Worksheet 2: Stochastic models

u1624423

February 9, 2022

All plots and calculations are performed in `Assignment2a.m`

Question 1

(a) $R_0 = \frac{\beta}{\gamma} = \frac{5}{1/3} = 15$ with no immunity. In the SEIRD model $R_e(t)$ is calculated identically to the SIR model, as demography is not involved. $R_e(0) = R_0$. With 90% prior immunity, $R_e(0) = R_0 = \frac{(1-0.9)\beta}{\gamma} = \frac{0.5}{1/3} = 1.5$, where 0.9 represents the 90% prior immunity.

(b) The ODE model produces the following infections.

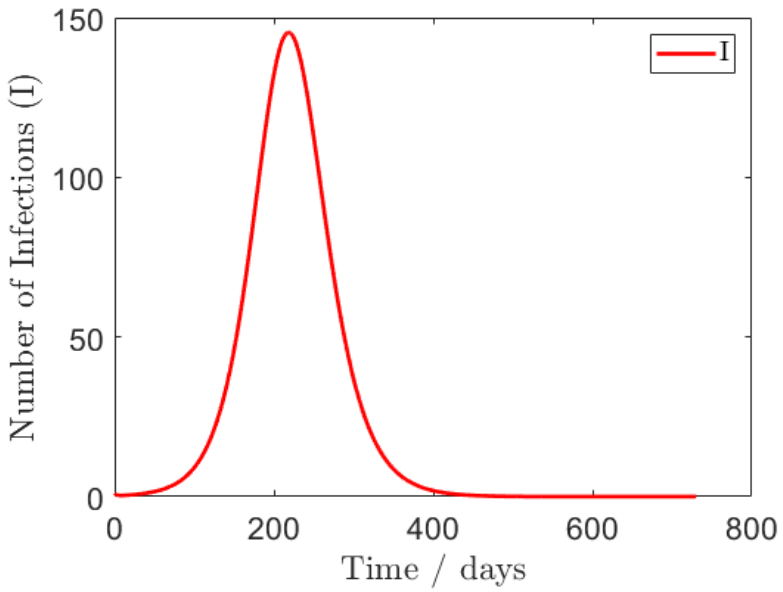

Figure 1: Infections over 2 years.

The number of deaths is cumulative, so we can just take the last element of the D class and round it, giving $deaths = 58$. Alternatively, we can calculate the total number of infections, I_{tot} by using trapz, then $deaths = p_D \gamma I_{tot} = 58$ (when rounded).

Question 2

(a)

Events	Rates	Change
New Infection Exposure	$\frac{\beta IS}{N}$	$E \to E+1,\ S \to S-1$
End latency period	σE	$I \to I+1,\ E \to E-1$
Recover	$(1-p_D)\gamma I$	$R \to R+1,\ I \to I-1$
Die	$p_D \gamma I E$	$D \to D+1,\ I \to I-1$

(b) The model is in `Gillespie_SEIRDmodel.m`

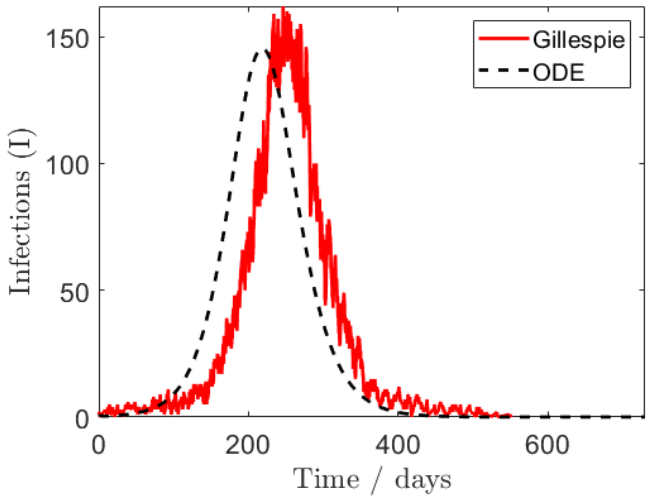

Figure 2: Gillespie simulation of Infections over 2 years.

(c) The duration of this simulation was 549 days, and the final size was 5767.

(d)

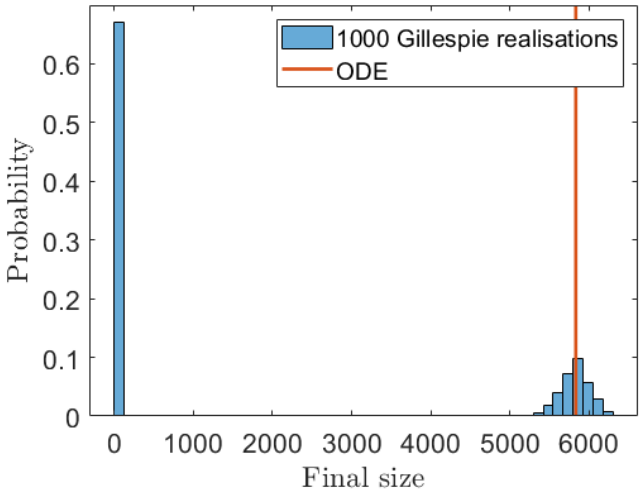

Figure 3: Histogram of final sizes of outbreaks for 1000 simulations.

