CHAPTER 3

Deriving Classic Models in Ecology and Evolutionary Biology

Chapter Goals:
- To introduce some classic models in ecology, evolution, and epidemiology
- To derive dynamical equations for analysis in subsequent chapters

Chapter Concepts:
- Exponential growth
- Logistic growth
- Linear and nonlinear models
- Population-genetic models
- Symmetric models
- Normalization
- Competition models
- Consumer-resource models
- SIR epidemiological models

3.1 Introduction

In this chapter, we describe a variety of models whose behavior we will explore in subsequent chapters. The models we have chosen are classics in ecology, evolution, and epidemiology. These classic models incorporate many of the elements commonly encountered when developing new models, and they illustrate many of the dynamical patterns that you are likely to encounter, from fairly predictable to weird and wonderful.

In section 3.2 we introduce the exponential and logistic models describing population growth. In section 3.3, the haploid and diploid single-gene models of evolution by natural selection are introduced. Section 3.4 then introduces the Lotka-Volterra competition and predator-prey models, as well as consumer-resource models. Section 3.5 introduces SIR epidemiological models for the spread of infectious diseases. We will describe the basic assumptions underlying all of these models and derive their dynamical equations using the steps described in Chapter 2 (see Table 3.1 for quick reference). Finally, in section 3.6 we discuss how one can work backward, starting with published equations for a model and deducing its underlying assumptions.

3.2 Exponential and Logistic Models of Population Growth

In any species, the number of individuals changes over time in response to resource availability, competition, predation, disease, weather, and chance events. The simplest models describing changes in population size are exponential and logistic growth. Both assume that the environment is constant, and both ignore any interactions with other species (no competing species, predators, parasites, etc.). The two models of growth differ in what they assume about the availability of resources (e.g., food, water, nesting sites, etc.). The exponential growth model assumes that the amount of resources available to each individual is constant, regardless of the population size, whereas the logistic growth model assumes that fewer resources are available to each individual as the population size increases.

3.2.1 Exponential Population Growth

The discrete-time exponential growth model assumes that each reproducing parent is replaced by a constant number of individuals, R, in the next time unit. Technically, this assumes that all individuals in the population are capable of reproduction (as in a hermaphroditic or asexual species). The model can also be applied to species with separate male and

female sexes, however, by assuming that the number of offspring is limited by the number of females and then counting females only.

TABLE 3.1
Models derived in this chapter. Type refers to (1) recursion equation in discrete time, (2) difference equation in discrete time, and (3) differential equation in continuous time.

Model	Type	Equation	
Exponential growth	1	$n(t+1) = R\,n(t)$	(3.1b)
	2	$\Delta n = (R-1)\,n(t)$	(3.2)
	3	$\frac{dn}{dt} = r\,n(t)$	(3.3)
Logistic growth	1	$n(t+1) = n(t) + r\,n(t)\left(1 - \frac{n(t)}{K}\right)$	(3.5a)
	2	$\Delta n = r\,n(t)\left(1 - \frac{n(t)}{K}\right)$	(3.5b)
	3	$\frac{dn}{dt} = r\,n(t)\left(1 - \frac{n(t)}{K}\right)$	(3.5c)
Haploid selection	1	$p(t+1) = \frac{W_A\,p(t)}{W_A\,p(t) + W_a\,q(t)}$	(3.8c)
	2	$\Delta p = \frac{(W_A - W_a)\,p(t)\,q(t)}{W_A\,p(t) + W_a\,q(t)}$	(3.9)
	3	$\frac{dp}{dt} = s\,p(t)\,q(t)$	(3.11b)
Diploid selection	1	$p(t+1) = p(t)^2\frac{W_{AA}}{W} + p(t)\,q(t)\frac{W_{Aa}}{W}$	(3.13a)

Competition equations	1	$n_1(t+1) = n_1(t) + r_1 n_1(t)\left(1 - \dfrac{n_1(t) + \alpha_{12} n_2(t)}{K_1}\right)$ $n_2(t+1) = n_2(t) + r_2 n_2(t)\left(1 - \dfrac{n_2(t) + \alpha_{21} n_1(t)}{K_2}\right)$	(3.14)
	3	$\dfrac{dn_1}{dt} = r_1 n_1(t)\left(1 - \dfrac{n_1(t) + \alpha_{12} n_2(t)}{K_1}\right)$ $\dfrac{dn_2}{dt} = r_2 n_2(t)\left(1 - \dfrac{n_2(t) + \alpha_{21} n_1(t)}{K_2}\right)$	(3.15)
Consumer-resource equations	3	$\dfrac{dn_1}{dt} = f(n_1) - g(n_1, n_2)$ $\dfrac{dn_2}{dt} = \varepsilon\, g(n_1, n_2) - h(n_2)$	(3.16)
SIR equations	3	$\dfrac{dS}{dt} = b - d\,S(t) - a\,c\,S(t)\,I(t) + \sigma R(t)$ $\dfrac{dI}{dt} = a\,c\,S(t)\,I(t) - \delta\, I(t) - \rho\, I(t)$ $\dfrac{dR}{dt} = \rho\, I(t) - \sigma R(t) - d\,R(t)$	(3.19)

To derive a general discrete-time exponential model, we allow births and deaths to affect the number of individuals in the population, $n(t)$. With two processes, we must specify an order to these events, and we shall assume that births are followed by deaths. Using b to denote the per capita number of births and d to denote the fraction of the population that dies, the life cycle and flow diagrams are then given by Figure 3.1.

This simple exponential model does not track the age of individuals (see Chapter 10 for models that do). In other words, individuals are treated equally regardless of whether they are one or several time units old. Consequently, the same exponential growth model can describe populations in which all parents die (nonoverlapping generations) or only some of them die (overlapping generations)—in either case, d measures the total fraction of the population that dies after the round of births.

We can now derive the recursion equation by applying Recipe 2.1 to Figure 3.1:

$$n'(t) = n(t) + b\,n(t) = (1 + b)\,n(t) \qquad \text{(after births)},$$

$$n''(t) = n'(t) - d\,n'(t) = (1 - d)\,n'(t) \qquad \text{(after deaths)}.$$

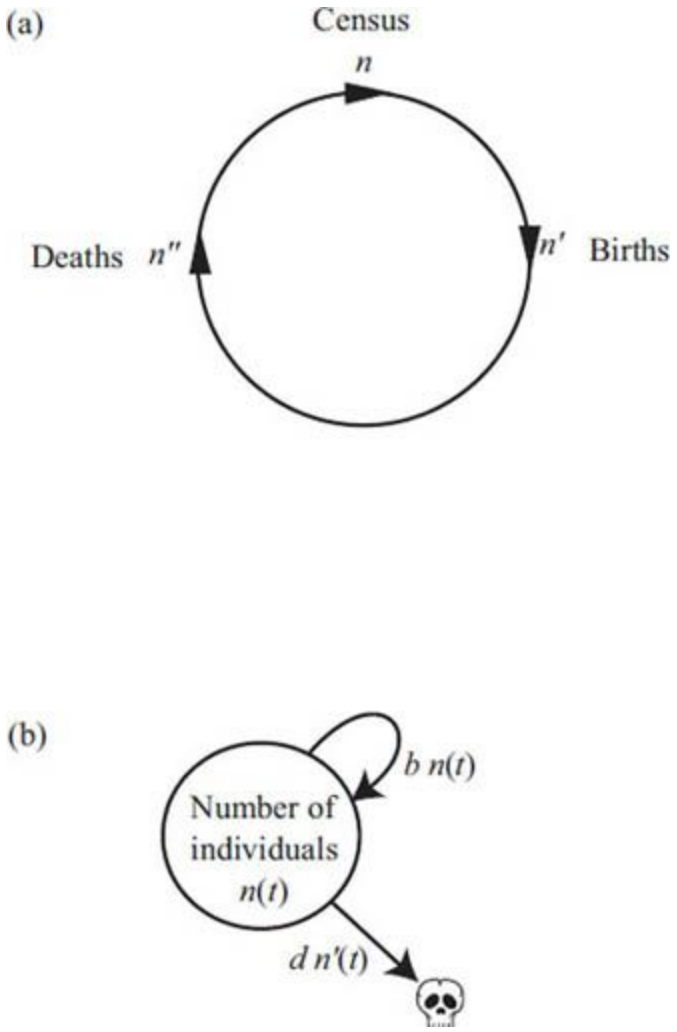

Figure 3.1: Exponential growth model. The (a) life cycle and (b) flow diagram for the exponential growth model.

As deaths represent the final event in the life cycle, we have $n(t + 1) = n''(t)$. Plugging the first equation into the second, we get

$$n(t + 1) = n''(t) = (1 - d)\,(1 + b)\,n(t). \tag{3.1a}$$

This recursion can be simplified by defining the compound parameter $R = (1 - d)\,(1 + b)$, so that

$$n(t + 1) = R\,n(t). \tag{3.1b}$$

R equals the number of surviving individuals per parent and is known as the *reproductive factor*. The key assumption of the exponential growth model is that R is constant, regardless of the population size.

We have already encountered a special case of the exponential growth model in Chapter 2. In the tree-branching example, each branch gave rise to b new branches in the next time step, and branches were never lost ($d = 0$). Thus, $R = 1 + b$, where R represents the parent branch plus b new branches (Figure 2.2a).

We can also describe the dynamics of the discrete-time exponential model using a difference equation, where $\Delta n = n(t + 1) - n(t)$:

$$\Delta n = (R - 1)\, n(t). \tag{3.2}$$

The quantity $(R - 1)$ is denoted by r in the biological literature and is the per capita change in the number of individuals from one generation to the next (known as the *growth rate*). Here we denote the growth rate by r_d, where the subscript d refers to the fact that this is the per capita change in the discrete-time model. From equation (3.1a) we have $r_d = R - 1 = b - d - b\,d$. If R equals one (each reproducing individual exactly replaces itself) then r_d equals zero (no change in the number of individuals), and the population size remains constant. Either of equations (3.1) and (3.2) can be used to predict the growth of a population over time (Chapter 4).

If births and deaths can occur at any point in time rather than during specific seasons, a continuous-time model of exponential growth is more appropriate. In the continuous-time exponential growth model, there is a per capita birth rate b and a per capita death rate d. When multiplied by a very small interval of time Δt, these give the number of births per individual and the fraction of the population that dies within that time interval (see Box 2.6). The flow diagram remains the same (Figure 3.1), except that, in any instant in time, only one event can happen. Therefore, we need not order the events, and we can drop the prime notation ($n'(t) = n(t)$). Applying Recipe 2.4 to the flow diagram, the differential equation for exponential population growth is

$$\frac{dn}{dt} = b\, n(t) - d\, n(t) = r_c\, n(t) \tag{3.3}$$

where $r_c = b - d$ is the per capita rate of change in the number of individuals in the continuous-time model.

The *exponential growth* model assumes that the variable changes at a rate that is proportional to its current value.

We can also derive the differential equation for exponential growth from the discrete-time recursion by shrinking the time interval over which events occur (Box 2.6). In a shorter unit of time Δt within the discrete-time model, we would expect the per capita number of births to be proportionately smaller (i.e., $b\,\Delta t$). The same is true of the fraction of the population that dies (i.e., $d\,\Delta t$). Consequently, after a short interval of time, the population size is

$$n(t + \Delta t) = (1 - d\,\Delta t)\,(1 + b\,\Delta t)\,n(t). \tag{3.4a}$$

Plugging $n(t + \Delta t)$ into the definition of a derivative (Appendix 2) and simplifying gives

$$\begin{aligned} \frac{dn(t)}{dt} &= \lim_{\Delta t \to 0}\left[\frac{n(t + \Delta t) - n(t)}{\Delta t}\right] \\ &= \lim_{\Delta t \to 0}\left[\frac{(1 - d\,\Delta t)(1 + b\,\Delta t)\,n(t) - n(t)}{\Delta t}\right] \\ &= \lim_{\Delta t \to 0}[(-d + b - d\,b\,\Delta t)\,n(t)]. \end{aligned} \tag{3.4b}$$

The term $-d\,b\,\Delta t$ disappears in the limit as Δt goes to zero, leaving us with $(-d + b)n(t)$, which is equivalent to the differential equation (3.3) derived from the flow diagram.

The derivation of equations (3.4) also helps to reveal why the discrete- and continuous-time parameters r_d and r_c differ by $-b\,d$. In the discrete-time model, deaths follow births within a single time step, and therefore newborns can potentially die before the next census. As we shrink down the time unit Δt, however, this possibility becomes negligible because very few births and deaths happen within a time unit.

As an example of exponential growth, in 1937, eight pheasants were introduced onto Protection island off the coast of Washington State, United States of America. Over the next five years the population grew exponentially (Lack 1954), nearly tripling in size every year ($R = 3$; Figure 3.2). If this trend had continued, we would be overrun by pheasants. Starting with eight pheasants, we would have $8 \times 3 = 24$ pheasants after one

year, $8 \times 3 \times 3 = 72$ pheasants after two years, $8 \times 3 \times 3 \times 3 = 216$ pheasants after 3 years, etc. Generalizing from this trend, there would be 8×3^i pheasants after i years of growth. With the average pheasant weighing about 1.5 kilograms, it would take only 50 years for the total mass of pheasants (i.e., $8 \times 3^{50} \times 1.5 = 8.6\ 10^{24}$ kilograms) to exceed the mass of the earth (5.98×10^{24} kilograms). Obviously exponential growth cannot continue unabated. In fact, Lack observed that "the increase was slowing down and was about to cease (the number in 1942 falls short of a threefold increase), but at this point the island was occupied by the military and many of the birds shot."

