Chapter 5
Individual Stochasticity and Implicit Age Dependence

5.1 Introduction

Demography is the study of the population consequences of the fates of individuals. As an individual organism develops through its life cycle it may increase in size, change its morphology, develop new physiological functions, exhibit new behaviors, or move to new locations. It may marry and divorce, become ill and recover, or change its employment status. It may change sex and/or change its reproductive status. These changes can be dramatic. This developmental process, and its attendant risks of death and opportunities for reproduction, determine the rates of birth and death that, in turn, determine population growth or decline.

Individuals are differentiated on the basis of age or, in general, life cycle stages. The movement of an individual through its life cycle is a random process, and although the eventual destination (death) is certain, the pathways taken to that destination are stochastic and will differ even between identical individuals; this is *individual stochasticity*. A stage-classified demographic model contains implicit age-specific information, which can be analyzed using Markov chain methods. The living stages in the life cycles are transient states in an absorbing Markov chain, in which death is an absorbing state.

This chapter presents Markov chain methods for computing the mean and variance of the lifetime number of visits to any transient state, the mean and variance of longevity, the net reproductive rate R_0, and the cohort generation time. It presents the matrix calculus methods needed to calculate the sensitivity and elasticity of all these indices to any life history parameters.

Chapter 5 is modified from Caswell, H. 2009. Stage, age, and individual stochasticity in demography. The Per Brinck Oikos Award Lecture 2008. Oikos 118:1763–1782. ©Hal Caswell

The Markov chain approach is then generalized to variable environments (deterministic environmental sequences, periodic environments, iid random environments, Markovian environments). Variable environments are analyzed using the vec-permutation method to create a model that classifies individuals jointly by the stage and environmental condition. Throughout, examples are presented using the North Atlantic right whale (*Eubaleana glacialis*) and an endangered prairie plant (*Lomatium bradshawii*) in a stochastic fire environment.

5.1.1 Age and Stage, Implicit and Explicit

The essence of demography is the connection between the fates of individual organisms and the dynamics of populations. There exist diverse mathematical frameworks in which this connection can be studied (Keyfitz 1967; Metz and Diekmann 1986; Nisbet and Gurney 1982; Caswell 1989; Tuljapurkar and Caswell 1997; Caswell et al. 1997; DeAngelis and Gross 1992; Ellner et al. 2016). Regardless of the type of equations used, demographic analysis must account for differences among individuals, and the ways in which those differences affect the vital rates.

Among the many ways that individuals may differ, age has long had a kind of conceptual priority. Age is universal in the sense that every organism becomes one minute older with the passage of one minute of time. Age is also often associated with predictable changes in the vital rates. However, in some organisms characteristics other than age provide more and better information about an individual. Ecologists recognized this long ago, and have developed demographic theory based on size, maturity, physiological condition, instar, spatial location, etc.—referred to in general as "stage-classified" demography. Human demographers, who were responsible for the classical age-classified theory, by no means deny the importance of other properties, such as employment, parity, or health status; see Land and Rogers (1982), Goldman (1994), Robine et al. (2003), and Willekens (2014) for a sample of the kinds of issues that arise.

Even when the demographic model is entirely stage-classified, however, age is still implicitly present. Individuals in a given stage may differ in age, and individuals of a given age may be found in many different stages, but each individual still becomes one unit of age older with the passage of each unit of time. Extracting this implicit age-dependent information makes it possible to calculate interesting age-specific properties, such as survivorship, longevity, life expectancy, generation time, and net reproductive rate (Cochran and Ellner 1992; Caswell 2001, 2006; Tuljapurkar and Horvitz 2006; Horvitz and Tuljapurkar 2008).[1]

---

[1] *Explicit* age and stage dependence is explored in Chap. 6; see also Caswell and Salguero-Gómez (2013) and Caswell et al. (2018).

In this chapter, I show how to calculate some of these implicit age-specific properties from any stage-classified model. The trick is to formulate the life cycle as a Markov chain, and to generalize the "life" cycle to include death as a stage. Because death is permanent, it is called an *absorbing state*, and the theory of absorbing Markov chains provides the starting point for our analysis (Feichtinger 1971; Caswell 2001).

A Markov chain is a stochastic model for the movement of a particle among a set of states (e.g., Kemeny and Snell 1976; Iosifescu 1980). The probability distribution of the next state of the particle may depend on the current state, but not on earlier states. In our context, a "particle" is an individual organism. The states correspond to the stages of the life cycle, plus death (or perhaps multiple types of death, for example deaths due to different causes). This structure is ideally suited to asking questions about individual stochasticity, because it accounts for all the possible pathways, and their probabilities, that an individual can follow through its life. I will focus on discrete-time models, but much of the theory can no doubt be generalized to continuous-time models.

The use of Markov chains in demographic analysis is not new. As far as I know, Feichtinger (1971, 1973) was the first to use discrete-time absorbing Markov chains in demography, paying particular attention to competing risks and multiple causes of death. At around the same time, Hoem (1969) applied continuous-time Markov chains in the analysis of insurance systems (with states such as "active," "disabled," and "dead"). Later, Cochran and Ellner (1992) independently proposed the use of Markov chains to generate age-classified statistics from stage-classified models, but minimized the use of matrix notation in their presentation. Influenced by Feichtinger's work, and relying heavily on Iosifescu's (1980) treatment of absorbing Markov chains, I extended the calculations using matrix notation (Caswell 2001; Keyfitz and Caswell 2005), introduced sensitivity analysis (Caswell 2006), and presented results for both time-invariant and time-varying models. At the same time, Tuljapurkar and Horvitz (2006) and Horvitz and Tuljapurkar (2008) developed the same approaches and presented a more extensive investigation of time variation.

5.1.2 Individual Stochasticity and Heterogeneity

Consider a newborn individual. As it develops through the stages of its life cycle, it may grow, shrink, mature, move, reproduce, and allocate resources among its biological processes. At each moment, it is exposed to various mortality risks. At each moment, it has some chance of reproducing. Because these processes are stochastic, the lives of any two individuals may differ. These random outcomes—this *individual stochasticity*—imply that the age-specific properties of an individual (say, longevity) are random variables—there is a distribution among individuals that should be characterized by its mean, moments, etc. (Caswell 2009).

It is critical to notice that the calculation of these moments explicitly assumes that every individual in a given stage experiences exactly the same rates and hazards. There is no heterogeneity among the individuals (or, at least, none that matters demographically), even though there is variation in their lifetime properties. Empirical studies of longevity or lifetime reproductive output find that the variation among individuals is usually large, but it is a mistake to jump to the conclusion that it is due to heterogeneity among individuals without first examining the variance that is inevitably created by individual stochasticity (e.g., Tuljapurkar et al. 2009; Steiner and Tuljapurkar 2012; Caswell 2011; Caswell and Kluge 2015; Caswell and Vindenes 2018; Hartemink et al. 2017; Hartemink and Caswell 2018; van Daalen and Caswell 2017).

5.1.3   Examples

The calculations will be demonstrated by means of two case studies. The first is a stage-classified model for the North Atlantic right whale (*Eubaleana glacialis*). Later, in Sect. 5.5.4, a stochastic matrix model for the threatened prairie plant *Lomatium bradshawii* will appear as part of a study of variable environments.

The North Atlantic right whale is a large, highly endangered baleen whale (Kraus and Rolland 2007). Once abundant in the north Atlantic, it was decimated by whaling, beginning as much as a thousand years ago (Reeves et al. 2007). By 1900 the eastern North Atlantic stock had been effectively eliminated, and the western North Atlantic stock hunted to near extinction. The population has recovered only slowly since receiving at least nominal protection in 1935, and now numbers only about 300 individuals. Right whales migrate along the Atlantic coast of North America, from summer feeding grounds in the Gulf of Maine and Bay of Fundy to winter calving grounds off the Southeastern U.S. They are killed by ship collisions and entanglement in fishing gear (Kraus et al. 2005), and may also be affected by pollution of coastal waters.

Individual right whales are photographically identifiable by scars and callosity patterns. Since 1980, the New England Aquarium has surveyed the population, accumulating a database of over 10,000 sightings (Crone and Kraus 1990). Treating the first year of identification of an individual as marking, and each year of resighting as a recapture, permits the use of mark-recapture statistics to estimate demographic parameters of this endangered population (Caswell et al. 1999; Fujiwara and Caswell 2001, 2002; Caswell and Fujiwara 2004).

Figure 5.1 shows a life cycle graph used by Caswell and Fujiwara (2004) as the basis of a stage-structured matrix population model for the right whale. The stages are calves, immature females, mature but non-reproductive females, mothers, and "resting" mothers (because of the long period of parental care and gestation, right whales do not reproduce in the year after giving birth). This life cycle is typical of large, long-lived monovular mammals and birds.
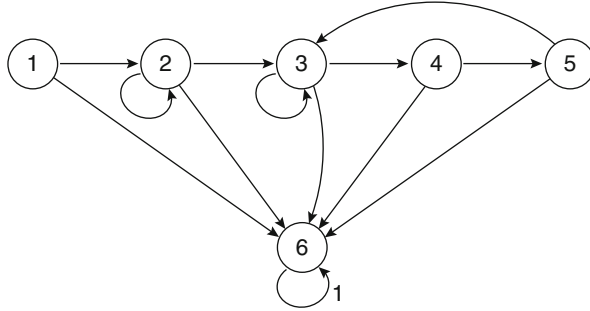
Fig. 5.1 Absorbing Markov chain transition graph for females of the North Atlantic right whale (*Eubalaena glacialis*). Projection interval is 1 year. Stages: 1 = calf, 2 = immature, 3 = mature, 4 = mother, 5 = post-breeding female, 6 = death. See Caswell and Fujiwara (2004) for explanation and parameter estimates

The model is parameterized in terms of survival probabilities $\sigma_1, \ldots, \sigma_5$, the probability of maturation γ_2, and the birth probability γ_3. The projection matrix is

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & F & 0 & 0 \\ \sigma_1 & \sigma_2(1 - \gamma_2) & 0 & 0 & 0 \\ 0 & \sigma_2\gamma_2 & \sigma_3(1 - \gamma_3) & 0 & \sigma_5 \\ 0 & 0 & \sigma_3\gamma_3 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4 & 0 \end{pmatrix} \tag{5.1}$$

The fertility term in the (1, 3) position is $F = 0.5\sigma_3\gamma_3\sqrt{\sigma_4}$, accounting for the sex ratio, the survival of mature females, their probability of giving birth if they survive, and the effect of survival of the mother on survival of the calf. For reasons related to parameter estimation, σ_5 is constrained to equal σ_3.

5.2 Markov Chains

The familiar life cycle graph (e.g., Fig. 5.1) corresponds to a projection matrix $\mathbf{A}$, in which a_{ij} gives the per-capita production of stage i individuals at $t + 1$ by a stage j individual at t. This production may occur by the transition of an individual from stage j to stage i, or by the production of one or more new individuals (by reproduction, fragmentation, etc.). So, we partition $\mathbf{A}$ into a matrix $\mathbf{U}$ describing transition probabilities of extant individuals and a matrix $\mathbf{F}$ describing the production of new individuals

$$\mathbf{A} = \mathbf{U} + \mathbf{F} \tag{5.2}$$

The column sums of $\mathbf{U}$ are all less than or equal to 1. Because individuals eventually die and pass out of the stages contained in $\mathbf{U}$, those stages are called transient states.

5.2.1   An Absorbing Markov Chain

If we include death explicitly (Fig. 5.1) and remove the arcs representing reproduction, we obtain the graph corresponding to the transition matrix for an absorbing Markov chain

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{U} & 0 \\ \hline \mathbf{m} & 1 \end{array} \right) \tag{5.3}$$

The element m_j of the vector $\mathbf{m}$ is the probability of mortality of an individual in stage j. Death is an absorbing state. I will assume that at least one absorbing state is accessible from any transient state in $\mathbf{U}$, and that the spectral radius of $\mathbf{U}$ is strictly less than 1. This guarantees that, with probability 1, every individual ends up in the absorbing state.

The right whale Fujiwara estimated $\mathbf{U}$ by applying multi-stage mark-recapture methods to the photographic identification catalog. Although the best model, out of a large number evaluated, included significant time variation in survival and birth rates, here I will analyze a single matrix obtained from a time-invariant model. The complete transient matrix $\mathbf{U}$ and the fertility matrix $\mathbf{F}$ are

$$\mathbf{U} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.90 & 0.85 & 0 & 0 & 0 \\ 0 & 0.12 & 0.71 & 0 & 1.00 \\ 0 & 0 & 0.29 & 0 & 0 \\ 0 & 0 & 0 & 0.85 & 0 \end{pmatrix} \tag{5.4}$$

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0.13 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.5}$$

5.2.2   Occupancy Times and the Fundamental Matrix

As the syllogism asserts, all men are mortal; absorbtion is certain. Our question is, how long does absorbtion take and what happens en route? From a demographic perspective, this is asking about the lifespan of an individual and the events that happen during that lifetime. The key to these questions is the *fundamental matrix* of the absorbing Markov chain. Consider an individual presently in transient state j. As time passes, it will visit other transient states, repeating some, skipping others,

until it eventually dies. Let v_{ij} denote the number of visits to, or the occupancy time in, transient state i that our individual, starting in transient state j, makes before being absorbed. The v_{ij} are random variables, reflecting individual stochasticity.

The entries of the matrix $\mathbf{U}$ give the probabilities of visiting each of the transient states after one time step. The entries of $\mathbf{U}^2$ give the probabilities of visiting each of the transient states after two time steps. Adding the powers of $\mathbf{U}$ gives the expected number of visits to each transient state, over a lifetime, in a matrix $\mathbf{N}$; i.e.,

$$\mathbf{N} = \left(E(v_{ij}) \right)$$
$$= \sum_{t=0}^{\infty} \mathbf{U}^t$$
$$= (\mathbf{I} - \mathbf{U})^{-1} . \tag{5.6}$$

The right whale The fundamental matrix for the right whale is calculated from (5.6) to be

$$\mathbf{N} = \begin{pmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 5.88 & 6.52 & 0.00 & 0.00 & 0.00 \\ 16.35 & 18.11 & 22.94 & 19.49 & 22.94 \\ \mathbf{4.74} & 5.25 & 6.65 & 6.65 & 6.65 \\ 4.02 & 4.46 & 5.65 & 5.65 & 6.65 \end{pmatrix} . \tag{5.7}$$

The first column corresponds to calves. On average, a calf will spend 1 year as a calf, 5.9 years as a juvenile, 16.3 years as a mature but non-breeding female, etc. Row 4 of $\mathbf{N}$ is of particular interest. Stage 4 represents mothers, so n_{4j} is the expected number of reproductive events that a female in stage j will experience during her remaining lifetime. Based on this model, a newborn calf could expect to give birth $n_{41} = 4.74$ times. A mature female could expect to give birth $n_{43} = 6.65$ times; the difference reflects the likelihood of mortality between birth and maturity.[2]

We would like to know how the entries of $\mathbf{N}$ vary in response to changes in the vital rates. To accomplish this, we need matrix calculus, which is the topic of the next section.

---

[2]Note that $n_{43} = n_{44} = n_{45} = n_{55}$ and $n_{53} = n_{54}$. This seems to be due to the fact, specific to these data, that the survival probability of stages 3 and 5 is indistinguishable from 1.0, and influences the results below.

