The Inclusion of Asymptomatic Cases in Basic Epidemic Models
MATH 308H Research Paper

Dayton Berezoski

1 Introduction

1.1 Background

With the current world pandemic (as of November 2020) due to the SARS-CoV-2 virus (COVID-19) many people are looking to different epidemic models to gain a better understanding of the trends of past data. This is to use this data to help us predict the timeline of the COVID-19 virus. Many of these models, like the SIR model, have been around for almost a hundred years now and has been used to accurately model epidemics such as the Spanish Flu. One of the things that makes the SIR model so relevant is its ability to have modifications put on it because of how simple it is, since it divides the population into one of three categories being either susceptible, infected or recovered (with the assumption that recovered people cannot be reinfected). One modification of this model is the SIRS model which assumes that the recovered patients can eventually become suspectable and therefore be reinfected. Another model branched off of the SIR model is the SEIRD model which accounts for a latency period for people exposed to the virus and a category for the people who have died from the virus.

The unique thing about the COVID-19 virus, is that it is able to infect its host without the host showing any signs of being infected, otherwise known as being asymptomatic. This is what ultimately makes the COVID-19 virus unique, it's that people can be hosts for the virus without them knowing it which can easily spread the virus to other people making them be hosts with or without symptoms.

1.2 Goals

The goal of this research paper is to try and modify the basic SIR epidemic model to account for the asymptomatic cases, combine the three models (SIR, SIRS, and SEIRD) to be used for accounting for asymptomatic hosts and then

compare the results with the trends of COIVD-19 data, secifically the trend of having large spikes of infection over time. Finally, a conclusion can be made to see if these new models have any relevance in the correct modeling of the COVID-19 virus or if the basic SIR method is still the best method for modeling the effects of the COVID-19 virus.

2 Inclusion of Asymptotic Cases on the SIR Model

2.1 Background

The SIR model was made in 1926 by Anderson Gray McKendrick, to model the 1918 Spanish Flu which devastated the world by infecting 500 million people and killing 50 million of the infected. Back then, it was very easy to tell if you had the Spanish Flu, so McKendrick did not include a section in the model for people who were asymptomatic. The simple, yet effective model was:

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - kI$$
$$\frac{dR}{dt} = kI$$

2.2 SIIR Model

The inclusion of the asymptomatic case into the SIR model is going to require that the infected group be split into two different groups. For reference, the traditional SIR model will continue to be referred to as the SIR model and the new model that includes two different infected cases will be called the SIIR model. Logically, the infected cases will be split into asymptomatic infected (I_a) and symptomatic infected (I_s). The new model is as follows:

$$\frac{dS}{dt} = -\alpha SI_a - \beta SI_s$$
$$\frac{dI_a}{dt} = \alpha SI_a - k_1 I_a - \lambda I_a$$
$$\frac{dI_s}{dt} = \beta SI_s - k_2 I_s + \lambda I_a$$
$$\frac{dR}{dt} = k_1 I_a + k_2 I_s$$

The first notable difference between the SIR and SIIR model is the inclusion of α and β coefficients for the different infected populations in the $\frac{dS}{dt}$ portion. This is because there is a difference in the number of people who become asymptomatic rather than symptomatic, roughly about 30% of people infected are asymptomatic according to Hartford Healthcare. Thus, $\alpha < \beta$ due to the increased chance of becoming symptomatic when catching the COVID-19 virus. Next is the difference between the k_1 and k_2 values which is only due to the CDC's differentiation

of quarantine times for symptomatic and asymptomatic people, 14 and 10 days respectively. This means that people who are asymptomatic go to the recovered state faster than the people who are symptomatic, therefore $k_1 > k_2$. Finally, there is the λ, which represents the people who go from being asymptomatic in the first few days of their infection to then developing symptoms, putting them in the symptomatic case. Hartford Healthcare estimates that the percent of asymptomatic cases that this happens to is 19.1%.

