

Part I

Introduction to Modeling

Chapter 1

Mathematical Modeling

1.1 The Modeling Cycle

A *mathematical model* is an equation or system of equations that describes a real-life system. For example, the *algebraic model*

$$v(t) = gt \tag{1.1}$$

is an equation that describes the velocity of an object dropped from rest near the surface of a planet.

Models often describe *dynamical systems*, that is, systems that change in time. Since continuous-time rates of change are derivatives, models often take the form of *differential equations* (equations that involve derivatives). For example, the algebraic model (1.1) can be written as the differential equation model

$$\frac{ds}{dt} = gt, \tag{1.2}$$

where $s(t)$ is the distance traveled by the object. The differential equation model (1.2) is a *continuous-time model*.

Some real-life processes are better described by *discrete-time models*. For example, if each plant in a population of annuals produces b seeds per year, p of which survive, then the number of plants at the next year is

$$N_{t+1} = pbN_t, \tag{1.3}$$

where N_t is the number of plants in the population at year t . Model (1.3) is called a *difference equation*, a *discrete-time map*, or a *recursion formula*.

Mathematical models are powerful tools for addressing scientific problems. This is because a mathematical model is the translation of a real-life problem into the precise language of mathematics, where it becomes amenable to many powerful techniques. Fig. 1.1 shows a conceptual diagram for the process of mathematical modeling. Important issues arise in each of the four modeling steps.

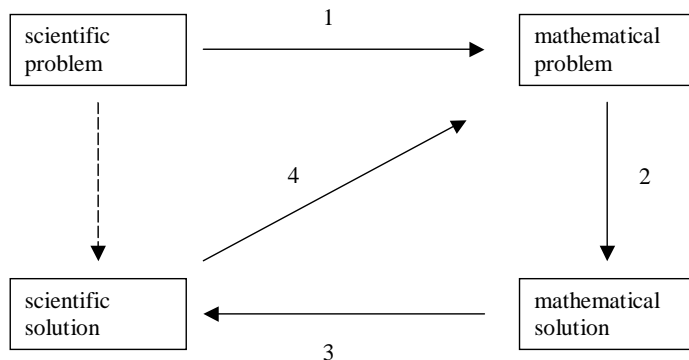


Figure 1.1: The Modeling Cycle

1.1.1 Step 1: Translation into mathematics

The first step in modeling is to translate the scientific problem into the precise language of mathematics. This process requires a crisp conceptual clarification that often proves valuable in and of itself.

Choosing variables and parameters We must decide what quantitative aspects of the system we wish to follow, and name those quantities with variables. Input variables are called *independent*. Time, location in space, and age are examples of common independent variables. *Dependent variables* are the functions of interest that depend on independent variables. In modeling, dependent variables are usually called *state variables*, because they record the state of the system. For example, suppose we want to study how a population is distributed in time and space. We could take $N(t, x)$ to be the number of animals at time t at location x . Here t and x are the independent variables, and N is the state variable. In equation (1.2), the state variable is s and the independent variable is t . We know this because the presence of the derivative ds/dt alerts us to the fact that $s(t)$ is the function under consideration.

Besides the input variables, the system in question may depend on other *extrinsic* (environmental) factors that are relatively constant on the time scale of the problem. Typical examples are temperature and humidity. We might include these factors in the equation as constants. Such constants are called *parameters*. (Some people call them “coefficients”, but this is not the best terminology, because the constants do not always appear in the equation as true coefficients; for example, they may appear as terms or as exponents). In equation (1.1), t is the independent variable, v is the state variable, and g is an extrinsic parameter. Earth has one value of g , while Mars has another. Other parameters may be *intrinsic*. In equation (1.3), N is the state variable, t is the independent variable, and b is an intrinsic parameter. The parameter p might be a compendium of both extrinsic and intrinsic factors.

Simplifying assumptions Once the variables have been identified, they must be linked together with equations. Constructing model equations always involves making simplifying assumptions about the real system: the object is close to the surface of the planet; the force of gravity at the surface of the planet is constant; gravity is the only force acting on the object (no friction); etc. Real systems probably cannot be described *exactly* by equations, and so mathematical models are always approximations to the real situation. We can add more and more mathematical complications (for example, the force due to friction of the falling object) to a model in an attempt to make it more realistic, but at some point the model becomes unwieldy and intractable. There is no advantage in replacing a real system that is too complicated to understand with a mathematical systems that is too complicated to understand. An important modeling mantra is therefore *KISS* (Keep it simple, stupid!). On the other hand, if the assumptions oversimplify the problem, the mathematical model will not be able to approximate the real system. Thus, in modeling there is always a trade-off between realism and tractability. The goal of modeling is to capture the essential behavior of a complicated real system with a simplified mathematical system. The “art” of modeling is to determine the main mechanisms driving the system, building the assumptions around those, and ignoring lesser influences.

