

Glenn Ledder

Mathematics for the Life Sciences

Calculus, Modeling, Probability, and
Dynamical Systems

February 12, 2013

Springer

Preface

Science is built up with facts, as a house is built with stones. But a collection of facts is no more a science than a heap of stones is a house.

Jules Henri Poincaré

There are several outstanding mathematical biology books at the advanced undergraduate and beginning graduate level, each with its own set of topics and point of view. Personal favorites include the books by Britton, Brauer and Castillo-Chavez, and Otto and Day. These books are largely inaccessible to biologists simply because they require more mathematical background than most biologists have. This book began as lecture notes for a course intended to help biologists bridge the gap between the mathematics they already know and what they need to know to read advanced books. The only prerequisite for the course was the first semester of the calculus sequence. Topics included mathematical modeling, probability, and dynamical systems. My original notes included a brief review of calculus, which I subsequently expanded into the first chapter of this book so that it could be used for courses that do not require a calculus prerequisite or by biologists whose calculus experience is but a distant memory. Most students will probably find this book to be more challenging than the typical calculus book, albeit in a different way. I do not make as many demands on students' computational skills, but I require a greater conceptual understanding and an ability to harness that conceptual understanding for service in mathematical modeling.

A Focus on Modeling

In its early days, science consisted of careful observation and experimentation, with a focus on collecting facts. However, as eloquently stated by the French mathematician, philosopher, and scientist Henri Poincare, this is not enough to make science work.

In contrast with science, mathematics is a purely mental discipline focused entirely on structures that we create in our minds. It can be very useful in science, but it has to be connected to science carefully if scientifically valid results are to be achieved. The connection is perhaps best made by a metaphor:

The muscles of mathematics are connected to the bones of experimental science by the tendons of mathematical modeling.

As you read through this book, you will see that mathematical modeling goes far beyond the “application” problems that mathematics text authors include so as to make mathematics appear relevant. The problem is that what little modeling work appears in these problems is generally done by the author rather than the students. At best, the experience of doing these problems only benefits science students if their science instructors are also good enough to do the modeling work for them.

This book is written from a modeling perspective rather than a mathematics or biology perspective. The lack of modeling content in the standard mathematics and science curricula means that the typical reader will have little or no modeling experience. Readers may find the modeling skills of Section 1.1 and Chapter 2 to be difficult to learn, but the effort to do so will be well rewarded in the remainder of the book and in any subsequent attempts to read biological literature with quantitative content. While it is unreasonable to expect readers of this book to become expert modelers, my primary goal is to make them sufficiently comfortable with mathematical modeling that they can successfully read scientific papers that have some mathematical content.

Pedagogy

There are a lot of connections between mathematics and biology, yet most students—and even many mathematicians and biologists—are unaware of these connections. One reason for this situation is that neither the historical development nor the pedagogical introduction of either subject involves the other.

Biology grew out of natural philosophy, which was entirely descriptive. Modern biology curricula generally begin with descriptive biology, either organismal or cellular. The mathematically-rich areas of genetics, ecology, and physiology make their appearance in advanced courses, after students have come to see biology as a non-mathematical subject.

Calculus and calculus-based mathematics were developed to meet the mathematical needs of physics, and it remains standard practice to use physics to motivate calculus-based mathematics. Other areas of mathematics, such as game theory and difference equations, were motivated to some extent by biology,¹ but these topics appear in specialized courses generally taken only by mathematics majors. Probability is another mathematical topic with strong connections to biology, but it is generally encountered in statistics courses that emphasize social science applications.

¹ Leonardo of Pisa, more commonly known as Fibonacci, developed his famous sequence as the solution of a difference equation model of population growth.

The basic premise of this book is that there is a lot of mathematics that is useful in some life science context and can be understood by people with a limited background in calculus, provided it is presented at an appropriate level and connected to life science ideas.

This is a mathematics book, but it is intended for non-mathematicians. Mathematicians like to have a mathematical definition for a concept and consider the meaning of the concept to be a consequence of the mathematical definition. Contrary to that plan, I prefer to begin with a functional definition and then present the mathematical definition as the solution of a problem. In probability, for example, I define each distribution according to its purpose rather than its mathematical representation and then present the mathematical representation as a result. This is pedagogically appropriate; there are infinitely many functions that satisfy the mathematical definition of a probability distribution and we should only be interested in those that have some practical value. The context should precede the mathematical definition.

A mathematics book for non-mathematicians needs to be clear about the extent to which rigor matters. A colleague of mine once started a talk to undergraduates with a joke: “An engineer, a physicist, and a mathematician are traveling in a train in Scotland when the train passes a black sheep standing along the track. The engineer concludes that sheep in Scotland are black. The physicist concludes that there is at least one black sheep in Scotland. The mathematician concludes that ...” Mathematicians have no trouble finishing the joke: the mathematician concludes that there is a sheep in Scotland that is black on one side. This insistence on rigor is both a strength and a weakness. It was long the common practice in calculus books (and such books are still popular) not to introduce the logarithm function until after the definite integral, even though the students have seen logarithm functions in precalculus. This example and others support my contention that “Mathematicians are people who believe you should not drive a car until you have built one yourself.”

It is my aim to provide a balanced approach to mathematical precision. Conclusions should be backed by solid evidence and methods should be supported by an understanding of why they work, but that evidence and understanding need not have the rigor of a mathematical proof. At the risk of stern rebukes from my mathematics colleagues, I will say up front that I believe that students should focus on how we use mathematical results to solve problems. For this goal, we need to know why mathematical results are true, but we do not need to know how we prove them to be true. An example is the limit result needed to derive the formula for the derivative of the exponential function. The proof of this result appears in most calculus books and is indeed a beautiful piece of mathematics; however, understanding it does not help us compute derivatives or apply them to solve problems. Graphs and numerical computations strongly hint at the correct limit result; while not rigorous, these methods are more convincing to anyone but a professional mathematician and use problem solving skills that will be useful in other contexts. Similarly, the derivation of the Akaike information criterion (AIC) is very difficult, else it would have been

done prior to its actual discovery in the 1970s; nevertheless, it is not difficult to explain AIC in general terms. The mathematical error of presenting it without proof is far less serious in this book than would be the modeling error of omitting it.

