E2.204: Stochastic Processes and Queuing Theory
Spring 2019

Course Instructor: Prof. Vinod Sharma (vinod@iisc.ac.in)
Teaching Assistants: Panju and Ajay
Indian Institute of Science, Bangalore
## Contents

1 **Poisson Processes** 4  
1.1 Introduction to stochastic processes 4  
1.2 Poisson process 5  
1.2.1 Definition 5  
1.2.1 (Contd.) Poisson Processes: Definition 7  
1.2.2 Properties of Poisson processes 10  
1.2.3 Generalization of Poisson processes 11  
1.3 Problems 13  

2 **Renewal Theory and Regenerative Processes** 14  
2.1 Renewal Process: Introduction 14  
2.2 Limit Theorems 16  
2.2.2 Blackwell's Theorem 18  
2.2.3 Renewal Equation 19  
2.2.4 Renewal theorems 21  
2.3 Regenerative Processes 25  
2.4 Problems 28  

3 **Discrete Time Markov Chains** 30  
3.1 Markov Chains: Definitions 30  
3.2 Class Properties of Transience and Recurrence 31  
3.3 Limiting distributions of Markov chains 32  
3.2 (Contd.) Limiting distributions of Markov chains 33  
3.4 Tests for transience, null recurrence and positive recurrence 35  
3.5 Reversible Markov Chains 35  
3.6 Example: M/GI/1 queue 37  
3.7 Rate of convergence to the stationary distribution 37  
3.8 Problems 40  

4 **Continuous-Time Markov Chains** 41  
4.1 Introduction 41  
4.2 Strong Markov property, Minimal construction 42  
4.3 Chapman Kolmogorov equations 44  
4.4 Irreducibility and Recurrence 46  
4.5 Time Reversibility 49  
4.6 Birth-Death process 50  
4.6.1 Reversibility of Birth-Death process 51  
4.6.2 Examples 51
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.7</td>
<td>Problems</td>
<td>53</td>
</tr>
<tr>
<td>5</td>
<td>Martingales</td>
<td>55</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>55</td>
</tr>
<tr>
<td>5.2</td>
<td>Optional Sampling Theorem</td>
<td>57</td>
</tr>
<tr>
<td>5.2 (Contd.)</td>
<td>Optional Sampling Theorem: Example</td>
<td>59</td>
</tr>
<tr>
<td>5.3</td>
<td>Martingale inequalities</td>
<td>59</td>
</tr>
<tr>
<td>5.4</td>
<td>McDiarmid’s Inequality: Applications</td>
<td>62</td>
</tr>
<tr>
<td>5.5</td>
<td>Martingale Convergence Theorem</td>
<td>63</td>
</tr>
<tr>
<td>5.6</td>
<td>Applications to Markov chain</td>
<td>65</td>
</tr>
<tr>
<td>5.7</td>
<td>Problems</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>Random Walks</td>
<td>70</td>
</tr>
<tr>
<td>6.1</td>
<td>Definitions</td>
<td>70</td>
</tr>
<tr>
<td>6.2</td>
<td>Ladder Heights, Maxima, GI/GI/1 Queue</td>
<td>71</td>
</tr>
<tr>
<td>6.2 (Contd.)</td>
<td>Ladder Epochs</td>
<td>73</td>
</tr>
<tr>
<td>7</td>
<td>Queuing Theory</td>
<td>75</td>
</tr>
<tr>
<td>7.1</td>
<td>GI/GI/1 Queue</td>
<td>75</td>
</tr>
<tr>
<td>7.2</td>
<td>Palm Theory, PASTA</td>
<td>79</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Rate conservation laws</td>
<td>82</td>
</tr>
<tr>
<td>7.2.2</td>
<td>PASTA</td>
<td>82</td>
</tr>
<tr>
<td>7.3</td>
<td>Product-form Networks</td>
<td>83</td>
</tr>
<tr>
<td>7.3.1</td>
<td>M/M/1 queue</td>
<td>83</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Tandem Queues</td>
<td>83</td>
</tr>
<tr>
<td>7.3.3</td>
<td>Open Jackson Networks</td>
<td>84</td>
</tr>
<tr>
<td>7.3.4</td>
<td>Closed queueing networks</td>
<td>85</td>
</tr>
<tr>
<td>7.4</td>
<td>Product-Form Networks: Quasireversible networks</td>
<td>86</td>
</tr>
<tr>
<td>7.4.1</td>
<td>Quasireversible Queues</td>
<td>86</td>
</tr>
<tr>
<td>7.4.2</td>
<td>Networks of Quasireversible Queues</td>
<td>88</td>
</tr>
<tr>
<td>7.4.2 (Contd.)</td>
<td>Networks of quasireversible queues (contd.)</td>
<td>90</td>
</tr>
<tr>
<td>7.5</td>
<td>Problems</td>
<td>93</td>
</tr>
</tbody>
</table>
Course structure

- Two sessions per week. Tuesdays and Thursdays between 5:15 p.m. to 6:45 p.m.
- One assignment per topic. There will be a quiz based on each assignment. Tutorial and quiz sessions will be held by TAs on Saturdays 10-11 a.m. after each topic is finished.
- One mid-term exam.
- One final exam.

Reference Books

Chapter 1

Poisson Processes

Lecture 1

Course E2.204: Stochastic Processes and Queueuing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

1.1 Introduction to stochastic processes

Review: Let $(\Omega, \sigma, P)$ be a probability space. A measurable mapping $X : \Omega \to \mathbb{R}$ is called a random variable (r.v.). $X(\omega)$ for $\omega \in \Omega$ is called a realization of $X$. $F_X(x) = P[X \leq x]$ is called the distribution function of r.v. $X$. $f_X(x) = dF_X(x)/dx$ is called the probability density function of $X$. The probability density function may not always exist. $E[X] = \int xdF_X(x)$ is the expectation of $X$. When probability density of $X$ exists $E[X] = \int xf(x)dx$.

**Stochastic processes:** $\{X_t : t \in \mathbb{R}\}$, where $X_t$ is a r.v. is called a continuous time stochastic process. $\{X_n : n \in \mathbb{N}\}$, where $X_n$ is a r.v. is called a discrete time stochastic process.

The function $t \mapsto X_t(\omega)$ is called a sample path of the stochastic process. For each $\omega \in \Omega$, $X_t(\omega)$ is a function of $t$. $F_t$ is the distribution of $X_t$. An analogous definition holds for discrete time stochastic processes. A stochastic process is described by the joint distribution of $(X_1, X_2, \ldots, X_n)$ for any $-\infty < t_1 < t_2 < \cdots < t_n$ and $n \in \mathbb{N}^+$.

A stochastic process $\{X_t\}$ is said to be stationary if for any $0 \leq t_1 < t_2 < \cdots < t_n$, the joint distribution of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ is identical to the joint distribution of $(X_{t_1+\tau}, X_{t_2+\tau}, \ldots, X_{t_1+\tau})$ for any $\tau \in \mathbb{R}$. A stochastic process $\{X_t\}$ is said to have independent increments if $(X_{t_2} - X_{t_1}), (X_{t_3} - X_{t_2}), \ldots, (X_{t_n} - X_{t_{n-1}})$ are independent. If joint distribution of $(X_{t_1+\tau} - X_{t_1+\tau}), (X_{t_2+\tau} - X_{t_{n-1}+\tau}), \ldots, (X_{t_1+\tau} - X_{t_2+\tau})$ does not depend on $\tau$, then $\{X_t\}$ is said to have stationary increments. If $\{X_t\}$ has both stationary and independent increments, it is called a stationary independent increment process.

**Point process:** A stochastic process $\{N_t, t \geq 0\}$ with $N_0 = 0, N_t$ a non-negative integer, non-decreasing with piece-wise constant sample paths is called a point process. $N_t$ counts the number of points or ‘arrivals’ in the interval $(0, t]$.

Let $A_n$ denote the interarrival time between $n^{th}$ and $(n-1)^{th}$ arrival. Let $S_0 = 0$ and $S_n = \sum_{k=1}^{n} A_k, \forall n \geq 1$. Then $S_n$ denotes the time instant of the $n^{th}$ arrival. $N_t = \max\{n : S_n \leq t\}$. A point process with at most one arrival at any time is called a simple point process. Mathematically, a simple point process $\{N_t\}$ is
described by following constraints for all $t$:

$$\Pr\{N_{t+h} - N_t \geq 2\} = o(h).$$

Here, the notation $o(g(x))$ means a class of functions such that if $f(x) \in o(g(x))$, then $\lim_{x \to 0} \frac{f(x)}{g(x)} = 0$.

### 1.2 Poisson process

#### 1.2.1 Definition

In the following we customarily take $N_0 = 0$. A point process $N_t$ is Poisson if any of the following conditions hold.

**Definition [1]:**

1. $\{A_k, k \geq 1\}$ are independent and exponentially distributed with parameter $\lambda$: $\Pr\{A_k \leq x\} = 1 - e^{-\lambda x}$.
   
   If $\lambda = 0$, $A_1 = \infty$ w.p.1 and $N_t = 0$ $\forall t$. If $\lambda = \infty$, $A_1 = 0$ w.p.1 and $N_t = \infty$ $\forall t$. Thus, we restrict to $0 < \lambda < \infty$. In this range for $\lambda$, $N_t$ is guaranteed to be simple because $\Pr\{A_k = 0\} = 0$ $\forall k$.

2. $N_t$ is simple.

3. $N_t$ has stationary independent increments.

**Definition [3]:**

1. $N_t$ has independent increment.

2. For $s < t$,

   $$\Pr\{N_t - N_s = n\} = \frac{(\lambda(t-s))^n}{n!} e^{-\lambda(t-s)}$$

**Definition [4]:**

1. $N_t$ has stationary and independent increments.

   (a) $\Pr\{N_{t+h} - N_t = 1\} = \lambda h + o(h)$

   (b) $\Pr\{N_{t+h} - N_t = 0\} = 1 - \lambda h + o(h)$

   (c) $\Pr\{N_{t+h} - N_t \geq 2\} = o(h)$

   We will show below that these definitions are equivalent. We need the following important characterization of exponential distribution.
Exponential r.v. is memoryless: Let \( X \) be an exponential r.v.

\[
\Pr\{X > t + s | X > t\} = \frac{\Pr\{X > t + s, X > t\}}{\Pr\{X > t\}} \\
= \frac{\Pr\{X > t + s\}}{\Pr\{X > t\}} \\
= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} \\
= e^{-\lambda s} \\
= \Pr\{X > s\}
\]

If \( X \) is interarrival time, this property of an exponential r.v. indicates that the remaining time till next arrival does not depend on time \( t \) since last arrival. Thus, the term memoryless is used.

**Theorem 1.2.1.** Exponential distribution is the unique distribution on \( \mathbb{R}^+ \) with the memoryless property.

**Proof.** If a r.v. \( X \) on \( \mathbb{R}^+ \) is memoryless, we show that \( X \) must be exponential. If \( X \) is memoryless, we have for all \( t, s \geq 0 \), \( \Pr\{X > t + s\} = \Pr\{X > t\} \Pr\{X > s\} \). Let \( f(t) = \Pr\{X > t\} \). We have the functional equation

\[
f(t+s) = f(t)f(s)
\]

Taking \( t = s \), we get \( f(2t) = f^2(t) \). By repeated application of Eq 1.1, \( m \) times, we get \( f(mt) = f^m(t) \) for positive integer \( m \). Equivalently, we have \( f(t/m) = f^{\frac{1}{m}}(t) \). Again by repeated application of Eq 1.1 \( n \) times, \( f\left(\frac{r}{m^n}t\right) = f^{\frac{r}{m^n}}(t) \) for any positive integers \( m \) and \( n \). So, we have \( f(rt) = f^r(t) \) for any positive rational number \( r \). We know that \( 0 \leq f(t) \leq 1 \) since \( 1 - f \) is probability distribution. So, we can write \( f(1) = e^{-\lambda} \) for some \( \lambda \geq 0 \). Therefore we have, \( f(r) = f(r \times 1) = f^r(1) = e^{-r\lambda} \) for any positive rational number \( r \).

For any \( x \in \mathbb{R} \), there is a sequence of rationals \( r_n \downarrow x \). Since \( f \) is right continuous, \( f(r_n) \to f(x) \). In other words, for any \( x \in \mathbb{R} \),

\[
f(x) = \lim_{r_n \to x} f(r_n) \\
= \lim_{r_n \to x} e^{-r_n \lambda} \\
= e^{-\lambda x}
\]

Thus, \( \Pr\{X > x\} = e^{-\lambda x} \) and \( X \) is an exponential random variable. \( \square \)

Now we show that definitions \([1-4]\) for Poisson process given above are equivalent.

**Proposition 1.2.2.** Definition \([3]\) \( \implies \) Definition \([2]\).

**Proof.** We need to show that \( N_t \) has stationary increments if \( N_t \) has independent increments and \( N_t - N_s \) is Poisson distributed with mean \( \lambda(t-s) \). Stationarity follows directly from the definition since the distribution of number of points in an interval depends only in the length of interval. The conditions for a simple process is also met which can be easily verified from the definition:

\[
\Pr\{N_{t+h} - N_t \geq 2\} = 1 - \left(\frac{\lambda h}{0!}e^{-\lambda h} + \frac{(\lambda h)^1}{1!}e^{-\lambda h}\right) \\
= o(h)
\]

\( \square \)
1.2.1 (Contd.) Poisson Processes: Definition

Proposition 1.2.3. Definition [1] of Poisson processes (see lecture-01) $\implies$ Definition [3].

Proof. Step 1: Density function of $n^{th}$ arrival $p_{S_n}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}$. This can be shown using induction on $n$. $S_n = \sum_{k=1}^n A_k$, where $A_k \sim \text{exp}(\lambda)$. For $n = 1$ and $S_1 = A_1$, the expression is true since it reduces to density function of an exponential random variable with mean $\lambda$. Now, assuming it is true for $S_n$, we will show that it is true for $S_{n+1}$.

$$p_{S_{n+1}}(t) = p_{S_n} * A_{n+1}$$

(density of sum of ind. r.v.s is their convolution)

$$= \int_0^t \left( \frac{\lambda^n \tau^{n-1}}{(n-1)!} e^{-\lambda \tau} \right) \left( \lambda e^{-\lambda (t-\tau)} \right) d\tau$$

$$= e^{-\lambda t} \lambda^{n+1} \int_0^t \frac{\tau^{n-1}}{(n-1)!} d\tau$$

$$= e^{-\lambda t} \lambda^{n+1} \frac{t^n}{n!}$$

Step 2:

$$\mathbb{P}\{N_t = n\} = \mathbb{P}\{S_n \leq t < S_{n+1}\}$$

$$= \int_0^t \mathbb{P}\{S_n \leq t | S_n = s\} p_{S_n}(s) ds$$

$$= \int_0^t \mathbb{P}\{A_{n+1} > t - s\} p_{S_n}(s) ds \quad (\{S_n = s, S_{n+1} > t\} \equiv \{A_{n+1} > t - s\})$$

$$= \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad (\text{Use the fact that } A_1 \text{ is exp}(\lambda)) \quad \square$$

Independent increments and stationary increments property follows from $A_k$ begin i.i.d. with $\text{exp}(\lambda)$.

Proposition 1.2.4. Definition [2] of Poisson processes (see lecture-01) $\implies$ Definition [1].

Proof. Step 1: Show that $A_1$ must be exponential.

$$\mathbb{P}\{A_1 > t+s|A_1 > t\} = \mathbb{P}\{N_{t+s} - N_t = 0|N_t = 0\}$$

$$= \mathbb{P}\{N_{t+s} - N_t = 0\} \quad (\text{increments are independent})$$

$$= \mathbb{P}\{N_t = 0\} \quad (\text{stationary increments})$$

$$= \mathbb{P}\{A_1 > s\}$$

$$\implies \mathbb{P}\{A_1 > t+s\} = \mathbb{P}\{A_1 > s\} \mathbb{P}\{A_1 > t\}$$

This leads to the functional equation whose only right continuous solution is that $A_1$ is an exponential r.v. (this was proved in lecture-01).
Step 2: Show that \( A_2 \) is independent of \( A_1 \) and has the same distribution.

\[
\begin{align*}
\mathbb{P}\{A_2 > t | A_1 = s\} &= \mathbb{P}\{N_{r+x} - N_r = 0 | N_t = 1\} \\
&= \mathbb{P}\{N_{r+x} - N_r = 0\} \quad \text{(independent increments)} \\
&= \mathbb{P}\{N_r = 0\} \quad \text{(stationary increments)} \\
&= \mathbb{P}\{A_1 > t\}
\end{align*}
\]

This shows that \( A_2 \) is independent of and also identically distributed as \( A_1 \). Similar argument holds for all other \( A_n \)'s.

\[\blacksquare\]


Proof. We have already shown that definition [3] implies stationary increments. We now show that it implies (2) in definition [4].

\[
\begin{align*}
\mathbb{P}\{N_{r+h} - N_r = 1\} &= \frac{\lambda he^{-\lambda h}}{1!} \\
&= \lambda h(1 - \lambda h + o(h)) \\
&= \lambda h + o(h)
\end{align*}
\]


\[
\begin{align*}
\mathbb{P}\{N_{r+h} - N_r \geq 2\} &= \sum_{k=2}^{\infty} \frac{(\lambda h)^2 e^{-\lambda h}}{k!} \\
&= o(h)
\end{align*}
\]

which shows [4] – (2) – (c). These together also proves [4] – (2) – (b). \[\blacksquare\]


Proof. Let \( f_n(t) = \mathbb{P}\{N_t = n\} \). We will first find \( f_0(t) \) by developing and solving a differential equation.

\[
\begin{align*}
f_0(t + h) &= \mathbb{P}\{N_t = 0, N_{r+h} - N_r = 0\} \\
&= \mathbb{P}\{N_t = 0\} \mathbb{P}\{N_{r+h} - N_r = 0\} \quad \text{(independent increments)} \\
&= f_0(t)(1 - \lambda h + o(h)) \quad \text{(using definition [4] – (2) – (b))} \\
\implies f_0'(t) &= -\lambda f_0(t) \quad \text{(rearranging and taking } h \downarrow 0) \\
f_0(t) &= e^{-\lambda t} \quad \text{(Solving diff equation using } N_0 = 0) \\
\end{align*}
\]

For \( n \geq 1 \), we have

\[
\begin{align*}
f_n(t + h) &= \mathbb{P}\{N_t = n, N_{r+h} - N_r = 0\} \\
&= \mathbb{P}\{N_t = n - 1, N_{r+h} - N_r = 1\} + o(h) \\
&= f_n(t)(1 - \lambda h) - f_{n-1}(t)\lambda h + o(h) \quad \text{(independent increments and definition [4] – (2) – (a,b))} \\
\implies f_n'(t) &= -\lambda f_n(t) - \lambda f_{n-1}(t) \quad \text{(taking } h \downarrow 0) \\
\implies \frac{d}{dt} \left( e^{\lambda t} f_n(t) \right) &= -\lambda e^{\lambda t} f_{n-1}(t) \quad \text{(multiplying by } e^{\lambda t} \text{ and rearranging)}
\end{align*}
\]
Solving the above equation for $n = 1$ using the initial condition $f_1(0) = 0$, we obtain

$$f_1(t) = \lambda t e^{-\lambda t}.$$ 

For general $n$, we can verify using induction on $n$ that

$$f_n(t) = \mathbb{P}\{N_t = n\} = \frac{(\lambda t)^n}{n!} e^{\lambda t}.$$

This verifies definition [3].
1.2.2 Properties of Poisson processes

Conditional distribution of points in an interval:

Theorem 1.2.7. Given that there are \( n \) points in the interval \((a, b]\), these \( n \) points are distributed uniformly in the interval \( I \). Their joint distribution is given by order statistics of \( n \) uniformly distributed point in the interval \( I \).

Proof outline: Take \( s_1 < s_2 < \cdots < s_n \) and \( h > 0 \) small enough such that \( s_1 + h < s_2, s_2 + h < s_3, \ldots, s_{n-1} + h < s_n \) in \((0, t]\).

\[
\mathbb{P}\{S_1 \in (s_1, s_1 + h], S_2 \in (s_2, s_2 + h], \cdots, S_n \in (s_n, s_n + h]|N_t = n\} = \mathbb{P}\{N_t = n\} = \frac{e^{-\lambda h} \cdot e^{-\lambda (s_2 - (s_1 + h))} \cdots e^{-\lambda (s_n - (s_{n-1} + h))}}{n!} \frac{(\lambda t)^n}{n!} e^{-\lambda t} + o(h^n) = \frac{n!}{t^n} h^n + o(h^n)
\]

Now, we have

\[
p_{S_1,S_2,\ldots,S_n}|N_t=n(s_1,s_2,\ldots,s_n) = \lim_{h \to 0} \frac{\mathbb{P}\{S_1 \in (s_1, s_1 + h], S_2 \in (s_2, s_2 + h], \cdots, S_n \in (s_n, s_n + h]|N_t = n\}}{h^n}
\]

where \( p_{S_1,S_2,\ldots,S_n}|N_t=n \) is the joint density of \( S_1, S_2, \ldots, S_n \) conditioned on \( N_t = n \). Therefore,

\[
p_{S_1,S_2,\ldots,S_n}|N_t=n(s_1,s_2,\ldots,s_n) = \frac{n!}{t^n}
\]

This is the density function of \( n \) ordered random variables uniformly distributed in the interval \((0, t]\). By stationarity, this property holds for any interval.

Superposition of independent Poisson processes:

Theorem 1.2.8. If \( n \) Poisson processes \( N^{(1)}_t, N^{(2)}_t, \ldots, N^{(n)}_t \) of rates \( \lambda_1, \lambda_2, \ldots, \lambda_n \) respectively are independent, then their superposition is a Poisson process of rate \( \sum_{k=1}^{n} \lambda_k \).

Proof. We use definition [3]. The superposition process is simple since component processes are simple (follows from union bound). Independent increments property also from the fact that component process are independent and have independent increments property. [3] – (2) follows from the fact that sum of independent Poisson random variables with mean \( \lambda_1 t, \lambda_2 t, \ldots, \lambda_n t \) is a Poisson random variable with mean \( \sum_{k=1}^{n} \lambda_k t \).
Splitting of Poisson processes:

**Theorem 1.2.9.** Let \( N_i \) be a Poisson process of the rate \( \lambda \). Suppose each point of the process \( N_i \) is marked independently as type \( i \) for \( i \in \{1, 2, \cdots, m\} \) with probability \( p_i \) such that \( \sum_{i=1}^{m} p_i = 1 \). Let the processes \( \{N_i^{(i)}\} \) for \( i \in \{1, 2, \cdots, m\} \) be comprised of only those points marked as type \( i \) respectively. Then, \( \{N_i^{(i)}\} \) are independent Poisson processes with respective rates \( p_i \lambda \).

**Proof.** We prove for the case of \( m = 2 \). The proof for the general case is similar. The processes \( N_i^{(1)} \) and \( N_i^{(2)} \) are simple and have independent increment property which follows from \( N_i \) being simple with independent increments. We also have

\[
P\{N_i^{(1)} = k_1, N_i^{(2)} = k_2\} = \sum_{k=0}^{\infty} P\{N_i^{(1)} = k_1, N_i^{(2)} = k_2|N_i = k\} P\{N_i = k\}
\]

\[
= P\{N_i^{(1)} = k_1, N_i^{(2)} = k_2|N_i = k_1 + k_2\} P\{N_i = k_1 + k_2\}
\]

\[
= \frac{(p_1 \lambda t)^k_1 e^{-p_1 \lambda t}}{k_1!} \times \frac{(p_2 \lambda t)^k_2 e^{-p_2 \lambda t}}{k_2!}.
\]

We can now appeal to definition [3] to conclude that \( N_i^{(1)} \) and \( N_i^{(2)} \) are independent with rates \( p_1 \lambda \) and \( p_2 \lambda \) respectively. \(\square\)

### 1.2.3 Generalization of Poisson processes

**Batch Poisson processes:**

Let \( \{N_i\} \) be a Poisson process of rate \( \lambda \). At each arrival, instead of just one customer, a batch of customers arrive. The number of customers at \( n^{th} \) arrival is \( X_n \). The sequence of \( X_n, \ n = 1, 2, \cdots \ ) is i.i.d and is also independent of \( \{N_i\} \). The overall process \( \{Y_i\} \) where \( Y_i \) is the total number of arrivals in \( (0, t] \) is called a batch Poisson process. \( Y_t = \sum_{k=1}^{N_t} X_k \). We can compute the distribution of \( Y_t \) as follows

\[
P\{Y_t = m\} = \sum_{n=0}^{\infty} P\{\sum_{k=1}^{N_t} X_k = m\} P\{N_t = n\}
\]

\[
= \sum_{n=0}^{\infty} \left( \frac{(\lambda t)^n e^{-\lambda t}}{n!} \right) P\{\sum_{k=1}^{N_t} X_k = m\}
\]

where \( P\{\sum_{k=1}^{N_t} X_k = m\} \) can be computed by convolution or from momentum generating functions. For the mean of \( Y_t \), we have \( E[Y_t] = (\lambda t) E[X_t] \). We note that batch Poisson process is a relaxation of requirement of simplicity in definition [2] of Poisson process.

**Nonstationary Poisson processes:**

The non-stationary (also called non-homogeneous) Poisson process \( \{N_t\} \) is defined as follows.

- \( \{N_t\} \) has independent increments.
- Let \( \lambda(t) \) be non negative function of \( t \).
  - \( P\{N_{t+h} - N_t = 1\} = \lambda(t) h \)
  - \( P\{N_{t+h} - N_t = 0\} = 1 - \lambda(t) h + o(h) \)
  - \( P\{N_{t+h} - N_t \geq 2\} = o(h) \)
Let \( m(t) = \int_0^t \lambda(s) \, ds \). We show that \( N_{t+s} - N_t \) is a Poisson random variable with mean \( m(t+s) - m(t) \).

Let \( f_n(s) = \mathbb{P}\{N_{t+s} - N_t = n\} \). For \( n = 0 \), we have for \( h \downarrow 0 \)

\[
    f_0(s+h) = \mathbb{P}\{N_{t+s+h} - N_t = 0, N_{t+s} - N_t = 0\} \\
    = \mathbb{P}\{N_{t+s+h} - N_{t+s} = 0\}\mathbb{P}\{N_{t+s} - N_t = 0\} \\ 
    = (1 - \lambda(t+s)h + o(h))f_0(s) \\ 
    \Rightarrow f_0'(s) = -\lambda(t+s)f_0(s) \\
    \Rightarrow f_0(s) = e^{-(m(t+s)-m(t))} 
\]

Using similar argument we can show by induction on \( n \) that

\[
    \mathbb{P}\{N_{t+s} - N_t = n\} = f_n(s) = \frac{(m(t+s)-m(t))^n e^{-(m(t+s)-m(t))}}{n!} 
\]

Spatial Poisson process:

So far we have defined Poisson processes on \( \mathbb{R}^+ \). Now, we generalize to \( \mathbb{R}^k, k \geq 2 \). Let \( A, B \subset \mathbb{R}^k \) and \( N_A \) is the number of points in \( A \).

1. \( N_A \) and \( N_B \) are independent for disjoint \( A \) and \( B \).

2. if \( |A| \) denotes the volume of \( A \),

\[
    \mathbb{P}\{N_A\} = \frac{(\lambda|A|)^n e^{-\lambda|A|}}{n!} 
\]

From this definition, it follows that for the process \( \{N\} \), \( \mathbb{P}\{N_{B_x(h)} \in \{0,1\}\} = 1 + o(h^k) \) for all \( x \in \mathbb{R}^k \) where \( B_x(h) \) is a ball of radius \( h \) around \( x \). This means that \( \{N\} \) is simple.
1.3 Problems

**Problem 1:** An item has a random lifetime with exponential distribution with parameter $\lambda$. When the item fails, it is immediately replaced by an identical item. Let $N_t$ be the number of failures till time $t$. Show that $\{N_t, t \geq 0\}$ is a Poisson process. Find the mean and variance of the total time $T$ when the fifth item fails.

**Problem 2:** Let $A_1, A_2, \ldots, A_n$ be disjoint intervals on $\mathbb{R}^+$ and $B = \bigcup_{k=1}^n A_k$. Let $a_1, a_2, \ldots, a_n$ be their respective lengths and $b = \sum_{i=1}^n a_i$. Then for $k = k_1 + k_2 + \ldots + k_n$, show for $N_t$ a Poisson process

$$ \mathbb{P}\{N_{A_1} = k_1, N_{A_2} = k_2, \ldots, N_{A_n} = k_n | N_B = k\} = \frac{k!}{k_1!k_2! \ldots k_n!} \left( \frac{a_1}{b}\right)^{k_1} \left( \frac{a_2}{b}\right)^{k_2} \ldots \left( \frac{a_n}{b}\right)^{k_n} $$

**Problem 3:** A department has three doors. Arrivals at each door form Poisson process with rates $\lambda_1 = 110$, $\lambda_2 = 90$ and $\lambda = 160$ customers/sec. 30% of the customers are male and 70% are female. The probability that a male customer buys a item is 0.6. The probability that a female buys an item is 0.1. An average purchase is worth Rs 4.50. Assume all the random variables are independent. What is the average worth of total sales in 10 hours? What is the probability of the third female who also buys an item arrives during the first 15 minutes? What is the expected time of her arrival?

