

Population growth I: birth and death processes

9.1 INTRODUCTION

It is clearly desirable that governments and some businesses be able to predict future human population numbers. Not only are the total numbers of male and female individuals of interest, but also the numbers in certain categories such as age groups. The subject which deals with population numbers and movements is called **demography**.

Some of the data of concern to human demographers is obtained from our filling out census forms. The type of data is exemplified by that in Tables 9.1 and 9.2. In Table 9.1 is given the total population of Australia at various times since 1881 and Table 9.2 contains some data on births and deaths and their rates. Notice the drastic fall in the birth rate in the last few decades compared with an almost steady death rate.

Table 9.1

Time	Population of Australia (thousands)*
3 April 1881	2250.2
5 April 1891	3177.8
31 March 1901	3773.8
3 April 1911	4453.0
4 April 1921	5435.7
30 June 1933	6629.8
30 June 1947	7579.4
30 June 1954	8986.5
30 June 1961	10548.3
30 June 1966	11599.5
30 June 1971	12937.2
30 June 1979	14417.2
30 June 1981	14923.3

*Obtained from Cameron (1982) and Australian Bureau of Statistics.

- (b) If the answer to (a) is yes, compute the equilibrium probability distribution \mathbf{p} .
- (c) If $X_0 = 3$, what is the eventual probability that the position of the particle is 3?
19. In the Markov chain model of random mating with mutation in a population of size N , find \mathbf{P} if $\alpha_1 = \alpha_2 = \alpha \neq 0$. Given an arbitrary initial probability distribution $\mathbf{p}(0)$, find $\mathbf{p}(1)$ and deduce that the stationary distribution is attained in one generation.
20. What will happen in the Markov chain model of random mating with mutation if $\alpha_1 \neq 0$ but $\alpha_2 = 0$?

Table 9.2

Time period	Annual (average) number		Annual rates per thousand	
	Births	Deaths	Births	Deaths
1956-60	222 459	86 488	22.59	8.78
1961-65	232 952	93 465	21.34	8.75
1966-70	240 325	107 263	19.95	8.90
1971-75	253 438	111 216	18.99	8.52
1976	227 810	112 662	16.37	8.10
1977	226 291	108 790	16.08	7.73
1978	224 181	108 425	15.73	7.61

Deterministic model

An accurate mathematical model for the growth of a population would be a very useful thing to have. There is a substantial literature on various models, as exemplified by the books of Bartlett (1960), Keyfitz (1968), Pollard (1973) and Ludwig (1978). A first division of such models is into **deterministic** versus **stochastic** ones. In the former category there are no chance effects.

Let $N(t)$ be the population size at time $t \geq 0$ and assume that the initial population size $N(0) > 0$ is given. The number of individuals at t is a non-negative integer, but it is convenient to assume that $N(t)$ is a differentiable function of time. A simple differential equation for N is obtained by letting there be $b\Delta t$ births and $d\Delta t$ deaths per individual in $(t, t + \Delta t]$ where $b, d \geq 0$ are the **per capita birth and death rates**. Then

$$N(t + \Delta t) = N(t) + N(t)b\Delta t - N(t)d\Delta t$$

and it follows, upon rearranging and taking the limit $\Delta t \rightarrow 0$, that

$$\frac{dN}{dt} = (b - d)N, \quad t > 0.$$

This differential equation is called the **Malthusian growth law** (after Thomas Malthus, whose essay on population appeared in 1798). Its solution is

$$N(t) = N(0)e^{(b-d)t},$$

and once we specify $N(0)$ the population size $N(t)$ is determined for all $t \geq 0$. Three qualitatively different behaviours are possible, depending on the relative magnitudes of b and d . These are illustrated in Fig. 9.1.

It is seen that

$$\lim_{t \rightarrow \infty} N(t) = \begin{cases} 0, & b < d \text{ (exponential decay)} \\ N(0), & b = d \text{ (constant population)} \\ \infty, & b > d \text{ (exponential growth)}. \end{cases}$$

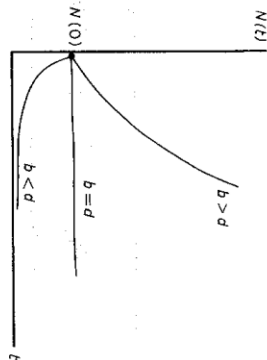


Figure 9.1 Population size as a function of time for various relative magnitudes of the birth and death rates in the Malthusian growth model.

Experience tells us that the times of occurrence of births and deaths are not predictable in real populations. Hence the population sizes $N = \{N(t), t \geq 0\}$ constitute a random process. As observed above, the number of individuals is a non-negative integer so N is a random process with a **discrete state space** $\{0, 1, 2, \dots\}$ and a continuous parameter set $\{t \geq 0\}$. The models we will consider do not contain very many elements of reality. They are nevertheless interesting because they may be solved exactly and provide a starting point for more complicated models.

9.2 SIMPLE POISSON PROCESSES

The random processes discussed in this section are of fundamental importance even though, as will be seen, their use as population growth models is very limited. The following process has a different parameterization from that defined in Chapter 3.

