

4

Count-Based PVA: Incorporating Density Dependence, Demographic Stochasticity, Correlated Environments, Catastrophes, and Bonanzas

The count-based population viability analysis that we outlined in Chapter 3 was based on the simplest possible stochastic population model (Equation 3.1) and on measures of extinction risk derived using a diffusion approximation. As we noted at the end of Chapter 3, the model and the diffusion approximation rest on the following assumptions:

1. The population growth rate is unaffected by population density.
2. The only source of variability in the population growth rate is environmental stochasticity (in particular, we neglected demographic stochasticity).
3. There are no trends in the mean and variance of the population growth rate over time.
4. Environmental conditions in successive years are uncorrelated; hence the population growth rate in a given year does not depend on whether the previous year was good or bad.
5. Environmental variability is moderate; that is, there are no catastrophes or bonanzas.
6. The measured variability in population growth rates is real, rather than a result of observation error.

In reality, virtually every population is certain to violate one or more of these assumptions. The population growth rate will surely change at some point as the size of the population changes, and its mean or variance may show a temporal trend due to a changing environment or the impact of manage-

ment. The growth rate of very small populations will vary due to both environmental and demographic stochasticity, as described in Chapter 2. Cycles in environmental conditions that last longer than one year, such as El Niño oscillations, may cause the environmental effects on the population growth rate to be correlated in two or more successive years, resulting in strings of good or bad years that can impact the viability of a population. Large fluctuations that are outside the range of normal environmental variation may occur due to rare floods, hurricanes, droughts, or other infrequent occurrences. Finally, some variability in observed growth rates is certainly due to imperfect measurements.

Unfortunately, even if we know that such violations are likely to occur, it is an inescapable fact that incorporating these effects into PVA models requires more and better data, which we will often lack for rare species. Hence we may have little choice but to use simpler methods and models, with the consolation that we can often understand the direction in which violations of the above assumptions will cause our viability estimates to err (as we discussed at the end of Chapter 2 and also in Chapter 3). However, when a greater quantity and quality of data are available and support the importance of these other factors, alternative models that avoid one or more of the assumptions of the simplest count-based methods can be employed. In this chapter, we explore some of these models, including ways to incorporate density dependence, demographic stochasticity, correlated environments, and extreme events such as bonanzas and catastrophes. When discussing density-dependent models, we deal in depth with the important topic of choosing among alternative models, and the tools we use here will reappear throughout the book. In Chapter 5, we deal with the problem of observation error and how to disentangle it from true variability in growth rates. As a caveat, we reiterate that the material in these two chapters is both more mathematically complex and less likely to be generally applicable (because the appropriate data aren't available) than are the methods presented in Chapter 3. Hence readers may wish to reserve this chapter for more in-depth study during a second pass through the book.

Density Dependence

No population can continue to grow indefinitely. As a population approaches the limits of its resources, or as the impacts of its natural enemies intensify, the annual population growth rate (N_{t+1}/N_t) will approach 1.0, so that population growth ceases. As we noted in Chapter 2, such negative density dependence is an important factor to consider in assessing population viability, for two opposing reasons. On one hand, negative density dependence places a cap on how far a population can move away from the extinction threshold, even if the population tends to grow when its size is small. The closer to the threshold a population is held, the greater is the chance that a

string of bad years will cause it to actually hit the threshold. On the other hand, if a population is experiencing negative density dependence, the growth rates we measure at the current population size might *underestimate* the growth rates the population would experience if it were reduced to a smaller size. That is, the population might be more buffered against extinction than growth rates measured at intermediate densities would imply.

However, we have also seen that the population growth rate could *decline* as population size decreases, due to positive density dependence (Allee effects). Lower population growth results from either a depressed birth rate or an elevated death rate at low density. Low birth rates may be caused by an inadequate number of potential mates within an individual's neighborhood, whereas high death rates can result from failure to achieve an adequate group size to fend off the attacks of predators. Both effects could result from reduced foraging success of small groups. Finally, inbreeding depression could underlie a reduction in growth rate at low population size. Clearly, Allee effects, if they occur, will be important to include in PVAs, because they will increase the extinction risk of small populations.

In the following sections, we discuss ways to account for density dependence in count-based PVAs. The bulk of our attention is devoted to negative density dependence because evidence for it is more widespread and hence it has been incorporated into count-based PVAs far more often than have Allee effects. Nevertheless, we briefly indicate how to include Allee effects as well.

Two Population Models with Negative Density Dependence

In this section, we describe two discrete-time population models that have played an important role in understanding and assessing the viability of populations exhibiting negative density dependence. In the following section, we review general theoretical insights that the first (and simpler) of these two models yields regarding the influence of negative density dependence on population viability. Following this, we present a case study of a PVA that makes use of the second (more realistic) model.

Perhaps the simplest way to incorporate negative density dependence is to introduce a population ceiling to the density-independent population growth model of Equation 3.1. Specifically, let

$$N_{t+1} = \begin{cases} \lambda_t N_t & \text{if } \lambda_t N_t \leq K \\ K & \text{if } \lambda_t N_t > K \end{cases} \quad (4.1)$$

We shall refer to Equation 4.1 as the *ceiling model*. As in Chapter 3, N_t is the size of the population in year t , and the population growth rate λ_t is assumed to vary from year to year due to environmental stochasticity alone. The parameter K is the population ceiling. In Equation 4.1, the population's growth rate does not depend on its size as long as it remains below the ceil-

ing; for any increment of growth that would take the population above the ceiling, we simply reset population size to K . Note that K is somewhat different from the carrying capacity in the familiar logistic equation, because in Equation 4.1 there is an abrupt rather than a gradual decline in the growth rate as population size increases. Nevertheless, many authors refer to K in Equation 4.1 as the “carrying capacity” to reflect the fact that population size is forced to remain within some bounds (in contrast to populations governed by Equation 3.1).

As we shall see in the next section, the mathematical simplicity of the ceiling model allows some clear insights into the viability of negatively density-dependent populations. Nevertheless, for many populations, the growth rate will change gradually rather than abruptly as population size increases. An alternative model that allows for a gradually changing growth rate is the *theta logistic model*:

$$N_{t+1} = N_t \exp \left\{ r \left[1 - \left(\frac{N_t}{K} \right)^\theta \right] + \varepsilon_t \right\} \quad (4.2)$$

(modified from Gilpin and Ayala 1973), variants of which have served as the basis for several published PVAs that incorporated negative density dependence (Saether et al. 1998a, 2000b). The mean population growth rate is e^r when N_t is very small relative to the carrying capacity K . Unlike the other models we've covered, environmental variability in the population growth rate is not represented by variance in the usual growth parameters themselves (e.g., λ or r) but by a separate variable, ε_t (the Greek letter epsilon), which has a mean of zero and a variance that reflects the degree of environmental stochasticity. The parameter θ (the Greek letter theta) determines how the population growth rate declines as population size increases. As shown in Figure 4.1, the advantage of the theta logistic model is its ability to describe several different patterns of negative density dependence depending upon the value of θ . In particular, if θ is large, the population growth rate remains high until population size gets very close to K , at which point the growth rate drops off precipitously. Thus the theta logistic model approaches the ceiling model as θ approaches infinity. The well-known Ricker model¹ (Ricker 1954), which has seen much use in fisheries biology, is a special case of the theta logistic model obtained by setting θ equal to 1. Note also that as the carrying capacity K goes to infinity, Equation 4.2 becomes $N_{t+1} = \lambda_t N_t$, where $\lambda_t = \exp\{r + \varepsilon_t\}$. That is, as density dependence gets weaker and weaker (as the carrying capacity gets higher and higher), Equation 4.2 reverts to the density-independent model of Equation 3.1.

¹In some sense, the Ricker model can be thought of as a discrete-time version of the familiar continuous-time logistic model.

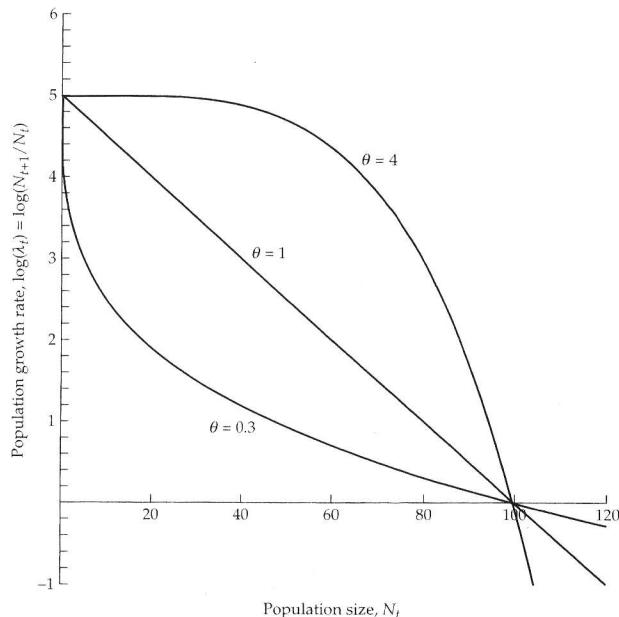


Figure 4.1 The parameter θ (theta) in the theta logistic model (Equation 4.2) governs the shape of negative density dependence. Specifically, the log of the population growth rate $\log(N_{t+1}/N_t)$ declines most steeply at low densities if $\theta < 1$, linearly with density if $\theta = 1$, or most steeply at high densities if $\theta > 1$. For all three values of θ , the carrying capacity, K , equals 100 and the log population growth rate at low density, r , equals 5.

Extinction Risk Predicted by the Ceiling Model

In Chapters 2 and 3, we discussed several metrics of extinction risk, including the geometric mean population growth rate, the mean time to reach the extinction threshold, and the probability of reaching the threshold by a given future time horizon. In the ceiling model, just as for the density-independent model (Equation 3.1), any population with a geometric mean growth rate less than 1.0 is doomed to eventual extinction. However, the two models differ in an important way when the geometric mean growth rate is greater than 1.0. Whereas the density-independent model predicts (unrealistically) that some population trajectories could grow to infinite size (so that the probability of ultimate extinction is less than 1; see Equation 3.6), the presence of an upper limit to population size in the ceiling model (or, for that

matter, in any other model with negative density dependence) causes *all* trajectories to hit the extinction threshold *eventually*. (That is, the probability of ultimate extinction is *always* 1.) Sooner or later, a sufficient number of bad years will occur in succession to drive the population down to the extinction threshold, regardless of the height of the ceiling. However, such a sequence of years may take a *very* long time to occur by chance if the population tends to hover near the ceiling and the ceiling itself is high.

To see exactly *how* long, we can examine another viability metric, the mean time to extinction. The principal advantage that the ceiling model possesses over other models with negative density dependence is its mathematical simplicity, which makes it possible to calculate an analytic expression for the mean time to reach the extinction threshold from any initial density.² We caution that the same limitations of the mean time to extinction that we discussed in Chapter 2 also apply to the ceiling model, namely that the probability distribution of extinction times is highly skewed so that the mean is not very representative of the behavior of most possible outcomes. Nevertheless, other measures of extinction risk are not easy to calculate for this model, so we use the mean extinction time to gauge how the *relative* extinction risk depends on the height of the ceiling and the values of μ and σ^2 , the mean and the variance of the log population growth rates, $\log \lambda_t$. (Note that μ and σ^2 are exactly the same parameters as in Chapter 3). If the variation in the λ_t 's is not too large, the ceiling model can be approximated by a diffusion model with a reflecting upper boundary at the ceiling (in addition to the absorbing lower boundary at the extinction threshold; see Chapter 3). From this diffusion approximation, an expression for the mean time to reach the quasi-extinction threshold N_x from a current population size of N_c is:

$$\bar{T} = \frac{1}{2\mu c} \left[e^{2ck} \left(1 - e^{-2cd} \right) - 2cd \right] \quad (4.3)$$

where $c = \mu/\sigma^2$, $d = \log(N_c/N_x)$, and $k = \log(K/N_x)$ (for details, see Lande 1993, Foley 1994, Middleton et al. 1995). If the population starts at the ceiling (i.e., $N_c = K$) and $N_x = 1$, then Equation 4.3 simplifies to

$$\bar{T} = \frac{1}{2\mu c} \left[K^{2c} - 1 - 2c \log K \right] \quad (4.4)$$

(Lande 1993). A MATLAB program to compute \bar{T} using Equations 4.3 and 4.4 is given in Box 4.1.

²Mangel and Tier (1993) developed a numerical method to calculate the mean and variance of the time to extinction for any density-dependent model, but their method does not yield a simple expression for the mean extinction time, such as Equation 4.3.