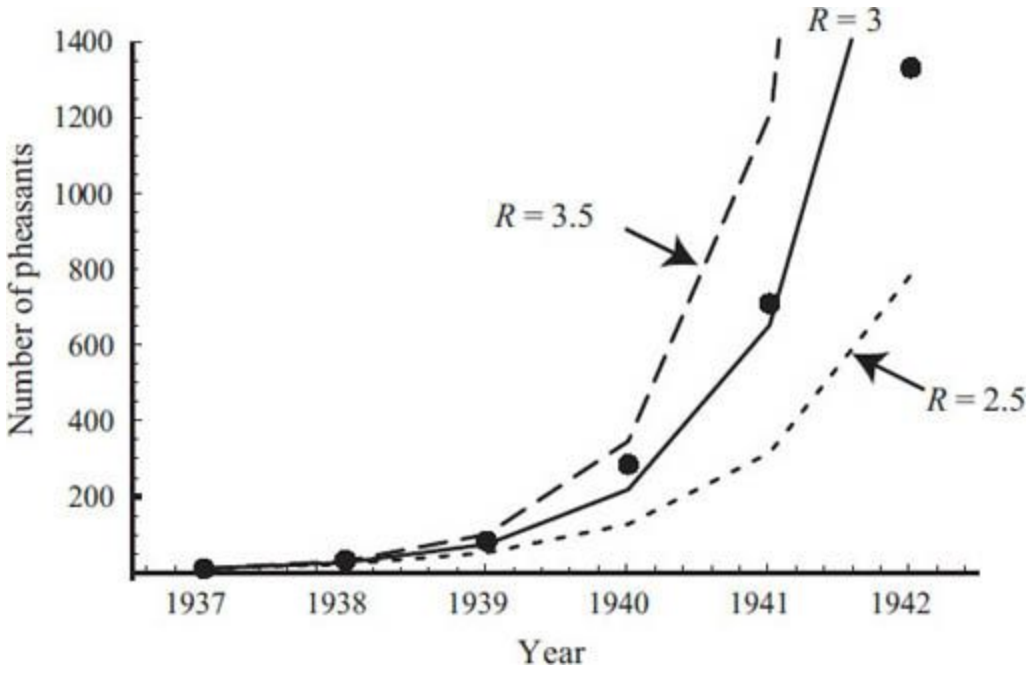

Figure 3.2: Example of exponential population growth in pheasants. The dots are based on spring census data (Lack 1954). The curves show the projected number of pheasants, $8 \times R^i$, with reproductive factors of $R = 2.5$ (short-dashed curve), $R = 3$ (solid curve), and $R = 3.5$ (long-dashed curve). These curves suggest that the reproductive factor of the pheasants was nearly 3 per year.

3.2.2 Logistic Population Growth

Many factors slow the rate of growth of a population, including declining resource availability, increased predation pressure, and a higher incidence of disease. The logistic model in discrete time describes these processes indirectly by assuming that R declines with increasing population

size. The exact form of this decline has been modeled in many ways (see Problem 3.4). The standard logistic model assumes that the number of surviving individuals per parent declines linearly with population size (Figure 3.3).

Writing the reproductive factor as $R(n)$ to emphasize its dependence on the population size, $R(n)$ starts at $(1 + r_d)$ when the population size is near zero and there is no competition for resources. The parameter r_d is known as the *intrinsic* rate of growth because it measures whether the population tends to grow ($r_d > 0$) or shrink ($r_d < 0$) when there is no competition for resources. As the population size increases, the reproductive factor decreases and eventually reaches a point where each individual exactly replaces itself; the population size at this point is called the *carrying capacity* K, because it is the maximum population size at which the population can sustain itself. From Figure 3.3, the equation for the line describing the reproductive factor as a function of population size is

$$R(n) = \underbrace{(1 + r_d)}_{\text{intercept}} + \underbrace{\left(-\frac{r_d}{K}\right)}_{\text{slope}} n(t).$$

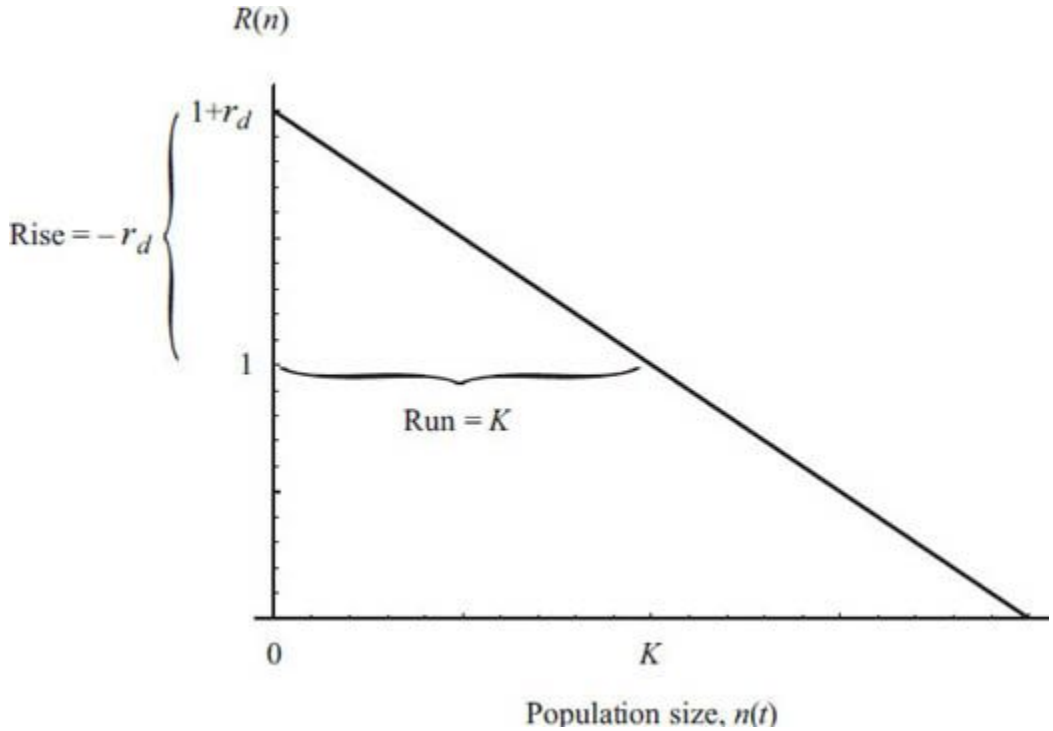

Figure 3.3: The number of surviving individuals per parent. The logistic model assumes that the number of surviving individuals declines linearly from a maximum of $1 + r_d$ when $n(t)$ is zero (the intercept) to one when $n(t)$ is K. The slope of the line is given by the rise ($-r_d$) over the run (K) between these two points.

Replacing R with $R(n)$ and gathering terms that depend on r_d, the recursion equation (3.1b) for population growth becomes

$$n(t + 1) = n(t) + r_d\, n(t)\left(1 - \frac{n(t)}{K}\right). \tag{3.5a}$$

Similarly, the difference equation (3.2) becomes

$$\Delta n = r_d\, n(t)\left(1 - \frac{n(t)}{K}\right). \tag{3.5b}$$

The *logistic growth* model assumes that the rate of change of a variable decreases linearly as the variable increases in value, with no change occurring if the variable is at the carrying capacity K.

The logistic model in continuous time may be derived by assuming that the per capita growth rate declines linearly with population size. Denoting this growth rate by $r(n)$, try to derive an expression for this function assuming that $r(n)$ declines linearly with $n(t)$, from a maximum equal to the intrinsic growth rate r_c when there are no competitors (i.e., $r(0) = r_c$) to zero when the population is at its carrying capacity (i.e., $r(K) = 0$). Doing so, and replacing the constant growth rate r_c in equation (3.3) with this function, gives the continuous-time logistic equation

$$\frac{dn}{dt} = r_c\, n(t)\left(1 - \frac{n(t)}{K}\right). \tag{3.5c}$$

As an example, Mable and Otto (2001) cultured haploid and diploid populations of the yeast *Saccharomyces cerevisiae* in separate flasks containing nutrients. “Haploid” means that there is one copy of every gene within the genome, while “diploid” means that there are two copies. Bacteria are haploid; animals and plants are typically diploid; and many single-celled organisms, fungi, and algae can persist in either state. By counting cell densities over time, population sizes were estimated for the

two ploidy levels (Figure 3.4 based on Figure 2 of Mable and Otto 2001). Although the populations grew nearly exponentially at first, growth rate decreased as the population size increased. The observed carrying capacity (K) is larger for the haploid cells, mainly because haploid cells are smaller than diploid cells and presumably require fewer resources per cell. Technically, yeast cells grow until the resources are depleted at which point the culture is said to be in "stationary phase" rather than at carrying capacity. But do the haploid and diploid cells have different intrinsic rates of increase, r_c? By fitting the logistic equation (3.5c) to the data, Mable and Otto (2001) estimated the parameters of the logistic equation as $K_{\text{haploid}} = 3.7 \times 10^8$, $r_{\text{haploid}} = 0.55$, $K_{\text{diploid}} = 2.3 \times 10^8$, and $r_{\text{diploid}} = 0.55$. (To find the best fitting parameters, a transformation was performed to turn the logistic equation into a line, and then fitting a line to the transformed data by a standard statistical technique known as linear regression. Details are available online in the *Mathematica* notebook for the figures in Chapter 3.) Despite the large difference in carrying capacities, the difference between the haploid and the diploid growth rate was negligible under these conditions.

Before moving on to the next section, we should introduce some terminology. The exponential growth equations (3.1), (3.2), and (3.3) represent *linear* models because the change that occurs per unit time depends on the variable only through a term that is proportional to $n(t)$. That is, these equations do not involve more complicated functions like $n(t)^2$, $1/n(t)$, $\exp(n(t))$, etc. In contrast, the logistic growth equations (3.5) represent *nonlinear* models because they are not linear functions of $n(t)$. This is more obvious if we multiply out the terms in the logistic equation; for example, equation (3.5b) can be rewritten as $\Delta n = r_d\, n(t) - r_d\, n(t)^2/K$, which involves the square of the variable, $n(t)^2$. As we shall see in Chapter 6, it is straightforward to obtain the *general solution* of linear models, meaning that there is an explicit equation that predicts exactly what the value of the variable will be at any future point in time. But it can be difficult or even impossible to find general solutions for nonlinear equations. Importantly, most models in biology are nonlinear, because we typically want to understand how cells/organisms/populations interact, and interactions generate nonlinear equations.

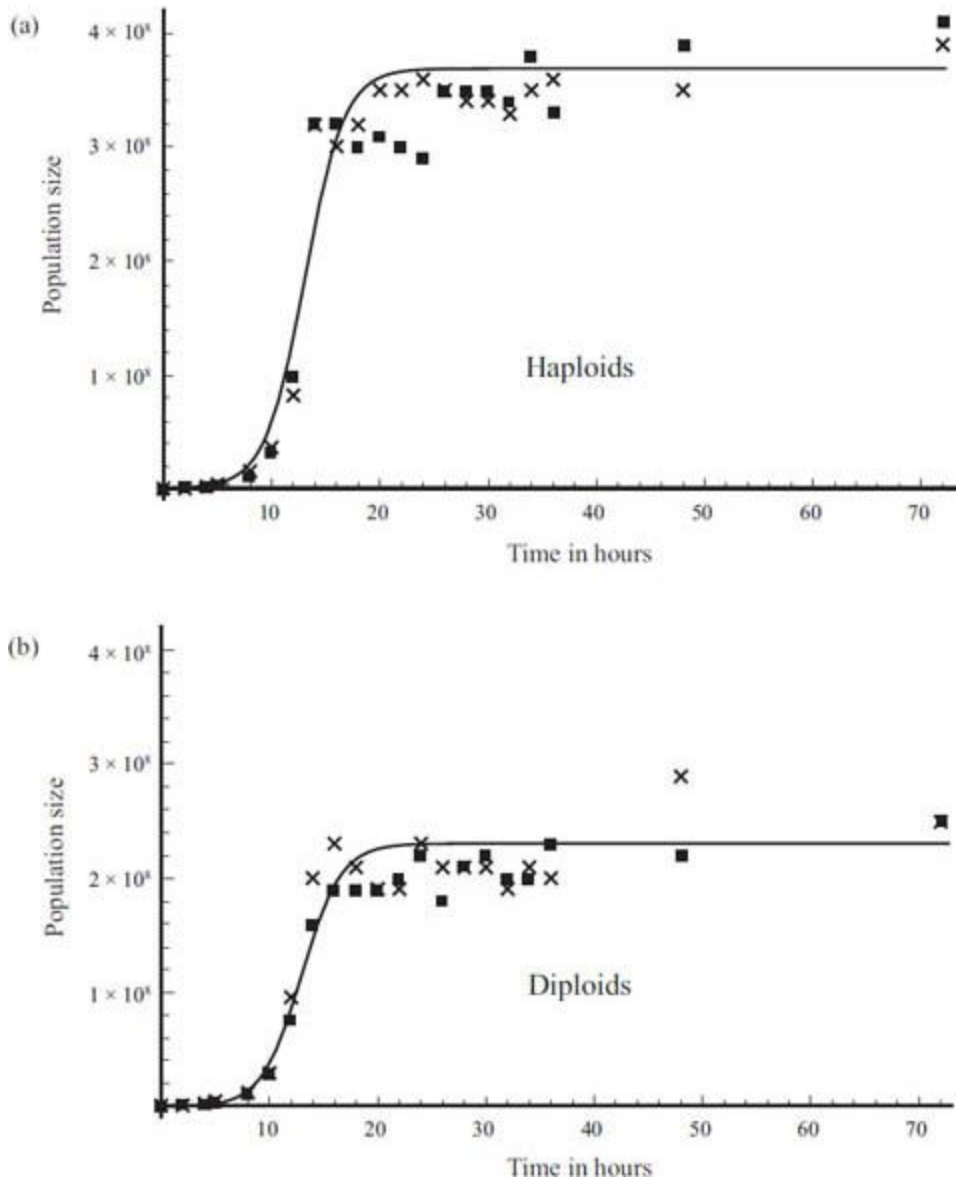

Figure 3.4: Example of logistic population growth in yeast. Population sizes from Mable and Otto (2001) are plotted over time based on data from (a) haploids and (b) diploids. Crosses and squares represent data from two different replicates. The solid curves are plots of the population size over time based on equations (3.5c) and the estimated parameter values (see Chapter 4).