Most of the sections are highly focused, often on one extended example. Mathematics experts know that we learn much more from a deep study of one problem than from many superficial examples. Many of my biological settings are in ecology, the area of biology I know best, but I have also tried to find settings of very broad interest such as environmental biology, conservation biology, physiology, and the biology of DNA. In particular, these areas are more likely to interest lower-division undergraduates, many of whom are pre-medicine majors rather than biology majors, and most of whom have very little knowledge of biology.

I have attempted to be brief, in the hope that readers will work harder to read a short presentation than a long one. I use examples as contexts in which to present ideas rather than instances where a formula is used to obtain an answer. Hence, the number of examples is limited, but each example is treated with some depth. Similarly, I include only a small number of figures. Each figure is essential to the presentation, and the reader should work hard to understand each one. Being able to explain² a figure represents a high level of understanding.

Broad modeling problems require a variety of mathematical approaches. Hence, there are topics that are ideal for problems that are distributed among the relevant sections rather than being incorporated into a single project. I have indicated these connections within the problems themselves and also called attention to them in each chapter introduction. It is possible to combine all of the problems on a given model into one large project if desired.

Technology

Some mathematical modeling work must be done by hand, while other work is greatly facilitated by the use of computers. I view both hand methods and computational methods as tools in my modeling toolbox. I try to identify the best tool for any particular task without a bias either for or against technology. I do have a bias against using computer algebra systems to do routine algebra and calculus; this stems from frequently encountering problems where valuable results can be found only with the careful use of algebraic substitution and simplification that requires a human touch. I could not resist the temptation to point some of these out in the text.

There are a multitude of platforms for doing mathematical modeling tasks on computers. None of these is ideal, and the choice of which to use is a matter of taste. Rather than trying to find the very best tool for each individual task, my preference is to work with one tool that is reasonably good for any task (save symbolic computation) and is readily available. By these criteria, my choice is R, which runs smoothly in any standard operating system and is popular among biologists. Matlab is also an excellent choice. Both R and Matlab are programming environments, as

² An *explanation* includes context and analysis in addition to mere *description*.

opposed to packaged software or programming languages. Spreadsheets and other packaged software provide easy access to mathematics because of their intuitive graphical interface; however, programming is limited and the details of formulas are hidden from view, making it impossible to see the overall structure of a program at a glance. Reusability is limited, as anyone who has tried to modify a spreadsheet created by another author can attest. By comparison, one can see an entire R or Matlab program at a glance and adapt prior work to a similar context with minimal changes. High-level languages, such as Java and C++, offer sophisticated programming capability, but they are difficult to learn compared to the languages used in programming environments such as R or Matlab. The choice between R and Matlab is a matter of personal taste. It is easier to get professional-quality graphics with Matlab, but R has more intuitive syntax that facilitates programming. Matlab requires an add-on toolbox for probability and statistics, while R requires supplementation for dynamical systems. I use R because students can get it for free and install it seamlessly in any operating system. R lacks the excellent documentation that comes with Matlab; however, I maintain a collection of R scripts for various algorithms presented in the text, and these are readily available from my web page. These scripts are designed to be simple rather than robust; that is, compared to professionally written programs, they are easier to understand but less efficient and they lack error detection machinery. Their presence allows students to evade the difficulty of having to learn R from scratch, replacing it with the much lesser difficulty of having to be able to read an R program and make minor modifications. A detailed set of R scripts with commentary can be found on the web site for the book.

Topics

Of course no book on mathematics for the life sciences can be complete. Some important areas do not appear here at all because they do not fall into the broad categories of mathematical modeling, probability, and dynamical systems. Several concessions have been made in the interest of accessibility. Some topics are given only a partial treatment as compared to the treatment they would receive in a higher-level course; for example, I do not find eigenvectors for complex eigenvalues, since these are not generally needed in biology. Others are presented in a roundabout way.

Finding the “correct” order of topics in this book was an insoluble problem. Mathematics is a hierarchical subject, but the hierarchy is not linear. Arguments can be given for significant rearrangements of the topics that are included here. Ultimately, the only reasonable solution was to group topics into related clusters. In particular, parts 2 and 3 could easily have been reversed. Those who read this book for their own benefit or to design a course should be flexible in the way they structure their study. Each part introduction contains a graph with sections as nodes and arrows indicating which sections are necessary background for others. A syllabus that moves frequently between chapters is entirely possible, but for me to have written the book in that way would have been to exclude other topic orders.

One feature of mathematical models that causes difficulties for students is the appearance of parameters, which are constants whose values are not necessarily assigned. Without parameters, a function is merely an example to be used for routine calculations. With parameters, a function can be a model, which can serve as an environment for theoretical experimentation. Even the reader with a solid background in calculus should study Section 1.1.

The remainder of Chapter 1 can serve as a review of calculus or a conceptually-oriented calculus primer. This chapter is not a complete treatment of calculus, which would require far more space than is available in one chapter. I present here only those aspects of calculus that provide the necessary background for the modeling, probability, and dynamical systems that make up the rest of the book. The reader who works through this chapter will be well equipped with the calculus background needed for the purpose at hand. The material in this book has been used successfully with life science graduate students who had no background in calculus. Anyone who requires a more complete understanding of calculus can consult any calculus book.

After the calculus primer comes a chapter on mathematical modeling, which is the necessary focus of any study of mathematics for those whose purpose is to use mathematics to better understand science. Even the most mathematical of topics, such as probability, are best seen by scientists from a viewpoint of mathematical modeling. Unfortunately for the science student, mathematical modeling has not been granted a place in the standard mathematics and science curricula. In mathematics books, we generally present mathematical ideas and then look for their applications to science. The result is a collection of idiosyncratic examples devoid of the analysis necessary for good mathematical modeling. In science books, the mathematics is usually presented as a collection of formulas, to be used as facts when required. Neither approach teaches modeling skills. If we are to use mathematics to improve our understanding of the natural and physical world, we must focus on the connections of mathematics to science.

