
Contents

1	Introduction	1
2	Simple epidemic models	5
2.1	SIS epidemic	5
2.1.1	Solving directly	8
2.1.2	Phase portraits	10
2.2	SIR epidemics	14
2.3	Lab work	16
2.3.1	Exercises	19
3	Calculating R_0	23
3.1	Threshold parameters	23
3.2	The method of Anderson & May	24
3.3	The Jacobian	25
3.3.1	The SIS model threshold	26
3.3.2	The endemic equilibrium for the SIS model	28
3.3.3	The SIR model threshold	28
3.3.4	The endemic equilibrium for the SIR model	30
3.4	Other methods for calculating R_0	30
3.4.1	Survival Function	31
3.4.2	Next-generation method	32
3.4.3	Average Age at Infection	35
3.4.4	The Final Size Equation	35
3.4.5	Calculation from the Intrinsic Growth Rate	36
3.5	Vectorial capacity	36
3.6	Lab work	37
3.6.1	Exercises	38
4	A vector-borne disease with lifelong immunity	41
4.1	Building a vector-borne disease model	41
4.2	Finding eigenvalues	43

4.3	Deriving a threshold condition	44
4.4	Total population behaviour	45
4.5	Lab work	46
4.5.1	Exercises	49
5	A vector-borne disease with temporary immunity	53
5.1	Temporary immunity	53
5.2	Lab work	55
5.2.1	Exercises	56
6	The spread of measles	59
6.1	The Conservation Law	59
6.1.1	Diffusion	62
6.2	Measles in a corridor	63
6.3	Lab work	65
6.3.1	Exercises	72
7	Solving Partial Differential Equations	75
7.1	Separation of variables	75
7.2	Boundary conditions	78
7.3	Initial conditions	79
7.4	Lab work	79
7.4.1	Exercises	80
8	Fitting curves to data	83
8.1	Model fitting vs. interpolation	83
8.2	Sources of error in the modelling process	87
8.3	Visual fitting with the original data	88
8.4	Transforming the data	90
8.4.1	Regression coefficients can be misleading	93
8.4.2	Transformations can also be misleading	94
8.5	Lab work	98
8.5.1	Exercises	100
9	Splines	103
9.1	Spline interpolation	103
9.2	Linear splines	104
9.3	Cubic splines	106
9.3.1	An example	108
9.3.2	Generalising	111
9.4	Lab work	111
9.4.1	Exercises	111

10 The discrete logistic equation	115
10.1 Developing a nonlinear model	115
10.1.1 Spatial limitations	115
10.1.2 Difference equations	116
10.2 Cobwebbing	119
10.2.1 Stability in the discrete logistic equation	121
10.2.2 More complex behaviour	122
10.3 Lab work	124
10.3.1 Exercises	126
11 Bifurcations	129
11.1 Stability in difference equations	129
11.1.1 The range $0 < r < 3$	130
11.1.2 The range $3 < r < 3.45$	131
11.1.3 The range $r > 3.45$	134
11.2 Other bifurcations	134
11.3 Lab work	135
11.3.1 Exercises	136
12 Application: AIDS and end-stage renal disease	139
12.1 Determining AIDS prevalence	140
12.2 Determining end-stage renal disease prevalence	144
12.3 Model fitting	147
12.4 Using the model to predict future outcomes	150
12.5 Lab work	152
12.5.1 Exercise	152
13 Application: Malaria with a time delay	155
13.1 A delay differential equation model of malaria	155
13.2 Lab work	159
13.2.1 Exercises	159
Partial fractions	161
Eigenvalues	163
The R_0 sleight of hand	165
Finding eigenvalues for the case of permanent immunity	167
Integrating factors	169
Finding eigenvalues for the case of temporary immunity	171
Taylor's Theorem	173

Stability of periodic orbits in the logistic equation	175
Solving the end-stage renal disease equations	177
Matlab overview	179
J.1 Variables and operators	179
J.2 M-Files	180
J.3 Figures.....	181
J.4 Loops	182
J.5 Solving ODEs.....	183
J.6 Solving delay differential equations	183
J.7 Glossary of other terms	184
References	185
Index	187

