Topic 5b: Continuous Time/Discrete-Space Stochastic Pocesses

Lecture 5.5 Master Equations

Review:

- CTDS stochastic processes can be visualized using a transition diagram
- CTDS stochastic processes can be formulated using a transition rate matrix
- The right-eigenvector, $\vec{v}_1$, determines the stationary distribution of the CTDS process
- Models of molecular evolution are useful examples of CTDS processes
- We can use molecular evolution models to formulate the probability a given tree gave rise to the observed sequences (the tree likelihood).
- Stochastic mapping using molecular models to predict ancestral states and simulate mutations on a tree.

So far we have considered only two methods for analyzing a CTDS processes: simulation and stationary distributions. The next few lectures will discuss alternative analysis approaches.

Definition

A **master equation** is a differential equation describing the dynamics of the state probability distribution, $P_n(t)$ for a CTDS stochastic process.

Solving the system of master equations for all states in the process then, gives the **exact solution** for the predicted outcome represents the density of an infinite number of simulation runs.

Let $P_n(t)$ be the probability that the system is in state n at time t.

$$\frac{dP_n(t)}{dt} = \underbrace{-P_n(t)\sum_{i\neq n} q_{n,i}}_{\text{flow out}} + \underbrace{\sum_{i\neq n} P_i(t)q_{i,n}}_{\text{flow in}}$$
$$= P_n(t)q_{n,n} + \sum_{i\neq n} P_i(t)q_{i,n} = \sum_i P_i(t)q_{i,n}$$

Example: 5.12 JC69

1. Formulate the system of master equations describing the dynamics of JC69 model.

The system is particularly easy in this case as the rates are independent of the state.

$$\frac{dP_n}{dt} = -3\mu P_n(t) + \sum_{i\neq n} \mu P_i(t) \quad n = \{1, 2, 3, 4\}$$

2. Assuming that the system starts in state i what is the probability of being in state j after t units of time?

We can fortunately solve this system of equations exactly:

$$P_n(t) = \left(\frac{1}{4} + \frac{3}{4}e^{-4\mu t}\right) P_n(0) + \left(\frac{1}{4} - \frac{1}{4}e^{-4\mu t}\right) \sum_{i\neq n} P_i(0)$$

$$P_{i,j}(t) = P_j(t|P_i(0) = 1) = \begin{cases} \left(\frac{1}{4} + \frac{3}{4}e^{-4\mu t}\right) & \text{if } i = j \\ \left(\frac{1}{4} - \frac{1}{4}e^{-4\mu t}\right) & \text{if } i \neq j \end{cases}$$

Given that this has a known general solution makes it MUCH faster and easier to calculate tree likelihoods under this model as we don't need to do matrix exponentiation to get $P_{i,j}(t)$.

3. After 1 unit of time, how close is the system to the stationary distribution?

Note that as $t \to \infty$ we have:

$$\hat{P}_n = \left(\frac{1}{4}\right) P_n(0) + \left(\frac{1}{4}\right) \sum_{i \neq n} P_i(0) = \frac{1}{4}$$

which is what we knew to be the stationary distribution. After one unit of time we have:

$$P_n(1) = \left(\frac{1}{4} + \frac{3}{4}e^{-4\mu}\right) p_n(0) + \left(\frac{1}{4} - \frac{1}{4}e^{-4\mu}\right) \sum_{i \neq n} p_i(0)$$

So using $\left|P_n(1) - \hat{P}_n\right|$ as a measure of closeness we have:

$$\left|P_n(1) - \hat{P}_n\right| = \frac{3}{4}e^{-4\mu} P_n(0) + \frac{1}{4}e^{-4\mu} P_i(0)$$

Often there is no known general solution and we must solve the dynamics numerically.

Example: 5.13 SIS model

Python: Lecture5_5.ipynb

Consider an SIS stochastic epidemic where transmissions occur at a mass-action rate of $\frac{\beta}{\kappa} * S * I$ with $\beta = 0.5$, hosts recover (becoming susceptible again) at a rate $\gamma = 0.1$ and the total population size is $\kappa = 100$.

1. Numerically solve the system of master equations for $t = [0, 50]$ assuming that initially there are $I(0) = 5$ infections.