Chapters 3 and 4 present the basic ideas and some applications of probability, including applications commonly classified as statistics. The treatment given here is organized differently from the treatment of this topic in statistics or probability books. Mathematicians generally use an axiomatic approach to introduce probability. My colleagues in biology helped me appreciate that the central topic of probability for scientists is that of the probability distribution, and this topic is best approached informally by thinking of a probability distribution as a mathematical model of a data set. My aim has been to get to probability distributions as quickly as possible while saving other topics, such as conditional probability, for later. The essentials of probability distributions form the subject of Chapter 3. Chapter 4 includes additional topics that build onto or supplement the basic material on probability distributions. The high point of these two chapters is Section 4.4, which looks at the question of how likely it is that a subpopulation used to provide sample data is distinct from some larger population.

The final three chapters introduce the mathematics of dynamical systems, which consist of one or more related quantities that change in time according to prescribed rules. These rules may be in the form of difference equations, where time is taken

as discrete, or differential equations, where time is taken as continuous. It is usual to make this the primary distinction within the area of dynamical systems; however, there are valuable connections to be made between the two kinds, particularly for models with only one dynamic quantity. For this reason, I have chosen to treat all dynamic models of one variable together in Chapter 5 before presenting multivariable discrete systems in Chapter 6 and multivariable continuous systems in Chapter 7. For reasons presented in the modeling chapter, I believe that continuous models are almost always preferable to discrete models. Nevertheless, the analysis of continuous models requires an understanding of some discrete mathematics. Hence, Chapter 6 precedes Chapter 7. The reader whose primary interest is in continuous dynamical systems needs the tools developed in Section 6.3, and these tools are more easily acquired with Sections 6.1 and 6.2 as background. The high points of the three chapters on dynamical systems are the graphical and analytical tools used for continuous systems; these are the topics of Section 7.3 and 7.5 respectively. The book contains two additional sections on discrete dynamical systems, presented as appendices.

Advice for the Reader

How one reads a book depends on what one wishes to get from the reading. I assume that the reader wants a working knowledge of mathematics that will enable him/her to read biological literature with quantitative content or to read a more advanced book on mathematical biology. At the same time, many readers will be interested in only a portion of the topics presented here. As noted above, each part begins with a schematic diagram showing the logical relationships among the topics of that part and any essential topics from earlier parts. In particular, the reader is cautioned not to skip Chapters 1 and 2 so as to get to some other topic more quickly. People who try to learn to play the organ without having already learned to play the piano are starting with an enormous handicap; the same is true for someone who attempts to learn probability or dynamical systems without an adequate mastery of calculus and mathematical modeling. Not every section in Chapters 1 and 2 is essential for the remainder of the book; however, parts of these chapters are indispensable background and should be mastered before moving on. In particular, understanding of parameters (Section 1.1) and basic concepts of mathematical modeling (Section 2.2) are essential.

It is natural to try to work a large number of problems as quickly as possible. However, this is not the best way to learn mathematics. When a mathematician learns something new (s)he works through a relatively small number of examples carefully rather than a large number of examples superficially. At a talk I heard on mathematics pedagogy, the speaker asked the audience, "Why do we ask our students to work problems? Is it because we want to know the answer?" Usually we don't care about the answer; we work problems to learn mathematics. Keep this in mind when you are working a problem: your goal is to learn mathematics, not to get the answer to the problem. There are only a small number of routine problems in

this book. Most problems are guided case studies and will take quite a bit of time for a thorough understanding. Carefully working a small number of these will benefit the reader more than a cursory look at a larger number.

Course Designs

There is no standard curriculum of mathematics for biology. Mathematical biology can be incorporated into a calculus course, or calculus can be incorporated into a mathematical biology course for students who have not had calculus. There are also mathematical biology courses with a calculus prerequisite, and these can be for students with or without backgrounds in linear algebra and differential equations. While many institutions treat probability/statistics as being distinct from mathematics, the difficulty of finding room for either in the program of a biology major suggests the possibility of incorporating some probability and some topics often included in a statistics course within the mathematical biology or calculus for biology course. I have tried to make this book suitable for a variety of plans.

Before listing possible course plans using the material in this book, it is important to start with a broad discussion of pacing. Books for calculus and other lower-division courses are generally written under the assumption that each section will require one day of class. At this pace, it is not difficult to put more than 30 sections into a standard 3-credit course. This is typically what is done in a differential equations course, but not a calculus course. At the University of Nebraska, we cover something like 32 sections in our first-semester calculus course; however, we structure this course with a lecture-recitation format and offer it for 5 credits. This means that our actual rate of coverage is more like 6 sections per credit hour than 10 sections per credit hour. My own mathematics for biology course was originally a 5-credit course with a one-semester calculus prerequisite, and I covered approximately 32 sections of this book, which is again an average of only 6 sections per credit hour. The next time I use this book for a course, it will be a 5-credit calculus-for-biology course with three hours of lecture and two hours of recitation/laboratory. I expect to do only the 21 sections of Chapters 1, 2, and 5 and perhaps the first three sections of Chapter 6. For students at the calculus level, I would certainly not try to do more than 16 sections for a 3-credit course. A slightly faster pace could be used with students who are more sophisticated or very highly motivated. I spend less than half of the total class time presenting lectures; in particular, I mark out days for laboratory-style activities such as collecting data from a virtual laboratory, writing a computer program to run a computer simulation, or working through one of the more difficult problems either in small groups or as a committee-of-the-whole. The standard requirement that mathematics courses cover as much material as possible sacrifices depth for breadth; a mathematics course for biology students should have some balance between the two, with some case studies coming at the expense of broad coverage.

A 2-Course Sequence of 4-Credit Courses

It should be possible to do almost the entire book with a total of 8 credit hours. I would do Chapters 1, 2, and 5 in a first-semester Calculus for Biology course and most of Chapters 3, 4, 6, and 7 in a second-semester Probability and Dynamical Systems course.

A 2-Course Sequence of 3-Credit Calculus-for-Biology Courses

Given two courses for students with no calculus background, I would use Chapters 1 and 2 for the first semester and then make the second semester a dynamical systems course that would include Chapters 5 through 7, probably without Section 6.4, and possibly with one or both topics of Appendix A. Both of these courses would be well focused, and the second course could be open to strong students with a background somewhat beyond one course in calculus.

