Teaching Calculus, Probability, and Statistics to Undergraduate Life Science Majors: A Unified Approach

Frederick R. Adler, University of Utah *

Institution	University of Utah
Size	about 2000 students
Institution Type	Large comprehensive state university offering bachelors through doctoral degrees
Student Demographic	Biology majors (recommended in lieu of the standard first two semesters of calculus, but taken only by a minority due mainly to scheduling constraints.
Department Structure	Mathematics and Biology are separate departments in the College of Science.

Abstract

The Department of Mathematics at the University of Utah has developed an integrated modeling, calculus, and probability course for life science majors that emphasizes the central role of dynamics in biological thinking, including statistical analysis. The course and associated textbook use the themes of growth, diffusion, and selection throughout, and we describe how diffusion is presented in several contexts. Although the University of Utah has a large group of faculty and students in mathematical biology, we discuss the challenges of serving all biology majors, integrating the course with the biology and mathematics curriculum, and motivating mathematically underprepared students.

COURSE STRUCTURE

- Weeks per term: Two semesters, about 30 weeks
- Classes per week: Three 1-hour lectures
- Labs per week: One 1-hour lab
- Average class size: 25-50 students

*adler@math.utah.edu

- Enrollment requirements: For freshman, students should have passed college algebra and trigonometry, or have Math SAT score of at least 630, Math ACT score of at least 28, or AP Calculus AB score of at least 3
- Faculty/dept per class, TAs: One math instructor with one graduate student TA to handle the computer labs
- Next course: The purpose of this course is to prepare students for the quantitative aspects of the biology curriculum
- Current website: `http://www.math.utah.edu/~borisyuk/1170/`
- Archived website: `http://www.math.utah.edu/~adler/oldcourses/math1170_2008/`

Scope and goals

In 1990, the University of Utah received a grant from the Howard Hughes Medical Institute; among the goals of the proposal was the development of a new calculus course and associated textbook for life science majors. The University of Utah had, and still has, a one-year mathematics requirement for biology majors. To this end, I was hired as a new faculty member into a joint position in the Departments of Mathematics and Biology to develop the course.

I was given one semester for preparation, and had a committee of experts in biology and mathematics to help with planning. The goal was to create a course that could serve the quantitative needs of the large majority of life science majors, but without closing off the option of taking additional mathematics courses.

After surveying the biology faculty, we found that modeling and statistics were the two most important topics (along with the basics of being able to read a graph effectively and other remedial skills). The course was designed to emphasize these elements by covering three areas under the quarter system then in effect:

1. Discrete-time dynamical systems
2. Differential equations
3. Probability and statistics

Within these broad headings, united by their focus on modeling, the fundamental ideas of functions, differentiation, and integration were to be seamlessly intertwined.

We chose to take a conventional approach toward teaching in two ways. First, the course would build from simple basic principles toward more complex ideas (rather than confronting students with realistic data to analyze right from the start). Second, much of the learning would be done in a traditional classroom setting, but supplemented with an additional hour each week in the computer lab.

Because no textbook then existed that matched this model and built from biological principles, I began to draft what became *Modeling the Dynamics of Life: Calculus and Probability for Life Scientists* (Adler, 1998; Adler, 2005).

The biological themes were simple and chosen so that they could be revisited in growth, diffusion, and selection. The goal was to develop materials that would focus on understanding the modeling process, with as much (or as little) attention on the mechanics of algebraic calculation as was necessary to achieve this goal.

The section on probability and statistics was to tie the work on deterministic models together with data analysis using the basic models of probability theory (the binomial, Poisson, exponential and normal distributions). Through computer experimentation, students were to see how the probabilistic nature of population growth, selection, and diffusion make patterns difficult to see without the proper tools, namely a combination of basic understanding and statistical insight.

Course Content

The course begins with a four-week section that introduces modeling by using discrete-time dynamical systems, building or renewing students' familiarity and facility with linear, exponential, and trigonometric functions and their graphs. Starting with the classic model of exponential growth, we derive a linear diffusion model of gas exchange in the lungs, and a non-linear model that shows how a selectively-favored allele spreads through a population. Students use graphical methods, such as cobwebbing, to visualize and understand dynamics.

