general). Gotelli (2001) is more detailed. Begon et al. (1996) gives an extremely thorough introduction to general ecology, including some basic ecological models. Case (1999) provides an illustrated treatment of theory, while Roughgarden (1997) integrates ecological theory with programming examples in MATLAB. Mangel (2006) and Otto and Day (2007), two new books, both give basic introductions to the "theoretical biologist's toolbox".

1.2.2 Other kinds of models

Ecologists sometimes want to "learn how to model" without knowing clearly what questions they hope the models will answer, and without knowing what kind of models might be useful. This is a bit like saying "I want to learn to do experiments", or "I want to learn molecular biology": do you want to analyze microsatellites? Use RNA inactivation to knock out gene function? Sequence genomes? What people usually mean by "I want to learn how to model" is "I have heard that modeling is a powerful tool and I think it could tell me something about my system, but I'm not really sure what it can do".

Ecological modeling has many facets. This book covers only one: statistical modeling, with a bias towards mechanistic descriptions of ecological patterns. The next section briefly reviews a much broader range of modeling frameworks, and gives some starting points in the modeling literature in case you want to learn more about other kinds of ecological models.

1.3 FRAMEWORKS FOR MODELING

This book is primarily about how to combine models with data and how to use them to discover the answers to theoretical or applied questions. To help fit statistical models into the larger picture, Table 1.1 presents a broad range of dichotomies that cover some of the kinds and uses of ecological models. The discussion of these dichotomies starts to draw in some of the statistical, mathematical and ecological concepts I suggested you should know. However, if a few are unfamiliar, don't worry — the next few chapters will review the most important concepts. Part of the challenge of learning the material in this book is a chicken-and-egg problem: in order to know why certain technical details are important, you need to know the big picture, but the big picture itself involves knowing some of those technical details. Iterating, or cycling, is the best way to handle this problem. Most of the material introduced in this chapter will be covered in more detail in later chapters. If you don't completely get it this time around, hang on and see

Scope and approach	
abstract	concrete
strategic	tactical
general	specific
theoretical	applied
qualitative	quantitative
descriptive	predictive
mathematical	statistical
mechanistic	phenomenological
pattern	process
Technical details	
analytical	computational
dynamic	static
continuous	discrete
population-based	individual-based
Eulerian	Lagrangian
deterministic	stochastic
Sophistication	
simple	complex
crude	sophisticated

Table 1.1 Modeling dichotomies. Each column contrasts a different qualitative style of modeling. The loose association of descriptors in each column gets looser as you work downwards.

if it makes more sense the second time.

1.3.1 Scope and approach

The first set of dichotomies in the table subdivides models into two categories, one (theoretical/strategic) that aims for general insight into the workings of ecological processes and one (applied/tactical) that aims to describe and predict how a particular system functions, often with the goal of forecasting or managing its behavior. Theoretical models are often mathematically difficult and ecologically oversimplified, which is the price of generality. Paradoxically, although theoretical models are defined in terms of precise numbers of individuals, because of their simplicity they are usually only used for qualitative predictions. Applied models are often mathematically simpler (although they can require complex computer code), but tend to capture more of the ecological complexity and quirkiness needed to make

detailed predictions about a particular place and time. Because of this complexity their predictions are often less general.

The dichotomy of mathematical *vs.* statistical modeling says more about the culture of modeling and how different disciplines go about thinking about models than about how we should actually model ecological systems. A mathematician is more likely to produce a deterministic, dynamic process model without thinking very much about noise and uncertainty (e.g. the ordinary differential equations that make up the Lotka-Volterra predator prey model). A statistician, on the other hand, is more likely to produce a stochastic but static model, that treats noise and uncertainty carefully but focuses more on static patterns than on the dynamic processes that produce them (e.g. linear regression)*.

The important difference between phenomenological (pattern) and mechanistic (process) models will be with us throughout the book. Phenomenological models concentrate on observed patterns in the data, using functions and distributions that are the right shape and/or sufficiently flexible to match them; mechanistic models are more concerned with the underlying processes, using functions and distributions based on theoretical expectations. As usual, there are shades of gray; the same function could be classified as either phenomenological or mechanistic depending on why it was chosen. For example, you could use the function $f(x) = ax/(b + x)$ (a Holling type II functional response) as a mechanistic model in a predator-prey context because you expected predators to attack prey at a constant rate and be constrained by handling time, or as a phenomenological model of population growth simply because you wanted a function that started at zero, was initially linear, and leveled off as it approached an asymptote (see Chapter 3). All other things being equal, mechanistic models are more powerful since they tell you about the underlying processes driving patterns. They are more likely to work correctly when extrapolating beyond the observed conditions. Finally, by making more assumptions, they allow you to extract more information from your data — with the risk of making the *wrong* assumptions.†

