## 6.3. Birth and Death Processes

An obvious generalization of the pure birth and pure death processes discussed in Sections 6.1 and 6.2 is to permit $X(t)$ both to increase and to decrease. Thus if at time $t$ the process is in state $n$ it may, after a random sojourn time, move to either of the neighboring states $n+1$ or $n-1$. The resulting *birth and death process* can then be regarded as the continuous time analog of a random walk (Section 3.5.3).

Birth and death processes form a powerful tool in the kit of the stochastic modeler. The richness of the birth and death parameters facilitates modeling a variety of phenomena. At the same time standard methods of analysis are available for determining numerous important quantities such as stationary distributions and mean first passage times. This section and later sections contain several examples of birth and death processes and illustrate how they are used to draw conclusions about phenomena in a variety of disciplines.

### 6.3.1. Postulates

As in the case of the pure birth processes we assume that $X(t)$ is a Markov process on the states $0, 1, 2, \ldots$ and that its transition probabilities $P_{ij}(t)$ are stationary, i.e.,

$$P_{ij}(t) = \Pr\{X(t+s) = j | X(s) = i\} \qquad \text{for all } s \geq 0.$$

In addition we assume that the $P_{ij}(t)$ satisfy

1. $P_{i,i+1}(h) = \lambda_i h + o(h)$ as $h \downarrow 0,\ i \geq 0$;
2. $P_{i,i-1}(h) = \mu_i h + o(h)$ as $h \downarrow 0,\ i \geq 1$;
3. $P_{i,i}(h) = 1 - (\lambda_i + \mu_i)h + o(h)$ as $h \downarrow 0,\ i \geq 0$;
4. $P_{ij}(0) = \delta_{ij}$;
5. $\mu_0 = 0,\ \lambda_0 > 0,\ \mu_i, \lambda_i > 0,\ i = 1, 2, \ldots$.

The $o(h)$ in each case may depend on $i$. The matrix

$$\mathbf{A} = \left\Vert \begin{matrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{matrix} \right\Vert \tag{6.17}$$



is called the *infinitesimal generator* of the process. The parameters $\lambda_i$ and $\mu_i$ are called, respectively, the infinitesimal birth and death rates. In Postulates 1 and 2 we are assuming that if the process starts in state $i$, then in a small interval of time the probabilities of the population increasing or decreasing by 1 are essentially proportional to the length of the interval.

Since the $P_{ij}(t)$ are probabilities we have $P_{ij}(t) \geq 0$ and

$$\sum_{j=0}^{\infty} P_{ij}(t) \leq 1. \tag{6.18}$$

Using the Markov property of the process we may also derive the so-called *Chapman–Kolmogorov equation*

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s). \tag{6.19}$$

This equation states that in order to move from state $i$ to state $j$ in time $t+s$, $X(t)$ moves to some state $k$ in time $t$ and then from $k$ to $j$ in the remaining time $s$. This is the continuous time analog of formula (3.11).

So far we have mentioned only the transition probabilities $P_{ij}(t)$. In order to obtain the probability that $X(t) = n$ we must specify where the process starts or more generally the probability distribution for the initial state. We then have

$$\Pr\{X(t) = n\} = \sum_{i=0}^{\infty} q_i P_{in}(t),$$

where

$$q_i = \Pr\{X(0) = i\}.$$

### 6.3.2. Sojourn Times

With the aid of the preceding assumptions we may calculate the distribution of the random variable $S_i$ which is the sojourn time of $X(t)$ in state $i$; that is, given that the process is in state $i$, what is the distribution of the time $S_i$ until it first leaves state $i$? If we let

$$\Pr\{S_i \geq t\} = G_i(t)$$

it follows easily by the Markov property that as $h \downarrow 0$

$$\begin{aligned} G_i(t+h) = G_i(t)G_i(h) &= G_i(t)[P_{ii}(h) + o(h)] \\ &= G_i(t)[1 - (\lambda_i + \mu_i)h] + o(h) \end{aligned}$$

or

$$\frac{G_i(t+h) - G_i(t)}{h} = -(\lambda_i + \mu_i)G_i(t) + o(1),$$

so that

$$G_i'(t) = -(\lambda_i + \mu_i)G_i(t). \tag{6.20}$$

If we use the conditions $G_i(0) = 1$ the solution of this equation is

$$G_i(t) = \exp[-(\lambda_i + \mu_i)t];$$

i.e., $S_i$ follows an exponential distribution with mean $(\lambda_i + \mu_i)^{-1}$. The proof presented here is not quite complete, since we have used the intuitive relationship

$$G_i(h) = P_{ii}(h) + o(h)$$

without a formal proof.

