MMNUCM Final Assignment

Lyn Saunders

Contents

1 Deterministic SIR model of epidemic.

Consider a version of the SIR model of an epidemic, given by

$$\begin{aligned} \frac{dS}{dt} &= -\beta\frac{IS}{N}, \\ \frac{dI}{dt} &= \beta\frac{IS}{N} - \gamma I - \mu I, \\ \frac{dR}{dt} &= \gamma I, \\ N &= S + I + R, \end{aligned}$$

where $S(t)$, $I(t)$ and $R(t)$ are the numbers of susceptible, infected and recovered individuals; β is a coefficient characterising the transmission of infection from an infected individual to a susceptible one; γ is the recovery rate and μ is the mortality rate due to infection. Let the parameter values be as follows: $\beta = 2$, $\gamma = 0.1$, $\mu = 0.01$.

1.1 Part a)

Solve the system with these parameters using MATLAB's ode45 with initial conditions $S(0) = 100$, $I(0) = 1$, and $R(0) = 0$ and plot the result.

See file ex1a.m for the corresponding code and Figure 1 for the resulting plot.

The function **sir_func** takes a vector **initial_conditions**, which hold the initial counts of the susceptible, infected and recovered populations, and the parameters β (infection rate), γ (recovery rate), and μ (mortality rate) as inputs and evaluates the differential equations of the SIR model of the epidemic using the differential equation solver **ode45** over a time span, **T**, of 200 days. The function returns the solution matrix X, recording the changing SIR populations evaluated over **T**, and produces a graphical representation of the simulation in Figure 1.

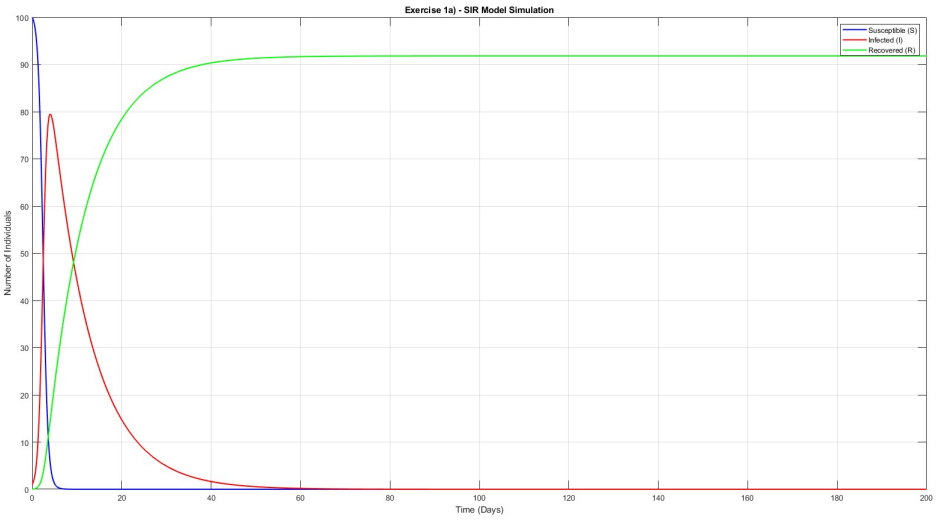


Figure 1: the resulting plot from solving the deterministic SIR model of epidemic with initial conditions $S(0) = 100$, $I(0) = 1$ and $R(0) = 0$, and using the parameters provided in the description for Question 1.

1.2 Part b)

Compute and plot the total number of deaths due to the epidemic for a range of values of β. Keep the other parameters and the initial conditions fixed.

See file ex1b.m for the corresponding code and Figure 2 for the resulting plot.

We again use **ode45** to solve the given system, but this time we have simulated the model over a 200 day period for different values of the parameter β using a **for** loop. We have then tracked the number of deaths that occur due to infection for each scenario and plotted the results in Figure 2.

