Instructions:

- Please turn in a single PDF file.
- Please share a link to your code, but do not attach the actual code:
- By the way, if you don't have a github, why not make one? It's easy to do, and a great place to stash your code and link to it. Pop by office hours or phone a friend if you need help. It's never too early (or late) to learn!
- Handwritten math (scanned and included in a PDF) is fine, but please avoid 10MB+ file sizes! Reduce your image quality as needed.
- Don't forget to list anyone you collaborated with: **Altaf Barelvi and Violet Ross**.

1. The goal of this problem is to get you over any barriers with (i) getting Python set up, (ii), getting the SIR model implemented in a Forward Euler solver, and (iii) getting matplotlib set up.

    Write a function in Python that uses the Forward Euler method to simulate the SIR model *in which the population is slowly growing*.[1]

    In your simulations, suppose that we have an initial population of size $N = 1000$, with $I_0 = 1$, $S_0 = 999$. Suppose that $\beta = 1$, $\gamma = 0.5$, and then $\mu_{\text{birth}} = 0.01$ and $\mu_{\text{death}} = \frac{1}{2}\mu_{\text{birth}}$.

    Produce a single plot that captures the dynamics of transmission over the appropriate amount of time for the population to grow by 50% to a total of $N = 1500$. Be sure to include a legend and your name in the title of the plot.

    The amount of time required for the population to grow by 50% is given by

    $$1000 \cdot (1 + \mu_{birth} - \mu_{death})^t = 1500$$
    $$(1 + 0.01 - 0.005)^t = 1.5$$
    $$t = \frac{\log(1.5)}{\log(1.005)} \approx 81.3 \text{ time units}.$$

    The dynamics of the SIR model with birth and death over this time period are shown in the plot below.

---

[1]Not to be turned in, but consider checking your work by first reproducing the three plots from Figure 1 of the Week 2 lecture notes in the absence of population growth. The parameters are: $N = 1000$, $I_0 = 1$, $S_0 = 999$, with $\beta = 1$, $\gamma = 0.5$, then $\beta = 1.5$, $\gamma = 0.5$, and finally $\beta = 2$, $\gamma = 0.5$
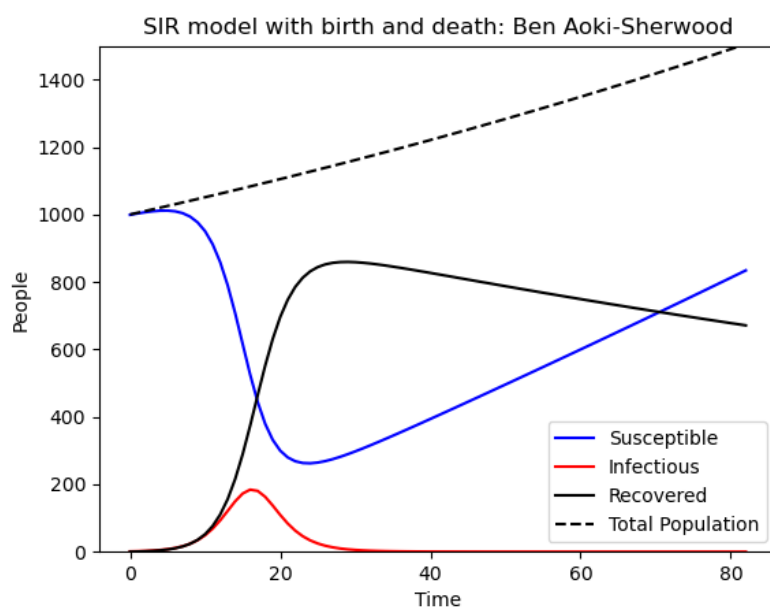
Figure 1: SIR dynamics with population growth. Parameters are $N = 1000$, $I_0 = 1$, $S_0 = 999$, $\beta = 1$, $\gamma = 0.5$, $\mu_{\text{birth}} = 0.01$ and $\mu_{\text{death}} = \frac{1}{2}\mu_{\text{birth}}$. The population grows to $N = 1500$ over the course of the simulation.