Parameterization If model equations are to describe a real system, they must be connected to real data through *parameterization*. Parameterization is the statistical process by which data are used to estimate the values of the parameters. For example, in equation (1.1) the value of g must be estimated from experimental data on falling objects. You may have done this experiment in a general physics lab.

The parameterized model equation(s) may now be taken, tentatively, as a surrogate for the real system. In fact, the model constitutes a hypothesis about the scientific problem. More than one model can be constructed, based on different hypotheses about what drives the system. These competing models serve as *alternative hypotheses* which are then tested in Step 3.

1.1.2 Step 2: Model analysis

The second step in modeling is that of mathematical analysis. All the powerful concepts and tools of mathematics may now be brought to bear on the scientific problem through its surrogate, the model. Analysis may require solving differential equations, drawing bifurcation diagrams, linearizing, or utilizing any of a host of other techniques we will learn in this book.

1.1.3 Step 3: Back-translation

The third step is to translate the mathematical results back into the framework of the original scientific problem. Before drawing scientific conclusions, however, we must ask whether the results correspond to extant data, and whether the model can accurately predict new data.

Model selection and validation Model selection is the process by which alternative parameterized models are compared to each other in an effort to select the best model. In this book we will discuss several statistical tools that are available for this task. Model validation is the process of evaluating a model by comparing its output to (new) data. If there is a pronounced lack of correspondence with data, then the modeling assumptions need to be revised. Model validation should be a procedure distinct from model fitting. Ideally, models should be validated on independent data sets, that is, on data sets that were not used to estimate the model parameters. Validation involves computing so-called *goodness-of-fit* statistics that measure how closely model output corresponds to data.

Test of model predictions A good model not only describes and explains, but also predicts. The best way to further test a validated model is to make *a priori* predictions, and then collect data to test those predictions. Ideally, the predictions should be unusual or unexpected. A successful test of such predictions makes a strong case that the model really does capture the essential behavior of the system.

1.1.4 Step 4: Revising model assumptions

The fourth step is ubiquitous in modeling. A model that doesn't correspond to data must be revised. One or more of the assumptions were false or were oversimplifications; important factors may have been left out of the system. In practice, model revision occurs more or less continually throughout the modeling process as new insights are gained and equations are tweaked.

You can see that modeling in biology requires a thorough integration of biology, mathematics, and statistics. The next three sections contain some important general ideas and terms from each of these disciplines.

1.2 Biology

Biology is the study of living systems. Mathematical models can describe living systems at subcellular, cellular, tissue, organ, organ system, organism, population, community, and ecosystem levels of organization. The level of organization chosen depends on the interest of the investigator and the question asked. For example, if an epidemiologist (biologist who studies the spread of disease) wants to know how fast a disease will spread through a region, the modeling process would focus on the population inhabiting the region. Modeling subcellular processes might be little help! In other words, before modeling begins, one must select an appropriate *scale* of interest.

Living systems are the most complex systems known. In order to model a living system at any scale, the system should be well observed. In Chapter @ we explain how we developed a model to accurately predict the number of gulls occupying a particular habitat at a particular time. Our efforts were successful because we and others spent a lot of time learning about the everyday

lives of these birds. Armed with this knowledge, we could ask the right questions, identify the independent variables most likely to influence gull behavior, and develop an efficient data collection protocol for model parameterization and testing. There is no substitute for an intimate observational knowledge of a living system whose mechanisms you want to understand better through modeling.

Some biologists assume that living systems are too complicated to be modeled mathematically. It may be that at some scales every living system is too complicated to model. It might be infeasible, for example, to model how each of the millions of neurons in your cerebral cortex are working in concert to process information conveyed by the words in this paragraph. A neurobiologist, however, might be able to model the activity of two or three interacting neurons, or entire sections of the brain. In short, mathematical modeling may not be able to answer all the interesting questions a biologist might ask.

Evolutionary biologists and population geneticists have long taken advantage of mathematical modeling. This is because people like G. H. Hardy, Sir Ronald Fisher, Sewell Wright, and J. B. S. Haldane, trained in mathematics, turned their attention to solving biological problems. Biologists in other subdisciplines have also discovered the power of modeling.

