Predicting population extinction or disease outbreaks with stochastic models

Linda J. S. Allen, Sophia R. Jang & Lih-Ing Roeger

Published online: 27 Dec 2016.

Submit your article to this journal ⧉

Article views: 813

View related articles ⧉

View Crossmark data ⧉

Letters in Biomathematics

Taylor & Francis
Taylor & Francis Group

EDUCATIONAL ARTICLE

🔓 OPEN ACCESS   Check for updates

Predicting population extinction or disease outbreaks with stochastic models

Linda J. S. Allen, Sophia R. Jang and Lih-Ing Roeger

Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, USA

ABSTRACT

Models of exponential growth, logistic growth and epidemics are common applications in undergraduate differential equation courses. The corresponding stochastic models are not part of these courses, although when population sizes are small their behaviour is often more realistic and distinctly different from deterministic models. For example, the randomness associated with births and deaths may lead to population extinction even in an exponentially growing population. Some background in continuous-time Markov chains and applications to populations, epidemics and cancer are presented with a goal to introduce this topic into the undergraduate mathematics curriculum that will encourage further investigation into problems on conservation, infectious diseases and cancer therapy. MATLAB programs for graphing sample paths of stochastic models are provided in the Appendix.

1. Introduction

Deterministic and stochastic models of populations and epidemics have a long history. Malthusian or exponential growth, named for Thomas Robert Malthus, was first discussed in *An Essay on the Principle of Population* (under the pseudonym Joseph Johnson) published in 1798 (Johnson, 1798). Malthus edited and published five more versions of this original essay. These essays were very influential at the time, impacting, for example, the work of Charles Darwin on the theory of natural selection and the work of Pierre-Francois Verhulst on logistic growth. In 1838, Verhulst mentioned Malthus' principle of geometric growth but argued for a constraint on the rate of growth which he modelled as proportional to the square of the population size (Verhulst, 1838). In epidemic models, William Ogilvy Kermack and Anderson Gray McKendrick were the first to introduce the three-compartment SIR epidemic model (Susceptible-Infectious-Recovered). Their first paper in 1927 entitled 'A contribution to the mathematical theory of epidemics' defined the concept of an epidemic threshold for disease outbreaks, now known as the basic reproduction number (Kermack & McKendrick, 1927). Some of the earliest stochastic models for population growth and extinction were published in the mid-1800s. Seemingly

CONTACT Linda J. S. Allen ✉ linda.j.allen@ttu.edu

simple questions about the survival of family names led to some new mathematical approaches in probability theory. These new approaches in turn led to the birth of a new field of mathematics, now known as branching processes. It was Bienaymé (1845) and Watson and Galton (1875) who first posed these questions. Other early contributors to stochastic population and epidemic theory on simple birth and death processes and probability of disease outbreaks include Feller (1939), Kendall (1949, 1956), Bartholomay (1958) and Whittle (1955). Kendall (1949, p. 230) stated:

> The classical theory of population growth treats the size of the population as a continuous variable, …. It is clear that a more refined analysis must take into account the role of chance effects in the development of the population, and starting with W. Feller in 1939 a good deal of attention has recently been paid to this question.

Despite the long history and the many advances in stochastic population theory, this theory has not been included in many undergraduate courses. The accessibility of computers makes it an easy task to demonstrate the differences between deterministic and stochastic modelling approaches in the classroom. The drawback, however, may be the required background in continuous-time Markov chain theory. Although discrete-time Markov chains are introduced in courses on Discrete Mathematics, the closely related continuous-time theory requires calculus and differential equations. With some background in continuous-time Markov chains, stochastic applications to populations, epidemics and cancer can be easily introduced into the undergraduate mathematics curriculum. Stochastic models are especially relevant when population sizes are small and subject to extinction, whereas deterministic models, differential equations or difference equations, are a good approximation to the overall dynamics when population sizes are large. These latter topics are covered in the traditional mathematics undergraduate courses.

Our goals are to provide a few well-known applications to populations and epidemics, and to a more recent application to cancer therapy based on continuous-time Markov chain theory. These applications can be incorporated into undergraduate mathematics courses such as Calculus, Differential Equations or Mathematical Modelling. Specifically, our goals are (1) to give a brief introduction to continuous-time Markov chain theory and the simple birth and death process, (2) to provide MATLAB algorithms to numerically simulate continuous-time Markov chain models and (3) to demonstrate how probability of extinction from the simple birth and death process can be used to predict extinction in problems associated with population growth or infectious diseases, when population sizes or number of infectious individuals are small or to predict time of extinction when treatment is designed to eliminate disease. In advanced courses, the theory and application of Markov chains can be studied in some depth whereas in introductory courses, a short presentation on Markov chains accompanied by MATLAB simulations can be used for illustration purposes.

2. Simple birth and death process

Formally, a stochastic process is a collection of random variables, $X(\omega, t)$. Capital X denotes a random variable that depends on ω and t. The value of ω is an element in the sample space Ω and t is time, $t \in [0, \infty)$. Therefore, a stochastic process is a collection

$$\{X(\omega, t) | \omega \in \Omega, t \in [0, \infty)\}.$$

The sample space Ω is the same for all random variables in the collection. Since we are interested in the size of a population, an element ω in the sample space is a subset of the nonnegative integers. For a fixed ω, $X(\omega, t)$ is a discrete random variable that changes with time, and is referred to as a *sample path* or a *stochastic realization* of the process. Each $\omega \in \Omega$, yields a new sample path.

For simplicity, the ω is often omitted and we write $X(t)$ instead of $X(\omega, t)$. We emphasize that each of the random variables, $X(t)$, is discrete-valued, taking values in the set of nonnegative integers,

$$X(t) \in \{0, 1, 2, \ldots\} \text{ or } X(t) \in \{0, 1, \ldots, N\}.$$

We use the term *state* of a random variable to mean the value of the random variable. For example, if $X(t) = 2$, then 2 is the state of the random variable X at time t.

A stochastic process is generally defined by relating the random variable $X(t + \Delta t)$ to the random variable $X(t)$, when Δt is small (Karlin & Taylor, 1975). These relations are defined by the *infinitesimal transition probabilities*. Here, we refer to them as *transition probabilities*. A transition probability for a birth is the probability associated with the change $X(t + \Delta t) = X(t) + 1$. A transition probability for a death is the probability associated with the change $X(t + \Delta t) = X(t) - 1$.

Denote the transition probability for a change in state from i to state j in a short time Δt as the following conditional probability:

$$p_{i,j}(\Delta t) = \mathbb{P}(X(t + \Delta t) = j | X(t) = i). \tag{1}$$

When the transition probabilities depend only on the length of time Δt between the transitions and not on the particular time t at which these transitions occur, as in (1), then the stochastic process is referred to as a *time-homogeneous process*. Let b be the per capita birth rate and d be the per capita death rate of the process. Then the transition probabilities for a small interval of time Δt,

$$p_{i,j}(\Delta t) = \begin{cases} bi\Delta t, & \text{if } j = i + 1 \\ di\Delta t, & \text{if } j = i - 1 \\ 1 - (b + d)i\Delta t, & \text{if } j = i. \end{cases} \tag{2}$$

Definition (2) states that when Δt is small and $X(t) = i$, the probability of a birth during the time interval Δt is approximately $bi\Delta t$, the probability of a death is approximately $di\Delta t$, the probability of no change is approximately $1 - (b + d)i\Delta t$ and all other changes are negligible. The sum of all the probabilities must equal one. Definition (2) also states that jumps occur in the population size when there is a birth or a death. The time that elapses between jumps (a birth or a death) is referred to as the *interevent time*.

