Population Modelling
III. SIR Epidemic Model

Peter McIntyre
School of Science UNSW Canberra

Contents

Note: Section 3 of this study uses programs for a TI-84/CE graphics calculator.

Population Modelling Series

Population Modelling I. Exponential Growth looks at simple exponential-growth models and problems; suitable for Years 7–9.
Population Modelling II. Non-Exponential Models looks at other types of population models, both continuous and discrete; suitable for Years 10–12.
Population Modelling III. SIR Epidemic Model works through a classic model of the spread of an epidemic; suitable for good students in Year 12 .

1 Introduction to Epidemic Modelling

Based on www.stat.columbia.edu/ regina/research/notes123.pdf.

The modelling of infectious diseases is a tool that is used to study the mechanisms by which diseases spread, to predict the future course of an outbreak and to evaluate strategies to control an epidemic.

Infectious agents have had decisive influences on the history of mankind. Fourteenth-century Black Death took the lives of about a third of Europe's population at the time. Thucydides described the Plague of Athens (430–428 BC): 1,050 of 4,000 soldiers on an expedition died of a disease. He gave a detailed account of symptoms, some so horrendous that the last one – amnesia – seems a blessing (Bailey, 1975). An interesting feature of this account is that there was no mention of person-to-person contagion, which we now suspect with most new diseases; it was not until the 19th century that this was beginning to be discussed.

The practical use of epidemic models relies heavily on the assumptions underlying the models. A reasonable model does not have to include all possible effects but should incorporate the main mechanisms that influence disease propagation in the simplest possible fashion. Great care should be taken before epidemic models are used for prediction of real phenomena. However, even simple models should, and often do, pose important questions about the underlying mechanisms of infection spread and possible means of control of the disease or epidemic.

The classical papers by Kermack and McKendrick (1927, 1932 and 1933) have had a major influence on the development of mathematical models for disease spread, and are still relevant in many epidemic situations. The first of these papers laid out a foundation for modelling infections which confer complete immunity after recovery (or, in the case of lethal diseases, death). If infected individuals are introduced into a large population, a basic problem is to describe the spread of the infection within the population as a function of time. One of the most important questions is whether the epidemic comes to an end only when all the initially susceptible individuals have contracted the disease or if some interplay of infectivity, recovery and mortality factors may result in epidemic "die out" with many susceptibles still present in the unaffected population.

In their first paper, Kermack and McKendrick started with the assumption that all members of the community were initially equally susceptible to the disease, and that complete immunity was conferred after the infection. The population was divided into three distinct classes: susceptibles S, healthy individuals who can catch the disease; infecteds I, those who have the disease and can transmit it; and removed R, individuals who have had the disease and are now immune to the infection (or removed from further propagation of the disease by some other means). Schematically, the individual goes through consecutive states S→I→R; such models are often called the SIR models.

Bailey, NTJ (1975). *The mathematical theory of infectious diseases and its applications.* 2nd ed. Hafner Press, New York.

Ethier, SN and Kurtz, TG (1986). *Markov Processes. Characterization and Convergence.* Wiley, New York.

Kermack, WO and McKendrick, AG (1927). Contributions to the mathematical theory of epidemics, part i. *Proceedings of the Royal Society of Edinburgh. Section A Mathematics* **115**, 700–721.

Kermack, WO and McKendrick, AG (1932). Contributions to the mathematical theory of epidemics, ii – the problem of endemicity. *Proceedings of the Royal Society of Edinburgh. Section A Mathematics* **138**, 55–83.

Kermack, WO and McKendrick, AG (1933). Contributions to the mathematical theory of epidemics, iii – further studies of the problem of endemicity. *Proceedings of the Royal Society of Edinburgh. Section A Mathematics* **141**, 94–122.

2 The Problem

Differential equations (continuous-time models) are used to model the spread of an epidemic through a population in Section 3; the model here is the SIR model. Several approaches are used to obtain information and to make predictions: numerical solutions are obtained with a TI graphics calculator; analytical and graphical techniques extract more information from the equations. The continuous SIR model is then applied to an epidemic scenario.

Difference equations (discrete-time models) are used to model the spread of an epidemic in Section 4. The discrete SIR model is given and applied to the same epidemic scenario.

2.1 Your Report

Compile your answers to the problems below, together with supporting plots[1] and any other relevant material, into a coherent report that can be understood by one of your Maths colleagues.

3 Continuous SIR Model

3.1 Numerical Approach

Infectious diseases such as influenza spread rapidly, then 'burn out' over a period of weeks or months. Natural births and deaths in the population can therefore be ignored on this timescale.

In the SIR model, the population is divided into three categories, **susceptible**, **infected** and **removed**, with $S(t)$, $I(t)$ and $R(t)$ being the respective numbers in each at time t days. Susceptible persons can catch the disease; infected persons are infectious and can therefore transmit the disease to susceptibles; removed persons have recovered, are then assumed immune to the disease and cannot spread it.

Let t be measured in days. The differential equations modelling this situation are

$$\frac{dS}{dt} = -\beta S I \tag{1}$$

$$\frac{dI}{dt} = \beta S I - \alpha I \tag{2}$$

$$\frac{dR}{dt} = \alpha I \tag{3}$$

where α (the recovery rate) and β (the transmission rate) are positive constants. S, I and R cannot be negative.

[1] You can print out graphics-calculator screens or save them as png files using the calculator-to-computer cable and TI-Connect/Ti-Connect CE software.

βSI in Eq. (1) is the number of new infections per day. These people move to the infected category, thereby decreasing S in Eq. (1) and increasing I in Eq. (2).

βSI models the transmission of the disease. The expression here is one of several possible choices. It is called the *Law of Mass Action*, which states that the rate of new cases is proportional to the product of S and I. Transmission occurs when an infected meets a susceptible, so intuitively, transmission should depend somehow on the numbers of each. In this form, the number of contacts each individual makes per day is assumed to be large in a large population, and small in a small population.

People recover from the disease according to the second term on the RHS of Eq. (2), and move to the removed category. Assuming all in R are alive, no equation term represents births or deaths in this simplified SIR model. Deaths from the disease could be included in the removed category if necessary.

To make the problem ready for a numerical solution, we need to specify values for the constants α and β, and for the initial conditions $S(0)$, $I(0)$ and $R(0)$. For our calculations, we assume $R(0)=0$, so that $S(0)+I(0)=N$, the total population, which is constant (see Question 6).