2. The goal of this problem is to show an important fact about transition rates in compartmental models. It is also a good chance to become refreshed on simple ODE solving and separation of variables. Finally, it makes good on a promise made in lecture notes to ask this homework question!

   Imagine that we are interested in SIR dynamics, but everyone starts out either infected or recovered, and no one starts out susceptible.

   a. Use this information to simplify the typical equation for $\dot{I}$.

      The standard SIR ODEs are

      $$\dot{S} = \frac{-\beta SI}{N}$$
      $$\dot{I} = \frac{\beta SI}{N} - \gamma I$$
      $$\dot{R} = \gamma I.$$

      Setting $S = 0$ yields $\dot{I} = -\gamma I$.

   b. Solve your simplified differential equation with the initial condition $I(0) = I_0$.

      Solving the ODE for I yields an exponential decay solution:

      $$\frac{dI}{dt} = -\gamma I$$
      $$\frac{dI}{I} = -\gamma dt$$
      $$\int_{I_0}^{I(t)} \frac{1}{I} dI = \int_0^t -\gamma dt$$
      $$\ln|I(t)| - \ln|I_0| = -\gamma t$$
      $$I(t) = I_0 e^{-\gamma t}.$$

   c. Manipulate your solution to derive the fraction of the initially infected people who are still infected at time t.

      The fraction of initially infected people who are still infected at time t is given by

      $$\frac{I(t)}{I_0} = e^{-\gamma t}.$$

   d. Discuss this equation. What does it do over time? How is it related to the fraction of infected people who have *left* the infected class?

3

This is a decaying exponential equation that starts at 1 when $t = 0$ and asymptotes towards 0 as $t \to \infty$. It is equal to 1 minus the fraction of infected people who have left the infected class.

e. This formula produces values between 0 and 1, and it tells us the probability that a randomly chosen infected person is still infected at time t. How does this relate to the cumulative distribution function (CDF) that describes the probability that someone is infected for less than or equal to t units of time? Take a derivative of the CDF to get a PDF for the duration of infection lengths is. Then, find out what this famous probability distribution is called, and write down its expected value.

If this equation gives the probability that someone is *still* infected at time t, then we take the complement to find the probability that someone is infected for $\leq t$ time units. Thus this CDF is

$$F(t) = 1 - e^{-\gamma t}.$$

Taking the derivative gives the PDF:

$$f(t) = \gamma e^{-\gamma t},$$

which is the PDF of the Exponential distribution, which has expected value $1/\gamma$.

f. Use your results to explain how the recovery rate γ is related to the typical amount of time a person remains infectious.

By the result of the previous part, the expected value of the duration of infection is $1/\gamma$.

3. The goal of this problem is to (i) figure out how to solve the final epidemic size equation, and (ii) test the equation's predictions.

   a. First, explain how an epidemic's total size, also called its cumulative incidence, is related to s_∞ and r_∞.

   The cumulative incidence is the total number of people who are ever infected. In the limit of the basic SIR model, this is equal to the reduction in the number of susceptibles $s_0 - s_\infty$ or the increase in the number of recovereds $r_\infty - r_0$.

   b. Recall that
   $$r_\infty = 1 - e^{-R_0 r_\infty} \tag{1}$$

   Though we can't solve this equation, we can use a valuable graphical technique as follows. First, we define a function $f(r_\infty) = r_\infty$ and define a second function $g(r_\infty) = 1 - e^{-R_0 r_\infty}$. Note that f is the left hand side of Eq. (1), while g is the right hand side. This allows us to plot f and plot g, both as a function of r_∞, to see where they cross—that is, where $f = g$.

   Create four plots for $R_0 \in \{0.9, 1.0, 1.1, 1.2\}$ with f in black and g in red. Use the **fsolve** function to find the intersection point, and use matplotlib's **scatter** function to plot a blue circle at the intersection.[2]

   Plots are shown below.

