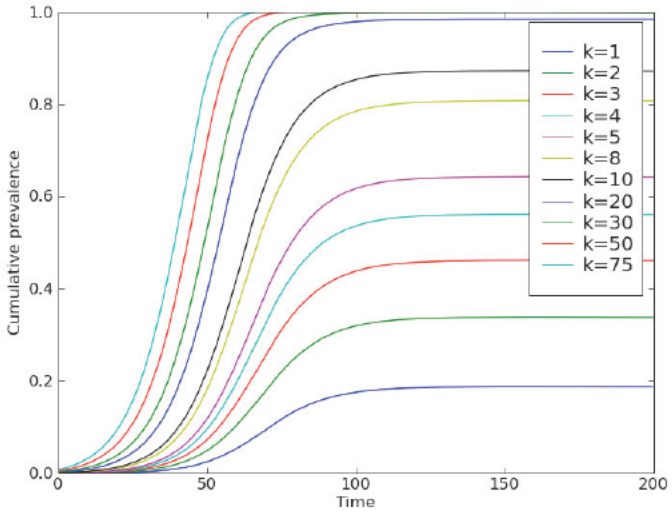
In order to better understand how the degree distribution affects the outbreak, we need to look at each individual degree to see how many become infected as time progresses. In figures 3.8 and 3.9, we see the cumulative prevalence plotted against time for the Poisson and power-law distributions, respectively. In each figure, $\rho = 0$. Each graph has many curves, and each curve represents the proportion of nodes of a certain degree that became infected over the course of the outbreak. These two graphs show what we basically expected, where the nodes of low degree become infected not only more slowly but also in less quantity than nodes of higher degree. However, again we see that the power-law distribution has nodes infected much faster than does the Poisson. Notice that the scale for the power-law time axis only extends to 40 time units while the Poisson extends to 200.

Now looking at figures 3.10 and 3.11, we see something a bit more unusual as the numbers of nodes of a given degree that became infected over the course of the outbreak are plotted instead of the proportion of infected nodes of a given degree.

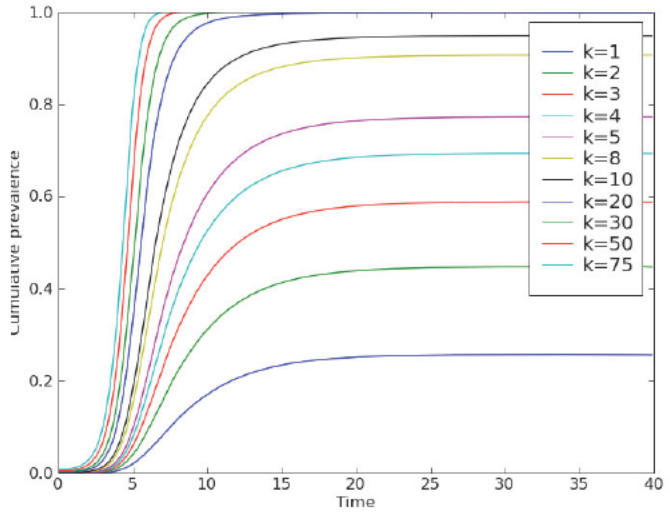
For Poisson in figure 3.10, we can see that the nodes of degree 3 have the most number of infected nodes at the end of the epidemic, followed by nodes of degree 2, 4, 5, 1 and then 8. Not shown in figure 3.8 are the number of nodes with degree greater then 8, but they were not included in the graph because the number of total infected was so low that it could not even be read. Also, the nodes with degree greater than 8 just continue decreasing in number of total infected as the degree decreases, simply because there are not that many nodes with high degree to begin with. The interesting thing to note about figure 3.10 is that in a Poisson distribution of average degree 2.5, there are more nodes of degree 2 than degree 3 because of inherent properties of the Poisson distribution. Thus the fact that a greater number of nodes of degree 3 are infected than of degree 2 is significant in that this extra contact makes nodes of degree 3 significantly more likely to be infected than nodes of degree 2, at least when $\rho = 0$. The same logic can be applied to nodes of degree 4 and 1.

Looking at figure 3.11 for the power-law distribution, we see that nodes proceed in order from lowest to highest in degree. The nodes with degree of 1 have the highest total number of infected, followed by nodes of decreasing degree. Degrees after a value of 8 are not included, but they do continue all the way in order to degree 75 having the lowest total number infected. This result corresponds with what we would expect, as there are more nodes of low degree than of high degree.

