Abstract
Biology and mathematics are inextricably linked. In this article, we show a few of the many areas in which this linkage might be made explicit. By doing so, teachers can deepen students’ understanding and appreciation of both subjects. In this article, we explore some of these areas, providing brief explanations of the mathematics and some of the applications in biology, with ideas for some of the investigations that teachers might wish to pose to their students.

Key Words: Math in biology; mathematical applications; quantification in biology.

Biology and mathematics are inextricably linked. By making biology more quantitative, teachers can deepen students’ understanding and appreciation of both biology and mathematics. With the “Next Generation Science Standards,” which calls for greater mastery and more engagement in scientific inquiry in the learning of all content, teachers must be familiar with the role that mathematics can play in helping students make sense of what they observe. Certainly, the idea of linking mathematics and biology is not a new one. Jungck (1997) discussed 10 key equations that, over history, have transformed the study of biology, from the relationships between size and shape, to logistical models of populations that take carrying capacity into consideration, to genetic coding. Texts, such as Britton’s (2003) work on the mathematics of biology, can make these connections explicit. Here, we explore a few mathematical ideas that biology teachers can use and adapt to connect these key content domains.

Precision vs. Accuracy
Though people often use the terms precision and accuracy interchangeably, they do not represent the same idea. Precision is an absolute term, suggesting how much error might be in a given measurement. For example, we might measure distance to the nearest kilometer, such that if we determine that a distance is, say, 7 km, we mean that it is between 6.5 and 7.5 km. By measuring to the nearest 10th of a kilometer, we show greater precision. For example, if we determine that the distance is 6.8 km, we are now implying that it is between 6.75 and 6.85 km.

This can lead to some non-intuitive results, suggesting a need for care. Say, for example, that points A, B, and C are on the same path in that order, and the distance from A to B is 2 km and the distance from B to C is 2 km. How far is it from A to C? Seems simple, right? We’ve always known that 2 + 2 = 4. However, the distance from A to B might be as little as 1.5 km or as large as 2.5 km. The same holds true for B to C. Therefore, all we know is that the distance from A to C is between 3 and 5 km! Consequently, 2 + 2 might equal 3 or it might equal 5!

Though accuracy is related to precision, it is a relative, not absolute, term. Specifically, accuracy refers to how close a measurement is to the actual measurement. Suppose that we determine a measure to be 7 km (measured to the nearest km). As discussed above, it is possible that the actual distance might be as little as 6.5 km or as great as 7.5 km. If we suppose the actual distance is 7.5 km, then our error was 0.5 km out of 7.5 km. Because 0.5/7.5 is ~6.7%, we could say that we had an error of 6.7%, or 93.3% accurate.

Suppose we measured the distance around the earth with the same precision – to the nearest kilometer. For simplicity, further suppose that we obtain 39,999 km but the actual distance is 40,000 km. Our error, in relation to the actual distance, is now 1/40,000, or 0.0025%. We are 99.9975% accurate. With the same degree of precision, we have wildly different degrees of accuracy.

How can these ideas play out in biology? Students should understand these ideas wherever measurement plays a role.
Whether we are measuring volumes, distances across a cell, or a population, students should have a sense of how precise they have been and how accurate their results might be.

- **Powers of 10 & Multiplication by 1**
  As in other sciences, in biology students must grapple with extremely large and extremely small numbers, making an understanding of powers of 10 and scientific notation essential. Consider two different samples of bacteria, one with 10^7 bacteria and the other with 10^9. To some students, it might seem that there's not that much difference between 10^7 and 10^9 but by graphing these values with a normal (arithmetic) scale, students will find that there are 10x as many in the latter. In a similar way, an earthquake measuring 6 on the Richter scale may not seem to be all that different from one measuring 8 on the Richter scale, but when one starts to understand the powers of 10 that are involved and realizes that the latter quake is 100x more powerful than the first, then one is beginning to use mathematics to make sense of real-world phenomena.

  In math class, students learn how to use powers of 10 to write numbers in scientific notation; you might use this skill to help them understand the importance of this more fully. Also, students know that multiplying by 1 does not change a value; however, they often do not know why this is so important. Here are a few areas where you might engage students with these ideas:

  - **Evolutionary time:** How big is a billion? How many generations might this entail?
  - **Human population on earth:** How big is 7,083,051,866? Is there a simpler way to express this? If everyone lived in a city with 100,000 people, how many cities would there be? What is the population density, on average, and what is the range of these densities in various countries throughout the world? (Similar questions can be asked for various species.)
  - **Converting units in the metric system:** A micron or micrometer is a millionth of a meter. How else can this be expressed? What part of a kilometer is a micron expressed in powers of 10? Conversion by unit analysis, in which students multiply by appropriate forms of 1, is far more instructive than pneumonic gimmicks. For example, because the numerators and denominators are of equal value, 1000 m/1 km = 1 km/1000 m = 1. We can use this idea to convert 2.5 m to km. Rather than trying to remember whether to multiply or divide by 1000, if we multiply 2.5 m by 1 km/1000 m, a useful form of 1, the distance stays the same but is now expressed in km instead of m. This idea of multiplying by 1 is essential both in mathematics and in science.