5.2.3 Sensitivity of the Fundamental Matrix

Let us apply matrix calculus to find the sensitivity of the fundamental matrix $\mathbf{N}$ (Caswell 2006). This result will appear in the sensitivity analysis of most other demographic quantities. Let $\boldsymbol{\theta}$ be a vector of parameters (of dimension $p \times 1$) on which the entries of the transition matrix $\mathbf{U}$ depend. The fundamental matrix satisfies

$$\mathbf{I} = \mathbf{N}\mathbf{N}^{-1}. \tag{5.8}$$

Differentiating both sides gives

$$\mathbf{0} = (d\mathbf{N})\mathbf{N}^{-1} + \mathbf{N}\left(d\mathbf{N}^{-1}\right). \tag{5.9}$$

Applying the vec operator and Roth's theorem to both sides gives

$$\text{vec } \mathbf{0} = \left[\left(\mathbf{N}^{-1}\right)^{\mathsf{T}} \otimes \mathbf{I}_s\right] d\text{vec } \mathbf{N} + (\mathbf{I}_s \otimes \mathbf{N}) \, d\text{vec } \mathbf{N}^{-1} \tag{5.10}$$

Solving for $d\text{vec } \mathbf{N}$ gives

$$d\text{vec } \mathbf{N} = \left[\left(\mathbf{N}^{-1}\right)^{\mathsf{T}} \otimes \mathbf{I}_s\right]^{-1} (\mathbf{I}_s \otimes \mathbf{N}) \, d\text{vec } \mathbf{U} \tag{5.11}$$

To simplify this, it helps to know two facts about the Kronecker product:

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \tag{5.12}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A}\mathbf{C} \otimes \mathbf{B}\mathbf{D}) \tag{5.13}$$

provided that the sizes of the matrices permit the indicated operations. Thus $d\text{vec } \mathbf{N}$ in (5.11) simplifies to

$$d\text{vec } \mathbf{N} = \left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) d\text{vec } \mathbf{U} \tag{5.14}$$

The identification theorem (2.47) implies

$$\frac{d\text{vec } \mathbf{N}}{d\text{vec }^{\mathsf{T}}\mathbf{U}} = \mathbf{N}^{\mathsf{T}} \otimes \mathbf{N} \tag{5.15}$$

and the chain rule permits us to write

$$\frac{d\text{vec } \mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) \frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{5.16}$$
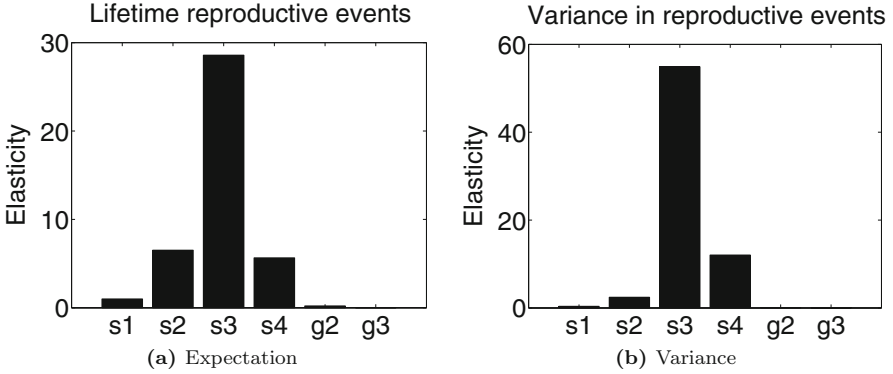
Fig. 5.2 Elasticity, to each of the vital rates, of reproductive outcomes in the right whale. (**a**) The elasticity of the expected lifetime number of reproductive events ($E(\nu_{41})$). (**b**) The elasticity of the variance in the lifetime number of reproductive events, $V(\nu_{41})$. Vital rates: s_1–s_4 are survival probabilities ($s_5 = s_3$ by assumption in this model); g_2 is the probability of maturation, and g_3 is the probability of reproduction

The left-hand side of (5.16) is a matrix, of dimension $s^2 \times p$, containing the sensitivity of every entry of $\mathbf{N}$ to every parameter in $\boldsymbol{\theta}$. The matrix $d\,\mathrm{vec}\,\mathbf{U}/d\boldsymbol{\theta}^{\mathsf{T}}$ is an $s^2 \times p$ matrix containing the sensitivities of all the elements of $\mathbf{U}$ to all the elements of $\boldsymbol{\theta}$. From (2.55), the elasticity of the fundamental matrix is given by

$$\frac{\epsilon\,\mathrm{vec}\,\mathbf{N}}{\epsilon\boldsymbol{\theta}^{\mathsf{T}}} = \mathcal{D}\,(\mathrm{vec}\,\mathbf{N})^{-1}\,\frac{d\,\mathrm{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}}\,\mathcal{D}\,(\boldsymbol{\theta}) \qquad (5.17)$$

The right whale As an example, we use (5.16) and (5.17) to calculate the elasticity of the expected lifetime number of reproductive events, $E(\nu_{41}) = n_{41}$, with respect to the survival probabilities $\sigma_1, \ldots, \sigma_4$, the maturation probability γ_2, and the breeding probability γ_3. Figure 5.2 shows that the number of breeding events is most elastic to mature female survival (σ_3), and less so to the survival of mature females or mothers (σ_2 and σ_4). Changes in the probability of giving birth, γ_3, have, remarkably enough, no impact on the expected number of reproductive events.

The elasticity of n_{41} to σ_3 (survival of mature females) is approximately 30. This implies that a 1% increase in σ_3 would produce about a 30% increase in the expected number of reproductive events.

5.3 From Stage to Age

The fundamental matrix summarizes the age-specific information implicit in the transient matrix $\mathbf{U}$, even if the model is stage-classified and age does not appear explicitly. We now extend this, to explore a series of age-specific demographic

indices and their sensitivity analyses. Some are well known (R_0, generation time), others little explored (variance in longevity, for example). They can, however, all be easily calculated from any stage-classified model.

5.3.1 Variance in Occupancy Time

The occupancy time in any transient state is a random variable; the fundamental matrix $\mathbf{N}$ gives its mean. Some individuals will visit that state more often, some less often, some not at all. This basic property of individual stochasticity can be described by the variance of v_{ij}. Iosifescu ([1980]), Theorem 3.1 gives a formula for all the moments of the v_{ij}; from this we can calculate the matrix of variances

$$\mathbf{V} = \left(V(v_{ij}) \right) = \left(2\mathbf{N}_{\mathrm{dg}} - \mathbf{I} \right) \mathbf{N} - \mathbf{N} \circ \mathbf{N} \tag{5.18}$$

(Caswell [2006]) where $\circ$ denotes the Hadamard, or element-by-element, product and $\mathbf{N}_{\mathrm{dg}}$ is a matrix with the diagonal elements of $\mathbf{N}$ on its diagonal and zeros elsewhere. The standard deviations of the occuancy times are the square roots of the elements of $\mathbf{V}$.

The right whale For the right whale, the matrix of variances calculated from (5.18) is

$$\mathbf{V} = \begin{pmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 36.18 & 35.95 & 0.00 & 0.00 & 0.00 \\ 466.44 & 484.80 & 503.32 & 494.86 & 503.32 \\ 35.80 & 36.98 & 37.54 & 37.54 & 37.54 \\ 33.28 & 34.94 & 37.54 & 37.54 & 37.54 \end{pmatrix}, \tag{5.19}$$

and the corresponding standard deviations are

$$\left(SD(v_{ij}) \right) = \begin{pmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 6.02 & 6.00 & 0.00 & 0.00 & 0.00 \\ 21.60 & 22.02 & 22.43 & 22.25 & 22.43 \\ 5.98 & 6.08 & 6.13 & 6.13 & 6.13 \\ 5.77 & 5.91 & 6.13 & 6.13 & 6.13 \end{pmatrix}. \tag{5.20}$$

The variance in the v_{ij} is the result of luck, not heterogeneity. That is, it is the variance among a group of individuals all experiencing exactly the same stage-specific transition and mortality probabilities in $\mathbf{U}$. As such, it can provide a null model for studies of heterogeneity in quantities such as the number of reproductive events. This idea has been explored independently, and in more detail, by Tuljapurkar and colleagues (Tuljapurkar et al. [2009]; Steiner and Tuljapurkar [2012]).

The sensitivity of the variance is derived in Appendix A.1 as

$$\frac{d \operatorname{vec} \mathbf{V}}{d \boldsymbol{\theta}^{\mathsf{T}}} = \left[2 \left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{I}_s \right) \mathcal{D} \left(\operatorname{vec} \mathbf{I}_s \right) + 2 \left(\mathbf{I}_s \otimes \mathbf{N}_{\mathrm{dg}} \right) \right.$$

$$\left. - \mathbf{I}_{s^2} - 2 \mathcal{D} \left(\operatorname{vec} \mathbf{N} \right) \right] \frac{d \operatorname{vec} \mathbf{N}}{d \boldsymbol{\theta}^{\mathsf{T}}} \tag{5.21}$$

Elasticities of $\mathbf{V}$ are calculated using (2.55).

Hint Before looking at Appendix A.1, to derive (5.21), write $\mathbf{N}_{\mathrm{dg}} = \mathbf{I} \circ \mathbf{N}$, differentiate (5.18), and use the fact that $\operatorname{vec}(\mathbf{A} \circ \mathbf{B}) = \mathcal{D}(\operatorname{vec} \mathbf{A}) \operatorname{vec} \mathbf{B} = \mathcal{D}(\operatorname{vec} \mathbf{B}) \operatorname{vec} \mathbf{A}$.

The right whale The elasticities of $V(\nu_{41})$, calculated from (5.21) and (5.17), are shown in Fig. 5.2b. They are roughly proportional to the elasticities of $E(\nu_{41})$; that is, the vital rates that have large effects on the expected number of reproductive events also have large effects on the variance.

5.3.2   Longevity and Life Expectancy

Longevity is an important demographic characteristic (Carey 2003). Mean longevity, or life expectancy, it is one of the most widely reported demographic statistics, used to compare populations, species, countries, regions, historical periods, etc., and to examine the effects of evolutionary, management, medical, and social processes. The longevity of an individual is the sum of the time spent in all of the transient states before final absorption. Let the random variable η_j denote the longevity of an individual currently in stage j. Then

$$\eta_j = \sum_i \nu_{ij}. \tag{5.22}$$

A vector $E(\boldsymbol{\eta})$ of expected longevities, or life expectancies, is obtained by summing the columns of $\mathbf{N}$:

$$E(\boldsymbol{\eta}^{\mathsf{T}}) = \mathbf{1}^{\mathsf{T}} \mathbf{N} \tag{5.23}$$

where $\mathbf{1}$ is a vector of ones. Often, life expectancy at birth is of primary interest. If stages are numbered so that birth corresponds to stage 1, then life expectancy at birth is

$$E(\eta_1) = \mathbf{1}^{\mathsf{T}} \mathbf{N} \mathbf{e}_1 \tag{5.24}$$

where $\mathbf{e}_1$ is a vector with 1 in the first entry and zeros elsewhere.

The sensitivity of life expectancy in age-classified models has been studied by Pollard (1982) and Keyfitz (1971); see Keyfitz and Caswell (2005, Section 4.3), Vaupel (1986), and Vaupel and Canudas Romo (2003).

For more general stage-classified models, the sensitivity of $E(\boldsymbol{\eta})$ is (Caswell 2006)

$$\frac{dE(\boldsymbol{\eta})}{d\boldsymbol{\theta}^\mathsf{T}} = \left(\mathbf{I}_s \otimes \mathbf{1}^\mathsf{T}\right)\left(\mathbf{N}^\mathsf{T} \otimes \mathbf{N}\right)\frac{d\operatorname{vec}\mathbf{U}}{d\boldsymbol{\theta}^\mathsf{T}} \tag{5.25}$$

Hint To obtain (5.25), differentiate both sides of (5.23), apply the vec operator, and use (5.16) for the derivative of $\mathbf{N}$. See Appendix A.2 for the derivation.

The right whale For the right whale, the vector of life expectancies is

$$E(\boldsymbol{\eta}^\mathsf{T}) = \begin{pmatrix} 32.0 & 34.3 & 35.2 & 31.8 & 36.2 \end{pmatrix} \tag{5.26}$$

Because mortality rates vary relatively little among stages, the life expectancies of the stages differ by only about 15%. Thus life expectancy for a calf implied by these data was 32 years. The elasticities of life expectancy to the vital rates are shown in Fig. 5.3. Life expectancy is most elastic to mature female survival σ_3, and less so to σ_2 and σ_3. This partly reflects the longer amount of time spent as a mature female, compared to an immature female or mother; see (5.7). The elasticity to the birth rate γ_3 is negative, because of the reduced survival of mothers. A 1% increase in γ_3 will lead to a 0.51% *decrease* in life expectancy. This is one possible measure of the cost of reproduction.
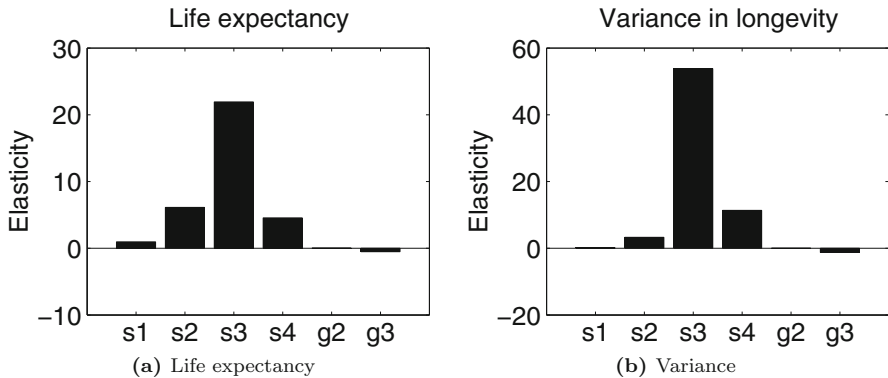

Fig. 5.3 Elasticities of longevity for the right whale. (**a**) The elasticity, to each of the vital rates, of life expectancy for a female right whale calf. (**b**) The elasticity of the variance in longevity for a female right whale calf. Parameters as in Fig. 5.2

5.3.3 Variance in Longevity

Like the occupancy time in a transient state, longevity is a random variable, the variability of which is a measure of individual stochasticity. Individuals differ in longevity depending on the pathways taken from birth to death. This variance has been explored by human demographers, using life table methods, as one way of studying the inequality in life span generated by a given mortality schedule, and how that inequality has changed over time (e.g., Wilmoth and Horiuchi 1999; Shkolnikov et al. 2003; Edwards and Tuljapurkar 2005; Van Raalte and Caswell 2013).

The variance of the time to absorbtion is

$$V(\boldsymbol{\eta}^\mathsf{T}) = \mathbf{1}^\mathsf{T} \mathbf{N} (2\mathbf{N} - \mathbf{I}) - E\left(\boldsymbol{\eta}^\mathsf{T}\right) \circ E\left(\boldsymbol{\eta}^\mathsf{T}\right). \tag{5.27}$$

(Caswell 2006; Iosifescu 1980).

The sensitivity of the variance in longevity is

$$\frac{dV(\boldsymbol{\eta})}{d\boldsymbol{\theta}^\mathsf{T}} = \left[2\left(\mathbf{N}^\mathsf{T} \otimes \mathbf{1}^\mathsf{T}\right) + 2\left(\mathbf{I}_s \otimes \mathbf{1}^\mathsf{T}\mathbf{N}\right) \right.$$

$$\left. - \left(\mathbf{I}_s \otimes \mathbf{1}^\mathsf{T}\right) - 2\mathcal{D}\left(E(\boldsymbol{\eta})\right)\left(\mathbf{I}_s \otimes \mathbf{1}^\mathsf{T}\right) \right] \left(\mathbf{N}^\mathsf{T} \otimes \mathbf{N}\right) \frac{d\mathbf{U}}{d\boldsymbol{\theta}^\mathsf{T}} \tag{5.28}$$

The first entry of (5.28) is the sensitivity of the variance in longevity starting in stage 1.

Hint To derive (5.28), differentiate (5.27) and apply the vec operator and Roth's theorem to each term, using (5.25) for the derivative of $E(\boldsymbol{\eta})$. See Sect. A.3 for details.

The right whale For the right whale, the variance and standard deviation of longevity are given by

$$V(\boldsymbol{\eta})^\mathsf{T} = \left(1157 \ 1167 \ 1172 \ 1163 \ 1172\right) \tag{5.29}$$

$$SD(\boldsymbol{\eta})^\mathsf{T} = \left(34.0 \ 34.2 \ 34.2 \ 34.1 \ 34.2\right) \tag{5.30}$$

The life expectancy at birth of 32 years has a standard deviation of about 34 years. Note that this result implies a very long positive tail of longevity. The interpretation of this result is tricky; I will return to it in Sect. 5.7.

The elasticities of the variance of longevity of a calf are shown in Fig. 5.3b. The variance in longevity is increased by increases in σ_3, less so by increases in σ_2 and σ_4. The pattern of the elasticities is strikingly similar to that of the elasticities of $E(\boldsymbol{\eta})$.

