12
Modeling

12.1 What is a model?

A model is a representation of reality that allows us to understand something better. There are many types of models, including *conceptual*, *mathematical*, and *physical* models. A physical model is a physical object or set of objects intended to represent something else that is too large, small, complex or otherwise inaccessible for direct investigation. A conceptual model is a collection of hypothesized relationships between different objects or variables, and is usually described in narrative. From an early age, we learn how to construct both physical and conceptual models. Children create conceptual models to help them understand cause and effect relationships that lead to either desirable or unwanted outcomes ('if I jump down one or two steps, it's fun, but if I jump down three or more steps it hurts my legs: jumping farther hurts more'). When my gradeschool son builds a spaceship from Legos, he is creating a physical model of a spaceship he has seen in a movie or book. These are not particularly sophisticated models, but they are nevertheless ways of representing some aspect of reality (or imagined reality).

As with Legos, mathematical models can serve mostly a desire for creative play. Like Lego models, it is perfectly possible to create a mathematical model that represents reality poorly, and is therefore not very useful. Perhaps we claim to have created a model of a car, but if we've only stacked rectangular bricks together and failed to add wheels, it is not a particularly good or useful model of a car. Thus, model construction and use should be done with the broader problem context in mind. The means should justify the desired ends.

In this book, we are interested in mathematical and conceptual models and the connections between them. Ultimately, our goal isn't necessarily to become mathematical modelers, but rather to be able to construct, use, and understand models that can assist with problem-solving. Indeed, many mathematical models originate from

Heuristic: Mathematical models are only as useful as the conceptual models on which they are based.

a desire to quantify the relationships in a conceptual model devised to address a problem. Several possible approaches to quantification lead to a handful of varieties of mathematical models. We'll focus our discussion on three distinct but related types of mathematical models that differ in their origins and implementation. The first two are grounded in theory, while the third often arises from statistical data analysis.

- **Analytical models** are usually developed from theory based on fundamental physical, chemical or biological principles. A hypothesis that a tree's height should scale with it's trunk diameter raised to the 2/3 power in order to retain structural integrity is such a model. These models are often the most general and abstract, and can sometimes be solved with paper and pencil. However, they can become hopelessly complex and un-solvable when one tries to incorporate realistic details and context. The idealizations necessary to make an analytical model solvable can also sometimes limit its utility.
- **Numerical models** may be created and motivated in the same manner as analytical models, but employ techniques for mathematical *approximation* that permit relaxation of analytical idealizations and introduction of detail without making the equations too difficult to solve. Numerical models can be solved by hand for very small systems, but are more appropriately implemented in computer programs.
- **Empirical models** may have analytical or numerical components, but contain parameters that must be quantified by experiment or systematic observation. Data must be incorporated and usually analyzed statistically in order to define parameter values. In some cases, regression is used to constrain the functional relationships between variables or to identify the value of coefficients. Thus, a fully empirical model is data-driven or data-calibrated.

We have already seen or worked with a few examples of models. The Logistic population growth model that we discussed briefly in Section 10.1.3 is a theoretical model that can be implemented either in numerical or analytical form. Even that model, however, has empirical components, since it's use in practical problem-solving requires some observational constraints on r and K. When we solved for total brook trout population in Section 11.2.1, we employed an empirical model known as the Leslie method, which is based on a conceptual model of the change in catch probability under declining population.

12.1.1 Example: The Universal Soil Loss Equation (USLE)

The widely-used Universal Soil Loss Equation (USLE) is an example
of an empirical model. The master equation for USLE is:

$$A = RKLSCP \tag{12.1}$$

where A is the soil loss (usually in tons/acre/year), R is a rainfall-
erosivity factor, K is a soil erodibility factor, L and S are the slope
length and angle factors, C is a ground-cover factor and P is a param-
eter that accounts for soil conservation practices or structures.

    The factors in USLE are quantities whose values cannot be mea-
sured directly. Instead, the numerical values are each derived from
a combination of carefully-designed field experiments where all but
one factor is held constant. The factor values are then derived from
measured differences in soil loss.