We assume the epidemic is over once I has dropped below 1.

The next few exercises aim to help us gain some intuition into the model. They use the TI-84/CE graphics-calculator program SIR/SIRCE, which solves the SIR system of differential equations numerically using the modified Euler's method.

1. Read the instructions for the SIR/SIRCE program in Appendix 1.

Run SIR/SIRCE and choose TIME PLOT.

Assume a total population $N=1000$.

Run the model for 50 days (Tmax).

Set $\alpha=0.2$ and $\beta=0.0008$.

As initial conditions, set $S(0) = 999$, so that $I(0) = 1$ (you need at least one infective). This models introducing the disease into essentially a fully susceptible population.

Start with a step length H of 1 day.

Check that you get $S(50)=20.8$ and $I(50)=1.0$, so that $R(50)=1000-S(50)-I(50)=978.3$, all rounded to 1 decimal place.

SIR/SIRCE graphs S, I and R as functions of time on the one plot.

What do the curves tell you about this epidemic?

2. Decreasing the step length increases the accuracy of the results but the calculations take longer. Decrease the step length by a factor of 10 to check the accuracy of your results in 1; running the program with just one step length does not tell you how accurate your results are.

Give values for $S(50)$, $I(50)$ and $R(50)$, rounded to an appropriate number of decimal places, and **accurate** to this number of decimal places.

3. The **size** of an epidemic is the total number of people infected during the epidemic.

What was it for the case above (after 50 days)?

How many people out of the population of 1000 remained free of the disease?

Hint: The numbers displayed after the plot will help you here.

Check the accuracy of your results by

- running the model for say 75 days to see if the values for S and I change significantly after 50 days;
- decreasing the step length to check the accuracy of the results.

4. Run SIR/SIRCE, keeping $\alpha = 0.2$ fixed and varying β.

Try some values of β centred on 0.0008, such that there are significant changes in the graphs.

Describe how the curves change for different values of β, particularly with regard to the duration of the epidemic, the time to peak I and the maximum value of I.

Remembering that β determines the rate of transmission of the disease, explain the changes.

5. Run SIR/SIRCE, keeping $\beta = 0.0008$ fixed and varying α.

Try some values of α centred on 0.2, such that there are significant changes in the graphs.

α is the reciprocal of the mean infective period, the mean time that a person who has caught the disease remains infectious.

Describe how the curves change for different values of α, particularly with regard to the duration of the epidemic, the time to peak I and the maximum value of I.

Remembering that α determines how long, on average, a person is infectious, explain the changes.

3.2 Analytical Approach

Section 3.1 showed that it is relatively easy to get approximate solution curves for an epidemic, and to understand how changing the parameter values affects the solution. Although we cannot solve Eqs. (1), (2) and (3) algebraically, we can find other algebraic ways of obtaining information about the epidemic, and identify some important summary quantities. This we do in this section.

You should do the questions here for general values of the parameters α, β and N, and general initial values $S(0)$ and $I(0)$. But by all means use numerical values to check your results.

6. Since there are no deaths from the disease in our model, and natural births and deaths are ignored, we expect that the total population, $N = S + I + R$, will not change.

Show that, according to the model, Eqs. (1), (2) and (3), N really is a constant. *Hint*: If N is a constant, what do we know about dN/dt?

Explain why solving three DEs is then unnecessary (solving two will be sufficient).

7. The rising phase of an epidemic (as seen in an $I(t)$ curve) is a worrying time, which lasts until the epidemic 'peaks' and enters a waning stage.

Show from Eq. (2) that the number of cases (the number of persons who have caught the disease $S(0)-S$) up to the time of peak I can be predicted in advance, if reliable estimates of α and β are available. *Hint*: Think max/min I values.

Compare this theoretical value with the approximate value from your graph in Question 2.

8. According to Eq. (1), the number of susceptibles never increases. Use Eq. (2) and your thinking from Question 7 to decide what happens if the disease is introduced into a population which has $S(0) < \alpha/\beta$. What if $S(0) > \alpha/\beta$?

Hint: Think about dI/dt at $t = 0$ in the case when there is/isn't an epidemic.

The *basic reproductive number*, R_0, is defined in the SIR model as $R_0 = \beta S(0)/\alpha$. What is the threshold value of R_0 for an epidemic?

3.3 Graphical Approach: Phase Plots

9. Now use PHASE PLOT in SIR/SIRCE — this generates an (approximate) graph of I versus S, called a **phase portrait**.

Take $\alpha = 0.2$, $\beta = 0.0008$, $N = 1000$ (as in Questions 1, 2 and 3) and several different sets of initial conditions. Start with $S(0) = 999$ ($I(0) = 1$). Then take $S(0) = 900$ and $S(0) = 800$ using Option 3 in the NEXT menu. Choose the number of points (time steps) so that the epidemic is over ($I < 1$) by the final point.

You may need to change the window in the program so that the graph fills the screen. Don't forget to check for accuracy before proceeding to the next step.

Print out the plot. Label points that are significant in the course of the epidemic, such as the initial point, the point of maximum I and the point at which the epidemic is taken as finished.

Give the value of R_0 for each curve.

Describe what these phase portraits tell you about the possible epidemics with these values of α, β and N.

10. As $S(t)$ always decreases as t increases, S can be used as a time-like variable.

Use the chain rule[2] to show from Eqs. (1) and (2) that

$$\frac{dI}{dS} = -1 + \frac{\alpha}{\beta}\frac{1}{S}.$$

Solve this DE[3] to get

$$I(t) = -S(t) + \frac{\alpha}{\beta}\ln\big(S(t)\big) + D, \quad (4)$$

where D is an arbitrary constant. In this case therefore, we can find an exact expression for the phase portrait.

[2] $\frac{dI}{dS} = \frac{dI}{dt}\frac{dt}{dS} = \frac{dI}{dt}/\frac{dS}{dt}$.
[3] or show that the given solution is a solution

Choose $\alpha = 0.2$, $\beta = 0.0008$, and use your calculator or other graphics to plot I as a function of S (as a Y(X) graph) for some different values of D.

What is a sensible window*?*

What determines the value of D in any given situation?

What value of D corresponds to the graph you generated in Question 9?

Do the two graphs then give the same values?

