STOCHASTIC POPULATION DYNAMICS

Sebastian Schreiber
## Contents

Chapter 1. Introduction 4  
  1. What and Who? 4  
  2. Tales from a simple model 5  
Exercises 12  

Part I. Environmental Stochasticity 13  

Chapter 2. Random matrix models 14  
  1. Random matrix products 14  
  2. Sensitivity of $r$ 21  
  3. General random products 27  
  4. Random affine models 27  
Exercises 33  

Chapter 3. Nonlinear single species models 35  
  1. Unstructured models 35  
  2. Structured models 42  
  3. Ergodicity and stationary distributions 45  
  4. Exercises 48  

Chapter 4. Multispecies models 49  
  1. The lottery model 49  
  2. Two species competition 52  
  3. Many species 58  

Chapter 5. Approximation methods 65  
  1. The scalar case 65  
  2. The Multivariate Case 68  

Index 70  
Bibliography 71
CHAPTER 1

Introduction

1. What and Who?

1.1. What are these notes about? Since the pioneering work of Lotka [1925] and Volterra [1926] on competitive and predator–prey interactions, Thompson [1924], Nicholson and Bailey [1935] on host–parasite interactions, and Kermack and McKendrick [1927] on disease outbreaks, nonlinear difference and differential equations have been used to understand population dynamics. However (with apologies to John Gay), lest biologists believe ones theory untrue, models need to keep probability in view. All populations experience demographic and environmental stochasticity. Demographic stochasticity stems from populations and communities consisting of a finite number of interacting individuals whose fates aren’t perfectly correlated. Some individuals die while others survive as if coin flips determined their fates. Environmental stochasticity stems from fluctuations in environmental conditions such as temperature and precipitation, or large stochastic shocks such as floods, fires, droughts, and hurricanes. As these environmental conditions influence survival, growth, and reproduction, these environmental fluctuations result in demographic fluctuations that influence population dynamics.

These stochastic forces can impact population dynamics in a diversity of ways. Most fundamentally, for populations without immigration, demographic stochasticity in a finite world implies the population will go extinct in finite time, contrary to predictions from most deterministic models. This simple observation leads to a host of questions: What is the likelihood of extinction next year versus 50 years from now? Under what conditions is extinction going to be more rapid? For interacting populations, which species or genotypes are likely to be lost first? Despite extinction being inevitable, populations may persist for exceptionally long periods, in which case the nature of these transient dynamics are of interest. One may ask: To what extent are the transient dynamics driven by underlying deterministic feedbacks? If so, how frequently, do stochastic forces causes the population state to drift between different attractors of the underlying deterministic dynamics?

When populations are sufficiently large to diminish the importance of demographic stochasticity, the interplay between deterministic feedbacks and environmental fluctuations can play a crucial role in determining species richness and genetic diversity [Gillespie, 1973, Chesson and Warner, 1981, Turelli, 1981, Chesson, 1994, Ellner and Sasaki, 1996, Abrams et al., 1998, Bjornstad and Grenfell, 2001, Kuang and Chesson, 2008, 2009]. For example, competition for limited resources [Gause, 1934] or sharing common predators [Holt, 1977] may result in species or genotypes being displaced. However, random forcing of these systems can reverse these trends and, thereby, enhance diversity [Gillespie and Guess, 1978, Chesson and Warner, 1981, Abrams et al., 1998]. Conversely, differential predation can mediate coexistence between competitors [Paine, 1966, Holt et al., 1994, Chesson and Kuang, 2008], yet environmental fluctuations can disrupt this coexistence mechanism. If species do coexist, one might ask: How do species interactions and environmental stochasticity determine the mean abundance of a species? Which species exhibit the greatest demographic fluctuations? Which species densities covary most closely?

The goal of these lectures is to describe a theory of stochastic population dynamics in a mathematically rigorous, but practical manner. The theory consists of a mixture of analytical characterizations of short-term and long-term behavior of the stochastic population dynamics (e.g. existence of a stationary distribution supporting all populations), analytical and numerical methods for verifying these
conditions (e.g. small noise approximations), numerical methods for exploring finer details of the stochastic dynamics (e.g. extinction likelihoods), and analytical approximations for these finer details (e.g. diffusion type approximations).

1.2. Who is this written for and how much do you need to know? There are two target audiences for these lecture notes: theoretical population biologists and applied mathematicians. To reach both audiences, the main presentation provides the minimal technical details to get the main punchlines. Proofs are given in the main text to the extent they provide useful insights into the main ideas. Additional technical details and background will be provided in the Appendices.

All ideas are motivated and illustrated by applications to real models from the biological literature. The impact of the theory on the models is developed in detail to highlight the biological implications and the many challenges that these models still contain.

The lectures assumes all readers are familiar with some basic concepts from dynamical systems and probability theory. Specifically, from dynamical systems theory, readers should be familiar with difference and differential equations, linearization and stability of equilibria, and attractors. From probability, familiarity with expectations, probability distributions, the law of large numbers, the central limit theorem, and basic finite state Markov chain theory will be useful. For the more advanced readers that want to get into the mathematical details, familiarity with real analysis, measure theory, Markov chains on general state spaces, large deviation theory, and martingales will be useful.

2. Tales from a simple model

To introduce some of the themes for these lectures, we consider two stochastic variants of the Beverton-Holt model for a single species. In this model \( n_t \) denotes the population density at time \( t \).

At low densities, individuals produce on average \( R \) offspring to replace themselves. Density-dependent feedbacks result in a reduction in the fitness of the individuals. For the Beverton-Holt, this reduction is of the form \( R/(1 + an_t) \) where \( a > 0 \) is a measure of the strength of intra-specific competition. Thus, the population dynamics are given by

\[
(1.1) \quad n_{t+1} = n_t \frac{R}{1 + an_t}.
\]

The dynamics of this classical are well-known. If \( R > 1 \) (i.e. at low densities individuals more than replace themselves) and \( n_0 > 0 \), then

\[
\lim_{t \to \infty} n_t = \frac{R - 1}{a}
\]

where \((R - 1)/a\) is the unique positive equilibrium of the model. Alternatively, if \( R \leq 1 \) and \( n_0 \geq 0 \), then \( \lim_{t \to \infty} n_t = 0 \). Notice that extinction in this latter case only occurs in the limit as time marches to \( \infty \). For all finite time, the population density is positive provide it initially was positive.

2.1. Demographic stochasticity. A fundamentally strange aspect of the classical Beverton-Holt model is \( n_t \) can be any non-negative real number. We skated around this issue by calling \( n_t \) a "density" but even densities can not take on any value in a finite world. Namely, for given habitat size \( S \), the possible densities are 0, 1/S, 2/S, \ldots. So what in the world does the model represent? Here we briefly consider this question.

Natural populations consist of a finite number, discrete individuals. To account for this feature in the model, let's consider an individual based model where a simple set of individual-based rules determine the population dynamics. To this end, let \( N_t \) be the number (not density!!) of individuals at time \( t \). These individuals live in some part of the world with a finite amount of space. Let \( S \) be the total amount (e.g. area or volume) of this space. Then \( N_t/S \) is the density of individuals. Consistent with the deterministic model, let's assume that each individual produces on average \( R/(1 + aN_t/S) \) number of offspring. But how is this average achieved? While there a countless number of answers to
2. TALES FROM A SIMPLE MODEL

this questions, lets keep the answer simple by assuming that each individual, independent of the others produces a Poisson number of offspring with mean \( R/(1 + aN_t/S) \). During the time step between \( t \) and \( t+1 \), let \( X_i \) be the number of offspring that individual \( 1 \leq i \leq N_t \) produces to replace themselves. Then, the population size in the next time step is

\[
N_{t+1} = X_1 + \cdots + X_{N_t}.
\]

Equivalently, \( N_{t+1} \) is a Poisson distributed random variable with mean \( N_tR/(1 + aN_t/S) \). What can one say about such a model?

Density-independence. Consider the very special case where there is no density-dependent feedback in the population dynamics i.e. \( a = 0 \). Then our stochastic model is an example of a classical branching process which we discuss later in greater detail (see XX.XXX). A basic result from this classical theory implies that if the mean number of offspring \( R \) is less than or equal to one, then \( N_t \) goes to 0 in finite time with probability one (see left hand panel of Figure 1). Unlike the deterministic model, extinction occurs in finite time.

On the other hand, if the mean number of offspring \( R \) is strictly greater than one, then there is a positive probability that \( N_t \) goes to 0 in finite time and a positive complementary probability that \( \lim_{t \to \infty} N_t/R \) exists and is positive (see right hand panel of Figure 1). In particular, the population size grows to infinity at a geometric rate on this later event. Unlike the deterministic model, extinction still occurs despite individuals replacing themselves on average. However, the probability of extinction occurring decreases exponentially with population size.

Of course, exponential growth can not continue unabated. In the words of economist Kenneth Boulding,

"Anyone who believes that exponential growth can go on forever in a finite world is either a madman or an economist."

Density-dependence. To avoid being called a madman (or perhaps worse an economist), suppose that there is a density-dependent feedback in the model i.e \( a > 0 \). Figure 2 illustrates several simulations of the dynamics when \( R = 1.1 \) for different habitat sizes. These simulations illustrate that as the habitat size \( S \) gets larger the stochastic dynamics of the densities \( N_t/S \) follow the deterministic dynamics more and more closely. In fact, what one can show (see Section XX.XXX) is that for any \( \epsilon > 0 \) and \( T > 0 \) and \( N_0/S = n_0 \), the stochastic dynamics \( N_t \) and the deterministic dynamics \( n_t \) satisfy

\[
\mathbb{P}[|N_t/S - n_t| \leq \epsilon] \geq 1 - \epsilon \quad \text{for} \quad 0 \leq t \leq T
\]
Figure 2. Multiple realizations of the density-dependent branching process with $S = 20$, 100, 1,000, and 10,000. Parameters: $N_0 = S/20$, $a = 0.1$, $R = 1.1$. 

provided that $S$ is sufficiently large. In words, for any interval of time $T$, the stochastic dynamics are highly likely to be close to the deterministic dynamics provided the habitat size is sufficiently large. This close correspondence, however, ultimately breaks up as the stochastic model goes to extinction in finite time with probability one i.e. $\mathbb{P}[N_t = 0$ for some $t] = 1$.

While extinction is inevitable for $N_t$, the time to extinction tends to increase exponentially with the habitat size $S$ and, consequently, can be an extremely long time for even moderate habitat sizes $s$(see upper left hand panel of Figure 3). Therefore, it is natural to ask what can we say about this “meta-stable” behavior? It turns out there is a fair amount we can say. For example, the metastable behavior is characterized by so-called quasi-stationary distributions which describe the long-term behavior of the model conditioned on non-extinction. One way of thinking about the quasi-stationary distribution is that there is a random variable, call $\hat{N}$, taking values on the natural numbers such that

$$\lim_{t \to \infty} \mathbb{P}[N_t = i | N_t \not= 0] = \mathbb{P}[\hat{N} = i]$$

Figure 2 plots quasi-stationary distributions for several habitat sizes. In these lectures, we will discuss how to compute these quasi-stationary distributions, to relate them to the dynamics of deterministic models when $S$ is large, how to compute “intrinsic” extinction statistics associated with these quasi-stationary distributions, and how to approximate them when $S$ is large but not too large. In particular, we will answer the following question raised by Peter Jaegers
2.2. Environmental stochasticity. While demographic stochasticity is an "intrinsic" form of stochasticity (i.e., generated within the system itself), environmental stochasticity is an "extrinsic" form of stochasticity (i.e., an "external forcing" of the system). If the habitat size $S$ is sufficiently large (really infinite), then we can ignore demographic stochasticity and focus on the external forcing of the system. For example, in our Beverton-Holt model, the population may be experiencing environmental fluctuations which influence the maximal reproductive output $R$ of all individuals, e.g., fluctuations in precipitation for annual plants. We can model this external forcing with the following non-autonomous difference equation:

\[
R_{t+1} = R_t \left(1 + \frac{a}{R_t}ight)
\]

where $R_{t+1}$ represents how the environmental conditions from time $t$ to time $t+1$ determine the fecundity of individuals at low population densities. Environmental fluctuations could lead to fluctuations in the density of the population, which in turn can affect the population's ability to maintain stable asymptotic behaviors.

"Any population allowing individual variation in reproduction, ultimately dies out—unless it grows beyond all limits, an impossibility in a bounded world. Deterministic population mathematics on the contrary allows stable asymptotics. Are these artifacts or do they tell us something interesting about quasi-stationary stages of real or stochastic populations?"

Figure 3. Intrinsic mean times to extinction and quasi-stationary distributions for the density-dependent branching process with $S = 10, 100,$ and $500$. Parameters: $a = 0.1$, $R = 1.1$.
2. TALES FROM A SIMPLE MODEL

Figure 4. Realizations of $n_{t+1} = R_{t+1} n_t$ with $R_1 = 2$ and $1/5$ with equal probability. The dashed line on the right is the expected values of the log density.

Intraspecific competition coefficient $a$, but for simplicity, lets ignore this possibility for now. If the $R_t$ are sequence of random variables, then (1.2) is an example of a stochastic difference equation. For now, lets assume that $R_1, R_2, \ldots$ are independent and identically distributed.

2.2.1. Density-independent model. In the absence of density dependent feedbacks i.e. $a = 0$, our stochastic difference equation (1.2) becomes

$$n_{t+1} = R_{t+1} n_t$$

The solution of this difference equation is (check for yourself!)

$$n_t = R_t R_{t-1} \ldots R_1 n_0$$

What can we say about the long-term behavior of this equation? If we are interested in the expected population density $E[n_t]$ and $n_0$ is given, then

$$E[n_t] = E[R_t R_{t-1} \ldots R_1 n_0]$$

$$= E[R_t R_{t-1} \ldots R_1] n_0$$

$$= E[R_t] E[R_{t-1}] \ldots E[R_1] n_0$$

by independence of the $R_t$

$$= E[R_1]^t n_0$$

as $R_t$ are identically distributed

Hence, the expected population density grows by a factor $E[R_1]$ every time step. This suggest that if $E[R_1] > 1$, then the population grows and persists, and otherwise it would decrease toward extinction. This turns out to be misleading. For example, suppose that $R_1 = 2$ and $1/5$ with equal probability. Then $E[R_1] = 1.1$ and we expect the population to grow. However, the left hand side of Figure 4 shows 20 simulations of $n_t$ and in all of them, the population size is tending toward zero. Why?

To understand the discrepancy between the simulations and the expectation that we computed in the previous example, we can make use of the Law of Large Numbers. Let $x_t = \log n_t$. Then

$$x_t = \log(R_t R_{t-1} \ldots R_1 n_0)$$

$$= \sum_{s=1}^t \log R_s + \log n_0$$

If $\log R_1$ has finite mean $r$ and finite variance, then the Law of Large Numbers implies

$$\lim_{t \to \infty} \frac{x_t}{t} = E[\log R_1] = r$$

with probability one.
Hence, if $r < 0$, then $\log n_t \to -\infty$ with probability one, and if $r > 0$, then $\log n_t \to \infty$. How does this help with our specific example where $R_t = 2$ and $1/5$ with equal probability? For this example, $\mathbb{E}[\log R_1] = \frac{1}{2} \log \frac{2}{5} < 0$. Hence, $n_t \to 0$ with probability one. Indeed the right hand side of Figure 4 illustrates this trend by plotting many realizations on a log scale. This difference stems from the fact that the geometric mean of $R_1$, $e^r$, is smaller than the arithmetic mean $\mathbb{E}[R_1]$. Thus, we have rediscovered the main conclusion of a PNAS paper by Lewontin and Cohen [1969] who wrote

Even though the expectation of population size may grow infinitely large with time, the probability of extinction may approach unity, owing to the difference between the geometric and arithmetic mean growth rates.

2.2.2. Density-dependent model. Again to avoid being called a madman or economist, lets add some density-dependence back into the model i.e. allow $a > 0$. Some simulations for this model with log-normally distributed values of $R_t$ are shown in Figure 5. What might we be able to say about this model? For populations to persist, it seems that it would necessary that they have a tendency to increase at low densities. But at low densities the model is given approximately the density-independent model $n_{t+1} = R_{t+1} n_t$ which exhibits a tendency of population increase only if $r = \mathbb{E}[\log R_1] > 0$. For the stochastic Beverton-Holt model (see XX.XX), the condition $r > 0$ implies that there is a positive stationary distribution. Namely, a positive random variable $\hat{n}$ such that if $n_0$ is distributed like $\hat{n}$, then $n_t$ is distributed like $\hat{n}$ for all time i.e. $\mathbb{P}[a \leq n_t \leq b|n_0 = \hat{n}] = \mathbb{P}[a \leq \hat{n} \leq b]$ for all $a < b$.

Similar to the basic limit theorems for finite-state Markov chains, this stationary distribution $\hat{n}$ characterizes the long-term behavior of the model in two ways. First, given $n_0 > 0$, $n_t$ converges to $\hat{n}$ in probability: for all intervals $[a, b]$,

$$\lim_{t \to \infty} \mathbb{P}[a \leq n_t \leq b] = \mathbb{P}[a \leq \hat{n} \leq b].$$

As $\hat{n}$ is a positive random variable, this implies that $n_t$ is unlikely to be close to 0 for large $t$. Furthermore, this result also implies that the long-term statistical behavior of $n_t$, from the ensemble perspective, is independent of $n_0$ (provide it is positive.)

The second way $\hat{n}$ characterizes the long-term behavior is from the “typical trajectory” perspective. Namely, given any interval $[a, b]$ and $n_0 > 0$, the asymptotic fraction of time spent in the interval by a “typical” realization of the stochastic process equals the probability that $\hat{n}$ is the interval. More precisely,

$$\lim_{t \to \infty} \frac{\# \{1 \leq s \leq t \text{ such that } a \leq n_s \leq b\}}{t} = \mathbb{P}[a \leq \hat{n} \leq b].$$

Note the important implication for simulations. To estimate the stationary distribution it suffices (with probability one) to run the model one time for a long time. Both of these forms of convergence are illustrated in Figure 5.

When the environmental fluctuations are sufficiently small, we can approximate the long-term behavior by linearizing the system around the stable equilibrium of the deterministic system (see Section XX.XX). These approximations yield another type of linear stochastic difference equation for which we can say a lot about their behavior. The approximation for the stationary distribution is shown as red curves in Figure 5.

Finally, what if $r < 0$? Then, $\lim_{t \to \infty} n_t = 0$ with probability one. However, unlike the case of demographic stochasticity, a density of zero is only approached in the infinite time horizon.

2.3. Where to hence? This introduction provided a quick overview about the types of results that are the focus of these lectures. The remainder of these lectures will flush out the details in a sufficiently general context to be applicable to models accounting for population structure (e.g. spatial, age, or stage structure) and interactions between multiple populations (e.g. predator-prey interactions in ecological communities or competition among multiple genotypes). Furthermore, we will examine models which simultaneously account for demographic and environmental stochasticity.
Figure 5. Runs of BH with environmental stochasticity. Histograms for first two rows correspond to a single run of the model. Final histogram corresponds to the distribution at time 5,000 for 1,000 runs of the model. Red curves show an analytically based approximation of the stationary distribution. Parameters: $N_0 = S/20$, $a = 0.1$, $R = 1.1$. 
Exercises

(1) For the demographic stochasticity version of the Beverton-Holt model, prove that for all \( \epsilon > 0 \), \( n_0 \geq 0 \), and \( T \geq 1 \), there exists \( S^* \) such that

\[
P[|n_t - N_t / S| \leq \epsilon | N_0 = n_0 S] \geq 1 - \epsilon \text{ for } 0 \leq t \leq T
\]

whenever \( S > S^* \). Hint: Use Chebyshev’s inequality.

(2) Recently, human activities have isolated the acorn woodpecker population in Water Canyon from other acorn woodpecker populations. This isolation coupled with their recent decline to ten breeding pairs has caused an outcry. People fear, extinction is just around the corner for this beloved bird. The Nature Conservancy has hired you as a consultant to evaluate various intervention strategies. As a first step to assess these strategies, consider a density-independent model of the form

\[
n_{t+1} = (j_{t+1} f_{t+1} + a_{t+1}) n_t
\]

where \( n_t \) is the density of females in year \( t \), \( j_{t+1} \) is the fraction of juveniles that survive to year \( t + 1 \), \( a_{t+1} \) is the fraction of adults that survive to year \( t + 1 \), and \( f_{t+1} \) is the average number of daughters produced per female in year \( t + 1 \). To tackle this issue, the Nature Conservancy has given you the following data from Stacey and Taper [1992]

\[
\begin{align*}
    j &\sim (0.56, 0.64, 0.3, 0.4, 0, 0.38, 0.18, 0.25, 0.44) \\
    a &\sim (0.53, 0.68, 0.71, 0.38, 0.54, 0.69, 0.66, 0.49, 0.61) \\
    f &\sim (3.38, 1.27, 2.77, 2.17, 0.05, 4.0, 2.37, 0.5, 1.6)/2
\end{align*}
\]

(a) Determine the \( r \) value for this population. Should the population be increasing or decreasing in the long-term?
(b) How much do you need to increase juvenile survivorship to ensure the population will have a positive \( r \)?
(c) How much do you need to increase adult survivalship to ensure the population will have a positive \( r \)?
(d) If there is only enough money to support one conservation effort, discuss which one you would support and why.
Part I

Environmental Stochasticity
CHAPTER 2

Random matrix models

Not only the adult, but the whole life cycle will be considered the organism. This is an ancient notion, for philosophers have often pointed out that an individual conventionally means an organism in a short instant of time...For example, if we refer to a “dog” we usually picture in our minds an adult dog momentarily immobilized in time as though by a photographic snapshot...[But] is the dog not a dog from the moment of fertilization of its egg, through embryonic and fetal development, through birth and puppyhood, through adolescence and sexual maturity, and finally through senescence? – John Tyler Bonner (1965)

1. Random matrix products

Consider a population consisting of individuals in \( k \) different states or stages (e.g. age, geographical locations, behavioral type). Let \( n_t = (n_{1,t}, \ldots, n_{k,t}) \) be the vector of densities of individuals in the \( k \) states. Each time step, surviving individuals for one state contribute to other states by transitioning to a different state (e.g. growing older, dispersing to another patch, switching behavioral states) or by producing offspring. Let \( A_{t+1} = (s_{ij,t+1}) \) be a \( k \times k \) matrix that captures this contributions. Specifically, the entry \( A_{ij,t+1} \) in the \( i \)-th row and \( j \)-th column of \( A_t \) is the contribution of individuals of state \( i \) to individuals in state \( j \). With this notation, we get the population dynamics are given by

\[
n_{t+1} = n_t A_{t+1}
\]

As \( A_{ij,t} \) correspond to contributions from one state to another, we assume that these entries are non-negative (we discuss the more general case in Section ??). \(^1\)

The solution of (2.1) is given by iteration:

\[
n_t = n_0 A_1 A_2 \ldots A_t
\]

What can we say about the long-term behavior of this solution? We consider two cases: a constant environment and a fluctuating, but stationary, environment.