The central portion of the first semester introduces derivatives and some of their applications. After learning the meaning of the derivative, its graphical interpretation, and computation rules for biologically important functions, students use the derivative to analyze stability of equilibria of discrete-time dynamical systems. Other applications include optimization, using the derivative to understand graphs of complex functions, and using the tangent line and Taylor polynomials to approximate functions. If time permits, we tie together approximation and discrete dynamics tie together with Newton's method for numerical solution of nonlinear equations.

The first semester concludes with an introduction to integration, starting with the antiderivative as the way to solve simple pure-time differential equations, where the rate of change depends only on time. Riemann sums and the interpretation of integrals as areas are delayed, and emphasize the importance of the fundamental theorem of calculus in linking definite and indefinite integrals. The principal methods of integration are covered, although in less detail than in a standard calculus course.

The second semester begins with autonomous differential equations and shows how they parallel discrete-time dynamical systems. The phase-line diagram, like cobwebbing for discrete-time dynamical systems, provides graphical understanding. The parallel with discrete-time dynamics continues with the application of the derivative to analyze stability, and the parallel with pure-time differential equations continues with their solution using separation of variables. The phase plane forms the culmination of the deterministic portion of the course, emphasizing the interplay between modeling and analysis in the dynamics of a neuron as described by the Fitzhugh-Nagumo equations.

The remainder of the second semester, 10-12 weeks, is devoted to probability and statistics, with an emphasis on modeling in data interpretation. The first section develops the concepts of probability, focusing on conditional probability and independence, and includes the visual display of probabilistic information. The second section introduces probability distributions through their derivation from discrete-time dynamical systems (the binomial and geometric distributions) and from differential equations (the Poisson and exponential distributions). This section concludes with the central limit theorem (presented without proof) and the normal distribution, along with the difficult concept of the probability density function and its link with the fundamental theorem of calculus.

The course concludes with an introduction to statistics designed to tie together the key ideas of modeling and calculus. Maximum likelihood forms the backbone, showing how maximization methods can be applied to data when interpreted in terms of probability distributions. Although necessarily brief, the key ideas of classical statistics, such as confidence limits and hypothesis testing, are introduced with their application to linear regression, analysis of variance, and contingency tables.

An Extended Example: Diffusion

The course treats diffusion in several ways. In the first chapter, before the introduction of calculus, students develop a discrete-time model of gas exchange in the lungs,

$$c_{t+1} = (1 - q)c_t + \gamma q,$$

where c_t is the concentration of some inert gas as a function of the fraction of air exchanged, q, and the ambient concentration. The students derive this from first principles (keeping track of air volumes), and then as a weighted average. This allows students to develop more realistic equations for a gas, like oxygen or carbon dioxide, that is used or created in the lung with each breath. After the development of the derivative, students use stability analysis to show that the lung will indeed approach the ambient concentration, and to study the effects of non-linearities.

The second portion of the course, on differential equations, introduces both Newton's law of cooling and the formally identical law for chemical diffusion between two containers in continuous time,

$$\frac{dC}{dt} = \beta(\Gamma - C),$$

where β is rate of chemical exchange and Γ is the ambient concentration. Students extend the derivation to more interesting biological processes and use graphical and algebraic methods to evaluate stability.

Finally, in the section on probability, students meet diffusion from the molecular perspective as a stochastic process. They see that the probabilities that describe the location of a molecule follow exactly the discrete-time dynamical system found with macroscopic reasoning about volumes. After deriving the binomial distribution, they see that the ensemble of molecules, assuming independence, obeys it, and in the limit the ensemble behaves like its expectation. They have then come full circle to see that the deterministic discrete-time dynamical systems and differential equations derived and studied in the first part of the course are the equations for the expectation of a stochastic process.

The familiarity of diffusion allows students to use their intuition to derive and understand mathematical models that can then challenge and extend that intuition. The multiplicity of modeling approaches illustrates that the tool chosen to study a problem depends on the problem and the question being asked, and shows that apparently different methods can be closely related.

Successes and Failures

Although no formal assessment has been done, students seem to enjoy the course, and some show evidence of having used the material in other courses. The University of Utah has a tradition of encouraging undergraduate research in biology (the initial Hughes grant funding the development of the course was focused on further strengthening these programs), and many students have been attracted to this course because of the benefits for research.

