STOCHASTIC DYNAMICS IN NETWORK STRUCTURES

A THESIS

Presented to

The Faculty of the Department of Mathematics

The Colorado College

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Arts

By

Maggie Bailey

May 2015

Stochastic Dynamics in Network Structures

Maggie Bailey

April 14, 2015

Abstract

Epidemics have long been studied in the realm of mathematical modeling, aiding in parameter estimation, determining effective intervention methods, and, in some cases, predicting epidemics. How do networks and network structure influence infectious diseases? Are certain networks better suited to naturally contain an epidemic where others are not? In this paper, we will explore one method of stochastic modeling and implement it into a network of sub-populations. We will conclude with interesting results on networks and their dynamics in disease propagation.

1 Introduction

Epidemics never fail to capture the world's attention due to their potential danger to populations across the globe. The Ebola outbreak in Guinea that began in the year 2014 was no exception. Media produced graphs predicting future numbers of cases that instilled fear in society. But, how likely is it that an outbreak spreads through large, urban areas and rural villages, a complex structure of networks? Will outbreaks always occur when a single individual becomes infected in a small village in West Africa or will the disease fizzle out? These questions can be answered using mathematical modeling which helps define various characteristics of a disease such as its transmission rate, effective control methods, and likely duration. Specifically, stochastic modeling captures the randomness that defines the beginning of an outbreak and illustrates the potential for an epidemic depending on the defined parameters. Often in stochastic modeling, an outbreak may never occur. With the recent Ebola epidemic in West Africa and its trans-Atlantic jump, the complex network through which Ebola has spread could be modeled using stochastic methods. This paper aims to study the theoretical spread of infectious diseases, such as Ebola, through various types of networks in order to answer questions about network structure and its effect on the spread of a disease.

The use of mathematical modeling has extreme importance in studying the characteristics of infectious diseases and their control methods. Daniel Bernoulli was the first mathematician to quantify the spread of smallpox using mathematical modeling in 1766. From this, standard epidemic modeling was born. The standard model is now widely used in exploring an assortment of infectious diseases such as Ebola, STD's, and foot and mouth disease (Lekone 2006; Bearman 2004; Ferguson 2001). Mathematical models help in discovering effective intervention methods to contain an epidemic, estimating parameter values, and determining which parameters have higher influence over transmission. Control methods range from

wide spread slaughtering of farm animals to slow the spread of foot and mouth disease within farms or vaccination of "core" individuals- those with a high number of connections- in infected populations.

Modeling the spread of disease often follows a conventional SEIR model or, depending on the disease, SIR model. These models exhibit the movement of individuals within a population, or populations, through the four stages of a disease: susceptible (S), exposed (E), infectious (I), and removed (R). The susceptible population consists of individuals that are not yet immune to a disease and have the potential to contract it. Exposed individuals are infected but are not yet infectious to others and so infectious individuals carry the disease and have the potential to spread it to susceptible individuals. Finally, an individual will move into the removed, or recovered, stage, signifying that an individual is either deceased or recovered from a disease and immune to it. In some cases, an individual may recuperate and become susceptible again. In this case, there are only two stages of the disease classifying it as an SIS model. Diseases of this type include the common cold where immunity is never attained after recovery. In my model, we will focus on the standard SEIR model consisting of the four stages an individual may move through during an epidemic. Many infectious diseases follow this trajectory including whooping cough and Ebola. (Black and McKane 2010; Lekone 2006)

Epidemic mathematical models often fall under two categories: stochastic or deterministic, deterministic being the simpler of the two.Deterministic and stochastic models use either discrete or continuous-time approximations. In discrete time, the length of each interval is non-variable ranging from one day to the length of the incubation period of a disease. (Allen and Burgin 2001) Deterministic models utilize a series of differential equations to estimate the relative sizes of populations. Both types of models may assume a closed population with homogeneous, random mixing. That is, every individual has an equal probability of contact with any other individual in the population. Homogeneous mixing generally captures the dynamics of an infectious diseases but it is not always the most accurate (Keeling 2005). In the case of STD modeling where contacts are biased and chosen based on specific characteristics, homogeneous mixing is not realistic. (Bearman 2004) While homogeneous, random mixing is not always the most realistic it is a convenient assumption to help simplify a model. Deterministic models follow a system of differential equations representing a population of N individuals:

$$\begin{aligned}
\frac{ds}{dt} &= -\beta si \\
\frac{de}{dt} &= \beta si - \lambda e \\
\frac{di}{dt} &= \lambda e - \gamma i \\
\frac{dr}{dt} &= \gamma i
\end{aligned}$$

Here, $s(t)+e(t)+i(t)+r(t) = N$. Thus $s(t)$, $e(t)$, $i(t)$, and $r(t)$ are the sizes of the respective populations at time t. An individual follows the trajectory through $S \to E \to I \to R$ once infected with the disease.