2.3   Analysis

The first thing to note is the growth of the different functions. For the susceptible portion of the population, that number will always be decreasing since it's derivative is always going to be negative. Skipping to the recovered, we can also see that the number will always be growing as its derivative will always be positive. Now, the complicated part comes in with the two types of infected. For the asymptomatic case, there is a simple critical point:

$$\frac{dI_a}{dt} = \alpha S I_a - k_1 I_a - \lambda I_a \Rightarrow S = \frac{k_1 + \lambda}{\alpha}$$

So with S values greater than $\frac{k_1 + \lambda}{\alpha}$, the number of asymptotic cases is increasing and when S is less than $\frac{k_1 + \lambda}{\alpha}$ the number of asymptotic cases is decreasing. Finally, the symptomatic case provides a bit more trouble as it is slightly more complicated than the previous function. None the less, there does exist a critical point that is dependent on both the values of T_a and I_s:

$$\frac{dI_s}{dt} = \beta S I_s - k_2 I_s + \lambda I_a \Rightarrow S = \frac{k_2}{\beta} + \frac{\lambda I_a}{\beta I_s}$$

Same principle here as before, when S is greater than $\frac{k_2}{\beta} + \frac{\lambda I_a}{\beta I_s}$, then the number of symptomatic cases are increasing and when S is less than $\frac{k_2}{\beta} + \frac{\lambda I_a}{\beta I_s}$, it is decreasing.

These two critical points resemble the inverse contact number $\left(\frac{k}{\beta}\right)$ from the SIR model where the first term of the two cases is the same as it is either $\frac{k_1}{\alpha}$ or $\frac{k_2}{\beta}$ (the β's between the SIR and SIIR models are different representations). From this we can see that the λ term in the original two differential equations makes the difference in critical points when comparing the SIR and SIIR models.
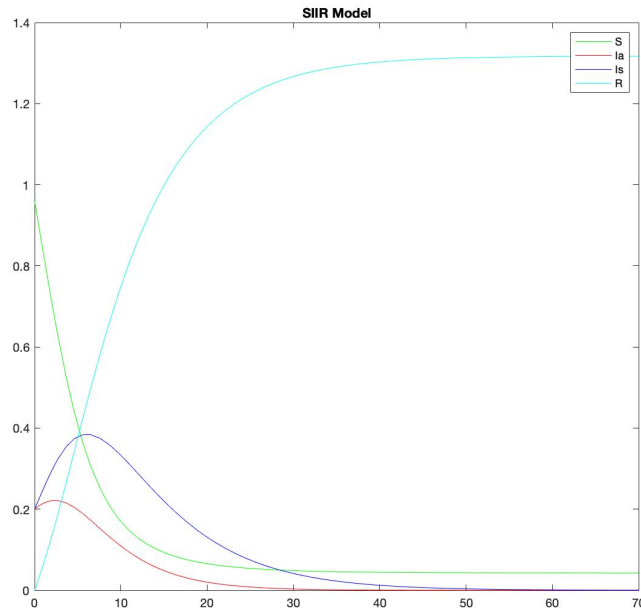
2.4   Example Graph

From the previous section, we can remember condense the relationship of some of the constants being:

$$\beta > \alpha$$
$$k_1 > k_2$$

Following these constraints, we can make a plot of the data using these values:

$$\beta = 0.35$$
$$\alpha = 0.30,\ \lambda = 0.05$$
$$k_1 = 0.15,\ k_2 = 0.14$$
$$S_0 = 0.96,\ I_{a0} = 0.2$$
$$l_{s0} = 0.2,\ R_0 = 0$$

Resulting in the plot:

From the graph we can see that both of the infected cases grow slightly before dying out. While this does adhere to the Fundamental Theorem of Epidemiology stating that as $t \to \infty$, the number of infected and susceptible will go to zero, it does not accurately model the COVID-19 data that has been collected so far. The data collected so far shows that there are many peaks in infected populations, but in here there is only one peak for both infected cases, corresponding to the aforementioned critical points. This can be attributed to the fact that the model we are currently using does not account for the fact that recovered patients can be reinfected again (the CDC estimates that it takes three months for this to happen) and is a shortcoming of this particular model.

Of course, as with all mathematical models, there are many different trends with the data depending on the constants and initial conditions given, but the

provided data was hand-picked to simulate natural trends and not the outliers that can come from these models.