**Problem 4:** The customers arrive at a facility as a Poisson process with rate $\lambda$. There is a waiting cost of $c$ per customer per unit time. The customers wait till they are dispatched. The dispatching takes place at times $T, 2T, \ldots$. At time $kT$ all customers in waiting will be dispatched. There is dispatching cost of $\beta$ per customer.

1. What is the total dispatch cost till time $t$.
2. What is the mean total customer waiting time till time $t$.
3. What value of $T$ minimizes the mean total customer and dispatch cost per unit time.

**Problem 5:** Let $N_t$ be a Poisson process with rate $\lambda$ and let the $n^{th}$ arrival epoch be $S_n$. Calculate $\mathbb{E}[S_3 | N_t = 3]$.

**Problem 6:** Let $N_1$ and $N_2$ be two Poisson processes with rates $\lambda_1$ and $\lambda_2$. Let the $n^{th}$ arrival epoch be $S_n^{(1)}$ and $S_n^{(2)}$ respectively. Calculate

1. $\mathbb{P}\{S_1^{(1)} < S_1^{(2)}\}$
2. $\mathbb{P}\{S_1^{(1)} < S_1^{(2)}\}$

**Problem 7:** Shocks occur to a system according to a Poisson process $N_t$ of intensity $\lambda$. Each shock causes some damage to the system and these damages accumulate. Let $Y_t$ be the damage caused by the $i^{th}$ shock. Assume $Y_t$s are independent of each other and $N_t$. $X_t = \sum_{k=1}^{N_t} Y_k$ is the total damage till time $t$. Suppose the system fails when $X_t > \alpha$, where $\alpha > 0$. If $\mathbb{P}\{Y_t = k\} = (1 - p)^{k-1}p$, $k = 1, 2, \ldots$ Calculate the mean time till system failure.

**Problem 8 (M/M/1 queue):** A Poisson process $N_t$ with rate $\lambda$ form the arrivals to a queue. Each customer requires an i.i.d. service time of exponential distribution with rate $\mu$. Let $q_t$ be the number of customers at time $t$, $D_t$ the number of customers departed till time $t$. Then $q_t = N_t - D_t$.

1. Calculate $\mathbb{P}\{D_t = m | N_t = n\}$
2. Calculate $\mathbb{P}\{q_t = m | N_t = n\}$
3. Calculate $\mathbb{P}\{q_t = m\}$, $\mathbb{E}[q_t^k]$ for $k = 1, 2, \ldots$
4. Let $q_n$ be the queue length when the $n^{th}$ customer arrives excluding itself. Calculate $\mathbb{P}[q_n = n | q_{n-1} = m]$.

**Problem 9:** Events occur according to a Poisson process with rate $\lambda$. These events are registered by a counter. However, each time an event is registered, the counter is blocked for the next $b$ units of time. Any new event that occurs when the counter is blocked is not registered by the counter. Let $R_t$ denote the number of registered events that occur by time $t$ (= number of events that occurred when the counter was not blocked).

1. Find the probability that the first $k$ events are all registered.
2. For $t \geq (n - 1)b$, find $\mathbb{P}\{R(t) \geq n\}$. 13
Chapter 2

Renewal Theory and Regenerative Processes

Lecture 4

Course E2.204: Stochastic Processes and Queueuing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

2.1 Renewal Process: Introduction

We know that the interarrival times for the Poisson process are independent and identically distributed exponential random variables. If we consider a counting process for which the interarrival times are independent and identically distributed with an arbitrary distribution function on $\mathbb{R}^+$, then the counting process is called a renewal process.

Let $X_n$ be the time between the $(n - 1)$th and $n$th event and $\{X_n, n = 1, 2, \ldots\}$ be a sequence of nonnegative independent random variables with common distribution $F$ and $X_n \geq 0$.

The mean time $\mu$ between successive events is given by

$$\mu = \mathbb{E}X_n = \int_0^\infty x dF(x).$$

We take $\mu > 0$. Let $S_0 = 0$ and $S_n = \sum_{i=1}^n X_i$ for $n \geq 1$, $S_n$ indicating the time of $n^{th}$ event. The number of events by time $t$, is given by

$$N(t) = \sup\{n : S_n \leq t\}.$$

**Definition 2.1.1.** The counting process $\{N(t) : t \geq 0\}$ is called a renewal process.

Note that the number of renewals by time $t$ is greater than or equal to $n$ if, and only if, the $n^{th}$ renewal occurs before or at time $t$. That is,

$$N(t) \geq n \Leftrightarrow S_n \leq t.$$
The distribution of \( N(t) \) can be written as \( \mathbb{P}\{N(t) \geq n\} = \mathbb{P}\{S_n \leq t\} \) and from this we can write \( \mathbb{P}\{N(t) = n\} \) as follows,

\[
\mathbb{P}\{N(t) = n\} = \mathbb{P}\{N(t) \geq n\} - \mathbb{P}\{N(t) \geq n + 1\} = \mathbb{P}\{S_n \leq t\} - \mathbb{P}\{S_{n+1} \leq t\}.
\]

By strong law of large numbers \( \frac{S_n}{n} \to \mu \) as \( n \to \infty \) with probability 1. Hence \( S_n \to \infty \) a.s. Also,

\[
\mathbb{P}\{N(t) < \infty\} = 1 - \lim_{n \to \infty} \mathbb{P}\{N(t) \geq n\} = 1 - \lim_{n \to \infty} \mathbb{P}\{S_n \leq t\} = 1.
\]

**Proposition 2.1.2.** \( \mathbb{E}\[N'(t)^r\] < \infty \) for \( r > 0, t \geq 0 \).

**Proof.** Construct a new process \( X'_k \) from \( X_k \) as follows,

\[
X'_k = \begin{cases} 
0, & X_k < \alpha \\
\alpha, & X_k \geq \alpha 
\end{cases}
\]

where \( X_k \) are the interarrival times of the original process. Let \( \beta = \mathbb{P}\{X_1 \geq \alpha\} \).

Let \( \{N'(t)\} \) be constructed from \( \{X'_k\} \) as interarrival times. Then it is clear that \( X'_k \leq X_k \) and \( N'(t) \geq N(t) \). Then,

\[
\mathbb{P}\{N'(\frac{\alpha}{2}) = n\} = \mathbb{P}\{X_1 < \alpha\}\mathbb{P}\{X_2 < \alpha\} \cdots \mathbb{P}\{X_{n-1} < \alpha\}\mathbb{P}\{X_n \geq \alpha\} 
\leq (1 - \beta)^n \beta.
\]

Thus,

\[
\mathbb{E}[(N'(\frac{\alpha}{2}))^r] = \sum_{n=0}^{\infty} n^r \mathbb{P}\{N'_\alpha = n\} 
\leq \sum_{n=0}^{\infty} n^r (1 - \beta)^n \beta < \infty.
\]

\[
\mathbb{E}[(N'(t))^r] = \mathbb{E} \left[ \sum_{k=1}^{\lfloor t \rfloor + 1} (N'_k)^r \right] 
\leq \left( \frac{t}{\alpha} + 1 \right)^r \mathbb{E}[N'(\frac{\alpha}{2})^r] < \infty.
\]

where \( \overline{N}_k = N'_\alpha - N'_{\alpha(k-1)} \).  

\( \Box \)
## Lecture 5

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019  
Instructor: Vinod Sharma  
Indian Institute of Science, Bangalore

### 2.2 Limit Theorems

We have

$$\lim_{t \to \infty} P \{ N_t \geq n \} = \lim_{t \to \infty} P \{ S_n \leq t \} = 1.$$  

Hence, \( \lim_{t \to \infty} N(t) = \infty \) a.s.

Let us denote by \( S_{N(t)} \), the time of the last renewal prior to or at time \( t \) and by \( S_{N(t)+1} \) the time of the first renewal after time \( t \).

**Proposition 2.2.1.** With probability 1,

$$\lim_{t \to \infty} \frac{N(t)}{t} \to \frac{1}{\mu}.$$  

**Proof.** Since, \( S_{N(t)} < t < S_{N(t)+1} \), we see that

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)}.$$  

By strong law of large numbers \( \frac{S_{N(t)}}{N(t)} \to \mu \) a.s. as \( t \to \infty \). Also,

$$\frac{S_{N(t)+1}}{N(t)} = \left[ \frac{S_{N(t)+1}}{N(t)+1} \right] \left[ \frac{N(t)+1}{N(t)} \right] \to \mu \text{ a.s.}$$

**Definition 2.2.2.** (Stopping Time) \( N \) a nonnegative integer valued random variable is a stopping time w.r.t. sequence \( \{X_k\} \) if \( \{N \leq n\} \) is a function of \( \{X_1, X_2, \ldots, X_n\} \).

**Theorem 2.2.3.** (Wald’s Lemma) If \( N \) is a stopping time w.r.t. \( \{X_k\} \) is i.i.d. and \( E[X_1] = \mu < \infty, E[N] < \infty \) then

$$E[\sum_{k=1}^{N} X_k] = E[N]E[X_1].$$  

**Proof.** Let

$$I_k = \begin{cases} 1, & N \geq k \\ 0, & N < k \end{cases}$$

Then we can write the following

$$\sum_{k=1}^{N} X_k = \sum_{k=1}^{\infty} X_k I_k.$$
Hence,
\[ E[\sum_{k=1}^{N} X_k] = \sum_{k=1}^{\infty} E[X_k I_k]. \]

Since \( I_k \) is determined by \( X_1, X_2, \ldots, X_{k-1} \), therefore \( I_k \) is independent of \( X_k \). Thus we obtain
\[ E[\sum_{k=1}^{N} X_k] = \sum_{k=1}^{\infty} E[X_k] E[I_k] = \sum_{k=1}^{\infty} E[X_k] \mathbb{P}\{N \geq k\} = \sum_{k=1}^{\infty} E[X_k] \mathbb{E}[N]. \]

\[ \square \]

**Theorem 2.2.4 (Elementary Renewal Theorem).**

\[ \lim_{t \to \infty} \frac{\mathbb{E}[N(t)]}{t} = \frac{1}{\mu}. \]

**Proof.** Let us denote \( \mathbb{E}[N(t)] \) as \( m(t) \) and we prove the result for \( \mu < \infty \). We know that \( S_{N(t)+1} > t \) and \( N(t)+1 \) is a stopping time. Taking expectations, by Wald’s lemma, we get,
\[ \mu (m(t) + 1) > t. \]

This implies
\[ \liminf_{t \to \infty} \frac{m(t)}{t} \geq \frac{1}{\mu}. \]

To get the other way, define a new renewal process from \( \{X'_k, k = 1, 2, \ldots\} \) where \( X'_k = \min(M, X_k) \), for a constant \( M > 0 \). Let \( S'_{n} = \sum_{i=1}^{n} X'_i \) and \( N'(t) = \text{sup} \{n, S'_n \leq t\} \). The interarrival times for this truncated renewal process are bounded by \( M \). Therefore
\[ S'_{N'(t)+1} \leq t + M. \]

Taking expectations on both sides we get,
\[ \mu_M (m'(t) + 1) \leq t + M \]

where \( \mu_M = \mathbb{E}[X_k'] \). Thus
\[ \limsup_{t \to \infty} \frac{m'(t)}{t} \leq \frac{1}{\mu_M}. \]

Since \( m(t) \leq m'(t) \),
\[ \limsup_{t \to \infty} \frac{m(t)}{t} \leq \frac{1}{\mu_M}. \]

for all \( M > 0 \). Taking \( M \to \infty, \mu_M \to \mu \) then implies that \( \limsup_{t \to \infty} m(t)/t \leq \mu. \) \( \square \)
2.2.2 Blackwell’s Theorem

Theorem 2.2.5 (Blackwell’s theorem). Let $F$ be the distribution of interarrival times.

1. If $F$ is non-lattice
   \[
   \lim_{t \to \infty} E[N(t+a) - N(t)] = \frac{a}{\mu}
   \]
   for all $a \geq 0$.

2. If $F$ is lattice with period $d$
   \[
   \lim_{n \to \infty} E[N((n+1)d) - N(nd)] = \frac{d}{\mu}.
   \]

Proposition 2.2.6. Blackwell’s theorem implies elementary renewal theorem.

Proof:

\[
E\left[ \frac{N_n}{n} \right] = \frac{1}{n} \sum_{k=1}^{n} (E[N_k] - E[N_{k-1}])
\]

From Blackwell’s theorem and definition of limits, for every $\varepsilon > 0$, there exists $n_0$ such that for $n > n_0$

\[
|E[N_k] - E[N_{k-1}] - \frac{1}{\mu}| < \varepsilon. \tag{2.1}
\]

Therefore for $n > n_0$,

\[
E\left[ \frac{N_n}{n} \right] = \frac{1}{n} \left( \sum_{k=1}^{n_0} (E[N_k] - E[N_{k-1}]) + \sum_{k=n_0+1}^{n} (E[N_k] - E[N_{k-1}]) \right).
\]

\[
\leq \frac{1}{n} \left( \sum_{k=1}^{n_0} (E[N_k] - E[N_{k-1}]) + (n - n_0 - 1) \left( \varepsilon + \frac{1}{\mu} \right) \right).
\]

Thus, taking $n \to \infty$,

\[
\limsup_{n \to \infty} \frac{E[N_n]}{n} \leq \frac{1}{\mu} + \varepsilon.
\]

Now take $\varepsilon \to 0$.

Similarly, by taking the opposite sign of the modulus in Eq 2.1, we get

\[
\liminf_{n \to \infty} \frac{E[N_n]}{n} \geq \frac{1}{\mu}.
\]

Thus,

\[
\lim_{n \to \infty} \frac{E[N_n]}{n} = \frac{1}{\mu},
\]

which is elementary renewal theorem. \qed
2.2.3 Renewal Equation

Definition 2.2.7 (Renewal Equation). A functional equation of the form

\[ Z(t) = z(t) + F * Z(t) \]

where \( z(t) \) is a function on \([0, \infty)\) and \(*\) denotes convolution is called a renewal equation. \( F \) and \( z \) are known and \( Z \) is the unknown function.

The renewal equation arises in several situations. We need to know the conditions for existence and uniqueness of the solution for renewal equation. The following theorem provides the answer.

Proposition 2.2.8. If \( F(\infty) = 1, F(0) < 1 \) and \( z(t) : [0, \infty) \to [0, \infty) \) is bounded on bounded intervals the renewal equation has a unique solution given by

\[ Z(t) = U(t) * z(t) \]

where \( U(t) = E[N(t)] = \sum_{k=0}^{\infty} F^k(t) \). Here, \( F^n \) denotes \( n \) fold convolution of \( F \) with itself.

Proof. Define \( U^n = \sum_{k=0}^{n} F^k \). Now, \( U^n \to U \) monotonically. Let \( Z^n(t) = U^n * z(t) = \sum_{k=0}^{n} F^k * z(t) \). We have

\[ Z^{n+1}(t) = z + \sum_{k=1}^{n} F^k * z(t) \]

\[ = z + F * Z^n \]

Let \( Z(t) := U * z(t) \). Since, \( U^n \to U \) monotonically, from monotone convergence theorem, we get \( Z^n \to Z \).

Fixing \( t \), we get from above equation

\[ \lim_{n \to \infty} Z^{n+1}(t) = \lim_{n \to \infty} Z^n + z(t) \]

\[ = Z(t) = z(t) + Z * z(t) \]

Thus, we see that \( Z(t) = U * z(t) \) is a solution of the renewal equation.

Uniqueness: Let \( Z_1 \) and \( Z_2 \) be two solutions. We have \( Z_1 - Z_2 = F *(Z_1 - Z_2) \). Let \( V = Z_1 - Z_2 \). Since \( U \) and \( z \) are bounded on bounded intervals, so are \( Z_1, Z_2 \) and \( V \). Iterating \( n \) times, we get \( V = F^n * V \).

Therefore

\[ |V| = |\int_0^t V(t-s)dF^n(s)| \]

\[ \leq M_1 \int_0^t dF^n(s) \]

\[ = M_1 P\{S_n \leq t\} \]

\[ = 0 \text{ as } n \to \infty. \]

where \( |V(s)| \leq M_1 \) for all \( 0 \leq s \leq t \).

Example of renewal equation: Consider residual time of a renewal process \( \{B_t\} \).

\[ P\{B_t \leq x\} = P\{B_t \leq x, X_1 > t\} + \int_0^t P\{B_t \leq x|X_1 = s\} F(ds) \]
Letting \( Z(t) = \mathbb{P}\{B_t \leq x\}, \) \( z(t) = \mathbb{P}\{B_t \leq x, X_1 > t\}, \) and noting that \( \mathbb{P}\{B_t \leq x | X_1 = s\} = \mathbb{P}\{B_{t-s} \leq x\} \), we have

\[
Z(t) = z(t) + F * Z(t),
\]

which is a renewal equation whose solution is \( \mathbb{P}\{B_t \leq x\} \).

Let \( A(t) \) be the age of the process. Then, we have \( \{B_t \leq x\} = \{A_{t+x} \leq x\} \). If \( \{B_t\} \) is stationary we have \( \mathbb{P}\{B_{t+s} \leq x\} = \mathbb{P}\{B_t \leq x\} \forall x, t, s \). Thus, we get \( \mathbb{P}\{A_{t+s+x} \leq x\} = \mathbb{P}\{A_{t+x} \leq x\} \). This means that \( \{A_t\} \) is also stationary. We can similarly show that \( \{A_t\} \) is stationary implies that \( \{B_t\} \) is also stationary.
2.2.4 Renewal theorems

**Definition 2.2.9 (Directly Riemann Integrable (d.r.i.)).** A function \( z: \mathbb{R}^+ \rightarrow \mathbb{R}^+ \) is called directly Riemann integrable if
\[
\lim_{h \to 0} \sum_{k=1}^{\infty} \inf_{(k-1)h \leq t \leq kh} z(t) = \lim_{h \to 0} \sum_{k=1}^{\infty} \sup_{(k-1)h \leq t \leq kh} z(t).
\]
Then the limit is denoted as \( \int_0^{\infty} z(t) \, dt \). For \( z: \mathbb{R}^+ \rightarrow \mathbb{R} \), define \( z^+: t \mapsto \max(z(t), 0) \) and \( z^-: t \mapsto -\min(0, z(t)) \).

**Proposition 2.2.10.** We take \( z \geq 0 \).

1. A necessary condition for \( z \) to be d.r.i is \( z \) is bounded and continuous a.e.
2. \( z \) is d.r.i. if (1) holds and any of the following holds.
   a. \( z \) is non-increasing and Lebesgue integrable.
   b. \( z \leq z' \) and \( z' \) is d.r.i.
   c. \( z \) has bounded support.
   d. \( \int_0^{\infty} z(t) \, dt < \infty \) for some \( h > 0 \).

**Theorem 2.2.11 (Key Renewal Theorem).** Let \( Z(t) = z(t) + F \ast Z(t) \) be the renewal equation. Then, if \( z \) is directly Riemann integrable (d.r.i.) for the solution \( Z(t) = \mathbb{E}[N(t)] \ast z(t) \) and \( F \) is non-lattice, the following limit holds
\[
\lim_{t \to \infty} \mathbb{E}[N(t)] \ast z(t) = \frac{1}{\mu} \int_0^{\infty} z(t) \, dt.
\]
Here, \( \mathbb{E}[N(t)] = \sum_{k=0}^{\infty} F^{*k}(t) \).

**Definition 2.2.12 (Delayed renewal process).** Let \( X_k \) for \( k = 0, 1, \ldots \) be independent. Let \( X_0 \sim F' \) and \( X_1, X_2, \ldots \) i.i.d with distribution \( F \). The renewal process defined using these r.v.s as interarrival times is called a delayed renewal process.

**Theorem 2.2.13.** If \( \mu < \infty \) and \( F \) is non-lattice, for a renewal process (or a delayed renewal process with arbitrary \( F' \)), the following all hold and are equivalent.

1. (Key Renewal Theorem): Let \( Z(t) = z(t) + F \ast Z(t) \) be the renewal equation and \( z \) be d.r.i. Then,
\[
\lim_{t \to \infty} Z(t) = \frac{1}{\mu} \int_0^{\infty} z(t) \, dt.
\]

2. \( \mathbb{P}[A_t \leq x] \to F_0(x) \) as \( t \to \infty \).
(3) \( \mathbb{P}\{B_t \leq x\} \rightarrow F_0(x) \) as \( t \rightarrow \infty \).

(4) (Blackwell’s theorem):

\[
\mathbb{E}[N_{t+a} - N_t] \xrightarrow{a}{\mu} \text{as } t \rightarrow \infty.
\]

where

\[
F_0(x) = \frac{1}{\mu} \int_0^x F(t) dt.
\]

If \( \mu = \infty \), the above results hold with \( F_0(x) = 0 \forall x \).

Proof.

- (1) \( \implies \) (2): Let \( Z(t) = \mathbb{P}\{A_t \leq x\} \). Then, \( Z(t) \) satisfies the renewal equation with \( z(t) = \mathbb{P}\{A_t \leq x, X_1 > t\} = 1\{t \leq x\}\mathbb{P}\{A_t > t\} \) is d.r.i. because it is bounded and continuous a.e. Thus, from key renewal theorem (1) we have, as \( t \rightarrow \infty \)

\[
\mathbb{P}\{A_t \leq x\} \rightarrow \frac{1}{\mu} \int_0^x 1\{t \leq x\} \mathbb{P}\{A_1 > t\} dt
= \frac{1}{\mu} \int_0^x F(t) dt
= F_0(x).
\]

- (2) \( \iff \) (3): We have the relationship for any \( t > 0 \), \( \{B_t \leq x\} = \{A_{t+x} \leq x\} \). Taking the limit as \( t \rightarrow \infty \) we get the equivalence.

- (2) \( \implies \) (4):

\[
\mathbb{E}[N_{t+a} - N_t] = \int_0^a \mathbb{E}[N_{t+a} - N_t | B_t = s] d\mathbb{P}_B(s) + \int_a^\infty \mathbb{E}[N_{t+a} - N_t | B_t = s] d\mathbb{P}_B(s)
= \int_0^a \mathbb{E}[N_{t+a} - N_t | B_t = s] d\mathbb{P}_B(s) + 0
= \int_0^a U(a-s) d\mathbb{P}_B(s)
\]

Now since \( \mathbb{P}\{B_t \leq x\} \rightarrow F_0(x) \) and \( U \) is bounded and continuous a.e.,

\[
\lim_{t \rightarrow \infty} \int_0^a U(a-s) d\mathbb{P}_B(s) = \lim_{t \rightarrow \infty} \int_0^a 1\{s \leq a\} U(a-s) d\mathbb{P}_B(s)
= \int_0^a 1\{s \leq a\} U(a-s) dF_0(s)
= \int_0^a U(a-s) dF_0(s)
= U \ast F_0(a)
\]

Noting that \( U(t) = \sum_{k=0}^\infty F^{*n}(t) \) and

\[
\frac{d}{dt} F_0(t) = \frac{1 - F(t)}{\mu},
\]

22
we get $U * F_0$ has density
\[
\frac{1}{\mu} U * (1 - F) = \frac{1}{\mu} (U - U * F) = \frac{1}{\mu} \left( \sum_{k=0}^{\infty} F^*k - \sum_{k=1}^{\infty} F^*k \right) = \frac{1}{\mu}.
\]

Therefore, we get $U * F_0(a) = a/\mu$ and hence
\[
\lim_{t \to \infty} E[N_t + a - N_t] = a/\mu.
\]

• (4) $\implies$ (1): For small enough $h$ and appropriate $n$ such that $nh < t \leq (n+1)h$, we write
\[
Z(t) = U * z(t) = \int_0^t z(t-a)dU(a) = \int_0^{t-nh} z(t-a)dU(a) + \int_{t-nh}^t z(t-a)dU(a)
\]
Since $z$ is d.r.i., and hence bounded, as $h \to 0$, the first term goes to zero. Taking
\[
\xi_h(k) = \max_{x \in [t-(k+1)h, t-kh]} z(x),
\]
and noting that $|U(t - (k+1)h) - U(t - kh)| \leq U(h)$, the second term
\[
\int_{t-nh}^t z(t-a)dU(a) \leq \sum_{k=0}^{n} \xi_h(k)[U(t - kh) - U(t - (k+1)h)]
\]
\[
= \sum_{k=0}^{M} \xi_h(k)[U(t - kh) - U(t - (k+1)h)] + \sum_{k=M+1}^{n} \xi_h(k)[U(t - kh) - U(t - (k+1)h)]
\]
\[
\leq \sum_{k=0}^{M} \xi_h(k)[U(t - kh) - U(t - (k+1)h)] + U(h) \sum_{k=M+1}^{n} \xi_h(k).
\]
Now take $t \to \infty$. Then, $n \to \infty$ also. Thus, by Blackwell’s theorem (4), the above converges to
\[
\frac{1}{\mu} \sum_{k=0}^{M} \xi_h(k) + U(h) \sum_{k=M+1}^{n} \xi_h(k).
\]
Now, take $M \to \infty$. Since $z$ is d.r.i., the second term goes to zero. Next, take $h \to 0$. Then, the first term goes to
\[
\frac{1}{\mu} \int_0^\infty z(t)dt.
\]
Therefore,

\[ \limsup_{t \to \infty} U \ast z(t) \leq \frac{1}{\mu} \int_{0}^{\infty} z(t) dt. \]

Similarly, we can show

\[ \liminf_{t \to \infty} U \ast z(t) \geq \frac{1}{\mu} \int_{0}^{\infty} z(t) dt. \]

Discrete time versions and when \( F \) is lattice for the above theorems also hold. \( \square \)
2.3 Regenerative Processes

Let \( \{X_t\} \) be a stochastic process and \( Y_0, Y_1, Y_2, \ldots \) be i.i.d with distribution \( F \). Let \( S_n = \sum_{k=0}^{n} Y_k \).

**Definition 2.3.1.** The process \( \{X_t\} \) is regenerative if there exists \( Y_0, Y_1, Y_2, \ldots \) i.i.d such that the process \( Z_{n+1} = \{X_{S_n+t}, t \geq 0\} \) is independent of \( Y_0, Y_1, Y_2, \ldots, Y_n \) and the distribution of \( Z_{n+1} \) does not depend on \( n \). \( \{X\} \) is a delayed regenerative process if distribution of \( Y_0 \) is different from \( Y_1 \).

**Examples:**

1. The process \( \{B_t\} \) corresponding to residual life in a renewal process is regenerative if we take \( Y_k = X_k \) where \( X_k \) is the \( k \)th inter-arrival time.

2. In a Markov chain the time instants when the Markov chain visits a particular state, say \( i_0 \), the process regenerates itself.

3. Consider a GI/GI/1 queue. The process \( \{q_t\} \), the queue length at time \( t \) is a continuous time regenerative process which regenerates when an arrival sees empty queue. The process \( \{W_n\} \), the waiting time of the \( n \)th customer is a discrete time regenerative process with the above arrival epochs.

**Theorem 2.3.2.** Let \( \{X_t, t \geq 0\} \) be a delayed regenerative process with \( \mu = \mathbb{E}[Y_1] < \infty \) and \( F \) is non-lattice. Let \( f \) be a bounded, continuous function a.s. Then,

\[
\lim_{t \to \infty} \mathbb{E}[f(X_t)] = \mathbb{E}_c[f(X_0)] = \frac{1}{\mu} \mathbb{E}_0 \left[ \int_0^{Y_1} f(X_s) ds \right]
\]

where \( \mathbb{E}_c \) is the expectation w.r.t. equilibrium or stationary distribution and \( \mathbb{E}_0 \) is the expectation w.r.t the process when \( S_0 = 0 \).

**Proof.** We show for \( S_0 = 0 \). Then,

\[
\mathbb{E}[f(X_t)] = \mathbb{E}[f(X_t), Y_1 > t] \int_0^{Y_1} \mathbb{E}[f(X_s)] ds F(s)
\]

\[
= \mathbb{E}[f(X_t), Y_1 > t] + \int_0^{Y_1} \mathbb{E}[f(X_{s-})] ds F(s)
\]

\[
= \mathbb{E}[f(X_t), Y_1 > t] + \mathbb{E}[f(X)] F(t)
\]

This is a renewal equation. Since \( f \) is bounded and \( X_t \) is right continuous, it follows that \( \mathbb{E}[f(X_t), Y_1 > t] \) is d.r.i. Therefore,

\[
\lim_{t \to \infty} \mathbb{E}[f(X_t)] = \frac{1}{\mu} \int_0^{\infty} \mathbb{E}_0[f(X_s), Y_1 > s] ds dt
\]

\[
= \frac{1}{\mu} \mathbb{E}_0 \left[ \int_0^{\infty} f(X_s) 1 \{Y_1 > s\} ds \right]
\]

\[
= \frac{1}{\mu} \mathbb{E}_0 \left[ \int_0^{Y_1} f(X_s) ds \right] \square
\]
The following theorem for lattice $F$ can be proved in the same way.