The three parameters are β, the transmission rate per unit time, λ, the incubation rate, and γ, the recovery rate.

From this system of equations, we can derive the value of R_0 which can give a short characterization of a disease. It can be thought of as the basic reproduction rate of a disease. That is, given a single infected individual, R_0 represents the number of secondary infections caused by the infected individual. (Anderson and May 1991) It is approximately equal to $\frac{\beta}{\gamma}$, where β is the transmission rate and γ the recovery rate of a disease. The R_0 value may differ in definition from one disease to the next however its significance holds. As it is a threshold parameter, it exhibits the likelihood of an outbreak in addition to the growth of an epidemic at the beginning of an outbreak in a population of mainly susceptibles. If $R_0 > 1$ an epidemic will likely occur. Conversely, if $R_0 \leq 1$ an epidemic will hopefully be avoided. (Trapman 2006) While R_0 has historically been one of the more important values in deterministic models, studies have examined other parameters that have similar importance in disease propagation. In terms of disease control, it has been found that the value θ, defined as the proportion of transmissions occurring prior to symptoms, is crucial in reducing the number of infections and determines actions for symptom-based public health. (Fraser 2004) It has also been found that R_0 is affected by small changes in the average number of contacts and the degree of clustering within a population. (Keeling 2005)

The system of deterministic equations above can be made more complicated to incorporate realistic properties of populations such as a natural birth and death rates. Additionally, this system of equations can be altered to model the implementation of control methods like vaccines or other driving forces within an epidemic. (Keeling and Eames 2005)

Deterministic models work well within the framework of large populations over a long period of time with previously gathered data. Unlike stochastic models, deterministic models produce the exact same output over multiple simulations given an initial condition. While deterministic models are simpler and suitable for various aspects of disease modeling, such as evaluating intervention methods, there are many drawbacks to this type of modeling. Deterministic models are not accurate in expressing disease duration, namely the beginning and end of epidemics since stochasticity largely affects this time period when the populations of interest are small. (Trapman 2006) Additionally, deterministic models are not exact when studying epidemics within networks of populations. Neither do they capture the variability that occurs with a single disease. That is, they do not capture the possibility of no outbreak. (Keeling and Ross 2008)

Where deterministic models may fail, stochastic models provide a better exhibition of disease propagation. Given an initial condition, stochastic models will rarely produce the same output through consecutive implementations of the same set of equations. Stochasticity is event driven. It focuses on the probability that an epidemic may occur while deterministic models always guarantee that there will be an outbreak when $R_0 > 1$. A single simulation of a stochastic model does not accurately portray

the possible range of infected individuals an epidemic may have thus many simulations are required to understand its dynamics. Unlike their deterministic counterparts, the methodology of stochastic models may vary greatly between diseases. The Reed-Frost model, first developed in 1927, is the most basic stochastic model. It is an SIR model where each state of infection only lasts one time period, t. A Markov chain determines the spread of infection. Infection can only pass during the time period when a susceptible comes into contact with an infected individual. The basic equation that defines this model is $C_{t+1} = S_t(P_{t+1})$ where $P_{t+1} = (1 - q^{C_t})$. This model is state dependent so the population of the next time step depends on the current population and is not affected by any populations previously.

Chowell uses a continuous time Markov chain to estimate parameter values for Ebola virus outbreaks in the Democratic Republic of Congo (1995) and Uganda (2000). Three possible events were considered: exposure, infection, and removal. Event times were defined as the movement of an individual from one state to the next with even increments distributed exponentially. (Chowell 2004) It has been noted, however, that exponential distribution for the exposed and infectious periods is not realistic since the probability of recovery depends heavily on the time since infection. In a stochastic whooping cough model, the exposed and infectious classes were split into stages that an individual traverses in sequential order. If there are M and L stages in the exposed and infectious periods, respectively, then we have a gamma-distributed period with rates $M\sigma$ and $L\gamma$ between each stage. (Black and McKane 2010) This alteration creates a more realistic model for the spread of whooping cough.