BOX 4.1 MATLAB code to plot the mean time to extinction for the ceiling model (Equation 4.1) as functions of the carrying capacity and initial population size.

```
% PROGRAM tbar_ceiling
% Plots mean time to extinction vs. K and starting population
% size for the ceiling model, using expressions for the mean
% time to extinction from Lande, Am. Nat. 142:911-927 (1993),
% Foley, Conservation Biology 1:124-137 (1994), and
% Middleton et al., Theor. Pop. Biol. 48:277-305 (1995).

mu=0.1; % mean log population growth rate
s2=0.1; % variance of log population growth rate
c=mu/s2;

% Plot mean time to extinction (N=1) vs. K for populations
% starting at K, according to Equation 4.4
K=[0:50:1000]; % vector of carrying capacities
K(1)=1; % set smallest K to 1
k=log(K);
Tbar=(K^(2*c)-1-2*c*k)/(2*mu*c); % Eq. 4.4

subplot(2,1,1); % upper of 2 plots on same page
plot(K,Tbar) % plot Tbar vs K
xlabel('Carrying capacity, K','FontSize',12)
ylabel('Mean time to extinction, Tbar','FontSize',12)
title('Populations starting at K','FontSize',14)

% Plot mean time to extinction (N=1) vs. Nc for populations
% starting at various Nc<=K, according to Equation 4.3
K=200; % carrying capacity
Nc=[0:10:K]; % vector of initial population sizes
Nc(1)=1; % set smallest Nc to 1
k=log(K);
d=log(Nc);
Tbar=(exp(2*c*k)*(1-exp(-2*c*d))-2*c*d)/(2*mu*c); % Eq. 4.3

subplot(2,1,2); % lower of 2 plots on same page
plot(Nc,Tbar) % plot Tbar vs Nc
xlabel('Initial population size, Nc','FontSize',12)
ylabel('Mean time to extinction, Tbar','FontSize',12)
title('Populations starting below K','FontSize',14)
```

Equation 4.4 is useful for exploring how increasing the “height” of the ceiling (for example by increasing the size of a reserve and therefore increasing its carrying capacity) affects the mean time to reach the extinction threshold. The shape of the curve relating mean extinction time to the carrying capacity depends upon the parameters μ and σ^2 (Lande 1993). From a conservation standpoint, we would ideally like the mean time to extinction to increase faster than linearly, so that we get more and more protection from extinction for each increment in reserve size. However, when μ is equal to or less than zero, the mean time to extinction always increases in a slower-than-linear fashion with an increase in the carrying capacity, as shown in Figure 4.2A. This “law of diminishing returns” makes intuitive sense. As we saw in Chapter 3, a negative value of μ (equivalent to $\lambda_G < 1.0$) implies that the population will decline over the long term, and we cannot very effectively delay the inevitable crossing of the extinction threshold simply by raising the ceiling. In contrast, if μ is sufficiently positive, growth of the population will tend to quickly push it back up to the ceiling following a bad year, so that raising the ceiling widens the gap between the zone of sizes in which the population will typically be found (that is, close to the ceiling) and the extinction threshold, thus making extinction less likely in the short term.

However, even if μ is positive, increasing the environmental variance σ^2 can cause the mean extinction time curve as K becomes large to go from faster-than-linear to linear and finally to slower-than-linear (Figure 4.2B; note that when σ^2 exactly equals 2μ , the curve becomes perfectly linear as K increases). Again, this makes sense given what we know from Chapter 3 about the effects of environmental variation. The more severe are the extremes of the population growth rate, the more likely is a string of extremely bad years that could drive the population to extinction, even if the ceiling is high. Thus the expression for the mean time to extinction in the ceiling model tells us that management efforts aimed at increasing the carrying capacity (e.g., by expanding the size of a reserve) are destined to fail if the population is declining over the long term or if it exhibits very high inter-annual variation in growth. Identifying strategies to increase the growth rate or ameliorate causes of variation in the growth rate may be a more productive management goal for such populations. On the other hand, if long-term population growth is relatively high and environmental variation low, increasing the carrying capacity can dramatically enhance a population’s protection from extinction (Figure 4.2B).³

In Chapters 2 and 3, we saw the value of viability metrics based on the extinction probability, particularly the cumulative distribution function (CDF)

³An important assumption of the ceiling model is that μ and σ^2 are independent not only of population size but also of the height of the ceiling. If increasing the size of a reserve both “raises” the ceiling and improves μ or σ^2 (e.g., by reducing edge effects or by buffering environmental stochasticity), then it may have substantial benefits even when μ is initially negative or σ^2 initially large.

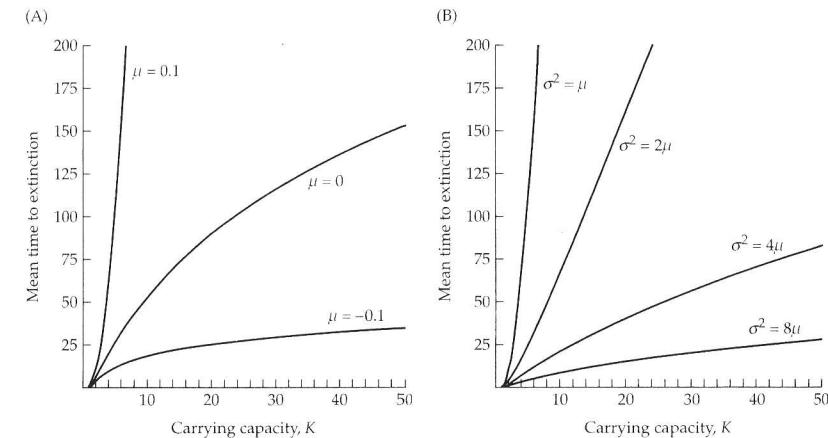


Figure 4.2 The mean time to extinction given by Equation 4.4 for populations starting at K in the ceiling model (Equation 4.1) with an extinction threshold $N_x = 1$. In panel A, σ^2 always equals 0.1, but μ changes from positive to zero to negative values. In panel B, μ is always 0.1, but σ^2 increases.

for the time to extinction (Equation 3.5). Unfortunately, there is no equivalent analytical expression for the extinction probabilities for any density-dependent model, including the relatively simple ceiling model.⁴ Nevertheless, a straightforward alternative to an analytical formula is to estimate the extinction time CDF using computer simulations. The idea is to simulate many random trajectories that a population might take (as we did to create the 20 paths shown in Figure 3.1, but typically for tens or hundreds of thousands of trajectories), seeing when each becomes extinct (if ever), and using these results to build a distribution of predicted extinction times.

⁴Middleton et al. (1995) derived a solution to the diffusion approximation for the ceiling model using a Fourier series, which technically involves the sum of an infinite number of terms. In principle, one could (as did Middleton and colleagues for particular combinations of μ , σ^2 , and K) use the Fourier series solution for the ceiling model to calculate a close approximation to the probability of extinction by time T , but the procedure required to do so is technically involved. (One needs to solve a set of implicit equations, substitute the results into the Fourier series, integrate the series term by term, and sum the resulting terms until the sum converges to a nearly constant value.) In practice, and given the peculiar way that density dependence is represented in the ceiling model, such a procedure is probably not worth the effort for specific PVA applications.

However, if we go to the work to make such a density-dependent simulation, we generally won't want to use the ceiling model. In the ceiling model, the assumption that population growth below the ceiling is density-independent was made primarily to allow the expressions for the mean time to extinction (Equations 4.3 and 4.4) to be calculated. Although this exercise helped give insight into how the existence of a ceiling affects population viability, the unusually sharp onset of negative density dependence in the ceiling model is probably not realistic for actual populations. Fortunately, with adequate data to estimate density-dependent effects, we can relax this confining stricture when we use computer simulations, as we demonstrate in the following section.

PVA for a Population with Negative Density Dependence: The Bay Checkerspot Butterfly

In this section, we demonstrate by means of a worked example how to use a more realistic negatively density-dependent model, in combination with computer simulations, to calculate the probability of extinction. The example we use is the Bay checkerspot butterfly, *Euphydryas editha bayensis*, which occurs in small populations restricted to serpentine outcrops in the San Francisco Bay area of California. We begin with the JRC population, one of three populations that were the subjects of a long-term population study conducted by Paul Ehrlich and his colleagues at Stanford University. The number of female butterflies in the population was estimated by capture-recapture methods each year from 1960 to 1986. The 27 consecutive estimates of population size are given in Table 4.1 (reproduced from Harrison et al. 1991) and shown in Figure 4.3. Note that the population size estimates undergo large fluctuations across the 27 years. There are two phases to conducting a density-dependent PVA with this kind of data. First we need to find the best model to describe the population's dynamics, and then we need to simulate the model to predict population viability.

TESTING FOR DENSITY DEPENDENCE AND IDENTIFYING THE BEST MODEL As we did for the Yellowstone grizzly bear population in Chapter 3, we can get a quick idea of whether the census data show evidence of negative density dependence by looking for a negative relationship between the log population growth rate and population size. Such a negative relationship does in fact appear likely for the JRC Bay checkerspot population (Figure 4.4), although there is clearly a lot of variability in the growth rate due to environmental variation, demographic stochasticity, and/or observation error. But is there *statistical* support for the presence of negative density dependence in the checkerspot data? And if so, what shape of the density dependence function (Figure 4.1) would best describe the data? We can answer both of these questions by fitting a suite of population models to the counts in Table 4.1 and determining which model best fits the data.

TABLE 4.1 Census data for two populations of the Bay checkerspot butterfly^a

Year, t	JRC population		JRH population	
	Estimated number of female butterflies, N_t	Log population growth rate, $\log(N_{t+1}/N_t)$	Estimated number of female butterflies, N_t	Log population growth rate, $\log(N_{t+1}/N_t)$
1960	90	0.6650	70	1.6094
1961	175	-1.4759	350	0.7621
1962	40	0.1178	750	0.0000
1963	45	1.3581	750	0.6242
1964	175	0.1335	1400	0.3567
1965	200	0.7538	2000	-0.1335
1966	425	0.0000	1750	-0.6650
1967	425	0.6325	900	-0.4463
1968	800	-1.1394	576	0.4135
1969	256	1.0243	871	-0.0603
1970	713	-1.2812	820	-1.2497
1971	198	2.2178	235	1.5871
1972	1819	-1.1517	1149	-1.1331
1973	575	-0.0140	370	-0.7374
1974	567	1.1657	177	0.5828
1975	1819	1.3795	317	1.1499
1976	7227	-2.1380	1001	-1.6617
1977	852	-1.3723	190	0.5849
1978	216	0.1219	341	-0.9266
1979	244	0.0901	135	-0.0770
1980	267	1.8818	125	0.9274
1981	1753	-0.5623	316	-1.0644
1982	999	0.5821	109	0.1127
1983	1788	-2.5260	122	-1.3700
1984	143	-0.5934	31	0.4372
1985	79	0.1738	48	-0.9808
1986	94		18	

^aData from Harrison et al. 1991.

To identify the best model, we advocate the use of Information Criterion statistics, which are computed with the aid of maximum likelihood methods. In the Appendix, we give a thumbnail sketch of the method of maximum likelihood for identifying best-fit model parameters. (To obtain more information on these methods than we have space to provide here, see Edwards 1972, Hilborn and Mangel 1997, and Burnham and Anderson 1998.) *Likelihood* has a technical meaning here: it is the probability of obtaining the observed data given a particular set of parameter values for a particular

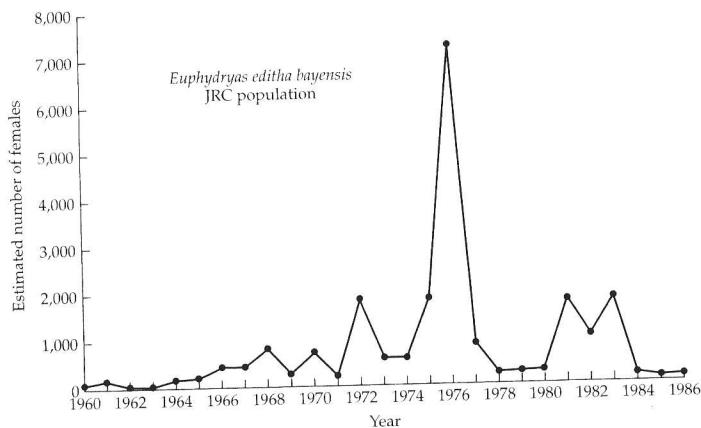


Figure 4.3 Census counts from the JRC population of the Bay Checkerspot butterfly, *Euphydryas editha bayensis* (data from Harrison et al. 1991).

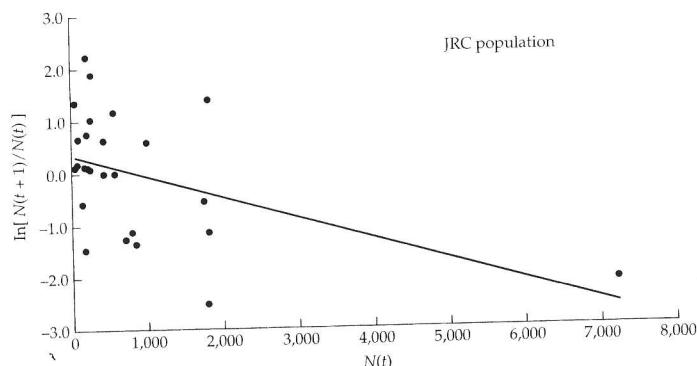


Figure 4.4 The log population growth rate for the JRC checkerspot population is negatively related to population size. The relationship remains significantly negative even if the outlying point at the far right is removed. Line is the prediction from the best-fit Ricker model.

model. For reasons of computational accuracy, it is usually better to calculate the log of the likelihood. The maximum log likelihood for a particular model is obtained when the values of the parameters that have the highest probability of generating the observed data are substituted into the model. Information Criterion statistics combine the maximum log likelihood for a model with the number of parameters it includes to provide a measure of the model's "support"—how good a description of the data it provides, given how complex it is. Support is higher for models with *higher* likelihoods and *fewer* parameters. More complex models are penalized because more parameters will almost always lead to a better fit to the data, but at the cost of less precision in the estimate of each parameter. Comparing the Information Criteria for two or more models will then help us to choose the model that provides the best description of the data.

Briefly, the procedure we will use to assess the presence and pattern of negative density dependence in the checkerspot data involves three steps:

Step 1: Fit three models to the data using nonlinear least-squares regression of $\log(N_{t+1}/N_t)$ against N_t , where N_t is the count obtained from population census t .⁵ The regression equations for the three models are:

Model	Regression Equation
The density independent model	$\log(N_{t+1}/N_t) = r$
The Ricker model	$\log(N_{t+1}/N_t) = r(1 - N_t/K)$
The theta logistic model	$\log(N_{t+1}/N_t) = r[1 - (N_t/K)^\theta]$

As we noted above, the first two models are simplifications of the theta logistic model.

Step 2: Compute the maximum log likelihood of each model using the least-squares estimates of the parameters and the residual variance (defined below).

Step 3: Calculate Information Criterion statistics for the three models, and use them to decide which model provides the best description of the data.