Examples of theoretical models include the Lotka-Volterra or Nicholson-Bailey predator-prey equations (Hastings, 1997); classical metapopulation models for single (Hanski, 1999) and multiple (Levins and Culver, 1971; Tilman, 1994) species; simple food web models (May, 1973; Cohen et al., 1990); and theoretical ecosystem models (Ågren and Bosatta, 1996). Ap-

*Of course, both mathematicians and statisticians are capable of more sophisticated models than the simple examples given here.

†For an alternative, classic approach to the tradeoffs between different kinds of models, see Levins (1966) (criticized by Orzack and Sober (1993); Levins's (1993) defense invokes the fluidity of model-building in ecology).

plied models include forestry and biogeochemical cycling models (Blanco et al., 2005), fisheries stock-recruitment models (Quinn and Deriso, 1999), and population viability analysis (Morris and Doak, 2002; Miller and Lacy, 2005).

Further reading: books on ecological modeling overlap with those on ecological theory listed on p. 6. Other good sources include Nisbet and Gurney (1982) (a well-written but challenging classic) Gurney and Nisbet (1998) (a lighter version) Haefner (1996) (broader, including physiological and ecosystem perspectives) Renshaw (1991) (good coverage of stochastic models), Wilson (2000) (simulation modeling in C), and Ellner and Guckenheimer (2006) (dynamics of biological systems in general).

1.3.2 Technical details

Another set of dichotomies characterizes models according to the methods used to analyze them or according to the decisions they embody about how to represent individuals, time, and space.

An analytical model is made up of equations solved with algebra and calculus. A computational model consists of a computer program which you run for a range of parameter values to see how it behaves.

Most mathematical models and a few statistical models are dynamic; the response variables at a particular time (the state of the system) feed back to affect the response variables in the future. Integrating dynamical and statistical models is challenging (see Chapter 11). Most statistical models are static; the relationship between predictor and response variables is fixed.

One can specify how models represent the passage of time or the structure of space (both can be continuous or discrete); whether they track continuous population densities (or biomass or carbon densities) or discrete individuals; whether they consider individuals within a species to be equivalent or divide them by age, size, genotype, or past experience; and whether they track the properties of individuals (individual-based or Eulerian) or the number of individuals within different categories (population-based or Lagrangian).

Deterministic models represent only the average, expected behavior of a system in the absence of random variation, while stochastic models incorporate noise or randomness in some way. A purely deterministic model allows only for qualitative comparisons with real systems; since the model will never match the data *exactly*, how can you tell if it matches closely

enough? For example, a deterministic food-web model might predict that introducing pike to a lake would cause a trophic cascade, decreasing the density of phytoplankton (because pike prey on sunfish, which eat zooplankton, which in turn consume phytoplankton); it might even predict the expected magnitude of the change. In order to test this prediction with real data, however, you would need some kind of statistical model to estimate the magnitude of the average change in several lakes (and the uncertainty), and to distinguish between observed changes due to pike introduction and those due to other causes (measurement error, seasonal variation, weather, nutrient dynamics, population cycles . . .).

Most ecological models incorporate stochasticity crudely, by simply assuming that there is some kind of (perhaps normally distributed) variation, arising from a combination of unknown factors, and estimating the magnitude of that variation from the variation observed in the field. We will go beyond this approach, specifying different sources of variability and something about their expected distributions. More sophisticated models of variability enjoy some of the advantages of mechanistic models: models that make explicit assumptions about the underlying causes of variability can both provide more information about the ecological processes at work and can get more out of your data.

There are essentially three kinds of random variability:

- *Measurement error* is the variability imposed by our imperfect observation of the world; it is always present, except perhaps when we are counting a small number of easily detected organisms. It is usually modeled by the standard approach of adding normally distributed variability around a mean value.
- *Demographic stochasticity* is the innate variability in outcomes due to random processes even among otherwise identical units. In experimental trials where you flip a coin 20 times you might get 10 heads, or 9, or 11, even though you're flipping the same coin the same way each time. Likewise, the number of tadpoles out of an initial cohort of 20 eaten by predators in a set amount of time will vary between experiments. Even if we controlled everything about the environment and genotype of the predators and prey, we would still see different numbers dying in each run of the experiment.
- *Environmental stochasticity* is variability imposed from "outside" the ecological system, such as climatic, seasonal, or topographic variation. We usually reserve environmental stochasticity for unpredictable variability, as opposed to predictable changes (such as seasonal or latitudinal changes in temperature) which we can incorporate into our

models in a deterministic way.

