**Deterministic and Stochastic Models of Influenza**

Ryan Ly

**Introduction**

Background (what is the flu?)

Last year, an estimated 80,000 individuals died from influenza in the United States alone. Influenza, more well-known as the flu, is a common contagious viral infection that affects millions of people worldwide every year. This infectious disease mainly affects the upper respiratory system, the common symptoms including fever, chills, cough, rhinorrhea, congestion, and sore throat. Symptoms typically occur 2 days after exposure and last about one week, with a contagious period of 3-5 days. Although the flu causes similar symptoms to other viruses, like the common cold, the severity of the symptoms can range from mild to severe, which can sometimes lead to hospitalization and/or death. Death, however, usually results from complications from the flu, such as pneumonia or exacerbation of other chronic diseases like asthma, rather than the flu virus itself. This makes the young, the elderly, and the immunocompromised individuals in society the most at risk of acquiring the flu.

However, contracting the influenza virus is not a very difficult task. What makes the flu dangerous is how easily it is spread: the influenza virus can spread from person to person through droplets, which are often transferred through the air or through direct contact with infected surfaces. This is why the flu always spreads in yearly outbreaks, typically beginning in the fall and peaking in winter. At these specific times of the year, people are more likely to stay indoors due to the colder weather and thus, have increased opportunities for close contact, which promotes disease transmission. This contagious characteristic of the flu has led to several large epidemics throughout history. In these epidemics, the influenza virus spread extremely rapidly and was hard to control. One of the most lethal outbreaks of the flu was the Spanish flu pandemic in 1918, in which 50-100 million people died. A majority of these outbreaks, including the Spanish flu, were caused by a specific type of flu, influenza A.

There are 3 main types of the flu: type A, type B, and type C. Type A is the most virulent of all of the flu types and variations in this species, specifically variations in the surface proteins, leads to many different serotypes. Some well-known serotypes include H1N1 (Figure 1), also known as the swine flu, H2N2, the Asian flu, and H5N1, the bird flu. The other two types of the flu are less common and induce less severe symptoms in humans. There are many more subtypes and strains of influenza, which are a result of mutations within the antigenic structure of the virus. Mutation of the virus occurs readily and is what makes it possible for people to get the flu more than once. The virus undergoes up to 8 x 10^-3 point mutations per year.

While a lot is known about the flu virus and its different strains and genetic structure, treatment of the flu is still extremely limited. Since it is a viral infection, treatments largely consist of just symptom control. There are few antiviral drugs (Oseltamivir/Tamiflu) that can decrease the severity of the symptoms of the flu if taken within a short time window after onset of symptoms, but the most effective form of treatment is not a drug, but prevention. One of the best preventative methods to protect against the flu is the flu vaccine. The flu vaccine is prepared annually and protects against several strains of the flu that are predicted to be most prominent that given season. Since there are so many different types of the virus, the flu vaccine needs to be updated every year. Additionally, promotion of good health habits, such as regularly washing hands, covering your mouth and nose when coughing or sneezing, staying at home when sick, etc, helps reduce the spread of the flu.

All of these precautions are needed because the influenza virus is highly contagious. From a public health standpoint, it is important to prevent a large outbreak of the virus, or worst case scenario, a pandemic outbreak of the disease. Pandemic influenza viruses have consistently emerged about every 8 to 41 years over the past several centuries. This is why it is advantageous to understand the dynamics of the spread of influenza in order to better control and prepare against it.