1.1. Temporally Homogeneous Case. An important special case occurs when the \( A_t \) are constant in time i.e. \( A_t = A \) for all \( t \). We say \( A \) is primitive if there exists \( t \) such that \( A^t \) only has positive entries. Namely, after \( t \) time steps, any subpopulation contributes positively to any subpopulation. Most (but not all) matrix models in demography are primitive. A classical result for primitive matrices is as follows:

\(^1\)We have chosen to write down models using left rather than right multiplication which is the conventional choice in many ecological modelling texts. Here are five reasons why:

1. The \( i-j \) entry of \( A_t \) corresponds to the contribution of state \( i \) to state \( j \).
2. Equations read naturally from left to right: future population state equals the current population state projected forward by \( A_{t+1} \).
3. Row vectors display nicer in a paragraph than column vectors
4. When plotting solutions in time with matplotlib, the \( t \)-th row is \( N_t \).
5. This is how probabilists do it.
Theorem 2.1. (Perron-Frobenius) If $A$ is a non-negative, primitive matrix, then there exists a pair of unique, positive row vectors $v$ and $w$, and a unique positive real number $\lambda$ such that

- $\sum_i v_i = 1$ and $\sum_i v_i w_i = 1$, and
- if $n_0$ is nonnegative and non-zero, then
  $$\lim_{t \to \infty} n_t / \lambda^t = (n_0 w^T)v$$

where $^T$ denotes the transpose of a vector. Loosely, we have the approximation
  $$n_t \approx \lambda^t (n_0 w^T)v$$

$\lambda$ is known as the dominant eigenvalue of $A$ with associated left and right eigenvectors $v$ and $w^T$ i.e. $vA = \lambda v$ and $Aw^T = \lambda w^T$. Each of these quantities has a biological interpretation. $v$ is the stable state distribution of the population as it describes the fraction of the population in each state in the long-term. $w$ is the vector of reproductive values as it describes the contributions of the initial number of individuals in each state to the long-term population size. Finally, $\lambda$ is the long-term ”growth rate” in the sense that it determine the factor by which the population expands or shrinks per time step in the long-term. About the reproductive values, Ronald Fisher wrote

We may ask, not only about the newly born, but about persons of any chosen age, what is the present value of their future offspring; and if the present value is calculate at the rate determined as before, the question has definite meaning – To what extent will persons of this age, on average, contribute to the ancestry of future generations? The question is one of some interest, since the direct action of Natural Selection must be proportional to this condition. – Ronald A. Fisher (1930)

Example 2.1. To illustrate the use of the Perron-Frobenius theorem, lets consider a matrix model of loggerhead turtles developed by Crouse et al. [1987]. Their model broke the population into 7 stages: hatchlings, yearlings, juveniles, sub-adults, first time reproducers, remigrants, mature adults. The matrix for their model is given by

$$A = \begin{pmatrix}
0 & 0.6747 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.737 & 0.0486 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.6610 & 0.0147 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.6907 & 0.0518 & 0 & 0 \\
127 & 0 & 0 & 0 & 0 & 0.8091 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & 0.8091 \\
80 & 0 & 0 & 0 & 0 & 0 & 0.8089
\end{pmatrix}$$

The top row of Figure 1 shows the stable state distribution and reproductive values. On the left of the bottom row, the dynamics of fraction $n_{i,t} / \sum_j n_{j,t}$ in each state converges to the stable state distribution within 25 years. On the right hand side, the population density normalized by the growth rate, $\sum_i n_{i,t} / \lambda^t$, converges to the reproductive value of a single adult.

1.2. Stationary Environments. Now assume $A_1, A_2, \ldots$ is a stationary, sequence of random matrices. Stationarity means that for all $t$ the shifted sequence $A_t, A_{t+1}, A_{t+2}, \ldots$ and the original sequence $A_1, A_2, A_3, \ldots$ have the same distribution. That is

$$\mathbb{P}[A_t \in A_1, \ldots, A_{t+n} \in A_n] = \mathbb{P}[A_1 \in A_1, \ldots, A_n \in A_n]$$

for all Borel sets $A_0, \ldots, A_n$ in $\mathbb{R}^{k^2}$ (i.e. the space of all $k \times k$ matrices). Ergodicity means, intuitively, the environmental sequence is not a mixture of two other stationary distributions. \footnote{More precisely, for any bounded Borel function $H : \mathbb{R}^{k^2} \times \mathbb{R}^{k^2} \times \ldots \to \mathbb{R}$, $\mathbb{P}[H(A_1, A_2, \ldots) = H(A_2, A_3, \ldots)] = 1$ implies that $H(A_1, A_2, \ldots)$ is almost surely constant.} An important
Figure 1. Stable state distribution and reproductive values for the loggerhead model. Convergence of the dynamics of $n_t$ is shown for $n_0 = (0, 0, 0, 0, 0, 1)$ i.e. one adult.

consequence of ergodicity is that if $\mathbb{E}[|h(A_1)|] < \infty$ for some function $h : \mathbb{R}^{k^2} \to \mathbb{R}$, then the temporal average of $h$, $\frac{1}{t} \sum_{i=1}^{t} h(A_i)$, converges to its expected value $\mathbb{E}[h(A_1)]$ with probability one.

This representation of the environmental dynamics allows for a diversity of possibilities as the following examples illustrate.

**Example 2.2. (Periodic environments)** Let $\{B_1, B_2, \ldots, B_n\}$ be non-negative matrices corresponding to the population dynamics in different years or seasons. $A_0$ is chosen with equal likelihood for these $n$ matrices and $A_{t+1} = B_{i+1 \mod n}$ on the event that $A_t = B_i$. Then the environmental dynamics correspond to randomly choosing an initial environmental state and successive environmental states changing periodically. This leads to a stationary and ergodic sequence of matrices.

**Example 2.3. (Independent and identically distributed environments)** $A_1, A_2, A_3, \ldots$ are a sequence of independent and identically distributed random matrices.

**Example 2.4. (Finite Markovian environments)** Suppose the environmental dynamics are characterized by a finite number of matrices $\{B_1, \ldots, B_m\}$ and the transition probabilities between these matrices are given by the transition matrix $P$ i.e. the $i$-$j$-th entry $P_{ij}$ of $P$ is the probability of going from the matrix $B_i$ to the matrix $B_j$. If $P$ is irreducible, there is a unique stationary distribution $\mu$ for $P$. Namely, there is a row vector $\mu$ such that $\mu P = \mu$ with $\sum_i \mu_i = 1$. Let $A_1, A_2, \ldots$ be the sequence of random matrices defined by $A_1$ equals $B_i$ with probability $\mu_i$ and $\mathbb{P}[A_{t+1} = B_j | A_t = B_i] = P_{ij}$. This random sequence of matrices is stationary and ergodic.
Example 2.5. (auto-regressive models) A standard model of a fluctuating environmental variable (e.g. precipitation or temperature) is of the form $x_{t+1} = a_{t+1}x_t + b_{t+1}$ where $x_t$ is the environmental variable at time $t$, $a_t$ are i.i.d. with $\mathbb{E}[\log a_t] < 0$, and $b_t$ are i.i.d. with $\mathbb{E}[b_t] < \infty$. Asymptotically (as $t \to \infty$) this sequence is stationary and ergodic. If the entries of $A_t$ have some continuous dependence on $x_t$ (e.g. survivorship or fecundity as a function of precipitation or temperature), then the $A_t$ are asymptotically stationary and ergodic. We will discuss these models, and their multivariate counterparts, more extensively in section XX.XX.

Quite remarkably, there is a stochastic analog of the Perron-Frobenius Theorem. The version we present is a slight extension of a result due to Arnold et al. [1994]. A proof of this extension will be given in section XX.XX. To state the theorem, we need to introduce two definitions. We say $A_t$ is primitive if there is a $s$ such that the random product $A_{t+1}A_{t+2}\ldots A_{t+s}$ has only positive entries. Define

$$m = \min_{i,j} B_{ij} \text{ and } M = \max_{i,j} B_{ij}$$

where $B = A_1A_2\ldots A_s$. Furthermore, define

$$\log^+ x = \max\{\log x, 0\}$$

to be the non-negative part of $\log x$.

Theorem 2.2. (Random Perron-Frobenius Theorem) Let $A_1, A_2, \ldots$ be a primitive, stationary, ergodic sequence of random variables such that $\mathbb{E}[\log^+ \frac{1}{m}]$ and $\mathbb{E}[\log^+ M]$ are finite. Then there exists two stationary sequences of positive, random vectors $v_t$ and $w_t$ and stationary sequence of scalars $\lambda_t$ such that

1. $\sum_i v_{i,t} = \sum_i v_{i,t}w_{i,t} = 1$ for all $i$ and $t$,
2. $v_tA_{t+1} = \lambda_{t+1}v_{t+1}$ for all $t$,
3. $A_{t+1}w_{t+1}^T = \lambda_{t+1}w_{t}^T$ for all $t$,
4. for any $n_0 > 0$ (i.e. all non-negative entries with one entry at least positive) and vector $c > 0$,

$$\lim_{t \to \infty} \frac{1}{t} \log nc^T = \mathbb{E}[\log \lambda_1] =: r,$$

and

5. there exists a $\tau > 0$ such that for any $n_0 > 0$,

$$\lim_{t \to \infty} \log \left| \frac{n_t}{\lambda_1\lambda_2\ldots\lambda_t} - v_t(n_0w_0^T) \right| \leq -\tau$$

with probability one.

The interpretation of this theorem is analogous to the Perron-Frobenius theorem. The stationary sequence of vectors $v_t$ correspond to the stable state distribution of the structured population. However, due to the temporal fluctuations, the stable-state distribution varies in time in a stationary manner. The final assertion of the theorem implies that the population distribution for any initial condition approaches the $v_t$ at an exponential rate. The $\lambda_t$ describe how the population changes in size as it follows the stable state distribution. In particular, in the long-term, the log population size grows linearly like $\mathbb{E}[\log \lambda_1]t$. The $w_t$ describe the time varying reproductive values of individuals. The final assertion of the theorem puts all the pieces together and implies that

$$n_t \approx (v_0w_0^T)\lambda_1 \ldots \lambda_tv_t \approx (v_0w_0^T)e^{rt}v_t$$

for $t$ sufficiently large where $r$ is the stochastic per-capita growth rate.

The next example illustrates one of the earliest ecological applications of this random Perron-Frobenius theorem.
Figure 2. Graphic depiction of the Silva et al. [1991] model for the neotropical, perennial grass *Andropogon semiberbis*. There are 4 states (the nodes): 1 tiller (state 1), 2−10 tillers (state 2), 11−20 tillers (state 3), > 20 tillers (state 4). Arrows indicate contributions of one stage to another. Burn populations in (a) and unburnt populations in (b).

**Example 2.6. (Burning grasses)** Silva et al. [1991] examined the effect of temporal fluctuations in fires on the persistence of tropical savanna grass *Andropogon semiberbis*. For this species, fires help remove “necromass” (dead biomass) and, thereby, improve survival, growth, and reproduction of individuals. Silva et al. [1991] were interested on how the frequency of fires influences the persistence of this species. They divided this size-structured population into 4 stages: individuals with 1 tiller (stage 1), individuals with 2−10 tillers (stage 2), individuals with 11−20 tillers (stage 3), and individuals with > 20 tillers (stage 4). Based on two year experiments in burnt and unburnt plots, they estimate the vital rates under these two sets of conditions (Figure 2). The transpose of the following matrix was used for the burnt years:

$$B=\begin{bmatrix} 0.08 & 1.63 & 2.42 & 4.4 \\ 0.21 & 0.64 & 0.35 & 0.16 \\ 0.019 & 0.43 & 0.24 \\ 0.03 & 0.23 & 0.48 \end{bmatrix}, 4 \text{ byrow=TRUE}$$

and the transpose of the following matrix was used for the unburnt (or protected) years:

$$U=\begin{bmatrix} 0.706 & 0.391 & 3.59 \\ 0.018 & 0.158 & 0.136 & 0.093 \\ 0.008 & 0.07 & 0.21 \\ 0.001 & 0.07 \end{bmatrix}, 4 \text{ byrow=TRUE}$$

To model the stochastic switching between burnt and unburnt years, Silva et al. [1991] assumed that the probability of going from a burnt year to an unburnt year is $p$, from a burnt year to a burnt year is $1 - p$, from an unburnt year to a burnt year is $q$, and from an unburnt year to an unburnt year is $1 - q$. Under these assumptions, the frequency of burnt years in the long term is

$$\frac{q}{q + p}$$
Figure 3. Convergence to the stable state distribution $v_t$ for 100 random initial conditions $n_0$ for Silva's model of Andropogon semiberbis. Parameters: $p = q = 0.5$

We can model these stochastic transitions between burnt years (state 1) and unburnt years (state 2) using the transition matrix

$$P = \begin{pmatrix} 1 - p & p \\ q & 1 - q \end{pmatrix}$$

where $P_{ij}$ is the probability of the environment going from state $i$ to state $j$.

Figure 3 illustrates the convergence of $n_t / \sum_i n_{i,t}$ to the stable stage distribution for 50 random initial conditions.

Figure 4 plots $r$ for different $p$ and $q$ values against the fraction $\frac{p}{q + p}$ of burnt years. This figure demonstrates that (i) the critical burning frequency is approximately 85% and (ii) temporal correlations have little effect on $r$.

Now suppose that $q = 0.9 = 1 - p$ (i.e. every year there is a 90% changes of a burning) in which case $r \approx 0.1$. Figure 5 plots the distribution of $v_t$ (left) and $w_t$ (right) values for each stage (stage 1 on top, stage 4 on bottom). This figure illustrates that in the long term, most individuals are in the first stage, the fewest individual are in the 4th stage, and the greatest coefficient of variation occurs in the 2nd stage. Intuitively, stage 4 has the greatest reproductive value due to these individuals having higher survivorship and fecundity than other stages, while stage 1 has the least reproductive value.

Example 2.7. (Well-mixed populations) Many species live in habitats that appear as discrete patches across the landscape e.g. butterflies in meadows, water fleas in ephemeral ponds, big horn sheep on mountain tops, annual plants on soil patches. Assume there are $k$ habitat patches and the density of an annual population in the $i$-th habitat patch is $n_{i,t}$ in year $t$. Let $R_{i,t+1}$ be the number of offspring produced by an individual living in patch $i$ that make it to year $t + 1$. Assume offspring disperse before reproducing and the fraction dispersing from patch $i$ to patch $j$ is $D_{ij}$. Under these
Figure 4. Stochastic growth rates $r$ for the burning grass model for a 20 by 20 grid of $p$ and $q$ values.

Figure 5. Stable stage distributions (left) and reproductive values (right) when $p = 0.9 = 1 - q$. 
assumptions, we get the following matrix model
\[ n_{t+1} = n_t R_{t+1} D \]
where \( R_{t+1} = \text{diag}(R_{1,t+1}, \ldots, R_{k,t+1}) \).

In general finding \( r \) for this model is quite challenging. However, in the special case that the populations are well-mixed i.e. the probability of going to patch \( j \) is independent of where an offspring is coming from, then finding \( r \) is straightforward. In this case, there are \( d_j \) such that \( D_{ij} = d_j \) for all \( i \) and \( j \). This implies that the beginning of each year, the fraction of the population in patch \( j \) is \( d_j \). Hence, the stable stage distribution is given by \( v_t = (d_1, d_2, \ldots, d_k) \) for all time. Thus
\[
\lambda_{t+1} = \sum_i d_i R_{i,t+1},
\]
and
\[
r = \mathbb{E}[\log \lambda_t] = \mathbb{E}[\log \sum_i d_i R_{i,t}].
\]

Exponentiating \( r \), we get \( e^r \) is the geometric mean of the spatial mean of the fitnesses \( R_{i,t} \). This expression for “metapopulation” fitness was first discovered by Metz et al. [1983] and rediscovered independently in the two-patch case by Jansen and Yoshimura [1998]. Provided this “metapopulation fitness” is greater than one, the population grows at an exponential rate. If it is less than one, then the metapopulation tends toward extinction at an exponential rate.

As log is a concave function, applying Jensen’s inequality twice to this expression (with respect to the spatial average and the temporal average) yields
\[
\sum_i d_i \log \mathbb{E}[R_{i,t}] > r > \sum_i d_i \mathbb{E}[\log R_{i,t}]
\]

The first inequality implies that \( r > 0 \) requires that the arithmetic mean \( \mathbb{E}[R_{i,t}] \) is greater than one in at least one patch. As the geometric means are less than the arithmetic means, the second inequality implies that \( r \) can be \( > 0 \) even if each patch, in and of itself, can not sustain the population i.e. \( \mathbb{E}[\log R_{i,t}] < 0 \). For example, if the fitness in patch is log normal with log-mean \( \mu \), log variance \( \sigma^2 \) and \( d_i = \frac{1}{k} \), then \( \mathbb{E}[R_{i,t}] = \exp(\mu + \sigma^2/2) \) and \( \mathbb{E}[\log R_{i,t}] = \mu \). In the limit of many patches, the law of large numbers implies that
\[
\lim_{k \to \infty} \mathbb{E}[\log \frac{1}{k} \sum_i R_{i,t}] = \mathbb{E}[\log e^{\mu + \sigma^2/2}] = \mu + \sigma^2/2
\]

Hence, if \( \mu < 0 \), \( \mu + \sigma^2 > 0 \), and \( k \) is sufficiently large, then \( r > 0 \) despite all patches being unable, in and of themselves, to support a population. Figure 6 illustrates this phenomena where there is critical number of habitat patches above which the metapopulation growth rate \( r \) is positive.

### 2. Sensitivity of \( r \)

While \( r = \mathbb{E}[\log \lambda_t] \) can be easy to compute for unstructured populations due to the commutativity of scalar products, explicitly computing \( r \) for structured populations is often analytically intractable. Two notable exceptions are Roerdink [1987]’s formula for \( r \) for model of biennial populations, random matrices that share a reproductive value or share a stable structure [Tuljapurkar, 1986, 1990], see, e.g., Example 2.7. One useful approach to getting some insights into how \( r \) depends on the nature of the
stochastic forcing is study small perturbations of random matrix products for which \( r \) is understood. For example, small random perturbations of a constant matrix model, or adding dispersal limitation to the metapopulation model in Example 2.7. What do we mean by small random perturbations? Suppose that you know the value of \( r \) for the random matrix model \( n_{t+1} = n_tA_{t+1} \) and you want to know what happens when you replace \( A_{t+1} \) by \( C_{t+1}(\varepsilon) = A_{t+1} + \varepsilon B_{t+1} \) where \( C_1(\varepsilon), C_2(\varepsilon), \ldots \) is a stationary, ergodic sequence of matrices. If \( C_t(\varepsilon) \) satisfy the assumptions of the Random Perron-Frobenius Theorem, then the stochastic growth rate \( r(\varepsilon) \) exists as a function of \( \varepsilon \). Given that one understands \( r(0) \), wouldn’t it be great if one could write down a Taylor’s expansion of the form

\[
r(\varepsilon) = r(0) + r'(0)\varepsilon + \frac{1}{2} r''(0)\varepsilon^2 + \ldots
\]

Remarkably, such a Taylor expansion is possibly provided that \( r(\varepsilon) \) is an analytic function. So when is \( r \) analytic? (Feel free to skip the rest of this paragraph if you are interested in this technical detail.) A sufficient condition for analyticity is that there is common underlying stochastic process driving both sequences of matrices. Namely, there is a stationary, ergodic sequence \( \xi_1, \xi_2, \ldots \) in separable metric space \( X \) (e.g. a finite set of points \( \{1, 2, \ldots, n\} \) or \( \mathbb{R}^n \) for some \( n \)) and continuous functions \( A, B : X \rightarrow \mathbb{R}^{k^2} \) such that \( A_t = A(\xi_t) \) and \( B_t = B(\xi_t) \). One can view the \( \xi_t \) as describing the environmental state of the system (e.g. the climatic conditions) and the functions \( A, B \) map environmental conditions to demographic effects i.e. vital rates. For example, if there are a finite number of environmental states \( \{1, \ldots, n\} \), then \( \xi_1, \xi_2, \ldots \) might be given by an irreducible Markov chain on these states and \( A(i) \) and \( B(i) \) are matrices of vital rates when the system is in the \( i \)-th environmental state. Or \( \xi_t \) are multivariate numerals and \( a_{ij}(\xi) \) describe how the vital rates depend on the value of the multivariate normal. Under this additional assumption \( r(\varepsilon) \) is analytic and we can compute derivatives.