We now provide more detail on each of these steps.

STEP 1: ESTIMATE MODEL PARAMETERS WITH NONLINEAR REGRESSION Any statistical software capable of performing nonlinear least-squares regression can be used in Step 1. For example, the SAS program in Box 4.2 uses the `nlin` procedure to perform the necessary regressions. The parts of the output produced by the SAS program in Box 4.2 that are important for our purposes are reproduced in Box 4.3. To calculate the maximum log likelihood for each model, we need the output labeled `Estimate` for each parameter

⁵Unlike in Chapter 3, in this chapter we will assume that censuses were taken at fixed time intervals.

BOX 4.2 SAS program^{1,2} to fit three models to the Bay checkerspot census data in Table 4.1

```

data jrc;
infile "c:chkrspot.prn" firstobs=2;
input year nt yt;
run;

proc nlin data=jrc;
  title 'JRC Population - theta logistic model';
  parameters r=0.5 to 1.5 by 0.5 K=400 to 800 by 200 theta=0.2 to 1 by 0.4;
  bounds r>=0, K>=0, theta>=0;
  model yt=r*(1-(nt/K)**theta);
  der.r=1-(nt/K)**theta;
  der.K=( r*theta*nt*(nt/K)**(theta-1) )/(K*K);
  der.theta=-r*log(nt/K)*(nt/K)**theta;
  output out=fits p=ypred residual=yres;
run;

proc plot data=fits;
  plot yt*nt='*' ypred*nt='-' / overlay;
  plot yres*nt;
run;

proc nlin data=jrc;
  title 'JRC Population - Ricker model';
  parameters r=0.5 to 1.5 by 0.5 K=400 to 800 by 200;
  bounds r>=0, K>=0;
  model yt=r*( 1-(nt/K) );
  der.r=1-(nt/K);
  der.K=(r*nt) / (K*K);

```

¹To use the program, the data in Table 4.1 must first be placed into a space-delimited data file named "chkrspot.prn"; in reading the data, the program skips the first row of the data file, which may contain column labels.

²Notes on nonlinear least-squares regression:

1. Users must supply SAS with the derivatives of the function to be fit to the data with respect to each of the parameters in the model (hence the `der.r`, `der.K`, and `der.theta` commands in `proc nlin`).
2. Parameter estimates are constrained to be non-negative using the "bounds" statement.
3. It is recommended that users repeat the NLIN procedure using different starting values, to be sure that the search converges to parameter estimates that correspond to the global minimum sum of squares.
4. It is a good idea to check the overlay plots of predicted and observed values produced by this program, to be sure that the best model fits the data reasonably well.

BOX 4.2 (continued)

```

output out=fits p=ypred residual=yres;
run;

proc plot data=fits;
  plot yt*nt='*' ypred*nt='-' / overlay;
  plot yres*nt;
run;

proc nlin data=jrc;
  title 'JRC Population - density-independent model';
  parameters r=0.5 to 1.5 by 0.5;
  bounds r>=0;
  model yt=r;
  der.r=1;
  output out=fits p=ypred residual=yres;
run;

proc plot data=fits;
  plot yt*nt='*' ypred*nt='-' / overlay;
  plot yres*nt;
run;

```

(these are known as the "least-squares estimates"), as well as two pieces of output that can be used to estimate the so-called residual variance, which we will call V_r . The residual variance V_r is the mean squared deviation between the observed values of $\log(N_{t+1}/N_t)$ (i.e., the y values in the nonlinear regression) and the values predicted by the model. The residual variance, in addition to being a component of the maximum log likelihood, will also give us an estimate of the environmental variance in the log population growth rate, equivalent to $\hat{\sigma}^2$ from Chapter 3 (see the following steps). The residual variance is obtained by dividing the Residual Sum of Squares from the regression output (called error sum of squares by many statistics packages) by the number of data points in the regression (in this case, 26 for the number of estimates of the log population growth rate in Table 4.1). For example, for the density-independent model, the residual variance is $36.3976/26 = 1.3999$. (Note that the residual variances are *not* the residual mean squares given in standard regression ANOVA tables such as the one in Box 4.3 (Burnham and Anderson 1998, pp. 15–17)). The least-squares

BOX 4.3 Key output from SAS program used to fit three models to data from the JRC Bay checkerspot population in Table 4.1. Key results are shown in bold.

JRC Population - density-independent model				
Non-Linear Least Squares Summary Statistics		Dependent Variable YT		
Source	DF	Sum of Squares	Mean Square	
Regression	1	0.000072745	0.000072745	
Residual	25	36.397556702	1.455902268	
Uncorrected Total	26	36.397629448		
(Corrected Total)	25	36.397556702		
Parameter Estimate Asymptotic Std. Error Asymptotic 95 % Confidence Interval Lower Upper				
R	0.0016726923	0.23663524902	-.48568282538	0.48902821000
JRC Population - Ricker model				
Non-Linear Least Squares Summary Statistics		Dependent Variable YT		
Source	DF	Sum of Squares	Mean Square	
Regression	2	8.520152294	4.260076147	
Residual	24	27.877477153	1.161561548	
Uncorrected Total	26	36.397629448		
(Corrected Total)	25	36.397556702		
Parameter Estimate Asymptotic Std. Error Asymptotic 95 % Confidence Interval Lower Upper				
R	0.3458095	0.24661973	-0.16318449	0.8548035
K	846.0152803	517.10358675	-221.22547418	1913.2560348

parameter estimates and the residual variance for each model are listed in Table 4.2.

STEP 2: CALCULATE MAXIMUM LOG LIKELIHOOD VALUES FOR EACH MODEL With least-squares estimates of the model parameters and the residual variance in

BOX 4.3 (continued)

JRC Population - theta logistic model				
Non-Linear Least Squares Summary Statistics		Dependent Variable YT		
Source	DF	Sum of Squares	Mean Square	
Regression	3	9.969340119	3.323113373	
Residual	23	26.428289328	1.149056058	
Uncorrected Total	26	36.397629448		
(Corrected Total)	25	36.397556702		
Parameter Estimate Asymptotic Std. Error Asymptotic 95 % Confidence Interval Lower Upper				
R	0.9940620	1.15375841	-1.39264991	3.3807739
K	551.3788238	317.91269892	-106.26840990	1209.0260575
THETA	0.4565906	0.43302473	-0.43918206	1.3523633

hand, we can now calculate maximum log likelihood values for each model. In fitting the models using least squares, we implicitly assumed that the deviations between the observed and predicted log population growth rates followed a normal distribution with a mean of zero (and in so doing, we assumed that small deviations are more probable than large ones, and that positive and negative deviations of a given size are equally probable). The log likelihood of a model assuming normally distributed deviations (see Appendix) is

$$\log L = -\frac{q}{2} \log(2\pi V_r) - \frac{1}{2V_r} \sum_{t=1}^q \left[\log\left(\frac{N_{t+1}}{N_t}\right) - P_t \right]^2 \quad (4.5)$$

where P_t is the value of $\log(N_{t+1}/N_t)$ predicted by the model for census interval t , q is the number of measurements of the log population growth rate in the data set, and V_r is the residual variance we calculated in Step 1. The predicted values for the three models are:

Model	P_t
Density-independent	r
Ricker	$r(1-N_t/K)$
Theta logistic	$r[1-(N_t/K)^\theta]$

(4.6)

TABLE 4.2 Parameter estimates, residual variances, maximum log likelihood, AIC_c values, and Akaike weights for three models fit to data from the JRC Bay checkerspot population in Table 4.1 (number of data points q = 26)^a

Model	Least-squares parameter estimates			Number of parameters (including V_r), p	Maximum log likelihood, log L_{max} (from equation 4.7)	AIC_c (from equation 4.8)	Akaike weights	
	r	K	θ					
Density-independent	0.001673	—	—	1.3999	2	-41.266	87.054	0.07
Ricker	0.3458	846.02	—	1.0722	3	-37.799	82.689**	0.62
Theta logistic	0.9941	551.38	0.4566	1.0165	4	-37.105	84.115	0.31

^aThe most parsimonious model is indicated by asterisks.

Note that the right-hand side of Equation 4.5 includes a sum of squared deviations between the observed and predicted values of the log population growth rate. It was exactly this sum that we minimized by estimating the best parameter values using least-squares regression.

To calculate the maximum log likelihood for a given model, we could simply substitute into Equation 4.5 the appropriate form for P_i in Equation 4.6, the least-squares estimates of the parameters and the residual variance from Table 4.2, and the observed values of N_t and N_{t+1} from Table 4.1. However, as it turns out, when the least-squares estimates of the parameters are substituted into Equation 4.5, the sum on the right hand side equals q times V_r (in other words, the residual sum of squares that we got from the model-fitting output in Box 4.3), so that Equation 4.5 greatly simplifies to

$$\log L_{\max} = -\frac{q}{2} [\log(2\pi V_r) + 1] \quad (4.7)$$

Here *max* signifies that the least-squares parameter estimates by definition produce the maximum value of the log likelihood. The maximum log likelihood values calculated using Equation 4.7 are given in the seventh column of Table 4.2. Note that the values are negative, because each represents the log of a small number (the product of the probabilities of each data point given the model, each of which is a number between zero and one). The model with the largest (that is, the least negative) maximum log likelihood is the theta logistic model. However, the theta logistic model achieves such a high likelihood by using the greatest number of parameters (r , K , θ , and V_r).

As noted above, we can usually improve the likelihood of a model simply by adding more parameters. But do the data really justify including additional parameters, each of which will be estimated with less and less precision? Information Criterion statistics provide a way to answer this question.

STEP 3: COMPARE THE SUPPORT FOR DIFFERENT MODELS USING INFORMATION CRITERION STATISTICS To find the simplest, best-fitting model, we need to account for the cost of using models that achieve high values of the maximum log likelihood by including many parameters. These costs include imprecision of parameter estimates and incorporation of spurious patterns from the data into future predictions. To measure and compare the support for different models, we recommend using the corrected Akaike Information Criterion (AIC_c ; Hurvich and Tsai 1989, Burnham and Anderson 1998), which is more appropriate than other Information Criterion statistics when the number of data points used to fit the models is small relative to the number of parameters (as will usually be the case in PVA). Specifically, for each model we calculate

$$AIC_c = -2 \log L_{\max} + \frac{2pq}{q-p-1} \quad (4.8)$$

where p is the number of parameters estimated (including the residual variance, V_r) and q is (once again) the number of data points. The most parsimonious model is the one with the *lowest* value of AIC_c . The first term on the right hand side of Equation 4.8 gets smaller as the maximum log likelihood increases, so that AIC_c favors models with a high likelihood. However,

the second term on the right hand side of Equation 4.8 gets *larger* as the number of parameters in the model increases, so that models with many parameters are less likely to have the lowest AIC_c value. In fact, the theta logistic model, which has the most parameters, does not have the lowest value of AIC_c (Table 4.2). Instead, the best model identified by the corrected Akaike Information Criterion is the Ricker model, in which the log population growth rate declines linearly with population size. To determine if this difference in AIC_c values is strong enough to firmly reject the more complicated theta logistic, we can also calculate Akaike weights, which quantify the probability that each of a suite of models is the best approximation to the truth (Burnham and Anderson 1998). For a set of R models, the weight of model i is:

$$w_i = \frac{\exp[-0.5(AIC_{c,i} - AIC_{c,best})]}{\sum_{i=1}^R \exp[-0.5(AIC_{c,i} - AIC_{c,best})]} \quad (4.9)$$

Here, $AIC_{c,i}$ is the corrected information criterion for model i and $AIC_{c,best}$ is the best (i.e., smallest) criterion of any model tested. The Akaike weights (Table 4.2) tell us that the Ricker model is twice as likely to be the best model as is the theta logistic. Unless we had some other compelling reason to use the more complicated model, this result supports use of the Ricker model. Also, note that the density-independent model has only a 7% chance of being the best description of the data, giving strong support to the inclusion of negative density dependence in estimation of extinction risks for this population. We now use maximum likelihood parameter estimates for the Ricker model from Table 4.2 to estimate extinction risk for the JRC Bay checkerspot butterfly population by means of simulations.

ESTIMATING EXTINCTION RISK FOR A DENSITY-DEPENDENT POPULATION BY COMPUTER SIMULATION To estimate extinction risk for a density-dependent population, we use a computer to predict the size of the population at one year intervals into the future, starting with the current population size. In making these predictions, we rely on the computer's ability to generate random numbers to introduce environmental variation into the population growth process. In this case, we are using Equation 4.2 with $\theta = 1$ (i.e., the Ricker model) as our population model, so we program the computer to generate a new value for ε_t each year and then substitute it, the parameter estimates, and the current population size N_t into the equation to yield a predicted value of N_{t+1} . Repeating this process over a number of years and for many replicate populations yields estimates of extinction risk.

To implement this algorithm, we must first decide what characteristics the environmentally driven deviations in the log population growth rate (that is, the ε_t 's in Equation 4.2) should exhibit. For a number of reasons it makes sense for the ε_t 's to be drawn from a normal distribution with a mean of

zero.⁶ By estimating V_r during the model-fitting step, we also have a way to estimate the variance of the ε_t 's. The variance of the ε_t 's is the environmentally driven variance in the log population growth rate, which we called σ^2 in Chapter 3. Hence to emphasize that these are the same quantity, we will use σ^2 to represent the true variance of the ε_t 's, and $\hat{\sigma}^2$ to represent an estimate of σ^2 . Whereas the maximum likelihood estimate of the residual variance, V_r , is the appropriate parameter to use when choosing among models with Information Criteria (Burnham and Anderson 1998), for finite data sets V_r gives a biased estimate of the environmental variance in the log population growth rate, σ^2 (Dennis et al. 1991). This is especially true when the number of annual transitions in the data set is small, as will often be the case for endangered populations. An unbiased estimate of σ^2 is

$$\hat{\sigma}^2 = \frac{qV_r}{q-1} \quad (4.10)$$

where q is (once again) the number of data points (Dennis 1989, Dennis et al. 1991, Dennis and Taper 1994). For the JRC Bay checkerspot population, the V_r estimate for the Ricker model from Table 4.2 yields $\hat{\sigma}^2 = 26 \times 1.0722 / (26 - 1) = 1.1151$. Hence to produce appropriate values for the ε_t 's, we multiply computer-generated normally distributed random numbers with a mean of zero and a variance of one (such as those produced by the `randn` function in MATLAB) by $\sqrt{\hat{\sigma}^2}$, to produce normally distributed random numbers with a mean of zero and a variance of $\hat{\sigma}^2$.

