10.2: The Birth-Death Model

A birth-death model is a continuous-time Markov process that is often used to study how the number of individuals in a population change through time. For macroevolution, these "individuals" are usually species, sometimes called "lineages" in the literature. In a birth-death model, two things can occur: births, where the number of individuals increases by one; and deaths, where the number of individuals decreases by one. We assume that no more than one new individual can form (or die) during any one event. In phylogenetic terms, that means that birth-death trees cannot have "hard polytomies" - each speciation event results in exactly two descendant species.

In macroevolution, we apply the birth-death model to species, and typically consider a model where each species has a constant probability of either giving birth (speciating) or dying (going extinct). We denote the per-lineage birth rate as $\lambda$ and the per-lineage death rate as $\mu$. For now we consider these rates to be constant, but we will relax that assumption later in the book.

We can understand the behavior of birth-death models if we consider the waiting time between successive speciation and extinction events in the tree. Imagine that we are considering a single lineage that exists at time $t_0$. We can think about the waiting time to the next event, which will either be a speciation event splitting that lineage into two (Figure 10.2A) or an extinction event marking the end of that lineage (Figure 10.2B). Under a birth-death model, both of these events follow a Poisson process, so that the expected waiting time to an event follows an exponential distribution (Figure 10.2C). The expected waiting time to the next speciation event is exponential with parameter $\lambda$, and the expected waiting time to the next extinction event exponential with parameter $\mu$. [Of course, only one of these can be the next event. The expected waiting time to the next event (of any sort) is exponential with parameter $\lambda + \mu$, and the probability that that event is speciation is $\lambda/(\mu + \lambda)$, extinction $\mu/(\mu + \lambda)$].

Figure 10.2. Illustration of the basic properties of birth-death models. A. Waiting time to a speciation event; B. Waiting time to an extinction event; C. Exponential distribution of waiting times until the next event; D. A
When we have more than one lineage “alive” in the tree at any time point, then the waiting time to the next event changes, although its distribution is still exponential. In general, if there are \( N(t) \) lineages alive at time \( t \), then the waiting time to the next event follows an exponential distribution with parameter \( N(t)(\lambda + \mu) \), with the probability that that event is speciation or extinction the same as given above. You can see from this equation that the rate parameter of the exponential distribution gets larger as the number of lineages increases. This means that the expected waiting times across all lineages get shorter and shorter as more lineages accumulate.

Using this approach, we can grow phylogenetic trees of any size (Figure 10.2D).

We can derive some important properties of the birth-death process on trees. To do so, it is useful to define two additional parameters, the net diversification rate \( (r) \) and the relative extinction rate \( (\epsilon) \):

\[
\begin{align*}
\epsilon &= \frac{\mu}{\lambda} \label{10.1B} \\
\end{align*}
\]

These two parameters simplify some of the equations below, and are also commonly encountered in the literature.

To derive some general properties of the birth-death model, we first consider the process over a small interval of time, \( \Delta t \). We assume that this interval is so short that it contains at most a single event, either speciation or extinction (the interval might also contain no events at all). The probability of speciation and extinction over the time interval can be expressed as:

\[
\begin{align*}
Pr_{\text{speciation}} &= N(t)\lambda \Delta t \label{10.2A} \\
Pr_{\text{extinction}} &= N(t)\mu \Delta t \label{10.2B} \\
\end{align*}
\]

We now consider the total number of living species at some time \( t \), and write this as \( N(t) \). It is useful to think about the expected value of \( N(t) \) under a birth-death model [we consider the full distribution of \( N(t) \) below]. The expected value of \( N(t) \) after a small time interval \( \Delta t \) is:

\[
\begin{align*}
\mathbb{E}[N(t + \Delta t)] &= N(t)\lambda \Delta t - N(t)\mu \Delta t \label{10.3} \\
\end{align*}
\]

We can convert this to a differential equation by subtracting \( N(t) \) from both sides, then dividing by \( \Delta t \) and taking the limit as \( \Delta t \) becomes very small:

\[
\begin{align*}
\frac{dN}{dt} &= N(\lambda - \mu) \label{10.4} \\
\end{align*}
\]

We can solve this differential equation if we set a boundary condition that \( N(0) = n_0 \); that is, at time 0, we start out with \( n_0 \).
lineages. We then obtain:

\[ N(t) = n_0 e^{(\lambda - \mu)t} \]

This deterministic equation gives us the expected value for the number of species through time under a birth-death model. Notice that the number of species grows exponentially through time as long as \( \lambda > \mu \), i.e. \( r > 0 \), and decays otherwise (Figure 10.3).

Figure 10.3. Expected number of species under a birth-death model with \( r = \lambda - \mu > 0 \) (top line), \( r = 0 \) (middle line), and \( r < 0 \) (bottom line). In each case the starting number of species was \( n_0 = 1000 \). Image by the author, can be reused under a CC-BY-4.0 license.