List of Figures

1.1	The mathematical modelling process.	3
2.1	An SI model.	6
2.2	An SIS model.	7
2.3	The time series of the infection.	10
2.4	The linear relationship between S and I	11
2.5	Equilibria for the two cases.	12
2.6	The stability of the equilibrium points.	13
2.7	An SIR model.	14
2.8	An SIR model with vaccination.	15
2.9	An SIR model with mutation.	15
2.10	Phase portrait for $b = 10$	18
2.11	The epidemic diagram for the SIMD model.	20
3.1	A example of a two-stage basic reproductive ratio.	34
4.1	A vector-host model for a disease with permanent immunity.	42
4.2	The roots of a quadratic.	44
4.3	Infected mosquitos vs. infected humans.	49
5.1	A vector-host model for malaria with temporary immunity.	54
5.2	The roots of a cubic.	55
6.1	The total flow of droplets in a segment of space of length Δx	60
6.2	The flow of droplets within a small length of space.	60
6.3	Approximating a thin area by a rectangle	62
6.4	The trajectory of an airborne measles droplet	63
6.5	Average trajectory of droplets over time.	65
8.1	Data from the 1918 influenza pandemic.	85
8.2	Interpolating the data using a smooth curve.	85

8.3	Fitting a parabola to the data points.	86
8.4	Each data point is thought of as an interval of confidence.	88
8.5	Minimising the sum of the absolute deviations	89
8.6	Minimising the largest absolute deviation from the fitted line.	90
8.7	Plot of collected data for new HIV infections in 1981.	91
8.8	Plot of y versus e^x for the original data.	91
8.9	Plot of $\ln y$ versus x for the transformed data.	92
8.10	Avian influenza data.	94
8.11	Log transformation of avian influenza data	94
8.12	Exponential vs. original curve	95
8.13	Biannual outbreaks of influenza.	96
8.14	The transformed data points and a line of best fit.	97
8.15	A plot of the curve $y = Ce^{1/x}$ based on the value $\ln C \approx -1.25$	97
9.1	A linear spline model	104
9.2	A cubic spline model.	107
9.3	The natural cubic spline.	110
10.1	The previous value vs. the updated value	118
10.2	Iterating the dynamics.	119
10.3	Cobwebbing	120
10.4	Logistic map for $r = 0.5$	121
10.5	Logistic map for $r = 2$	122
10.6	Logistic map for $r = 3.2$	123
10.7	Logistic map for $r = 4$	123
11.1	Bifurcation diagram for $0 < r < 3$	131
11.2	Equilibrium points of $g(x)$	132
11.3	Bifurcation diagram for the appearance of period 2 points.	133
12.1	AIDS deaths among African Americans in the US	141
12.2	AIDS prevalence among African Americans in the US	142
12.3	Linear fit to pre- and post-HAART AIDS data	143
12.4	Linear fit to the entire AIDS data set.	144
12.5	End-stage renal disease mortality	146
12.6	End-stage renal disease prevalence	147
12.7	Linear fit to pre- and post-HAART end-stage renal disease data	148
12.8	Linear fit to the entire end-stage renal disease data.	149
12.9	HAART-blocking effects on disease progression.	151
13.1	A example of a two-stage basic reproductive ratio.	157
13.2	Infected humans and mosquitos.	159

List of Tables

8.1	Collected data for new HIV infections in 1981.	90
8.2	Transformed data	92
8.3	Initial infections and cull size for an avian influenza outbreak..	93
8.4	Diameter of a ponderosa pine versus volume.....	98
8.5	Mean number of bugs per infected site	100
9.1	Data for MRSA which increases sharply one year.	105
9.2	A linear spline model.....	106
9.3	A natural cubic spline model.	110
13.1	Variable and parameters for the malaria model with delay	156

Introduction

Infectious diseases have been a part of the human condition since time immemorial. Some, such as herpes or chicken pox, usually have mild symptoms and vanish of their own accord. Others, such as HIV, tuberculosis or malaria, are responsible for millions of deaths each year. Diseases have been the source of fear and superstition throughout the ages; sociological upheavals following the Black Plague led to the demise of the church as an all-powerful institution, the destruction of the serf system and subsequent creation of labour movements, and the creating of colonialism. Malaria is likely the greatest killer of human beings of all time, being thought to have killed one in two humans who ever lived. The 1918 influenza pandemic killed 50-100 million people in the space of six months (about 5.5% of the entire world; this would be equivalent to 350 million deaths today).

Mathematical epidemiology has its roots in 1760, when Daniel Bernoulli formulated and solved a model for smallpox. In 1906, Hamer used a discrete-time model of measles to understand recurrent epidemics. In 1911, Sir Ronald Ross used mathematical models to help demonstrate that malaria was spread by mosquitos (later winning a Nobel prize for his work). This work was later followed up by public-health physician William Kermack and biologist Anderson McKendrick in 1927, who showed that mortality rates were much more dependent on the year of birth than the year of death, as previously thought. This was particularly impressive, since Kermack had been blinded three years earlier and did all the mathematics in his head.