Definition A collection of random variables $N = \{N(t), t \geq 0\}$ is called a **simple Poisson process with intensity λ** if the following hold:

- (i) $N(0) = 0$.
- (ii) For any collection of times $0 \leq t_0 < t_1 < \dots < t_{n-1} < t_n < \infty$ the random variables $N(t_k) - N(t_{k-1}), k = 1, 2, \dots, n$, are mutually independent.
- (iii) For any pair of times $0 \leq t_1 < t_2$, the random variable $N(t_2) - N(t_1)$ is Poisson distributed with parameter $\lambda(t_2 - t_1)$.

Condition (i) is a convenient initialization. The quantity $\Delta N_k = N(t_k) - N(t_{k-1})$ is the change or increment in the process in the time interval

$(t_{k-1}, t_k]$. Condition (ii) says that the increments in disjoint (non-overlapping) time intervals are independent. Condition (iii) gives the probability law of the increments. Then

$$\Pr \{N(t_2) - N(t_1) = k\} = \frac{(\lambda(t_2 - t_1))^k e^{-\lambda(t_2 - t_1)}}{k!}, \quad k = 0, 1, 2, \dots$$

and in particular, choosing $t_1 = 0$ and $t_2 = t > 0$ we have, since $N(0) = 0$,

$$\Pr \{N(t) = k\} = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \quad k = 0, 1, 2, \dots \quad (9.1)$$

The mean and variance of $N(t)$ are therefore

$$E(N(t)) = \text{Var}(N(t)) = \lambda t.$$

Sample paths

What does a typical realization of a simple Poisson process look like? To get some insight, let Δt be a very small time increment and consider the increment

$$\Delta N(t) = N(t + \Delta t) - N(t).$$

The probability law of this increment is

$$\Pr \{\Delta N(t) = k\} = \frac{(\lambda \Delta t)^k e^{-\lambda \Delta t}}{k!} = \begin{cases} 1 - \lambda \Delta t + o(\Delta t), & k = 0 \\ \lambda \Delta t + o(\Delta t), & k = 1 \\ o(\Delta t), & k \geq 2. \end{cases} \quad (9.2)$$

We see therefore that when Δt is very small, $N(t + \Delta t)$ is most likely to be the same as $N(t)$, with probability $\lambda \Delta t$ that it is one larger, and there is a negligible chance that it differs by more than one from $N(t)$. We may conclude that sample paths are right-continuous step functions with discontinuities of magnitude unity – see Fig. 9.2. Note that equation (9.2) can be used as a definition and the probability law of the increments derived from it – see Exercise 2.

The simple Poisson process is called **simple** because all of its jumps have the same magnitude – unity in the above case. This contrasts with **compound Poisson processes** in which jumps may be any of several magnitudes (see for example Parzen, 1962).

To relate the Poisson process to a growth model we imagine that each time a new individual is born, $N(t)$ jumps up by unity. Thus $N(t)$ records the number of births in $(0, t]$, i.e. up to and including time t and $\{N(t), t \geq 0\}$ may be regarded as a birth process. If in Fig. 9.2 we place a cross on the t -axis at each

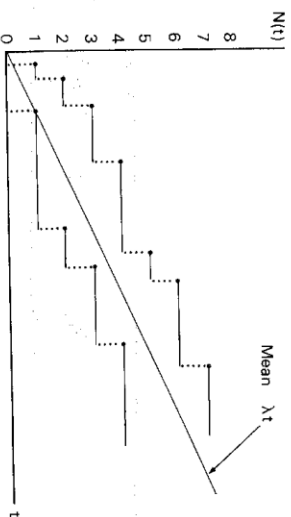


Figure 9.2 Two representative sample paths for a simple Poisson process. The mean value function λt is also indicated.

jump of $N(t)$ we obtain a collection of points. Thus there is a close relationship between the simple Poisson processes of this section and the Poisson point processes of Section 3.4. This is elaborated on in the exercises.

Poisson processes, though clearly limited as population growth models, find many other applications. Some were mentioned in Chapter 3. Furthermore they may result when several sparse point processes, not themselves Poisson, are pooled together. This makes them useful in diverse areas (Tuckwell, 1981).

9.3 MARKOV CHAINS IN CONTINUOUS TIME

In Chapter 8 we encountered Markov random processes which take on discrete values and have a discrete parameter set. Markov processes of that kind are called Markov chains. However, Markov processes in **continuous time** and with discrete state space are also called Markov chains. Usually we mention that they have continuous parameter set when we refer to them in order to distinguish them from the former kind of process. If $X = \{X(t), t \geq 0\}$ is such a process, the Markov property may be written, with $t_0 < t_1 < t_2 \dots < t_{n-1} < t_n$,

$$\Pr \{X(t_n) = s_n | X(t_{n-1}) = s_{n-1}, X(t_{n-2}) = s_{n-2}, \dots, X(t_1) = s_1, X(t_0) = s_0\} = \Pr \{X(t_n) = s_n | X(t_{n-1}) = s_{n-1}\},$$

where $\{s_0, s_1, \dots, s_n\}$ is any set in the state space of the process. The quantities $P(s_n, t_n | s_{n-1}, t_{n-1}) = \Pr \{X(t_n) = s_n | X(t_{n-1}) = s_{n-1}\}$, $0 \leq t_{n-1} < t_n < \infty$, (9.3) are called the **transition probabilities** of the continuous time Markov chain. If

the transition probabilities depend only on time differences they are said to be **stationary** and the corresponding process is called **temporally homogeneous**. Simple Poisson processes are examples of continuous time Markov chains with stationary transition probabilities which may be written

$$p_k(t, j, s) = \Pr\{N(t) = k | N(s) = j\} \\ = \frac{(\lambda t - s)^k e^{-\lambda(t-s)}}{(k-j)!}, \quad k-j = 0, 1, 2, \dots$$

where $0 \leq s < t$ and $j \leq k$ are two non-negative integers. Further examples are discussed in the following sections.

