

Rachael Ren

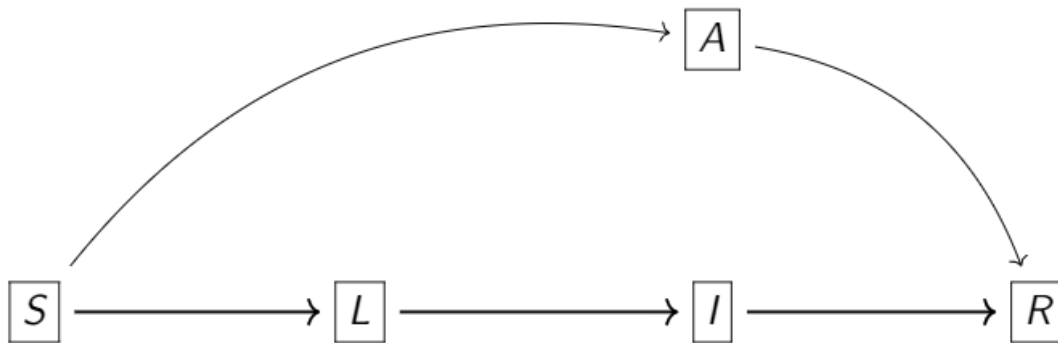
Spring 2020

Disease Modeling

This quarter, I studied SIR modeling and built a model for COVID-19 under the mentorship of Anna Neufeld. Our project was titled “Disease Modeling”, but we decided to focus on COVID-19 because of our mutual interest in the ongoing pandemic.

For the first few weeks of the quarter, I read from a textbook titled *Epidemics: Models and Data in R*. I was not familiar with SIR modeling before this quarter, but the textbook helped me gain a general understanding of the model and the math and basic epidemiology behind it. Anna and I also did online research on COVID-19 to study the characteristics of the virus and pandemic.

After reading and researching for a few weeks, we started to build our own model for COVID-19. We modified the SIR model structure and differential equations to reflect the characteristics of the virus and pandemic. Major modifications we made included adding a subpopulation for latent (L) and asymptomatic (A) individuals since they play a significant role in spreading the virus.



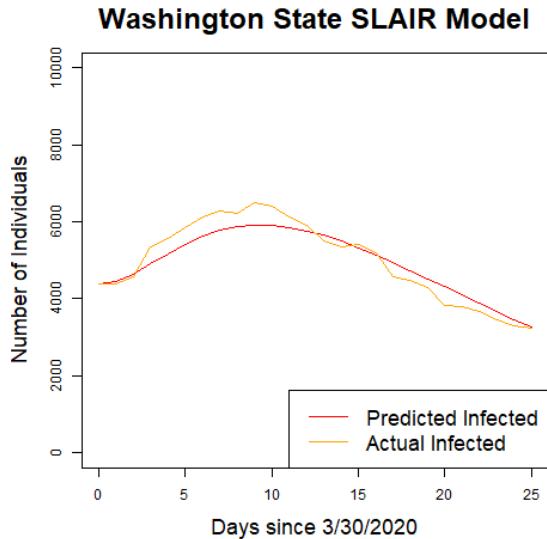
*Our modified SIR model structure. Each box represents a subpopulation. We assume that every individual in our population is in one of these subpopulations. The arrows represent the movement between the subpopulations, which are determined by differential equations.*

Once we set up our modified model and its differential equations, our next goal was to find the value of unknown parameters. To accomplish this, I modified the SIR model R code from the textbook to fit our model and attempted to fit our predicted infected (I) curve to the actual Washington State infected curve. I used the `expand.grid()` command to generate sets of plausible parameter values and ran them through our model to determine which set minimized the least-squares error. After making some adjustments to our model, we were able to produce the following best-fit line:

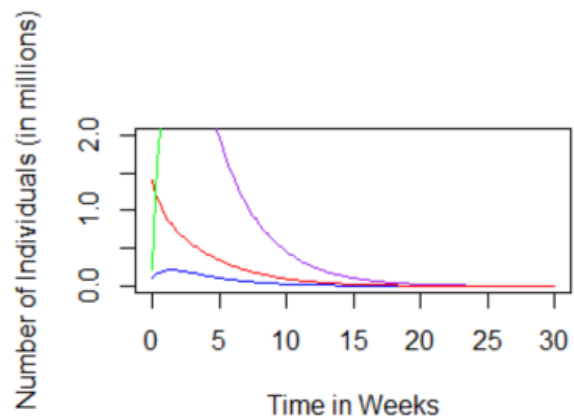
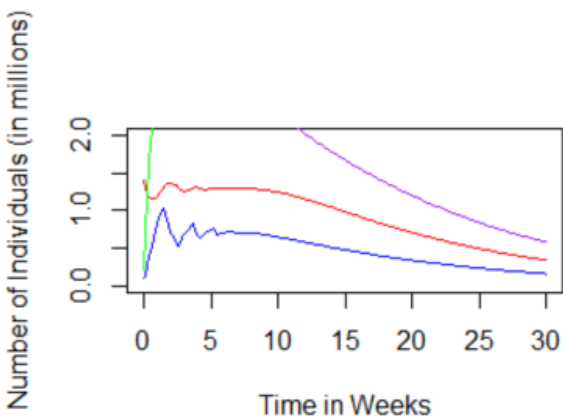
Rachael Ren

Spring 2020

Disease Modeling



One application I wanted to explore using our model was the effect of social distancing. I simulated turning on and off social distancing by changing the values of parameters based on a threshold of the ratio of infected to susceptible individuals.



Both of these graphs use the US population. The red curves represent the predicted infected curve. For the graph on the left, the threshold  $13/3282$  was used, which was the ratio of infected to susceptible individuals at the time at which these graphs were generated. For the graph on the right, the threshold  $1/1000000$  was used, which is the IHME recommended threshold for turning on and off social distancing.

Although we were not able to get an in-person DRP experience this quarter, I still thoroughly enjoyed my DRP project and learning from Anna Neufeld. Anna was extremely accommodating to my schedule and was receptive to my ideas and interests. I would highly recommend the DRP to anyone who is interested in learning more about a specific statistics topic and applying it to real-world problems.