The stochastic models explained above typically represent a single populace and could be applied to networks and network theory. Network theory is based in the social sciences and in graph theory (Keeling 2005). In disease dynamics, hosts and contacts are translated into vertices and edges on a graph. In order to construct a realistic network, a great amount of research and data collection must be done on the pattern of contacts within a population. In large populations, this is extremely time consuming. Past research in network theory focuses on the spatial structure of individuals in a single population. (Keeling 2005, Moore 2000) Additionally, there has been a great deal of research on network structures for the spread of STD's. It has been found that for the spread of STD's, the structure of a given network is extremely important for propagation, especially concerning core groups, or sub-populations of individuals with a high number of contacts. Furthermore, this network structure affects intervention methods. For example, target vaccination or quarantine may be most effective in slowing the spread of STDs. (Bearman 2004)

The process of disease transmission differs for air-borne diseases and influenzas so the network structure is assumed to be different. To trace the spread of disease in these complicated networks, researchers have used infection tracing, contact tracing, and diary-based studies. (Keeling 2005) This tracing has led to various networks structures that focus on the position of individuals connected by edges in a graph. Random, lattice, small world, spatial, and scale-free are five basic networks structures where the posi-

tion of an individual is defined differently for each network. For example, in random networks contacts between individuals are not based upon the spatial location of an individual. On the other hand, lattices are highly structured and contacts are made only between individuals that are directly connected by an edge. Small world networks are regarded as more realistic in social networks and epidemiological settings, where contacts are determined spatially with some random short paths to other individuals in the population making the spread of disease rapid in most cases. (Keeling 2005) Each network model is defined by specific probabilistic equations, often a Poisson or normal distribution.

Keeling has also developed a pairwise model for SIS diseases. Here, higher ordered structure is ignored along with network interactions in its entirety. Instead, connections between every type of pair are the focus. Keeling rewrites the standard differential equations to examine SI, SS, and II pairs. The number of different pairs is incorporated as a variable rather than a number of individuals since the number of connections can be affected by infections within the pair, outside the pair, or due to immunity. An advantage of this type of modeling is that only samples of each pair type within the population are required to get an accurate estimate of population movement rather than data collection of the movement of an entire population. However, pairwise modeling does not take into account higher ordered structures within a population and are not as accurate as network modeling. (Keeling 2005)

Research has been done on the global spread of diseases in multiscale networks such as those structured around airports and surrounding cities. One study found that the method of transportation has an effect on the spread of disease and it is therefore important to consider when analyzing epidemic invasion and timing. Commuting has a small effect on the percent of global outbreaks but does affect the timing of the disease. Commuting also synchronizes the timing of the heights of epidemics. Unsurprisingly, short-range commuting has a larger effect on the local spread of disease while airport transportation has a larger effect on the global spread of disease. (Balcan 2009)

While most research on networks has focused on the spatial structure of individuals, the structure of connected populations may also be important in the spread of disease. My thesis focuses on exploring one method of stochastic modeling that shows epidemic dynamics in a variety of networks. Unlike many previous studies, my model focuses on the networks of sub-populations instead of networks of individuals. I first examine modeling one village to gather base data and then move into networks of higher numbers of villages. I examine the structure of each network and how the spread of disease changes based upon the structure of small sub-populations, the size of each sub-populations, and various parameter changes. I hope to gain insight into network structure and disease propagation and possibly determine which structures are better suited to naturally contain the spread of disease.

2 Model and Methods

I implemented my mathematical model in the computer program Mathematica. My model focuses on the early stages of an epidemic and communities of smaller populations within various types of simple network structures. As I will show, the stochastic model does not function as well for larger populations over longer periods of time as it begins to resemble a deterministic model, losing random characteristics. However, some element of stochasticity is preserved even with larger populations. In my model, the susceptible population of each village starts with a certain number of individuals, ranging from 100 to 500 individual, and a single individual in the infected population. The exposed and removed populations initially have size equal to zero.

My model uses probabilistic equations to determine the number of individuals that move populations in a single time set. The movement of individuals from each stage of the disease, S, E, I and R, is determined by selecting a random number from a binomial distribution. Thus, the probability that x individuals successfully leave one stage to move into the next is equal to $\binom{n}{x} * p^x(1-p)^{n-x}$. Here n is equal to the population of the village and x is equal to the parameters defined in the model (β, σ, and γ). A success in the susceptible population is occurs when a contact between an infected individual and a susceptible produces an infection in the susceptible individual. The probability of a success equates to the parameter values that determine a disease. Similar to most $SEIR$ models, these parameters are β, the transmission rate, γ the latent period of the disease, and σ, the recovery rate. In my model, the probability of transmission changes as the infectious population increases. They probability of infection is p where $p = 1 - (1-\beta)^i$ and i is the number of infectious individuals at time t. Thus an individual moves from the susceptible stage to the exposed stage with probability p, and an individual spends an average of $\frac{1}{\gamma}$ and $\frac{1}{\sigma}$ days in the exposed and infected stages, respectively, before moving into the removed stage.