This simple birth and death process is an example of a continuous-time Markov chain. The term *chain* implies the random variable $X(t)$ is discrete-valued and the term *Markov* (named after Andrey Andreyevich Markov) implies the future state at $t + \Delta t$, given the present state at time t, is independent of the past. That is, for time $\Delta t > 0$,

$$\mathbb{P}(X(t + \Delta t) | X(t)) = \mathbb{P}(X(t + \Delta t) | X(t), X(s)) \text{ for any } s < t.$$

This is referred to as the *memoryless property*. The only continuous probability distribution with the memoryless property is the exponential distribution. The interevent time in a Markov chain is a continuous random variable with an exponential distribution. This property is demonstrated in the birth and death process in Equation (4). It is important to note that if the initial state $X(0) = i$ is known, then the transition probabilities are actually the probabilities $p_j(t)$ associated with the random variables $X(t)$:

$$p_j(t) = p_{i,j}(t) = \mathbb{P}(X(t) = j | X(0) = i), \;\; j = 0, 1, 2, 3, \dots . \tag{3}$$

2.1. Sample paths

Because the values of the random variables are discrete, any sample path $X(t)$ of a continuous-time Markov chain is not a continuous function of time. For example, if the process is in state i and there is a birth, then the process jumps to state $i + 1$ or if there is a death, the process jumps to state $i - 1$. From Definition (2), if $X(0) = i \geq 1$ and if the interevent time is $\tau > 0$, then a change of state occurs at τ. It follows that the left-hand limit of the process at τ is i and the right-hand limit is either $i - 1$ or $i + 1$. The stochastic process is continuous from the right but not from the left (see Figure 1). That is,

$$X(\tau^-) = \lim_{\Delta t \to 0^-} X(\tau + \Delta t) = i$$

and

$$X(\tau^+) = \lim_{\Delta t \to 0^+} X(\tau + \Delta t) = i - 1 \text{ or } i + 1.$$

The probability of a birth, $X(\tau^+) = i + 1$, is

$$\frac{b}{b+d} = \lim_{\Delta t \to 0^+} \frac{bi\Delta t}{(b+d)i\Delta t}$$

and the probability of a death, $X(\tau^+) = i - 1$, is $d/(b+d)$.

To numerically simulate a birth or a death the uniform distribution U on $[0, 1]$ is applied. In general, for k different events, each representing a jump in the process and each with a given positive probability $q_j, j = 1, \dots, k$, the unit interval $[0, 1]$ is divided into k subintervals of length q_j, $\sum_{j=1}^{k} q_j = 1$. Then, if a uniform random number $u \in U$ lies in the subinterval q_j, event j occurs. In the birth and death process there are only two events, $q_1 = b/(b+d)$ and $q_2 = d/(b+d)$. If $u < b/(b+d)$, there is a birth but if not, there is a death. For example, given $X(0) = 2$, $b = 2$, $d = 1.5$ and $u \in U$, the MATLAB code is given below. Notice in MATLAB that the initial state is written as $X(1) = 2$ and initial time as $t(1) = 0$, since array indices in MATLAB start at one.

```
t(1)=0; X(1)=2; b=2; d=1.5;
u=rand;
if u<b/(b+d)
   X(2)=X(1)+1; % birth
else
   X(2)=X(1)-1; % death
end
```
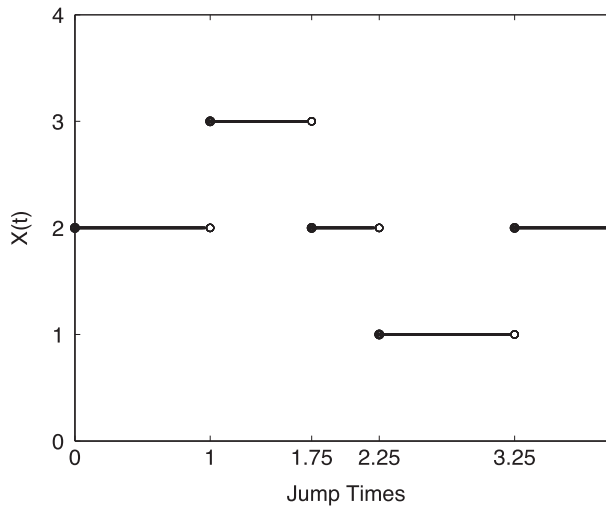
Figure 1. A sample path is continuous from the right but not from the left. In this example, the jump times are 1, 1.75, 2.25 and 3.25 and the interevent times are $\tau_1 = 1$, $\tau_2 = .75$, $\tau_3 = .5$ and $\tau_4 = 1$.

Notes: Beginning at $X(0) = 2$, a birth occurs, $X(1) = 3$, then death $X(1.75) = 2$, followed by another death $X(2.25) = 1$ and finally a birth $X(3.25) = 2$.

In addition to the birth and death probabilities, numerical simulation of a sample path requires computation of the values for the interevent times. The interevent time T_i is a continuous random variable with values $\tau \in (0, \infty)$. As noted earlier, for Markov chains, the interevent time is exponentially distributed. Given that the process is in state i at time t, $X(t) = i$, the interevent time is exponentially distributed with parameter $(b + d)i$ (the sum of the rates corresponding to all possible events). In particular, the probability density function of T_i is $f_i(\tau) = (b + d)i \exp(-(b + d)i\tau)$ and the cumulative distribution is $F_i(\tau) = 1 - \exp(-(b + d)i\tau)$. The mean and standard deviation of the exponential distribution are equal to $[(b + d)i]^{-1}$. Therefore, as the population size i increases, the mean interevent time decreases. The Markov property of the interevent time can be easily demonstrated. The interevent time T_i does not depend on the length of time t to reach state i but only on the current state i of the process (independent of past history or memoryless property):

$$\mathbb{P}(T_i \geq t + \tau | T_i \geq t) = \frac{\exp(-(b + d)i(t + \tau))}{\exp(-(b + d)it)} = \exp(-(b + d)i\tau) = \mathbb{P}(T_i \geq \tau). \quad (4)$$

The uniform distribution U and the cumulative distribution F_i are used to compute the interevent time of a sample path. We derive an identity for T_i in terms of U that depends on the following properties of U:

$$a = \mathbb{P}(U \leq a) = \mathbb{P}(1 - U \leq a), \quad a \in [0, 1].$$

Applying the first property to $F_i(\tau)$ leads to $F_i(\tau) = \mathbb{P}(U \leq F_i(\tau))$. But $\mathbb{P}(U \leq F_i(\tau)) = \mathbb{P}(F_i^{-1}(U) \leq \tau)$. The definition $F_i(\tau) = \mathbb{P}(T_i \leq \tau)$ and the preceding identity show that T_i and $F_i^{-1}(U)$ have the same cumulative distribution, namely, $F_i(\tau)$. Computation of the
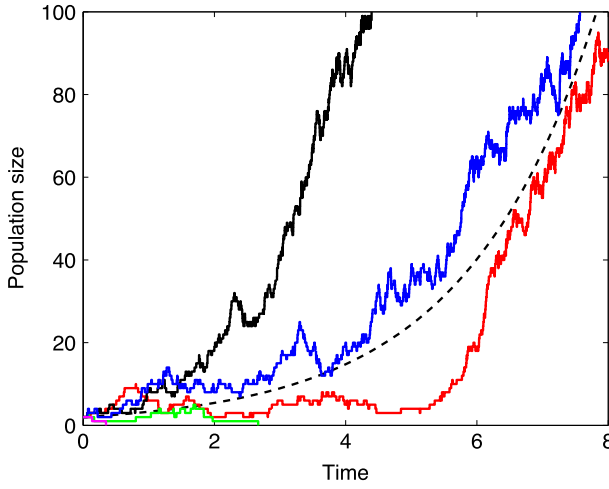
Figure 2. The exponential solution (black dashed curve) and five sample paths of the simple birth and death process (black, red, blue, green and magenta curves) are plotted for parameter values $b = 2$, $d = 1.5$ and initial value $X(0) = 2 = x(0)$.
Note: Although the sample paths are discontinuous, when graphed with MATLAB the discrete jumps are connected.

inverse $F_i^{-1}(U)$ yields a formula for the interevent time in terms of the uniform random variable U:

$$T_i = F_i^{-1}(U) = -\frac{\ln(1-U)}{(b+d)i} = -\frac{\ln(U)}{(b+d)i},$$

where we have used the second property of U. The preceding calculation is part of an algorithm for the interevent time which is often referred to as the Gillespie algorithm (Gillespie, 1977). A uniform random number $u \in U$ generates a value $\tau = \ln(u)/((b+d)i)$ for the interevent time. For example, in MATLAB code, the time of the first jump, given $X(0) = 2$, $b = 2$ and $d = 1.5$ is computed as follows:

```
t(1)=0;  X(1)=2;  b=2;  d=1.5;
u=rand;
t(2)=t(1)-log(u)/((b+d)*X(1));
```

The deterministic analogue of the simple birth and death process is the well-known Malthusian exponential growth model. Expressed as a differential equation, the model is

$$\frac{dx(t)}{dt} = (b-d)x(t), \quad x(0) = x_0 > 0 \tag{5}$$

with solution $x(t) = x_0 \exp((b-d)t)$. If $b > d$, there is exponential growth, solutions approach infinity, but if $b < d$, there is exponential decay; solutions approach zero. In fact, the exponential growth model (5) is the mean of the simple birth and death process.