- **Exponents & Logarithms**
  Exponents and logarithms, which are related to powers of 10, also have applications in biology. Like addition and subtraction, exponents and logs are inverses of each other. In fact, a good way to think of a log is simply as an exponent. Consider the statement that 10^3 = 1000. An equivalent statement is that the log_{10}1000 = 3 (read “log, base 10, of 1000, equals 3”). Note that the log is 3, and 3 is the exponent in the first statement. Because we are working in a base 10 system, we usually omit stating that we are in base 10, stating simply that \( \log 1000 = 3 \) or that \( \log 0.001 = -3 \). The log is the power that 10 must be raised to in order to obtain the number. Base 10 logs are often referred to as common logs. In other situations, we use base e and the symbol ln, for natural log. In other cases, the base must be explicitly stated.

  pH is a measure of acidity or basicity of a solution, particularly in living systems. The pH represents the opposite of the log of the hydrogen ion concentration. That is, \( pH = -\log[H^+] \). pH values typically range from 1 to 14. A pH of 1 means that the solution has \( 10^{-1} \) (0.1) moles/L of hydrogen ions and is a very strong acid. A pH of 14 has \( 10^{-14} \) ions of hydrogen and is a very strong base. Water has equal amounts of acid and base and has a pH of 7. Battery acid has a pH of 0, which means it contains an entire mole \((6.02 \times 10^{23} \text{ ions of hydrogen})\) per liter of water. Vinegar has a pH of 3, milk 6.6, blood 7.4, and household NaOH 13.7. pH can be estimated with a solution such as phenol, with pH paper, or more accurately with a pH meter. More details about pH and its derivation can be found at [http://en.wikipedia.org/wiki/PH](http://en.wikipedia.org/wiki/PH).

  Sound level (or measure of sound pressure as a force) is often measured in decibels, a log scale. Because of the wide range of human hearing, it would be more difficult to interpret an arithmetic scale. An increase of 10 decibels suggests a 10-fold increase in level, an increase of 20 decibels suggests a 100-fold increase, and so forth. The unit “decibel” was coined in recognition of communications pioneer Alexander Graham Bell. More information about the physics of the decibel and decibel readings can be found at [http://en.wikipedia.org/wiki/Decibel](http://en.wikipedia.org/wiki/Decibel).

- **Selecting the Best Data Display**
  In mathematics, students study various displays, including, but not limited to, pie charts, bar graphs, histograms, box-and-whisker plots, trend lines, and scatter plots. Each of these has strengths and weaknesses, though it is important for students to recognize that sometimes certain graphs are inappropriate. The first step is recognizing how many variables are under consideration and what types of variables are being measured.

  Pie charts and bar graphs are appropriate for problems dealing with single variables of a categorical nature. Pie charts are excellent for viewing parts of the whole, while bar graphs show relative sizes of the frequencies within each of the categories.

  Histograms and box-and-whisker plots also concern single variable problems, but the variables are numerical in nature. Like bar graphs, histograms are useful for seeing relative sizes within the data. In histograms, we typically set up ranges of values and place each data point in one of the ranges. For example, if we were considering distances, we could set up ranges of 0–10 km, 10–20 km, 20–30 km, … Box-and-whisker plots are excellent for viewing the range of values and key measures of the data, specifically the minimum and maximum, the median, and the first and third quartiles.

  Trend lines concern bivariate data, with the independent variable representing time. Although variations exist, the dependent variable is usually numeric in nature, such as the number of individuals in a defined population. By viewing a trend line,
we can see what has happened to the population numbers over time. A scatterplot is also used for bivariate, numerical data and can be used to depict the relationship between the variables. For example, students could create a scatterplot that shows the relationship between the mass of various quantities of a substance and the volume of those quantities in an exploration leading to an understanding of density. Once we have the scatterplot, we typically seek some type of mathematical model that describes the relationship.

**Proportions**

Proportions can serve as a powerful tool for estimating populations, particularly of animals in the wild. A strategy of mark-recapture connects biology to mathematics. From a certain population, a given number are captured and marked in some fashion and then released back into their environment. After giving the animals sufficient time to mix back into the habitat, several of the animals are recaptured. The ratio of the marked to the number recaptured provides the information needed to estimate the entire population:

\[
\text{# of total tagged : total # in population = # of tagged in recaptured sample : total # recaptured}
\]

To improve on the accuracy of this technique, results can be combined from several trials. Statistics teachers may want to construct confidence intervals based on multiple trials to estimate the population.