5.3.4 Cohort Generation Time

Generation time measures the typical age at which offspring are produced, or the age at which the typical offspring is produced. It appears in the IUCN criteria for classifying threatened species (IUCN Species Survival Commission 2001) as well as in various evolutionary considerations. There are several definitions of generation time (Coale 1972); here we will examine the cohort generation time, defined as the mean age of production of offspring in a cohort of newborn individuals. From the definition it is clear why calculation of generation time is a problem in stage-classified models, in which the age of parents does not appear. Moreover, in stage-classified models, individuals may be born into several stages (e.g., cleisthogamous vs. chasmogamous seeds; LeCorff and Horvitz 2005), each with a different subsequent pattern of development, survival, and fertility. There could be a different generation time for each type of offspring, and if individuals may produce more than one type of offspring, the average age at which they are produced could differ from one kind of offspring to another.

Thus, we expect to have a generation time that measures the mean age of production of offspring of type i by an individual born in stage j. Write this as a vector $\boldsymbol{\mu}^{(j)}$. Then it can be shown (Sect. A.5) that

$$\boldsymbol{\mu}^{(j)} = \mathcal{D} \left(\mathbf{FNe}_j \right)^{-1} \mathbf{FNUNe}_j \tag{5.31}$$

The sensitivity of $\boldsymbol{\mu}^{(j)}$ is obtained by a methodical application of matrix calculus to (5.31). To simplify notation, define
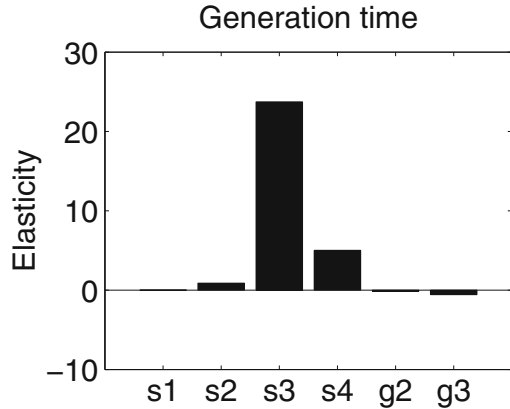
$$\mathbf{X} = \mathcal{D} \left(\mathbf{FNe}_j \right) \tag{5.32}$$

$$\mathbf{r} = \mathbf{FNUNe}_j \tag{5.33}$$

The resulting sensitivity of $\boldsymbol{\mu}^{(j)}$ is

$$\begin{aligned}
\frac{d\boldsymbol{\mu}^{(j)}}{d\boldsymbol{\theta}^\mathsf{T}} = & -\left(\mathbf{r}^\mathsf{T} \otimes \mathbf{I} \right) \left(\mathbf{X}^{-1} \otimes \mathbf{X}^{-1} \right) \mathcal{D} \left(\mathrm{vec} \, \mathbf{I} \right) \\
& \times \left[\left(\mathbf{1e}_j^\mathsf{T} \mathbf{N}^\mathsf{T} \otimes \mathbf{I} \right) \frac{d\mathrm{vec}\,\mathbf{F}}{d\boldsymbol{\theta}^\mathsf{T}} + \left(\mathbf{1e}_j \otimes \mathbf{F} \right) \frac{d\mathrm{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^\mathsf{T}} \right] \\
& + \left\{ \left[\left(\mathbf{NUNe}_j \right)^\mathsf{T} \otimes \mathbf{I} \right] \frac{d\mathrm{vec}\,\mathbf{F}}{d\boldsymbol{\theta}^\mathsf{T}} + \left[\left(\mathbf{UNe}_j \right)^\mathsf{T} \otimes \mathbf{F} \right] \frac{d\mathrm{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^\mathsf{T}} \right. \\
& + \left. \left[\left(\mathbf{Ne}_j \right)^\mathsf{T} \otimes \mathbf{FN} \right] \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^\mathsf{T}} + \left[\mathbf{e}_j^\mathsf{T} \otimes \mathbf{FNU} \right] \frac{d\mathrm{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^\mathsf{T}} \right\} \tag{5.34}
\end{aligned}$$

Fig. 5.4 The elasticity, to
each of the vital rates, of the
cohort generation time for a
newborn calf right whale.
Parameters as in Fig. 5.2

Hint To derive (5.34), it helps to note that, for any vector $\mathbf{z}$, one can write $\mathcal{D}(\mathbf{z}) = \mathbf{I} \circ \mathbf{z1}^\mathsf{T}$. Apply this to $\mathbf{X}$, differentiate all the terms in $\mu^{(j)}$, and apply the vec operator. With any luck, you will come out to this answer. See Sect. A.5.1 for derivation.

The right whale The elasticities of the generation time $\boldsymbol{\mu}^{(1)}$ of a calf are shown in Fig. 5.4. Changes in early survival (σ_1 and σ_2) have little effect. Adult survival σ_3 and, to a lesser extent, σ_4 increase the generation time by extending the reproductive lifespan. The maturation probability γ_2 and the birth probability γ_3 have negative effects on generation time, because they speed up reproduction.

5.4 The Net Reproductive Rate

In age-classified demography, the net reproductive rate R_0 measures lifetime reproductive output. It also appears in epidemiology, where it measures the potential of a disease to spread (e.g., Diekmann et al. 1990; van den Driessche and Watmough 2002). The classical net reproductive rate satisfies three conditions:

C_1: R_0 measures the expected lifetime production of offspring.
C_2: R_0 measures the rate of increase per generation (in contrast to the rate of increase per unit of time, which is given by λ or r).
C_3: R_0 is an indicator function for population persistence. If $R_0 > 1$ then an individual will, on average, produce more than enough offspring to replace itself, the next generation will be larger than the present generation, and the population will grow. If $R_0 < 1$, each generation is smaller than the one before, and the population will decline to extinction.

In classical demography (Lotka 1939; Rhodes 1940),

$$R_0 = \int_0^\infty \ell(x)m(x)dx \qquad (5.35)$$

where $\ell(x)$ is survivorship to age x and $m(x)$ is the maternity function. It is not difficult to show that R_0 defined in this way satisfies conditions C_1, C_2, and C_3.

In stage-classified models, however, the calculation of R_0 must account for the multiple pathways that an individual may follow through the life cycle, and the production of multiple kinds of offspring along each of these pathways. Rogers (1974; see also Lebreton 1996) considered R_0 in the context of an age-classified population distributed across a set of spatial regions. However, these calculations assume that age-specific survival and fertility schedules are available for each region. A more general solution was provided by Cushing and Zhou (1994) for stage-classified populations with no age-specific information. Their analysis produces an index that satisfies as many as possible of the conditions C_1, C_2, and C_3. de Camino-Beck and Lewis (2007, 2008) have derived graph-theoretic ways to calculate R_0.

Consider an initial cohort at $t = 0$ with structure $\mathbf{x}_0$, and call this the first generation. This cohort will produce offspring according to $\mathbf{F}\mathbf{x}_0$. The survivors of the cohort at $t = 1$ will produce offspring according to $\mathbf{F}\mathbf{U}\mathbf{x}_0$. The survivors at $t = 2$ will produce offspring $\mathbf{F}\mathbf{U}^2\mathbf{x}_0$, and so on. The second generation is composed of all the offspring of the first generation, obtained by summing over the lifetime of the cohort

$$\mathbf{x}(1) = \left(\mathbf{F} \sum_{i=0}^{\infty} \mathbf{U}^i \right) \mathbf{x}_0$$

$$= (\mathbf{F}\mathbf{N})\,\mathbf{x}_0 \tag{5.36}$$

Iterating this process leads to a model for the growth from one generation to the next

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{N}\mathbf{x}(k) \tag{5.37}$$

Cushing and Zhou (1994) define R_0 as the per-generation growth rate, given by the dominant eigenvalue ρ of $\mathbf{F}\mathbf{N}$,

$$R_0 = \rho[\mathbf{F}\mathbf{N}] \tag{5.38}$$

Thus the Cushing-Zhou measure of R_0 clearly satisfies condition C_2. Cushing and Zhou (1994) also prove (their Theorem 3) that R_0 defined in this way is less than, equal to, or greater than 1 if and only if λ is less than, equal to, or greater than one, respectively, thus satisfying condition C_3.

The relation between lifetime offspring production and R_0 (condition C_1) is more complicated when the life cycle contains multiple types of offspring. If only a single type of offspring is produced (call it stage 1), then $\mathbf{F}$ will have nonzero entries only in its first row, and $\mathbf{F}\mathbf{N}$ will be upper triangular, with its dominant eigenvalue appearing in the $(1, 1)$ position. i.e., the sum of the fertilities of each stage weighted by the expected time spent in that stage. This is precisely the expected lifetime offspring production, so for the case of a single type of offspring, the Cushing-Zhou R_0 also satisfies C_1.

However, if the life cycle contains multiple types of offspring (say stages $1, \ldots, h$), the upper left $h \times h$ corner of $\mathbf{FN}$ will contain the expected lifetime production of offspring of types $1, \ldots, h$ by individuals starting life as types $1, \ldots, h$. Since such a life cycle contains more than one kind of expected lifetime production of offspring, R_0 cannot satisfy C_1 in the sense of being *the* expected lifetime reproduction. Instead, R_0 is calculated from all these expectations (as the dominant eigenvalue of this $h \times h$ submatrix). It determines per-generation growth and population persistence as a function of the expected lifetime production of all types of offspring in a way that satisfies C_2 and C_3.

The right whale The right whale produces only a single type of offspring. The fundamental matrix $\mathbf{N}$ is given by (5.7), the fertility matrix is given by (5.5), and the generation growth matrix is

$$\mathbf{FN} = \begin{pmatrix} 2.18 & 2.42 & 3.06 & 2.60 & 3.06 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.39}$$

The dominant eigenvalue of $\mathbf{FN}$ is its (1, 1) entry

$$R_0 = \sum_j f_{1j} E(\nu_{j1}) = 2.18 \tag{5.40}$$

It is interesting to compare $R_0 = 2.18$ with $E(\nu_{14}) = 4.74$. Only female offspring are counted in R_0, whereas $E(\nu_{14})$ counts reproductive events regardless of the sex of the offspring produced. Still, R_0 is less than half of $E(\nu_{14})$, because of the less than perfect survival of calves from t to $t + 1$.

5.4.1 Net Reproductive Rate in Periodic Environments

Periodic time-varying models (Caswell 2001, Chapter 13) are an interesting special case of the multiple offspring type problem. In a periodic model, apparently identical offspring (e.g., seeds) produced at different phases of the cycle (e.g., seasons) are, in effect, of different types of. To the extent that they face different environments, they will differ in their expected offspring production, and R_0 will differ depending on the phase of the cycle in which it is calculated.

The net reproductive rate in a periodic environment was calculated by Hunter and Caswell (2005a) in a study of the sooty shearwater, a pelagic seabird nesting on offshore islands in New Zealand. In that study, the year was divided into two short phases, during which breeding and harvest of chicks occur, and a longer phase encompassing the rest of the year. Let $\mathbf{B}_i = \mathbf{U}_i + \mathbf{F}_i$ be the projection matrix in phase i of the cycle. Without loss of generality, consider an environment with a period of 2

(e.g., winter and summer). The population is projected over a year, starting in phase 1, by

$$\mathbf{A}_1 = \mathbf{B}_2 \mathbf{B}_1 \tag{5.41}$$

which is decomposed as

$$\begin{aligned} \mathbf{A}_1 &= (\mathbf{U}_2 + \mathbf{F}_2)\,(\mathbf{U}_1 + \mathbf{F}_1) \\ &= \mathbf{U}_2\mathbf{U}_1 + \mathbf{U}_2\mathbf{F}_1 + \mathbf{F}_2\mathbf{U}_1 + \mathbf{F}_2\mathbf{F}_1 \end{aligned} \tag{5.42}$$

The first term includes only transitions, whereas the last three terms all describe some aspect of reproduction. Thus the annual matrix is $\mathbf{A}_1 = \widehat{\mathbf{U}} + \widehat{\mathbf{F}}$, where

$$\widehat{\mathbf{U}}_1 = \mathbf{U}_2\mathbf{U}_1 \tag{5.43}$$

$$\widehat{\mathbf{F}}_1 = \mathbf{U}_2\mathbf{F}_1 + \mathbf{F}_2\mathbf{U}_1 + \mathbf{F}_2\mathbf{F}_1 \tag{5.44}$$

and

$$R_0^{(1)} = \rho\left[\widehat{\mathbf{F}}_1\left(\mathbf{I} - \widehat{\mathbf{U}}_1\right)^{-1}\right] \tag{5.45}$$

where the superscript 1 indicates that this is the net reproductive rate of a generation beginning in season 1. The corresponding matrices for a generation starting in season 2 are obtained from

$$\mathbf{A}_2 = \mathbf{B}_1 \mathbf{B}_2 \tag{5.46}$$

and lead to a net reproductive rate $R_0^{(2)}$. It is easily verified that $R_0^{(1)} \neq R_0^{(2)}$ in general. This contrasts with the population growth rate λ, which is independent of cyclic permutation of the seasons. However, since λ is the same for $\mathbf{A}_1$ and $\mathbf{A}_2$, it must be the case that $R_0^{(1)}$ and $R_0^{(2)}$ are both greater than or less than 1 together.

An alternative formulation of R_0 in periodic environments was published at the same time as Caswell (2009), by Bacaër (2009). He wrote the model, using methods equivalent to those in Sect. 5.5 below, by jointly classifying individuals by stage and by their phase within a seasonal cycle. Let $\mathbf{A}_i = \mathbf{U}_i + \mathbf{F}_i$ be the projection matrix in season i. Then, for example with three seasons, the projection matrix would take the block-circulant form

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0 & 0 & \mathbf{A}_3 \\ \mathbf{A}_1 & 0 & 0 \\ 0 & \mathbf{A}_2 & 0 \end{pmatrix} \tag{5.47}$$

(with similar formulations for $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{F}}$). After some manipulations, Bacaër shows that $\mathbf{R}_0$ is the dominant eigenvalue of the matrix[3]

---

[3]It might be easier to apply the Cushing-Zhou theorem directly to $\tilde{\mathbf{A}}$ and write

$$R_0 = \rho\left(\tilde{\mathbf{F}}\left(\mathbf{I} - \tilde{\mathbf{U}}\right)^{-1}\right) \tag{5.48}$$

but Bacaër does not do this.

$$\begin{pmatrix} \mathbf{F}_1 & 0 & 0 \\ 0 & \mathbf{F}_2 & 0 \\ 0 & 0 & \mathbf{F}_3 \end{pmatrix} \begin{pmatrix} -\mathbf{U}_1 & \mathbf{I} & 0 \\ 0 & -\mathbf{U}_2 & \mathbf{I} \\ \mathbf{I} & 0 & -\mathbf{U}_3 \end{pmatrix}^{-1}. \tag{5.49}$$

Bacaër (2009) proves that R_0 calculated in this way satisfies condition C_3, providing an indicator for population growth ($R_0 > 1$) or decline ($R_0 < 1$). However, this definition of R_0 does not satisfy C_1 because it does not distinguish the different lifetime reproductive output of individuals born in different seasons.

Cushing and Ackleh (2012) returned to this issue. They argue that the standard approach for studying dynamics of periodic models is to study the "periodic composite map", which is the map for the entire cycle composed of the product of the phase-specific matrices, as in (5.41), which projects over the entire cycle, rather than from one season to the next. They separate transitions and reproduction as in Eqs. (5.43) and (5.44), and prove that R_0 calculated in this way satisfies C_1 (with a different lifetime reproductive output for each starting season) and C_3 (so that the values of R_0 in each season agree in their determination of positive or negative growth). Cushing and Ackleh (2012) also explore the net reproductive rate in nonlinear models, in which R_0 calculated at zero density determines whether the extinction equilibrium is stable.

In the end, it is valuable to have two different ways of calculating R_0, but it highlights the need to carefully specify which properties one wants the index to have.

