IE 536 Concise Notes,[1] v1.1[2]

Text in blue will be given on a test; ALL OTHER MATERIAL SHOULD BE COMMITTED TO MEMORY.

---

[1] This document is based in part on original materials shared by Profs. Mark Lewis, Joel Nachlas, Raghu Pasupathy, and Sid Resnick. Other material is drawn liberally and sometimes without reference from various texts — primarily the class text *Introduction to Probability Models* (Ross, 2010) [3], as well as from *An Introduction to Stochastic Modeling* (Taylor & Karlin, 1984, 1998) [4], *Introduction to Modeling and Analysis of Stochastic Systems* (Kulkarni, 2011) [1], and *Adventures in Stochastic Processes* (Resnick, 2005) [2].

[2] **This document has been proofread, but may have typos.** If you find a typo, email hunter63[AT]purdue[DOT]edu.

[3] All material in the IE 230 Concise Notes is also considered review material; IE 230 blue text will also be given on a test.

1 Review[1] and Preliminaries

Definition 1.1 (Matrix Multiplication)**.** If A and B are $m \times m$ matrices whose elements in the ith row and jth column are a_{ij} and b_{ij}, respectively, then $A \cdot B$ is defined to be the $m \times m$ matrix whose element in the ith row and jth column is $\sum_{k=1}^{m} a_{ik} b_{kj}$.

Definition 1.2 (Little-Oh)**.** A function $f(\cdot)$ is $o(h)$ ("little-oh-of-h") if $\lim_{h \to 0} \frac{f(h)}{h} = 0$.

Definition 1.3 (Sample Space)**.** The sample space Ω is the set of all possible outcomes of an experiment.

Definition 1.4 (Random Variable)**.** A random variable $Y \colon \Omega \to \mathcal{Y}, \mathcal{Y} \subseteq \mathbb{R}$ is a function that assigns a real number $Y(\omega) \in \mathcal{Y}$ to each outcome $\omega \in \Omega$.

Notation. Let $f \colon \mathbb{R} \to \mathbb{R}$ be a function. Usually, f or $f(\cdot)$ refers to the function, and $f(x)$ refers to the value of the function at $x \in \mathbb{R}$. Likewise, Y refer to the random variable, and $Y(\omega)$ refers to the value of the random variable at $\omega \in \Omega$. (Thus, Y is a random variable, $Y(\omega)$ is a number.)

Result 1.5. Let Y be a random variable that, with probability one (w.p.1), takes only non-negative values. Further, let $F_Y(y)$ denote its cdf. Then

$$\mathrm{E}[Y] = \begin{cases} \sum_{n=0}^{\infty} P\{Y > n\} = \sum_{n=0}^{\infty} (1 - F_Y(n)) & \text{if } Y \text{ takes on integer values } n \in \{0, 1, 2, \ldots\} \\ \int_0^{\infty} P\{Y > y\} dy = \int_0^{\infty} (1 - F_Y(y)) dy & \text{if } Y \text{ is continuous.} \end{cases}$$

Result 1.6 (Law of Total Probability)**.** Let A be an event in the sample space Ω, and let $Y \colon \Omega \to \mathcal{Y}$, $\mathcal{Y} \subseteq \mathbb{R}$ be a random variable. Then

$$P\{A\} = \begin{cases} \sum_y P\{A \,|\, Y = y\} P\{Y = y\} & \text{if } Y \text{ is discrete} \\ \int_{-\infty}^{\infty} P\{A \,|\, Y = y\} f_Y(y) dy & \text{if } Y \text{ is continuous.} \end{cases}$$

Remark 1. *You may be used to seeing this result as written in the IE 230 Concise Notes: If $A, B_1, B_2, \ldots, B_n$ are events in Ω such that $B_1, B_2, \ldots, B_n$ partition Ω, then $P\{A\} = \sum_{i=1}^{n} P\{A \,|\, B_i\} P\{B_i\}$. Notice that instead of indexing the conditioning events from $1, 2, \ldots, n$, the event $Y = y$ is indexed over all values of y, thus forming a partition. (Further, the event $Y = y$ can be considered shorthand for $B_y = \{\omega \in \Omega \colon Y(\omega) = y\}$.)*

Result 1.7 (Iterated Expectations)**.** Let X and Y be random variables. Then

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X \,|\, Y]] = \begin{cases} \sum_y \mathrm{E}[X \,|\, Y = y] P\{Y = y\} & \text{if } Y \text{ is discrete} \\ \int_{-\infty}^{\infty} \mathrm{E}[X \,|\, Y = y] f_Y(y) dy & \text{if } Y \text{ is continuous.} \end{cases}$$

Notation. Notice that $\mathrm{E}[X \,|\, Y]$ denotes a random variable, while $\mathrm{E}[X \,|\, Y = y]$ denotes a number.

2 Introduction to Stochastic Processes

Definition 2.1 (Stochastic Process, Informal)**.** A stochastic process $\{X(t), t \in \mathcal{T}\}$ is an ordered collection of random variables, where $\mathcal{T} \subseteq [0, \infty)$ is the index set and t is usually interpreted as time.

Definition 2.2 (State Space)**.** The state space $\mathcal{X}$ is the set of all possible values that the ordered collection of random variables $\{X(t), t \in \mathcal{T}\}$ can assume.

Definition 2.3 (Stochastic Process, More Formal)**.** A stochastic process $X \colon \Omega \times \mathcal{T} \to \mathcal{X}$, $\mathcal{X} \subset \mathbb{R}$ is a function that assigns a real number $X(\omega, t)$ to each outcome $\omega \in \Omega$ and index $t \in \mathcal{T}$.

Notation. For each $t \in \mathcal{T}$, the random variable $X(\cdot\,, t)$ is usually written as $X(t)$, which is the state of the process at time t. We often use $X(t)$ when the index is unspecified or continuous, and X_t when the index is discrete, e.g., $t \in \mathcal{T} = \{0, 1, 2, \ldots\}$, although this rule is not consistently applied.

[1] All material in the IE 230 Concise Notes is also considered review material; IE 230 blue text will also be given on a test.

Definition 2.4 (Sample Path)**.** For a an outcome $\omega \in \Omega$, the function $X(\omega, \cdot)$ is called a sample path. (That is, we fix the outcome ω so that $X(\omega, \cdot)$ is a function of time only; see Definitions 1.4 and 2.3).

Definition 2.5 (Markov Property)**.** A stochastic process $\{X(t), t \in \mathcal{T}\}$ has the Markov property if, for any set of n time points $t_1 < t_2 < \ldots < t_n$ in the index set, we have

$$P\{X(t_n) \leq x_n \,|\, X(t_{n-1}) = x_{n-1}, \ldots, X(t_2) = x_2, X(t_1) = x_1\} = P\{X(t_n) \leq x_n \,|\, X(t_{n-1}) = x_{n-1}\}.$$

Table 1: Stochastic Processes Classified by Time and State Space Type

	Discrete State Space	Continuous State Space
Discrete Time	Discrete Time Markov Chains	Time Series (Not Covered)
Continuous Time	Poisson Processes, Continuous Time Markov Chains, Renewal Processes	Brownian Motion

3 Discrete Time Markov Chains

3.1 Introduction to Discrete Time Markov Chains

Definition 3.1 (Discrete Time Markov Chain)**.** A discrete time Markov chain (DTMC) $\{X_n, n = 0, 1, 2, \ldots\}$ is a Markov process whose (time) index set is $\mathcal{T} = \{0, 1, 2, \ldots\}$, implying 'discrete time,' and whose state space $\mathcal{X}$ is a countable set, implying 'chain.' Thus, for all states $i_0, \ldots, i_{n-1}, i, j \in \mathcal{X}$ in and times $n = 0, 1, 2, \ldots$,

$$P\{X_{n+1} = j \,|\, X_n = i, X_{n-1} = i_{n-1}, \ldots X_0 = i_0\} = P\{X_{n+1} = j \,|\, X_n = i\}.$$

Definition 3.2 (Time Homogeneous DTMC)**.** A DTMC $\{X_n, n = 0, 1, 2, \ldots\}$ is time homogeneous if the one-step transition probabilities satisfy $P\{X_{n+1} = j \,|\, X_n = i\} = P\{X_{m+1} = j \,|\, X_m = i\}$ for all times $n, m \geq 0$ and states $i, j \in \mathcal{X}$. (We study only time homogeneous DTMC's. Henceforth, this property is implied.)

