Tutorial: Stochastic Modeling in Biology
Applications of Discrete- Time Markov Chains

Linda J. S. Allen
Texas Tech University
Lubbock, Texas U.S.A.

NIMBioS
Knoxville, Tennessee
March 16-18, 2011

OUTLINE

Part I: Discrete-Time Markov Chains - DTMC

- **Summary of Notation**
- **Applications:**

(1) Proliferating Cells [MATLAB PROGRAM]

(2) Restricted Random Walk [MATLAB PROGRAM]

(3) Simple Birth and Death Process

(4) SIS Epidemic Process [MATLAB PROGRAM]

Part II: Discrete-Time Branching Processes

- **Summary of Notation**
- **Applications:**

(1) Cell Cycle

(2) Age-Structured Population

References for these Notes

1. Allen, EJ 2007 Modeling with Itô Stochastic Differential Equations. Springer, Dordrecht, The Netherlands

2. Allen, LJS 2010 An Introduction to Stochastic Processes with Applications to Biology. 2nd Ed. CRC Press/Chapman & Hall, Boca Raton, Fl

3. Allen, LJS 2008 Chapter 3: An Introduction to Stochastic Epidemic Models. Mathematical Epidemiology, Lecture Notes in Mathematics. Vol. 1945. pp. 81-130, F Brauer, P van den Driessche, and J Wu (Eds) Springer

4. Athreya, KB and P Ney 197. Branching Processes. Springer, Berlin

5. Caswell, H 2001 Matrix Population Models. 2nd Ed. Sinauer Assoc. Inc., Sunderland, MA

6. Karlin S and H Taylor 1975 A First Course in Stochastic Processes. 2nd Ed. Acad. Press, NY

The Choice of Discrete or Continuous Random Variables with a Discrete or Continuous Index Set Defines the Type of Stochastic Model.

Discrete Time Markov Chain (DTMC): $n \in \{0, 1, 2, \ldots\}$, X_n is a discrete random variable. The term chain implies that the random variable is discrete.

$$X_n \in \{0, 1, \ldots, N\} \text{ or } \{0, 1, \ldots\}$$

Continuous Time Markov Chain (CTMC): $t \in [0, \infty)$, $X(t)$ is a discrete random variable.

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

Diffusion Process, Stochastic Differential Equation (SDE): $t \in [0, \infty)$, $X(t)$ is a continuous random variable.

$$X(t) \in [0, N] \text{ or } [0, \infty)$$

Part I:
Discrete-Time Markov Chains
Notation and Terminology

Discrete random variable: X_n, $n \in \{0, 1, 2, \ldots\}$ **with state space**

$$\{0, 1, 2, \ldots, N\} \text{ or } \{0, 1, 2, \ldots\}.$$

Markov property:

$$\text{Prob}\{X_n = i_n | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\}$$

$$= \text{Prob}\{X_n = i_n | X_{n-1} = i_{n-1}\}.$$

Probability mass function of X_n: $\{p_i(n)\}_{i=0}^{\infty}$, **where**

$$p_i(n) = \text{Prob}\{X_n = i\}.$$

Transition Matrix

One-step transition probability:

$$p_{ji}(n) = \mathrm{Prob}\{X_{n+1} = j | X_n = i\}.$$

Stationary or Homogeneous Transition Probabilities:

$$p_{ji}(n) \equiv p_{ji}$$

Transition matrix:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots \\ p_{31} & p_{32} & p_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

Column sum is one, $\sum_j p_{ji} = 1$, a stochastic matrix. Assume the transition probabilities are homogeneous.

$n-$Step Transition Matrix

$n-$ **Step transition probabilities:**

$$p_{ji}^{(n)} = \text{Prob}\{X_n = j | X_0 = i\}.$$

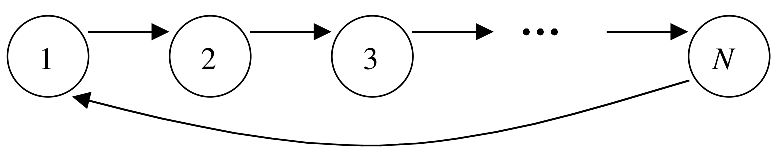
n**-step transition matrix**

$$P^{(n)} = \left(p_{ji}^{(n)}\right) = P^n$$

Irreducible if only one communication class; otherwise reducible.

Irreducible and periodic:

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



First Passage Time and Recurrent Chains

First return to state i is at the nth time step:

$$f_{ii}^{(n)} = \text{Prob}\{X_n = i, X_m \neq i, m = 1, 2, \ldots, n-1 | X_0 = i\}.$$

Transient state i: $\sum_{n=1}^{\infty} f_{ii}^{(n)} < 1$ **Recurrent state i:** $\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$.

State i recurrent (transient) iff

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty \ (< \infty).$$

Mean recurrence time:

$$\mu_{ii} = \sum_{n=1}^{\infty} n f_{ii}^{(n)}.$$

Positive recurrent: $\mu_{ii} < \infty$. **Null recurrent:** $\mu_{ii} = \infty$.