5.4.2 Sensitivity of the Net Reproductive Rate

Since R_0 is obtained as an eigenvalue, its sensitivity to parameter changes is easy to derive. Let $\mathbf{x}$ and $\mathbf{y}$ be the right and left eigenvectors of $\mathbf{FN}$ corresponding to R_0. Then (Caswell 2006) the sensitivity of R_0 is
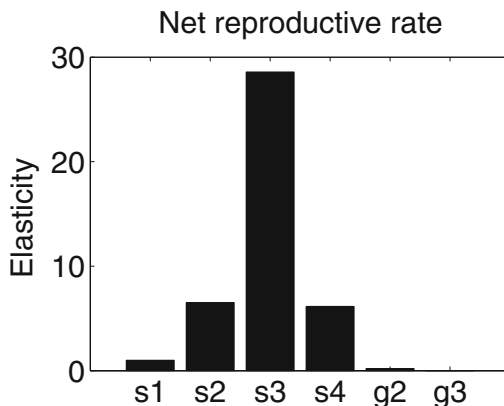
$$\frac{dR_0}{d\boldsymbol{\theta}^\mathsf{T}} = \left(\mathbf{y}^\mathsf{T}\mathbf{N}^\mathsf{T} \otimes \mathbf{x}^\mathsf{T}\right)\frac{d\,\mathrm{vec}\,\mathbf{F}}{d\boldsymbol{\theta}^\mathsf{T}} + \left(\mathbf{y}^\mathsf{T}\mathbf{N}^\mathsf{T} \otimes \mathbf{x}^\mathsf{T}\mathbf{FN}\right)\frac{d\,\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^\mathsf{T}} \tag{5.50}$$

The first term captures the effects of changing fertility, the second term captures effects of changes in survival and transitions. The derivation of (5.50) is given in Appendix A.4.

Hint To derive (5.50), write $R_0 = \rho[\mathbf{FN}]$ and write dR_0 in terms of the right and left eigenvectors of $\mathbf{FN}$ and the differential of $\mathbf{FN}$. Then expand $d(\mathbf{FN}) = (d\mathbf{F})\mathbf{N} + \mathbf{F}d(\mathbf{N})$ and apply the vec operator and the chain rule.

The right whale The elasticity of R_0 is shown in Fig. 5.5; R_0 is most elastic to σ_3, less so to σ_2 and σ_4. Remarkably, the elasticity of R_0 to the birth probability γ_3 is zero (actually, $\sim 10^{-9}$). This is a case where lifetime reproductive output is

Fig. 5.5 The elasticity, to
each of the vital rates, of the
net reproductive rate (R_0) for
the right whale. Parameters as
in Fig. 5.2

affected strongly by survival, slightly by maturation, but not at all by the probability
of breeding given survival. This seems to be a consequence of the lower survival
probability of mothers; an increase in γ_3 increases the probability of reproduction,
but reduces the lifetime over which that reproduction will be realized.

5.4.3 Invasion Exponents, Selection Gradients, and R_0

Selection on life history traits can be studied in terms of the invasion exponent,
which measures the rate at which a mutation, introduced at low densities, will
increase in the environment created by a resident phenotype (Metz et al. 1992;
Ferriére and Gatto 1993); for a recent introduction see Otto and Day (2007). The
selection gradient on a trait is the derivative of the invasion exponent with respect to
the value of the trait. If the derivative is positive, selection favors an increase in the
trait, and vice-versa. The invasion exponent in a density-independent model is given
by $\log \lambda$. In a density-dependent model, the invasion exponent is given by the growth
rate at equilibrium, $\lambda[\hat{\mathbf{n}}]$. The net reproductive rate R_0 is not, strictly speaking,
an invasion exponent, but because it measures expected lifetime reproduction, it is
attractive as a measure of fitness (see, e.g., the discussion in Kozlowski 1999). Using
R_0 as a measure of fitness will lead to erroneous conclusions unless the selection
gradients, measured in terms of λ and of R_0, give the same answers, i.e., unless
$dR_0/d\theta \propto d \log \lambda/d\theta$.

   For an age-classified model, we write R_0 in terms of the net maternity function
$\phi(x, \theta) = \ell(x, \theta)m(x, \theta)$ where both survival and reproduction depend on some
parameter θ. Then

$$R_0(\theta) = \int_0^\infty \phi(x, \theta)dx \tag{5.51}$$

The growth rate $r = \log \lambda$ is the solution to

$$1 = \int_0^\infty \phi(x, \theta) e^{-r(\theta)x} dx \qquad (5.52)$$

Differentiating (5.51) and (5.52) gives

$$\frac{dR_0}{d\theta} = \int_0^\infty \frac{d\phi(x, \theta)}{d\theta} dx \qquad (5.53)$$

$$\frac{dr}{d\theta} = \frac{\int_0^\infty e^{-rx} \frac{d\phi(x,\theta)}{d\theta} dx}{\int_0^\infty x\phi(x, \theta) e^{-rx} dx} \qquad (5.54)$$

Equation (5.54) is Hamilton's (1966) famous result; the denominator is the genera-
tion time measured as the average age of reproduction in the stable age distribution
(see Chap. 3).

   When $R_0 = 1$ and $r = 0$, it follows from (5.53) and (5.54) that the gradients
$dr/d\theta$ and $dR_0/d\theta$ are proportional. Use of either will lead to the same conclusions
about selection. But when $r \neq 0$, this is not the case. If $r > 0$, then $dr/d\theta$ is
reduced for traits that operate at later ages, because $d\phi/dx$ is weighted by e^{-rx}. It
is an open problem to generalize this result to stage-classified models, and prove
that

$$\frac{d \log \lambda}{d\theta^{\mathsf{T}}} \propto \frac{dR_0}{d\theta^{\mathsf{T}}} \qquad (5.55)$$

when $\lambda = R_0 = 1$. In a few cases I have examined, it appears to be true numerically.
As the following example shows, it is certainly the case that when $\lambda \neq 1$, the
derivatives are not generally proportional.

The right whale   The lack of proportionality between the selection gradients in
terms of λ and of R_0 means that evolutionary conclusions will differ depending
on which is used, especially when tradeoffs exist between two or more traits.
For example, for the right whale, $\lambda = 1.025$ and $R_0 = 2.183$. Figure 5.6
shows the sensitivity of λ and of R_0; while the patterns are similar, they are *not*
proportional, and the use of R_0 as an invasion exponent would result in erroneous
predictions. Suppose a trait existed that would increase the birth probability γ_3
at the cost of a reduction in calf survival σ_1, with the cost measured by $c =
-d\sigma_1/d\gamma_3$. An increase in this trait would be favored by selection provided
that

$$c < \frac{\partial \lambda / \partial \gamma_3}{\partial \lambda / \partial \sigma_1} = 0.96 \qquad (5.56)$$
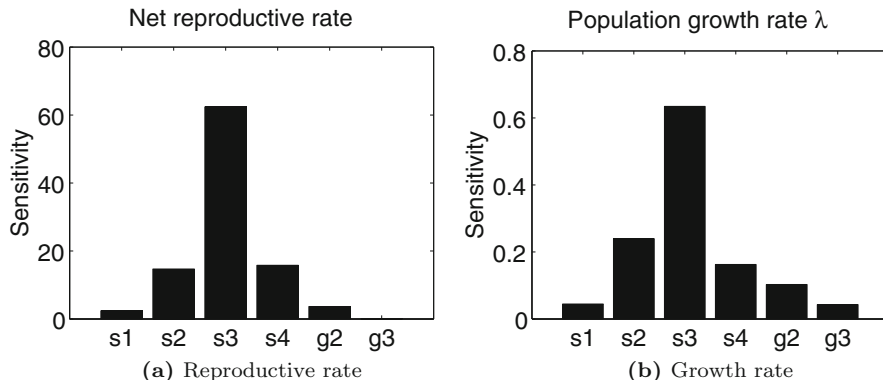
Fig. 5.6 (**a**) The sensitivity, to each of the vital rates, of the net reproductive rate R_0 for the right whale. (**b**) The sensitivity of population growth rate λ. The derivative of λ is the selection gradient; use of the derivative of R_0 leads to erroneous predictions unless the population is at equilibrium. Parameters as in Fig. 5.2

But if expected lifetime reproduction was used as an invasion exponent, the analysis would conclude that selection would favor an increase in the trait only if

$$c < \frac{\partial R_0 / \partial \gamma_3}{\partial R_0 / \partial \sigma_1} = 0.0 \qquad (5.57)$$

That is, according to R_0, any cost whatsoever of increased birth rate would prevent selection from favoring it. According to λ (and correctly, in this case), selection would favor increased birth rate provided that the cost was not too great. In spite of the superficial similarity of the patterns in Fig. 5.6, the evolutionary implications are quite different, reflecting the impact of *timing* of life history events on λ. The sensitivities of λ to σ_2 and γ_2, which influence early survival and the age at maturity, are larger than the sensitivities of R_0 to the same parameters.

5.4.4   Beyond R_0: Individual Stochasticity in Lifetime Reproduction

Variation among individuals is fundamental to population biology. As argued here, two sources of variation must be distinguished: heterogeneity and individual stochasticity Heterogeneity refers to genuine differences among individuals, because of which the individuals experience different vital rates. Individual stochasticity refers to the apparent differences that result from the random outcome of identical vital rates, applied to identical individuals. We have seen above that individual stochasticity is always present. That is particularly true of lifetime

reproductive output (LRO). The net reproductive rate is the *expectation* of LRO, but what can we say about the variance among individuals.

Empirical measurement shows that LRO is usually highly variable among individuals and positively skewed. Typically, a few individuals produce many offspring while most produce few, or none at all (Clutton-Brock 1988; Newton 1989). If this variance reflected heterogeneity among individual properties, and if the heterogeneity had a genetic basis, the variance would provide material for natural selection (the "opportunity for selection" of Crow 1958). Population and quantitative genetics are replete with methods to measure such genetic variation; e.g., Lande and Arnold (1983) and Endler (1986).

However, variance among individuals in LRO is not evidence of heterogeneity, genetic or otherwise; some is due to individual stochasticity. Only after evaluating the extent of individual stochasticity can data on LRO be interpreted as evidence for heterogeneity (Caswell 2011; Tuljapurkar et al. 2009; Steiner et al. 2010; Steiner and Tuljapurkar 2012). Caswell (2011) developed a method to calculate the mean, variance, and higher moments of lifetime reproductive output for any age- or stage-classified life cycle, using Markov chains with rewards; see van Daalen and Caswell (2015, 2017) for full details. In these models[4] the movement of the individual through its life cycle is described by an absorbing Markov chain; mortality appears as transitions to an absorbing (dead) state. At each step, the individual accumulates a "reward." In our context, the reward is the production of offspring. The reproductive reward is a random variable with a specified set of moments. The reward accumulated by the inevitable death of the individual is its LRO. Although every individual experiences the same vital rates—there is no heterogeneity—each individual may experience a different life and thus a different lifetime reproductive output.

Stage-specific reproductive output is specified by a set of reward matrices $\mathbf{R}_k$. The (i, j) element of $\mathbf{R}_k$ is the kth moment of the reproductive output associated with the transition from stage j to stage i. Given the reward matrices, the Markov chain transition matrix $\mathbf{P}$, and the reasonable assumption that the dead do not reproduce, all the moments of LRO can be calculated (van Daalen and Caswell 2017).

Let $\tilde{\boldsymbol{\rho}}_k$ be a vector containing the kth moments of LRO for individuals starting in each transient (living) stage. Then, it has been shown (van Daalen and Caswell 2017) that, e.g., the first two moments of LRO are

$$\tilde{\boldsymbol{\rho}}_1 = \mathbf{N}^\mathsf{T}\mathbf{Z}\left(\mathbf{P} \circ \mathbf{R}_1\right)^\mathsf{T}\mathbf{1}_{s+1} \tag{5.58}$$

$$\tilde{\boldsymbol{\rho}}_2 = \mathbf{N}^\mathsf{T}\left[\mathbf{Z}(\mathbf{P} \circ \mathbf{R}_2)^\mathsf{T}\mathbf{1}_{s+1} + 2(\mathbf{U} \circ \mathbf{R}_1)^\mathsf{T}\tilde{\boldsymbol{\rho}}_1\right] \tag{5.59}$$

---

[4]Markov chains with rewards have a long history in stochastic process theory; see Howard (1960), Puterman (1994), and Sheskin (2010).

where s is the number of stages in the life cycle, $\circ$ denotes the Hadamard product, $\mathbf{N} = (\mathbf{I} - \mathbf{U})^{-1}$ is the fundamental matrix of the Markov chain and $\mathbf{Z}$ is a matrix that selects the living states. From these moment vectors we can calculate all the statistics of LRO. In addition, the full sensitivity analysis, calculating the derivatives of any of the moments of LRO to any parameters affecting any of the transition, mortality, or reward matrices, has been presented by van Daalen and Caswell (2017).

One of the most significant findings of this line of research has been that, in many cases, individual stochasticity can account for most or all of the observed phenotypic variance in LRO (Steiner and Tuljapurkar 2012; van Daalen and Caswell 2017). It appears that the contribution of stochasticity to variance in lifetime reproductive output has been underappreciated.

5.5   Variable and Stochastic Environments

The variance due to individual stochasticity can be examined in the case of variable environments (Caswell 2006; Tuljapurkar and Horvitz 2006; Horvitz and Tuljapurkar 2008; see also Chap. 8). Several cases can be considered:

- Deterministic aperiodic environments. These usually appear as specific historical sequences; e.g., the specific sequence of vital rates exhibited by the right whale between 1980 and 1998 (Caswell 2006). That sequence is fixed, and is neither random nor periodic.
- Periodic environments. A periodic model may describe seasonal variation within a year, or may approximate inter-annual variability in events such as floods, fires, or hurricanes.
- Stochastic iid environments. In such environments, successive states are drawn independently from a fixed probability distribution; hence the identifier iid, short for "independent and identically distributed."
- Markovian stochastic environments. In a Markovian environment the probability distribution of the next environmental state may depend on the current state. This permits study of the effects of environmental autocorrelation. Markovian environments include periodic and iid environments as special cases.

See Tuljapurkar (1990) for a thorough discussion of types of stochastic environments.

When studying variable environments, it is important to distinguish *period* and *cohort* calculations. Period calculations are based on the vital rates in a given year. They describe the results of the hypothetical situation where the conditions of year t are maintained indefinitely, and compare those to the results for conditions in year $t + 1$, etc. Period calculations are a way to summarize the effects of changing environment. But an individual born in year t does not live its life under the conditions of year t. It spends its first year of life under the conditions in year t, its second year under the conditions of year $t + 1$, and so on. Results calculated in this way are called *cohort* calculations, because they describe a cohort born in

year t and living through the environmental sequence starting then. Period-specific calculations are easy; simply apply the time-invariant calculation to the vital rates of each year and tabulate the results. Cohort calculations, however, must account for all the possible environmental sequences through which a cohort may pass. Caswell (2006) and Tuljapurkar and Horvitz (2006) independently introduced two different, complementary approaches to doing so. I will present the former approach here.

5.5.1   A Model for Variable Environments

In a variable environment, the transient matrix $\mathbf{U}$ is a time-varying matrix $\mathbf{U}(t)$. We can define a fundamental matrix by

$$\mathbf{N} = \mathbf{I} + \mathbf{U}(0) + \mathbf{U}(1)\mathbf{U}(0) + \mathbf{U}(2)\mathbf{U}(1)\mathbf{U}(0) + \cdots \tag{5.60}$$

The (i, j) element of $\mathbf{N}$ is the expected occupancy time in transient state i by an individual starting in transient state j at time 0, and experiencing the specific sequence of environments $\mathbf{U}(0), \mathbf{U}(1), \ldots$. Thus there will be a different matrix $\mathbf{N}$ for each possible environmental sequence.

Tuljapurkar and Horvitz (2006), whose paper I highly recommend, work directly from (5.60) to develop the means and variances of $\mathbf{N}$, $\boldsymbol{\eta}$, and survivorship, in periodic, iid, and Markovian environments. Here, we consider an approach in which an individual is jointly classified by stage and environment, using the vec-permutation model developed by Hunter and Caswell (2005b).

Suppose that there are q environmental states $\epsilon = 1, \ldots, q$ and s stages, $g = 1, \ldots, s$. Corresponding to environment i is a $s \times s$ transient matrix $\mathbf{U}_i$. Assemble the matrices $\mathbf{U}_i$ into a block-diagonal matrix

$$\mathbb{U} = \begin{pmatrix} \mathbf{U}_1 & & \\ & \ddots & \\ & & \mathbf{U}_q \end{pmatrix} \tag{5.61}$$

of dimension $sq \times sq$.