Definition 3.3 (One-Step Transition Probability)**.** For a Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$, the one-step transition probability is the probability that X_{n+1} is in state j, given that X_n is in state i, is

$$P_{ij} := P\{X_{n+1} = j \,|\, X_n = i\} \quad \text{for each for each } n \in \{0, 1, 2, \ldots\} \text{ and } i, j \in \mathcal{X}.$$

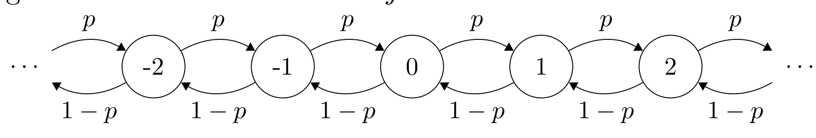
Notation. Without loss of generality, label the states with indices $i, j \geq 0$. Then, $\mathbb{P}$ is the matrix of one-step transition probabilities P_{ij}, where $\sum_{j=0}^{\infty} P_{ij} = 1$ for $i = 0, 1, 2, \ldots$ (each row sums to one) and

$$\mathbb{P} := \begin{matrix} 0 \\ 1 \\ \vdots \\ i \\ \vdots \end{matrix} \begin{pmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

Result 3.4. A (time homogeneous) Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$ is completely defined once its transition probability matrix *and* the probability distribution of its initial state, $\mathbf{a}_0$, are specified, where

$$\mathbf{a}_0 = (a_{i_0}, a_{i_1}, \ldots) = (P\{X_0 = i_0\}, P\{X_0 = i_1\}, \ldots) \text{ for each state } i_0, i_1, \ldots \in \mathcal{X}.$$

Example 3.5 (Random Walk)**.** A one-dimensional random walk is a DTMC whose state space is the integers $\mathcal{X} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ and, for $p \in (0, 1)$, the transition probabilities are $P_{i,i+1} = p = 1 - P_{i,i-1}$. We can visualize this DTMC using the following diagram, where the probability of going from each state i to state j is written along an arc from state i to state j:



3.2 Calculating the Probability of a Particular Sample Path

Result 3.6 (Probability of a Sample Path[1]). For a Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$, the probability of observing a particular sample path is

$$P\{X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = P_{i_{n-1},i_n} P_{i_{n-2},i_{n-1}}, \ldots, P_{i_0,i_1} a_{i_0},$$

where $i_n, i_{n-1}, i_{n-2} \ldots, i_1, i_0 \in \mathcal{X}$ are states and $a_i = P\{X_0 = i\}$ is the probability the chain starts in state i.

3.3 Chapman-Kolmogorov Equations

Definition 3.7 (n-Step Transition Probability). The probability that the Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$ currently in state i will be in state j after $n \geq 0$ steps is

$$P_{ij}^{(n)} \coloneqq P\{X_{k+n} = j \,|\, X_k = i\}.$$

Result 3.8 (Chapman-Kolmogorov Equations[1]). Without loss of generality, label the states of the DTMC with indices $i, j, k \geq 0$. The Chapman-Kolmogorov equations, stated as

$$P_{ij}^{(n+m)} = \sum_{k=0}^{\infty} P_{ik}^{(n)} P_{kj}^{(m)} \quad \text{for all } n, m \geq 0 \text{ and all states } i, j,$$

provide a method for computing the n-step transition probabilities by noting that a process going from state i to state j in $n + m$ steps must pass through an intermediary state k in n steps. In matrix notation, the Chapman-Kolmogorov equations are $\mathbb{P}^{(n+m)} = \mathbb{P}^{(n)} \cdot \mathbb{P}^{(m)}$ for all $n, m \geq 0$.

Result 3.9 (Calculating the n-Step Transition Probability Matrix[1]). The n-step transition probability matrix of a DTMC is $\mathbb{P}^{(n)} = \mathbb{P}^n$, which represents $\mathbb{P}$ multiplied with itself n times.

Definition 3.10 (Transient Distribution). For a Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$, using the notation that $a_i = P\{X_0 = i\}$ for a state i, let $\boldsymbol{a}_n = (P\{X_n = i_0\}, P\{X_n = i_1\}, ...)$ denote a vector containing the (unconditional) probability distribution of X_n at time n. The vector $\boldsymbol{a}_n$ is called the transient distribution of the Markov chain at time n.

Result 3.11. The transient distribution of a DTMC is obtained by conditioning on the initial state. That is,

$$P\{X_n = j\} = \sum_{i=0}^{\infty} P\{X_n = j | X_0 = i\} P\{X_0 = i\} = \sum_{i=0}^{\infty} a_i P_{ij}^{(n)}.$$

In matrix notation, we have $\boldsymbol{a}_n = \boldsymbol{a}_0 \mathbb{P}^{(n)} = \boldsymbol{a}_0 \mathbb{P}^n$.

3.4 Classification of States

3.4.1 Absorbing States

Definition 3.12 (Absorbing State). An absorbing state is a state that, once entered, cannot be left. That is, if i is an absorbing state of a DTMC, then $P_{ii} = 1$ and $P_{ij} = 0$ for all states $j \neq i$.

Result 3.13. Let $\{X_n, n = 0, 1, 2, \ldots\}$ be a Markov chain for which the set of states $\mathcal{A} \neq \emptyset$ is the set of all absorbing states. Then for non-absorbing states $i, j \notin \mathcal{A}$, $P_{ij}^{(n)}$ represents the probability that the chain will go from i to j in n steps without ever entering any states in $\mathcal{A}$,

$$P_{ij}^{(n)} = P\{X_n = j, X_k \notin \mathcal{A}, k = 1, \ldots, n \,|\, X_0 = i\}.$$

For states $i, j \notin \mathcal{A}$, the probability the chain goes from state i to j in n steps, given it is not "absorbed," is

$$P\{X_n = j \,|\, X_0 = i, X_k \notin \mathcal{A}, k = 1, \ldots, n\} = \frac{P\{X_n = j, X_k \notin \mathcal{A}, k = 1, \ldots, n \,|\, X_0 = i\}}{P\{X_k \notin \mathcal{A}, k = 1, \ldots, n \,|\, X_0 = i\}} = \frac{P_{ij}^{(n)}}{\sum_{r \notin \mathcal{A}} P_{ir}^{(n)}}.$$

[1]You should be able to prove this result and justify all the steps.

Remark 2. *Suppose you want to know the probability a Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$ with transition probabilities P_{ij} enters any of a specified set of states $\mathcal{A}$ by time n. Create a new Markov chain identical to $\{X_n, n = 0, 1, 2, \ldots\}$ but with one-step transition probabilities Q_{ij} that reset the states in $\mathcal{A}$ to be absorbing,*

$$Q_{ij} = \begin{cases} 1 & \text{if } i \in \mathcal{A}, j = i \\ 0 & \text{if } i \in \mathcal{A}, j \neq i \\ P_{ij} & \text{otherwise.} \end{cases}$$

Then, analyze the new chain using the information in Result 3.13.

3.4.2 Transient and Recurrent States

Definition 3.14 (Recurrent State)**.** Let f_i be the probability that, starting in state i, the DTMC ever returns to state i. State i is recurrent if $f_i = 1$. That is, starting in state i, the DTMC returns to state i w.p.1.

Definition 3.15 (Transient State)**.** State i is transient if $f_i < 1$; that is, starting in state i, the DTMC escapes state i forever with probability (w.p.) $(1 - f_i) > 0$.

Result 3.16. Let V_i denote the number of visits to transient state i until the DTMC escapes transient state i forever (which is equivalent to the number of time periods the DTMC spends in transient state i). Then given the DTMC starts in state i, V_i is a geometric(p) random variable, where $p = (1 - f_i)$ is the probability of escaping state i forever (that is, the probability of "success"). That is,

$$P\{V_i = n \,|\, X_0 = i,\ i \text{ is transient}\} = f_i^{n-1}(1 - f_i), \quad n = 1, 2, \ldots$$

is the probability that, starting in state i, the DTMC will be in state i exactly n times, and

$$\mathrm{E}[V_i \,|\, X_0 = i,\ i \text{ is transient}] = 1/(1 - f_i) < \infty.$$

Result 3.17. [1]For any state i, $\sum_{n=0}^{\infty} P_{ii}^{(n)}$ equals the expected number of visits to state i, starting in state i.

Result 3.18. State i is recurrent if and only if, starting in state i, the expected number of visits to state i is infinite. Thus, state i is recurrent if $\sum_{n=0}^{\infty} P_{ii}^{(n)} = \infty$ and transient if $\sum_{n=0}^{\infty} P_{ii}^{(n)} < \infty$.

3.4.3 Classes and Class Properties

Definition 3.19 (Accesible)**.** State j is accessible from state i if $P_{ij}^{(n)} > 0$ for some $n \geq 0$.

Definition 3.20 (Communicating States)**.** Two states i and j that are accessible to each other communicate, which is denoted $i \leftrightarrow j$, and is equivalent to $j \leftrightarrow i$.

Remark 3. *Any state communicates with itself since $P_{ii}^{(0)} = 1$. Further, by the Chapman-Kolmogorov equations, if $i \leftrightarrow k$ and $k \leftrightarrow j$, then $i \leftrightarrow j$.*

Definition 3.21 (Communicating Class)**.** Two states that communicate are in the same *class*.

Remark 4. *By the properties of communication, any two classes of states are either identical or disjoint.*

Definition 3.22 (Irreducible)**.** A DTMC is irreducible if there is only one class, that is, if all states communicate with each other.

Result 3.23. If state i is recurrent, and state i communicates with state j, then state j is recurrent. Therefore, recurrence and transience are *class properties.*

Remark 5. *For a finite-state DTMC, not all states can be transient (there must be at least one recurrent state). All states of a finite, irreducible DTMC are recurrent.*

[1] You should be able to prove this result and justify all the steps.

Definition 3.24 (Period)**.** A state i has period d if any return to state i must occur in multiples of d time steps, e.g., $P_{ii}^{(n)} = 0$ unless $n = k \times d \geq 1$ for some integer k, and d is the largest integer with this property. That is, $d = \gcd\{n\colon P_{ii}^{(n)} > 0\}$, where gcd stands for the "greatest common divisor."

Definition 3.25 (Aperiodic)**.** A state with period $d = 1$ is said to be aperiodic. A state i is aperiodic if there exists n such that for all $n' \geq n$, $P_{ii}^{(n')} > 0$.