A 3-Credit Calculus-for-Biology Course

In a 3-credit calculus for biology course, I would expect to complete all of Chapter 1 in about half of the semester or perhaps a little more. I would probably try to do some dynamical systems rather than a complete treatment of Chapter 2. It would be possible to do Sections 1, 2, 5, and 6 from Chapter 2 along with all of Chapter 5. I would present only a minimal version of Section 2.5, the point being to do whatever is required to set up Section 2.6.

A 3-Credit Empirical Modeling and Probability Course with a Calculus I Prerequisite

One could teach a course on empirical modeling and probability as an alternative to a standard statistics course. For such a course, I would do Sections 1.1, 2.1 through 2.4, 2.7, all of Chapter 3, and as much of Chapter 4 as could be done without rushing.

A 3-credit Dynamical Systems Course with a Calculus I Prerequisite

A course on dynamical systems could not reasonably assume an adequate modeling background, so it would be necessary to start with Sections 1.1, 2.1, 2.2, 2.5, and 2.6, with 2.1 done in a cursory manner. It would then be possible to do all of the material in Chapters 5 through 7 except for Section 6.4. If any extra time is available, Section A.1 would round out the course.

Glenn Ledder
Lincoln, Nebraska
January 2013

Contents

Part I Calculus and Modeling

1	A BRIEF SUMMARY OF CALCULUS	5
1.1	Working with Parameters	7
1.1.1	Scaling Parameters	9
1.1.2	Nonlinear Parameters	12
1.1.3	Bifurcations	13
1.2	Rates of Change and the Derivative	20
1.2.1	Rate of Change for a Function of Discrete Time	21
1.2.2	Rate of Change for a Function of Continuous Time	22
1.2.3	The Derivative	25
1.2.4	Slope of a Tangent to a Graph	26
1.3	Computing Derivatives	30
1.3.1	Two Notations	30
1.3.2	Elementary Derivative Formulas	32
1.3.3	General Derivative Rules	33
1.3.4	Partial Derivatives	37
1.4	Local Behavior and Linear Approximation	40
1.4.1	Tangent Lines	40
1.4.2	Local Extrema	41
1.4.3	Linear Approximation	46
1.5	Optimization	50
1.5.1	The Marginal Value Theorem	54
1.6	Related Rates	62
1.6.1	Differential Equations	63
1.6.2	The Chain Rule	65
1.7	Accumulation and the Definite Integral	71
1.7.1	Estimating Total Volume from Flow Rate Data	71
1.7.2	The Definite Integral	73
1.7.3	Applications of the Definite Integral	76
1.8	Properties of the Definite Integral	81

1.8.1	The Fundamental Theorem of Calculus	83
1.8.2	Computing Definite Integrals with the Fundamental Theorem	85
1.9	Computing Antiderivatives and Definite Integrals	89
1.9.1	Substitution	90
1.9.2	Constructing Antiderivatives with the Fundamental Theorem	92
1.9.3	Obtaining a Graph of f from a Graph of f'	94
	References	98
2	MATHEMATICAL MODELING	99
2.1	Mathematics in Biology	101
2.1.1	Biological Data	102
2.1.2	Overall Patterns in a Random World	103
2.1.3	Determining Relationships	105
2.2	Basic Concepts of Modeling	106
2.2.1	Mechanistic and Empirical Modeling	109
2.2.2	Aims of Mathematical Modeling	112
2.2.3	The Narrow and Broad Views of Mathematical Models	114
2.2.4	Accuracy, Precision, and Interpretation of Results	115
2.3	Empirical Modeling I: Fitting Linear Models to Data	119
2.3.1	The Basic Linear Least Squares Method ($y = mx$)	120
2.3.2	Adapting the Method to the General Linear Model	124
2.3.3	Implied Assumptions of Least Squares	126
2.4	Empirical Modeling II: Fitting Semilinear Models to Data	129
2.4.1	Fitting the Exponential Model by Linear Least Squares	129
2.4.2	Linear Least Squares Fit for the Power Function Model $y = Ax^p$	130
2.4.3	Semilinear Least Squares	133
2.5	Mechanistic Modeling I: Creating Models From Biological Principles	140
2.5.1	Constructing Mechanistic Models	140
2.5.2	Dimensional Analysis	143
2.5.3	A Mechanistic Model for Resource Consumption	144
2.5.4	A More Sophisticated Model for Food Consumption	146
2.5.5	A Compartment Model for Pollution in a Lake	148
2.5.6	“Let the Buyer Beware”	149
2.6	Mechanistic Modeling II: Equivalent Forms	152
2.6.1	Notation	152
2.6.2	Algebraic Equivalence	153
2.6.3	Different Parameters	154
2.6.4	Visualizing Models with Graphs	155
2.6.5	Dimensionless Variables	156
2.6.6	Dimensionless Forms	157
2.7	Empirical Modeling III: Choosing Among Models	161
2.7.1	Quantitative Accuracy	162
2.7.2	Complexity	163

2.7.3	The Akaike Information Criterion	165
2.7.4	Choosing Among Models	166
2.7.5	Some Recommendations	168
	References	171

Part II Probability

3	PROBABILITY DISTRIBUTIONS	177
3.1	Characterizing Data	178
3.1.1	Types of Data	178
3.1.2	Displaying Data	180
3.1.3	Measures of Central Tendency	182
3.1.4	Measuring Variability	184
3.2	Concepts of Probability	188
3.2.1	Experiments, Outcomes, and Random Variables	190
3.2.2	Probability Distributions	192
3.2.3	Sequences and Complements	194
3.3	Discrete Probability Distributions	197
3.3.1	Distribution Functions	201
3.3.2	Expected Value, Mean, and Standard Deviation	203
3.4	The Binomial Distribution	207
3.4.1	Bernoulli Trials and Binomial Distributions	208
3.5	Continuous Probability Distributions	215
3.5.1	Cumulative Distribution Functions	217
3.5.2	Probability Density Functions	220
3.5.3	Expected Value, Mean, and Variance	223
3.6	The Normal Distribution	226
3.6.1	The Standard Normal Distribution	227
3.6.2	Standard Probability Intervals	231
3.7	The Poisson and Exponential Distributions	234
3.7.1	The Poisson Distribution	235
3.7.2	The Exponential Distribution	238
3.7.3	Memory in Probability Distributions	240
	References	246
4	WORKING WITH PROBABILITY	249
4.1	An Introduction to Statistical Inference	250
4.2	Tests on Probability Distributions	257
4.2.1	A Graphical Test for Distribution Type	258
4.2.2	The Cramer-von Mises Test	260
4.2.3	Outliers	262
4.3	Probability Distributions of Samples	267
4.3.1	Sums and Means of Two Random Variables	268
4.3.2	General Characteristics of Samples	268
4.3.3	Means and Standard Deviations of Sample Sums and Means	270