In the binomial distribution, the number of trials is equal to the number of individuals in each stage at a given time step, measured in days. An individual moving into the removed stage is assumed to have either recovered from the disease or died. I assumed homogeneous mixing so that every individual has the same probability of becoming infected. In other words, no individual is assumed have immunity to the disease when $t = 0$. While this is not completely realistic, it has been assumed for simplicity reasons in building the model.

I first studied the spread of disease throughout one village to build a baseline data set for comparison to network structures. To test the parameters, the population varied from 100 to 1000 individuals in the susceptible population with a single individual initially in the infected population. The model is structured by first defining the parameters described above as $\beta = 0.0008, \gamma = 1/6$, and $\sigma = 1/6$. The γ and σ values are based upon previous studies done on the Ebola virus. (Lekone 2006)

I then translated the single village model into a network of villages. The network structures examined

are based upon theoretical network structures for villages and adapted from previous studies in networks structure of individuals. (Keeling 2005) Three network structures are evaluated: a cross like structure of five villages with a "central hub" or downtown area, where the central population is larger than the surrounding sub-populations; a pentagon structure where five villages are in a circular outline and only connected to the neighboring villages; and finally, a pentagon shaped network where each village is connected to every other (Figure 1). The structure of the various networks is determined by a 5x5 matrix of values where the entry a_{ij} is the probability an individual from community i moves to community j. The probability of migration, 0.01, remains constant for each network so that it does not affect the focus of the study. After implementing the network models once for base data, I changed the population sizes and "central hub" locations to explore the effects of network structure in epidemics. These values are recorded, graphed, and analyzed using statistical methods in Minitab and Mathematica.

While intervention methods are a popular topic of study, I did not implement them here. However, in the future intervention methods in networks of sub-populations is an important topic to be explored.

3 Results

3.1 Base Network- Single Village

In a single network with parameter values $\beta = .0008$, $\gamma = \frac{1}{6}$, $\sigma = \frac{1}{6}$ and a population of 100, very few epidemics occurred over one thousand simulations. A histogram of the data (Figure 2) follows an exponential distribution curve with only four instances of epidemics where more than 10 people are infected in a single run of the model. On average, the number of people affected in a single village with these parameter values was 1.83%. Figure 2 shows a histogram of the frequency of outbreaks, y, occurring with x number of individuals infected. When the population is increased to 300 individuals with the same β, γ, and σ values, the number of epidemics occurring increases significantly. A population of 300 individuals is shown in Figure 3. Here, a similar exponential distribution curve, like that in Figure 2, is seen for x values less than 100 but the histogram resembles a normal distribution as the number of people infected approaches the size of the total population, 1500 individuals. Once the population is increased to 500 individuals, the histogram changes significantly with very few instances where epidemics do not occur but many where most of the population, more than 85%, is infected.

3.2 Central Hub Network

In a central hub network of five villages with parameters $\beta = .0008, \gamma = \frac{1}{6}$, and $\sigma = \frac{1}{6}$ where populations in the four surrounding villages are equal to 100 individuals, the central hub 500 individuals, and the infected individual starting in the central hub, a curve analogous to the exponential distribution left of a normal distribution is produced as in the single base network (Figure 4). Similarly, when the population is

doubled in each village, the curve becomes more drastic with higher percentages of each village becoming infected. Changing the initial village of the infected individual greatly changes the way in which the virus spreads. If patient zero is initially in the north village, or without loss of generality, one of the surrounding villages instead of the central hub, then the number of epidemics decreases (Figure 5).

When the location of the central hub is changed to one of the surrounding villages, numbers one through four, the way in which the epidemic propagates changes as well. When patient zero starts in the new central hub, the histogram once again follows the exponential distribution partnered with a normal distribution graph pattern. And similarly, with patient zero initially located in a non-central hub village with the central hub located in village one (the north village), the histogram solely resembles an exponential distribution curve with the number of outbreaks notably decreased. Unsurprisingly, lowering the transmission parameter, β, to 0.0006 decreases the number of outbreaks but the frequency pattern of the histogram is the same shape as when the β value is equal to 0.0008 (Figure 6).