According to Postulates 1 and 2, during a time duration of length $h$ a transition occurs from state $i$ to $i+1$ with probability $\lambda_i h + o(h)$ and from state $i$ to $i-1$ with probability $\mu_i h + o(h)$. It follows intuitively that, given that a transition occurs at time $t$, the probability that this transition is to state $i+1$ is $\lambda_i/(\lambda_i + \mu_i)$ and to state $i-1$ is $\mu_i/(\lambda_i + \mu_i)$. The rigorous demonstration of this result is beyond the scope of this book.

It leads to an important characterization of a birth and death process, however, wherein the description of the motion of $X(t)$ is as follows: The process sojourns in a given state $i$ for a random length of time whose distribution function is an exponential distribution with parameter $(\lambda_i + \mu_i)$. When leaving state $i$ the process enters either state $i+1$ or state $i-1$ with probabilities $\lambda_i/(\lambda_i + \mu_i)$ and $\mu_i/(\lambda_i + \mu_i)$, respectively. The motion is analogous to that of a random walk except that transitions occur at random times rather than at fixed time periods.



The traditional procedure for constructing birth and death processes is to prescribe the birth and death parameters $\{\lambda_i, \mu_i\}_{i=0}^{\infty}$ and build the path structure by utilizing the preceding description concerning the waiting times and the conditional transition probabilities of the various states. We determine realizations of the process as follows. Suppose $X(0) = i$; the particle spends a random length of time, exponentially distributed with parameter $(\lambda_i + \mu_i)$, in state $i$ and subsequently moves with probability $\lambda_i/(\lambda_i + \mu_i)$ to state $i+1$ and with probability $\mu_i/(\lambda_i + \mu_i)$ to state $i-1$. Next the particle sojourns a random length of time in the new state and then moves to one of its neighboring states, and so on. More specifically, we observe a value $t_1$ from the exponential distribution with parameter $(\lambda_i + \mu_i)$ that fixes the initial sojourn time in state $i$. Then we toss a coin with probability of heads $p_i = \lambda_i/(\lambda_i + \mu_i)$. If heads (tails) appears we move the particle to state $i+1$ $(i-1)$. In state $i+1$ we observe a value $t_2$ from the exponential distribution with parameter $(\lambda_{i+1} + \mu_{i+1})$ that fixes the sojourn time in the second state visited. If the particle at the first transition enters state $i-1$, the subsequent sojourn time $t_2'$ is an observation from the exponential distribution with parameter $(\lambda_{i-1} + \mu_{i-1})$. After the second wait is completed, a Bernoulli trial is performed that chooses the next state to be visited, and the process continues in the same way.

A typical outcome of these sampling procedures determines a realization of the process. Its form could be

$$X(t) = \begin{cases} i, & 0 < t < t_1, \\ i+1, & t_1 < t < t_1 + t_2, \\ i, & t_1 + t_2 < t < t_1 + t_2 + t_3. \\ \vdots & \end{cases}$$

Thus by sampling from exponential and Bernoulli distributions appropriately, we construct typical sample paths of the process. Now it is possible to assign to this set of paths (realizations of the process) a probability measure in a consistent way so that $P_{ij}(t)$ is determined satisfying (6.18) and (6.19). This result is rather deep and its rigorous discussion is beyond the level of this book. The process obtained in this manner is called the minimal process associated with the infinitesimal matrix $\mathbf{A}$ defined in (6.17).

The preceding construction of the minimal process is fundamental since the infinitesimal parameters need not determine a unique stochastic process obeying (6.18), (6.19), and Postulates 1 through 5 of Section 6.3.1. In fact there could be several Markov processes that possess the same infinitesimal generator. Fortunately, such complications do not arise in the modeling of common phenomena. In the special case of birth and death processes for which $\lambda_0 > 0$, a sufficient condition that there exists a unique Markov process with transition probability function $P_{ij}(t)$ for which the infinitesimal relations (6.18) and (6.19) hold is that

$$\sum_{n=0}^{\infty} \frac{1}{\lambda_n \theta_n} \sum_{k=0}^{n} \theta_k = \infty, \tag{6.21}$$

where

$$\theta_0 = 1, \qquad \theta_n = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n}, \qquad n = 1, 2, \ldots.$$

In most practical examples of birth and death processes the condition (6.21) is met and the birth and death process associated with the prescribed parameters is uniquely determined.

### 6.3.3. Differential Equations of Birth and Death Processes

As in the case of the pure birth and pure death processes the transition probabilities $P_{ij}(t)$ satisfy a system of differential equations known as the backward Kolmogorov differential equations. These are given by

$$\begin{aligned} P'_{0j}(t) &= -\lambda_0 P_{0j}(t) + \lambda_0 P_{1j}(t), \\ P'_{ij}(t) &= \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i) P_{ij}(t) + \lambda_i P_{i+1,j}(t), \qquad i \geq 1, \end{aligned} \tag{6.22}$$

and the boundary condition $P_{ij}(0) = \delta_{ij}$.