In a *linear model*, the change per unit time is a linear function of the variables; models that involve more complicated functions of the variables are called *nonlinear models*.

It is confusing that equations (3.1), (3.2), and (3.3) are sometimes referred to as linear equations and, at other times, referred to as equations for exponential growth. These different terms reflect different ways of viewing a dynamical model: we can focus either on *changes* to the system

as a function of the variables themselves, or on the *value* of the variable as a function of time. Consider equation (3.3). This equation for dn/dt tells us that the rate at which the variable changes is a linear function of the variable itself, $n(t)$. We therefore call this a linear differential equation. But just because a differential equation is linear does not mean that the value of the variable is a linear function of time. We must analyze the model further to determine the exact value of the variable as a function of time. For the exponential growth model, the value of the variable as a function of time is given by the general solution $n(t) = n(0)e^r\ c^t$ (see Chapters 4 and 6). Thus, the value of the variable $n(t)$ is an exponential function of the independent variable t.

3.3 Haploid and Diploid Models of Natural Selection

In the above section, survival and reproduction were assumed to be the same for every member of a population. What if genetic variation in these characteristics exists within a population? This question is addressed by population-genetic models that track the frequency of different variants within a population over time. In this section, we explore two models of evolution by natural selection, one that applies to a haploid population where each individual carries one *allele* (one variant) of a gene, and one that applies to a diploid population where each individual carries two alleles. Figure 3.5 illustrates the life cycle for the two models, specifying where selection acts relative to reproduction.

> *Population-genetic models* describe how variants of a gene (alleles) change in frequency over time.

3.3.1 Haploid Models of Natural Selection

Consider a population of two types of individuals (A and a), which both breed true (i.e., type A produces only A offspring, and type a produces only a offspring). This scenario describes the case of a haploid population with two alleles (A and a) at one genetic locus. The model also describes two types of individuals (A and a) within an asexual population. First, consider

the number of each type of individual: n_A and n_a. To be concrete, we assume that the population is censused immediately after juveniles are born, although equivalent equations can be derived using other census points (e.g., by counting haploid adults). Each type survives and reproduces according to Figure 3.1, but we now allow R to differ between A and a individuals. It is traditional in evolutionary biology to denote the reproductive factor by W, and therefore, using equation, (3.1) we have

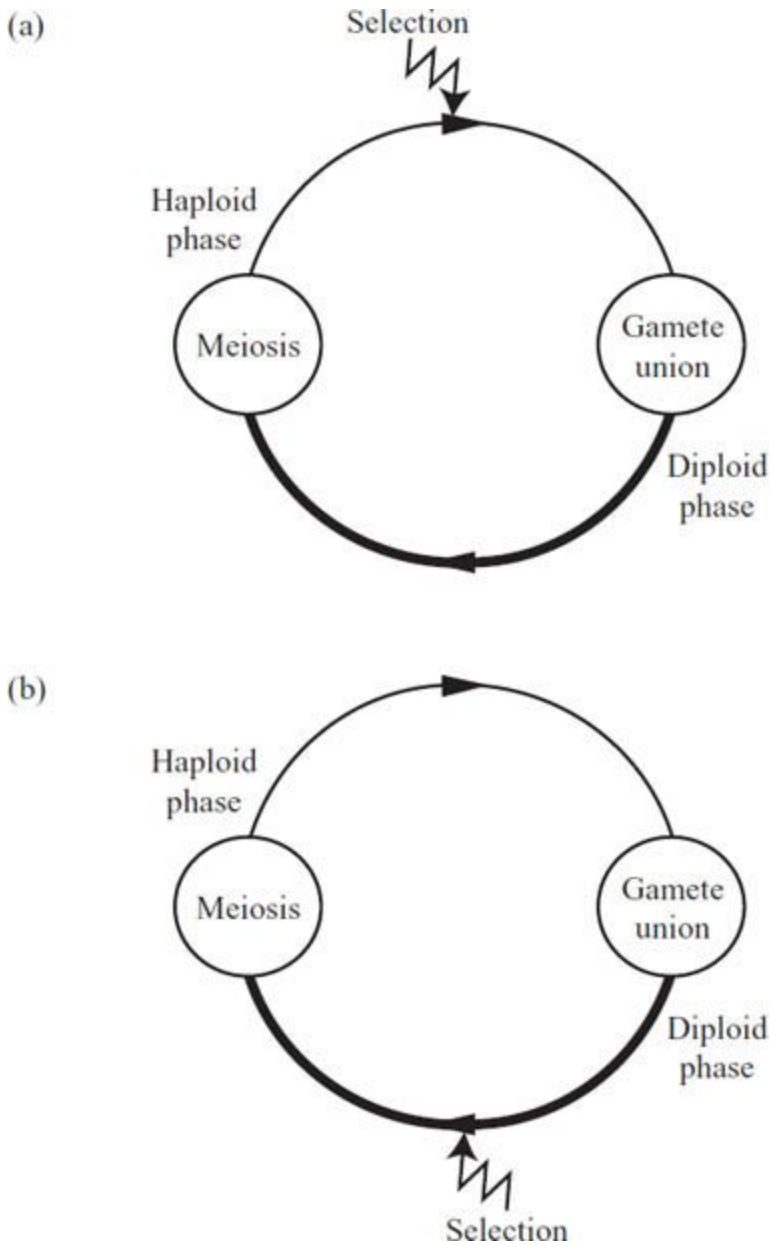

Figure 3.5: Life-cycle diagram for models of selection. All sexual life cycles can be conceptualized as having a haploid and a diploid phase, although there might be no growth in one phase (for

example, the haploid phase of the human life cycle consists only of the single cells, sperm and eggs). Selection can act in the (a) haploid phase or (b) diploid phase.

$$n_A(t+1) = W_A\, n_A(t), \tag{3.6a}$$

$$n_a(t+1) = W_a\, n_a(t). \tag{3.6b}$$

In terms of the birth and death parameters, we have $W_A = (1 - d_A)\,(1 + b_A)$ and $W_a = (1 - d_a)\,(1 + b_a)$ both of which must be positive (or zero).

In model (3.6), it is arbitrary which allele we label A. If we exchange the label A for a (and vice versa) in equations (3.6), we end up with the same two equations. Such models are called *symmetric*. Observing that a model is symmetric is helpful because it provides an excellent way to check a set of equations; exchanging the parameters and variables for the symmetric terms should not affect the equations or their predictions.

> In a *symmetric* model, the labels of the variables are arbitrary; interchanging the variable and parameter names does not alter the form of the equations.

Most evolutionary models focus on the frequency of each type rather than their absolute numbers. Typically, p and q are used to denote the proportions of A and a alleles. By definition, a proportion is equal to the number of a type divided by the total number of individuals ($n_A + n_a$):

$$p = \frac{n_A}{n_A + n_a}, \tag{3.7a}$$

$$q = \frac{n_a}{n_A + n_a}. \tag{3.7b}$$

Because they are proportions, p and q must lie between zero and one, and their sum must always equal one. Thus, if we know the value of one of these variables, the other can be obtained from the fact that $p + q = 1$. Consequently, the dynamics of the allele frequencies can be described using only a single variable, e.g., p.

By plugging equation (3.6) into (3.7a), the frequency of allele A in the next time unit $p(t + 1)$ can be found:

$$p(t+1) = \frac{n_A(t+1)}{n_A(t+1) + n_a(t+1)} \tag{3.8a}$$

$$= \frac{W_A\, n_A(t)}{W_A\, n_A(t) + W_a\, n_a(t)}. \tag{3.8b}$$

Equation (3.8b) gives the allele frequency at time $t + 1$, but it does so in terms of the number of each type at time t. To obtain a recursion equation that can predict the dynamics of the allele frequency over the course of several generations, we need to write the right-hand side of (3.8b) in terms of $p(t)$. There are many ways to accomplish this transformation using the relationship (3.7) between allele frequencies and allele numbers. The easiest is to use a trick: if we divide the top and bottom by $(n_A + n_a)$, then each term involves only p or q:

$$\begin{aligned} p(t+1) &= \frac{W_A \dfrac{n_A(t)}{n_A(t) + n_a(t)}}{W_A \dfrac{n_A(t)}{n_A(t) + n_a(t)} + W_a \dfrac{n_a(t)}{n_A(t) + n_a(t)}} \\ &= \frac{W_A\, p(t)}{W_A\, p(t) + W_a\, q(t)} \\ &= \frac{W_A\, p(t)}{W_A\, p(t) + W_a(1 - p(t))} \end{aligned} \tag{3.8c}$$

where, in the last line, we use the fact that the allele frequencies sum to one to rewrite $q(t)$ as $1 - p(t)$.

While most experienced modelers would use this trick (because they can "see" that it would simplify the equation), such insight only comes with experience. An alternative and more general method toward the same end is to solve equation (3.7) for the old variable n_A in terms of the new variable p. Multiplying both sides of equation (3.7a) by $n_A + n_a$, we get $p\,(n_A + n_a) = n_A$. Gathering the terms involving the old variable n_A on the left, we get $p\, n_A - n_A = -p\, n_a$. Finally, dividing both sides by the factor $(p - 1)$, we get

$$n_A = \frac{-p}{p - 1} n_a = \frac{p}{1 - p} n_a.$$

Plugging this term at time t into the right-hand side of equation (3.8b) gives

$$p(t + 1) = \frac{W_A \frac{p(t)}{1 - p(t)} n_a(t)}{W_A \frac{p(t)}{1 - p(t)} n_a(t) + W_a n_a(t)}.$$

By factoring out $n_a(t)$ and multiplying the top and bottom by $1 - p(t)$, we obtain equation (3.8c).

In this model, the numbers of surviving individuals per parent for the two alleles are W_A and W_a. These numbers are known as the *absolute fitnesses* of the two types. Although equation (3.8c) contains both of these parameters, we can use a technique similar to that used when deriving (3.8c) to reduce the number of parameters in this model. Again, it might not be obvious how to do this, but with experience you will be able to see that dividing the top and bottom of (3.8c) by W_a produces an equation with only one parameter:

$$p(t + 1) = \frac{V_A p(t)}{V_A p(t) + (1 - p(t))} \qquad (3.8d)$$

where $V_A = W_A/W_a$ is the *relative fitness* of allele A (i.e., the fitness of allele A relative to allele a).

The fact that the dynamics of allele frequency depend only on the relative fitnesses of the two alleles reveals important information about the causes of evolutionary change. For example, suppose that resource availability fluctuates over time in a way that causes the number of surviving offspring per parent to equal $\sigma(t)\, W_A$ and $\sigma(t)\, W_a$ for alleles A and a, where $\sigma(t)$ is a factor that varies over time and modulates the otherwise constant reproductive factors W_A and W_a. Despite this added environmental complexity, the allele frequency dynamics continue to be described by equation (3.8d), because $\sigma(t)$ cancels out when calculating the relative

fitness V_A. Similarly, if the density of the population ($n_A + n_a$) affects the number of surviving offspring per parent carrying allele A by the same factor as those parents carrying allele a, the allele frequency dynamics will remain the same. This is a key insight: evolutionary change can be studied without reference to the ecological context of a population as long as the ecological context affects the reproductive output of each allele by the same factor. This result does not hold, however, if the ecological context differs between the alleles, e.g., if they differ in their sensitivity to density-dependent competition (see Problems 3.10 and 3.17).