Remark 6. *Periodicity is a class property: if state i has period d and $i \leftrightarrow j$, then j has period d.*

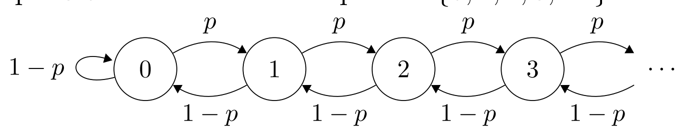
Definition 3.26 (Positive Recurrent State)**.** If state i is recurrent, then it is positive recurrent if, starting in state i, the expected time until the process returns to state i is finite.

Definition 3.27 (Ergodic)**.** Positive recurrent, aperiodic states are called ergodic.

Remark 7. *Positive recurrence, like recurrence, is a class property. In a finite-state DTMC, all recurrent states are positive recurrent.*

Remark 8. *Recurrent states that are not positive recurrent are called* null recurrent*.*

Example 3.28. Consider the DTMC defined by a one-dimensional random walk on the non-negative integers. That is, modify Example 3.5 so that the state space is $\{0, 1, 2, 3, \ldots\}$ and $P_{00} = 1 - p$, as follows:



This chain is irreducible. It is positive recurrent if and only if $p < 1/2$. If $p = 1/2$, it is null recurrent, and if $p > 1/2$, it is transient.

3.5 Limiting Probabilities

Definition 3.29 (Limiting Distribution)**.** Without loss of generality, label the states of the Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$ with indices $i, j \geq 0$. Assuming it exists, for each $j \geq 0$, let

$$\pi_j = \lim_{n\to\infty} P\{X_n = j | X_0 = i\} = \lim_{n\to\infty} P_{ij}^{(n)},$$

and, if it exists, let $\boldsymbol{\pi} = (\pi_0, \pi_1, \pi_2, \ldots)$ denote the limiting distribution of $\{X_n, n \geq 0\}$.

Theorem 3.30. If a DTMC is irreducible, positive recurrent, and aperiodic, then π_j exists and does not depend on the initial state i. Furthermore, π_j is the unique nonnegative solution to the system of equations

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad j \geq 0, \quad (\text{"balance" equations}); \qquad 1 = \sum_{j=0}^{\infty} \pi_j \quad (\text{"normalizing" equation}).$$

In matrix notation, these equations are written as $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbb{P}$, $\sum_{j=0}^{\infty} \pi_j = 1$.

Definition 3.31 (Stationary Distribution)**.** A distribution $\tilde{\boldsymbol{\pi}} = (\tilde{\pi}_0, \tilde{\pi}_1, \tilde{\pi}_2, \ldots)$ is called a stationary distribution if setting the initial distribution $\mathbf{a}_0 = \tilde{\boldsymbol{\pi}}$ implies the transient distribution $\mathbf{a}_n = \tilde{\boldsymbol{\pi}}$ for all $n \geq 0$.

Theorem 3.32. A distribution $\tilde{\boldsymbol{\pi}}$ is a stationary distribution if and only if it satisfies $\tilde{\boldsymbol{\pi}} = \tilde{\boldsymbol{\pi}}\mathbb{P}$, $\sum_{j=0}^{\infty} \tilde{\pi}_j = 1$.

Corollary 3.33. If the limiting distribution $\boldsymbol{\pi}$ exists, it is also a stationary distribution.

Result 3.34. If a DTMC is irreducible, the solution to $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbb{P}$, $\sum_{j=0}^{\infty} \pi_j = 1$ exists if and only if the Markov chain is positive recurrent.

Remark 9. *If a DTMC is irreducible and there does not exist a solution to $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbb{P}$, $\sum_{j=0}^{\infty} \pi_j = 1$, then it is either transient or null recurrent, and $\pi_j = 0$ for each state j.*

Result 3.35. If a DTMC is irreducible and positive recurrent, the solution to $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbb{P}$, $\sum_{j=0}^{\infty} \pi_j = 1$ is unique, $\boldsymbol{\pi}$ is a stationary distribution, and π_j is the long-run proportion of time the process is in state j.

Remark 10. *If a DTMC is irreducible and positive recurrent, the solution to $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbb{P}$, $\sum_{j=0}^{\infty} \pi_j = 1$ exists and is unique, but is* not necessarily *a limiting distribution. (The limit may not exist due to periodicity.)*

Result 3.36. [2, p. 138] Let $\mathbf{1}_{m\times n}$ be an $m \times n$ matrix all of whose entries equal one, let $\mathbb{I}$ be the identity matrix, and let A^{-1} denote the matrix inverse for any matrix A. If $\boldsymbol{\pi}$ is the stationary distribution of an irreducible DTMC with transition probability matrix $\mathbb{P}$, then

$$\boldsymbol{\pi} = \mathbf{1}_{1\times m}(\mathbb{I} - \mathbb{P} + \mathbf{1}_{m\times m})^{-1}.$$

Result 3.37. If a DTMC is irreducible and positive recurrent, then starting in state j, the expected number of steps to return to j is $m_{jj} = 1/\pi_j$, where $m_{jj} \geq 1$.

Remark 11. *Since π_j is the long-run proportion of time an irreducible, positive recurrent DTMC spends in state j, then in expectation, the chain spends one unit of time in state j per m_{jj} units of time.*

3.6 Rewards

Result 3.38. Let $\{X_n, n = 0, 1, 2, \ldots\}$ be an irreducible Markov chain with stationary probabilities $\pi_j, j \geq 0$, and let r be a bounded function on the state space. Then w.p.1, the expected reward per unit time is

$$\lim_{m\to\infty} \tfrac{1}{m} \textstyle\sum_{n=0}^{m-1} r(X_n) = \sum_{j=0}^{\infty} r(j)\pi_j.$$

3.7 Mean Time Spent in Transient States

Definition 3.39 (Matrix of One-Step Transition Probabilities for Transient States)**.** Let $\{X_n, n = 0, 1, 2, \ldots\}$ be a finite-state Markov chain, and let $\mathcal{X}_T = \{1, 2, \ldots, t\}$ be the set of transient states. Define the matrix containing one-step transition probabilities for the transient states as

$$\mathbb{P}_T := \begin{pmatrix} P_{11} & P_{12} & \ldots & P_{1t} \\ \vdots & \vdots & \vdots & \vdots \\ P_{t1} & P_{t2} & \ldots & P_{tt} \end{pmatrix}.$$

Remark 12. *Notice that at least one row of $\mathbb{P}_T$ does not sum to one.*

Result 3.40. [1]For transient states $i, j \in \mathcal{X}_T$, let $\delta_{ij} := 1$ represent the initial visit if $i = j$; otherwise, $\delta_{ij} := 0$. The expected number of time periods the DTMC spends in state j given that it starts in state i is

$$\mathrm{E}[V_j \mid X_0 = i,\ i, j \text{ are transient}] = s_{ij} = \delta_{ij} + \textstyle\sum_{k=1}^{t} P_{ik}s_{kj},$$

where the random variable V_j is defined in Result 3.16. In matrix notation, letting $\mathbb{I}$ be the identity matrix and $\mathbb{S}$ be a matrix whose (i,j)th entry is s_{ij}, we have $\mathbb{S} = \mathbb{I} + \mathbb{P}_T\,\mathbb{S}$, which implies $\mathbb{S} = (\mathbb{I} - \mathbb{P}_T)^{-1}$.

Result 3.41. Let f_{ij} be the probability that, starting in transient state i, the DTMC ever visits transient state j. Then

$$f_{ij} = \frac{s_{ij} - \delta_{ij}}{s_{jj}}.$$

Remark 13. *This result is consistent with Result 3.16 since $f_{ii} = (s_{ii} - 1)/s_{ii}$ implies $s_{ii} = 1/(1 - f_{ii})$.*

3.8 Expected First Passage Times

Theorem 3.42. [1]Let $\{X_n, n = 0, 1, 2, \ldots\}$ be a finite-state DTMC, and $T_j = \min\{n \geq 0 : X_n = j\}$. Then

$$\mathrm{E}[T_j \mid X_0 = i] = 1 + \textstyle\sum_{k\neq j} \mathrm{E}[T_j \mid X_0 = k]P_{ik}.$$

[1]You should be able to prove this result, which relies on a first step analysis (see Ch. 3 §4 of [4]), and justify all the steps.

3.9 Special Case: Branching Processes

Definition 3.43 (Branching Process)**.** A branching process $\{X_n, n = 0, 1, 2, \ldots\}$ is a DTMC designed to study the size of a population capable of producing offspring of the same kind. It has the following properties:

- The state space is $\mathcal{X} = \{0, 1, 2, \ldots\}$.
- The state of the process at time n, X_n, represents the number of individuals in the nth generation. The individuals in generation n are the offspring of generation $n-1$ for all generations $n \geq 1$.
- Each individual produces J offspring independent of all other individuals, where J is a discrete random variable whose support is a subset of $\{0, 1, 2, \ldots\}$ and no one number of offspring is certain, that is, $P\{J = j\} < 1$ for all $j = 0, 1, 2, \ldots$. The expected value and variance of a single individual's number of offspring are $\mu := \text{E}[J] = \sum_{j=0}^{\infty} jP\{J = j\}$ and $\sigma^2 := \text{Var}(J) = \text{E}[(J - \mu)^2] = \sum_{j=0}^{\infty}(j - \mu)^2 P\{J = j\}$.
- The classes are $\{0\}$, which is recurrent and absorbing, and $\{1, 2, 3, \ldots\}$

Result 3.44. If $P\{J = 0\} > 0$, then the population dies out or its size tends to infinity, and states $\{1, 2, 3, \ldots\}$ are transient.

