CSE 8803 EPI: Data Science for Epidemiology, Fall 2020

Lecturer: B. Aditya Prakash
Scribe: Samantha Markley

Lecture # 3
August 31, 2020

---

1 Summary of Lecture Content

This lecture covers foundational material on the use of models in epidemiology.

The lecture began by covering important concepts regarding models in general, such as how to effectively review models and conflicting properties that arise, so that their applications and limitations in an epidemiological setting can be better understood.

After this overview, we moved into discussing specific models used in epidemiology, beginning with Bernoulli's smallpox model. This model is a simple differential equation model that Bernoulli implemented to weigh the risks and benefits of variolation for smallpox. This model introduces how models are used in epidemiology in a simplistic yet effective way. Next, we began discussing modern epidemiology models.

The most fundamental model used in epidemiology today is the SIR model, which was explained in depth. From this model, we then discussed several extensions and variations. These models are all considered compartmental models and can be modified extensively to fit the particular parameters of the disease and situation being studied.

This lecture concluded with a discussion of the Threshold Phenomenon, which is an important concept in understanding models and how infectious diseases spread throughout a population.

2 Models: General Information and Implications

Models, in an epidemiological setting, can be useful for several purposes including abstracting meaning from noisy or limited data, predicting future events, and guiding decision making and intervention. When reviewing model results or designing a model, there are several important questions to keep in mind in order to determine what exactly the model is telling [4]. These include the following:

- The purpose or time-frame of the model
- The modeling assumptions used
- How uncertainty is computed
- What data the model is fit to
- If the model is general or designed for a specific situation

All of these questions are especially important when considering epidemiology models, as they are being used in real time to make policy decisions. Furthermore, it is also important to consider the conflicting properties of models, like accuracy, transparency, and flexibility, when applying them to real-life situations in order to avoid mismatches and erroneous applications of a model.

The focus for the lecture today is mechanistic models.

3 Bernoulli's Smallpox model

Daniel Bernoulli's smallpox model is one of the oldest cases of a mechanistic model. He was particularly interested in smallpox because it was one of the leading causes of death at the time and variolation, although well known, was not common in France, so he wanted to compare the benefits of variolation versus the immediate risk of dying.

His model began by dividing the population into a groups of those who are susceptible and those who have immunity under the assumption that an infection either causes death or life-long immunity. Resulting in the simple equation:

$$l(t) = x(t) + z(t) \tag{1}$$

Where t is any age group, *l(t)* is the probability of survival at age t, *x(t)* is the probability of never getting infected, and *z(t)* is the probability of obtaining immunity after infection. It is also assumed, for simplification purposes, that the probability of someone within *x(t)* getting infected is a constant, b. From this, we can compose the following differential equations and solution:

$$\frac{dx}{dt} = -b * x(t) \tag{2}$$

$$\frac{dw}{dt} = (1 - a) * b * x(t) \tag{3}$$

$$x(t) = \frac{w(t)}{(1 - a)e^{bt} + a} \tag{4}$$

To apply this solution, Bernulli looked at the data to estimate that parameters a and b were both 0.125. These values were then plugged into the equation and allowed him to make the conclusions that, while not considering the risk of variolation, if all children are variolated at birth, the population will be 14% larger at age 26 and life expectancy will increase 3.17 years. Furthermore, with the inclusion of the risk of variolation, the life expectancy gain would only be reduced by one month and secondary infection effects would be minimal.

Although the benefit of variolation appears to be clear, there is still a risk, such that the state of the population as a whole will improve at the risk of the individual.

4 SIR ODE Model

4.1 Basic Description

The SIR model is one of the most simple models used in epidemiology today and is an example of a compartmental model, which divides the population into different "compartments" that interact with each other. SIR, specifically, begins by having the population divided into susceptible (S), infected (I), and recovered (R) groups and assumes that any infected person can infect any susceptible person, the total population remains constant, and that it is deterministic. This model includes the parameters β and δ which are rate of infection and rate of cure, respectively. The model is as follows:

$$\frac{dI}{dt} = \beta SI - \delta I \tag{5}$$

$$\frac{dR}{dt} = \delta I \tag{6}$$

$$\frac{dS}{dt} = -\beta SI \tag{7}$$

Equation 5 represents the rate of change in the number of infections by subtracting the number of infected people who have been cured from the number of new infections, which is obtained by determining the fraction of successful attacks from all attacks that occur. Equation 6 represents the change in the number of people entering the recovered compartment. Finally, equation 7 computes the rate of change of people who are susceptible that become infected. The sum of all of these equations results in zero, which matches the assumption that the population remains constant.