The course was to be taught to a single section of students in the pilot phase and thenscaled up to serve all biology majors. This has not occurred, primarily because of scheduling difficulties, and the course remains at one or two sections. Strong advising from the biology department has maintained enrollment although the course is not required and is widely considered to be harder than the ordinary calculus sequence. The majority of the students have not taken calculus before, and are comparable in mathematical background to their peers who enroll in the ordinary calculus sequence.

We have found that separating biologists from physicists and engineers has promoted a positive and collaborative atmosphere, which extends to the instructors. Over the years, the course has been successfully staffed by both faculty and advanced graduate students, thanks to the large mathematical biology group at the University of Utah. The shared motivation of life science students and the choice of biologically relevant topics provides the best argument for such a course, with its implementation depending on an institution's requirements and goals.

The course has faced many problems. The pressure to develop a second edition of the book took energy away from improving the teaching of the course, leading to a period of relative stagnation. Work on the second edition coincided with a switch to semesters that broke up the elegant three part structure in an unnatural way, with differential equations being divided over two semesters.

There is little doubt that the course is more difficult than the standard calculus sequences because of the challenges of modeling, the open-ended material, and the diversity of topics. Given the wide range of mathematical backgrounds and abilities among students, the goals of the course must be adjusted along a sliding scale. The weakest students at least gain familiarity with concepts such as models, differential equations, and statistics. The strongest students can apply what they have learned to work in other courses and in research. Those with the most mathematical enthusiasm and ability can in principle move on to advanced calculus, linear algebra, or mathematical biology, but the department advisors have discouraged this, assuming that students from this course will lack the background needed for more advanced work.

In response to pressure from advisors and reviewers, and in part for completeness, I added several topics to the third edition:

- Double-log graphs and an introduction to allometry,
- Implicit differentiation and related rates,
- Infinite series, Taylor series, and improper integrals,
- Integration by partial fractions,
- Trigonometric substitutions,
- Computing volumes of solids of revolution.

Whether these additions will help integrate this course with the rest of the mathematics curriculum will depend on the idiosyncrasies of particular institutions.

Modernizing the course will probably take two directions. First, the introductory part of the course needs to be closer to real biological problems. One of the difficulties is balancing elucidation of general principles with the complexity of real biology. The themes of growth, diffusion, and selection could be better introduced with real data, ideally student-generated. A bit of preliminary statistical analysis would help motivate the models and their analysis. Second, the lack of bioinformatics is frustrating. Mathematicians believe that a course for beginning undergraduates must avoid methods presented without rigorous development,[1] and methods for dealing with complex genetic data cannot be developed from scratch at this level. However, carefully chosen examples should reveal the methods, challenges, and excitement of modern biology.

Even in the biology-friendly atmosphere of the University of Utah Department of Mathematics, integrating this course with the rest of the curriculum has been challenging. We have failed to integrate the material learned with the rest of biology curriculum, due to the inability of the University of Utah to enforce prerequisites, and to the challenges of incorporating any new material in already over-stuffed biology courses.

Looking Ahead

Today's college educator must contend with two conflicting problems. Entering students generally have poor quantitative skills of all sorts, from number sense through algebra and including lack of real computer programming experience and knowledge of statistics. However, real biological problems, whether in the realm of research or in medicine, are becoming ever more complex. Those of us trained in the step-by-step logic of mathematics find it unacceptable to present students with a series of black boxes that can be used to solve problems. But the jump from solving and understanding a linear discrete-time dynamical system to appreciating the logic required to make inferences from genetic data seems too large to make in a single year. Much lip service is paid to integrating quantitative material into biology courses, but few institutions have succeeded in doing so in a serious way due to the limitations of both faculty and students.

The different sizes, emphases, and personalities of institutions make a one-size fits allmodel for integrating calculus-level mathematics with a life science curriculum inappropriate. Each institution needs to use its strengths (such as the large mathematical biology group at the University of Utah) to advance quantitative education and initiate the long-awaited generational shift in biological thinking.

References

Adler, F. A., 1998: *Modeling the Dynamics of Life: Calculus and Probability for Life Scientists*, Brooks/Cole.

[1] Glenn Ledder has described mathematicians as "people who believe that you should not drive a car until you have built one yourself."

Adler, F. A., 2011: *Modeling the Dynamics of Life: Calculus and Probability for Life Scientists, Third Edition*, Brooks/Cole.