3 Inclusion of Asymptotic Cases on the SEIRDS Model

3.1 Background

The SEIRDS model is a combination of the SIRS and SEIRD models (and therefore build off of the SIR model) and contains many more cases than the previous SIR model. This model divides people into susceptible, exposed, infected, recovered, and dead. The model is as follows:

$$\frac{dS}{dt} = -\beta SI + \eta R$$
$$\frac{dE}{dt} = \beta SI - \sigma E$$
$$\frac{dI}{dt} = \sigma E - (1 - \alpha)kI - \alpha\rho I$$
$$\frac{dR}{dt} = (1 - \alpha)kI - \eta R$$
$$\frac{dD}{dt} = \alpha\rho I$$

3.2 SEIIRDS Model

The inclusion of the asymptomatic case in the SEIRDS model is going to require two separate cases for the infected population, I_a and I_s like before. This new model will be referred to as the SEIIRDS model, while the previous one will continue to be called the SEIRDS model. The new model is as follows:

$$\frac{dS}{dt} = -\alpha SI_a - \beta SI_s + \eta R$$
$$\frac{dE}{dt} = \alpha SI_a + \beta SI_s - (\sigma_1 + \sigma_2)E$$
$$\frac{dI_a}{dt} = \sigma_1 E - (1 - \tau)k_1 I_a - \tau\rho I_a - \lambda I_a$$
$$\frac{dI_s}{dt} = \sigma_2 E - (1 - \tau)k_2 I_s - \tau\rho I_s + \lambda I_a$$
$$\frac{dR}{dt} = (1 - \tau)(k_1 I_a + k_2 I_s) - \eta R$$
$$\frac{dD}{dt} = \tau\rho(I_a + I_s)$$

The first thing to notice is the replacement of α from the SEIRDS model to τ in the SEIIRDS model, which is due to the fact that we already use α in the SIIR model and want to keep the two consistent. Like with the SIIR model, the susceptible function has the $-\beta SI$ term split into two to account for the different types of infected. However, the α and β coefficients are different here than they were in the SIR model. Here the α and β represent the chance of being in contact with an asymptomatic and symptomatic person, respectively which we can call the contact chance. Since the asymptomatic person is generally unknown to whether they have the virus or not, they will have a higher contact chance as most of the symptomatic hosts will be quarantining away from others. Therefore, $\alpha > \beta$. We

will see later that the α and β from the SIR model where it is the infected being either asymptomatic or symptomatic is replaced by σ_1 and σ_2. That function also contains ηR which is to account for the people who have recovered being able to be reinfected. η should be relatively large, as the CDC estimates that people who have recovered only have immunity from the COVID-19 virus for about three weeks, meaning that everyone in recovered should return to susceptible within a few months.

The exposed function is there to account for the people who have been exposed, but cannot determine if they have the virus or not. One reason why this is so important to the COVID-19 virus, is because this specific virus can lay dormant in the for up to 14 days with the median being 4-5 days, known as the incubation period. The first two terms are the odds of susceptible people being exposed to an infected person, while the last terms are the exposed going to the infected cases. The σ_1 and σ_2 represents the people who do get infected and whether or not they are symptomatic. This is the same thing as the α and β from the SIIR model where the infected had to be split into two different cases. Therefore, $\sigma_2 > \sigma_1$

The two infected cases are similar, so they will be described generally with anything specific to one case being explicitly mentioned. The first term is just putting the exposed people that were infected into the right function. The second term contains $(1 - \tau)$, where τ is the death coefficient, therefore $(1 - \tau)$ represents the recovered people. The kI part of that term is the same as the SIR model where the respective k is how long it takes for a person to recover from the virus, and the $(1 - \tau)$ part is just to split the infected into recovery and death. Logically the next term $-\tau \rho I$, is the people who do not recover from the virus and die because of it. Again τ is there to split the infected into recovered or dead and here ρ is how long it takes for someone to dies from the time they are infected with the virus. As a loose definition of the constants, the CDC argues that the overall death rate of the virus is around 0.2% with some experts arguing that it is higher at 0.5%. Of course, this will vary by age and other factors, but the underlying conclusion is that the death rate is relatively low, so τ will be too. k_1 and k_2's relationship will be reused from the SIIR model where $k_1 > k_2$. Finally, ρ is how long it takes for someone to dies from the COVID-19 virus, which from a study by the Lancet takes a median of 18 days. Thus $k_1 > k_2 > \rho$, because the rate at which people die from COVID (going from infected cases to dead case) is lower than the average rate of recovery.

The recovered and death cases' variables and coefficients have already been explained by the previous cases, in these cases the signs are opposite.

