COVID-19 Version 1.2 Student Notes

**Objective:** The module objective is to use a “sandbox” spreadsheet to access a preprogrammed model of the COVID-19 pandemic to explore how projections of the epidemic’s progress depend on the values of key disease parameters and parameters that represent the combination of public policy decisions and societal behavior.

**Overview**

Rather than focusing on the complicated model, this spreadsheet activity takes the model as given in order to concentrate on the most important issues involved in addressing the pandemic. The spreadsheet file contains five sheets. Three of these contain the code to run simulations with the model. The user should simply leave these sheets as is, although they can be referenced for more details on the outcomes.

The top sheet is a “sandbox,” meaning that it is a simple environment for exploration. This is the “model user interface.” The user provides data (cells highlighted in yellow) for three model scenarios. The results are automatically computed, with five key outcomes (cells highlighted in blue) and four graphs. Varying just one parameter in the three scenarios provides an experiment to identify the effect of that parameter. A separate file contains a list of several experiments and questions to address for each. Data for these experiments appears in the workbook as a separate sheet to facilitate experiment setup.

**Preliminary Discussion**

This module uses an October 2020 revision of a model created by Glenn Ledder (University of Nebraska-Lincoln Department of Mathematics) in March 2020 to try to capture the population dynamics of the COVID-19 pandemic, using what facts were known at the time and current best estimates for quantitative aspects of the pandemic. These facts have changed as more knowledge has become available, necessitating improvements in the model and parameter estimates. The results of sandbox experiments have to be judged in this light. To paraphrase a quotation by Albert Einstein: No one believes a model except the person who created it; everyone believes data except the person who collected it. We should not `believe’ either the model or the data that is used in it. The results of any investigation are an accurate portrayal of the behavior of the *model*, which is a still a useful thing. A good model can provide a reasonable approximation of the behavior of the real-world phenomena being modeled, but the results should be compared to observation and data whenever possible. When we find discrepancies between model results and real-world behavior, we have an opportunity to improve the model. If a model shows good qualitative agreement with a real scenario, then we can use it to make qualitative judgments of the effects of public health policies and societal behavior.

**Epidemic Models**

The basic idea of epidemic models is to track changes in the numbers of people in a variety of classes by using formulas for the processes that move people among the classes. The best known epidemic model is the SIR model, where the letters stand for *susceptible*, consisting of those who are at risk of catching the disease; *infectious*, consisting of those who can pass the disease to others; and *removed*, consisting of all individuals who are neither at risk nor infectious. Our COVID-19 model is similar, but it has more classes: SEAIHRD, where S is as in SIR, E is *exposed*, for individuals who have been infected but cannot yet pass the disease to others (*latent* is a better term), A and I are *asymptomatic* and symptomatic *infectious*, H is *hospitalized*, R is *recovered*, and D is *deceased*.

**Parameters**

In addition to equations for calculating the state variables, models require *parameters*, which are quantities that are fixed in a given scenario but can vary among different scenarios. The model has 16 parameters, which is too many to study all at once. Columns A through C in the sandbox sheet contain 11 parameters that are not varied in the three scenarios: these are measures of the fraction of infectives who are asymptomatic, the fraction of symptomatic infectives who require hospitalization, the fraction of hospitalized patients who die, the relative infectivities of asymptomatics, isolated symptomatics, and hospitalized patients as compared to undiagnosed symptomatics, the incubation period, the average recovery times for different classes, and the initial number of symptomatic infectives out of a base population of one million. Most of these parameters play only a small role in determining the outcomes.

The other 5 parameters are specified for each scenario so that their effects can be studied individually.

1. *δ* is a contact factor that represents the effect of physical distancing. A value of 1 means that there is no physical distancing, and a value of 0.5 means that contact rates for individuals (not isolated ones) are half of normal.
2. *pC* is the fraction of symptomatic individuals who are confirmed cases. Roughly speaking, this parameter is a measure of how much testing is being conducted. A value of 0.1 means that only 10% of symptomatic patients are tested and diagnosed.
3. *pCA* is the fraction of asymptomatic individuals who are isolated (*confined*). This parameter serves as a measure of contact tracing, since that is the only reasonable way to find asymptomatic carriers.
4. *T2* is the initial infection doubling time, which is the best way to measure the rate of spread. Values of 3 to 5 days have been reported.
5. *v* is the fraction of the population that is initially immune. This is 0 for scenarios that begin at first exposure of a community. Including this parameter allows the model to be used to investigate herd immunity, which is the reduction of the spreading power of the disease caused by having a relatively large fraction of the population consist of people who cannot get the disease because they have acquired immunity or been vaccinated.

**Outcomes**

The model simulations generate a large amount of data. In order to understand the model predictions, we need to identify a small number of outcomes to report. For example, rather than reporting the requirements for ventilators, ICU beds, and total hospital beds, just one of these quantities can serve to represent the stress on the health care system. The outcomes that appear on the sandbox sheet are

1. The percentage of people still susceptible at the end of the scenario. This is important because it measures the overall risk to society if physical distancing were to end.
2. The population that dies of COVID-19 during the scenario (which we can express as a total number for the United States and/or as a percentage of the population).
3. The maximum number of patients who require hospitalization per hundred thousand people. This is important because the most immediate threat of a serious illness is the capacity of the health care system to treat all patients who need professional care. For reference, cell C3 contains an average for the number of hospital beds per hundred thousand people in the United States. Of course this number varies by locality, so the model predicts dire consequences if the number of patients requiring hospitalization comes close to the number of hospital beds.
4. The maximum number of new infections. This quantity can be compared to real data.

The sandbox also includes graphs of the number of hospitalizations per hundred thousand people, the number of new infections per hundred thousand people, the number of thousands of US deaths, and the percentage of people who are still susceptible. These graphs give a visual representation of the epidemic progress and the differences among the three experiment scenarios.