9.4 THE YULE PROCESS

Among the deficiencies of the Poisson process as a model for the growth of populations is that no matter how large the existing population, the chance of a birth in any time interval is always the same. The **simple birth process** we are about to describe was proposed by Yule (1924) as a model for the appearance of new species. Its applicability as a population growth model is limited except perhaps in a few cases such as algae undergoing relatively unchecked reproduction in large lakes (Pielou, 1969). The Yule process is sometimes called the **Yule-Furry Process** due to a related application in physics by Furry (1937). Again we denote by $N(t)$ the number of individuals in existence at time t .

Assumptions on births to individuals

We begin with the following assumptions concerning the births which occur to the individual members of the population.

- (i) Births occur to any individual independently of those to any other individual.
 - (ii) In any small time interval of length Δt , the probability of an offspring to any individual is $\lambda \Delta t + o(\Delta t)$, the probability of no offspring is $1 - \lambda \Delta t + o(\Delta t)$ and the probability of more than one offspring is $o(\Delta t)$.
- Notice that there are no deaths - individuals persist indefinitely and are forever capable of producing offspring.

Population birth probabilities

Suppose that there are known to be n individuals at time t so that $N(t) = n$. Under the above assumptions (i) and (ii), the number of births in the whole population in $(t, t + \Delta t]$ is a binomial random variable with parameters n and $\lambda \Delta t$. We can drop reference to $o(\Delta t)$ as such terms eventually make no

contribution. Then

$$\Pr\{\text{\# births in } (t, t + \Delta t] | N(t) = n\} \\ = \binom{n}{k} (\lambda \Delta t)^k (1 - \lambda \Delta t)^{n-k}, \quad k = 0, 1, \dots, n$$

If $k = 0$ we have

$$\Pr\{0 \text{ births in } (t, t + \Delta t] | N(t) = n\} = (1 - \lambda \Delta t)^n \\ = 1 - \lambda n \Delta t + o(\Delta t), \tag{9.4}$$

whereas for $k = 1$,

$$\Pr\{1 \text{ birth in } (t, t + \Delta t] | N(t) = n\} = \lambda n \Delta t (1 - \lambda \Delta t)^{n-1} \\ = \lambda n \Delta t + o(\Delta t), \tag{9.5}$$

Also, for $k \geq 2$ we find

$$\Pr\{k \text{ births in } (t, t + \Delta t] | N(t) = n\} = o(\Delta t). \tag{9.6}$$

Differential-difference equations satisfied by the transition probabilities

We consider a continuous time Markov chain $N = \{N(t), t \geq 0\}$ with initial value

$$N(0) = n_0 > 0,$$

subject to the evolutionary laws (9.4)-(9.6). Define the transition probabilities

$$p_n(t) = \Pr\{N(t) = n | N(0) = n_0\}, \quad n_0 > 0, \quad t > 0,$$

and note that these are stationary. Our aim is first to find equations governing the evolution in time of p_n and then to solve them.

To obtain a differential equation for p_n we seek a relation between $p_n(t)$ and $p_n(t + \Delta t)$. If $N(t + \Delta t) = n > n_0$ then, ignoring the possibility of more than one birth, we must have

$$N(t) = \begin{cases} n \text{ and no births in } (t, t + \Delta t], \\ n-1 \text{ and one birth in } (t, t + \Delta t]. \end{cases}$$

Dropping reference to the initial state, the law of total probability gives

$$\Pr\{N(t + \Delta t) = n\} = \Pr\{N(t + \Delta t) = n | N(t) = n\} \\ \times \Pr\{N(t) = n\} + \Pr\{N(t + \Delta t) = n | N(t) = n-1\} \\ \times \Pr\{N(t) = n-1\}. \tag{9.7}$$

In symbols this is written, using (9.4) and (9.5),

$$p_n(t + \Delta t) = (1 - \lambda n \Delta t) p_n(t) + \lambda (n-1) \Delta t p_{n-1}(t) + o(\Delta t), \quad n > n_0.$$

This rearranges to

$$\frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} = \lambda[(n-1)p_{n-1}(t) - np_n(t)] + \frac{q(\Delta t)}{\Delta t}$$