At the most basic level, mathematical modeling provides biologists a way to quantitatively *describe* the behavior of living systems. At a higher level, modeling provides a tool to *identify* factors that drive living systems. Finally, once these factors have been identified, biologists can *test predictions* about how these systems will function in the future.

Our interest and experience concern *ecological modeling*. Because many of the examples used in this book come from ecology, we now provide a brief description of this subdiscipline.

1.2.1 Ecology

Ecologists are concerned with interactions among organisms and their environments. Both organisms and environments are very complex; moreover, they constantly undergo change. So ecologists face significant challenges as they try to uncover ecological patterns.

An interesting challenge that faces all ecologists is the problem of *scale*. Patterns apparent at one scale may disappear at higher or lower levels of organization; still other patterns may emerge at different levels. Patterns at some scales exhibit *emergent properties*. This is another way of saying the whole often seems more than the sum of the parts—qualities emerge at higher levels that might not be predicted on the basis of the known properties of lower levels.

Ecologists view ecosystems as consisting of progressively more inclusive scales:

Individual organism—Individuals live more or less autonomous lives. Members of *solitary species* interact little with conspecifics, whereas members of *social species* interact frequently with one another. Moreover, individuals are

difficult to distinguish. The assemblage of medusae and polyps in a Portugese man-o-war, the collection of individuals forming a sphere of *Volvox*, and a grove of aspen trees generated through vegetative reproduction come to mind.

Population—Members of a single species that inhabit a particular geographic region.

Species—A group of interbreeding or potentially interbreeding organisms genetically isolated from other such groups. (Note: This is only one of many competing species concepts. Species concepts generate lots of debate among biologists.)

Community—An assemblage of populations that occupy the same geographic region.

Ecosystem—A community plus all the non-living components of the system. These non-living (abiotic) components include things like energy, moisture, and nutrients.

Physiological ecologists (or *environmental physiologists*) deal with how individual organisms cope with environmental change. For example, they have investigated how camels conserve water in the heat of the desert, and how seals and porpoises conserve oxygen during long underwater dives.

Population ecologists are interested in the factors that influence population growth.

Community ecologists are concerned with interactions between and among species that live in the same community. These interactions include food web relations, predator-prey interactions, interspecific competition, mutualism, parasitism, speciation, and coevolution, among others.

Ecosystem ecologists examine how energy flows through and how nutrients cycle within ecosystems. Thus, ecosystem ecologists work with energy budgets, accounting schemes by which the precise nature of community interactions can be teased apart, and determine pathways taken by nutrients through ecosystems.

Behavioral ecologists study the causation, function, ontogeny, and evolution of the behaviors of animals in their ecological setting.

Considerable overlap occurs among the areas of ecology. For example, a population ecologist needs to understand the life history characteristics of individual organisms in the population she studies, as well as how community interactions such as competition and predator-prey relations impact this population.

Ecologists, like other biologists, are often tempted to confuse *correlation* with *causation*. The fact that we can mathematically model a relationship between particular independent and dependent variables is no proof that the two variables are causally linked. For example, we might develop a model based on temperature change that quite accurately predicts when white-crowned sparrows migrate south for the winter. Yet we know from experimental studies that photoperiod, not temperature, is the environmental factor that causes white-crowned sparrows to migrate. Because temperature covaries with the light-dark cycle, we see how easy it would be to arrive at an invalid conclusion concerning

the cause of migration. Thus, mathematical modeling can sometimes lead us to incorrect conclusions if we are not careful. On the other hand, mathematical modeling can often suggest possible causal relationships that we might not have guessed otherwise.

Another point worth making in this context is the difference between *proximate* and *ultimate* causes. In the example just given, the proximate cause of migration is photoperiod change, which alters the neuroendocrine system of white-crowned sparrows. The ultimate cause in this example was natural selection which, over generations, favored the fitness of birds that nested where food was abundant during the spring and moved to warmer regions when nesting areas turned frigid.

1.3 Mathematics

The most basic mathematical concept in modeling is that of the *function*. A relationship between an input t and an output $x(t)$ is called a function if and only if the value of the output $x(t)$ is completely determined by the input t . Symbolically, $x(t)$ is a function if and only if

$$t_1 = t_2 \implies x(t_1) = x(t_2),$$

where the arrow means “implies”. In words, if the inputs are the same, then the outputs must be the same. Said another way, you can’t get two different outputs from the same input. Functions are called *deterministic*. (A relation in which you can get two different outputs from the same input is called *stochastic*; we will turn our attention to stochasticity in the next section.)