Overview

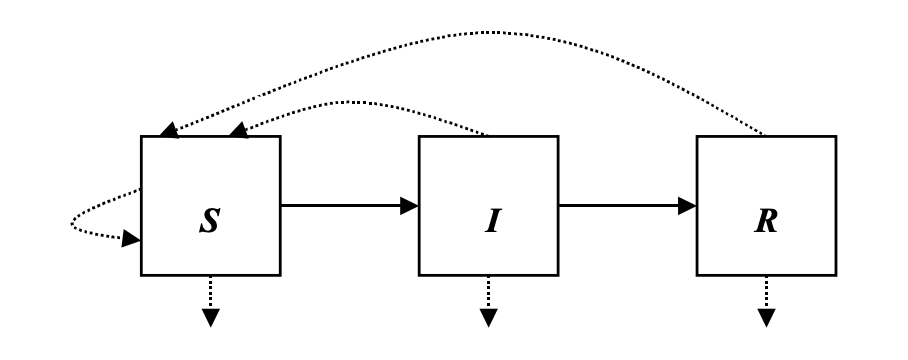
Mathematical models have been used to study many scientific processes and also have the potential to provide valuable insight into the dynamics of influenza. In this paper, I will review the well-known deterministic discrete SIR model and then extend the analysis of influenza beyond this model, studying several other different mathematical models. These models include the deterministic continuous SIR model and its stochastic counterparts, the continuous time Markov chain model and stochastic differential equation model. I will compare the deterministic and stochastic approaches and discuss properties of each model in the context of providing unique information about the dynamics of the influenza virus.

Discrete SIR Model

One of the most well-known mathematical models of infectious disease is the SIR model. The SIR is a compartmental model that breaks down a population into three compartments: the susceptible (S), the infected (I), and the recovered (R). In the context of the flu, during each flu season, every individual is susceptible to contracting the virus and once contract, becomes infected, moving from the susceptible compartment to the infected compartment. Then, after a short period of time, about 1-2 weeks (the duration of the flu), the infected individuals recover and attain immunity against the flu for the rest of the season. In order to effective the following season, a different mutation of the flu will need to be spread since the number of susceptible individuals will be much smaller the season following the outbreak. This model works well for modeling other diseases in which individuals who become infected and acquire immunity, including chickenpox, mumps, and measles.

The SIR model is deterministic, meaning that there is no randomness in the system and that the system will behave the same way for a given set of initial conditions and parameters. While it may not be as realistic as a stochastic model, which will be explored later, it provides a good starting point for understanding flu dynamics as it is easy to interpret. The discrete form of the SIR model is shown below, where n is the number of weeks after the beginning of the flu season:

[INSERT EQUATIONS]



The simplest version of this model makes simplifying assumptions that make analyzing the dynamics much easier. It assumes that the population, N, is constant, with no births or deaths. Represented mathematically, N = S + I + R. To further simplify this system of equations, R can be rewritten as a function of the other variables: R = N – S – I, and so the system of three equations can become two equations that only depend on two variables, S and I:

[INSERT EQUATIONS]

The SIR model has several important parameters that need to be defined. For each iteration (each week of the evolution of the disease), [beta/N] is the contact rate between susceptible and infected individuals in the population. The number of susceptible decreases at a contact rate [beta/N] and, as a result, the number of infected individuals increases at the same rate. [Gamma] is the rate at which infected individuals are cured. The number of infected individuals decreases at a recovery rate [gamma] and the number of recovered increases at this same rate.

Some other important terms that are derived from this set of equations is the average length of infectious period of the disease, 1/gamma, and the basic reproduction ratio, R0 = beta/gamma. The basic reproduction ratio is a measure for how rapidly the virus will spread and how much of the population will be affected by the virus. Intuitively, an R0 > 1 implies that the disease is spreading (outbreak) since more people are getting the disease than recovering, and an R0 < 1 implies that the disease is dying out since less people are getting the disease than recovering. The larger the magnitude of these values, the stronger the effect.

Because the discrete SIR model is a simple system of equations, one can determine the equilibria by setting the equations for S and I = 0. Linearizing about the equilibria provides information about the stability of the different equilibria (with the second eigenvalue determining the stability):

[INSERT EQUATIONS]

Continuous SIR Model

The continuous SIR model is almost identical to the discrete SIR model with the difference being that the continuous SIR model is defined on a continuous time scale instead of a discrete one. The S, I, and R state variables, however, still remain discrete. Like with the discrete model, we are assuming a simple SIR model with no births or deaths, meaning a constant total population size. The total population size, N = S(t) + I (t)+ R(t). The disease-free equilibrium occurs when S = N and when I = R = 0.



Similarly, since this is a simple system of ordinary differential equations, one can determine the equilibria and stability by setting each equation equal to 0:

[INSERT EQUATIONS]

Numerical Simulation

The continuous SIR model is a deterministic model and so a given set of initial conditions and parameters will yield distinct behavior. Numerical simulations will be computed for later stochastic models, but for this model, it is more valuable to fit the model to real-world data.