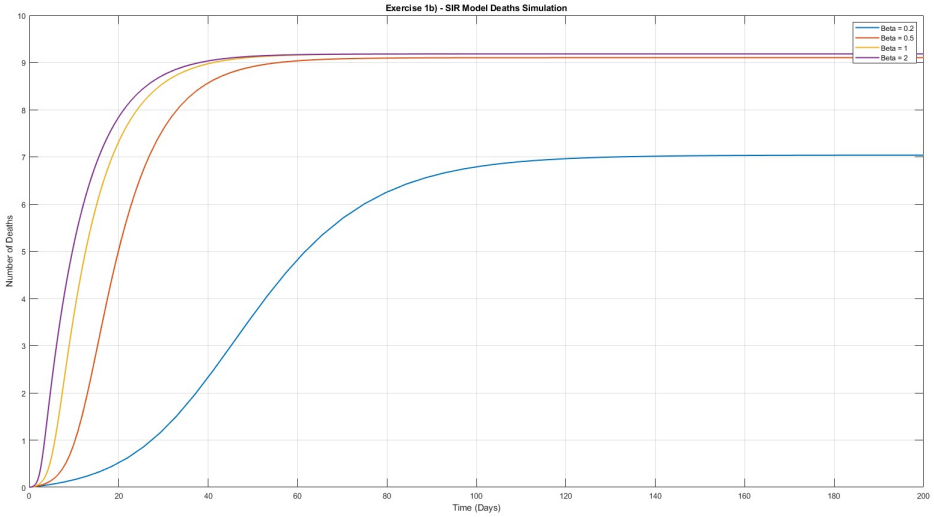


Figure 2: here we have plotted the total number of deaths due to the epidemic for the following values of β; 0.2, 0.5, 1, and 2.

Taking a look at our plot we see that for $\beta = 0.2$, the total number of deaths approximates to 7, whereas for the range of values 0.5, 1, and 2 we see that the number of deaths in each of these cases all approximate to 9. The consistency of the number of deaths that occur for $\beta \in [0.5, 2]$, in comparison to how many occur for $\beta = 0.2$ could be due to the population dynamics of the system reaching an equilibrium at around $\beta = 0.5$.

1.3 Part c)

Describe how the equations could be altered to include the effects of vaccination or reinfection.

Implementing a vaccination process into our SIR model would result in individuals being 'fast-tracked' from the susceptible population to the recovered population. The affected equation would be of the form $\frac{dR}{dt} = \gamma I + \pi S$, where $\pi > 0$ would be the vaccination parameter.

Introducing reinfection into the SIR model would lead to a portion of the recovered population, multiplied by a reinfection parameter $\sigma > 0$, being reintroduced into the susceptible population. The resulting equation would look like the following: $\frac{dS}{dt} = -\beta \frac{IS}{N} + \sigma R$.

2 Stochastic SIR model of epidemic.

2.1 Part a)

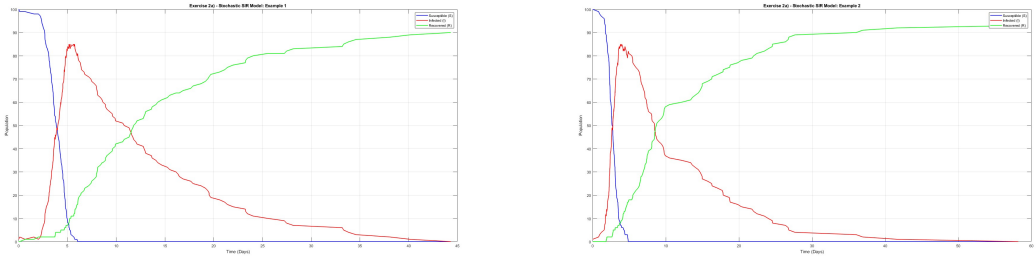
Create MATLAB code which simulates an epidemic as a sum of three independent Poisson processes:

Poisson Process	Rate
Infection of susceptible individual	$\frac{\beta SI}{N}$
Death of infected individual	μI
Recovery of infected individual	γI

Simulate this process with parameter values and initial conditions from Question 1. Provide graphical output showing two independent examples of the process and compare it with the deterministic model of Question 1.

See file ex2a.m for the corresponding code and Figure 3 for the resulting plots.

