Some Lessons form Fifteen Years of Educational Initiatives at the Interface between Mathematics and Biology: The Entry-Level Course

Louis J. Gross[1], Departments of Ecology and Evolutionary Biology and Mathematics,
University of Tennessee-Knoxville

Name of Institution	University of Tennessee - Knoxville
Size	About 28,000 students
Institution Type	Large comprehensive state university
Student Demographic	One of two possible mathematics sequences for biology majors, and required for some agriculture majors
Department Structure	Sequence is based in the Mathematics Department but is sometimes taught by faculty and students based in the Division of Biology

Abstract

In 1992, I led a workshop devoted to undergraduate quantitative education for life science students. One outcome was a suggested set of topics for the entry-level mathematics course for life science undergraduates. With NSF support, I developed a two-semester course that followed this recommended topic coverage and represented a considerable departure from the emphasis on calculus that was (and is) prevalent for undergraduates in the sciences. This two-semester course sequence involves an hypothesis testing and data-rich view of mathematics, includes computer-based projects, and has been taught by instructors with widely varying backgrounds and levels of biological expertise over the past eighteen years. It is the main mathematics sequence for biological science undergraduates at UTK, with approximately 300 students enrolled each semester, and was included as one of the case studies in the National Research Council's Bio2010 report. I discuss the goals of the course (a focus on conceptual development rather than on skills), provide details on the topic coverage and student assessment methods, note how the course sequence has changed since it was initiated, and discuss the challenges arising from the diversity of backgrounds of instructors.

Course Structure

- Weeks per term: 15-week semester
- Classes per week/type/length: two 50-minute periods and one 75-minute lab
- Labs per week/length: one 75-minute lab/recitation
- Average class size: 150 students with 35-40 students per section
- Enrollment requirements: Placement exam evaluation based upon high school algebra and trigonometry.

[1] gross@NIMBioS.org

- Faculty/dept per class, TAs: One math/biology faculty member with several TAs teaching the labs/recitations.

- Next course: Calculus-based statistics or Models in Biology.

- Website: http://www.tiem.utk.edu/~gross/math151.html

Background

The history of courses in mathematics in the U.S. designed specifically for undergraduate students with interests in the biological sciences dates back at least to the early 1970s with the text by Batschelet (1971). A number of texts and courses appeared in this period, with the majority emphasizing elementary calculus and its prerequisites (Arya and Lardner 1979, Cullen 1983, DeSapio 1977, Gentry 1978, Levin 1975). Many of courses had disappeared by the early 1990s, with those remaining mainly at institutions with research emphases in mathematical biology. In 1991, I began a project funded by the National Science Foundation, with a goal of developing a quantitative curriculum for life science students. The initial focus was on entry-level mathematics courses and started with a survey of quantitative course requirements for life science students at U.S. universities. The results indicated that about 80% of institutions had some calculus requirements, while the others required only pre-calculus or had no explicit requirement.

In 1992, I convened a workshop, attended by many leading mathematical and quantitative biologists with interests in undergraduate education, to discuss all aspects of undergraduate mathematics preparation for life science students. Among the recommendations from the workshop was a strong suggestion that students are not well served by being exposed only to a calculus course, because a broader array of quantitative concepts were necessary for students to be successful in careers in modern biology. At UTK we had been teaching an elementary calculus for life sciences course since the late 1970s. As an outcome of this workshop, I revised the course to incorporate a more diverse set of topics and moved the focus towards conceptual understanding and away from hand calculations based standard calculus rules.

Structure of the Entry-Level Course

The 1992 workshop provided specific suggestions on topic coverage for an entry-level course, including the addition of descriptive statistics, matrix algebra, discrete probability, and difference equations. Additional suggestions included ensuring that the calculus portion included an introduction to differential equations and incorporated the notions of equilibria and stability. Many additional suggestions were provided to incorporate specific biological concepts as part of this course, with the objective being to encourage problem-based learning. The full workshop report (Gross 1992) also provided suggestions for upper division courses and methods to link the mathematical and biological portions of the curriculum.

Based on the recommendations of the workshop, I implemented a pilot version of the course in place of the two-semester, 3-credit hour course we had been teaching at UTK. There had been concern expressed at the workshop that it would be extremely difficult to squeeze in all the suggested material into two 4-credit courses, let alone 3-credit ones. Our experiences over the fifteen years that this course sequence has been offered it that is possible. The pilot course followed closely the suggestions of the workshop, which had focused mostly on specific mathematical topics. Over the next several years, I developed a set of concepts, rather than explicit mathematical methods, that served as a focus for much of the material included in the course sequence. These have been expanded upon elsewhere (NRC 2003), so I will simply list them here:

- Rate of Change
- Scale

- Equilibria
- Stability
- Structure
- Interactions
- Stochasticity
- Visualization

The mathematics needed to formalize these concepts provided the motivation for the revisions to the entry-level course that have occurred since the pilot version was offered. The first semester of the course covers the discrete mathematics topics through limits of sequences, and the second semester covers the calculus up to first order linear and simple non-linear differential equations.