Taking the limit $\Delta t \rightarrow 0$ we obtain the required equation

$$\frac{dp_n}{dt} = \lambda[(n-1)p_{n-1} - np_n] \quad n = n_0 + 1, n_0 + 2, \dots \quad (9.8)$$

since $q(\Delta t)/\Delta t \rightarrow 0$ as $\Delta t \rightarrow 0$ by definition. When $n = n_0$ there is no possibility that the population was $n_0 - 1$ so (9.7) becomes

$$\Pr\{N(t + \Delta t) = n_0\} = \Pr\{N(t) = n_0\} \Pr\{N(t) = n_0\}$$

which leads to

$$\frac{dp_{n_0}}{dt} = -\lambda n_0 p_{n_0} \quad (9.9)$$

Initial conditions

Equations (9.8) and (9.9) are first-order differential equations in time. To solve them the values of $p_n(0)$ are needed and since an initial population of n_0 individuals was assumed, we have

$$p_n(0) = \begin{cases} 1, & n = n_0 \\ 0, & n > n_0 \end{cases}$$

Solutions of the differential-difference equations

The solution of (9.9) with initial value unity is

$$p_{n_0}(t) = e^{-\lambda n_0 t}$$

Armed with this knowledge of p_{n_0} we can now find p_{n_0+1} . The differential equation (9.8) is, with $n = n_0 + 1$,

$$\frac{dp_{n_0+1}}{dt} + \lambda(n_0 + 1)p_{n_0+1} = \lambda n_0 p_{n_0} \quad (9.10)$$

This will be recognized as a linear first-order differential equation in standard form (see any first-year calculus text). Its integrating factor is $e^{\lambda(n_0+1)t}$ and in Exercise 10 it is shown that

$$p_{n_0+1}(t) = n_0 e^{-\lambda n_0 t} (1 - e^{-\lambda t}) \quad (9.11)$$

Having obtained p_{n_0+1} , we can solve the equation for p_{n_0+2} , etc. In general we obtain for the probability that there have been a total number of k births at t ,

$$p_{n_0+k}(t) = \binom{n_0+k-1}{n_0-1} e^{-\lambda n_0 t} (1 - e^{-\lambda t})^k \quad k = 0, 1, 2, \dots \quad (9.12)$$

as will be verified in Exercise 11.

It can be seen that $p_{n_0}(t)$ is an exponentially decaying function of time. If we define T_1 as the time of the first birth and observe that $p_{n_0}(t)$ is the probability of no birth in $(0, t]$ we get

$$\Pr\{T_1 > t\} = e^{-\lambda n_0 t}$$

or equivalently

$$\Pr\{T_1 \leq t\} = 1 - e^{-\lambda n_0 t}$$

Hence T_1 is exponentially distributed and has mean

$$E(T_1) = \frac{1}{\lambda n_0}$$

We notice that the larger n_0 is, the faster does $p_{n_0}(t)$ decay towards zero - as it must because the larger the population, the greater the chance for a birth. Note also that $p_{n_0}(t)$ is never zero for $t < \infty$. Thus there is always a non-zero probability that the population will remain unchanged in any finite time interval.

A plot of $p_n(t)$ versus t when $\lambda = n_0 = 1$ is shown in Fig. 9.3. Also shown are the graphs of $p_{n_0+1}(t)$ and $p_{n_0+2}(t)$ which rise from zero to achieve maxima before declining to zero at $t = \infty$. In Exercise 12 it is shown that for these parameter values $p_{1+k}(t)$ has a maximum at $t = \ln(1+k)$.

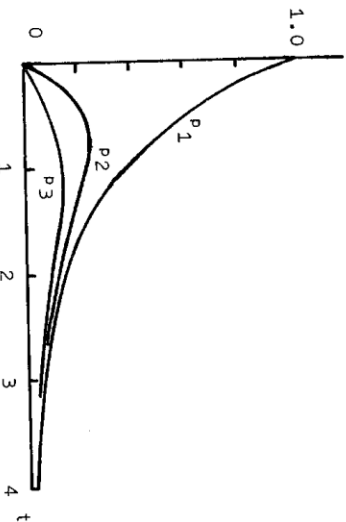


Figure 9.3 Probabilities of no births, one birth and two births as functions of time in the Yule process with $\lambda = 1$ and initially $n_0 = 1$ individual.

9.5 MEAN AND VARIANCE FOR THE YULE PROCESS

For the simple birth process described in the previous section we will show that the mean population size $\mu(t) = E(N(t)) | N(0) = n_0$ at time t is

$$\mu(t) = n_0 e^{\lambda t}.$$

This is the same as the Malthusian growth law with no deaths and a birth rate $b = \lambda$.

Proof By definition

$$\mu(t) = \sum_{n=n_0}^{\infty} n p_n(t).$$

Differentiate to get

$$\frac{d\mu}{dt} = \sum_{n=n_0}^{\infty} n \frac{dp_n}{dt} = n_0 \frac{dp_{n_0}}{dt} + \sum_{n=n_0+1}^{\infty} n \frac{dp_n}{dt}.$$

Substituting from the differential and differential-difference equations (9.9) and (9.8),

$$\frac{d\mu}{dt} = -\lambda n_0^2 p_{n_0} + \lambda \sum_{n=n_0+1}^{\infty} [(n^2 - n)p_{n-1} - n^2 p_n].$$

The coefficient of p_{n-1} is now rewritten to contain a perfect square:

$$\begin{aligned} \frac{d\mu}{dt} &= -\lambda n_0^2 p_{n_0} + \lambda \sum_{n=n_0+1}^{\infty} (n-1)^2 p_{n-1} \\ &\quad + \lambda \sum_{n=n_0+1}^{\infty} (n-1)p_{n-1} - \lambda \sum_{n=n_0+1}^{\infty} n^2 p_n. \end{aligned}$$