For the stochastic model we ran two independent simulations of the epidemic as a sum of three independent Poisson processes, with rates described by the parameters β, μ, and γ and the SIR populations over the time threshold of 1000 days. Since the Poisson processes are independent, we are able to combine the rates into the scalar **total_rate**, and using this combined rate we describe the 'time to next event', τ, as being exponentially distributed with mean $\frac{1}{\textbf{total\_rate}}$. Within a **while** loop we determine, in time steps of τ, which of the three possible events (infection, death or recovery) will occur according to their Poisson rates. During each time step we scale the **total_rate** by multiplying it by **rand** (a random number between 0 and 1) to get our general event probability (**event_prob**). Using this probability we use an **if, else** loop to determine, for each time step within the threshold, whether the next event will be an infection (**if event_prob** $<$ **rs**), a death (**elseif event_prob** $<$ **rs**+**ri**), or a recovery (**else**), where **rs**, **ri** and **rr** describe the Poisson rates of infection, recovery and death, respectively. At the end of each time step the population vectors are updated according to which event occurred, and each independent example of the simulation is plotted in its own figure in MATLAB.



(a) First example (b) Second example

Figure 3: here we have generated two independent examples of the stochastic SIR model of epidemic as a sum of three independent Poisson processes.

Comparing these plots with that of Figure 1 we see that the shape of the curves that describe the susceptible, infected and recovered populations become more 'jagged' in the stochastic examples, as opposed to the smoother curves produced by the deterministic model. This is due to the inherent

nature of stochastic modelling, which takes into account the invariability and uncertainty of a model, unlike deterministic modelling, which assumes accuracy in the given parameters.

2.2 Part b) section i)

For a suitably large sample size, calculate the mean duration of the epidemic (i.e the time needed for the number of infected individuals to drop to zero) and plot the histogram.

See file ex2b_i.m for the corresponding code and Figure 4 for the resulting plot.

For this exercise we use the same code as in part a) but with a few changes; we now use the **for** loop to simulate the epidemic 1000 times, and after each simulation the time at which the infected population reaches 0 is recorded and stored in the vector **Tfinal**. The mean of this vector is calculated and rounded to 3 significant figures before being displayed in the command window. Additionally, a histogram depicting the mean duration of the epidemic over the 1000 simulations is plotted, an example of which can be seen in Figure 4.

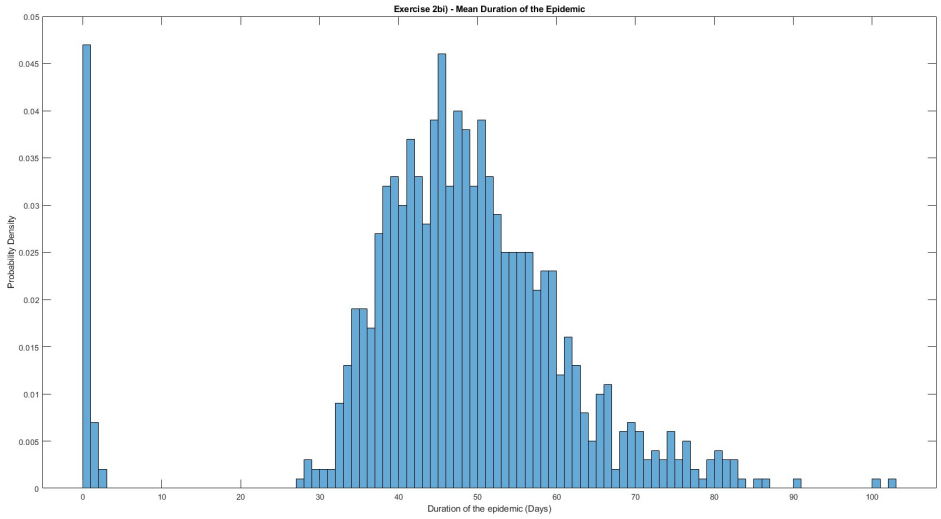


Figure 4: for this particular example we calculated the mean duration as 47 days, rounded to the nearest day.

As seen in the above figure, our stochastic model gives a spike at the beginning of the histogram, indicating that within our 1000 simulations quite a few epidemics have terminated very quickly, particularly within the first day. We may pin the cause of this phenomena to the high ratio of the initial susceptible population with the initial infected population, or perhaps the chosen value of the infection rate β.

2.3 Part b) section ii)

Calculate the mean of the total number of deaths due to the epidemic and compare it with the prediction of the deterministic model of Question 1.

See file ex2b_ii.m for the corresponding code.

Using the same principal of stochasticity in our simulation, we alter the code previously used in part i) of Question 2b to record the final total population after each simulation in the vector **Nfinal** and use this to calculate the number of deaths that occurred in each simulation.