Box 4.4 provides a MATLAB program to compute the probability that a population described by the theta logistic model will fall below a specified quasi-extinction threshold at a suite of future times. In the section of the code labeled **SIMULATION PARAMETERS**, the appropriate values for the JRC Bay checkerspot population have been entered. There are several things to note about these parameters. First, we have set `theta` equal to 1, because the Ricker model was judged, using AIC_c , to provide a more parsimonious fit to the data than the theta logistic model (Table 4.2). However, the program allows for situations in which the best-fit value of θ in Equation 4.2 is not equal to 1.⁷ Second, the starting population size (`N0`) is the last census count in Table 4.1. Third, we do not attempt to predict very far into the future

⁶Because the only effect we want the ε_t values to have is to add variability to population growth, their mean should be zero. It is reasonable to assume that their distribution is normal if the additive impact of many environmental variables, each with small effect, determines the overall variability in log population growth, since by the Central Limit theorem, a random variable that is the sum of many other variables will be approximately normally distributed. Note that using least squares to estimate V_r assures that we will obtain the proper variance for a normal distribution of ε_t 's using Equation 4.10.

⁷And the program can easily be modified to use other density-dependent population models.

BOX 4.4 MATLAB code to predict the probability of extinction using the theta logistic model (Equation 4.2).

```
% PROGRAM theta_logistic.m
% Calculates by simulation the probability that a population
% following the theta logistic model and starting at Nc
% will fall below the extinction threshold Nx by time tmax

***** SIMULATION PARAMETERS *****
r=0.3458; % intrinsic rate of increase
K=846.017; % carrying capacity
theta=1; % nonlinearity in density dependence
sigma2=1.1151; % environmental variance
Nc=94; % starting population size
Nx=20; % quasi-extinction threshold
tmax=20; % time horizon
NumReps=50000; % number of replicate population trajectories
*****  
  

sigma=sqrt(sigma2);
randn('state',sum(100*clock)); % seed the random number generator  
  

N=Nc*ones(1,NumReps); % all NumRep populations start at Nc
NumExtant=NumReps; % all populations initially extant
Extant=[NumExtant]; % vector for number of extant pops. vs. time
for t=1:tmax,
    % For each future time,
    % compute new pop. sizes from
    N=N.*exp( r*( 1-(N/K).^theta )... % the theta logistic model
        + sigma*randn(1,NumExtant) ); % with random environmental
        % effects.
    for i=NumExtant:-1:1, % Looping over all extant populations,
        if N(i)<=Nx, % if at or below quasi-extinction threshold,
            N(i)=[]; % delete the population.
        end;
    end;
    NumExtant=length(N); % Count remaining extant populations
    Extant=[Extant NumExtant]; % and store the result.
end;  
  

% Compute quasi-extinction probability as the fraction of
% replicate populations that have hit the threshold by each
% future time, and plot quasi-extinction probability vs. time
ProbExtinct=(NumReps-Extant)/NumReps;
```

BOX 4.4 (continued)

```
plot([0:tmax],ProbExtinct)
xlabel('Years into the future');
ylabel('Cumulative probability of quasi-extinction');
axis([0 tmax 0 1]);
```

(t_{max} is small), for reasons we articulated in Chapter 3. Fourth, the number of replicate populations (NumReps) must be large to get accurate estimates of the probability of extinction.

Some output from the MATLAB code in Box 4.4 is shown in Figure 4.5. With 50,000 replicate populations, there is very close agreement among separate runs of the model regarding the fraction of populations hitting the quasi-extinction threshold by each future time. (Note, however, that these predictions do not account for uncertainty in the parameter estimates; see the following discussion.) Not surprisingly, the population is more likely to fall below a

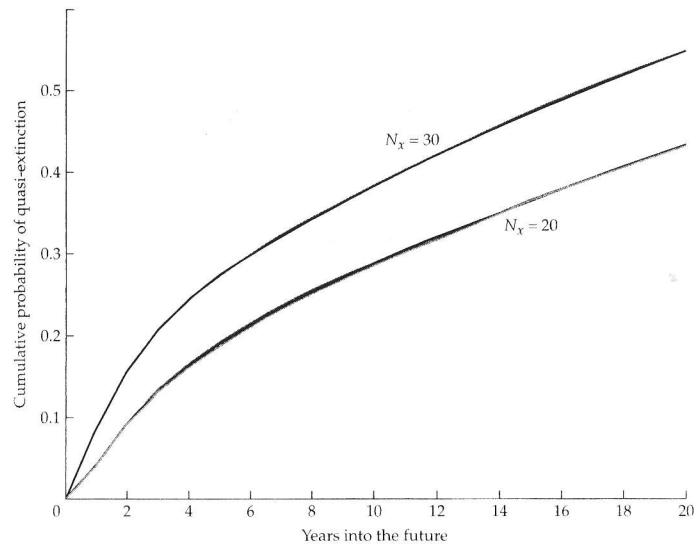


Figure 4.5 Probability of quasi-extinction vs. time for the JRC checkerspot population predicted by the simulation model in Box 4.4. Three runs of the program, each with 50,000 replicate trajectories, are shown for each of two quasi-extinction thresholds.

quasi-extinction threshold of 30 than a threshold of 20. Even with the lower threshold, the probability of extinction is rather high (> 0.4) at only 20 years. To see how including negative density dependence affected the outcome of our analysis, we can reanalyze the data from the JRC population using the methods in Chapter 3, which assume, inappropriately in this case, that the population growth rate is density-independent. This yields estimates for the parameters μ and σ^2 of 0.001673 and 1.4459, respectively. (Note that these same values are obtained from the least-squares estimates of r and σ^2 for the density-independent model in Table 4.2, after correcting the estimate of σ^2 for bias, using Equation 4.10.) Substituting these values into the CDF for the density-independent model (Equation 3.4) yields a predicted probability of 0.772 that the population will decline from 94 to 20 individuals in 20 years. Thus, including negative density dependence results in a predicted extinction probability at 20 years that is only about half the value predicted by a density-independent model. One reason for this difference is that the density-independent model fails to account for the increase in the average population growth rate as the population becomes very small, as we discussed in Chapter 2. Also notice that, by failing to account for the fact that some of the variation among the $\log(N_{t+1}/N_t)$ values is due to density dependence, we wind up with a higher estimate for the environmental variance (1.4559 for the density-independent model versus 1.1151 for the Ricker model), and, as we've seen repeatedly, higher variability increases extinction risk.

That being said, the JRC population actually went extinct in 1996, 10 years after the last census listed in Table 4.1 (McGarrahan 1997; fortunately, other, much larger populations of the subspecies are still extant). Of course, extinction of a single population is a unique event, whereas the probability of extinction reflects the outcomes of many possible realizations of population growth. Rather than ask whether an estimated probability of extinction correctly predicts the observed fate of a single population, it is more appropriate to ask whether the *relative* probabilities of extinction estimated for two or more populations correctly predict their *relative* extinction risk.

Consider the JRH Bay checkerspot population; census counts from this second population at Stanford University's Jasper Ridge Preserve are also listed in Table 4.1. The JRH population went extinct earlier than did the JRC population. As an exercise, readers should verify, using the same likelihood and Information Criterion approach we employed earlier, that the density-independent model provides the best fit to the counts from the JRH population. The estimates of μ and σ^2 for the JRH population are -0.05224 and 0.8410, respectively. Based on the density-independent model, these parameter estimates yield a predicted probability of 0.916 that the population would decline from 18 to 10 individuals in 20 years. This risk is much higher than the quasi-extinction probability predicted for the JRC population by the best-fit Ricker model, even though for the JRC population we used a higher threshold of 20 individuals. Thus the best estimates of the extinction probabilities successfully predict the order of extinction for the two popula-

tions. We should not make too much of this success, as with only two populations, it could easily have occurred by chance. Nevertheless, these results illustrate the value of quantitative estimates of extinction risk for comparing the relative viabilities of multiple populations.

Finally, we mention that individuals from both the JRC and JRH populations were periodically removed for studies performed by Paul Ehrlich and his associates at Stanford University. In our analysis, we have not explicitly considered the effects of this removal, which may have altered both the mean and variance of the population growth rate. For an assessment of how this removal may have affected population viability, see Harrison et al. 1991.

One important concern that we have ignored throughout this section is uncertainty in our parameter estimates. In Chapter 3 we suggested that extinction time predictions should take into account uncertainty in estimates of μ and σ^2 by generating a range of extinction predictions across the confidence intervals for these parameters. Although in this section we have only conducted simulations to predict the extinction time CDF for the best-fit parameter values, attempts to incorporate parameter uncertainty should be undertaken with density-dependent models as well. This is more time-consuming when each extinction probability must be generated by many thousands of simulation runs, and this entire process must be repeated many times for different parameter combinations.⁸ Nevertheless, the increased speed of desktop computers makes this approach feasible for most PVA practitioners.

Positive Density Dependence, or Allee Effects

As we discussed in Chapter 2, Allee effects may be very important to include in PVAs, because they can greatly increase the extinction risk of small populations. Allee effects are well-documented at the individual level (e.g., Birkhead 1977, Widén 1993), and at some point become a logical certainty for all but asexual species. However, as we also discussed in Chapter 2, in spite of some excellent efforts to document positive density dependence at low population sizes, evidence of population-level Allee effects has remained extremely weak. These results leave us in something of a quandary as to

⁸Moreover, estimates of parameters in density-dependent models, such as r , K , and θ in Equation 4.2, will not in general be independent, as were estimates of μ and σ^2 for the density-independent model (Equation 3.1; Dennis et al. 1991). Thus we cannot simply draw values for each parameter independently. Instead, the most generally applicable approach to placing confidence limits on an extinction probability predicted by a density-dependent count-based model will be to conduct a nonparametric bootstrap. Briefly, sample with replacement q pairs of adjacent counts, N_t and N_{t+1} , from the data set, compute log population growth rates and estimate values for all model parameters (e.g., using least-squares regression as in the checkerspot example), and then calculate by simulation the probability of extinction. Now repeat this entire procedure many times. The limits of the central 95% of the resulting values provide a 95% confidence interval for the probability of extinction. Bootstrap confidence intervals may also need to be corrected for bias (see Dixon 2001 and Ellner and Fieberg 2002).

whether we should worry about including Allee effects in count-based PVAs. Information on how Allee effects influence the reproductive success or survival of individuals can be more explicitly incorporated into the demographic PVA models covered in Chapters 6 through 9 than in the count-based models described here. If census data do show evidence of Allee effects, we can include them in count-based PVAs in two ways. First, as we discussed in Chapter 2, if we can estimate a threshold population size below which Allee effects become strong enough to virtually guarantee population extinction, we can simply set the quasi-extinction threshold at or above the population size at which Allee effects become important. Above the threshold, we would not include positive density dependence in the model, although we might still include negative density dependence as described earlier. The second approach is to explicitly include Allee effects in the population model on which the PVA is based. A number of different models with Allee effects have been proposed, both in continuous time (Dennis 1989, Lewis and Kareiva 1993, Lande et al. 1994, Courchamp et al. 1999) and discrete time (Burgman et al. 1993, Stephan and Wissel 1994, Myers et al. 1995, Veit and Lewis 1996, McCarthy 1997, Amarasekare 1998a,b) formulations. As far as we know, none of these models have yet been applied to any rare species. Because discrete time models lend themselves most easily to census data, we briefly describe one discrete-time model with an Allee effect generated by mate-finding problems (Burgman et al. 1993, McCarthy 1997), and we use it to look for evidence of an Allee effect in the Bay checkerspot data (Table 4.1).

Let us assume that the larger a population becomes, the greater the chance that each individual will be able to find a mate. If so, then an appropriate function to describe the number of potential offspring produced per individual, O_t , in relation to population size, N_t , is $O_t = e^r N_t / (A + N_t)$, where A is a parameter controlling the population size at which Allee effects are felt. Justification for using this functional form comes from explicit consideration of the processes by which individuals search for mates (Dennis 1989, Veit and Lewis 1996, McCarthy 1997). When N_t is close to zero, the number of potential offspring per individual will be close to zero, because most individuals will fail to find a mate. But as N_t becomes large relative to A (so that $A + N_t \approx N_t$), and all individuals are able to find mates, the number of potential offspring per individual approaches a maximum of e^r . The parameter A represents the population density at which potential per-capita reproduction is one-half its maximum value. Thus A determines the range of sizes over which the population will "feel" the Allee effect.

We have used the term *potential offspring* above because negative density dependence will typically reduce the actual number of offspring each mated individual can produce, or the survival of those offspring, at high density. Assume that the fraction of the potential per-capita reproduction that is actually achieved is $e^{-\beta N_t}$, where the parameter β reflects the strength of negative density dependence (that is, high values of β imply that actual reproduction falls off steeply as population size increases). Multiplying population size in

year t by the number of potential offspring per individual and by the fraction of potential reproduction that is actually achieved yields the population model⁹

$$N_{t+1} = O_t e^{-\beta N_t} N_t = \frac{N_t^2}{A + N_t} e^{r - \beta N_t} \quad (4.11)$$

We can add environmentally driven deviations to the population growth rate in the same way as in Equation 4.2 to yield a stochastic model

$$N_{t+1} = \frac{N_t^2}{A + N_t} e^{r - \beta N_t + \varepsilon_t} \quad (4.12)$$

where ε_t is the environmental perturbation in year t .