We can describe this process with a 1D state space counting the number of infected hosts between 0 and 100. If the number of infected hosts is given by the state n, the number of susceptible hosts is given by $\kappa - n$.

To derive the master equation note that the system can only transition between neighbouring states through recovery and infection. The system leaves state n via infection (at rate $\frac{\beta}{\kappa}n(\kappa - n)$ going to state $n + 1$) and recovery (at rate γn going to state $n - 1$) and enters state n from state $n - 1$ by infection (at rate $\frac{\beta}{\kappa}(n - 1)(\kappa - n + 1)$) and via recovery from state $n + 1$ (at rate $\gamma(n + 1)$).

The system of master equations is:

$$\frac{dP_n(t)}{dt} = -P_n(t)\left(\gamma n + \frac{\beta}{\kappa}n(\kappa - n)\right) + P_{n-1}(t)\frac{\beta}{\kappa}(n - 1)(\kappa - n + 1) + P_{n+1}(t)\gamma(n + 1)$$
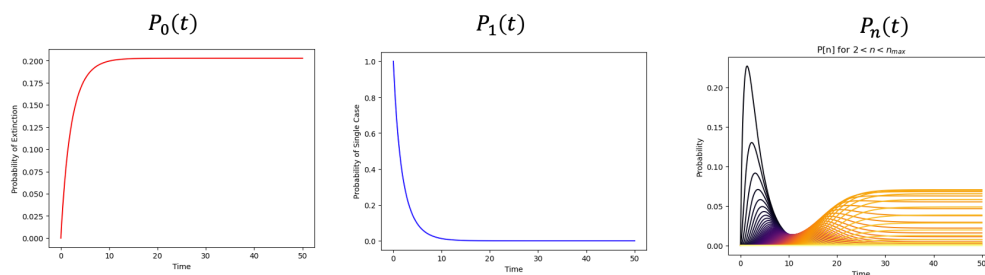
## 2. Use the master equations to calculate the probability that the disease goes extinct by time t.

The disease goes extinct if $I(t) = n = 0$.

When $n = 0$ the Master equation reduces to:

$$\frac{dP_0(t)}{dt} = P_1(t)\gamma \quad \text{where} P_0(0) = 0 \text{ and } P_1(0) = 1$$

$$P_0(t) = \int_0^t \gamma P_1(x)dx$$

3. Simulate 100 trajectories from the same stochastic process. How do your results compare to the solution from the master equations? What are some advantages/disadvantages to simulations versus master equations?
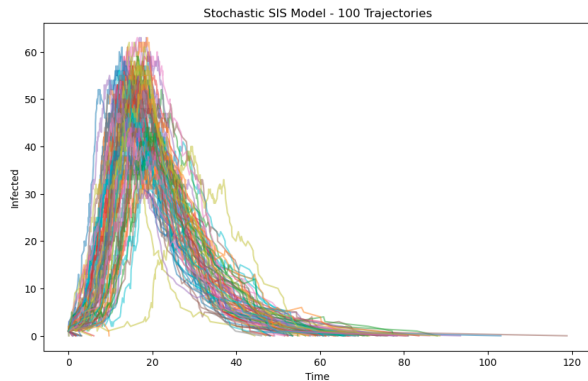
In many ways the trajectories are faster, they also give us an idea of the dynamics of individual outcomes in a way that the master equation do not. But they are finite in number and hence do not provide an exact result. Plus we would have to infer from this data the probability of extinction for example. If this probability is small it can be hard to obtain accurately from trajectories.

The state space of many biological stochastic processes may be more conveniently viewed in 2 dimensions rather than a single dimension (note that a 2 D discrete state space can always be mapped into 1D but the mapping might be complicated an inconvenient). As an example of this consider the SIR epidemic model

---

Example: 5.14 SIR model

Python: Lecture5_5.ipynb

Consider an SIR stochastic epidemic where transmissions occur at a mass-action rate of $\frac{\beta}{\kappa} * S * I$ with $\beta = 0.5$, hosts recover (gaining permanent resistance) at a rate $\gamma = 0.1$ and the total population size is $\kappa = 20$.