Put $m = n - 1$ in the first two sums on the right and get

$$\frac{d\mu}{dt} = \left(-\lambda n_0^2 p_{n_0} + \lambda \sum_{m=n_0}^{\infty} m^2 p_m - \lambda \sum_{m=n_0+1}^{\infty} m^2 p_m \right) + \lambda \sum_{m=n_0}^{\infty} m p_m.$$

The terms in brackets cancel to leave the first-order differential equation

$$\frac{d\mu}{dt} = \lambda \sum_{m=n_0}^{\infty} m p_m = \lambda \mu, \tag{9.13}$$

which has the initial condition

$$\mu(0) = n_0. \tag{9.14}$$

Integrating (9.13) and using (9.14) gives the required result. Similarly, the second moment of $N(t)$ may be found and hence the variance. However, it is quicker to use the properties of the negative binomial distribution.

Mean and variance from the negative binomial distribution

Consider a sequence of Bernoulli trials with probability p of success and probability $q = 1 - p$ of failure at each trial. Let the random variable X_r be the number of trials up to and including the r th success, $r = 1, 2, \dots$. Then it is easily seen that the distribution of X_r is given by

$$\Pr \{X_r = k\} = \binom{k-1}{r-1} p^r q^{k-r}, \quad k = r, r+1, r+2, \dots \tag{9.15}$$

the smallest possible value of X_r being r since there must be at least r trials to obtain r successes. The mean and variance of X_r are found to be (see Exercise 13),

$$\begin{aligned} E(X_r) &= \frac{r}{p} \\ \text{Var}(X_r) &= \frac{r q}{p^2}. \end{aligned}$$

We now put $j = n_0 + k$ in (9.12) to get

$$\Pr \{N(t) = j\} = \binom{j-1}{n_0-1} e^{-\lambda n_0 t} (1 - e^{-\lambda t})^{j-n_0}, \quad j = n_0, n_0 + 1, n_0 + 2, \dots$$

This is seen to be a negative binomial distribution as in (9.15) with parameters

$$\begin{aligned} r &= n_0, \\ p &= e^{-\lambda t}, \quad q = 1 - e^{-\lambda t}. \end{aligned}$$

Thus we quickly see that the mean and variance of the population in the Yule process are

$$\begin{aligned} E(N(t)) &= n_0 e^{\lambda t} \\ \text{Var}(N(t)) &= n_0 (1 - e^{-\lambda t}) e^{2\lambda t} \end{aligned}$$

Approximations

(i) *Large t* For large t the variance is asymptotically

$$\text{Var}(N(t)) \underset{t \rightarrow \infty}{\sim} n_0 e^{2\lambda t},$$

and the standard deviation of the population is

$$\sigma(t) \sim \sqrt{n_0} e^{\lambda t}.$$

Thus the mean and standard deviation grow, as indicated in Fig. 9.4, and eventually their ratio is constant.

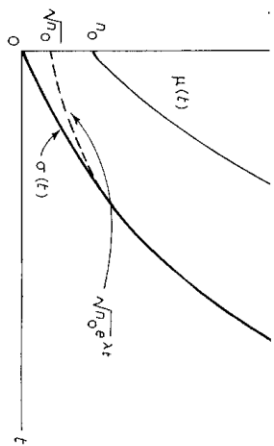


Figure 9.4 Mean and standard deviation of the population as functions of time in a Yule process.

(ii) *Large n_0* Let us suppose that there is just one individual to start with. The probability that there is still only one individual at time t , an event we will call a 'failure', is

$$\Pr\{\text{population is unchanged at } t\} = e^{-\lambda t} = \Pr\{\text{'failure'}\}.$$

Now, if there are $n_0 > 1$ individuals to start with, since individuals act independently,

$$\Pr\{\text{population is unchanged at } t\} = \Pr\{n_0 \text{ failures in } n_0 \text{ trials}\} \\ = (e^{-\lambda t})^{n_0} = e^{-\lambda n_0 t}.$$

this being an exact result. Also, when $n_0 = 1$ let the mean and variance of the population be denoted by μ_1 and σ_1^2 which are given by

$$\mu_1 = e^{\lambda t}, \sigma_1^2 = e^{2\lambda t}(1 - e^{-\lambda t}).$$

When $n_0 > 1$ we may at any time divide the population into n_0 groups, those in each group being descendants of one of the original individuals. The number in each group is a random variable and it follows that the population at time t is the sum of n_0 independent and identically distributed random variables each having mean μ_1 and variance σ_1^2 . By the central limit theorem we see that for large n_0 and any t , the random variable $N(t)$ is approximately normal with mean $n_0 \mu_1$ and variance $n_0 \sigma_1^2$. Hence we may estimate with reasonable accuracy the probability that the population lies within prescribed limits (see exercises).

9.6 A SIMPLE DEATH PROCESS

In this rather macabre continuous time Markov chain, individuals persist only until they die and there are no replacements. The assumptions are similar to

those in the birth process, now each

individual, if still alive at time t , is removed in $(t, t + \Delta t]$ with probability $\mu \Delta t + o(\Delta t)$. Again we are interested in finding the transition probabilities

$$p_n(t) = \Pr\{N(t) = n | N(0) = n_0\}, \quad n = n_0, n_0 - 1, \dots, 2, 1, 0.$$

We could proceed via differential-difference equations for p_n but there is a more expeditious method.