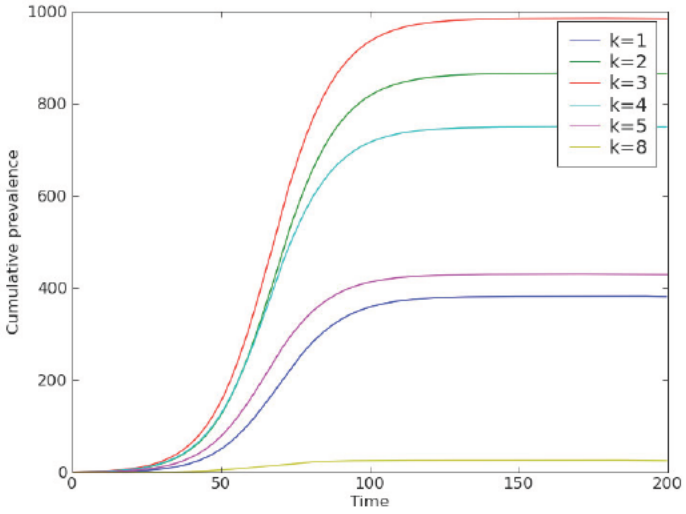
Poisson: Prevalence vs Time $\rho = 0$ (figure 3.8)



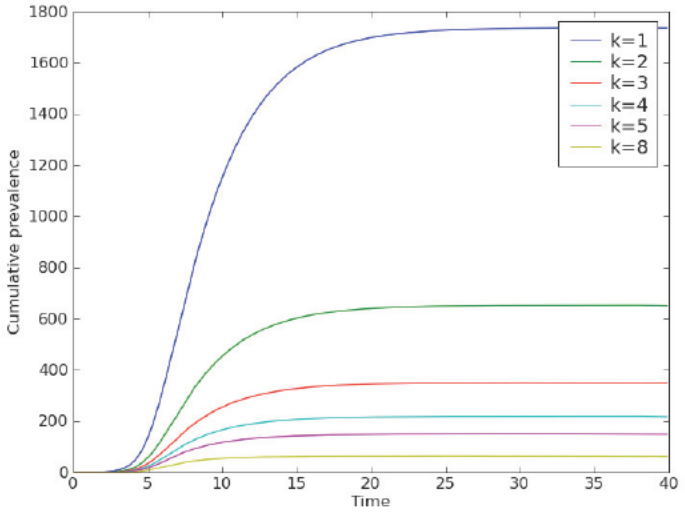
Power-law: Prevalence vs Time $\rho = 0$ (figure 3.9)



Poisson: Prevalence vs Time $\rho = 0$ (figure 3.10)



Power: Prevalence vs Time $\rho = 0$ (figure 3.11)



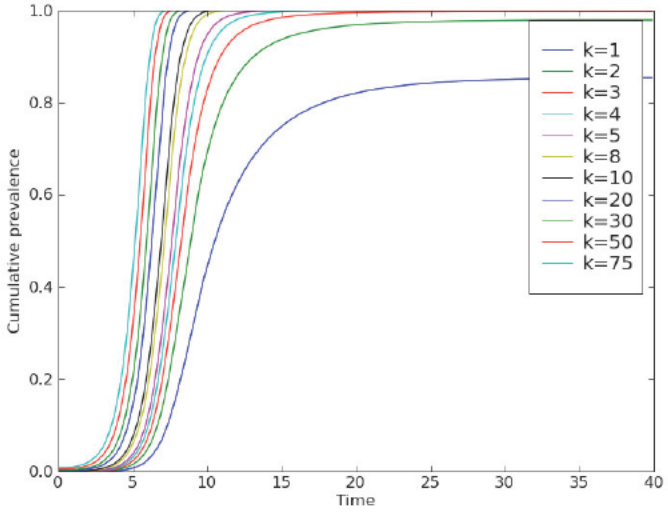
In order to make a comparison, we must look at these same values, given that $\rho = 1000\mu$. As we can see in figures 3.12 and 3.13, the proportion of nodes with a certain degree that were infected in the outbreak is as we expect it would be for both the Poisson and power-law distributions. The nodes of lowest degree have the least percent infected while nodes of higher