We use MATLAB to generate five sample paths of the simple birth and death process and plot them against the corresponding deterministic solution in Figure 2 (the MATLAB code is in Appendix 2). To compare the dynamics of the exponential growth model to the simple birth and death process, we assume both models have the same initial size and the same birth and death rates. In the example in Figure 2, $X(0) = 2 = x(0)$, and $b = 2$ and $d = 1.5$.

In the simple birth and death process, the zero state is an *absorbing state,* corresponding to population extinction. That is, if $X(t) = 0$, then $X(t + \tau) = 0$ for $\tau \in [0, \infty)$. Absorption into the zero state is illustrated in two of the five sample paths in Figure 2. Although extinction cannot occur in the deterministic model when $b > d$, the probability of extinction in the stochastic model is always positive. In the next section, we give an analytical expression for the probability of extinction $p_0(t)$.

2.2. Probability of extinction

The probability of extinction can be derived from differential equations that follow from the transition probabilities, Definition (2) and techniques from probability generating functions (e.g. Allen, 2010; Bailey, 1990). The differential equations are known as the *Kolmogorov differential equations* to honour the contributions of the mathematician Andrey Nikolaevich Kolmogorov. The derivation is relegated to Appendix 1. The derivation for the probability of extinction can also be found in the classic textbook by Feller (1968).

The derivation in Appendix 1 yields an expression for the probability generating function,

$$\mathcal{P}(s, t) = \sum_{j=0}^{\infty} p_j(t) s^j, \ \ s \in [0, 1].$$

The probability generating function is useful for generating the probabilities $p_j(t)$ and the moments $\mathbb{E}([X(t)^m])$. Evaluating $\mathcal{P}$ at $s = 0$ gives $p_0(t)$ and differentiating $\mathcal{P}$ with respect to s and evaluating at $s = 0$ gives $p_1(t)$. The expectation $\mathbb{E}(X(t)) = \sum_{j=1}^{\infty} j p_j(t)$ is found by differentiating $\mathcal{P}$ with respect to s and evaluating at $s = 1$.

The closed form expression for $\mathcal{P}(s, t)$, derived in Appendix 1, is well-known (e.g. Allen, 2010; Bailey, 1990). An underlying assumption in this derivation is that each individual in the simple birth and death process acts independently, so that the probability generating function for $X(0) = i$ is just the product of i probability generating functions for $X(0) = 1$. In particular,

$$\mathcal{P}(s, t) = \begin{cases} \left(\dfrac{e^{(d-b)t}(bs - d) - d(s - 1)}{e^{(d-b)t}(bs - d) - b(s - 1)} \right)^i, & \text{if } b \neq d \\ \left(\dfrac{1 - (bt - 1)(s - 1)}{1 - bt(s - 1)} \right)^i, & \text{if } b = d. \end{cases} \quad (6)$$

Differentiation of $\mathcal{P}$ with respect to s leads to

$$\mathbb{E}(X(t)) = \left. \frac{\partial \mathcal{P}(s, t)}{\partial s} \right|_{s=1} = i e^{(b-d)t}, \ \ X(0) = i.$$

Evaluating $\mathcal{P}$ at $s = 0$ leads to a formula for the probability of extinction by time t:

$$p_0(t) = \begin{cases} \left(\dfrac{d - d e^{(d-b)t}}{b - d e^{(d-b)t}} \right)^i, & \text{if } b \neq d \\ \left(\dfrac{bt}{1 + bt} \right)^i, & \text{if } b = d. \end{cases} \quad (7)$$

In the limit, the probability of extinction is

$$
\lim_{t \to \infty} p_0(t) =
\begin{cases}
\left(\dfrac{d}{b}\right)^i, & \text{if } b > d \\
1, & \text{if } b \leq d.
\end{cases}
\tag{8}
$$

If the difference $b - d$ is large, then the limit in (8) is reached more rapidly than if the difference $b - d$ is small (see Figure 3). In the case that the birth rate is less than the death rate, the probability of extinction approaches one. That is, the limiting stationary distribution of $X(t)$ as $t \to \infty$ is concentrated at zero.

For the example in Figure 2, the value of $p_0(10)$ is very close to the asymptotic limit. Applying the limiting formula in (8) to the simple birth and death process illustrated in Figure 2, the probability of population extinction is

$$
\left(\frac{d}{b}\right)^2 = \left(\frac{1.5}{2.0}\right)^2 = .5625.
$$

Recall that two out of five sample paths reached extinction in the computations.

The extinction results of the simple birth and death process are applicable to a wide range of population and epidemic processes. The limiting formula in (8), where $b > d$, approximates the initial dynamics of population and epidemic processes when there is exponential growth away from an absorbing state, demonstrated in Examples 1 and 2. In Example 3, we illustrate an application of the simple birth and death process to cancer therapy, when the birth rate is less than the death rate. In this example, the probability of extinction $p_0(t)$ approximates the time until cancer cells are eliminated.

3. Logistic growth

The well-known Verhulst logistic model (Verhulst, 1838) is given by the following non-linear differential equation

$$
\frac{dx(t)}{dt} = rx(t)\left(1 - \frac{x(t)}{K}\right), \quad x(0) = x_0 > 0,
\tag{9}
$$

where $r > 0$ is the intrinsic growth rate and $K > 0$ is the carrying capacity of the population. The growth declines at a rate proportional to the square of the population size. Notice that the birth and death rates in (9) are not explicitly defined. The solution of (9) can be written in terms of the parameters and the initial population size x_0 as

$$
x(t) = \frac{x_0 K}{x_0 + (K - x_0)e^{-rt}}, \quad t > 0.
$$

It follows that $\lim_{t \to \infty} x(t) = K$ so there is no possibility of extinction. However, if the population size is small, close to zero, the logistic Equation (9) is approximately an exponential growth model,

$$
\frac{dx(t)}{dt} \approx rx(t), \quad r > 0.
$$

To account for randomness in births and in death, we formulate a logistic Markov chain, $\{X(t), t \in [0, \infty)\}$, a continuous-time general birth and death process. The state space is a subset of the nonnegative integers and may be finite $X(t) \in \{0, 1, \ldots, N\}$ or infinite $X(t) \in \{0, 1, 2, \ldots\}$. We assume that in a small time interval Δt, only a birth, a death or no change may occur.

Let λ_i and μ_i be the birth and death rates given $X(t) = i$. The infinitesimal transition probabilities, $\mathbb{P}(X(t + \Delta t) = j | X(t) = i) = p_{i,j}(t)$ are similar to (2) but are not linear,

$$
p_{i,j}(\Delta t) = \begin{cases} \lambda_i \Delta t, & \text{if } j = i + 1 \\ \mu_i \Delta t, & \text{if } j = i - 1 \\ 1 - (\lambda_i + \mu_i)\Delta t, & \text{if } j = i. \end{cases} \tag{10}
$$

In the logistic model $dx/dt = \lambda_x - \mu_x$, where

$$
\lambda_x - \mu_x = rx - \frac{r}{K}x^2, \ \lambda_0 = \mu_0 = 0. \tag{11}
$$

In addition, the birth rate equals the death rate at $x = K$ since K is a steady-state solution,

$$
\lambda_K = \mu_K. \tag{12}
$$

The preceding assumptions (11) and (12) are satisfied for many choices of the birth and death rates (Allen, 2010). A reasonable assumption is that the birth and death rates are quadratic functions,

$$
\lambda_x = b_1 x + b_2 x^2 \geq 0, \ \mu_x = d_1 x + d_2 x^2 \geq 0. \tag{13}
$$

Parameters $b_1 > 0$, $d_1 \geq 0$ and $d_2 > 0$ but b_2 may be positive, negative or zero to ensure

$$
r = b_1 - d_1 > 0 \text{ and } K = \frac{b_1 - d_1}{d_2 - b_2} > 0.
$$

There are four parameters b_1, b_2, d_1 and d_2 in the stochastic logistic model but only two parameters r and K in the deterministic model. Therefore, there are infinitely many stochastic logistic models with birth and death rates explicitly defined that correspond to the same deterministic logistic model. For example, $b_1 = r + c$, $d_1 = c$, $b_2 = 0$ and $d_2 = r/K$, then λ_x is linear in x but μ_x is quadratic and both are nonnegative for $x \in \{0, 1, 2, \ldots\}$. Another example, is $b_1 = r$, $d_1 = 0$, $b_2 = -r/(2K)$ and $d_2 = r/(2K)$, where both λ_x and μ_x are quadratic and nonnegative for $x \in \{0, \ldots, 2K\}$. In this latter example, births and deaths are density-dependent.