Summary of Classification Schemes

Markov chains or classes can be classified as

Periodic or Aperiodic

Then further classified as

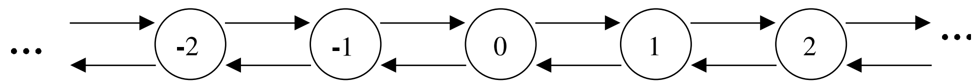
Transient or Recurrent

Then recurrent MC can be classified as

Null recurrent or Positive recurrent.

Infinite Random Walk is Transient if $p \neq q$ and Null Recurrent if $p = \frac{1}{2} = q$.

The chain is irreducible and periodic of period 2.



$p =$probability moving right, $q =$probability moving left, $p + q = 1$.

$$p_{00}^{(2n+1)} = 0 \quad \text{for} \quad n = 0, 1, 2, \ldots.$$

$$p_{00}^{(2n)} = \frac{(2n)!}{n!n!} p^n q^n \sim \frac{(4pq)^n}{\sqrt{\pi n}}.$$

Transient: $p \neq q$, $4pq < 1$, $\sum_{n=0}^{\infty} p_{00}^{(2n)} < \infty$

Null Recurrent: $p = \frac{1}{2} = q$, $4pq = 1$, $\sum_{n=0}^{\infty} p_{00}^{(2n)} = \infty$, $p_{00}^{(2n)} \to 0$, $\mu_{00} = \infty$

Finite DTMC

In finite DTMC, there are NO null recurrent states and not all states are transient. An irreducible MC is

Periodic or Aperiodic

Positive Recurrent

A stationary probability distribution, π is an eigenvector of P with eigenvalue one:

$$P\pi = \pi$$

$$\lim_{n\to\infty} P^n p(0) = \pi.$$

Example: Transition Matrix: $P = \begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix}$. Stationary probability distribution: $P\pi = \pi$, where

$$\pi = (2/5, 3/5)^T.$$

Mean recurrence times: $\mu_{11} = 5/2$ and $\mu_{22} = 5/3$.

Biological Application of DTMCs

(1) Proliferating Epithelial Cells [MATLAB PROGRAM]

(2) Restricted Random Walk [MATLAB PROGRAM]

(3) Simple Birth and Death Process

(4) SIS Epidemic Process [MATLAB PROGRAM]

(1) Proliferating Epithelial Cells

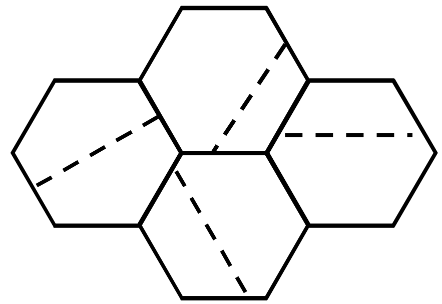


Figure 1: Cell division results in two new vertices and three new sides per cell.

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.

Gibson et al. 2006 *Nature*

First Step Transition Matrix

A single cell has s_n sides. Each daughter receives at least two sides from the parent, leaving $s_n - 4$ for daughter cells. Let r_{n+1} be the number of sides distributed to each daughter cell, uniformly and randomly according to $b(s_n - 4, 1/2)$.

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

First step transition matrix:

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

Second Step Transition Matrix

Each neighbor gains one side after cell division because a new junction is created. Second step transition matrix:

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \quad SM = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 1/2 & 1/4 & 1/8 & 1/16 & \cdots \\ 0 & 1/2 & 1/2 & 3/8 & 1/4 & \cdots \\ 0 & 0 & 1/4 & 3/8 & 3/8 & \cdots \\ 0 & 0 & 0 & 1/8 & 1/4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

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

$$\boxed{P = SM}$$

Transient class: $\{4\}$. Recurrent class: $\{5, 6, \ldots\}$.

Approximate Truncated Transition Matrix is a Stochastic Matrix

Matrix P is truncated to an $m \times m$ matrix; the truncated matrix is not a stochastic matrix. Approximating the entries to four decimal places yields a stochastic matrix of dimension $m \times m$. The finite Markov chain applied to the set $\{5, 6, \ldots, m\}$ is positive recurrent. The approximate stationary probability distribution for cells up to 10 sides, $(5, 6, \ldots, 10)$ is

$$(0.2888, 0.4640, 0.2085, 0.0359, 0.0027, 0.0001)^T.$$

The largest proportion of cells are hexagonal in shape. In addition, the expected value of π is $E(\pi) = 6$.

See the MatLab program.

(2) Random Walk on $\{0, 1, \ldots, N\}$ or Gambler's Ruin

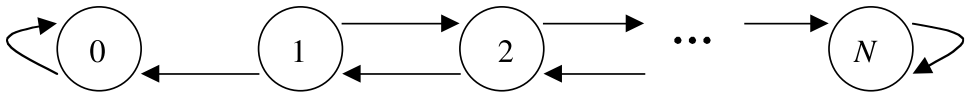


Figure 2: Movement right has probability p; movement left has probability q.

Absorbing boundaries, $\{0\}$, $\{N\}$ are recurrent states and $\{1, 2, \ldots, N-1\}$ are transient.
Transition matrix:

$$P = \begin{pmatrix} 1 & q & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & q & \cdots & 0 & 0 & 0 \\ 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & p & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & p & 1 \end{pmatrix} = \begin{pmatrix} 1 & A & 0 \\ 0 & T & 0 \\ 0 & B & 1 \end{pmatrix},$$

where T is an $N-1 \times N-1$ matrix of transient states.