Remark 14. *$P_{i0} = P\{J = 0\}^i$ is the probability all i individuals in generation n have no offspring, so that $X_{n+1} = 0$. Thus, $P_{i0} = P\{J = 0\}^i$ is a lower bound on the probability that we escape state i forever, that is, no later generation will ever consist of i individuals.*

Remark 15. *Henceforth, we start with one individual, $X_0 = 1$. Notice that we can analyze $X_0 = 1$ and when $X_0 = m$, we treat each of the m individuals as a new branching process with $X_0 = 1$.*

Result 3.45. ([1]for $\text{E}[X_n]$) Let $X_0 = 1$. The expected value of the number of individuals in the nth generation is $\text{E}[X_n] = \mu^n$. If $\mu = 1$, then $\text{Var}(X_n) = n\sigma^2$; if $\mu \neq 1$, $\text{Var}(X_n) = \sigma^2\mu^{n-1}(1 - \mu^n)/(1 - \mu)$.

Result 3.46. ([1]except $\mu = 1$ case.) Let π_0 denote the probability that the population eventually dies out,

$$\pi_0 := \lim_{n\to\infty} P\{X_n = 0 | X_0 = 1\},$$

let $p_j := P\{J = j\}$, and assume $X_0 = 1$. If the expected number of offspring of each individual is less than or equal to one, $\mu \leq 1$, then $\pi_0 = 1$. If $\mu > 1$, then π_0 is the unique solution to $\pi_0 = \sum_{j=0}^{\infty}(\pi_0)^j p_j$ in $[0, 1)$.

Remark 16. *When $\mu > 1$, [2] provides a proof for why a root at $\pi_0 = 1$ is* not *the probability that the population eventually dies out, starting on p. 21.*

4 Poisson Processes

4.1 Properties of Exponential Random Variables

Definition 4.1 (Exponential Distribution)**.** A continuous random variable X has an exponential distribution with parameter $\lambda, \lambda > 0$, if its pdf is $f_X(x) = \lambda e^{-\lambda x}, x \geq 0$ or, equivalently, its cdf is $F_X(x) = 1 - e^{-\lambda x}, x \geq 0$.

Result 4.2. ([1]for $\text{E}[X]$) If a random variable $X \sim \text{exponential}(\lambda)$, $\lambda > 0$, then $\text{E}[X] = 1/\lambda$, $\text{Var}(X) = 1/\lambda^2$.

Definition 4.3 (Memoryless Property)**.** A random variable X is said to be memoryless if, for all $s, t \geq 0$,

$$P\{X > s + t \mid X > t\} = P\{X > s\}.$$

Result 4.4. [1]The exponential distribution is memoryless.

Remark 17. *The exponential distribution is the only continuous distribution that satisfies the memoryless property. (Recall that the geometric distribution, which is a discrete distribution, is is also memoryless.)*

[1]You should be able to prove this result and justify all the steps.

Definition 4.5 (Hazard Function)**.** The hazard function (or failure rate) of a random variable X is

$$h(t) = \frac{f_X(t)}{(1 - F_X(t))}$$

Result 4.6. [1]The hazard function for an exponential(λ) random variable is constant, with $h(t) = \lambda$.

Result 4.7 (Property A)**.** If $X_1, X_2, \ldots, X_n$ are independent and identically distributed (iid) exponential(λ), $\lambda > 0$, random variables, then the pdf of $X_1 + X_2 + \ldots + X_n$ is

$$f_{X_1+X_2+\ldots+X_n}(t) = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t}, \quad t > 0,$$

which is an Erlang(n, λ) distribution with $\Gamma(n) = (n-1)!$. It is also a gamma(n, λ) distribution, where the gamma family is a generalization of the Erlang family that allows n to take on non-integer values.

Result 4.8 (Property B[1])**.** If $X_1, X_2, \ldots, X_n$ are mutually independent random variables with $X_i \sim$ exponential(λ_i), $\lambda_i > 0$ for $i = 1, 2, \ldots, n$, then

$$\min(X_1, X_2, \ldots, X_n) \sim \text{exponential}\left(\textstyle\sum_{i=1}^n \lambda_i\right).$$

Result 4.9 (Property C[1])**.** Consider $X_1 \sim$ exponential(λ_1), $\lambda_1 > 0$ and $X_2 \sim$ exponential(λ_2), $\lambda_2 > 0$, where X_1 and X_2 are independent. Then

$$P\{X_1 < X_2\} = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Result 4.10 (Property D)**.** Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables with $X_i \sim$ exponential(λ_i), $\lambda_i > 0$ for each $i = 1, 2, \ldots, n$. Then combine Properties B and C to see that the probability X_i is the smallest is

$$P\{X_i = \min_j X_j\} = P\{X_i < \min_{j \neq i} X_j\} = \frac{\lambda_i}{\lambda_i + \sum_{j \neq i} \lambda_j} = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}.$$

Result 4.11 (Property E[1])**.** Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables with $X_i \sim$ exponential(λ_i), $\lambda_i > 0$ for each $i = 1, 2, \ldots, n$. Then $\min_{1 \leq i \leq n} X_i$ is independent of the rank ordering of the X_i's. That is,

$$P\{X_{i_1} < \ldots < X_{i_n} \mid \min_{1 \leq i \leq n} X_i > t\} = P\{X_{i_1} < \ldots < X_{i_n}\}.$$

Result 4.12 (Property F)**.** Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables with $X_i \sim$ exponential(λ_i), $\lambda_i > 0$ for each $i = 1, 2, \ldots, n$ and $\lambda_i \neq \lambda_j$ for all $i \neq j$. Then the sum $\sum_{i=1}^n X_i$ is hypoexponential (or generalized Erlang) with pdf

$$f_{X_1+\ldots+X_n}(t) = \sum_{i=1}^n \Big(\prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}\Big) \lambda_i e^{-\lambda_i t}, \quad t > 0.$$

4.2 Counting Processes

Definition 4.13 (Counting Process)**.** A stochastic process $\{N(t), t \geq 0\}$ is a counting process if $N(t) \geq 0$ and

(i) $N(t)$ is integer-valued;
(ii) if $s < t$, then $N(s) \leq N(t)$;
(iii) if $s < t$, then $N(t) - N(s)$ equals the number of events that occur in the interval $(s, t]$.

Definition 4.14 (Independent Increments)**.** A counting process $\{N(t), t \geq 0\}$ has *independent increments* if the number of events that occur in disjoint time intervals are independent (regardless of the interval lengths).

Definition 4.15 (Stationary Increments)**.** A counting process $\{N(t), t \geq 0\}$ has *stationary increments* if the distribution of the number of events that occur in the interval $(s, s+t)$, $N(s+t) - N(s)$, is a function of the length of the interval t *only* (and thus not a function of s).

[1]You should be able to prove this result and justify all the steps.

4.3 Three Equivalent Definitions of a Poisson Process

Definition 4.16 (PP Definition I)**.** The counting process $\{N(t), t \geq 0\}$ is a Poisson Process with rate λ, $\lambda > 0$ (PP(λ)), if $N(0) = 0$ and

(i) the process has independent increments;

(ii) the number of events in any interval of length t is Poisson distributed with mean λt. That is, for all $s, t \geq 0$, $\mathrm{E}[N(s+t) - N(s)] = \lambda t$ and

$$P\{N(s+t) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \ldots$$

Definition 4.17 (PP Definition II; see Defintion 1.2 for $o(\cdot)$ notation)**.** The counting process $\{N(t), t \geq 0\}$ is a Poisson Process with rate λ, $\lambda > 0$ (PP(λ)) if $N(0) = 0$ and

(i) the process has independent increments;

(ii) $P\{N(s+h) - N(s) = 1\} = \lambda h + o(h)$; and

(iii) $P\{N(s+h) - N(s) \geq 2\} = o(h)$.

Remark 18. *Notice that both Definitions I and II imply that a PP(λ) has stationary increments. For intuition on showing that Definition I implies II, expand $P\{N(h) = 1\}$ in a Taylor series about zero. That Definition II implies I follows from the Poisson approximation to the binomial.*

Result 4.18 (Exponential Interarrivals ([1]for T_1, T_2))**.** Let $\{T_n, n = 1, 2, \ldots\}$ be the sequence of interarrival times for a PP(λ), $\lambda > 0$. Then $T_n, n = 1, 2, \ldots,$ are i.i.d exponential(λ) random variables with mean $1/\lambda$.

Result 4.19. Let $S_n = \sum_{i=1}^n T_i, n = 1, 2, \ldots$ be the arrival time of the nth event, which is also the waiting time until the nth event. Since $T_i, n = 1, 2, \ldots$ are iid exponential(λ), $\lambda > 0$, then $S_n \sim$ gamma(n, λ) with

$$f_{S_n}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}, \; t \geq 0,$$

where $\mathrm{E}[S_n] = n/\lambda$ and $\mathrm{Var}(S_n) = n/\lambda^2$.