As required for the stochastic model, we used the same initial conditions and parameter values as we did in out deterministic model. Thus, when comparing the results from this Question we will compare them with the results from Question 1b) (Section 1.2) which utilises $\beta = 2$. In that exercise we discovered that the number of deaths due to the epidemic after 200 days for the deterministic SIR model was 9.18 (evaluated to 2 decimal places), which we can also be seen in Figure 2, whereas using our stochastic model we find that the mean total number of deaths due to the epidemic for 1000 samples ranges between 8.5 and 8.8. This difference showcases the inherent nature of stochastic models, which account for uncertainty and variability, in comparison to deterministic models, which assume constant parameters and provide an expected outcome.

2.4 Part b) section iii)

Calculate the mean time needed for the epidemic to reach its peak (i.e. the maximum number of infected individuals).

See file ex2b_iii.m for the corresponding code and Figure 5 for the resulting histogram.

Once again we used the same base code established for section i) of Question 2b and, after each simulation, recorded the maximum number of infected individuals in the vector **I_peak**. By aligning the indices of the vector **I_peak** and the **times** vector, which records the time steps at which the populations change, we are able to find the times at which the epidemic reached its peak in each simulation, and evaluate them to 2 significant figures. We then find the mean over all of these times, and consequentially plot the histogram of our results. An example of this can be seen in Figure 5.

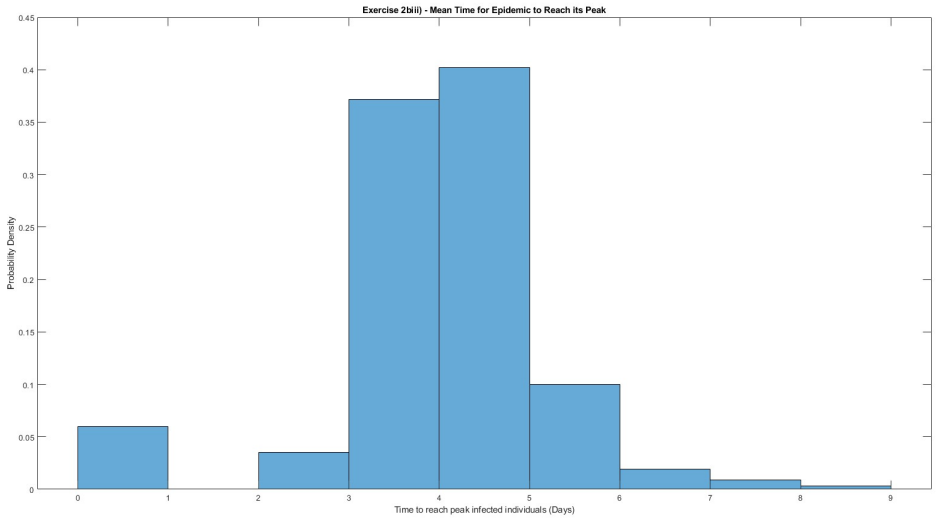


Figure 5: for this particular example we calculated the mean time needed for the epidemic to reach its peak over 1000 samples is 4 days (rounded to two significant figures).

As mentioned previously in section i) of Question 2b (Section 2.2) we see a spike in the beginning of the histogram, due to many simulations wherein the epidemic terminated early.

2.5 Part b) section iv)

Approximate the probability that the number of infected individuals at the peak will be more than 10.

See file ex2b_iv.m for the corresponding code.

Omitting the sections of the code for section iii) of Question 2b after the ending of the **for** loop we now add an **if, else** loop nested in an additional **for** loop, which runs over the same 1000 simulations of the epidemic to determine for how many simulations out of our large sample the maximum number of infected individuals at the peak will be more than 10. We do this by initialising a (1,k) vector (where k iterates over the 1000 simulations) of zeroes labelled **yes** and, using the **if, else** loop, changes the input at k^{th} index to a 1 if $\mathbf{I\_peak(1,k)} > 10$. To find the overall probability that the number of infected individuals at the peak will be more than 10 we simply sum up the elements of **yes** and divide by the total number of simulations.