**Theorem 2.3.3.**

1. For discrete time regenerative processes \{\(X_n\)\} and $F$ is aperiodic,
   \[
   \lim_{n \to \infty} \mathbb{E}[f(X_n)] = \mathbb{E}_e[f(X_n)] = \frac{1}{\mu} \mathbb{E}_0 \left[ \sum_{k=1}^{Y_1} f(X_k) \right].
   \]

2. For discrete time regenerative processes \{\(X_n\)\} and $F$ has period $d$,
   \[
   \lim_{n \to \infty} \frac{1}{d} \sum_{k=0}^{d-1} \mathbb{E}[f(X_{n+k})] = \mathbb{E}_e[f(X_n)] = \frac{1}{\mu} \mathbb{E}_0 \left[ \sum_{k=1}^{Y_1} f(X_k) \right].
   \]

\(\square\)

Taking $f(X_t) = 1\{X_t \leq x\}$, for the non-lattice case we have
\[
\lim_{t \to \infty} \mathbb{P}\{X_t \leq x\} = \frac{\mathbb{E} \left[ \int_0^{Y_1} 1\{X_t \leq x\} \right]}{\mathbb{E}[Y_1]}.
\]

Similar results hold for the lattice case.

**Example (GI/GI/1 queue):** If regenerative lengths $\tau$ of \{\(W_n\)\} in GI/GI/1 queue satisfies $E[\tau] < \infty$ and it is aperiodic, then $W_0$ has unique stationary distribution and $W_n \to W_\infty$ where $W_\infty$ is a r.v. with the stationary distribution. We can show that if $E[A] < E[s]$, then the above conditions are satisfied. Also, if the queue starts empty with an arrival, then, if $\tau$ is the regeneration length of \{\(q_t\)\}, then $\tau = \sum_{k=0}^{\infty} A_k$. Since, $\tau$ is a stopping time w.r.t \{\(A_n, s_n\)\} sequence, by Wald’s lemma, $E[\tau] = E[A_1]E[\tau] < \infty$ whenever $E[\tau] < \infty$. Thus when $E[A_1] < E[s_1]$, \{\(q_t\)\} also has a unique stationary distribution and starting from any initial distribution, it converges in distribution to the limiting distribution.

**Theorem 2.3.4 (Strong law for regenerative processes).**

- If $F$ is non-lattice (with arbitrary distribution of $Y_0$) and $E[\int_1^t |H(X_t)| dt] < \infty$,
  \[
  \lim_{t \to \infty} \frac{1}{t} \int_0^t f(X_s)ds = \mathbb{E}[X_e] \text{ a.s.}
  \]

- If $F$ is lattice, with similar conditions,
  \[
  \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \mathbb{E}[X_e] \text{ a.s.}
  \]

**Proof.** We show only for the non-lattice case. The proof for other cases is similar.

\[
\frac{1}{t} \int_0^t f(X_s)ds = \frac{1}{t} \int_0^{S_1} f(X_s)ds + \frac{1}{t} \int_{S_1}^{S_2} f(X_s)ds + \cdots + \frac{1}{t} \int_{S_{N-1}}^{S_N} f(X_s)ds + \frac{1}{t} \int_{S_N}^{t} f(X_s)ds
\]

\[
= \frac{1}{t} (U_0 + U_1 + U_2 + \cdots + U_N + \Delta).
\]

26
where $U_1, U_2, \ldots$ are i.i.d. with $U_i = \int_{t_i}^{t_i+1} f(X_s)ds$, $\Delta = \int_{t_0}^t f(X_s)ds$ and $U_0 < \infty$ a.s. We have,

$$\lim_{t \to \infty} \frac{1}{t} (U_1 + U_2 + \cdots + U_N) = \lim_{t \to \infty} \frac{1}{N} (U_1 + U_2 + \cdots + U_N) \frac{N_t}{t} = \frac{\mathbb{E}[U_1]}{\mathbb{E}[Y_1]} \text{ a.s.}$$

by ordinary S.L.L.N. and elementary renewal theorem.

To complete the proof, we need to show that $\lim_{t \to \infty} \Delta/t \to 0$ a.s. as $t \to \infty$. We have,

$$\frac{\Delta}{t} \leq \frac{1}{t} \int_{S_{N_t}} |f(X_s)| ds \leq \frac{1}{t} \int_{S_{N_t}}^{S_{N_t+1}} |f(X_s)| ds.$$

Let

$$V_k = \int_{S_k}^{S_{k+1}} |f(X_s)| ds.$$

We have

$$\frac{V_{N_t+1}}{t} = \frac{V_{N_t+1} N_t + 1}{N_t + 1}.$$

Now, $(N_t/t) \to 1/\mathbb{E}[Y_1]$ a.s. We need to show that $\lim_{t \to \infty} V_{N_t}/N_t = 0$. This will be true if $V_n/n \to 0$ a.s. as $n \to \infty$. But,

$$\mathbb{P} \left\{ \bigcup_{n=N}^{\infty} \left\{ \frac{V_n}{n} > \epsilon \right\} \right\} \leq \sum_{n=N}^{\infty} \mathbb{P} \left\{ \frac{V_1}{\epsilon} > n \right\} \to 0$$

if

$$\sum_{n=0}^{\infty} \mathbb{P} \left\{ \frac{V_1}{\epsilon} > n \right\} < \infty.$$

This holds when $\mathbb{E}[V_1] < \infty$. \hfill \Box
2.4 Problems

Notation: \((X_1, X_2, \ldots)\) iid non-negative random variables with distribution \(F\). \(S_n = \sum_{k=1}^{n} X_k\). \(N(t)\) is number of arrivals till time \(t\) (excluding the one at 0). \(m(t) = \mathbb{E}[N(t)]\). \(\mu = \mathbb{E}[X_1]\)

**Problem 1:** Show that \(p\{X_{N(t)+1} + 1 \geq x\} \geq \bar{F}(x)\), also show that \(\mathbb{E}[(X_{N(t)+1})^m] \geq \mathbb{E}[X^m]\) for any positive integer \(m\). Compute \(p\{X_{N(t)+1} + 1 \geq x\}\) for \(X_i \sim \exp(\lambda)\).

**Problem 2:** Prove the renewal equation

\[
m(t) = F(t) + \int_0^t m(t-x)dF(x).
\]

**Problem 3:** If \(F\) is uniform on \((0,1)\) then show that for \(0 \leq t \leq 1\)

\[
m(t) = \exp(t) - 1.
\]

**Problem 4:** Consider a single server bank in which potential customers arrive at a Poisson rate \(\lambda\). However a customer only enters the bank if the server is free when the customer arrives. Let \(G\) denote the service distribution.

a) What fraction of time the server is busy?

b) At what rate customers enter the bank?

c) What fraction of potential customers enter the bank?

**Problem 5:** Find the renewal equation for \(\mathbb{E}[A(t)]\), then also find \(\lim_{t \to \infty} \mathbb{E}[A(t)]\).

**Problem 6:** Consider successive flips of a fair coin.

a) Compute the mean number of flips until the pattern \(HHT HHT T\) appears.

b) Which of the two patterns \(HHTT, HTHT\) requires a larger expected time to occur?

**Problem 7:** Consider a delayed renewal process \(\{N_D(t), t \geq 0\}\), whose first interarrival time has distribution \(G\) and the others have distribution \(F\). Let \(m_D(t) = \mathbb{E}[N_D(t)]\).

a) Prove that

\[
m_D(t) = G(t) + \int_0^t m(t-x)dG(x),
\]

Where \(m(t) = \sum_{n=1}^{\infty} F^{*n}(t)\).

b) Let \(A_D(t)\) denote the age at time \(t\). Show that if \(F\) is non-lattice with \(\int x^2dF(x) < \infty\) and \(t\tilde{G}(t) \to 0\) as \(t \to \infty\), then

\[
\mathbb{E}[A_D(t)] \to \frac{1}{2} \int_0^\infty x^2dF(x).
\]

**Problem 8:** Consider a \(GI/GI/1\) queue: Interarrival times \(\{A_n\}\) are iid and service times \(\{S_n\}\) are iid with \(\mathbb{E}[S_n] < \mathbb{E}[A_n] < \infty\). Let \(V(t)\) be the virtual service time in the queue at time \(t\). Show that

a) \(v \triangleq \lim_{t \to \infty} \frac{1}{t} \int_0^t v(s)ds\)
exists a.s. and is constant. Under condition $\mathbb{E}[S_1] < \mathbb{E}[A_1]$ the mean regeneration length for this process is finite.

b) Let $D_n$ be the amount of time $n$th customer waits in the queue. Define

$$W_Q = \lim_{n \to \infty} \frac{D_1 + D_2 + \cdots + D_n}{n}.$$ 

Show $W_Q$ exists a.s. and is constant.

c) Show $V = \lambda \mathbb{E}[Y] W_Q + \lambda \mathbb{E}[Y^2]/2$.

Where $1/\lambda = \mathbb{E}[A_n]$ and $Y$ has distribution of service time.
Chapter 3

Discrete Time Markov Chains

Lecture 9

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

3.1 Markov Chains: Definitions

Definition: Let $S$ be a countable set. A discrete time stochastic process $\{X_k\}$ is a Markov chain with state space $S$ if

$$P\{X_n = j | X_{n-1} = i, X_{n-2} = i_{n-2}, \ldots, X_1 = i_1, X_0 = i_0\} = P\{X_n = j | X_{n-1} = i\}.$$  

When $P\{X_n = j | X_{n-1} = i\}$ does not depend on $n$ is called a homogeneous Markov chain. Now onward, we assume all Markov chains are homogeneous, unless mentioned otherwise. The matrix $P$ where $P(i, j) = P\{X_n = j | X_{n-1} = i\}$ is called the transition matrix of the Markov chain $\{X_n\}$. We write $P^n$ to denote matrix multiplication of $P$ with itself $n$ times. We use $P_i\{\cdot\}$ to mean $P\{\cdot | X_0 = i\}.$

Strong Markov property: If $\tau$ is a stopping time and $\tau < \infty$ with probability 1, then Markov chain has strong Markov property if

$$P\{X_{\tau+1} = j | X_{\tau} = i, X_{\tau-1} = i_{\tau-1}, \ldots, X_1 = i_1, X_0 = i_0\} = P\{X_1 = j | X_0 = i\}. \quad \forall n \in \mathbb{N} \text{ and } i, j \in S.$$

Theorem 3.1.1. Every Markov chain has strong Markov property.
Proof.
\[
\begin{align*}
\mathbb{P}\{X_{t+1} = j|X_t = i, X_{t-1} = i_{t-1}, \ldots, X_1 = i_1, X_0 = i_0\} \\
= \sum_{m=1}^{\infty} \mathbb{P}\{X_{t+1} = j|X_t = i, X_{t-1} = i_{t-1}, \ldots, X_1 = i_1, X_0 = i_0, \tau = m\} \mathbb{P}\{\tau = m|X_t = i, X_{t-1} = i_{t-1}, \ldots, X_1 = i_1, X_0 = i_0\} \\
= \sum_{m=1}^{\infty} \mathbb{P}\{X_{m+1} = j|X_m = i, X_{m-1} = i_{m-1}, \ldots, X_1 = i_1, X_0 = i_0, \tau = m\} \mathbb{P}\{\tau = m|X_t = i, X_{t-1} = i_{t-1}, \ldots, X_1 = i_1, X_0 = i_0\} \\
= \sum_{m=1}^{\infty} \mathbb{P}\{X_{m+1} = j|X_m = i, X_{m-1} = i_{m-1}, \ldots, X_1 = i_1, X_0 = i_0\} \mathbb{P}\{\tau = m|X_t = i, X_{t-1} = i_{t-1}, \ldots, X_1 = i_1, X_0 = i_0\} \\
= \mathbb{P}\{X_1 = j|X_0 = i\}. 
\end{align*}
\]

The finiteness of \(\tau\) is used in the first equality to expand the probability as an infinite summation. The fact that \(\tau\) is a stopping time has been used in the third equality. \(\square\)

**Classification of states:** Let \(\tau_i = \min\{n \geq 1 : X_n = i\}\). \(\tau_i\) is a stopping time. Let \(N_i\) be the total number of times a state \(i\) is visited by the Markov chain.

- If for a state \(i\), \(\mathbb{P}\{\tau_i < \infty\} < 1\), it is called a transient state. For a transient state, \(\mathbb{P}\{N_i = m\} = \mathbb{P}\{\tau < \infty\}^m \mathbb{P}\{\tau < \infty\}\) and \(\mathbb{E}[N_i] = 1/\mathbb{P}\{\tau < \infty\} < \infty\). If \(\mathbb{P}\{\tau_i < \infty\} = 1\), state \(i\) is called a recurrent state. If further \(\mathbb{E}[\tau_i] < \infty\), it is called a positive recurrent state. When \(i\) is recurrent but \(\mathbb{E}[\tau_i] = \infty\), it is called a null recurrent state.

- For a recurrent state \(i\), \(\mathbb{P}\{N_i = \infty\} = 1\) and \(\mathbb{E}[N_i] = \infty\). Since, \(\mathbb{E}[N_i] = \mathbb{E}[\sum_{n=0}^{\infty} 1(X_n = i)] = \sum_{n=0}^{\infty} \mathbb{P}^{n}\{i, i\}\), an equivalent criterion for recurrence is \(\sum_{n=0}^{\infty} \mathbb{P}^{n}\{i, i\} = \infty\). The period of state \(i\) is denoted by \(d(i)\). If \(d(i) = 1\), \(i\) is called aperiodic.

-**Communicating classes:** A state \(j\) is said to be reachable from state \(i\) if there exists an \(n\) such that \(\mathbb{P}^{n}\{i, j\} > 0\). We denote this by \(i \rightarrow j\). If \(i \rightarrow j\) and \(j \rightarrow i\), we say that states \(i\) and \(j\) are communicating states and denote this by \(i \leftrightarrow j\). A subset \(A\) of state space is called closed if for all \(j \in A^{c}\), and \(i \in A\), \(j\) is not reachable from \(i\). A subset \(A\) of state space is called a closed communicating set if it is closed and \(i \leftrightarrow j, \forall i, j \in A\). Communication is an equivalence relation. A Markov chain with its state space \(S\) a communicating class is called an irreducible chain.

**Example:** \(S = \{0, 1, 2, 3, 4\}\).

\[
\begin{bmatrix}
0.2 & 0.3 & 0 & 0.5 & 0 \\
0 & 0.3 & 0.7 & 0 & 0 \\
0 & 0.5 & 0.5 & 0 & 0 \\
0 & 0 & 0 & 0.3 & 0.7 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

Here, \(\{3, 4\}\) and \(\{1, 2\}\) are closed sets. \(\{1, 2\}\) is closed communicating set. \(\{4\}\) is an absorbing state and \(\{0, 3\}\) are transient states. \(\{1, 2\}\) are recurrent states (positive).

### 3.2 Class Properties of Transience and Recurrence

**Proposition 3.2.1 (Periodicity is a class property).** If \(i \leftrightarrow j\), and \(i\) has period \(d\), \(j\) also has period \(d\).
Proof. \( i \rightarrow j \implies \exists n : P^n(i, j) > 0 \) and \( j \rightarrow i \implies \exists m : P^m(j, i) > 0 \). We have

\[
P^{n+m}(j, j) = \sum_k P^n(j, k)P^m(k, j) \geq P^n(j, i)P^m(i, j) > 0
\]

\[
P^{n+s+m}(j, j) \geq P^n(j, i)P^s(i, i)P^m(i, j) > 0
\]

The last inequality is true whenever \( P^n(i, i) > 0 \). From these two inequalities and definition of period, \( d(j) \) divides \( n + m \) and \( n + s + m \). Therefore, \( d(j) \) divides \( s \) whenever \( P^n(i, i) > 0 \). In particular, \( d(j) \) divides \( d(i) \). Using exactly the same argument with roles of \( i \) and \( j \) interchanged, we can show \( d(i) \) divides \( d(j) \). Thus, \( d(i) = d(j) \).

\[\Box\]

Proposition 3.2.2 (Recurrence is a class property). If \( i \leftrightarrow j \), and \( i \) is recurrent, then \( j \) is also recurrent.

Proof. Since \( i \) is recurrent, \( \sum_n P^n(i, i) = \infty \). Since \( i \leftrightarrow j \), \( \exists m, n : P^n(i, j) > 0, P^m(j, i) > 0 \). Therefore,

\[
\sum_k P^{n+k+m}(j, j) \geq P^n(j, i) \left( \sum_k P^k(i, i) \right) P^m(i, j) = \infty.
\]

This shows that state \( j \) is also recurrent. \[\Box\]

If \( i \) is transient and \( j \) is recurrent, then as the above example shows, \( i \rightarrow j \) is possible but \( j \rightarrow i \) is not possible, as we now show. If \( i \rightarrow j \), then \( j \rightarrow i \) is ruled out by Prop 3.2.2. If \( i \rightarrow j \) is not true, but \( j \rightarrow i \) is true, then there is \( m \) such that \( P^n(j, i) > 0 \) without \( j \) visiting itself. But then, \( P_j(\tau_j = \infty) \geq P^n(j, i) > 0 \). Thus \( j \) will not be recurrent.

Thus, the state space \( S \) can be partitioned into disjoint sets where one set could include all the transient states and then there are disjoint communicating closed classes. In the above example this partition is \{0, 3\}, \{1, 2\} and \{4\}.

### 3.3 Limiting distributions of Markov chains

Let \( \mu_{jj} = \mathbb{E}_j[\tau_j] \). The time periods at which the Markov chain enters state \( j \) are renewal epochs and \( \mu_{jj} \) is the expected time between renewals. From regeneration process theorem,

\[
\lim_{n \to \infty} P^n(i, j) = \frac{1}{\mu_{jj}} \quad \text{(when } j \text{ is aperiodic)},
\]

\[
\lim_{n \to \infty} P^{nd}(j, j) = \frac{d}{\mu_{jj}} \quad \text{(when } j \text{ has period } d > 1)\).
\]

If the initial state \( i \) is positive recurrent and aperiodic, then visits to state \( i \) are regeneration epochs and we have

\[
\lim_{n \to \infty} P^n(i, j) = \frac{\mathbb{E}_i \left[ \sum_{k=1}^{\tau_i} 1(X_k = j) \right]}{\mu_{ii}}.
\]

In the above, if state \( j \) is transient or null recurrent, \( \mu_{jj} = \infty \) and the corresponding limits equal zero.

\[32\]
Lecture 10

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

3.2 (Contd.) Limiting distributions of Markov chains

In the following, we assume that the Markov chain is irreducible and aperiodic.

Let \( \pi(j) = \lim_{n \to \infty} P^n(i, j) \). If \( j \) is a transient or null recurrent, \( \pi(j) = 0 \). When state \( j \) is positive recurrent, \( \pi(j) > 0 \) and the Markov chain converges in distribution to the limiting distribution \( \pi \).

Let \( X_0 \) be \( \pi \). The distribution of \( X_1 \) is then \( \pi P \). If \( \pi = \pi P \), the distribution of \( X_n, n \geq 1 \) is \( \pi P^n = \pi \). This suggests that the solution of \( \pi = \pi P \) such that \( \sum \pi(i) = 1 \) could be a stationary distribution of the Markov chain.

**Proposition 3.2.1.** The solution \( \pi \) of the equation \( \pi = \pi P \) such that \( \sum \pi_i = 1 \) is the stationary distribution of the Markov chain \( \{X_n\} \) with transition probability matrix \( P \).

**Proof.** We need to show that if \( X_0 \) has the distribution \( \pi \), then the distribution of \( X_n \) is also \( \pi \) and joint distribution of \( \{X_{k+1}, X_{k+2}, \ldots, X_{k+m}\} \) does not depend on \( k \). The distribution of \( X_n \) being equal to \( \pi \) has been deduced in the discussion above. We now show the remaining. We have

\[
\mathbb{P}\{X_{k+1} = i_1, X_{k+2} = i_2, \ldots, X_{k+m} = i_m\} = \sum_{i_0} P(X_k = i_0) \times P(i_0, i_1) \times \cdots \times P(i_{m-2}, i_{m-1}) \times P(i_{m-1}, i_m).
\]

Also, \( X_{k+1} = i_1, X_{k+2} = i_2, \ldots, X_{k+m} = i_m \).

\[\square\]

**Proposition 3.2.2 (Positive recurrence is a class property).** If \( i \leftrightarrow j \) and \( i \) is positive recurrent, then \( j \) is also positive recurrent.

**Proof.** Let \( i \) be positive recurrent. Then \( \mu_{ii} > 0 \) and \( i \to j \) implies that \( \mathbb{E}_i \left[ \sum_{k=1}^{\infty} 1(X_k = j) \right] > 0 \) because there is a path with positive probability from \( i \) to \( j \) without visiting \( i \). Thus,

\[
\lim_{n \to \infty} P^n(i, j) = \frac{\mathbb{E}_i \left[ \sum_{k=1}^{\infty} 1(X_k = j) \right]}{\mu_{ii}} = \frac{1}{\mu_{jj}} > 0.
\]

Also, \( j \) is recurrent by Prop 3.1.2. Hence, \( \lim_{n \to \infty} P^n(j, j) = 1/\mu_{jj} > 0 \). This means that \( \mu_{jj} < \infty \). Thus, \( j \) is also a positive recurrent state. \( \square \)

**Proposition 3.2.3.** If \( S \) is finite and irreducible, then \( S \) is positive recurrent.

**Proof.** All the states in a finite state Markov chain cannot be transient because at least one state will be visited infinitely often with probability 1. Then, since, the Markov chain is irreducible, all states must be either null recurrent or positive recurrent (Prop 3.2.2). Suppose that all the states are null recurrent. Then we have

\[
\lim_{n \to \infty} P^n(i, j) = 0 \forall i, j \in S \quad (3.1)
\]
But, \( \sum_j P^n(i, j) = 1, \forall n. \) Therefore \( \lim_{n \to \infty} \sum_j P^n(i, j) = \sum_j \lim_{n \to \infty} P^n(i, j) = 1. \) This shows that Eq 3.1 cannot be true. Thus, all states are positive recurrent.

Let \( A \) be a subset of state space which is a closed communicating class. Let \( f_i(A) \) be the probability that Markov chain enters \( A \) starting in state \( i \). Then,

\[
f_i(A) = \sum_{j \in A} P(i, j) + \sum_{j \in A^c} P(i, j)f_j(A)
\]

Here, we need to sum over only transient states in the second summand because \( f_j \) for \( j \in A^c \) and not transient is zero. This will result in a set of linear equations in \( f_i(A) \) if we want to compute \( \lim_{n \to \infty} \mathbb{P}_i \{ X_n = j \} \) for \( j \in A \). We can regard \( A \) as an irreducible Markov chain and compute its stationary distribution \( \pi_A(j) \forall j \in A \). Then, \( \lim_{n \to \infty} \mathbb{P}_i \{ X_n = j \} = f_i(A)\pi_A(j). \) In general for a Markov chain, there is one unique stationary distribution corresponding to every closed communicating class if the class is positive recurrent. Any convex combination of these stationary distributions will also be a stationary distribution.

Consider the example in previous class: \( S = \{0, 1, 2, 3, 4\}. \)

\[
P = \begin{bmatrix}
0.2 & 0.3 & 0 & 0.5 & 0 \\
0 & 0.3 & 0.7 & 0 & 0 \\
0 & 0.5 & 0.5 & 0 & 0 \\
0 & 0 & 0 & 0.3 & 0.7 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

In this example, \( A = \{1, 2\} \) and \( B = \{4\} \) are two closed communicating classes. \( \{0, 3\} \) are transient states. \( \{3, 4\} \) is a closed set but not a closed communicating class. Considering \( A \) as an irreducible Markov chain, \( \pi_A(1) = 7/13 \) and \( \pi_A(2) = 6/13 \). \( f_1(A) = f_2(A) = 1 \) and \( f_3(A) = f_4(A) = 0 \). \( f_0(A) = 0.2f_0(A) + 0.3f_1(A) + 0.2f_2(A) + 0.5f_3(A). \) The stationary distribution corresponding to \( A \) is \( [0.7/13, 6/13, 0, 0] \) and the stationary distribution corresponding to \( B \) is \( [0, 0, 0, 0, 1] \). All the convex combinations of these form the whole class of stationary distributions for the Markov chain.
3.4 Tests for transience, null recurrence and positive recurrence

**Theorem 3.4.1.** Let $S$ be irreducible and $f : S \to \mathbb{R}$.

1. Let $f(i) \to \infty$ as $i \to \infty$. If $E[f(X_1)|X_0 = i] \leq f(i)$ for all $i$ outside a finite set $S_0 \subset S$, then the Markov chain is recurrent.

2. Let $f : S \to \mathbb{R}^+$ and $S_0 \subset S$ be finite. If
   
   (a) $E[f(X_1)|X_0 = i] < \infty \forall i$,
   
   (b) For $i \not\in S_0$, $E[f(X_1)|X_0 = i] - f(i) < -\epsilon$ for some $\epsilon > 0$

   the Markov chain is positive recurrent.

3. Let $f : S \to \mathbb{R}$ and $S_0 \subset S$ be finite. The Markov chain is transient if

   (a) $f$ is bounded and $E[f(X_1)|X_0 = i] \geq f(i) \forall i \in S_0$

   (b) $f(i) > f(j) \forall j \in S_0$, for some $i \not\in S_0$

The proof of this theorem requires Martingale methods. Thus we will prove it after we have studied Martingales.

**Example:** Consider a slotted queuing system in which one service time is equal to one slot. Let $q_k$ be the queue length at the end of $k$th slot. Let $A_k$ denote the number of arrivals in the $k$th slot. Then, $q_{k+1} = (q_k - 1)^+ + A_k$ and for $i > 0$

$$P\{q_{k+1} = j|q_k = i, q_{k-1} = i_{k-1}, \ldots q_0 = i_0\} = P\{A_k = j - i - 1\}.$$ 

Thus, $\{q_k\}$ is a Markov chain with state space $S = \{0, 1, 2, \ldots\}$. If $P\{A_1 \leq 1\} > 0$, every state is aperiodic. If further, $P\{A_1 > 1\} > 0$, the Markov chain is also irreducible.

Consider $f(i) = i$. Let $S_0 = \{0\}$. For $i > 0$, $E[f(q_1)|q_0 = i] - f(i) = E[A_1] - 1$. Thus, we see from case (1) in Theorem 3.4.1, if $E[A_1] \leq 1$, $\{q_k\}$ is recurrent. From (2) in Theorem 3.4.1, we have positive recurrence if $E[A_k] < 1$.

3.5 Reversible Markov Chains

We can easily check that

$$P\{X_{k-1} = j|X_k = i, X_{k+1} = i_{k+1}, X_{k+2} = i_{k+2}, \ldots\} = P\{X_{k-1} = j|X_k = i\}.$$
Thus, the reversed Markov chain is also a Markov chain. For an irreducible stationary Markov chain \( \{X_k\} \) with stationary distribution \( \pi \),

\[
\mathbb{P}\{X_{k-1} = j | X_k = i, X_{k+1} = i_{k+1}, X_{k+2} = i_{k+2}, \ldots \} = \frac{\mathbb{P}\{X_{k-1} = j, X_k = i \}}{\mathbb{P}\{X_k = i \}} = \frac{\mathbb{P}\{X_k = i | X_{k-1} = j \} \mathbb{P}\{X_{k-1} = j \}}{\pi(i)} = \frac{P(j,i)\pi(j)}{\pi(i)}.
\]

Let us define

\[
P^*(i,j) = \frac{P(j,i)\pi(j)}{\pi(i)}
\]

\(P^*\) is the transition probability of the reversed Markov chain.

**Reversible Markov Chain:** The stationary irreducible Markov chain \( X_k \) is called *reversible* if \( P^*(i,j) = P(i,j) \). In the other words, for a reversible Markov chain, we have, \( P(i,j)\pi(i) = P(j,i)\pi(j) \).

**Proposition 3.5.1 (Test for reversibility).** For an irreducible Markov chain with stationary distribution \( \pi \), for all paths \( i \rightarrow i_1 \rightarrow i_2 \cdots \rightarrow i_k \rightarrow i \),

\[
\mathbb{P}_\pi\{i \rightarrow i_1 \rightarrow i_2 \cdots \rightarrow i_k \rightarrow i\} = \mathbb{P}_\pi\{i \rightarrow i_k \rightarrow i_{k-1} \cdots \rightarrow i_1\}
\]

under stationarity is a necessary and sufficient condition for reversibility of the Markov chain.

**Proof.** Necessity: Assume \( P^* = P \). Then,

\[
\mathbb{P}_\pi\{i \rightarrow i_1 \rightarrow i_2 \cdots \rightarrow i_k \rightarrow i\} = \pi(i)P(i,i_1)P(i_1,i_2)\ldots P(i_{k-1},i_k)P(i_k,i) = P(i_1,i)\pi(i_1)P(i_1,i_2)\ldots P(i_{k-1},i_k)P(i_k,i) = P(i_1,i)P(i_2,i_1)\pi(i_2)\ldots P(i_{k-1},i_k)P(i_k,i) = \cdots = P(i_1,i)P(i_2,i_1)\ldots P(i_k,i_{k-1})\pi(i_k)P(i_k,i) = P(i_1,i)P(i_2,i_1)\ldots P(i_k,i_{k-1})P(i_{k-1},i)\pi(i) = \mathbb{P}_\pi\{i \rightarrow i_k \rightarrow i_{k-1} \cdots \rightarrow i_1\}.
\]

Sufficiency: Consider the path \( i \rightarrow i_1 \rightarrow i_2 \cdots \rightarrow i_k \rightarrow j \rightarrow i \) and its reverse path. Then,

\[
\sum_{i_1,i_2,\ldots,i_k} \mathbb{P}_\pi\{i \rightarrow i_1 \rightarrow i_2 \cdots \rightarrow i_k \rightarrow j \rightarrow i\} = \sum_{i_1,i_2,\ldots,i_k} \mathbb{P}_\pi\{i \rightarrow j \rightarrow i_k \rightarrow i_{k-1} \cdots \rightarrow i_1\}
\]

Thus,

\[
P^k(i,j)P(j,i) = P(i,j)P^k(j,i)
\]

Now, taking the limit as \( k \rightarrow \infty \), we get \( \pi(j)P(j,i) = P(i,j)\pi(i) \), which is \( P^* = P \).  