For most systems of DEs, we cannot find an exact expression like Eq. (4), and must therefore use a numerical approach, like we did in Question 9.

11. A special case of Eq. (4) is the **final-size equation**

$$I_\infty - I_0 = S_0 - S_\infty + \frac{\alpha}{\beta} \ln\left(\frac{S_\infty}{S_0}\right), \tag{5}$$

where $S_0 = S(0)$, $I_0 = I(0)$, and I_∞ and S_∞ are the respective limits as $t \to \infty$.

Explain how this equation follows from Eq. (4).

It is usual to assume that $I_0 = I_\infty = 0$ (absence of the disease before and after the epidemic).

In Question 3 you found the final size of an epidemic for $\alpha = 0.2$, $\beta = 0.0008$ and $N = 1000$.

Here, predict the size using Eq. (5) and compare the two values.

Hint: You will have to use graphics or a numerical solver to find S_∞ from Eq. (5).

3.4 Case Study: On a Pacific Island

The population on an idyllic Pacific island is stable at $N = 200$. The birth and death rates are small, and can be ignored. The island's post office employs five people,[4] who are all clustered around when a Christmas hamper from the USA is opened for customs inspection. Unfortunately the white powder in the hamper is not artificial snow. Within a day, during which the workers mix as usual with other persons on the island, all five workers are infected with a severe and mysterious virus.

You are the only person on the island with some knowledge of epidemiology, gained from Maths in Year 12. The island's Chief Medical Officer needs to know how many people might need to be treated for the virus and whether to call for emergency hospital facilities. The island's hospital can cope with an absolute maximum of 50 patients at any one time. The Medical Officer asks for your help in predicting the course of the epidemic.

Based on the post-office experience, you assume that the incubation period is less than a day, and can be ignored. The mean infectious period, $1/\alpha$ in the SIR model, needs to be estimated, as does β.

1. Based on other viruses, you take the mean infectious period, $1/\alpha$, to be 5 days and the transmission parameter $\beta = 0.0025$.
What is the approximate duration of the epidemic ($I < 1$) with these parameters? Vary *Tmax* to find this accurately.
Could the hospital cope with the number of patients predicted by this model?

2. After a week you have some actual data. You find that the mean infectious period is actually 4 days, not 5, and that $I(8) = 30$ and $S(8) = 140$.

Change the value of α, and experiment with different values of β to match this $I(8)$ value.
What is R_0 for this model?

What then are your predictions (theoretically and/or numerically based) for

(a) when the epidemic will peak?

(b) the duration of the epidemic?

(c) the number of people who are untouched by the disease at the end?

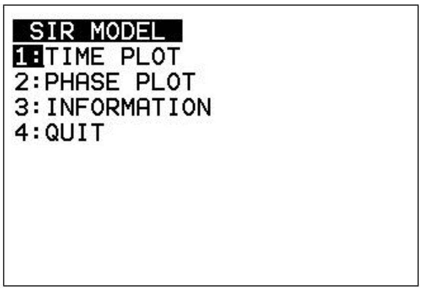
Will the hospital be able to cope?

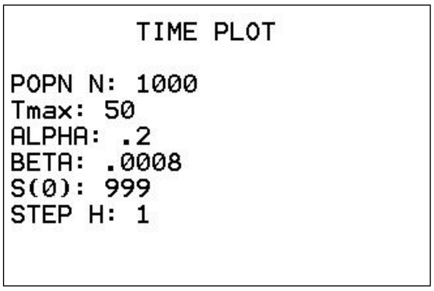
[4] The island's economy is based around issuing stamps, banking and processing asylum seekers.

3.5 Appendix 1: SIR/SIRCE program

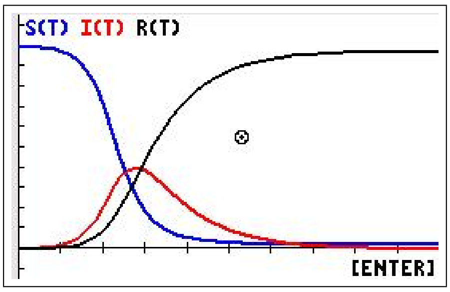
Using the programs

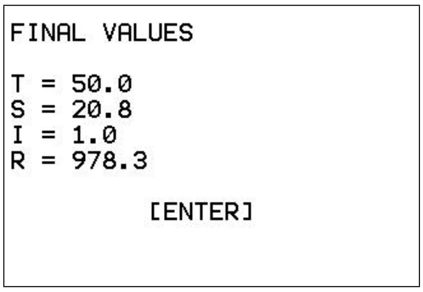
1. Run the program. TIME PLOT (below left) plots S, I and R as functions of time. PHASE PLOT plots I as a function of S. INFORMATION gives the SIR equations.



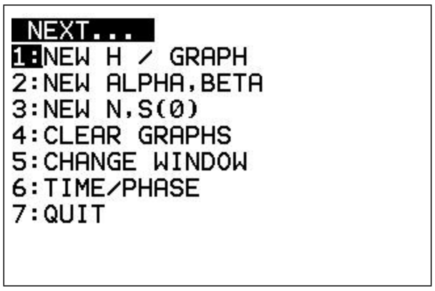


2. Input the appropriate parameters on the prompts (above right), including a step length H, pressing enter after each. Decreasing H will in general improve the accuracy but the calculations will take longer. Starting with a step length of 1 day (H = 1) over 50 days is usually about right. The resulting plot is shown below left.





3. Press enter when the graph has been plotted to see the final calculated values for S, I and R (above right); press enter again to display the NEXT menu (below).



Here you can:

- change the H value;
- change the values of α and β;
- change the values of N and $S(0)$;
- clear the graphs for another plot; if you don't, the new plot will overlay the current one;
- change the window within the program
- change from a TIME PLOT to a PHASE PLOT, or vice versa, with the same parameters;
- QUIT.

4 Discrete SIR Model

4.1 The Model

If we use a discrete-time model, rather than a continuous-time model (the differential equations of Section 3), to describe the spread of such a disease, we obtain a SIR model with the difference equations[5]

$$S_{n+1} = S_n - \beta S_n I_n \tag{6}$$

$$I_{n+1} = I_n + \beta S_n I_n - \alpha I_n \tag{7}$$

$$R_{n+1} = R_n + \alpha I_n, \tag{8}$$

where S_n, I_n and R_n are the numbers of susceptible persons, infected persons and recovered persons, respectively, after n time intervals, and α and β are constants. The time interval in the case here is 1 day.