However, when each village is equally sized and the infected individual starts in the north village, the number of epidemics produced behaves like that of the 100 individual sized village (Figure 7). When the infected individual starts in the center village of the network, the number of epidemics is slightly less than when patient zero is initially in the north village. With equal sized villages, the number of epidemics produced seems to have no impact from the starting location of patient zero.

3.3 Pentagon Network

The pentagon network follows a similar trajectory as the other networks when evaluated over several sizes of populations and starting points for patient zero. With equal sized villages, frequencies of outbreaks follow a similar exponential distribution curve as the base network of 100 individuals. As the population is increased with a β value of .0008, the frequency of outbreaks increases and we see an exponential distribution curve for no outbreaks and a normal distribution curve when outbreaks occur (see Figure 4).

3.4 Connected Pentagon Network

In a connected network of five villages, where each village has an equal population size and is connected to every other, the number of epidemics follows a similar exponential distribution curve as we have seen in the other two network structures (Figure 8). When the population is increased, the number of epidemics also increases and again we see an exponential distribution with an added normal distribution curve on the right as x, the number of infected individuals, approaches higher values. A comparable histogram is produced when two villages have a higher population and the other three are equated to 100 or 200 individuals.

4 Discussion

After running the model over several types of network structures, I found an analogous pattern across all histograms of the networks: an exponential distribution on the left coupled with a normal distribution on the right. I first analyzed this pattern as whole and attempted to make sense of it as one curve. I then split the graphs into two- analyzing the exponential distribution and the normal distribution individually.

The normal distribution curve occurred whenever the initial infected individual had a direct interaction with the largest population in the network structure. That is, patient zero must initially be in a large population or the central hub. This graph pattern rarely appeared when patient zero was placed in the neighboring sub-population to the central village. The exponential and normal distributions also occurred when the population sizes were increased significantly, by double the amount or more. This suggests that in order for an outbreak to occur, patient zero must interact with the larger population initially in order for the disease to propagate. An increase in population size may also cause this pattern. However, increasing the population size takes a step away from the stochastic process since I wanted to focus on smaller populations.

The exponential distribution curve was generated when there was no central hub included in the network structure, meaning that all villages were of equal population size, or in histograms where the normal distribution curve was also produced for higher x values. Running several fit tests to the exponential distribution curve was rather inconclusive due to the massive quantity of single patient epidemics, or instances where there were no secondary infections. Without the large number of unsuccessful propagation, the curve appears to follow an exponential distribution very closely. (Figure 9)

Testing the network structure produced very interesting results, however they may be influenced by other factors than the network structures themselves. I attempted to control all other factors of the network by setting the same parameter values for all three network structures and the base network. Yet the population sizes may have had a higher impact on propagation than the network structure. When I increased population sizes, the number of epidemics increased greatly suggesting that population size has an impact on the probability of an epidemic. Conversely, when I decreased the size of the population, the number of epidemics also fell. Because the same curve appeared in every network structure, it could also be concluded that the network structures I studied have little importance in the spread of disease. This result may be seen as a positive result since it suggests that network structures do not increase the probability of an epidemic. In fact, network structures may have little importance in disease propagation. This could signify that disease control may focus on other intervention methods concerning individuals within a population rather than their movement to neighboring sub-populations.

In the future, it may be useful to study the migration matrix in more detail by changing the probability of movement between sub-populations. Additionally, my network structures may be too simple to draw conclusions about the influence of network structure on epidemics. Using my model for n villages would

allow for more complicated network structures. I would also like to examine the parameter values in more depth, specifically how varying the parameter values in various network structures impact epidemic frequency.

5 Conclusion

Applying stochastic modeling to epidemiology generates important results especially when dealing with networks of sub-populations. Here, I studied the effect of network structures on the spread of disease to determine if certain networks naturally contained or encouraged disease propagation.

The networks of sub-populations I studied have an interesting effect on the spread of disease. The starting position of patient zero is especially important in networks with villages of unequal population sizes. When patient zero starts in a larger population of a network, or the central hub, the number of outbreaks is much greater than if the infected individual starts in a neighboring village with a smaller population. This suggests that the central hub of a network is important for disease propagation. Additionally, when the population sizes are increased in each network, the number of outbreaks escalates. Thus, the population size of a sub-population may have a greater effect on outbreak frequency than the network structure itself- where higher populations cause a higher number of epidemics.