Of course, growth in the stochastic logistic process differs significantly from the simple birth and death process when population sizes are large, but their behaviour is similar near the origin. In fact, when population sizes are small, the asymptotic limit in (8) is a good approximation to the probability of extinction.

Near the origin, the linear approximation of the logistic model has birth and death rates $b = b_1$ and $d = d_1$. The formula in (8), given $X(0) = x_0$, yields an estimate for the probability of population extinction for the stochastic logistic model. The estimate
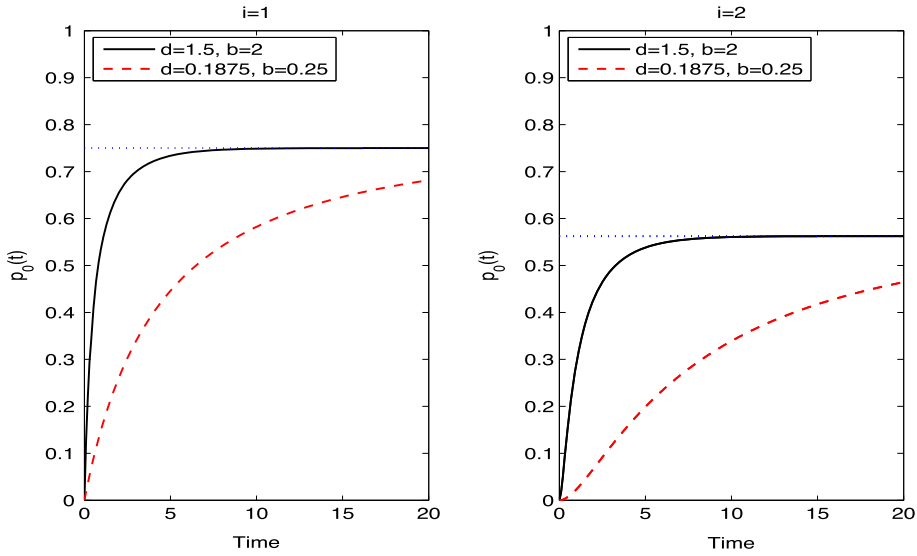
Figure 3. The probability of extinction, $p_0(t)$, is plotted when the initial population size $X(0) = i$ is either $i = 1, 2$ for two sets of values for d and b; $d = 1.5, b = 2$ and $d = .1875, b = .25$.

Notes: The limiting value for the probability of extinction is $(d/b)^i$. For the figure on the left $X(0) = 1$, the limit is $d/b = .75$ and for the figure on the right $X(0) = 2$, the limit is $(d/b)^2 = .5625$. When $d = 1.5, b = 2$, then $b - d = .5$ and the limit is reached quickly. When birth and death rates are decreased by a factor of 1/8, $d = .1875, b = .25$, then $b - d = .0625$ and the time until extinction is increased by a factor of 8.

depends on the initial population size x_0. The limit is denoted as $\mathbb{P}_0(x_0)$:

$$
\mathbb{P}_0(x_0) =
\begin{cases}
\left(\dfrac{d_1}{b_1}\right)^{x_0}, & \text{if } b_1 > d_1 \\
1, & \text{if } b_1 \leq d_1.
\end{cases}
$$

The case $b_1 \leq d_1$ never occurs in the logistic model since $r > 0$. The probability of extinction is always less than one but greater than zero for the stochastic logistic model.

3.1. Example 1: population extinction

Suppose the birth rate is linear and the death rate is a quadratic function,

$$
\lambda_x = b_1 x \text{ and } \mu_x = d_1 x + r\frac{x^2}{K}, \quad x = 0, 1, \cdots,
$$

where the two conditions (11) and (12) are satisfied, $b_1, d_1 > 0$ and

$$
\lambda_x - \mu_x = rx\left(1 - \frac{x}{K}\right), \quad x = 0, 1, \ldots.
$$

Three sample paths of the non-linear Markov chain model are graphed along with the deterministic logistic solution in Figure 4 when $K = 250$, $b_1 = 1$, $d_1 = .5$ and
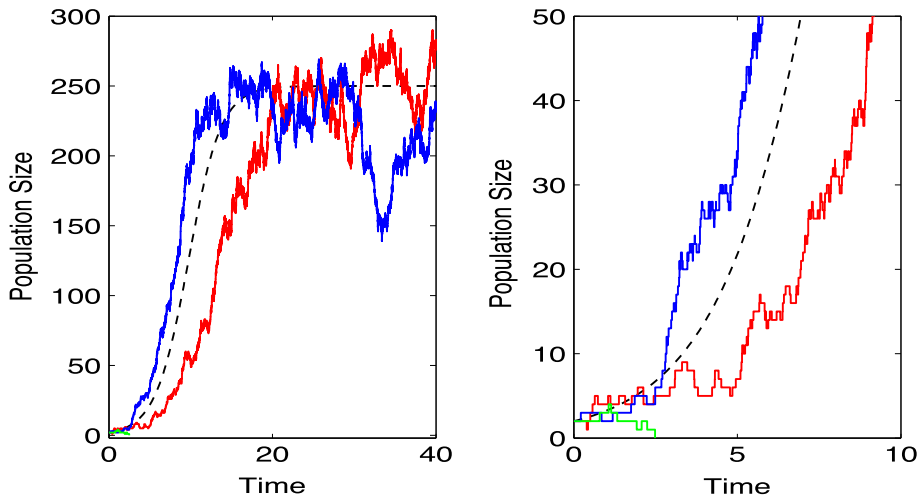
Figure 4. The deterministic solution (black dashed curve) along with three sample paths (red, blue and green curves) of the logistic model are plotted for $b_1 = 1$, $d_1 = .5$, $K = 250$ and $X(0) = 2$.
Notes: The figure on the left shows two of the sample paths that reach the carrying capacity K and the figure on the right shows a close-up of the same sample paths on the interval $[0, 10]$ with one hitting zero before $t = 5$. The probability of population extinction is $\mathbb{P}_0(2) = .25$.

$X(0) = x_0 = 2$. One of the three sample paths hits zero. The analytical approximation of the probability of extinction for this example is

$$\mathbb{P}_0(x_0) = \left(\frac{d_1}{b_1}\right)^{x_0} = \left(\frac{.5}{1.0}\right)^2 = .25. \tag{14}$$

Extinction occurs rapidly in this example.

To check the analytical estimate given in (14), we numerically simulated ten thousand sample paths of the stochastic logistic model using the MATLAB code in Appendix 2, until either the population size hits zero or the population reaches a size of 50. Values other than 50 can be chosen but 50 is reasonable since it is a sufficiently large value with exponential growth prior to 50 ($< K/2$). It is highly unlikely for the population to hit zero after reaching a size of 50. The simulations continue until the population size enters one of two subsets, either $\Omega_0 = \{X | X(t) = 0 \text{ for the first time } t \in (0, \infty)\}$ or $\Omega_{50} = \{X | X(t) = 50 \text{ for the first time } t \in (0, \infty)\}$. Then we count the number of sample paths in each of these sets. The union of these sets contains all ten thousand sample paths, $|\Omega_0 \cup \Omega_{50}| = 10^4$. For example, in one particular example, the numerical estimate $|\Omega_0|/10^4 = .2547$ which compares well with the analytical estimate given in (14).