4.2 Case Study: On a Pacific Island

1. Use the same data that you used in Section 3.4 for the corresponding continuous model. Draw graphs (time plots) of S_n and I_n vs n (time in days). Details of how to do this on a TI graphics calculator are in Appendix 2.

What is the approximate duration of the epidemic ($I < 1$)? How many persons don't catch the disease?

When will the peak of infection occur according to the discrete model? Will the hospital cope?

For what values of S is I increasing? decreasing?

Show, from the equation for I above, that the theoretical maximum value[6] of I in the discrete model occurs when $S = \alpha/\beta$.

Compare this value with the value you obtained from your graph. Use [trace] to do this. *Why might not these values be exactly the same?*

2. Use the revised value, $\alpha = 0.25$, as in Question 2 of Part 4. Again, experiment with different β to find the value that gives $I_8 = 30$.

Hint: For each β value, use either a graph and [trace] or [table][7] (easier) to find I_8 until you obtain the actual value (or as close as possible).

When you have found the right β value, use the model to predict how many days it will take before the disease dies out and when the peak of infection will occur. *Will the hospital be able to cope according to this model? How many people will avoid catching the disease?*

3. Compare the results from the discrete model with those from the continuous model.

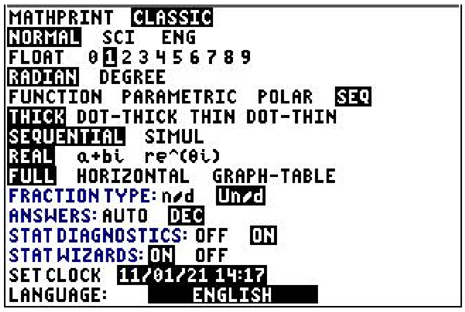
[5]See Population Modelling II for other examples.
[6]At a maximum or minimum, there is no change in I, so that $I_{n+1} = I_n$. This corresponds to the condition $dI/dt = 0$ in the continuous case.
[7]see the end of Appendix 2

4.3 Appendix 2: Sequence Graphing

Population Modelling II. Non-Exponential Models has an introduction to discrete models, in particular the Discrete Logistic Model.

The TI-84/CE calculator has three built-in sequences u, v and w (on the [7], [8] and [9] keys).[a] To access them, press [mode] and select *Seq* with the cursor and [enter].
Select 1 decimal place as well (second line).

[a]most graphics calculators have this facility



Now press [y=] and you will see where to define the sequence functions. Note that the sequences here on a TI-84 have $u(n)$ defined in terms of $u(n-1)$, $v(n-1)$, etc. The equations for S and I written in this form are

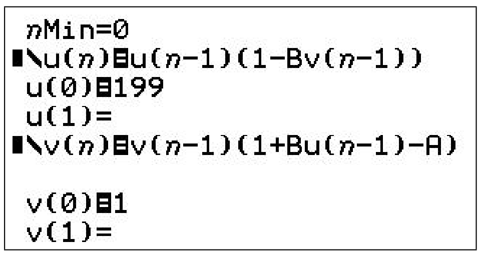
$$\begin{aligned} S_n &= S_{n-1} - \beta S_{n-1} I_{n-1} = S_{n-1}\big(1 - \beta I_{n-1}\big). \\ I_n &= I_{n-1} + \beta S_{n-1} I_{n-1} - \alpha I_{n-1} = I_{n-1}\big(1 + \beta S_{n-1} - \alpha\big). \end{aligned}$$

With $S_n \to u(n)$, $I_n \to v(n)$, $\alpha \to A$ and $\beta \to B$, we have the equations in calculator variables as

$$\begin{aligned} u(n) &= u(n-1)\big(1 - Bv(n-1)\big). \\ v(n) &= v(n-1)\big(1 + Bu(n-1) - A\big). \end{aligned}$$

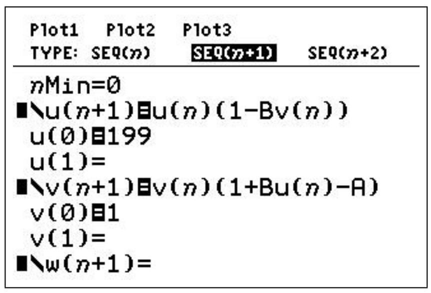
Enter these equations into the calculator as shown below.

In Seq mode, the [X,T,θ,n] key now gives n, the independent variable for the sequence functions. You can't use the letter N.



On the TI-84CE calculator, the sequences can also be input in the original form, Eqs. (6)–(8) (note the second line on the screen below).

$$\begin{aligned} u(n+1) &= u(n)\big(1 - Bv(n)\big). \\ v(n+1) &= v(n)\big(1 + Bu(n) - A\big). \end{aligned}$$

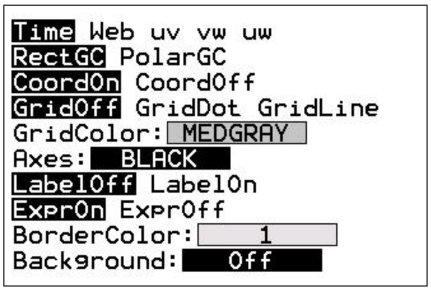


Set nMin=0, that is we start with S_0, I_0, etc. Then S_1 is the S value after 1 day, etc.

The initial conditions are contained in u(nMin) and v(nMin). Note the curly brackets here.

Time plots

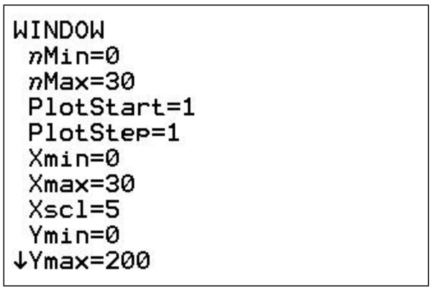
Press 2nd format and select *Time* as shown below.[8] The X axis is n (time in days) and the Y axis S and I.

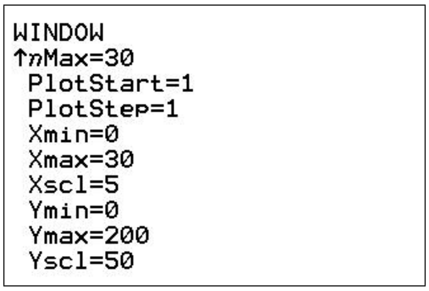


The final step before plotting is to choose a window. As well as the usual window settings for the X and Y axes, we have to specify nMax, the maximum n value we want.