The exponential distribution occurred when there were no outbreaks in a network. This distribution was often coupled with a normal distribution representing the total outbreaks when patient zero started in the central hub. It also appeared when all villages were of equal size, reflecting results from the single village model.

The addition of the exponential distribution coupled with a normal distribution suggests that there is no medium response to outbreaks. That is, either an epidemic occurs, with the majority of the population becoming infected, or it doesn't. Medium sized outbreaks arose when when the population size was increased in small increments or when the parameters were changed by small quantities. In the future, it will be helpful to examine the effects of the migration matrix in greater detail.

Overall, the network structure seems to have had no significant effect on the spread of disease. This could have positive effects on the control of disease in various networks, suggesting that different networks will not encourage disease propagation. Therefore epidemiologists should be able to focus on maintaining the disease via individual containment methods rather than focusing on suppressing the disease through preventing migration between small sup-populations.

6 Figures

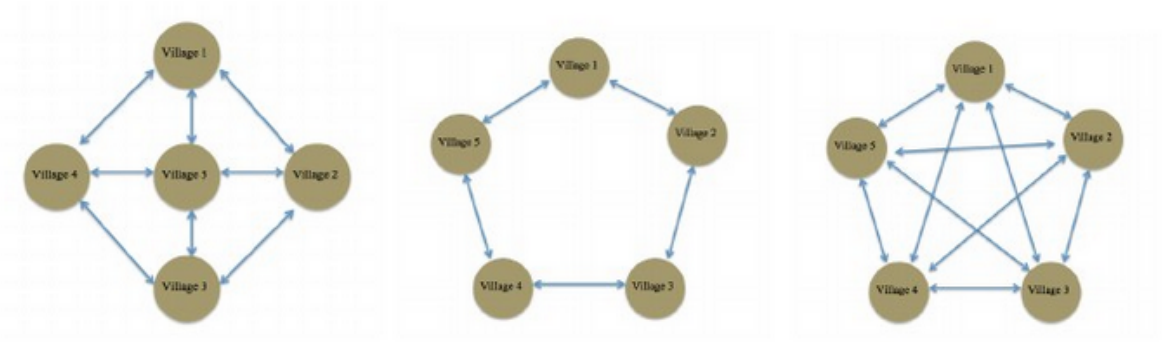

Figure 1: Network Structures- Central Hub Network (left) ; Pentagon Network (center); Connected Pentagon (middle).

In Figures 2-8 the x axis represents the number of individuals infected in a single outbreak and the y axis represents the frequency of epidemics with x number of individuals infected.

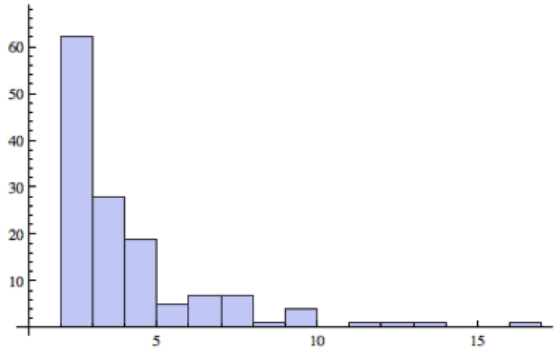

Figure 2: Base Network Histogram (single village)

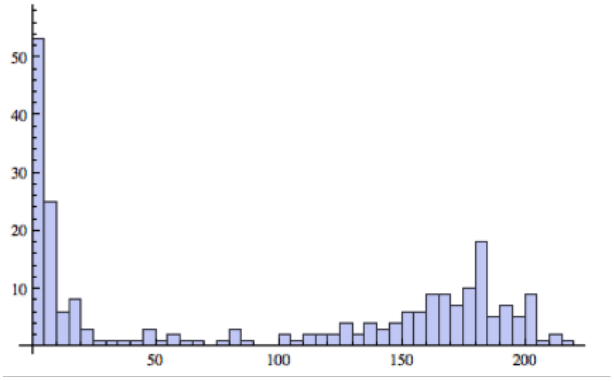

Figure 3: Base Network (300 individuals)