This example illustrates the importance of a sufficiently large population size to prevent extinction. However, with only two parameters, the obvious conclusion is that extinction can be prevented by increasing births or by decreasing deaths. In practice, models for threatened or endangered species require much more complex models that depend on the particular species, the spatial environment, and the available options. Developing more comprehensive models to accurately portray the species dynamics in their natural environment pose significant mathematical, statistical and computational challenges (Green et al., 2005).

4. SIR epidemic model

The Kermack–McKendrick SIR epidemic model is a system of differential equations for the spread of an infectious disease in a population (Kermack & McKendrick, 1927). In the SIR model, an infected individual recovers and develops immunity which protects against reinfection. The model applies to diseases such as measles, mumps, rubella and chickenpox, where immunity occurs after infection. The total population N of individuals is divided into three groups: (S) susceptible, (I) infected and infectious, and (R) recovered. In the simplest case the population size is constant $N = S(t) + I(t) + R(t)$. There are no births or deaths, only infection and recovery. The system of differential equations for the SIR epidemic model is as follows:

$$\frac{dS(t)}{dt} = -\frac{\beta}{N}S(t)I(t),$$
$$\frac{dI(t)}{dt} = \frac{\beta}{N}S(t)I(t) - \gamma I(t), \tag{15}$$
$$\frac{dR(t)}{dt} = \gamma I(t),$$

where the initial conditions satisfy $S(0) > 0, I(0) > 0, R(0) \geq 0$ and $S(0)+I(0)+R(0) = N$. The parameter β is the transmission rate, the number of contacts per unit time that result in an infection of a susceptible individual, γ is the recovery rate, and $1/\gamma$ is the average length of the infectious period.

The dynamics of this model are well known and we summarize them here. Eventually, the disease dies out: $\lim_{t\to\infty} I(t) = 0$, $\lim_{t\to\infty} S(t) = S(\infty)$, and $\lim_{t\to\infty} R(t) = R(\infty)$. The effective reproduction number $\mathcal{R}$ and the basic reproduction number $\mathcal{R}_0$ are defined as

$$\mathcal{R} = \frac{S(0)}{N} \cdot \frac{\beta}{\gamma} \text{ and } \mathcal{R}_0 = \frac{\beta}{\gamma}.$$

If $\mathcal{R} \leq 1$, then there is no epidemic; solution $I(t)$ decreases monotonically to zero. If $\mathcal{R} > 1$, then $I(t)$ increases first before decreasing to zero; an epidemic occurs. The basic reproduction number $\mathcal{R}_0$ (introduced in 1927 by Kermack and McKendrick) is the number of secondary infections caused by introduction of one infectious individual into an entirely susceptible population.

Since $R(t) = N - S(t) - I(t)$, the system (15) can be reduced to a system of two differential equations:

$$\frac{dS(t)}{dt} = -\frac{\beta}{N}S(t)I(t),$$
$$\frac{dI(t)}{dt} = \frac{\beta}{N}S(t)I(t) - \gamma I(t). \tag{16}$$

Near the disease-free equilibrium $(S, I) = (N, 0)$, the infectious differential equation, dI/dt, in (16) is approximately

$$\frac{dI(t)}{dt} \approx (\beta - \gamma)I(t).$$

Therefore, near the disease-free equilibrium, the stochastic model for the infectious population will behave similarly to a birth and death process with birth rate $b = \beta$ and death rate $d = \gamma$.

Now, we construct a stochastic SIR model. The same notation $S(t)$ and $I(t)$ are used for the random variables. The SIR Markov chain model is a multivariate process, $\{(S(t), I(t)) | t \in [0, \infty)\}$, where $S(t)$ and $I(t)$ are discrete random variables for the number of susceptible and infectious individuals at time t. The values of the random variables satisfy

$$S(t), I(t) \in \{0, 1, 2, \ldots, N\} \text{ and } S(t) + I(t) \leq N.$$

Let $\mathbb{P}((S(t+\Delta t), I(t+\Delta t)) = (k, j) | (S(t), I(t)) = (s, i)) = p_{(s,i),(k,j)}(\Delta t)$. The infinitesimal transition probabilities, given $(S(t), I(t)) = (s, i)$, are

$$p_{(s,i),(k,j)}(\Delta t) = \begin{cases} \dfrac{\beta}{N} si \Delta t, & \text{if } (k, j) = (s-1, i+1), \\ \gamma i \Delta t, & \text{if } (k, j) = (s, i-1), \\ 1 - \left(\dfrac{\beta}{N} si + \gamma i \right) \Delta t, & \text{if } (k, j) = (s, i). \end{cases} \qquad (17)$$

The process is time-homogeneous.

Near the disease-free equilibrium, the Markov chain model can be approximated by a simple birth and death process (similar to the exponential growth model). Death of an infected individual corresponds to a recovery, $d = \gamma$, and birth of an infected individual corresponds to a new infection, $b = \beta$. Then

$$\frac{d}{b} = \frac{\gamma}{\beta} = \frac{1}{\mathcal{R}_0}.$$

The probability of disease extinction follows from (8) and the simple birth and death process. Given an initial infective population size of $I(0) = i_0$, the probability of no epidemic outbreak is approximately

$$\mathbb{P}_0(i_0) = \begin{cases} \left(\dfrac{1}{\mathcal{R}_0} \right)^{i_0}, & \text{if } \mathcal{R}_0 > 1, \\ 1, & \text{if } \mathcal{R}_0 \leq 1. \end{cases} \qquad (18)$$

The preceding formula is equivalent to the expression originally derived by Whittle (1955). An estimate for the probability of an outbreak is approximately $1 - \mathbb{P}_0(i_0)$.

4.1. Example 2: disease outbreaks

Suppose the infectious period on the average is 5 days ($\gamma = 1/5$), the total population size is $N = 800$ and the basic reproduction number is $\mathcal{R}_0 = 2$ ($\beta = 2/5$). Three sample paths of the SIR Markov chain model along with the deterministic solution are plotted in Figure 5 when $I(0) = i_0 = 2$ (The MATLAB program is given in Appendix 2). Applying (18), an estimate for the probability of no outbreak is

$$\mathbb{P}_0(i_0) = \left(\frac{1}{\mathcal{R}_0} \right)^{i_0} = \left(\frac{1}{2} \right)^2 = .25. \qquad (19)$$

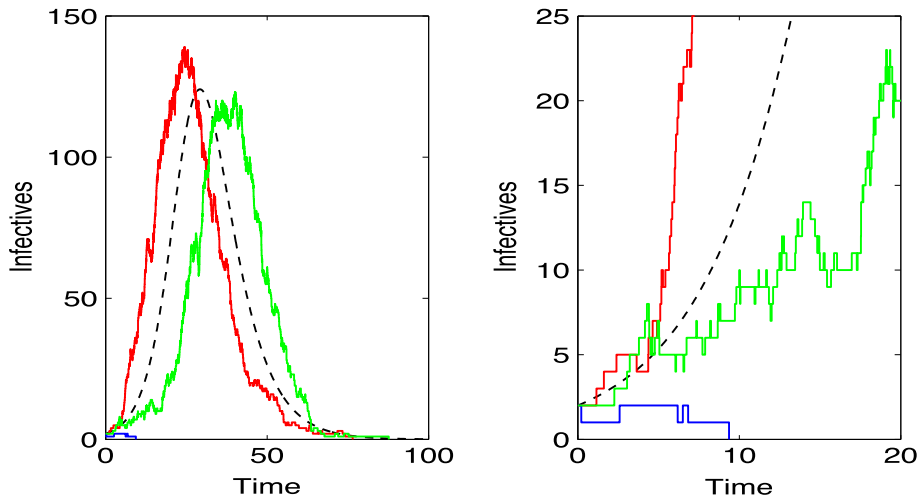An estimate for the probability of a disease outbreak is $1 - \mathbb{P}_0(2) = .75$.

Figure 5. The deterministic solution (black dashed curve) along with three sample paths (red, blue and green curves) of the infectives in the SIR model are plotted for parameter values $\beta = 2/5$, $\gamma = 1/5$ and $N = 800$ so that $\mathcal{R}_0 = 2$. The initial values are $S(0) = 798$, $I(0) = 2$ and $R(0) = 0$.