4.3.4	Distributions of Sample Means	271
4.4	Inferences About Populations	276
4.4.1	Inferences About Sample Means	277
4.4.2	Inferences About Sample Proportions	279
4.5	Estimating Parameters for Distributions	285
4.5.1	Confidence Intervals	285
4.5.2	Estimating Success Probability with a Likelihood Function	287
4.5.3	Estimating Population Size by Mark-and-Recapture	289
4.5.4	A Final Observation	291
4.6	Conditional Probability	295
4.6.1	The Concept of Conditional Probability	296
4.6.2	Formulas for Conditional Probability	297
4.6.3	The Partition Rule	299
4.7	Conditional Probability Applied to Diagnostic Tests	302
4.7.1	The Standard Test Interpretation Problem	305
	References	310

Part III Dynamical Systems

5	DYNAMICS OF SINGLE POPULATIONS	315
5.1	Discrete Population Models	316
5.1.1	Discrete Exponential Growth	318
5.1.2	The Discrete Logistic Model	320
5.1.3	Simulations	321
5.1.4	Fixed Points	322
5.2	Cobweb Analysis	326
5.2.1	Cobweb Diagrams	327
5.2.2	Stability Analysis	329
5.3	Continuous Population Models	332
5.3.1	Exponential Growth	333
5.3.2	Logistic Growth	334
5.3.3	Equilibrium Solutions	335
5.3.4	A Renewable Resource Model	336
5.3.5	Equilibria of the Renewable Resource Model	338
5.3.6	A Graphical Method	339
5.4	Phase Line Analysis	346
5.4.1	The Phase Line	346
5.4.2	The Phase Line for the Holling Type 3 Renewable Resource Model	347
5.4.3	Comparison of Graphical Methods	349
5.5	Linearized Stability Analysis	351
5.5.1	Stability Analysis for Discrete Models—a Motivating Example	351
5.5.2	Stability Analysis for Discrete Models—the General Case	353
5.5.3	Stability Analysis for Continuous Models	356

5.5.4	Similarities and Differences	357
	References	361
6	DISCRETE DYNAMICAL SYSTEMS	363
6.1	Discrete Linear Systems	364
6.1.1	Simple Structured Models	364
6.1.2	Finding the Growth Rate and Stable Distribution	367
6.1.3	General Properties of Discrete Linear Structured Models	368
6.2	A Matrix Algebra Primer	374
6.2.1	Matrices and Vectors	375
6.2.2	Population Models in Matrix Notation	377
6.2.3	The Central Problem of Matrix Algebra	378
6.2.4	The Determinant	379
6.2.5	The Equation $\mathbf{Ax} = \mathbf{0}$	380
6.3	Long-term Behavior of Linear Models	383
6.3.1	Eigenvalues and Eigenvectors	383
6.3.2	Solutions of $\mathbf{x}_{t+1} = \mathbf{Mx}_t$	386
6.3.3	Long-term Behavior	387
	References	393
7	CONTINUOUS DYNAMICAL SYSTEMS	395
7.1	Pharmacokinetics and Compartment Models	396
7.1.1	Compartment Models	396
7.1.2	A Simplified Model	398
7.1.3	The Dimensionless Model	401
7.2	Enzyme Kinetics	406
7.2.1	Nondimensionalization	408
7.2.2	Simulation	409
7.2.3	The Briggs-Haldane Approximation	410
7.3	Phase Plane Analysis	415
7.3.1	Equilibrium Solutions	416
7.3.2	Solution Curves in the Phase Plane	416
7.3.3	Nullclines	417
7.3.4	Nullcline Analysis	418
7.3.5	Using Small Parameters	421
7.3.6	Nullcline Analysis in General	422
7.4	Stability in Linear Systems	430
7.4.1	Linear Systems	431
7.4.2	Eigenvalues and Stability	432
7.4.3	The Routh-Hurwitz Conditions	435
7.4.4	The Routh-Hurwitz Conditions for a Three-Dimensional System	436
7.5	Stability in Nonlinear Systems	440
7.5.1	Approximating a Nonlinear System at an Equilibrium Point	441
7.5.2	The Jacobian and Stability	443

7.6	Primary HIV Infection	448
7.6.1	Nondimensionalization	449
7.6.2	Reduction to Two Variables	450
7.6.3	Equilibria and Stability	451
7.6.4	Phase Plane Analysis	453
	References	457
A	Additional Topics in Discrete Dynamical Systems	459
A.1	Discrete Nonlinear Systems	459
A.1.1	Linearization for Discrete Nonlinear Systems	461
A.1.2	A Structured Population Model with One Nonlinearity	463
A.2	Markov Chains	472
A.2.1	Some Scientific Background	473
A.2.2	A Model for DNA Change	475
A.2.3	Equilibrium Analysis of Markov Chain Models	476
A.2.4	Analysis of the DNA Change Model	477
A.3	Boolean Algebra Models	481
A.3.1	Boolean Algebra	482
A.3.2	Boolean Functions and Boolean Networks	484
A.3.3	Using a Boolean Network for a Consistency Check	486
	References	488
B	The Definite Integral Via Riemann Sums	489
C	A Runge-Kutta Method for Numerical Solution of Differential Equations	491
	Hints and Answers to Selected Problems	493
	Index	513