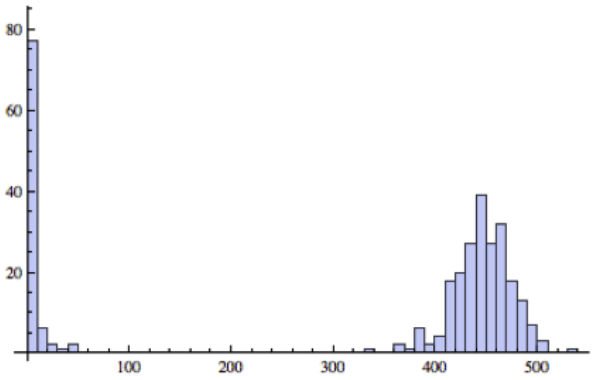

Figure 4: Central Hub Network with infected individual initially in the central hub (village 5)

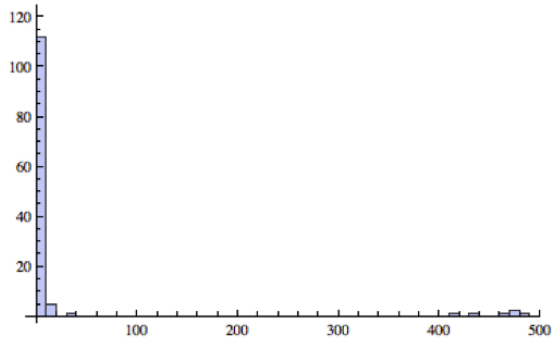

Figure 5: Central Hub Network with patient zero initially in village 1 (100 individuals) where village 5 is the central hub (500 individuals)

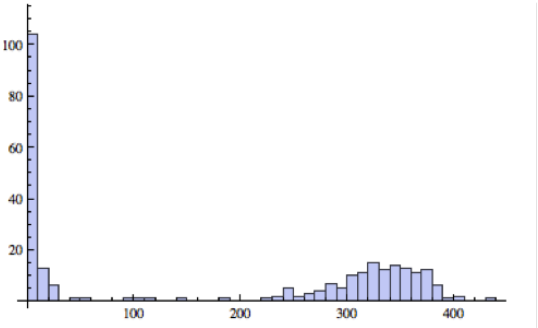

Figure 6: Central Hub Network- $\beta = .0006$; largest population in village 5 (center village) where patient zero starts in village 5.

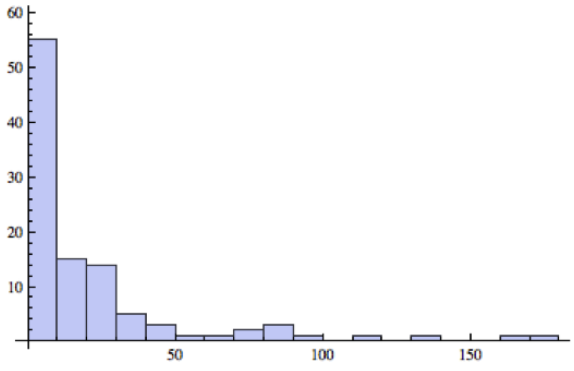

Figure 7: Central Hub Network with equal sized villages (100 individuals each).

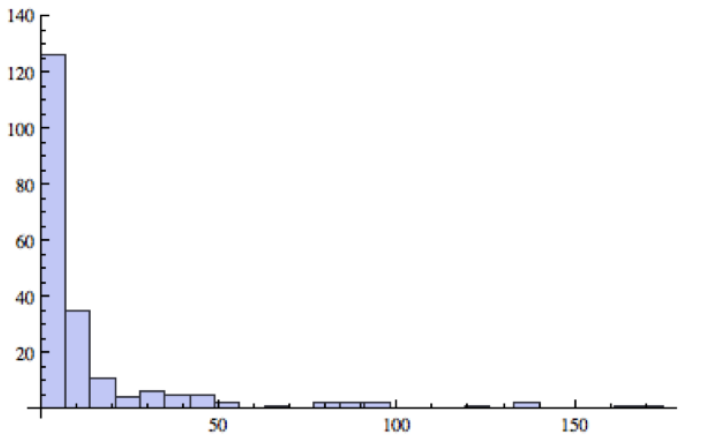

Figure 8: Connected Pentagon Network: equal sized villages (100 individuals).

Figures 9 and 10 display a fit test for the exponential distribution seen on the left side of a histogram (Figure 9) and the normal distribution often seen on the right side of a the histogram (Figure 10).