One can determine the set of initial conditions and parameters, namely [BETA] and [GAMMA], that fit the real-world data best. The dataset used to fit the continuous SIR model will be the latest influenza data form CDC form 2017-2018, which provides weekly measures of the number of positive flu specimens for each tested strain versus the number of individuals sampled. Within the context of this fit, it can be assumed that the total population, N, will be the total number of individuals sampled and that the sum of positive flu specimens for each strain will be the total number of infected individuals, I, for a given week. The best fit will be calculated by minimizing the sum of squared errors between the model and the data at each of the discrete data points:

[INSERT PLOT]

For this best fit, the best fitting values for [BETA] is 1.1837 and for [GAMMA] is 0.8691. The total sum of squared errors is 3699300. Visually, the model fits the data relatively well and the magnitude of the sum of squared errors can be attributed to the large population sizes used for the fit (a small discrepancy between the model and data will be magnified for a squared error). These values, along with the initial values of S0 = 106487 and I0 = 161, will be used for the later numerical simulations.

The deterministic and continuous SIR models provide a good understanding of the dynamics of influenza, but it will be more insightful to model the dynamics stochastically. Stochastic models will be more realistic since the spread of the virus is intrinsically probabilistic.

Continuous Time Markov Chain Model

The continuous time Markov chain model (CTMC) extends the continuous SIR model by including a stochastic process. The CTMC epidemic processes are defined on a continuous time scale, t ∈ [0, ∞), and the states S(t), I(t), and R(t) are discrete random variables, i.e., S(t), I(t), R(t) ∈ {0, 1, 2, . . ., N}. Furthermore, since R can be rewritten in terms of S and I, this model is a bivariate process with independent random variables S and I. The dynamics of the random variables S and I will be determined by transition probabilities of two events: infection and recovery.

The two possible events at a given time t: a susceptible individual becomes infected (k-1, j+1) or an infected individual recovers (k, j-1). The sum of the these transitions equals one because these transitions represent all possible changes in the state i during the time interval ∆t. Also, because the CTMC model is a Markov process, the transition probability at time tn+1 only depends on the most recent time tn. The transition probabilities are shown below:

[INSERT EQUATIONS]

These transition probabilities can be applied to all of the possible ordered pairs of states at a given time to obtain a difference equation that can be used to predict future dynamics (at time t + delta t):

[INSERT EQUATION]

Finally, a differential equation can be derived by subtracting P(i(t)) from both sides taking the limit as ∆t goes to 0. This differential equation known as the forward Kolmogorov differential equation. There is also a backward Kolmogorov differential equation that can be used to estimate the probability of reaching a given state.

[INSERT EQUATION]

An important thing to note is that I = 0 is an absorbing state for this model. This means that once the number of infectious individuals reaches 0, it continues to remain at 0 for the remaining future states, which results in a disease-free state. Logically, if there are no more infectious individuals left in a population, no additional individuals can be infected. States that are otherwise not absorbing are referred to as transient states.

Numerical Simulation

Like with other stochastic processes, it is much easier to numerically simulate stochastic processes rather than finding an analytical solution for the transition probabilities. The Gillespie algorithm will used to simulate the CMTC model, using uniform random numbers to simulate the change in state.

One random number is used to determine interevent time, which is the amount of time in between different events (the amount of time that the system remains unchanged at a given state). It will be assumed, due to it being Markov process, that the interevent time follows an exponential distribution so that: \*\*\*

The second random number is used to determine which event occurs at time t, with the random numbers subdivided into intervals corresponding with the transition probabilities of each event:

[INSERT PROBABILITIES]

Using the same parameters as determined from the continuous SIR model best fit, three different simulations were calculated:

[INSERT PLOT]

Comparing the CTMC model with the continuous SIR model, it is clear that both model the same general trends. However, unlike the deterministic model in which only one outcome is possible for a given set of parameters, the stochasticity captured in the CTMC model provides many possible simulation trajectories.