The idea of time reversal of Markov chains and reversibility will be considered in continuous time as well. It will be extensively used at the end of the course to study queuing networks.
3.6 Example: M/GI/1 queue

Consider an $M/GI/1$ queue. Let $\lambda$ be the Poisson arrival rate, $S_k$ the service time of the $k^{th}$ customer and $E[S]$ the mean service time. Let

- $q_k =$ queue length just after the $k^{th}$ departure.
- $\hat{q}_k =$ queue length just before the $k^{th}$ arrival.
- $q_t =$ queue length at arbitrary time $t$.
- $W_k =$ waiting time of the $k^{th}$ customer.

The process $\{q_k\}$ satisfies $q_{k+1} = (q_k - 1)^+ + A_{k+1}$ where $A_{k+1}$ is the number of arrivals during the service of the $(k+1)^{th}$ customer. Since $\{A_k\}$ is i.i.d., $\{q_k\}$ is a Markov chain. The state space $S = \{0, 1, 2, \ldots\}$ and it is easy to check that it is aperiodic and irreducible. By choosing $f(t) = i$, we can deduce using the test for positive recurrence that $\{q_k\}$ is positive recurrent when $E[A_1] = \lambda E[S] < 1$. Thus, we conclude that when $\lambda E[S] < 1$, the process $\{q_k\}$ has a stationary distribution. We will see later that when $\lambda E[S] = 1$, $\{q_k\}$ is recurrent and when $\lambda E[S] > 1$, it is transient.

The process $\{\hat{q}_k\}$ however, is not a Markov chain if $S_k$ is not exponentially distributed. But $\{\hat{q}_k\}$ is a regenerative process with regeneration epochs occurring when $k^{th}$ arrival sees an empty queue. That is, regenerative epochs for $\{\hat{q}_k\}$ occur when $\hat{q}_k = 0$. Let $\hat{\tau}$ be the regeneration length of $\{\hat{q}_k\}$.

To obtain the conditions for existence for stationary distribution for $\{\hat{q}_k\}$, we can relate it to the process $\{q_k\}$. The process $\{q_k\}$ is also a regenerative process with regeneration epochs occurring when a departure leaves behind an empty queue. That is, the regeneration epochs correspond to $q_k = 0$. Let $\tau$ be the regeneration time of $\{q_k\}$. Now, we can see that $\tau = \hat{\tau}$. Since, $\{q_k\}$ has stationary distribution when $\lambda E[S] < 1$, $E[\tau] < \infty$. So, $E[\hat{\tau}] < \infty$. Therefore, the process $\{\hat{q}_k\}$ is stationary iff $\{q_k\}$ is stationary.

The process $\{W_k\}$ also has the same regeneration epochs as $\{q_k\}$ and hence has unique stationary distribution when $\lambda E[S] < 1$.

The process $\{q_k\}$ is also a regenerative process with regeneration epochs the arrival times that see an empty queue. Let $T$ be a regeneration length of $\{q_k\}$. Then, considering one regenerative cycle, we can write $T = \sum_{i=1} T_i$ where $T_i$ is the inter-arrival time between $(k-1)^{th}$ and $k^{th}$ arrival. Now, $\hat{T}$, which is equal to the number of services in one regeneration cycle is a stopping time for $\{a_k, S_k\}$ where $S_k$ is the service time of the $k^{th}$ arrival. Thus, we can use Wald’s lemma to conclude that $E[T] = E[\hat{T}] E[a_1]$. Thus, $\{q_k\}$ also has stationary distribution whenever $\lambda E[S] < 1$.

3.7 Rate of convergence to the stationary distribution

The following results are stated without proofs.
Definition 3.7.1 (Total Variation distance). The total variation distance between two probability distributions $\mu$ and $\pi$ on $S$,

$$||\mu - \pi||_{TV} = \frac{1}{2} \sum_{x \in S} (\mu(x) - \pi(x)).$$

The following results have been classically known. We consider an irreducible Markov chain.

- If $E[\tau^\alpha] < \infty$ for some $\alpha > 1$, then $||X_k - \pi||_{TV} < c_1 k^{-\alpha + 1}$.
- If $E[\beta^\gamma] < \infty$ for some $\beta > 1$, then $||X_k - \pi||_{TV} \leq \exp(-\lambda k)$ for some $0 < \lambda < \beta$.

These results have been extensively used in the literature to obtain rates of convergence to stationary distributions for different queueing systems.

Now, we consider a finite state space $S$. Let

$$K_n^x(y) = \frac{P_n(x,y)}{\pi(y)}$$

where $P$ is the transition probability matrix of the Markov chain. Then, $K_n^x(y) \to 1$ as $n \to \infty \forall x, y \in S$.

Definition 3.7.2. $L^p$ distance between distributions $P_n(x,\cdot)$ and $\pi$,

$$||K_n^x - 1||_{p,\pi}^p = \sum_{y \in S} |K_n^x(y) - 1|^p \pi(y) \text{ for } 1 \leq p < \infty.$$

Also, $L^\infty(\nu,\mu) = \sup_{x \in S} |\nu(x) - \mu(x)|$.

The following are known.

$$||v - \mu||_{TV} = \frac{1}{2} ||\frac{v}{\mu} - 1||_{1,\mu} \leq \frac{1}{2} ||\frac{v}{\mu} - 1||_{2,\mu}. \quad (3.2)$$

Definition 3.7.3 (Mixing times).

$$\tau_1(\epsilon) = \min \{ n : \sup_x ||P_n^r(x,\cdot) - \pi||_{TV} \leq \epsilon \}.$$
$$\tau_2(\epsilon) = \min \{ n : \sup_x ||K_n^r(x,\cdot) - 1||_{2,\pi} \leq \epsilon \}.$$
$$\tau_\infty(\epsilon) = \min \{ n : \sup_x |P_n(x,\cdot) - \pi|_\infty \leq \epsilon \}.$$

Let $\pi_* = \min_x \pi(x)$ and

$$||P^*|| = \sup_{f: \mathbb{F} \to \mathbb{R}, \mathbb{E}[f] = 0} \frac{||P^* f||_2}{||f||_2},$$

where $P^*$ is the complex conjugate of $P$.

Proposition 3.7.4.

$$\tau_2(\epsilon) \leq \frac{1}{1 - ||P^*||} \log \left( \frac{1}{\epsilon \sqrt{\pi_*}} \right).$$
Proposition 3.7.5.

$$\tau_2(\varepsilon) \leq \frac{2}{\lambda_{pp^*}} \log \left( \frac{1}{\varepsilon \sqrt{\pi_*}} \right)$$

where $\lambda_{pp^*} = 1 - \lambda$ and $\lambda$ is the largest eigenvalue of $PP^*$ less than 1.

These results provide an exponential rate of convergence to the stationary distribution for a finite state Markov chain. From $\tau_2(\varepsilon)$ we get an upper bound on $\tau_1(\varepsilon)$ through Eq (3.2). We also have

$$\tau_2(\varepsilon) \leq \tau_\infty(\varepsilon) \leq \tau_2 \left( \varepsilon \sqrt{\frac{\pi_*}{1 - \pi_*}} \right).$$

If the Markov chain is reversible, the upper bound can be tightened to $2\tau_2(\sqrt{\varepsilon})$. Thus, for most applications getting $\tau_2(\varepsilon)$ is sufficient.

Mixing times have recently been used in Markov chain Monte Carlo (MCMC) algorithms, random graphs and many other applications.
3.8 Problems

Problem 1: There are a total of $N$ balls in urns $A$ and $B$. At step $k$, one of the $N$ balls is picked at random (with probability $1/N$). Then, one of the urns $A$ or $B$ is chosen. The probability of picking urn $A$ is $p$. The ball picked is put in the chosen urn. Let $X_n$ denote the number of balls in urn $A$ after step $n$. Show that $\{X_n\}$ is a Markov chain. Determine its state space and transition probability matrix. Find if it is irreducible or not. Find $\lim_{n \to \infty} P^n(i,j)$ for all $i$ and $j$.

Problem 2: Show that for a finite state aperiodic irreducible Markov chain, $P^n(i,j) > 0 \forall i,j$ for all $n$ large enough.

Problem 3: Let a Markov chain has

1. Show that if state $j$ can be reached from state $i$, it can be reached in almost $r - 1$ steps.
2. If $j$ is a recurrent state, show that $\exists \alpha$ such that for $n > r$, the probability that first return to state $j$ (from state $j$) occurs after $n$ transitions is less than $\alpha^n$.

Problem 4: Let $P$ be the transition probability matrix with additional requirement that $\sum P(i,j) = 1$ (such a $P$ is called a doubly stochastic matrix). Then, show that if $P$ is finite state irreducible, then its stationary probability satisfies $\pi(i) = \pi(j) \forall i,j$.

Problem 5: Consider a Markov chain with state space $E = \{0,1,2,3,4,5\}$ and transition probability matrix

\[
P = \begin{bmatrix}
1/3 & 2/3 & 0 & 0 & 0 & 0 \\
2/3 & 1/3 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/4 & 3/4 & 0 & 0 \\
0 & 0 & 1/5 & 4/5 & 0 & 0 \\
1/4 & 0 & 1/4 & 0 & 1/4 & 1/4 \\
1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6
\end{bmatrix}
\]

Find all the closed sets. Find all the transient states. Calculate $\lim_{n \to \infty} P^n(5,i), i = \{0,1,2,3,4,5\}$.

Problem 6: For a Markov chain prove that

1. $P[X_n = j|X_{n_1} = i_1, X_{n_2} = i_2, \ldots, X_{n_k} = i_k] = P[X_n = j|X_{n_k} = i_k]$ whenever $n_1 < n_2 < \cdots < n_k < n$.
2. $P[X_k = i_k|X_j = i_j \forall j \neq k] = P[X_k = i_k|X_{k-1} = i_{k-1}, X_{k+1} = i_{k+1}]$.

Problem 7: Consider a recurrent Markov chain starting at state 0. Let $m_i$ denote the expected number of time periods it spends in state $i$ before returning to state 0. Use Wald’s equation to show that $m_j = \sum_{i} m_i P_{ij}$, $j > 0$, $m_0 = 1$.

Problem 8: Let $X_1, X_2, \ldots$ be independent r.v.s such that $P[X_j = j] = \alpha_j, j \geq 1$. Say that a record occurs at time $n$ if $X_n > \max(X_1, X_2, \ldots, X_{n-1})$ where $X_0 = -\infty$. If a record occurs at time $n$, $X_n$ is called a record value. Let $R_i$ denote the $i^{th}$ record value.

1. Argue that $\{R_i, i \geq 1\}$ is a Markov chain and compute its transition probabilities.
2. Let $T_i$ denote that time between $i^{th}$ and $(i+1)^{th}$ record. Is $\{T_i\}$ a Markov chain? What about $\{R_i, T_i\}$? Compute transition probabilities wherever appropriate.
3. Let $S_n = \sum_{i=1}^{n} T_i$, $n \geq 1$. Argue that $\{S_n\}$ is a Markov chain. Compute its transition probability matrix.
Chapter 4

Continuous-Time Markov Chains

Lecture 13

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

4.1 Introduction

Consider a continuous-time stochastic process \{X(t), t \geq 0\} taking values in a countable (can be finite) set $S$. A process \{X(t), t \geq 0\} is a Continuous-Time Markov Chain (CTMC) if for all $s, t \geq 0$, and $i, j, x(u) \in S, 0 \leq u \leq s$,

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

If in addition,

$$P\{X(t+s) = j | X(s) = i\} \triangleq P_t(i,j)$$

is independent of $s$, then the continuous-time Markov chain is said to have stationary transition probabilities. All the Markov chains we consider will be assumed to have stationary transition probabilities.

Further, we will restrict to pure jump processes: the sample paths of the process are piecewise constant, right continuous. We will see that such versions of the processes can usually be constructed.

By Markov property, for a pure jump process the sojourn time $T_i$ in state $i$ satisfies,

$$P\{T_i > s + t | T_i > s\} = P\{T_i > t\}$$

for all $s,t \geq 0$. Hence, the random variable $T_i$ is memoryless and must thus be exponentially distributed, say with parameter $\lambda_i$. If $\lambda_i = 0$ then

$$P[T_i \geq t] = e^{-\lambda t} = 1$$

for all $t$ and the state $i$ is called absorbing. If $\lambda_i = \infty$ then

$$P[T_i \geq t] = 0$$
for all $t$ and the state $i$ is called instantaneous. We will assume $\lambda_i < \infty$ for all states $i$. For a pure jump process this will hold.

For a Markov jump process,

- the amount of time it spends in a state $i$ before making a transition into a different state is exponentially distributed with mean, say, $\frac{1}{\lambda_i}$, and
- when the process leaves state $i$, it next enters state $j$ with probability $P_{ij}$.

$P_{ij}$ satisfies, for $i$ not an absorbing state,

$$P_{ii} = 0, \quad \sum_j P_{ij} = 1, \quad \forall i,$$

and if $i$ is an absorbing state, $P_{ii} = 1$.

In other words, a Continuous-Time Markov Chain (CTMC) is a stochastic process that moves from state to state in accordance with a (discrete-time) Markov chain, but is such that the amount of time it spends in each state, before proceeding to the next state, is exponentially distributed independently of the next state visited.

### 4.2 Strong Markov property, Minimal construction

A random variable $\tau$ is a stopping time with minimal construction if the event $\{\tau \leq t\}$ can be determined completely by the collection $\{X(u) : u \leq t\}$. A stochastic process $X$ has strong Markov property if for any almost surely finite stopping time $\tau$,

$$P[X(\tau + s) = j|X(u), u \leq \tau, X(\tau) = i] = P[X(\tau + s) = j|X(\tau) = i] = P_{s(i,j)}.$$

**Lemma 4.2.1.** A continuous time jump Markov chain $X$ has the strong Markov property.

**Proof.** Let $\tau$ be an almost surely finite stopping time with conditional distribution $F$ on the collection of events $\{X(u) : u \leq s\}$. Then,

$$\Pr\{X(\tau + s) = j|X(u), u \leq \tau, X_\tau = i\} = \int_0^\tau dF(t) \Pr\{X(t + s) = j|X(u), u \leq t, \tau = t, X_\tau = i\} = P_{s(i,j)}.$$

We give a minimal construction of a CTMC with given $\lambda_i, s$ and $P_{ij,s}$. Construct a DTMC $Y_0, Y_1, \ldots$, with parameters $P_{ij}$ and construct exponential random variables $T_1, T_2, \ldots$ independent of each other, where $T_n \sim \exp(\lambda(Y_n))$.

Let $S_n = \sum_{k=1}^n T_k$ and $S_0 = 0$.

Define $X_t = Y_j$ if $S_j \leq t < S_{j+1}$. If $\omega(\Delta) \triangleq \sup_{s \leq s_0} < \infty$, then $X_t = \Delta$ for $t \geq \omega(\Delta)$, where $\Delta$ is an element which is not in the state space $S$. On the extended state space $S \cup \{\Delta\}$ we define $P_t(\Delta, \Delta) = 1$ for all $t > 0$ and $P(\Delta, \Delta) = 1$. There are other possibilities to define the MC after $\omega(\Delta)$ but the above construction makes $P_t(i,j), i, j \in S$ minimal. When $\omega(\Delta) < \infty$, we say the MC has explosion.

One can show that for any initial condition $i$,

$$\mathbb{R}(\omega) = \{\omega : \omega(\Delta) < \infty\} = \{\omega : \sum_{k=0}^{\infty} \frac{1}{\lambda(Y_k(\omega))} < \infty\} \quad a.s$$

Easier to verify the conditions for non explosion of the MC are the following.

42
Lemma 4.2.2. Any of the following conditions are sufficient for $P\{\omega : \omega(\Delta) < \infty\} = 0$:

1. $\sup_i \lambda (i) < \infty$,
2. $S$ is finite,
3. $\{Y_n\}$ is recurrent.

Proof. If $\omega(\Delta) < \infty$, then $\frac{1}{\lambda(Y_k(\omega))} \to 0$ as $k \to \infty$.

If (1) is true, $\lambda(Y_k(\omega)) \leq \bar{\lambda}$, therefore, $\lambda(Y_k(\omega)) \to \infty$ is not possible.

If (2) is true, since $\lambda(i) < \infty \forall i$, $\sup_i \lambda(i) < \infty$.

If (3) is true $Y_k(\omega) = i$ for an infinite number of $k$ w.p.1. Therefore,

$$\lambda(Y_k(\omega)) \not \to \infty, \quad a.s.,$$

$$P[\lambda(Y_k(\omega)) \to \infty] = 0.$$
4.3 Chapman Kolmogorov equations

We define the generator matrix $Q$ for MC $\{X_t\}$ as $Q_{ii} = -\lambda_i$ and $Q_{ij} = \lambda_i P_{ij}$ for $j \neq i$. Also, the fact that MC stays in state $i$ with $\exp(\lambda_i)$, implies that

$$\frac{1 - P_i(t)}{t} \to \lambda_i \quad \text{as} \quad t \to 0,$$

and then,

$$\lim_{t \downarrow 0} \frac{P_{ij}(t)}{t} = \lim_{t \downarrow 0} \frac{1 - P_i(t)}{t} P_{ij} = Q_{ij} \quad \text{for all} \quad i \neq j.$$

Theorem 4.3.1 (Backward equation). For a homogeneous CTMC with transition matrix $P(t)$ and generator matrix $Q$, for the minimal construction,

$$\frac{dP(t)}{dt} = QP(t), \quad t \geq 0.$$

Proof. Using semigroup property of transition probability matrix $P(t)$ for a homogeneous CTMC, we can write

$$\frac{P(t + h) - P(t)}{h} = \frac{(P(h) - I)}{h} P(t).$$

Taking limits $h \downarrow 0$ and exchanging limits and summation, justified below, on the RHS we get

$$\frac{dP_{ij}(t)}{dt} = \sum_{k \neq i} Q_{ik} P_{kj}(t) - \lambda_i P_{ij}(t).$$

Now we justify the exchange of limit and summation. For any finite subset $F \subset S$, we have

$$\liminf_{h \downarrow 0} \sum_{k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \geq \sum_{k \in F \setminus \{i\}} \liminf_{h \downarrow 0} \frac{P_{ik}(h)}{h} P_{kj}(t) = \sum_{k \in F \setminus \{i\}} \lambda_k P_{kj}(t).$$

Since, above is true for any finite set $F \subset E$, taking supremum over increasing sets $F$, we get the lower bound. For the upper bound, we observe for any finite subset $F \subseteq E$

$$\limsup_{h \downarrow 0} \sum_{k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \leq \limsup_{h \downarrow 0} \left( \sum_{k \in F \setminus \{i\}} \frac{P_{ik}(h)}{h} P_{kj}(t) + \sum_{k \in F \setminus \{i\}} \lambda_k \frac{1 - P_{kj}(h)}{h} \right)$$

$$= \limsup_{h \downarrow 0} \left( \sum_{k \in F \setminus \{i\}} \frac{P_{ik}(h)}{h} P_{kj}(t) + \frac{1 - P_{kj}(h)}{h} \right) - \sum_{k \in F \setminus \{i\}} \frac{P_{ik}(h)}{h}$$

$$= \sum_{k \in F \setminus \{i\}} Q_{ik} P_{kj}(t) + \left( \lambda_i - \sum_{k \in F \setminus \{i\}} Q_{ik} \right).$$
Now take $F \not\supset S$, then the term $(\lambda_i - \sum_{k \in F \setminus \{i\}} Q_{ik})$ goes to zero.

Theorem 4.3.2 (Forward equation). For a homogeneous CTMC with transition matrix $P(t)$ and generator matrix $Q$, we have for the minimal construction,

$$\frac{dP(t)}{dt} = P(t)Q.$$  

Proof. Using semigroup property of transition probability matrix $P(t)$ for a homogeneous CTMC, we can write

$$P(t+h) - P(t) \approx P(t)(P(h) - I).$$

Taking limits $h \downarrow 0$, if we can justify the interchange of limit and summation on RHS,

$$\frac{dP_{ij}(t)}{dt} = \sum_{k \neq j} P_{ik}(t)Q_{kj} - \lambda_j P_{ij}(t).$$

Corollary 4.3.3. For a homogeneous CTMC with finite state space $E$, the transition matrix $P(t)$ and generator matrix $Q$, we have

$$P(t) = e^{tQ} = I + \sum_{n \in \mathbb{N}} \frac{t^n Q^n}{n!}, \quad t \geq 0.$$
4.4 Irreducibility and Recurrence

Let \( \{X_t\} \) be a Markov chain, \( \{Y_0, Y_1, \ldots\} \) be its jump Markov chain, \( \{T_0, T_1, \ldots\} \) the sojourn times. Let \( P \) be the transition matrix for \( \{Y_n\} \), \( \hat{P} \to P \).

- \( i \to j \) in \( \{Y_n\} \) if \( \exists n_1 \) s.t. \( P_{ij}^{n_1} > 0 \).
- \( i \to j \) in \( \{X_t\} \) if \( \exists t_1 > 0 \) s.t. \( P_t(i, j) > 0 \).

Since \( P_t(i, i) \to 1 \) as \( t \to 0 \) for all \( i \), \( P_t(i, i) > \varepsilon \) for all \( t \in [0, \delta] \), for some \( \varepsilon > 0 \) and \( \delta > 0 \). Then \( P_t(i, i) \geq (P_t(i, i))^n \) implies that \( P_t(i, i) > 0 \) for all \( t \in [0, \delta n] \) and hence \( P_t(i, i) > 0 \) for all \( t \). This also follows directly from \( P_t(i, i) \geq P[T_0 > t] > 0 \) for \( T_0 \sim \exp(\lambda) \).

**Proposition 4.4.1.** The following statements are equivalent, under minimal construction

1. \( i \to j \) in \( \{Y_n\} \).
2. \( i \to j \) in \( \{X_t\} \).
3. \( P_t(i, j) > 0 \), \( \forall t > 0 \).

**Proof.** (1) \( \Rightarrow \) (2). (3)

\[ \exists n_1 \text{ s.t. } P_{n_1}(i, j) > 0, \text{ therefore } i \xrightarrow{T_{i_1}} i_1 \xrightarrow{T_{i_2}} \cdots \xrightarrow{T_{i_{n_1-1}}} j \text{ where } i, i_1, \ldots, i_{n_1-1} \text{ are not absorbing states.} \]

\[ T_0 \sim \exp(\lambda_0) \text{ where } \infty > \lambda(i) = \lambda(1), \ldots, \lambda(i_{n_1-1}) > 0, \lambda(j) < \infty. \]

\[ T_0 \sim \exp(\lambda_0) \text{ where } \infty > \lambda_ii = \lambda q_{ii} / \theta_0 > 0. \]

\[ P_t(i, i) \geq P|[\sum_{k=0}^{n-1} T_k] < \frac{1}{2}, T(j) > \frac{1}{2}] > 0, \quad \forall t > 0. \]

Hence \( P_t(i, j) > 0 \).

(3) \( \Rightarrow \) (2) is clear from the definition itself.

(2) \( \Rightarrow \) (1)

\[ P(\omega(\Delta) > t_1) > P_t(i, j) > 0. \text{ Therefore, there is a finite path } i \to i_1 \to \cdots \to i_n \to j, \text{ such that,} \]

\[ P(i \to i_1 \to \cdots \to i_n \to j) > 0, P_t(i_1, \ldots, P_{i_{n-1}i_n}, P_{it} > 0. \text{ This implies } i \to j \text{ in Markov chain } \{Y_n\}. \]

In particular this also implies that closed irreducible classes are same in \( \{Y_n\} \) and \( \{X_t\} \) and \( \{Y_n\} \) is irreducible \( \Leftrightarrow \{X_t\} \) is irreducible.

Let \( \omega(i) = \inf \{ t > 0, \text{ s.t. } X_t = i \text{ and } \lim_{t \to \infty} X_t \neq i \} \). This is the first time MC visits state \( i \), after exiting from \( i \) if \( X_0 = i \).

**Definition 4.4.2.** A state \( i \) is transient if \( P(\omega(i) < \infty) < 1 \). It is recurrent if \( P(\omega(i) < \infty) = 1 \). A recurrent chain is positive recurrent if \( \mathbb{E}[\omega(i)] < \infty \), otherwise null recurrent.

**Theorem 4.4.3.** State \( i \) is recurrent for \( \{Y_n\} \) iff it is for \( \{X_t\} \) in a minimal construction.
Proof. Let \( i \) be recurrent for \( \{ Y_k \} \). Let \( \tau(i) = \inf \{ k > 0 \text{ s.t. } Y_k = i \} \).
Then \( P[\tau(i) < \infty] = 1 \) and \( \omega(i) = \sum_{k=0}^{\tau(i)-1} T_k \), where \( T_k \) is the sojourn time of \( X_t \) in state \( Y_k \) Then

\[
P_i[\omega(i) < \infty] = P_i \left[ \sum_{k=0}^{\tau(i)-1} T_k < \infty \right]
= \sum_{n=1}^{\infty} P_i \left[ \sum_{k=0}^{n-1} T_k < \infty | \tau(i) = n \right] P[\tau(i) = n],
\]

But

\[
P_i \left[ \sum_{k=0}^{n-1} T_k < \infty | \tau(i) = n \right] = \sum_{i_1, i_2, \ldots, i_{n-1}} P_i \left[ \sum_{k=0}^{n-1} T_k < \infty | \tau(i) = n, Y_j = i_j, j = 1, 2, \ldots, n-1 \right]
= 1.
\]

Thus

\[
P_i[\omega(i) < \infty] = 1.
\]

Therefore, \( P_i[\tau(i) < \infty] = P_i[\omega(i) < \infty] = 1 \).

The above theorem implies that \( i \) is transient in \( \{ X_t \} \) iff it is in \( \{ Y_k \} \).

If \( \{ X_t \} \) is irreducible then we have seen that \( Y_k \) is irreducible. If \( i \) is recurrent/transient in \( \{ X_t \} \) then so is it in \( \{ Y_k \} \). Then in \( \{ Y_k \} \) every state is recurrent/transient. Thus it is so in \( \{ X_t \} \). Therefore, in \( \{ X_t \} \) also recurrence/transience is a class property. However positive recurrence of \( \{ Y_k \} \) does not imply \( \{ X_t \} \) and vice versa.

Let \( \{ X_t \} \) be irreducible and \( i \) be recurrent in \( \{ X_t \} \). We take visit times to \( i \) as regenerative epochs with \( \omega(i) \) as a regeneration length. Then from delayed regenerative process limit theorem for a bounded function \( f \),

\[
E_i[f(X_t)] \to E[f(X_0)],
\]

\[\text{where } \pi \text{ is a stationary measure for } X = \{ X_t \}. \text{Taking } X_0 = i, \]

\[
E_\pi[f(X_0)] = \frac{E_i[f_0] f(X_0) dt}{E_i[\omega(i)]}.
\]

If \( i \) is null recurrent then \( E_i[\omega(i)] = \infty \) and \( E_i[f(X_0)] = 0 \). If \( i \) is positive recurrent then \( \pi \) can be normalized to get a stationary (unique) distribution for \( X \). Also, then

\[\pi(j) = \frac{E_i[f_0 1_{(X_n = j)} dt]}{E_i[\omega(i)]} > 0, \ \forall j. \quad (4.1)\]

Furthermore, as \( t \to \infty \),

\[
P_t(i, j) \to \pi(j)
\]

and also, \( P_t(k, j) \to \pi(j) \) for all \( k \) and \( j \). If \( i \) is null recurrent, then for all \( j \), as \( t \to \infty \)

\[
P_t(i, j) \to 0, \quad P_t(k, j) \to 0 \quad \forall j, k.
\]
**Proposition 4.4.4.** If \( i \) is transient then \( P_t(j,i) \to 0 \) as \( t \to \infty \) for any \( j \in S \).

**Proof.** Now \( P_t[\omega(i) < \infty] = p < 1 \). Let \( N \) be number of times state \( i \) is visited. Then \( P[N = n] = p^n(1-p) \) and \( P[N < \infty] = 1 \). Let \( \tilde{\omega}(i) \equiv [\omega(i) | \omega(i) < \infty] \) a when \( X_0 = i \), and \( Z_k \) be i.i.d. \( \exp(\lambda_k) \).