The case of one individual

Let us assume $n_0 = 1$. Now $p_1(t)$ is the probability that this single individual is still alive at t and we see that

$$p_1(t + \Delta t) = p_1(t)(1 - \mu \Delta t) + o(\Delta t) \quad (9.16)$$

since $1 - \mu \Delta t$ is the probability that the individual did not die in $(t, t + \Delta t]$. From (9.16) it quickly follows that

$$\frac{dp_1}{dt} = -\mu p_1, \quad t > 0.$$

The solution with initial value $p_1(0) = 1$ is just

$$p_1(t) = e^{-\mu t}$$

The initial population size is $N(0) = n_0 > 1$

If there are n_0 individuals at $t = 0$, the number alive at t is a binomial random variable with parameters n_0 and $p_1(t)$. Therefore we have immediately

$$p_n(t) = \binom{n_0}{n} e^{-\mu n t} (1 - e^{-\mu t})^{n_0 - n}, \quad n = n_0, n_0 - 1, \dots, 1, 0.$$

Also

$$E\{N(t)\} = n_0 e^{-\mu t},$$

which corresponds to a Malthusian growth law with $d = \mu$ and $b = 0$, and

$$\text{Var}\{N(t)\} = n_0 e^{-\mu t} (1 - e^{-\mu t}).$$

Extinction

In this pure death process the population either remains constant or it decreases. It may eventually reach zero in which case we say that the population has gone *extinct*. The probability that the population is extinct at time t is

$$\Pr\{N(t) = 0 | N(0) = n_0\} = (1 - e^{-\mu t})^{n_0} \rightarrow 1 \quad \text{as } t \rightarrow \infty.$$

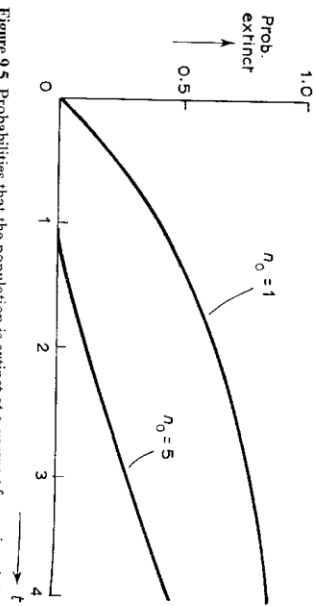


Figure 9.5 Probabilities that the population is extinct at t , versus t for various initial populations with $\mu = 0.5$.

Thus extinction is inevitable in this model. In Fig. 9.5 are shown the probabilities of extinction versus time for various initial populations.

9.7 SIMPLE BIRTH AND DEATH PROCESS

We now combine the ideas of the Yule process and the simple death process of the previous section. Let there be n_0 individuals initially and $N(t)$ at time t . In $(t, t + \Delta t]$ an individual has an offspring with probability $\lambda \Delta t + o(\Delta t)$ and dies with probability $\mu \Delta t + o(\Delta t)$. Using the same kind of reasoning as in Section 9.4 for the population birth probabilities we find that

$$\begin{aligned} \Pr\{\text{one birth in } (t, t + \Delta t] | N(t) = n\} &= \lambda n \Delta t + o(\Delta t) \\ \Pr\{\text{one death in } (t, t + \Delta t] | N(t) = n\} &= \mu n \Delta t + o(\Delta t) \\ \Pr\{\text{no change in population size in } (t, t + \Delta t] | N(t) = n\} &= 1 - (\lambda + \mu)n \Delta t + o(\Delta t) \end{aligned}$$

The ways to obtain a population size n at time $t + \Delta t$ are, if $n \geq 1$,

$$\begin{cases} N(t) = n - 1 \text{ and one birth in } (t, t + \Delta t] \\ N(t) = n \text{ and no change in } (t, t + \Delta t] \\ N(t) = n + 1 \text{ and one death in } (t, t + \Delta t]. \end{cases}$$

Hence

$$P_n(t + \Delta t) = P_{n-1}(t)(\lambda(n-1)\Delta t + P_n(t)[1 - (\lambda + \mu)n\Delta t] + P_{n+1}(t)\mu n \Delta t + o(\Delta t).$$

It quickly follows that

$$\frac{dP_n}{dt} = \lambda(n-1)P_{n-1} - (\lambda + \mu)nP_n + \mu(n+1)P_{n+1}, \quad n \geq 1. \quad (9.17)$$

If $n = 0$ we have simply

$$\frac{dP_0}{dt} = \mu P_1, \quad (9.18)$$

and the initial conditions are

$$P_n(0) = \begin{cases} 1, & n = n_0 \\ 0, & n \neq n_0 \end{cases}$$

The system of equations (9.17) and (9.18) cannot be solved recursively as could the equations for the simple birth (Yule) process as there is no place to get started.