Running the code we find that the probability that the number of infected individuals will be more than 10 for 1000 samples ranges from 0.93 to 0.95. This result corresponds to our results found in part a) of the same Question (Section 2.1); looking at the two independent examples of the stochastic SIR model of the epidemic we see clearly that the number of infected individuals at its peak lies between 80 and 90.

2.6 Part b) section v)

If there is a vaccine which is 100% effective, what level of vaccination (a percentage of vaccinated population) is needed to reduce the probability computed in part (iv) by a factor of 2?

See file ex2b_v.m for the corresponding code, which may take a little longer to run than the previous files.

In the final section for Question 2b we take all of our code from section iv) of part b) of the same Question and make one major modification to the existing main body of the code before nesting it in another **for** loop. The substantial modification we make is to the initial susceptible population, by including a term that accounts for the vaccinated population: we write S = 100*(1 - **vax_percent**). Then we input this code into a **for** loop, which iterates the entire process computed in section iv) 100 times, updating the **vax_percent** by 1% with each iteration. After each iteration over a new vaccination rate the probabilities that, with said vaccination rate, the maximum number of infected individuals is more than 10 is stored in the vector **effected_probabilities**. By aligning indices, we can find when the probability from iv) is halved by labelling it as our **target_probability**, which can be calculated as the **effected_probabilities(1)** divided by 2. We then use the index at which probability is equal to our target probability to find our required vaccination rate.

Running the corresponding file in MATLAB we find that the level of vaccination needed to halve the probability computed in part iv) is approximately 90%, evaluated to the nearest integer.

3 Modified Stochastic Model.

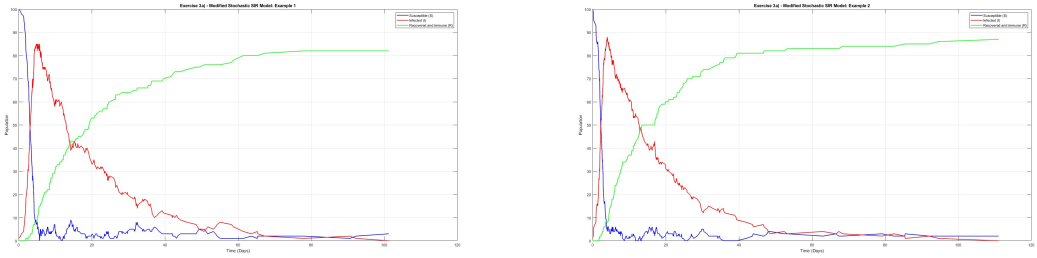
3.1 Part a)

The model of Question 2 is based on the assumption that recovered individuals are immune to infection. Assume that each recovered individual, with equal probability, is immune or susceptible and modify your code for Question 2 to take this assumption into account.

Simulate this modified process with the same parameter values and initial conditions as in Question 2 and produce graphical output showing two independent examples of the process.

See file ex3a.m for the corresponding code and Figure 6 for the resulting plots.

In order to modify our stochastic model to account for immunity we alter the code from part a) of Question 2 to nest an **if, else** loop within the final **else** statement. Since each recovered individual has equal probability of gaining immunity or becoming susceptible again we determine the events by calculating if the **immune_prob**, which is a scalar described by a uniform random distribution, is less or more than 0.5 (the former case refers to the susceptible event, making the latter the immune event). The SIR populations are updated after each event and corresponding time step, as in the un-modified model, and plots for two independent simulations are generated, as can be seen in Figure 6.



(a) First example (b) Second example

Figure 6: here we have generated two independent examples of the modified stochastic SIR model of epidemic as a sum of three independent Poisson processes.

3.2 Part b) section i)

Using the modified model, and for a suitably large sample size, calculate the mean duration of the epidemic (i.e. the time needed for the number of infected individuals to drop to zero) and plot the histogram.

See file ex3b_i.m for the corresponding code and Figure 7 for the resulting plot.

In a similar fashion to part a) of the same Question, we modified the code from part i) of Question 2b to account for the added immunity and reinfection.