1. Numerically solve the system of master equations for $t = [0, 50]$ assuming that initially there are $I(0) = 5$ infections.
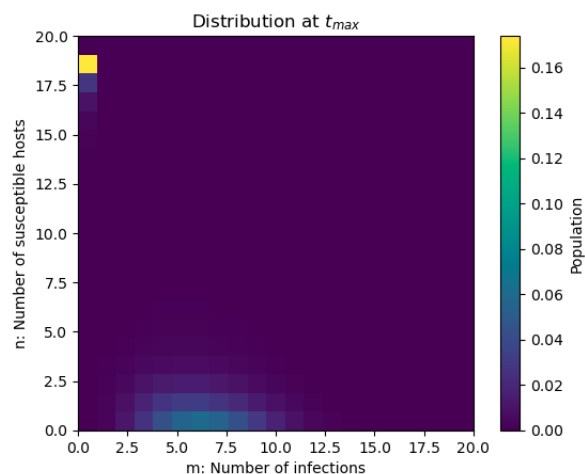
Let $P_{n,m}(t)$ be the probability that the system has $S(t) = n$ and $I(t) = m$ individuals present at time t. Then the process is described by the following events:

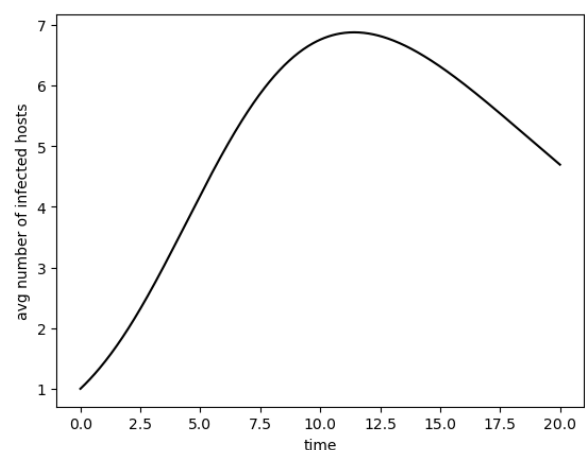| Event Name	Rate	Effect $\{\Delta n, \Delta m\}$
Transmission	$\frac{\beta}{\kappa}nm$	$\{-1, 1\}$
Recovery	γm	$\{0, -1\}$

The master equations are given by:

$$\frac{dP_{n,m}(t)}{dt} = -\left(\frac{\beta}{\kappa}nm + \gamma m\right)P_{n,m}(t) + \frac{\beta}{\kappa}(n+1)(m-1)P_{n+1,m-1} + \gamma(m+1)P_{n,m+1}$$
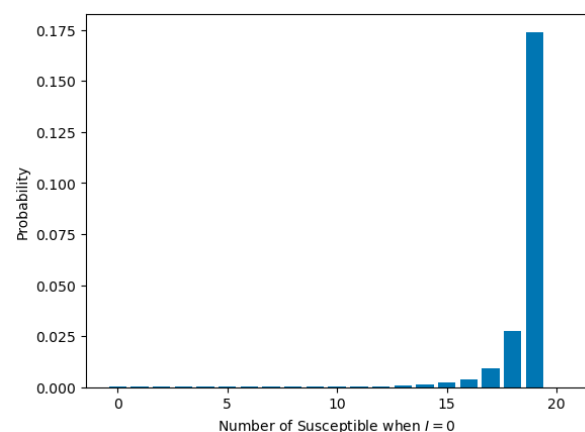
We can look at the distribution at the end of the integration period:

Or we can look at the average number of infected hosts as calculated by $\sum_{n,m} m P_{n,m}$

Or the number of susceptible hosts left when the disease goes extinct (first column of the matrix above)

Lecture 5.6 Ensemble Moment Approximations

In the previous lecture we saw how to derive a system of differential equations describing the dynamics of the state-space distribution $P_n(t)$. While providing a convenient and understandable solution in many cases, the state space $n \in \mathcal{S}$ may be very (or even infinity) large in which case approximations must be used to obtain a tractable system. Even in small or well approximated systems, subsequent interpretation of the master equation results can be cumbersome. In such cases the Ensemble Moment Approximation (EMA) can provide a useful alternative or complementary approach.

The EMA provides an approximation for the dynamics of the moments (either centered or non-centered) of the state space distribution.