Notes: The figure on the left illustrates two sample paths with an epidemic outbreak and the figure on the right is a close-up of the same sample paths on the interval $[0, 20]$ with one sample path hitting zero before $t = 10$. The probability of a disease outbreak is $1 - \mathbb{P}_0(2) = .75$.

In the deterministic model, the total number of cumulative cases is 639 with a maximum total number of cases at any time reaching about 125. To check the analytical estimate in (19), ten thousand sample paths are numerically simulated until either the infectious population hits zero or reaches a size of 25 (MATLAB program in Appendix 2). Values other than 25 can be chosen but 25 is reasonable since there is exponential growth prior to 25, before the peak value is reached in the deterministic model. In realistic outbreaks, when the number of cases reaches 25, this would be considered a major outbreak. We count the number of sample paths that lie in one of two subsets: $\Omega_0 = \{I | I(t) = 0$ for the first time $t \in (0, \infty)\}$ or $\Omega_{25} = \{I | I(t) = 25$ for the first time $t \in (0, \infty)\}$. The numerical estimate for probability of disease extinction is $|\Omega_0|/10^4 = .2531$ for one set of ten thousand sample paths which is in good agreement with the analytical estimate in Equation (19).

Prevention of an outbreak is a major public health concern. Methods for prevention include prophylactic interventions (antibiotic and antiviral drugs), vaccination, quarantine and isolation. Studies on the emergence of new diseases such as SARS and Ebola and the re-emergence of old diseases such as drug resistant tuberculosis require more sophisticated models and the expertise of public health and medical professionals, statisticians and mathematicians. Mathematical models play an important role in testing the feasibility and the outcomes of various disease prevention and intervention strategies (e.g. Anderson & May, 1992; Brauer & Castillo-Chavez, 2010).

5. Cell populations

A cell divides and reproduces two daughter cells. During the cell cycle, mutations may occur, so that the daughter cells are not identical to the parent cell. The mutation can be neutral meaning the genetic sequences passed on to a daughter cell are neither beneficial nor detrimental to survival or reproduction. However, if the mutation impacts survival

or reproduction then 'fitness' has been altered. In an advantageous mutation, fitness is increased and either the reproduction cycle is shortened or the survival time is lengthened. On the other hand, in a deleterious mutation, fitness is decreased and the cell line does not survive. Most acquired mutations that occur in somatic stem cells (not egg or sperm) are neutral or deleterious. However, mutations that lead to cancer or tumour cells are generally advantageous, reproducing more rapidly than healthy cells. Cancer cells have the ability to renew and regenerate, a property typical of normal stem cells (Hardavella, George, & Sethi, 2016). These features of cancer cells have led to what is known as the cancer stem cell hypothesis: the existence of cancer stem cells that may be 'responsible for cancer initiation, progression, metastasis, recurrence and drug resistance' (Hardavella et al., 2016).

The simple birth and death process is directly applicable to cellular reproduction, where birth is cell division; two cells are reproduced from one via cell division, $p_{i,i+1}(\Delta t) = bi\Delta t$. Cell death is the transition $p_{i,i-1}(\Delta t) = di\Delta t$ (Equation (2)). Additional reasons that the birth and death process is a good approximation for cellular reproduction are that cell lines reproduce independently and the fitness of individual cell lines is constant, not frequency-dependent or density-dependent (Iwasa, Michor, & Nowak, 2004). Here, we consider the simple birth and death process as a model for two distinct cell lines, that is, normal healthy cells and cancer cells under chemotherapy treatment. The birth and death rates differ between the two cell lines when chemotherapy treatment targets the cancer cells (Sehl, Zhou, Sinsheimer, & Lange, 2011).

5.1. Example 3: cancer therapy

Suppose there are two cell lines: a normal healthy cell population H and a cancer cell population C. Healthy cells divide into two daughter cells at rate b_h or die at a rate d_h, whereas cancer cells divide at rate b_c or die at rate d_c. With chemotherapy treatment, both healthy cells and cancer cells are affected, but the treatment is designed to target the cancer cells. Therefore, we consider the case that with targeted treatment, the death rate for cancer cells is greater than for normal healthy cells.

In a recent model for cancer stem cell therapy, Sehl et al. (2011), apply a simple birth and death process to study the effects of treatment on cancer stem cells. The following birth and death rates were assumed in two examples for treatment of chronic leukaemia, $b_h = .02$/week, $b_c = b_h, d_h = .08$/week, with two different death rates for cancer stem cells: $d_c = .31$/week and $d_c = .59$/week (Sehl et al., 2011). In the absence of treatment, the death rate of healthy stem cells is about .002/day (Sehl et al., 2011). At the beginning of therapy, the number of cancer cells in the population is large, $X_h(0) = 4,400$ and $X_c(0) = 17,600$ (Sehl et al., 2011). Because $b_h < d_h$ and $b_c < d_c$, the probability of extinction approaches one (Equation (8)),

$$\lim_{t \to \infty} p_0(t) = 1,$$

for both healthy and cancer stem cells. The rate of convergence depends on $e^{(b-d)t}$. During treatment, an important question is the time to stop treatment. If chemotherapy treatment is prolonged, both healthy and cancer stem cells will be eliminated. Chemotherapy must be continued until most of the cancer stem cells are eliminated, but should cease before the healthy cell population has decreased to dangerously low levels. With cessation of treatment, the hope is that the healthy cells will resume normal growth with $b_h > d_h$.
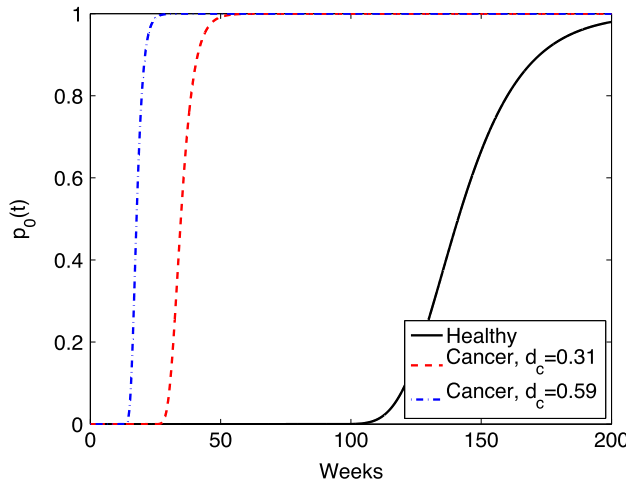
Figure 6. The cumulative distribution for the time to extinction for healthy cells H and cancer stem cells C under chemotherapy treatment.

Notes: Parameters for the birth and death rates (per week) are $b_h = .02 = b_c$, $d_h = .08$ and two different cancer cell death rates $d_c = .31$ and $d_c = .59$. The initial number of cells for each population is $X_h(0) = 4,400$ and $X_c(0) = 17,600$ (Sehl et al., 2011).

In a simple birth and death process, when the death rate exceeds the birth rate, $d > b$, the expression for the probability of extinction $p_0(t)$ in (7) is the cumulative distribution for the time to extinction. For the particular parameter values in the preceding examples, Figure 6 is a plot of $p_0(t)$ for healthy cells and for two cancer cell lines with different death rates. There is a distinct difference in time to extinction. Since $d_c \gg d_h$, most of the cancer stem cells are eliminated before the healthy cells are damaged (while $p_0(t)$ for the healthy cells is still small). The timing of treatment is a crucial part of chemotherapy. A higher death rate d_h of cancer cells shortens the time of treatment and reduces the damage to healthy cells. Increased understanding of these processes and the effects of treatment have been obtained by applying more complex stochastic models along with methods from multitype branching processes (see, e.g. Durrett, 2015; Kimmel & Axelrod, 2002).

6. Additional resources

Population extinction, disease outbreaks and cancer therapy represent three applications where the theory of birth and death processes provides greater insight into the extinction process. More complex models with several random variables representing multiple age classes, multiple disease stages or multiple types of cancer cells have been studied via methods from branching processes (e.g. Allen & van den Driessche, 2013; Athreya & Ney, 1972; Dorman, Sinsheimer, & Lang, 2004; Durrett, 2015; Haccou, Jager, & Vatutin, 2005; Jagers, 1975; Kimmel & Axelrod, 2002; Michor, Nowak, & Iwasa, 2006; Novozhilov, Karev, & Koonin, 2006). Interesting historical accounts about model development and applications to population dynamics and branching processes can be found in the book by Bacaër (2011) and in the Lecture Notes by Jagers (2009).