In modeling dynamical systems, we are interested in the state of the system (call it x) as a function of time t . That is, we are interested in the behavior of $x(t)$ as t changes. We can list corresponding values of x and t in a table, and we can also graph x vs. t . This list of pairs of numbers (or its graph) is called a *time series*. If there are two state variables, say $x(t)$ and $y(t)$, we must show two time series graphs together: x vs. t and y vs. t . In general, the two state variables may depend on each other, so the two time series graphs must be interpreted together. This is a visually difficult task. A better visual tool in this case is that of *state space*, or *phase space*, in which the state variables are graphed against each other in the $x - y$ plane. The current state of the system is represented by a point $(x(t), y(t))$ that moves around in the plane as time progresses, tracing out an *orbit*. Calculus provides us with many powerful tools for studying systems whose state changes more or less continuously throughout time. The main tool, of course, is the derivative dx/dt , which is the instantaneous rate of change of $x(t)$ with respect to time t .

A dynamical system is at *equilibrium* if it does not change over time. The time series graph for such a system is a horizontal line. In state space, a system at equilibrium is represented by an orbit that consists of a single stationary point. Mathematically speaking, equilibria are constant solutions of

the model. Consider, for example, the logistic population model, which we will derive in a later chapter:

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right).$$

Here $N(t)$ is the population size at time t , and $r, K > 0$ are constant parameters. A solution $N(t)$ of the logistic model is a constant solution if and only if it satisfies

$$\frac{dN}{dt} = 0$$

for all time t . Thus, the equilibria are the (constant) roots of the *equilibrium equation*

$$0 = rN_e \left(1 - \frac{N_e}{K} \right),$$

that is,

$$N_e = 0, K.$$

For a discrete time example of equilibria, consider the plant population model (1.3). A solution of this difference equation is a constant solution if and only if it satisfies

$$N_{t+1} = N_t = N_e$$

for all time t . Thus, the equilibria are the (constant) roots of the equilibrium equation (or *fixed point equation*)

$$N_e = pbN_e.$$

If $b, p > 0$, the only such constant N_e is the extinction equilibrium $N_e = 0$.

Another important notion is that of *stability*. We will give precise definitions of stability later in the book. For now, consider a planar pendulum. There are two equilibrium states: angle zero (straight down) with zero velocity, and angle π (straight up) with zero velocity. If the pendulum is at rest in the “down” position and is perturbed slightly away from equilibrium, it will remain close to equilibrium (in fact, the system asymptotically returns to equilibrium in this case). The “down” equilibrium is called *stable*, because if the system starts close to it, the system stays close to it. If, on the other hand, the pendulum is at rest in the “up” position and is perturbed, the pendulum will move away from the “up” equilibrium. This equilibrium is called *unstable*.

Consider now a large hemispherical bowl, sitting upright on a table, with a tennis ball inside it. If the ball is placed on the bottom of the bowl with zero velocity, it will remain there. This state is therefore an equilibrium (“if it starts there, it stays there”). If the tennis ball is released, perhaps with a small velocity, near the bottom of the bowl, it will remain in the vicinity of the bottom of the bowl with small velocity, and so the equilibrium state is stable (“if it starts close, it stays close”). But in fact, the tennis ball actually approaches the bottom of the bowl and its velocity approaches zero, so the equilibrium

state is said to be *asymptotically stable*. Now suppose the bowl is turned upside down. The top point of the bowl gives rise to an unstable equilibrium for the ball, because it is not true that “if you start close, you stay close”. Now remove the bowl and set the tennis ball down on the table top with zero velocity. If the table is flat, the ball remains where you placed it. Therefore, it is at equilibrium. If you move the ball a small distance away, still on the table top with perhaps a small velocity, it will probably not return to the old position, but it will not stray far. Thus, it is true that “if you start close, you stay close” to the original position; hence the equilibrium is stable. However, since the ball does not return to the original state, the stable equilibrium is not asymptotically stable. We say it is *neutrally stable*. In fact, every point on the table top gives rise to a neutrally stable equilibrium.