To derive these we have, from Equation (6.19),

$$\begin{aligned} P_{ij}(t+h) &= \sum_{k=0}^{\infty} P_{ik}(h) P_{kj}(t) \\ &= P_{i,i-1}(h) P_{i-1,j}(t) + P_{i,i}(h) P_{ij}(t) + P_{i,i+1}(h) P_{i+1,j}(t) \\ &\quad + \Sigma'_k P_{ik}(h) P_{kj}(t), \end{aligned} \tag{6.23}$$

where the last summation is over all $k \neq i-1,\ i+1$. Using Postulates 1,



2, and 3 of Section 6.3.1 we obtain

$$\begin{aligned} \Sigma'_k P_{ik}(h) P_{kj}(t) &\leq \Sigma'_k P_{ik}(h) \\ &= 1 - [P_{i,i}(h) + P_{i,i-1}(h) + P_{i,i+1}(h)] \\ &= 1 - [1 - (\lambda_i + \mu_i)h + o(h) + \mu_i h + o(h) + \lambda_i h + o(h)] \\ &= o(h), \end{aligned}$$

so that

$$P_{ij}(t+h) = \mu_i h P_{i-1,j}(t) + [1 - (\lambda_i + \mu_i)h] P_{ij}(t) + \lambda_i h P_{i+1,j}(t) + o(h).$$

Transposing the term $P_{ij}(t)$ to the left-hand side and dividing the equation by $h$, we obtain, after letting $h \downarrow 0$,

$$P'_{ij}(t) = \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i) P_{ij}(t) + \lambda_i P_{i+1,j}(t).$$

The backward equations are deduced by decomposing the time interval $(0, t+h)$, where $h$ is positive and small, into the two periods

$$(0, h), \qquad (h, t+h),$$

and examining the transition in each period separately. In this sense the backward equations result from a "first step analysis," the first step being over the short time interval of duration $h$.

A different result arises from a "last step analysis" which proceeds by splitting the time interval $(0, t+h)$ into the two periods

$$(0, t), \qquad (t, t+h)$$

and adapting the preceding reasoning. From this viewpoint, under more stringent conditions, we can derive a further system of differential equations

$$\begin{aligned} P'_{i0}(t) &= -\lambda_0 P_{i,0}(t) + \mu_1 P_{i,1}(t), \\ P'_{ij}(t) &= \lambda_{j-1} P_{i,j-1}(t) - (\lambda_j + \mu_j) P_{ij}(t) + \mu_{j+1} P_{i,j+1}(t), \qquad j \geq 1, \end{aligned} \tag{6.24}$$

with the same initial condition $P_{ij}(0) = \delta_{ij}$. These are known as the forward Kolmogorov differential equations. To derive these equations we interchange $t$ and $h$ in Equation (6.23), and under stronger assumptions in addition to Postulates 1, 2, and 3 it can be shown that the last term is again $o(h)$. The remainder of the argument is the same as before. The usefulness of the differential equations will become apparent in the examples that we study in the next section.

A sufficient condition that (6.24) hold is that $[P_{kj}(h)]/h = o(1)$ for $k \neq j, j-1, j+1$ where the $o(1)$ term apart from tending to zero is uniformly bounded with respect to $k$ for fixed $j$ as $h \to 0$. In this case it can be proved that $\Sigma'_k P_{ik}(t)P_{kj}(h) = o(h)$.

**Example** *Linear Growth with Immigration* A birth and death process is called a linear growth process if $\lambda_n = \lambda n + a$ and $\mu_n = \mu n$ with $\lambda > 0$, $\mu > 0$, and $a > 0$. Such processes occur naturally in the study of biological reproduction and population growth. If the state $n$ describes the current population size, then the average instantaneous rate of growth is $\lambda n + a$. Similarly, the probability of the state of the process decreasing by one after the elapse of a small duration $h$ of time is $\mu nh + o(h)$. The factor $\lambda n$ represents the natural growth of the population owing to its current size while the second factor $a$ may be interpreted as the infinitesimal rate of increase of the population due to an external source such as immigration. The component $\mu n$ which gives the mean infinitesimal death rate of the present population possesses the obvious interpretation.