Acknowledgements

Disclosure statement

References

Allen, L. J. S. (2010). *An introduction to stochastic processes with applications to biology* (2nd ed.). Boca Raton, Fl: CRC Press.

Allen, L. J. S., & van den Driessche, P. (2013). Relations between deterministic and stochastic thresholds for disease extinction in continuous- and discrete-time infectious disease models. *Mathematical Biosciences, 243*, 99–108.

Anderson, R. M., & May, R. M. (1992). *Infectious diseases of humans: Dynamics and control.* Oxford: Oxford University Press.

Athreya, K. B., & Ney, P. E. (1972). *Branching process.* NewYork, NY: Springer-Verlag.

Bacaër, N. (2011). *A short history of mathematical population dynamics.* London: Spring-Verlag.

Bailey, N. T. J. (1990). *The elements of stochastic processes with applications to the natural sciences.* New York, NY: Wiley.

Bartholomay, A. F. (1958). On the linear birth and death processes of biology as Markoff chains. *Bulletin of Mathematical Biophysics, 20*, 97–118.

Bienaymé, I.-J. (1845). De la loi de multiplication et de la durée des familles [The law of multiplication and the duration of families]. *Société Philomathique de Paris, 5*, 37–39.

Brauer, F., & Castillo-Chavez, C. (2010). *Mathematical models in population biology and epidemiology* (2nd ed.). New York, NY: Springer.

Dorman, K. S., Sinsheimer, J. S., & Lang, K. (2004). In the garden of branching processes. *SIAM Review, 46*, 202–229.

Durrett, R. (2015). *Branching process models of cancer.* Mathematical Biosciences Institute Lecture Notes 1.1. Cham: Springer International Publishing.

Feller, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung [The Foundations of Volterra's Theory of the Struggle for Life in a Probabilistic Treatment]. *Acta Biotheoretica, 5*, 11–40.

Feller, W. (1968). *An introduction to probability theory and its applications* (3rd ed.). Wiley Series in Probability and Mathematical Statistics. New York, NY: Wiley.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry, 81*, 2340–2361.

Green, J. L., Hastings, A., Arzberger, P., Ayala, F. J., Cottingham, K. L., Cuddington, K., …Neubert, M. (2005). Complexity in ecology and conservation: Mathematical, statistical, and computational challenges. *BioScience, 55*, 501–510.

Haccou, P., Jager, P., & Vatutin, V. A. (2005). *Branching processes: Variation, growth, and extinction of populations*, Cambridge Studies in Adaptive Dynamics. Cambridge: Cambridge University Press.

Hardavella, G., George, R., & Sethi, T. (2016). Lung cancer stem cells-characteristics, phenotype. *Translational Lung Cancer Research, 5*, 272–279.

Iwasa, Y., Michor, F., & Nowak, M. A. (2004). Evolutionary dynamics of invasion and escape. *Journal of Theoretical Biology, 226*, 205–214.

Jagers, P. (1975). *Branching processes with biological applications.* London: Wiley.

Jagers, P. (2009, January 18–24). Some notes on the history of branching processes, from my perspective. Lecture at the Oberwolfach Symposium on 'Random Trees'.

Johnson, J. (1798). *An essay on the principle of population.* London: St. Paul's Church-yard.

Karlin, S., & Taylor, H. (1975). *A first course in stochastic processes* (2nd ed.). New York, NY: Academic Press.

Kendall, D. G. (1949). Stochastic processes and population growth. *Journal of the Royal Statistical Society, Series B., 11*, 230–282.

Kendall, D. G. (1956). Deterministic and stochastic epidemics in closed populations. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. (Vol. 4, pp. 149–165). Berkeley: University of California Press.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A, 11*, 700–721.

Kimmel, M., & Axelrod, D. E. (2002). *Branching processes in biology*. New York, NY: Springer.

Michor, F., Nowak, M. A., & Iwasa, Y. (2006). Evolution of resistance to cancer therapy. *Current Pharmaceutical Design, 12*, 261–271.

Novozhilov, A. S., Karev, G. P., & Koonin, E. V. (2006). Biological applications of the theory of birth-and-death processes. *Briefings in Biomathematics, 7*, 70–85.

Sehl, M., Zhou, H., Sinsheimer, J. S., & Lange, K. L. (2011). Extinction models for cancer stem cell therapy. *Mathematical Biosciences, 234*, 132–146.

Verhulst, P.-F. (1838). Notice sur la loi que la population poursuit dans son accroissement [Note on the law of population growth]. *Correspondance Mathématique et Physique, 10*, 113–121.

Watson, H. W., & Galton, F. (1875). On the probability of the extinction of families. *Journal of the Anthropological Institute of Great Britain and Ireland, 4*, 138–144.

Whittle, P. (1955). The outcome of a stochastic epidemic: A note on Bailey's paper. *Biometrika, 42*, 116–122.

Appendix 1. Derivation of probability of extinction

We derive the forward Kolmogorov differential equations, often referred to as the *master equations*. Consider the changes that occur during time intervals $[0, t]$ and $[t, t + \Delta t]$. In the time interval $[0, t]$, we assume the process has already made a transition to either states $j - 1$, j or $j + 1$ with probabilities $p_{i,j-1}(t)$, $p_{i,j}(t)$ or $p_{i,j+1}(t)$, respectively. Thus, we only consider a birth, death or no change in the time interval $[t, t + \Delta t]$. The probability of a transition in a small interval of time Δt comes directly from the probabilities defined in (2). For example, it follows from the Markov property (independence of past history) that the probability of transitioning from i to $j - 1$ in time t and from $j - 1$ to j in the time interval Δt is the product of $p_{i,j-1}(t)$ and $b(j - 1)\Delta t$. Considering all possible cases, $j - 1$, j and $j + 1$ we obtain the following equation for the probability $p_{i,j}(t + \Delta t)$:

$$p_{i,j}(t + \Delta t) = p_{i,j-1}(t)b(j - 1)\Delta t + p_{i,j}(t)[1 - (bj + dj)\Delta t] + p_{i,j+1}(t)d(j + 1)\Delta t$$

for $i, j = 0, 1, 2, 3, \ldots$. Subtracting $p_{i,j}(t)$ from both sides of the preceding expression, dividing by Δt, and letting $\Delta t \to 0$ lead to the forward Kolmogorov differential equations for $j = 0, 1, 2, \ldots$, $i = 1, 2, \ldots$ (Allen, 2010; Karlin and Taylor, 1975):

$$\frac{\mathrm{d}p_{i,j}(t)}{\mathrm{d}t} = b(j - 1)p_{i,j-1}(t) - (b + d)jp_{i,j}(t) + d(j + 1)p_{i,j+1}(t), \tag{A1}$$

where $p_{i,j}(t) = 0$ if i or j are negative integers. If $X(0) = i \geq 1$, then we drop the dependence on i because of the identity (3) to obtain a differential equation for $p_j(t), j = 0, 1, 2, \ldots$.

$$\frac{\mathrm{d}p_j(t)}{\mathrm{d}t} = b(j - 1)p_{j-1}(t) - (b + d)jp_j(t) + d(j + 1)p_{j+1}(t). \tag{A2}$$

where $p_{-1}(t) = 0$.

   The forward Kolmogorov differential equations are easier to solve if they are expressed in terms of generating functions. The probability generating function (pgf) of $X(t)$ is formally defined in terms of the probabilities in (3) as the expectation of $s^{X(t)}$. Therefore, a new variable s is introduced.