We are also interested in the stochastic behavior of the model – that is, how much should we expect \( N(t) \) to vary from one replicate to the next? We can calculate the full probability distribution for \( N(t) \), which we write as \( p_n(t) = Pr[N(t) = n] \) for all \( n \geq 0 \), to completely describe the birth-death model’s behavior. To derive this probability distribution, we can start with a set of equations, one for each value of \( n \), to keep track of the probabilities of \( n \) lineages alive at time \( t \). We will denote each of these as \( p_n(t) \) (there are an infinite set of such equations, from \( p_0 \) to \( p_\infty \)). We can then write a set of difference equations that describe the different ways that one can reach any state over some small time interval \( \Delta t \). We again assume that \( \Delta t \) is sufficiently small that at most one event (a birth or a death) can occur. As an example, consider what can happen to make \( n = 0 \) at the end of a certain time interval. There are two possibilities: either we were already at \( n = 0 \) at the beginning of the time interval and (by definition) nothing happened, or we were at \( n = 1 \) and the last surviving lineage went extinct. We write this as:

(eq. 10.6)

\[ p_0(t + \Delta t) = p_1(t) \mu \Delta t + p_0(t) \]

Similarly, we can reach \( n = 1 \) by either starting with \( n = 1 \) and having no events, or going from \( n = 2 \) via extinction.

(eq. 10.7)

\[ p_1(t + \Delta t) = p_1(t)(1 - (\lambda + \mu))\Delta t + p_2(t)2\mu \Delta t \]

Finally, any for \( n > 1 \), we can reach the state of \( n \) lineages in three ways: from a birth (from \( n - 1 \) to \( n \)), a death (from \( n + 1 \) to \( n \)), or neither (from \( n \) to \( n \)).

(eq. 10.8)
\[ p_n(t + \Delta t) = p_{n-1}(t)(n-1)\lambda\Delta t + \mu p_n(t) + p_{n+1}(t)(n\mu + 1(1-n(\lambda + \mu))\Delta t \]

We can convert this set of difference equations to differential equations by subtracting \( p_n(t) \) from both sides, then dividing by \( \Delta t \) and taking the limit as \( \Delta t \) becomes very small. So, when \( n = 0 \), we use 10.6 to obtain:

(eq. 10.9)

\[
\frac{dp_0(t)}{dt} = \mu p_1(t)
\]

From 10.7:

(eq. 10.10)

\[
\frac{dp_1(t)}{dt} = 2\mu p_2(t) - (\lambda - \mu) p_1(t)
\]

and from 10.8, for all \( n > 1 \):

(eq. 10.11)

\[
\frac{dp_n(t)}{dt} = (n-1)\lambda p_{n-1}(t) + (n+1)\mu p_n(t) - n(\lambda + \mu) p_n(t)
\]

We can then solve this set of differential equations to obtain the probability distribution of \( p_n(t) \). Using the same boundary condition, \( N(0) = n_0 \), we have \( p_0(t) = 1 \) for \( n = n_0 \) and 0 otherwise. Then, we can find the solution to the differential equations 10.9, 10.10, and 10.11. The derivation of the solution to this set of differential equations is beyond the scope of this book (but see Kot (2001) for a nice explanation of the mathematics). A solution was first obtained by Bailey (1964), but I will use the simpler equivalent form from Foote et al. (1999). For \( p_0(t) \) – that is, the probability that the entire lineage has gone extinct at time \( t \) – we have:

(eq. 10.12)

\[
p_0(t) = a_0^n
\]

And for all \( n \geq 1 \):

(eq. 10.13)

\[
\sum_{j=1}^{\min(n_0,n)} \binom{n_0}{j} \binom{n-1}{j-1} \alpha^{n_0 - j} \beta^{n-j} (1-\alpha)(1-\beta)^j
\]

Where \( \alpha \) and \( \beta \) are defined as:

(eq. 10.14)
$$\alpha=\frac{\epsilon(e^{rt}-1)}{(e^{rt}-\epsilon)}$$

$$\beta =\frac{(e^{rt}-1)}{(e^{rt}-\epsilon)}$$

$\alpha$ is the probability that any particular lineage has gone extinct before time $t$.

Note that when $n_0 = 1$ – that is, when we start with a single lineage - equations 10.12 and 10.13 simplify to (Raup 1985):

(eq. 10.15)

$$p_0(t)=\alpha$$

And for all $n \geq 1$:

(eq. 10.16)

$$p_n(t)=(1-\alpha)(1-\beta)^{n-1}$$

In all cases the expected number of lineages in the tree is exactly as stated above in equation (10.5), but now we have the full probability distribution of the number of lineages given $n_0$, $t$, $\lambda$, and $\mu$. A few plots capture the general shape of this distribution (Figure 10.4).

Figure 10.4. Probability distributions of $N(t)$ under A. pure birth, B. birth death after a short time, and C. birth-death after a long time. Image by the author, can be reused under a CC-BY-4.0 license.

There are quite a few comparative methods that use clade species richness and age along with the distributions defined in 10.14 and 10.15 to make inferences about clade diversification rates (see chapter 11).