4.2 Force of Infection

The force of infection is also helpful in understanding the kinds of transmission that can occur in these models. The equation for force of infection, with the parameter λ representing the number of infected people, is:

$$F = \lambda S \tag{8}$$

There are generally two types of transmission of infectious diseases. The first is mass action transmission, $\lambda = \beta I$, which simply shows that each newly infected person successfully infects a particular number of other people. The other form in density-dependent transmission, $\lambda = \beta I/N$. This transmission is similar to mass action but suggests that an increase in density will cause an increase in the transmission.

4.3 SIR Solution

Solving the SIR model results in no closed form solution but the functions can be computed numerically and graphed to understand the general shape of the curves, shown in Figure 1, such that *S(t)* decreases as more people become infected and move out of the susceptible group. Likewise, *R(t)* increases as every infected person eventually recovers, and, finally, I*(t)* is unimodal, as it represents the total number of infections at any given time. It is also important to note that it is not necessary for every susceptible person to become infected, so

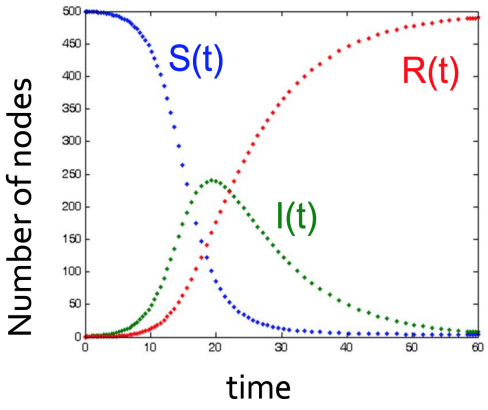

Figure 1: Numerical solution of SIR model [2]

there will remain a susceptible population as the graph approaches infinity. To investigate how different starting parameters (λ, β, S, I, and R) can change the solution to the graph, visit Epirecipes, which provides an interactive Python notebook designed specifically for the SIR model [2].

5 Extensions and Variations of SIR

Compartmental models follow the basic scheme pictured in Figure 2. By using these compartments, a variety of different combinations and can also be extended to account for different constraints such as birth and death rates, changing contact rates, making things stochastic, multiple diseases, etc. The framework of these models is very extensive and can be modified to fit many different complexities. The main driver is the S compartment moving to the I compartment, since it is an infectious disease and this movement models the spread.

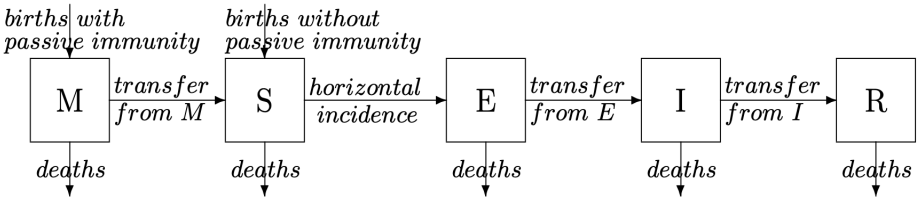

Figure 2: Compartment model scheme [3]

5.1 SIS Model (Susceptible and Infected Compartments)

The SIS model is a typical model used for endemics. It follows the same perfecting mixing assumption as the SIR model but differs from the SIR model since the disease does not die out, instead the infected and susceptible populations remain near constant after initial introduction of the disease to the population. The model is as follows:

$$\frac{dS}{dt} = -\beta SI + \delta I \tag{9}$$

$$\frac{dI}{dt} = \beta SI - \delta I \tag{10}$$

The decrease in the susceptible population, represented in equation 9, is the same as in SIR, but the group that is cured after infection is added back into the population since they do not acquire immunity. The change in the infected population, equation 10, is simply the opposite of the change in the susceptible population. Again, the sum of these equations results in a total rate of change of zero, so the population is remaining constant, and there is no closed form solution but you can do numerical computations to achieve the graph shown in Figure 3.

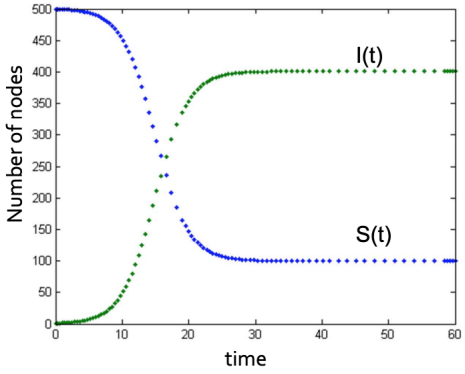

Figure 3: Numerical solution of SIS model [2]

5.2 SIR with Birth and Death Rates

In this model, the population is fixed but there are people dying, not due to disease, and being born at rate μ. In this example, the birth and death rates are assumed to be the same but this is not necessary.

$$\frac{dS}{dt} = -\beta SI + \mu(I + R) \tag{11}$$

$$\frac{dI}{dt} = \beta SI - \gamma I + \mu I \tag{12}$$

$$\frac{dR}{dt} = \gamma I + \mu R \tag{13}$$

Equation 11 shows what makes it infectious, such that I is infecting S with the addition of people becoming susceptible due to birth. Equation 12 shows inflow from the S compartment and outflow to the R compartment and some people that are just dying, and equation 13 simply shows the inflow to R.