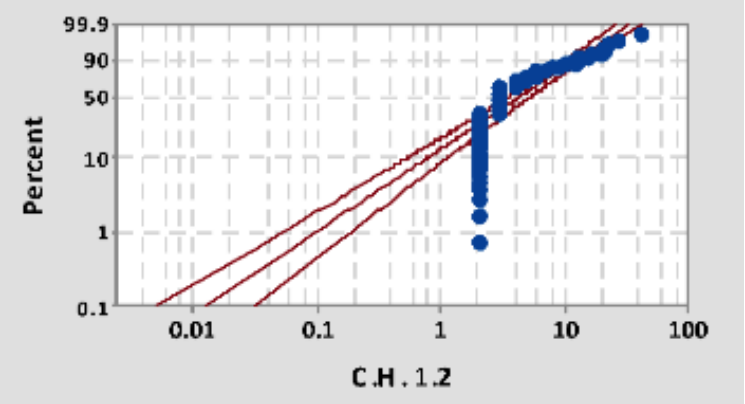

Figure 9: Exponential distribution fit test for Central Hub Network (villages 1-4 have 100 susceptible individuals and village 5 has 500 susceptible individuals).

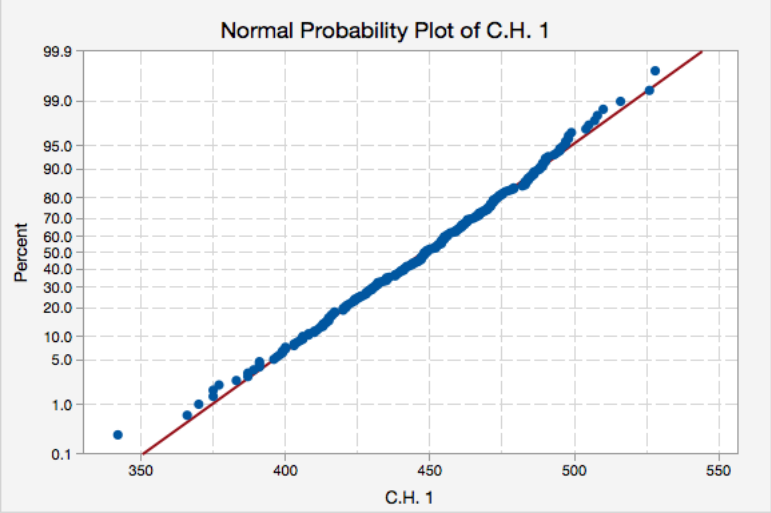

Figure 10: Normal distribution fit test for Central Hub Network with patient zero initially in the central

hub.

References

[1] Allen, Linda JS, and Amy M. Burgin. "Comparison of deterministic and stochastic SIS and SIR models in discrete time." *Mathematical biosciences*163.1 (2000): 1-33.

[2] Anderson, Roy M., and Robert McCredie May. *Infectious diseases of humans.* Vol. 1. Oxford: Oxford university press, 1991.

[3] Balcan, Duygu, et al. "Multiscale mobility networks and the spatial spreading of infectious diseases." *Proceedings of the National Academy of Sciences* 106.51 (2009): 21484-21489.

[4] Bearman, Peter S., James Moody, and Katherine Stovel. "Chains of affection: The structure of adolescent romantic and sexual networks1." *American Journal of Sociology* 110.1 (2004): 44-91.

[5] Black, Andrew J., and Alan J. McKane. "Stochasticity in staged models of epidemics: quantifying the dynamics of whooping cough." *Journal of The Royal Society Interface* 7.49 (2010): 1219-1227.

[6] Chowell, Gerardo, et al. "The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda." *Journal of Theoretical Biology* 229.1 (2004): 119-126.

[7] Ferguson, Neil M., Christl A. Donnelly, and Roy M. Anderson. "The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions." *Science* 292.5519 (2001): 1155-1160.

[8] Fraser, Christophe, et al. "Factors that make an infectious disease outbreak controllable." *Proceedings of the National Academy of Sciences of the United States of America* 101.16 (2004): 6146-6151.

[9] Keeling, Matt. "The implications of network structure for epidemic dynamics." *Theoretical population biology* 67.1 (2005): 1-8.

[10] Keeling, Matt J., and Ken TD Eames. "Networks and epidemic models." *Journal of the Royal Society Interface* 2.4 (2005): 295-307.

[11] Keeling, Matthew James, and Joshua V. Ross. "On methods for studying stochastic disease dynamics." *Journal of The Royal Society Interface* 5.19 (2008): 171-181.

[12] Lekone, Phenyo E., and Bärbel F. Finkenstädt. "Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study." *Biometrics* 62.4 (2006): 1170-1177.

[13] Trapman, P. "On stochastic models for the spread of infections." *ISBN-13*(2006): 978-90.