degree have a large percent infected. However, the time scales now both extend to 40 time units as the Poisson network becomes infected just as quickly as the power-law network. Now looking at figures 3.14 and 3.15, we see drastic changes in the graphs from figures 3.10 and 3.11. In figure 3.14, we can see that the Poisson distribution is different than in figure 3.10. Now, nodes of degree 2 have the highest total number infected, followed by nodes of degree 3, 1, 4, 5, and then 8. After this, the total number infected continues to decrease as the degree of the nodes decreases. This seems to make sense, as there are more nodes of degree 2 so that there should be more that are infected than nodes of degree 3, and the two values have switched positions from in figure 3.10. The same applies to nodes of degree 4 and 1. In figure 3.10, nodes of degree 4 had a higher total number infected, while in figure 3.14, nodes of degree 1 have a higher number infected because there are more total nodes of degree 1 to become infected in the first place. As we can see, ρ has a substantial effect on the types of nodes (their degree values) that become infected. An interesting thing to look into further would be the inflection points of figure 3.14. Up until specific times, the nodes of various degrees are infected at different totals.

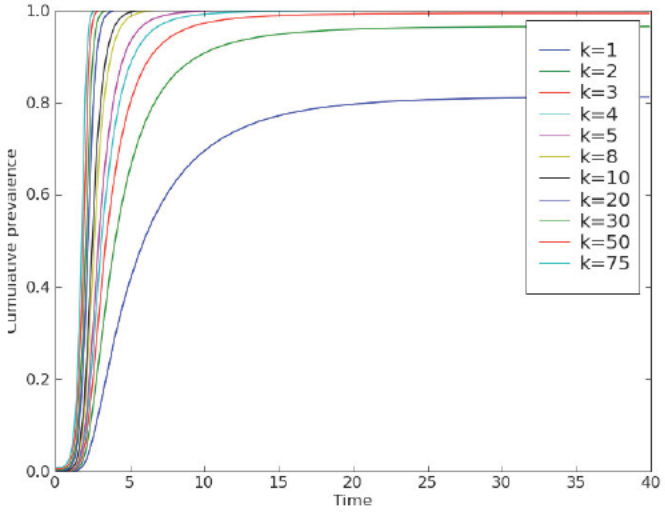
Looking at the power-law distribution in figure 3.15, we see a similar graph shape as in figure 3.11. The nodes of degree 1 have the highest cumulative prevalence, followed by nodes with systematically decreasing degrees. However, there is a much larger jump from the number infected of degree 1 to degree 2 than in figure 3.11. In figure 3.11, the cumulative prevalence for degree 1 is about 1750 while in figure 3.15 it is about 5,500. Similarly nodes of degree 2 jump from 600 in figure 3.11 to about 1200 in figure 3.15. However, as the degree value increases, the total number infected look about the same, no matter the value of ρ. From this, we can assume that almost all nodes of high degree become infected in the outbreak, and this hypothesis

can be confirmed by figures and 3.9 and 3.13 as we see that almost all nodes of high degree become infected.

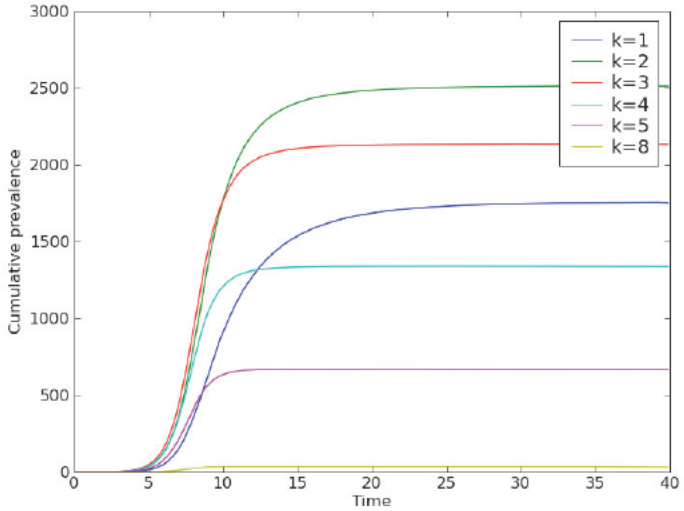
Poisson: Prevalence vs Time $\rho = 1000\mu$ (figure 3.12)