The latter two categories, demographic and environmental stochasticity, make up *process variability** which unlike measurement error affects the future dynamics of the ecological system. Suppose we expect to find three individuals on an isolated island. If we make a measurement error and measure zero instead of three, we may go back at some time in the future and still find them. If an unexpected predator eats all three individuals (process variability), and no immigrants arrive, any future observations will find no individuals. The conceptual distinction between process and measurement error is most important in dynamic models, where the process error has a chance to feed back on the dynamics.

The distinctions between stochastic and deterministic effects, and between demographic and environmental variability, are really a matter of definition. Until you get down to the quantum level, any "random" variability can in principle be explained and predicted. What determines whether a tossed coin will land heads-up? Its starting orientation and the number of times it turns in the air, which depends on how hard you toss it (Keller, 1986). What determines exactly which and how many seedlings of a cohort die? The amount of energy with which their mother provisions the seeds, their individual light and nutrient environments, and encounters with pathogens and herbivores. Variation that drives mortality in seedlings — e.g. variation in available carbohydrates among individuals because of small-scale variation in light availability — might be treated as a random variable by a forester at the same time that it is treated as a deterministic function of light availability by a physiological ecologist measuring the same plants. Climatic variation is random to an ecologist (at least on short time scales) but might be deterministic, although chaotically unpredictable, to a meteorologist. Similarly, the distinction between demographic variation, internal to the system, and environmental variation, external to the system, varies according to the focus of a study. Is the variation in the number of trees that die every year an internal property of the variability in the population or does it depend on an external climatic variable that is modeled as random noise?

1.3.3 Sophistication

I want to make one final distinction, between simple and complex models and between crude and sophisticated ones. One could quantify simplicity vs. complexity by the length of the description of the analysis, or the number

*Process variability is also called *process noise* or *process error* (Chapter 10).

of lines of computer script or code required to implement a model. Crudity and sophistication are harder to recognize; they represent the conceptual depth, or the amount of *hidden* complexity, involved in a model or statistical approach. For example, a computer model that picks random numbers to determine when individuals give birth and die and keeps track of the total population size, for particular values of the birth and death rates and starting population size, is simple and crude. Even simpler, but far more sophisticated, is the mathematical theory of random walks (Okubo, 1980) which describes the same system but — at the cost of challenging mathematics — predicts its behavior for *any* birth and death rates and any starting population sizes. A statistical model that searches at random for the line that minimizes the sum of squared deviations of the data is crude and simple; the theory of linear models, which involves more mathematics, does the same thing in a more powerful and general way. Computer programs, too, can be either crude or sophisticated. One can pick numbers from a binomial distribution by virtually flipping the right number of coins and seeing how many come up heads, or by using numerical methods that arrive at the same result far more efficiently. A simple R command like `rbinom`, which picks random binomial deviates, hides a lot of complexity.

The value of sophistication is generality, simplicity, and power; its costs are opacity and conceptual and mathematical difficulty. In this book, I will take advantage of many of R's sophisticated tools for optimization and random number generation (since in this context it's more important to have these tools available than to learn the details of how they work), but I will avoid many of its sophisticated statistical tools, so that you can learn from the ground up how statistical models really work and make your models work the way you want them to rather than being constrained by existing frameworks. Having reinvented the wheel, however, we'll briefly revisit some standard statistical frameworks like generalized linear models and see how they can solve some problems more efficiently.

1.4 FRAMEWORKS FOR STATISTICAL INFERENCE

This section will explore three different ways of drawing statistical conclusions from data — frequentist, Bayesian, and likelihood-based. While the differences among these frameworks are sometimes controversial, most modern statisticians know them all and use whatever tools they need to get the job done; this book will teach you the details of those tools, and the distinctions among them.

To illustrate the ideas I'll draw on a seed predation data set from Duncan and Duncan (2000) that quantifies how many times seeds of two different species disappeared (presumably taken by seed predators, although we can't

be sure) from observation stations in Kibale National Park, Uganda. The two species (actually the smallest- and largest-seeded species of a set of eight species) are *Polyscias fulva* (`pol`: seed mass < 0.01 g) and *Pseudospondias microcarpa* (`psd`: seed mass ≈ 50 g).