In figure 3, as expected, there is a normal distribution around the ODE model's final size.

(e) Any exposed individual will become infected, so the probability of someone becoming infected is the same as in the SIR model, and we can similarly treat any death or recovery as an exit from the infected class, so the probability of recovering is also the same as the SIR model's. The probabilities are given by:

$$\text{Probability of Infection occuring, } P(I) = \frac{\beta}{\beta+\gamma},$$
$$\text{Probability of Recovery occuring, } P(R) = \frac{\gamma}{\beta+\gamma}.$$

In the case there are 0 extra infections, there is only 1 possible case, that is that the intial infected individual recovers and doesn't infect another individual. Similarly, when there is 1 extra infection, there is also only 1 possible case. However, when there are 2 extra infections, there are 2 possible combinations of what can happen, and when there are 3 extra cases there are 5 possible combination. The equation for the probability of a certain number of extra infections is, where C is the number of combinations possible for n which is the number of extra infections:

$$P(n) = P(I)^n P(R)^{n+1} C$$

Given this, we can sum up $P(0) + P(1) + P(2) + P(3) = 0.066654127091169$ which we must scale by N/N_i, where N_i is the immune population, giving us a probability of 0.66654127091169, which is not too far from the model's calculated value of 0.6220 (over 1000 runs). The value we get analytically makes sense because the probability of extinction is given by $P(\text{extinction}) = 1/R_0 = 2/3$, where our $R_0 = 1.5$ as calculated in question 1. Most times when a large outbreak does not occur, it is within the first few timesteps, so the final size is very small (i.e. less than about 4). The probability of extinction matches up very well with the model which has about 67% of cases end with a final size less than 126, which the vast majority of which are in the sizes 1-4.

Question 3

(a) The new system of ODEs is as follows:

$$\frac{dS}{dt} = ((1-v) + (1-vp_e))\mu N - \frac{\beta SI}{N} - \mu S, \qquad \frac{dR}{dt} = (1-p_D)\gamma I - \mu R,$$
$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E - -\mu E, \qquad \frac{dD}{dt} = p_D \gamma I,$$
$$\frac{dI}{dt} = \sigma E - \gamma I - \mu I, \qquad \frac{dV}{dt} = p_e v \mu N - \mu V.$$

The new events table is:

Events	Rates	Change
New Infection Exposure	$\frac{\beta IS}{N}$	$E \to E+1$, $S \to S-1$.
End latency period	σE	$I \to I+1$, $E \to E-1$.
Recover	$(1-p_D)\gamma I$	$R \to R+1$, $I \to I-1$..
Deaths from measles	$p_D \gamma I E$	$D \to D+1$, $I \to I-1$.
Natural deaths of Susceptible	μS	$S \to S$
Natural deaths of Exposed	μE	$E \to E-1$, $S \to S+1$.
Natural deaths of Infected	μI	$I \to I-1$, $S \to S+1$.
Natural deaths of Recovered	μR	$R \to R-1$, $S \to S+1$.
Natural deaths of Successfully vaccinated	μV	$V \to V-1$, $S \to S+1$.
Immune at birth	$p_e v \mu N$	$V \to V+1$
Not immune at birth	$((1-v)+(1-p_e)v)\mu N$	$S \to S+1$

In the last two events on the events table add to V or S from the births, which are assumed to be at the same rate of natural deaths. In the ODE system, the μN is the total number of births, which is equal to $\mu(S+E+I+R+V)$, the total deaths. The $((1-v)+(1-vp_e))$ term is the proportion of those born who are left susceptible due to either not being vaccinated, $(1-v)$, or being unsuccessfully vaccinated, $(1-vp_e)$.

(b) See `Tauleap_SEIRDmodel.m` for the Tau Leap model. The Tau Leap simulations oscillate sinusoidally for the most part, much like the ODE model, but without a vaccine, it is much less likely for measles to be eliminated within this population within 20 years. Also, the model often predicts much larger or smaller amounts of infections than the ODE suggests, but the timings of the peaks frequently match up between the Tau Leap simulations and the ODE. These over/underestimates are expected occasionally as the random nature of the Tau Leap algorithm can lead to quite chaotic outcomes.