Equation 4.11 illustrates an important feature of models with Allee effects. Because reproduction falls as population size declines, the growth rate of small populations may be insufficient to allow them to avoid extinction (Figure 4.6). In a deterministic model, populations above the so-called Allee threshold will increase (at least initially), whereas populations below the threshold will always go extinct. If we add environmental stochasticity as in Equation 4.12, some populations below the threshold may be pushed above it by serendipitously favorable environmental conditions, and so avoid extinction for a time. On the other hand, environmental variability can push populations initially above the threshold below it, greatly increasing their extinction risk.

Just as we looked for evidence of negative density dependence by fitting density-independent and density-dependent models to census data, we can also compare the fits of a model with an Allee effect, such as Equation 4.12, and one without it (cf. Myers et al. 1995). For example, it is easy to modify the SAS code in Box 4.2 to fit Equation 4.12 instead of the theta logistic model to the census data from the JRC Bay checkerspot population.¹⁰ One

⁹Equations 4.11 and 4.12 arise most naturally as a model for an organism that lives only one year or less, so that each individual contributes to next year's population only through reproduction, not survival. The model can easily be modified to allow overlapping generations, but at the cost of additional parameters that must be estimated.

¹⁰Specifically, change the call to the nlin procedure as follows:

```
proc nlin data=jrc best=1;
  title 'JRC Population - Allee effect model';
  parameters r=0.1 to 1.5 by 0.1 b=0.1 to 1 by 0.1 A=10 to 100 by 10;
  bounds r>=0, b>=0, A>=0;
  model yt=log(nt)-log(A+nt)+r-b*nt;
  der.r=1;
  der.b=-nt;
  der.A=-1/(A+nt);
  output out=fits pypred residual=yres;
run;
```

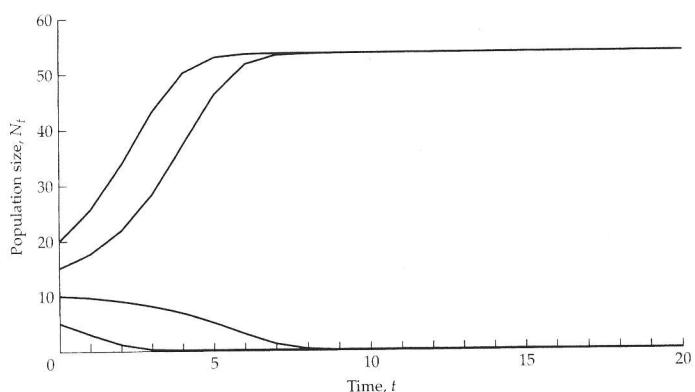


Figure 4.6 In a deterministic model with an Allee effect (Equation 4.11), populations starting below the Allee threshold (which occurs at a population size of approximately 12 in this graph) are doomed to extinction (parameter values: $r = 2$, $A = 50$, $\beta = 0.025$).

could then compute the AIC_c value for the Allee effect model using Equations 4.7 and 4.8, and compare it to the AIC_c value for the Ricker model (which, of course, lacks an Allee effect).¹¹ For the JRC Bay checkerspot data in Table 4.1, fitting the Allee effect model via least squares regression identifies a value of 0 as the maximum likelihood estimate of the parameter A , assuming that we constrain the estimate to be nonnegative¹² (try it!). When $A = 0$, the Allee effect model of Equation 4.12 reduces to the Ricker model (with $\beta = r/K$). Thus even without computing the likelihood or AIC_c values, we can see that the checkerspot data show no evidence of an Allee effect. Parsimony thus argues for the use of a model with negative density dependence only, as we have done. If you do find good evidence for Allee effects in a data set, the most practical way to derive extinction time estimates is to follow the same model fitting and simulation approach we detailed for negatively density-dependent dynamics in the preceding section.

¹¹An advantage of using Information Criterion statistics to compare models is that they do not assume (as do likelihood ratio tests, for example) that the models are nested (i.e., that one model can be converted to another by setting one or more parameters equal to 0 or 1; for example, the theta logistic model (Equation 4.2) and the Allee effect model (Equation 4.12) are *not* nested). For details, see Burnham and Anderson 1998.

¹²As we did in the SAS code listed in Footnote 10.

Combined Effects of Demographic and Environmental Stochasticity

As we discussed in Chapter 2, very small populations may go extinct due to random variation among individuals in their survival or in the number or sex ratio of their offspring. By omitting demographic stochasticity in all we have discussed until now, we have ignored a factor that may often deliver the coup-de-grâce to populations that have been depressed to low levels by environmental stochasticity.

That being said, it is not possible to estimate the magnitude of demographic stochasticity using count data alone, as demographic stochasticity is intimately connected to variation in the fates of individual organisms, not just the total number of individuals in a population. In essence, to estimate the amount of variability that demographic stochasticity introduces into the population growth rate, we need the type of data on survival and reproduction of individuals that is typically used to construct the projection matrix models we discuss in Chapters 6 through 9. However, when such individual-level data are available from a population that has also been the focus of a long-term census, the two types of data can be used in conjunction to incorporate both demographic and environmental stochasticity into the count-based PVA models we have described in this and the preceding chapter. Consequently, we give an overview here of how such a tandem approach might work in those rare cases in which census counts and individual-level information are available for the same endangered population. It is indicative of how rarely this situation will arise that the only published examples of this approach rely on intensive studies of small, isolated populations of otherwise common (and therefore not endangered) species. We cite most of these papers in describing how to approach this problem, and if you want to use this method you should consult them for more details.

Engen et al. (1998) proposed a method to estimate the separate contributions of demographic and environmental stochasticity to the variance in the population growth rate (also see Saether et al. 1998a,b, 2000a,b, Tufto et al. 2000, Saether and Engen 2002). We must first collect data on a set of marked individuals over several years, recording the number of offspring each marked individual produces each year and whether those offspring, as well as the marked parent, survive to the following year (for methods to conduct such a demographic study, see Chapter 6). The procedure of Engen and colleagues begins by calculating the contribution each marked individual in year t makes to the population in year $t + 1$. This contribution equals the number of that individual's offspring produced in year t that survive to year $t + 1$ plus 1 if the focal individual itself survives to the next year. For example, the contribution of an individual producing 5 surviving offspring would be 6 if that individual survives and 5 if not. Let us call the contribution of individual i in year t C_{it} , and the mean of those contributions

$$\bar{C}_t = \left(\sum_{i=1}^{m_t} C_{it} \right) / m_t$$

where m_t is the number of individuals whose contributions in year t are known. An estimate of the demographic variance in year t is the sample variance of the C_{it} 's, given by

$$V_d(t) = \frac{1}{m_t - 1} \sum_{i=1}^{m_t} (C_{it} - \bar{C}_t)^2 \quad (4.13)$$

Notice that Equation 4.13 is measuring *variation among individuals in the same year*, which is the driving force behind demographic stochasticity.

Data from each year in which the survival and reproduction of marked individuals were measured yield separate estimates of the demographic variance when they are substituted into Equation 4.13. These separate estimates should then be regressed against population size in each year, to see if the demographic variance is itself density dependent. (For example, high density may depress reproduction or survival of all individuals, reducing variation among them; Saether et al. 1998a.) If the regression is significant, the regression equation provides a way to update the demographic variance given the population size each year. If the regression is not significant, the average of the separate estimates of $V_d(t)$ across years may be used as an overall estimate of the demographic variance.¹³

Once we have calculated demographic variance estimates for each of n years using Equation 4.13, we can use the census data to obtain n estimates of the environmental variance from the expression

$$\sigma^2(t) = \left[\frac{N_{t+1}}{N_t} - f(N_t) \right]^2 - \frac{V_d(t)}{N_t} \quad (4.14)$$

where $f(N_t)$ is the (possibly density-dependent) population growth rate in year t predicted by the best-fit population model $N_{t+1} = N_t f(N_t)$. For example, if we were using the theta-logistic model (Equation 4.2), $f(N_t)$ would be $\exp[r(1 - (N_t/K)^\theta)]$ with the values of r , K , and θ determined by nonlinear regression, as we did above. The term in square brackets in Equation 4.14 is the squared deviation between the observed population growth rate from year t to year $t + 1$ and the growth rate predicted by the model. This deviation is due to both environmental and demographic stochasticity. Subtracting from it our estimate of the contribution of demographic stochasticity to the variance in the population growth rate ($V_d(t)/N_t$) yields an estimate of

¹³Saether et al. 1998a, 2000b, and Saether and Engen 2002 advocate weighting the yearly estimates of $V_d(t)$ by m_t when computing the overall demographic variance. See White 2000 for an alternative weighting.

the variance in population growth due to environmental stochasticity alone. As with the $V_d(t)$ estimates, the $\sigma^2(t)$ estimates should be regressed against N_t to check for density dependence in the environmental variance. (For example, environmental perturbations to the population growth rate may be larger when population size is large and, consequently, refuges against extreme environmental conditions are scarce.) If they are density-independent, the separate estimates of $\sigma^2(t)$ should be averaged to obtain an overall estimate of the environmental variance; if density-dependent, the regression equation should be used to generate an appropriate value for $\sigma^2(t)$ each year of a computer simulation. Note that the estimate of σ^2 we obtain from Equation 4.14 is not identical to that used in Chapter 3 and earlier in this chapter. Those earlier calculations assumed that the observed variation in the log population growth rate was due solely to environmental stochasticity, even though in reality it always included some (hopefully very small) variability due to demographic stochasticity. Equation 4.14 removes this demographic variability, yielding an estimate of σ^2 that is a smaller, more accurate representation of true environmental stochasticity.

Notice that in Equation 4.14, the contribution of demographic stochasticity to variation in the population growth rate is the variation in an individual's contribution to population growth divided by the size of the population, $V_d(t)/N_t$, which becomes smaller and smaller as population size N_t increases. This makes sense, since as populations become larger, the variances in the contributions made by different individuals will tend to cancel each other out. In other words, Equation 4.14 predicts that the effects of demographic stochasticity will be important only for small populations, just as we would expect from our discussion of demographic stochasticity in Chapter 2 (also see Gabriel and Burger 1992 and Lande 1993).

If both demographic and count data are available so that V_d and σ^2 can be estimated separately, computer simulations can be used to predict extinction risk under both environmental and demographic stochasticity. For completeness, let us assume that linear regressions of $V_d(t)$ and $\sigma^2(t)$ against N_t are both statistically significant, with the regression equations given by $V_d(t) = a_d N_t + b_d$ and $\sigma^2(t) = a_e N_t + b_e$. We will also assume that the density-dependent population growth rate $\lambda(N_t)$ is lognormally distributed, with a mean given by the theta logistic model (that is, $M[\lambda_t] = \exp[r(1 - (N_t/K)^\theta)]$), where $M[\cdot]$ denotes the mean) and a variance given by $V[\lambda_t] = \sigma^2(t) + V_d(t)/N_t$, which is the sum of components due to environmental and demographic stochasticity. If $\lambda(N_t)$ is lognormally distributed with mean $M[\lambda_t]$ and variance $V[\lambda_t]$, then we can create random values for $\lambda(N_t)$ by computing values of $\exp[X_t]$, where X_t is normally distributed with mean and variance given by $M[X_t] = \log M[\lambda_t] - \frac{1}{2}V[\lambda_t]$ and $V[X_t] = \log[1 + V[\lambda_t]/(M[\lambda_t])^2]$ (see Equation 8.9). The MATLAB program demstoch listed in Box 4.5 uses the theta logistic model to simulate density-dependent population growth in the face of both demographic and

BOX 4.5: A MATLAB program to simulate growth of a density-dependent population with both environmental and demographic stochasticity. It would be straightforward to hybridize this code with the program “theta_logistic” in Box 4.4 to compute the probability of extinction under both types of stochasticity.

```
% PROGRAM demstoch;
%   Simulates the theta-logistic model with environmental and
%   demographic stochasticity. Assumes the population growth
%   rate at time t is lognormally distributed with mean
%   exp(r*(1-(Nt/K).^theta)) and variance Vd/N+sigma^2, where
%   both the demographic variance Vd and the environmental
%   variance sigma^2 may themselves be density-dependent -
%   i.e., Vd=ad*Nt+bd and sigma^2=ae*Nt+be

%*****SIMULATION PARAMETERS *****
NumPops=10; % number of trajectories to simulate
Nc=10; % initial population density
r=0.1; % finite rate of increase
K=15; % carrying capacity
theta=1; % pattern of density dependence
ad=0; % slope of demographic variance vs. Nt
bd=1; % intercept of demographic variance vs. Nt
ae=0; % slope of environmental variance vs. Nt
be=0.1; % intercept of environmental variance vs. Nt
tmax=50; % duration of simulation
%*****
```

```
randn('state',sum(100*clock)); % seed the random number generator
N=Nc*ones(1,NumPops); % start all trajectories at Nc

Ns=[N]; % store initial densities
for t=1:tmax % for each future time,
    Mlam=exp(r*(1-(N/K).^theta)); % compute means of lambdas,
    Vlam=(ad*N+bd)/N + ae*N+be; % variances of lambdas,
    CV2=Vlam./((Mlam.^2)); % squared coefficients of
    % variation of lambdas,
    Mx=log(Mlam)-0.5*log(CV2+1); % means of Xt's, and
    SDx=sqrt(log(CV2+1)); % standard deviations of Xt's,
    N=N.*exp(Mx+SDx.*randn(1,NumPops)); % then project populations
    % one year ahead using
    % lambda=exp(Xt),
    % and store the results.
    Ns=[Ns; N];
end;
```

BOX 4.5 (continued)

```
% For all trajectories, plot Nt vs t, with a log scale for the y
% axis
semilogy(Ns)
xlabel('Time (years)');
ylabel('Population density');
axis([1 tmax .01 100]);
```

environmental stochasticity, the magnitudes of which are determined by the simulation parameters a_d , b_d , a_e , and b_e estimated as we have just described. The variance in the population growth rate is computed in the line:

$$V_{\text{lam}} = (ad * N + bd) / N + ae * N + be;$$

At first glance, it may appear as though partitioning the observed variation in population growth into components due to demographic and environmental stochasticity was unnecessary, since the two parts of V_{lam} simply reconstitute the observed total variability with which we started. However, the importance of making this breakdown comes as population size drops. Although we have previously assumed the total variance in the population growth rate would remain fixed at the single observed value, the fact that the first part of V_{lam} has N in the denominator ensures that the total variance will increase due to the growing importance of demographic stochasticity at small population size (and conversely shrink with increasing numbers).