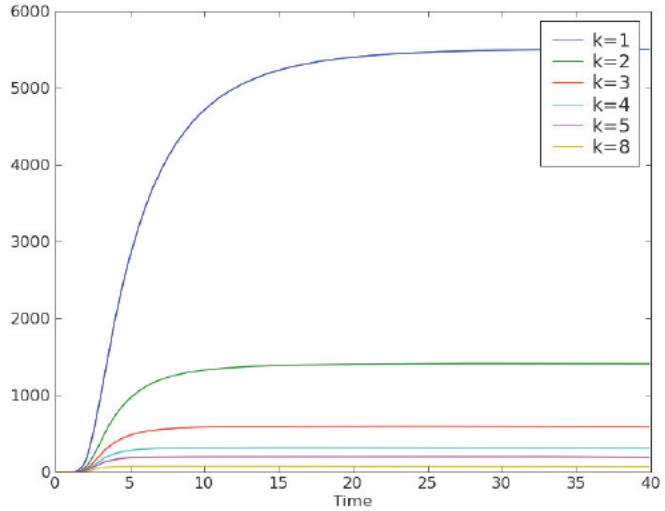
Power: Prevalence vs Time $\rho = 1000\mu$ (figure 3.13)



Poisson: Prevalence vs Time $\rho = 1000\mu$ (figure 3.14)



Power: Prevalence vs Time $\rho = 1000\mu$ (figure 3.15)



DISCUSSION

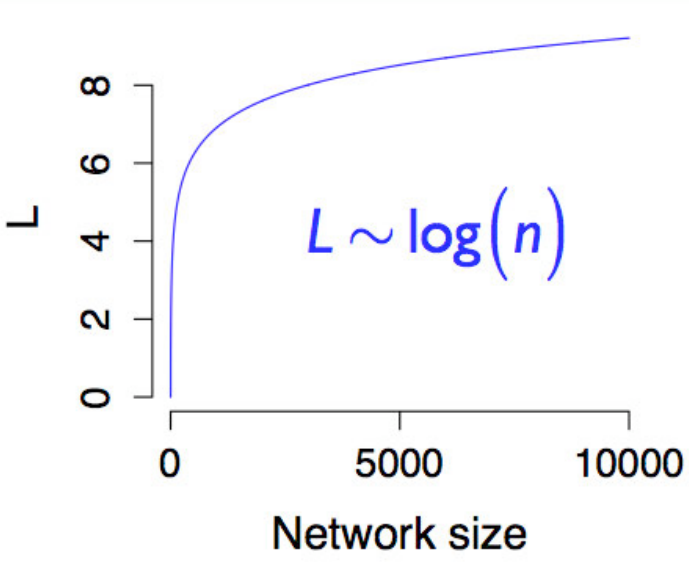
The network structure of a power-law degree distribution typically has a hub-and-spoke model so that when disease finally does reach a hub, there will be a burst of infections as the spokes that surround the hub become infected simultaneously. However, when that hub recovers, disease spread will slow, as the spokes don't have many contacts with which to spread disease. Now the Poisson model, on the other hand, will spread disease more slowly but also more steadily. Perhaps this will occur because one single node cannot be held responsible for a substantial proportion of infections in the population, but each node could be held responsible for a few. However, the actual size of the epidemic is larger for the Poisson distribution for all values of ρ. We can see this when comparing figure 3.3 and 3.4. In figure 3.3, we see how the Poisson distribution (at each value of ρ) reaches the end of the epidemic more slowly than does the power-law in figure 3.4, but we also see that the total epidemic size is larger for Poisson than power-law at each value of ρ.

Looking at the time to peak incidence versus ρ in figure 3.7, we see that for the Poisson distribution the time decreases dramatically as ρ increases, while in the power-law distribution the time remains relatively the same no matter the value of ρ. Thus ρ doesn't really effect the power-law distribution, except that the time to peak incidence does decrease a bit. The fact that the Poisson network spreads disease more slowly at low values of ρ than does the power-law network stems from the shorter path lengths of the power-law network. Disease spreads much more slowly in a Poisson network because no one node has very many contacts, and disease must travel a longer distance before reaching the outskirts of the network. Comparatively, the power-law network has fairly short path lengths between any two nodes as the hubs allow you to move quickly through the network. Hubs have many contacts so that once you reach a hub, you