Part I
Calculus and Modeling

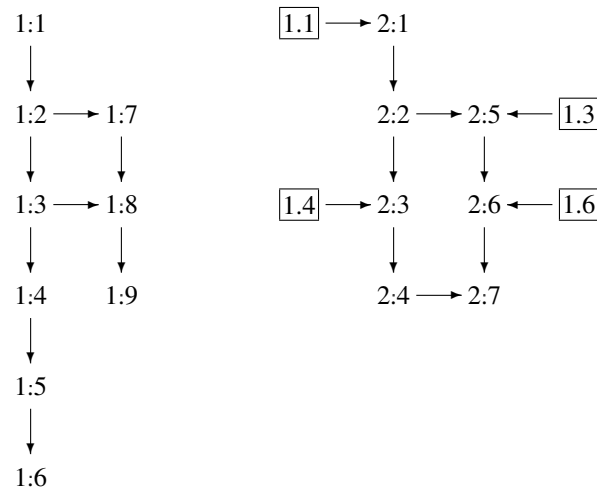
The first two chapters contain core material in mathematics: calculus and the basic elements of mathematical modeling. Section 1.1 is an essential preparatory section for the entire book. A significant part of the power of mathematics lies in its capacity for generalization. A single symbol can represent a range of numerical values, allowing the mathematical work to be done on a whole class of problems rather than an individual example. One cannot read any quantitative work in biology or any other science without an ability to understand how symbols are used in a particular context. This is a topic that most readers will find difficult, but one that is essential to master.

Aside from the opening section, Chapter 1 contains five sections on differential calculus and three on integral calculus. Each of these topics includes material on concepts, techniques, and applications. Contrary to the view of most students, calculus is largely a conceptual topic. Concepts are needed to understand the important applications of calculus; hence, the reader should spend enough time on the conceptual material for a thorough understanding. In particular, the derivative concept is essential to understanding dynamical systems, while the definite integral concept is essential for continuous probability distributions. The amount of effort to be expended on techniques is a matter of taste. Differentiation techniques are needed in the rest of the book, but all such cases are fairly elementary. Integration techniques are not needed for the rest of the book. Both of these can be done using computer algebra systems if desired. The applications that appear in Chapter 1 can generally be considered as ends in themselves, and can be accorded as much or as little interest as the reader desires. There are a few exceptions. Related rates (in Section 1.6) are vital background for nondimensionalization, which appears in Section 2.6, and linear approximation (in Section 1.4) is vital background for the nonlinear dynamics that appears in Chapters 5 and 7 and Appendix A.1.

In the preface, I described mathematical modeling as the tendons that connect the muscle of mathematics to the bones of science. Colleagues who teach science and engineering courses often say that their students do not seem to be able to do the mathematics necessary for their subject. The real problem is not so much an inability to do mathematics but an inability to harness the power of mathematics in a scientific context. More or different mathematics will not address this problem; it requires attention to mathematical modeling, which is largely absent from courses in either mathematics or science. This is the purpose of Chapter 2. The first two sections provide basic terminology and ideas. There does not seem to be a lot of material in these sections, and there are very few associated problems. The reader should plan to reread these sections several times while working through the rest of the book, as the ideas in them are hard to understand well without prior experience in mathematical modeling. The remainder of the chapter is divided into sections on mechanistic modeling, which starts with assumptions about the scientific setting, and empirical modeling, which starts with examination of data. The three sections on empirical modeling are of value to anyone who collects and analyzes data, but they are not essential background for the rest of the book. The basic ideas of mechanistic modeling (Section 2.5) are helpful to try to work through, but the reader does not need to be expert on this subject. Most biologists need to be able to read

and understand mechanistic models but do not need to be able to create them. The reader of Part III of this book is often asked to interpret mathematical models mechanistically, but is never asked to construct one. In contrast, nondimensionalization (Section 2.6) is a vital skill for anyone who wants to do any work with dynamical systems. It would have been impossible to write a useful introduction to dynamical system using complete versions of well-known models without relying on the power of nondimensionalization to simplify model analysis.

The accompanying sketch shows the interdependencies of the sections in Part I. Sections 1.1, 1.2, 1.3, and 2.2 are necessary background for the remainder of the book. Sections 1.7 and 1.8 are needed for Part II, while Sections 2.5 and 2.6 are needed for Part III. Sections 1.4, 1.5, 1.6, and 1.9 are important topics in calculus, but are not necessary for the rest of the book; similarly, the empirical modeling topics in Sections 2.3, 2.4, and 2.7 are not needed later.



Part II

Probability

Chapters 3 and 4 provide a foundation in probability and a small amount of inferential statistics. In most books, probability appears in the context of statistics. I prefer to think of probability as a subject in its own right. Descriptive statistics is essential background for probability and appears in Section 3.1. I can then present probability as the study of mathematical models of data sets. This viewpoint places the focus on probability distributions, which play a prominent role in the characterization of data and are needed to run simulations for stochastic dynamical system models.

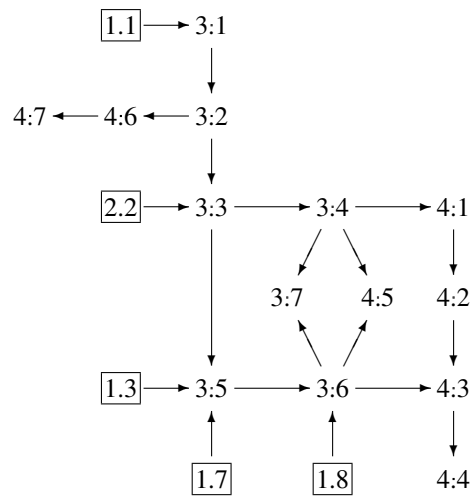
Much of statistics is concerned with the assessment of inferences, and one can see Chapter 3 and Sections 4.1 through 4.3 as providing the necessary background for the brief treatment of inferential statistics that occupies section 4.4. A full treatment of this subject lies beyond the scope of this book. I do not present analysis of variation, which asks questions such as “Does this characteristic explain variation in that characteristic?” My treatment of inferential statistics is limited to questions such as “Is the population from which this data was collected different from the general population?” which can be addressed using the standard distributions that characterize simple data sets. Section 4.4 focuses on using probability models, with minimal attention to the statistical rules of thumb that convert probability results into answers of “yes” or “no.”

The material on conditional probability (Sections 4.6 and 4.7) is normally presented before probability distributions and could be studied immediately after Section 3.2. I have placed this material at the end of part II so as not to interrupt the flow toward Section 4.4. I also believe that it is better pedagogy to focus on independent random variables in detail before examining dependence.