Finally, we list a few phrases in English that you should always be able to translate directly into mathematics. We say *y is proportional to x* if and only if we can write

$$y = ax$$

for some constant a . We say *y is inversely proportional to x* if and only if we can write

$$y = \frac{a}{x}$$

for some constant a . *Population growth rate* is the change in population size with respect to time. In continuous-time, this is the derivative dN/dt . Thus, if a population’s growth rate is proportional to the square root of the population size N , and inversely proportional to the temperature T , we write

$$\frac{dN}{dt} = a \frac{\sqrt{N}}{T}.$$

Per capita growth rate is defined as

$$\frac{1}{N} \frac{dN}{dt}.$$

1.4 Statistics

Deterministic models are approximations to real systems; a good model captures the *signal* (main deterministic trend) in the data. Nonetheless, the data will likely deviate somewhat from the model prediction. This deviation from the signal is called *noise* or *stochasticity*. The two main types of noise in biological data are *process error* and *measurement error* (*observational error*). Process error occurs because the real system is more complicated than the mathematical model. Measurement error occurs because the real system cannot be measured exactly. Stochasticity in ecological data can be handled with statistical methods, many of which will be addressed in this book.

There are two main types of process error in ecology: *environmental stochasticity*, and *demographic stochasticity*. Stochastic events in a population

can be likened to the toss of a fair coin. Imagine that a single coin is tossed for a population of animals. The outcome of the toss, although random, is the same for each individual member of the population. This is environmental stochasticity. Such extrinsic events as weather cause this type of noise. Now imagine that each animal in the population tosses its own coin. This time there is a random outcome for each individual. This is demographic stochasticity. Individual variability in intrinsic parameters such as birth and death rates cause this type of noise.

Systems in classical physics have relatively little stochasticity, and their mathematical models are so precise that some people call them “laws”. Some social science systems, on the other hand, have a lot of stochasticity—so much so that the signal may be swamped out by noise and mathematical modeling may be impossible. In ecology, deterministic and stochastic forces are more or less equally important. Therefore, noise should—ideally—be incorporated explicitly into a deterministic model to produce a stochastic version of the model. As we explore the behavior of stochastic models, we shall see that the interaction of deterministic and stochastic forces can give rise to a rich class of emergent dynamic phenomena that cannot occur in purely deterministic or purely random systems.

Stochasticity is modeled mathematically through the notions of *random variable* and *distribution*. Suppose you count the number of seabirds on a beach. If the number of birds is large, repeated counts of the exact same group of birds will probably yield different results due to observational error. We can use the notion of a *random variable* X to stand for the outcomes of trial observations (counts). Each particular observation $X = x$ is a *realization* of the random variable. You would want to know if some observations x are more likely than others, because you would want to know, for example, whether the count errors are biased. Mathematically, you would be asking, “what is the distribution of the random variable X ”? The answer to this question is, of course, situation-dependent. One has to make assumptions about how measurements, and hence errors, are distributed. Ideally, such assumptions can be tested experimentally.

The main distribution used in this book is the *normal distribution*. We will continue our discussion by supposing the bird count random variable X is normally distributed about the true number of birds μ . The normal distribution *probability density function* (PDF) is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (1.4)$$

(Fig. 1.2) Here $f(x)$ is *not* the probability of observing x birds. Rather, the probability that the observational count X will be between the values a and b is the area under the normal curve that lies between $x = a$ and $x = b$; that is,

$$P[a \leq X \leq b] = \int_a^b f(x) dx.$$

The total area under the normal curve gives the probability that the count lies between $-\infty$ and $+\infty$, which is, of course, equal to one. Therefore, it must be

true that

$$P[-\infty < X < +\infty] = \int_{-\infty}^{\infty} f(x) dx = 1 \quad (1.5)$$

(exercise 10).

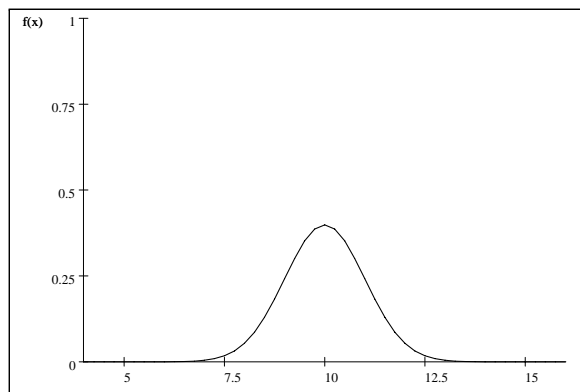


Figure 1.2: The Normal Distribution

The *expected value* of the random variable X is the sum of all possible outcomes of X , weighted by the probability of obtaining that outcome. For the normal distribution, $E[X] = \mu$; that is

$$E[X] = \int_{-\infty}^{\infty} xf(x) dx = \mu \quad (1.6)$$

(exercise 11). The inflection points of the curve f occur at $x = \mu \pm \sigma$ (exercise 12). The parameter σ is called the *standard deviation*. It measures the spread of the distribution about the mean μ . The square of the standard deviation, σ^2 , is called the *variance* of the distribution.