Let \( \tilde{\omega}_k(i) \sim i.i.d. \tilde{\omega}(i) \). Then
\[
P_t(i,i) \leq P_t[\sum_{k=1}^{N} (\tilde{\omega}_k(i) + Z_k) > t] \to 0 \quad \text{as} \quad t \to \infty.
\]

Also, for \( j \neq i \),
\[
P_t(j,i) \leq P_t[\tilde{\omega}(i) + \sum_{k=1}^{N} (\tilde{\omega}_k(i) + Z_k) > t] \to 0,
\]
where \( \tilde{\omega}(i) \equiv [\omega(i) | \omega(i) < \infty] \) when \( X_0 = j \). \( \square \)

Now assume \( X \) is recurrent, irreducible, and \( Y = \{Y_k\} \) is positive recurrent with the unique stationary distribution \( \mu \). Let \( N(j) \) be the number of visits to state \( j \) between two visits of \( i \). Let \( \tau(i) \) be the intervisit time to state \( i \) in \( \{Y_k\} \). Then
\[
\mu(j) = \mathbb{E}[N(j)]/\mathbb{E}[\tau(i)].
\]

Also, from 4.1
\[
\frac{\mathbb{E}[\sum_{j=1}^{N(j)} T_k(j)]}{\mathbb{E}[\sum_{k=1}^{N(j)} T_k(j)]} = \frac{\mathbb{E}[\sum_{j=1}^{N(j)} T_k(j)]}{\mathbb{E}[\sum_{k=1}^{N(j)} T_k(j)]} \quad (4.2)
\]
where \( T_k(j) \) is the sojourn time in state \( j \) on \( k \)th visit to state \( j \). The RHS of 4.2 is
\[
\frac{\mathbb{E}[N(j)]/\lambda_j}{\sum_j \mathbb{E}[N(j)]/\lambda_j} = \frac{\mu(j)\mathbb{E}[\tau(i)]/\lambda_j}{\sum_j \mu(j)\mathbb{E}[\tau(i)]/\lambda_j} = \frac{\mu(j)/\lambda_j}{\sum_j \mu(j)/\lambda_j}.
\]
Thus, if \( \sum_j \mu(j)/\lambda_j < \infty \) then \( \pi(j) > 0 \) for all \( j \) and \( X \) is positive recurrent. More directly, if \( \mathbb{E}[\tau_i] < \infty \), then \( \mathbb{E}[\omega(i)] = \mathbb{E}[\tau_i] \sum_j \mu(j)/\lambda_j \) implies that if \( \sum_j \mu(j)/\lambda_j < \infty \) then \( \mathbb{E}[\omega(i)] < \infty \) and hence \( i \) is positive recurrent. Thus, also \( X \).

Take derivative at \( t = 0 \), we get \( \pi Q = 0 \). More generally we have

**Proposition 4.4.5.** An irreducible positive recurrent, nonexplosive MC \( X \) is positive recurrent iff we can find a probability measure \( \pi \) s.t. \( \pi Q = 0 \). Then \( \pi \) is the unique stationary distribution of \( X \).

More along the lines mentioned above, we also have

**Proposition 4.4.6.** A sufficient condition for positive recurrence of an irreducible chain \( X \) is that \( \exists \) a distribution \( \pi \) s.t. \( \pi Q = 0 \) and \( \sum_j \pi(j) / \lambda_j < \infty \). Then \( \{Y_k\} \) is also positive recurrent with the unique stationary distribution \( \mu(j) = \pi(j) / \lambda_j \).
4.5 Time Reversibility

Let $\{X_t\}$ be irreducible and positive recurrent and $\pi$ is its stationary distribution. We consider $\{X_t\}$ under stationarity. For fixed $T$ consider $Y_t = X_{T-t}$. It is a Markov chain with transition function,

$$
\tilde{P}_t(i,j) = \frac{P[Y_t = j | Y_0 = i]}{P[X_T = i]} = \frac{P[X_{T-t} = j | X_T = i]}{P[X_T = i]} = \frac{P[X_T = j | X_{T-t} = i]}{P[X_T = i]} \frac{P(j,i)\pi(j)}{\pi(i)}.
$$

The stationary distribution of the reverse process is same as for the forward process, because the fraction of the time spent by the MC in state $i$ in both directions is same.

**Definition 4.5.1.** $\{X_t\}$ is time reversible if

$$
\tilde{P}_t(i,j) = \frac{P(j,i)\pi(j)}{\pi(i)} = P(i,j).
$$

Taking derivative at $t = 0$,

$$
Q_{ij} = \frac{\pi(j)}{\pi(i)} Q_{ji}, \quad \forall i, j.
$$

The above equation is called detailed balance.
4.6 Birth-Death process

Consider a system whose state $X_t$ at any time is represented by the number of people in the system at that time. Suppose that whenever there are $n$ people in the system, then (i) new arrivals enter the system at an exponential rate $\beta_n$, and (ii) people leave the system at an exponential rate $\delta_n$. Such a system $\{X_t\}$ is called a birth-death process (B-D). The parameters $\{\beta_n\}_{n=0}^{\infty}$ and $\{\delta_n\}_{n=0}^{\infty}$ are called, respectively, the arrival (or birth) and departure (or death) rates.

The state space of a birth-death process is $\{0, 1, \ldots\}$. The transitions from state $n$ may go only to either state $n-1$ (if $n > 0$) or state $n+1$. Thus, it is a Markov chain with its jump chain $\{Y_k\}$ having the transition matrix,

$$
P_{i,i+1} = \frac{\beta_i}{\beta_i + \delta_i}, \quad i > 0, \quad P_{i,i-1} = \frac{\delta_i}{\beta_i + \delta_i}, \quad i > 0.
$$

An example of a birth-death process is the $\{q_t\}$ process of an $M/M/1$ queue.

We use conditions for transience of DTMC given earlier to give conditions for recurrence of a B-D process.

Recurrence of $\{X_t\} \iff$ Recurrence of $\{Y_k\}$. Thus, we look for a bounded solution $h : S \setminus \{0\} \to \mathbb{R}$ with

$$
h(j) = \sum_{k \neq 0} P_{jk} h(k), \quad j \neq 0,
$$

Then,

$$
h(1) = \frac{\beta_2}{\beta_2 + \delta_2} h(2).
$$

Writing $p_i = \frac{\beta_i}{\beta_i + \delta_i}$ and $q_i = 1 - p_i = \frac{\delta_i}{\beta_i + \delta_i}$,

$$(p_j + q_j) h(j) = q_j h(j-1) + p_j h(j+1)
$$

Solving this iteratively, we get

$$
h(2) - h(1) = \frac{q_1}{p_1},
$$

$$
h(j + 1) - h(j) = h(1) \frac{q_j q_{j-1} \cdots q_1}{p_j p_{j-1} \cdots p_1}
$$
For this to be a bounded function, we need

\[ \sum_{j=1}^{\infty} \frac{q_j q_{j-1} \cdots q_1}{p_j p_{j-1} \cdots p_1} < \infty. \]

This is necessary and sufficient condition for \( \{Y_n\} \) to be transient. Therefore,

\[ \sum_{j=1}^{\infty} \frac{q_j q_{j-1} \cdots q_1}{p_j p_{j-1} \cdots p_1} = \infty \iff \{Y_n\} \text{ is recurrent} \iff \{X_t\} \text{ is positive recurrent}. \]

Now we give conditions for positive recurrence of a B-D process. Solving the equation \( \pi Q = 0 \), we get

\[ \pi(n) = \frac{\beta_n \beta_{n-1} \cdots \beta_1}{\delta_{n+1} \delta_n \cdots \delta_2} \pi(0) \]

From this we can conclude that

\[ \sum \pi(i) < \infty \iff \sum \frac{\beta_i \beta_{i-1} \cdots \beta_1}{\delta_{i+1} \delta_i \cdots \delta_2} < \infty. \]

This is a necessary and sufficient condition for positive recurrence of the birth-death process.

### 4.6.1 Reversibility of Birth-Death process

**Proposition 4.6.1.** A stationary birth-death process is reversible.

**Proof.** We need to show that

\[ \pi(i) Q_{ij} = \pi(j) Q_{ji} \]

But

\[ \frac{\beta_i \beta_{i-1} \cdots \beta_1}{\delta_{i+1} \delta_i \cdots \delta_2} Q_{i,i+1} = \frac{\beta_{i+1} \beta_{i-1} \cdots \beta_1}{\delta_{i+2} \delta_{i+1} \delta_i \cdots \delta_2} Q_{i+1,i} \]

because,

\[ Q_{i,i+1} = \beta_i + 1, \quad Q_{i+1,i} = \delta_{i+2}. \]

\[ \square \]

### 4.6.2 Examples

**Example 1:** In the M/M/1 queue \( \beta_n = \lambda, \delta_n = \mu \). For recurrence,

\[ \sum_{j=1}^{\infty} \left( \frac{\mu}{\mu+\lambda} \right)^j = \sum_{j=1}^{\infty} \left( \frac{\lambda}{\mu} \right)^j = \infty. \]

That is \( \frac{\lambda}{\mu} \leq 1 \) is the necessary and sufficient condition for an M/M/1 queue to be recurrent.
For positive recurrence,
\[ \sum_{k=1}^{\infty} \frac{\beta_k \beta_{k-1} \ldots \beta_1}{\delta_k \delta_{k-1} \ldots \delta_2} < \infty \quad \Rightarrow \quad \sum_{k=1}^{\infty} \left( \frac{\lambda}{\mu} \right)^k < \infty, \quad \frac{\lambda}{\mu} < 1 \]

Then,
\[ \pi(n) = \left( \frac{\lambda}{\mu} \right)^n \pi(0) \quad \text{and} \quad \sum_{n=0}^{\infty} \pi(n) = 1 \]

Therefore,
\[ \pi(n) = \rho^n (1 - \rho) \quad \text{where} \quad \rho = \frac{\lambda}{\mu}. \]

**Example 2:** The $M/M/\infty$ queue has an $\infty$ number of servers. Whenever a customer arrives it joins an idle server and gets service with an exponential distribution $\exp(\mu)$. After completion of service it leaves the system. Let $q_i$ be the number of customers in the system at time $t$. Its $Q$ matrix is given by
\[ q_{i,i+1} = \lambda, \quad q_{i,i-1} = i \mu \quad \text{for} \quad i > 0. \]

For recurrence, we need
\[ \sum_j q_j q_{j-1} \ldots q_1 \frac{p_j p_{j-1} \ldots p_1}{p_j p_{j-1} \ldots p_1} = \infty, \]
\[ \sum_j j \mu (j-1) \mu \ldots \mu \left( \frac{q}{\lambda} \right)^j \frac{1}{\lambda^j} = \sum_j j! \left( \frac{q}{\lambda} \right)^j \frac{1}{\lambda^j} \]
\[ = \infty. \]

This holds if $\frac{q}{\lambda} \neq 0$. Positive recurrence also holds because
\[ \sum_{k=1}^{\infty} \frac{\beta_k \beta_{k-1} \ldots \beta_1}{\delta_k \delta_{k-1} \ldots \delta_2} = \sum_k \frac{\lambda^k}{(k+1)\mu k \mu \ldots \mu} \]
\[ = \sum_{k=0}^{\infty} \frac{\lambda^k}{\mu^k (k+1)!} < \infty. \]
4.7 Problems

**Problem 1:** Consider a population in which each individual independently gives birth at an exponential rate \( \lambda \) and dies at an exponential rate \( \mu \). In addition, new individuals enter according to a Poisson process with rate \( \theta \). Let \( \{X(t)\} \) denote the number of individuals in the population at time \( t \).

1. Show that \( \{X(t)\} \) is a Markov chain.
2. Find the generator matrix of \( \{X(t)\} \).
3. Find the conditions for stationary distribution to exist. Also, find the stationary distribution under these conditions.
4. Find \( \mathbb{E}[X(t)|X(0)] \).

**Problem 2:** Let \( A \) be a subset of the state space of Markov chain \( \{X(t)\} \). Let \( T_i(t) \) be the amount of time spent in \( A \) in time \([0,t]\) given that \( X(0) = i \). Let \( Y_1, Y_2, \ldots, Y_n \) be i.i.d. with \( \exp(1/\lambda) \) independent of \( \{X(t)\} \). Let \( t_i(n) = \mathbb{E}[T_i(Y_1 + Y_2 + \cdots + Y_n)] \).

1. Derive a set of linear equations for \( t_i(1) \), \( \forall i \).
2. Derive a set of linear equations for \( t_i(n) \) in terms of \( t_i(1) \) and \( t_i(n-1) \).
3. When \( n \) is large, for what values of \( \lambda \) is \( t_i(n) \) a good approximation of \( \mathbb{E}[T_i(t)] \).

**Problem 3:** Consider a CTMC \( \{X(t)\} \) with stationary distribution \( \pi \) and generator matrix \( Q \).

1. Compute the probability that its sojourn time in state \( i \) is greater that \( \alpha > 0 \).
2. Consider the jump chain \( \{Y_n\} \). Compute its transition matrix \( P \). Find the mean of the first time it comes back to state \( i \) if \( X(0) = i \).
3. Use the above two to find \( \mathbb{E}[T_i|X(0) = i] \), where \( T_i \) is the first time \( \{X(t)\} \) has its sojourn time in state \( i \) greater that \( \alpha > 0 \).

**Problem 4:** Consider an M/M/1/2 queue. Arrival rate \( \lambda = 3 \) per hour and service times are i.i.d. \( \exp(4) \). Let \( q(t) \) be the number of customers in the system at time \( t \).

1. Find the generator matrix for \( \{q(t)\} \).
2. Find the proportion of customers that enter the queue.
3. If the service rate is increased to 8, find (2) above.
4. Find the conditions for stationary distribution for \( \{q(t)\} \).
5. Compute the mean queue length and mean delay of a customer entering the system.

**Problem 5:** If \( \{X(t)\} \) and \( \{Y(t)\} \) are independent, reversible Markov chains, show that \( \{X(t), Y(t)\} \) is also a reversible Markov chain.

**Problem 6:** Customers move among \( r \) servers circularly (after completion of service at service \( i \), the customer moves to the server \((i+1) \mod r\)). Service times at server \( i \) is \( \exp(\mu_i) \). Consider the process \( \{q(0), q(1), \ldots, q(r-1)\} \) where \( q(i) \) denotes the number of customers in server \( i \) for \( i \in \{0, 1, \ldots, (r-1)\} \). Show this process is reversible. Find its stationary distribution.

**Problem 7:** Consider an M/M/\( \infty \) system with arrival rate \( \lambda \) and service rate \( \mu \).
1. Let \( q(t) \) be the number of customers in the system at time \( t \). Find the generator matrix. Find the conditions for stationary distribution. Find the stationary distribution under these conditions.

Now, consider this system as follows: whenever a customer arrives, it joins the lowest numbered server that is free. In other words, when a customer arrives, it enters server 1 if it is free. Otherwise, it enters server 2 if it is free and so on.

1. Find the fraction of time server 1 is free under stationarity.
2. By considering the \( M/M/2 \) loss system, find the fraction of time server 2 is busy.
3. Find the fraction of time server \( c \) is busy for arbitrary \( c \).
4. What is the overflow rate from server \( c \) to \( c+1 \). Is it a renewal process? Is it a Poisson process? Show wherever applicable.

**Problem 8:** Consider an ergodic CTMC \( \{X(t)\} \) with generator matrix \( Q \) and stationary distribution \( \pi \). Let \( E \) be a subset of the state space. Let \( G = E^c \). Under stationarity,

1. compute \( \mathbb{P}\{X(t) = i | X(t) \in B\}, i \in B \),
2. compute \( \mathbb{P}\{X(t) = i | X(t) \in B, X(t^-) \in G\}, i \in B \) and
3. show that

\[
\sum_{i \in G} \sum_{j \in B} \pi_{ij} q_{ij} = \sum_{i \in B} \sum_{j \in G} \pi_{ij} q_{ij}.
\]
Chapter 5

Martingales

Lecture 17

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

5.1 Introduction

Martingales are a very versatile tool in stochastic processes. We will use it to get several useful results in this course.

A discrete time stochastic process is \( \{X_k, k \geq 0\} \) is a martingale w.r.t. filtration \( \mathcal{F}_n = \{Y_0, Y_1, \ldots, Y_n\} \) if

1. If \( X_k \) is a function of \( Y_1, Y_2, \ldots, Y_k \).
2. \( E[|X_k|] < \infty, \quad \forall \, k \geq 0 \).
3. \( E[X_{k+1}|Y_1, Y_2, \ldots, Y_k] = X_k \quad a.s. \)

If the equality in third condition is replaced by \( \leq \) or \( \geq \), then the process is called a supermartingale or a submartingale, respectively.

Example 5.1.1. Let \( Z_1, Z_2, \ldots \) be i.i.d., and \( \mathbb{E}[Z_k] = 0, S_n = \sum_{k=0}^n Z_k \).

\[
\mathbb{E}[S_{n+1}|Z_1, \ldots, Z_n] = \mathbb{E}[S_n + Z_{n+1}|Z_1, \ldots, Z_n] \\
= S_n + \mathbb{E}[Z_{n+1}|Z_1, \ldots, Z_n] \\
= S_n + \mathbb{E}[Z_{n+1}] \\
= S_n.
\]

Hence, \( S_n \) is a martingale w.r.t. \( \{Z_1, Z_2, \ldots, Z_n\} \).
Example 5.1.2. Let $Z_1, Z_2, \ldots$ be i.i.d., and $\mathbb{E}[Z_k] = 0$, $\operatorname{var}(Z_k) = \sigma^2 < \infty$, $S_n = \sum_{k=1}^{n} Z_k$, $S_0 = 0$, $Z_0 = 0$, $X_n = S_n^2 - n\sigma^2$. Then

\[
\mathbb{E}[X_{n+1}|Z_1, \ldots, Z_n] = \mathbb{E}[S_{n+1}^2 - (n+1)\sigma^2|Z_1, \ldots, Z_n] \\
= \mathbb{E}[(S_n + Z_{n+1})^2|Z_1, \ldots, Z_n] - (n+1)\sigma^2 \\
= S_n^2 - n\sigma^2 \\
= X_n.
\]

Hence, $X_n$ is a martingale.

Example 5.1.3. Let $Z_0 = 0, Z_1, Z_2, \ldots$ be i.i.d., and $\mathbb{E}[e^{\theta Z_k}] < \infty$ for some $\theta > 0$, $S_n = \sum_{k=0}^{n} Z_k$. $X_n = \frac{e^{\theta S_n}}{\mathbb{E}[e^{\theta Z_1}]^n}$.

\[
\mathbb{E}[X_{n+1}|Z_1, \ldots, Z_n] = \frac{\mathbb{E}[e^{\theta S_{n+1}}|Z_1, \ldots, Z_n]}{\mathbb{E}[e^{\theta Z_1}]^{n+1}} \\
= e^{\theta S_n} \frac{\mathbb{E}[e^{\theta Z_{n+1}}|Z_1, \ldots, Z_n]}{\mathbb{E}[e^{\theta Z_1}]^{n+1}} \\
= \frac{e^{\theta S_n}}{\mathbb{E}[e^{\theta Z_1}]^n} = X_n.
\]

Hence, $X_n$ is a martingale.

Example 5.1.4. $Y_0 = 1, Y_1, Y_2, \ldots$ independent, $\mathbb{E}[Y_i] = 1$, $X_0 = 1$, $X_n = \prod_{k=1}^{n} Y_k$. Then

\[
\mathbb{E}[X_{n+1}|Y_1, Y_2, \ldots, Y_n] = \mathbb{E}\left[\prod_{k=1}^{n} Y_k|Y_1, Y_2, \ldots, Y_n\right] \\
= \prod_{k=1}^{n} Y_k \mathbb{E}[Y_{k+1}] \\
= \prod_{k=1}^{n} Y_k \\
= X_n.
\]

Hence, $X_n$ is a martingale.

Example 5.1.5. Let $\{X_k, k \geq 0\}$ be a martingale w.r.t. filtration $\mathcal{F}_n = \{Y_0, Y_1, \ldots, Y_n\}$. Let $\phi : \mathbb{R} \to \mathbb{R}$, $Z_n = \phi(X_n)$,

\[
\mathbb{E}[Z_{n+1}|Y_0, Y_1, \ldots, Y_n] = \mathbb{E}[\phi(X_{n+1})|Y_0, Y_1, \ldots, Y_n]
\]

If $\phi$ is a convex function, then by Jensen’s inequality

\[
\mathbb{E}[\phi(X_{n+1})|Y_0, Y_1, \ldots, Y_n] \geq \phi(\mathbb{E}[X_{n+1}|Y_0, Y_1, \ldots, Y_n]) = \phi(X_n)
\]

Hence, $\phi(X_n)$ is a submartingale.
If \( \{X_n\} \) is a submartingale and \( \phi \) convex, non decreasing,
\[
\mathbb{E}[Z_{n+1}|Y_0, Y_1, \ldots, Y_n] = \mathbb{E}[\phi(X_{n+1})|Y_1, \ldots, Y_n] \\
\geq \phi(\mathbb{E}[X_{n+1}|Y_1, \ldots, Y_n]) \\
\geq \phi(X_n).
\]
Hence, \( \phi(X_n) \) is a submartingale.

### 5.2 Optional Sampling Theorem

If \( \{X_n\} \) is a martingale, w.r.t. \( \{Y_n\} \) then, for \( n > k \)
\[
E[X_n] = E[E[X_n|Y_1, Y_2, \ldots, Y_k]] = E[X_k] = E[X_{k-1}] = \cdots = E[X_0]
\]
For a submartingale, \( E[X_{k+1}] \geq E[X_k] \geq \cdots \geq E[X_0] \).

**Proposition 5.2.1.** If \( \{X_n\} \) is a martingale w.r.t. \( \{Y_n\} \), \( T \) a stopping time w.r.t. \( \{Y_n\} \), then \( \mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0] \) for all \( n \geq 0 \).

**Proof.**

\[
X_{T\wedge n+1} - X_{T\wedge n} = (X_{n+1} - X_n)1_{\{T>n\}}
\]
Taking expectations,
\[
\mathbb{E}[X_{T\wedge n+1} - X_{T\wedge n}] = \mathbb{E}[(X_{n+1} - X_n)1_{\{T>n\}}] = \mathbb{E}[\mathbb{E}[(X_{n+1} - X_n)1_{\{T>n\}}|\mathcal{F}_n]]
\]
Since \( T \) is a stopping time \( \{T\leq n\} \) is a function of \( (Y_1, Y_2, \ldots, Y_n) = \mathcal{F}_n \). Thus, \( \{T\leq n\}^c \) is also a function of \( \mathcal{F}_n \).

Thus,
\[
\mathbb{E}[\mathbb{E}[(X_{n+1} - X_n)1_{\{T>n\}}|\mathcal{F}_n]] = \mathbb{E}[1_{\{T>n\}}\mathbb{E}[(X_{n+1} - X_n)|\mathcal{F}_n]] = 0.
\]
Therefore,
\[
\mathbb{E}[X_{T\wedge n+1}] = \mathbb{E}[X_{T\wedge n}] = \cdots = \mathbb{E}[X_{T\wedge 0}] = \mathbb{E}[X_0].
\]

For a submartingale, the above proof gives
\[
\mathbb{E}[X_{T\wedge n+1}] \geq \mathbb{E}[X_{T\wedge n}] \geq \cdots \geq \mathbb{E}[X_{T\wedge 0}] \geq \mathbb{E}[X_0].
\]

Since,
\[
\lim_{n \to \infty} T(\omega) \wedge n = T(\omega),
\]
\[
\lim_{n \to \infty} X_{T\wedge n}(\omega) = X_T(\omega) \text{ a.s.}
\]
If
\[
\mathbb{E}[X_{T\wedge n}] \to \mathbb{E}[X_T] \text{ as } n \to \infty, \quad (5.1)
\]
then from above proposition
\[
\mathbb{E}[X_0] = \mathbb{E}[X_T]. \quad (5.2)
\]
Conditions for Eq (5.1) to hold are,
• $T \leq n_0$ a.s. for some $n_0 < \infty$. Then $X_{T \wedge n} = X_T$ a.s. when $n \geq n_0$. Thus $E[X_{T \wedge n}] = E[X_T]$, $\forall n \geq n_0$.

• $X_n \leq Z$ a.s. $\forall n \geq n_0$ and $E[z] < \infty$. Then $X_{T \wedge n} \leq Z$ a.s. $\forall n$. Thus $X_{T \wedge n} \to X_T$ a.s., and dominated convergence theorem, implies $E[X_{T \wedge n}] \to E[X_T]$.

**Proposition 5.2.2.** If $\{X_n\}$ is a martingale, $E[|X_n - X_{n-1}| \mid \mathcal{F}_n] \leq c < \infty, \forall n \geq 1$ and $E[T] < \infty$, then $E[X_T] = E[X_0]$.

**Proof.** Define $Z = |X_0| + |X_1 - X_0| + \cdots + |X_T - X_{T-1}|$. We have

$$X_{T \wedge n} = (X_{T \wedge n} - X_{(T \wedge n) - 1}) + (X_{(T \wedge n) - 1} - X_{(T \wedge n) - 2}) + (X_{1} - X_0) + X_0$$

$$\leq |X_{T \wedge n} - X_{(T \wedge n) - 1}| + |X_{(T \wedge n) - 1} - X_{(T \wedge n) - 2}| + |X_1 - X_0| + |X_0| \leq Z$$

Thus if we show $E[Z] < \infty$, we will have the proof from the above result. But,

$$E[Z] = E[|X_0|] + E[|X_1 - X_0|] + \cdots + E[|X_T - X_{T-1}|]$$

$$= E[|X_0|] + \sum_{k=1}^{\infty} E[|X_k - X_{k-1}| 1_{\{T \geq k\}}]$$

$$= E[|X_0|] + \sum_{k=1}^{\infty} E[|X_k - X_{k-1}| 1_{\{T \geq k\}} | \mathcal{F}_{k-1}]$$

$$= E[|X_0|] + \sum_{k=1}^{\infty} E[1_{\{T \geq k\}}] E[|X_k - X_{k-1}| | \mathcal{F}_{k}]$$

$$\leq E[|X_0|] + c E[T]$$

$$< \infty.$$
5.2 (Contd.) Optional Sampling Theorem: Example

Example 5.2.1. Let $Y_0 = 0, Y_1, Y_2, \ldots$ i.i.d., $E[|Y_i|] < \infty$, $E[Y_i] = 0$, $S_0 = 0$, $S_n = \sum_{i=1}^{n} Y_i$, $T$ stopping time w.r.t. $\{Y_0, Y_1, Y_2, \ldots\}$, and $E[T] < \infty$. Then

$$E[|S_{n+1} - S_n| \mid \mathcal{F}_n] = E[|Y_{n+1}| \mid \mathcal{F}_n] = E[|Y_{n+1}|] < \infty.$$ 

hence from previous optional sampling theorem

$$E[S_T] = E[S_0].$$

5.3 Martingale inequalities

Theorem 5.3.1. (Doob’s inequality for submartingales) If $\{X_n\}$ is a submartingale, $M_n = \sup_{0 \leq k \leq n} X_k$, then for $\alpha > 0$

$$P[M_n \geq \alpha] \leq \frac{E[X_n]}{\alpha}.$$

Proof. Let $T$ be a stopping time defined as $T = n \wedge \inf \{k : M_k \geq \alpha\}$.

$$E[X_n] = \sum_{k=0}^{n} E[X_k 1_{T=k}]$$

$$= \sum_{k=0}^{n} E[E[X_k 1_{T=k}] \mid \mathcal{F}_k]$$

$$= \sum_{k=0}^{n} E[1_{T=k} E[X_k \mid \mathcal{F}_k]]$$

$$\geq \sum_{k=0}^{n} E[1_{T=k} X_k]$$

$$= E[X_T].$$

Thus,

$$P[M_n \geq \alpha] = P[X_T \geq \alpha] \leq \frac{E[X_T]}{\alpha} \leq \frac{E[X_n]}{\alpha}.$$

If $\{X_k\}$ is a martingale, $E[|X_k|^\beta] < \infty$, for some $\beta \geq 1$, then $\{|X_k|^\beta\}$ is also a submartingale. Thus from the above theorem

$$P \left[ \sup_{1 \leq k \leq n} |X_k| \geq \alpha \right] = P \left[ \sup_{1 \leq k \leq n} |X_k|^\beta \geq \alpha^\beta \right] \leq \frac{E[|X_n|^\beta]}{\alpha^\beta}.$$
Example 5.3.2. \(Y_0, Y_1, Y_2, \ldots, \mathbb{E}[Y_1] = 0, \mathbb{E}[|Y_1| < \infty]\). Then \(S_n\) is a martingale. Thus, from above theorem
\[
P \left[ \sup_{0 \leq k \leq n} |S_k| \geq \alpha \right] \leq \frac{\mathbb{E}[|S_n|]}{\alpha}.
\]
If \(\mathbb{E}[|Y_1|^\beta] < \infty, \beta \geq 1\), then \(P[\sup_{0 \leq k \leq n} |S_k| \geq \alpha] \leq \frac{\mathbb{E}[|S_n|^\beta]}{\alpha^\beta}\).
For \(\beta = 2\),
\[
\frac{\mathbb{E}[|S_n|^2]}{\alpha^2} = \frac{n\sigma^2}{\alpha^2} \quad \text{where} \quad \sigma^2 = \text{var}(Y_1).
\]
This is called Kolmogorov’s inequality.