Given the recursion equation for the discrete-time haploid model, we can derive a continuous-time differential equation following the method of Box 2.6. To determine the allele frequency change, we calculate the difference in allele frequency over one generation by subtracting the current allele frequency $p(t)$ from both sides of equation (3.8c):

$$\begin{aligned}\Delta p &= p(t+1) - p(t) = \frac{W_A\, p(t)}{W_A\, p(t) + W_a\,(1 - p(t))} - p(t) \\ &= \frac{(W_A - W_a)\, p(t)\,(1 - p(t))}{W_A\, p(t) + W_a\,(1 - p(t))},\end{aligned} \tag{3.9}$$

which is factored by placing each term over the common denominator $W_A\, p(t) + W_a\,(1 - p(t))$. It is useful to define the *selection coefficient* for the discrete-time model, $s_d = (W_A - W_a)/W_a$, which is the proportional difference in fitness between allele A and allele a. We can use this definition to replace W_A in equation (3.9) with $W_A = W_a s_d + W_a$, giving us

$$\begin{aligned}\Delta p &= \frac{W_a s_d\, p(t)(1 - p(t))}{(W_a s_d + W_a)\, p(t) + W_a\,(1 - p(t))} \\ &= \frac{s_d\, p(t)(1 - p(t))}{1 + s_d\, p(t)}.\end{aligned} \tag{3.10}$$

Equation (3.10) describes the amount of allele frequency change over one generation. We can derive the differential equation describing the rate of allele frequency change in continuous time by supposing that the number

of births and the fraction of the population that dies in a time interval, Δt, are $b\,\Delta t$ and $d\,\Delta t$, respectively, just as we did in the exponential growth model (3.4a). The reproductive outputs of A and a individuals over this interval become $W_A(\Delta t) = (1 - d_A\Delta t)(1 + b_A\Delta t)$ and $W_a(\Delta t) = (1 - d_a\Delta t)(1 + b_a\Delta t)$, and therefore the selection coefficient will be a function of the length of the time interval Δt. The selection coefficient $s_d(\Delta t)$ can be written explicitly as

$$s_d(\Delta t) = \frac{W_A(\Delta t) - W_a(\Delta t)}{W_a(\Delta t)}$$

$$= \frac{((b_A - d_A) - (b_a - d_a))\Delta t - (b_A\, d_A - b_a\, d_a)\Delta t^2}{1 + (b_a - d_a)\Delta t - b_a\, d_a\, \Delta t^2}.$$

We can derive a differential equation for the allele frequency using equation (3.10) and the definition of the derivative (Appendix 2):

$$\frac{dp}{dt} = \lim_{\Delta t \to 0} \frac{\Delta p}{\Delta t} = \lim_{\Delta t \to 0} \left(\frac{s_d(\Delta t)}{\Delta t} \frac{p(t)(1 - p(t))}{1 + s_d(\Delta t)\, p(t)} \right). \qquad (3.11a)$$

The term $s_d(\Delta t)$ enters equation (3.11a) in two places. It first enters as $s_d(\Delta t)/\Delta t$. Dividing $s_d(\Delta t)$ by Δt and letting Δt go to zero, we get $(b_A - d_A) - (b_a - d_a)$, which is the selection coefficient for the continuous-time model, s_c. The second place it enters is $1 + s_d(\Delta t)\, p(t)$, which goes to one as Δt goes to zero. Consequently, as we shrink the time interval in equation (3.11a), we are left with

$$\frac{dp}{dt} = s_c p(t)(1 - p(t)). \qquad (3.11b)$$

The same differential equation can be derived from the flow diagram for each of the alleles (Figure 3.1) using the quotient rule of calculus (Box 3.1).

The differential equation (3.11b) describes how the allele frequency changes over time and provides insight into the action of natural selection (Figure 3.6). First, when allele A or a is absent from the population (i.e.,

when p or q equals zero), dp/dt will be zero, and p remains constant. In other words, if there is no genetic variation, there will be no change in allele frequency due to selection. Second, the rate of evolutionary change is maximized when the alleles are equally frequent (Figure 3.6; Problem 3.19), i.e., when genetic variation is most abundant.

At first sight, the differential equation (3.11b) describing the change in allele frequency in continuous time looks substantially different from the difference equation (3.10) describing the change in allele frequency over a discrete generation. Yet these equations are not as different as they seem. When selection is weak (i.e., when s_d is near zero), the denominator in equation (3.10), $1 + s_d p(t)$, will be very nearly one. Making this approximation, the difference and differential equations become identical (see Primer 1 for an overview of approximations). Consequently, when selection is weak the evolution of allele frequencies should be similar, regardless of whether they are modeled in discrete or continuous time (Figure 3.6).

3.3.2 Diploid Models of Natural Selection

As diploid organisms, it is natural for us to wonder if evolutionary change in diploids differs from evolutionary change in haploids. For example, does the efficiency with which natural selection changes allele frequencies differ between haploids and diploids? To answer this question, we need to develop a model of selection for diploids. Even with only two alleles present (again, A and a at frequencies p and q), we must keep track of three genotypic combinations within a diploid population: AA homozygotes, Aa heterozygotes, and aa homozygotes. You might wish to repeat the derivations used above, starting with the number of each diploid genotype and converting these equations into the frequencies of each allele as we did in equations (3.6)–(3.11) (see Supplementary Material 3.1). In the following, we take a different approach, which allows us to introduce several shortcuts that are often used in deriving evolutionary models.

The first shortcut involves censusing the population at the stage in the life cycle that is simplest to describe. In a sexual population, diploid organisms are produced by the union of haploid gametes. Because there are only two types of gametes (A and a) rather than three diploid genotypes, it

is simpler to census at the gamete stage, letting the frequency of A and a gametes be $p(t)$ and $q(t)$, respectively. As long as every diploid individual is formed by gametes uniting at random with one another (i.e., as long as there is no selfing, assortative mating, sexual selection, spatial structure, surviving individuals from the previous generation, etc.), then the frequency of each diploid genotype can be calculated from the gamete frequencies alone. (Fortunately, the same holds when individuals, not gametes, unite at random; see Problem 3.7.) When two gametes unite, the chance that the first gamete (the egg, say) carries allele A is $p(t)$. The chance that the second gamete (the sperm, say) carries allele A is also $p(t)$. Thus, of the $p(t)$ offspring formed from an A egg, a fraction $p(t)$ will involve an A sperm. But a fraction $p(t)$ of $p(t)$ is just $p(t)^2$ (e.g., 1/6 of 1/6 is 1/36). Thus, $p(t)^2$ of the offspring will be AA. This logic may be repeated for all of the possible combinations of gametes as illustrated in Figure 3.7, demonstrating that the frequencies of the diploid genotypes at the beginning of a generation (among the "zygotes") are $p(t)^2$ for AA, $2\,p(t)q(t)$ for Aa, and $q(t)^2$ for aa. These are the so-called Hardy-Weinberg proportions. As proportions, they should sum to one, which indeed they do: $p(t)^2 + 2\,p(t)q(t) + q(t)^2 = (p(t) + q(t))^2 = (1)^2 = 1$.

Box 3.1: Haploid Selection and the Quotient Rule

In the text we considered how a discrete-time haploid model of selection can be derived from the exponential model of population growth (Figure 3.1) by considering two types of individuals, those that carry allele A and those that carry allele a. The same method can be used to derive a continuous-time haploid model of selection. Again, we consider the number of each type of individual, n_A and n_a, and we let each type reproduce according to Figure 3.1 (except that only one event happens at any instant in time and $n' = n$). Let the growth rate $r = (b - d)$ depend on the allele carried by an individual, r_A or r_a. We can write down differential equations for the dynamics of the two types (see equation 3.3):

$$\frac{dn_A}{dt} = r_A n_A(t), \tag{3.1.1a}$$

$$\frac{dn_a}{dt} = r_a n_a(t). \tag{3.1.1b}$$

Equations (3.1.1) describe how the numbers of each type change, but what if we want to know how the allele frequency changes over time? The first step is to rewrite dp/dt using equation (3.7) as $d(n_A/(n_A + n_a))/dt$. (Here, we simplify the notation by dropping the (t) and calling the variables p, n_A, and n_a.) We can use the quotient rule (A2.13) to evaluate this derivative:

$$\begin{aligned}\frac{dp}{dt} &= \frac{d\left(\frac{n_A}{n_A + n_a}\right)}{dt} \\ &= \frac{\frac{dn_A}{dt}(n_A + n_a) - n_A\frac{d(n_A + n_a)}{dt}}{(n_A + n_a)^2}.\end{aligned} \tag{3.1.2}$$

Because the derivative of a sum, $d(n_A + n_a)/dt$, is the sum of the derivatives, $dn_A/dt + dn_a/dt$ (Rule A2.3), equation (3.1.2) equals

$$\frac{dp}{dt} = \frac{\frac{dn_A}{dt}(n_A + n_a) - n_A\frac{dn_A}{dt} - n_A\frac{dn_a}{dt}}{(n_A + n_a)^2}.$$

Canceling out terms in the numerator leaves

$$\frac{dp}{dt} = \frac{n_a\frac{dn_A}{dt} - n_A\frac{dn_a}{dt}}{(n_A + n_a)^2}.$$

Now, we can use equations (3.1.1) to replace the derivatives on the right-hand side:

$$\frac{dp}{dt} = \frac{n_a(r_A n_A) - n_A(r_a n_a)}{(n_A + n_a)^2}$$

$$= (r_A - r_a)\frac{n_A n_a}{(n_A + n_a)^2}.$$

Finally, we can use (3.7) to rewrite the product $n_A\, n_a/(n_A + n_a)^2$ as pq, leaving us with

$$\frac{dp}{dt} = (r_A - r_a)\, pq. \qquad (3.1.3)$$

If we set $q = 1 - p$ and define the selection coefficient s_c as the difference in growth rates, $s_c = r_A - r_a$, equation (3.1.3) becomes equation (3.11b).

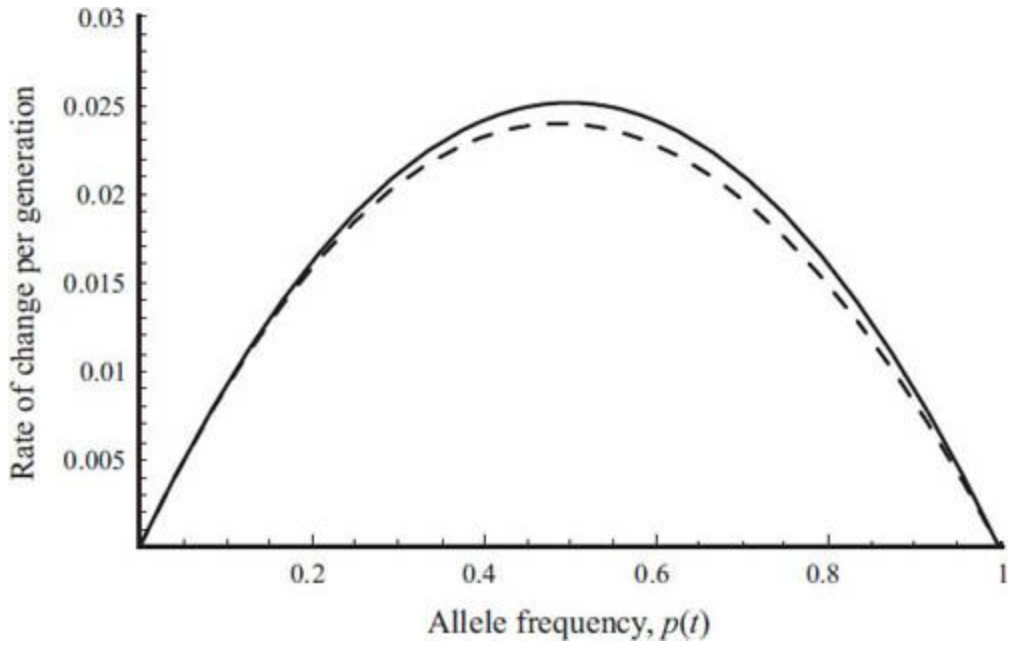

Figure 3.6: Allele frequency change with selection. The thick curve illustrates the instantaneous rate of change in the frequency of allele A, dp/dt, for the continuous-time model of haploid selection, given by the differential equation (3.11b). The dashed curve illustrates the per generation amount of change Δp in the discrete-time model, given by the difference equation (3.10). The two equations predict similar rates of change as long as selection is not too strong ($s = 0.1$ shown here).

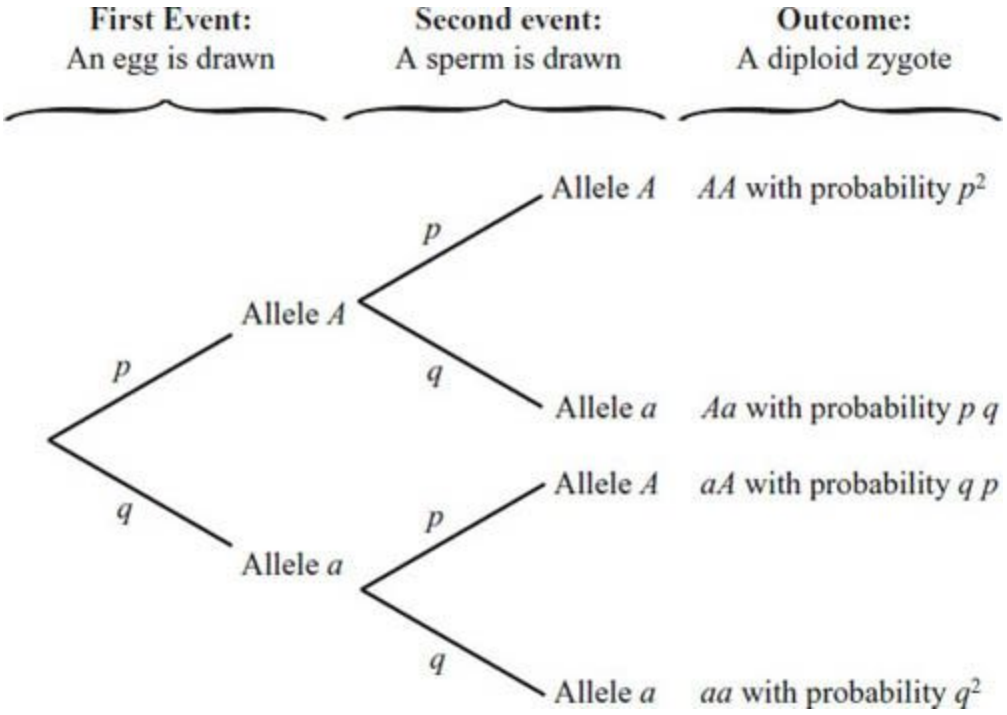

Figure 3.7: A "decision" tree for the union of two gametes. The outcome of this decision tree is a zygote (far right), which is formed by drawing an egg (first event) followed by a sperm (second event) from a large gamete pool. At each branching point in a decision tree, all of the possible alternatives are illustrated as lines, with the chance of observing each alternative written above the line. For each gamete, the chance of drawing allele *A* is *p*, and the chance of drawing allele *a* is *q*. As long as the events are independent of one another, then the probability of observing a particular outcome is equal to the product of the probabilities over all lines leading to that outcome. Because we assume that the gametes unite at random, the types of eggs and sperm are, by definition, independent of one another. If there are alternative ways of achieving the same outcome, the total probability of observing the outcome is the sum of the probabilities for each alternative. For example, an *Aa* heterozygote may be produced by the combination of an *A* egg and an *a* sperm, or vice versa, each of which occurs with probability *pq*. Thus, the total probability of observing a heterozygote is 2*pq*.