If we substitute the above values of $\lambda_n$ and $\mu_n$ in (6.24) we obtain

$$P'_{i0}(t) = -aP_{i0}(t) + \mu P_{i1}(t),$$

$$P'_{ij}(t) = [\lambda(j-1) + a]P_{i,j-1}(t) - [(\lambda + \mu)j + a]P_{ij}(t)$$
$$+ \mu(j+1)P_{i,j+1}(t), \qquad j \geq 1.$$

Now if we multiply the $j$th equation by $j$ and sum, it follows that the expected value

$$E[X(t)] = M(t) = \sum_{j=1}^{\infty} jP_{ij}(t)$$

satisfies the differential equation

$$M'(t) = a + (\lambda - \mu)M(t),$$

with initial condition $M(0) = i$, if $X(0) = i$. The solution of this equation is

$$M(t) = at + i \qquad \text{if } \lambda = \mu,$$



and

$$M(t) = \frac{a}{\lambda - \mu}\{e^{(\lambda-\mu)t} - 1\} + ie^{(\lambda-\mu)t} \qquad \text{if } \lambda \neq \mu. \tag{6.25}$$

The second moment or variance may be calculated in a similar way. It is interesting to note that $M(t) \to \infty$ as $t \to \infty$ if $\lambda \geq \mu$, while if $\lambda < \mu$ the mean population size for large $t$ is approximately

$$\frac{a}{\mu - \lambda}.$$

These results suggest that in the second case, wherein $\lambda < \mu$, the population stabilizes in the long run in some form of statistical equilibrium. Indeed it can be shown that a limiting probability distribution $\{\pi_j\}$ exists for which $\lim_{t\to\infty} P_{ij}(t) = \pi_j$, $j = 0, 1, \ldots$. Such limiting distributions for general birth and death processes are the subject of the next section.

## Exercises 6.3

**1.** Particles are emitted by a radioactive substance according to a Poisson process of rate $\lambda$. Each particle exists for an exponentially distributed length of time, independent of the other particles, before disappearing. Let $X(t)$ denote the number of particles alive at time $t$. Argue that $X(t)$ is a birth and death process, and determine the parameters.

**2.** Patients arrive at a hospital emergency room according to a Poisson process of rate $\lambda$. The patients are treated by a single doctor on a first come, first served basis. The doctor treats patients more quickly when the number of patients waiting is higher. An industrial engineering time study suggests that the mean patient treatment time when there are $k$ patients in the system is of the form $m_k = \alpha - \beta k/(k+1)$, where $\alpha$ and $\beta$ are constants with $\alpha > \beta > 0$. Let $N(t)$ be the number of patients in the system at time $t$ (waiting and being treated). Argue that $N(t)$ might be modeled as a birth and death process with parameters $\lambda_k = \lambda$ for $k = 0, 1, \ldots$ and $\mu_k = 1/m_k$ for $k = 1, 2, \ldots$. State explicitly any necessary assumptions.

two machines can operate at any time. The amount of time that an operating machine works before breaking down is exponentially distributed with mean 5. The amount of time that it takes a single repairman to fix a machine is exponentially distributed with mean 4. Only one repairman can work on a failed machine at any given time. Let $X(t)$ be the number of machines in operating condition at time $t$.

(a) Calculate the long run probability distribution for $X(t)$.
(b) If an operating machine produces 100 units of output per hour, what is the long run output per hour of the system?

**14.** A birth and death process has parameters

$$\lambda_k = \alpha(k+1) \qquad \text{for } k = 0, 1, 2, \ldots,$$

and

$$\mu_k = \beta(k+1) \qquad \text{for } k = 1, 2, \ldots.$$

Assuming that $\alpha < \beta$, determine the limiting distribution of the process. Simplify your answer as much as possible.

## 6.5. Birth and Death Processes with Absorbing States

Birth and death processes in which $\lambda_0 = 0$ arise frequently and are correspondingly important. For these processes, the zero state is an absorbing state. A central example is the linear growth birth and death process without immigration (cf. Section 6.3.3). In this case $\lambda_n = n\lambda$ and $\mu_n = n\mu$. Since growth of the population results exclusively from the existing population, it is clear that when the population size becomes zero it remains zero thereafter; i.e., 0 is an absorbing state.

### 6.5.1. Probability of Absorption into State 0

It is of interest to compute the probability of absorption into state 0 starting from state $i$ $(i \geq 1)$. This is not, a priori, a certain event since



conceivably the particle (i.e., state variable) may wander forever among the states $(1, 2, \ldots)$ or possibly drift to infinity.

Let $u_i$ $(i = 1, 2, \ldots)$ denote the probability of absorption into state 0 from the initial state $i$. We can write a recursion formula for $u_i$ by considering the possible states after the first transition. We know that the first transition entails the movements

$$i \to i+1 \qquad \text{with probability } \frac{\lambda_i}{\mu_i + \lambda_i},$$

$$i \to i-1 \qquad \text{with probability } \frac{\mu_i}{\mu_i + \lambda_i}.$$

Invoking the familiar first step analysis we directly obtain

$$u_i = \frac{\lambda_i}{\mu_i + \lambda_i} u_{i+1} + \frac{\mu_i}{\mu_i + \lambda_i} u_{i-1}, \qquad i \geq 1, \tag{6.37}$$

where $u_0 = 1$.