1.4.1 Classical frequentist

Classical statistics, which are part of the broader *frequentist* paradigm, are the kind of statistics typically presented in introductory statistics classes. For a specific experimental procedure (such as drawing cards or flipping coins), you calculate the probability of a particular outcome, which is defined as *the long-run average frequency of that outcome in a sequence of repeated experiments*. Next you calculate a *p-value*, defined as the probability of that outcome *or any more extreme outcome* given a specified null hypothesis. If this so-called *tail probability* is small, then you reject the null hypothesis: otherwise, you fail to reject it. But you don't accept the alternative hypothesis if the tail probability is large, you just fail to reject the null hypothesis.

The frequentist approach to statistics (due to Fisher, Neyman and Pearson) is useful and very widely used, but it has some serious drawbacks — which are repeatedly pointed out by proponents of other statistical frameworks (Berger and Berry, 1988). It relies on the probability of a series of outcomes that didn't happen (the tail probabilities), and which depend on the way the experiment is defined; its definition of probability depends on a series of hypothetical repeated experiments that are often impossible in any practical sense; and it tempts us to construct straw-man null hypotheses and make convoluted arguments about why we have failed to reject them. Probably the most criticized aspect of frequentist statistics is their reliance on p-values, which when misused (as frequently occurs) are poor tools for scientific inference. It seems to be human nature to abuse p-values, acting as though alternative hypotheses (which are usually what we're really interested in) are "true" if we can reject the null hypothesis with $p < 0.05$ and "false" if we can't. In fact, when the null hypothesis is true we still find $p \leq 0.05$ one time in twenty (we falsely reject the null hypothesis 5% of the time, by definition). If $p > 0.05$ the null hypothesis could still be false but we have insufficient data to reject it. We could also reject the null hypothesis, in cases where we have lots of data, even though the results are biologically insignificant — that is, if the estimated effect size is ecologically irrelevant (e.g. a 0.01% increase in plant growth rate with a 30°C increase in temperature). More fundamentally, if we use a so-called *point null hypothesis* (such as "the slope of the relationship between plant productivity and temperature is zero"), common sense tells us that the null

hypothesis *must* be false, because it can't be exactly zero — which makes the p value into a statement about whether we have enough data to detect a non-zero slope, rather than about whether the slope is actually different from zero. Working statisticians will tell you that it is better to focus on estimating the values of biologically meaningful parameters and finding their confidence limits rather than worrying too much about whether p is greater or less than 0.05 (Yoccoz, 1991; Johnson, 1999; Osenberg et al., 2002) — although Stephens et al. (2005) remind us that hypothesis testing can still be useful.

Looking at the seed data, we have the following 2×2 table:

	pol	psd
any taken (t)	26	25
none taken	184	706
total (N)	210	731

If t_i is the number of times that species i seeds disappear and N_i is the total number of observations of species i then the observed proportions of the time that seeds disappeared for each species are (pol) $t_1/N_1 = 0.124$ and (psd) $t_2/N_2 = 0.034$. The overall proportion taken (which is not the average of the two proportions since there are different total numbers of observations for each species) is $(t_1 + t_2)/(N_1 + N_2) = 0.054$. The ratio of the predation probabilities (proportion for pol/proportion for psd) is $0.124/0.034 = 3.62$. The ecological question we want to answer is "is there differential predation on the seeds on these two species?" (Given the sample sizes and the size of the observed difference, what do you think? Do you think the answer is likely to be statistically significant? How about biologically significant? What assumptions or preconceptions does your answer depend on?)

A frequentist would translate this biological question into statistics as "what is the probability that I would observe a result this extreme, or more extreme, given the sampling procedure?" More specifically, "what proportion of possible outcomes would result in observed ratios of proportions greater than 3.62, *or* smaller than $1/3.62 = 0.276$?" (Figure 1.1). Fisher's exact test (`fisher.test` in R) calculates this probability, as a one-tailed test (proportion of outcomes with ratios greater than 3.62) or a two-tailed test (proportion with ratios greater than 3.62 or less than its reciprocal, 0.276); the two-tailed answer in this case is 5.26×10^{-6}. According to Fisher's original interpretation, this number represents the strength of evidence against the null hypothesis, or (loosely speaking) for the alternative hypothesis — that there is a difference in seed predation rates. According to the Neyman-Pearson decision rule, if we had set our acceptance cutoff at $\alpha = 0.05$, we