The second short-cut is to recognize that selection alters the frequency of each genotype by an amount proportional to its fitness. This is clearly true when dealing with the numbers of each genotype (see equations (3.6)), but it also holds if we convert to genotype frequencies by dividing by the total population size. This short-cut allows us to skip directly to a description of the diploid genotype frequencies after selection, which are proportional to $p(t)^2 W_{AA}$, $2p(t)q(t)W_{Aa}$, and $q(t)^2 W_{aa}$. These quantities no longer sum to one, however, but instead sum to what is known as the mean fitness of the population:

$$\overline{W} = p(t)^2 W_{AA} + 2\,p(t)q(t)\,W_{Aa} + q(t)^2 W_{aa}. \qquad (3.12)$$

To ensure that the frequencies sum to one and that the frequency of each genotype after selection is proportional to its fitness, we divide by the mean fitness to get $p(t)^2 W_{AA}/\overline{W}$, $2p(t)q(t)W_{Aa}/\overline{W}$, and $q(t)^2 W_{aa}/\overline{W}$. This process of dividing a set of non-negative numbers by their sum is called *normalization*; it converts the numbers into frequencies. These calculations are summarized in Table 3.2. Such a table of events is very handy, especially when the number of genotypes becomes large, in which case flow diagrams become hopelessly complex.

Finally, the diploid adults that have survived selection undergo meiosis to produce gametes. With Mendelian segregation, the percentage of gametes that carry the *A* allele is 100% from *AA* individuals, 50% from *Aa* individuals, and 0% from *aa* individuals (ignoring mutation). This information is provided in the second to last column of Table 3.2. Multiplying this column by the frequency of each adult, gives us the total frequency of the *A* gamete in the next generation:

> To convert a set of numbers into the frequency of each type, we *normalize* the numbers by dividing by their sum; this ensures that the frequencies sum to one.

$$p(t+1) = p(t)^2 \frac{W_{AA}}{\overline{W}} + \frac{1}{2}\left(2\,p(t)\,q(t)\frac{W_{Aa}}{\overline{W}}\right). \tag{3.13a}$$

Similarly, the total frequency of the *a* gamete is:

$$q(t+1) = \frac{1}{2}\left(2\,p(t)\,q(t)\frac{W_{Aa}}{\overline{W}}\right) + q(t)^2 \frac{W_{aa}}{\overline{W}}. \tag{3.13b}$$

Because this model is also symmetric, we can check equation (3.13a) by replacing *A* with *a* and *p* with *q*, which correctly has the form of equation (3.13b). Another check is that the gamete frequencies in the next generation should sum to one, which they do: $p(t+1) + q(t+1) = (p(t)^2 W_{AA} + p(t)\ q(t)W_{Aa} + p(t)\ q(t)W_{Aa} + q(t)^2 W_{aa})/\overline{W}$, which equals one by the definition (3.12) of the mean fitness.

In the above derivation, the fitnesses (W) act as "weights" that increase or decrease the representation of each genotype relative to the others. In biological terms, W could represent differences in survival ability

("viability") or differences in an individual's ability to make offspring ("fertility") or both. Furthermore, we could multiply each of the W's by any common factor σ without changing the behavior of the model, because the factor would cancel out when dividing by the mean fitness. Thus, as in the haploid model of natural selection, changes in allele frequencies among diploid organisms depend only on the relative success of each genotype under the assumptions we have made (e.g., no mutation, random mating, large population size).

TABLE 3.2
Genotype frequencies in the diploid model of natural selection

Gametes uniting to form a zygote	Frequency before selection	Frequency weighted by fitness	Frequency after selection	Gametes produced	
				A	a
$A \times A$	p^2	$p^2 W_{AA}$	$p^2 \frac{W_{AA}}{\overline{W}}$	1	0
$A \times a$	$p\,q$	$p\,q\,W_{Aa}$	$p\,q \frac{W_{Aa}}{\overline{W}}$	$\frac{1}{2}$	$\frac{1}{2}$
$a \times A$	$q\,p$	$q\,p\,W_{Aa}$	$q\,p \frac{W_{Aa}}{\overline{W}}$	$\frac{1}{2}$	$\frac{1}{2}$
$a \times a$	q^2	$q^2\,W_{aa}$	$q^2 \frac{W_{aa}}{\overline{W}}$	0	1

Although the haploid and diploid models of natural selection are similar in form, the efficiency of natural selection differs substantially. As described in Problem 3.9, the diploid model requires twice the amount of selection to accomplish the same change in allele frequency as the haploid model. At an intuitive level, diploid selection is less effective because the benefits of the most fit allele are partially masked in heterozygotes by the expression of the least fit allele.

3.4 Models of Interactions among Species

Species do not exist in isolation from one another. The simple models of exponential and logistic growth fail to capture the fact that the growth of a species must depend on its interactions with other species. Do the species compete with or aid each other in obtaining resources? Do they fight for territories or provide a habitat for one another? Does a species consume or provide fodder for another species? A complete model describing every species interaction would be prohibitively complex except in the simplest of biological communities. Yet some insight can be obtained by focusing on specific types of interactions and how these interactions affect population growth. Below we focus on competition models (section 3.4.1) and consumer-resource models (section 3.4.2).

3.4.1 The Lotka-Volterra Model of Competition

Let us begin by introducing the Lotka-Volterra model of competition, which builds upon the one-species logistic model. In the logistic equation, intraspecific competition becomes more severe as the population size increases, resulting in a lower per capita growth rate (Figure 3.3). The Lotka-Volterra model extends the logistic model by allowing multiple species to compete for resources, resulting in both inter- and intraspecific competition.

To keep things simple, we consider only two competing species, whose numbers $n_1(t)$ and $n_2(t)$ change over time. The two species can have different intrinsic growth rates r_1 and r_2, and different carrying capacities K_1 and K_2. To account for competition between the two species, the Lotka-Volterra model assumes that each individual of species i experiences competition as if its own species had a population size of $n_i(t) + \alpha_{ij}\ n_j(t)$. (Here we use i to refer to either species 1 or 2, depending on which species is the current focus of our attention, and use j to refer to the competing species. This saves us from having to write down equations for both species each time.) The parameter α_{ij} is known as the *competition coefficient*, which converts the strength of competition exerted by an individual of species j on an individual of species i into the equivalent amount of competition that would be exerted if both individuals were of species i. Assuming that the reproductive factor of species i declines linearly with the amount of competition, we now have

Competition models describe how the number of individuals of each species changes when more than one species uses the same resource.

$$R_i = \underbrace{(1 + r_i)}_{\text{intercept}} + \underbrace{\left(-\frac{r_i}{K_i}\right)}_{\text{slope}} (n_i(t) + \alpha_{ij}\, n_j(t))$$

Incorporating this reproductive factor into the discrete-time recursion equations $n_i(t + 1) = R_i n_i(t)$ and gathering together terms involving r_i, the Lotka-Volterra model becomes

$$n_1(t + 1) = n_1(t) + r_1\, n_1(t)\left(1 - \frac{n_1(t) + \alpha_{12}\, n_2(t)}{K_1}\right), \tag{3.14a}$$

$$n_2(t + 1) = n_2(t) + r_2\, n_2(t)\left(1 - \frac{n_2(t) + \alpha_{21}\, n_1(t)}{K_2}\right). \tag{3.14b}$$

Similarly, the continuous-time differential equations for the Lotka-Volterra model are

$$\frac{dn_1}{dt} = r_1\, n_1(t)\left(1 - \frac{n_1(t) + \alpha_{12} n_2(t)}{K_1}\right), \tag{3.15a}$$

$$\frac{dn_2}{dt} = r_2\, n_2(t)\left(1 - \frac{n_2(t) + \alpha_{21} n_1(t)}{K_2}\right). \tag{3.15b}$$

As a check, if the species do not interact in any way (i.e., $\alpha_{12} = \alpha_{21} = 0$), then equations (3.14) and (3.15) should reduce to the logistic equations (3.5a) and (3.5c), which indeed they do. Another check is provided by the fact that this model is also symmetric: which species is labeled 1 and which is labeled 2 is arbitrary. Thus, we can interchange the indices referring to species 1 and 2 in the recursion equation (3.14a) and regain equation (3.14b), and vice versa. The same holds for the differential equations (3.15a) and (3.15b).

Although we have been discussing competition, species j could facilitate the growth of species i, in which case α_{ij} would be negative. For

example, species 1 could excavate holes that are used as nesting sites by species 2. The signs of α_{12} and α_{21} therefore reflect the relationship between the two species:

α_{12}	α_{21}	Relationship
−	−	Mutualistic
−	0	Commensal
0	−	Commensal
+	−	Parasitic
−	+	Parasitic
+	+	Competitive

Equations (3.14) and (3.15) can thus be used to describe the dynamics of populations experiencing many different types of species interactions.

3.4.2 Consumer-Resource Models

Throughout the above, we have treated resources as constant. While this might be reasonable for some physical resources (e.g., light striking a patch of land or nutrients flowing down a river past a stationary aquatic community), in many cases the level of resources will be impacted by the growth of the species that we are modeling. For example, nutrients such as nitrogen might decline as a plant population grows, and the number of plankton might decline as a fish population grows. To account for such phenomena, we must construct models that explicitly track the dynamics of a resource as well as the population consuming the resource. Such models are referred to as *consumer-resource* models or as *predator-prey* models if the resource is itself an organism (a prey).

Consumer-resource models describe the dynamics of a resource and the dynamics of a population that consumes this resource.

The general structure of a consumer-resource model is illustrated in Figure 3.8 for the case with one resource $n_1(t)$ and one consumer $n_2(t)$. From this flow diagram, we can derive the dynamical equations for this model in continuous time by summing the rate of change over all arrows for each circle (Recipe 2.4):

$$\begin{aligned} \frac{dn_1}{dt} &= f(n_1(t)) - g(n_1(t),n_2(t)), \\ \frac{dn_2}{dt} &= \varepsilon\, g(n_1(t),n_2(t)) - h(n_2(t)). \end{aligned} \tag{3.16}$$

Similar equations can be derived in discrete time, once we have specified the order in which events occur.

We have chosen to illustrate the flow diagram and derive these differential equations in very general terms by not specifying the functions $f(n_1)$, $g(n_1,n_2)$, and $h(n_2)$. The advantage of developing such a general model is that it makes it possible to derive many different consumer-resource models by appropriate choices for the functions $f(n_1)$, $g(n_1,n_2)$, and $h(n_2)$, as described next.

The term $f(n_1)$ represents the rate of change of the resource through means other than consumption (i.e., assuming $n_2(t) = 0$). This resource-renewal function might involve inflow and outflow of abiotic resources (as in a chemostat; see Problem 2.4) or immigration, emigration, births, and deaths of biotic resources. Furthermore, the rate of change of the resource might depend on the current density of resources, or it might not. Table 3.3 lists some possible choices for $f(n_1)$. To simplify the notation, we drop the (t) notation, but we must remember that n_1 and n_2 are functions of time.

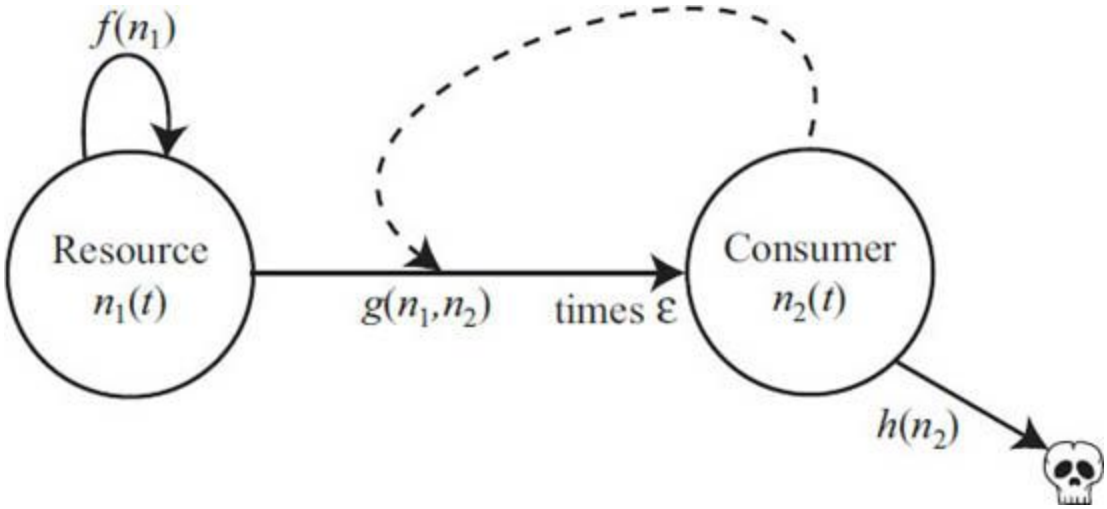

Figure 3.8: Flow diagram for consumer-resource model. A general flow diagram describing the levels of a resource n_1 and a consumer n_2, where the rate of renewal of resources, $f(n_1)$, the rate of resource consumption, $g(n_1,n_2)$, and the rate of loss of consumers, $h(n_2)$, are left unspecified.