Probability of Absorption

$a_k =$ probability of reaching state 0 from state k
$b_k =$ probability of reaching state N from k:

$$a_k = pa_{k+1} + qa_{k-1}$$

$$\boxed{-qa_{k-1} + a_k - pa_{k+1} = 0}$$

$a_0 = 1$, $a_N = 0$. Expressed in matrix form:

$$\boxed{aD = \mathbf{c}}$$

where $a = (a_0, a_1, \dots, a_N)$, $\mathbf{c} = (1, 0, \dots, 0)$, and

$$D \quad = \quad \begin{pmatrix} 1 & -q & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -q & 0 & \cdots & 0 & 0 \\ 0 & -p & 1 & -q & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -p & 1 \end{pmatrix} = \begin{pmatrix} 1 & -A & 0 \\ 0 & I - T & 0 \\ 0 & -B & 1 \end{pmatrix}.$$

$$\boxed{a = \mathbf{c}D^{-1}}$$

An Explicit Solution for the Probability of Absorption (Extinction)

An explicit solution is

$$a_k = \begin{cases} \dfrac{(q/p)^N - (q/p)^k}{(q/p)^N - 1}, & p \neq q \\ \dfrac{N-k}{N}, & p = q \end{cases}$$

Let $N \to \infty$.

$$a_k = \begin{cases} 1, & p \leq q \\ \left(\dfrac{q}{p}\right)^k, & p > q \end{cases}$$

Table 1: Let $k = 50$ and $N = 100$

Prob.	a_{50}	b_{50}	τ_{50}
$q = 0.50$	0.5	0.5	2500
$q = 0.51$	0.880825	0.119175	1904
$q = 0.55$	0.999956	0.000044	500
$q = 0.60$	1.00000	0.00000	250

Three Sample Paths

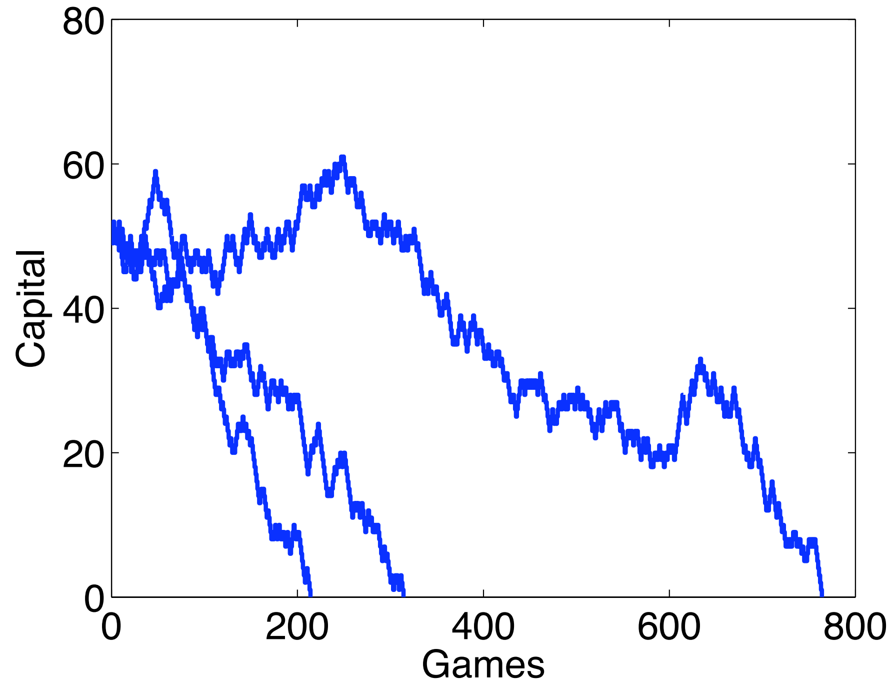


Figure 3: $N = 100$, $q = 0.55$, $k = 50$

Expected Time Until Absorption (Extinction)

$\tau_k = E(T_k)$ **expected time until absorption beginning from state k.**

$$\tau_k = p(1 + \tau_{k+1}) + q(1 + \tau_{k-1})$$

$$\boxed{-p\tau_{k-1} + \tau_k - q\tau_{k=1} = 1}$$

In matrix form:

$$\tau(I - T) = (1, 1, \ldots, 1)$$

$$\boxed{\tau = (1, 1, \ldots, 1)(I - T)^{-1}}$$

Table 2: Let $k = 50$ and $N = 100$

Prob.	a_{50}	b_{50}	τ_{50}
$q = 0.50$	0.5	0.5	2500
$q = 0.51$	0.880825	0.119175	1904
$q = 0.55$	0.999956	0.000044	500
$q = 0.60$	1.00000	0.00000	250

(3) Simple Birth and Death Process

The probability of a birth or a death is not constant but depends on the size of the population. Let X_n, $n = 0, 1, 2, \ldots$ denote the size of the population. The birth and death probabilities are b_i and d_i for a population of size i, $b_0 = 0 = d_0$, $b_i, d_i \geq 0$, for $i = 1, 2, \ldots$. During the time interval Δt, $n \to n + 1$, at most one event occurs, either a birth or a death. Assume

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

for $i = 1, 2, \ldots$, $p_{00} = 1$. The probabilities $b_i \equiv b_i(\Delta t)$ and $d_i = d_i(\Delta t)$ depend on the time interval Δt.