The problem sets include several case studies that are split over multiple sections:

Section	3.1, 3.4	3.5	3.6	3.6,3.7	4.2	4.4	4.5
Dopamine and psychosis	3.1.3				4.2.1	4.4.2	4.5.1
Fruitfly egg production	3.1.4				4.2.2	4.4.3	4.5.2
Weight gain in rats	3.1.5				4.2.3	4.4.4	4.5.3
Blood pressure medication	3.1.6					4.4.6	
Resting pulse rates	3.1.7		3.6.9		4.2.5		
Left-handed presidents	3.4.5					4.4.8	
Chest sizes of soldiers		3.5.2	3.6.6				
Cuckoo egg measurements		3.5.3	3.6.7				
Butterfat in cow milk		3.5.4	3.6.12	3.6.13	4.2.4	4.4.5	4.5.4
Iris sepal lengths		3.5.5 3.5.6	3.6.10	3.6.11	4.2.7		4.5.5
Malaria parasite counts				3.7.6			4.5.9
Spiders under boards				3.7.7			4.5.10
Nerve pulse times				3.7.13	4.2.8		

The accompanying sketch shows the interdependencies of the sections in Part II and connections to prerequisite topics.



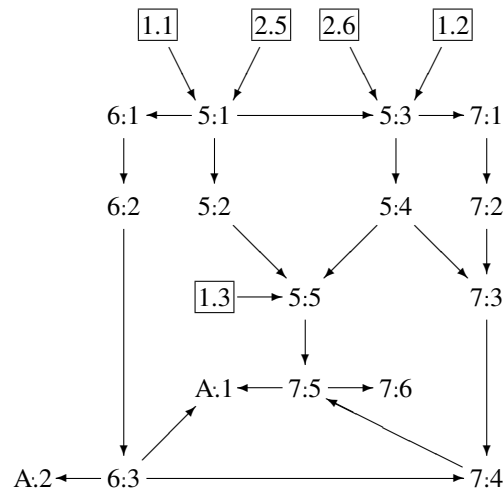
Part III
Dynamical Systems

Chapters 5 through 7 provide a treatment of dynamical systems, beginning with one-variable discrete and continuous equations in Chapter 5, progressing to discrete linear systems in Chapter 6, and concluding with nonlinear continuous systems in Chapter 7. It is easy to see how to interpret discrete systems and use them for simulations. Their advantages end there, and the remaining advantages are with continuous models. Continuous systems have superior graphical methods and simpler mathematical properties. These advantages, which will become apparent in Chapter 5, more than offset the initial advantages of discrete models. In general, one should only use discrete models when synchronicity of events dictates discrete time.

The importance of good nondimensionalization of models is a recurring theme. This is one additional advantage of continuous models, which allow for nondimensionalization of time. At minimum, nondimensionalization reduces the number of parameters requiring estimated values for simulation or study in analysis. Beyond that, it can sometimes be used to reduce the number of essential components in a model. As will be seen in Chapter 7 in particular, analysis of models becomes more difficult as the number of components increases, while graphical methods are generally limited to one-component discrete models and two-component continuous models. Any reader who has skipped Section 2.6 so as to get to this point sooner is strongly urged to go back and study that section in detail before continuing.

The reader should have noticed that there is no mention of discrete nonlinear models in the description of the chapters in this part. I have not omitted these models entirely, but have relegated them to an appendix. This choice makes the material accessible to those who need it, while deemphasizing it according to the author's professional judgment.

The accompanying sketch shows the interdependencies of the sections in Part III and connections to prerequisite topics.



Index

- Accuracy, 115
- Akaike information criterion, 165–169
- Asymptotic stability, 421
- Atlantic croakers, 138

- Basic reproductive number, 15
- Bayes' formula, 306
- Bernoulli trial, 209
- Beverton-Holt model, 324
- Bieber fever, 16
- Bifurcation, 13–14, 453
- Binomial distribution, 209–212, 251
- Birthday problem (probability), 196
- Boolean function, 484
- Boolean network, 484
- Boolean variable, 482
- Briggs-Haldane approximation, 411
- BUGBOX-population, 373
- BUGBOX-predator, 103
- Butterfat content of milk, 224, 233

- Carbon dioxide, 21
- Central limit theorem, 272
- CG ratio, 193
- Chain rule, 36, 65–67
- Characteristic equation, 435
- Chauvenet's test, 262
- Cheetahs, 392
- Chemostat, 150, 414
- Cobweb diagram, 327–330, 349
- Codon, 189
- Colorado potato beetles, 325
- Compartment analysis, 148, 396–400
- Complement rules (probability), 195
- Conceptual model, 109
- Conditional probability, 296–300
- Confidence intervals, 285–286

- Continuous probability distribution, 218
- Cramer-von Mises test, 260–261
- Critical point, 43
- Cuckoo eggs, 224, 232
- Cumulative distribution function, 202, 217–220

- Decay model, 12, 141–143
- Deepwater Horizon oil spill, 71
- Definite integral, 73–76
- Degree days, 78
- Demographic stochasticity, 104, 208
- Derivative, 25–26
- Derivative formulas, 32
- Development rate, 78
- Differential equations, 63
- Dimensional analysis, 143–144
- Dimensional consistency, 143
- Dimensionless variables, 156–159
- Discrete exponential model, 318–320
- Discrete logistic model, 320–321
- DNA, 188
- Dogs and calculus, 57, 69
- Dominant gene, 207
- Drug absorption, 438

- Eigenvalues, 432
- Empirical distribution function, 258
- Empirical model, 109
- Enzyme kinetics, 406–411
- Epidemic model, 150
- Equilibrium solution, 335, 338, 416
- Error propagation, 49
- Euler's method, 334
- Euler-Lotka equation, 77, 88, 117
- Eutrophication, 344
- Event (probability), 191

