

3 Biological Applications of Discrete-Time Markov Chains

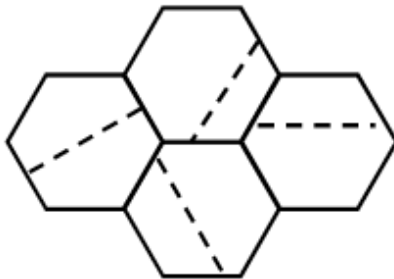
Angela Peace

Biomathematics II
MATH 5355 Spring 2017

Lecture notes follow: Allen, Linda JS. An introduction to stochastic processes with applications to biology. CRC Press, 2010.

Proliferating Epithelial Cells

Gibson et al. 2006 Nature



Proliferating epithelial cells in animal tissues have a polygonal shape with most cells being hexagonal (six-sided). An infinite MC is approximated by a finite positive recurrent MC to show the highest probability among all of the polygonal shapes is six-sided.

- ▶ k sided cell (k sides and k vertices) divides into 2 daughter cells.
- ▶ Cell division results in two new vertices and three new sides per cell.

Proliferating Epithelial Cells

Gibson et al. 2006 Nature

DTMC with following assumptions

1. minimal # of sides is 4
2. cells don't resort, vertices and sides stay intact
3. daughter cells share a common side
4. cells have \sim uniform cell cycle times $n \rightarrow n + 1$
5. cells divide within a side, not a vertex
6. mitosis randomly distributes tricellular junctions to both daughter cells

The DTMC is a 2-step process:

1. Cell division is considered
2. The change in # of sides in neighboring cells is considered

Proliferating Epithelial Cells

Gibson et al. 2006 Nature

Step 1: Consider cell division

- ▶ a single cell has s_n sides at time n
- ▶ the random variable r_{n+1} is the # of sides distributed to 1 daughter
- ▶ then the other daughter gets $s_n - r_{n+1}$
- ▶ Each daughter receives at least two sides from the parent
- ▶ So $s_n - 4$ sides are distributed among the daughter cells.
 - ▶ assume sides are distributed uniformly and randomly
 - ▶ use binomial distribution $b(s_n - 4, 1/2)$
 - ▶ each daughter also gets 2 sides from the new interface
- ▶ The probability of a transition from an i -sided cell to a j -sided cell is:

$$\text{Prob}\{2 + r_{n+1} = j | s_n = i\} = p_{ji} = \binom{i-4}{j-4} \frac{1}{2^{i-4}}$$

Proliferating Epithelial Cells

Gibson et al. 2006 Nature

$$\text{Prob}\{2 + r_{n+1} = j | s_n = i\} = p_{ji} = \binom{i-4}{j-4} \frac{1}{2^{i-4}}$$

- ▶ A 4-sided cell divides into 2 4-sided cells:
 - ▶ $p_{44} = 1$
- ▶ A 5-sided cell divides into 1 4-sided cell and 1 5-sided cell:
 - ▶ $p_{45} = 1/2$
 - ▶ $p_{55} = 1/2$
- ▶ a 6-sided cell divides into 2 5-sided cells or 1 4-sided and 1 6-sided cell:
 - ▶ $p_{46} = 1/4$
 - ▶ $p_{56} = 1/2$
 - ▶ $p_{66} = 1/4$

Proliferating Epithelial Cells

Gibson et al. 2006 Nature

Step 1: Cell division: the # of sides per cell after division gives the following Transition matrix:

$$M = \begin{pmatrix} 1 & 1/2 & 1/4 & 1/8 & 1/16 & \dots \\ 0 & 1/2 & 1/2 & 3/8 & 1/4 & \dots \\ 0 & 0 & 1/4 & 3/8 & 3/8 & \dots \\ 0 & 0 & 0 & 1/8 & 1/4 & \dots \\ 0 & 0 & 0 & 0 & 1/16 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Proliferating Epithelial Cells

Gibson et al. 2006 Nature

Step 2: consider the change in the # of sides for the neighboring cells

- ▶ each neighbor gains 1 side after division because a new junction is created
- ▶ 1 side is added after cell division: $p_{i+1,i} = 1$
- ▶ the transition matrix for the 2nd step:

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Proliferating Epithelial Cells

Gibson et al. 2006 Nature

The 2-step process becomes:

$$p(n+1) = SMp(n) = Pp(n)$$

Here $P = SM$ is the transition matrix for the DTMC.