Start by plotting 30 points, corresponding to running the system through 30 days. Note that we must also set Xmax to 30.

Since $N = 200$ is the maximum value for both S and I, setting Ymax to 200 seems like a good starting point.





Store the values of α and β in memories A and B respectively. Press graph to plot graphs of S and I versus n.

Adjust the window, if necessary, until the graphs more or less fill the screen.

Use trace to explore the values. You can go directly to the point with a particular n value by just typing in the n value and pressing enter.

Adjust nMax and Xmax (both corresponding to Tmax) in window so that the disease runs its full course ($I < 1$).

You can also see the values of S and I in a table by pressing table (2nd graph).

If the n values don't start at 1 and/or don't increment in steps of 1, fix this in tblset (2nd window).

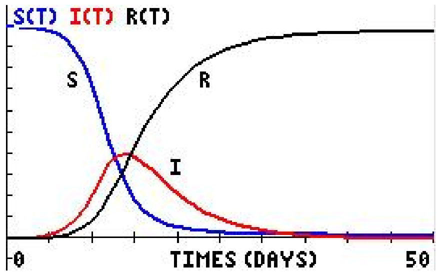
[8] uv would give a phase plot

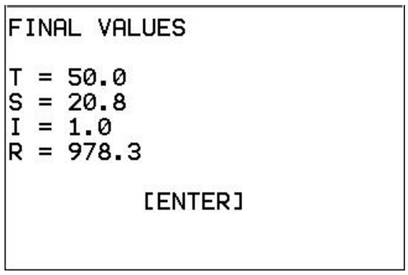
5 Solutions

Diagrams here were done on a TI-84CE graphics calculator.

5.1 Numerical Solution of an SIR Model

1. The final values are as given in the question.

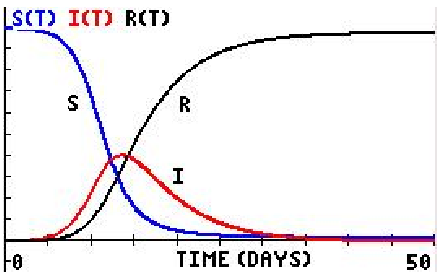


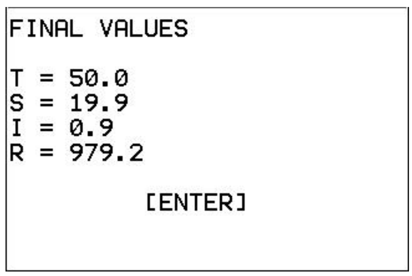


We can see from the graphs, that I increases to a maximum, then decreases gradually to 0.

S decreases, slowly at first, then more rapidly, eventually levelling out to some small value (about 20, given that $S(50) = 20.8$).

2. Using the SIR program with the parameters as in Question 1 but with a step length H of 0.1 gives the following graph and final values.





The graphs for S and I, and the values for $S(50)$ and $I(50)$ are the same as in Question 1 when rounded to integer values.

Running SIR with step lengths H of 1, 0.1 and 0.001 days gives the following values after 50 days. We expect the accuracy of the numerical method to increase as we decrease H.

H	$S(50)$	$I(50)$	$R(50)$
1	20.8	1.0	978.3
0.1	19.9	0.9	979.2
0.01	19.9	0.9	979.2

The values change slightly as we decrease the step length, but as we are going to round to integer values, the changes are not significant.

We have, accurate to the nearest integer (0 decimal places):

$S(50) = 20$

$I(50) = 1$

$R(50) = 979$

3. The following table gives the values of S, I and R after 75 days to check if there is any significant change after 50 days. The accuracy of these values is also checked by running the program with steps of 1 day and 0.1 day.

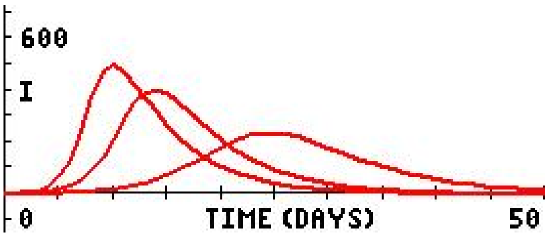
H	$S(75)$	$I(75)$	$R(75)$
1	20.7	0.0	979.3
0.1	19.8	0.0	980.2

The overall result remains the same: 980 people have caught the disease (the size of the epidemic); 20 people do not catch the disease.

After 75 days, the one person still infective after 50 days has recovered.

4. We first keep $\alpha = 0.2$ and vary β. The table below shows the values of $S(50)$, $I(50)$ and $R(50)$ for each value of β.

$\alpha = 0.2$; $S(0) = 999$; $I(0) = 1$; $R(0) = 0$; $H = 1$. Values rounded to nearest integer.



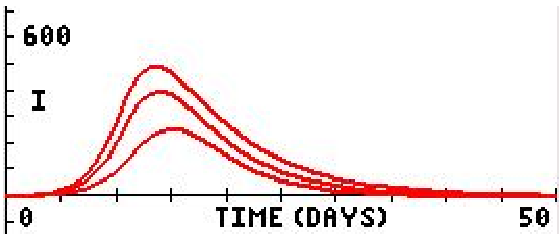
β	$S(50)$	$I(50)$	$R(50)$
0.0005	113	14	873
0.0008	21	1	978
0.0011	5	0	995

The left-hand curve in the figure above has $\beta = 0.0011$, the middle curve $\beta = 0.0008$ and the right-hand curve $\beta = 0.0005$.

From the table and graphs, we see that, the larger the β value, the larger the maximum value of I, the earlier it is reached and the shorter the duration of the epidemic.

The larger the β, the faster the spread of the disease, so we would anticipate the above findings.

5. Now we keep $\beta = 0.0008$ and vary α. The other parameters are as in Question 4.



α	$S(50)$	$I(50)$	$R(50)$
0.15	5	4	991
0.2	21	1	978
0.3	89	0	911

The top curve in the figure above has $\alpha = 0.15$, the middle curve $\alpha = 0.2$ and the bottom curve $\alpha = 0.3$.