Another method for deriving (6.37) is to consider the "embedded random walk" associated with a given birth and death process. Specifically we examine the birth and death process only at the transition times. The discrete time Markov chain generated in this manner is denoted by $\{Y_n\}_{n=0}^{\infty}$, where $Y_0 = X_0$ is the initial state and $Y_n$ $(n \geq 1)$ is the state at the $n$th transition. Obviously, the transition probability matrix has the form

$$\mathbf{P} = \left\| \begin{matrix} 1 & 0 & 0 & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & \cdots \\ 0 & q_2 & 0 & p_2 & \cdots \\ \vdots & \vdots & & & \end{matrix} \right\|,$$

where

$$p_i = \frac{\lambda_i}{\lambda_i + \mu_i} = 1 - q_i \qquad \text{for } i \geq 1.$$

The probability of absorption into state 0 for the embedded random walk is the same as for the birth and death processes since both processes execute the same transitions. A closely related problem (gambler's ruin) for a random walk was examined in Section 3.6.1.

We turn to the task of solving (6.37) subject to the conditions $u_0 = 1$ and $0 \le u_i \le 1$ $(i \ge 1)$. Rewriting (6.37) we have

$$(u_{i+1} - u_i) = \frac{\mu_i}{\lambda_i}(u_i - u_{i-1}), \qquad i \ge 1.$$

Defining $v_i = u_{i+1} - u_i$, we obtain

$$v_i = \frac{\mu_i}{\lambda_i} v_{i-1}, \qquad i \ge 1.$$

Iteration of the last relation yields the formula $v_i = \rho_i v_0$, where

$$\rho_0 = 1 \quad \text{and} \quad \rho_i = \frac{\mu_1 \mu_2 \cdots \mu_i}{\lambda_1 \lambda_2 \cdots \lambda_i} \qquad \text{for } i \ge 1,$$

and with $u_{i+1} - u_i = v_i$, then

$$u_{i+1} - u_i = v_i = \rho_i v_0 = \rho_i(u_1 - u_0) = \rho_i(u_1 - 1) \qquad \text{for } i \ge 1.$$

Summing these last equations from $i = 1$ to $i = m - 1$ we have

$$u_m - u_1 = (u_1 - 1) \sum_{i=1}^{m-1} \rho_i, \qquad m > 1. \tag{6.38}$$

Since $u_m$, by its very meaning, is bounded by 1 we see that if

$$\sum_{i=1}^{\infty} \rho_i = \infty \tag{6.39}$$

then necessarily $u_1 = 1$ and $u_m = 1$ for all $m \ge 2$. In other words, if (6.39) holds then ultimate absorption into state 0 is certain from any initial state.

Suppose $0 < u_1 < 1$; then, of course,

$$\sum_{i=1}^{\infty} \rho_i < \infty.$$

Obviously, $u_m$ is decreasing in $m$ since passing from state $m$ to state 0 requires entering the intermediate states in the intervening time. Furthermore, it can be shown that $u_m \to 0$ as $m \to \infty$. Now letting $m \to \infty$ in (6.38) permits us to solve for $u_1$; thus



$$u_1 = \frac{\sum_{i=1}^{\infty} \rho_i}{1 + \sum_{i=1}^{\infty} \rho_i}$$

and then from (6.38) we obtain

$$u_m = \frac{\sum_{i=m}^{\infty} \rho_i}{1 + \sum_{i=1}^{\infty} \rho_i}, \qquad m \ge 1.$$

### 6.5.2. Mean Time Until Absorption

Consider the problem of determining the mean time until absorption, starting from state $m$.

We assume that condition (6.39) holds so that absorption is certain. Notice that we cannot reduce our problem to a consideration of the embedded random walk since the actual time spent in each state is relevant for the calculation of the mean absorption time.

Let $w_i$ be the mean absorption time starting from state $i$ (this could be infinite). Considering the possible states following the first transition, instituting a first step analysis, and recalling the fact that the mean waiting time in state $i$ is $(\lambda_i + \mu_i)^{-1}$ (it is actually exponentially distributed with parameter $\lambda_i + \mu_i$), we deduce the recursion relation

$$w_i = \frac{1}{\lambda_i + \mu_i} + \frac{\lambda_i}{\lambda_i + \mu_i} w_{i+1} + \frac{\mu_i}{\lambda_i + \mu_i} w_{i-1}, \qquad i \ge 1, \tag{6.40}$$

and where $w_0 = 0$. Letting $z_i = w_i - w_{i+1}$ and rearranging (6.40) leads to

$$z_i = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i} z_{i-1}, \qquad i \ge 1. \tag{6.41}$$