The transitions among environmental states are defined by a $q \times q$ column-stochastic matrix $\mathbf{D}$. Use the matrix $\mathbf{D}$ to construct a block-diagonal environmental transition matrix

$$\mathbb{D} = \begin{pmatrix} \mathbf{D} & 0 & \cdots & 0 \\ 0 & \mathbf{D} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \mathbf{D} \end{pmatrix} \tag{5.62}$$

of dimension $sq \times sq$.

Suppose that there are 4 environmental states. In an aperiodic deterministic environment,

$$\mathbf{D} = \begin{pmatrix} 0\ 0\ 0\ 0 \\ 1\ 0\ 0\ 0 \\ 0\ 1\ 0\ 0 \\ 0\ 0\ 1\ 1 \end{pmatrix} \tag{5.63}$$

That is, the environment moves deterministically from state 1 to state 2 to state 3 to state 4. Setting $d_{44} = 1$ solves the problem of what to do at the end of the sequence, by the (possibly satisfactory) trick of letting the final state repeat indefinitely. In a periodic environment,

$$\mathbf{D} = \begin{pmatrix} 0\ 0\ 0\ 1 \\ 1\ 0\ 0\ 0 \\ 0\ 1\ 0\ 0 \\ 0\ 0\ 1\ 0 \end{pmatrix} \tag{5.64}$$

In an iid environment in which environment i occurs with probability π_i,

$$\mathbf{D} = \begin{pmatrix} \pi_1\ \pi_1\ \pi_1\ \pi_1 \\ \pi_2\ \pi_2\ \pi_2\ \pi_2 \\ \pi_3\ \pi_3\ \pi_3\ \pi_3 \\ \pi_4\ \pi_4\ \pi_4\ \pi_4 \end{pmatrix} \tag{5.65}$$

In a Markovian environment, $\mathbf{D}$ is a column stochastic transition matrix describing the transition probabilities. I will assume that the environmental Markov chain is ergodic, with a stationary probability distribution denoted by $\boldsymbol{\pi}$. This gives the long-term frequency of occurrence of each environmental state.

The state of the cohort can be specified by a matrix $\mathbf{X}$, of dimension $s \times q$, with rows corresponding to stages and columns to environments, and where $x_{ij}(t)$ is the expected number of individuals in stage i and environmental state j at time t.

$$\mathbf{X}(t) = \begin{pmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & & \vdots \\ x_{s1} & \cdots & x_{sq} \end{pmatrix} \tag{5.66}$$

We rearrange $\mathbf{X}$ into a vector by applying the vec operator to $\mathbf{X}^{\mathsf{T}}$,

$$\text{vec}\,\mathbf{X}^{\mathsf{T}} = \begin{pmatrix} x_{11} & \cdots & x_{1q} \big| \cdots \big| x_{s1} & \cdots & x_{sq} \end{pmatrix}^{\mathsf{T}} \tag{5.67}$$

The first block of entries gives stage 1 individuals in environments 1 through q. The second block gives stage 2 individuals in environments 1 through q, and so on.

To describe the dynamics of the cohort, suppose that individuals first move among stages, according to the vital rates determined by the current environment, and then the environment changes to a new state according to **D**. Then

$$\text{vec}\,^{\mathsf{T}}\mathbf{X}(t+1) = \mathbb{D}\,\mathbf{K}_{s,q}\,\mathbb{U}\,\mathbf{K}_{s,q}^{\mathsf{T}}\,\text{vec}\,^{\mathsf{T}}\mathbf{X}(t) \tag{5.68}$$

The matrix $\mathbf{K}_{s,q}$ is the vec-permutation matrix (Henderson and Searle 1981; Hunter and Caswell 2005b), commutation matrix (Magnus and Neudecker 1979), which permutes the entries of a vector so that

$$\text{vec}\,^{\mathsf{T}}\mathbf{X} = \mathbf{K}_{s,q}\,\text{vec}\,\mathbf{X} \tag{5.69}$$

(see Sect. 2.2.3). Like all permutation matrices, its transpose is also its inverse. Its role here is to rearrange the population vector into a form appropriate for multiplication by the block-diagonal matrices $\mathbb{B}$ and $\mathbb{D}$.

Working from right to left, (5.68) first rearranges the vector, then applies the block-transition matrix $\mathbb{U}$, then reverses the rearrangement of the vector, and finally applies the environmental transition block matrix $\mathbb{D}$ to obtain the expected cohort at $t+1$. This gives a transition matrix for the joint process,

$$\widetilde{\mathbf{U}} = \mathbb{D}\,\mathbf{K}_{s,q}\,\mathbb{U}\,\mathbf{K}_{s,q}^{\mathsf{T}} \tag{5.70}$$

that incorporates the demographic transitions within each environment and the patterns of time variation among environments.[5] Here and in what follows, the tilde distinguishes the matrix from the environment-specific matrices.

Matrices of similar form, but not using this formalism, were introduced by Horvitz to study populations in habitat patches where the habitat patches change state over time, for example in recovering from disturbance (Horvitz and Schemske 1986; Pascarella and Horvitz 1998). Horvitz introduced the term "megamatrix" to describe these models. A megamatrix, in the sense of Horvitz, is a special case of (5.70) when the population is classified by stages within environmental states, the demographic matrices are applied first, and the environmental transition matrices $\mathbf{D}_i$ are identical for all stages, as is the case in (5.62).

---

[5]Note that (5.68) computes the *expected* population at $t+1$ from the *expected* population at t. It might be tempting to do this with the projection matrix $\mathbb{A}$ and use the eigenvalues of $\tilde{\mathbf{A}}$ to calculate the stochastic population growth rate. However, this would give the growth rate of the mean population, but not the stochastic growth rate (which is always less than or equal to the growth rate of the mean population). For calculations such as moments of longevity, which are explicitly properties of the expected population, the difference does not arise.

5.5.2   The Fundamental Matrix

Since $\widetilde{\mathbf{U}}$ is the transient matrix of an absorbing Markov chain, the fundamental matrix in the time-varying environment is

$$\widetilde{\mathbf{N}} = \left(\mathbf{I}_{sp} - \widetilde{\mathbf{U}}\right)^{-1} \tag{5.71}$$

The elements of $\widetilde{\mathbf{N}}$ give the expected occupancy times in each stage, in each environment, as a function of the starting stage and starting environment.

Notation alert Developing a complete system of notation for $\widetilde{\mathbf{N}}$ would obscure more than it would clarify. Pictures can help. As I present the fundamental matrix and some of the properties calculated from it, I will use diagrams for a simple case with three stages and two environments. I will often indicate the dimension of matrices and vectors with subscripts. I will use g to denote stages ($g = 1, 2, \ldots, s$) and ϵ to denote environments ($\epsilon = 1, \ldots, q$). I will use superscripts on $\widetilde{\mathbf{N}}$ and quantities derived from it, to distinguish different ways of combining information across environmental states (see Table 5.1).

Recall that in a constant environment, v_{ij} was the number of visits to stage i, starting in stage j. Now we must consider the visits to stage i in environment ϵ, starting in stage j and environment ϵ_0, so we write

$$\widetilde{\mathbf{N}} = E\left(v_{ij,\epsilon}|\epsilon_0\right) \tag{5.72}$$

Table 5.1 Superscript notation for time-varying models. The tilde indicates quantities calculated from the complete transient matrix $\widetilde{\mathbf{U}}$ in (5.70). Occupancy and times to absorbtion depend on the initial and final demographic and environmental states. The superscripts ($\ddagger$, $\S$, $\heartsuit$) indicate choices of summing and averaging over the environmental states. The superscripts are shown here for the fundamental matrix $\widetilde{\mathbf{N}}$

Symbol	Definition	Description	Equation
$\widetilde{\mathbf{N}}$	$E\left(v_{ij,\epsilon}|\epsilon_0\right)$	Expected visits to state i in environment ϵ, starting from state j in environment ϵ_0	(5.72)
$\widetilde{\mathbf{N}}^{\ddagger}$	$E\left(v_{ij}|\epsilon_0\right)$	Expected visits to state i, summed over environments, starting from state j in environment ϵ_0	(5.73)
$\widetilde{\mathbf{N}}^{\ddagger\ddagger}$		Rearrangement of the rows and columns of $\widetilde{\mathbf{N}}^{\ddagger}$	(5.74)
$\widetilde{\mathbf{N}}^{\S}$	$E\left(v_{ij,\epsilon}\right)$	Expected visits to state i and environmental state ϵ, averaged over initial environmental states	(5.75)
$\widetilde{\mathbf{N}}^{\S\S}$		Rearrangement of the rows and columns of $\widetilde{\mathbf{N}}^{\S}$	(5.76)
$\widetilde{\mathbf{N}}^{\heartsuit}$	$E\left(v_{ij}\right)$	Expected visits to state i summed over environments, starting from state j and averaged over initial environmental states	(5.77)

The structure of $\widetilde{\mathbf{N}}$ when $s = 3$ and $q = 2$ is

		$g = 1$		$g = 2$		$g = 3$	
$\widetilde{\mathbf{N}}_{sq \times sq}$:		$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$
	$g = 1 \;\; \epsilon = 1$						
	$\epsilon = 2$						
	$g = 2 \;\; \epsilon = 1$						
	$\epsilon = 2$						
	$g = 3 \;\; \epsilon = 1$						
	$\epsilon = 2$						

From $\widetilde{\mathbf{N}}$ we can obtain the expected occupancy time in each stage, regardless of the environment in which those visits occur, by aggregating rows. The resulting matrix $\widetilde{\mathbf{N}}^{\ddagger}$ is

$$\widetilde{\mathbf{N}}^{\ddagger} = E\left(v_{ij}|\epsilon_0\right)$$
$$= \left(\mathbf{I}_s \otimes \mathbf{1}_{q \times 1}^{\mathsf{T}}\right) \widetilde{\mathbf{N}} \tag{5.73}$$

where $\mathbf{1}_{q \times 1}$ is a vector of ones. The structure of $\widetilde{\mathbf{N}}^{\ddagger}$ is

		$g = 1$		$g = 2$		$g = 3$	
$\widetilde{\mathbf{N}}^{\ddagger}_{s \times sq}$:		$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$
	$g = 1$						
	$g = 2$						
	$g = 3$						

If it is useful to group stages within initial environments, rather than grouping environments within stages, $\widetilde{\mathbf{N}}^{\ddagger}$ can be rearranged as

$$\widetilde{\mathbf{N}}^{\ddagger\ddagger} = \widetilde{\mathbf{N}}^{\ddagger} \, \mathbf{K}_{s,q} \tag{5.74}$$

with the structure

		$\epsilon_0 = 1$			$\epsilon_0 = 2$		
$\widetilde{\mathbf{N}}^{\ddagger\ddagger}_{s \times sq}$:		$g = 1$	$g = 2$	$g = 3$	$g = 1$	$g = 2$	$g = 3$
	$g = 1$						
	$g = 2$						
	$g = 3$						

The matrices $\widetilde{\mathbf{N}}^{\ddagger}$ and $\widetilde{\mathbf{N}}^{\ddagger\ddagger}$ both display expected occupancy of each stage as a function of initial state and environment. To describe the fates of individuals without specifying their initial environment, we take an expectation over the stationary distribution $\boldsymbol{\pi}$ of initial environments. This gives

$$\widetilde{\mathbf{N}}^{\S} = E\left[\nu_{ij,\epsilon}\right]$$

$$= \widetilde{\mathbf{N}}\left(\mathbf{I}_s \otimes \boldsymbol{\pi}\right) \qquad (5.75)$$

The structure of $\widetilde{\mathbf{N}}^{\S}$ is

$$
\widetilde{\mathbf{N}}^{\S}_{sq \times s} :
\begin{array}{cc}
 & \\
\end{array}
$$

		$\epsilon_0 = \bar{\epsilon}$		
		$g=1$	$g=2$	$g=3$
$g=1$	$\epsilon=1$			
	$\epsilon=2$			
$g=2$	$\epsilon=1$			
	$\epsilon=2$			
$g=3$	$\epsilon=1$			
	$\epsilon=2$			

The rows of $\widetilde{\mathbf{N}}^{\S}$ can be rearranged to display stages within environments, giving

$$\widetilde{\mathbf{N}}^{\S\S} = \mathbf{K}_{s,q}^{\mathsf{T}}\ \widetilde{\mathbf{N}}^{\S} \qquad (5.76)$$

with the structure

		$\epsilon_0 = \bar{\epsilon}$		
		$g=1$	$g=2$	$g=3$
$\epsilon=1$	$g=1$			
	$g=2$			
	$g=3$			
$\epsilon=2$	$g=1$			
	$g=2$			
	$g=3$			

$$\widetilde{\mathbf{N}}^{\S\S}_{sq \times s} :$$

Finally, aggregating over destination environments *and* averaging over initial environments gives a matrix containing the expected occupancy of stages as a function of initial stage, averaged over environments

$$\widetilde{\mathbf{N}}^{\heartsuit} = E\left[\nu_{ij}\right]$$

$$= \left(\mathbf{I}_s \otimes \mathbf{1}_{q\times 1}^{\mathsf{T}}\right) \widetilde{\mathbf{N}}\ \left(\mathbf{I}_s \otimes \boldsymbol{\pi}\right) \qquad (5.77)$$

The structure of $\widetilde{\mathbf{N}}^\heartsuit$ is

$$
\widetilde{\mathbf{N}}^\heartsuit_{s\times s}: \quad
\begin{array}{cc}
 & \begin{array}{c} \epsilon_0 = \bar{\epsilon} \\ g=1 \ g=2 \ g=3 \end{array} \\
\begin{array}{c} g=1 \\ g=2 \\ g=3 \end{array} & \boxed{}
\end{array}
$$

The matrix $\widetilde{\mathbf{N}}^\heartsuit$, obtained by the simple calculation (5.77), is "the" fundamental matrix for the variable environment. It could be compared directly to the fundamental matrix in a constant environment (e.g., the environment defined by one of the environmental states).

5.5.3 Longevity in a Variable Environment

Life expectancy, as a function of initial stage and initial environment is obtained by summing the columns of $\widetilde{\mathbf{N}}$,

$$
E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right) = E\left[\boldsymbol{\eta}^{\mathsf{T}}|\epsilon_0\right]
$$
$$
= \mathbf{1}^{\mathsf{T}}_{sq\times 1}\widetilde{\mathbf{N}} \tag{5.78}
$$

The structure of $E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right)$ is

$$
E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right): \quad
\begin{array}{ccc}
g=1 & g=2 & g=3 \\
\epsilon_0=1 \ \epsilon_0=2 & \epsilon_0=1 \ \epsilon_0=2 & \epsilon_0=1 \ \epsilon_0=2 \\
\boxed{} & \boxed{} & \boxed{}
\end{array}
$$

Averaging this conditional life expectancy over the stationary distribution $\boldsymbol{\pi}$ of initial environments gives

$$
E\left(\widetilde{\boldsymbol{\eta}}^\heartsuit\right) = E\left(\widetilde{\boldsymbol{\eta}}\right)(\mathbf{I}_s \otimes \boldsymbol{\pi}) \tag{5.79}
$$

This measure of life expectancy in a variable environment is directly comparable to $E(\boldsymbol{\eta})$ calculated from the same life history in a constant environment.

5.5.3.1 Variance in Longevity

In a constant environment, the variance among individuals in longevity is due to individual stochasticity. In a time-varying environment, the variance contains an

additional component due to differences among individuals as a function of their environment at birth. Applying (5.27) to $\widetilde{\mathbf{N}}$ we obtain the variances conditional on the initial environment:

$$V\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}|\epsilon_0\right] = E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right)\left(2\widetilde{\mathbf{N}} - \mathbf{I}_{sq}\right) - E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right) \circ E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right) \tag{5.80}$$

As indicated by the notation, $V\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}|\epsilon_0\right]$ is a conditional variance of $\widetilde{\boldsymbol{\eta}}$, given the initial environment ϵ_0. The initial environment is distributed according to the stationary distribution $\boldsymbol{\pi}$, so the unconditional longevity $\boldsymbol{\eta}$ follows a finite mixture distribution with mixing distribution $\boldsymbol{\pi}$.