From the table and graphs, we see that, the larger the α value, the smaller the maximum value of I. The time to maximum I and the duration of the epidemic are both slightly longer with larger α, but the differences are small compared to those seen when changing β in Question 4.

The larger the α, the faster people recover from the disease, so we would expect there to be fewer with the disease at any one time, compared with cases in which α is smaller.

5.2 Analytical Approach

6. We are given that

$$\begin{aligned} N &= S + I + R. \\ \therefore \frac{dN}{dt} &= \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} \\ &= -\beta SI + \beta SI - \alpha I + \alpha I \qquad \text{from Eqs. (1), (2) and (3)} \\ &= 0. \end{aligned}$$

Therefore, N is a constant. If we know two of the variables, say S and I, we can calculate the third by $R = N - S - I$. Therefore, we only need to work with the DEs involving S and I.

7. Equation (2) is

$$\begin{aligned} \frac{dI}{dt} &= \beta SI - \alpha I \\ &= I\big(\beta S - \alpha\big) \\ &= 0 \quad \text{when } I \text{ is a maximum.} \end{aligned}$$

As $I \neq 0$ at the maximum, we must then have

$$S = \frac{\alpha}{\beta}.$$

The total number of cases up to this time is then $S(0) - S = S(0) - \alpha/\beta$.

In Questions 1 and 2, $\alpha = 0.2$ and $\beta = 0.0008$, giving the theoretical value of S at maximum I of 250, and the total number of cases by then as $999 - 250 = 749$.

Reading from the graph in Question 1 or 2 (by moving the cursor around), we see that the maximum value of I occurs after about 14 days, with the corresponding S value close to 250, in agreement with the theoretical value.

8. We have

$$\begin{aligned} \frac{dI}{dt} &= \beta SI - \alpha I \qquad \text{for all time } t \\ &= I\big(\beta S - \alpha\big) \qquad \text{for all time } t \\ &= I(0)\big(\beta S(0) - \alpha\big) \qquad \text{at } t = 0. \end{aligned}$$

If $\beta S(0) - \alpha < 0$ or, equivalently, $S(0) < \alpha/\beta$, then $dI/dt < 0$ at $t = 0$.

Since S decreases with time, $S(t) \leq S(0)$ for all t. This means that, for all time, $\beta S - \alpha < 0$, so that $dI/dt < 0$.

Therefore I decreases with time from its intial value, and there is no epidemic.

Conversely, if $S(0) > \alpha/\beta$, I increases with time from its initial value, and there is an epidemic.

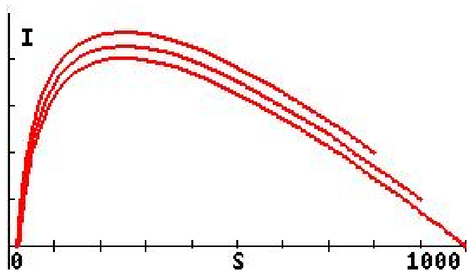
The condition $S(0) \lesseqgtr \alpha/\beta$ is equivalent to the basic reproductive number $R_0 \lesseqgtr 1$. The threshold for an epidemic is therefore $R_0 = 1$, with an epidemic if $R_0 > 1$.

5.3 Graphical Approach: Phase Plots

9. Running SIR/SIRCE with $\alpha = 0.2$, $\beta = 0.0008$, $N = 1000$ for 50 days with a step $H = 0.5$ gives the following phase portrait $I(S)$.

The graphs have $S(0) = 999$, $I(0) = 1$ (bottom curve); $S(0) = 900$, $I(0) = 100$ (middle curve) and $S(0) = 800$, $I(0) = 200$ (top curve). The respective R_0 values are 4.0, 3.6 and 3.2.

The window is $[0, 1000, 100] \times [0, 550, 100]$.[9]



The graph with a step $H = 0.05$ is very little different, so the curves with $H = 0.5$ are reasonably accurate.

The first plot starts at $(S, I) = (999, 1)$ and initially moves up (I increasing) and to the left (S decreasing). Eventually I peaks at a theoretical value of $S = \alpha/\beta = 250$, and then decreases to 0 at $t \approx 50$. S always decreases. The final point on the graph ($t = 50$) is $(S, I) = (20, 0)$.

The three epidemic graphs differ in initial values, but show very similar behaviour, being roughly parallel until near the end of the epidemic. They appear to have almost the same final S value.

10. We have from Eqs. (1) and (2),

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta S I \\
\frac{dI}{dt} &= \beta S I - \alpha I. \\
\therefore \quad \frac{dI}{dS} &= \frac{dI/dt}{dS/dt} \\
&= \frac{\beta S I - \alpha I}{-\beta S I} \\
&= -1 + \frac{\alpha}{\beta}\frac{1}{S}.
\end{aligned}
$$

Integrating this with respect to S gives

$$
I(S) = -S + \frac{\alpha}{\beta}\ln(S) + D,
$$

where D is an arbitrary constant.

[9] changed within the program from the initial values

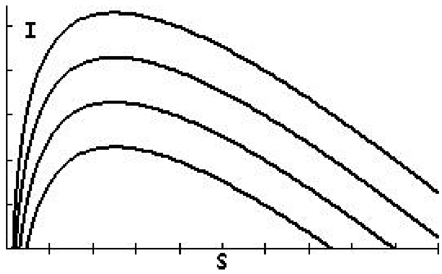
If we use t as the independent variable, we can write this as

$$I(t) \;=\; -S(t) \,+\, \frac{\alpha}{\beta}\,\ln\big(S(t)\big) \,+\, D. \tag{9}$$

The plots here are I versus S. From Question 9, a sensible window is $[0, 1000, 100] \times [-50, 550, 100]$

Some graphs of $I(S)$ with $\alpha = 0.2$, $\beta = 0.0008$ and various D values are graphed below. They resemble the phase plots in Question 9.

The D values are, top to bottom, -600, -700, -800 and -900.



The value of D, for given α, β and N, is set by the initial conditions $S(0)$ and $I(0)$.

Setting $S(0) = 999$, $I(0) = 1$, $\alpha = 0.2$ and $\beta = 0.0008$ in Eq. (9) gives

$$\begin{aligned} 1 &= -999 + 250\ln(999) + D. \\ \therefore\; D &= 1 + 999 - 250\ln(999) \\ &= -726.7 \qquad \text{rounded to 1 decimal place.} \end{aligned}$$

The values for $I(t)$ generated by Eq. (9) are the same values to reasonable accuracy as those generated by SIR.