5.3 SEIR Model (Susceptiable, Exposed, Infected, and Recovered Compartments)

This model includes the exposed state and is currently being used widely for COVID and influenza. The exposed state includes those who have been exposed to the disease but are not yet exhibiting symptoms or contagious. The length of the exposed state can vary greatly and is currently a major issue in controlling COVID. The model is as follows:

$$\frac{dS}{dt} = -\beta SI \tag{14}$$

$$\frac{dE}{dt} = \beta SI - \sigma E \tag{15}$$

$$\frac{dI}{di} = \sigma E - \gamma I \tag{16}$$

$$\frac{dR}{dt} = \gamma I \tag{17}$$

6 Threshold Phenomenon

The implicit solution for SIR can be found by setting up a system of equations that depend on the value of R_0, which is known as the reproductive number. This represents the average number of secondary cases caused by one individual.

$$R_0 = \frac{N\beta}{\delta} = \frac{\beta}{\delta} \tag{18}$$

The curing rate is δ, so the average time for someone in I to go into R is $\frac{1}{\delta}$, and every time you are in the I compartment, you are attacking others at the rate β.

To investigate the threshold phenomenon we need to rewrite the equation for $\frac{dI}{dt}$.

$$\frac{dI}{dt} = \beta SI - \delta I = I(\beta S - \delta) \tag{19}$$

This implies that $\frac{dI}{dt}$ will be less than zero if βS is less than δ, which means *S(0)* is less than $\frac{\beta}{\delta}$. If $\frac{dI}{dt}$ is less than zero, this means that the rate of change in infections per unit time is negative, so the disease will die out.

If R_0 is less than 1, then *S(0)* is less than $\frac{\delta}{\beta}$ and the epidemic dies out. On the other hand, the number of infections R_∞ will only be large if R_0 is greater than 1. Reducing the number of susceptible people to below $\frac{1}{R_0}$, you are able to reduce the epidemic. A high R_0 value will cause the epidemic to reach a higher peak number infection earlier in time in comparison to a lower R_0 value. This is the basic concept of "flattening the curve." It is also important to note that reducing R_0 can not only change the peak, but it can also change the area under the curve, or the total number of infections. The impact of these variations in R_0 can be seen in Figure 3.

The R_0 value varies greatly across different diseases and plays a role in vaccinations and understanding how diseases spread. However, the estimation of R_0 takes time to determine through accumulating data and can vary across different populations.

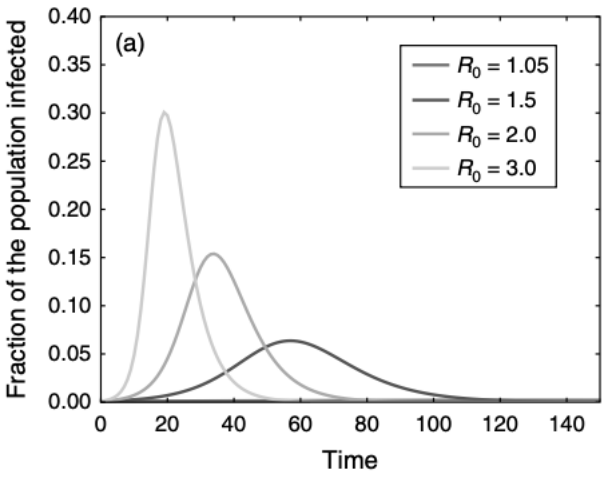

Figure 4: Impact of R_0 on the magnitude of an epidemic [1]

References

[1] N. B. Dimitrov and L. A. Meyers. Mathematical approaches to infectious disease prediction and control. In *Risk and optimization in an uncertain world*, pages 1–25. INFORMS, 2010.

[2] S. Frost. Epirecipes: Sir model, python using scipy. http://epirecip.es/epicookbook/chapters/sir/python, 2018.

[3] H. W. Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000.

[4] I. Holmdahl and C. Buckee. Wrong but useful — what covid-19 epidemiologic models can and cannot tell us. *New England Journal of Medicine*, 383(4):303–305, 2020.