We can use the program `demstoch` to examine how environmental and demographic stochasticity combine to determine extinction risk (Figure 4.7). In Figure 4.7A, individuals vary little in their contributions to population growth (i.e., the demographic variance V_d is small), so demographic stochasticity is weak. Even so, the predominating effect of environmental variation can still cause some of the replicate trajectories (all of which start at a population density of 10) to fall low enough that we would probably label them “quasi-extinct” (note the log scale of the y -axis). However, if we ratchet up the variation in individual contributions to population growth, as in Figure 4.7B, we see qualitatively different behavior; any trajectories that happen to be pushed below a certain level by environmental variation embark on a precipitous, runaway decline that looks nearly deterministic in its predictability. The reason for this virtually inescapable slide to extinction is that, as population density falls, the variance in the population growth rate increases due to demographic stochasticity, and higher variance then leads to a lower geometric mean population growth rate (see Chapter 2), which causes the population to decline still more, further increasing the variance due to demographic stochasticity and further reducing the geometric mean, and so on, in an

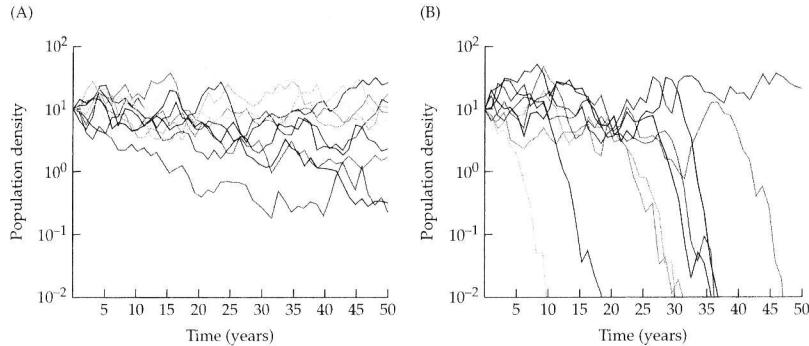


Figure 4.7 Population dynamics in the face of both environmental and demographic stochasticity, as simulated by the program demstoch (Box 4.5). Both panels show ten population trajectories simulated with the theta logistic model with parameters $(r, K, \theta) = (0.1, 15, 1)$ and a (density-independent) environmental variance of 0.1 (i.e., $a_e = 0$ and $b_e = 0.1$). In panel B (in which $a_d = 0$ and $b_d = 1$, so that $V_d = 1$ for all t), the (density-independent) demographic variance is ten times higher than in panel A (in which $a_d = 0$, $b_d = 0.1$, $V_d = 0.1$). When contributions to population growth vary greatly among individuals, populations that fall to low numbers are doomed to extinction (panel B).

inevitable downward spiral. Thus environmental and demographic stochasticity conspire in a “one-two punch,” the former acting to push populations below some fuzzy threshold at which the latter takes fatal hold. The height of this threshold is proportional to the demographic variance V_d (see Lande 1998, 2002). It is precisely this threshold-like behavior of demographic stochasticity that argues for the use of a quasi-extinction threshold, rather than measuring the risk of outright extinction, when lack of data on the variation in individual contributions (i.e., the C_{it} ’s) forces us to assess extinction risk with a model that includes environmental stochasticity only. The challenge, of course, is to decide what that threshold should be without an estimate of the demographic variance. In such cases, the general guidelines for setting the quasi-extinction threshold that we gave in Chapter 2 are probably about the best we can do.

Although the method we have described in this section lets us estimate the effects of demographic stochasticity and to extrapolate its effects to low population sizes, one caveat needs to be mentioned about this procedure. It makes the implicit assumption that all variation among individuals in their contributions to population growth in a single year is, in fact, due to random chance, rather than to fixed differences between individuals related to their size, age, genotype, habitat, and so on. In other words, there is an assumption that all

individuals are identical to each other, so that they have the same *expected* contribution to population growth in any one year. As we discussed in Chapter 2, fixed differences among individuals are most likely to inflate our estimate of the importance of demographic stochasticity, because variation due to fixed differences, unlike that caused by demographic stochasticity, will not increase as population size declines (Kendall and Fox 2002, Fox and Kendall 2002). Thus we should keep in mind that if a substantial amount of the variation in the C_{it} ’s in Equation 4.13 is due to fixed inter-individual differences, we may overestimate the magnitude of demographic stochasticity at low population sizes, and thus overestimate the risk of extinction. In the terminology of Chapter 3, falsely attributing variation caused by fixed differences to demographic stochasticity will tend to make our viability assessments pessimistic.

Aside from this concern, this method does allow the simultaneous treatment of both environmental and demographic stochasticity in a count-based PVA. In Equation 4.13, the contribution each individual makes to the population in the next year accounts for both the surviving offspring it produces and its own survival. Thus two major factors contributing to demographic stochasticity have been accounted for (although variation in offspring sex ratio has not; for models that include sex ratio variation but no environmental stochasticity, see Legendre et al. 1999). Other approaches focus on either survival or reproduction. For example, Kendall (1998) proposed a maximum likelihood method to disentangle effects of demographic and environmental stochasticity when data on survival, but not reproduction, of individuals are available. However, as Kendall’s method does not involve a combination of demographic and count data, we defer discussion of it until Chapter 8.

Environmental Autocorrelation

Up to now, we have assumed that the impacts of environmental factors on population growth are independent from one year to the next. For example, in the computer simulation in Box 4.4, we drew a new random number for the environmental deviation, ϵ_t , in each year, without regard to the value in the previous year. However, temporal environmental autocorrelation may cause the deviations in the population growth rate to be correlated among years, although this is not necessarily the case. For example, some correlated environmental conditions may have only a weak or insignificant effect on birth and death rates in a population, or changes that are caused by several environmental factors or that affect different population processes may cancel each other out. On the other hand, if environmental events in one year lead to long lasting changes in the age, size, or stage structure of a population, uncorrelated environmental events can still lead to autocorrelation in population growth rates. For example, bad weather conditions in one year could lead to depressed population growth rates for several years thereafter because of a reduction in the fraction of the population that is of prime

reproductive age. Untangling autocorrelations in population growth rates that are due to direct autocorrelation in the environment versus interactions between the environment and the population's composition is not possible with census data alone, and in this chapter we will refer to both effects as environmental autocorrelation. In either case, if population growth is correlated in time, the risk of extinction may be poorly predicted by PVAs that assume environmental deviations are independent. In this section, we briefly review the impact of environmental autocorrelation on population viability. In addition, we summarize how to test for autocorrelation in census data, and how to incorporate it in viability assessments if it occurs.

Hypothetically, environmental factors could be positively or negatively correlated in time. Positive correlations occur when adjacent years in a sequence tend to be more similar than would be expected in a completely random sequence (for example, wet years tend to follow wet years, and dry years dry). Negative correlations occur if adjacent years differ more than would be expected by chance (dry years follow wet, and vice versa). In reality, positive correlations in environmental variables such as temperature are commonly encountered (Steele 1985, Halley 1996), but convincing examples of negative correlations are hard to find,¹⁴ so we will only discuss positive correlations below.

In a density-independent model, positive autocorrelation elevates extinction risk relative to population dynamics with the same mean and variance in growth but no correlation (Figure 4.8). The reason is simple: strong positive correlation means that strings of bad years will be less likely to be broken up by the occasional good year that could pull the population back from the brink of extinction. A similar effect is seen in the ceiling model (Equation 4.1). Foley (1994) suggested that when the environmental effects on the population growth rate are correlated, the “effective” environmental variance in the log population growth rate is

$$[(1 + \rho)/(1 - \rho)] \sigma^2 \quad (4.15)$$

where ρ (rho) is the correlation coefficient between the environmental effects on population growth in adjacent years and σ^2 is the environmental variance estimated as in Chapter 3. (See the following discussion for a way to estimate ρ .) Note that when ρ is positive, the effective environmental variance is greater than σ^2 . To calculate the mean time to extinction in a correlated environment, Foley recommended using Expression 4.15 in place of σ^2 in Equation 4.4. We have already seen that a high degree of environmen-

¹⁴The grizzly bear data in Chapter 3 provide one possible example of (weak) negative autocorrelation.

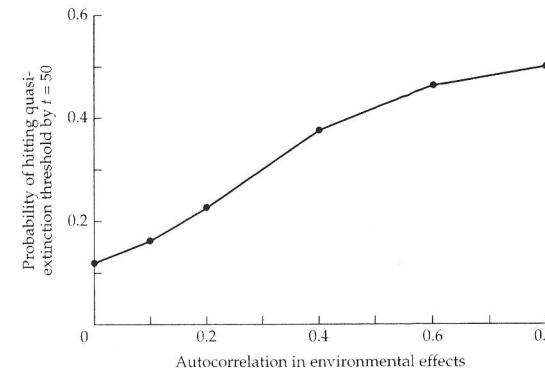


Figure 4.8 Positive autocorrelation in environmental effects increases extinction risk in a density-independent population. To make autocorrelation easier to add, we've modified the familiar density-independent model (Equation 3.1). The model is $N_{t+1} = e^{\mu + \epsilon_t} N_t$, where μ and σ^2 are as in Chapter 3, and the

environmental effect is $\epsilon_t = \rho \epsilon_{t-1} + \sqrt{\sigma^2} \sqrt{1 - \rho^2} z_t$, where ρ is the correlation between environmental effects in adjacent years and z_t is a random number drawn from a normal distribution with a mean of zero and a variance of 1 (see Equation 4.16 for an explanation of ϵ_t). Starting population size and extinction threshold were 10 and 5, respectively. Ten-thousand replicate trajectories were computed for each level of environmental autocorrelation. Parameter values: $\mu = \sigma^2 = 0.01$.

tal variation reduces the mean time to extinction (Figure 4.2B), so by increasing the effective environmental variation, autocorrelation increases extinction risk in the ceiling model just as it does in the density-independent model (also see Johst and Wissel 1997).

The story is more complicated for fully density-dependent models such as the Ricker model (Ripa and Lundberg 1996, Petchey et al. 1997, Heino 1998, Ripa and Heino 1999). In that case, positive autocorrelation can either increase or decrease extinction risk, depending on the values of the other parameters in the model and the range over which the strength of autocorrelation is varied. A key determinant of how temporal correlation will affect extinction risk is whether the recruitment curve—that is, a curve plotting N_{t+1} against N_t —has a positive or negative slope at the equilibrium population size (Petchey et al. 1997, Ripa and Heino 1999). If the slope is positive, then positive autocorrela-

tion will increase extinction risk; if the slope is negative, increasingly positive autocorrelation can decrease extinction risk over some range of values¹⁵ (Figure 4.9). Two arguments can help to explain this divergent effect of environmental correlation. First, if the slope is positive, positive autocorrelation increases the variance in population size near the equilibrium, whereas the opposite is true if the slope is negative (Roughgarden 1975, Ripa and Heino 1999), and high variance increases extinction risk. This is the same effect we saw with the density-independent and ceiling models. The second argument involves the meaning of the negative slope of the recruitment curve in Figure 4.9B. The negative slope at the equilibrium indicates that the intrinsic dynamics of the population are “overcompensatory”; that is, the population will tend to repeatedly overshoot and then undershoot the equilibrium. Thus following a good year that pushes the population above the equilibrium, density dependence will push the population below the equilibrium in the next year (dramatically so if the slope is strongly negative). A second good year made more likely by positive environmental autocorrelation will tend to resist the population’s inherent tendency to crash, thus keeping it close to the equilibrium and farther from the extinction threshold. So, when a population’s own dynamics tend to cause “boom and bust” cycles, positive autocorrelation may help to stabilize dynamics and reduce extinction risk.

In summary, whether temporal autocorrelation will increase or decrease a population’s risk of extinction depends on whether density dependence is operating, and if so, on the “shape” of density dependence near the population equilibrium (Figure 4.9). Thus we cannot make blanket statements about the impact of temporal environmental autocorrelation on population viability, and we must be careful to choose a model (or models) that incorporates density dependence in an appropriate way when including environmental correlations in a PVA.

How do we test for environmental correlations in a set of population counts? We advocate the following approach, which is slightly more elaborate

¹⁵The effect of positive autocorrelation on extinction risk is sensitive to the way environmental variation is entered in the model. In Figure 4.8, environmental variation is incorporated by multiplying the new population size predicted by the deterministic part of the model (i.e., $N_{t+1} = N_t \exp\{r(1 - N_t / K)\}$) by a lognormal random deviate (i.e., $\exp\{\varepsilon_t\}$). In contrast, Ripa and Lundberg (1996) entered environmental variation into the Ricker model either by letting K vary randomly (that is, the strength of density dependence varies over time) or by introducing additive noise that is proportional to population size. In both cases, they found that, rather than decreasing and then increasing the extinction risk (as in Figure 4.8), increasingly positive environmental autocorrelation strictly decreased the probability of extinction. Unfortunately, it may be very difficult to determine which stochastic population model best fits a set of census data, and when exploring the consequences of environmental autocorrelation, it may be wise to use multiple models to see if their predictions agree (see Chapter 12).