Remark 19. *Notice that $N(t) \geq n \Leftrightarrow S_n \leq t$. Further, using the cdf of an Erlang(n, λ) random variable, we can show that*

$$P\{S_n > t\} = \sum_{i=0}^{n-1} e^{-\lambda t}\frac{(\lambda t)^i}{i!} = \sum_{i=0}^{n-1} P\{N(t) = i\}.$$

Definition 4.20 (PP Definition III)**.** Let $\{N(t), t \geq 0\}$ be a counting process with inter-arrival time sequence $\{T_n, n = 1, 2, \ldots\}$, arrival times $S_n = \sum_{i=1}^n T_i, n \geq 1$, and $N(t) = \max\{n \geq 0 : S_n \leq t\}, t \geq 0$. Then $\{N(t), t \geq 0\}$ is PP(λ) if $\{T_n, n = 1, 2, \ldots\}$ is a sequence of iid exponential(λ), $\lambda > 0$ random variables.

Theorem 4.21. Let $\{N(t), t \geq 0\}$ be a PP(λ), $\lambda > 0$. Then it has the Markov property at each time t:

$$P\{N(t+s) = j \mid N(s) = i, N(u) = i', 0 \leq u \leq s, i' \leq i\} = P\{N(t+s) = j \mid N(s) = i\}.$$

4.4 Superposition of Poisson Processes and Thinning of a Poisson Processes

Result 4.22 (Superposition)**.** Let $\{N_A(t), t \geq 0\}$ be a PP(λ_A) and $\{N_B(t), t \geq 0\}$ be a PP(λ_B), where these two processes are independent of each other. Let $N(t) = N_A(t) + N_B(t)$ be the total number of events by time t, so that $\{N(t), t \geq 0\}$ is the superposition of these two processes. Then $\{N(t), t \geq 0\}$ is a PP($\lambda_A + \lambda_B$).

Result 4.23 (Thinning)**.** Let $\{N(t), t \geq 0\}$ be a PP(λ) in which an arrival is type A w.p. $p \in (0, 1)$ and type B w.p. $(1-p)$, independent of everything else. Let $N_A(t)$ and $N_B(t)$ be the number of type A and type B events occurring by time t, respectively. Then $\{N_A(t), t \geq 0\}$ and $\{N_B(t), t \geq 0\}$ are independent Poisson processes, where $\{N_A(t), t \geq 0\}$ is a PP(λp) and $\{N_B(t), t \geq 0\}$ is a PP($\lambda(1-p)$).

[1]You should be able to prove this result and justify all the steps.

4.5 Conditional Distribution of the Arrival Times

Definition 4.24 (Order Statistics)**.** Let $Y_1, Y_2, \dots, Y_n$ be n random variables. Then $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$ are the order statistics corresponding to $Y_1, Y_2, \dots, Y_n$ if $Y_{(k)}$ is the kth smallest value among $Y_1, Y_2, \dots, Y_n$.

Remark 20. *As special cases, the first order statistic is the minimum* $Y_{(1)} = \min\{Y_1, Y_2, \dots, Y_n\}$ *and the last order statistic is the maximum* $Y_{(n)} = \max\{Y_1, Y_2, \dots, Y_n\}$.

Theorem 4.25. ([1]for $n = 1$ case) Given that $N(t) = n$, the n arrival times $S_1, S_2, \dots, S_n$ of a PP(λ) have the same distribution as the order statistics corresponding to n independent uniform$(0, t)$ random variables.

4.6 Nonhomogeneous Poisson Processes

Definition 4.26 (NHPP Definition I)**.** The counting process $\{N(t), t \geq 0\}$ is a nonhomogeneous Poisson Process with intensity function $\lambda(t)$, $t \geq 0$ $\big(\text{NHPP}(\lambda(t))\big)$, if $N(0) = 0$ and

(i) the process has independent increments;

(ii) the number of events in any interval of length t starting at any time $s \geq 0$, $N(s+t) - N(s)$, is Poisson distributed with mean $\int_s^{s+t} \lambda(y)dy$. That is, for all $s, t \geq 0$, $\text{E}[N(s+t) - N(s)] = \int_s^{s+t} \lambda(y)dy$ and

$$P\{N(s+t) - N(s) = n\} = \exp\big(-\textstyle\int_s^{s+t} \lambda(y)dy\big) \frac{\big(\int_s^{s+t} \lambda(y)dy\big)^n}{n!}, \quad n = 0, 1, 2, \dots$$

Definition 4.27 (NHPP Definition II)**.** The counting process $\{N(t), t \geq 0\}$ is a nonhomogeneous Poisson Process with intensity function $\lambda(t)$, $t \geq 0$ $\big(\text{NHPP}(\lambda(t))\big)$ if $N(0) = 0$ and

(i) the process has independent increments;

(ii) $P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$; and

(iii) $P\{N(t+h) - N(t) \geq 2\} = o(h)$.

Remark 21. *Notice that we still have independent increments for an NHPP($\lambda(t)$), but do not have stationary increments. (This fact makes the process is more flexible for modeling.) Further, these definitions reduce to a PP(λ) if $\lambda(t) = \lambda$ for all $t \geq 0$, so that $N(t)$ has a Poisson distribution with mean* $\text{E}[N(t)] = \int_0^t \lambda dy = \lambda t$.

Result 4.28. Let $\{N(t), t \geq 0\}$ be a PP(λ) and suppose an event at time t, independent of events prior to t, is counted as Type A w.p. $p(t) \in (0, 1)$. Then the process of type A events, $\{N_A(t), t \geq 0\}$ is an NHPP($\lambda p(t)$).

Result 4.29. Let $\{N_A(t), t \geq 0\}$ be an NHPP($\lambda(t)$), where $\lambda(t) \leq \lambda$ for all $t \geq 0$. Further, let $\{N_B(t), t \geq 0\}$ be an NHPP($\lambda - \lambda(t)$), independent of $\{N_A(t), t \geq 0\}$. Then $\{N_A(t) + N_B(t), t \geq 0\}$ is a PP(λ), and $\{N_A(t), t \geq 0\}$ can be regarded as the process of time-sampled events w.p. $p(t) = \lambda(t)/\lambda$.

Definition 4.30 (Cumulative Intensity Function)**.** For an NHPP($\lambda(t)$), the cumulative intensity function is

$$\Lambda(t) := \text{E}[N(t)] = \textstyle\int_0^t \lambda(y)dy.$$

(Notice that $\text{E}[N(s+t) - N(s)] = \Lambda(s+t) - \Lambda(s)$.)

Theorem 4.31. The random variables $S_1, S_2, \dots$ are event times corresponding to an NHPP with cumulative intensity function $\Lambda(t)$ if and only if $\Lambda(S_1), \Lambda(S_2), \dots$ are the event times corresponding to a PP($\lambda = 1$).

Theorem 4.32. [2]Let $S_1, S_2, \dots$ be random variables representing the event times of an NHPP with continuous cumulative intensity function $\Lambda(t)$. Consider the events in the interval $[0, t^*]$. Conditional on $N(t^*) = n$, the event times $S_1, S_2, \dots$ have the same distribution as the order statistics of a random sample from the distribution with cdf

$$F(t) = \frac{\Lambda(t)}{\Lambda(t^*)} \text{ for } t \in [0, t^*].$$

[1]You should be able to prove this result and justify all the steps.

[2]This result is included for your interest.

5 Continuous Time Markov Chains

5.1 Introduction to Continuous Time Markov Chains

Definition 5.1 (Sojourn Time)**.** Let $\{X(t), t \geq 0\}$ be a stochastic process with state space $\mathcal{X}$. The sojourn time in state $i \in \mathcal{X}$ is the amount of time the process spends in state i before it transitions to a different state $j \in \mathcal{X}, j \neq i$.

Definition 5.2 (CTMC Definition I)**.** A continuous time Markov chain (CTMC) $\{X(t), t \geq 0\}$ is a Markov process whose (time) index set is $\mathcal{T} = \{t \in \mathbb{R} \colon t \geq 0\}$, implying 'continuous time,' and whose state space $\mathcal{X}$ is a countable set, implying 'chain.' Thus, for all states $i, j \in \mathcal{X}$ and times $t, s \geq 0$ of a CTMC,

$$P\{X(s+t) = j \mid X(s) = i, X(u), 0 \leq u < s\} = P\{X(s+t) = j \mid X(s) = i\}.$$

Definition 5.3 (Time Homogeneous CTMC)**.** A CTMC $\{X(t), t \geq 0\}$ is *time homogeneous* if the probability that the CTMC, presently in state i, will be in state j after t time units satisfies

$$P_{ij}(t) \coloneqq P\{X(s+t) = j \mid X(s) = i\} = P\{X(t) = j \mid X(0) = i\}$$

for all times $s, t \geq 0$ and states $i, j \in \mathcal{X}$. (We study only time homogeneous CTMC's.)

Definition 5.4 (CTMC Definition II)**.** A stochastic process $\{X(t), t \geq 0\}$ on state space $\mathcal{X}$ is a CTMC if each time the process enters a state $i \in \mathcal{X}$,

(i) the sojourn time $T_i \sim$ exponential(ν_i), $\nu_i > 0$; and

(ii) when the process leaves state i, it enters state j w.p. P_{ij}, where $P_{ii} = 0$ and $\sum_j P_{ij} = 1$ for all $i \in \mathcal{X}$.

Remark 22. *If we consider a CTMC only at transition times, we can define a corresponding DTMC with transition probabilities P_{ij}, Let this DTMC be called the* embedded DTMC*. Also, notice that unlike in a DTMC, in a CTMC we require $P_{ii} = 0$. This requirement results from the way we define the sojourn times.*

Result 5.5. A (time homogeneous) CTMC is completely defined once the following have been determined: (a) the transition probabilities P_{ij} for all $i, j \in \mathcal{X}$; (b) the mean sojourn times $1/\nu_i$ for all $i \in \mathcal{X}$; and (c) the initial state information.