Since the middle of the twentieth century, mathematical epidemiology has grown exponentially. The appearance of HIV/AIDS saw a further explosion in the growth of models, as ways of predicting the spread of the disease became crucial. Mathematical models were used to design the triple-drug cocktail, responsible for saving the lives of millions of people. In recent years, mathematics has been used to determine intervention strategies for the foot and mouth outbreak in Britain, the SARS epidemic, human papillomavirus vaccination, and West Nile virus. Mathematics can determine the critical proportion of

individuals who must be vaccinated against a disease, the minimum amount of drug required to treat an infection, the frequency of use of an insecticide to control mosquitos or the number of doses of a drug regimen that may be missed before drug resistance emerges.

Mathematical modelling has become a tool that is now widely used to study problems and questions in public health. Such models have focussed on the spread of an infectious agent in a population, the pathogenesis of an infectious agent in a host and its effect on the immune system, the growth and spread of tumour cells in a body, and even the economics of vaccine and drug delivery at the population or individual level. However, public-health workers rarely have the mathematical skills to develop informative models for public-health policy.

In recent years, the rise of detailed electronic surveillance of infectious disease and improvements in computing power, electronic data management, rapid diagnostic tests, genetic sequence analysis, and the ability to share and deposit data over the internet has given rise to a vastly improved ability to monitor the spread of diseases. Statistical methods have been integrated into the modelling process, allowing estimation of key parameters and the testing of hypotheses using available data.

Mathematical modelling is analogous to map-making. Without a decent map, even finding our way home is difficult, let alone exploring new avenues. Like any good map, we aren't trying to create a perfect replica of reality; rather, we hope to elucidate key features and ignore what isn't relevant. Often, what we leave out is as important as what we include. To solve biological problems, first we translate the biology into a mathematical model. (Despite appearances, usually borne out of math phobia, this is often the most challenging part of the process.) Once we are satisfied with the model, we perform mathematical analysis. Assuming we do this correctly, we can be confident that what happens here is a series of entirely logical steps. Our mathematical conclusion then needs to be translated back to a biological conclusion. Since the process of mathematical analysis is rigorous, it follows that the strength of our conclusions are wholly dependent upon the formulation of the mathematical model. See Figure 1.1. Thus, it is vital to understand the mechanisms, as well as the limitations, of this process.

More formally, mathematical modelling of infectious disease is a framework designed to convey ideas about the components of host-pathogen interaction. It requires a clear understanding of the interaction between the infectious agent and the host within an individual, the mode and rate of transmission between individuals, and host-population characteristics, such as demography and behaviour. Models can be used to determine the dominant factors that generate observed patterns and phenomena. They can aid data collection, interpretation and parameter estimation. They can identify possible approaches for disease control and for assessing the potential impact of different intervention measures. Importantly, they can do all this without costly and time-consuming experiments.

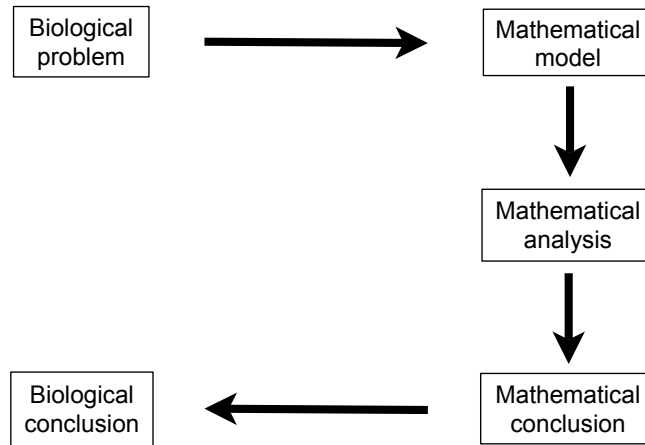


Fig. 1.1. The mathematical modelling process.

This monograph facilitates the learning of mathematical biology for public-health researchers and mathematical modellers, who may have a range of skills and knowledge drawn from mathematics, clinical medicine, sociology, economics, decision-making, risk analysis, psychology, industrial production capacity and politics. There needs to be a symbiosis between these areas; people working in these disciplines need to communicate to each other effectively so that key questions in health can be addressed.

This monograph specifically introduces mathematical modelling to those in the public-health sector who have had no formal training in mathematics. It allows people working in a range of disciplines to understand the strengths and weaknesses of models. Mathematical models are formulated that represent current diseases; the tools needed to analyse these models are introduced as necessary. Through MATLAB exercises, numerical simulations are developed that provide the reader with the ability to formulate control strategies, test hypothetical interventions and explore disease-management options. The exposition of the material is mainly addressed to researchers and graduate students interested in the application of mathematics to biological problems.

One of the objectives is to familiarise public-health workers with questions and outcomes that can be produced by simple mathematical models. Ultimately, policy-makers need to understand the importance and limitations of models, in order to be confident in basing decisions on them.