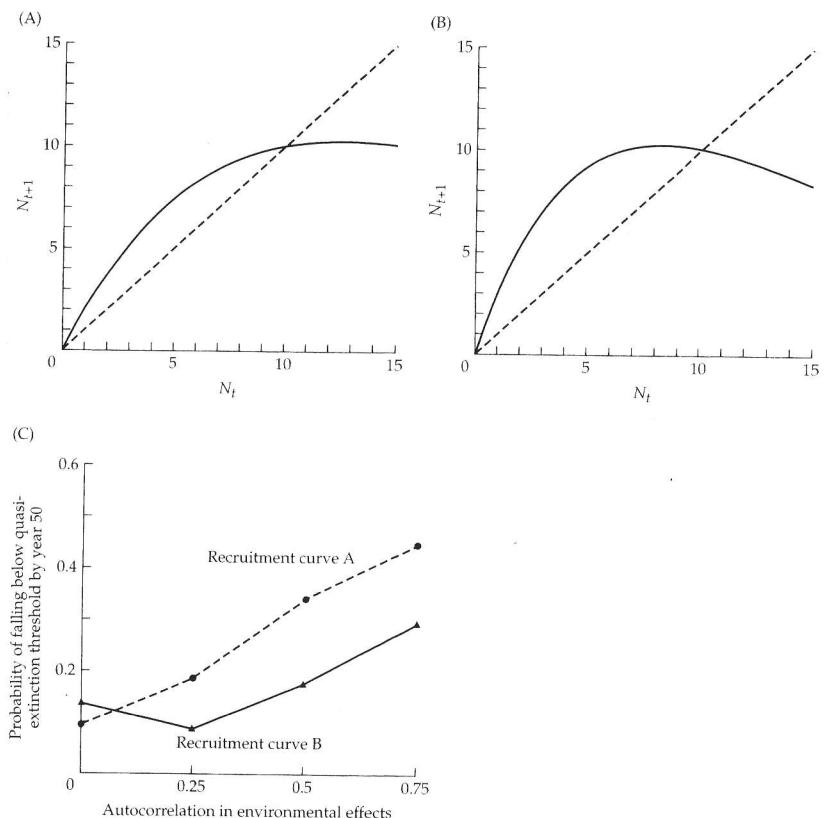


Figure 4.9 In a model with negative density dependence, temporal correlation in the environmental effects can increase or decrease extinction risk, depending in part on the slope of the recruitment curve (N_{t+1} vs. N_t) at the equilibrium. In graphs A and B, the solid line is the recruitment curve for the Ricker model — $N_{t+1} = N_t \exp\{r(1 - N_t / K)\}$ — and the dotted line connects all points where $N_{t+1} = N_t$; the equilibrium population size K lies at the intersection of the two lines (here, $K = 10$). When $r < 1$ as in graph A (where $r = 0.8$), the slope of the recruitment curve at the equilibrium is positive; when $r > 1$ as in graph B (where $r = 1.2$), the slope of the recruitment curve at the equilibrium is negative. Graph C shows the result of simulating the stochastic Ricker model $N_{t+1} = N_t \exp\{r(1 - N_t / K) + \varepsilon_t\}$ with the (autocorrelated) environmental effect ε_t defined in Equation 4.16 ($\sigma^2 = 0.05$; initial population size is K). Positive environmental autocorrelation increases extinction risk for recruitment curve A, but initially decreases it for recruitment curve B.

than that discussed at the end of Chapter 3, because of the need to incorporate density-dependent effects into the test.

1. Calculate the observed log population growth rates—that is, $\log(N_{t+1}/N_t)$ —using all of the population censuses. (As usual, there will be one less log growth rate than the number of censuses; for example, see the third column of Table 4.1.)
2. Fit density-independent and density-dependent models to the data and use AIC_c criteria to choose the best model to describe population growth.
3. Using the best-supported model, calculate the predicted log population growth rate for each inter-census interval. First, substitute each observed value of N_t into the best-fit population model to calculate its associated predicted value of N_{t+1} . Second, divide the predicted value of N_{t+1} by the observed value of N_t . The log of this ratio is the predicted log population growth rate for year t . Repeat for all values of N_t .
4. Subtract the predicted log growth rates from the observed log growth rates; these differences represent the environmental deviations. If the best model is density independent, then these are simply the differences between each observed log growth rate and μ (see *Using regression diagnostics to test for temporal autocorrelation in the population growth rate* in Chapter 3).
5. Make a spreadsheet with two columns, the first containing the deviations for time intervals 1 through $q - 1$, and the second containing deviations for intervals 2 through q (where q is the number of estimates of the population growth rate in the data set; note that both columns will have $q - 1$ rows). Finally, calculate the Pearson correlation coefficient between the environmental deviations in successive years (e.g., using the function `corrcoef` in MATLAB or the function `CORREL` in Excel); this coefficient estimates the strength of “first-order” environmental autocorrelation (i.e., the correlation between adjacent years).

For example, Figure 4.10 shows the environmental deviations for the JRC Bay checkerspot population in year t versus the deviations in the previous year, using the parameter values for the Ricker model in Table 4.2 to calculate the predicted log population growth rates. The calculated correlation coefficient ($r = -0.189$, $n = 26$) does not differ significantly from zero. Thus there is no evidence that the impact of environmental factors on the JRC population is correlated from one year to the next, and we are justified in omitting such correlations from our computer simulation (Box 4.4).

In theory, one could test for correlations between environmental deviations that are separated by more than one year using time series analysis. (See Royama 1992 for an introduction to the application of time series analysis to population ecology.) These methods account for the fact that, if deviations in adjacent years are correlated, the correlation between years t and $t + 1$ and between years $t + 1$ and $t + 2$ will generate an apparent correlation between years t and $t + 2$, even if there is no direct correlation between them.

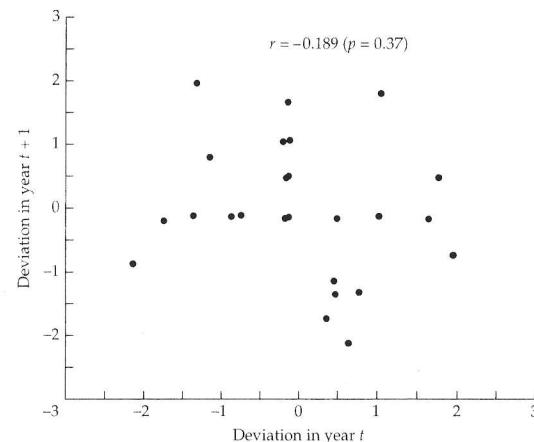


Figure 4.10 Deviations between observed and predicted log population growth rates for the JRC checkerspot population are not significantly correlated between adjacent years.

However, as the sample size decreases with an increase in the number of years between two environmental deviations we wish to compare, testing for long-term environmental correlations will usually require censuses that are longer than those available for most endangered populations.¹⁶

If we do find a significant correlation between the environmental deviations in successive years, we can incorporate environmental correlation into our viability assessment in the following way. Basically, we modify the computer simulation to make the environmental deviation in one year dependent upon the value of the deviation in the previous year. We must do this in such a way that, despite the fact that adjacent deviations are not independent, the expected variance of the environmental deviations still equals σ^2 as estimated from the data. We can accomplish this by generating the environmental deviations (for example, the ε_t terms in Equation 4.2) using the equation

$$\varepsilon_t = \rho \varepsilon_{t-1} + \sqrt{\sigma^2(1-\rho^2)} z_t \quad (4.16)$$

where ρ is the correlation coefficient between successive environmental deviations and z_t is a random number drawn from a normal distribution with a mean of zero and a variance of 1 (Foley 1994, Ripa and Lundberg

¹⁶Halley 1996 has argued that environmental variation may be better represented by “1/f noise,” which differs from autoregressive noise and may have different effects on population viability (Halley and Kunin 1999; also see Morales 1999). As it remains unclear whether 1/f noise provides a better description of environmental impacts on population growth, we follow the more traditional autoregressive approach here.

BOX 4.6 MATLAB code to calculate the probability of quasi-extinction for the Ricker model with temporally autocorrelated environmental effects.

```
% PROGRAM ricker_corr
%   Uses the stochastic Ricker model
%   N(t+1)=N(t)exp(r(1-N(t)/K)+e(t))
%   with correlation rho between environmental deviations e(t)
%   in adjacent years to calculate (by simulation) the
%   probability that a population starting at Nc will fall below
%   the quasi-extinction threshold Nq by time tmax

%*****SIMULATION PARAMETERS *****
tmax=50;          % time horizon
NumReps=10000;    % number of replicate populations to simulate
Nc=10;            % current population size
Nx=5;             % quasi-extinction threshold
r=0.8;            % intrinsic rate of increase
K=10;              % equilibrium population size
sigma2=0.05;       % total environmental variance
rho=0.1;           % first-order environmental autocorrelation
%*****
```

```
sigma=sqrt(sigma2);
randn('state',sum(100*clock)); % Seed the random number generator
beta=sqrt(1-rho^2); % Beta scales new random numbers so that
                     % total environmental variance=sigma2.

NumExt=0;           % No populations are extinct initially.
for i=1:NumReps, % For each replicate population,
    Nc=;            % start at current population size,
    eold=randn;     % generate an initial environmental deviation
    for j=1:tmax, % then looping over time,
        enew=rho*eold + ... % make new (auto-correlated)
        sigma*beta*randn; % environmental deviate and
        N=N*exp(r*(1-N/K)+enew); % use the Ricker model to
                                    % project one year ahead.
        eold=enew; % Save old environmental deviation.
        if N<Nx % Stop projecting if threshold is hit.
            break;
        end;
    end;
    if N<Nx
        NumExt=NumExt+1; % Store number of extinct populations.
    end;
end;
```

BOX 4.6 (continued)

```
NumExt/NumReps % Compute probability of hitting Nq at or before
                % tmax
```

1996). Each e_t is the sum of a term due to correlation with the previous environmental deviation (ρe_{t-1}) and a new random term, scaled by the factor $\sqrt{\sigma^2 \sqrt{1 - \rho^2}}$ to assure that the variance of a long string of e_t 's generated in this way equals σ^2 . Box 4.6 provides MATLAB code that incorporates environmental correlations into the Ricker model. This program was used to generate the results shown in Figure 4.9, and can also be easily modified to simulate correlations for a density-independent model of similar form (see the legend to Figure 4.8).

Catastrophes, Bonanzas, and Other Highly Variable Environmental Effects

Of the factors discussed in this chapter, catastrophes and bonanzas are perhaps the most difficult to include in a PVA in a defensible way, even though catastrophes may also be the greatest cause of concern for many species. The reason for this is that catastrophes and bonanzas are by their very nature infrequent events; consequently, we will usually have even less information about their frequencies and magnitudes than we will about density dependence, environmental correlations, and even demographic stochasticity. In practice, we may be forced to use indirect measures, such as the frequency of extreme climatic conditions, to infer the impact of catastrophes/bonanzas on population growth.

In reality, both catastrophes and bonanzas are likely to exert a range of effects on population growth. That is, not all severe droughts (or favorably wet years) are equally severe (or favorable). If we were omniscient, we could envision incorporating the full range of effects of catastrophes or bonanzas in the following way. In projecting the population forward by one year in a computer simulation, we could first decide whether the next year is to be a "typical" or an "extreme" year, using data on the long-term frequency of extreme environmental conditions (where we would use some criterion to determine the cutoff between "typical" and "extreme," such as whether a transition is an outlier in a regression analysis—see Chapter 3). If the year is to be typical, we would simply draw a population growth rate from the typical distribution of environmental variation. If the year is extreme, we would draw a value from a second distribution of conditions observed in extreme

years. The problem is that censuses of threatened and endangered populations are almost always too short in duration to see more than one or two extreme events (if any), so that we typically have too little information to be able to construct a distribution for the severity of bonanzas or catastrophes or to estimate their true frequencies. Given this limitation, we argue that, if we aim to include catastrophes or bonanzas at all, we should use the actual severities observed in the data (or empirical estimates from other populations or related species), rather than incorporating a distribution of severities that will have poor empirical justification. However, unless your data on these rare events is very good, we also suggest that you run multiple models, ranging across reasonable values for the frequency and severity of both catastrophes and bonanzas (based upon your data). This approach will give a better sense of how much these rare events actually matter for extinction risk, and it will guard against results that depend critically upon very uncertain estimates.

Another problem is to decide whether to include catastrophes only, bonanzas only, or both. For example, in the analysis of the Yellowstone grizzly bear population in Chapter 3, only a single transition (the 1983–1984 transition) out of the 38 transitions in the data set was identified as an outlier. This transition represented an unusually large increase in population size, and so we might view it as a bonanza. But what should we do when constructing a simulation to account for extreme events? We could only include bonanzas occurring at a frequency of about once every 38 years in the simulation, because we never actually observed a catastrophe in the census (where a catastrophe would be an unusually *low* outlier identified by the same regression diagnostics used to identify the bonanza). On the other hand, if we assume that *either* bonanzas or catastrophes occur about once every 38 years, we would not be surprised to see only a single bonanza and no catastrophes in a 38-year period, *even if catastrophes actually do occur*. Because there is no clear answer to this puzzle, we advocate trying both possibilities. That is, run one set of simulations in which only events actually observed are included. For example, if the data show one unusually high outlier (i.e., a bonanza) and two unusually low ones (i.e., catastrophes), draw the magnitudes of extreme events from among these three possibilities. Then run a second set of simulations in which bonanzas and catastrophes are equally likely over the long run (e.g., by duplicating the single bonanza in the set of extreme values from which the simulation chooses). These two analyses should provide reasonable bounds on the possible outcomes, and both can be defended on the basis of the argument presented at the start of this paragraph.