The term $g(n_1,n_2)$ represents the rate of consumption of the resource by the consumer. In the simplest case, a "mass-action" rate of consumption is assumed, as in the flu model of Chapter 2. That is, the total rate of contact between consumers and resources within the community is assumed to equal $c\, n_1 n_2$. At each contact, the probability that the consumer successfully uses the resource is a. Hence, $g(n_1,n_2) = a\ c\ n_1 n_2$, which is known as a "linear" or "type-I" consumption rate (sometimes referred to as a type-I functional response). Table 3.3 lists other common choices for $g(n_1,n_2)$. After having gathered a resource, the consumer converts it into the biomass needed for offspring production. In the flow diagram and in equation (3.16), the conversion factor by which resource units are turned into consumers is given by ε. For example, one prey might represent only a fraction ε of the resources needed to produce one predator.

Finally, $h(n_2)$ represents the rate at which the number of consumers changes in the absence of resources (i.e., assuming $n_1 = 0$). Typically, it is assumed that consumers die off at a constant per capita rate δ, so that $h(n_2) = \delta n_2$. But again, the rate of change of consumers might involve immigration, emigration, and even births if the consumer can utilize alternative resources. Furthermore, this rate might or might not depend on the current density of consumers. Table 3.3 lists two possible alternatives for $h(n_2)$.

TABLE 3.3
Consumer-resource models. Examples of functions that can be used in the consumer-resource model (3.16), where n_1 refers to the level of resources (e.g., number of prey) and n_2 refers to the level of consumers (e.g., number of predators).

Function	Description
$f(n_1) = \theta$	Inflow of resources at a constant rate
$f(n_1) = -\psi$	Outflow of resources at a constant rate
$f(n_1) = r\, n_1$	Constant per capita growth of resource species
$f(n_1) = r n_1\left(1 - \frac{n_1}{K}\right)$	Per capita growth of resource species declines linearly with resource level (logistic)
$f(n_1) = r\, n_1\, e^{-\alpha n_1}$	Per capita growth of resource species declines exponentially with resource level
$g(n_1, n_2) = a\, c\, n_1 n_2$	Linear (type I) rate of resource consumption
$g(n_1,n_2) = \frac{a\, c\, n_1}{b + n_1} n_2$	Saturating (type II) rate of resource consumption
$g(n_1,n_2) = \frac{a\, c\, n_1^k}{b + n_1^k} n_2$	Generalized (type III) rate of resource consumption
$h(n_2) = \delta\, n_2$	Constant per capita death rate of consumer
$h(n_2) = (\delta\, n_2)\, n_2$	Per capita death rate of consumer increases linearly with consumer population size

By combining the functions $f(n_1)$, $g(n_1,n_2)$, and $h(n_2)$ in different ways, a large number of consumer-resource models are possible. Next, we present

equations for two of these models, which we will return to in later chapters.

We start with the model that incorporates the simplest choices for each of these functions: a constant inflow (immigration) of resources at rate $f(n_1) = \theta$, a type-I rate of resource consumption $g(n_1,n_2) = a\ c\ n_1 n_2$, and a constant death rate of consumers $h(n_2) = \delta\ n_2$. Equations (3.16) then become

$$\begin{aligned} \frac{dn_1}{dt} &= \theta - a\,c\,n_1(t)\,n_2(t), \\ \frac{dn_2}{dt} &= \varepsilon\,a\,c\,n_1(t)\,n_2(t) - \delta\,n_2(t). \end{aligned} \tag{3.17}$$

Equation (3.17) can be used, for example, to model the inflow into a lake of a nutrient (e.g., nitrogen) that is required for and limits the growth of an algal species, where $n_1(t)$ and $n_2(t)$ represent the levels of nutrients and algae, respectively. By setting $\theta = 0$, equation (3.17) can be used to model the growth of a community when resources begin at some level $n_1(0)$ and are not replenished. Interestingly, as we shall see in Chapters 4 and 9, the logistic equation and equation (3.17) without immigration and death ($\theta = \delta = 0$) generate identical predictions about the level of consumers over time.

The second consumer-resource model that we consider is a classic one in ecology, known as the Lotka-Volterra predator-prey model. The only difference from the previous model (3.17) is that the resources are prey that undergo exponential growth in the absence of predators, $f(n_1) = r\ n_1$. Equations (3.16) then become

$$\begin{aligned} \frac{dn_1}{dt} &= r\,n_1(t) - a\,c\,n_1(t)\,n_2(t), \\ \frac{dn_2}{dt} &= \varepsilon\,a\,c\,n_1(t)\,n_2(t) - \delta\,n_2(t). \end{aligned} \tag{3.18}$$

As we shall see in Chapter 4, the Lotka-Volterra predator-prey model predicts that the number of prey and number of predators cycle over time, demonstrating that species interactions can lead to interesting dynamical

behavior that might help explain the cyclic dynamics of several species (e.g., Fussmann et al. 2000; Krebs et al. 1995).

Although we have focused on continuous-time consumer-resource and predator-prey models, discrete-time models are straightforward to develop given the tools that you have now learned. As an example, we develop the discrete-time version of the Lotka-Volterra predator-prey model in Supplementary Material 3.2.

3.5 Epidemiological Models of Disease Spread

The final model that we derive describes the spread of a disease within a population. This is an extension of the flu model that we explored in the previous chapter, but we now add more realism by accounting for births, deaths, recovery, and the gain and loss of immunity. In the previous chapter we considered only two types of individuals, those susceptible to the flu and those infected with the flu. We now allow for a third class of individuals: those who have recovered from the disease and are currently resistant. This model is known as the SIR model in epidemiology, which stands for Susceptible-Infected-Recovered. Consequently, we use the variables $S(t)$, $I(t)$, and $R(t)$ to denote these three types of individuals at time t. We focus on the continuous-time model; the discrete-time version is similar, but with the variables updated after each event.

SIR epidemiological models track the dynamics of infectious diseases by modeling the number of susceptible, infected, and recovered individuals within a population.

We begin by using the "mass-action" assumption for the rate at which susceptible individuals contract the disease as in Chapter 2: $a\ c\ S(t)\ I(t)$, where an infected individual contacts susceptible individuals at a rate c per susceptible individual, and a is the probability of disease transmission upon contact (see Figure 2.4c). As illustrated in the flow diagram (Figure 3.9), infected individuals can now recover from the disease at a per capita rate ϱ. We assume that, while infected, these individuals developed antibodies to the pathogen that enable them to resist reinfection. Resistance need not be

permanent, however, and individuals that have recovered from the disease become susceptible again at a per capita rate σ.

We also incorporate the possibility that individuals die and that infected individuals have a different death rate δ than healthy individuals, d. Finally, we assume that new susceptible individuals enter the population at a rate θ through immigration. This assumption is reasonable if immigrants arrive from a location that has not yet been exposed to the disease. The model can also be generalized to account for reproduction from within the population, by replacing θ with the desired growth rate (e.g., by $b(S(t) + I(t) + R(t))$ assuming that infection does not alter the birth rate of individuals and that all offspring are born susceptible).

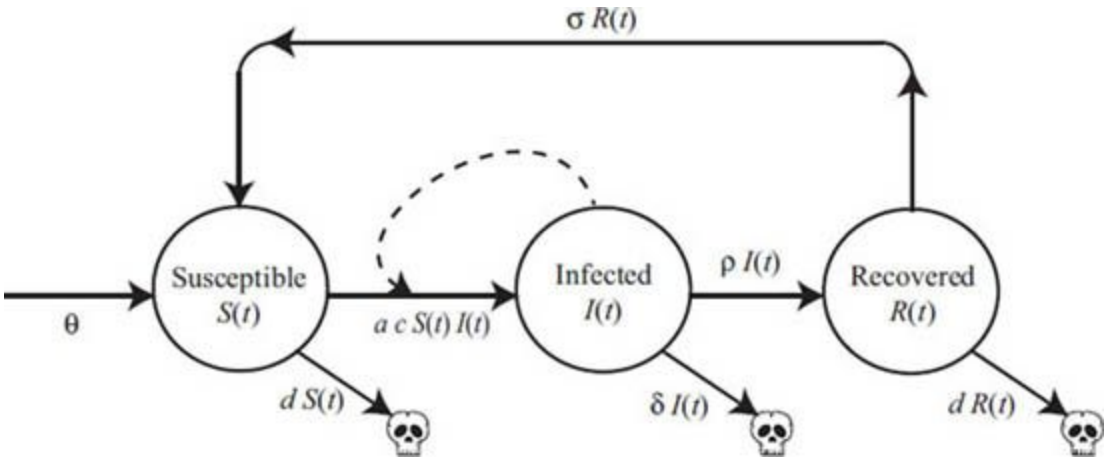

Figure 3.9: A flow diagram for the susceptible-infected-recovered (SIR) model

Using Figure 3.9, the differential equations for the SIR model are

$$\frac{dS}{dt} = \theta - d\,S(t) - a\,c\,S(t)\,I(t) + \sigma\,R(t), \tag{3.19a}$$

$$\frac{dI}{dt} = a\,c\,S(t)\,I(t) - \delta\,I(t) - \rho\,I(t), \tag{3.19b}$$

$$\frac{dR}{dt} = \rho\,I(t) - \sigma\,R(t) - d\,R(t) \tag{3.19c}$$

(Recipe 2.4). As a check, we should be able to regain the differential equations for our previous flu model if we remove the recovered category and let the parameters that describe the added processes of immigration,

death, and recovery equal zero ($\theta = \delta = d = \varrho = \sigma = 0$). Equations (3.19) then become

$$\frac{dS}{dt} = -a\,c\,S(t)\,I(t),$$
$$\frac{dI}{dt} = a\,c\,S(t)\,I(t).$$

Translating to the variable names used in Chapter 2 (where $S(t)$ was called $s(t)$ and $I(t)$ was called $n(t)$), we do indeed regain equations (2.10).

It is not always reasonable to assume that the transmission of disease between infected and susceptible individuals follows a law of mass action (McCallum et al. 2001). We can generalize the transmission term in epidemiological models by decomposing the rate of disease transmission from an infected individual into the product of three factors: (i) the rate of contact with other individuals in the population, (ii) the probability that, if a contact occurs, the individual contacted is susceptible to the disease, and (iii) the probability that, if a susceptible individual is contacted, the disease is transmitted.

Let us first show how the mass-action assumption is just a special case of this general approach. Under mass action, the rate at which an infected individual contacts other members of the population (regardless of their disease status) is $cN(t)$ per infected individual, where $N(t) = S(t) + I(t) + R(t)$ is the total population size. The probability that a contacted individual is susceptible to the disease equals $S(t)/N(t)$ (the fraction of the population that is susceptible). These two factors, corresponding to factors (i) and (ii) above, can be multiplied together into a single factor $c\,S(t)$, representing the rate of contact with susceptible individuals per infected individual. Multiplying by the total number of infected individuals and by the probability of transmission, a, gives $a\ c\ S(t)\ I(t)$, which is the total transmission rate that we used for mass action.

For sexually transmitted diseases, it is often more realistic to assume that the rate of contact between an infected individual and other individuals does not depend strongly on the population size, because people seek out sexual relationships regardless of the population size. In this case, the rate of contact with other individuals in the population will be nearly constant, c. Given that susceptible individuals represent a fraction $S(t)/N(t)$ of the

population, the rate of contact with susceptible individuals is then c $S(t)/N(t)$ per infected individual. Accounting for all of the infected individuals within the population and for the probability of transmission per contact, a, the total rate of transmission becomes a c $S(t)$ $I(t)/N(t)$. This is sometimes termed "proportional transmission", "frequency-dependent transmission," or "standard incidence." Proportional transmission was assumed by Blower et al. (2000) in their model of HIV (see Box 2.5).

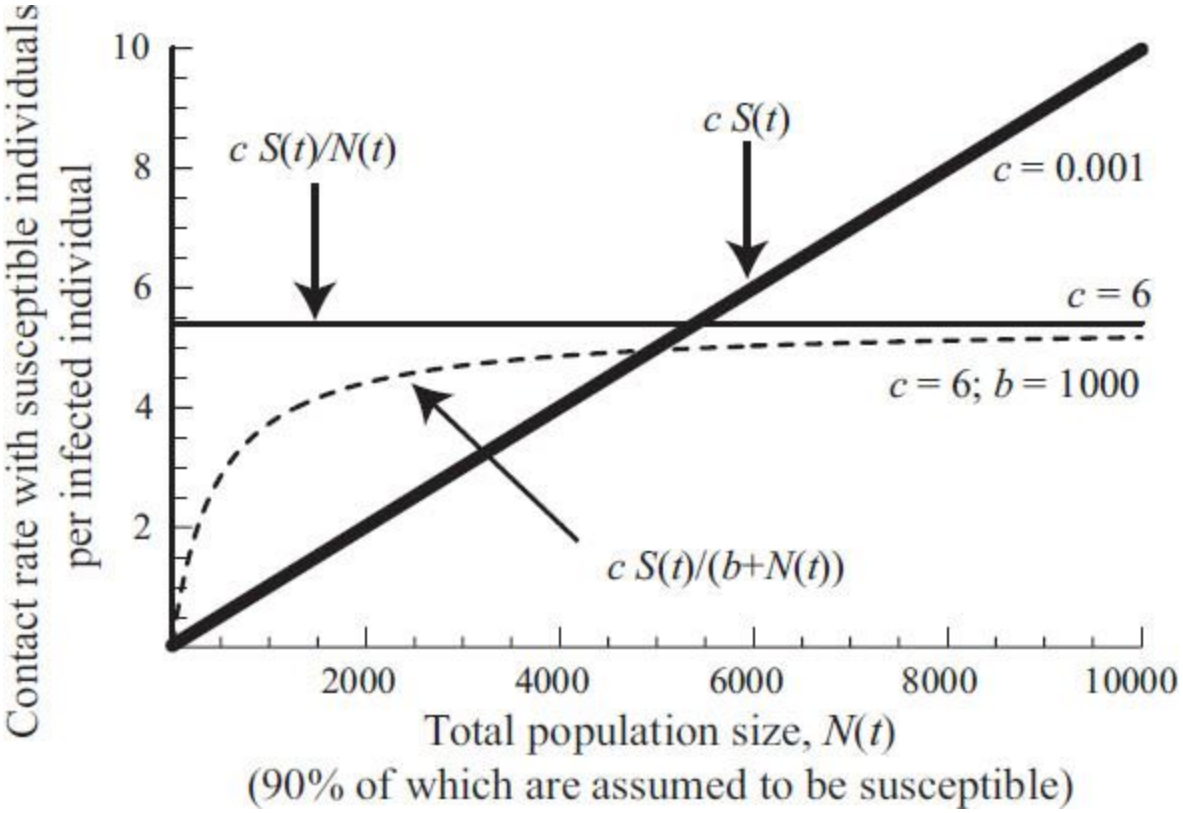

Figure 3.10: Contact rate between infected and susceptible individuals. When modeling disease transmission, the rate at which a particular infected individual contacts susceptible individuals depends on the assumptions of the model. Here, we compare the assumption of mass action (thick line, c $S(t)$), the assumption of a constant contact rate regardless of population size (thin line, $cS(t)/N(t)$), and the assumption of a diminishing-returns contact rate as the population size increases (dashed curve, c $S(t)/(b + N(t))$. In each case, we assume that 90% of the population is susceptible ($S(t) = 0.9$ $N(t)$).