The Transition Matrix for a Birth and Death Process

The transition matrix P has the following form:

$$\begin{pmatrix} 1 & d_1 & 0 & \cdots \\ 0 & 1-(b_1+d_1) & d_2 & \cdots \\ 0 & b_1 & 1-(b_2+d_2) & \cdots \\ 0 & 0 & b_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} 1 & A \\ 0 & T \end{pmatrix}.$$

During the time interval Δt, either the population size increases by one, decreases by one, or stays the same size. This is a reasonable assumption if the time interval is sufficiently small.

$$\sup_i \{b_i + d_i\} \leq 1$$

Probability of Extinction

If $b_i = 0$ for $i \geq N$ and $d_i = 0$ for $i > N$ and $b_i, d_i > 0$ elsewhere, then the population size is finite. There are two communication classes, $\{0\}$ and $\{1, 2, \ldots, N\}$. The first one is positive recurrent and the second one is transient. There exists a unique stationary probability distribution π, $P\pi = \pi$, where $\pi_0 = 1$ and $\pi_i = 0$ for $i = 1, 2, \ldots, N$. Eventually, population extinction occurs from any initial state:

$$\lim_{n\to\infty} P^n p(0) = \pi = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

But if $b_i, d_i > 0$ for $i = 1, 2, \ldots,$ then the probability of extinction may be less than one.

When the Probability of Extinction equals One, the Expected Time to Extinction Can be Calculated.

Let τ_k = the expected time until extinction for a population with initial size k.

$$\tau_k = b_k(1+\tau_{k+1}) + d_k(1+\tau_{k-1}) + (1-(b_k+d_k))(1+\tau_k)$$

$$\boxed{-d_k\tau_{k-1} + (b_k+d_k)\tau_k - b_k\tau_{k+1} = 1}$$

and $-d_N\tau_{N-1} + d_N\tau_N = 1$. This can be expressed in matrix form:

$$\boxed{\tau D = \mathrm{c}}$$

where $\tau = (\tau_0, \tau_1, \ldots, \tau_N)$, $\mathrm{c} = (0, 1, \ldots, 1)$, and D is

$$\begin{pmatrix} 1 & -d_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & b_1+d_1 & -d_2 & 0 & \cdots & 0 & 0 \\ 0 & -b_1 & b_2+d_2 & -d_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & d_N \end{pmatrix} = \begin{pmatrix} 1 & -\mathrm{A} \\ 0 & I-T \end{pmatrix}.$$

$$\boxed{\tau = \mathrm{c}D^{-1}}$$

An Example of a Simple Birth and Death Process with $N = 20$.

Suppose the maximal population size is $N = 20$, where the birth and death probabilities are linear: $b_i \equiv 0.03i\Delta t$, for $i = 1, 2, \ldots, 19$, $d_i \equiv 0.02i\Delta t$, for $i = 1, 2, \ldots, 20$, $\Delta t = 1$ a simple birth and death process.

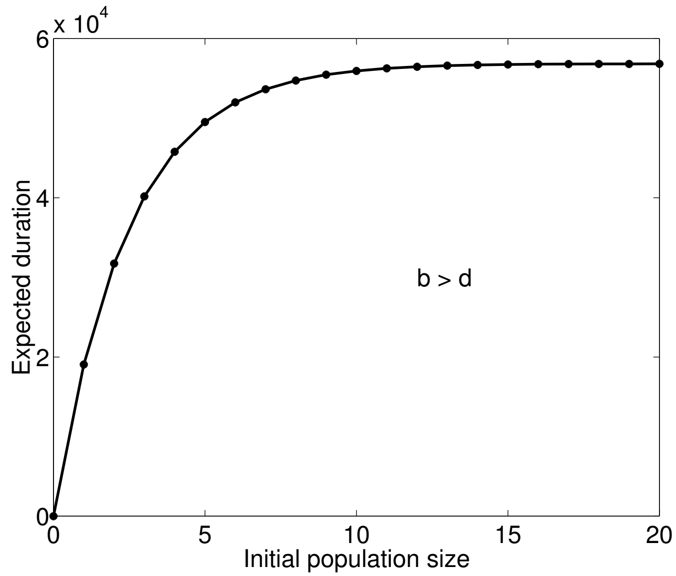
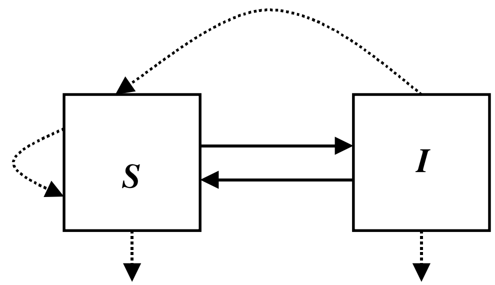


Figure 4: Expected time until population extinction τ when the maximum population size is $N = 20$, $b_i = 0.03i\Delta t$, and $d_i = 0.02i\Delta t$.

If $\Delta t = 1$ day, $6 \times 10^4 \approx 160$ years.