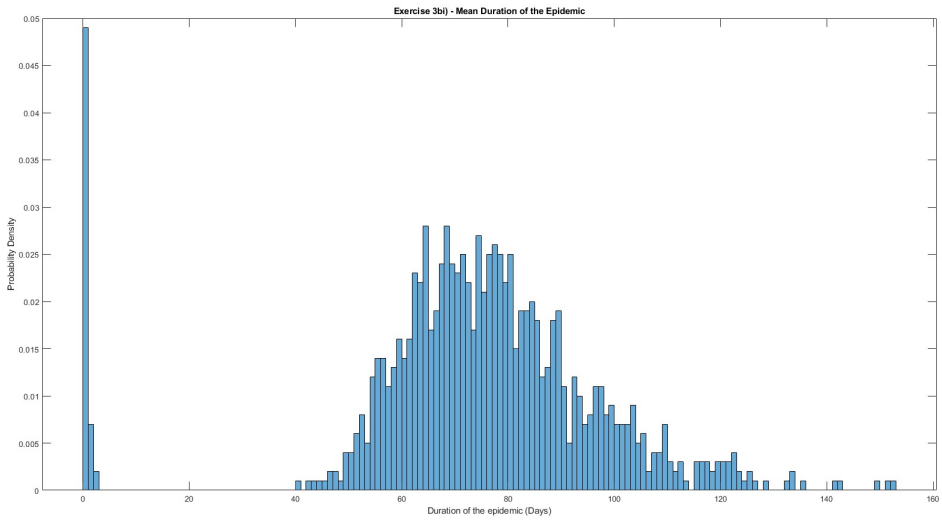


Figure 7: for this particular example we calculated the mean duration as 74 days, rounded to the nearest day.

Despite the modifications to the model, we see that similarly to the simulations from section i) of Question 2b (Section 2.2) we find a spike in the beginning of the histogram, indicating that for some simulations the epidemics ended within a day of their starting. Apart from the general shape of the histogram being similar to that produced by the un-modified model in section i) of Question 2b (Figure 4), we find that the mean duration of the epidemic is approximately 1.5x longer when recovered individuals have the chance of becoming reinfected.

3.3 Part b) section ii)

Using the modified model, and for a suitably large sample size, calculate the mean of the total number of deaths due to the epidemic and compare it with the prediction of the deterministic model of Question 1.

See file ex3b_ii.m for the corresponding code.

In a similar fashion to part a) of the same Question, we modified the code from part ii) of Question 2b to account for the added immunity and reinfection.

Using the same initial conditions and parameter values that we assigned for Questions 1 and 2 (Section 1), we found that, for a sample size of 1000, the mean total number of deaths due to the epidemic for our modified stochastic SIR model is around 15, which is larger than the estimates found in our deterministic and original stochastic models. We credit this increase to the fact that, in our modified model, we have a portion of the recovered individuals re-entering the susceptible population, and which consequentially results in some individuals becoming re-infected and possibly dying. Hence our observation of a higher mean number of deaths is expected.

3.4 Part b) section iii)

Using the modified model, and for a suitably large sample size, calculate the mean time needed for the epidemic to reach its peak (i.e. the maximum number of infected individuals).

See file ex3b_iii.m for the corresponding code and Figure 8 for the resulting histogram.

In a similar fashion to part a) of the same Question, we modified the code from part iii) of Question 2b to account for the added immunity and reinfection.

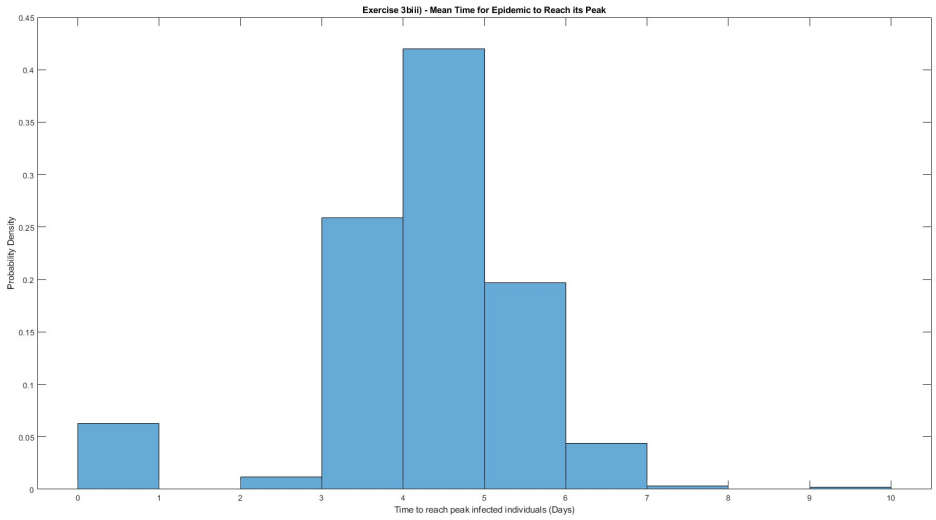


Figure 8: for this particular example we calculated the mean time needed for the epidemic to reach its peak over 1000 samples is 4.2 days (rounded to two significant figures).