Remark 23. *The Poisson process is a CTMC with state space $\mathcal{X} = \{0, 1, 2, \ldots\}$, transition probabilities $P_{i,i+1} = 1$ for all $i \in \mathcal{X}$, sojourn times having mean $1/\lambda$ for all $i \in \mathcal{X}$, and initial state $X(0) = 0$.*

5.2 Special Case: Birth-Death Processes

Definition 5.6 (Birth-Death Process)**.** A birth-death process is a CTMC with state space $\mathcal{X} = \{0, 1, 2, \ldots\}$ for which transitions from state i may go only to state $i+1$ (birth) or $i-1$ (death).

Result 5.7. For a birth-death process currently in state $i \in \{0, 1, 2, \ldots\}$, let $T_i^B \sim$ exponential(λ_i) be the time until the next birth, and if $i \in \{1, 2, \ldots\}$, let $T_i^D \sim$ exponential(μ_i) be the time until the next death; if $i = 0$, define $\mu_0 \coloneqq 0$ and $T_0^D \coloneqq \infty$. Then the sojourn times $\{T_i, i = 0, 1, 2, \ldots\}$ satisfy

$$T_i = \min\left(T_i^B, T_i^D\right) \sim \text{ exponential}(\nu_i = \lambda_i + \mu_i),$$

where ν_i can be interpreted as the rate of leaving state i, and the parameters $\{\lambda_i, i = 0, 1, 2, \ldots\}$ and $\{\mu_i, i = 1, 2, \ldots\}$ are the birth and death rates, respectively.

Result 5.8. For a birth-death process, $P_{0,1} = 1$, and for states $i \in \{1, 2, \ldots\}$, we have

$$P_{i,i+1} = P\{T_i^B < T_i^D\} = \frac{\lambda_i}{\lambda_i + \mu_i} \quad \text{and} \quad P_{i,i-1} = P\{T_i^D < T_i^B\} = \frac{\mu_i}{\lambda_i + \mu_i}.$$

Remark 24. *Birth-death processes can be viewed as a random walk on the non-negative integers in continuous time with $P_{01} = 1$ and an exponential "clock" governing transitions.*

Remark 25. *We have already seen that a PP(λ) is a CTMC; notice that it is also a birth-death process in which for all $i \in \{0, 1, 2, \ldots\}$, we have $\mu_i = 0$ (deaths never occur), $\lambda_i = \lambda$ (the birth rate is constant), $\nu_i = \lambda$, $P_{i,i+1} = 1$, and $P_{i,i-1} = 0$.*

Example 5.9 (Yule Process)**.** The Yule process is an example of a continuous time branching process on the state space $\mathcal{X} = \{1, 2, \ldots\}$. Consider a population in which each member acts independently and takes time $T \sim$ exponential(λ) to give birth, and no member of the population can die. Then if $X(t)$ is the size of the population at time t, then $\{X(t), t \geq 0\}$ is a pure birth process where, whenever the population has i members, the time until the next arrival is $T_i \sim$ exponential($i\lambda$) and $\mathrm{E}[T_i] = 1/(i\lambda)$.

5.2.1 Queueing Systems as Birth-Death Processes

Definition 5.10 (Kendall's Notation)**.** Queueing systems are described by the notation $A/S/c/K$ where

- A is the interarrival time distribution (common distribution notation inculdes 'M' for a memoryless or Markovian distribution, or 'G' for a General distribution);
- S is the service time distribution, in the same notation as the interarrival time distribution;
- c is the number of servers;
- K is the capacity of the system, which includes the number of customers in service and the number in queue. If not specified, the capacity of the system is infinite.

Other quantities may also be specified, such as the size of jobs to be processed or the queueing discipline; if not specified, the population of jobs is infinite and the queueing discipline is first in, first out.

Example 5.11 (M/M/1 Queue)**.** An M/M/1 queue has exponential (memoryless) interarrival times, exponential (memoryless) service times, one server, and infinite capacity with an infinite population of arrivals and first in, first out service. Thus arrivals occur as a PP(λ), and departures (or service endings) occur as a PP(μ). If $X(t)$ is the number in the M/M/1 system at time t, $\{X(t), t \geq 0\}$ is a birth-death process with

$$\lambda_i = \lambda \text{ for all } i \in \{0, 1, 2, \ldots\}; \quad \mu_i = \mu \text{ for all } i \in \{1, 2, \ldots\}.$$

Hence $T_i \sim \exp(\nu_i = \lambda + \mu)$ for all $i \in \{1, 2, \ldots\}$, and $P_{0,1} = 1$, $\nu_0 = \lambda$,

$$P_{i,i+1} = \frac{\lambda}{\lambda + \mu},\ i \in \{0, 1, 2 \ldots\}, \quad P_{i,i-1} = \frac{\mu}{\lambda + \mu},\ i \in \{1, 2, \ldots\}.$$

Example 5.12 (M/M/s Queue)**.** Now consider a queue identical to the M/M/1 queue, except with s servers, each serving at a rate μ. Note that there is still just one queue for all s servers (e.g., an airline counter with s agents and one line). If $X(t)$ is the total number of individuals in the M/M/s system at time t, $\{X(t), t \geq 0\}$ is a birth/death process with birth (arrival) rate and death (departure) rate

$$\lambda_i = \lambda \text{ for all } i \in \{0, 1, 2 \ldots\} \quad \text{and} \quad \mu_i = \begin{cases} i\mu, & i \in \{1, 2, \ldots, s\} \\ s\mu, & i \in \{s+1, s+2, \ldots\} \end{cases}$$

respectively. For brevity, let $1 \leq i \leq s$ denote $i \in \{1, 2, \ldots, s\}$ and $i > s$ denote $i \in \{s+1, s+2, \ldots\}$. Then $T_i \sim$ exponential(ν_i) where

$$\nu_i = \begin{cases} \lambda + i\mu, & 1 \leq i \leq s \\ \lambda + s\mu, & i > s \end{cases}$$

and letting $P_{0,1} = 1$ and $\nu_0 = \lambda$, we have

$$P_{i,i+1} = \begin{cases} \dfrac{\lambda}{\lambda + i\mu}, & 1 \leq i \leq s \\ \dfrac{\lambda}{\lambda + s\mu}, & i > s \end{cases} \qquad P_{i,i-1} = \begin{cases} \dfrac{i\mu}{\lambda + i\mu}, & 1 \leq i \leq s \\ \dfrac{s\mu}{\lambda + s\mu}, & i > s, \end{cases}$$

5.2.2 Calculating $P_{ij}(t)$ for a Pure Birth Process

Result 5.13. Consider a pure birth process with constant birth rate $\lambda_i = \lambda$ for all $i \geq 0$ and interarrival times $T_i \sim$ iid exponential(λ); i.e., we consider a Poisson process.[1] Then

$$P_{ij}(t) = P\{X(t) < j+1 \,|\, X(0) = i\} - P\{X(t) < j \,|\, X(0) = i\} = P\{N(t) = j - i\}.$$

5.3 The Generator Matrix

Definition 5.14 (Instantaneous Transition Rates)**.** Recall that for a CTMC, ν_i is the rate at which the process is moving out of state $i \in \mathcal{X}$ given that it is currently in state i, and P_{ij} is the probability that the CTMC currently in state i next transitions to state $j \in \mathcal{X}$. For any pair of states $i, j \in \mathcal{X}$, let q_{ij} be the instantaneous rate at which the process transitions to state j, given that it is currently in state i, where

$$q_{ij} = \nu_i P_{ij}.$$

Remark 26. *Since $P_{ii} = 0$ for all states $i \in \mathcal{X}$, it follows that $q_{ii} = 0$ for all $i \in \mathcal{X}$.*

Result 5.15. For states $i, j \in \mathcal{X}$ of a CTMC, we have $\nu_i = \sum_j q_{ij}$ and

$$P_{ij} = \frac{q_{ij}}{\nu_i} = \frac{q_{ij}}{\sum_j q_{ij}}.$$

Definition 5.16 (Generator Matrix)**.** The instantaneous transition rates are written in a matrix, called the generator matrix $\mathbb{Q}$, where

$$\mathbb{Q} \coloneqq \begin{pmatrix} -\nu_0 & q_{01} & \cdots & q_{0j} & \cdots \\ q_{10} & -\nu_1 & \cdots & q_{1j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ q_{j0} & q_{j1} & \cdots & -\nu_j & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{pmatrix}.$$

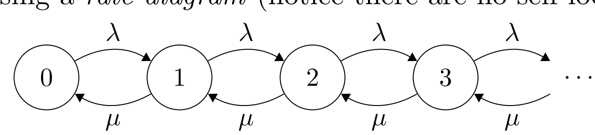
Remark 27. *Notice that the rows of the generator matrix sum to zero. For example, the process leaves state 0 at rate ν_0, and it enters other states at rate $\sum_j q_{0j} = \nu_0$.*

Result 5.17. A CTMC is completely defined by the generator matrix $\mathbb{Q}$ and the initial state distribution.