(4) SIS Epidemic Model



Deterministic Model:

$$\frac{dS}{dt} = -\frac{\beta}{N} IS + (b + \gamma)I$$

$$\frac{dI}{dt} = \frac{\beta}{N} IS - (b + \gamma)I = \frac{\beta}{N} I(N - I) - (b + \gamma)I$$

$S(t) = N - I(t)$, **where** $N =$ **constant total population size.**

Basic Reproduction Number: $$\mathcal{R}_0 = \frac{\beta}{b + \gamma}$$

If $\mathcal{R}_0 \leq 1$, **then** $\lim_{t\to\infty} I(t) = 0$.

If $\mathcal{R}_0 > 1$, **then** $\lim_{t\to\infty} I(t) = N\left(1 - \frac{1}{\mathcal{R}_0}\right) > 0$.

SIS Epidemic Process

Since $S(t) = N - I(t)$ and N is constant, only $I(t)$ is modeled.

Let $I(t)$ be the random variable for the number infectious at time $t = 0, \Delta t, 2\Delta t, \ldots$.

$$p_i(t) = \text{Prob}\{I(t) = i\}, \; i = 0, 1, 2, \ldots, N.$$

Transition Probabilities:

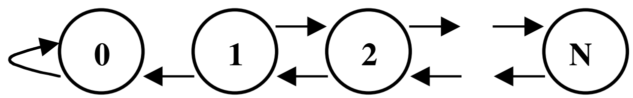
$$p_{ji}(\Delta t) = \text{Prob}\{I(t + \Delta t) = j | I(t) = i\}.$$

Transition Probabilities

$$p_{ji}(\Delta t) = \begin{cases} \dfrac{\beta i(N-i)}{N}\Delta t, & j = i+1 \\ (b+\gamma)i\Delta t, & j = i-1 \\ 1 - \left[\dfrac{\beta i(N-i)}{N} + (b+\gamma)i\right]\Delta t, & j = i \\ 0, & j \neq i+1, i, i-1. \end{cases}$$

Similar to a birth and death process:

$$p_{ji}(\Delta t) = \begin{cases} b(i)\Delta t, & j = i+1 \\ d(i)\Delta t, & j = i-1 \\ 1 - [b(i)+d(i)]\Delta t, & j = i \\ 0, & j \neq i+1, i, i-1. \end{cases}$$



Recurrent class: $\{0\}$ Transient class: $\{1, \ldots, N\}$.

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

Three Sample Paths

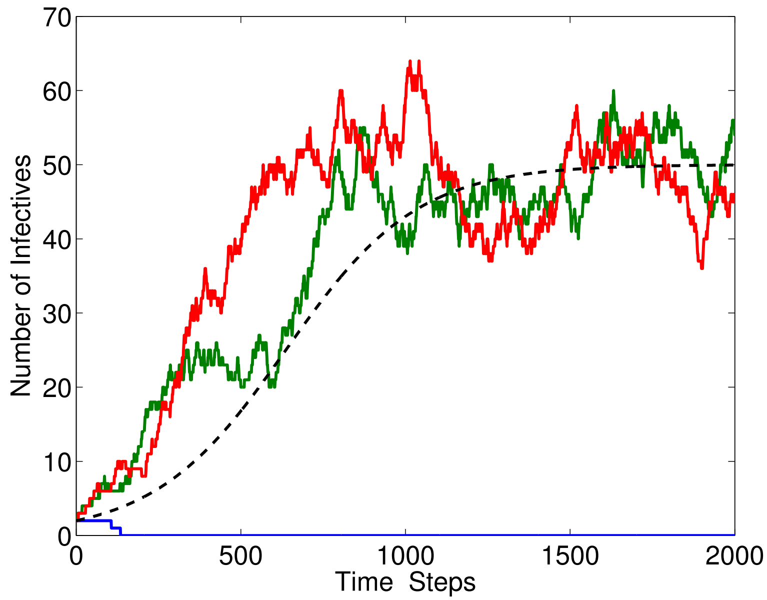


Figure 5: Three sample paths of the DTMC SIS epidemic model are graphed with the deterministic solution (*dashed curve*); $\Delta t = 0.01$, $N = 100$, $\beta = 1$, $b = 0.25$, $\gamma = 0.25$, and $I(0) = 2$.