Once again we observe the spike at the beginning of our histogram, as is expected due to the nature of stochastic modelling. With regards to the shape of the histogram, as well as the mean time calculated from the model, we don't see much of a difference between the stochastic model that accounts for immunity and the un-modified model.

3.5 Part b) section iv)

Using the modified model, and for a suitably large sample size, Approximate the probability that the number of infected individuals at the peak will be more than 10.

See file ex3b_iv.m for the corresponding code.

In a similar fashion to part a) of the same Question, we modified the code from part iv) of Question 2b to account for the added immunity and reinfection.

Running the corresponding code in MATLAB we find that, similarly to that of the un-modified stochastic model from part iv) of Question 2b (Section 2.5), the probability that the number of infected individuals will be more than 10 for 1000 samples ranges from 0.93 to 0.95. This indifference between the results gathered from the two models can only be expected as with the added consideration of reinfected we are adding more infected individuals to the population, and thus the probability that the number of infected individuals will be more than 10, which was already quite high, will either remain the same or increase, naturally.

3.6 Part b) section v)

If there is a vaccine which is 100% effective, what level of vaccination (a percentage of vaccinated population) is needed to reduce the probability computed in part (iv) by a factor of 2?

See file ex3b_v.m for the corresponding code, which may take a little longer to run than the previous files.

In a similar fashion to part a) of the same Question, we modified the code from part v) of Question 2b to account for the added immunity and reinfection.

Running the corresponding file in MATLAB we find that, similarly to the result computed in part v) of Question 2b (Section 2.6), the level of vaccination needed to halve the probability computed in part iv) is approximately 90%, evaluated to the nearest integer. This result is reasonable and, arguably, expected, due to the incredibly similar probabilities computes in section iv) of both Questions.

3.7 Part c)

Compare your results for the two stochastic models. In your opinion, which model is better? What other modifications can you propose?

Considering the strengths and limitations of both models I would personally argue that the model from Question 3, though not perfect, is the better model of the two. Whilst the model from Question 2 draws its strengths from its simplicity, its ultimate drawback is its assumption of post-recovery immunity for all individuals. By modifying this to include immunity variation for recovered individuals, and therefore arrive at the model from Question 3, we have already enhanced the models reliability. However, with this change comes further complexity to the model and its analysis, as well as the need for more data.

Possible modifications could include incorporating how **seasonal variation** could influence the susceptibility of a population; for example, if the epidemic in question was a similar strain to the influenza virus it would be worth modifying the equations to account for the higher virus activity during the winter months [1]. Additionally, considering the **duration** at which an individual, or a population, is immune to the epidemic after vaccination would prove beneficial to enhancing the models reliability [2]. Extending the model to encompass **spatial dynamics**, such as densely-populated vs sparsely-populated regions, and the movement between these regions, would also be worth considering [3].

References

[1] Gabriele Neumann and Yoshihiro Kawaoka, *Seasonality of influenza and other respiratory viruses*, *EMBO Molecular Medicine*, vol. 14, no. 4, p. e15352, 2022. `https://doi.org/10.15252/emmm.202115352`

[2] M. Rabiu and S.A. Iyaniwura, *Assessing the potential impact of immunity waning on the dynamics of COVID-19 in South Africa: an endemic model of COVID-19*, *Nonlinear Dynamics*, vol. 109, pp. 203–223, 2022. https://doi.org/10.1007/s11071-022-07225-9

[3] John C. Lang, Hans De Sterck, Jamieson L. Kaiser, Joel C. Miller, *Analytic models for SIR disease spread on random spatial networks*, *Journal of Complex Networks*, vol. 6, no. 6, pp. 948–970, December 2018. https://doi.org/10.1093/comnet/cny004