Questions

1. How many communication classes does this DTMC have?
2. Are they transient or recurrent?
3. Determine an approximation for the stationary probability distribution
4. What number of sides is most common?

General Birth and Death Process

- ▶ X_n denotes the size of the population at time n
- ▶ State space can be finite or infinite
- ▶ $b_i > 0$ is the birth probability
- ▶ $d_i > 0$ is the death probability
- ▶ assume only 1 event occurs each time interval $n \rightarrow n + 1$

$$p_{ji} = \text{Prob}\{X_{n+1} = j | X_n = i\} = \begin{cases} b_i & \text{if } j = i + 1 \\ d_i & \text{if } j = i - 1 \\ 1 - (b_i + d_i) & \text{if } j = i \\ 0 & \text{if } j \neq i - 1, i, i + 1 \end{cases}$$

for $i = 1, 2, \dots$, $p_{00} = 1$ and $p_{j0} = 0$ for $j \neq 0$.

(For finite case: $p_{N+1,N} = b_N = 0$)

General Birth and Death Process

The transition matrix:

$$\begin{pmatrix} 1 - b_0 & d_1 & 0 & \cdots & 0 & 0 \\ b_0 & 1 - (b_1 + d_1) & d_2 & \cdots & 0 & 0 \\ 0 & b_1 & 1 - (b_2 + d_2) & \cdots & 0 & 0 \\ 0 & 0 & b_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 - (b_{N-1} + d_N - 1) & d_N \\ 0 & 0 & 0 & \cdots & b_{N-1} & 1 - d_N \end{pmatrix}$$

Questions

1. Find the stationary probability distribution π
2. Describe π when $b_0 = 0$.

General Birth and Death Process

Assume $b_0 = 0$ and the MC is finite, then

$$\lim_{n \rightarrow \infty} \text{Prob}\{X_n = 0\} = \lim_{n \rightarrow \infty} p_0(n) = 1$$

eventually population extinction occurs from any state.

Question

1. How long will it take for extinction to occur?
 - ▶ computational method using Fundamental matrix
 - ▶ analytical method from Nisbet and Gurney 1982

General Birth and Death Process

Expected time to extinction: computational method

Partition matrix P :

$$P = \begin{pmatrix} 1 & A \\ 0 & T \end{pmatrix}$$

$$= \begin{pmatrix} 1 & d_1 & 0 & \cdots & 0 & 0 \\ 0 & 1 - (b_1 + d_1) & d_2 & \cdots & 0 & 0 \\ 0 & b_1 & 1 - (b_2 + d_2) & \cdots & 0 & 0 \\ 0 & 0 & b_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 - (b_{N-1} + d_N - 1) & d_N \\ 0 & 0 & 0 & \cdots & b_{N-1} & 1 - d_N \end{pmatrix}$$

The expected time to extinction can be calculated:

$$\tau = \mathbf{1}(I - T)^{-1} = \mathbf{1}F$$

where $\mathbf{1}$ is a row vector of ones of length $1 - N$.

General Birth and Death Process

Expected time to extinction: analytical method

Nisbet and Gurney 1982 derived an analytical expression for τ_k .

Theorem 3.1

Suppose $\{X_n\}_{n=0}^N$ is a general birth and death process with $X_0 = m \geq 1$ satisfying $b_0 = d_0 = 0$, $b_i > 0$ for $i = 1, 2, \dots, N-1$, and $d_i > 0$ for $i = 1, 2, \dots, N$. The expected time until population extinction is

$$\tau_k = \begin{cases} \frac{1}{d_1} + \sum_{i=2}^N \frac{b_i \cdots b_{i-1}}{d_1 \cdots d_i} & \text{for } m = 1 \\ \tau_1 + \sum_{s=1}^{m-1} \left[\frac{d_1 \cdots d_s}{b_1 \cdots b_s} \sum_{i=s+1}^N \frac{b_1 \cdots b_{i-1}}{d_1 \cdots d_i} \right] & \text{for } m = 2, \dots, N. \end{cases}$$

General Birth and Death Process

Simple birth and death process example

Suppose the maximal population size is $N = 20$ in a birth and death process. Let $b_i = bi$ and $d_i = di$ for $i = 1, 2, \dots, 20$ where b and d are constants. For each of the following cases determine the expected time until population extinction for all initial population sizes.