**Squaring a Binomial: Hardy-Weinberg Equilibrium**

Algebra students often square binomials. For example, \((x + y)^2 = x^2 + 2xy + y^2\). If we know that the value of \(x + y = 1\), then we also know that \((x + y)^2 = 1\). Hardy and Weinberg used this idea to generate some valuable understandings of traits that humans and other species have. If we have a simple trait determined by two alleles, a dominant one (indicated by a capital letter) and a recessive one (indicated by a lower case letter) and if \(x\%\) of the alleles are dominant, then \((100 - x)\%\) of the alleles must be recessive, because combined they must total 100%. Combining these mathematical ideas can help students better understand population genetics.

For a simple trait with alleles \(P\) and \(p\), all the alleles in a population can be represented as \(P + p = 1\). If you square \(P + p\), obtaining \(P^2 + 2Pp + p^2\), we must also get 1, because 1 squared equals 1. You can use the result to estimate the proportions of homozygous dominant individuals (PP), heterozygous individuals (Pp), and homozygous recessive individuals (pp) in the population. For example, suppose you find that 36% of your population of interest is homozygous recessive (pp). This means that \(p^2 = 0.36\). Taking the square root, we find that \(p = 0.6\), which suggests that \(P = 0.4\). Thus, \(P^2 = 0.16\), which suggests that 16% of the population is homozygous dominant (PP). And \(2Pp = 2*0.4*0.6 = 0.48\), which suggests that 48% of the population is heterozygous. Thus, with some simple math, you can determine the likely proportions of all phenotypes and genotypes in the population.

**Using a t-Test to Determine Whether Two Populations Are Really Different**

Often, we need to determine whether two populations differ significantly. To do so, we can take samples from the different populations and run a hypothesis test to see if the samples are different enough to allow us to conclude that the populations must be different.

Though there is always error when we generalize from samples to populations, we can quantify this error. We determine ahead of time how much error we are willing to tolerate by establishing an \(\alpha\) (alpha) value, setting the value based on the context and the consequence of being wrong. We consider an event with probability less than \(\alpha\) implausible, and an event with probability of \(\alpha\) or higher plausible. Often \(\alpha\) is set at 0.05, in which case we are willing to accept a 5% chance of error of deciding that two populations are different when, in fact, they are not. This is called a Type I error.

Suppose we wish to determine whether the mean growth from watering a certain plant species daily is greater than the mean growth from watering the species weekly. Assuming we found that this is true from our sample, can we conclude that daily watering produces greater growth?

The \(t\)-distribution is appropriate for comparing the two means. To begin, we establish two hypotheses. Let \(\mu_1\) represent the mean growth with daily watering and \(\mu_2\) represent the mean growth with weekly watering. We hope to establish that \(\mu_1 > \mu_2\);

This is called our research hypothesis (\(H_r\)).

Our goal is to show that the closest possible case, \(\mu_1 = \mu_2\), which we call the null hypothesis (\(H_o\)), is implausible. We run our test assuming that \(H_o\) is true. If this leads to results that are not plausible (\(>\alpha\)), then we reject \(H_o\), which indicates that our research hypothesis must be true. If we are unable to show that the null hypothesis is implausible, we have been unsuccessful in our efforts.

Our samples allow us to generate sample means and a standard error. A \(t\)-statistic, which is a measure of how many standard errors away from a difference of zero our results are, is then calculated. A large \(t\) value suggests that the means are relatively far apart, while a small \(t\) value suggests that the means are not all that different. Each \(t\) value determines a \(p\) value, which represents the probability that our null hypothesis is true, given our sample data. A low \(p\) value (\(<\alpha\)) allows us to reject \(H_o\) and accept our research hypothesis. A higher \(p\) value tells us we have not proved our research hypothesis.

**Other Areas for Connecting Mathematics to Biology**

Quantifying many aspects of biology can enrich the learning in both biology and mathematics. We have barely scratched the surface in this article. Here are just a few additional mathematical areas that can be useful in biology:

- Determining which measures of central tendency to use
- Statistics/logic: linking cause and effect between variables
- Using effect size to determine magnitude of impact
- Estimation: Do the figures make sense?
Significant figures: How many numbers need to be reported in a measurement?

We welcome hearing from those of you who have integrated mathematics into biology. By sharing your ideas, you will enrich all of us.

References


ROBERT M. HORTON is Professor of Mathematics Education at Clemson University, 400 Tillman Hall, Clemson, SC 29634; e-mail: bhorton@clemson.edu. WILLIAM H. LEONARD is Emeritus Professor of Science at Clemson University and the ABT Editor; e-mail: leonard@clemson.edu.