Example 5.18 (M/M/1 Queue)**.** Recall the M/M/1 queue from Example 5.11: arrivals are PP(λ), departures are PP(μ), there is one server, and $X(t)$ is the number in the M/M/1 system at time t. Then

$$\mathbb{Q} = \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ \mu & -(\lambda+\mu) & \lambda & 0 & 0 & \cdots \\ 0 & \mu & -(\lambda+\mu) & \lambda & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

We can visualize this CTMC using a *rate diagram* (notice there are no self-loops):



To visualize the rate diagram, consider a particle that stays on a node i for an exponential(ν_i) amount of time. Then, it moves to one of the other nodes with probability proportional to the rate on the arc, where here, $P_{01} = \lambda/\lambda = 1$, and for all states $i \in \{1, 2, \ldots\}$, we have $\nu_i = \lambda + \mu$, $P_{i,i-1} = \mu/(\lambda+\mu)$, and $P_{i,i+1} = \lambda/(\lambda+\mu)$.

[1] For a general treatment of $P_{ij}(t)$ for pure birth processes, see p. 382 of *Introduction to Probability Models* (Ross, 2010).

5.4 Calculating $P_{ij}(t)$

Remark 28. *Usually, we are unable able to write an expression for $P_{ij}(t)$ directly. Instead, we derive a set of differential equations that the $P_{ij}(t)$'s satisfy, where we take the derivative of $P_{ij}(t)$ with respect to time t. Recall that by the definition of a derivative,*

$$P'_{ij}(t) = \lim_{h \to 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h}.$$

Result 5.19. For a CTMC, the instantaneous rate of moving from state i to state j, where $i \neq j$, is

$$\lim_{h \to 0} \frac{P_{ij}(h)}{h} = q_{ij}.$$

Result 5.20. For a CTMC, notice that $1 - P_{ii}(h)$ is the probability of *not* being in state i after a time interval of length h. Then the instantaneous rate of moving out of state i is

$$\lim_{h \to 0} \frac{1 - P_{ii}(h)}{h} = \nu_i.$$

5.4.1 Chapman-Kolmogorov Equations

Result 5.21 (Chapman-Kolmogorov Equations)**.** Without loss of generality, label the states of the CTMC with indices $i, j, k \geq 0$. The Chapman-Kolmogorov equations are

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s) \quad \text{for all } s, t \geq 0 \text{ and all states } i, j.$$

5.4.2 Kolmogorov's Forward and Backward Equations

Theorem 5.22 (Kolmogorov's Backward Equations (KBE's))**.** Without loss of generality, label the states of the CTMC with indices $i, j, k \geq 0$. For all states i, j and times $t \geq 0$,

$$P'_{ij}(t) = \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t),$$

where the $\nu_i P_{ij}(t)$ term can be considered the "base rate" from the definition of the derivative, and the $\sum_{k \neq i} q_{ik} P_{kj}(t)$ term can be considered the amount of change in addition to the base rate.

Theorem 5.23 (Kolmogorov's Forward Equations (KFE's))**.** Without loss of generality, label the states of the CTMC with indices $i, j, k \geq 0$. For all states i, j and times $t \geq 0$,

$$P'_{ij}(t) = \sum_{k \neq j} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t).$$

Remark 29. *The KBE's and KFE's each imply that we have a system of differential equations to solve: one equation for every i, j state pair. When deciding whether to use the KBE's or KFE's, we should use whichever results in an easier system to solve. Since $P_{ij}(t)$ is a function of time, the result of solving KBE's or KFE's numerically is a sequence of matrices corresponding to each t used.*

Remark 30. [1] *To solve the KBE's or KFE's by hand, there are two common types of differential equations:*

1. *$P'_{ij}(t) = aP_{ij}(t)$ for constant a, which implies $P_{ij}(t) = e^{at+c}$ for some constant c;*
2. *$P'_{ij}(t) + aP_{ij}(t) = g(t)$ for constant a and function g, which implies $P_{ij}(t) = e^{-at}\left(\int e^{at} g(t) dt + c\right)$ for some constant c.*

We find the constant c using the conditions of the problem.

[1] You should be able to prove this result and justify all the steps.

Result 5.24. Let $\tilde{q}_{ij}$ denote the (i,j)th entry of the $\mathbb{Q}$ matrix. Then the KBE's and KFE's become

$$P'_{ij}(t) = \textstyle\sum_k \tilde{q}_{ik} P_{kj}(t) \quad \text{and} \quad P'_{ij}(t) = \sum_k \tilde{q}_{kj} P_{ik}(t) \quad \text{for all states } i,j \text{ and times } t \geq 0,$$

respectively. Let $\mathbb{P}(t)$ be a matrix whose (i,j)th element is $P_{ij}(t)$ and $\mathbb{P}'(t)$ be a matrix whose (i,j)th element is $P'_{ij}(t)$. Then we write these equations in matrix notation as

$$\text{KBE: } \mathbb{P}'(t) = \mathbb{Q}\mathbb{P}(t) \text{ for all } t \geq 0, \qquad \text{KFE: } \mathbb{P}'(t) = \mathbb{P}(t)\mathbb{Q} \text{ for all } t \geq 0.$$

Remark 31. (See also [3, §6.8].) *Solving the differential equations yields* $\mathbb{P}(t) = e^{\mathbb{Q}t} := \sum_{n=0}^{\infty} \mathbb{Q}^n (t^n/n)$,*which is not especially stable on a computer. (Since* $\mathbb{Q}$ *has positive and negative elements, the roundoff errors may be undesirably large.) One approximation method (there are others) is to use the identity*

$$e^{\mathbb{Q}t} = \lim_{n\to\infty} \left(\mathbb{I} + \mathbb{Q}\frac{t}{n}\right)^n$$

to approximate $\mathbb{P}(t)$ *when* n *is very large, since* n *large enough, all elements of* $(\mathbb{I} + \mathbb{Q}\frac{t}{n})$ *are positive.*

5.4.3 Uniformization

Remark 32. *If* $\nu_i = \nu$ *for all states* i*, we can view a CTMC as a DTMC in which the transition times occur according to a PP*(ν)*.*

Result 5.25. [1]Let $\{X(t), t \geq 0\}$ be a CTMC for which the rate of leaving state i is the same for all states i. That is, $\nu_i = \nu$ for all $i \in \mathcal{X}$. Let $N(t)$ be the number of state transitions by time t, and let $P_{ij}^{(n)}$ be the n-stage transition probability of the embedded DTMC having one-step transition probabilities P_{ij}. Then $\{N(t), t \geq 0\}$ is a PP(ν), and

$$P_{ij}(t) = \sum_{n=0}^{\infty} P_{ij}^{(n)} e^{-\nu t} \frac{(\nu t)^n}{n!}$$

Remark 33. *We can use this result to approximate* $P_{ij}(t)$ *for a general CTMC by introducing fictitious transitions from a state to itself.*

Result 5.26. Consider a CTMC for which $\nu_i \leq \nu$ for all states i, and define a new embedded DTMC with

$$P^*_{ij} = \begin{cases} 1 - \nu_i/\nu, & j = i \\ (\nu_i/\nu)P_{ij}, & j \neq i. \end{cases}$$

Thus, all transitions occur according to a (faster) PP(ν), but some transitions are fake self-transitions. Then

$$P_{ij}(t) = \sum_{n=0}^{\infty} P_{ij}^{*(n)} e^{-\nu t} \frac{(\nu t)^n}{n!}.$$

Remark 34. *With uniformization, the transitions out of state* i *are real transitions to a new state with probability* ν_i/ν *and fake self-transitions with probability* $1 - \nu_i/\nu$*. Thus in each state* i*, real transitions occur according to a thinned PP*$\big((\nu_i/\nu)\nu\big)$*.*

5.5 Limiting Probabilities

Definition 5.27. Without loss of generality, albel the states of the CTMC with indices $i, j \geq 0$. Assuming it exists, for each $j \geq 0$, let

$$P_j := \lim_{t\to\infty} P_{ij}(t),$$

denote the limiting probability that the CTMC will be in state j.

[1]You should be able to prove this result and justify all the steps.

Result 5.28. Consider a CTMC with the following properties: (a) all states communicate; that is, starting in state i, there is positive probability of ever being in state j for i, j; and (b) the chain is positive recurrent, that is, starting in state i, the mean time to return to state i is finite. Then P_j exists and is independent of the initial state i. Furthermore, P_j can be found by solving the system of equations

$$\nu_j P_j = \sum_{k \neq j} q_{kj} P_k, \quad j \geq 0 \quad (\text{“balance” equations}); \qquad 1 = \sum_{j=0}^{\infty} P_j \quad (\text{“normalizing” equation})$$

Result 5.29. If P_j exists for a CTMC, then it is also the long-run portion of time the process is in state j.

Remark 35. *A closer look at the balance equations, which state that the unconditional rate of leaving a state j must equal the unconditional rate of entering state j:*

$$\underbrace{\overbrace{\nu_j}^{\substack{\textit{rate leaving } j, \\ \textit{starting in } j}} \overbrace{P_j}^{\substack{\textit{proportion} \\ \textit{of time in } j}}}_{\textit{unconditional rate leaving } j} = \underbrace{\sum_{k \neq j} \underbrace{\overbrace{q_{kj}}^{\substack{\textit{rate going} \\ \textit{from } k \textit{ to } j}} \overbrace{P_k}^{\substack{\textit{proportion} \\ \textit{of time in } k}}}_{\textit{unconditional rate moving } k \textit{ to } j}}_{\textit{unconditional rate entering } j}$$

Result 5.30. If P_j exists for a CTMC, then it is also the stationary distribution for the CTMC. That is, if we choose the initial distribution to be $\mathbf{a}_0 = (P_0, P_1, P_2, \ldots, P_j, \ldots)$, then the probability of being in state j at time t is P_j for all t.