Presumably, in very small populations, the contact rate for sexually transmitted diseases should depend on the population size, because it becomes difficult to find compatible partners. Thus, we might want to modify the above transmission model to allow the contact rate to be

proportional to the total population size for low population sizes but to reach a roughly constant value as the population size increases. As described in Primer 1, one potential function with this property is $c\,N(t)/(b + N(t))$ where b is a positive parameter. Multiplying this contact rate by the fraction of individuals susceptible to the disease, $S(t)/N(t)$, the rate of contact with susceptible individuals becomes $c\,S(t)/(b + N(t))$ per infected individual. Multiplying by the number of infected individuals and by the probability of transmission per contact, the total rate of transmission becomes $a\,c\,S(t)\,I(t)/(b + N(t))$. These three different types of interaction represent the most commonly used transmission functions and are compared in Figure 3.10.

3.6 Working Backward—Interpreting Equations in Terms of the Biology

Most of this book is aimed at teaching the process of model construction and analysis. This process involves starting with an idea, constructing mathematical equations to describe the idea, and then analyzing these equations in order to extract meaningful biological information. As biologists, however, we are often required to work through this process in reverse. For example, many biological papers that contain models have very terse and sometimes incomplete descriptions of the assumptions and methods used. As a result, it is crucial to be able to "read" equations and decipher the assumptions that are implicit in their derivation. This not only is useful for gaining a better understanding of such papers but is also a crucial skill for ensuring that a model accurately describes the processes that it is meant to. Additionally, models can often be applied to a much broader set of questions than those stated in the papers in which they appear, and therefore it is also important to be able to understand what the equations mean so that you can determine when it is safe to port them over to an entirely different question. As we will see, there are often multiple biological interpretations that are consistent with a single model.

Suppose you are reading a paper that models the population dynamics of a single species subject to exploitation, such as hunting or fishing. The paper claims to model a population in which there is logistic growth in the

absence of exploitation, along with a constant hunting pressure. Here are three different possible differential equations for such a model, each of which makes slightly different assumptions about these processes:

$$\frac{dn}{dt} = r\,n(t)\left(1 - \frac{n(t)}{K}\right) - \theta \tag{3.20a}$$

where r is the intrinsic growth rate parameter, K is the carrying capacity parameter, and θ is the constant harvesting rate parameter;

$$\frac{dn}{dt} = r\,n(t)\left(1 - \frac{n(t)}{K}\right) - H\,n(t) \tag{3.20b}$$

where H is the constant harvesting rate parameter;

$$\frac{dn}{dt} = r\,n(t)\left(1 - \frac{n(t)}{K}\right) - H\,(n(t) - P) \tag{3.20c}$$

where P and H are constant harvesting rate parameters.

These three equations have two terms; the first term represents within-population reproduction and the second term represents the loss of individuals through exploitation. According to the first term in all three models, the growth of the population in the absence of exploitation is described by the logistic equation (3.5c). Therefore, if there is no exploitation, we would expect the population to increase in size and plateau at a value of K (Figure 3.4).

All three equations also purport to model a constant harvesting rate, but precisely what is meant by a constant harvest rate is quite different for each model. There are many different ways in which the harvesting rate might be set, but we consider two possibilities: (i) the government regulates the number of organisms that can be taken from the population per time unit (e.g., year), or (ii) the government regulates the amount of "hunter-hours" that the population must endure per time unit. Let us consider each equation in turn to decipher the assumptions that were made about these processes.

We can see from equation (3.20a) that the *total* loss rate of individuals through harvesting is constant and independent of population size. In other words, the rate of loss of individuals through harvesting is the same

regardless of the current population size. This means that, of the two different possibilities for setting the harvest rate, this model more closely represents the assumption that the government sets the number of organisms that can be harvested from the population. Indeed, equation (3.20a) can be used to model any process in which a population grows logistically and is subject to a constant total rate of loss or gain of individuals. For example, it can be used to model density-independent emigration or immigration instead of harvesting.

What is meant by a constant harvest rate in equation (3.20b)? We can see from this equation that the loss of individuals through harvesting occurs at a constant per capita rate. Consequently, each individual in the population suffers a constant probability of being harvested, regardless of the population size. This contrasts with equation (3.20a) in which the *total* harvest rate is constant (i.e., each individual is less likely to be harvested as the population size increases). This model more closely represents the assumption that the government sets the total number of hunter-hours allowed, because we would expect the total loss rate to increase as the population size increases under this scheme. More generally, equation (3.20b) can be used to model any process in which a population is subject to a constant per capita rate of loss or gain of individuals.

The expression for the harvest rate in equation (3.20c) is more complex. Mathematically, we can see that the total loss rate of individuals through harvesting is proportional to $n(t) - P$. What assumptions about harvesting are consistent with this formulation? To answer this question we need to consider how the total loss rate is affected by the parameters and by the population size $n(t)$. The term $n(t) - P$ is positive as long as the population size is larger than P, meaning that there is a loss of individuals in this case through exploitation. If the population size is smaller than P, however, then the "exploitation" term actually results in an influx of individuals into the population. Consequently, this model represents a situation in which there is exploitation of the population provided that the population size is large enough, but there is stocking of the population if the population size decreases below the critical value P. As with most models, however, there are also other possible interpretations or applications. For example, if we multiply out the exploitation term in (3.20c), we obtain

$$\frac{dn}{dt} = r\,n(t)\left(1 - \frac{n(t)}{K}\right) - H\,n(t) + H\,P. \tag{3.21}$$

This result suggests an alternative interpretation of equation (3.20c): the population is subject to a constant per capita harvesting rate of H and is being stocked at a constant total rate of $H\,P$, where the amount of harvesting depends on the population size, but the stocking does not.

As one final example, consider a model of a sexually transmitted disease that tracks the number of susceptible, S, and infected, I, individuals within a sexually active adult population. The rate at which an infected individual transmits the disease is equal to the rate of unprotected sexual contact with all other individuals, c, times the probability that the partner is susceptible, $S/(S + I)$, times the probability that sex between an infected and susceptible individual results in disease transmission, a (see section 3.5). In addition, new individuals are added to the population at a constant total rate b (via immigration or maturation) and are removed from the population at a constant per capita rate d (via emigration or death). The following equations might be used:

$$\frac{dS}{dt} = -a\,c\,I\frac{S}{S + I} + b - d\,S, \tag{3.22a}$$

$$\frac{dI}{dt} = a\,c\,I\frac{S}{S + I} - d\,I. \tag{3.22b}$$

In our verbal description of this model, we did not specify whether the new arrivals are disease-free or infected (or a mixture of the two). The differential equations do, however, indicate what assumption was made about the disease status of immigrants. By being able to read and interpret the equations, we can identify this assumption. The rate at which new individuals are added to the population was said to equal b, and we can see from equations (3.22) that this term affects only the differential equation for S. As a result, we can conclude that all new arrivals are assumed to be susceptible. Different assumptions about the pool of immigrants would generate different equations (Problem 3.18).

3.7 Concluding Message

In this chapter we introduced classic models in ecology, evolution, and epidemiology (Table 3.1). As we have witnessed, models often involve generalizations of, or alterations to, previous models. We hope that you are gathering a sense of the interconnectedness of many dynamical models. This interconnectedness makes it easier to develop new models because you can use a previous model as a foundation upon which to build. As we shall see in later chapters (especially Chapters 6 and 9), this interconnectedness also makes it easier to analyze new models, by drawing upon known results from previous models. We turn next to methods for analyzing the equations that we have developed in this chapter.

Problems

Problem 3.1: In equation (3.1), it was assumed that parents first give birth and then all individuals (parents and offspring alike) have a probability d of dying. (a) How will the number of surviving individuals per parent, R, differ if deaths happen first and then births? (b) What will the number of surviving individuals per parent, R, equal if births happen first but then all parents (but no offspring) die, so that the model describes a population with nonoverlapping generations?

Problem 3.2: In the logistic model, we assumed that the number of surviving offspring per parent, R, declines linearly with population size n. Show that growth is still described by equation (3.5a) if we assume instead that (a) only the birth rate b depends linearly on n or (b) only the death rate d depends linearly on n. (c) Show, however, that if both b and d depend linearly on n, then $R(n)$ is not a linear function of n. [Recall the definition $R = (1 - d)(1 + b)$ and assume that $R(0) = 1 + r$ and $R(K) = 1$.]

Problem 3.3: According to the recursion equation (3.5a) for the logistic model,

$$n(t+1) = n(t) + r\,n(t)\left(1 - \frac{n(t)}{K}\right),$$

it is possible for $n(t + 1)$ to be negative even if $n(t)$, r, and K are all positive. (a) Solve for $n(1)$ by hand using $r = 1$ and $K = 100$ starting from the population sizes $n(0) = 50$, 100, 200, and 500. (b) By rearranging the recursion equation, determine the population size $n(t)$ above which $n(t + 1)$ becomes negative and the population goes extinct. That is, find n^* in terms of r and K such that $n(t + 1) < 0$ whenever $n(t) > n^*$. Check that your answer to part (b) is consistent with your answer to part (a).

Problem 3.4: Many alternatives to the logistic equation have been described, each of which incorporates different assumptions about how density affects the per capita growth term $R(n)$. (a) Write the recursion and difference equations for n under the assumption that $R(n)$ decreases exponentially from $1 + r$ as the population size increases, so that $R(n)$ can be written as $(1 + r)$

$e^{-an(t)}$. (b) According to the recursion equation that you derive, is it possible for $n(t + 1)$ to be negative if $n(t)$ is positive (justify your answer)? (This model is known as the Ricker model.)

Problem 3.5: In the equations for logistic growth, (3.5), if a population has a high intrinsic growth rate r and grows rapidly when the population is very small, then it must also decline rapidly when the population is very large and above the carrying capacity. (a) Use equation (3.5) to prove this assertion for specific choices of r, K, and n. (b) Describe how you might generalize the logistic model so that a species that grows rapidly does not necessarily decline rapidly as well. (c) Illustrate your argument with an appropriate differential equation.

Problem 3.6: In the derivation of the haploid model of selection, we let the number of surviving individuals in the next generation equal W_A for allele A and W_a for allele a. Consider, instead, the case where these alleles alter the *growth rate* r by a factor, W_A for allele A and W_a for allele a. That is, the number of each allele changes over time according to

$$n_A(t + 1) = (1 + W_A r)\, n_A(t),$$
$$n_a(t + 1) = (1 + W_a r)\, n_a(t).$$

(a) Derive the recursion equation for the allele frequency $p(t + 1)$ as a function of $p(t)$ for this model. (b) Can r be factored out of this recursion equation? (c) Show that if we measure selection by the constant $s = (W_A r - W_a r)/(1 + W_a r)$, i.e., as the difference in growth rates divided by the absolute fitness of the a allele, we regain the recursion equation $p(t + 1) = p(t) + \Delta p$, where Δp is given by equation (3.10).

Problem 3.7: In the diploid model of selection in the text, we assumed that gametes unite at random to produce zygotes (Table 3.2). A more realistic model for many animal populations is that diploid individuals, not their gametes, mate at random. In a discrete-time model, let the frequency of diploid zygotes at time t be $d(t)$, $h(t)$, and $r(t)$ for AA, Aa, and aa individuals, respectively. Similarly, let W_{AA}, W_{Aa}, and W_{aa} equal the relative fitnesses of the three genotypes. Accounting for such fitness differences, the mating table with random mating among diploid adults is given by Table 3.4. Here, we have combined all matings involving the same genotypes into the same row; for example, the second row combines matings involving an AA female × Aa male with matings involving an Aa female × AA male. (a) Complete the mating table by filling in the proportion of each offspring genotype produced by each mating pair. (b) Calculate recursion equations for $d(t)$, $h(t)$, and $r(t)$ by taking the product of the "Frequency of mating pair" column with each "Offspring" column, in turn, and summing across rows. (c) Use these recursions to prove that Hardy-Weinberg proportions are achieved in the next generation. Specifically, show that $d(t + 1) = p'(t)^2$, $h(t + 1) = 2\, p'(t)\, q'(t)$, and $r(t + 1) = q'(t)^2$ where $p'(t) = (d(t)w_{AA} + \frac{1}{2}h(t)w_{Aa})/\overline{W}$ and $q'(t) = (\frac{1}{2}h(t)w_{Aa} + r(t)w_{aa})/\overline{W}$ are the frequencies of the A and a alleles among the surviving parents (after selection).