3.3 Analysis

The growth of the different functions in the SEIIRDS model are much more complicated than in the SIIR model, but I will try my best to give a reasonable conclusion on all of the functions.

The most simplistic growth is the death case which is always growing, since

$$\tau\rho(I_a + I_s) > 0, \forall t(t \geq 0)$$

The growth of the rest of the functions can be represented by it's critical point:

$$\frac{dS}{dt} = -\alpha S I_a - \beta S I_s + \eta R \Rightarrow R = \frac{\alpha S I_a + \beta S I_s}{\eta}$$
$$\frac{dE}{dt} = \alpha S I_a + \beta S I_s - (\sigma_1 + \sigma_2)E \Rightarrow E = \frac{\alpha S I_a + \beta S I_s}{\sigma_1 + \sigma_2}$$
$$\frac{dI_a}{dt} = \sigma_1 E - (1-\tau)k_1 I_a - \tau\rho I_a - \lambda I_a \Rightarrow E = \frac{(1-\tau)k_1 I_a + \tau\rho I_a + \lambda I_a}{\sigma_1}$$
$$\frac{dI_s}{dt} = \sigma_2 E - (1-\tau)k_2 I_s - \tau\rho I_s + \lambda I_a \Rightarrow E = \frac{(1-\tau)k_2 I_s + \tau\rho I_s - \lambda I_a}{\sigma_2}$$
$$\frac{dR}{dt} = (1-\tau)(k_1 I_a + k_2 I_s) - \eta R \Rightarrow R = \frac{(1-\tau)(k_1 I_a + k_2 I_s)}{\eta}$$

As a generalized summary, whenever the R or E on the left is greater than the expression on the right, the function is growing and whenever it is less the function is decreasing. It is important to notice that the unlike the SIR model's critical points there is the ability for multiple critical points due to the fact that the recovered can go back to being susceptible and repeating the cycle again. This behavior models the COIVD-19 virus' pattern of having many different peaks in the number of people it has infected.
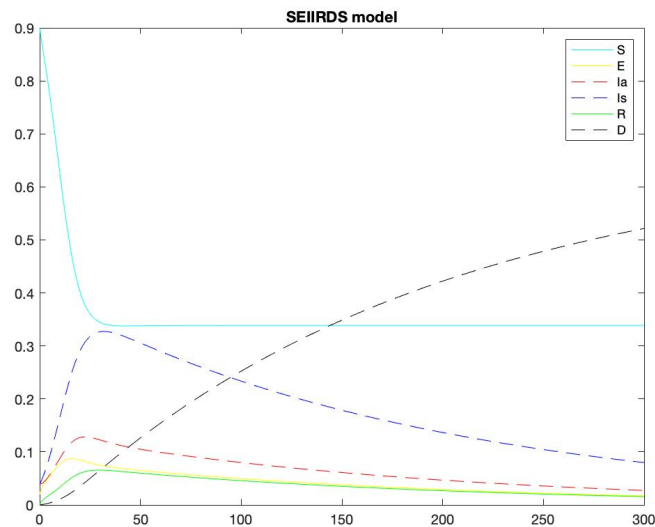
3.4 Example Graph

From the previous section, we can remember condense the relationship of some of the constants being:

$$\alpha > \beta$$
$$\sigma_2 > \sigma_1$$
$$k_1 > k_2 > \rho$$

Following these constraints, we can make a plot of the data using these values:

