Workshop on models of an epidemic, Due: Wednesday, July 17, 2024

Group members: Xia (Alice) Zhu, Finn Jensen, Avirul Islam, Matthew Ramli

1. For the continuous deterministic approach, we can describe our model as a system of equations. We used ode45 in Matlab to model the evolution of S, R, and I populations over time. Use a = 2ln(2) per week, b = ln(2) per week, S = 999, I = 1, R = 0 (hence N = 1000). Integrate over t = [0,20] in units of weeks.

Here is the resulting plot:

A graph of a function

Description automatically generated with medium confidence

Description of plot qualitatively: We start with N = 1000, i.e. the entire population being susceptible to an arbitrary/unnamed disease. The basic reproduction number (a/b) is the quotient of the infection rate (a) and the half-life of an infected state (b), which is 2 in this case. This describes the number of people that an infected individual can infect before he/she recovers without immunity or intervention. We see maximum of infections at t = ~10 weeks, then infections dropping once the number of recovered people increases. This is because we assume that people who recover do not get infected again, due to development of natural immunity. We expect population of infected people to decrease roughly around when the number of recovered is greater than the number of susceptible people. This is roughly when herd immunity has been established. We expect the number of infected to increase exponentially (i.e. 1 infects 2, 2 infects 4, etc.)

Hold ‘b’ (half-life of infected state) same – i.e. same recovery rate, change ‘a’ to represent e.g. different virulence of virus, different social distancing practices, or quarantine measures. Notes: a>b exponential growth, a<b exponential decay; epidemic reaches peak when I(t) reaches maximum. We are assuming a and b are both constant with time.

|  |  |  |
| --- | --- | --- |
| a = 1\*ln(2)  b = ln(2)  R0 = a/b  = 1  Represents a virus less virulent than COVID-19 |  | if ‘a’ is small, no one or very few people get infected. |
| a = 3\*ln(2)  b = ln(2)  R0 = a/b  = 3  The Ro for COVID-19 is ~2-3 |  | There are more infected people as ‘R0’ increases; large R0 means large a with respect to b, so more people are infected by a person before that person can recover. The susceptible and recovered curves intersect a roughly when # infected reaches its peak, meaning infected pool starts to diminish and herd immunity is eventually achieved. |
| a= 5\*ln(2)  b = ln(2)  R0 = a/b  = 5  Some new variants of the virus are found to have R0 1.5x higher than that of COVID-19 (i.e. R0 = ~4-5) |  | Again, max # of infected people (I) roughly coincides with intersection of susceptible (S) and recovered (R) curves. However, greater ‘a’ parameter results in # infections shooting up and achieving maximum more quickly, probably due to faster spread of the virus. Maximum I here is greater than maximum I for less virulent virus. |
| a = 10\*ln(2)  b = ln(2)  R0 = a/b  = 10  See what would happen if we had an extremely virulent virus |  | As you increase virus infection rate, you reach max infected pool (I) faster. Here, herd immunity is also achieved more quickly. |

1. Stochastic simulations

Using a while loop and a random number generator, we stochastically evolved the S, I, and R pools over time using the 1) probability of a susceptible individual becoming infected (p1), 2) probability of an infected individual recovering (p2), and 3) probability of the pools staying the same, i.e. no change (p3).

* Probability of one susceptible individual getting infected for a given timestep = a/N\*I(t)\*dt. Multiply this by S(t) to get the total probability of any individual in a pool of suspected individuals getting infected: p1 = a/N\*I(t)\*S(t)\*dt.
* p2 = probability of infected individual recovering = b\*I(t)\*dt.
* p3 = probability of no changes occurring to the I(t), S(t), and R(t) pools = 1 – p1 – p2.

Model is provided in ‘epidemic\_stochastic.m’ file.

1. Check consistency between the deterministic and the probabilistic approaches. I’m using same basic reproduction numbers (R0) as before to compare deterministic and stochastic models:

|  |  |  |  |
| --- | --- | --- | --- |
| basic reproduction number | deterministic | stochastic | discussion |
| R0 = 1 (less virulent than COVID-19) |  |  | These plots look similar |
| R0 = 3 (COVID-19) |  |  | For the stochastic plot, infection curve does not reach a maximum then decrease. This is because our x-axis is “timesteps”, and our total time = 1 week (corresponds to nsteps = 10,000). Extending total time to 20 weeks would give us the same curves as the deterministic model. |
| R0 = 5 (new variants of COVID-19) |  |  | For stochastic plot, infection curve does not reach a maximum then decrease because we are only plotting up to 1 week’s worth of evolutions. |
| R0 = 10 (much more virulent than COVID-19) |  |  | Here in the stochastic plot we can see the ‘infections’ curve beginning to peak, around where susceptible and recovered curves intersect. Infections curve peaks earlier due to greater R0. |

1. Represent each individual in the population explicitly. Use code on the ‘individual agent model’ from course webpage to model the epidemic as well:
2. Get a sense of average results or distribution of s, i, r (n = 100 iterations):

for n = 1000; nc = 20; pt = 0.1; iterations = 100

Every time the simulate1D code is run, the ‘results’ matrix that is produced varies in length (# of rows). In order to average all of the results together, I made all 100 results matrices the same length before averaging:

* First, I found the maximum results matrix length out of all 100 iterations
* Then, I appended additional final rows to the matrices that were shorter than the maximum length
* Then I summed all of the matrices and saved in a matrix called ‘resultssum’
* Finally, I took the average of all the results matrices by dividing ‘resultssum’ by 100 (# of iterations) and converted all the numbers to integer form

Please see ‘individualagentmodel.m’ for full code.