You may have noticed that we have assumed X to be a *continuous random variable*, while in the bird-counting example it is actually a *discrete random variable*. That is, you can count only an integer number of birds, while the normal distribution allows x to be any real number. We could have used a bell-shaped histogram for the distribution of a discrete random variable, but we often simplify the problem with continuous approximations.

1.5 Epistemology: How We Know

There are two main kinds of inference: *deduction* and *induction*. Deduction infers a particular conclusion from a general statement:

All Andrews University students have cars.
 Matthew is an Andrews University student.
 (conclusion) Therefore, Matthew has a car.

Notice that if the first two statements are true, then we all agree that the conclusion is guaranteed to be true; deductive arguments are *conclusive*. However, if one (or both) of the first two statements is false, then the conclusion is not guaranteed to be true; it might be true or it might be false.

Induction infers a general conclusion from a set of particular statements or observations:

All the Andrews students I have observed have cars.
 (conclusion) Therefore, all Andrews students have cars.

Notice that despite your observations, the conclusion might still be false, depending on the sample of data you observed. The only way you can be completely sure of your conclusion is if your induction was *exhaustive*, that is, if you observed *all* Andrews students having cars. Inductive arguments are not in general conclusive unless you are able to observe all possible instances of the data.

Mathematics is pure deduction. Make no mistake: the activity of doing mathematics is not purely deductive. Trial and error, hunches, and experimentation lead mathematicians to pose the conjectures that they then try to prove deductively. The final result, however, must be purely deductive, or it is not mathematics. Inductive arguments are not permitted as mathematical proof. (In your mathematics classes you may have learned a special kind of proof called “Mathematical Induction” (MI). MI is actually a kind of deduction. See exercise 14.) Mathematics is the only discipline in which arguments are conclusive. In one sense deduction cannot create “new” knowledge, since the conclusions must be implicit in the statements with which you begin. In another sense, deduction *can* create new knowledge, because it is a very powerful tool that teases information out of more general statements—information that you might never have guessed was latent in the original statements. The “problem”, of course, is that the general statements themselves must have come from even more general statements, and so on. You can see that in pure deduction, there must be Original Statements which must be taken without proof! Indeed, the edifice of mathematics rests on unproven *axioms*. Now, pure mathematicians do not ask whether or not those axioms are true with respect to the real world. They only want to know that they are logically consistent. The axiomatic method is a fascinating and tremendously powerful tool.

In general, scientists want to know whether or not their starting statements correspond to observation in the real world. Unlike pure mathematicians, they use deduction to develop correspondences between logical inferences and observations of nature. Science, therefore, utilizes both deduction and induction. Science inductively infers hypotheses from data, draws deductive conclusions from these hypotheses, and then tests the conclusions against more data

(Fig. 1.3). The inclusion of observation and induction in scientific methodology makes science more powerful than pure mathematics as a tool for understanding real systems. But there is a price to pay for this increase in power, and that is a decrease in certainty. In science, there is no proof in the mathematical sense of the word.

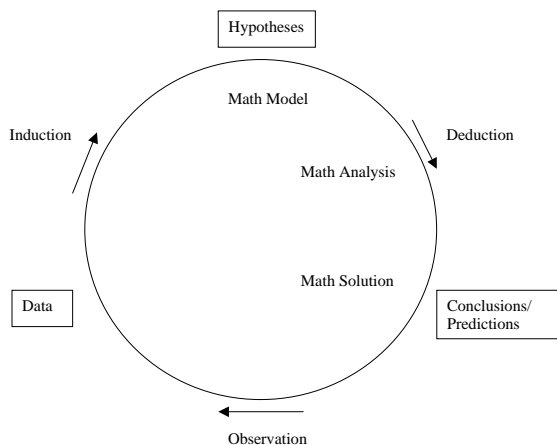


Figure 1.3: Scientific Methodology

The most powerful science is not purely observational, but also uses deduction. Observation guides theory, and theory guides observation. Note that the so-called *hard sciences* (physics is the prime example) employ mathematics in the deductive step.