Because the CTMC model is a probabilititic model, it can be used to calculate the probability of severe outbreak. Using the branching approximation process, in which the model is linearly approximated near the disease-free equilibrium, \*\*\*. The branching approximation also assumes that each infected individual is independent from another individual with the same probabilities of infecting and recovering, and that there is a large susceptible population.

[INSERT EQUATIONS]

Using this approximation, it is possible to estimate the probabilities of two possible scenarios, a minor outbreak, in which there is a slight increase in the number of infected individuals, and a major outbreak, in which there is a substantial increase in the number of cases.

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

Stochastic Difference Equation

The final model studied in this paper is the stochastic difference equation model (SDE). This model is similar to the CTMC model except that the model follows a diffusion process in which the random variables are continuous instead of discrete. The SDE model uses the same transition probabilities as derived in the CTMC model and uses the probabilities to calculate the mean and covariance matrix, assuming that the random variables are independent and identically distribution for small enough delta t and approximate a normal distribution for large enough n (small enough time steps). This approximation is valid due to the central limit theorem.

The change in the random variables can be approximated by a system of Ito SDEs:

Each of the differential equations has a term, W, that represents independent Wiener process. Independent Wiener processes, also called Brownian motion process, is a continuous-time stochastic process in which W is a random variable with a Gaussian distribution with zero mean and unit variance.

In an applied implementation, the Wiener process can be represented by a random number generated from a normal distribution. One can observe that removing the Wiener process terms returns the same ODEs in the original continuous SIR model. Essentially, these added Wiener processes add the element of chance into the SIR model.

Numerical Simulation

Simulating SDEs can be done using the Euler-Maruyama method, which uses a finite-difference approximation:

For large population sizes, SDE numerical simulation computations perform much faster than for CMTC models since it easier to solve numerically.

Control of Influenza

As mentioned in a previous section, there are several somewhat effective methods to help control the rate of infection and spread of the influenza virus: vaccination, public health education/quarantining, and antiviral medication. In terms of the mathematical model, each of these methods affects a specific parameter in the model:

Flu vaccine: Decreases S0 by moving S0 to R0.

Quarantine/Public Health Education: Decreases the contact rate for susceptible individuals and thus decreases beta.

Antiviral (Oseltamivir/Tamiflu): Increases the rate of recovery for infected individuals by shortening the time with symptoms (decreasing the period of infectivity) and thus increases gamma.

Below are plots comparing the effect of each control method (only plotting the infected individuals) versus their control. These results are summarized in the subsequent table.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Vaccine** | **Quarantine** | **Antiviral** |
| **Continuous SIR** |  |  |  |
| **CTMC** |  |  |  |
| **SDE** |  |  |  |

Evaluation

Which is the best model for influenza? The simplest model

The stochastic models have an advantage of determining the probabilities of the different possible outcomes.

Further Analysis

Although it is beyond the scope of this paper, it would be interesting to see how varying the birth rate will influence each model. The Matlab code for each of the given models is included in the appendix and can take an input for the rate of birth, b. Changing the value of b can increase the number of susceptible individuals at any given time, which can increase the \*\*\*. Analysis of the dynamics of influenza focusing on the effect of having a non constant population, which is more realistic, can provide insightful results.

Conclusion

These different methods for modeling the dynamics of influenza, the deterministic SIR models, the CTMC model, and the SDE models, all have inherent advtanages.

The impact of vaccination, Tamiflu, and quarantining can be seen in each model

SUMMARY

Further

These stochastic models can be extended to simulate the dynamics of other infectious disease and determine probabilities for epidemics. For example, one can study malaria or \*\*\*.

The difficult part of simulating the dynamics is to first determine the governing equations and the transition probabilities that are derived from those equations.

In order to improve upon the SIR model,

There does not seem to be a long-term solution to the threat of influenza on a global health scale. Although there are no immediate cures or treatment for the flu, public health entities can continue to promote vaccination to decrease the rate of infection and spread of disease. While it may be more cost effective to shift the focus to public health education, vaccination is not a significant cost and can have a much stronger effect, but only if the predicted formula accurately covers individuals against the specific virus for that season.

Until a better solution emerges, it seems that influenza will continue to be an active concern each winter. So in the meantime, let us make sure to wash our hands frequently and get vaccinated!