As the process of finding the first derivative \( r'(0) \) is fairly straight-forward and provides good practice with understanding the key components of the Random Perron-Frobenius theorem, we go through the details for computing it. To this end, let \( w_t, v_t, \) and \( \lambda_t \) be as given by the Perron-Frobenius Theorem.
for the unperturbed matrix model $n_{t+1} = n_t A_{t+1}$. Let $g_t(\varepsilon) = \mathbb{E} \left[ \log v_0 C_1(\varepsilon) \ldots C_t(\varepsilon) w_t^T \right]$. The Perron-Frobenius theorem applied to $C_t(\varepsilon)$ implies that

$$r(\varepsilon) = \lim_{t \to \infty} g_t(\varepsilon).$$

Notice that

$$v_0 C_1 \ldots C_t w_t^T = v_0 (A_1 \ldots A_t) w_t^T + \varepsilon \sum_{s=1}^t v_0 A_1 \ldots A_{s-1} B_s A_{s+1} \ldots A_t w_t^T + O(\varepsilon^2)$$

$$= \lambda_1 \ldots \lambda_t + \varepsilon \sum_{s=1}^t \lambda_1 \ldots \lambda_{s-1} \lambda_{s+1} \ldots \lambda_t v_{s-1} B_s w_s^T + O(\varepsilon^2)$$

$$= \lambda_1 \ldots \lambda_t \left( 1 + \varepsilon \sum_{s=1}^t v_{s-1} B_s w_s^T / \lambda_s \right) + O(\varepsilon^2)$$

As $\log(1 + x) = x + O(x^2)$, we get that

$$\log(v_0 C_1 \ldots C_t w_t^T) = \sum_{s=1}^t \log \lambda_s + \varepsilon \sum_{s=1}^t v_{s-1} B_s w_s^T / \lambda_s + O(\varepsilon^2)$$

Taking expectation, dividing by $t$, and taking the limit at $t \to \infty$ yields

$$\mathbb{E} \left[ \log \lambda_1 \right] + \varepsilon \mathbb{E} \left[ v_0 B_1 w_1^T / \lambda_1 \right] + O(\varepsilon^2).$$

Hence,

$$r(0) = \mathbb{E} \left[ v_0 B_1 w_1 / \lambda_1 \right]$$

and

$$r(\varepsilon) = r(0) + \varepsilon \mathbb{E} \left[ v_0 B_1 w_1 / \lambda_1 \right] + O(\varepsilon^2).$$

**Example 2.8.** (Perturbations of deterministic models) If $A_t = A$ is a constant matrix model, then $v_0 = v$ and $w_1 = w$ are the left and right eigenvectors normalized so that $\sum_i v_i = \sum_i w_i w_i = 1$, and $\lambda_t = \lambda$ is the associated eigenvector. In this case, $r'(0) = v \mathbb{E} [B_1 w_1^T / \lambda]$. If $B_1$ has all zeros except a value of one in the $i$–$j$–th entry, then $r'(0) = v_i w_j / \lambda$. In words, the increase in $r$ due to an increase in the contribution of individuals in stage $i$ to stage $j$ is proportional to the long-term frequency of stage $i$ individuals (i.e. $v_i$) and the reproductive value of stage $j$ individuals (i.e. $w_j$). This corresponds to the sensitivity of $r$ to increases in the $i$–$j$–th entry of the unperturbed matrix $A$.

For example, Figure 7 shows all of the sensitivities of $r \approx -0.05653757$ for the loggerhead model from Example ???. Notice that the extremely large sensitivities corresponds to the transitions from yearlings (stage 2) to adults (stages 5 through 7). Hence, if even a small percentage of yearlings were able to become mature adults within a year, there would be a large increase in the population rate $r$. However, this change isn’t biologically feasible, consistent with the fact that the corresponding entries in $A$ are zero. Hence, to look at “biologically feasible” sensitivities, one should only consider the ones that correspond to non-zero entries of the matrix. Alternatively, one can look at the elasticity of $r$ to changes in $A$ i.e. the relative change in $r$ due to a relative change in $A$. These are given by $w_j v_i A_{ij} / \lambda$ and plotted in the right hand side of Figure 7. These elasticities suggest that a relative increase in survival of sub-adults (stage 3) would produce the largest increase in $r$.

**Example 2.9.** (Should I stay or should I go?) Lets consider a two patch version of the metapopulation matrix model from Example 2.7 with $d_1 = d_2 = 1/2$:

$$A_t = \begin{pmatrix} R_{1,t} & 0 \\ 0 & R_{2,t} \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$
2. SENSITIVITY OF $r$

Figure 7. Sensitivities for $r$ for the loggerhead model on the left. Elasticities on the right. Each group of bars corresponds to the sensitivities/elasticities for the corresponding row of the matrix.

For this model, $v_t = (1/2, 1/2)$ for all $t$, $\lambda_t = (R_{1,t} + R_{2,t})/2$, and (check for yourself!) $w_t = (R_{1,t+1}, R_{2,t+1})/\lambda_{t+1}$.

To understand the effect of dispersal limitation on the stochastic growth rate, we consider the matrix model given by

$$C_t(\varepsilon) = \begin{pmatrix} R_{1,t} & 0 \\ 0 & R_{2,t} \end{pmatrix} \begin{pmatrix} \frac{1}{2} + \varepsilon & \frac{1}{2} - \varepsilon \\ \frac{1}{2} - \varepsilon & \frac{1}{2} + \varepsilon \end{pmatrix} = A_t + \varepsilon \begin{pmatrix} R_{1,t} & -R_{1,t} \\ -R_{2,t} & R_{2,t} \end{pmatrix} = B_t$$

i.e. an additional fraction $\varepsilon \geq 0$ of individuals stay in their natal patch. The effect of this additional patch fidelity on the stochastic growth rate is determined by

$$r'(0) = \mathbb{E} \left[ \frac{vB_tw_t^T}{\lambda_t} \right] = \mathbb{E} \left[ \frac{R_{1,t} - R_{2,t}}{\lambda_t} (1/2, -1/2) w_t^T \right] = \frac{1}{2} \mathbb{E} \left[ \frac{R_{1,t} - R_{2,t}}{\lambda_t} \frac{R_{1,t+1} - R_{2,t}}{\lambda_{t+1}} \right] = 2\mathbb{E} \left[ \frac{R_{1,t} - R_{2,t}}{R_{1,t} + R_{2,t}} \frac{R_{1,t+1} - R_{2,t+1}}{R_{1,t+1} + R_{2,t+1}} \right]$$

If we define $W_{i,t} = \frac{R_{i,t}}{R_{1,t} + R_{2,t}}$ as the relative fitness of patch $i$ in year $t$, then our approximation is given by

$$r(\varepsilon) \approx r(0) + 2\varepsilon \mathbb{E}[(W_{1,t} - W_{2,t})(W_{1,t+1} - W_{2,t+1})]$$

If the $W_{i,t}$ are identically distributed and independent in space, then this approximation has two implications. If the difference in relative fitness is positively auto-correlated in time, then $r'(0) > 0$ and greater site fidelity increases the stochastic growth rate. If the relative fitness is negatively auto-correlated in time, then $r'(0) < 0$ and greater site fidelity decreases the stochastic growth rate. Hence, this analysis provides an answer to a query posed by the Clash in the 1980s:
Should I stay or should I go now?
Should I stay or should I go now?
If I go there will be trouble
An’ if I stay it will be double
So come on and let me know

Namely, go if relative fitness differences are negatively auto-correlated, and stay if they are positively auto-correlated. Figure 8 illustrates how well the approximation (lines) works. A more general formula for any number of patches is derived in [Schreiber, 2010].

The approximation (2.3) is useful for examining what happens if the random perturbation $B_t$ has non-zero mean. However, often stochastic demographers are interested in what effect does just increases variation in some of the vital rates (i.e. entries of $A_t$) have on the long-term growth rate $r$. In this case, one would want $E[B_t] = 0$. If the $B_t$ are independent of the $v_t, w_t$ for $A_t$, then $r'(0) = 0$ and one needs to consider higher order correction terms. In the exercises, you are asked to verify the following classic result from stochastic demography [Tuljapurkar, 1990].

Theorem 2.3. Variance sensitivity Assume $A_1, A_2, \ldots$ are i.i.d., $B_1, B_2, \ldots$ are i.i.d., and the $A_t$’s and $B_t$’s are independent of one another. If $E[B_t] = 0$ and $r(\varepsilon)$ is the stochastic growth rate of $A_t + \varepsilon B_t$, 

---

**Figure 8.** Stochastic growth rates $r$ and the first order approximation for the two patch model. Temporal correlations for three curves are 0.5, 0, −0.5.
then
\[ r'(0) = 0 \text{ and } r''(0) = -\text{Var}\left[ v_0 B_1 w_1^T / \lambda_1 \right]. \]

Hence,
\[ (2.4) \quad r(\varepsilon) = r(0) - \frac{1}{2} \text{Var}\left[ v_0 B_1 w_1^T / \lambda_1 \right] \varepsilon^2 + O(\varepsilon^3) \]

A key implication of this result is that temporally uncorrelated fluctuations in demographic rates decrease the per-capita growth rate \( r \) of populations. However, as suggested by Example 2.9, temporal auto-correlated fluctuations can increase population growth rates.

**Example 2.10.** *(Tuljapurkar’s approximation)* An important special case of (2.4) is when \( A_t \) is a constant matrix \( A \). Then \( v_t = v \), \( w_t = w \), and \( \lambda_t = \lambda \) are constant. Define \( S_{ij} = v_i w_j / \lambda \) to be the sensitivity of \( r = \log \lambda \) to changes in the \( i-j \)-th entry of \( A \). If \( C_t(\varepsilon) = A + \varepsilon B_t \) where \( B_t \) are i.i.d. with mean 0, then (try this for yourself!)
\[ r(\varepsilon) \approx r - \frac{1}{2\lambda^2} \sum_{i,j,k,t} S_{ij} S_{kt} \text{Cov}[b_{ij,t}, b_{kt,t}]. \]

Hence \( r \) is most sensitive to vital rates that exhibit the greatest variation i.e. the largest \( S_{ij}^2 \text{Var}[b_{ij,t}] \) terms. Negative correlations between vital rates can reduce, but never eliminate, the negative effect of environmental fluctuations on the population growth rate.

**Example 2.11.** *(Burning grasses revisited)* Let \( E(ij) \) be the \( 4 \times 4 \) matrix with a 1 in the \( i-j \) entry and zeros elsewhere. To see the effect of increasing the mean value of any of the entries of \( A_t \), we consider \( B_t = E(ij) \) for all \( i \) and \( j \) values. The left hand side of Figure 9 plots \( r'(0) \) for all of these perturbations. The first block bars corresponds to contributions from stage 1, the second group of bars corresponds to contributions from stage 2, etc. These sensitivities indicate that increasing the fecundity of stage 4 individuals provides the biggest boost to \( r \). To see the effect of increasing the variation in any of the entries, we use the perturbation \( B_t = \pm E(i,j) \) where \( \pm \) occur with equal probabilities. These sensitivities are plotted on the right hand side of Figure 9. These bar plots show that variation in the fecundity of the stage 4 individuals has the most negative impact on the long-term growth rate \( r \).
3. General random products

It will be useful to consider random matrix models $n_{t+1} = n_t A_{t+1}$ where $A_t$ are allowed to have negative as well as positive entries. For example, these type of matrices arise when linearizing a deterministic model around an equilibrium. For this general class of models, one can say something about the “stability” of the zero solution. Assume $A_t$ are stationary and ergodic and define $\|A\| = \max_{i,j} |A_{ij}|$.

**Theorem 2.4. (Stability of random matrix products)** If $\mathbb{E}[\log^+ \|A_1\|] < \infty$, then there exists a real number $r$ (possibly $-\infty$) such that

$$\lim_{t \to \infty} \frac{1}{t} \log \|A_1 \ldots A_t\| = r$$

with probability one. In particular, if $r < 0$, then

$$\lim_{t \to \infty} n_t = 0 \text{ with probability one}$$

for all $n_0$. Alternatively, if $r > 0$ and $n_0$ is a random variable uniformly distributed on the unit sphere, then

$$\lim_{t \to \infty} \sum_i |n_{i,t}| = \infty \text{ with probability one.}$$

$r$ is often called the dominant Lyapunov exponent. As with the projection matrices, $r < 0$ implies stability of the origin as with probability one all solutions converge exponentially fast to the origin. Alternatively, $r > 0$ implies that “most” initial conditions result in $n_t$ going to infinity at an exponential rate.

4. Random affine models

The random matrix models from the previous sections are good described closed populations who remain at sufficiently low densities that density-dependent feedbacks are negligible. Many populations are open and experience immigration. To account for this immigration, we need to allow for random inputs which often uncorrelated to local densities. Hence, we are interested in models of the form

$$n_{t+1} = n_t A_{t+1} + b_{t+1}$$

where are $A_t$ and $b_t$ are given by ergodic, stationary sequences of random matrices and vectors, respectively. When the $A_t$ and $b_t$ are non-negative matrices and vectors, $b_t$ correspond to random inputs of individuals in all the stages to the focal population i.e. the population being modeled. More generally, $A_t$ and $b_t$ may have positive and negative entries e.g. they come from the linearization of a density-dependent model (see Chapter XX.XX). Model (2.5) is called a generalized auto-regressive process. For the rest of the section, I assume that the $A_t$ and $b_t$ are i.i.d. I still need to check which of the results extend to the stationary case.

Iterating (2.5) for $t$ time steps and starting with $n_0$ yields

$$n_t = n_0 A_1 \ldots A_t + b_t + b_{t-1} A_t + \ldots + b_1 A_2 \ldots A_t$$

Let $r$ be dominant Lyapunov exponent associated with the product $A_1 \ldots A_t$. If $r < 0$, then the term $n_0 A_1 \ldots A_t$ goes to zero with probability one. Hence, asymptotically, we are only left with the term

$$b_t + b_{t-1} A_t + \ldots + b_1 A_2 \ldots A_t$$

By our assumption of the sequences being i.i.d., the distribution of this term is the same as the distribution of

$$b_t + b_2 A_1 + b_3 A_1 A_2 \ldots + b_t A_1 \ldots A_{t-1}$$
Hence, provided the limit is well-defined with probability one, we would expect that $n_t$ converges in distribution to

$$n = \lim_{t \to \infty} b_1 + b_2 A_1 + b_3 A_1 A_2 + \cdots + b_t A_1 \cdots A_{t-1}$$

As the distribution of this limiting expression unchanged by multiplying by a new $A$ and adding a new $b$, it is stationary.

To state a theorem that justifies this heuristic, define a subspace $V$ of $\mathbb{R}^k$ to be invariant for (2.5) if $P[X_1 \in V|X_1 = x] = 1$ whenever $x \in V$. Namely, if you start in the subspace you stay in the subspace for all time with probability one. For most of the models we consider, the only invariant subspace is $\mathbb{R}^k$ itself.

**Theorem 2.5.** Assume that $E[\log^+ ||A_1||]$ and $E[||b_1||]$ are finite, $\mathbb{R}^k$ is the only invariant subspace for (2.5), and the dominant Lyapunov exponent $r$ for $A_1, A_2, \ldots$ is negative. Then (2.6) exists almost-surely. Moreover, for any $n_0, n_t$ converges in distribution to $\hat{n}$ i.e. for any Borel set $\mathcal{A} \subset \mathbb{R}^k$,

$$\lim_{t \to \infty} P[n_t \in \mathcal{A}] = P[\hat{n} \in \mathcal{A}]$$

and

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} h(n_s) = E[h(\hat{n})] \text{ with probability one}$$

for any continuous function $h : \mathbb{R}^k \to \mathbb{R}$ such that $E[|h(\hat{n})|] < \infty$.

**Example 2.12.** (A simple immigration model) Let $n_t$ be the density of an unstructured population (i.e. $k = 1$) for which the fitness of an individual is $R_{t+1}$ and $I_{t+1}$ immigrates in by time $t + 1$. Then $n_{t+1} = n_t R_{t+1} + I_{t+1}$. Provided that $r = E[\log R_t] < 0$ and $E[I_t]$ is finite, then $n_t$ converges with probability one to a stationary distribution. Consider the special case where $R_t$ is log-normally distributed with log mean $r$ and log variance $\sigma^2$ and $I_t = I$ for all time (i.e. constant immigration). Then $r < 0$ ensures the existence of a stationary distribution. Figure 10 illustrates these distributions for $r = -0.05, \sigma^2 = 0.01, 0.02, 0.1$, and $I = 1$. For $\sigma^2 = 0.1$, the distribution is much broader than the smaller $\sigma^2$ values; populations reaching densities of a million quite frequently. We will see why this value of $\sigma^2$ is special shortly.
Example 2.13. Higher order auto-regressive processes AR(\(k\)) Auto-regressive models of order \(k\) are used extensively in modelling weather, stock markets, etc. If \(x_t\) is the scalar variable of interest (e.g. temperature), then AR(\(k\)) models are of the form

\[ x_{t+1} = a + \sum_{s=1}^{k} a_s x_{t+1-s} + \eta_{t+1} \]

where \(a, b_i\) are constants and \(\eta_1, \eta_2, \ldots\) are i.i.d. with zero mean i.e. a white noise processes. Hence, for these processes, the state of the system in the next time step depends on the state of the system for the previous \(k\) time steps.

To turn this model into the type we have been talking about, define

\[ n_{t+1} = n_t \begin{pmatrix} b_1 & 1 & 0 & \ldots & 0 & 0 \\ b_2 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{k-1} & 0 & 0 & \ldots & 0 & 1 \\ b_k & 0 & 0 & \ldots & 0 & 0 \end{pmatrix} + \begin{pmatrix} a + \eta_{t+1} \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \]

If the \(|b_i| < 1\), then this auto-regressive process has a stationary distribution.

Figure 11 (upper panels) shows the daily average temperature at Bodega Marine Laboratory during 2014. Plotting tomorrow’s temperature against today’s temperature demonstrates a fairly linear relation which suggest a first order auto-regressive process could be used to model this data. In fact, AIC model selection yields a third auto-regressive model with coefficients \(b_1 = 0.7898, b_2 = -0.0595, b_3 = 0.1711, a = 25.02924\) and white noise with variance \(\sigma^2 = 1.284\). Simulating this AR(3) process yields the time series shown in the lower panels of Figure 11.

Example 2.14. (Stochastic ferns) Random affine models can be used to generate all types of surprisingly fun figures. For example, suppose that \(A_t\) and \(b_t\) are given by

\[ \begin{pmatrix} 0.4 \\ -0.3733 \end{pmatrix} \begin{pmatrix} 0.06 \\ 0.6 \end{pmatrix} \]

with probability 0.2933 and

\[ \begin{pmatrix} -0.8 \\ -0.1876 \end{pmatrix} \begin{pmatrix} 0.1371 \\ 0.8 \end{pmatrix} \]

with the complementary probability. As there are only a finite number of non-zero matrices and vectors being chosen, \(\mathbb{E}[\log^+ \|A_t\|], \mathbb{E}[\|b\|]\) are finite. Furthermore, as \(\|A_t\| \leq 0.8\), we get that

\[ r = \lim_{t\to\infty} \frac{1}{t} \log \|A_1 \ldots A_t\| \leq \log 0.8 < 0 \]

with probability one. Hence, iterating this random system should converge to a stationary distribution. Indeed, it converges to a surprisingly cool fern as shown in Figure 12.