The unconditional variance of $\boldsymbol{\eta}$, taking account of both sources of variability, is

$$V\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right] = V\left[E(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}|\epsilon_0)\right] + E_\pi\left[V(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}|\epsilon_0)\right] \tag{5.81}$$

where E_π denotes the expectation over the stationary distribution $\boldsymbol{\pi}$ of initial environments (Rényi 1970, p. 275, Theorem 1). This can be rearranged as

$$V\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right] = E_\pi\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}} \circ \widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right] - E_\pi\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right] \circ E_\pi\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right] + E_\pi\left[V(\boldsymbol{\eta}^{\mathsf{T}}|\epsilon_0)\right]$$

$$= \left[E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right) \circ E\left(\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}\right)\right](\mathbf{I}_s \otimes \boldsymbol{\pi}) - \left[E\left(\widetilde{\boldsymbol{\eta}}^{\heartsuit}\right) \circ E\left(\widetilde{\boldsymbol{\eta}}^{\heartsuit}\right)\right]^{\mathsf{T}}$$

$$+ V\left[\widetilde{\boldsymbol{\eta}}^{\mathsf{T}}|\epsilon_0\right](\mathbf{I}_s \otimes \boldsymbol{\pi}) \tag{5.82}$$

(e.g., Frühwirth-Schnatter 2006, p. 10). This variance decomposition has developed into a powerful tool for the analysis of heterogeneity in demography (Edwards 2011; Hartemink and Caswell 2018; Hartemink et al. 2017; Caswell et al. 2018; Jenouvrier et al. 2018).

The choice of the mixing distribution $\boldsymbol{\pi}$ is important. Hernandez-Suarez et al. (2012) present an alternative where $\boldsymbol{\pi}$ is the stationary distribution of births across environments, rather than the distribution of environments itself.

5.5.4 A Time-Varying Example: Lomatium bradshawii

Lomatium bradshawii is an endangered herbaceous perennial plant, found in only a few isolated populations in prairies of Oregon and Washington. These habitats were, until recent times, subject to natural and anthropogenic fires, to which *L. bradshawii* seems to have adapted. Fall-season fires increase plant size and seedling recruitment, but the effect fades within a few years. Populations in burned areas have higher growth rates and lower probabilities of extinction than unburned populations (Caswell and Kaye 2001).

A stochastic demographic model for *L. bradshawii* was developed by Caswell and Kaye (2001), Kaye et al. (2001), and Kaye and Pyke (2003) based on data from an experimental study using controlled burning. Individuals were classified into six stages based on size and reproductive status: yearlings, small and large vegetative plants, and small, medium, and large reproductive plants. The environment was classified into four states defined by fire history: the year of a fire and 1, 2, and 3+ years post-fire. Projection matrices were estimated in each environment; the example here is based on one of the two sites (Rose Prairie) in the original study. The matrices are given in Caswell and Kaye (2001).

L. bradshawii performs well under recently burned conditions, but less well in sites that have not been recently burned. For example, the values of λ are

$$
\begin{array}{lcccc}
\text{Years since fire:} & 0 & 1 & 2 & \geq 3 \\
\text{Growth rate } \lambda : & 1.18 & 1.12 & 0.48 & 0.88
\end{array}
$$

Caswell and Kaye (2001) found a minimum frequency of fire (0.4–0.5) below which the stochastic growth rate was negative and the population would be unable to persist. Effects of autocorrelation were small, but positive autocorrelation reduced the stochastic growth rate.

As an example of a time-varying analysis, let us examine *L. bradshawii* in a Markovian environment. Let f be the long-term frequency of fire, and ρ the temporal autocorrelation. Then the transition matrix for environmental states is

$$
\mathbf{D} = \begin{pmatrix}
p & q & q & q \\
1-p & 0 & 0 & 0 \\
0 & 1-q & 0 & 0 \\
0 & 0 & 1-q & 1-q
\end{pmatrix}
\tag{5.83}
$$

where $q = f(1-\rho)$ and $p = \rho + q$.

Figure 5.7a shows the life expectancy $E(\widetilde{\boldsymbol{\eta}}|\epsilon_0)$ of *L. bradshawii* as a function of initial stage and initial environmental state, from (5.78). Life expectancy increases with the stage (size) of a plant. A seedling has its greatest life expectancy in the year of a fire, less in an environment three or more years post-fire. A large flowering plant, in contrast, has its greatest life expectancy in an environment three or more years post-fire. When the environment-dependence is averaged over the stationary distribution of environmental states, there is a smooth increase in life expectancy from ∼2.5 years for a seedling to 8 years for a large flowering plant (Fig. 5.7b). The standard deviation of longevity also increases with stage, in a pattern very similar to that of the expectation.

These patterns in the mean and variance of longevity (Fig. 5.7) depend on the stochastic properties of the environment—in this case, the frequency f and autocorrelation ρ of fires. Even with an environmental model this simple, the effects of f and ρ can be complicated. I know of no previous attempts to examine their effects on longevity. To do so, I calculated life expectancy with $f = 0.5$ for autocorrelation $-1 < \rho < 1$, and with $\rho = 0$ for fire frequency $0 < f < 1$.
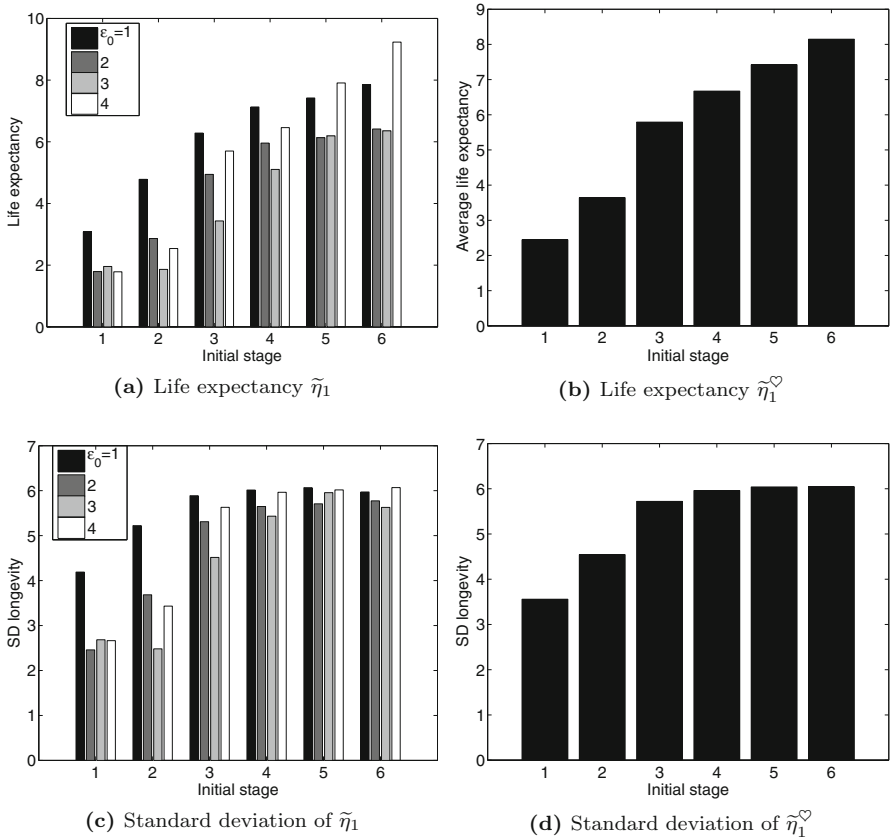
(a) Life expectancy $\widetilde{\eta}_1$

(b) Life expectancy $\widetilde{\eta}_1^{\heartsuit}$

(c) Standard deviation of $\widetilde{\eta}_1$

(d) Standard deviation of $\widetilde{\eta}_1^{\heartsuit}$

Fig. 5.7 The expectation and standard deviation of longevity for *Lomatium bradshawii* in a stochastic fire environment. (**a**) Expected longevity conditional on initial environment (ϵ_0). (**b**) Expected longevity averaged over the stationary distribution of initial environments. (**c**) The standard deviation of longevity conditional on initial environment. (**d**) The standard deviation of longevity over the stationary distribution of initial environments. The frequency of fire is 0.5 and the temporal autocorrelation $\rho = 0.7$

The life expectancy of early life cycle stages increases monotonically with fire frequency (Fig. 5.8a), but the life expectancy of large reproductive plants is greatest at either low or high fire frequencies. The standard deviation of longevity increases with f (Fig. 5.8b). As $f \rightarrow 1$, the standard deviation of longevity is approximately twice the mean.

The autocorrelation of fires has little effect on the life expectancy of seedlings, but a larger effect on that of large plants. For the latter, life expectancy is maximized as $\rho \rightarrow -1$ (alternating fire and non-fire years) or as $\rho \rightarrow 1$ (long periods of fires
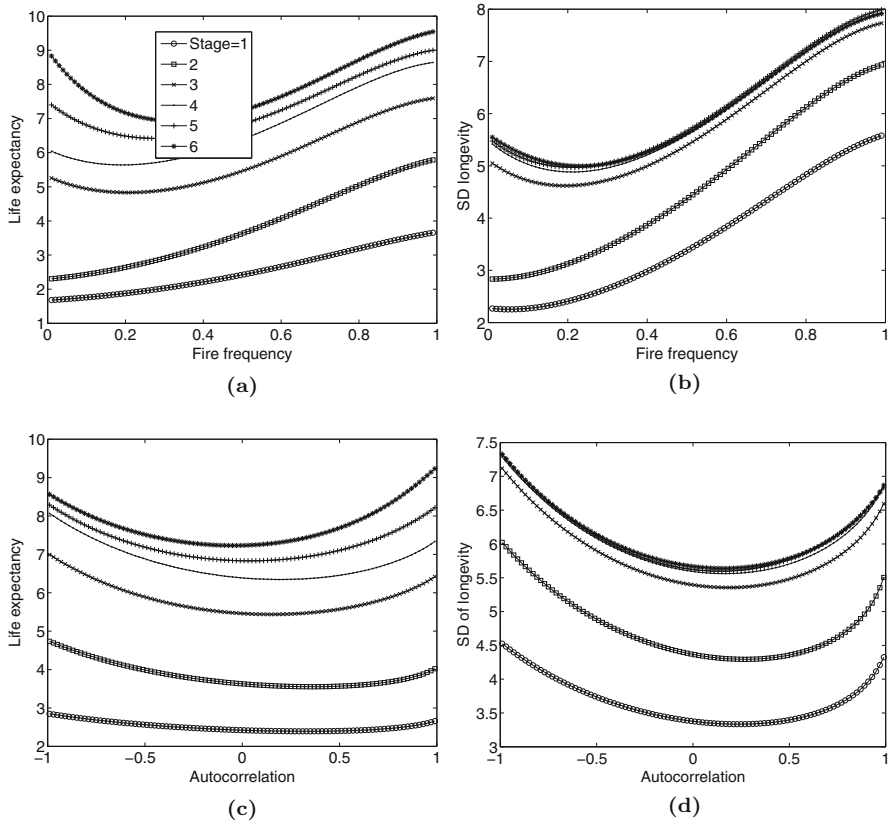
Fig. 5.8 The expectation and standard deviation of longevity, averaged over the stationary distribution of initial environments, for *Lomatium bradshawii*, as a function of the initial stage, the fire frequency f, and the temporal autocorrelation ρ. Parameters as in Fig. 5.7. (**a**) Life expectancy $\widetilde{\boldsymbol{\eta}}^{\heartsuit}$. (**b**) Standard deviation of longevity. (**c**) Life expectancy $\widetilde{\boldsymbol{\eta}}^{\heartsuit}$. (**d**) Standard deviation of longevity

alternating with long periods without fire). The standard deviation of longevity also shows a strong U-shaped response to ρ for all stages. The generality of this pattern is unknown.

5.6   The Importance of Individual Stochasticity

The concept of individual stochasticity strikes to the heart of one of the most fundamental problems in population biology: the sources of variability among individuals. Heterogeneity—genuine differences among individuals—translates into differences in the age- or stage-specific vital rates to which they are subject.

Heterogeneity may arise from genetics, from physiological effects, from health conditions, or from unknown causes ("frailty," "quality"). Stochasticity results from the random outcomes of probabilistic processes. Markov chains naturally treat individual trajectories (i.e., individual lives) as realizations of an underlying stochastic process, and so much of this chapter has been focused on the analysis of individual stochasticity. The distinction is particularly important in evolutionary demography, where variance in lifetime reproductive output is routinely treated as variance in fitness, or a component of fitness. See Sect. 5.4.4 for some recent work on this problem.

Individual stochasticity is an important component of demography, for both human and non-human populations. It complements environmental stochasticity (externally imposed random changes in vital rates) and demographic stochasticity (randomness in the growth of populations due to stochastic survival and reproduction) (Caswell and Vindenes 2018). Individual stochasticity reflects randomness in the pathways that individuals take through the life cycle. It expresses itself in inter-individual variation in occupancy times, longevity, lifetime reproductive output, and other outcomes. The availability of methods based on Markov chains promises to change the way population biologists approach the analysis of variance among individuals (Caswell 2011; Tuljapurkar et al. 2009; Steiner and Tuljapurkar 2012; van Daalen and Caswell 2015; van Daalen and Caswell 2017).

5.7   Discussion

Taking advantage of the Markov chain formulation of the life cycle opens up a wealth of demographic information. The age-classified information extracted from a stage-classified model can form a valuable component of behavioral studies, especially if the model (like the right whale example) includes reproductive behavior as part of the life cycle structure. Longevity provides a powerful way to compare mortality schedules among species, populations, or environmental conditions, but it has been inaccessible to stage-classified analysis prior to the development of Markov chain methods. The generation time characterizes an important population time scale, with implications in conservation (IUCN Species Survival Commission 2001), but there has been no way to compute it from stage-classified models.

Stage-classified life cycles may have consequences that are not yet appreciated, but must be considered when interpreting the results. For example, any stage-classified model eventually leads to an age-independent mortality rate (Horvitz and Tuljapurkar 2008), and so is of limited use in the study of senescence. This fact has consequences for life expectancy and variance in longevity that are not well understood (at least by me). For the right whale, expected longevity at birth is 32 years with a standard deviation of 34 years. It is unlikely that there are appreciable numbers of whales alive at even one standard deviation above this mean. The high survival probability and the assumption of age-independence lead to the high standard deviation. Those of us who work with stage-classified models are

accustomed to this, but discount its importance because it (often) has little effect on λ. It will be important to determine the stochastic consequences of simplifying assumptions in the life cycle graph.

This chapter does not begin to exhaust the information that can be extracted from the Markov chain formulation of a stage-classified model. Three examples of particular interest are the occupancy of sets of states, the problem of competing risks, and the calculation of passage times. It is often of interest to calculate the statistics of occupancy of sets of states (e.g., all reproductive classes, or all stages in some particular health condition). We have seen how to calculate the moments of the occupancy time of single states. The mean occupancy time of a set of states is the sum of the mean occupancy times of each state, but that is not true for the variance or higher moments. Roth and Caswell (2018) derived a general expression for all the statistics, and the complete distribution, of occupancy time for any set of states. If more than one absorbing state exists (e.g., death at different stages, or from different causes), then the risks of absorbtion compete, because an individual can only be absorbed (i.e., die) once. It is possible to calculate the probability of absorbtion in each state, and to explore the effects of changing one risk on the probability of experiencing another (Caswell and Ouellette 2018). Passage times refer to the time required to get from one stage to another in the life cycle. An important passage time is the birth interval: the time from one birth to the next. This can only be calculated for individuals that do reproduce a second time (otherwise the interval is infinite), and so it requires developing a chain that is conditional on successfully reaching the reproductive state (Caswell 2001). In species that produce only one or a few offspring, reproduction cannot be adjusted in response to the environment by changing offspring number, and so changes in the birth interval are particularly important in such species.