Iterating this relation gives

$$z_1 = \frac{1}{\lambda_1} + \frac{\mu_1}{\lambda_1} z_0,$$

$$z_2 = \frac{1}{\lambda_2} + \frac{\mu_2}{\lambda_2} z_1 = \frac{1}{\lambda_2} + \frac{\mu_2}{\lambda_2 \lambda_1} + \frac{\mu_2 \mu_1}{\lambda_2 \lambda_1} z_0,$$

$$z_3 = \frac{1}{\lambda_3} + \frac{\mu_3}{\lambda_3 \lambda_2} + \frac{\mu_3 \mu_2}{\lambda_3 \lambda_2 \lambda_1} + \frac{\mu_3 \mu_2 \mu_1}{\lambda_3 \lambda_2 \lambda_1} z_0,$$

and finally

$$z_m = \sum_{i=1}^{m} \frac{1}{\lambda_i} \prod_{j=i+1}^{m} \frac{\mu_j}{\lambda_j} + \left(\prod_{j=1}^{m} \frac{\mu_j}{\lambda_j}\right) z_0.$$

(The product $\prod_{m+1}^{m} \mu_j/\lambda_j$ is interpreted as 1.) Using the notation

$$\rho_0 = 1 \quad \text{and} \quad \rho_i = \frac{\mu_1 \mu_2 \cdots \mu_i}{\lambda_1 \lambda_2 \cdots \lambda_i}, \qquad i \geq 1,$$

the expression for $z_m$ becomes

$$z_m = \sum_{i=1}^{m} \frac{1}{\lambda_i} \frac{\rho_m}{\rho_i} + \rho_m z_0,$$

or, since $z_m = w_m - w_{m+1}$ and $z_0 = w_0 - w_1 = -w_1$, then

$$\frac{1}{\rho_m}(w_m - w_{m+1}) = \sum_{i=1}^{m} \frac{1}{\lambda_i \rho_i} - w_1. \tag{6.42}$$

If $\sum_{i=1}^{\infty} (1/\lambda_i \rho_i) = \infty$, then inspection of (6.42) reveals that necessarily $w_1 = \infty$. Indeed, it is probabilistically evident that $w_m < w_{m+1}$ for all $m$ and this property would be violated for $m$ large if we assume to the contrary that $w_1$ is finite.

Now suppose $\sum_{i=1}^{\infty} (1/\lambda_i \rho_i) < \infty$; then letting $m \to \infty$ in (6.42) gives

$$w_1 = \sum_{i=1}^{\infty} \frac{1}{\lambda_i \rho_i} - \lim_{m \to \infty} \frac{1}{\rho_m}(w_m - w_{m+1}).$$

It is more involved but still possible to prove that

$$\lim_{m \to \infty} \frac{1}{\rho_m}(w_m - w_{m+1}) = 0,$$

and then

$$w_1 = \sum_{i=1}^{\infty} \frac{1}{\lambda_i \rho_i}.$$

We summarize the discussion of this section in the following theorem:

**Theorem 6.1** *Consider a birth and death process with birth and death parameters $\lambda_n$ and $\mu_n$, $n \geq 1$, where $\lambda_0 = 0$ so that 0 is an absorbing state. The probability of absorption into state 0 from the initial state $m$ is*

$$u_m = \begin{cases} \dfrac{\sum_{i=m}^{\infty} \rho_i}{1 + \sum_{i=1}^{\infty} \rho_i} & \text{if } \sum_{i=1}^{\infty} \rho_i < \infty, \\ 1 & \text{if } \sum_{i=1}^{\infty} \rho_i = \infty. \end{cases} \tag{6.43}$$

*The mean time to absorption is*

$$w_m = \begin{cases} \infty & \text{if } \sum_{i=1}^{\infty} \dfrac{1}{\lambda_i \rho_i} = \infty, \\ \sum_{i=1}^{\infty} \dfrac{1}{\lambda_i \rho_i} + \sum_{k=1}^{m-1} \rho_k \sum_{j=k+1}^{\infty} \dfrac{1}{\lambda_j \rho_j} & \text{if } \sum_{i=1}^{\infty} \dfrac{1}{\lambda_i \rho_i} < \infty, \end{cases} \tag{6.44}$$

*where $\rho_0 = 1$ and $\rho_i = (\mu_1 \mu_2 \cdots \mu_i)/(\lambda_1 \lambda_2 \cdots \lambda_i)$.*



**Example** *Population Processes* Consider the linear growth birth and death process without immigration (cf. Section 6.3.3) for which $\mu_n = n\mu$ and $\lambda_n = n\lambda$, $n = 0, 1, \ldots$. During a short time interval of length $h$, a *single individual* in the population dies with probability $\mu h + o(h)$ and gives birth to a new individual with probability $\lambda h + o(h)$, and thus $\mu > 0$ and $\lambda > 0$ represent the *individual* death and birth rates, respectively.