Given a random affine model with a unique stationary distribution \(\hat{n}\), what can say about the distribution of \(\hat{n}\)? Using independence and stationarity, we can fairly easily compute the mean and co-variance of this distribution, assuming that they exist. By stationarity and independence, if \(\mathbb{E}[\hat{n}]\) exists, then it satisfies

\[ \mathbb{E}[\hat{n}] = \mathbb{E}[\hat{n}] \mathbb{E}[A_1] + \mathbb{E}[b_1] \]

\[ (I - \mathbb{E}[A_1]) \mathbb{E}[\hat{n}] = \mathbb{E}[b_1]. \]

If the spectral radius of \(\mathbb{E}[A_1]\) is less than one (this is the necessary condition for the expectation existing), then \(I - \mathbb{E}[A_1]\) is invertible and

(2.7) \[ \mathbb{E}[\hat{n}] = (I - \mathbb{E}[A_1])^{-1} \mathbb{E}[b_1] \]
To compute the co-variance matrix provided, we begin by assuming the expectation exists and define $\tilde{y} = \tilde{n} - \mathbb{E}[\tilde{n}]$ and $\beta_1 = \mathbb{E}[\tilde{n}]A_1 + b_1 - \mathbb{E}[\tilde{n}]$. Then (check the details for yourself!) $\mathbb{E}[\beta_1] = 0$ and $\mathbb{E}[\tilde{y}] = 0$. Stationarity and independence imply

$$
\begin{align*}
\mathbb{E}[\tilde{y}^T \tilde{y}] &= \mathbb{E}[(\tilde{y}A_1 + \beta_1)^T(\tilde{y}A_1 + \beta_1)] \\
&= \mathbb{E}[A_1^T] \mathbb{E}[\tilde{y}^T \tilde{y}] \mathbb{E}[A_1] + \mathbb{E}[\beta_1^T \beta_1] \\
\text{Cov}[\tilde{n}] &= \mathbb{E}[A_1^T] \text{Cov}[\tilde{n}] \mathbb{E}[A_1] + \text{Cov}[\beta_1]
\end{align*}
$$

To solve this system of linear equations, we can take advantage of two matrix operations: the vec operation and the Kronecker product $\otimes$. The vec operation $\text{vec}(A)$ takes a matrix $A$ and creates a long vector by concatenating the column vectors. Two key properties of the vec operation are $\text{vec}(A + B) = \text{vec}(A) + \text{vec}(B)$ (obvious), and $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$ (less obvious). Using these two properties, we get

$$
\begin{align*}
\text{vec}(\text{Cov}[\tilde{n}]) &= \text{vec}(\mathbb{E}[A_1^T] \text{Cov}[\tilde{n}] \mathbb{E}[A_1] + \text{Cov}[\beta_1]) \\
&= \text{vec}(\mathbb{E}[A_1^T] \text{Cov}[\tilde{n}] \mathbb{E}[A_1]) + \text{vec}(\text{Cov}[\beta_1]) \\
&= (\mathbb{E}[A_1]^T \otimes \mathbb{E}[A_1]^T) \text{vec}(\text{Cov}[\tilde{n}]) + \text{vec}(\text{Cov}[\beta_1]) \\
(I - \mathbb{E}[A_1]^T \otimes \mathbb{E}[A_1]^T) \text{vec}(\text{Cov}[\tilde{n}]) &= \text{vec}(\text{Cov}[\beta_1])
\end{align*}
$$

Figure 11. Actual (top) and simulated (bottom) daily average temperatures at Bodega Marine Lab in 2015. Simulated data is from an AR(3) model whose parameters are described in the main text.
Provided that the spectral radius of $\mathbb{E}[A_1^T] \otimes \mathbb{E}[A_1^T]$ is less than one, then we can take the inverse and get the following formula for the co-variance matrix:

$$(2.8) \quad \text{vec}(\text{Cov}[\hat{n}]) = (I - \mathbb{E}[A_1^T] \otimes \mathbb{E}[A_1^T])^{-1} \text{vec}(\text{Cov}[\beta_1])$$

**Example 2.15.** (The simple immigration model revisited) Let $n_{t+1} = n_t R_{t+1} + I$ where $R_1, R_2, \ldots$ is a log-normal sequence of i.i.d. random variables with log mean $r$ and log variance $\sigma$, and $I \geq 0$ corresponds to constant rain of immigration. This process has a stationary distribution provided that $r = \mathbb{E}[\log R_t] < 0$. Provided that $\mathbb{E}[R_t] = \exp(r + \sigma^2/2) < 1$, the mean of this stationary distribution is given by (verify this for yourself!)

$$\mathbb{E}[\hat{n}] = I/(1 - \exp(r + \sigma^2/2)).$$

As $\beta_1 = \mathbb{E}[\hat{n}] R_1 + I - \mathbb{E}[\hat{n}]$, $\text{Var}[\beta_1] = \mathbb{E}[\hat{n}]^2 \text{Var}[R_1]$ and $\text{Var}[R_1] = (e^{\sigma^2} - 1) e^{2r + \sigma^2}$, the variance of $\hat{n}$ is given by

$$\text{Var}[\hat{n}] = \mathbb{E}[\hat{n}](e^{\sigma^2} - 1) \exp(2r + \sigma^2)/(1 - \exp(2r + \sigma^2))$$

As $\mathbb{E}[\log R_t] = r < 0$ doesn’t imply that $\mathbb{E}[R_t] = \exp(r + \sigma^2/2) < 1$, we get this immigration model has a stationary distribution with infinite moments whenever $r < 0 \leq r - \sigma^2/2$ e.g. $-r = \sigma^2 = 0.05$ as illustrated in Figure 10.

**Example 2.16.** (AR models revisited) Consider the AR($k$) model described in Example 2.13. UPCOMING

**Example 2.17.** (Open two stage population) Consider a population with two stages, juveniles and adults. Juveniles survive with probability 0.1 in which case they become adults. Adults survive with probability 0.5 to the next year. New juveniles only arrive through immigration. Let $I_t$ be the number
of juveniles that arrive in year $t$. Then get

$$n_{t+1} = n_t \begin{pmatrix} 0 & 0.5 \\ 0 & 0.1 \end{pmatrix} + I_{t+1} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Assume that $\log I_t$ is normally distributed with $\mu$ and variance $\sigma^2$. Then

$$\mathbb{E}[A_1] = \begin{pmatrix} 0 & 0.5 \\ 0 & 0.1 \end{pmatrix} \quad \text{and} \quad \mathbb{E}[b_1] = \begin{pmatrix} \exp(\mu + \sigma^2) \\ 0 \end{pmatrix}$$

Therefore

```r
A=rbind(c(0,0.1),c(0,0.5))
mu=0.1
sigma=0.1
b=c(exp(mu+sigma^2/2),0)
Covb=rbind(c((exp(sigma^2)-1)*exp(2*mu+sigma^2),0),c(0,0))
require(fBasics)
## Loading required package: fBasics
## Loading required package: timeDate
## Loading required package: timeSeries
## Warning: package 'timeSeries' was built under R version 3.2.3
##
#### Rmetrics Package fBasics
## Analysing Markets and calculating Basic Statistics
## Copyright (C) 2005-2014 Rmetrics Association Zurich
## Educational Software for Financial Engineering and Computational Science
## Rmetrics is free software and comes with ABSOLUTELY NO WARRANTY.
## https://www.rmetrics.org --- Mail to: info@rmetrics.org
```
Exercises

(1) Verify the expression for the reproductive values in Example 2.9.

(2) Let $d_1, d_2 \geq 0$, $d_1 + d_2 = 1$,

$$A_t = \begin{pmatrix} R_{1,t} & 0 \\ 0 & R_{2,t} \end{pmatrix} \begin{pmatrix} d_1 & d_2 \\ d_1 & d_2 \end{pmatrix}$$

and

$$C_t = \begin{pmatrix} R_{1,t} & 0 \\ 0 & R_{2,t} \end{pmatrix} \begin{pmatrix} d_1 - \varepsilon & d_2 + \varepsilon \\ d_1 & d_2 \end{pmatrix}$$

Find $r'(0)$. Interpret the result in terms of the stochastic growth rate of a two patch model.

(3) Prove that (2.4) holds.

(4) Verify the variance sensitivity formula in Example 2.10.

(5) Verify the expressions for $\mathbb{E}[\hat{n}]$ and $\text{Var}[\hat{n}]$ in Example 2.15.

(6) Find $\mathbb{E}[\hat{n}]$ and $\text{Cov}[\hat{n}]$ for Example 2.14.

(7) For the auto-regressive model of rainfall in Example 2.13, let $\hat{n} = (\hat{n}_1, \hat{n}_2, \hat{n}_3)$ be the stationary distribution. Find $\mathbb{E}[\hat{n}]$ and $\text{Cov}[\hat{n}]$.

(8) Desert tortoise, *Gopherus agassizii*, was listed as endangered in 1989, and a draft recovery plan was issued in 1993. Declines are occurring throughout the range of the species, but they appear to be particularly severe in the Western Mojave. Direct human impacts on the tortoise include habitat degradation and habitat loss, hunting (up to 14% of mortality in some areas), and getting run over by cars or off-road vehicles. Indirect impacts on the populations include habitat degradation by sheep or cattle grazing, predation by ravens (which attack yearlings and juveniles, and are associated with human activity), and an upper respiratory tract infection that may have been introduced by release of pet tortoises into the wild. This goal of this homework is to compare two or more management scenarios (e.g. reducing human disturbance, removing ravens) and determine which will have a greater effect in conserving the desert tortoise populations. Human disturbance mainly affects larger individuals, while raven predation is limited to smaller ones. To achieve this goal, you need to develop a stage-structured model using the following data and analyze this model using the techniques developed in class.

In 1978, K. H. Berry classified Desert Tortoise sizes as follows:

<table>
<thead>
<tr>
<th>class</th>
<th>name</th>
<th>maximum carapace length (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Yearling</td>
<td>&lt;60</td>
</tr>
<tr>
<td>1</td>
<td>Juvenile 1</td>
<td>60–99</td>
</tr>
<tr>
<td>2</td>
<td>Juvenile 2</td>
<td>100–139</td>
</tr>
<tr>
<td>3</td>
<td>Immature 1</td>
<td>140–179</td>
</tr>
<tr>
<td>4</td>
<td>Immature 2</td>
<td>180–207</td>
</tr>
<tr>
<td>5</td>
<td>Subadult</td>
<td>208–239</td>
</tr>
<tr>
<td>6</td>
<td>Adult 1</td>
<td>&gt;240</td>
</tr>
</tbody>
</table>

Government reports and previously unanalyzed mark-recapture data at 8 Bureau of Land Management permanent study plots in the Western Mojave provide the following information about growth, survival, and reproduction for the turtle populations. Growth is defined as the probability, conditional on survival, of moving from one size class to the next largest one; of the $>1400$ individuals transitions in the data, no tortoises either shrank or were estimated to grow more than one size class in a year.
Since no data are available on the demographic rates for yearlings and size class one tortoises, you should make the optimistic assumption that they are the same as those for size class 2 individuals.

For the Mojave, there were no direct observations of individual fecundity (e.g. egg or hatchling production). Instead, you can use four estimates for fecundity that were computed in various ways:

<table>
<thead>
<tr>
<th>Yearling production</th>
<th>Low</th>
<th>Med-low</th>
<th>Med-high</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yearling production of</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stage 5</td>
<td>0.042</td>
<td>0.42</td>
<td>1.3</td>
<td>2.22</td>
</tr>
<tr>
<td>Yearling production of</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stage 6</td>
<td>0.069</td>
<td>0.69</td>
<td>1.98</td>
<td>3.38</td>
</tr>
<tr>
<td>Yearling production of</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stage 7</td>
<td>0.069</td>
<td>0.69</td>
<td>2.57</td>
<td>4.38</td>
</tr>
</tbody>
</table>

That amount of variation is certainly not ideal but it’s not necessarily catastrophic, because the model predictions that you care about may not be affected even by that much parameter uncertainty.

(a) Write down a matrix model for each of the low and high fecundity estimates. Compute the dominant eigenvalue, stable age distribution, and the reproductive values for each of these models. Discuss the similarities and differences of these quantities for the low and high fecundity scenarios.

(b) To see the effect of stochasticity in this system, assume that the conservation biologists have implemented a strategy that increases survivorship of all stages by 5.5%. Furthermore, assume the probability of going from a low fecundity year to a high fecundity year is \( p \) and the probability of going from a high fecundity year to a low fecundity year is \( q \).

(i) High and low years are uncorrelated when \( p = 1 - q \). Assuming that \( p = 1 - q \), compute and plot the stochastic growth rate as \( p \) goes from 0 to 1. Discuss for what frequencies of low fecundities the population can persist.

(ii) A frequency of 50% low fecundity years occurs whenever \( \frac{p}{p+q} = 1/2 \). Equivalently \( p = q \). Assuming \( p = q \), compute and plot the stochastic growth rate as function \( p \). Discuss how varying \( p \) corresponds to varying the temporal correlation of the environmental states and discuss how this temporal correlation influences the stochastic growth rate and thereby persistence.
CHAPTER 3

Nonlinear single species models

Some negative feedback between \( r \) and population density (that is, density dependence) is a necessary (but not sufficient) condition for population regulation. – Peter Turchin [1999]

1. Unstructured models

Our study of non-linear, single species models begins with unstructured models where \( n_t \) is the density of the population at time \( t \). To simultaneously account for density-dependent feedbacks and environmental fluctuations, let \( \mathcal{E} \) be a separable metric space (e.g. \( \mathbb{R}^n \) for some \( n \)) that represents all possible environmental states. For example, if the important environmental variables are temperature and precipitation, then \( \mathcal{E} \) might be \( \mathbb{R} \times [0, \infty) \). Alternatively, if we model the fluctuations in \( n \) demographic parameters directly without explicitly describing their dependence on environmental factors, then \( \mathcal{E} \) is some subset of \( \mathbb{R}^n \). We will write \( \xi \in \mathcal{E} \) for a particular environmental state.

Let \( f(n, \xi) \) be the fitness of an individual when the population density is \( n \) and the environmental state is \( \xi \). Let \( \xi_1, \xi_2, \ldots \in \mathcal{E} \) be a sequence of i.i.d. random variables determining the environment. The population dynamics are given by

\[
(3.1) \quad n_{t+1} = n_t f(n_t, \xi_{t+1}).
\]

1.1. Negative density-dependence. Many population experience negative feedbacks of density on survival and reproduction. For example, as densities get higher, there might be fewer resources per individual which lowers survivorship and reproductive output of individuals. To account for these negative feedbacks, we assume that \( f(n, \xi) \) is a decreasing function of \( n \). We will relax this assumption substantially when examining structured population models later in this chapter.

For models of this type, we may be interested in when populations are regulated: their densities tend remain bounded away from zero and infinity. To identify when regulation occurs, we can use linearizations of the population dynamics when the population is at either low or high population densities. In the presence of negative density dependence, the per-capita growth rate of the population is greatest at low densities. Hence, for the population to tend away from extinction, it should increase when at low densities. So, consider that \( n_0 \approx 0 \) but not zero. Then as long \( n_t \) remains at low densities, we can approximate the dynamics by

\[
n_{t+1} \approx n_t f(0, \xi_{t+1})
\]

We know from the previous chapter that for this stochastic linear difference equation \( r = \mathbb{E}[\log f(0, \xi_1)] \) determines whether the population has a tendency to increase (i.e. \( r > 0 \)) or decrease (i.e. \( r < 0 \)).

To know whether the population tends to decrease at high densities, we can define

\[
f(\infty, \xi) = \lim_{n \to \infty} f(n, \xi)
\]

for all \( \xi \in \mathcal{E} \). This limit is well defined (possibly 0) as \( f \) is a decreasing function of \( n \) for all \( \xi \). To avoid an absurd heuristic, lets assume that this limit is strictly positive for all \( \xi \). If \( n_t \) is sufficiently large,

\[\text{Need to see if we can extend the theorems in this section to the stationary case}\]
then we can approximate the dynamics by the linear stochastic difference equation

\[ n_{t+1} \approx n_t f(\infty, \xi_{t+1}). \]

In this case, we expect that

\[ s := \mathbb{E}[\log f(\infty, \xi_1)] \]

(possibly \(-\infty\)) determines whether the population continues to increase (i.e. \(s > 0\)) or starts to decrease (i.e. \(s < 0\)).

Based on these linear approximations, we have three scenarios. First, if \(r < 0\), then under the best conditions (i.e. low densities) the population is decreasing and, consequently, we expect the population to tend toward extinction. Second, if \(s > 0\), then under the worst conditions (i.e. high densities) the population is increasing and, consequently, we expect the population to grow without bound. Finally, if \(r > 0\) and \(s < 0\), then the population tends to increase when rare and decrease when very common. Hence, in this case, we expect the population to persist in manner that doesn’t cause madness. Indeed, the following theorem justifies these conclusions.

**Theorem 3.1.** Schreiber [2012] Assume \(f(n, \xi)\) is a positive decreasing function in \(n\) for all \(\xi \in \mathcal{E}\) and \(\mathbb{E}[\log^+ f(0, \xi_1)] < \infty\). Then

- **Extinction:** if \(r = \mathbb{E}[\log f(0, \xi_1)] < 0\), then \(\lim_{t \to \infty} n_t = 0\) with probability whenever \(n_0 \geq 0\),
- **Unbounded growth:** if \(s = \lim_{n \to \infty} \mathbb{E}[\log f(n, \xi_1)] > 0\), then \(\lim_{t \to \infty} n_t = \infty\) with probability whenever \(n_0 > 0\), and
- **Stochastic persistence:** if \(r > 0\) and \(s < 0\), then for all \(\epsilon > 0\) there exists \(M > 0\) such that

\[ \liminf_{t \to \infty} \frac{\#\{1 \leq s \leq t : 1/M \leq n_s \leq M\}}{t} \geq 1 - \epsilon \text{ with probability one} \]

whenever \(n_0 > 0\).

Stochastic persistence, the final case, implies that the typical trajectory spends most of its time in a sufficiently large compact interval excluding the extinction state 0. In the words of Peter Chesson [1982]

This criterion requires that the probability of observing a population below any given density, should converge to zero with density, uniformly in time. Consequently it places restrictions on the expected frequency of fluctuations to low population levels. Given that fluctuations in the environment will continually perturb population densities, it is to be expected that any nominated population density, no matter how small, will eventually be seen. Indeed this is the usual case in stochastic population models and is not an unreasonable postulate about the real world. Thus a reasonable persistence criterion cannot hope to do better than place restrictions on the frequencies with which such events occur.

The proofs for the first two cases of Theorem 3.1 are fairly straightforward. Consequently, we discuss them briefly. Suppose that \(r < 0\). As \(f(n, \xi)\) is a decreasing function of \(n\), we have \(f(0, \xi) \geq f(n, \xi)\) for any \(n \geq 0\). Hence,

\[ n_{t+1} \leq n_t f(0, \xi_{t+1}) \]

for all \(t\). Iterating this inequality yields

\[ n_t \leq n_0 f(0, \xi_1) \ldots f(0, \xi_t) \]

Hence, the law of large numbers implies

\[ \limsup_{t \to \infty} \frac{1}{t} \log n_t \leq \lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} \log f(0, \xi_s) = r < 0 \text{ with probability one.} \]
Figure 1. Grain beetle dynamics. Left panel shows fitting of the survivorship curve \( a/(1 + bn^c) \) to survivorship data of the flour beetle species *Tribolium castaneum*: \( a = 0.8, b = 0.0148920, \) and \( c = 4.2470522. \) The right hand panels show simulations with \( r = \mu + \log 0.8 = 0.1 \) and \( \sigma = 0 \) (middle), \( \sigma = 0.1 \) (right).

Thus \( n_t \) converges to zero exponentially fast with probability one. In the exercises, you are asked to prove the complementary result in the case that \( s > 0. \)

To illustrate Theorem 3.1, we apply it to stochastic versions of the Ricker model, a seed-bank model for annual plants, and a model of checkerspot butterflies.

**Example 3.1.** (Generalized Beverton-Holt model) Bellows [1981] used a model of the form \( f(n, \xi) = \lambda a/(1 + bn^c) \) where \( \xi = (\lambda, a, b, c) \) to described the demography of various species of grain beetles. In this model, \( a > 0 \) is the probability of an individual surviving to adulthood and reproducing, \( \lambda > 0 \) corresponds to the low-density reproductive output of a reproducing individual, \( b > 0 \) measures the strength of competition, and \( c > 0 \) determines the strength of density-dependence (e.g. \( b \leq 1 \) is under-compensatory and \( b > 1 \) is over-compensatory). Suppose \( \lambda_t \) are log-normally distributed with log mean \( \mu \) and log standard deviation \( \sigma \), but \( a, b \) and \( c \) are constant. The low-density per-capita growth rate is \( r = \mu + \log a. \) The high-density per-capita growth rate is (check for yourself!) \( s = -\infty. \)

Figure 1 shows a fitting of the survivorship curve \( a/(1 + bn^c) \) to survivorship data of the flour beetle species *Tribolium castaneum*: \( a = 0.8, b = 0.0148920, \) and \( c = 4.2470522. \) Hence, provided that \( \mu > -\log 0.8, \) the corresponding population dynamics model would exhibit persistence. The right hand panels of Figure 1 show simulations with \( r = \mu + \log 0.8 = 0.1 \) and \( \sigma = 0 \) (middle), \( \sigma = 0.1 \) (right).

**Example 3.2.** (An annual plant model) Consider an annual plant where \( n \) is the density of seeds in the ground. Each year a fraction \( g \) of seeds germinate. Of the seeds that don’t germinate, a fraction \( s \) survive to the next year. At low densities, germinating seeds produce a yield of \( Y \) seeds. Competition among st germinating individuals reduces yield by a factor of \( 1/(1 + agn) \) where \( a > 0 \) is a competition coefficient. Then our model is given by

\[
 f(n, \xi) = \frac{gY}{1 + agn} + (1 - g)s \text{ where } \xi = (a, g, s, Y) 
\]

The low-density and high density per-capita growth rates are given by

\[
 r = \mathbb{E}[\log f(0, \xi_1)] = \mathbb{E}[\log(g_1Y_1 + (1 - g_1)s_1)] \text{ and } s = \lim_{n \to \infty} \mathbb{E}[\log f(n, \xi_1)] = \mathbb{E}[\log(1 - g_1)s_1] < 0 
\]

As the latter quantity \( s \) is always negative, this model always exhibits bounded dynamics. Persistence requires that \( r \) is positive.
Consider the special case in which only $Y$ fluctuates between $Y_{\text{max}}$ and 0 with probabilities $p$ and $1 - p$ respectively and the remaining parameters are constant. One could view this as each year being either a drought year or high rainfall year. Under these conditions, we get

$$r = p \log(gY_{\text{max}} + (1 - g)s) + (1 - p) \log((1 - g)s).$$

Figure 2 illustrates the dynamics from this model when $r > 0$.