Lemma 5.3.3. Let \(\mathbb{E}[X] = 0, \text{ and } P[|X - a| \leq b] = 1\) for some constants \(a\) and \(b\). Then
\[
\mathbb{E}[e^{\theta X}] \leq \exp \left( \frac{\theta^2 b^2}{2} \right).
\]

Theorem 5.3.4 (Azuma inequality). Let \(Y_n\) be a martingale, \(\mathbb{P}[,Y_{n-1}| \leq d_n] = 1\), then
\[
P[|Y_n - Y_0| \geq \alpha] \leq 2\exp \left( -\frac{\alpha^2}{2\sum_{i=1}^{n} d_i^2} \right).
\]

Proof. Since \(\mathbb{E}[e^{(\theta Y_{n-1})}] < \infty\) for all \(\theta\),
\[
P[Y_n - Y_0 \geq \alpha] \leq \frac{\mathbb{E}[e^{\theta(Y_n - Y_0)}]}{e^{\theta \alpha}}.
\]
Also,
\[
\mathbb{E}[\exp(\theta(Y_n - Y_0))] = \mathbb{E}[\mathbb{E}[e^{\theta(Y_n - Y_{n-1}) + \theta(Y_{n-1} - Y_0)}]|\mathcal{F}_{n-1}]
= \mathbb{E}[e^{\theta(Y_{n-1} - Y_0)}\mathbb{E}[e^{\theta(Y_{n-1} - Y_{n-1})}|\mathcal{F}_{n-1}]]
\leq \mathbb{E}[\theta(Y_{n-1} - Y_0)]e^{\left(\frac{\alpha^2 d_i^2}{2}\right)}.
\]
by the above lemma, iterating,
\[
\mathbb{E}[\exp(\theta(Y_n - Y_0))] \leq e^{\left(\frac{\alpha^2}{2} \sum_{i=1}^{n} d_i^2\right)}
\]
(5.4)
From 5.3, 5.4,
\[
P[|Y_n - Y_0| \geq \alpha] \leq \frac{e^{\left(\frac{\alpha^2}{2} \sum_{i=1}^{n} d_i^2\right)}}{e^{(\theta \alpha)}} \quad \text{holds for any } \theta > 0.
\]
Choosing \(\theta = \frac{\alpha}{\sum_{i=1}^{n} d_i^2}\), we get the tightest upper bound,
\[
P[Y_n - Y_0 \geq \alpha] \leq 2\exp \left( -\frac{\alpha^2}{2\sum_{i=0}^{n} d_i^2} \right).
\]
\[\square\]
Consider independent random variables $X_1, X_2, \ldots, X_n$, $E[X_i] = \mu$, $S_n = \sum_{i=1}^n X_i$.

**Definition 5.3.5.** $F: \mathbb{R}^n \to \mathbb{R}$ is called Lipschitz-$c$ if

$$|F(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_m) - F(x_1, x_2, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_m)| \leq c, \forall i, \forall x_1, x_2, \ldots, x_m, y_i.$$

**Proposition 5.3.6 (McDiarmid’s Inequality).** Under above conditions, if $F$ is Lipschitz-$c$,

$$P[|F(X_1, X_2, \ldots, X_m) - E[F]| \geq \alpha] \leq 2e^{(-2\alpha^2/c^2)}.$$

**Proof.** Let $Z = f(X_1, X_2, \ldots, X_m)$, $Z_i = E[Z|X_1, X_2, \ldots, X_i]$. Then $\{Z_i\}$ is a martingale.

Also,

$$|Z_{i+1} - Z_i| = |E[f(x_1, x_2, \ldots, x_{i+1}, X_{i+2}, X_{i+3}, \ldots, X_m)|X_1 = x_1, \ldots, X_i = x_i, X_{i+1} = x_{i+1}] - E[f(x_1, x_2, \ldots, x_i, X_{i+1}, \ldots, X_m)|X_1 = x_1, \ldots, X_i = x_i]|$$

$$\leq E[|f(x_1, x_2, \ldots, x_{i+1}, X_{i+2}, X_{i+3}, \ldots, X_m) - f(x_1, x_2, \ldots, x_i, X_{i+1}, X_{i+2}, X_{i+3}, \ldots, X_m)|] \leq c.$$

Thus,

$$P[|Z_{i+1} - Z_i| \leq c|\mathcal{F}_i] = 1.$$

and the result follows by Azuma’s inequality, because $Z_0 = E[F(X_1, X_2, \ldots, X_m)]$. 


5.4 McDiarmid’s Inequality: Applications

**Example 5.4.1 (Machine learning: Classification problem).** Let the training samples $(X_1, Y_1), \ldots, (X_n, Y_n)$ be iid. $X_i \in \mathbb{R}^d$, $Y_i \in \{1, 2, \ldots, N\}$. $h$ is the classifier, $h(x) \rightarrow \{1, 2, \ldots, N\}$. $1_{(h(X) \neq Y)}$ denotes the error. The probability of error for a given classifier is given by

$$R(h) = P[h(X) \neq Y]$$

and its estimate from the training sample is

$$\hat{R}_n(h) = \frac{1}{n} \sum_{i=1}^{n} 1_{(h(X_i) \neq Y_i)}.$$ 

Define

$$f((x_1, y_1), \ldots, (x_n, y_n)) = \frac{1}{n} \sum_{i=1}^{n} 1_{(h(x_i) \neq y_i)}.$$ 

Removing $i$th component and replacing with another,

$$|f((x_1, y_1), \ldots, \hat{x}_i, \ldots, x_n) - f((x_1, y_1), \ldots, \hat{x}_i, \ldots, x_n)| \leq \frac{1}{n}.$$ 

Then by McDiarmid’s Inequality,

$$P[|R_n(h) - E[R_n(h)]| \geq \lambda] \leq 2 \exp \left( -\frac{2\lambda^2}{n} \right) = 2 \exp(-2\lambda^2 n),$$

where $E[R_n(h)] = R(h)$.

**Example 5.4.2.** If $X_1, X_2, \ldots, X_n \sim P$. We want an estimate of $P$. We estimate $P$ by

$$\hat{P}_n(A) = \frac{1}{n} \sum_{i=1}^{n} 1_{(X_i \in A)}.$$ 

Since,

$$E[\hat{P}_n(A)] = P(A),$$

it is an unbiased estimate. Define

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} 1_{(x_i \in A)}.$$ 

Since $|f(x_1, x_2, \ldots, x_n) - f(x_1, \ldots, \hat{x}_i, \ldots, x_n)| \leq \frac{1}{n}$, by McDiarmid’s Inequality,

$$P[|f(x_1, x_2, \ldots, x_n) - E[f]| \geq \epsilon] = P[|\hat{P}_n(A) - P(A)| \geq \epsilon] \leq 2 \exp \left( -\frac{2\epsilon^2 n}{n} \right).$$
Hoeffding’s inequality

Let $X_1, X_2, \ldots, X_n$ be i.i.d., $a \leq |X_i| \leq b$ a.s., and $S_0 = X_1 + X_2 + \cdots + X_n$, $\mu = \mathbb{E}[X_1]$.

Define

$$f(x_1, x_2, \ldots, x_n) = X_1 + X_2 + \cdots + X_n.$$ 

Since,

$$|f(x_1, x_2, \ldots, x_n) - f(x_1, \ldots, y_i, \ldots, x_n)| = |x_i - y_i| \leq |b - a|,$$

by McDiarmid’s Inequality,

$$P[|S_n - n\mu| \geq \epsilon] \leq 2\exp \left( -2\frac{\epsilon^2 n}{n(b-a)^2} \right)$$

$$P[|S_n - n\mu| \geq \frac{\epsilon}{n}] = P[|S_n - n\mu| \geq \delta] \leq 2\exp \left( -2\frac{\delta^2 n}{(b-a)^2} \right).$$

5.5 Martingale Convergence Theorem

Let $\{X_n\}$ be a submartingale and $\alpha < \beta$. Let $U_n$ denote the number of times $X_k$ goes from below $\alpha$ to above $\beta$ in time $n$. We will need the following equality.

Lemma 5.5.1 (Upcrossing inequality).

$$\mathbb{E}[U_n] \leq \frac{\mathbb{E}[|X_n|] + \alpha}{\beta - \alpha}.$$ 

Proof. Let $Y_n = \max(0, X_n - \alpha)$. $Y_n$ is also a submartingale. Let $U_n$ be the number of times $Y_n$ goes from 0 to above $\beta - \alpha$. This number is same as of $\{X_n\}$ upcrossing from $\alpha$ to $\beta$. Let $T_1$ be the first time when $Y_k = 0$, $T_2 = \min\{k > T_1 \text{ such that } Y_k \geq \beta - \alpha\}$. Similarly, define $T_k$ as the sequence of stopping times up to time $n$, with $T_n$ the maximum possible.

$$\mathbb{E}[Y_n] = \mathbb{E}[Y_{T_n}] = \mathbb{E}\left[ \sum_{k=0}^{n} (Y_{T_k} - Y_{T_{k-1}}) \right]$$

$$= \sum_{k: \text{even}} \mathbb{E}[Y_{T_k} - Y_{T_{k-1}}] + \sum_{k: \text{odd}} \mathbb{E}[Y_{T_k} - Y_{T_{k-1}}].$$

All are bounded stopping times, $Y_n$ is a submartingale, $T_k < T_{k+1} < \cdots < n$. Also $\mathbb{E}[Y_{T_k}] \geq \mathbb{E}[Y_{T_{k-1}}]$ for $k$ odd and $\mathbb{E}[Y_{T_k} - Y_{T_{k-1}}] \geq \beta - \alpha$ for $k$ even. Therefore,

$$\mathbb{E}[Y_n] \geq \mathbb{E}[U_n](\beta - \alpha).$$

Hence,

$$\mathbb{E}[U_n] \leq \frac{\mathbb{E}[Y_n]}{(\beta - \alpha)} = \frac{\mathbb{E}[\max(0, X_n - \alpha)]}{(\beta - \alpha)}$$

$$\leq \frac{\mathbb{E}[|X_n - \alpha|]}{(\beta - \alpha)}$$

$$\leq \frac{\mathbb{E}[|X_n|] + \alpha}{(\beta - \alpha)}.$$ 

$\Box$
**Theorem 5.5.2.** \( \{X_n\} \) submartingale and \( \sup_k E[|X_k|] \leq M < \infty \). Then \( X_n \to X \) almost surely and \( E[|X|] < \infty \).

**Proof.** Take \( \alpha < \beta \). Let \( \liminf X_n(\omega) = X_\omega(\omega) \), \( \limsup X_n(\omega) = X^*(\omega) \). If \( X_\omega(\omega) < \alpha < \beta < X^*(\omega) \), then this sequence will not converge.

Let \( U_n \) be the number of upcrossings of \( \{X_n\} \) from below \( \alpha \) to above \( \beta \) in time \( n \). Thus, from the above lemma,

\[
E[U_n] \leq \frac{E[|X_n|] + \alpha}{\beta - \alpha} \leq \frac{M - \alpha}{\beta - \alpha}
\]

\( U_n(\omega) \) increases as \( n \) increases, and it will converge to \( U(\omega) < \infty \) or reach \( \infty \). Thus,

\[
E[U_n] \nearrow E[U] \leq \frac{M - \alpha}{\beta - \alpha} < \infty.
\]

Therefore, \( P[U(\omega) < \infty] = 1 \), and \( P[X_\omega < \alpha < \beta < X^*] = 0 \). Hence, for rational \( \alpha, \beta \),

\[
P[X_\omega < X^*] \leq P[U_{\alpha<\beta}{X_\omega < \alpha < \beta < X^*}] \leq \sum_{\alpha<\beta} P[X_\omega < \alpha < \beta < X^*] = 0.
\]

This implies \( X_\omega \to X \) a.s.. Also,

\[
M \geq \liminf E[|X_n|] \geq E[|X|], \text{ by Fatou’s lemma.}
\]

If \( X_n \) is a martingale, then \( \{|X_n|\} \) is also a submartingale and \( E[|X_n|] = E[E[|X_n| | F_{n-1}]] \geq E[|X_{n-1}|] \). Therefore,

\[
\sup_k E[|X_k|] = \lim_{n \to \infty} E[|X_n|].
\]

**Lemma 5.5.3.** If \( \{X_k\} \) is submartingale, then for \( X_k^+ = \max(0, X_k) \),

\[
\sup_k E[|X_k|] < \infty \iff \sup_k E[X_k^+] < \infty.
\]

**Proof.** \( \Rightarrow \) \( X_k^+ \leq |X_k| \), hence

\[
\sup_k E[X_k^+] \leq \sup_k E[|X_k|].
\]

\( \Leftarrow \) \( |X_k| = 2X_k^+ - X_k \). Thus,

\[
E[|X_k|] = 2E[X_k^+] - E[X_k] \leq 2E[X_k^+] - E[X_0].
\]

\[
\sup_k E[|X_k|] \leq 2\sup_k E[X_k^+] - E[X_0].
\]

If RHS is finite, then LHS is finite.

Thus, if a submartingale is upper bounded, \( X_k \leq M_1 < \infty \) a.s., then \( X_k^+ \leq M_1 \Rightarrow \sup_k E[X_k^+] \leq M_1 \). Therefore, if a submartingale is upperbounded then it converges a.s..

If \( X_k \) is a supermartingale, then \( -X_k \) is submartingale. Therefore, if \( \sup_k (-X_k) \leq M_1 \), \( X_k \geq -M_1 \), \( \forall k \). Therefore, if \( X_k \) is a supermartingale and lower bounded then it converges.

If \( X_k \) is a martingale then it is a supermartingale and a submartingale. Therefore an upper or lower bounded martingale converges a.s.
Lecture 20

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

Example 5.5.4. Let $Z$ be a random variable with $\mathbb{E}[|Z|] < \infty$, and $\{Y_n\}$ a sequence of random variables. Define $X_0 = 1$, $X_n = \mathbb{E}[Z|Y_0, Y_1, \ldots, Y_n]$, then $\mathbb{E}[X_{n+1}|Y_0, Y_1, \ldots, Y_n] = \mathbb{E}[Z|Y_0, Y_1, \ldots, Y_n] = X_n$. Also,

$$\mathbb{E}[|X_n|] = \mathbb{E}[\mathbb{E}[|Z||Y_0, Y_1, \ldots, Y_n]]
\leq \mathbb{E}[\mathbb{E}[|Z||Y_0, Y_1, \ldots, Y_n]]
= \mathbb{E}[|Z|]
< \infty \quad \text{for all } n \geq 1.$$

Hence $\{X_n\}$ is a martingale w.r.t. $\{Y_n\}$ (called Doob’s Martingale). Also, $X_n \to X_\infty = \mathbb{E}[Z|Y_0, Y_1, \ldots]$.

Example 5.5.5. Let $\{X_n\}$ be a MC with transition matrix $P$. If $h$ is a function such that

$$h(i) = \sum_{j \in \mathcal{S}} p_{ij} h(j) = \mathbb{E}[h(X_1)|X_0 = i].$$

Define $Y_n = h(X_n)$. Then,

$$\mathbb{E}[Y_{n+1}|X_0, X_1, \ldots, X_n] = \mathbb{E}[h(X_{n+1})|X_0, X_1, \ldots, X_n]
= \mathbb{E}[h(X_{n+1})|X_n]
= h(X_n) = Y_n.$$

Therefore $\{Y_n\}$ is a martingale w.r.t. $\{X_n\}$ . If the equality is replaced with $\leq$ then it is submartingale.

Suppose $h$ is bounded. Since $Y_n = h(X_n)$ is a submartingale, $Y_n \to Y_\infty$ a.s. and $\mathbb{E}[Y_\infty] < \infty$. Assume, $\{X_n\}$ is irreducible and recurrent. Consider $i, j \in \mathcal{S}, i \neq j$. State-$i$ occurs infinitely often w.p.1 and state-$j$ also occurs infinitely often w.p.1. Thus, $Y_n = h(X_n) = h(i)$ and $h(j)$ infinitely often with w.p.1. Therefore, for $Y_n$ to converge a.s.,

$$h(i) = h(j), \quad \forall i, j \in \mathcal{S}.$$

5.6 Applications to Markov chain

In this section we use martingale convergence theorems to get conditions for recurrence and transience of Markov chains.

Theorem 5.6.1. Let $\{X_n\}$ be an irreducible, MC with state space $\mathcal{S}$ and transition matrix $P$. It is transient if and only if $\exists$ a state-$i$ and $h : \mathcal{S} \setminus \{i\} \to \mathbb{R}$, $h$ is bounded, non-zero and satisfies

$$h(j) = \sum_{k \neq i} p_{jk}(h(k)) \quad \forall j \neq i.$$

Proof. Suppose $\{X_n\}$ is transient. Fix a state $i$. Let $T(i)$ be the first time chain enters state $i$. Define $h(j) = P_j[T(i) = \infty]$. It is bounded. Since $\{X_n\}$ is transient, $P_j[T(i) = \infty] > 0$. Also,

$$P_j[T(i) = \infty] = \sum_{k \neq i} P_{jk} P_k[T(i) = \infty].$$
Now we assume that such an \( h \) exists and we show that MC is transient. Define, \( \tilde{h} \) on \( S \) such that \( \tilde{h}(j) = h(j) \quad \forall \ j \neq i, \tilde{h}(i) = 0 \). Thus, when \( j \neq i, \)
\[
\mathbb{E}[\tilde{h}(X_1)|X_0 = j] = \sum_k P_{jk}\tilde{h}(k) = \sum_{k\neq i} P_{jk}h(k) = h(j) = \tilde{h}(j) = \tilde{h}(X_0).
\]
When \( j = i, \)
\[
\mathbb{E}[\tilde{h}(X_1)|X_0 = i] = \sum_{k\neq i} P_{ik}h(k) \geq \tilde{h}(i).
\]
Therefore,
\[
\mathbb{E}[\tilde{h}(X_1)|X_0] \geq \tilde{h}(X_0).
\]
Thus from the previous example, \( \tilde{h}(X_n) = Y_n \) is a submartingale and it is bounded. Hence, it converges to \( Y_\infty \) a.s.

If \( \{X_n\} \) is recurrent then as shown above, \( \tilde{h} \) is a constant. But \( \tilde{h}(i) = 0 \) and \( \tilde{h} \) is non-zero. Therefore it cannot be recurrent.

**Theorem 5.6.2.** Let \( \{X_n\} \) be irreducible. If \( \exists h : S \rightarrow \mathbb{R} \) such that \( h(i) \rightarrow \infty \) as \( i \rightarrow \infty \) and there is a finite set \( E_0 \subset S \) such that \( \mathbb{E}[h(X_i)|X_0 = i] \leq h(i), \forall i \notin E_0, \) then \( \{X_n\} \) is recurrent.

**Proof.** We can if needed add a constant to make \( h \geq 0 \). Let \( T \) be entrance time to set \( E_0, X_0 = i, \ i \notin E_0. \ Y_n = h(X_n)1_{\{T \geq n\}}, \mathcal{F}_n = \{X_0,X_1,\ldots,X_n\}. \)

Then,
\[
\mathbb{E}[Y_{n+1} | \mathcal{F}_n] = \mathbb{E}[h(X_{n+1})1_{\{T \geq n+1\}} | \mathcal{F}_n] \\
\leq \mathbb{E}[h(X_{n+1})1_{\{T > n\}} | \mathcal{F}_n] \\
= 1_{\{T > n\}} \mathbb{E}[h(X_{n+1}) | \mathcal{F}_n] \\
\leq 1_{\{T > n\}} h(X_n) \\
= Y_n.
\]
Therefore, \( Y_n \) is a nonnegative supermartingale and \( Y_n \rightarrow Y_\infty \) a.s. with \( P[Y_\infty < \infty] = 1. \)

Suppose \( X_n \) is transient. Then \( X_n \) will be out of any finite set \( \{i : h(i) \leq a\} \) after some time. Thus,
\[
h(X_n) \rightarrow \infty \quad \text{a.s.}
\]
But \( Y_\infty < \infty \quad \text{a.s.} \). Therefore,
\[
P[T < \infty] = 1, \quad \forall i \notin E_0.
\]
Thus, finite set \( E_0 \) is being visited infinitely often w.p.1. Since \( E_0 \) is a finite set, at least one of the states \( i \in E_0 \) is being visited infinitely often w.p.1., That state is recurrent. Then \( \{X_n\} \) is recurrent. Hence a contradiction.

Under slightly stronger conditions, we get positive recurrence of the MC.

**Theorem 5.6.3.** Let \( \{X_n\} \) be irreducible. \( h : S \rightarrow \mathbb{R} \) s.t. \( h \) is lower bounded (make it \( \geq 0 \) by adding a constant) and \( E_0 \) is finite such that \( \mathbb{E}[h(X_i)|X_0 = i] \leq h(i) - \varepsilon \quad \forall i \notin E_0, \) for some \( \varepsilon > 0, \) and \( \mathbb{E}[h(X_i)|X_0 = i] < \infty, \quad \forall i \in E_0. \) Then \( \{X_n\} \) is positive recurrent.

66
Proof. Let $T$ be the entrance time to set $E_0$, $X_0 = i$, $i \notin E_0$, $Y_n = h(X_n)1_{\{T > n\}}$. Then
\[
\mathbb{E}[Y_{n+1} | \mathcal{F}_n] = \mathbb{E}[h(X_{n+1})1_{\{T > n+1\}} | \mathcal{F}_n] \\
\leq \mathbb{E}[h(X_{n+1})1_{\{T > n\}} | \mathcal{F}_n] \\
= 1_{\{T > n\}} \mathbb{E}[h(X_{n+1}) | \mathcal{F}_n] \\
\leq 1_{\{T > n\}} [h(X_0) - \epsilon] \\
= Y_n - \epsilon 1_{\{T > n\}}.
\]

Take expectations on both sides,
\[
\mathbb{E}[Y_{n+1}] \leq \mathbb{E}[Y_n] - \epsilon P[T > n] \\
\leq \mathbb{E}[Y_{n-1}] - \epsilon P[T > n-1] - \epsilon P[T > n] \\
\vdots \leq \mathbb{E}[Y_0] - \epsilon \sum_{k=0}^{\infty} P[T > k].
\]

Take $n \to \infty$,
\[
0 \leq \mathbb{E}[Y_0] - \epsilon \sum_{k=0}^{\infty} P[T > k] = \mathbb{E}[Y_0] - \epsilon \mathbb{E}_0[T].
\]

Thus,
\[
\mathbb{E}_i[T] \leq \frac{\mathbb{E}[Y_0]}{\epsilon} = \frac{h(i)}{\epsilon} < \infty, \quad \forall i \notin E_0.
\]

For $i \in E_0$,
\[
\mathbb{E}_i[T] = \sum_{j \in E_0} p_{ij} + \sum_{j \notin E_0} p_{ij} \mathbb{E}_j[T + 1] \\
= 1 + \sum_{j \notin E_0} p_{ij} \mathbb{E}_j[T] \\
\leq 1 + \sum_{j \notin E_0} \frac{1}{\epsilon} p_{ij} h(j) \\
\leq 1 + \frac{1}{\epsilon} \mathbb{E}_i[h(X_1)] \\
< \infty.
\]

Therefore, starting from any initial state, mean time to reach the finite set $E_0$ is finite. We can show that this implies that $\{X_n\}$ is positive recurrent. \hfill \qed

Theorem 5.6.4. $\{X_n\}$ is irreducible, $\exists$ a bounded function $h : S \to \mathbb{R}$ and a finite set $E_0 \subset S$ s.t.
\[
\mathbb{E}[h(X_1) | X_0 = i] \geq h(i), \quad i \notin E_0
\]
and $h(i) > h(j)$ for some $i \notin E_0$ and all $j \in E_0$. Then $\{X_n\}$ is transient.

Proof. Take $Y_n = h(X_n; T)$, $T$ the entrance time to $E_0$, $X_0 = i \notin E_0$. We can show that $Y_n$ is a submartingale. Since $h$ is bounded, $Y_n \to Y_\infty$ a.s. Also,
\[
\mathbb{E}[Y_n] \geq \mathbb{E}[Y_0] = h(i).
\]
and the fact that $Y_\infty < h(i)$ on $\{T < \infty\}$ $\Rightarrow P[T = \infty] > 0$. Therefore, $\{X_n\}$ is transient. \hfill \qed

67
Example 5.6.5. Consider a discrete time queue with $X_n$ the number of requests in the queue in the beginning of slot $n$. Then $X_{n+1} = (X_n + Y_n)^+$ where $Y_n = A_n - S_n$. $A_n$ is the number of arrivals in slot $n$, $S_n$ is number of requests we can serve in slot $n$. Then $Y_n$ is i.i.d., integer valued. Taking $h(i) = i$ we show that if $\mathbb{E}[Y_n] < 0$ then we have positive recurrence.

We have

$$\mathbb{E}[X_1 | X_0 = i] - i = \mathbb{E}[(i + Y_0)^+] - i.$$ 

As $i \to \infty$, $|\mathbb{E}[(i + Y_0)^+] - \mathbb{E}[i + Y_0]| \to \infty$. Therefore if $\mathbb{E}[Y_0] < -\varepsilon$ for some $\varepsilon > 0$ then $\mathbb{E}[X_1 | X_0 = i] - i < \varepsilon/2$ for all $i$ large enough. Then we get positive recurrence of $\{X_k\}$ from theorem 5.6.3.
5.7 Problems

Problem 1: Let $\delta_1, \delta_2, \ldots$ be independent with $E[\delta_i] = 0$ Let $X_1 = \delta_1$ and $X_{n+1} = X_n + \delta_{n+1} f_n(X_1, X_2, \ldots X_n)$. Suppose $X_n$ are integrable. Show that $\{X_n\}$ is a martingale.

Problem 2: Let $\{X_n\}$ be a martingale with $E[X_1] = 0$ and $E[X_2^2] < \infty$.
1. Show that $E[(X_{n+r} - X_r)^2] = \sum_{k=1}^r E[(X_{n+k} - X_{n+k-1})^2]$.
2. Assume $\sum_n E[(X_n - X_{n-1})^2] < \infty$. Prove that $X_n$ converges with probability 1.

Problem 3: If $\{X_n\}$ is martingale bounded either above or below, then show that $\sup_n E[|X_n|] < \infty$.

Problem 4: Let $\{Y_n\}$ be i.i.d. with $P\{Y_n = 1\} = p = 1 - q = P\{Y_n = -1\}$. Let $S_0 = 0$, $S_n = Y_1 + Y_2 + \cdots + Y_n$, $T = \inf\{S_n = -a \text{ or } S_n = b\}$. When $p \neq q$ show that $E[T] = \frac{b}{p-q} - \frac{a+b}{p-q} 1 - \frac{(p/q)^b}{1 - (p/q)^b}$.

Problem 5: Suppose $\{X_n\}$ is martingale. Let for some $\alpha$, $E[|X_n|^\alpha] < \infty$ for all $n$. Show that $E \left[ \max_{0 \leq k \leq n} |X_k| \right] \leq \frac{\alpha}{1 - \alpha} E[|X_n|^\alpha]^{\frac{1}{\alpha}}$.

Problem 6: Show that a submartingale $\{X_n\}$ can be represented as $X_n = Y_n + Z_n$ where $\{Y_n\}$ is a martingale and $0 \leq Z_1 \leq Z_2 \leq \ldots$. Hint: Take $X_0 = 0$, $\delta_n = X_n - X_{n-1}$ and $Z_n = \sum_{k=1}^n E[\delta_k | \mathcal{F}_{k-1}]$.

Problem 7: Let $\{X_i\}$ be i.i.d. with $P\{X_i = 0\} = P\{X_i = 2\} = 1/2$. Check if $\{X_i\}$ is a martingale and if we can apply martingale stopping theorem.

Problem 8: There are $n$ red balls, $n$ yellow balls and $m$ boxes. A red ball is kept in box $j$ with probability $p_j$ and a yellow ball with probability $q_j$ independently. Let $X$ be the number of boxes with one red and one yellow ball. Calculate $E[X] = \mu$ and an exponential upper bound for $P[|X - \mu| > b]$.
Chapter 6

Random Walks

Lecture 21

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

6.1 Definitions

Definition 6.1.1 (Random Walks). Let $X_1, X_2, \ldots$ be i.i.d., $S_0 = 0$ and $S_n = \sum_{k=1}^{n} X_k$. Then, $S_n$ is called random walk.

If $\mu = \mathbb{E}[X_1] (-\infty \leq \mathbb{E}[X_1] \leq \infty)$ is defined, then by strong law of large numbers (SLLN) $S_n/n \to \mathbb{E}[X_1]$ a.s. as $n \to \infty$. According to law of iterated logarithms (LIL), if $\text{Var}(X_1) = \sigma^2 < \infty$,

$$\limsup_{n \to \infty} \frac{S_n - n\mu}{\sigma \sqrt{n \log \log n}} = +1 \text{ a.s.}$$
$$\liminf_{n \to \infty} \frac{S_n - n\mu}{\sigma \sqrt{n \log \log n}} = -1 \text{ a.s.}$$

We can use martingales theory to analyze $S_n$. $S_n$ is also a Markov chain. So we can use Markov chain theory (although it may not have countable state space). If $X_k \geq 0$, we can use renewal theory. In this chapter, we use random walk theory and will also show how to use renewal theory when $X_1$ takes positive as well as negative values.