A   Appendix: Derivations

This appendix contains step-by-step derivations of many of the results in this chapter, especially for sensitivities. Taking advantage of the freedom from length limits, I have tried to show the derivations step-by-step. Recall the definitions of the Hadamard product

$$\mathbf{A} \circ \mathbf{B} = \left(a_{ij} b_{ij} \right), \tag{5.84}$$

the Kronecker product

$$\mathbf{A} \otimes \mathbf{B} = \left(a_{ij} \mathbf{B} \right), \tag{5.85}$$

the vec operator

$$\text{vec} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}, \tag{5.86}$$

and Roth's theorem

$$\text{vec} \left(\mathbf{ABC} \right) = \left(\mathbf{C}^\mathsf{T} \otimes \mathbf{A} \right) \text{vec} \, \mathbf{B}. \tag{5.87}$$

A.1   Variance in Occupancy Times

The occupancy time in transient state i, starting from transient state j, is ν_{ij}. The matrix of variances of the ν_{ij} is

$$\mathbf{V} = \left(V(\nu_{ij}) \right) = \left(2\mathbf{N}_{\text{dg}} - \mathbf{I}_s \right) \mathbf{N} - \mathbf{N} \circ \mathbf{N} \tag{5.88}$$

(Caswell 2006, derived from Theorem 3.1 of Iosifescu 1980) where $\mathbf{N}_{\text{dg}}$ is a matrix with the diagonal elements of $\mathbf{N}$ on its diagonal and zeros elsewhere; it can be written

$$\mathbf{N}_{\text{dg}} = \mathbf{I}_s \circ \mathbf{N} \tag{5.89}$$

Differentiating both sides of (5.88) gives

$$d\mathbf{V} = 2(\mathbf{I}_s \circ d\mathbf{N})\mathbf{N} + 2(\mathbf{I}_s \circ \mathbf{N})(d\mathbf{N}) - d\mathbf{N}$$
$$- (d\mathbf{N}) \circ \mathbf{N} - \mathbf{N} \circ (d\mathbf{N}) \tag{5.90}$$

The next step is to apply the vec operator to both sides. The vec of a Hadamard product can be written in two ways:

$$\text{vec} \left(\mathbf{A} \circ \mathbf{B} \right) = \mathcal{D} \left(\text{vec} \, \mathbf{A} \right) \text{vec} \, \mathbf{B} = \mathcal{D} \left(\text{vec} \, \mathbf{B} \right) \text{vec} \, \mathbf{A}. \tag{5.91}$$

Using this result and Roth's theorem (5.87) gives

$$d\text{vec} \, \mathbf{V} = 2 \left(\mathbf{N}^\mathsf{T} \otimes \mathbf{I}_s \right) \mathcal{D} \left(\text{vec} \, \mathbf{I}_s \right) d\text{vec} \, \mathbf{N} + 2 \left[\mathbf{I}_s \otimes (\mathbf{I} \circ \mathbf{N}) \right] d\text{vec} \, \mathbf{N}$$
$$- d\text{vec} \, \mathbf{N} - 2\mathcal{D} \left(\text{vec} \, \mathbf{N} \right) d\text{vec} \, \mathbf{N} \tag{5.92}$$

Factoring out $d\text{vec}\,\mathbf{N}$ and using the chain rule gives the final result

$$
\frac{d\text{vec}\,\mathbf{V}}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left[2\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{I}_s\right)\mathcal{D}\left(\text{vec}\,\mathbf{I}_s\right) + 2\left(\mathbf{I}_s \otimes \mathbf{N}_{\text{dg}}\right) \right.
$$
$$
\left. -\mathbf{I}_{s^2} - 2\mathcal{D}\left(\text{vec}\,\mathbf{N}\right) \right]\frac{d\text{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{5.93}
$$

A.2   Life Expectancy

Let η_i be the time to absorbtion (i.e., death) of an individual currently in stage i. The vector $E(\boldsymbol{\eta})$ of expected values of the η_i satisfies

$$
E(\boldsymbol{\eta})^{\mathsf{T}} = \mathbf{1}^{\mathsf{T}}\mathbf{N} \tag{5.94}
$$

where $\mathbf{1}$ is a vector of ones. Differentiating both sides gives

$$
dE(\boldsymbol{\eta})^{\mathsf{T}} = \mathbf{1}^{\mathsf{T}}(d\mathbf{N}) \tag{5.95}
$$

Applying the vec operator gives

$$
dE(\boldsymbol{\eta}) = \left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\right)d\text{vec}\,\mathbf{N} \tag{5.96}
$$

Applying the identification theorem and the chain rule, and using (5.16) for the sensitivity of the fundamental matrix, gives

$$
\frac{dE(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left(\mathbf{I}_s^{\mathsf{T}} \otimes \mathbf{1}^{\mathsf{T}}\right)\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right)\frac{d\text{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{5.97}
$$

This gives the derivative of the entire vector of life expectancies. Suppose that stage 1 corresponds to birth. The life expectancy at birth is then

$$
E(\eta_1) = \mathbf{1}^{\mathsf{T}}\mathbf{N}\mathbf{e}_1 \tag{5.98}
$$

where $\mathbf{e}_1$ is a vector with 1 in the first position and zeros elsewhere. Following the same derivation gives

$$
\frac{dE(\eta_1)}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left(\mathbf{e}_1^{\mathsf{T}} \otimes \mathbf{1}^{\mathsf{T}}\right)\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right)\frac{d\text{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{5.99}
$$

A.3 Variance in Longevity

The variance of the time to absorbtion satisfies

$$V(\boldsymbol{\eta})^{\mathsf{T}} = \mathbf{1}^{\mathsf{T}} \mathbf{N} \left(2\mathbf{N} - \mathbf{I}\right) - E\left(\boldsymbol{\eta}^{\mathsf{T}}\right) \circ E\left(\boldsymbol{\eta}^{\mathsf{T}}\right) \tag{5.100}$$

(Caswell 2006, derived from Theorem 3.2 of Iosifescu 1980). Differentiating gives

$$dV(\boldsymbol{\eta})^{\mathsf{T}} = 2\mathbf{1}^{\mathsf{T}}(d\mathbf{N})\mathbf{N} + 2\mathbf{1}^{\mathsf{T}}\mathbf{N}(d\mathbf{N}) - \mathbf{1}^{\mathsf{T}}(d\mathbf{N}) - 2E\left(\boldsymbol{\eta}^{\mathsf{T}}\right) \circ dE\left(\boldsymbol{\eta}^{\mathsf{T}}\right) \tag{5.101}$$

Applying the vec operator and Roth's theorem (5.87), using (5.91) for the vec of the Hadamard product, gives

$$dV(\boldsymbol{\eta}) = \left[2\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{1}^{\mathsf{T}}\right) + 2\left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\mathbf{N}\right)\right.$$
$$\left. - \left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\right)\right] d\mathrm{vec}\,\mathbf{N} - 2\mathcal{D}\left(E(\boldsymbol{\eta})\right) dE(\boldsymbol{\eta}) \tag{5.102}$$

Substituting (5.96) for $dE(\boldsymbol{\eta})$ gives

$$dV(\boldsymbol{\eta}) = \left[2\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{1}^{\mathsf{T}}\right) + 2\left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\mathbf{N}\right)\right.$$
$$\left. - \left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\right) - 2\mathcal{D}\left(E(\boldsymbol{\eta})\right)\left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\right)\right] d\mathrm{vec}\,\mathbf{N} \tag{5.103}$$

Using (5.16) for the sensitivity of $\mathbf{N}$, the identification theorem, and the chain rule finally leads to

$$\frac{dV(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left[2\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{1}^{\mathsf{T}}\right) + 2\left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\mathbf{N}\right)\right.$$
$$\left. - \left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\right) - 2\mathcal{D}\left(E(\boldsymbol{\eta})\right)\left(\mathbf{I}_s \otimes \mathbf{1}^{\mathsf{T}}\right)\right]\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right)\frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{5.104}$$

A.4 Net Reproductive Rate

The net reproductive rate R_0 is given by the dominant eigenvalue of $\mathbf{FN}$. Let $\mathbf{y}$ and $\mathbf{x}$ be the right and left eigenvectors, respectively, of $\mathbf{FN}$, corresponding to R_0. The matrix calculus version of the standard eigenvalue perturbation result (e.g., Caswell 1978) gives

$$dR_0 = \mathbf{x}^{\mathsf{T}} d(\mathbf{FN})\mathbf{y}$$
$$= \mathbf{x}^{\mathsf{T}} \left[(d\mathbf{F})\mathbf{N} + \mathbf{F}(d\mathbf{N})\right]\mathbf{y} \tag{5.105}$$

Applying the vec operator to both sides gives

$$d R_0 = \left(\mathbf{y}^\mathsf{T}\mathbf{N}^\mathsf{T} \otimes \mathbf{x}^\mathsf{T}\right) d\text{vec}\,\mathbf{F} + \left(\mathbf{y}^\mathsf{T} \otimes \mathbf{x}^\mathsf{T}\mathbf{F}\right) d\text{vec}\,\mathbf{N} \tag{5.106}$$

Applying the chain rule and the result (5.14) for $d$vec $\mathbf{N}$ gives the sensitivity of R_0 in terms of effects of the parameter vector $\boldsymbol{\theta}$ on the fertility matrix $\mathbf{F}$ and the transient matrix $\mathbf{U}$:

$$\frac{d R_0}{d\boldsymbol{\theta}^\mathsf{T}} = \left(\mathbf{y}^\mathsf{T}\mathbf{N}^\mathsf{T} \otimes \mathbf{x}^\mathsf{T}\right)\frac{d\text{vec}\,\mathbf{F}}{d\boldsymbol{\theta}^\mathsf{T}} + \left(\mathbf{y}^\mathsf{T} \otimes \mathbf{x}^\mathsf{T}\mathbf{F}\right)\left(\mathbf{N}^\mathsf{T} \otimes \mathbf{N}\right)\frac{d\text{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^\mathsf{T}} \tag{5.107}$$

A.5   Cohort Generation Time

To derive the cohort generation time, we begin at time $t = 0$ with an individual newly born in stage j. This tiny cohort is described by an initial vector $\mathbf{e}_j$. The expected survivors of this cohort at time t are $\mathbf{U}^t\mathbf{e}_j$. The expected offspring produced by these survivors at time t are $\mathbf{F}\mathbf{U}^t\mathbf{e}_j$. Summing over the lifetime of the cohort gives a vector of expected lifetime reproduction, of all types of offspring,

$$E(\text{total offspring}) = \sum_{t=0}^{\infty}\mathbf{F}\mathbf{U}^t\mathbf{e}_j$$

$$= \mathbf{F}\left(\sum_{t=0}^{\infty}\mathbf{U}^t\right)\mathbf{e}_j$$

$$= \mathbf{F}\mathbf{N}\mathbf{e}_j \tag{5.108}$$

Let $\mathbf{m}^{(j)}(t)$ be the vector of offspring production at time t, expressed as a *proportion* of the lifetime total of the individual starting in stage j. Then

$$\mathbf{m}^{(j)}(t) = \mathcal{D}\left(\mathbf{F}\mathbf{N}\mathbf{e}_j\right)^{-1}\left(\mathbf{F}\mathbf{U}^t\mathbf{e}_j\right) \tag{5.109}$$

If no offspring of some stage, say stage i, are produced, then set $m_i^{(j)}(t) = 0$.

The cohort generation time $\boldsymbol{\mu}^{(j)}$ is the expectation of the distribution defined by $\mathbf{m}^{(j)}(t)$:

$$\boldsymbol{\mu}^{(j)} = \sum_{x=0}^{\infty}x\mathbf{m}^{(j)}(x)$$

$$= \sum_x \mathcal{D}\left(\mathbf{F}\mathbf{N}\mathbf{e}_j\right)^{-1}x\mathbf{F}\mathbf{U}^x\mathbf{e}_j$$

$$= \mathcal{D}\left(\mathbf{F}\mathbf{N}\mathbf{e}_j\right)^{-1}\mathbf{F}\left(\sum_x x\mathbf{U}^x\right)\mathbf{e}_j. \tag{5.110}$$

The summation can be simplified

$$\sum_x x\mathbf{U}^x = \mathbf{0} + \mathbf{U} + 2\mathbf{U}^2 + 3\mathbf{U}^3 + \cdots$$

$$= \mathbf{U}\left[\mathbf{0} + \mathbf{U} + 2\mathbf{U}^2 + \cdots + \mathbf{I} + \mathbf{U} + \mathbf{U}^2 + \cdots\right]$$

$$= \mathbf{U}\left[\mathbf{N} + \sum_x x\mathbf{U}^x\right]. \tag{5.111}$$

Solving this gives

$$\sum_x x\mathbf{U}^x = \mathbf{NUN}. \tag{5.112}$$

Putting all the pieces together gives the generation time

$$\boldsymbol{\mu}^{(j)} = \mathcal{D}\left(\mathbf{FNe}_j\right)^{-1} \mathbf{FNUNe}_j. \tag{5.113}$$

A.5.1   Sensitivity of Generation Time

To differentiate (5.113) may seem complicated. To make life easier, define some notation,

$$\mathbf{X} = \mathcal{D}\left(\mathbf{FNe}_j\right) \tag{5.114}$$

$$\mathbf{r} = \mathbf{FNUNe}_j \tag{5.115}$$

in terms of which (5.113)

$$\boldsymbol{\mu}^{(j)} = \mathbf{X}^{-1}\mathbf{r}. \tag{5.116}$$

Differentiate,

$$d\boldsymbol{\mu}^{(j)} = d\left(\mathbf{X}^{-1}\right)\mathbf{r} + \mathbf{X}^{-1}d\mathbf{r} \tag{5.117}$$

and apply the vec operator

$$d\boldsymbol{\mu}^{(j)} = \left(\mathbf{r}^{\mathsf{T}} \otimes \mathbf{I}\right)d\mathrm{vec}\,\mathbf{X}^{-1} + \mathbf{X}^{-1}d\mathrm{vec}\,\mathbf{r}. \tag{5.118}$$

The same steps that led to Eq. (5.14) for $d\mathrm{vec}\,\mathbf{N}$, and noting that $\mathbf{X}$ is symmetric, leads to

$$d\mathrm{vec}\,\mathbf{X}^{-1} = -\left(\mathbf{X}^{-1} \otimes \mathbf{X}^{-1}\right)d\mathrm{vec}\,\mathbf{X}. \tag{5.119}$$

The differential of vec $\mathbf{X}$ is obtained by writing

$$\mathbf{X} = \mathbf{I} \circ \left(\mathbf{FNe}_j \mathbf{1}^\mathsf{T} \right). \tag{5.120}$$

Differentiating and using the rule (5.91) for the vec of a Hadamard product gives

$$d\text{vec}\,\mathbf{X} = \mathcal{D}\left(\text{vec}\,\mathbf{I}\right)\left[\left(\mathbf{1e}_j^\mathsf{T}\mathbf{N}^\mathsf{T} \otimes \mathbf{I}\right)d\text{vec}\,\mathbf{F} + \left(\mathbf{1e}_j \otimes \mathbf{F}\right)d\text{vec}\,\mathbf{N}\right] \tag{5.121}$$

Differentiating $\mathbf{r}$ and applying the vec operator gives

$$d\text{vec}\,\mathbf{r} = \left[\left(\mathbf{NUNe}_j\right)^\mathsf{T} \otimes \mathbf{I}\right]d\text{vec}\,\mathbf{F} + \left[\left(\mathbf{UNe}_j\right)^\mathsf{T} \otimes \mathbf{F}\right]d\text{vec}\,\mathbf{N}$$
$$+ \left[\left(\mathbf{Ne}_j\right)^\mathsf{T} \otimes \mathbf{FN}\right]d\text{vec}\,\mathbf{U} + \left[\mathbf{e}_j^\mathsf{T} \otimes \mathbf{fNU}\right]d\text{vec}\,\mathbf{N} \tag{5.122}$$

Whew!