$$\alpha = 0.6, \ \beta = 0.2, \ k_1 = 0.15, \ k_2 = 0.1$$
$$\eta = 0.7, \ \tau = 0.1, \ \lambda = 0.05, \ \sigma 1 = 0.3$$
$$\sigma 2 = 0.35, \ \rho = 0.07, \ S_0 = 0.92, \ E_0 = 0.02$$
$$I_{a0} = 0.04, \ I_{s0} = 0.04, \ R_0 = 0, \ D_0 = 0$$
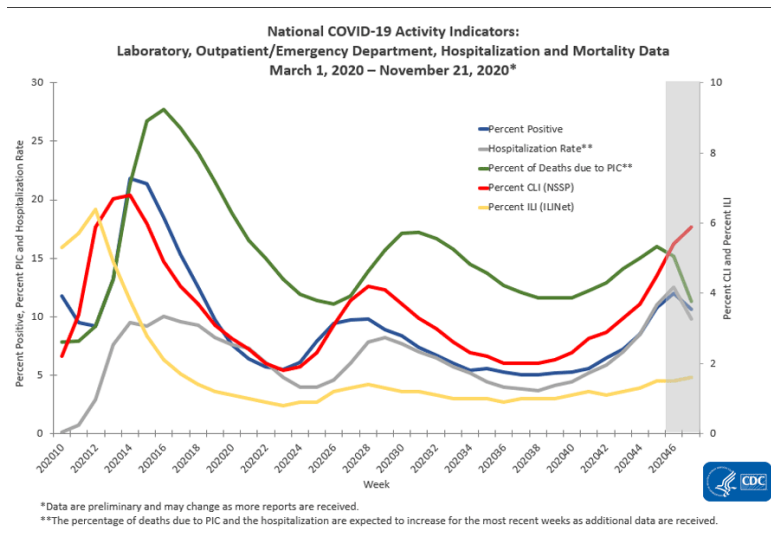
Resulting in the plot:

From the graph we can see that again, the number of infected people goes down to zero as time goes on, adhering to the Fundamental Theorem of Epidemiology. We can also see that the number of susceptibles is slightly increasing after it hits its low point due to the recovered going back to susceptibles, but it is not growing that fast due to the number of susceptibles being exposed and the death toll taking people out of the cycle entirely. This is not the exact results that I wanted where the graph simulated the spikes in the COVID-19 infection numbers, but that does not mean that it is impossible to do so with this model. The extreme complexity of the model makes it a challenge to get the exact results wanted, but it is possible to get those results with this graph, however the way in obtaining those results is as arbitrary as picking random values that adhere to our rules until we get what we want.

4 Conclusion

4.1 Results and Findings

To summarize, the asymptomatic case was added to the SIR and SEIRDS models in an attempt to simulate model the trend lines of the COVID-19 virus as seen below. While the new SIIR and SEIIRDS models did seem a bit more specific

National COVID-19 Activity Indicators:
Laboratory, Outpatient/Emergency Department, Hospitalization and Mortality Data
March 1, 2020 – November 21, 2020*

*Data are preliminary and may change as more reports are received.
**The percentage of deaths due to PIC and the hospitalization are expected to increase for the most recent weeks as additional data are received.

to what we were looking for, the overall trends could have been modeled with the standard SIR and SEIRDS models. However, this does not mean the models have no use, as on small scale levels the models could see more accuracy and if the disparity between the asymptomatic and symptomatic cases increase, there will be more uses for this model. The biggest problem so far, is that the difference in the asymptomatic and symptomatic cases are currently too small to be used on such a large scale, as their growths and declines seem to mirror one another.

4.2   Future Research

Future research and directions these models could go into is the refinement of the coefficients. In this report the coefficients were chosen arbitrarily based on the relationships they have with each other and nothing more. While in actuality the numbers are very complex, as someone going from exposed to infected depends on things such as what masks they were using, length of exposure, underlying health conditions, distance from the infected and etc. Finding better numbers to use as coefficients could lead to more accurate results based off of these new models.

Another topic to consider is the different demographics that the virus effects. The data used to come up with these relations were based on very general ideas surrounding the COVID-19 virus, while the demographics end up making a huge difference in data. For example, the death rate for elderly people who catch COVID-19 is much higher than the death rate of teenagers who catch COVID-19. This larger difference in results also could have contributed to less than ideal results in this report and should be considered if this idea is furthered.

9

4.3 Final Remarks

While the new models do not seem to show too much increase in accuracy against the old models, it does modify the models to deal with COVID-19. COVID-19 is a unique virus in the fact that it does not always make its hosts show symptoms and that fact alone makes it hard to wipe out. While the example graphs did not seem to present too many new ideas, that could be chalked up to the coefficient values based only on relations. While these models might look rough at the moment, they could be built upon to further credit predictions of COVID-19 in the future.

5 Works Cited

https://www.cdc.gov/flu/pandemic-resources/1918-pandemic-h1n1.html

https://hartfordhealthcare.org/about-us/news-press/news-detail?articleId=29806&publicid=743

https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html

https://www.usatoday.com/story/news/factcheck/2020/06/05/fact-check-cdc-estimates-covid-19-death-rate-0-26/5269331002/

https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html

https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30566-3/fulltext