When we compare the *epistemologies* (ways of knowing) of various disciplines, we can see that they are nested (Fig. 1.4).

We also see that the more rigorous the method of inference, the more certain one can be, but the less one can learn about nature (Fig. 1.5). The least rigorous methods of “inference” include extra-rational (not *irrational!*) ways of knowing. These epistemologies are the most “powerful”, in the sense that they can address the greatest range of questions about reality. But they are also the least “compelling”, in the sense of being able to rationally convince someone else. (See exercise 23.)

In conclusion, systems of thought – mathematics, science, philosophy, theology, etc. – have different epistemologies. One cannot do science (using induction) and call it mathematics, even though it may be perfectly good science. One cannot do theology (using revelation) and call it science, even though it may be perfectly good theology. Nevertheless, all of these disciplines have a common goal: understanding reality. They are not so much about *certainty* as

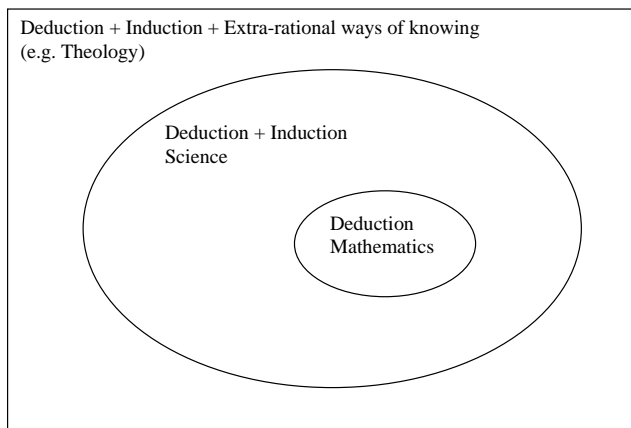


Figure 1.4: Nested Epistemologies

they are about *understanding*. They *model* reality. In particular, mathematical modeling is not about absolute certainty. It is about understanding. In the words of Alfred J. Lotka (1925), an early pioneer in mathematical ecology, “An empirical formula is therefore not so much the solution of a problem as the challenge to such a solution. It is a point of interrogation, an animated question mark.”

We hope you will find this book to be an “animated question mark”.

1.6 Exercises

1. Consider the population model

$$\frac{dx}{dt} = (b - d)x$$

Identify the independent variable(s), the state variable(s), and the parameter(s).

2. Consider the LPA model of flour beetle (*Tribolium*) dynamics

$$\begin{aligned} L(t+1) &= bA(t)e^{-c_{el}L(t)-c_{ea}A(t)} \\ P(t+1) &= (1 - \mu_l)L(t) \\ A(t+1) &= P(t)e^{-c_{pa}A(t)} + (1 - \mu_a)A(t) \end{aligned}$$

Identify the independent variable(s), the state variable(s), and the parameter(s).

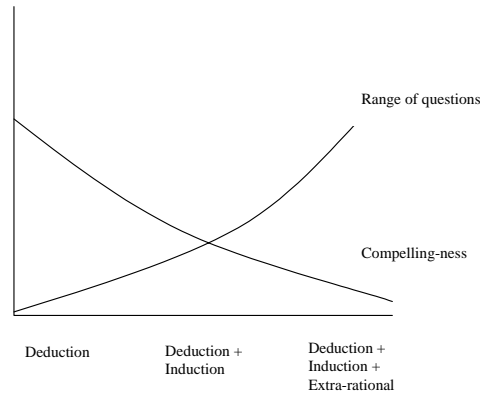


Figure 1.5: Range of Questions and Compellingness

3. Consider the logistic population model

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right)$$

Identify the independent variable(s), the state variable(s), and the parameter(s).

4. Consider the dynamical system

$$\frac{dx}{dt} = 2x(1-x)(2x-4)^2$$

Find all the equilibria.

5. Consider the discrete-time dynamical system known as the Ricker model:

$$x_{t+1} = bx_t e^{-cx_t}$$

Here $b, c > 0$. This famous model was first used in fisheries. Find all the equilibria.

6. Consider the model

$$\begin{aligned} \frac{dx}{dt} &= x + xy \\ \frac{dy}{dt} &= x + y \end{aligned}$$

Find all the equilibria. (Hint: the equilibria are constant solution pairs (x_e, y_e) .)

7. Malthusian, or exponential, growth is modeled by

$$\frac{dx}{dt} = rx$$

Is the extinction equilibrium $x_e = 0$ asymptotically stable, neutrally stable, or unstable? (Hint: consider the cases $r < 0$, $r = 0$, and $r > 0$ separately.)