Now let’s look at an application of these models to a real ecological puzzle.

**Example 3.3. (Extinction of checkerspot butterflies)** In the 1990s, two populations of checkerspot butterflies went extinct in Northern California. The population densities for one of these populations is shown in the left hand side of Figure 3. These extinctions were observed to coincide with a change in precipitation variability in the 1970s. As shown in the right hand side of Figure 3, the standard deviation in precipitation was approximately 50% higher after 1971 than before 1971.

To evaluate whether this shift in precipitation variability may have caused the extinction of the checkerspots, McLaughlin et al. [2002] developed a stochastic difference equation of the following type

$$n_{t+1} = n_t \exp(a - bn_t + c\xi_{t+1}^-\xi_t^-)$$

where $\xi_t$ is precipitation in year $t$. Using linear regression on a log-scale yields a model whose fit for one-year predictions are shown as red diamonds in Figure 3.

The model with random draws from the pre-1971 precipitation data yields $r = E[a + c\xi_1^-\xi_t^-] = 0.04$. Hence, the population is expected to persist with this form of climatic variability as illustrated in the simulations on the left hand side of Figure 5. The model with random draws from the post-1971 precipitation data yields $r = -0.049$. Hence, the population is expected to tend toward extinction with this form of climatic variability as illustrated in the simulations on the right hand side of Figure 5.

Running the two scenarios and comparing extinction risk.

**1.2. Positive density-dependence.** Let’s now assume that $f(n, \xi)$ is an increasing function of density i.e. $f$ exhibits positive density dependence. Common causes of this positive density-dependent feedback include predator saturation, cooperative predation, increased availability of mates, and conspecific enhancement of reproduction [Courchamp et al., 1999, Stephens et al., 1999, Gascoigne and Lipcius, 2004, Courchamp et al., 2008, Gascoigne et al., 2009, Kramer et al., 2009].
To understand the dynamics for models with positive density-dependence, we can, once again, define the low-density and high-density per-capita growth rates, $r$ and $s$, as in the case of negative density-dependence. However, unlike the negative density-dependent case, we have $s > r$. The signs of $r$ and $s$ determine three scenarios. If $r > 0$, then the population tends to increase under the worst conditions and therefore the population should increase without bound. If $s < 0$, then the population tends to decrease under the best conditions and, consequently, it should asymptotically go extinct. Finally, if $r < 0$ and $s > 0$, then depending on initial conditions we may expect that asymptotic extinction or blow might happen.
Figure 5. Simulated checkerspot population dynamics with pre-1971 precipitation data (left) and post-1971 precipitation data (right)

To state a theorem about this trichotomy, we say that \( \{0, \infty\} \) is accessible from a set \( B \subset (0, \infty) \) if for any \( M > 0 \), there exists \( \gamma > 0 \) such that

\[
P\left( \{ n_t \in [0, 1/M] \cup [M, \infty) \text{ for some } t \geq 0 \} \mid n_0 = x \right) > \gamma
\]

for all \( x \in B \). Intuitively, accessibility of 0 and \( \infty \) from \( B \) means that if \( n_0 \) lies in \( B \) then there is a positive probability that it gets arbitrarily small or arbitrary large after enough time.

**Theorem 3.2.** Assume \( f(n, \xi) \) is an increasing function of \( n \) for all \( \xi \in E \). Define  

\[
s = \lim_{n \to \infty} \mathbb{E}[\log f(n, \xi_1)]
\]

and \( r = \mathbb{E}[\log f(0, \xi_1)] \). Then

**Extinction:** if \( s < 0 \), then \( \lim_{t \to \infty} n_t = 0 \) with probability one whenever \( n_0 \geq 0 \).

**Unbounded growth:** if \( r > 0 \), then \( \lim_{t \to \infty} n_t = \infty \) with probability one whenever \( n_0 > 0 \).

**Conditional persistence:** if \( r < 0 \) and \( s > 0 \), then for any \( 0 < \delta < 1 \), there exist \( m, M > 0 \) such that

\[
P\left( \lim_{t \to \infty} n_t = \infty \mid n_0 = x \right) \geq 1 - \delta \text{ and } P\left( \lim_{t \to \infty} n_t = 0 \mid n_0 = y \right) \geq 1 - \delta,
\]

for all \( x \in [M, \infty) \) and all \( y \in (0, m] \).

Moreover, if \( \{0, \infty\} \) is accessible, then

\[
P\left( \left\{ \lim_{t \to \infty} n_t = 0 \text{ or } \infty \right\} \mid n_0 = x \right) = 1
\]

for all \( x \in (0, \infty) \).

The final statement of the theorem about conditional persistence has three implications. First, if the initial population density is sufficiently small, then extinction is highly likely. Second, if the initial population density is sufficiently large, then unbounded growth is highly likely. Third, if the environmental stochasticity pushes population densities eventually to either high or low densities (i.e. accessibility of \( \{0, \infty\} \)), then extinction or growth without bound occurs with probability one.

While models with positive density-dependence aren’t realistic in the long term (negative density-dependence must ultimately kick in), they are still useful as means to understand how positive density-dependence influences the establishment of populations. The next example illustrates their use from an empirically inspired study.
Figure 6. Probability of \( n_{100} > 1000 \) for \( n_{t+1} = n_t \exp(0.1(n_t - C) + \xi_{t+1}) \) where \( \xi_t \) are normal with mean 0 and standard deviation \( \sigma \) and \( n_0 = 5 \).

Example 3.4. (Allee effects in Gypsy moths) Liebhold and Bascompte [2003] used models with only positive density-dependence to examine numerically the joint effects of Allee effects, environmental stochasticity, and externally imposed mortality on the probability of successfully exterminating an invasive species such as the gypsy moth. Their fitness function was

\[
f(x, \xi) = \exp(\gamma(x - C) + \xi)
\]

were \( C \) is the deterministic Allee threshold, \( \gamma \) is the “intrinsic rate of natural increase,” and \( \xi \) are chosen to be normal random variables with mean 0.

By assuming that \( \xi_t \) are normally distributed, \( \{1, \infty\} \) is accessible from any initial positive density. Furthermore, as \( r = \mathbb{E}[-\gamma C + \xi_t] = -\gamma C < 0 \) and \( s = +\infty \) for this model, our results imply both extinction and unbounded growth occur with positive probability and, thereby, provide a rigorous mathematical foundation for Liebhold and Bascompte [2003]’s numerical analysis. Figure 6 recreates one of the figures from their paper. The paper estimates the probability of establishment (i.e. \( n_t \to \infty \)) for a range of threshold values and noise levels. Quite interestingly, this figure illustrates that higher levels of environmental stochasticity can increase the probability of establishment when the initial density lies below the threshold value (i.e. values of \( C \geq 5 \) in Figure). Conversely, environmental stochasticity reduces the probability of establishment for initial densities greater than the threshold density (i.e. values of \( C < 5 \) in the Figure).


**Figure 7.** Simulations of the sterile insect release model with $\mu = 0.1$ and $S = 100$. Dashed line indicates the density at which the geometric mean of the fitness is one.

**Example 3.5.** (Sterile insect releases) Let $n_t$ be the density of females of an insect pest, $S$ the number of sterile males, and $\xi_t$ the number of daughter produced by a mated male. Let $\xi_t$ are log-normally distributed with log mean $\mu$ and log standard deviation $\sigma$. If we assume there is a 50-50 sex ratio and mating is panmictic, then

$$n_{t+1} = n_t \xi_{t+1} \frac{n_t}{n_t + S}$$

where $\frac{n_t}{n_t + S}$ is the probability of a female mating a non-sterile male. Then $f(n, \xi) = \xi n/(n + S)$, $r = E[\log 0] = -\infty$ and $s = E[\log \xi_t] = \mu$. As $\xi_t$ are log normally distributed, $\{0, \infty\}$ are accessible. Thus, if $\mu > 0$, there is a positive probability of converging to 0 or $\infty$ for any positive initial density $n_0$.

For this model, geometric mean of fitness at density $n$ is $e^{\mu n/(n + S)}$. When this mean equals 1, the population has no inherent tendency to increase or decrease. Indeed, in the absence of noise, this critical density, $n^* = S/(e^\mu - 1)$, is an equilibrium of the deterministic model. In the absence of noise, $n_t$ converges to zero whenever $n_0 < n^*$. In the presence of noise, however, $n_t$ may grow without bounded even if $n_0 < n^*$. Hence, environmental stochasticity can rescue populations expected to go extinct (Figure 7).

2. Structured models

To account for population structure, let $n_t = (n_{1,t}, n_{2,t}, \ldots, n_{k,t})$ be vector of densities of individuals in $k$ different stages or states at time $t$. Let $A(n, \xi)$ be $k \times k$ non-negative matrices that vary continuously with $n$ and $\xi$. Then the population dynamics are given by

$$n_{t+1} = n_t A(n_t, \xi_{t+1})$$


2. STRUCTURED MODELS
where \( \xi_1, \xi_2, \xi_3, \ldots \) are an ergodic stationary sequence of random variables on a separable metric space \( \mathcal{E} \). As in the structured case, we will be interested in the conditions that ensure a population is regulated i.e. tends to stay bounded away from extinction and infinitely large population densities.

### 2.1. Bounded population dynamics

Unlike the unstructured case, there is no unique way to define the population growth rate at infinity i.e. there are many ways that to approach an infinite population size. Hence, we need another approach to ensure that populations tend to remain bounded. To this end, we introduce functions, \( V : \mathbb{R}_+^k \to \mathbb{R} \), that become arbitrarily large for arbitrarily large population sizes, and such that \( \log V \) on average decreases when the population gets large. For the first requirement, we require that \( V \) is proper i.e \( \lim_{i \to \infty} \sum_i n_i \to \infty \). For example, any weighed combination of the population densities, \( V(n) = \sum_i w_i n_i \) is proper with \( w_i > 0 \), is proper. The following theorem clarifies what we mean by \( \log V \) decreasing on average for large population sizes.

**Theorem 3.3.** Benaïm and Schreiber [2009] If there exists a proper function \( V : \mathbb{R}_+^k \to \mathbb{R} \) and non-negative, random variables \( \alpha, \beta : \mathcal{E} \to \mathbb{R} \) such that

1. \( V(nA(n, \xi)) \leq \alpha(\xi)V(n) + \beta(\xi) \);
2. \( \mathbb{E}[\log \alpha] < 0 \);
3. \( \mathbb{E}[\log^+ \beta] < \infty \)

then there exists \( C > 0 \) such that

\[
\liminf_{t \to \infty} \mathbb{P} \left[ \sum_i n_{i,t} \leq M \right] \geq 1 - \frac{C}{M}
\]

and

\[
\liminf_{t \to \infty} \frac{\# \{ 1 \leq s \leq t : \sum_i n_{i,t} \leq M \}}{t} \geq 1 - \frac{C}{M} \text{ with probability one.}
\]


**Example 3.6.** (Metapopulations revisited) Consider a \( k \)-patch model to one where offspring disperse and adults are sedentary (i.e. don’t disperse) and iteroparous (i.e. potentially reproduce multiple times in their life time). Let \( n_{i,t} \) be the population density in patch \( i \), with \( i = 1, 2, \ldots, k \). Let \( \xi_{i,t+1} = \text{exp}(\sum_k (a_i n_{k,t}) D_{ij} + \text{diag}(s_1, \ldots, s_k)) \)

To check stochastic boundedness, let \( V(n) = \sum_i n_i \) be the total population size. For non-negative \( x \), the maximum of \( x \exp(-ax) \) is \( 1/(ae) \) at \( x = a \). Therefore,

\[
V(nA(n, \xi)) = \sum_{i=1}^k n_i \xi_i \exp(-a_i n_i) + s_i n_i
\]

Define \( \alpha(\xi) = \max_i \{ s_i \} \) and \( \beta_i(\xi) = \sum_i \xi_i/(a_i e) \). As \( \mathbb{E}[\log \alpha(\xi)] = \log \max_i s_i < 0 \), Theorem 3.3 applies provided that \( \mathbb{E}[\log^+ \sum_i \xi_i/(a_i e)] \leq \infty \). This latter requirement is true for many distributions used in ecological models e.g. log-normally or gamma distributed \( \xi_{i,t} \) or any finite distribution of \( \xi_i \) values.
EXAMPLE 3.7. (A stochastic Leslie matrix model) Let \( n_i \) be the density of individuals in age class \( i \). Let \( s_i < 1 \), for \( i = 2, \ldots, k \), be the probability of individuals of age \( i - 1 \) surviving to age \( i \). Now, suppose that “size” of an individual of age \( i \) is \( w_i > 0 \), and the amount of resources acquired by an individual of age \( i \) is proportional to \( w_i \). Assuming that resource availability decreases with the total “size” \( \sum \omega w_i n_i \) of the population, we can model the per-capita fecundity of individuals of age \( i \) by \( w_i(\xi/\alpha)v \sum \omega w_i n_i / \sum w_i n_i \). Under these assumptions, we have

\[
A(n, \xi) = \begin{pmatrix}
\frac{\omega _1 \xi}{\sum \omega_i n_i} & s_2 & 0 & 0 & \cdots & 0 \\
\frac{\omega _2 \xi}{\sum \omega_i n_i} & 0 & s_3 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{\omega _k \xi}{\sum \omega_i n_i} & 0 & 0 & \cdots & 0 & s_k \\
\end{pmatrix}
\]

Lets consider the case that the \( \xi_i \) are log-normally distributed with log means \( \mu_i \) and log standard deviations \( \sigma_i \). Define \( V(n) = \sum \omega w_i n_i, s = \max_{2 \leq i \leq k} \sigma_i \). Then

\[
V(nA(n, \xi)) = w_1 \xi \sum_{i=1}^{k} \frac{w_i n_i}{1 + \sum \omega_i n_i} + \sum_{i=2}^{k} w_i s_i n_{i-1}
\]

\[
\leq w_1 \xi \frac{V(n)}{1 + V(n)} + s V(n)
\]

Suppose there is \( \nu^* > 0 \) such that \( \mathbb{E}[\log(nA_\xi n/\nu + 1)] < 1 \) e.g. \( \xi \) is log-normally distributed or is drawn from a finite, discrete distribution. Theorem 3.3 applies for \( \alpha(\xi) = w_1 \xi / (1 + \nu^*) + 1 \) and \( \beta(\xi) = w_1 \xi \). (verify yourself!)

2.2. Stochastic persistence and extinction. To understand when populations are likely to persist, we can examine the population growth rate when rare. Namely, if \( n_0 \approx 0 \) but not zero, then \( n_{t+1} \approx n_t A(0, \xi_{t+1}) \) which implies

\[
n_t \approx n_0 A(0, \xi_1) A(0, \xi_2) \ldots A(0, \xi_t)
\]

provided that \( n_0, n_1, \ldots, n_{t-1} \) remain close to zero. For this linear approximation, the Random Perron Frobenius Theorem in Chapter 2 implies that the population tends grow exponentially at a rate \( r \) which satisfies

\[
r = \lim_{t \to \infty} \frac{1}{t} \log \|A(0, \xi_1) \ldots A(0, \xi_t)\| \text{ with probability one}
\]

where \( \|A\| = \max_{ij} |A_{ij}| \). Hence, if \( r > 0 \), then the population increases when rare and we expect it to persist. Indeed, one can prove the following theorem which generalizes results from [Benaim and Schreiber, 2009, Roth and Schreiber, 2014].

THEOREM 3.4. \(^2\) Assume

- \( A(n, \xi) \) depend continuously on \( n \) and \( \xi \) and are primitive (i.e. there exists \( s > 0 \) such that \( A(n_1, \xi_1) \ldots A(n_s, \xi_s) \) has strictly positive entries with probability one for any \( n_1 \)),

- the conditions of Theorem 3.3 hold for some choice of \( V, \alpha, \) and \( \beta \), and

- \( \mathbb{E}[\log \|A(0, \xi_1)\|] < \infty \) so that \( r \) is well-defined.

If \( r > 0 \), then \( n_{t+1} = n_t A(n_t, \xi_{t+1}) \) is stochastically persistent and bounded: for all \( \epsilon > 0 \) there exists \( M > 0 \) such that

\[
\liminf_{t \to \infty} \frac{\# \{ 1 \leq s \leq t : 1/M \leq \min_i n_{i,s} \text{ and } \max_i n_{i,s} \leq M \} \geq 1 - \epsilon \text{ with probability one}}{t}
\]

\(^2\)This “theorem” is currently a conjecture. Special cases of this theorem are proved in [Benaim and Schreiber, 2009, Roth and Schreiber, 2014].
whenever \( n_0 > 0 \). If \( r < 0 \), then the extinction state is a stochastic attractor: for all \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that

\[
P \left[ \lim_{t \to \infty} n_t = 0 \bigg| n_0 \right] \geq 1 - \varepsilon
\]

whenever \( \sum_i n_{i,0} \leq \delta \). Furthermore, if \( \{0\} \) is accessible (i.e. for all \( \delta > 0 \) and \( n_0 > 0 \), \( P[\sum_i n_{i,t} \leq \delta \text{ for some } t|n_0] = 1 \)), then \( \lim_{t \to \infty} n_t = 0 \) with probability one.

In the case of \( r < 0 \) without additional assumptions (e.g. accessibility of 0), asymptotic extinction is highly likely only if the initial, total population density is sufficiently small. Accessibility ensures that population densities get sufficiently low, sufficiently frequently, that the process tends exponentially toward the origin. In the exercises, you are asked to show that \( r < 0 \) for models with only negative density-dependence (with or without the accessibility assumption) imply that \( \lim_{t \to \infty} n_t = 0 \) with probability one.

**Example 3.8.** *(Return of the metapopulation model)* Consider the \( k \) patch model presented in Example 3.6. This model satisfies the assumptions of Theorem 3.4, for example, if the \( \xi_{i,t} \) are log-normally distributed. Moreover, as the entries of \( A(n, \xi) \) are non-increasing functions of the entries of \( n \), \( r < 0 \) implies that the meapopulation converges exponentially to extinction for all initial conditions. In Example 2.9, we found a first order approximation of \( r \) for the case of \( k = 2 \) patches and non-overlapping generations i.e. \( s_i = 0 \) for all \( i \). This approximation showed that \( r \) could be positive even if the expected fitness in every patch was less than one. Specifically, this occurred when patches were sufficiently uncorrelated in space, positively auto-correlated in time, and partially connected by dispersal. Schreiber [2010] shows that similar conclusions hold for any number of patches. Figure 8 illustrates this phenomena for a system with 10 patches: convergence to extinction when there is no dispersal (left), and stochastic persistence when one quarter of the individuals disperse (right).

### 3. Ergodicity and stationary distributions

When \( r > 0 \), Theorem 3.4 implies that fraction of time spent at arbitrarily low densities is arbitrarily small. This theorem, however, does not ensure that there is a unique positive stationary distribution. For this stronger conclusion, we need either additional structure on the nonlinearities or on the the noise of the system.
3.1. Irreducibility in i.i.d. environments. Irreducibility occurs when for all initial conditions, there is a tendency to visit the same set of states. More precisely, given $\epsilon > 0$, we say that $\{n_t\}$ is irreducible over $(\epsilon, \infty)^k$ if there exists a probability measure $\Phi$ on $(\epsilon, \infty)^k$ such that for all $n_0 \in (\epsilon, \infty)^k$ and every Borel set $A \subset (\epsilon, \infty)^k$ there exists $T \geq 1$ (depending on $n_0$ and $A$) such that
\[
P[n_T \in A|n_0] > 0
\]
whenever $\Phi[A] > 0$.

Under the assumption of irreducibility and an i.i.d. environment, $r > 0$ implies the existence of a unique stationary distribution.

**Theorem 3.5.** (Uniqueness) Assume that $\{n_t\}$ is irreducible over $(\epsilon, \infty)^k$ for all sufficiently small $\epsilon > 0$. If $r > 0$, then there exists a unique invariant probability measure $\pi$ such that $\pi(\{0\}) = 0$ and
\[
\lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} h(n_s) = \int h(n)\mu(dn) \quad \text{with probability one}
\]
for $n_0 > 0$ and all continuous $h : \mathbb{R}^k_+ \to \mathbb{R}$ with $\int |h(n)|\mu(dn) < \infty$.

Theorem 3.5 ensures the asymptotic distribution of one realization of the population dynamics is given by the positive stationary distribution $\pi$. Hence, $\pi$ provides information about the long-term frequencies that a population trajectory spends in any part of the population state space.

To gain information about the distribution of $n_t$ across many realizations of the population dynamics, we need a stronger irreducibility condition. This stronger condition requires that after a fixed amount of time independent of initial condition, any positive population state can move close to any other positive population state. More precisely, we say that $\{n_t\}$ is strongly irreducible over $(\epsilon, \infty)^k$ if there exists a probability measure $\Phi$ on $(\epsilon, \infty)^k$, an integer $T \geq 1$ and some number $0 < h \leq 1$ such that for all $n_0 \in (\epsilon, \infty)^k$ and every Borel set $A \subset (\epsilon, \infty)^k$
\[
P[n_T \in A|n_0] \geq h\Phi(A).
\]
To state the next result given probability measures $\mu, \nu$ on $(0, \infty)^k$ define
\[
\|\mu - \nu\| = \sup_{B} |\mu(B) - \nu(B)|
\]
where the supremum is taken over all Borel sets $B \subset (0, \infty)^k$.