Substitution of $a = 0$ and $i = m$ in Equation (6.25) determines the mean population size at time $t$ for a population starting with $X(0) = m$ individuals. This mean population size is $M(t) = me^{(\lambda - \mu)t}$, exhibiting exponential growth or decay according as $\lambda > \mu$ or $\lambda < \mu$.

Let us now examine the extinction phenomenon and determine the probability that the population eventually dies out. This phenomenon corresponds to absorption in state 0 for the birth and death process.

When $\lambda_n = n\lambda$ and $\mu_n = n\mu$, a direct calculation yields $\rho_i = (\mu/\lambda)^i$ and then

$$\sum_{i=m}^{\infty} \rho_i = \sum_{i=m}^{\infty} (\mu/\lambda)^i = \begin{cases} \dfrac{(\mu/\lambda)^m}{1 - (\mu/\lambda)} & \text{when } \lambda > \mu, \\ \infty & \text{when } \lambda \leq \mu. \end{cases}$$

From Theorem 6.1, the probability of eventual extinction starting with $m$

individuals is

$$\Pr\{\text{Extinction}|X(0)=m\}=\begin{cases}(\mu/\lambda)^m & \text{when } \lambda>\mu,\\ 1 & \text{when } \lambda\le\mu.\end{cases} \tag{6.45}$$

When $\lambda=\mu$, the process is sure to vanish eventually. Yet in this case the mean population size remains constant at the initial population level. Similar situations where mean values do not adequately describe population behavior frequently arise when stochastic elements are present.

We turn attention to the mean time to extinction assuming extinction is certain, that is, when $\lambda\le\mu$. For a population starting with a single individual, then, from (6.44) with $m=1$ we determine this mean time to be

$$\begin{aligned}\sum_{i=1}^{\infty}\frac{1}{\lambda_i\rho_i}&=\frac{1}{\lambda}\sum_{i=1}^{\infty}\frac{1}{i}\left(\frac{\lambda}{\mu}\right)^i\\ &=\frac{1}{\lambda}\sum_{i=1}^{\infty}\int_0^{(\lambda/\mu)}x^{i-1}\,dx\\ &=\frac{1}{\lambda}\int_0^{(\lambda/\mu)}\sum_{i=1}^{\infty}x^{i-1}\,dx\\ &=\frac{1}{\lambda}\int_0^{(\lambda/\mu)}\frac{dx}{(1-x)}\\ &=-\frac{1}{\lambda}\ln(1-x)\Big|_0^{(\lambda/\mu)}\\ &=\begin{cases}\dfrac{1}{\lambda}\ln\left(\dfrac{\mu}{\mu-\lambda}\right) & \text{when } \mu>\lambda,\\ \infty & \text{when } \mu=\lambda.\end{cases}\end{aligned} \tag{6.46}$$

When the birth rate $\lambda$ exceeds the death rate $\mu$, a linear growth birth



and death process can, with strictly positive probability, grow without limit. In contrast, many natural populations exhibit density dependent behavior wherein the individual birth rates decrease or the individual death rates increase or both changes occur as the population grows. These changes are ascribed to factors including limited food supplies, increased predation, crowding, and limited nesting sites. Accordingly, we introduce a notion of environmental *carrying capacity* $K$, an upper bound that the population size cannot exceed.

Since all individuals have a chance of dying, with a finite carrying capacity, all populations will eventually become extinct. Our measure of population fitness will be the mean time to extinction, and it is of interest to population ecologists studying colonization phenomena to examine how the capacity $K$, the birth rate $\lambda$, and the death rate $\mu$ affect this mean population lifetime.

The model should have the properties of exponential growth (on the average) for small populations, as well as the ceiling $K$ beyond which the population cannot grow. There are several ways of approaching the population size $K$ and staying there at equilibrium. Since all such models give more or less the same qualitative results, we stipulate the simplest model in which the birth parameters are

$$\lambda_n=\begin{cases}n\lambda & \text{for } n=0,1,\ldots,K-1,\\ 0 & \text{for } n\ge K.\end{cases}$$

Theorem 6.1 yields $w_1$, the mean time to population extinction starting with a single individual, as given by

$$w_1=\sum_{i=1}^{\infty}\frac{1}{\lambda_i\rho_i}=\sum_{i=1}^{\infty}\frac{\lambda_1\lambda_2\cdots\lambda_{i-1}}{\mu_1\mu_2\cdots\mu_i}=\frac{1}{\mu}\sum_{i=1}^{K}\frac{1}{i}\left(\frac{\lambda}{\mu}\right)^{i-1}. \tag{6.47}$$