11. Setting $t = 0$ in Eq. (9) and using the notation given in the question results in

$$I_0 \;=\; -S_0 \,+\, \frac{\alpha}{\beta}\,\ln\big(S(0)\big) \,+\, D. \tag{10}$$

Similarly, letting $t \to \infty$ in Equation (9) gives

$$I_\infty \;=\; -S_\infty \,+\, \frac{\alpha}{\beta}\,\ln\big(S_\infty\big) \,+\, D. \tag{11}$$

Subtracting Eq. (10) from Eq. (11) gives

$$I_\infty - I_0 \;=\; S_0 - S_\infty + \frac{\alpha}{\beta}\,\ln\left(\frac{S_\infty}{S_0}\right),$$

as required.

If we now let $I_0 = I_\infty = 0$, we have

$$0 \;=\; S_0 - S_\infty + \frac{\alpha}{\beta}\,\ln\left(\frac{S_\infty}{S_0}\right).$$

Given $S_0 = S(0)$, we can find S_∞ and hence find $R_\infty = 1000 - S_\infty$, the number of persons who caught the disease.

With $S_0 = 999$, $\alpha = 0.2$ and $\beta = 0.0008$, we solve

$$0 = 999 - S_\infty + 250 \ln\left(\frac{S_\infty}{999}\right)$$

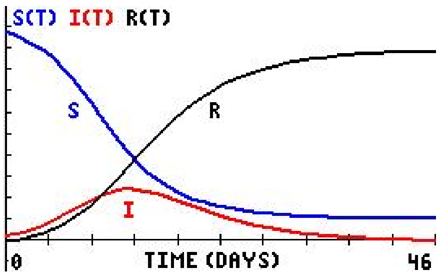
numerically for S_∞ (graph the RHS and find the *relevant* zero).

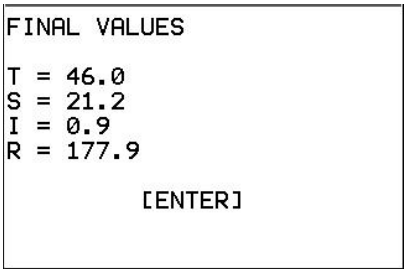
This gives $S_\infty = 20$, rounded to one decimal place.

The size of the epidemic is then $1000 - 20 = 980$, that is 980 persons, the same value we found in Question 3.

5.4 On a Pacific Island

1. Use SIR/SIRCE with $N = 200$, $S(0) = 195$, $I(0) = 5$, $\alpha = 0.2$, $\beta = 0.0025$ and a time of 46 days (after some experimentation). The step length $H = 1$ in the following graph, but $H = 0.1$ gives the same values when rounded to the nearest integer.





We have $S(46) = 21.2$, $I(46) = 0.9$ and $R(46) = 177.9$.

The duration of the epidemic is about 46 days, the maximum number of infected persons (reading from the graph) is about 49 on about Day 13.

If we change Xmax (Tmax) to give a more accurate value for $I_{\max}$ (with $H = 0.1$), we find that the maximum number of infected persons is 48.7, rounded to 49, on Day 13.

The hospital will therefore just cope.

An algebraic method to find $I_{\max}$

If we evaluate Eq. (9) at $S = S_0$, we obtain

$$I(0) = -S(0) + \frac{\alpha}{\beta} \ln\big(S(0)\big) + D. \tag{12}$$

Next, evaluate Eq. (9) at $S = S_{\max}$, the value of S when I is a maximum, $S = S_{\max} = \alpha/\beta$,

$$I_{\max} = -S_{\max} + \frac{\alpha}{\beta} \ln\big(S_{\max}\big) + D. \tag{13}$$

Subtracting Eq. (12) from Eq. (13) gives the following various expressions for $I_{\max}$.

$$\begin{aligned}
I_{\max} - I(0) &= S(0) - S_{\max} + \frac{\alpha}{\beta}\ln\left(\frac{S_{\max}}{S(0)}\right) \\
&= S(0) - \frac{\alpha}{\beta} + \frac{\alpha}{\beta}\ln\left(\frac{\alpha}{\beta S(0)}\right) \\
&= S(0) + \frac{\alpha}{\beta}\left(\ln\left(\frac{\alpha}{\beta S(0)}\right) - 1\right) \\
&= S(0) + \frac{S(0)}{R_0}\left(\ln\left(\frac{1}{R_0}\right) - 1\right) \\
&= S(0)\left(1 - \frac{\ln(R_0)+1}{R_0}\right).
\end{aligned}$$

For the problem here, $S(0)=195$, $I(0)=5$, $\alpha=0.2$ and $\beta=0.0025$.

Therefore, $R_0=\beta S(0)/\alpha\approx 2.44$, so that $I_{\max}\approx 48.7$, the same value we found from the SIR calculations.

2. Use SIR/SIRCE with $N=200$, $S(0)=195$, $I(0)=5$ and $\alpha=0.25$. Try various β values around 0.002 to find which value gives $I(8)=30$.

The table below shows the results (all with $H=0.1$).

$\boldsymbol{\beta}$	$\boldsymbol{I(8)}$	$\boldsymbol{S(8)}$
0.002	12.9	169.9
0.003	38.2	124.1
0.00276	30.8	137.7
0.00275	30.5	138.2
0.00274	30.2	138.8
0.00273	29.9	139.3
0.00272	29.6	139.8

The value of β that gives $I(8)$ closest to 30 is $\beta=0.00273$. Then,

$$R_0 = \frac{\beta S(0)}{\alpha} = \frac{0.00273\times 195}{0.25} \approx 2.13.$$

An algebraic method to find β

We can use the algebraic theory from the previous question to find a value for β, given that $S(0)=195$, $I(0)=5$, $\alpha=0.25$, $S(8)=140$ and $I(8)=30$.