**Theorem 3.6.** (Convergence in distribution) In addition to the assumptions of Theorem 3.5 assume that $\{n_t\}$ is strongly irreducible over $(\epsilon, \infty)^k$ for all sufficiently small $\epsilon > 0$. Then the distribution of $n_t$ converges to $\pi$ as $t \to \infty$ whenever $n_0 \in (0, \infty)^k$; that is
\[
\lim_{t \to \infty} \|P[n_t \in \cdot|n_0] - \pi\| = 0 \quad \text{for all } n_0 \in (0, \infty)^k.
\]

**Example 3.9.** (The LPA model) To study the dynamics of flour beetles (genus *Tribolium*), built a stochastic, nonlinear, stage-structured model. For these pests of grains, individuals are in one of three stages: larvae ($L_t$), pupae ($P_t$), and adults ($A_t$). In each week, each adult, on average, produce $b$ eggs of which a fraction $\exp(-c_1 A_t - c_2 L_t)$ escape cannibalism and become larvae. Each week, $\mu_1$ of larvae survive and become pupae, and a fraction $\exp(-c_3 A_t)$ of pupae survive to become adults. Adults survive to the next week with probability $\mu_2$. This yields the deterministic model
\[
L_{t+1} = bA_t \exp(-c_1 A_t - c_2 L_t)
\]
\[
P_{t+1} = L_t (1 - \mu_1)
\]
\[
A_{t+1} = P_t \exp(-c_3 A_t) + (1 - \mu_2) A_t
\]
To account for stochastic fluctuations, \( \xi \) multiplied each of the equations by \( \exp(\xi_{i,t+1}) \) where \( \xi_t = (\xi_{1,t}, \xi_{2,t}, \xi_{3,t}) \) is a multivariate normal random variable with mean 0 and covariance matrix \( \Sigma \). This leads to a stochastic difference equation of the form

\[
\begin{pmatrix}
L_{t+1} & P_{t+1} & A_{t+1}
\end{pmatrix} = 
\begin{pmatrix}
0 & (1 - \mu_1) \exp(\xi_{2,t+1}) & 0 \\
0 & 0 & \exp(-c_3 A_t + \xi_{3,t+1}) \\
b \exp(-c_1 A_t - c_2 L_t + \xi_{1,t+1}) & 0 & (1 - \mu_2) \exp(\xi_{3,t+1})
\end{pmatrix}
\]

In the exercises, you are asked to verify that (1) these equations are stochastically bounded, (2) they are strongly irreducible, and (3) there exists a critical \( b \) below which the population goes extinct and above which it persists.

### 3.2. Compensatory negative density-dependence.

Independent of the noise structure, there exist assumptions on the nonlinearities ensuring the existence of a unique stationary distributions. Biologically, these assumptions corresponds to the populations experiencing compensating, negative density-dependence.

The first of the additional assumptions is purely technical, but is meet for most models.

**A1: Smoothness:** The map \( (n, \xi) \rightarrow A(n, \xi) \) is Borel, \( F_\xi(n) = nA(n, \xi) \) is twice continuous differentiable for all \( \xi, n \in \mathbb{R}_+^k \), and

\[
\mathbb{E} \left( \sup_{\|n\| \leq 1} \log^+ \left( \|F_\xi(n)\| + \|DF_\xi(n)\| + \|D^2F_\xi(n)\| \right) \right) < +\infty
\]

To account for competition, we assume all of the entries should be non-increasing functions of the population densities. Moreover, for each population state, we assume one of its entries is a decreasing function of its density or the density of one the other population states.

**A2: Negative density-dependence:** The matrix entries \( A_{ij}(n, \xi) \) satisfy

\[
\frac{\partial A_{ij}}{\partial n_l}(n, \xi) \leq 0
\]

for all \( n \) and \( \xi \). Moreover, for each \( i \) there exists some \( j \) and \( l \) such that this inequality is strict for all \( n \) and \( \xi \).

Finally, we assume that the density-dependence is compensatory.
A3: Compensatory negative density dependence: All entries of the derivative $DF_\xi(n)$ of $F_\xi(n)$ are non-negative for all $n$ and $\xi$.
This assumption ensures that the population dynamics are monotone. In other words, if $n_0 \geq n$, then $F_\xi(n_0) \geq F_\xi(n)$.

**Theorem 3.7.** (Benaïm and Schreiber [2009]) If assumptions A1-A3 hold and $r > 0$, then there exists a unique stationary distribution $\pi$, such that $\pi[\{0\}] = 0$,

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} h(n_s) = \int h(n) \pi(dn) \text{ with probability one}$$

and

$$\lim_{t \to \infty} \|\mathbb{P}[n_t \in \cdot | n_0] - \pi\| = 0$$

for $n_0 > 0$ and all continuous $h : \mathbb{R}_+^k \to \mathbb{R}$ with $\int |h(n)| \pi(dn) < \infty$.

**Example 3.10.** (Revenge of the Leslie matrix models) Verify for yourself that Example 3.7 satisfies the assumptions of the Theorem.

4. Exercises

(1) Prove the second statement of Theorem 3.1.
(2) Construct an example of a model where $f(n, \xi)$ is strictly decreasing with $n$ but $\lim_{t \to \infty} n_t = \infty$ with probability one whenever $n_0 > 0$.
(3) Consider $n_{t+1} = n_t A(n_t, \xi_{t+1})$ which satisfies the first and third assumption of Theorem 3.4.
   (a) Assume $A_{ij}(n, \xi)$ are non-increasing functions of the coordinates of $n$. Prove if $r < 0$, then $\lim_{t \to \infty} n_t = 0$ with probability one.
   (b) Assume $A_{ij}(n, \xi)$ are non-decreasing functions of the coordinates of $n$. Prove if $r > 0$, then $\lim_{t \to \infty} n_t = \infty$ with probability one.
CHAPTER 4

Multispecies models

The diversity of the plankton [is] explicable primarily by a permanent failure to achieve equilibrium as the relevant external factors changes. – G.E. Hutchinson [1961]

Nearly all species interact with other species. These interactions can be classified by the net effect of each species on the average fitness of the other species. When the net effects are mutually harmful, the species are exhibiting competitive interactions. Mutually beneficial interactions occur for mutualists. Finally, if one species benefits while the other is harmed, then the species are likely predators and their prey, or pathogens/parasites and their hosts. Intuitively, we expect competitive interactions to inhibit the persistence of one or both of the competitors, mutualistic interactions to facilitate persistence of each species, and predator/prey or host/parasite interactions to facilitate persistence of one and inhibit persistence of the other. A basic issue of practical and theoretical interest is unraveling how environmental fluctuations interact with these nonlinear feedbacks to determine the composition of ecological communities and their dynamics. We examine some of these issues in this chapter.

1. The lottery model

To get a feeling for some basic principles underlying the analysis of multi-species models, we begin with the lottery model. Introduced by Chesson and Warner [1981], this model examines the dynamics of two species competing for space. Space that is so limiting that it is usually fully occupied by the competing species. Let \( n_{i,t} \) denote the fraction of space occupied by species \( i \). At the beginning of year \( t + 1 \), each individual of species \( i \) produces \( \xi_{i,t+1} \) offspring and a fraction \( \delta \) (independent of species for simplicity) of the reproducing adults die, thereby freeing up space. Each offspring is as likely as the other to “capture” one of the freed up locations. The fraction of freed up space captured by species 1 is equal to the fraction of the total offspring \( \xi_{1,t+1}n_{1,t+1} + \xi_{2,t+1}n_{2,t+1} \) produced by the community that belong to species 1:

\[
\frac{\xi_{1,t+1}n_{1,t+1}}{\xi_{1,t+1}n_{1,t+1} + \xi_{2,t+1}n_{2,t+1}}
\]

Under these assumptions, the competitive dynamics are

\[
\begin{align*}
n_{1,t+1} &= (1 - \delta)n_{1,t} + \frac{\delta}{\xi_{1,t+1}n_{1,t+1} + \xi_{2,t+1}n_{2,t+1}}
\end{align*}
\]

While this model appears to be two dimensional, in fact, it is one dimensional as \( n_{1,t} + n_{2,t} = 1 \) (Fig. 1).

We begin by examining the dynamics of this model in two special cases. First, lets assume that \( \xi_{i,t} = \xi_i \) are constant and positive. Then (see the exercises) one can prove that the species with the larger \( \xi_i \) value, say species 1, excludes the other i.e.

\[
\lim_{t \to \infty} n_t = (1, 0)
\]

whenever \( n_{1,0} > 0 \).

For the second case, we assume \( \xi_{i,1}, \xi_{i,2}, \ldots \) are ergodic stationary sequences of positive random variables, and \( \delta = 1 \) i.e. all individuals die after giving birth (semelparity). If we define \( x_t = n_{1,t}/n_{2,t} \)
1. THE LOTTERY MODEL

Figure 1. State-space for the lottery model $n_1 + n_2 = 1$. (1, 0) and (0, 1) correspond to equilibria of the model due to the no cats, no kittens principle.

(assuming $n_{2,t} > 0$), then $y_t$ satisfies the linear, stochastic difference equation

$$x_{t+1} = \frac{\xi_{1,t+1}}{\xi_{2,t+1}} x_t$$

Hence,

$$\log x_t = \log x_0 + \sum_{s=1}^{t} \log \frac{\xi_{1,s}}{\xi_{2,s}}$$

Provided that $\mathbb{E}[\log^+ \eta_t] < 0$, the Birkhoff Ergodic Theorem implies

$$\lim_{t \to \infty} \frac{1}{t} \log x_t = \mathbb{E}[\log \xi_{1,t} - \log \xi_{2,t}]$$

with probability one. Hence, the species with the larger geometric mean of $\xi_{i,t}$ displaces the other species.

The first special case is an example of Gause’s competitive exclusion principle: two species competing for the same resource in the same way can not coexist at equilibrium. Due to the fact that many species do coexist despite competing, ecologists have been studying coexistence mechanisms for nearly a century. As stated in the quote by Hutchinson at the beginning of this chapter, fluctuations in environmental conditions might allow for coexistence especially if different species are favored at different times. The second special case of the lottery model shows that fluctuations, in and of themselves, are insufficient to mediate coexistence if the populations have non-overlapping generations. This raises the question, can coexistence occur if there are fluctuations and overlapping generations?

To answer this question, we can, as in the single species models, look at what happens when each species becomes rare. When species 1 is rare, we have $n_{1,t} \approx 0$ and $n_{2,t} \approx 1$ and we can approximate
the dynamics of species 1 by the stochastic linear difference equation
\[ n_{1,t+1} \approx (1 - d)n_{1,t} + d \frac{\xi_{1,t+1}n_t}{\xi_{2,t+1}} = \left(1 - d\right) + d \frac{\xi_{1,t+1}}{\xi_{2,t+1}} n_{1,t} \]

Hence, the per-capita growth rate of species 1 when rare is
\[ r_1 = \mathbb{E} \left[ \log \left( 1 - d + \frac{\xi_{1,t}}{\xi_{2,t}} \right) \right] \]

Similarly, the per-capita growth rate of species 2 when rare is
\[ r_2 = \mathbb{E} \left[ \log \left( 1 - d + \frac{\xi_{2,t}}{\xi_{1,t}} \right) \right] \]

Intuitively, we would expect that if \( r_1 > 0 \) and \( r_2 > 0 \) (i.e. both species tend to increase when rare), then the species should coexist. Alternatively, if \( r_1 < 0 \) or \( r_2 < 0 \), then the species with the negative \( r_i \) are prone to extinction whenever they become rare. Indeed, one can prove the following theorem using methods from the single species models (see Exercises). Chesson [1982] provided the first proof of several of the statements in this Theorem.

**Theorem 4.1.** If \( r_1 > 0 \) and \( r_2 > 0 \), then the species coexistence in the sense of stochastic persistence in distribution and empirical measures i.e. for all \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that
\[ \mathbb{P}[n_{1,t} \in (\delta, 1 - \delta)] \geq 1 - \varepsilon \]

whenever \( n_{1,0} \in (0, 1) \) and \( t \) is sufficiently large, and
\[ \liminf_{t \to \infty} \frac{\#\{1 \leq s \leq t : n_{1,t} \in (\delta, 1 - \delta)\}}{t} \geq 1 - \varepsilon \]

with probability one whenever \( n_{1,0} \in (0, 1) \).

If \( r_1 < 0 \), then for all \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that
\[ \mathbb{P}[\lim_{t \to \infty} n_{i,t} = 0|n_0] = 1 \]

whenever \( n_{i,0} \leq \delta \).

But can \( r_1 \) and \( r_2 \) both be positive? In general understanding the sign of the \( r_i \) can be challenging. However, in the special case that the \( \xi_{i,t} \) are log-normally distributed and \( d \) is small, explicit interpretable expressions are possible. To see why, let \( r_i(d) \) denote \( r_i \) as a function of \( d \). These functions (at least for nice \( \xi_{i,t} \)) are analytic. We have (check for yourself!) \( r_i(0) = 0 \) and
\[ r_i'(0) = \mathbb{E} \left[ \frac{\xi_{i,t}}{\xi_{j,t}} \right] - 1 \text{ where } j \neq i \]

Hence, if \( \mathbb{E} \left[ \frac{\xi_{i,t}}{\xi_{j,t}} \right] > 1 \), then \( r_i(d) > 0 \) for sufficiently small \( d \).

Now, assume that \( \xi_{i,t} \) are log-normally distributed with log means \( \mu_i \), log variances \( \sigma_i^2 \) and correlation \( \rho \). The first term of \( r_i'(0) \), \( \mathbb{E} \left[ \frac{\xi_{i,t}}{\xi_{j,t}} \right] \), is the expectation of a log-normal random variable with log mean \( \mu_1 - \mu_2 \) and log variance \( \sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2 \). Hence,
\[ r_i'(0) = \exp(\mu_1 - \mu_2 + \sigma_1^2/2 - \rho\sigma_1\sigma_2 + \sigma_2^2/2) - 1 \]

This term is positive if
\[ \sigma_1^2/2 - \rho\sigma_1\sigma_2 + \sigma_2^2/2 > \mu_2 - \mu_1 \]

Similarly, \( r_2'(0) \) is positive if
\[ \sigma_1^2/2 - \rho\sigma_1\sigma_2 + \sigma_2^2/2 > \mu_1 - \mu_2 \]
Equivalently, both terms are positive if

\[ \frac{\sigma_1^2}{2} - \rho \sigma_1 \sigma_2 + \frac{\sigma_2^2}{2} > |\mu_1 - \mu_2| \]

which occurs if (i) the variances are larger relative to the mean log fitness difference \(|\mu_1 - \mu_2|\) and (ii) if the correlation between the two species fitnesses isn’t too close to one.

Thus, our analysis has identified two of the key ingredient’s of the “storage effect” due to Chesson and Warner [1981]. Namely, coexistence occurs if each species has a period in which it outperforms the other (i.e. the fitness variance is sufficiently high and the correlation is sufficiently low), and each species can store the gains during these favorable periods until the next favorable period (i.e. \(d\) is sufficiently low). The third key ingredient is “hidden” in this model but is uncovered in the next section. Figure 2 illustrates coexistence (left) and exclusion dynamics (right). Figure 3 illustrates how the conditions for coexistence depend on \(d\) and \(\sigma_i^2 = \sigma_i^2\) for \(i = 1, 2\). Consistent with our analysis, coexistence requires that \(\sigma^2\) is sufficiently large and \(d\) is sufficiently small.

2. Two species competition

And NUH is the letter I use to spell Nutches, Who live in small caves, known as Niches, for nutches. These Nutches have troubles, the biggest of which is the fact there are many more Nutches than Niches. Each Nutch in a Nich knows that some other Nutch Would like to move into his Nich very much. So each Nutch in a Nich has to watch that small Nich or Nutches who haven’t got Niches will snitch. –Dr. Seuss, On Beyond the Zebra

The lottery model identifies that under certain conditions, temporal fluctuations can allow for coexistence. A key component of this coexistence mechanisms is that there are years (i.e. due to different environmental conditions) where each species does better than the other species. As suggested by the quote of Dr. Seuss, this temporal partitioning of good years amongst the species can be viewed as each
species having a niche in time. In the lottery model, however, the total community density \( n_1 + n_2 = 1 \) remains constant for all time. While there are some competitive communities where this assumption nearly holds (e.g. the forest communities on Barro Colorado Island), many communities total densities fluctuate in time and, consequently, are best described by truly multivariate models. We begin by examining two species competition models to clarify some of the issues that arise in analyzing their dynamics.

Let \( n_t = (n_{1,t}, n_{2,t}) \) is the vector of densities of the two interacting species at time \( t \). As with the single species models, each species has a fitness \( f_i(n, \xi) \) that depends on the environmental state \( \xi \in \mathcal{E} \) and community state \( n = (n_1, n_2) \). In a fluctuating environment, the dynamics of these species are given by the coupled difference equations

\[
\begin{align*}
  n_{1,t+1} &= n_{1,t} f_1(n_t, \xi_{t+1}) \\
  n_{2,t+1} &= n_{2,t} f_2(n_t, \xi_{t+1})
\end{align*}
\]

where \( f_i \) are continuous, positive functions. We will assume throughout this section that \( \xi_1, \xi_2, \ldots \) are a sequence of independent and identically distributed random variables.

Let's assume the species are competing intraspecifically and interspecifically i.e. \( f_i \) are decreasing functions of both \( n_1 \) and \( n_2 \). Let \( e_1 = (1, 0) \) and \( e_2 = (0, 1) \) be the standard basis elements. To focus on their interactions, let's assume that each species stochastically persists in the absence of the other and the dynamics converge to a unique stationary distribution. Namely, for species 1 (and similarly for species 2), the single species model

\[
\begin{align*}
  n_{1,t+1} &= n_{1,t} f_1(n_{1,t} e_1, \xi_t) \\
  n_{2,t+1} &= n_{2,t} f_2(n_{2,t} e_2, \xi_t)
\end{align*}
\]
satisfies conditions of either Theorems 3.7 or 3.6 of Chapter 3. Thus, there is a positive, random variable \( \hat{n}_1 \) such that \( n_{1,t} \) converges to \( \hat{n}_1 \) in distribution and empirically i.e.

\[
\lim_{t \to \infty} \mathbb{E}[h(n_{1,t})|n_0] = \mathbb{E}[h(\hat{n}_1)]
\]

and

\[
\lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} h(n_{1,s}) = \mathbb{E}[h(\hat{n}_1)] \text{ with probability one}
\]

for all continuous functions \( h : [0, \infty) \to \mathbb{R} \) such that \( \mathbb{E}[|h(\hat{n}_1)|] < \infty \). Let \( \hat{n}_2 \) be corresponding stationary distribution for species 2. In particular, we have \( r_i(0) = \mathbb{E}[\log a_i(0, \xi)] > 0 \) for each of the species.

To determine whether the species coexist or not, we examine what happens when one species rare, say species 2. If \( n_{2,t} \approx 0 \), then we can approximate the dynamics by the partially coupled system

\[
n_{1,t+1} \approx n_{1,t}f_1(n_{1,t}e_1, \xi_{t+1})
\]

\[
n_{2,t+1} \approx n_{2,t}f_2(n_{1,t}e_1, \xi_{t+1})
\]

where, recall, that \( e_1 = (1,0) \). The second equation is (approximately) a stochastic linear difference equation and the per-capita growth rate of species 2 is given approximately by

\[
\lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} \log f_2(n_{1,s}e_1, \xi_{s+1})
\]

whenever the limit exists. As \( n_{1,t} \) converges empirically to \( \hat{n}_1 \) and \( \xi_t \) are ergodic, we would expect that this limit converges with probability one

\[
\mathbb{E}[\log f_2(\hat{n}_1e_1, \xi_t)] =: r_2(\hat{n}_1)
\]

where \( \xi_t \) is independent of \( \hat{n}_1 \). Indeed, this fact is true (see, Section XX). If \( r_2(\hat{n}_1) > 0 \), we expect species 2 to increase when rare. If \( r_2(\hat{n}_1) < 0 \), then we expect species 2 to decrease when rare. We can similarly define \( r_1(\hat{n}_2) \).

Intuitively, if each species tends to increase when rare i.e. \( r_1(\hat{n}_2) > 0 \) and \( r_2(\hat{n}_1) > 0 \), then we expect the species to coexist. This “mutual invasibility” condition was introduced by Turelli [1978] who wrote

the theory is based on a heuristic analytical approximation that provides conditions under which a rare invading species can increase in the presence of a community of established competitors.

Under the aforementioned assumptions and two additional assumptions \(^1\), Chesson and Ellner [1989] proved the following result.

**Theorem 4.2.** (Coexistence of competitors) If \( r_1(\hat{n}_2) > 0 \) and \( r_2(\hat{n}_2) > 0 \), then the system stochastically persistent in distribution i.e. for all \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that

\[
\mathbb{P}[n_t \in (\delta, \infty)^2|n_0] \geq 1 - \varepsilon
\]

for all \( n_0 \gg 0 \) and \( t \) sufficiently large.