The goals of the current course are to

1. Develop the ability to analyze problems quantitatively that arise in the biological areas of interest to the students.

2. Illustrate the utility of mathematical models to provide answers to biological problems.

3. Develop appreciation of the diversity of mathematical approaches useful in the life sciences.

4. Provide experience using computer software to analyze data, investigate mathematical models, and provide some exposure to programming.

The course methods to meet these goals include

1. Encouraging hypothesis formulation and testing for both the biological and mathematical topics covered.

2. Encouraging investigation of real-world biological problems through the use of data in class, for homework, and in examinations.

3. Reducing rote memorization of mathematical formulas and rules through the use of software such as MATLAB and Maple.

4. Providing biological motivation for each main mathematical component of the course through a central example that is returned to regularly.

I will here give a brief example of each of these methods. For hypothesis formulation, the first course begins with each class section going to the field in groups of three students to measure leaf lengths and widths (we do not provide details as to how to do this, but leave it up to each group) and to make a hypothesis concerning the relationships between the variables they are measuring. The class then goes to a computer lab, where we provide an introduction to basic MATLAB commands, have the groups enter data, and produce histograms and scatter plots to evaluate their hypotheses. This requires a complete lab session (1 1/4 hours) but serves an essential purpose of pointing out that mathematics connects to data, that hypothesis evaluation is part of science, and that mathematics can assist in the scientific process.

For real-world biology, the class is presented with a recently published paper (typically from *Science*, *Nature*, or *Proceedings of the National Academy of Sciences*) that includes mathematics related to the topic of the day. Although few of the students are capable of reading the paper in detail, they point out the importance of mathematical ideas in biological research, and some of the data are used directly to motivate the mathematics as well as appearing on tests and quizzes in the course. The course web page includes links to the papers and a brief summary I compose that discusses the results and how they relate to the course topics.

On the use of software, there are several objectives beyond reducing the need for hand calculation. One is the ability to include much larger data sets or deal with much more complex (and realistic) projects than could be included otherwise. For example, one project asks students to make a hypothesis about whether their height changes overnight, which the class evaluates by collecting data on themselves for several nights. The data are then combined and analyzed using MATLAB to illustrate histograms and regressions, point out the importance of looking carefully at data for bad values (e.g., data which imply a student gained 10 cm in height overnight!), and to discuss the obligation to deal carefully with human data (through discussion of the Institutional Review Boards that oversee human data collection at every institution). We have chosen to use MATLAB as the software for the initial portion of the course (we move to Maple for much of the calculus portion) because it introduces students to basic coding and logic and the idea of an algorithm.

For the matrix algebra section students are shown images of landscapes, asked how to characterize them, and asked to develop hypotheses about how they might change as the images are retaken over several decades. This leads naturally to the use of vectors to characterize the distribution of states across the landscape and provides an introduction to the ecological succession, which is typically familiar to students. It serves to motivate students to develop the notion of a Markov chain, which they can readily do themselves, though not with any detail on how the mathematics might work. The concept of an eigenvector as the long-term fraction of the landscape in each state (such as urban, forest, or agricultural) then arises naturally and the notion of stability is readily motivated through consideration of different initial landscape state distributions.

Student Assessment

Over time we have developed three major and one minor assessment method to evaluate student performance in the course sequence. There is a formal set of written exams covering the basic concepts and a comprehensive final exam that account for 60% of the course grade. The exams often include examples taken from data or research papers discussed in class. Students are encouraged to use standard calculators as they wish in the exams, though we do not emphasize their use in the course assignments. The exams are structured and evaluated, however, so that the use of a calculator is not necessary–the focus is on basic concepts and calculations or graphs that can be done readily by hand. Students find calculators helpful if their arithmetic skills are weak (and our focus in the sequence is not at all on arithmetic skills).

The course includes a set of weekly quizzes that cover the material assigned for homework (which we do not collect or grade). The quizzes are supplemented currently by a required solution at the board of a homework problem that counts as an additional quiz. These count for 20% of the course grade. Finally, we have a set of computer-based projects. Some require students to collect their own data, but most involve writing a formal laboratory report on a set of simulation experiments carried out using MATLAB or Maple. Sometimes we supply the code and the students modify it in a few places, and sometimes the students must develop their own code. The more complex of the projects are essentially research-level (such as the impact of random environmental variation on the dynamics and long-term structure of matrix population models), the motivation for which is readily understood by the students, although carrying out a full analysis would require extensive graduate-level mathematics. A collection of material, including syllabi, exams, basic lecture notes, and descriptions of the scientific journal publications used as examples, is available at the project website.

Challenges

Since its development in 1993, the sequence has been taught by approximately eighty different people. Most teachers have been from the Mathematics Department, faculty as well as part-time instructors and graduate students, while some have been from the Biology Department. The majority have had some exposure to mathematical biology (through our long-standing graduate education programs in this field) but quite a few have not. This has created challenges because there is less emphasis on the biological aspects of the course when taught by someone with little biological background. To deal with this we have developed a course supplement and instructors guide to provide motivating examples for the course. Additional lecture notes that develop the biological connections for the course and provide the various links to and explanations of research papers used to motivate sections of the course are provided on the course web site. To assist biologically naïve instructors, modules were developed to motivate many of the concepts in a general biology course through a connection with mathematics (Harrell et al. 2002).