Finally, substituting (5.119), (5.121) and (5.122) into (5.117), we obtain

$$\frac{d\boldsymbol{\mu}^{(j)}}{d\boldsymbol{\theta}^\mathsf{T}} = -\left(\mathbf{r}^\mathsf{T} \otimes \mathbf{I}\right)\left(\mathbf{X}^{-1} \otimes \mathbf{X}^{-1}\right)\mathcal{D}\left(\text{vec}\,\mathbf{I}\right)$$
$$\times \left[\left(\mathbf{1e}_j^\mathsf{T}\mathbf{N}^\mathsf{T} \otimes \mathbf{I}\right)\frac{d\text{vec}\,\mathbf{F}}{d\boldsymbol{\theta}^\mathsf{T}} + \left(\mathbf{1e}_j \otimes \mathbf{F}\right)\frac{d\text{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^\mathsf{T}}\right]$$
$$+ \left\{\left[\left(\mathbf{NUNe}_j\right)^\mathsf{T} \otimes \mathbf{I}\right]\frac{d\text{vec}\,\mathbf{F}}{d\boldsymbol{\theta}^\mathsf{T}} + \left[\left(\mathbf{UNe}_j\right)^\mathsf{T} \otimes \mathbf{F}\right]\frac{d\text{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^\mathsf{T}}\right.$$
$$\left.+ \left[\left(\mathbf{Ne}_j\right)^\mathsf{T} \otimes \mathbf{FN}\right]\frac{d\text{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^\mathsf{T}} + \left[\mathbf{e}_j^\mathsf{T} \otimes \mathbf{FNU}\right]\frac{d\text{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^\mathsf{T}}\right\}. \tag{5.123}$$

This may be an impressive formula, but it is straightforward to compute, given the derivatives of $\mathbf{U}$, $\mathbf{F}$, and $\mathbf{N}$ with respect to $\boldsymbol{\theta}$.

Bibliography

Bacaër, N. 2009. Periodic matrix population models: growth rate, basic reproduction number, and entropy. Bulletin of Mathematical Biology **71**:1781–1792.

Carey, J. R. 2003. Longevity: The biology and demography of life span. Princeton University Press, Princeton, New Jersey, USA.

Caswell, H. 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. Theoretical Population Biology **14**:215–230.

Caswell, H. 1989. Matrix Population Models: Construction, Analysis, and Interpretation. 1st edition. Sinauer Associates, Sunderland, MA, USA.

Caswell, H. 2001. Matrix Population Models: Construction, Analysis, and Interpretation. 2nd edition. Sinauer Associates, Sunderland, MA.

Caswell, H., 2006. Applications of Markov chains in demography. Pages 319–334 *in* MAM2006: Markov Anniversary Meeting. Boson Books, Raleigh, North Carolina.

Caswell, H. 2009. Stage, age and individual stochasticity in demography. Oikos **118**:1763–1782.

Caswell, H. 2011. Beyond R_0: Demographic models for variability of lifetime reproductive output. PloS ONE **6**:e20809.

Caswell, H., A. M. de Roos, R. M. Nisbet, and S. Tuljapurkar, 1997. Structured population models: many methods, a few general principles. Pages 3–18 *in* Structured population models in marine, terrestrial and freshwater systems. Chapman and Hall, New York, New York, USA.

Caswell, H., C. de Vries, N. Hartemink, G. Roth, and S. F. van Daalen. 2018. Age×stage-classified demographic analysis: a comprehensive approach. Ecological Monographs **88**:560–584.

Caswell, H., and M. Fujiwara. 2004. Beyond survival estimation: mark-recapture, matrix population models, and population dynamics. Animal Biodiversity and Conservation **27**:471–488.

Caswell, H., M. Fujiwara, and S. Brault. 1999. Declining survival probability threatens the North Atlantic right whale. Proceedings of the National Academy of Sciences **96**:3308–3313.

Caswell, H., and T. N. Kaye. 2001. Stochastic demography and conservation of an endangered perennial plant (*Lomatium bradshawii*) in a dynamic fire regime. Advances in Ecological Research **32**:1–51.

Caswell, H., and F. A. Kluge. 2015. Demography and the statistics of lifetime economic transfers under individual stochasticity. Demographic Research **32**:563–588.

Caswell, H., and N. Ouellette, 2018. Cause-of-death analysis: matrix formulation and sensitivity analysis. In prep.

Caswell, H., and R. Salguero-Gómez. 2013. Age, stage and senescence in plants. Journal of Ecology **101**:585–595.

Caswell, H., and Y. Vindenes. 2018. Demographic variance in heterogeneous populations: Matrix models and sensitivity analysis. Oikos **127**:648–663.

Clutton-Brock, T. H. 1988. Reproductive success: studies of individual variation in contrasting breeding systems. University of Chicago Press, Chicago, Illinois.

Coale, A. J. 1972. The growth and structure of human populations: a mathematical approach. Princeton University Press, Princeton, New Jersey.

Cochran, M. E., and S. Ellner. 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. Ecological Monographs **62**:345–364.

Crone, M. J., and S. D. Kraus. 1990. Right whale (*Eubalaena glacialis*) in the Western North Atlantic: a catalog of identied individuals. New England Aquarium, Boston, Massachusetts.

Crow, J. F. 1958. Some possibilities for measuring selection intensities in man. Human Biology **30**:1–13.

Cushing, J. M., and A. S. Ackleh. 2012. A net reproductive number for periodic matrix models. Journal of Biological Dynamics **6**:166–188.

Cushing, J. M., and Y. Zhou. 1994. The net reproductive value and stability in matrix population models. Natural Resources Modeling **8**:297–333.

de Camino-Beck, T., and M. A. Lewis. 2007. A new method for calculating net reproductive rate from graph reduction with applications to the control of invasive species. Bulletin of Mathematical Biology **69**:1341–1354.

de Camino-Beck, T., and M. A. Lewis. 2008. On net reproductive rate and the timing of reproductive output. American Naturalist **172**:128–139.

DeAngelis, D. L., and L. J. Gross. 1992. Individual based models and approaches in ecology. Chapman and Hall, New York, New York.

Diekmann, O., J. A. P. Heesterbeek, and J. A. J. Metz. 1990. On the definition and the computation of the basic reproduction number R_0 in models for infectious diseases in heterogeneous population. Journal of Mathematical Biology **28**:365–382.

Edwards, R. D. 2011. Changes in world inequality in length of life: 1970–2000. Population and Development Review **37**:499–528.

Edwards, R. D., and S. Tuljapurkar. 2005. Inequality in life spans and a new perspective on mortality convergence across industrialized countries. Population and Development Review **31**:645–674.

Ellner, S. P., D. Z. Childs, and M. Rees. 2016. Data-driven modelling of structured populations. Springer, New York, New York.

Endler, J. A. 1986. Natural selection in the wild. Princeton University Press, Princeton, New Jersey.

Feichtinger, G. 1971. Stochastische Modelle demographischer Prozesse. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin.

Feichtinger, G. 1973. Markovian models for some demographic processes. Statistiche Hefte **14**:310–334.

Ferriére, R., and M. Gatto. 1993. Chaotic population dynamics can result from natural selection. Proceedings of the Royal Society of London B **251**:33–38.

Frühwirth-Schnatter, S. 2006. Finite mixture and Markov switching models. Springer-Verlag, New York, New York.

Fujiwara, M., and H. Caswell. 2001. Demography of the endangered North Atlantic right whale. Nature **414**:537–541.

Fujiwara, M., and H. Caswell. 2002. Estimating population projection matrices from multi-stage mark-recapture data. Ecology **83**:3257–3265.

Goldman, N. 1994. Social factors and health: the causation-selection issue revisited. Proceedings of the National Academy of Sciences USA **91**:1251–1255.

Hamilton, W. D. 1966. The moulding of senescence by natural selection. Journal of Theoretical Biology **12**:12–45.

Hartemink, N., and H. Caswell. 2018. Variance in animal longevity: contributions of heterogeneity and stochasticity. Population Ecology **60**:89–99.

Hartemink, N., T. I. Missov, and H. Caswell. 2017. Stochasticity, heterogeneity, and variance in longevity in human populations. Theoretical Population Biology **114**:107–117.

Henderson, H. V., and S. R. Searle. 1981. The vec-permutation matrix, the vec operator and Kronecker products: a review. Linear and Multilinear Algebra **9**:271–288.

Hernandez-Suarez, C., J. Rabinovich, and K. Hernandez. 2012. The long-run distribution of births across environments under environmental stochasticity and its use in the calculation of unconditional life-history parameters. Theoretical Population Biology **82**:264–274.

Hoem, J. M. 1969. Markov chain models in life insurance. Sonderdruck aus Blatter der Deutschen Gelellschaft fur Versicherungsmathematik **9**:91–107.

Horvitz, C. C., and D. W. Schemske, 1986. Seed dispersal and environmental heterogeneity in a neotropical herb: a model of population and patch dynamics. Pages 169–186 *in* A. Estrada and T. H. Fleming, editors. Frugivores and seed dispersal. Dr. W. Junk, Dordrecht, Netherlands.

Horvitz, C. C., and S. Tuljapurkar. 2008. Stage dynamics, period survival, and mortality plateaus. American Naturalist **172**:203–215.

Howard, R. A. 1960. Dynamic programming and Markov processes. Wiley, New York, New York.

Hunter, C. M., and H. Caswell. 2005*a*. Selective harvest of sooty shearwater chicks: effects on population dynamics and sustainability. Journal of Animal Ecology **74**:589–600.

Hunter, C. M., and H. Caswell. 2005*b*. The use of the vec-permutation matrix in spatial matrix population models. Ecological Modelling **188**:15–21.

Iosifescu, M. 1980. Finite Markov Processes and Their Applications. Wiley, New York, New York.

IUCN Species Survival Commission. 2001. IUCN Red List categories and criteria: version 3.1. IUCN, Gland, Switzerland and Cambridge, UK.

Jenouvrier, S., L. M. Aubry, C. Barbraud, H. Weimerskirch, and H. Caswell. 2018. Interacting effects of unobserved heterogeneity and individual stochasticity in the life cycle of the Southern fulmar. Journal of Animal Ecology **87**:212–222.

Kaye, T. N., K. L. Pendergrass, K. Finley, and J. B. Kauffman. 2001. The effect of fire on the population viability of an endangered prairie plant. Ecological Applications **11**:1366–1380.

Kaye, T. N., and D. A. Pyke. 2003. The effect of stochastic technique on estimates of population viability from transition matrix models. Ecology pages 1464–1476.

Kemeny, J. G., and J. L. Snell. 1976. Finite Markov Chains. Second edition. Undergraduate Texts in Mathematics, Springer-Verlag, New York, New York, USA.

Keyfitz, N. 1967. Reconciliation of population models: matrix, integral equation and partial fraction. Journal of the Royal Statistical Society **130**:61–83.

Keyfitz, N. 1971. Linkages of intrinsic to age-specific rates. Journal of the American Statistical Association **66**:275–281.

Keyfitz, N., and H. Caswell. 2005. Applied mathematical demography. 3rd edition. Springer, New York, New York.

Kozlowski, J. 1999. Adaptation: a life history perspective. Oikos **86**:185–194.

Kraus, S. D., M. W. Brown, H. Caswell, C. W. Clark, M. Fujiwara, P. K. Hamilton, R. D. Kenney, A. R. Knowlton, S. Landry, C. A. Mayo, et al. 2005. North Atlantic right whales in crisis. Science **309**:561–562.

Kraus, S. D., and R. M. Rolland. 2007. The urban whale: North Atlantic right whales at the crossroads. Harvard University Press, Cambridge, Massachusetts, USA.

Land, K. C., and A. Rogers, editors. 1982. Multidimensional Mathematical Demography. Academic Press, New York, New York, USA.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution pages 1210–1226.

Lebreton, J. 1996. Demographic models for subdivided populations: the renewal equation approach. Theoretical Population Biology **49**:291–313.

LeCorff, J., and C. C. Horvitz. 2005. The contrasting roles of chasmogamous and cleistogamous reproduction in population growth and population spread in an ant-dispersed neotropical herb *Calathea micans* (Marantaceae). Ecological Modelling **188**:41–51.

Lotka, A. J. 1939. Théorie analytique des associations biologiques. Part II. Analyse démographique avec application particulière à l'espèce humaine. Number 780 in Actualités Scientifiques et Industrielles, Hermann et Cie, Paris, France.

Magnus, J. R., and H. Neudecker. 1979. The commutation matrix: some properties and applications. Annals of Statistics **7**:381–394.

Metz, J. A. J., and O. Diekmann. 1986. The dynamics of physiologically structured populations. Springer-Verlag, Berlin, Germany.

Metz, J. A. J., R. M. Nisbet, and S. A. H. Geritz. 1992. How should we define 'fitness' for general ecological scenarios? Trends in Ecology and Evolution **7**:198–202.

Newton, I. 1989. Lifetime reproductive success in birds. Academic Press, San Diego, California.

Nisbet, R. M., and W. S. C. Gurney. 1982. Modelling fluctuating populations. Wiley, New York, New York, USA.

Otto, S. P., and T. Day. 2007. A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press, Princeton, New Jersey, USA.

Pascarella, J. B., and C. C. Horvitz. 1998. Hurricane disturbance and the population dynamics of a tropical understory shrub: megamatrix elasticity analysis. Ecology **79**:547–563.

Pollard, J. H. 1982. The expectation of life and its relationship to mortality. Journal of the Institute of Actuaries **109**:225–240.

Puterman, M. L. 1994. Markov decision processes: Discrete dynamic stochastic programming. John Wiley, New York, New York.

Reeves, R. R., T. D. Smith, and E. A. Josephson, 2007. Near-annihilation of a species: right whaling in the North Atlantic. Pages 39–74 *in* S. D. Kraus and R. M. Rolland, editors. The urban whale: North Atlantic right whales at the crossroads. Harvard University Press, Cambridge, Massachusetts, USA.

Rényi, A. 1970. Probability Theory. North-Holland, Amsterdam.

Rhodes, E. C. 1940. Population mathematics – I. Journal of the Royal Statistical Society **103**:61–89.

Robine, J.-M., C. Jagger, C. D. Mathers, E. M. Crimmins, and R. M. Suzman. 2003. Determining health expectancies. Wiley, Chichester, England.

Rogers, A. 1974. The multiregional net maternity function and multiregional stable growth. Demography **11**:473–481.

Roth, G., and H. Caswell. 2018. Occupancy time in sets of states for demographic models. Theoretical Population Biology **120**:62–77.

Sheskin, T. J. 2010. Markov chains and decision processes for engineers and managers. CRC Press.

Shkolnikov, V. M., E. E. Andreev, and A. Z. Begun. 2003. Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples. Demographic Research **8**:305–358.

Steiner, U. K., and S. Tuljapurkar. 2012. Neutral theory for life histories and individual variability in fitness components. Proceedings of the National Academy of Sciences **109**:4684–4689.

Steiner, U. K., S. Tuljapurkar, and S. H. Orzack. 2010. Dynamic heterogeneity and life history variability in the kittiwake. Journal of Animal Ecology **79**:436–444.

Tuljapurkar, S. 1990. Population Dynamics in Variable Environments. Springer-Verlag, New York, New York, USA.

Tuljapurkar, S., and H. Caswell. 1997. Structured-Population Models in Marine, Terrestrial, and Freshwater Ecosystems. Chapman and Hall, New York, New York.

Tuljapurkar, S., and C. C. Horvitz. 2006. From stage to age in variable environments: life expectancy and survivorship. Ecology **87**:1497–1509.

Tuljapurkar, S., U. K. Steiner, and S. H. Orzack. 2009. Dynamic heterogeneity in life histories. Ecology Letters **12**:93–106.

van Daalen, S., and H. Caswell. 2015. Lifetime reproduction and the second demographic transition: Stochasticity and individual variation. Demographic Research **33**:561–588.

van Daalen, S. F., and H. Caswell. 2017. Lifetime reproductive output: individual stochasticity, variance, and sensitivity analysis. Theoretical Ecology **10**:355–374.

van den Driessche, P., and J. Watmough. 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical Biosciences **180**:29–48.

Van Raalte, A. A., and H. Caswell. 2013. Perturbation analysis of indices of lifespan variability. Demography **50**:1615–1640.

Vaupel, J. W. 1986. How change in age-specific mortality affects life expectancy. Population Studies **40**:147–157.

Vaupel, J. W., and V. Canudas Romo. 2003. Decomposing change in life expectancy: a bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. Demography **40**:201–216.

Willekens, F. 2014. Multistate Analysis of Life Histories with R. Springer, New York, New York.

Wilmoth, J. R., and S. Horiuchi. 1999. Rectangularization revisited: variability of age at death within human populations. Demography **36**:475–495.