8. The per capita growth rate of a certain population is proportional to both the population size and the temperature. It is inversely proportional to the humidity. Translate this statement into mathematics.
9. Identify the following random variables as environmental or demographic: wind speed, number of offspring, humidity, temperature, death due to predatory encounter.
10. Verify equation (1.5). Hint: Show that

$$\begin{aligned} \left(\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \right)^2 &= \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy \\ &= \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du \int_{-\infty}^{\infty} e^{-\frac{1}{2}v^2} dv \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}[u^2+v^2]} dudv \end{aligned}$$

where $u = (x - \mu) / \sigma$ and $v = (y - \mu) / \sigma$. Now change from rectangular to polar coordinates and carry out the integration.

11. Verify equation (1.6) for $\mu = 0$. That is, verify that

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2\sigma^2}} dx = 0.$$

12. Prove that the inflection points of the normal curve occur at $\mu \pm \sigma$.
13. A very large group of fixed objects is to be counted. There are so many objects that repeated counts give different tallies. Suppose that a large number of repeated counts are recorded and found to be normally distributed. Would you expect σ to be independent of group size? Explain.
14. Mathematical Induction (MI) is a particular kind of mathematical argument used to prove theorems of the form, "For all natural numbers $n = 1, 2, 3, \dots$, the statement $Q(n)$ is true." MI is like climbing an infinitely tall ladder, where each rung is a natural number. Provided one can always move to the next rung from any other rung, and provided one can

get on the first rung to begin with, one can ascend the whole ladder. MI uses a three-step method. First, one proves that the statement is true for $n = 1$, that is, $Q(1)$. Second, one assumes the statement is true for an arbitrary number $n = k$, that is, $Q(k)$. Third, one shows that the statement is therefore true for $n = k + 1$, that is, that $Q(k + 1)$. Here is an example: We will use MI to prove the statement “The sum of the first n natural numbers is $1 + 2 + 3 + \cdots + n = n(n + 1)/2$. Proof: (i) Clearly this is true for $n = 1$. (ii) Assume it is true for $n = k$, that is, assume $1 + 2 + 3 + \cdots + k = k(k + 1)/2$. (iii) Then $1 + 2 + 3 + \cdots + k + (k + 1) = k(k + 1)/2 + (k + 1) = (k + 1)(k/2 + 1) = (k + 1)([k + 1] + 1)/2$. QED

- (a) Use MI to prove that the sum of the first n squares is $n(n + 1)(2n + 1)/6$.
 - (b) Is Mathematical *Induction* a good name for MI? Explain.
15. Deduction is conclusive. Induction generally is not. Therefore, scientific reasoning (induction + deduction) is not conclusive; yet it is still a compelling and powerful type of argument. The following questions are highly nontrivial. Consideration of the questions in exercises 16-19 may help.
 - (a) Why do we believe good scientific reasoning is compelling?
 - (b) Why do we believe deduction is conclusive?
 - (c) Why is the word “believe” used in (a) and (b)?
 16. Can one use deduction to conclusively prove that deduction is conclusive?
 17. Suppose you observe $x = 4,269$ birds on the beach, while the model predicted $x = 0$. You conclude from this contradiction that there is something wrong with the model. Why did your observation take precedence over the model as the standard for reality? Can you prove that your observation accurately depicts physical reality? Explain.
 18. Criticize or support the following statements: Every system of thought must begin with unprovable presuppositions. However, a good system of thought minimizes its presuppositions.
 19. Criticize or support the following statements: Every system of thought begins with unprovable presuppositions; thus no system of thought can yield absolute certainty about “reality”. Therefore, all systems of thought are equally compelling.
 20. Is there really a difference between *extra-rational* and *irrational*? Explain.
 21. Consider Fig. 1.4. Can science logically contradict mathematics? Can philosophy or theology logically contradict science (or mathematics)? (Note that it can be very difficult to prove that an apparent contradiction is really a logical contradiction. Many apparent contradictions are actually *paradoxes* that can be logically resolved.)

22. Consider Fig. 1.4. Can the epistemologies of the outer levels ever be utilized legitimately at the inner levels? For example, can a mathematician use induction when doing mathematics? Can a mathematician or scientist use biblical revelation when doing mathematics or science? Explain your answer carefully.
23. Can one hold extra-rational knowledge with internal certainty even though the information cannot be transmitted to anyone else using rational inference? Explain. Give examples.