1. Run program with different nc and pt values to demonstrate their effect.

|  |  |  |
| --- | --- | --- |
| Try | Plot | Result |
| nc = 1, 5, 10  pt = 0.1 | nc = 1, pt = 0.1    nc = 5, pt = 0.1    nc = 10, pt = 0.1 | * As you increase ‘nc’, the length of the results matrix increases. In other words, it takes longer for # of infected people to reach zero. |
| nc = 20 (integer)  pt = 0.1    vs  nc = 4  pt = 0.5  \*same product as above, i.e. 2 infectious contacts per infected individual per timestep | nc = 20, pt = 0.1    nc = 4, pt = 0.5 | * These two scenarios look very similar in terms of distribution of s, i, and r * Also in how quickly # infected plummets to zero |
| nc = 5  pt = 0.1, 0.5, 1 (0 < pt < 1) | nc = 5, pt = 0.1    nc = 5, pt = 0.5    nc = 5, pt = 1 | * As you increase pt, which is the probability of s becoming i, we see # of infected people reaching a maximum and then decreasing. If pt is low enough, then # of i does not exponential increase; it just stays at a low number for a while until zero is reached * i.e. different ‘domains’ of behaviour? and certain values of the parameters trigger the shift from one domain to another… |

1. Does this increased resolution of the model, allowed by its individual agent nature, make any difference?

Try for the conditions of nc = 10, pt = 0.5:

|  |  |  |  |
| --- | --- | --- | --- |
| Population size  \* Try this for different population sizes and a large number of realizations may be needed to reach firm conclusions | Mean field model, for comparison (a/b = R0 = 3, representative of COVID-19) | Individual agent model | Discussion |
| n = 10 |  |  | For the individual agent models, as population decreases, # of infected peaks earlier |
| n = 100 |  |  |  |
| n = 1000 |  |  |  |
| n = 10,000 |  |  | Individual agent model: As population size increases, peaks appear later and maximum(i) is larger  Stochastic model: we are only plotting 1 week, but you can see that the maximum(i) peak is more delayed than in the other plots |

1. Discuss the pros and cons of the different approaches.

Comparison Table #1

|  |  |  |
| --- | --- | --- |
|  | Pros | Cons |
| Deterministic approach | * faster, less computationally expensive | * only true when stochastic model’s # of timesteps approaches infinity; not appropriate for all cases |
| Stochastic approach | * incorporates uncertainty into results, more realistic | * slower to compute |

Comparison Table #2

|  |  |  |
| --- | --- | --- |
|  | Pros | Cons |
| Mean field model | * useful for population-based results, analysis, policymaking | * cannot use individual characteristics, e.g. # of infectious contacts before recovery, probability of susceptible becoming infected |
| Individual agent model | * can argue is more realistic because individual characteristics/attributes are retained | * need high number of runs (replicates) * cannot use population parameters/characteristics, e.g. doubling time of infected, half-life of infected state |

As with any debate regarding different approaches, e.g. Eulerian vs Lagrangian, top-down vs bottom-up… I think there are cases where one is more appropriate than the other. If a government is trying to monitor the spread of COVID in an entire population, I think the population model is more suitable for this purpose and for setting policies in government, e.g. when people can go outdoors again. If we are trying to trace a string of cases back to the epicenter (or origin), then monitoring the spread of disease at the individual level might be more appropriate.

1. Make a list of the different ways that mathematical modelling is used in the fight against the COVID-19 pandemic.

|  |  |
| --- | --- |
| #1 | Like we did in HW2 Individual Question 1, we can use mathematical modelling to determine optimal pool size for pooled testing. This can help save resources and time when testing for prevalence of COVID-19 in a population. |
| #2 | Like we are doing in this epidemic model workshop, mathematical modelling at the population level can be used to model the evolution of an epidemic; in other words, at any given time, we can estimate how many people are susceptible, infected, and have recovered from the disease, respectively. We can also predict how long it will take for herd immunity to be achieved, based on factors such as i) virulence, ii) quarantine, iii) social distancing. Using this information, the public can be informed as to how to combat the virus, when it is safe to go outside again, etc. |
| #3 | Individual agent models are useful for tracking the spread of disease from person to person. I think this can be useful for contact tracing, to determine where the disease began. |