Example 5.15 Moment notation

How would one express the first two raw moments and first two centered moments of the state space distribution? In a

two dimensional state-space, indicate the additional moments.

Given that the derivation of this approximation involves the use of alot of expectations we use a shorthand bracket notation:
$E[x] = \langle x \rangle$

In terms of this notation the first two raw and centered moments are:

$$\mu_n(t) = \langle n \rangle = \sum_n n P_n(t) \quad \text{and} \quad \langle n^2 \rangle = \sum_n n^2 P_n(t)$$
$$0 \quad \text{and} \quad \text{Var}_n(t) = \langle n^2 \rangle - \mu_n(t)^2$$

If the state space was 2 dimensional (with dimensions n and m) there would be four raw moments

$$\mu_n(t) = \langle n \rangle = \sum_{n,m} n P_{n,m}(t) \quad \text{and} \quad \langle n^2 \rangle = \sum_{n,m} n^2 P_{n,m}(t)$$
$$\mu_m(t) = \langle m \rangle = \sum_{n,m} m P_{n,m}(t) \quad \text{and} \quad \langle m^2 \rangle = \sum_{n,m} m^2 P_{n,m}(t)$$
$$\text{Var}_n(t) = \langle n^2 \rangle - \mu_n(t)^2 \quad \text{and} \quad \text{Var}_m(t) = \langle m^2 \rangle - \mu_m(t)^2$$
$$\text{Cov}_{n,m}(t) = \langle nm \rangle - \mu_n(t)\mu_m(t)$$

To derive the EMA we use the simple relationship that

$$\frac{d}{dt} \langle x \rangle = \left\langle \sum_{e \in \text{events}} \lambda_e \Delta_e x \right\rangle$$

where the sum is over all the possible events that can occur, λ_e is the rate at which event e occurs and $\Delta_e x$ is the effect of event e on the quantity x. Here x may be an element of the state space $x = \{n, m\}$ or a product of elements of the state space (e.g., $x = \{n^2, nm, m^2\}$).

Example 5.16 Logistic Population Growth
The logistic model is an ecological model of population growth capturing infraspecific competition. In the deterministic population limit it is given by:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

where r is the intrinsic growth rate and K is the carrying capacity.

1. Specify a corresponding CTDS-MC of logistic population growth.
While the above equation is perfectly legitament ODE model, there is no such biological process as "intrinsic growth" or "carrying capacity", so we need to rewrite the above equation in a way the reflects its emergence from biological processes: namely births and deaths.

Important: there is no one way of doing this propoerly. In fact there are at least three different models of births and deaths that will ultimately result in the above ODE. But for now let's just pick one of these options.

Let's assume that births occur at a constant per-capita rate bN but deaths occur at a density-dependent rate $d(1 + \gamma N)N$ where the value of α describes the rate at which the death rate increases due to additional competition. Note that if $N \rightarrow 0$ then the death rate becomes dN.

Given this model of birth and death we have:

$$\frac{dN}{dt} = bN - d(1 + \gamma N)N$$

Discussion: What is the biological interpretation of γ?

To equate the $b, d,$ and γ is this new model to the original logistic growth equation we expand both equations and equate like terms (powers of N).

$$\underbrace{rN - \frac{rN^2}{K}}_{\text{logistic}} = \underbrace{(b-d)N - d\gamma N^2}_{\text{DD birth-death}}$$

So $r = (b - d)$ and $K = \frac{b-d}{d\gamma}$

Discussion: Does this equivalency make sense? How does K change and γ change?

With the new model we can specify the dynamics of our single state variable (population size $N \in \mathbb{N}$) with a 2-event CTDSMC

Event	Rate: λ_e	$\Delta_e N$
Birth	bN	$+1$
Death	$dN(1 + \gamma N)$	-1

2. Use the EMA to derive the dynamics of the expected population size through time.

$$\frac{d}{dt}\langle N \rangle = \left\langle \sum_{e \in \text{events}} \lambda_e \Delta_e N \right\rangle$$
$$= \langle bN * 1 + dN(1 + \gamma N) * (-1) \rangle$$

Recall from expectation rules that when we take an expectation the easiest way to start simplifying is to start expanding the terms.

$$\frac{d}{dt}\langle N \rangle = \left\langle (b-d)N - d\gamma N^2 \right\rangle$$

The b, d and γ are constants so we can factor them out and split the expectation on the sum/difference to get:

$$\frac{d}{dt}\langle N \rangle = (b-d)\langle N \rangle - d\gamma \langle N^2 \rangle$$

Note that this equation has a VERY similar form as our original ODE except now we have $\langle N \rangle$ instead of N.