- Expected value, 204, 223
- Experiment (probability), 190
- Exponential distribution, 222, 238–242
- Exponential growth model, 333
- Falcons, 372
- False positive, 305
- Fast variable, 421
- Felsenstein model, 481
- Fiddler crabs, 68
- Fish growth, 138
- Fisher's iris data, 217, 225, 227, 233
- Fishery model, 318, 341
- Fixed point, 322, 485
- Flour beetles, 470
- Focal length, 68
- Fundamental theorem of calculus, 83–87, 93
- Gambler's fallacy, 206
- Geometric distribution, 214
- Global extrema, 50
- Grape harvest dates, 128
- Half-life, 12
- Hardy-Weinberg principle, 116
- Hare-lynx data, 429
- Hassell model, 325
- Histogram, 181
- HIV, 448–453
- Holling type 2 model, 147, 154, 343, 460
- Holling type 3 model, 336, 341
- Holling's human experiment, 102
- Immune system, 453
- Independence (probability), 295
- Influenza, 15
- Insects
 - Hemiptera, 244
 - Staphylinidae, 244
- Jacobian, 461
- Jacobian matrix, 442
- Jukes-Cantor distance, 479
- Jukes-Cantor model, 476
- Jury conditions, 462–463
- Kimura model, 481
- Law of averages, 206
- Lead poisoning, 396–402, 431, 438
- Lead poisoning tests, 303
- Least squares method, 120–126, 129–132
- Left-handed presidents, 214
- Likelihood function, 287–292
- Limit, 25
- Linear approximation, 46–47, 441–442, 461
- Linear programming, 61
- Linear systems of differential equations, 431–438
- Lineweaver Burk method, 138
- Local extrema, 41–46
- Loggerhead sea turtles, 392
- Logistic growth model, 334–336
- Lotka-Volterra model, 107, 115
- Macrophage, 453
- Malaria, 52, 243
- Marginal value, 47
- Marginal value theorem, 55
- Mark Twain, 284
- Mark-and-recapture, 289–291
- Markov chain, 476
- Mass action, law of, 406
- Mathematical model
 - approximation, 109
 - characterization, 110
 - definition, 107
 - derivation, 110
 - parameterization, 111
 - selection, 110
 - simulation, 111
 - validation, 111
- Matrix
 - $\mathbf{Ax} = \mathbf{0}$, 380
 - characteristic polynomial, 384
 - determinant, 379
 - eigenvalues, 383–385
 - eigenvectors, 383–386
 - identity, 376
 - main diagonal, 376
 - population projection, 377
 - product, 376
 - singular, 380–381
- Mean, 182, 205, 223
- Measure it with your hand, mark it with a pencil, cut it with a laser fallacy, 115, 194
- Mechanistic model, 109
- Median, 182
- Mendel's pea experiments, 207
- Michaelis-Menten reactions, 138, 147, 406–411, 416, 442
- Model selection, 166
- Moose diet, 61
- Multiplication rule (probability), 299
- Negative binomial distribution, 246
- Nerve pulses, 245

- Nicholson-Bailey model, 469
- Nitrogen in water, 412
- No-egress region, 421
- Nondimensionalization, 157–159, 401, 408, 449
- Normal distribution, 226–232
- Nucleotides, 188
- Nullcline, 417
- Nullcline analysis, 418–424
- Occam's razor, 164
- Optimal fitness, 97
- Optimal foraging, 18, 54, 58
- Optimal harvesting, 341, 343
- Optimal size, 51, 64, 80, 158
- Organism growth, 68, 97
- Outcomes (probability), 190
- Outliers, 183, 262–263
- Parameters, 8
- Parasitoids, 242, 468
- Partial derivatives, 37
- Partition rule (probability), 299
- Pea aphids, 391
- Pharmacokinetics, 396, 438
- Phase line, 346–349
- Phase plane, 416
- Phase portrait, 416
- Phylogenetic distance, 472–480
- Plankton, 412
- Plant density, 244
- Plant growth, 69, 79
- Poisson approximation for the binomial distribution, 245
- Poisson distribution, 235–238
- Pollution in a lake, 148–149
- Pollution in lakes, 344
- Population growth, 19, 29, 76, 88
- Potato beetles, 325
- Precision, 115
- Predation, 67
- Predator-prey model, 429
- Probability density function, 220
- Probability distribution, 193
- Probability distribution function, 201
- Probability distributions of sample means, 268–272
- Product rule, 34
- Quotient rule, 34
- Random variable, 191, 198
- Rate of change
 - continuous, 22–25
 - discrete, 21–22
- Reaction velocity, 407
- Recessive gene, 207
- Recruitment, 317
- Red blood cells, 372
- Relative rate of change, 333
- Renewable resource model, 336–340, 347
- Residual sum of squares, 123
- Resource consumption, 26, 144–147
- Resource management, 342
- Ricker model, 324
- Riemann sums, 489–490
- Roots of quadratic polynomials, 435, 440
- Rosenzweig-MacArthur population model, 414
- Routh-Hurwitz conditions, 435–438
- Sample space, 190
- Schaefer model, 341
- Science in the early 1900s, 180
- Second derivative test, 44
- Self-limiting population, 443
- Self-limiting population model, 413
- Semilinear least squares method, 133–136
- Sensitivity (of medical tests), 305
- Sequences (probability), 194–195
- Set theory notation, 190
- Sickle cell disease, 52
- Simulation
 - continuous, 334, 398, 409, 491–492
 - discrete, 321
- SIR disease model, 150, 403
- SIS disease model, 403, 405
- Smallpox, 16
- Specificity (of medical tests), 305
- Spiders under boards, 243
- Stability
 - continuous, 347, 356–357, 421
 - continuous linear systems, 432–438
 - continuous nonlinear systems, 442–445
 - discrete, 323, 329, 351–356
 - discrete nonlinear systems, 462–467
- Standard deviation, 184, 205
- Statistics, 232
- Sterile insect model, 325
- Structured population models, 364–370, 387–388
- Student's yeast data, 244
- Substitution rule, 91
- Sunflowers, 67
- Survival probability, 202
- Swimming speed of fish, 56, 160
- Systems of linear algebraic equations, 375, 386

T cells, 448
Tangent line, 40–41
Tangent slope, 26
Tay-Sachs disease, 308
Teasel, 389
Tomato plants, 213
Tribolium, 470

Uniform distribution
 discrete, 193

Variance, 205, 223
Von Bertalanffy equation, 68, 97

Weevil eggs on beans, 243
Weightlifting, 17, 139
White lies, damn lies, and statistics, 284

Yahtzee, 214
Yeast, 244, 413, 443