Looking at the distribution of deaths due to measles, we can see that the most likely occurrence is that there are many deaths, (over 1000). This can be explained by looking at a histogram of the duration of the outbreaks, we can see that the most common occurrences are long outbreaks, that don't even end within the 20 year period we are modelling over.

Approximately 32% of the outbreaks do not end within 20 years, meaning approximately 68% of the simulations do result in measles being eliminated within 20 years. It is quite obvious that introducing a vaccine with a significant success rate would decrease the duration and number of deaths of the outbreak simulations. Furthermore, it is also apparent that for local elimination to occur, it is only really possible at the point where

the sinusoidally shaped Tau Leap is at a one of the minima turning points.

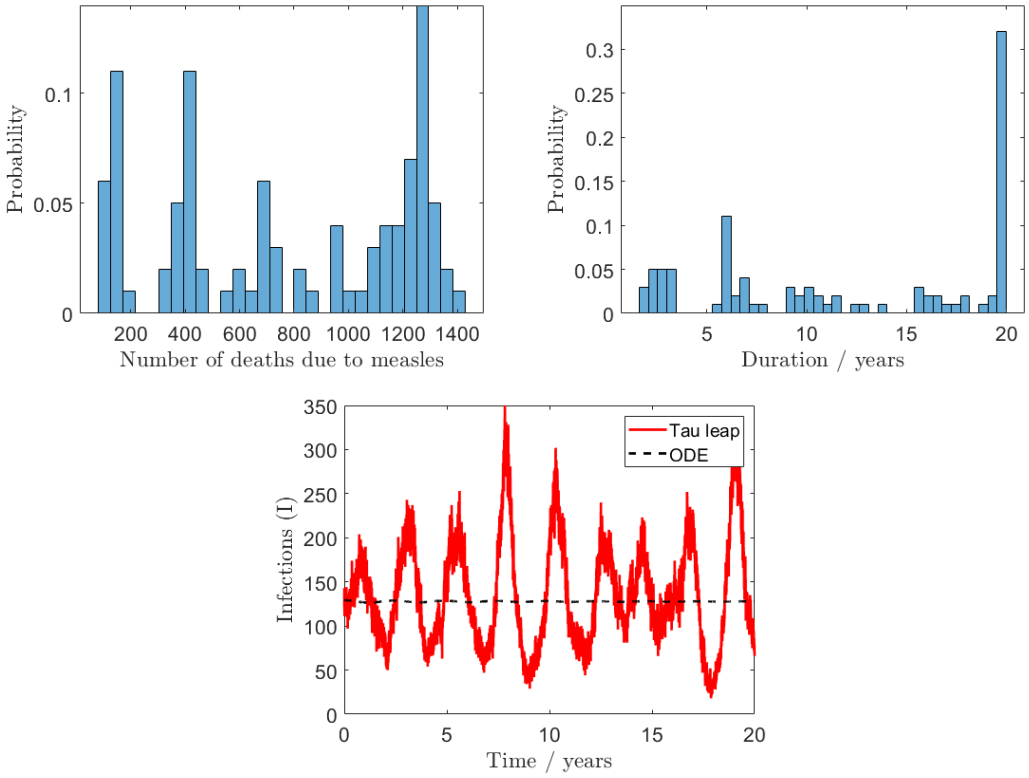

Figure 4: Duration and death distributions over 1000 Tau Leap simulations, with one tau leap simulation plotted compared to the ODE as an example.

The correction step to ensure biological feasibility, i.e. to correct when there is a negative class size, excess occurrences can be undone. For example, if S were to become negative, excess exposures to infection can be done, by putting the excess from E back into S. However, there is a complication that must be considered, which is that multiple events can happen for some classes. Using the infection class, I, as an example, people can be removed from this class one of 3 ways, by recovering, by dying from measles or by dying naturally. To choose how to return the excess to the I class, let the number of people in I be x. When x is negative,

$$\text{event total} = \text{recovery events} + \text{measles death events} + \text{ natural death events}$$

$$d = min\left(binornd\left(-x, \frac{\text{measles death events}}{\text{event total}}\right),\ \text{measles death events}\right)$$

$$r = min\left(binornd\left(-x-d, \frac{\text{recovery events}}{\text{event total}}\right),\ \text{recovery events}\right)$$

$$n = -x - d - r$$

Where *binornd* randomly chooses numbers from a binomial distribution, and where d, r and n are the numbers of people to remove from the D, R and S classes respectively to put back into the I class. We take the minimum so we don't take from a class with

less than 0 people and in that case, we return from nonempty classes (automatically by the equations above). Similar methodology is applied to the other classes as well.