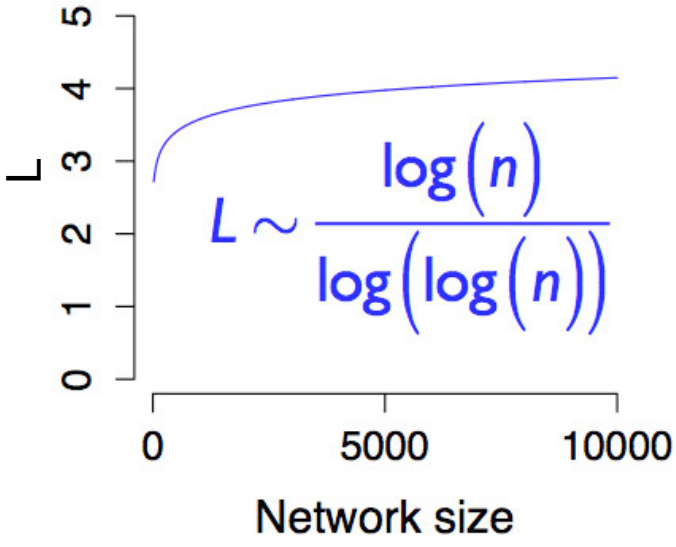
can easily move to a different part of the network simply by passing through that hub. Since hubs provide a fast route to a significant portion of the network, the location of patient zero doesn't matter as much because it won't take the disease many time steps to reach a hub, thus spreading disease quickly.

In Newman 2003, he actually characterizes the average shortest path length on each of these networks. In figures 4.1 and 4.2, we can see the curves of the average shortest path length plotted against the network size for Poisson and power-law distributions, respectively. Both path lengths plateau at fairly low lengths, with the Poisson average path length at about 9 and the power-law average path length at about 4. Also, the Poisson path length is proportional to the $\log(n)$ where n is the network size, and the power-law path length is proportional to $\frac{\log(n)}{\log(\log(n))}$ (Newman 2003). So even from these equations, we can see that the power-law network will have shorter path lengths than the Poisson distribution for the same size network, and thus the disease will spread more quickly on the power-law network.

Poisson: average shortest path length (figure 4.1)
Figure taken from Lauren Meyers Bio 337 Class



Power-law: average shortest path length (figure 4.2)
Figure taken from Lauren Meyers Bio 337 Class



An interesting problem in the future would be to address how simultaneously changing ρ and μ could affect the characteristics of the outbreak. Changing μ actually changes the length of the infectious period so that individuals have less time to switch partners when mixing. Since different diseases have different lengths of infectious periods, it makes sense to see how these two parameters interact.

Given what we've found to be true of power-law and Poisson networks, we could use these networks to compare the effects of contact patterns on the spread of diseases. Power-law distributions can often represent sexual contact patterns and therefore STD's, while Poisson distribution may be more appropriate for external contact pattern networks such as the flu (Bansal, Grenfell, and Meyers 2007). However, the differences found in time to total epidemic, total epidemic size, and degree value of those who become sick, point to the importance of the contact structure that underlies the spread of disease.

Especially as the value of ρ changes, we see how it affects the two distributions differently. The power-law distribution already has some nodes with large degrees (as it is heterogeneous) whereas the Poisson distribution has none (as it is homogenous). So although increasing the mixing rate will give a power-law network more nodes with high degree, the path lengths are already so short that this doesn't have a huge effect on the characteristics of the disease spread. For example, the time to peak incidence doesn't change as the value of ρ increases, as seen in figure 3.7. For the homogenous Poisson distribution, increasing the mixing rate leads to high degree nodes, and these nodes then make the shortest average path length even shorter. Thus the mixing rate has a huge impact on the time to peak incidence as seen in figure 3.7. Despite these differences in timing, the total epidemic size is basically the same for both

distributions. As seen in table 3.1, the Poisson distribution generally produces a higher epidemic size than the power-law distribution, but even these numbers are fairly close.

WORKS CITED

Bansal, S., B.T. Grenfell, and L.A. Meyers (2007) *When individual behavior matters: homogeneous and network models in epidemiology*. Journal of the Royal Society Interface **4**: 879-891.

M.E.J. Newman. *The spread of epidemic disease on networks*. Physical Review E 66:016128, 2002.

M. E. J. Newman. *The structure and function of complex networks*. SIAM Review **45**, 167-256, 2003.

Erik Volz and Lauren Meyers. *SIR dynamics in random networks with heterogeneous connectivity*. Springer-Verlag, 2007.