In particular,

$$\mathcal{P}(s,t) = \mathbb{E}(s^{X(t)}) = \sum_{j=0}^{\infty} p_j(t)s^j, \quad s \in [0,1].$$

The pgf 'generates' the probabilities $p_j(t)$ as well as the mean and variance associated with the random variable $X(t)$. For example, the first two probabilities are

$$\mathcal{P}(0,t) = p_0(t) \quad \text{and} \quad \left.\frac{\partial \mathcal{P}(s,t)}{\partial s}\right|_{s=0} = p_1(t)$$

and the mean of $X(t)$ is

$$\left.\frac{\partial \mathcal{P}(s,t)}{\partial s}\right|_{s=1} = \sum_{j=1}^{\infty} jp_j(t).$$

If we multiply Equation (A2) by s^j and sum from 0 to ∞, then the infinite system of forward Kolmogorov differential equations reduces to a single differential equation,

$$\frac{\partial \mathcal{P}(s,t)}{\partial t} = (b+d)\left[\frac{d+bs^2}{b+d} - s\right]\frac{\partial \mathcal{P}(s,t)}{\partial s}. \tag{A3}$$

Boundary conditions associated with this partial differential equation are $\mathcal{P}(1,t) = 1$ and $\mathcal{P}(s,0) = s$, i.e. $X(0) = 1$. The simplification given in (A3) requires two facts, summation and differentiation can be interchanged and

$$\frac{\partial \mathcal{P}(s,t)}{\partial s} = \sum_{j=1}^{\infty} jp_j(t)s^{j-1}$$

(the summation converges absolutely on $s \in [0,1]$). The first-order partial differential equation in (A3) can be solved by the method of characteristics. Details of the method of characteristics applied to (A3) can be found in Allen (2010). A major assumption in the simple birth and death process is that each individual behaves independently. Thus, if $X(0) = i$, then $\mathcal{P}(s,0) = \mathbb{E}(s^i) = (\mathbb{E}(s))^i$, the expectation of the product is the product of expectations. Thus, the solution of (A3) for the simple birth and death process for $X(0) = i$ is

$$\mathcal{P}(s,t) = \begin{cases} \left(\dfrac{e^{(d-b)t}(bs-d) - d(s-1)}{e^{(d-b)t}(bs-d) - b(s-1)}\right)^i, & \text{if } b \neq d \\[4mm] \left(\dfrac{1 - (bt-1)(s-1)}{1 - bt(s-1)}\right)^i, & \text{if } b = d. \end{cases}$$

The simple birth and death process has one of two asymptotic outcomes as $t \to \infty$, either the process approaches zero (absorption) or the process approaches infinity. The time until extinction depends on the value of $\exp((d-b)t)$. This can be seen by computing the probability of extinction. Evaluating $\mathcal{P}(s,t)$ at $s = 0$ leads to a formula for $p_0(t)$:

$$p_0(t) = \begin{cases} \left(\dfrac{d - de^{(d-b)t}}{b - de^{(d-b)t}}\right)^i, & \text{if } b \neq d \\[4mm] \left(\dfrac{bt}{1+bt}\right)^i, & \text{if } b = d. \end{cases}$$

Appendix 2. MATLAB programs

```
% FIGURE 2: Exponential Growth & Birth and Death
set(0,'DefaultAxesFontSize',18)
set(gca,'fontsize',18);
X0=2; b=2; d=1.5; % Initial value and parameters
x=[0:.1:8];
y=X0*exp((b-d).*x);% Deterministic solution
plot(x,y,'k--','Linewidth',2);
axis([0,8,0,100]);
hold on
for k=1:5 % Five Sample Paths
    clear t n
    t(1)=0; n(1)=X0;
    j=1;
    while n(j)>0 & n(j)<100 % Stop when hits zero or reaches size=100
        u1=rand; u2=rand; % Two uniform random numbers
        t(j+1)=-log(u1)/(b*n(j)+d*n(j))+t(j); % Interevent time
        if u2<b/(b+d)
            n(j+1)=n(j)+1; % Birth
        else
            n(j+1)=n(j)-1; % Death
        end
        j=j+1;
    end
    stairs(t,n,'-','Linewidth',2,'color',rand(1,3));
end
hold off
xlabel('Time'); ylabel('Population size');
axis([0,8,-.5,100]);

% FIGURE 4: Example 1 Logistic  Growth
set(0,'DefaultAxesFontSize',18)
set(gca,'fontsize',18);
b1=1.0; d1=0.5;r=b1-d1; K=250; X0=2; time=40;% Parameters and Initial values
x=[0:.1:time];
y=X0*K./(X0+(K-X0)*exp(-r*x));
for k1=1:2
  subplot(1,2,k1)
  plot(x,y,'k--','linewidth',2);
  axis([0,40/k1^2,0,300/k1^(log(6)/log(2))]);
  xlabel('Time'); ylabel('Population size');
  hold on
end
for k2=1:3 % 3 sample paths
  clear  t  n
  t(1)=0; n(1)=X0;
  j=1;
  while n(j)>0. & t(j)<time; % Stop when size hits zero or time reaches 40
    u1=rand;u2=rand; % Two uniform random numbers
    lam=b1*n(j);
    mu=d1*n(j)+r*n(j)^2/K;
    tot=lam+mu;
```

```
    t(j+1)=t(j)-log(u1)/(tot); % Interevent time
    if u2<lam/tot ;
      n(j+1)=n(j)+1;
    else
      n(j+1)=n(j)-1;
    end
      j=j+1;
    end
    for k3=1:2
      subplot(1,2,k3)
      stairs(t,n,'-','linewidth',2,'color',rand(1,3));
    end
end
% Estimate Probability of Extinction in Logistic Growth
count=0;
for k4=1:10000 % Number of sample paths
  clear   n
  n=X0;
  while n>0 & n<50; % Stop when size either hits zero or 50
    u=rand;
    lam=b1*n;
    mu=d1*n+r*n^2/K;
    tot=lam+mu;
    if u<lam/tot
      n=n+1; % Birth
    else
      n=n-1; % Death
    end
  end
  if n==0
     count=count+1;
  end
end
probext=count/10000
estext=(d1/b1)^(X0)

%  FIGURE 5: Example 2 SIR Epidemic Model
set(0,'DefaultAxesFontSize',18)
set(gca,'fontsize',18);
gam=0.2; beta=0.4; N=800; i0=2 % Parameters and initial values
dt=.05; tim=100; time=tim/dt; i(1)=i0; s(1)=N-i0; sumc=0;
for tt=1:time %Euler's method for solving ODE
    i(tt+1)=i(tt)+dt*((beta/N)*i(tt)*s(tt)-gam*i(tt));
    s(tt+1)=s(tt)+dt*(-(beta/N)*i(tt)*s(tt));
    sumc=sumc+dt*(beta/N)*i(tt)*s(tt);
end
TotalCases=round(sumc)+i0 % Count number of cases in ODE
for k1=1:2
  subplot(1,2,k1)
  plot([0:dt:tim],i,'k--','linewidth',2);
  axis([0,100/k1^(log(5)/log(2)),0,150/k1^(log(7)/log(2))]);
  xlabel('Time'); ylabel('Infectives');
  hold on
end
```

```
for k2=1:3 % Sample paths for SIR Markov chain
   clear  t i s
   t(1)=0; i(1)=i0; s(1)=N-i0;
   j=1;
   while i(j)>0 & t(j)<time % Stop when infectives hit zero or  time=100
      u1=rand;u2=rand;
      t(j+1)=-log(u1)/((beta/N)*i(j)*s(j)+gam*i(j))+t(j);
      if (u2<=(beta/N)*s(j)/(beta/N*s(j)+gam))
         i(j+1)=i(j)+1;
         s(j+1)=s(j)-1;
      else
         i(j+1)=i(j)-1;
         s(j+1)=s(j);
      end
        j=j+1;
   end
   for k3=1:2
       subplot(1,2,k3)
       stairs(t,i,'-','linewidth',2,'color',rand(1,3));
   end
end
% Estimate Probability of Epidemic Extinction in SIR model
count=0;
for k4=1:10000 % Number of sample paths
   clear  t s i
   i=i0; s=N-i0;
   j=1;
   while i>0 & i<25 % Stop when size either hits zero or reaches 25
      u1=rand; u2=rand;
      if (u2<=(beta/N)*s/(beta/N*s+gam))
         i=i+1;
         s=s-1;
      else
         i=i-1;
         s=s;
      end
      j=j+1;
    end
    if i==0
      count=count+1;
    end
end
probext=count/10000
estext=(gam/beta)^i0
```