In fact, I think one can also prove that

\[
\lim_{t \to \infty} \inf \frac{\#\{1 \leq s \leq t : n_t \in (\delta, \infty)^2\}}{t} \geq 1 - \varepsilon
\]

\(^1\)Let \( e_1 = (1,0) \) and \( e_2 = (0,1) \). Then the additional assumptions are: (i) \( \mathbb{P}[f_i(xe_i, \xi_t) = 1] < 1 \) for all \( x > 0 \), and (ii) there are positive constants \( c_i > 0 \) such that \( \mathbb{E}[\inf_{0<x<\infty} \log f_i(xe_i + \hat{n}_j e_j, \xi_t)] > -\infty \) where \( \hat{n}_j \) has the stationary distribution and is independent of \( \xi_t \).
with probability one for \( n_{i,0} > 0 \) for \( i = 1, 2 \). Again, these statements mean there is a small probability that any species will be at a low density, and the fraction of time any species spends at a low density is small. It also worth noting that the results from Chapter 3 give sufficient conditions for stochastic boundedness of both species in terms of the proper functions \( V \).

**Example 4.1.** (*Exploitative competition and the storage effect*) Consider two species competing for a single limiting resource (e.g. space, water, nutrients). Individuals from species extract the resource at an “intensity” \( a_i \). This intensity is determined by the length of time individuals extract the resource as well as how efficient they are in extracting the resource. Assume the resource availability is a decreasing function of the total rate at which individuals of either species extract the resource. Namely, there is a decreasing function, call it \( R \), such that the resource level equals \( R(\sum_j a_j n_j) \). For species \( i \), the resource extracted per individual is \( a_i R(\sum_j a_j n_j) \). Let \( b_i \) be the number of offspring produced per unit of extracted resource. Let \( s_i \) be the probability that an adult (i.e. not one of the new offspring) survives to the next year. Under these assumptions, we get

\[
f_i(\xi, n) = a_i b_i R(\sum_j a_j n_j) + s_i \text{ where } \xi = (a_1, a_2, b_1, b_2, s_1, s_2)
\]

For various choices of \( R \) (e.g. \( R(x) = \frac{1}{1+x^2} \) or \( (1 - \exp(-x))/x \)), one can show (see the exercises) that without environmental fluctuations in the parameters, the species with the larger value of \( a_i b_i/(1-s_i) \) competitively excludes the other species.

To understand how environmental fluctuations may alter this outcome, we follow Chesson [1988] and consider the degenerate (but very informative) case where the species are competitively equivalent in the absence of these fluctuations. Namely, \( a_1 = a_2 = a, b_1 = b_2 = b, s_1 = s_2 = s \) for all time. In this case, the deterministic dynamics are

\[
n_{i,t+1} = n_{i,t} (baR(a(n_{1,t} + n_{2,t})) + s) \text{ with } i = 1, 2.
\]

As \( n_{1,t+1}/n_{2,t+1} = n_{1,t}/n_{2,t} \) for all \( t \), all radial lines in the positive orthant are invariant. Consequently, provided \( baR(0) + s > 1 \), there exists a line of equilibria connecting the two axes. Regarding these neutral dynamics, Chesson [1988] wrote

Classically, when faced with a deterministic model of this sort ecologists have concluded that only one species can persist when the likely effects a stochastic environment are taken into account. The reason for this conclusion is the argument that environmental perturbations will cause a random walk to take place in which eventually all but one species becomes extinct. Our conclusion supports this conclusion in only a narrow range of circumstances.

To see how Chesson arrived at the final sentence, lets consider injecting stochasticity in the extraction intensities. Specifically, the \( a_{i,t} \) are independent and identically distributed in time. Moreover, to make things tractable and to add no systematic advantage of one species of the other, we follow Chesson and assume that \( a_{1,t} \) and \( a_{2,t} \) are exchangeable: \( \mathbb{P}[(a_{1,t}, a_{2,t}) \in \mathcal{A}] = \mathbb{P}[(a_{2,t}, a_{1,t}) \in \mathcal{A}] \) for any Borel set \( \mathcal{A} \subset \mathbb{R}^2 \).

We can write

\[
r_2(\hat{n}_1) = \mathbb{E}[\log (ba_{2,t}R(a_{1,t}\hat{n}_j) + s)] = \mathbb{E}[h(a_{1,t}, a_{2,t})]
\]

where

\[
h(a_1, a_2) = \mathbb{E}[\log (ba_2 R(a_1 \hat{n}_1) + s)]
\]

By the fundamental theorem of calculus,

\[
\int_{a_1}^{a_2} \int_{a_1}^{a_2} \frac{\partial^2 h}{\partial a_1 \partial a_2} (x, y) \, dx \, dy = \int_{a_1}^{a_2} \frac{\partial h}{\partial a_2} (a_2, y) - \frac{\partial h}{\partial a_1} (a_1, y) \, dy
\]

\[
= h(a_2, a_2) - h(a_2, a_1) - [h(a_1, a_2) - h(a_1, a_1)]
\]
As \( r_1(\hat{n}_1) = r_2(\hat{n}_2) = 0 \), we have \( \mathbb{E}[h(a_{2,t}, a_{2,t})] = \mathbb{E}[h(a_{1,t}, a_{1,t})] = 0 \). Hence,
\[
\mathbb{E}[h(a_{2,t}, a_{2,t}) - h(a_{2,t}, a_{1,t}) - [h(a_{1,t}, a_{2,t}) - h(a_{1,t}, a_{1,t})]] = -2r_2(\hat{n}_1) = -2r_1(\hat{n}_2)
\]
and
\[
r_2(\hat{n}_1) = -\frac{1}{2} \mathbb{E} \left[ \int_{a_{1,t}}^{a_{2,t}} \int_{a_1}^{a_2} \frac{\partial^2 h}{\partial a_1 \partial a_2} (x, y) \, dx \, dy \right]
\]
Computing this mixed partial yields
\[
\frac{\partial^2}{\partial a_1 \partial a_2} \log (ba_2 R(a_1 \hat{n}_1) + s) = \frac{\partial}{\partial a_1} \frac{bR(a_1 \hat{n}_1)}{ba_2 R(a_1 \hat{n}_1) + s} = \frac{bR'(a_1 \hat{n}_1) n_1 s}{(ba_2 R(a_1 \hat{n}_1) + s)^2}
\]
which is negative provided that \( s > 0 \). Hence, provided there are fluctuations and they are not perfectly correlated (i.e. \( a_{1,t} \neq a_{2,t} \) with positive probability), we get \( r_2(\hat{n}_1) = r_1(\hat{n}_2) > 0 \) and the species coexist. Hence, as pointed out by Chesson, contrary to naive expectations, adding fluctuations in the \( a_{i,t} \) does lead to a random walk toward extinction. Rather it facilitates coexistence.

We can also perform the same analysis with respect to fluctuations in the \( b_0 \) or the \( ss \) (try it for yourselves!). However, in each of these cases, the mixed partials of the corresponding \( h \) function equal zero. Hence, fluctuations in these parameter have no effect on the \( r_i \) (namely, they remain zero).

This highlights that the coexistence mechanism for this model requires three things: (i) there are times at which each species has a higher per-capita growth rate than the other (i.e. \( a_{i,t} \) vary but don’t perfectly co-vary), (ii) growth realized during good years can be “stored” through bad years (i.e. \( s > 0 \)), and (iii) during “good” years for each species, its increase in their per-capita growth rate is higher in those years which are worse for their competitor, than years which are also good for their competitor (i.e. the mixed partial is negative). Hence, there is more nuance to this effect than as first envisioned by Hutchinson in the quote at the start of this chapter. Chesson [1988] calls this trifecta “the storage effect.”

What happens when \( r_1(\hat{n}_2) < 0 \) or \( r_2(\hat{n}_1) < 0 \)? Under two additional technical assumptions, Chesson and Ellner [1989] proved

**Theorem 4.3. (Exclusion)** If \( r_i(\hat{n}_j) < 0 \) for some \( i = 1, 2 \) with \( j \neq i \), then \( \mathbb{P}[\lim_{t \to \infty} n_{i,t} = 0 | n_0] \) can be made arbitrarily close to 1 whenever \( n_{j,0} > 0 \) and \( n_{i,0} \) is sufficiently small.

If \( r_2(\hat{n}_1) > 0 \), \( r_1(\hat{n}_2) < 0 \), and \((0, \varepsilon) \times [\varepsilon, \infty)\) is accessible from \((0, \infty)^2\) for all \( \varepsilon > 0 \) sufficiently small, then \( \lim_{t \to \infty} n_{1,t} = 0 \) with probability one whenever \( n_{2,0} > 0 \).

If \( r_2(\hat{n}_1) < 0 \), \( r_1(\hat{n}_2) > 0 \), \((0, \varepsilon) \times [\varepsilon, \infty)\) is accessible from \((0, \infty)^2\) for all \( \varepsilon > 0 \) sufficiently small, and an additional technical assumption (see A.6 in [Chesson and Ellner, 1989]), then
\[
\mathbb{P} \left[ \lim_{t \to \infty} n_{i,t} = 0 \text{ for some } i \right] = 1
\]
whenever \( n_{i,0} > 0 \) for \( i = 1, 2 \).

**Example 4.2. (Reproductive interference for competing species)** This example is based on work in progress with Sharon Strauss Let \( \lambda_i \) be the maximal number of offspring produced by a mated individual of species \( i \), \( \alpha_i \) the strength of intraspecific competition, and \( \beta_j \) the strength of interspecific competition. In the presence of conspecific and heterospecific individuals, the expected number of offspring produced by a mated individual of species \( i \) is
\[
\frac{\lambda_i}{1 + \alpha_i n_i + \beta_j n_j}
\]
with \( j \neq i \).
2. TWO SPECIES COMPETITION

Figure 4. Dynamics of coexistence and exclusion for the stochastic exploitative competition model. No storage \((s = 0)\) on the left, lots of storage \((s = 0.8)\) on the right.

We assume the fraction of mated individuals increases with conspecific density and decreases with heterospecific density:

\[
\frac{n_i}{n_i + b_j n_j}
\]

where \(b_j\) measures the strength of reproductive interference of species \(j\) on species \(i\). One interpretation of this fraction is that the populations are panmictic and probability of mating with a heterospecific or conspecific is proportional to their population densities. Alternatively, equation (4.4) corresponds to the fraction of ovules or eggs that get fertilized (at the individual or population scale) by conspecific pollen or sperm.

We assume that environmental fluctuations cause fluctuations in the low-density reproductive rates. Specifically, \(\lambda_{i,t}\) are given by log-normally distributed random variables with log means \(\mu_i\), log variances \(\sigma_i^2\), and log correlation \(\rho\). Putting together these components together yields the following model

\[
\begin{align*}
    n_{1,t+1} &= n_{1,t} \frac{n_{1,t}}{n_{1,t} + b_2 n_{2,t}} \frac{\lambda_{1,t}}{1 + \alpha_1 n_{1,t} + \beta_2 n_{2,t}} \\
    &= \text{\% mated} & \text{\# potential offspring} \\
    n_{2,t+1} &= n_{2,t} \frac{n_{2,t}}{n_{2,t} + b_1 n_{1,t}} \frac{\lambda_{2,t}}{1 + \alpha_2 n_{2,t} + \beta_1 n_{1,t}}
\end{align*}
\]

(4.5)

If \(\mu_i > 0\), the single species systems have a unique positive stationary distribution by results at the end of Chapter 3 (both theorems apply!). Furthermore, the dynamics are stochastically bounded using \(V(n_1, n_2) = n_1 + n_2\). Provided that \(\rho < 1\), the sets \([\varepsilon, \infty) \times (0, \varepsilon)\) are accessible for all \(\varepsilon > 0\). As \(r_i(\hat{n}_j) = -\infty\) for \(i = 1, 2\) and \(j \neq i\), the previous theorem implies that whenever \(\rho < 1\), one species is
3. Many species

Partly to escape these complications, we count like the Australian Arunta Tribe, “one, two, many, ” and move on...directly to multispecies communities. – Robert May [1981]

For two species communities, the “mutual invasibility” paradigm does a good job of characterizing whether species coexist or not. However, does this condition extend to communities with three or more species? There are two natural candidates for extending the mutual invasibility condition. The first is that every missing species can invade a stationary distribution supporting a subset of species. While we will show this condition is sufficient for coexistence, it is generally too strong of a condition.

---

4Of course once allows for Allee effects or obligate mutualisms, this paradigm is not very useful. Many opportunities for new results in this area but they are likely to be much more nuanced if one allows for contingent coexistence.
For example, in a system consisting of a plant, herbivore, and predator, the herbivore might be able to have a positive per-capita growth rate when introduced at low densities in the plant-only system, but the predator (assuming its not an omnivore) would not. However, the predator could have a positive per-capita growth rate when introduced at low densities in the plant-herbivore system (assuming these two species coexist).

This simple gedanken experiment suggests another way to extend the “mutual invasibility” condition. Namely, for every stationary distribution supporting a subset of species, at least one of the missing species can invade i.e. have a positive per-capita growth rate at low densities. While this condition is necessary for the form of coexistence that we focus on, it is not sufficient in general. For example, May and Leonard [1975] introduced a deterministic model of three competing species engaged in a “rock-paper-scissors” game: species 2 displaces species 1 in pairwise competition, species 3 displaces species 2, and species 1 displaces species 3. In this system, the only subcommunities consist of a single species and one of the missing species can always invade these subcommunities. However, May and Leonard [1975] showed numerically that trajectories of this system can cycle toward extinction (See Figure 7 for the stochastically counterpart). It wasn’t until many years later that mathematicians (see, Hofbauer and Sigmund 1998 for a delightful exposition on this theory in the context of Lotka-Volterra systems) developed a theory to identify when these “rock-paper-scissors” systems are extinction prone or allow for coexistence. Here we introduce the stochastic analog of this deterministic theory.

For this theory, we study the dynamics of $k$ interacting populations in a random environment. Let $n_i^t$ denote the density of the $i$-th population at time $t$ and $n_t = (n_1^t, \ldots, n_k^t)$ the vector of population densities at time $t$. As before, we let $\xi_t$ represents the state of the environment (e.g. temperature, nutrient availability) at time $t$. The fitness $f_i(n_t, \xi_{t+1})$ of population $i$ at time $t$ depends on the population state and environmental state at time $t + 1$. Under these assumptions, we arrive at the following stochastic difference equation:

$$n_{i,t+1} = n_{i,t} f_i(n_t, \xi_{t+1}).$$

Following Schreiber et al. [2011], we make four assumptions:

**A1:** There exists a compact set $S$ of $\mathbb{R}_+^k = \{n \in \mathbb{R}^k : x_i \geq 0 \}$ such that $n_t \in S$ for all $t \geq 0$.

**A2:** $\{\xi_t\}_{t=0}^\infty$ is a sequence of i.i.d random variables taking values in a separable metric space $E$, having distribution $m$.

**A3:** $f_i(n, \xi)$ are strictly positive functions, continuous in $n$ and measurable in $\xi$.

**A4:** For all $i$, $\sup_{n \in S} \int (\log f_i(n, \xi))^2 m(d\xi) < \infty$

Assumption **A1** ensures that the populations remain bounded for all time. Currently there is active work being done to relax this assumption e.g. the work of Alex Henning on SDE counterparts to this theory. Assumptions **A2** and **A3** imply that $\{X_t\}_{t=0}^\infty$ is a Markov chain on $S$ and that $\{X_t\}_{t=0}^\infty$ is *Feller*.

Assumption **A4** is a technical assumption meet by many models.

The expected per-capita growth rate at state $n$ of population $i$ is

$$r_i(n) = \mathbb{E}[\log f_i(n, \xi_t)].$$

As discussed before, when $r_i(n) > 0$, the $i$-th population tends to increase when the current population state is $n$. When $r_i(n) < 0$, the $i$-th population tends to decrease when the current population state is $n$. Of course, the other species rarely will be a fixed state $n$. Rather they will fluctuate around some stationary distribution. Recall that a random vector $\hat{n}$ (independent of $\xi_t$ for all $t$) has a stationary distribution for (4.7) if

$$\mathbb{P}[n_{t+1} \in \mathcal{A}|n_t = \hat{n}] = \mathbb{P}[\hat{n} \in \mathcal{A}]$$

for every Borel set $\mathcal{A} \subset S$. Furthermore, this stationary distribution is ergodic if $\hat{n}$ can not be written as a non-trivial convex combination of other random vectors with a stationary distribution. We define

---

5 For this process, we can define the operator $P h(n) = \mathbb{E}[h(n_{t+1})|n_t = n]$ for any integrable function function $h : S \rightarrow \mathbb{R}$. The Markov chain $n_t$ is *Feller* if $P h$ is continuous whenever $h$ is continuous.
the invasion rate of species $i$ with respect to $\hat{n}$ as
\[ r_i(\hat{n}) = \mathbb{E}[\log f_i(\hat{n}, \xi_i)] \]
The following theorem clarifies why $r_i(\hat{n})$ is called an invasion rate. A proof is given in [Schreiber et al., 2011, Appendix A].

**Theorem 4.4.** Let $\hat{n}$ have a distribution which is stationary and ergodic for (4.7). If $n_0 = \hat{n}$, then
\[ \lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} \log f_i(n_s, \xi_{s+1}) = r_i(\hat{n}) \]
with probability one for each species $i$.

Before stating the main stochastic persistence theorem, we make a useful observation about the $r_i(\hat{n})$ for species “supported” by $\hat{n}$. Let $\hat{n}$ be as in the proposition. We say species $i$ is supported if $\mathbb{P}[\hat{n}_i > 0] > 0$. Ergodicity implies that $\mathbb{P}[\hat{n}_i > 0] = 1$ (otherwise one could decompose $\hat{n}$ into two pieces one of which has $\hat{n}_i = 0$ with probability one). What can we say about $r_i(\hat{n})$? Choose $m > 0$ such that $\mathbb{P}[1 - m > \hat{n}_i > m] > 1/10$. Let $n_0 = \hat{n}$. By Birkhoff’s ergodic theorem, $n_{i,t}$ enters the interval $[m, 1/m]$ infinitely often i.e. there are random times $t_1 < t_2 < t_3 < \ldots$ such that $m \leq n_{i,t_k} \leq 1/m$ with probability one for all $k$. Hence, with probability one
\[ 0 = \lim_{k \to \infty} \frac{1}{t_k} \log n_{i,t_k} \]
\[ = \lim_{k \to \infty} \frac{1}{t_k} \sum_{s=0}^{t_k-1} \log f_i(n_s, \xi_{s+1}) \]
\[ = r_i(\hat{n}) \]
where the first equality follows from $n_{i,t_k}$ lying in the interval $[m, 1/m]$ and $t_k \to \infty$ with probability one. Intuitively, a species supported by $\hat{n}$ (i.e. $\mathbb{P}[\hat{n}_i > 0] > 0$) is on average neither growing (as it dynamics is bounded from above) or on average decreasing (as it isn’t tending toward extinction). Hence, its average per-capita growth rate is zero.

Now we state the main persistence theorem. For all $\delta \geq 0$, define $S_\delta = S \cap (\delta, \infty)^k$ i.e. community states corresponding to all species having densities strictly greater than $\delta$. Then $S \setminus S_\delta$ corresponds to community states where at least one species is extinct.

**Theorem 4.5.** If there exists $p \in [0, 1]^k$ such that
\[ \sum_i p_i r_i(\hat{n}) > 0 \]
for all ergodic and stationary $\hat{n}$ supported by $S \setminus S_\delta$. Then for all $\epsilon > 0$ there exists $\delta > 0$ such that
\[ \liminf_{t \to \infty} \frac{\#\{1 \leq s \leq t : n_s \in S_\delta \}}{t} \geq 1 - \epsilon \]
almost surely whenever $n_0 \in S_\delta$.

**Outline of proof:** Upcoming!! Currently, working on a proof of the following partial converse: if there exists $p \in [0, 1]^k$ such that
\[ \sum_i p_i r_i(\hat{n}) < 0 \]
for all ergodic and stationary $\hat{n}$ supported by $S \setminus S_\delta$, then (4.7) is extinction prone i.e. for all $\epsilon > 0$ there exists $\delta > 0$ such that
\[ \mathbb{P}[\lim_{t \to \infty} \min_i n_{i,t} = 0 | n_0] \geq 1 - \epsilon \]
whenever $n_0 \in S \setminus S_\delta$. 
Example 4.3. (Two species models) In the exercises, you are asked to verify that Theorem 4.5 yields the mutual invasibility condition for competing species models from the previous section. Here, we examine a model of a predator-prey interaction where \( n_1 \) is the prey density and \( n_2 \) is the predator density. For the prey dynamics, we consider a classical stochastic Beverton-Holt model:

\[
n_{1,t+1} = n_{1,t} \frac{\xi_{t+1}}{1 + a n_{1,t}}
\]

where \( \xi_{t+1} \) corresponds to the low density fitness of the prey, and \( a > 0 \) determines the strength of intraspecific competition. The average, low density per-capita growth rate is \( r_1(0) = \mathbb{E}[\log \xi_t] \) which we assume is positive. By results from Chapter 3, we get there is a unique positive stationary distribution \( r \) in the presence of intraspecific competition. The average, low density per-capita growth rate is \( \bar{r} \) as desired.

If predators are randomly search for the prey with intensity \( b > 0 \), then the probability of a prey escapes predation is \( \exp(-bn_{2,t}) \). If each prey eaten leads to the production of \( c > 0 \) offspring (i.e. the conversion efficiency) and “adult” predators survive to the next year with probability \( 1 - \beta \) (where \( \beta \) is the probability of escape), then the right hand side of the predator equation is of the form

\[
\lambda_{n_{1,t}} h(n_{2,t}) + sn_{2,t}
\]

where \( n_{1,t} \) corresponds to the prey density, \( \lambda_{n_{1,t}} \) is the low density fitness of the prey, and \( a > 0 \) determines the strength of intraspecific competition. The average, low density net per-capita growth rate is \( r_2(0,0) = \mathbb{E}[\log \xi_t] \) which we assume is positive. By results from Chapter 3, we get there is a unique positive stationary distribution \( r \) for this single species model. Let \( \tilde{n}_1 \) be a random variable with this distribution which is independent of the \( \xi_t \) for all \( t \).