The strategy we just articulated demands that we estimate the magnitudes of both catastrophes and bonanzas. How can we decide how severe a catastrophe might be for the Yellowstone grizzly bear population, when we never actually observed one? The simplest way to answer this question is to

assume that a catastrophe would be just as bad relative to a typical year as the single observed bonanza was good. In the Yellowstone grizzly bear data (Table 3.1), the log population growth rate for the outlying 1983–1984 transition was 0.2683 (which translates into a population growth rate of 1.3077, or a 31% increase over 1 year). If we exclude this year and call the remaining log growth rates typical, their mean is 0.01467. Then, calculating the deviation between the log growth rate during the bonanza year and the mean log growth rate for typical years, and subtracting this deviation from the mean, we arrive at a catastrophic log growth rate (-0.2390) that is just as far below the mean as the bonanza growth rate is above it. Note that we have done this calculation using the log population growth rates, rather than the growth rates themselves, because it was on the log scale that the bonanza year was identified as an outlier. As $\exp\{-0.2390\}$ equals 0.7874, our hypothetical catastrophe represents a 21% decline in population size over one year. To reiterate the logic, we are simply saying here that it would not be implausible for the bear's population growth rate to occasionally drop as far below the average as it has actually been observed to climb above it, and to be cautious we would like to assess the consequences of such infrequent but unfavorable years in case they actually do occur.

We now explore the impact of these estimated bonanzas and catastrophes on the Yellowstone grizzly bear population, using a simple modification of the MATLAB programs we've been using (Box 4.7). The program allows for two different types of years. In "typical" years, the population growth rate is drawn from a lognormal distribution; that is, $\lambda_t = \exp\{x_t\}$ where x_t follows a normal distribution with a mean of μ and a variance of σ^2 . Each year has a probability of 1 minus `probout` of being a typical year, and a probability `probout` of being an outlying (i.e., extreme) year. Thus to mimic the frequency of extreme growth rates for the Yellowstone grizzly bear, we set `probout` to $1/38$. In extreme years, the program draws the growth rate at random from the values stored in the vector `outliers`. To explore the effect of bonanzas only, we place the single value 1.3077 in `outliers`. To account for both bonanzas and catastrophes, we add the value 0.7875 for the hypothetical catastrophe calculated in the preceding paragraph.

Our best estimate for the probability that the Yellowstone grizzly bear population would decline from 99 to 20 adult females at some point in the next 20 years, assuming no catastrophes or bonanzas,¹⁷ is only 8.5×10^{-5} . Adding in only bonanzas at an average frequency of once every 38 years reduces this probability to 4.2×10^{-5} (based on the average of five runs of the program in Box 4.7, with 100,000 realizations in each run; obviously, to calculate such a small probability with greater precision, we would need to

¹⁷That is, we deleted the 25th transition in Table 3.1, computed μ and σ^2 , and substituted the results into Equation 3.5.

BOX 4.7. MATLAB code to calculate extinction risk in the presence of catastrophes and bonanzas.¹

```
% PROGRAM extremes
%   Simulates effect of catastrophes and/or bonanzas on
%   extinction risk. Calculates the probability that a
%   population starting at Nc will hit the extinction threshold
%   Nx at or before t=tmax, using a density-independent model.
%   With probability "probout" each year, the population growth
%   rate is drawn at random from the vector "outliers", which
%   contains growth rates during catastrophes and/or bonanzas

% ***** SIMULATION PARAMETERS *****
tmax=20; % time horizon
numreps=1000000; % number of realizations to simulate
Nc=99; % initial population size
Nx=20; % extinction threshold
mu=0.01467; % mean log population growth rate
sigma2=0.01167; % environmental variance in log pop. growth
% rate
probout=0.0263; % annual probability of an outlying growth
% rate

% vector of observed extreme growth rates
outliers=[0.7875, 1.3077];
%*****
```

```
rand('state',sum(100*clock)); % Seed uniform random number
% generator
randn('state',sum(100*clock)); % Seed normal random number
% generator
sigma=sqrt(sigma2);
numout=length(outliers); % number of outlying growth rates
```

¹Note that this program provides an alternative to using results based on diffusion approximations in Chapter 3. That is, rather than assume the log population growth rate is normally distributed and that large environmental perturbations do not occur, users can easily modify this program to choose population growth rates *only* from among those actually observed. Specifically, set "probout" to 1 and put all the observed population growth rates (*not* log population growth rates) in "outliers". (Of course, other modifications such as eliminating the "if rand>probout" statement would make the program faster.) Also see the program "randdraw" in Box 2.1.

BOX 4.7 (continued)

```
% Fill the matrix "lambdas" with population growth rates
for i=1:tmax % Looping over future times
    for j=1:numreps % and all realizations,
        if rand>probout % if year is 'typical'
            lambdas(i,j)=exp(mu+sigma*randn); % draw a 'typical'
            % lambda, but
        else
            lambdas(i,j)=... % draw lambda
            outliers(floor(numout*rand)+1); % from "outliers".
        end;
    end;
end;

N=Nc*ones(1,numreps); % Start all realizations at Nc,
for t=1:tmax % project all realizations to
    N=N.*lambdas(t,:); % t=tmax,
    N=N.*(N>Nx); % set to zero any realization at or
end; % below Nx,
ProbExt= sum( (N<=Nx) )/numreps % and compute
% probability of
% quasi-
% extinction.
```

simulate many more realizations per run). When we include the possibility that both bonanzas and catastrophes may occur, the average probability of hitting the quasi-extinction threshold increases to 1.7×10^{-4} , which is still a rather small risk. In retrospect, it is not very surprising that including catastrophes and bonanzas of the observed (or at least justifiable) severity has little effect on the estimated short-term extinction risk. The chance that one of these extreme events occurs even once in a 20 year period is low, given the observed frequency of occurrence of once in 38 years, and although they represent outliers compared to the bulk of the observed population growth rates, their magnitude is not really all that extreme. Nonetheless, other populations may very well experience more extreme or more frequent catastrophes or bonanzas, which could have a strong impact (negative or positive) on population viability. We have now seen how such events can be incorporated into a PVA if their frequency and severity can be estimated.

While we've been discussing extreme years that are uncommon, for some populations typical environmental variability may be large enough that the

usual assumption of normally distributed log growth rates may not be a good description of the true variation. This may be especially true for short-lived species living in highly variable environments, where most years are either a boom or a bust. Desert annuals, rodents, and some insect species fall into this category, with the annual population growth rate showing peaks at high and low values with few years in between. In these cases, more realistic predictions of extinction probabilities can sometimes be generated by using the observed growth rates themselves, rather than trying to estimate a mean and variance for the growth rate and drawing the simulated growth rates from a normal distribution. This is relatively easy to do in the absence of density dependence, and more difficult if density dependence does operate. Here is a procedure to assess whether extreme variation is common enough to deal with in this way:

1. Test for density dependence, as discussed earlier in the chapter.
2. If the best model is density-independent, use the methods in Chapter 3 to remove points that are extreme outliers.
3. Use a Lilliefors test to determine if the remaining log population growth rates are normally distributed. The Lilliefors test is a version of the one-sample Kolmogorov-Smirnov test that evaluates whether a set of numbers is likely to have come from a normal distribution. This routine is part of most statistics packages, often as an option under a Kolmogorov-Smirnov procedure.
4. If the test indicates there is very little chance that the log growth rates are truly normal (say, $p < 0.05$), you should reject the usual approach of generating log growth rates from a normal distribution with mean and variance μ and σ^2 . Instead, simply draw a growth rate at random from the sample of observed rates in each year of a stochastic simulation.¹⁸

If your data do show significant density dependence, you can follow the same general procedure just outlined, but instead test the deviations between observed and predicted log population growth rates for normality (again using the Lilliefors test). However, now your simulation must add to next year's log population size predicted by the density-dependent model a random draw from the observed *deviations*.

Concluding Remarks

This chapter has reviewed methods to incorporate density dependence, demographic stochasticity, environmental correlations, and catastrophes and bonanzas into count-based PVAs. Even though we've tried to indicate how

¹⁸See the footnote to Box 4.7 for suggestions on how to do this.

to include different combinations of these effects in the same PVAs, we haven't done all the possible combinations. However, given the models and methods we have covered, it should be easy for you to put together many other combinations. As usual, your ability to include combinations of these factors will be limited not so much by an inherent difficulty in doing so, but more by the lack of data needed to estimate the strength of each factor for a single population. If such information is available for your focal population, you can view the fragments of MATLAB code presented in this chapter as building blocks from which a more comprehensive, multi-factor viability analysis can be constructed.

There are two additional problems we have not discussed in this chapter. The first is temporal trends in population parameters, that is, deterministic changes in the mean and variance of the population growth rate or the strength or pattern of density dependence over time. If you have the data to estimate how these changes will occur, simple modifications of the simulation models we have discussed will let you predict population dynamics for these situations (see *Temporal environmental trends* in Chapter 3 for a discussion of this issue). The second problem is much more difficult: how to untangle the effects of observation error from true variability in population numbers and hence growth rates. The next chapter is devoted to this problem.

Appendix: An Overview of Maximum Likelihood Parameter Estimation

The likelihood is simply the probability of obtaining a particular set of data given specified values of the parameters in a particular model. It makes intuitive sense that to find the best parameter values, we should seek to maximize this probability (i.e., we maximize the likelihood). To calculate the probability of obtaining the entire data set, we must first specify the probability that a single data point differs by a given amount from the value the model predicts it should be, given the parameter values and the values of the independent variables. For example, consider the checkerspot data in Table 4.1. For an observed value of N_t , the value of the log population growth rate $\log \lambda_t = \log(N_{t+1}/N_t)$ predicted by the Ricker model is $r(1 - N_t/K)$ (see Equation 4.6). The deviation between the observed and predicted log population growth rates is $D_t = \log \lambda_t - r(1 - N_t/K)$. Let us consider such deviations to follow a normal distribution with a mean of zero and a variance of V_r (the residual variance)¹⁹. The probability of obtaining a deviation of size D_t is then governed by the normal probability density:

¹⁹Although we use normally distributed deviations here, we use maximum likelihood methods based on other probability distributions in Chapters 7, 8, and 11.

$$p(D_t | r, K, V_r, N_t) = \frac{1}{\sqrt{2\pi V_r}} \exp\left\{-\frac{(D_t)^2}{2V_r}\right\} = \frac{1}{\sqrt{2\pi V_r}} \exp\left\{-\left[\log \lambda_t - r(1 - N_t/K)\right]^2 / 2V_r\right\} \quad (\text{A1})$$

Here, we use the notation $p(D_t | r, K, V_r, N_t)$ to emphasize that the probability of obtaining a deviation of size D_t depends on the values of the parameters r and K , the residual variance, and the independent variable N_t . The values of $\log \lambda_t$ and N_t are fixed features of the data set; the goal of maximum likelihood estimation is to identify the values of r and K that maximize (A1) over all the observed values of $\log \lambda_t$ and N_t . Some values of r and K will cause the value of $r(1 - N_t/K)$ to fall closer to one particular data point than to others. We want to choose the values of r and K resulting in predicted values that fall as close as possible to the greatest number of observed data points. Doing so will minimize the residual variance.

To find the values of r and K that provide the best fit to all the data points, we first compute the probability of obtaining the entire data set by computing the product of the probabilities of obtaining each data point:

$$L = \prod_{t=1}^q p(D_t | r, K, V_r, N_t) \quad (\text{A2})$$

where q is the number of observations of the population growth rate in the data set (one less than the number of censuses). L is the likelihood of obtaining the data given the model. Because L is a product of many, potentially small, probabilities, it will typically be a very small number, which can be difficult to calculate accurately on a computer. Therefore, because the log of a small number is larger (in absolute value) than the number itself, it is convenient to take the logs of both sides of Equation A2 to yield the log likelihood function:

$$\log L = \sum_{t=1}^q \log[p(D_t | r, K, V_r, N_t)] = \sum_{t=1}^q \left[\log(2\pi V_r)^{-1/2} - \frac{1}{2V_r} \left[\log \lambda_t - r \left(1 - \frac{N_t}{K}\right) \right]^2 \right] \quad (\text{A3})$$

Because the first term in the right-hand sum is independent of t , and there are q such terms, Equation A3 is the same as Equation 4.5.

Thus the best values of r and K are those that maximize Equation A3. There are several ways to find these values. The simplest way is to make a contour plot of the log likelihood versus r and K , and look for "peaks" in the likelihood surface. With finer and finer grids of r and K values, one can hone in on the best parameter values (the so-called maximum likelihood estimates), which are those that correspond to the highest peak in the likelihood surface. This graphical method for identifying the maximum likelihood estimates works well only if there are only one or two parameters to be estimated. Or one could take the derivatives of Equation A3 with respect to each of the parameters (e.g., r and K), set them equal to zero, and solve the resulting

equations simultaneously to give the maximum likelihood estimates. To be sure that these values correspond to a maximum, not a minimum, in the log likelihood function (i.e., Equation A3), one should check that the second derivatives of Equation A3 are negative at these values. In the many cases in which the system of equations that results from setting the derivatives to zero is too complex to solve analytically, one can use gradient-based methods, which rely on expressions for the derivatives of Equation A3 to always move "uphill" on the log likelihood surface towards a maximum. This is the method that the SAS nlin procedure uses (see Box 4.2). Finally, one could use nongradient-based methods such as the simplex algorithm (described well in Press et al. 1995), which is implemented in MATLAB's fminsearch function.²⁰ Because in general the log likelihood function may have more than one maximum, it is important to check, once a maximum is found, that it is the global, not a local, maximum. For example, one should start gradient-based and nongradient-based searches at numerous starting values for the parameters, and choose the resulting set of parameter values (if there is more than one) that has the highest log likelihood.

²⁰Note that fminsearch is a minimization routine, and so one must use it to find the parameter values that minimize the negative of the log likelihood function.