If, instead of evaluating Eq. (13) at $I_{\max}$, we evaluate it at $t=8$, following the same steps as above we have

$$I(8) - I(0) = S(0) - S(8) + \frac{\alpha}{\beta} \ln\left(\frac{S(8)}{S_0}\right).$$

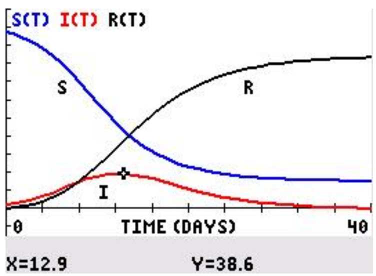
With the above values of the parameters, we then have

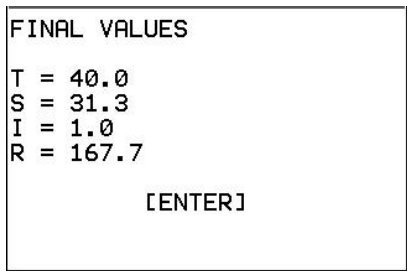
$$30 - 5 = 195 - 140 + \frac{0.25}{\beta} \ln\left(\frac{140}{195}\right),$$

giving $\beta=0.00276$, close to the value $\beta=0.00273$ we found by experimentation.

Note that we used both the given values, $S(8)=140$ and $I(8)=30$, in these calculations, whereas we used only $I(8)=30$ as the criterion in our experimentation above. The difference in the two β values is negligible in our calculations.

With $\alpha=0.25$, $\beta=0.00273$ and plotting over 40 days, we obtain the following graph.





(a) Reading from the graph, the epidemic peaks at about 12 or 13 days, with the maximum number of infected persons about 39.

If we again change Xmax to give a more accurate value for $I_{\max}$ (with $H=0.1$), we find that the maximum number of infected persons is 39 on Days 12 and 13. We could also use the algebraic theory of Question 1 to find $I_{\max}$.

(b) The epidemic lasts a little over 40 days $\big(I(40)=1.0\big)$.

(c) We find that $S(40)=31.3$ and $S(50)=30.9$, so that 31 people don't catch the disease.

The hospital can cope comfortably.

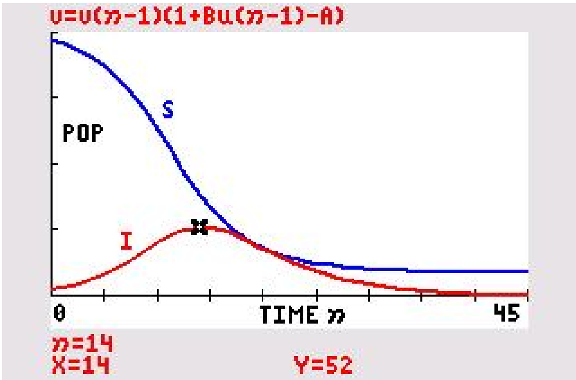
5.5 Discrete Approach: On a Pacific Island

1. $\alpha = 0.2$, $\beta = 0.0025$. A suitable window is

```
WINDOW
 nMin=0
 nMax=45
  PlotStart=1
  PlotStep=1
  Xmin=0
  Xmax=45
  Xscl=5
  Ymin=0
↓Ymax=200
```

```
WINDOW
↑nMax=45
  PlotStart=1
  PlotStep=1
  Xmin=0
  Xmax=45
  Xscl=5
  Ymin=0
  Ymax=200
  Yscl=50
```

with resulting time graph S, I vs n (time). The cursor is set on the maximum I value.



For these values of α and β (using trace on the calculator graph):

(a) the maximum I, the peak of infection, of 52 occurs after 14 days. The number of susceptibles then is 76.9.

(b) The epidemic lasts 44 days, with 18 persons not catching the disease.

(c) I increases (as n increases) for $77 < S < 195$ and decreases (as n increases) for $18 < S < 76$.

(d) Setting $I_{n+1} = I_n = I^*$ for maximum I_n, we have from Eq. (7),

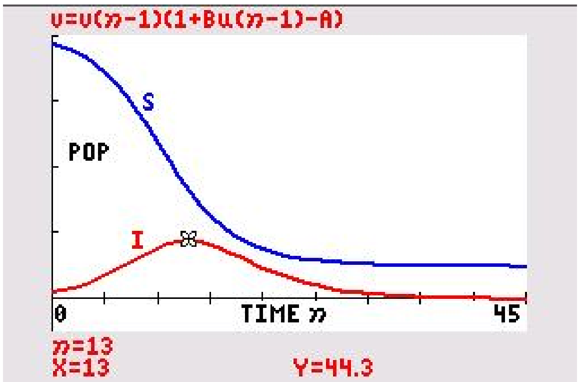
$$\begin{aligned} I^* &= I^* + \beta S_n I^* - \alpha I^*. \\ \therefore\ I^*\left(\beta S_n - \alpha\right) &= 0. \\ \therefore\ \beta S_n - \alpha &= 0 \qquad \text{since } I^* \neq 0 \text{ at the maximum.} \end{aligned}$$

Therefore, at the peak of infection (maximum I_n), we find the same theoretical result as in the continuous case,

$$S_n = \frac{\alpha}{\beta} = \frac{0.2}{0.0025} = 80.$$

From the graph, the value of S at $I_{\max}$, Day 14, is 76.9, close to the theoretical value of 80. The error in the theoretical value is ± 1 day, as S_n is treated as a continuous variable here. Note that $S(13) = 88.2$.

2. A mean infectious period $1/\alpha = 4$ gives $\alpha = 0.25$. We are given $S_8 = 147$, $I_8 = 30$. The value of $\beta = 0.00281$ gives $S_8 = 143.1$, $I_8 = 30.0$.



For these values of α and β (again using trace),

(a) the maximum I of 44 occurs after 13 days. The corresponding value of S, $S = 81.5$, is again close to the theoretical value $\alpha/\beta = 89$. (The S value on Day 14 is 93.)

(b) the disease runs for 37 days ($I_{37} < 1$)

(c) 25 persons (S_{37}) avoid catching the disease.

The hospital therefore just copes.

3. Draw up a table of corresponding values for Questions 1 and 2.

In all cases, $N = 200$, $S(0) = 195$, $I(0) = 5$.

Question 1 $\alpha = 0.2$, $\beta = 0.0025$, $\alpha/\beta = 80$

Quantity	Continuous	Discrete
I_{max}/Day/S	49/13/69.0	52/14/76.9
Duration	46	44
Uninfected	21	18

Question 2 $\alpha = 0.25$

Quantity	Continuous	Discrete
β	0.00273	0.00281
α/β	91.6	89.0
I_{max}/Day/S	39/12/94.2	44/13/81.5
Duration	41	37
Uninfected	31	25