Equation (6.47) isolates the distinct factors influencing the mean time to population extinction. The first factor is $1/\mu$, the mean lifetime of an individual since $\mu$ is the individual death rate. Thus, the sum in (6.47) represents the mean *generations* or mean lifespans to population extinction, a dimensionless quantity that we denote by

$$M_g=\mu w_1=\sum_{i=1}^{K}\frac{1}{i}\theta^{i-1},\qquad \text{where } \theta=\frac{\lambda}{\mu}. \tag{6.48}$$

Next we examine the influence of the birth-death or reproduction ratio

$\theta = \lambda/\mu$ and the carrying capacity $K$ on the mean time to extinction. Since $\lambda$ represents the individual birth rate and $1/\mu$ is the mean life of a single member in the population, we may interpret the reproduction ratio $\theta = \lambda(1/\mu)$ as the mean number of offspring of an arbitrary individual in the population. Accordingly, we might expect significantly different behavior when $\theta < 1$ as opposed to when $\theta > 1$, and this is indeed the case. A carrying capacity of $K = 100$ is small. When $K$ is of the order of 100 or more, we have the following accurate approximations, their derivations being sketched in Exercises 1 and 2 at the end of this section:

$$M_g \cong \begin{cases} \dfrac{1}{\theta}\ln\left(\dfrac{1}{1-\theta}\right) & \text{for } \theta < 1, \\ .5772157 + \ln K & \text{for } \theta = 1, \\ \dfrac{1}{K}\left(\dfrac{\theta^K}{\theta - 1}\right) & \text{for } \theta > 1. \end{cases} \tag{6.49}$$

The contrast between $\theta < 1$ and $\theta > 1$ is vivid. When $\theta < 1$, the mean generations to extinction $M_g$ is almost independent of carrying capacity $K$ and approaches the asymptotic value $\theta^{-1}\ln(1-\theta)^{-1}$ quite rapidly. When $\theta > 1$, the mean generations to extinction $M_g$ grows exponentially in $K$. Some calculations based on (6.49) are given in Table 6.1.

**Table 6.1** Mean generations to extinction for a population starting with a single parent and where $\theta$ is the reproduction rate and $K$ is the environmental capacity.

| $K$ | $\theta = .8$ | $\theta = 1$ | $\theta = 1.2$ |
|---|---|---|---|
| 10 | 1.96 | 2.88 | 3.10 |
| 100 | 2.01 | 5.18 | 4140899 |
| 1000 | 2.01 | 7.48 | $7.59 \times 10^{76}$ |

**Example** *Sterile Male Insect Control* The screwworm fly, a cattle pest in warm climates, was eliminated from the southeastern United States by the release into the environment of sterilized adult male screwworm flies. When these males, artificially sterilized by radiation, mate with native females, there are no offspring, and in this manner part of the reproductive capacity of the natural population is nullified by their presence. If the sterile males are sufficiently plentiful so as to cause even a small decline in the population level, then this decline accelerates in succeeding generations even if the number of sterile males is maintained at approximately the same level, because the ratio of sterile to fertile males will increase as the natural population drops. Because of this compounding effect, if the sterile male control method works at all, it works to such an extent as to drive the native population to extinction in the area in which it is applied.

Recently, a multibillion dollar effort involving the sterile male technique has been proposed for the control of the cotton boll weevil. In this instance, it was felt that a pretreatment with a pesticide could reduce the natural population size to a level such that the sterile male technique would become effective. Let us examine this assumption, first with a deterministic model, and then in a stochastic setting.

For both models we suppose that sexes are present in equal numbers, that sterile and fertile males are equally competitive, and that a constant number $S$ of sterile males is present in each generation. In the deterministic case, if $N_0$ fertile males are in the parent generation and the $N_0$ fertile females choose mates equally likely from the entire male population, then the fraction $N_0/(N_0 + S)$ of these matings will be with fertile males and will produce offspring. Letting $\theta$ denote the number of offspring of either sex in a fertile mating, we calculate the size $N$ of the next generation according to

$$N_1 = \theta N_0\left(\frac{N_0}{N_0 + S}\right). \tag{6.50}$$

For a numerical example, suppose that there are $N_0 = 100$ fertile males and an equal number of fertile females in the parent generation of the native population, and that $S = 100$ sterile male insects are released. If $\theta = 4$, meaning that a fertile mating produces four males and four females for the succeeding generation, then the number of either sex in the first generation is

$$N_1 = 4(100)\left(\frac{100}{100 + 100}\right) = 200;$$

the population has increased and the sterile male control method has failed.