Also note that our ODE for the first moment $\langle N \rangle$ depends on both the first moment ($(b-d)\langle N \rangle$ term) and the second moment ($-d\gamma \langle N^2 \rangle$ term). Which means that in order to derive the dynamics for the mean we need an ODE describing $\langle N^2 \rangle$

3. Derive the ODE for $\langle N^2 \rangle$.

We use the same recipie:

$$\frac{d}{dt}\langle N^2 \rangle = \left\langle \sum_{e \in \text{events}} \lambda_e \underbrace{\Delta_e N^2}_{N_f^2 - N_0^2 \neq (N_f - N_0)^2} \right\rangle$$

where $\Delta_e N^2$ is the change in the square of the state variable NOT the change in the state variable squared.

$$\Delta_e N^2 = N_f^2 - N_0^2 \neq (N_f - N_0)^2 = (\Delta_e N)^2$$

So let's add to our table

Event	Rate: λ_e	$\Delta_e N$	$\Delta_e N^2$
Birth	bN	$N + 1 - N = 1$	$(N+1)^2 - N^2 = 2N + 2$
Death	$dN(1 + \gamma N)$	$N - 1 - N = -1$	$(N-1)^2 - N^2 = -2N + 2$

Event	Rate: λ_e	$\Delta_e N$	$\Delta_e N^2$

Substituting in these values into our ODE we have:

$$\frac{d}{dt}\langle N^2 \rangle = \langle bN(2N+2) + dN(1+\gamma N)(-2N+2) \rangle$$
$$= \langle 2N^2(b+(\gamma-1)d) + 2N(b+d) - 2\gamma dN^3 \rangle$$
$$= 2(b+(\gamma-1)d)\langle N^2 \rangle + 2(b+d)\langle N \rangle - 2\gamma d\langle N^3 \rangle$$

So this equation depends on the third raw moments. This is a problem!!! What do we do?

4. Now consider the Lotka-Volterra model of inter-specific competition between two species. Derive the corresponding EMA for the mean dynamics of this model.

$$\frac{dN_1}{dt} = b_1 N_1 - d_1 N_1 \left(1 + \gamma_{1,1} N_1 + \gamma_{1,2} N_2\right)$$
$$\frac{dN_2}{dt} = b_2 N_2 - d_2 N_2 \left(1 + \gamma_{2,1} N_1 + \gamma_{2,2} N_2\right)$$

Take Home Challenge: Derive the relationship between b_i, d_i and γ_i and the parameters of the classic Lotka-Volterra model:

$$\frac{dN_1}{dt} = r_1 N_1 \left(1 - \frac{N_1 + \alpha_2 N_2}{K_1}\right)$$
$$\frac{dN_2}{dt} = r_2 N_2 \left(1 - \frac{N_2 + \alpha_1 N_1}{K_2}\right)$$

Here we have a 4 event CTDS process with two state variables N_1 and N_2

Event	Rate: λ_e	$\Delta_e N$
Birth of 1	$b_1 N_1$	$\{+1, 0\}$
Death of 1	$d_1 N_1(1 + \gamma_{1,1} N_1 + \gamma_{1,2} N_2)$	$\{-1, 0\}$
Birth of 2	$b_1 N_1$	$\{0, +1\}$
Death of 2	$d_2 N_2(1 + \gamma_{2,1} N_1 + \gamma_{2,2} N_2)$	$\{0, -1\}$

So

$$\frac{d\langle N_1 \rangle}{dt} = \langle b_1 N_1 * 1 + d_1 N_1 \left(1 + \gamma_{1,1} N_1 + \gamma_{1,2} N_2\right) * (-1) \rangle$$
$$\frac{d\langle N_2 \rangle}{dt} = \langle b_2 N_2 * 1 + d_2 N_2 \left(1 + \gamma_{2,1} N_1 + \gamma_{2,2} N_2\right) * (-1) \rangle$$

Take Home Challenge: Simplify the ODEs above using expectation rules. What higher order moments do these first moemnts depend on?