1. $b = 0.02 < 0.03 = d$
2. $b = 0.025 = d$
3. $b = 0.03 > 0.02 = d$

```

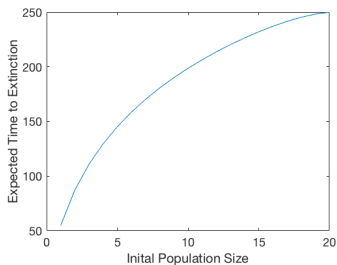
%simple birth death process
%Use to answer example 3.5 on page 123
b=0.02;
d=0.03;
N=20;

%Setup transition matrix P:
bvec=b*(0:1:N-1);
dvec=d*(1:1:N);
lowdiag=diag(bvec, -1);
updiag=diag(dvec,1);
P=diag(ones(1,N+1));
for i=2:N
    P(i,i)=P(i,i)-lowdiag(i+1,i)-updiag(i-1,i);
end
P(N+1,N+1)=1-updiag(N, N+1);
P=P+lowdiag+updiag;

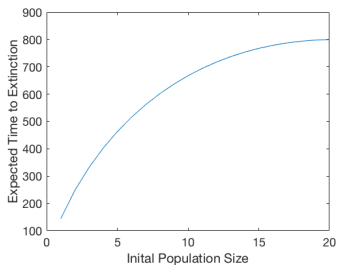
%Calculate expected time to extinction:
T=P(2:N+1, 2:N+1);
F=(eye(N)-T)^-1;
tau=ones(1,N)*F;

plot(tau)
xlabel('Initial Population Size')
ylabel('Expected Time to Extinction')
set(gca, 'FontSize', 18)

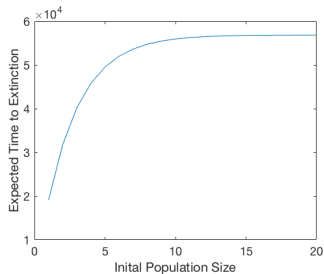
```



$$b = 0.02 < 0.03 = d$$



$$b = 0.025 = d$$



$$b = 0.03 > 0.02 = d$$

Logistic Growth Process

- ▶ Assumptions on the general birth and death probabilities b_i and d_i to make the process follow logistic growth
- ▶ Recall deterministic logistic model:

$$\frac{dy}{dt} = \tilde{r}y \left(1 - \frac{y}{K}\right), \quad y(0) = y_0 > 0$$

- ▶ where \tilde{r} is the intrinsic growth rate and K is the carrying capacity
- ▶ $\lim_{t \rightarrow \infty} y(t) = K$
- ▶ The right hand side equals the birth minus the death rate

Logistic Growth Process

- ▶ For a stochastic logistic growth process we assume:

$$b_i - d_i = ri \left(1 - \frac{i}{K} \right)$$

- ▶ for $i = 0, 1, 2, \dots, N$, where $r = \tilde{r}\Delta t$, $N > K$, and the time interval Δt is sufficiently small so that $\max_{i \in \{i=0, \dots, N\}} \{b_i + d_i\} \leq 1$. (The time interval Δt is the interval n to $n + 1$.)
- ▶ $b_i - d_i = 0$ when $i = 0$ or when $i = K$.
- ▶ We can assume that b_i and d_i are either linear or quadratic function of i .
- ▶ These give us 2 cases for the probabilities needed to make the process birth and death process logistic

Logistic Growth Process

Stochastic Logistic Growth Model

Case a:

$$b_i = r \left(i - \frac{i^2}{2K} \right) \text{ and } d_i = r \frac{i^2}{2K}$$

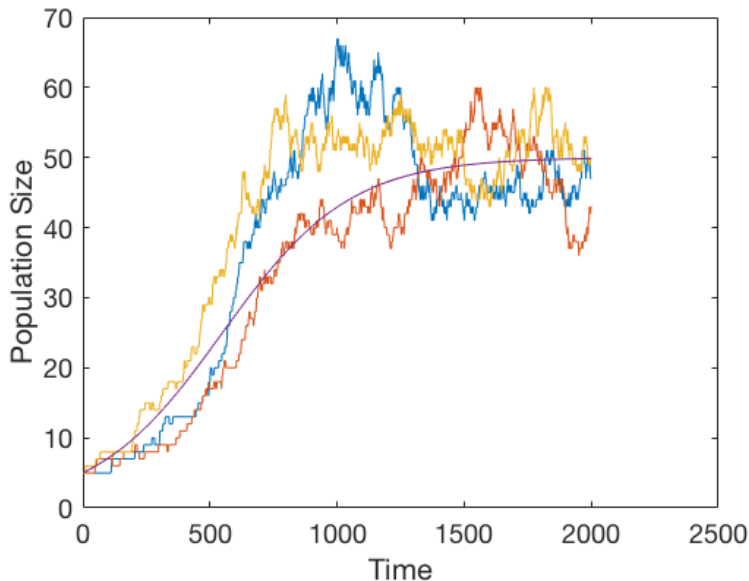
for $i = 0, 1, 2, \dots, 2K$

Case b:

$$b_i = \begin{cases} ri, & i = 0, 1, 2, \dots, N \\ 0, & i \geq N \end{cases} \text{ and } d_i = r \frac{i^2}{K}$$

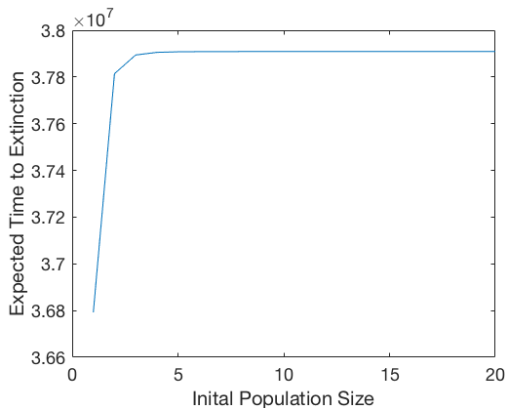
for $i = 0, 1, 2, \dots, N$

Logistic Growth Process



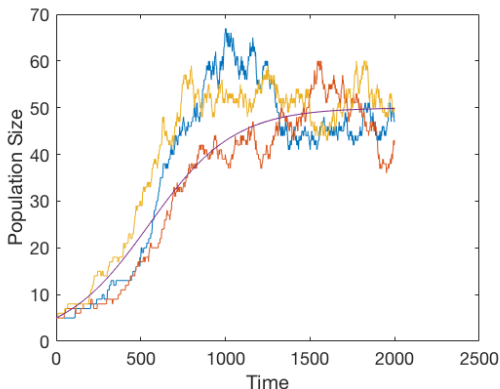
Logistic Growth Process

- ▶ Unlike the deterministic Logistic growth model, In the limit the stochastic logistic growth process does not approach K .
- ▶ It is still a birth and death process, and extinction is an absorbing state
- ▶ For large population size, the time to extinction is very large



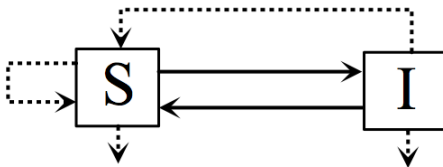
Logistic Growth Process

- ▶ For large N the stochastic model follows the deterministic model closely
- ▶ Before extinction (which may take a loooong time) the probability distribution is approximately stationary for a long period of time.
- ▶ This is called the *quasistationary distribution*



SIS Epidemic Model

- ▶ S Susceptible individuals
- ▶ I Infected individuals
- ▶ SIS model:
 - ▶ S individuals may become infected I .
 - ▶ Infected individuals eventually recovery but do not gain any immunity
 - ▶ I individuals become S
 - ▶ Assume the total # of births equals the total # of deaths so that the population remains constant: $S + I = N$.
 - ▶ All newborns are born S .



SIS Epidemic Model

Assumptions

- ▶ time interval n to $n + 1$ is sufficiently small that at most 1 event occurs
 - ▶ S individual becomes I
 - ▶ I individual recovers and becomes S
 - ▶ An individual gives birth to a new S and a corresponding death of either S or I occurs
- ▶ S individual becomes I with probability $\beta \frac{I}{N}$
 - ▶ β is the # of contacts made by one I individual that results in a new infection during the time interval n to $n + 1$
 - ▶ $\beta \frac{S}{N}$ of these contacts can results in a new infection
 - ▶ $\beta \frac{SI}{N}$ is the total # of new infections by the entire class of I individuals
- ▶ Individuals are born or die with probability b
- ▶ I individuals recover with probability γ

SIS Epidemic Model

Deterministic Model

Let S_n and I_n be the # of S and I individuals at time n .