The first course in the sequence is the most difficult to teach for those with little biological background (the second course focuses on calculus). It has been taught recently to about 200 students in a lecture that meets twice weekly and breaks into smaller sections for a once-a-week lab. Instructors have included other faculty members and graduate students with an extensive mathematical biology background. The large-lecture, small-lab section format has allowed only limited small-group discussion except in the weekly lab sessions. With only two meetings a week in lecture, topic coverage proceeds at a much more rapid pace than might be preferable.

An additional challenge has been the lack of an appropriate text for the sequence. Until recently we used one of the older texts (Cullen 1983) supplemented by notes and on-line materials. Although several new texts have appeared, they have a different topic coverage (often too extensive for this course) and have little emphasis on data and its relation to course topics. Thus we have developed a new text, used it for two years, and expect publication in 2013 (Bodine et al., submitted).

Lessons and Suggestions

In many ways the course is substantially more challenging, both for the students and the instructors, than the standard science and engineering calculus sequence or the standard two-semester calculus for business and social science students. It covers a vast array of topics very quickly, and students who are not well-prepared find it quite difficult. However, they represent a majority of the students majoring in pre-health sciences and are extremely motivated, realize the importance of working hard in order to be accepted by various professional schools, and often appreciate the connection between this course and their other biology courses. The class typically includes students with varying levels of biological experience (some of them delay taking the course until much later in the undergraduate program than we prefer). I view this as beneficial because they can point out to their peers how the topics relate to those covered in more advanced biology courses. An example is the Hardy-Weinberg equilibrium that students see in their general genetics course, but which we derive mathematically.

Despite the fact that students taking this course have had a pre-calculus background (mostly in high school) and have taken a mathematical skills placement exam to be recommended for the course, their comprehension of some pre-calculus concepts is weak. So, we start with descriptive statistics, including semi-log and log-log graphs, with application to biological growth and decay processes and allometry. This refreshes students understanding of exponentials and logarithms, motivated through biological observations, and introduces new scaling concepts. Providing a new perspective, based upon observations, of a topic that many students did not grasp completely may let students develop an understanding of non-linear scaling and why logarithms appear so often in biology.

Few students enter the course with any understanding of programming or the logic that underlies computer coding. It requires considerable effort on their part to follow basic MATLAB codes given them, and they have great difficulty developing their own programs. Initially, we gave students much of the MATLAB codes required for the course, but we have since reassessed the situation and now require the students to develop their own codes for at least the simpler computer projects. Otherwise we found that students had little understanding of the steps in the code. We include basic guides to MATLAB commands associated with each project.

Conclusion

We have had no way to assess the course's impact. Well over two thousand students have taken the courses over the past fifteen years, but we have not tracked students to compare them with those who take the standard science and engineering calculus sequence. Students complete evaluations but none have been compiled to allow for comparisons either across time or between students in different courses. Those beginning to teach such a course such should design assessment, find funding to implement it over a number of years, and track student success and behavior. Student information systems are now commonly available that would allow a comparison of subsequent undergraduate success between students following a different mathematics course sequence and those in a sequence such as the one described here. At this point, all I can point to is anecdotal information, including one student who told me that he was instructed by his doctor to take the course, as it had been beneficial in their own medical career. As gratifying as this is, we need to do better in assessing the impact of courses that fall at the interface between mathematics and biology.

References:

Arya, J. C., and R. W. Lardner, 1979: *Mathematics for the Biological Sciences.* Prentice-Hall, 712 pp.

Batschelet, E., 1971: *Introduction to Mathematics for Life Sciences.* Springer-Verlag, 643 pp.

Bodine, E., S. Lenhart, and L. Gross, (submitted): *Mathematics for the Life Sciences*. Princeton University Press.

Cullen, M. R., 1983: *Mathematics for the Biosciences***.** PWS, 751 pp.

DeSapio, R., 1977: *Calculus for the Life Sciences.* Freeman, 739 pp.

Gentry, R. D., 1978: *Introduction to Calculus for the Biological and Health Sciences.* Addison-Wesley, 686 pp.

Gross, L. J., cited 2010: Summary Report 1992: Workshop on a Quantitative Sciences Curriculum for Life Science Students. [Available online at http://www.tiem.utk.edu/~gross/WorkshopReport92.txt.]

Harrell, S., M. Beals, and L. J. Gross, cited 2010: General Biology Modules (2002). [Available online at http://www.tiem.utk.edu/~gross/bioed/modulelist.html.]

Levin, S. A., 1975: *Mathematics for the Biological Sciences.* Cornell University, 654 pp.

National Research Council, 2003: *BIO2010: Transforming Undergraduate Education for Future Research Biologists***.** National Academies Press, 192 pp.