There are three possibilities

1. $S_n \to -\infty$ a.s. as $n \to \infty$.
2. $S_n \to \infty$ a.s. as $n \to \infty$.
3. $\limsup_{n \to \infty} S_n = \infty$ a.s. and $\liminf_{n \to \infty} S_n = -\infty$ a.s.

If $\mathbb{E}[X_1]$ exists ($-\infty \leq \mathbb{E}[X_1] \leq \infty$), then by SLLN, (1) holds if $\mu < 0$ and (2) holds if $\mu > 0$. By LIL, (3) holds if $\sigma^2 < \infty$ and $\mu = 0$. 

70
Definition 6.1.2. If for all finite intervals $I \subset \mathbb{R}$, $\sum_{n=0}^{\infty} \mathbb{P}\{S_n \in I\} = \sum_{n=1}^{\infty} \mathbb{E}[1\{S_n \in I\}] < \infty$, then the random walk $S_n$ is called transient. Otherwise, it is called recurrent.

When $\mu < 0$ or $\mu > 0$, the random walk is transient. When $\mu = 0$, the random walk is recurrent.

### 6.2 Ladder Heights, Maxima, GI/GI/1 Queue

**Definition 6.2.1 (Ladder epochs and heights).** Let $T_1 = \inf\{n : S_n > 0\}$ and $T_k = \inf\{n > T_{k-1}, S_n > S_{T_{k-1}}\}$. The process $\{T_k\}$ is called strictly ascending ladder epochs and $\{S_{T_k}\}$ is called strictly ascending ladder heights.

Let $T_1^- = \inf\{n : S_n \leq 0\}$ and $T_k^- = \inf\{n > T_{k-1}, S_n \leq S_{T_{k-1}}\}$. The process $\{T_k^\prime\}$ is called weakly descending ladder epochs and $\{S_{T_k^-}\}$ is called weakly descending ladder heights.

When $S_n \to -\infty$ a.s., then after some time $S_n$ will not go below 0 and hence $\mathbb{P}\{T_\cdot < \infty\} < 1$. Also, when $S_n \to -\infty$ a.s., then $\mathbb{P}\{T < \infty\} < 1$.

Ladder heights and ladder epochs form renewal processes. Let $M_n = \sup_{1 \leq k \leq n} S_k$ and $m_n = \inf_{1 \leq k \leq n} S_k$.

Since, $M_n \geq S_n$ and is monotonically increasing,

1. If $S_n \to \infty$ a.s., then $M_n \uparrow \infty$ a.s. but $m_n \downarrow m > -\infty$ a.s.
2. If $S_n \to -\infty$ a.s., then $M_n \uparrow M < \infty$ a.s. and $m_n \to -\infty$ a.s.
3. If $S_n$ oscillates, then $M_n \uparrow \infty$ a.s. and $m_n \downarrow -\infty$ a.s.

**Proposition 6.2.2.** For GI/GI/1 queue, $W_n \sim M_n$.

**Proof.** Let $X_k = S_k - A_k$.

\[
W_{k+1} = (W_k + X_k)^+ = \max(0, W_k + X_k) = \max(0, \max(0, W_{k-1} + X_{k-1}) + X_k) = \max(0, \max(0, W_{k-1} + X_{k-1}) + X_k) = \max(0, X_k, X_k + X_{k-1}, X_k + X_{k-1}, X_k + X_{k-1}, X_k + X_{k-2}, \ldots, X_k) = \max(0, X_1, X_1 + X_2, X_1 + X_2 + X_3, \ldots, S_k) = M_k.
\]

We should note that $M_n \neq W_n$ a.s. and $M_n$ is monotonically increasing, but $W_n$ is not. Moreover, $(W_{n+1}, W_n) \neq (M_{n+1}, M_n)$ even though $W_{n+1} \sim M_{n+1}$. Now,

\[
(S_0, S_1, \ldots, S_n) \sim (0, X_1, X_1 + X_2, \ldots, X_1 + X_2 + \cdots + X_n) \sim (0, X_n, X_n + X_{n-1}, \ldots, X_n + X_{n-1} + \cdots + X_1) = (0, S_n - S_{n-1}, S_n - S_{n-2}, \ldots, S_n - S_0)
\]

$\max\{0, S_n - S_{n-1}, S_n - S_{n-2}, \ldots, S_n - S_0\} \sim S_n - m_n$. This shows that $(M_n, M_n - S_n) \sim (S_n - m_n, -m_n)$.

We can also write $M_n = S_{N(n)}$ where $N(n)$ is number of ascending ladder epochs till time $n$. Let $Z_k = S_{T_k} - S_{T_{k-1}}$. If $\mu > 0$, then we will show that $\mathbb{E}[T_1] < \infty$. 

71
Proposition 6.2.3. If $\mu > 0$, then

$$
\frac{M_n}{n} \to \mathbb{E}[X_1] \text{ a.s. as } n \to \infty.
$$

Proof. From renewal theory $N(n)/n \to 1/\mathbb{E}[T_1]$ a.s. Thus,

$$
\lim_{n \to \infty} \frac{M_n}{n} = \lim_{n \to \infty} \frac{\sum_{k=1}^{N(n)} Z_k}{n}
\quad = \lim_{n \to \infty} \frac{\sum_{k=1}^{N(n)} Z_k N(n)}{n}
\quad = \frac{\mathbb{E}[Z_1]}{\mathbb{E}[T_1]} \text{ a.s.}
\quad = \frac{\mathbb{E}[\sum_{k=1}^{T_1} X_k]}{\mathbb{E}[T_1]} \text{ a.s.}
\quad = \mathbb{E}[X_1] \text{ a.s.}
$$

Similarly, if $\mu < 0$, $m_n/n \to \mathbb{E}[X_1]$ a.s.

GI/GI/1 queue: Take $X_k = s_k - A_k$. If $\mu = \mathbb{E}[X_1] = \mathbb{E}[s_1 - A_1] > 0$, $M_n \uparrow \infty$ a.s. Therefore, since $W_n \sim M_n$, $\lim_{n \to \infty} \mathbb{P}\{W_n \leq x\} = 0$ for all $x$. Also, $W_n/n \to \mathbb{E}[X_1]$ a.s.

When $\mu < 0$, $M_n \to M$ a.s. where $M$ is a proper r.v. and $\mathbb{P}\{W_n \leq x\} \to \mathbb{P}\{M \leq x\}$. Then, the queue is stable. Also, $N(n) \to N$ a.s. where $N$ is a finite r.v. and

$$
M \triangleq \sum_{k=1}^{N} z_k,
$$

and $\mathbb{P}\{N = n\} = p^n(1 - p)$ where $p = \mathbb{P}\{T_1 < \infty\} < 1$. Also, conditioned on $N \geq k$, $z_1, z_2, \ldots, z_k$ are i.i.d. and do not depend on $k$. 72
6.2 (Contd.) Ladder Epochs

Let $T$ be the first strictly ascending ladder epoch and $T^-$ be the first weakly descending ladder epoch.

**Lemma 6.2.1.** If $\mu > 0$, $\mathbb{E}[T] < \infty$.

**Proof.**

$$
\mathbb{E}[T] = \sum_{k=0}^{\infty} \mathbb{P}\{T > k\} \\
= \sum_{k=0}^{\infty} \mathbb{P}\{S_1 \leq 0, S_2 \leq 0, S_3 \leq 0, \ldots, S_k \leq 0\} \\
= \sum_{k=0}^{\infty} \mathbb{P}\{X_1 \leq 0, X_1 + X_2 \leq 0, X_1 + X_2 + X_3 \leq 0, \ldots, X_1 + X_2 + \cdots + X_k \leq 0\} \\
= \sum_{k=0}^{\infty} \mathbb{P}\{S_k - S_{k-1} \leq 0, S_k - S_{k-2} \leq 0, S_k - S_{k-3} \leq 0, \ldots, S_k \leq 0\} \\
= \sum_{k=0}^{\infty} \mathbb{P}\{k \text{ is a weakly descending ladder epoch}\} \\
= \mathbb{E} \left[ \sum_{k=0}^{\infty} 1\{k \text{ is a weakly descending ladder epoch}\} \right] \\
= \mathbb{E}[N] \\
= \frac{1}{p}
$$

where $N$ is the number of weakly descending ladder epochs and $p = \mathbb{P}\{T^- = \infty\} > 0$ when $\mu > 0$. \qed

The following is a good application of martingale theory to random walks, which will then be used to obtain a useful result in queuing theory.

Assume there exists a $\theta \neq 0$ such that $\mathbb{E}[\exp(\theta X_1)] = 1$. Then, $\mathbb{E}[\exp(\theta S_{n+1})|\mathcal{F}_n] = \exp(\theta S_n)$ where $\mathcal{F}_n = \{X_1, X_2, \ldots, X_n\}$. Thus, $\exp(\theta X_n)$ is a martingale. Let $T = \inf\{S_n \leq -b \text{ or } S_n \geq a\}$ for some $a > 0$ and $b > 0$. We want to show $\mathbb{E}[\exp(\theta S_T)] = 1$. We can use optional sampling theorem if $\mathbb{E}[T] < \infty$ and $\text{sup}_n \mathbb{E}[|\exp(\theta S_{n+1}) - \exp(\theta S_n)||\mathcal{F}_n] < \infty$. 

73
We show the conditions now. We have

\[
E \left[ e^{\theta S_{n+1}} - e^{\theta S_n} \mid \mathcal{F}_n \right] = e^{\theta S_n} \mathbb{E} \left[ e^{\theta X_{n+1}} - 1 \right] \\
\leq e^{\theta S_n} \mathbb{E} \left[ e^{\theta S_{n+1}} - 1 \right] \\
= 2e^{\theta S_n} \leq 2e^{\theta a} < \infty,
\]

for \( n < T \).

Next, we show that \( E[T] < \infty \) when \( X_1 \) is not degenerate. Let \( c = a + b \). Since, \( X_1 \) is not degenerate, there exists an integer \( N \) and \( \delta > 0 \) such that \( \mathbb{P}\{|S_n| > c\} > \delta \). Define \( S'_1 = S_N, S'_2 = S_{2N} - S_N, \ldots \) Then,

\[
\mathbb{P}\{T \geq kN\} \leq \mathbb{P}\{|S'_1| \leq c\}\mathbb{P}\{|S'_2| \leq c\} \ldots \mathbb{P}\{|S'_n| \leq c\} = (1 - \delta)^k.
\]

Thus,

\[
E[T] = \sum_{n=0}^{\infty} \mathbb{P}\{T > n\} \\
\leq N \sum_{k=0}^{\infty} \mathbb{P}\{T > kN\}
\]

because \( \mathbb{P}\{T > k\} \) is decreasing with \( k \). Thus, \( E[T] < \infty \).

If \( p_a = \mathbb{P}\{S_T \geq a\} \), by optional sampling theorem,

\[
1 = \mathbb{E}[e^{\theta S_T}] \\
= \mathbb{E}[e^{\theta S_T} | S_T \leq -b](1 - p_a) + \mathbb{E}[e^{\theta S_T} | S_T \geq a]p_a \\
\geq \mathbb{E}[e^{\theta S_T} | S_T \geq a]p_a \\
\geq e^{\theta a} p_a.
\]

Thus, \( p_a = \mathbb{P}\{S_T \geq a\} \leq \exp(-\theta a) \). This upper bound is independent of \( b \). Taking \( b \to \infty \), we obtain that \( \mathbb{P}\{\sup_{k \geq 0} S_k \geq a\} \leq \exp(-a\theta) \).

**Application to GI/GI/1 queue.** If \( \mu < 0 \), the waiting times \( W_n \to W_\infty \) in distribution and \( W_\infty \) has the distribution of \( M = \sup_{k \geq 0} S_k \). Therefore, \( \mathbb{P}\{W_\infty > a\} \leq \exp(-\theta a) \) if there exists a \( \theta \neq 0 \) such that \( \mathbb{E}[\exp(\theta X_1)] = 1 \).
Chapter 7

Queuing Theory

Lecture 23

Course E2.204: Stochastic Processes and Queueing Theory (SPQT) Spring 2019
Instructor: Vinod Sharma
Indian Institute of Science, Bangalore

7.1 GI/GI/1 Queue

Consider GI/GI/1 queue. Let \{A_n\} be i.i.d interarrival times and \{s_n\} be i.i.d. service times. Let \(\mu = E[s_1] - E[A_1]\). Let the waiting time of \(n^{th}\) arrival be \(W_n\). The process \{\(W_n\}\} is a regenerative process in which the arrivals seeing the queue empty are the regeneration epochs. Let \(\tau\) be the regeneration length. If \(\mu < 0\), we have seen that \(W_n\) converges to a stationary distribution and hence \(E[\tau] < \infty\). Also, \(P\{\tau = 1\} = P\{s_0 < A_1\} > 0\). Therefore, \(\tau\) is aperiodic. If \(\mu \geq 0\), \(W_n \rightharpoonup \infty\) as \(n \to \infty\) and \(E[\tau] = \infty\). If \(\mu = 0\), then \(E[\tau] = \infty\) but \(P\{\tau < \infty\} = 1\).

Let \(V_t\) be total work (service times) of the customers in the queue, including the residual service time of the customer in service. The customers seeing the empty queue are regeneration epochs. Let \(\tau\) be the regeneration length. We have \(E[\tau] = E[\tau]E[A_1]\). If \(\mu < 0\), then also, \(P\{\tau = 1\} > 0\). Thus, If \(A_1\) is non-lattice, \(\tau\) is also non-lattice. Thus, \(V_t \rightharpoonup V_\infty\) in distribution as \(t \to \infty\) if \(\mu < 0\).

Furthermore,

\[
E[V_\infty] = \frac{E[S_1^2 + E[W_1]E[S_1]]}{2E[A_1]}
\]

The last line follows from Figure 7.1. The quantity \(\int_0^\tau V_t dt\) is the area under the curve which can be split into several triangles and parallelogram as shown.

Let the time between consecutive regeneration epochs of \(\{V_t\}\) be called a cycle. During this time, the duration when the queue is empty is called an idle period and the rest is called busy period. Then,
\[ E[\text{busy period}] = E[\sum_{k=0}^{\tau-1} s_k] \] if \( \tau = 0 \) is a regeneration epoch. Hence,

\[
P\{V_\infty = 0\} = \frac{E[\int_0^\tau 1[V_t = 0]dt]}{E[\tau]} = \frac{E[\tau] - E[\text{BusyPeriod}]}{E[\tau]} = \frac{E[\tau]E[A_1] - E[\tau]E[s_1]}{E[\tau]E[A_1]} = 1 - \frac{E[s_1]}{E[A_1]}.\]

Let \( q_t \) be the queue length at time \( t \). If \( \mu = E[X_1] = E[s_1] - E[A_1] < 0 \), \( q_t \to q_\infty \) in distribution as \( t \to \infty \). Let \( S_k \) be the sojourn time of the \( k^{th} \) customer and 0 be a regeneration epoch. Then, \( W_k \) and the sojourn time \( S_k = W_k + s_k \) also have stationary distributions and have the same regeneration epochs. Also, the regeneration epochs of \( V_t \) and \( q_t \) are the same with length \( \tau \). Then, (assuming \( E_\pi[q_\infty] \) and \( E_\pi[s_1] \) are finite, this requires \( E[s_1^2] < \infty \))

\[
E_\pi[q_\infty] = \frac{E[\int_0^\tau q_t dt]}{E[\tau]} = \frac{E[\sum_{k=0}^{\tau-1} s_k]}{E[\tau]E[A_1]} = \frac{E[s_1]}{E[A_1]} = \lambda E_\pi[S_1].
\]
where $\lambda$ is the arrival rate and $E_\pi[S_1]$ is the mean sojourn time under stationarity. Figure 7.2 shows the evolution of queue length in one regeneration cycle. $\int_0^\tau q(t)\,dt$ is the area under the curve for $q(t)$. This is equal to $\sum_{k=0}^{\tau-1} S_k$ as the break of the area in the Figure 7.2 shows.

The above is an example of a general result called Little’s law: $E_\pi[\text{Number in the system}] = (\text{Arrival rate}) \times E_\pi[\text{Sojourn time}]$. This holds for a general queuing system with the same proof if $\int_0^\tau q(t)\,dt = \sum_{k=0}^{\tau-1} S_k$ is valid in that system. We will see many examples of this in next few lectures.

**GI/GI/1-Last Come First Serve (LCFS):** When a new customer arrives, the service of the current customer is stopped and servicing of the latest customer begins. After completion of a service, the server resumes service of the customer it was serving before to complete the remaining service.

An example of this type of queueing is a stack in a computer system.

**Priority queues:** There are different classes of customers and each class is assigned a priority. The customer with the highest priority in the queue is served before others.

All the above schemes have an important property - work conservation:

1. The server is never idle when there is work in the system.
2. Workload will never be increased by policies and queuing schemes.

Irrespective of the policy the queue becomes empty and gets an arrival to the empty queue, at the same time in all the work-conserving queues. Thus, the regeneration epochs for $W_n$, $q_i$, or $V_i$ in the different queues remain same. Hence, $E[\tau] < \infty$, $E[\tau] < \infty$ and has stationary distributions for all these processes for any of the work conserving policy if $E[X_1] < 0$. But, the stationary distribution of $w$, $q$ etc. may be different for different queues. All these queuing systems satisfy Little’s law.

**Restriction of Markov chain to a subset of states:**

Consider $M/M/1/N$ queue with finite buffer of length $N$. The queue length process $\{q_i\}$ is a finite
state space, irreducible Markov chain. It is always positive recurrent with stationary distribution \( \pi_N \). The stationary distribution satisfies \( \pi_N Q = 0 \) where \( Q \) is the rate matrix given by \( Q(i,i+1) = \lambda \) for \( 0 \leq i < N \), \( Q(i,i-1) = \mu \) for \( 0 < i \leq N \) and \( Q(0,1) = \lambda \). Its stationary distribution can also be obtained from that of \( M/M/1 \) queue as

\[
\pi_N(n) = \begin{cases} 
\frac{\pi(n)}{\sum_{k=0}^{N-1} \pi(k)} & \text{for } n \in \{1, 2, \ldots, N\}, \\
0 & \text{otherwise}
\end{cases}
\]

using the following argument.

In general, if \( S \) is the state space of a Markov chain \( \{X_t\} \) with rate matrix \( Q \) and stationary distribution \( \pi \), we can limit the Markov chain to a subset \( A \subset S \) (by modifying the \( Q \) matrix such that the chain is not allowed to exit \( A \) as in the \( M/M/1/N \) queue above) and obtain the corresponding stationary distribution as

\[
\pi_A(i) = \begin{cases} 
\frac{\pi(i)}{\sum_{j \in A} \pi(j)} & \text{if } i \in A, \\
0 & \text{if } i \notin A.
\end{cases}
\]
7.2 Palm Theory, PASTA

Consider a GI/GI/1 queue. Let $T_n$ be the $n^{th}$ arrival epoch, $\{V_t\}$ the workload process and $W_n$ the waiting time of the $n^{th}$ arrival. Then $W_n = V_{T_n} \overset{d}{=} \omega$ a.s. But, $W_\infty \neq V_\infty$ in general. Also, in renewal processes, we have seen inspection paradox where $X_{N(t)} \neq X_n$. This shows that the distribution of the process sampled at some random points may be different from the distribution of the process. In the following, we relate the two distributions.

Let $X = \{X_t, -\infty < t < \infty\}$ be a stochastic process and $T = \{\ldots, T_{-1}, T_0, T_1, \ldots\}$ with $\ldots < T_{-1} < 0 \leq T_0 < T_1 < \ldots$ be a point process. Let $N(t)$ be the number of points of $T$ in the interval $[0, t]$. Let $Z = \{X, T\}$ and $\theta$ be the shift operator defined as $\theta X = X_{\theta T}$ where $(\theta x)_s = X_{s+T}$ and $(\theta T)_n = T_{N(t)+n} - s$. $Z$ is a stationary process if $\theta_s Z$ does not depend on $s$: $P\{Z \in A\} = P\{\theta_s Z \in A\}$ for all measurable $A$.

Now, we ask the question if $Z$ is stationary, is $\{X_t\}$ stationary? If yes, when is the distribution of $\{X_t\}$ same as that of $\{X_s\}$? This is answered in the following theorem called Poisson Arrivals See Time Averages (PASTA).

**Theorem (PASTA).** If $X_t$ is right continuous, $\{X_t, s < t\}$ and $\{N_t - N_s, s \geq t\}$ are independent and $\{N_t\}$ is a Poisson process, $X_{T_t}$ is stationary and has the same distribution as $X_t$.

We will prove this theorem later in this lecture. Let us consider an application of PASTA to $\{V_t\}$ in $M/G/1$ queue. The arrival process is Poisson and the conditions of the theorem hold. Thus, $W_n$ and $V_t$ have the same distribution under stationarity.

Consider the following quantity:

$$\lambda(t) = \frac{E[N(t, t+h)]}{h}.$$  

For a Poisson process $\lambda(t) = \lambda = \text{rate of the Poisson process}$.

**Proposition 7.2.1.** If the process is stationary, $\lambda(t)$ does not depend on $t$ or $h$.

**Proof.** Let $\phi(h) = E[N(t, t+h)]$. By stationarity, $\phi$ does not depend on $t$. We have

$$\phi(h_1 + h_2) = E[N(t, t+h_1 + h_2)] = E[N(t, t+h_1)] + E[N(t+h_1, t+h_1+h_2)] = \phi(h_1) + \phi(h_2).$$

This shows that $\phi$ is linear for a stationary process. So, $\phi(h) = h\phi(1)$. Thus, we have $\lambda(t) = hE[N(t, t+1)]$.

With $h = 1$, we can interpret $\lambda$ as the mean number of arrivals in unit time. Hence, $\lambda$ is called the intensity of the process $N$.  


Define a probability measure \( P_0 \) as

\[
P_0 \{ Z \in F \} = \frac{E \left[ \sum_{t \leq t + h} 1 \{ \theta_t Z \in F \} \right]}{\lambda h}.
\]

The distribution \( P_0 \) is called the Palm distribution of \( Z \).

**Proposition 7.2.2.** The process \( \{ Z \} \) under \( P_0 \) is event stationary: \( P_0 \{ \theta_t Z \in F \} = P_0 \{ Z \in F \} \).

**Proof.** We have

\[
P_0 \{ \theta_t Z \in F \} = \frac{E \left[ \sum_{n=1}^{N(h)} 1 \{ \theta_{t+n} Z \in F \} \right]}{\lambda h} \leq \frac{E \left[ \sum_{n=1}^{N(h)+1} 1 \{ \theta_{t+n} Z \in F \} \right]}{\lambda h} \leq \frac{E \left[ \sum_{n=1}^{N(h)} 1 \{ \theta_{t+n} Z \in F \} \right]}{\lambda h} + \frac{1}{\lambda h} = P_0 \{ Z \in F \} + \frac{1}{\lambda h}.
\]

Letting \( h \to \infty \), we get \( P_0 \{ \theta_t Z \in F \} \leq P_0 \{ Z \in F \} \). Similarly, for \( F^c \) we get, \( P_0 \{ \theta_t Z \in F^c \} \leq P_0 \{ Z \in F^c \} \) which implies \( P_0 \{ \theta_t Z \in F \} \geq P_0 \{ Z \in F \} \). Thus, we have the result.

This proposition shows that under \( P_0 \), the process \( Z \) is event stationary.

Next, define a probability measure \( P_1 \) from \( P_0 \) as

\[
P_1 \{ Z \in A \} = \frac{E_0 \left[ \int_0^{T_1} 1 \{ \theta_s Z \in A \} \right]}{kE_0[T_1]},
\]

where \( E_0 \) is the expectation with respect to \( P_0 \). By event stationarity, this does not depend on \( k \). If \( P_0 \) is the Palm distribution of \( Z \) under \( P \), then \( P_1 = P \). This is called the Palm inversion formula. This is generalization of the formula for regenerative processes.

**Proposition 7.2.3.** The process \( Z \) is time stationary under \( P_1 \).

**Proof.** We want to show \( P_1 \{ \theta Z \in A \} = P_1 \{ Z \in A \} \) \( \forall k \).

\[
P_1 \{ \theta Z \in A \} = \frac{E_0 \left[ \int_0^{T_1} 1 \{ \theta_s Z \in A \} \right]}{kE_0[T_1]} = \frac{E_0 \left[ \int_0^{T_1} 1 \{ \theta_{s+1} Z \in A \} \right]}{kE_0[T_1]} \leq \frac{E_0 \left[ \int_0^{T_1} 1 \{ \theta Z \in A \} \right]}{kE_0[T_1]} \leq \frac{E_0 \left[ \int_0^{T_1} 1 \{ \theta Z \in A \} \right]}{kE_0[T_1]} + \frac{s}{kE_0[T_1]}.
\]
By taking $k \to \infty$, we get $\Pr_1 \{ \theta, Z \in A \} \leq \Pr_1 \{ Z \in A \}$. Similarly for $A^c$, we get $\Pr_1 \{ \theta, Z \in A^c \} \leq \Pr_1 \{ Z \in A^c \}$ which implies $\Pr_1 \{ \theta, Z \in A \} \geq \Pr_1 \{ Z \in A \}$. This shows $\Pr_1 \{ \theta, Z \in A \} = \Pr_1 \{ Z \in A \}$.

**Lemma 7.2.4.** $\Pr_0 \{ T_0 = 0 \} = 1$.

**Proof.** We have

$$
\Pr_0 \{ T_0 = 0 \} = \frac{\mathbb{E}[\sum_{0 < T_i \leq 1} 1 \{ (\theta_i T_0) = 0 \}]}{\lambda} = \frac{\mathbb{E}[N(1)]}{\lambda} (\theta_i T_0) = 0 \text{ always by definition of shift} = \frac{\lambda}{\lambda} = 1
$$

**Lemma 7.2.5.** $\mathbb{E}_0[T_1] = \frac{1}{\lambda}$.

**Proof.**

$$
\lambda \mathbb{E}_0[T_1] = \mathbb{E} \left[ \sum_{k=1}^{N(1)} (T_{k+1} - T_k) \right] = \mathbb{E}[(T_{N(1)+1} - T_0) 1 \{ T_0 \leq 1 \}] = 1
$$

**Lemma 7.2.6.**

$$
\lim_{h \to 0} \frac{\Pr \{ T_0 \leq h \}}{h} = \lambda
$$

**Proof.**

$$
\frac{\Pr \{ T_0 \leq h \}}{h} = \frac{\mathbb{E}_{0[T]} 1 \{ (\theta, T) \leq h \}}{h \mathbb{E}_0[T_1]} = \lim_{h \downarrow 0} \frac{\lambda \mathbb{E}_0[\min(T_1, h)]}{h} = \lambda
$$

We can also show that $\Pr \{ T_i \leq h \}/h \to 0$ as $h \downarrow 0$. Thus,

$$
\Pr_0 \{ Z \in F \} = \lim_{h \downarrow 0} \frac{\mathbb{E} \left[ \sum_{0 < T_i \leq h} 1 \{ \theta_i Z \in F \} \right]}{h \lambda} = \lim_{h \downarrow 0} \frac{\mathbb{E} \left[ 1 \{ \theta_i Z \in F \}; T_0 \leq h \right]}{\Pr \{ T_0 \leq h \}} = \lim_{h \downarrow 0} \Pr \{ 1 \{ \theta_i Z \in F \} | T_0 \leq h \}
$$

This justifies the heuristic interpretation of $\Pr_0 \{ Z \in F \}$ as $\Pr \{ Z \in F | T_0 = 0 \}$. 

81
7.2.1 Rate conservation laws

Let \( \{X_t\} \) be a stochastic process whose sample paths have jumps at \( \{\ldots, T_{-2}, T_{-1}, T_0, T_1, T_2, \ldots\} \) with intensity \( \lambda \). The jump size at \( T_k \) is \( U_k \) (\( U_k = 0 \) is allowed) and between the jumps, \( \{X_t\} \) is differentiable with \( dX_t/\text{d}t = Y_t \) (sample pathwise). The rate conservation law states that

\[ \lambda \mathbb{E}_0[U_0] + \mathbb{E}[Y_0] = 0. \]

**Proof.** Sample pathwise

\[ X_t - X_0 = \sum_{i:0<T_i\leq t} U_i + \int_0^t Y_s \text{d}s. \]

Taking expectation and by stationarity,

\[ 0 = \mathbb{E}[X_t] - \mathbb{E}[X_0] = \lambda \mathbb{E}[U_0] + \mathbb{E} \left[ \int_0^t Y_s \text{d}s \right]. \]

Again, by stationarity of \( Y_t \), \( \mathbb{E}[\int_0^t Y_s \text{d}s] = \mathbb{E}[Y_0] \).