---

[2]Hint, use the fsolve docs, and note that, because fsolve wants to find roots (points where a function is zero), you can create an auxiliary function $h(r_\infty) = f(r_\infty) - g(r_\infty)$ which will be equal to zero at the point where f and g intersect!
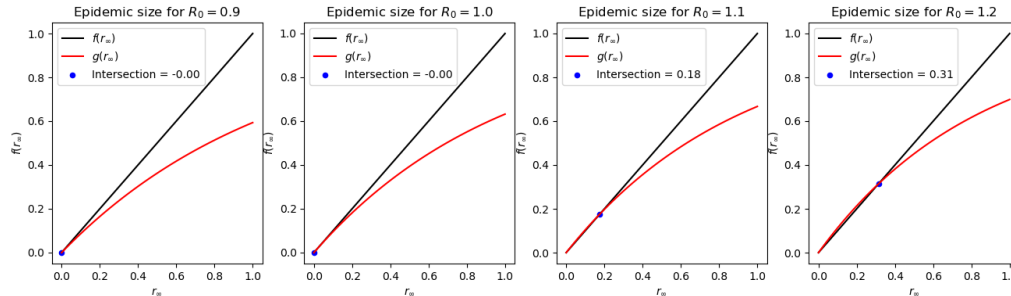
Figure 2: Graphical solutions for epidemic size for $R_0 = [0.9, 1.0, 1.1, 1.2]$. The right hand side $f(r_\infty) = r_\infty$ is shown in black and left hand side $g(r_\infty) = 1 - e^{-R_0 r_\infty}$ is shown in red. The intersection point is shown as a blue circle.

c. Comment on what you see in the plots in the context of what we have learned about R_0. What do you see in your figures? What happens when $R_0 < 1$?

When $R_0 \leq 1$, there is no epidemic because a single primary infection results in less than 1 secondary infection on average. In these cases (left two panels), the epidemic size is 0. As R_0 increases above, the epidemic size increases.

d. Finally, test the predictions made by this final-size equation by using your SIR code and $\beta = 1$, $\gamma = 0.5$ by creating a new version of that epidemic with a green dotted line at the height of r_∞. Does this final size prediction work?[3]

If $\beta = 1$ and $\gamma = 0.5$, then $R_0 = \beta/\gamma = 2$. The numerical solution to the transcendental equation for $R_0 = 2$ is $r_\infty \approx 0.7968$. The dynamics of the basic SIR model are plotted below with this estimate for r_∞ shown as a green dotted line. I used initial condition $S_0 = 999, I_0 = 1$. The final size prediction is close to where the recovered curve asymptotes, but the numerical solution slightly overshoots (this is due to numerical error: reducing the timestep brings the asymptote closer to the predicted value).

---

[3]Take care, as the units of our SIR plots and the units of our final size prediction are not the same! You may have to do a conversion...
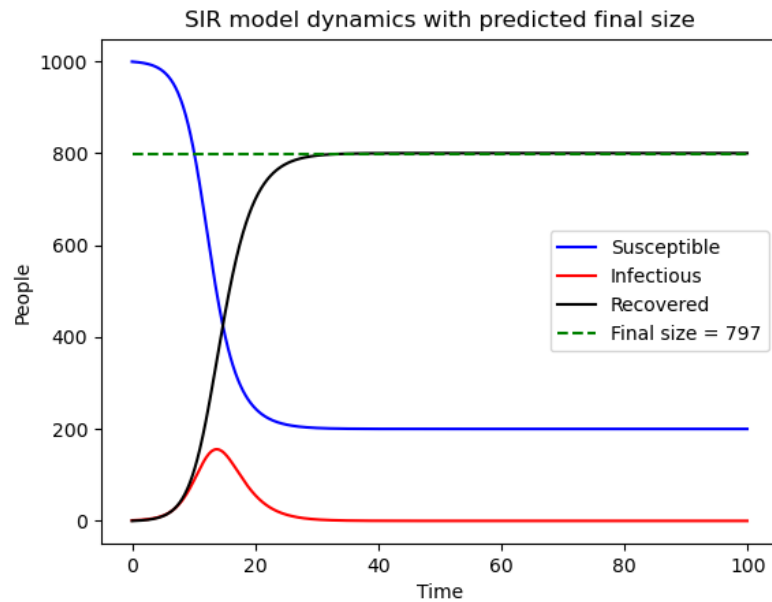
Figure 3: SIR dynamics with $R_0 = 2$. The green dotted line shows the predicted final size $r_\infty \approx 0.7968$.