The dynamics during the time interval Δt are modeled with a system of **difference equations**.

$$S_{n+1} = S_n - \beta \frac{S_n I_n}{N} + I_n(b + \gamma)$$

$$I_{n+1} = \beta \frac{S_n I_n}{N} + I_n(1 - b - \gamma)$$

where $n = 0, 1, 2, \dots$, and $S_0, I_0 > 0$ with $S_0 + I_0 = N$ and $0 < \beta \leq 1$ and $0 < b + \gamma \leq 1$.

SIS Epidemic Model

Deterministic Model

Since $S_n + I_n = N$ the systems can be reduced down to a single equation.
Let $S_n = N - I_n$

$$\begin{aligned}I_{n+1} &= \beta \frac{S_n I_n}{N} + I_n(1 - b - \gamma) \\&= I_n \left(\beta \frac{N - I_n}{N} + 1 - b - \gamma \right) \\&= I_n \left(1 + \beta - b - \gamma - \beta \frac{I_n}{N} \right)\end{aligned}$$

where $0 \leq I_n \leq N$.

SIS Epidemic Model

Deterministic Model

- ▶ There are 2 equilibria solutions:

$$I_{n0} = 0 \text{ and } I_n^* = N \left(1 - \frac{b + \gamma}{\beta} \right)$$

- ▶ The basic reproductive number is

$$R_0 = \frac{\beta}{b + \gamma}$$

- ▶ If $R_0 \leq 1$ then $\lim_{n \rightarrow \infty} I_n = 0$
- ▶ If $R_0 > 1$ then $\lim_{n \rightarrow \infty} I_n = I_n^*$

SIS Epidemic Model

Stochastic Model

- ▶ I_n is a discrete random variable for the # of infected individuals at time n .
- ▶ Set $\{0, 1, 2, \dots, N\}$ is the state space
- ▶ Assume Δt (time interval n to $n + 1$) is sufficiently small that there is at most 1 change in I_n .
 - ▶ If $I_n = i$ then I_{n+1} is either $i, i + 1$, or $i - 1$.
- ▶ The one-step transition probabilities are

$$p_{i+1,i} = \text{Prob}\{I_{n+1} = i + 1 | I_n = i\} = \beta i(N - i)/N = \Pi_i$$

$$p_{i-1,i} = \text{Prob}\{I_{n+1} = i - 1 | I_n = i\} = (b + \gamma)i$$

$$\begin{aligned} p_{ii} &= \text{Prob}\{I_{n+1} = i | I_n = i\} = 1 - \beta i(N - i)/N - (b + \gamma)i \\ &= 1 - \Pi_i - (b + \gamma)i \end{aligned}$$

for $i = 0, 1, \dots, N$ and $p_{ji} = 0$ if $j \neq i - 1, i + 1, i$ and $p_{00} = 1$.

SIS Epidemic Model

Stochastic Model

The transition matrix:

$$P = \begin{pmatrix} 1 & (b + \gamma) & 0 & \cdots & 0 \\ 0 & 1 - \Pi_1 - (b + \gamma) & 2(b + \gamma) & \cdots & 0 \\ 0 & \Pi_1 & 1 - \Pi_2 - 2(b + \gamma) & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & N(b + \gamma) \\ 0 & 0 & 0 & \cdots & 1 - N(b + \gamma) \end{pmatrix}$$

where $\max_i \{\Pi_i + i(b + \gamma)\} \leq 1$.

- ▶ 2 communication classes: $\{0\}$ and $\{1, 2, \dots, N\}$
- ▶ $\{0\}$ is absorbing and $\{1, 2, \dots, N\}$ is transient
- ▶ $\lim_{n \rightarrow \infty} P^n p(0) = (1, 0, \dots, 0)^T$. Eventually there are no infected individuals.
- ▶ It may take a long time until the epidemic ends. In this case we can consider the disease endemic (quasistationary distribution)