The probability generating function of $N(t)$

By definition, the probability generating function of $N(t)$ is

$$\phi(s, t) = \sum_{n=0}^{\infty} P_n(t) s^n.$$

This can be shown (see Exercise 16) to satisfy the first-order partial differential equation

$$\frac{\partial \phi}{\partial t} = (\lambda s - \mu)(s-1) \frac{\partial \phi}{\partial s}, \quad (9.19)$$

which is to be solved with the initial condition

$$\phi(s, 0) = s^{n_0} \quad (9.20)$$

It may be shown (see, for example, Pollard, 1973; Bailey, 1964) and it will be verified in Exercise 17, that the solution of (9.19) and (9.20) is

$$\phi(s, t) = \left(\frac{\mu - \psi(s)e^{-(\lambda - \mu)t}}{\lambda - \psi(s)e^{-(\lambda - \mu)t}} \right)^{n_0}, \quad (9.21)$$

where

$$\psi(s) = \frac{\lambda s - \mu}{s - 1}. \quad (9.22)$$

The probability of extinction

A few sample paths of the simple birth and death process are shown in Fig. 9.6. The state space is the set of all non-negative integers $\{0, 1, 2, \dots\}$ and the state 0

$\lambda e^{-\mu t} (9.17)$

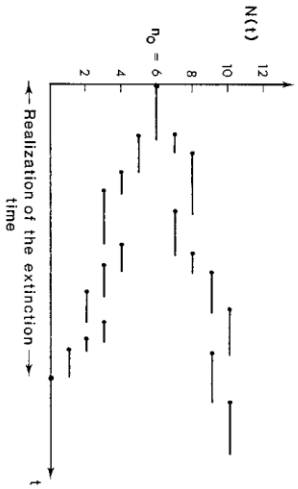


Figure 9.6 Two representative sample paths of a birth and death process. Here $N(0) = 6$ and one path is shown on which the population goes extinct.

is clearly absorbing. A sample path may terminate at 0 which corresponds to extinction of the population.

We can easily find the probability that extinction has occurred at or before time t from the probability generating function. This is just

$$\Pr\{N(t) = 0 | N(0) = n_0\} = \phi(0, t)$$

From (9.22) we have $\psi(0) = \mu$ and thence, from (9.21),

$$\phi(0, t) = \left(\frac{\mu(1 - e^{-(\lambda - \mu)t})}{\lambda - \mu e^{-(\lambda - \mu)t}} \right)^{n_0}, \quad \lambda \neq \mu. \quad (9.23)$$

When $\lambda = \mu$ the following expression is obtained by taking the appropriate limit in (9.23):

$$\phi(0, t) = \left(\frac{\lambda t}{\lambda t + 1} \right)^{n_0}, \quad \lambda = \mu. \quad (9.24)$$

In the limit $t \rightarrow \infty$, $\phi(0, t)$ approaches the probability that the population ever goes extinct. Denote this quantity by P_{ext} . Then from (9.23) and (9.24) we find

$$P_{\text{ext}} = \begin{cases} 1, & \lambda \leq \mu \\ \left(\frac{\mu}{\lambda}\right)^{n_0}, & \lambda > \mu \end{cases}$$

Thus extinction is inevitable if the probability of a birth to any individual is less than or equal to the probability of death in any small time interval. It may seem surprising that extinction is certain when $\lambda = \mu$. To understand this we note that 0 is an absorbing barrier which is always a finite distance from the

value of $N(t)$. The situation is similar in the random walk on $[0, 1, 2, \dots]$ with $p = q$ where we found that absorption at 0 is certain (see Section 7.6).

In the cases $\lambda \leq \mu$ where extinction is certain, we may define the random variable T which is the **extinction time**. Evidently the distribution function of T is

$$\Pr\{T \leq t\} = \phi(0, t)$$

since this is the probability that extinction occurs at or before t . When $\lambda = \mu$ the expected extinction time is infinite (see Exercise 18) but it is finite when $\lambda < \mu$. When $\lambda > \mu$ we may still talk of the random variable T , the extinction time. However, we then have

$$\Pr\{T < \infty\} = \left(\frac{\mu}{\lambda}\right)^{n_0},$$

so we must also have

$$\Pr\{T = \infty\} = 1 - \left(\frac{\mu}{\lambda}\right)^{n_0}.$$

Clearly in these cases T has no finite moments and, because its probability mass is not all concentrated on $(0, \infty)$ we say it is not a 'proper' random variable.

9.8 MEAN AND VARIANCE FOR THE BIRTH AND DEATH PROCESS

The expected number of individuals at time t is

$$m(t) = E[N(t) | N(0) = n_0] = \sum_{n=0}^{\infty} n p_n(t) = \sum_{n=1}^{\infty} n p_n(t).$$

We will find a differential equation for $m(t)$. We have

$$\frac{dm}{dt} = \sum_{n=1}^{\infty} n \frac{dp_n}{dt}$$

and on substituting from the differential-difference equation (9.17) we get

$$\frac{dm}{dt} = \sum_{n=1}^{\infty} n[\lambda(n-1)p_{n-1} - (\lambda + \mu)n p_n + \mu(n+1)p_{n+1}]$$

which rearranges to

$$\begin{aligned} \frac{dm}{dt} = & \lambda \sum_{n=1}^{\infty} (n-1)^2 p_{n-1} + \lambda \sum_{n=1}^{\infty} (n-1)p_{n-1} - (\lambda + \mu) \sum_{n=1}^{\infty} n^2 p_n \\ & + \mu \sum_{n=1}^{\infty} (n+1)^2 p_{n+1} - \mu \sum_{n=1}^{\infty} (n+1)p_{n+1}. \end{aligned}$$