Moment Closure

As exemplified above, when there is non-linearity in the system the dynamics of the first moments will depend on the second raw moments. So in order to solve the problem with a small number of ODEs we need to make an assumption.

Rule 1: Assume that the nth central moment is of order $\mathcal{O}(\epsilon^{n-1})$.

So in this case $Var(x) \approx \epsilon$, $Skew(x) \approx \epsilon^2$ etc.

Rule 2: If we want to understand the dynamics of the nth central moment then we need to approximate to order $\mathcal{O}(\epsilon^n)$ assuming the $(n+2)^{nd}$ moment is 0.

So if we want to understand the dynamics of the mean μ_x, we have to model the mean and the varaince and we assume the skew is 0. If we want to understand the mean and the variance we have to model the mean, variance, and skew and assume the kurtosis is 0. So you always include one more moment than you want to approximate well.

From a partical standpoint, if we want to understand the dynamics of the mean we are effectively modelling the dynamics of the populaton with a gaussian (a distribution that has a mean and variance but no skew).

Important Note: Just because the central moment is small $\langle (X - \mu)^n \rangle \approx 0$ DOES NOT imply the raw moment is small $\langle X^n \rangle \approx 0$

Example 5.17 Logistic Population Growth cont
Python: Lecture5_6.ipynb
Suppose we want to understand how stochasticity impacts mean population size.
1. Derive an EMA for the mean dynamics (including the second moment and closing at the third moment)
So we want to assume that $\text{Skew}(N) \approx 0$

From above we have:

$$\frac{d}{dt} \langle N \rangle = (b - d) \langle N \rangle - d\gamma \langle N^2 \rangle$$
$$\frac{d}{dt} \langle N^2 \rangle = 2(b + (\gamma - 1)d) \langle N^2 \rangle + 2(b + d) \langle N \rangle - 2\gamma d \langle N^3 \rangle$$

Take Home Challenge: Using the moment rules from Topic 1 show that:

$$\langle x^3 \rangle = Skew(x) + 3 \langle x^2 \rangle \langle x \rangle - 2 \langle x \rangle^3$$

or given our approximation that $Skew(x) \approx 0$

$$\langle x^3 \rangle = 3 \langle x^2 \rangle \langle x \rangle - 2 \langle x \rangle^3$$

$$\frac{d}{dt} \langle N \rangle = (b - d) \langle N \rangle - d\gamma \langle N^2 \rangle$$
$$\frac{d}{dt} \langle N^2 \rangle = 2(b + (\gamma - 1)d) \langle N^2 \rangle + 2(b + d) \langle N \rangle - 2\gamma d \left(3 \langle N^2 \rangle \langle N \rangle - 2 \langle N \rangle^3 \right)$$

We now have 2 ODEs for 2 variables $\langle N \rangle$ and $\langle N^2 \rangle$

2. Evaluate the EMA above numerically. Plot the deterministic expectation, the mean dynamics, and plus or minus 1 standard deviation of this mean. Under what conditions is the approximation valid?
We have to choose parameters. Let's assume that $b = 0.1$ and $d = 0.05$ and $\gamma = 0.005$. In this way the carrying capacity of the system is: $\frac{b-d}{\gamma d} = 200$

If we start with exactly 5 individuals in the population then $\langle N \rangle (0) = 20$ and $Var(N(0)) = \langle N^2 \rangle (0) - (\langle N \rangle (0))^2 = 0$ (there is no initial variability).
Hence:

$$\langle N^2 \rangle (0) = 20^2$$

To obtain the variance from $\langle N \rangle$ and $\langle N^2 \rangle$ we have:

$$Var(N) = \langle N^2 \rangle - \langle N \rangle^2$$

The results are roughly valid for this parameter condition. Note that if we start with far fewer indivdiauls $N(0) = 2$ then alot of the trajectories goes extinct. This dramatically increases the variance and violates our assumption.

3. Compare the results of the EMA to the deterministic solution.

Deterministic Versus EMA Dynamics