4. (Grad / EC): In class, we showed that the SIR model's disease-free equilibrium is stable when $s < \frac{1}{R_0}$ and unstable otherwise. Using $N = 10^6$, and $\varepsilon = \frac{1}{N}$ as your perturbation, produce a single figure *using your simulation code and its output* that illustrates this point. Write a caption that explains the principle of stability, and explain how your figure illustrates it.[4]

To analyze stability while varying s, simulated SIR dynamics with specified parameters and $R_0 = 2$ for values of $s \in [0, 1]$. For each set of initial conditions, I apply the perturbation $i_0 = \epsilon, s_0 = s \cdot (1 - \epsilon), r_0 = 1 - s - \epsilon$ and simulate the dynamics for many timesteps to get the final perturbed susceptible fraction r_∞. I plot the initial equilibrium value of s on the x-axis and the difference in the initial equilibrium susceptible fraction and final perturbed susceptible fraction, $s - s_\infty$, on the y-axis. As can be seen in the plot, for values of $s \leq 1/R_0 = 0.5$, the perturbation results in a final susceptible fraction that is very close to the initial value of s. These are *stable* equilibria. For values of $s > 0.5$, the perturbation results in a final susceptible fraction that is much lower than the initial value; these are *unstable* equilibria.

---

[4]There are many possible ways to make such a figure, and write its caption! One could use all of the outputs of the simulation, for instance, but one could also use just those that illustrate the intended point.
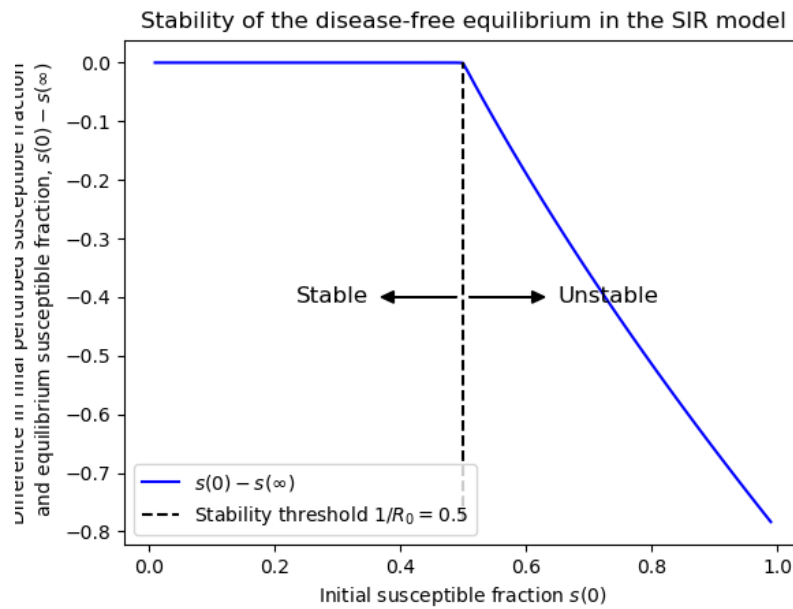
Figure 4: The *stability* of an equilibrium point in a dynamical system refers to whether a perturbation to that point eventually returns to equilibrium or diverges away from it. This figure illustrates the stability of the disease-free equilibrium ($s = 1 - r, i = 0, r = r$) of the SIR model as a function of the susceptible fraction s. For values of s lower than the stability threshold $1/R_0 = 0.5$ (black dashed line), a small perturbation (adding a single infected) does not change the final susceptible fraction much; these are stable equilibria in which there is not an epidemic. For values of s greater than the threshold, a small perturbation results in a large change in the final susceptible fraction; these are unstable equilibria in which there is an epidemic.