Probability Distribution

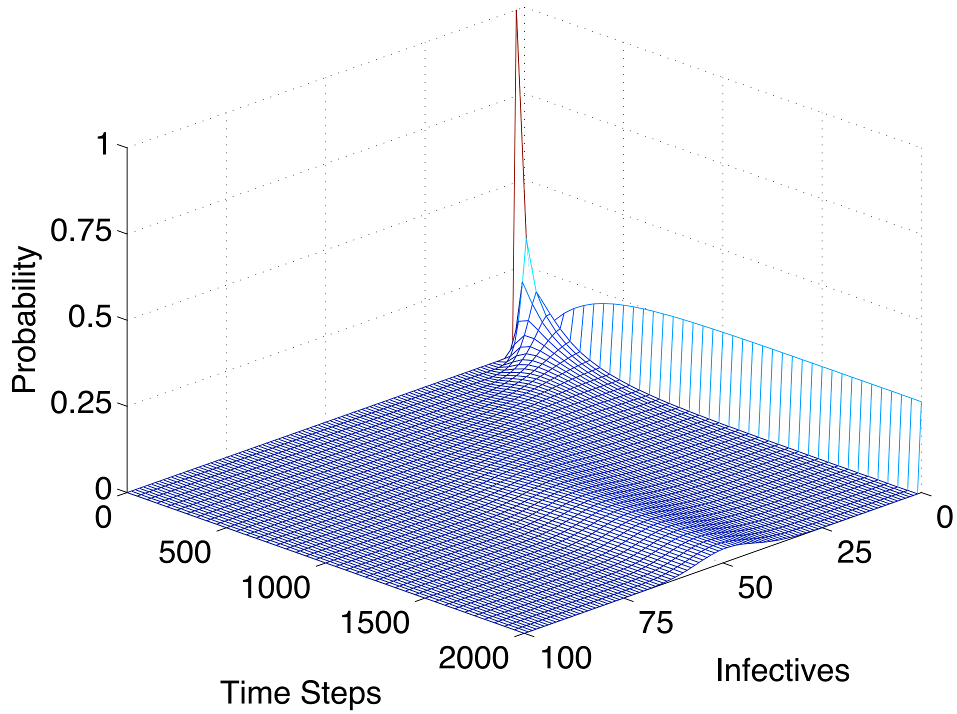


Figure 6: Probability distribution of the DTMC SIS epidemic model. $\Delta t = 0.01$, $N = 100$, $\beta = 1$, $b = 0.25$, $\gamma = 0.25$, $I(0) = 2$, $\mathcal{R}_0 = 2$. Quasistationary distribution-conditioned on nonextinction, $(1/\mathcal{R}_0)^2 = 0.25$

See MATLAB **program.**

Part II:

Discrete-Time Branching Processes (BP)

Single-Type and Multi-Type BP

Single-Type BP: The term "single-type" refers to the fact that all individuals are of one type such as the same gender, same cell type, or same genotype or phenotype.

(1) Cell Cycles, Active and Quiescent

Multi-type BP: Instead of only one type, there are several types of individuals. A population may be divided according to age, size, or developmental stage, representing different types. In genetics, genes may be classified as wild or mutant types.

(2) Age-Structured Population

Single-Type Galton-Watson BP

In 1873, Galton sent a problem to the Educational Times regarding the survival of family names. When he did not receive a satisfactory answer, he consulted Watson, who rephrased the problem in terms of generating functions. These types of problems are called Galton-Watson BP.

Assumptions:

(i) Each individual in generation n gives birth to Y offspring of the same type, where Y is a discrete random variable. Offspring probabilities:

$$\boxed{p_j = \text{Prob}\{Y = j\}, \quad j = 0, 1, 2, \ldots.}$$

(ii) Each individual in the population gives birth independently of all other individuals.

(iii) The same offspring distribution applies to all generations.

A Stochastic Realization or Sample Path of a BP

Let $X_0 = 1$.

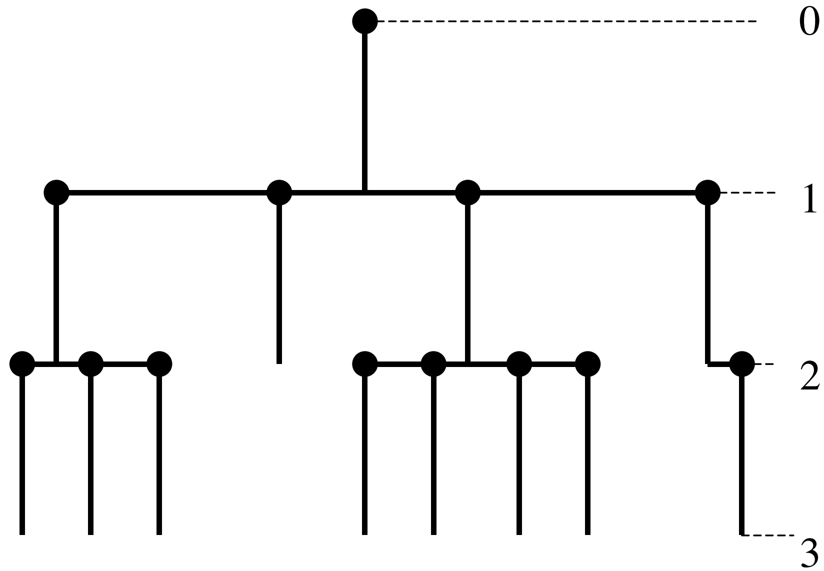


Figure 7: A sample path or stochastic realization of a branching process $\{X_n\}_{n=0}^{\infty}$. In the first generation, four individuals are born, $X_1 = 4$. The four individuals give birth to three, zero, four, and one individuals, respectively, making a total of eight individuals in generation 2, $X_2 = 8$.