Result 5.31. Consider a birth-death process with birth rates $\{\lambda_i, i = 0, 1, 2, \ldots\}$ and death rates $\{\mu_i, i = 0, 1, 2, \ldots\}$. Then a necessary and sufficient condition for the limiting probabilities to exist is that

$$\sum_{i=1}^{\infty} \frac{\lambda_{i-1}\lambda_{i-2}\cdots\lambda_1\lambda_0}{\mu_i\mu_{i-1}\cdots\mu_2\mu_1} < \infty. \tag{C.BD}$$

[1]Further, if the condition in line (C.BD) above is satisfied, then

$$P_i = \frac{\lambda_{i-1}\lambda_{i-2}\cdots\lambda_1\lambda_0}{\mu_i\mu_{i-1}\cdots\mu_2\mu_1\left(1 + \sum_{i=1}^{\infty} \frac{\lambda_{i-1}\lambda_{i-2}\cdots\lambda_1\lambda_0}{\mu_i\mu_{i-1}\cdots\mu_2\mu_1}\right)}, \quad i \geq 1.$$

Example 5.32 (M/M/1 Queue)**.** For an M/M/1 queue, apply Result 5.31 to see that the arrival rate λ must be less than the service rate μ. Otherwise, if $\lambda > \mu$, the number in the system tends to infinity.

6 Renewal Processes

Definition 6.1 (Renewal Process)**.** Let $\{N(t), t \geq 0\}$ be a counting process and let X_n be the time between the $(n-1)$st and the nth event of this process, $n \geq 1$. If $\{X_n, n = 1, 2, \ldots\}$ is a sequence of nonnegative iid random variables with $\mathrm{E}[X_n] = \mu < \infty$ for all $n = 1, 2, \ldots$, then $\{N(t), t \geq 0\}$ is a renewal process.

Notation. Let $S_0 = 0$, $S_n = \sum_{i=1}^{n} X_i$ denote the arrival times.

Remark 36. *We can think of a renewal process as a generalization of a Poisson process in which the inter-arrival times are no longer iid exponentials; instead, they are iid with some other continuous distribution. Due to the memoryless property of the exponential, the Markov property holds at every time t in a Poisson process. Here, the Markov property holds only at the arrival times, or “renewal” times.*

Remark 37. *Renewal processes are a special type of generalized Markov model. Generalized Markov models are continuous time stochastic processes in which the Markov property holds at a set of (possibly random) time points $0 = S_0 \leq S_1 \leq S_2 \leq \ldots$.*

[1]You should be able to prove this result and justify all the steps.

6.1 Limit Theorems

Result 6.2. By the strong law of large numbers, in the long run, the expected time between renewals is μ. That is, as $n \to \infty$,

$$\frac{S_n}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n} \to \mu \quad \text{w.p.1.}$$

Theorem 6.3 (Long-Run Renewal Rate)**.** Let $\{N(t), t \geq 0\}$ be a renewal process generated by a sequence of nonnegative iid random variables $\{X_n, n = 1, 2, \ldots\}$ with common mean $0 < \mu < \infty$. Then

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mu} \quad \text{w.p.1,}$$

where $1/\mu$ is the *rate* of the renewal process.

Theorem 6.4 (Elementary Renewal Theorem)**.** Let $\{N(t), t \geq 0\}$ be a renewal process generated by a sequence of nonnegative iid random variables $\{X_n, n = 1, 2, \ldots\}$ with common mean $0 < \mu < \infty$. Then

$$\lim_{t\to\infty} \frac{\mathrm{E}[N(t)]}{t} = \frac{1}{\mu} \quad \text{w.p.1}$$

Remark 38. *The Elementary Renewal Theorem is not a direct result of Theorem 6.3.*

6.2 Applications to Queueing Theory

Example 6.5 (GI/GI/1 Queue)**.** Consider a single-server queue where GI stands for "general independent." That is, arrivals occur according to a renewal process with inter-arrival times $\{X_n, n = 1, 2, \ldots\}$ that are iid with distribution function F having mean $1/\lambda$. Then the long-run arrival rate is

$$\lim_{t\to\infty} \frac{N(t)}{t} \to \frac{1}{1/\lambda} = \lambda \quad \text{w.p.1.}$$

Further, the service times are iid random variables having distribution G and mean $1/\mu$, so that the long-run service rate is μ. The queue is stable if $\lambda < \mu$.

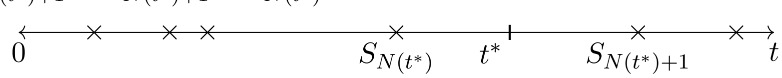
Theorem 6.6 (Little's Law)**.** Consider a queueing system and define the following quantities. Let $X(s)$ be the number of customers in the system at time s, and define L as the expected number of customers in the system in steady state. Let W_m be the amount of time the mth arriving customer spends in the system, and define W as the expected time spent in the system by a customer in steady state. Let $N_a(t)$ be the number of customers who arrive and enter the system by time t, and define λ as the long-run average rate at which arriving customers join the system. Formally,

$$L = \lim_{t\to\infty} \frac{1}{t}\int_0^t X(s)ds, \quad W = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} W_m, \quad \lambda = \lim_{t\to\infty} \frac{N_a(t)}{t}.$$

If the quantities L, W, and λ exist and are finite, then they satisfy $L = \lambda W$.

6.3 The Inspection Paradox

Remark 39. *When thinking about the inspection paradox, it helps to keep the following scenario in mind. Suppose you have a box containing an infinite number of lightbulbs, each of whose lifetimes is an iid copy of a random variable X with cdf F. As soon as a lightbulb goes out, it is instantaneously replaced with a new one from the box. Thus the number of lightbulbs that have failed by time t, $\{N(t), t \geq 0\}$, is a renewal process. Now suppose you fix a time t^*, and "inspect" the lifetime of the lightbulb currently in use at that time, which is $X_{N(t^*)+1} = S_{N(t^*)+1} - S_{N(t^*)}$:*



Since F is the lifetime distribution for all lightbulbs, it seems that this should be the lifetime distribution for the one in operation at t^*. However, this turns out not to be the case: as we see below, the lightbulb in use at time t^* tends to have a longer lifetime than an ordinary lightbulb.

Result 6.7. Let $\{N(t), t \geq 0\}$ be a renewal process with iid interarrival times $\{X_n, n = 1, 2, \ldots\}$ having cdf F, so that for ordinary renewal intervals,

$$P\{X_n > x\} = 1 - F(x).$$

Further, fix an inspection time t^*, and consider the length of the renewal interval that contains t^*, which is $X_{N(t^*)+1} = S_{N(t^*)+1} - S_{N(t^*)}$. Then

$$P\{X_{N(t^*)+1} > x\} \geq 1 - F(x).$$

Remark 40. *The intuition behind this result is that, if we think of line being covered with renewal intervals, it is more likely that a larger interval, rather than a shorter interval, covers the fixed time point t^*.*

Example 6.8. Suppose the lightbulb lifetime distributions are iid exponential(λ) random variables, so that $\{N(t), t \geq 0\}$ is a PP(λ). Then we can directly show that while $\mathrm{E}[X_n] = 1/\lambda$ for any ordinary renewal interval, we have $\mathrm{E}[X_{N(t^*)+1}] = 1/\lambda + (1/\lambda)(1 - e^{-\lambda t^*})$ which, for large t^*, is approximately $2/\lambda$.

7 Brownian Motion

Definition 7.1 (Brownian Motion)**.** A stochastic process $\{X(t), t \geq 0\}$ is called a Brownian motion process if $X(0) = 0$ and

(i) $\{X(t), t \geq 0\}$ has stationary and independent increments;
(ii) for every $t > 0$, $X(t) \sim N(0, \sigma^2 t)$, that is, $X(t)$ is normally distributed with mean 0 and variance $\sigma^2 t$.

Definition 7.2 (Standard Brownian Motion)**.** If a stochastic process $\{B(t), t \geq 0\}$ is a Brownian motion process with $\sigma^2 = 1$, then $\{B(t), t \geq 0\}$ is called a standard Brownian motion process.

Remark 41. *A Brownian motion process can be converted in to a standard Brownian motion process by letting $B(t) = X(t)/\sigma$ for all $t \geq 0$.*

Definition 7.3 (Brownian Motion with Drift, Definition I)**.** A stochastic process $\{X(t), t \geq 0\}$ is a Brownian motion process with drift coefficient μ and variance parameter σ^2 if $X(0) = 0$ and

(i) $\{X(t), t \geq 0\}$ has stationary and independent increments;
(ii) for every $t > 0$, $X(t) \sim N(\mu t, \sigma^2 t)$, that is, $X(t)$ is normally distributed with mean μt and variance $\sigma^2 t$.

Definition 7.4 (Brownian Motion with Drift, Definition II)**.** Let $\{B(t), t \geq 0\}$ be a standard Brownian motion process. Then $X(t) = \sigma B(t) + \mu t$ is a Brownian motion process with drift coefficient μ and variance parameter σ^2.

References

[1] V. G. Kulkarni. *Introduction to Modeling and Analysis of Stochastic Systems.* Springer, 2nd edition, 2011.

[2] S. I. Resnick. *Adventures in Stochastic Processes.* Birkhauser, Boston, 2005.

[3] S. M. Ross. *Introduction to Probability Models.* Academic Press, 10th edition, 2010.

[4] H. M. Taylor and S. Karlin. *An Introduction to Stochastic Modeling.* Academic Press, 3 edition, 1998.