    The great value of the USLE and it's kin is that it is sufficiently
easy to use that farmers with little formal training in math or com-
puting can easily get satisfactory results. Most factor values can
either be looked up in tables or measured on the ground or from
maps.

    The ease of use comes at a cost, however. Because factor values
are derived from experiments, they are strictly valid only within
the range of conditions considered within the experiments. In other
words, if applied in settings where – for example – rainfall intensity
is twice as large as the largest observed in experiments, the reliability
of results is uncertain. Fully empirical models can therefore some-
times be unreliable in conditions outside the range of the conditions
under which factor values were determined.

12.1.2 Example: probability of deer-automobile encounters (Problem 3.3)

As we have already seen, simple theoretical models can sometimes
be sufficient to explore a range of system behaviors, even when func-
tional relationships are uncertain. These models will inevitable by
limited in power by the simplifying assumptions or idealizations
used, but when the science or management problem permit a solu-
tion with substantial uncertainty, this approach is still warranted.

    Let's assume that deer in our county are randomly distributed in
space, and that they have no particular reason to either avoid or seek
out roads. Call the total area of the county A_c and the proportion of
the area occupied by roads f, so that the area of roads $A_r = fA_c$.
Let's assume that there are N_0 deer in the county. It follows that – if
the deer are randomly distributed – there will be approximately fN_0
deer on the road at any moment. What is that number according to
the numbers we produced earlier for Story County, IA? The value of

f was estimated to be approximately 0.0076, so if there are say 1000 deer in the county, we should expect either 7 or 8 of them on the road at any given time. That seems reasonable, but that isn't what we're after. We'd like to know about how likely collisions are between deer and automobiles. So we need to work in something about the number and distance of car trips through the road system, right? This is left as an exercise for the student, as there are many possible ways to approach this.

12.2   *Dealing with higher mathematics*

Many powerful mathematical models have been devised to explore and describe phenomena in nature. Some of the most powerful are those that allow predictions of unobserved or future events or patterns. These can directly inform management decisions provided that managers trust and understand their results. Unfortunately, many of these powerful models employ mathematical concepts and methods that are beyond the typical undergraduate training in math. Does that mean that most people are doomed to never understand or use these models? Absolutely not! There isn't any inherent reason that students need to take calculus, linear algebra, or differential equations courses before they can comprehend the gist of a model constructed with those skills. It certainly helps to have at least a conceptual grasp of some key concepts in calculus, but that doesn't translate to a pre-requisite.

12.2.1   *Example: prairie dog plague (Problem 3.4)*

Since this problem deals with hypothetical future events, it may not be possible to glean the answer directly from past work or from observation. Instead, we can construct a simple model of the prairie dog community with random, probabilistic interactions among well-mixed individuals.

A common way to model disease transmission is with a **compartment model** often called SIR. We consider individuals in a population to be in one of three (or four) states: Susceptible (S), Infected (I), and Removed (R) or Recovered. Individuals move from compartment S to compartment I by disease transmission. Infected individuals in compartment I then either recover and move to compartment R, or are removed from the population by death or isolation. These transfers between compartments are often described with a system of

differential equations:

$$\frac{dS}{dt} = -\beta SI \tag{12.2}$$

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{12.3}$$

$$\frac{dR}{dt} = \gamma I \tag{12.4}$$

These differential equations are not easily solved in most cases, but we can use them as a basis for a *numerical simulation* of disease dynamics if we are able to estimate the parameters β and γ. A numerical representation of the first equation might look something like this, for example:

$$S_{t+1} = S_t - \beta S_t I_t \tag{12.5}$$

$$I_{t+1} = I_t + \beta S_t I_t - \gamma I_t \tag{12.6}$$

$$R_{t+1} = R_t + \gamma I_t \tag{12.7}$$

This says that in a given time increment, susceptible individuals are moved from the S compartment to the I (infected) compartment at a rate that is proportional to the product of the numbers of individuals in each compartment and the transmission rate constant β. You can see in the first and second equations above that when a number of idividuals infected according to the βSI term in lost from the S compartment (because it is negative), it is gained (positive) in the I compartment. All individuals are accounted for in moving into or out of the I compartment. Similarly, individuals move from the I compartment to the R compartment at a rate governed by the rate constant γ. Selection of these rate constants to a large extent governs the behavior of the model, and thus the predicted fate of the prairie dog colony. But implementing management options informed by positive model outputs is where the biggest challenge arises.