Lauren Ancel Meyers. *Contact Network Epidemiology: Bond percolation applied to infectious disease prediction and control*. Bulletin of the American Mathematical Society. Vol 44, No. 1, 63-86, 2007.

Erik Volz and Lauren Meyers. *Susceptible-infected-recovered epidemics in dynamic contact networks*. Proc. R. Soc. B 274: 2925-2933, 2007.

Wikipedia, *probability-generating function*. April 2009. http://en.wikipedia.org/wiki/Probability_generating_function

APPENDIX

The following code is all in the Python programming language. It is included to help anyone trying to reproduce the graphs found in the Results section. To begin the program, you must import the packages scipy, pylab, and math.

```
from scipy import *
from scipy.integrate import *
from pylab import *
import math
```

Then you need to create g, g', and g'' for both Poisson and power-law. Poisson distributions have a special shortcut to create these three PGF's, but power-law distributions follow the generic pattern, and any other distribution can be substituted into the PGF's for the power-law distribution equation. See Table 5.1 for these PGF's for both Poisson and power-law.

Table 5.1

	Poisson	Power-law (Generic)

g'	def calc_g(lamda,x):
	g_val=math.e**(lamda*(x-1))
	return g_val
		def calc_g(**p_vec**,x):
		g_val=0
		for k in range(len(p_vec)):
		g_val = g_val + p_vec[k]*x**(k)
		return g_val

g''	def calc_g1(lamda,x):
	g_val=lamda*math.e**(lamda*(x-1))
	return g_val
		def calc_g1(**p_vec**,x):
		g_val=0
		for k in range(len(p_vec)):
		g_val = g_val + (k)*p_vec[k]*x**(k-1)
		return g_val

g'''	def calc_g2(lamda,x):
	g_val=(lamda**2)*math.e**(lamda*(x-1))
	return g_val
		def calc_g2(**p_vec**,x):
		g_val=0
		for k in range(len(p_vec)):
		g_val = g_val + (k)*(k-1)*p_vec[k]*x**(k-2)
		return g_val

In order to create the p_vec that is found in table 5.1, we need to generate it. The function needed to to do this is:

```
def power_law(alpha,maxdeg):
    dist=[0]*(maxdeg+1)
    alpha=float(alpha)
    denominator=0.0
    for i in range(1,maxdeg+1):
        denominator+=i**(-1*alpha)
    for k in range(1,maxdeg+1):
        dist[k]=k**(-1*alpha)/denominator
    return dist
```

Finally, the derv function in Python that integrates the dynamic equations found in table 2.7.

```
def derv(x,t,lam,rr,mm,pp):

    #y[0]= change of theta
    #y[1]= change of p_infec
    #y[2]= change of p_suscep
    #y[3]= proportion of S
    #y[4]= change of M_I
    #y[5]= change of I

    y=zeros(6);
    y[0]=-rr*x[1]*x[0]
    y[1]=rr*x[2]*x[1]*x[0]*calc_g2(lam,x[0])/calc_g1(lam,x[0])-rr*x[1]*(1-x[1])-x[1]*mm+pp*(x[4]-x[1])
    y[2]=rr*x[2]*x[1]*(1-x[0]*calc_g2(lam,x[0])/calc_g1(lam,x[0]))+pp*(x[0]*calc_g1(lam,x[0])/calc_g1(lam,1)-
x[2])
    y[3]=-rr*x[1]*x[0]*calc_g1(lam,x[0])
    y[4]=-mm*x[4]+rr*x[1]*(x[0]**2*calc_g2(lam,x[0])+x[0]*calc_g1(lam,x[0])/calc_g1(lam,1))
    y[5]=rr*x[1]*x[0]*calc_g1(lam,x[0])-mm*x[5]
    return(y)

out=odeint(derv, init, time, args=(lamda,r,mu,ro))
```

BIOGRAPHY

Megan Watson was born in Denton, Texas, on March 19th, 1987, where she lived until graduating high school in 2005. She then enrolled in the Plan II Honors program and the Dean's Scholars Honors Program in Mathematics at the University of Texas at Austin. In college, she researched probability at East Tennessee State University, interned at the food pantry of Eastside Community Connection, participated in Alpha Delta Pi Sorority, and tutored all ages of children. She graduated Phi Beta Kappa in 2009, and for the next two years she plans to move to Houston as an energy trader for Citigroup. This summer, Ms. Watson plans to hike the Camino de Santiago de Compostela pilgrimage in the north of Spain.