Definition of Probability Generating Function

Probability generating function (pgf) of X

$$f(t) = E(t^X) = \sum_{j=0}^{\infty} p_j t^j, \quad \text{some } t \in \mathbb{R}.$$

As the name implies, the pgf generates the probabilities associated with the distribution

$$f(0) = p_0, \quad f'(0) = p_1, \quad f''(0) = 2! p_2.$$

Mean:

$$f'(1) = \sum_{j=1}^{\infty} j p_j = E(X) = m.$$

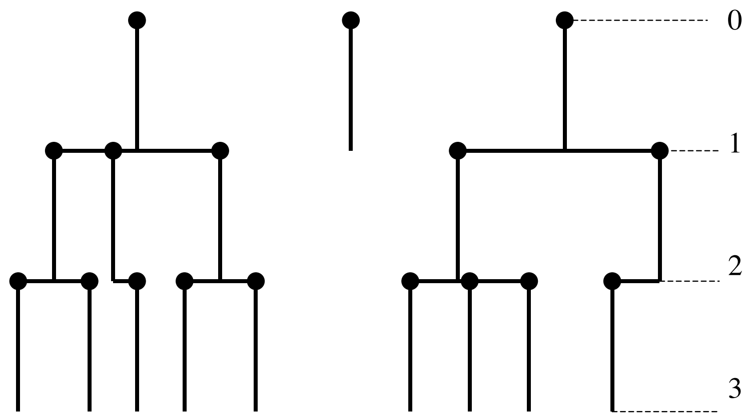
PGF h_n of the Galton-Watson BP X_n

Let the pgf of X_n be h_n and the offspring pgf $f(t) = \sum_{k=0}^{\infty} p_k t^k$.

$X_0 = 1$: $h_0(t) = t, \quad h_1(t) = \sum_{k=0}^{\infty} p_k t^k = f(t)$.

$$h_n(t) = f(f(\cdots(f(t))\cdots)) = f^n(t).$$

$X_0 = N$: $$h_0(t) = t^N, \quad h_n(t) = [f^n(t)]^N$$



Extinction Theorem in Branching Processes.

Theorem 1. Assume $X_0 = N$ and the offspring distribution $\{p_k\}_{k=0}^{\infty}$ satisfies $p_0 > 0$ and $0 < p_0 + p_1 < 1$.

(i) If $m \leq 1$, then

$$\lim_{n\to\infty} \text{Prob}\{X_n = 0\} = \lim_{n\to\infty} h_n(0) = \lim_{n\to\infty} [f^n(0)]^N = 1$$

(ii) If $m > 1$, then there exists unique $q < 1$ such that $f(q) = q$

$$\lim_{n\to\infty} \text{Prob}\{X_n = 0\} = \lim_{n\to\infty} h_n(0) = \lim_{n\to\infty} [f^n(0)]^N = q$$

Extinction in Branching Processes.

The Galton-Watson BP is referred to as supercritical if $m > 1$, critical if $m = 1$, and subcritical if $m < 1$.

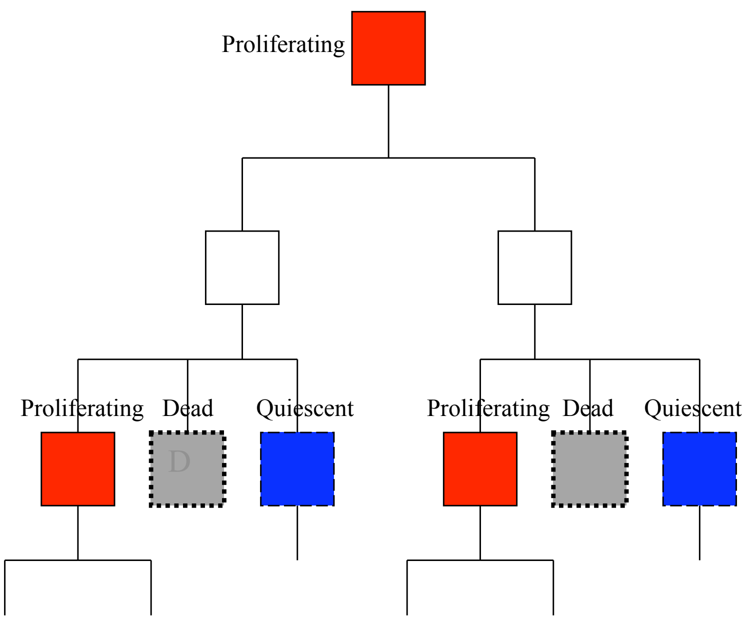
If the process is subcritical or critical, then the probability of extinction is certain.

But if the process is supercritical, then there is a positive probability, $1 - q^N$, that the population will survive. As the initial population size increases, the probability of survival also increases.

(1) Cell Cycle: Active and Quiescent

Each cell after completing its life cycle, doubles in size, then divides into two progeny cells of equal sizes. After cell division, some cells die, some remain inactive or quiesce and some keep dividing or proliferating.

(1) Cell proliferation, probability p_2
(2) Cell death, probability p_0
(3) Cell quiescence, probability p_1, $p_0 + p_1 + p_2 = 1$.