A relabeling of indices with $n' = n - 1$ in sums involving p_{n-1} and with $n' = n + 1$ in sums involving p_{n+1} yields

$$\begin{aligned} \frac{dn}{dt} = & -(\lambda + \mu) \sum_{n=1}^{\infty} n^2 p_n + \lambda \sum_{n=0}^{\infty} n^2 p_n + \mu \sum_{n=2}^{\infty} n^2 p_n \\ & + \lambda \sum_{n=0}^{\infty} n p_n - \mu \sum_{n=2}^{\infty} n p_n \end{aligned}$$

In the first three sums here, terms from $n, n', n'' = 2$ and onward cancel and leave $-(\lambda + \mu)p_1 + \lambda p_1 = -\mu p_1$. Thus

$$\begin{aligned} \frac{dn}{dt} = & -\mu p_1 + \lambda \sum_{n=0}^{\infty} n p_n - \mu \sum_{n=2}^{\infty} n p_n \\ = & (\lambda - \mu) \sum_{n=0}^{\infty} n p_n \end{aligned}$$

or simply

$$\frac{dn}{dt} = (\lambda - \mu)n.$$

With initial condition $m(0) = n_0$, the solution is

$$m(t) = n_0 e^{(\lambda - \mu)t}.$$

This is the same as the deterministic result (Malthusian law) of Section 9.1 with the birth rate b replaced by λ and the death rate d replaced by μ . The second moment of $N(t)$,

$$M(t) = \sum_{n=0}^{\infty} n^2 p_n(t)$$

can be shown to satisfy

$$\frac{dM}{dt} = 2(\lambda - \mu)M + (\lambda + \mu)M, \quad M(0) = n_0^2, \quad (9.25)$$

as will be seen in Exercise 19. The variance of the population in the birth and death process may then be shown to be

$$\text{Var}(N(t) | N(0) = n_0) = n_0 \frac{(\lambda + \mu) e^{(\lambda - \mu)t} [e^{(\lambda - \mu)t} - 1]}{(\lambda - \mu)}, \quad \lambda \neq \mu.$$

In the special case $\lambda = \mu$,

$$\text{Var}(N(t) | N(0) = n_0) = 2\lambda n_0 t.$$

An alternative method of finding the moments of $N(t)$ is to use the moment generating function (see Exercise 20).

Birth and death processes have recently become very important in studies of how ions move across cell membranes. In the simplest model there are just two states for an ion channel—open and closed. The channel stays in each state for an exponentially distributed time before making a transition to the other state. It is hoped that a study of such continuous time Markov chain models will elucidate the mechanisms by which molecules of the membrane interact with various drugs. For details of this fascinating application see Colquhoun and Hawkes (1977), Hille (1984) and Tuckwell (1988).

REFERENCES

- Bailey, N.T.J. (1964). *The Elements of Stochastic Processes*. Wiley, New York.
 Bartlett, M.S. (1960). *Stochastic Population Models*. Methuen, London.
 Cameron, R.J. (ed.) (1982). *Year Book Australia*. Australian Bureau of Statistics, Canberra.
 Colquhoun, D. and Hawkes, A.G. (1977). Relaxations and fluctuations of membrane currents that flow through drug operated channels. *Proc. R. Soc. Lond. B*, **199**, 231–62.
 Furry, W.H. (1937). On fluctuation phenomena in the passage of high-energy electrons through lead. *Phys. Rev.*, **52**, 569–81.
 Hille, B. (1984). *Ionic Channels of Excitable Membranes*. Sinauer, Sunderland, Mass.
 Kozmin, N. (1968). *Introduction to the Mathematics of Populations*. Addison-Wesley, Reading, Mass.
 Ludwig, D. (1978). *Stochastic Population Theories*. Springer, New York.
 Parzen, E. (1962). *Stochastic Processes*. Holden-Day, San Francisco.
 Paton, E.C. (1969). *An Introduction to Mathematical Ecology*. Wiley, New York.
 Pollard, J.H. (1973). *Mathematical Models for the Growth of Human Populations*. Cambridge University Press, London.
 Tuckwell, H.C. (1981). Poisson processes in biology. In *Stochastic Nonlinear Systems*, Springer-Verlag, Berlin, pp. 162–71.
 Tuckwell, H.C. (1988). *Stochastic Processes in the Neurosciences*. SIAM, Philadelphia.
 Yule, G.U. (1924). A mathematical theory of evolution based upon the conclusions of Dr J.C. Willis. *Phil. Trans. Roy. Soc. Lond. B*, **213**, 21–87.

EXERCISES

- Using the birth and death rates for 1966 given in Table 9.2 and the 1966 population of Australia given in Table 9.1, estimate the 1971 population. Compare with the actual population in 1971. Is the discrepancy in the direction you would expect? Why?
- In a simple Poisson process, let $p_k(t) = \Pr\{N(t) = k | N(0) = 0\}$. Use the relations

$$\Pr\{\Delta N(t) = k\} = \begin{cases} 1 - \lambda \Delta t + o(\Delta t), & k = 0, \\ \lambda \Delta t + o(\Delta t), & k = 1, \\ o(\Delta t), & k \geq 2, \end{cases}$$