At first glance this equation doesn’t look like it is of the right form i.e. there isn’t a factor of \( n_{2,t} \) in the first term of the predator equation. However, if we define \( h(x) = (1 - \exp(-bx))/x \) (which is analytic! check for yourself), then the right hand side of the predator equation is of the form \( cn_{2,t}n_{1,t}h(n_{2,t}) + sn_{2,t} \) as desired.

In the absence of the prey, the predator dynamics are given by exponential decay to extinction i.e \( n_{2,t+1} = sn_{2,t} \) with \( s < 1 \). Hence, there is no positive stationary distribution on the predator axis. Hence, the only ergodic, stationary distributions correspond to a Dirac measure at \( (0,0) \) and the distribution of \( (\tilde{n}_1,0) \). The persistence criterion states that we need weights \( p_1, p_2 \) such that

\[
\sum_i p_i r_i(0,0) > 0 \quad \text{and} \quad \sum_i p_i r_i(\tilde{n}_1,0) > 0
\]

We have \( r_1(0,0) > 0 \) by assumption, \( r_2(0,0) = \log s < 0 \), and \( r_1(\tilde{n}_1,0) = 0 \) as species 1 is supported by this stationary distribution. Hence, the conditions become

\[
p_1 r_1(0,0) + p_2 \log s > 0 \quad \text{and} \quad p_2 r_2(\tilde{n}_1,0) > 0
\]

The second condition implies that we need \( r_2(\tilde{n}_1,0) > 0 \). Assume this is true. Then setting \( p_1 = 1 \) and choosing \( p_1 > 0 \) sufficiently small ensures that both conditions are meet. Hence, \( r_2(\tilde{n}_1,0) > 0 \) suffices for stochastic persistence of both species. On the other hand, suppose that \( r_2(\tilde{n}_1,0) < 0 \). Setting \( p_2 = 1 \) and choosing \( p_1 > 0 \) sufficiently small ensures the extinction prone condition is meet and the species do not stochastically persist.

What is \( r_2(\tilde{n}_1,0) \)? As \( h(0) = b \), we get

\[
r_2(\tilde{n}_1,0) = \mathbb{E}[\log(c\tilde{n}_1b + s)]
\]

To get a sense of how this invasion rate depends on the mean and variance of \( \tilde{n}_1 \), lets write \( \tilde{n}_1 = \bar{n}_1 + \sigma X \) where \( \bar{n}_1 = \mathbb{E}[\tilde{n}_1] \), \( \sigma^2 = \text{Var}[\tilde{n}_1] \) and \( \sigma X = \tilde{n}_1 - \bar{n}_1 \). If we assume that \( \sigma > 0 \) is sufficiently small (in the next chapter we will see when this is true), we get the following second order approximation by Taylor’s theorem (try this yourself!):

\[
r_2(\tilde{n}_1,0) \approx \log(bc\bar{n}_1 + s) - \frac{1}{2} \left( \frac{bc\sigma}{bc\bar{n}_1 + s} \right)^2
\]
Figure 6. Reduction in the predator invasion rate due to the change in the mean and variance of the prey density.

Equivalently,

\[ r_2(\hat{n}_1, 0) \approx \log \mathbb{E}[bc\hat{n}_1 + s] - \frac{1}{2} \text{CV}[bc\hat{n}_1 + s]^2 \]

where CV denotes the coefficient of variation of a random variable.

This expression implies that variation in the fitness of the prey might inhibit the persistence of the predator in two ways. First, the fitness variability via Jensen’s inequality typically decreases the mean prey density \( \overline{n} \). This results in a reduction in the first term of the predator’s per-capita growth rate. Second, the fitness variability increases the variance in the prey’s stationary distribution. This leads to an additional reduction through the second term of the approximation. This second reduction stems for the concavity of the log function i.e. Jensen’s inequality once more! Figure 6 computes the magnitude of these two effects when increasing the variance experienced by the prey.

Example 4.4. Lottery model of evolutionary games The classical lottery model, as discussed at the beginning of this chapter, assumes that individuals reproduce at rates independent of the frequencies of the other populations. Here, we examine a lottery model that accounts for frequency dependent interactions among different genotypes in an asexual population. Let \( n_{i,t} \) denote the fraction of space occupied by genotype \( i \) at time \( t \). Individuals are randomly interacting and the effect of these interactions on their reproduction is determined by a payoff matrix, \( A_t \), whose \( i-j \)-th entry corresponds to the payoff to strategy \( i \) when playing strategy \( j \). The number of offspring produced by strategy \( i \) is given by \( \sum_j A_{ji,t+1} n_{j,t} \). After producing offspring individuals die and the frequencies in the next generation are given by the frequencies of the offspring. Namely,

\[ n_{i,t+1} = \frac{\sum_j A_{ji,t+1} n_{j,t} n_{i,t}}{\sum_j A_{j\ell,t+1} n_{j,t} n_{\ell,t}} \]

The state space for this model is the \( k - 1 \) dimensional simplex: \( S = \{ x \in \mathbb{R}_+^k : \sum_i x_i = 1 \} \).
Now consider the evolutionary game of rock, paper, scissors (see lecture slides for empirical examples) where strategy 2 beats strategy 1, strategy 3 beats strategy 2, and strategy 1 beats strategy 3. Assume when species $i$ wins, losses, or has a draw, it gets payoff of $w_{i,t}$, $\ell_{i,t}$, and $d_{i,t}$, respectively. Then the payoff matrix becomes

$$A_t = \begin{pmatrix}
    d_{1,t} & \ell_{1,t} & w_{1,t} \\
    w_{2,t} & d_{2,t} & \ell_{2,t} \\
    \ell_{3,t} & w_{3,t} & d_{3,t}
\end{pmatrix}$$

Assume $w_{i,t} > d_{j,t} > \ell_{k,t}$ with probability one for all $t$ and $1 \leq i, j, k \leq 3$. Then one can show (see the exercises!!) that for the pairwise dynamics, you get the winners out competing the losers. Hence, there are only three ergodic stationary distributions corresponding to the $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ steady states.

Assume the $w_{i,t}$ all have the same distribution, $d_{i,t}$ all have the same distribution, and $\ell_{i,t}$ have the same distribution. Furthermore, assume there are no cross correlations. For any given strategy $i$, we can compute the per-capita growth rate $r_w$ of the strategy that beats that strategy and the per-capita growth rate $r_\ell$ of the strategy that gets between by the strategy. These are given by

$$r_w = \mathbb{E}[\log \frac{w_t}{d_t}] > 0 \text{ and } r_\ell = \mathbb{E}[\log \frac{\ell_t}{d_t}] < 0$$

where $w_t$, $d_t$, and $\ell_t$ have the common distributions and uncorrelated. Now for the strategies to coexist, we need to find weights $p_i > 0$ such that $\sum_i p_i r_i(e_j) > 0$ for all $j$ (Recall $e_j$ are standard basis element of $\mathbb{R}^3$.) By symmetry, we should be able to choose all the $p_i$ to be equal, say 1. In which case, we need $r_w + r_\ell > 0$. Equivalently, stochastic persistence requires that

$$\mathbb{E}[\log w_t] + \mathbb{E}[\log \ell_t] > 2\mathbb{E}[\log d_t].$$

Conversely, the reversed inequality implies that the strategies are extinction prone. Both forms of dynamics are illustrated in Figure 7.

**Example 4.5.** *(Stochastic Lotka-Volterra difference equations)*
Figure 7. UPCOMING
CHAPTER 5

Approximation methods

Finding closed form descriptions of the stationary distributions for the nonlinear difference equations considered here is typically, at best, difficult or impossible. Hence, to get some analytical traction on these stochastic models, we avail ourselves to a fundamental tactic of the applied mathematician: (i) identify models whose dynamics we understand and (ii) use approximation methods to understand the effects of small perturbations on the dynamics. We will use this tactic to examine how feedbacks between deterministic and stochastic forces determine the mean and co-variance structure of stationary distributions \( \hat{n} \) and the per-capita growth rates \( r_i(\hat{n}) \) of species not supported by \( \hat{n} \).

We begin illustrating the basic ideas for the stationary distribution of a single species model and the implications for per-capita growth rate of another species being introduced at low densities. Then we extend these results to multivariate models. We conclude by providing some theorems that allow one to analytically verify the persistence conditions from the previous chapter. The work being presented here is part of the in progress, thesis work of William Cuello. Hence, please let us know if you have you seen similar results developed elsewhere in a mathematically rigorous fashion.

1. The scalar case

Many of the results in this chapter extend to arbitrary stochastic difference equations and are easier to present in this level of generality. Hence, we begin by considering a general, scalar stochastic difference equation of the form

\[
\begin{align*}
    n_{t+1} &= g(n_t, \xi_{t+1})
\end{align*}
\]

where \( \xi_t \) are i.i.d. scalars (for simplicity right now) and \( g \) is at least three times continuously differentiable with respect to both of its arguments.

We make two standing assumptions for this section. First, \( \xi_t = \bar{\xi} + \sigma Z_t \) where \( Z_t \) has mean zero, variance 1, and compact support (i.e. there exists an interval \([a, b]\) such that \( \mathbb{P}[a \leq Z_t \leq b] = 1 \)). We use the parameter \( \sigma \) to control the magnitude of the noise. Second, we assume that \( \hat{n} \) is a linearly stable equilibrium for the deterministic model

\[
\begin{align*}
    n_{t+1} = g(n_t, \bar{\xi})
\end{align*}
\]

i.e the model corresponding to \( \sigma = 0 \). Namely, \( \hat{n} = g(\hat{n}, \bar{\xi}) \) and \( \left| \frac{\partial g}{\partial n}(\hat{n}, \bar{\xi}) \right| < 1 \).

Intuitively, if \( \sigma > 0 \) is sufficiently small, we should expect that \( n_t \) fluctuates around \( \hat{n} \) provided that \( n_0 \) is sufficiently close to \( \hat{n} \). Indeed, in the next section, we shall see that for sufficiently small \( \sigma > 0 \), there exists a random variable \( \hat{n} \) such that such that \( n_t \) converges in distribution and empirically to \( \hat{n} \) whenever \( n_0 \) is sufficiently close \( \hat{n} \).

To understand the distribution of \( \hat{n} \) when \( \sigma \) is sufficiently small, we will linearize the system around \((\bar{n}, \bar{\xi})\) using the first-order approximation

\[
\begin{align*}
    f(n, \xi) &\approx \bar{n} + \frac{\partial g}{\partial n}(\bar{n}, \bar{\xi})(n - \bar{n}) + \frac{\partial g}{\partial \xi}(\bar{n}, \bar{\xi})(\xi - \bar{\xi}).
\end{align*}
\]

65
1. THE SCALAR CASE

Our approximation \( x_t \) for \( n_t - \bar{n} \) is given by the ARM(1) model

\[
x_{t+1} = g_n x_t + g_\xi \sigma Z_{t+1}
\]

Our assumption of linear stability implies that \( |g_n| < 1 \). Hence, by the results from Chapter 2, we get that \( x_t \) converges to a distribution with mean 0 and variance \( \sigma^2 g_\xi^2 / (1 - g_n^2) \).

What does this linear approximation tell us about \( \hat{n} \)? In the next section, we show that

\[
\mathbb{E}[\hat{n}] = \bar{n} + O(\sigma^2) \\
\text{Var}[\hat{n}] = \frac{\sigma^2 g_\xi^2}{1 - g_n^2} + O(\sigma^3)
\]

Let's see how well these approximations work.

**Example 5.1.** *(Beverton Holt model revisited)* Consider the stochastic Beverton-Holt model:

\[
n_{t+1} = \frac{n_t \xi_{t+1}}{1 + n_t}
\]

where \( \xi_t = \bar{\xi} + \sigma Z_t \). Provided that \( \bar{\xi} > 1 \), the deterministic model

\[
n_{t+1} = \frac{n_t \bar{\xi}}{1 + n_t}
\]

has a unique positive equilibrium \( \bar{n} = \bar{\xi} - 1 \) which is linearly stable as

\[
g_n = \frac{\bar{\xi}}{(1 + \bar{n})^2} = \frac{1}{\bar{\xi}} < 1.
\]

We also have

\[
g_\eta = \frac{1}{1 + \bar{n}} = \frac{1}{\bar{\xi}}.
\]

Hence,

\[
\mathbb{E}[\hat{n}] = \bar{n} - 1 + O(\sigma^2) \\
\text{Var}[\hat{n}] = \frac{\sigma^2}{\bar{\xi}^2 - 1} + O(\sigma^3)
\]

The approximation, as shown in Figure 1, of the variance is exceptionally good. However, for estimating \( \mathbb{E}[\hat{n}] \), the error of order \( \sigma^2 \) (all that we are guaranteed) is substantial.

Motivated by this example, we derive a second order correction term for \( \mathbb{E}[\hat{n}] \). To this end, we use the 2nd order approximation of \( g \) around \((\bar{n}, \bar{\xi})\):

\[
g(n, \xi) = \bar{n} + g_n (n - \bar{n}) + g_\xi (\xi - \bar{\xi}) + \frac{g_{nn} (n - \bar{n})^2}{2} + g_{n\xi} (n - \bar{n}) (\xi - \bar{\xi}) + \frac{g_{\xi\xi}}{2} (\xi - \bar{\xi})^2 + O(\sigma^3)
\]

where \( g_{nn}, g_{n\xi}, g_{\xi\xi} \) are the second order derivatives of \( g \) evaluated at \( n = \bar{n}, \xi = \bar{\xi} \). By stationarity, if \( n_0 = \bar{n} \), then

\[
\mathbb{E}[\hat{n}] = \mathbb{E}[g(\hat{n}, \xi_t)] \\
\quad = \bar{n} + g_n \mathbb{E}[\hat{n} - \bar{n}] + g_\xi \mathbb{E}[\xi_t - \bar{\xi}] + \frac{g_{nn}}{2} \mathbb{E}[(n - \bar{n})^2] + g_{n\xi} \mathbb{E}[(n - \bar{n})(\xi_t - \bar{\xi})] + \frac{g_{\xi\xi}}{2} \mathbb{E}[(\xi_t - \bar{\xi})^2] \\
\quad + O(\sigma^3)
\]

\[
(1 - g_n) \mathbb{E}[\hat{n} - \bar{n}] = \frac{g_{nn}}{2} \frac{\sigma^2 g_\xi^2}{1 - g_n^2} + \frac{g_{\xi\xi}}{2} \sigma^2 + O(\sigma^3)
\]

Hence, we get the higher order approximation

\[
\mathbb{E}[\hat{n}] = \bar{n} + \frac{\sigma^2}{2(1 - g_n)} \left( \frac{g_{nn} g_\xi^2}{1 - g_n^2} + g_{\xi\xi} \right) + O(\sigma^3)
\]
Implementing this approximation for the stochastic Beverton Holt model (check for yourself!) yields the red curve in the right hand Figure 1.

We can use these analytical approximations of the expectation and variance of \( \hat{n} \) to get estimates of \( \mathbb{E}[h(\hat{n})] \) for any smooth function \( h \). Namely,

\[
\mathbb{E}[h(\hat{n})] = h(\bar{n}) + \frac{\sigma^2}{2} \left[ \frac{h'(\bar{n})}{1 - g_n} \left( g_{nn}g_n^2 + g_{\xi \xi} \right) + h''(\bar{n}) \frac{g_n^2}{1 - g_n^2} \right] + O(\sigma^3)
\]

To illustrate the utility of this approximation, we can use to understand the stochastic persistence of the predator-prey model introduced in the previous chapter.

**Example 5.2. (Predator-prey dynamics)** Consider the predator-prey model

\[
\begin{align*}
n_{1,t+1} &= n_{1,t} \frac{\xi_{t+1}}{1 + n_t} \exp(-bn_t) \\
n_{2,t+1} &= cn_{1,t}h(n_t) + sn_{2,t}
\end{align*}
\]

where \( h(x) = (1 - \exp(-bx))/x \). Provided that \( \mathbb{E}[\log \xi_t] > 0 \), we had shown that there is a unique positive stationary distribution \( \hat{n}_1 \) for the prey and the per-capita growth rate of the predator at this stationary distribution is

\[
r_2(\hat{n}_1, 0) = \mathbb{E}\left[\log(b\hat{n} + s)\right]
\]
Define $\bar{f} = bc\hat{n} + s$. Using equation (5.2) and our calculations from the previous example, we can get the estimate

$$r_2(\hat{n}_1, 0) = \log \bar{f} + \frac{\sigma^2}{2} \left[ \frac{bc}{\bar{f}} \frac{1}{1 - g_n} \left( \frac{g_{nn}g_{n}^2}{1 - g_{n}^2} + g_{n}g_{\xi} \right) - \left( \frac{bc}{\bar{f}} \right)^2 \frac{g_{n}^2}{1 - g_{n}^2} \right] + O(\sigma^3)$$

$$= \log \bar{f} + \frac{\sigma^2}{2} \left[ \frac{bc}{\bar{f}} \frac{\xi}{\xi - 1} \left( \frac{-2}{\xi^4 - \xi^2} \right) - \left( \frac{bc}{\bar{f}} \right)^2 \frac{1}{\xi^2 - 1} \right] + O(\sigma^3)$$

$$= \log \bar{f} - \frac{\sigma^2}{2(\xi^2 - 1)} \frac{bc}{\bar{f}} \left[ \frac{2}{\xi^2 - \xi} + \frac{bc}{\bar{f}} \right] + O(\sigma^3)$$

Hence, fluctuations in the prey’s fitness reduces the per-capita growth rate of the predator in two ways: the reduction in $E[\hat{n}]$ (the first term of the correction) and variance in $\hat{n}$ (the second term of the correction).

2. The Multivariate Case

This section is in progress... Basic idea is similar to the scalar case. As in the scalar case, we consider the more general stochastic model

$$n_{t+1} = g(n_t, \xi_{t+1})$$

where we assume

**Small noise:** $\xi_t = \bar{\xi} + \sigma Z_t$ lies in $\mathbb{R}^m$ for some $m$, $E[Z_t] = 0$, $\text{Var}[Z_t] = 1$, and $Z_t$ lies in a fixed compact set with probability one.

**Stable equilibrium:** $\bar{n}$ is a linearly stable equilibrium for the deterministic model

$$n_{t+1} = g(n_t, \bar{\xi})$$
William has proven the following proposition

**Proposition 5.1.** If $\sigma > 0$ is sufficiently small, then there exists a unique stationary distribution $\hat{n}$ such that $n_t$ converges to $\hat{n}$ in distribution and empirically provided $n_0$ is sufficiently close to $\bar{n}$.

As before, we use the linear approximation

$$g(n, \xi) \approx \bar{n} + (n - \bar{n})g_n + (\xi - \bar{\xi})g_\xi$$

where $g_n$ is the $k \times k$ derivative matrix of $g$ with respect to $n$ and $g_\xi$ is the $k \times m$ derivative matrix of $g$ with respect to $\xi$; both evaluated at $(\bar{n}, \bar{\xi})$. The dynamics of $n_t - \bar{n}$ are approximated by the AR(1) process:

$$x_{t+1} = x_t g_n + (\xi_{t+1} - \bar{\xi}) g_\xi.$$

As the spectral radius of $g_n$ is less than one, this AR(1) process has a unique stationary distribution with mean 0 and co-variance matrix $\Sigma^2$ as described in the final section of Chapter 2.

One can prove the following properties of these approximations.

**Proposition 5.2.**

$$\mathbb{E}[\hat{n}] = \bar{n} + O(\sigma^2) \text{ and } \text{Cov}[\hat{n}] = \Sigma^2 + O(\sigma^3)$$

To improve our approximation of $\mathbb{E}[\hat{n}]$, we can do as before and use the second-order approximation of $g$:

$$g(n, \xi) \approx \bar{n} + (n - \bar{n})g_n + (\xi - \bar{\xi})g_\xi + \frac{1}{2}g_{\xi\xi}(\xi - \bar{\xi}, \xi - \bar{\xi}) + \frac{1}{2}g_{nn}(n - \bar{n}, n - \bar{n}) + g_{\xi n}(\xi - \bar{\xi}, n - \bar{n})$$

where $g_{nn}, g_{\xi\xi}, g_{n\xi}$ are the second-order derivatives of $g$ which are 2-tensors and, consequently, act on pairs of vectors.
Index

accessible, 45
  $0, \infty$, 40
auto-regressive processes, 27
demographic stochasticity, 4
density-dependence
  negative, 35
  positive, 39
environmental stochasticity, 4
ergodic, 17
exchangeable, 56
irreducibility, 46
iteroparous, 43
lottery model, 49
Lyapunov exponent
  dominant, 27
per-capita growth rate
  sensitivity, 21
    first-order, 23
    second-order, 26
  stochastic, 17
Perron-Frobenius
  constant environment, 14
  random environment, 17
primitive
  constant environment, 14
  stationary environment, 17
proper function, 43
regulated, 35
reproductive values
  constant environment, 15
  random environment, 17
sedentary, 43
stable state distribution
  constant environment, 15
  random environment, 17
  stationary, 17
stationary distribution
  species support, 60
stochastically bounded
  from above, 43
Bibliography