(c) Generally speaking, Tau Leap is preferable to Gillespie for various reasons. Since Tau Leaping has typically longer time intervals, less updates to the model are made, meaning that less calculation is made and it allows for a more efficient simulation. Additionally, in each time step of Tau Leap, multiple classes are updated, rather than just a single one as in Gillespie's algorithm. Both of these two factors contribute to faster computation times for a Tau Leap model, meaning larger systems can be considered. This directly applies to our models as we are performing Tau Leap over 20 years, whereas our Gillespie model is only running over 2 years. We can see the difference in time, using `tic` and `toc`, the time to run 100 Tau Leaps, with a more complex model, modelling for a date range 10 times as long as the Gillespie model, is $176.689272 seconds$ and for 100 Gillespie runs on the model in question 2 it takes $33.823024 seconds$. This means that for a model of 10 times the size, the Gillespie model would likely take over $300 seconds$, a significantly longer time than Tau Leap. N.B. these are run on my machine and will vary with different hardware. These timings alone are sufficient evidence to suggest that performing Tau Leap for this new model is preferable.

Another advantage to the Tau Leap algorithm is that we can set discrete time intervals to be whatever we want, which means that various analysis on the model generated is significantly easier, whereas with Gillespie, the time intervals aren't even necessarily equidistant. This means Tau Leap is generally more convenient than Gillespie which can be troublesome to use for purposes other than visualisation, although it is still possible.

(d) With $v = 0.6$, we can see a noticeable change in the shape of the ODE, it dampens much faster than before and the Tau Leap seems to match the amplitude of the ODE much more closely. Additionally, there is a much greater probability of the Tau Leap simulation results in elimination within a shorter period of time.

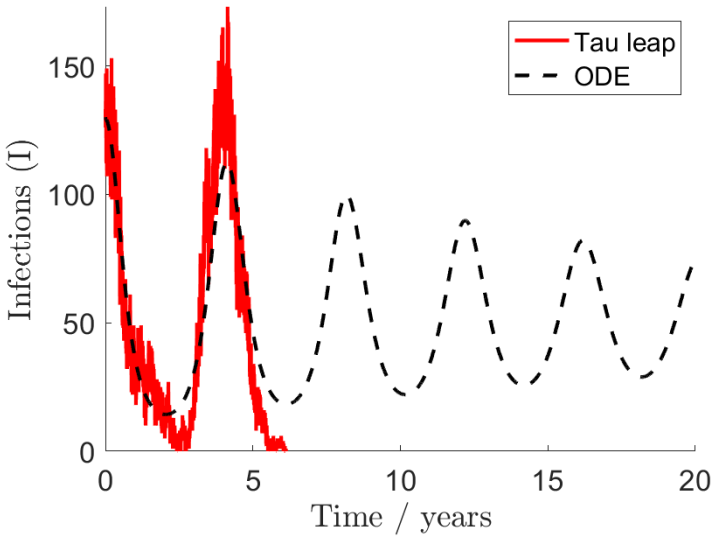

Figure 5: One Tau Leap simulation with 60% vaccination rate for infants

With a vaccine rate of 60%, the probability of local elimination within 5 years is approximately 17% (of 1000 simulations, 173 of them ended within 5 years). To ensure a 95% probability of elimination within 5 years, we have to use trial and error to work out what value of v we would need. In this model, we I found that a vaccination rate of approximately 84% is required to reach a 95% chance of elimination within 5 years (tested over 1000 runs multiple times). It is also worth mentioning that, as an increased v decreases the length of the simulation, the simulations run faster as they are shorter.

(e) There are many issues with the feasability of running a measles vaccination campaign in rural DRC. Key factors to this are to do with the widespread poverty in the DRC, unlike the UK, most people do not own cars or have easy means of transport. There is no NHS and essentially no emergency care is available outside of foreign intervention. People are not necessarily able to get to a hospital that has facilities to provide treatment and the vaccine for measles, many people don't live near any of the 4 MSF hospitals in rural DRC. The longer measles goes untreated, the less survivable it is, and it can lead to and stack up with other complications.

Another issue is that the measles vaccine, according to the WHO, only has an 85% chance to immunise, which is significantly lower than the model, which suggests 95%. A much lower amount of people would be able to get a second vaccine than those who have only had one. With our model, we need 84% of infants to be vaccinated to have a 95% chance of eliminating measles within 5 years, but our model is very different from the real case. In the DRC, people don't necessarily have the means to vaccinate their children at birth, and additionally, there are many other diseases that can be fatal, such as malaria. These alongside many other factors mean the model we have is most likely quite overtly optimistic.

The vaccine costs about $1USD, which very cheap. It would be quite easy for a wealthy country, such as the UK or USA, to provide an abundance of vaccines. If multiple countries' governments were to aid, it would cost them proportionally very little to provide more than enough vaccines to the DRC. The population is about 90 million, so even if everybody required 2 vaccines, it would only be $180 million USD. The problem is transporting the vaccines and distributing them, building infrastructure such as hospitals, providing other resources for other care, which would end up costing a lot more.