12.3   Power-Law Scaling

CONSIDER THIS SEEMINGLY INNOCUOUS QUESTION: are larger animals heavier than smaller animals?

> YOU: Hmmm, well, yeah I think so?! An adult bear weights more than a snowshoe hare, for instance.

OK, great, but how would we know if this is true more generally? And what exactly do we mean by larger? Does that mean taller? Larger volume? This brings up a few issues that become important when we're talking about real quantities rather than abstract variables. Unambiguously defining quantities can be an important first

step in communicating quantitative information. In the next section we'll be specific about what information is required to fully define a quantity. For now let's agree that we're satisfied with relating the mass of an animal to its volume. Do animals that take up more space (i.e., have greater volume) also weigh more? Maybe we can say it another way: is the weight or mass of an animal proportional to its body size? We could write this in symbols:

$$M \propto V? \tag{12.8}$$

The symbol '$\propto$' between M (body mass) and V (volume) means "proportional to". So this isn't an equation yet because we're not sure anything is equal. And of course it's nonsense that an animal's weight is equal to its volume. There must be some other parameter that transforms an animal's volume into a mass. Let's call it c, and try it out in an equation:

$$M = cV \tag{12.9}$$

But what is c? As we said above, we'd prefer to have some meaning for the symbols we throw around in equations. Let's use one of our old algebraic tools for manipulating equations and "solve the equation for c". By that we mean get c onto one side of the equation all by itself. To get there, we just need to divide both sides of the equation by V, yielding:

$$\frac{M}{V} = c \tag{12.10}$$

Now recall that the definition of density is mass per unit volume. That's exactly what we have on the left-hand side of the equation! So our equation now says that c, the parameter we used to transform volume into mass, is the same as density! So for an individual animal, the parameter that relates mass to volume is density. As we have done previously, let's assume that most animals have a density close to that of water so this proportionality parameter c doesn't vary significantly among species. So to the extent that it is correct to say that most animal's body density is close to that of water, we can argue that larger animals do indeed weigh more, in general.

This is probably not a very profound revelation to you[1]. But with only a few more small leaps in logic, we can get somewhere considerably more interesting. For more than a century, biologists have been intrigued by a remarkable relationship between the basal metabolic rate and body mass for animals of a wide range of sizes and shapes. Amazingly, if one assembles a large set of data and plots it on a graph with a logarithmic scale, mice, humans and elephants and most of the rest fall along a straight line! An equation that describes
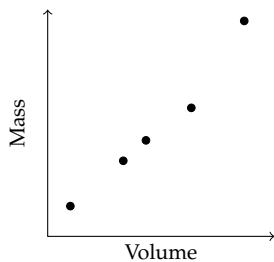

Figure 12.1: Plot of some hypothetical measurements of animal mass and volume.

[1] An entertaining and well-composed article on some not-so-obvious consequences of size differences in animals is *On Being the Right Size*, byt J.B.S. Haldane, published in *Harper's Magazine*, March 1926.

this relationship and the line on the graph looks like this:

$$B = B_0 M^b \qquad\qquad (12.11)$$

where B is the basal metabolic rate, M is body mass as before, and B_0 and b are constants (we'll see what they mean later!). This equation is yet another power law, and equations with this form pop up surprisingly often in ecology once you start looking. We'll get more into functions and power laws later on. But for now, some important points should be made:

- The argument that there should be a proportionality between body mass and metabolic rate was originally conceived theoretically on the basis that energy given off by an animal to its surroundings might depend mostly on the animal's surface area, while its mass scales with volume.

- Measurements by many researchers over more than a century have been compared against this theoretical prediction, with varying degrees of success. In most cases however, the power-law relationship holds.

- By comparing theoretical predictions with real data, one can discover truly novel and interesting things about physiological similarities or differences between different organisms – insights we might not have ever developed without the quantitative analyses.

  We'll look into this in more detail a bit later.