TABLE 3.4
Genotype frequencies in the diploid model of natural selection with random mating among adults

Mating pair	Frequency of mating pair (after selection)	Offspring produced		
		AA	*Aa*	*aa*
$AA \times AA$	$\left(d(t)\frac{W_{AA}}{\overline{W}}\right)\left(d(t)\frac{W_{AA}}{\overline{W}}\right)$			
$AA \times Aa$	$2\left(d(t)\frac{W_{AA}}{\overline{W}}\right)\left(h(t)\frac{W_{Aa}}{\overline{W}}\right)$			
$AA \times aa$	$2\left(d(t)\frac{W_{AA}}{\overline{W}}\right)\left(r(t)\frac{W_{aa}}{\overline{W}}\right)$			
$Aa \times Aa$	$\left(h(t)\frac{W_{Aa}}{\overline{W}}\right)\left(h(t)\frac{W_{Aa}}{\overline{W}}\right)$			
$Aa \times aa$	$2\left(h(t)\frac{W_{Aa}}{\overline{W}}\right)\left(r(t)\frac{W_{aa}}{\overline{W}}\right)$			
$aa \times aa$	$\left(r(t)\frac{W_{aa}}{\overline{W}}\right)\left(r(t)\frac{W_{aa}}{\overline{W}}\right)$			

Problem 3.8: (a) Show that you can rewrite $\overline{W}$ using the composite fitness terms $W_{A*} = p(t)W_{AA} + q(t)W_{Aa}$ and $W_{a*} = p(t)W_{Aa} + q(t)W_{aa}$. (b) Show that the diploid recursion equation (3.13a) can be written in a form equivalent to the haploid recursion equation (3.8c) with W_{A*} and W_{a*} taking the places of W_A and W_a. (c) What do W_{A*} and W_{a*} represent?

Problem 3.9: Show that if the fitness of a diploid individual is the product of the fitness effects of each of its alleles (i.e., $W_{AA} = W_A W_A$, $W_{Aa} = W_A W_a$, and $W_{aa} = W_a\ W_a$) then the recursion equation (3.13) for natural selection in diploids reduces to that observed in the haploid model (3.8c). To accomplish the same change in allele frequency per generation as the haploid model, the diploid model requires that fitness be equivalent to two rounds of haploid selection (e.g., $W_A W_A$). Thus, selection is half as effective in diploid organisms as in haploid organisms (see Crow and Kimura 1970 for the more general case with arbitrary heterozygous fitness).

Problem 3.10: [Challenging] We described the Lotka-Volterra model of competition, equation (3.14), assuming that the two competing entities were species. But the same model can be applied to the case where $n_1(t)$ and $n_2(t)$ are the numbers of two different geno-types within a single species (e.g., the numbers of haploids carrying alleles *A* and *a*, respectively). (a) Using equation (3.14), obtain recursions for the frequency of allele *A*, $p(t)$, and the total population size $n(t)$. (Write these recursions in terms of the variables $p(t)$ and $n(t)$, only.) (b) By equating your recursion for $p(t + 1)$ to equation (3.8c),

$$p(t\ +\ 1)\ =\ \frac{W_A\, p(t)}{W_A\, p(t)\ +\ W_a\,(1\ -\ p(t))},$$

specify what the relative fitnesses W_A and W_a equal in this model. (c) Show that, even after simplification, these relative fitnesses depend on both the allele frequencies and the population size. (d) For what special values of the parameters will the relative fitnesses be independent of the allele frequency?

Problem 3.11: The consumer-resource model described by equations (3.16) is very general. Show that it also describes the Lotka-Volterra model of competition by determining the forms of the functions $f(n_1)$, $g(n_1,n_2)$, and $h(n_2)$ that cause equations (3.16) to reduce to equations (3.15).

Problem 3.12: The Lotka-Volterra model of predator-prey dynamics discussed in the text assumes a single prey species and a single predator species, but it can be readily generalized to include more species. Figure 3.11 illustrates a slightly more complex food web, with two species of prey, n_0 and n_1. Derive the differential equations for this two-prey, single-predator Lotka-Volterra model.

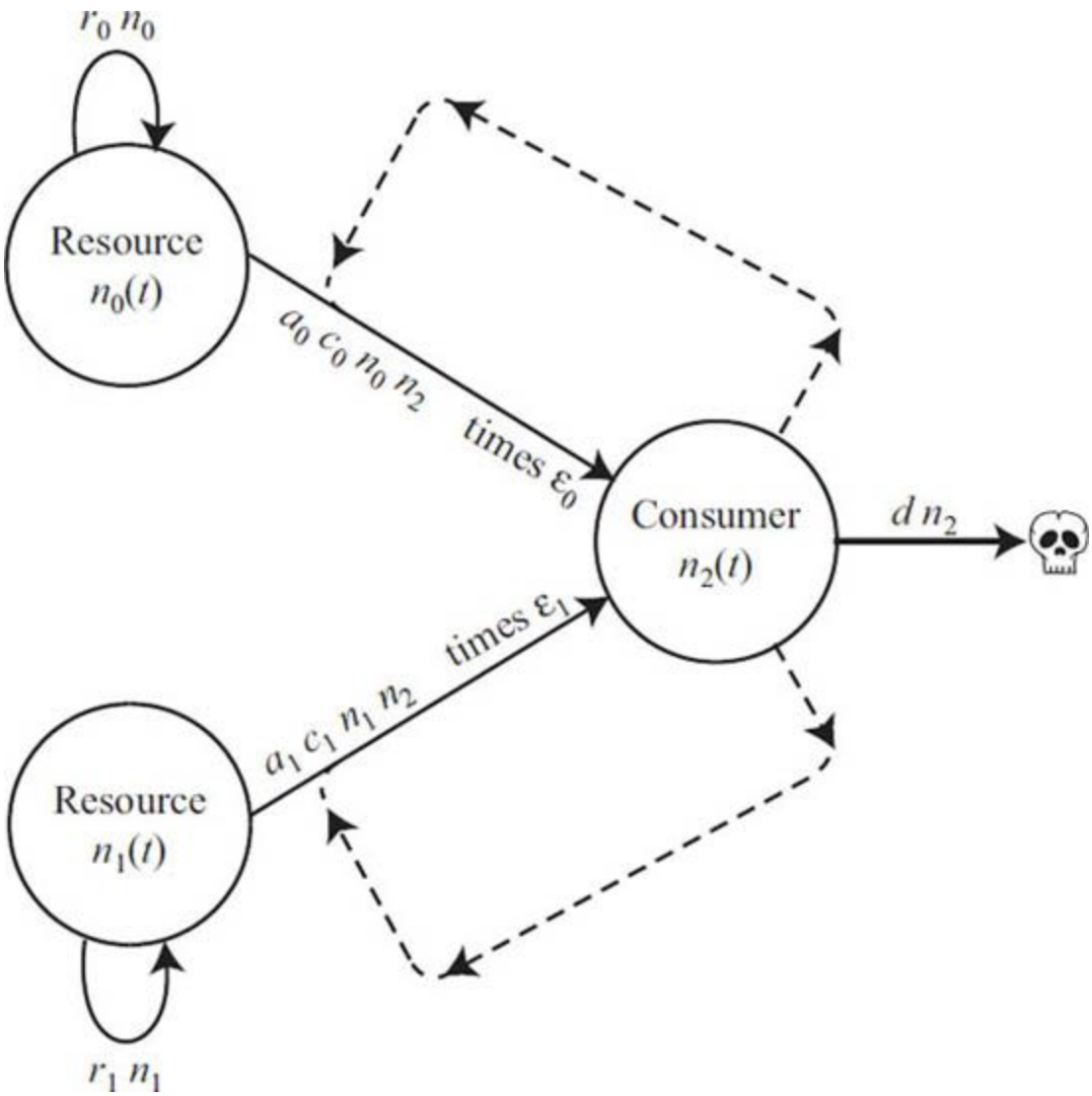

Figure 3.11: A flow diagram for the Lotka-Volterra predator-prey model with two prey species.

Problem 3.13: Head lice are readily spread among children, especially in day care and school settings. Individuals do not develop resistance to head lice even after infection; instead, following treatment, they become susceptible again. (a) Modify the flow diagram for the SIR model (Figure 3.9) to describe the spread of lice in a population of susceptible, $S(t)$, and infected, $I(t)$, children. (b) Write down differential equations for $S(t)$ and $I(t)$.

Problem 3.14: Mating within a population is often not random. If a population is spatially structured or if kin prefer to mate with one another, then gametes carrying the same genes unite more often than one would expect under random mating. This is because the parents are related to one another and therefore are likely to carry the same genes, a phenomenon known as *inbreeding*. The simplest way to account for inbreeding in population-genetics models is to say that each egg has a chance $1 - f$ of being fertilized by a sperm drawn randomly from the gamete pool and a chance f of being fertilized by a sperm carrying the same allele. Convince yourself that the probability of producing an AA zygote will then equal $(1 - f)\,p(t)^2 + f\,p(t)$. (a) Calculate the probability of producing an Aa zygote and an aa zygote. (b) How much less common are heterozygous zygotes than expected at Hardy-Weinberg? Using these results, modify the second through fourth columns of Table 3.2, making sure that the second column sums to one and that the fourth column sums to one. (c) Based on this modified table, derive the recursion equation for $p(t + 1)$, generalizing (3.13a) to the case of nonrandom mating. (d) Using this recursion, prove that the allele frequency does not change over time if the genotypes are equally fit ($W_{AA} = W_{Aa} = W_{aa} = 1$).

Problem 3.15: In Chapter 2, we developed a model of the spread of the flu described by the recursion equation (2.11):

$$n(t + 1) = n(t) + a\,c\,n(t)\,(N - n(t)),$$

where a, c, and N are constants (infection probability, per capita contact number per time step, and total population size, respectively). This equation is mathematically equivalent to the logistic equation (3.5a). That is, one can write r and K as functions of a, c, and N that when plugged into the logistic equation (3.5a) give (2.11). What must these functions for r and K be?

Problem 3.16: Defining $W_{AA} = 1 + s$, $W_{Aa} = 1 + hs$, and $W_{aa} = 1$, show that the recursion equation (3.13a) can be used to obtain the following continuous-time version of the diploid model of selection:

$$\frac{dp}{dt} = s\,p(1 - p)\,(p + h(1 - 2p)),$$

by shrinking the time step over which changes are observed (see Box 2.6 and the derivation of equation (3.11)). Check that this differential equation is equivalent to

$$\frac{dp}{dt} = p(1 - p)\,(p(W_{AA} - W_{Aa}) + (1 - p)(W_{Aa} - W_{aa})).$$

Problem 3.17: Suppose the population size at time t is $N(t)$, and that the absolute fitnesses $W_{ij}(N(t))$ are arbitrary functions of the population size at time t. Typically we expect absolute fitness to decrease as the population size gets large enough, and it might do so in different ways for different genotypes. (a) Starting with the numbers of each genotype, show that the dynamics of

the population size and the frequency of allele A obey the following pair of recursion equations:

$$N(t+1) = \overline{W}(N(t), p(t))\, N(t),$$

$$p(t+1) = \frac{p(t)\, W_{AA}(N(t)) + (1-p(t))\, W_{Aa}(N(t))}{\overline{W}(N(t),p(t))} p(t).$$

Assume that gametes unite at random, so that the population of newly formed zygotes is in Hardy-Weinberg proportions at time t. (b) What must be true about the form of the fitnesses $W_{ij}(N(t))$ so that the population size drops out of the recursion equation for the allele frequency? [Hint: If you get stuck, read through Sup. Mat. 3.1.]

Problem 3.18: (a) Modify equations (3.22) to reflect an assumption that a fraction f of individuals immigrating into the population already carry the sexually transmitted disease. (b) Describe one way to check your differential equations using equations (3.22).

Problem 3.19: Use the fact that the derivative of a function equals zero at a maximum (Appendix 2) to determine the allele frequency p that maximizes the rate of evolutionary change $dp/dt = s\,p\,(1-p)$ in the haploid model of selection (see Figure 3.6).

Further Reading

For more detailed information about classic models in ecology and evolution, consult

- Case, T. J. 2000. *An Illustrated Guide to Theoretical Ecology*. Oxford University Press, Oxford.
- Roughgarden, J. 1996. *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. Prentice-Hall, Upper Saddle River, N.J.

References

Blower, S. M., H. B. Gershengorn, and R.M. Grant. 2000. A tale of two futures: HIV and antiretroviral therapy in San Francisco. *Science* 287:650–654.

Crow, J. F., and M. Kimura. 1970. *An Introduction to Population Genetics Theory*. Harper & Row, New York.

Fussmann, G. F., S. P. Ellner, K. W. Shertzer, and N. G. J. Hairston. 2000. Crossing the Hopf bifurcation in a live predator-prey system. *Science* 290:1358–1360.

Krebs, C. J., S. Boutin, R. Boonstra, A.R.E. Sinclair, J.N.M. Smith, M.R.T. Dale, K. Nartin, and R. Turkington. 1995. Impact of food and predation on the snowshoe hare cycle. *Science* 269:1112–1115.

Lack, D. L. 1954. *The Natural Regulation of Animal Numbers*. Clarendon Press, Oxford.

Mable, B. K., and S. P. Otto. 2001. Masking and purging mutations following EMS treatment in haploid, diploid and tetraploid yeast (*Saccharomyces cerevisiae*). *Genet. Res*. 77:9–26.

McCallum, H., N. Barlow, and J. Hone. 2001. How should pathogen transmission be modeled? *Trends Ecol. Evol*. 16:295–300.