Kimmel and Axelrod, 2002

The Cell Cycle is a Galton-Watson Process

Let X_n be the number of proliferating cells at time n. The offspring pgf is

$$
\begin{aligned}
f(t) &= (p_0+p_1)^2 + 2p_2(p_0+p_1)t + p_2^2t^2 \\
&= (p_2t+p_0+p_1)^2
\end{aligned}
$$

The mean of the proliferating cells is

$$\boxed{m = f'(1) = 2p_2.}$$

If $p_2 \le 1/2$, then with probability one the cell line will go extinct.
If $p_2 > 1/2$ and $X_0 = 1$, the probability of extinction is determined by the fixed point $q = f(q)$:

$$\boxed{\lim_{n\to\infty} \text{Prob}\{X_n = 0\} = \frac{(1-p_2)^2}{p_2^2}.}$$

Multi-type Galton Watson BP

A multi-type GWbp $\{\vec{X}(n)\}_{n=0}^{\infty}$ is a collection of vector random variables $\vec{X}(n)$, where each vector consists of k different types,

$$\boxed{\vec{X}(n) = (X_1(n), X_2(n), \ldots, X_k(n)).}$$

The offspring pgf of X_i is $f_i(t_1, t_2, \ldots, t_k)$. The mean number of j-type offspring by an i-type parent is

$$m_{ji} = \left.\frac{\partial f_i(t_1, t_2, \ldots, t_k)}{\partial t_j}\right|_{t_1=1, t_2=1, \ldots, t_k=1}.$$

Expectation matrix:

$$\mathcal{M} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1k} \\ m_{21} & m_{22} & \cdots & m_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ m_{k1} & m_{k2} & \cdots & m_{kk} \end{pmatrix}.$$

Extinction Theorem for Multi-type Galton-Watson BP

Theorem 2. Let the initial sizes for each type be $X_i(0) = N_i$, $i = 1, 2, \ldots, k$. Suppose the generating functions f_i for each of the k types are nonlinear functions of t_j with some $f_i(0, 0, \ldots, 0) > 0$, the expectation matrix $\mathcal{M}$ is regular, and λ is the dominant eigenvalue of matrix $\mathcal{M}$.

(i) If $\lambda \leq 1$, then the probability of ultimate extinction is one,

$$\lim_{n\to\infty} \text{Prob}\{\vec{X}(n) = \vec{0}\} = 1.$$

(ii) If $\lambda > 1$, then the probability of ultimate extinction is less than one,

$$\lim_{n\to\infty} \text{Prob}\{\vec{X}(n) = \vec{0}\} = q_1^{N_1} q_2^{N_2} \cdots q_k^{N_k},$$

where $(q_1, q_2, \ldots, q_k)$ is the unique fixed point of the k generating functions $f_i(q_1, \ldots, q_k) = q_i$ and $0 < q_i < 1$, $i = 1, 2 \ldots, k$.

(2) Age-Structured Population

The number of females in k different age classes are modeled.
Type 1 represents newborn females.
A female of age i gives birth to r females with probability $b_{i,r}$, then survives, with probability $p_{i+1,i}$ to the next age $i+1$.
The mean number of female offspring by a female of age i is

$$b_i = b_{i,1} + 2b_{i,2} + 3b_{i,3} + \cdots .$$

Age k is the oldest age.

$$f_i(t_1, t_2, \ldots, t_k) = [p_{i+1,i}t_{i+1} + (1 - p_{i+1,i})] \sum_{r=0}^{\infty} b_{i,r}t_1^r, \quad i = 1, \ldots, k-1$$

$$f_k(t_1, t_2 \ldots, t_k) = b_{k,0} + b_{k,1}t_1 + b_{k,2}t_1^2 + \cdots = \sum_{r=0}^{\infty} b_{k,r}t_1^r.$$

Expectation Matrix is a Leslie Matrix

Expectation matrix:

$$\mathcal{M} = \begin{pmatrix} b_1 & b_2 & \cdots & b_{k-1} & b_k \\ p_{21} & 0 & \cdots & 0 & 0 \\ 0 & p_{32} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{k,k-1} & 0 \end{pmatrix},$$

In demography, matrix $\mathcal{M}$ is known as a Leslie matrix.

An Example with Two Ages or Stages.

Type 1 and Type 2

$$
\begin{aligned}
f_1(t_1,t_2) &= [(1/2)t_2 + 1/2][1/2 + (1/6)t_1 + (1/6)t_1^2 + (1/6)t_1^3] \\
f_2(t_1,t_2) &= 1/4 + (1/4)t_1 + (1/4)t_1^2 + (1/4)t_1^3.
\end{aligned}
$$

The mean number of offspring:

$$b_1 = b_{1,1} + 2b_{1,2} + 3b_{1,3} = (1/6) + 2(1/6) + 3(1/6) = 1$$

$$b_2 = b_{2,1} + 2b_{2,2} + 3b_{2,3} = (1/4) + 2(1/4) + 3(1/4) = 3/2.$$

Expectation Matrix

Expectation matrix:

$$\mathcal{M} = \begin{pmatrix} 1 & 3/2 \\ 1/2 & 0 \end{pmatrix}$$

has a dominant eigenvalue equal to $\lambda = 3/2$.

The fixed point of $f_i(q_1, q_2) = q_i$, $i = 1, 2$ **is**

$$(q_1, q_2) \approx (0.446, 0.443).$$

Acknowledgements

- Professors Lou Gross and Suzanne Lenhart for the invitation to co-organize this tutorial at NIMBioS.

- Co-organizers and lecturers, Professors Edward Allen, Jie Xiong, and Jose Miguel Ponciano, for their hard work in preparation for this tutorial.

- Participants that made this tutorial possible.