**Example: GI/GI/1 queue:** Let \( q_n^A \) be the queue length seen by \( n \)th arrival and \( q_n^D \) be the queue length left behind by \( n \)th departure. If \( \mathbb{E}[A_1] > \mathbb{E}[S_1] \), then \( \mathbb{P}\{q_n^A \leq k\} = \mathbb{P}\{q_n^D \leq k\} \). This follows from rate conservation law as follows: Define \( X_t = 1\{q_t \geq k\} \). We have \( U_t = +1 \) or \(-1\) and \( Y_t = 0 \) a.s. Therefore, \( \lambda \mathbb{E}_0[U_0] + 0 = 0 \). Thus, \( \mathbb{P}_0\{U_0 = 1\} = \mathbb{P}_0\{U_0 = -1\} \), which implies \( \mathbb{P}_0\{q_n^A \leq k\} = \mathbb{P}_0\{q_n^D \leq k\} \).

7.2.2 PASTA

**Theorem 7.2.8.** Let \( \{X_t\} \) be a stochastic process with \( X_t \in \mathbb{R}^d \) and right continuous sample paths. Let \( N_t \) be a Poisson process with rate \( \lambda \) such that \( \{X_s, s < t\} \) is independent of \( \{N_s, s \geq t\} \) for all \( s \) and \( t \). Then,

\[ \text{(time stationary) } \mathbb{P}\{X_0 \in A\} = \mathbb{P}_0\{X_0^- \in A\} \quad \text{(event stationary) } \]

**Proof.**

\[
\begin{align*}
\mathbb{P}_0\{X_0^- \in A\} &= \frac{\mathbb{E}\left[ \sum_{0<T_i\leq 1} 1\{(\theta_{i-}X_{0^-}) \in A\} \right]}{\lambda} \\
&= \frac{\mathbb{E}\left[ \int_0^1 1\{(\theta_tX_{0^-}) \in A; t \text{ is an event time}\} \text{d}t \right]}{\lambda} \\
&= \frac{\int_0^1 \mathbb{P}\{X_{t^-} \in A; t \text{ is an event time}\} \text{d}t}{\lambda} \\
&= \frac{\int_0^1 \mathbb{P}\{X_{t^-} \in A\} \lambda \text{d}t}{\lambda} \\
&= \int_0^1 \mathbb{P}\{X_{t^-} \in A\} \lambda \text{d}t \\
&= \mathbb{P}\{X_0 \in A\}. 
\end{align*}
\]
7.3 Product-form Networks

In this section, we study queueing networks that have explicit closed form expression for stationary distribution. Also, the stationary distribution of the whole network is the product of marginal stationary distributions of the individual queues.

7.3.1 M/M/1 queue:

Let $q_t$ be the queue length of an $M/M/1$ queue with arrival rate $\lambda$ and service rate $\mu$. It is a birth-death (B-D) process. If the arrival rate $\lambda < \mu$ then $q_t$ is positive recurrent and its stationary distribution $\pi$ is given by

$$\pi(n) = (1 - \rho)\rho^n,$$

where $\rho = \lambda / \mu$. Every B-D process under stationarity is time reversible. Therefore $\{q_t\}$ is time reversible when $\rho < 1$. We consider this queue under stationarity.

Define the reversed process $\hat{q}_t = q_{T-t}$ for some $T$. By reversibility, it has the same distribution as that of $q_t$. Therefore, it can also be considered as the queue length process of an $M/M/1$ queue with arrival rate $\lambda$ and service rate $\mu$. In $\hat{q}_t$ the departure epochs are the arrival epochs in $q_t$. Also, the arrival epochs are the departure epochs of $q_t$. Thus, the departure process of a stationary $M/M/1$ queue is also a Poisson process with rate $\lambda$. Also, because of Poisson arrivals in $q_t$, arrivals from time $t$ onward are independent of the $q_{t-}$. This implies (applying to $\hat{q}_t$) the departures till time $t$ in $q_t$ are independent of $q_{t-}$. These results are rather counter-intuitive.

7.3.2 Tandem Queues

Consider a tandem of $N$ queues. External arrivals enter queue 1 according to a Poisson process of rate $\lambda$. After service in queue $i$, a customer enters queue $i+1$, $i < N$. A customer departs from the system after completing service at queue $N$. The service times at queue $i$ are i.i.d. with $\exp(\mu_i)$. Let $q_i(t)$ denote the queue length at queue $i$ at time $t$. If $\lambda < \mu_1$, queue 1 is stable. Thus, as seen above, under stationarity, the departure process from queue 1 is also a Poisson process of rate $\lambda$. So, $q_1(2)$ is also ergodic if $\lambda < \mu_2$. Continuing this way, each of the queue is stable if $\lambda < \min_i(\lambda_i)$ and the stationary distribution of $q_i(i)$ is given by

$$\pi_i(n) = \rho_i^n(1 - \rho_i).$$

Also as explained above, $q_i(1)$ is independent of the departures till time $t$ (past departures). Therefore, it is independent of arrivals to queue 2 till time $t$. Hence, $q_i(2)$ is independent of $q_i(1)$. Thus, extending this way to other queues, the joint distribution $q_t = (q_1(1), q_1(2), \ldots, q_N(N))$ is given by

$$P[q_t(1) = n_1, q_t(2) = n_2, \ldots, q_t(N) = n_N] = \prod_{i=1}^{N} \rho_i^n(1 - \rho_i).$$
and \( q_t(1), q_t(2), \ldots, q_t(N) \) are independent of each other. However, for \( t_1 < t_2, q_{t_1}(1) \) is not independent of \( q_{t_2}(2) \).

### 7.3.3 Open Jackson Networks

In this system, there are \( N \) nodes. Each node \( i \) consists of 1 server with exponential (\( i.i.d. \)) service times at rate \( 0 < \mu_i < \infty \). At each node, there is an external arrival process according to a Poisson process with rate \( \lambda_i, 0 \leq \lambda < \infty \). After completion of service at queue \( i \), with probability \( p_{ij} \), a customer goes to queue \( j \) independently of routing of other customers. The customer leaves the network with probability \( p_{i0} \) from node \( i \)

\[
P_{i0} = 1 - \sum_{j=1}^{N} p_{ij}.
\]

This is called Markovian routing.

Let \( q_t(i) \) be the queue length at \( i^{th} \) node at time \( t \). Then, \( q_t = (q_t(1), q_t(2), \ldots, q_t(N)) \) is a Markov chain with \( q_t(i) \in \{0, 1, 2, \ldots\} \). We observe that \( q_t \) is irreducible.

Let \( \bar{\lambda}_i \) be the total arrival rate to node \( i \). Then,

\[
\bar{\lambda}_i = \lambda_i + \sum_{j=1}^{N} p_{ji} \bar{\lambda}_j.
\]

There is a unique solution to this set of \( N \) equations denoted by \( (\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_N) \). Let

\[
\rho_i = \frac{\bar{\lambda}_i}{\mu_i}.
\]

We show below that if \( \rho_i < 1 \), for \( i = 1, 2, 3, \ldots, N \), then \( \{q_t\} \) is positive recurrent and has a unique stationary distribution

\[
\pi[q_t(1) = n_1, q_t(2) = n_2, \ldots, q_t(N) = n_N] = \prod_{i=1}^{N} \rho_i^{n_i} (1 - \rho_i)
\]  \hspace{1cm} (7.1)

Hence, \( q_t(1), q_t(2), \ldots, q_t(N) \) are also independent of each other.

This MC is non-explosive and irreducible. If for \( \pi \bar{Q} = 0 \), a solution exists with \( \pi(i) > 0, \sum_i \pi(i) = 1 \), then MC is positive recurrent and \( \pi \) is its unique stationary distribution. We can easily check that Eq. (7.1) satisfies \( \pi \bar{Q} = 0 \).

If \( Q \) is time reversible,

\[
\pi(i)q_{ij} = \pi(j)q_{ji} \forall i, j
\]

Generally, this MC is not time reversible. But we can reverse as

\[
\tilde{q}_t = q_{t-T}, \quad T \text{ a fixed constant}.
\]

This corresponds to a queue length process of a Jackson network, with external input to queue \( i \) as Poisson with rate \( \tilde{\lambda}_i p_{i0} \) and service times \( i.i.d. \exp(\mu_i) \) with routing probabilities,

\[
\tilde{p}_{ij} = p_{ij} \frac{\tilde{\lambda}_j}{\tilde{\lambda}_i}.
\]

Therefore, the departures at each node in \( \{q_t\} \) that exit the system form a Poisson process independent of departures at other nodes that exit the system.

Also, \( q_t \) is independent of future arrivals implies that \( q_t \) is independent of past departures from the network.
7.3.4 Closed queueing networks

In this system, there are no arrivals from outside the network and no departures from the network. A fixed number of $M < \infty$ customers move around in the network. The service times and routing are same as that of a Jackson network. The queue length $q_t = (q_t(1), q_t(2), \ldots, q_t(N))$ is a MC. It is a finite state irreducible MC with state space $S = \{(n_1, n_2, \ldots, n_N) : \sum_{i=1}^{N} n_i = M\}$. It is always stable and has a unique stationary distribution.

The total arrival rate at each node $i$ is

$$\tilde{\lambda}_i = \sum_{j=1}^{N} p_{ji} \lambda_j,$$

By solving these $N$ equations we can get a unique solution upto a constant. The stationary distribution is given by

$$\pi(n_1, n_2, \ldots, n_N) = K \prod_{i=1}^{N} \rho_i^{n_i}, \quad \text{for} \quad \sum_{i=1}^{N} n_i = M, \quad (7.2)$$

where $\rho_i = \tilde{\lambda}_i/\mu_i$ and $K$ is a normalizing constant. It can be checked that Eq (7.2) is a solution for $\pi Q = 0$.

This network has the following bottleneck property. Let $\rho_1 = \max_i (\rho_i)$. If $M \to \infty$, $q_t(1) \to \infty$ in distribution. For other queues

$$\pi(q_t(2) = n_2, \ldots, q_t(N) = n_N) = \prod_{i=2}^{N} \rho_i^{n_i}(1 - \rho_i).$$
7.4 Product-Form Networks: Quasireversible networks

Till now we studied queuing networks with Markovian routing and exponential service times. Now, both of these assumptions will be generalized.

7.4.1 Quasireversible Queues

Consider an $M/M/1$ queue with multiple classes of customers. Let $C$ denote the set of classes, $\lambda_c$ be the Poisson arrival rate for class $c$, $\mu_c$ be the service rate of class $c$. The arrival process for different classes are independent. The traffic intensity of class-$c$ is $\rho_c = \frac{\lambda_c}{\mu_c}$. Its total traffic intensity is $\rho = \sum_{c \in C} \rho_c$.

Let $q_t$ be the number of customers in the queue. If $\rho < 1$, then it has a unique stationary distribution $\pi(n) = (1 - \rho)^n \rho^n$. The probability that a customer is of class-$c$ is $\frac{\rho_c}{\rho}$, independent of others. Let $\{X_t\}$ be the process which gives the class of each customer in the queue at time $t$. $\{X_t\}$ is a Markov chain. The state space $S$ is countable. Its stationary distribution is

$$P[X_t = (c_1, c_2, \ldots, c_n)] = (1 - \rho)^n \prod_{i=1}^{n} \frac{\rho_c}{\rho}.$$

(7.3)

Although it is not reversible, its reversed process $\tilde{X}_t(t)$ also represents a multiclass $M/M/1$ queue, where the last customer $c_n$ leaves the queue first. Its service distributions and arrival processes are same as in the original process. Thus, the departure process of $X_t$ is again Poisson with rate $\lambda_c$ for class $c$ and the Poisson departure processes of different classes are independent. Also, arrivals from $t$ onward, $X_t$ and departures till time $t$ are independent of each other.

It so turns out that the above properties of a multiclass $M/M/1$ queue are the key features needed for product-form stationary distributions of $X_t$. Thus, we abstract these out to study more general classes of queuing systems.

**Definition 7.4.1.** A system is called *quasireversible* if the future arrival processes from time $t$ onward, $X_t$ and past departure processes till time $t$ are independent. These are also independent for different classes.

The reversed process $\tilde{X}_t(t)$ is quasireversible if $\{X_t\}$ is quasireversible. Let $N_c^+(t)$ be arrival process of class-$c$ and $N_c^-(t)$ be the departure process of class-$c$ from the system.

**Proposition 7.4.2.** For a quasireversible system,

1. the arrival processes of different classes are independent Poisson process and
2. the departure processes of different classes are also independent Poisson processes.

**Proof.** For a point process to be Poisson, it should have independent stationary increments. Let $(t_0, t_1], (t_1, t_2], \ldots, (t_{n-1}, t_n]$ be time intervals with $t_0 < t_1 < t_2 < \cdots < t_n$. Let $z_i$ be number of arrivals of class-$c$ during interval $(t_i, t_{i+1}]$. We want to show, for any continuous and bounded $f$,

$$\mathbb{E}[f_0(z_0), f_1(z_1), \ldots, f_{n-1}(z_{n-1})] = \prod_{i=0}^{n-1} \mathbb{E}[f_i(z_i)].$$
We have

\[ \mathbb{E}[f_0(z_0), f_1(z_1), \ldots, f_{n-1}(z_{n-1})] = \mathbb{E}[\mathbb{E}[f_0(z_0), f_1(z_1), \ldots, f_{n-1}(z_{n-1})|F_{t_1}]] \\
= \mathbb{E}[f_0(z_0)|\mathbb{E}[f_1(z_1), \ldots, f_{n-1}(z_{n-1})|X_{t_1}]] \\
= \mathbb{E}[f_0(z_0)|\mathbb{E}[f_1(z_1), \ldots, f_{n-1}(z_{n-1})|X_{t_1}]] \\
= \mathbb{E}[f_0(z_0)|\mathbb{E}[f_1(z_1), \ldots, f_{n-1}(z_{n-1})|X_{t_1}]] \quad (X_t \text{ is a Markov chain}) \\
\]

Continuing this way by conditioning on \( F_{t_2}, F_{t_3}, \ldots, F_{t_n} \), we obtain the result. We can show the other claim similarly.

\[ \square \]

**Examples (Single queue):**

1. \( M/M/1/FCFS \) is quasireversible.
2. \( M/GI/1/FCFS \) is not quasireversible unless service times are exponential.
3. \( M/GI/\infty \) is quasireversible.
4. \( M/GI/1/PS \) is quasireversible.
5. \( M/GI/1/LCFS \) is quasireversible.

In (2) – (4), \( q_t \) is not a Markov chain. However, \( X_t = (q_t, r_t) \) is a Markov chain, where \( r_t \) is residual service time of the customers in service, a real number. But, this is not a countable state Markov chain which we have been assuming so far. To overcome this problem, we use phase type distributions.

**Phase type distribution:**

Let \( R_t \) be a finite state Markov chain with state space \( \{1, 2, 3, \ldots, m + 1\} \) and generator matrix

\[ Q = \begin{bmatrix} Q_m & q_0 \\ q_1 & q_2 \end{bmatrix}, \]

where \( Q_m \) is an \( m \times m \) matrix.

Define \( \tau = \inf\{t : R_t = m + 1\} \). This will be a service time of a customer. Now, \( X_t = (q_t, R_t) \) is a finite state Markov chain. We have,

\[ P[\tau > t] = \alpha \exp \mu t - 1, \]

where \( \alpha \) is the distribution of \( R_0 \) and \( 1 = [1, 1, \ldots, 1]^T \). This is called phase type distribution with parameters \((\alpha, Q_m)\).

Any distribution on \( \mathbb{R}^+ \) can be arbitrarily closely approximated by a phase type distribution. In the following, we will take the general distribution of service times as a phase type distribution. Then, \((q_t, R_t)\) will be a countable state Markov chain.

Consider \( X_t \) in \( M/GI/1/LCFS \). The arrival process is a Poisson process. Thus, \( X_t \) and future arrivals are independent of each other. Let \( \hat{X}_t \) be the reversed process.

**Proposition 7.4.3.** In \( M/GI/1/LCFS \) system, the reversed process \( \hat{X}_t \) also represents a \( M/GI/1/LCFS \) system.

Thus, \( X_t \) and past departures are independent. Hence, \( M/GI/1/LCFS \) is a quasireversible system. Similarly, we can show that other queues in the above example are quasireversible.

For \( M/M/1, M/GI/1/LCFS \) and \( M/GI/PS \), if \( \rho < 1 \), \( \{q_t\} \) has a unique stationary distribution

\[ \pi(n) = (1 - \rho)\rho^n, \quad \forall n \geq 0. \]

87
For a multiclass queue, we have Eq (7.3) as stationary distribution. For $M/G/\infty$, for any $0 < \rho < \infty$,

$$
\pi(n) = \frac{\rho^n}{n!} e^{-\rho}, \forall n \geq 0.
$$

For all these cases, we observe that the stationary distribution depends on service distribution only through its mean. This property is called insensitivity.

All the examples given above for quasireversible queues are special cases of a quasireversible queue called the symmetric queue.

The above results are shown for phase type service types. Using continuity arguments, these results can be extended to general service times.

### 7.4.2 Networks of Quasireversible Queues

We now consider a multiclass queueing network of queues where each queue is quasireversible in isolation with Poisson input. Let $C$ be a countable set of classes of customers. Arrivals of different classes are independent Poisson processes. Let $\lambda_i^c$ be the external arrival rate of class $c$ customers at node $i$. Let $p_{ij}^{cd}$ be the probability that a class $c$ customer after service from node $i$ goes to node $j$ as a customer of class $d$. Denote by $\bar{\lambda}_i^c$ the total arrival rate of customers of class $c$ at node $i$. Then,

$$
\bar{\lambda}_i^c = \lambda_i^c + \sum_j p_{ij}^{cd} \bar{\lambda}_j^d. \quad (7.4)
$$

Let $1/\mu_i^c$ be the mean service time for a class $c$ customer at node $i$. Let $\rho_i^c = \bar{\lambda}_i^c / \mu_i^c$. The total traffic intensity at node $i$ is $\rho_i = \sum_c \rho_i^c$. Then, if $\rho_i < 1$, $\forall i$, the system has product form distribution as in Eq (7.3) with corresponding $\bar{\lambda}_i^c$ and $\mu_i^c$. The proof for this can be obtained by verifying that the distribution in Eq (7.3) solves $\pi Q = 0$. We can also show that the system is quasireversible by verifying Eq (7.6) and (7.7) below.

Now, we provide conditions to verify quasireversibility of general Markovian queueing systems. The state of the system $X_t$ is a Markov chain. Let $S$ denote the state space and $Q$ the generator matrix of $\{X_t\}$. Let $N_i^c(t)$ be the arrival process of class $c$ and $N_i^-(t)$ be the departure process of class $c$ from the network. Let $\mu_i^c$ and $\mu_i^-$ be the arrival rate and the departure rate of class $c$ customers respectively.

Define for $i, j \in S$

$$
A^c = \{(i, j) \ s.t \ i \rightarrow j \text{ represents an arrival for class } c\}, \text{ and }
D^c = \{(i, j) \ s.t \ i \rightarrow j \text{ represents a departure of class } c\}
$$

For example, in an $M/M/1$ queue with $C$ classes, $X_t = (c_1, c_2, \ldots, c_n)$, an arrival of class $c$ is

$$(c_1, c_2, \ldots, c_n) \rightarrow (c_1, c_2, \ldots, c_n, c) \in A^c$$

and a departure of class $c_1$ is

$$(c_1, c_2, \ldots, c_n) \rightarrow (c_2, \ldots, c_n) \in D^{c_1}.$$  

Then, if $X_t = i, i \in S$, arrival rate of class $c = \sum_{j(i, j) \in A^c} q_{ij}$. Arrival rate of class $c$ under stationarity ($\pi$ is the stationary distribution) is

$$
\sum_i \pi(i) \sum_{j(i, j) \in A^c} q_{ij}. \quad (7.5)
$$
But, $\{X_t\}$ is a quasi reversible process. Therefore, $X_t$ is independent of future arrivals. Thus, $\sum_{j:(i,j)\in A^c} q_{ij}$ does not depend on $i$. Thus Eq (7.5) equals

$$
\sum_{j:(i,j)\in A^c} q_{ij} \sum_i \pi(i) = \sum_{j:(i,j)\in A^c} q_{ij} = \mu_i^+. \tag{7.6}
$$

This is independent of $i$.

Now, consider the reversed process $\tilde{X}_t$. This is also quasireversible. Its stationarity distribution is also $\pi$ with $\tilde{Q}$ given by

$$
\tilde{q}_{ij} = \frac{\pi(j)q_{ji}}{\pi(j)}. 
$$

The arrival rate of the class $c$ customers in $\tilde{X}_t$ is

$$
\tilde{q}_{ij} = \sum_{j:(i,j)\in D^c} \frac{\pi(j)q_{ji}}{\pi(j)} = \mu_i^-. \tag{7.7}
$$

This also does not depend on $i$.

We can show that if Eq (7.6) and (7.7) hold for a Markovian queueing system, then it is quasireversible.
7.4.2 Networks of quasireversible queues (contd.)

Let $X_t = (X_t(1), X_t(2), \ldots, X_t(N))$ denote the state of the network with $N$ quasireversible queues and $X_t(i)$ denotes the state of the $i^{th}$ queue at time $t$. The queues can be of different type as long as they are quasireversible in isolation.

In the last lecture, we showed the following.

If $\rho_i < 1 \forall i \in 1, 2, \ldots, N$ (if the queue needs it for stability, e.g., $M/G/\infty$ does not), then $X_t$ has the stationary distribution

$$\pi(x(1), x(2), \ldots, x(N)) = \prod_i \pi_i(x(i))$$

where $\pi_i$ is the stationary distribution of queue $i$.

Now, we claim that $X_t$ itself is a quasireversible process by showing that future arrivals, $X_t$ and past departures are independent. Consider the reversed process $\tilde{X}_t = X_{T-t}$ for some fixed time $T$ with generator $\tilde{Q}$. We have

$$\pi(i)\tilde{Q}(i, j) = \pi(j)Q(j, i).$$

Here, $\tilde{Q}$ corresponds to the another quasireversible system with parameters

$$\hat{p}_{ij}^{cd} = \frac{X_j}{\lambda_i} p^{dc}_{ji}.$$

This shows that future arrivals, $X_t$ and the past departures are independent. Also, the departures of all classes from the network form independent Poisson processes.

Sojourn times:

We can compute the mean queue length or the mean number of customers in the system under stationarity from $\pi$. The mean sojourn time $E[S]$ can be deduced by applying Little’s law to the whole system: $E[S] = \lambda E[q]$, where $E[q]$ is the mean number of customers in the system and $\lambda = \sum \lambda_i$ is the total external arrival rate into the system. Further, Little’s law can be applied to each class of customers individually. The mean sojourn time of a class $c$ customer $E[S_c] = \lambda_c E[q_c]$ where $E[q_c]$ is the mean number of customers of class $c$ in the system.

Next, consider a tandem of $N M/M/1$ queues with service rate $\mu_i$ at queue $i$ and arrival rate $\lambda_i$ under stationarity. Let the random variable $S'$ denote the sojourn time of a customer in queue $i$. Under stationarity, we know that the departure process of queue $1$ is Poisson with rate $\lambda$. This is also the arrival process of queue $2$. Thus, the second queue is also an $M/M/1$ queue. This way we can show that the arrival and departure processes for all queues are Poisson with rate $\lambda$. If an arriving customer at queue $i$ sees $n$ customers already in the queue, then $S' = \sum_{k=1}^{n+1} s_k$ where $s_k$ are the service times, which are i.i.d. with exponential distribution with mean $1/\mu_i$. By PASTA, the probability that an arriving customer sees
n customers already in the queue is equal to the stationary probability of \( q_i(i) = n \). Therefore,

\[
\mathbb{P}\{S^i \leq x\} = \sum_{n=0}^{\infty} \mathbb{P}\{S^i \leq x|q_i = n\} \mathbb{P}_{\pi}\{q_i = n\} = \sum_{n=0}^{\infty} \left( \sum_{k=1}^{n+1} s^k \leq x \right) \rho^n (1 - \rho_i).
\]

We can show (e.g., by taking moment generating function) that the above quantity is an exponential distribution with mean \( 1/(\mu_i - \lambda) \). This is applicable to all the queues. The total sojourn time is \( S = \sum_{i=1}^{N} S^i \). Furthermore, it can also be shown that \( S^1, S^2, \ldots, S^N \) are independent random variables.

The above results hold for a general network of quasireversible queues. This is summarized below.

**Definition 7.4.4.** An *overtake free path* for a class \( c \) is a path \( 1 \to 2 \to \ldots \to N_1 \) of queues it can pass through sequentially with a positive probability such that if any customer can go through its two queues with positive probability, then it must pass through the all intermediate nodes as well in the same order.

**Theorem 7.4.5.** For a network of quasireversible queues, the sojourn times \( \{S^1, S^2, \ldots, S^{N_1}\} \) under stationarity of customers of a class \( c \) on an overtake free path \( 1 \to 2 \to \ldots \to N_1 \) are independent and \( S^i \) is exponentially distributed with mean \( 1/(\mu_i - \lambda_i) \). \( \Box \)

### Total arrival process at a queue:

Consider an \( M/M/1 \) queue with feedback. A customer after service re-enters the queue with probability \( p \) and exits the system with probability \( 1 - p \). Let the external arrival rate be \( \lambda \) and the aggregate arrival rate (including from feedback) be \( \lambda' \). We have the relation \( \lambda' = \lambda + p\lambda \) from which we find

\[
\lambda' = \frac{\lambda}{1 - p}.
\]

We show that the aggregate arrival process is not a Poisson process. Let \( N_t \) denote the aggregate arrival process. For small enough \( \varepsilon > 0 \),

\[
\mathbb{P}\{N_{t+\varepsilon} \geq N_t + 1\} = \lambda \varepsilon + o(\varepsilon) + \mu \varepsilon \mathbb{P}_{\pi}\{q_i > 0\}
\]

where \( \mu \varepsilon \mathbb{P}_{\pi}\{q_i > 0\} \) is the probability of a customer finishing service at time \( t \). We also have

\[
\mathbb{P}\{N_{t+\varepsilon} \geq N_t + 1|N_t \geq N_{t-\varepsilon} + 1\} = \lambda \varepsilon + o(\varepsilon) + \mu \varepsilon.
\]

The above equation follows from the fact that the event \( \{N_t \geq N_{t-\varepsilon} + 1\} \) implies \( \{q_i > 0\} \). From these two equations, we see that \( \mathbb{P}\{N_{t+\varepsilon} \geq N_t + 1|N_t \geq N_{t-\varepsilon} + 1\} \neq \mathbb{P}\{N_{t+\varepsilon} \geq N_t + 1\} \). This shows that \( N_t \) is not an independent increment process and hence cannot be Poisson.

As a generalization of this result, we get

**Theorem 7.4.6.** In an open quasireversible network, the aggregate arrival process at a node is not Poisson if there is non-zero probability of a customer entering that node.

### Queue lengths seen by an arriving customer:

Consider again an \( M/M/1 \) queue. Define

- \( q_t = \) queue length at time \( t \)
- \( q_n^A = \) queue length just before \( n^{th} \) arrival
- \( q_n^D = \) queue length just after \( n^{th} \) departure
- \( \tilde{q}_n^A = \) queue length just after \( n^{th} \) arrival

91
Figure 7.3: A network of quasireversible queues with deterministic routing for class 1

Under stationarity, we have shown that for an $M/GI/1$ queue $q_{tn}^A = d_t$ (PASTA) and $q_{tn}^D = d_t^A$ (rate conservation law) for $GI/GI/1$ queue. But, $P[q_{tn}^A = 0] = 0$ as there is always at least one customer soon after a new arrival. Also, $P[q_{tn}^A = 0] \neq P(q_t = 0)$. Now, consider a tandem of two $M/M/1$ queues. Let $S_n(i)$ be the time instant of an arrival into queue $i$. $q_{S_n(2)}^D + (1)$ denotes the length of queue 1 just after a departure from queue 1. $q_{S_n(2)}^A - (2)$ denotes the length of queue 2 just before $n^{th}$ arrival into queue 2. An argument as in preceding paragraph shows that the stationary distribution of $q_{S_n(2)}^A - (2)$ is not the same as that of $q_t(2)$. But, we can show that the distribution of $(q_{S_n(2)}^D + (1), q_{S_n(2)}^A - (2))$ under stationarity is

$$\pi(n_1, n_2) = \rho_1^{n_1}(1 - \rho_1)\rho_2^{n_2}(1 - \rho_2).$$

The above discussion is true in an open network of quasireversible queues.

**Theorem 7.4.7.** Let $X_t = (X_t(1), X_t(2), \ldots, X_t(N))$ be the state of an open network of $N$ quasireversible queues. Let $\pi$ be the stationary distribution of $X_t$. Under stationarity, for a customer moving from queue $i$ to queue $j$ at time $S_n$,

$$(X_{S_n}(1), \ldots, X_{S_n}(i), \ldots, X_{S_n} - (j), \ldots, X_{S_n}(N)) \overset{d}{=} \pi.$$

**Non-Markovian routing:**

We give an example to show how non-Markovian routing can be taken care of in this framework. Consider a network of 4 quasireversible queues in Figure 7.3. Class 1 customers enter queue 1 and class 2 customers enter queue 2. After service in queue 4, the class 2 customers exit the system with probability $p$ and with probability $1 - p$ re-enter queue 2. The class 1 customers follow deterministic routing - they enter the network in queue 1 and after service in queue 4, they re-enter queue 1 exactly 3 times before exiting the system. This kind of routing cannot directly be modeled in a way we have been doing so far. However, it is possible to bring the network into our framework by introducing additional classes. We introduce new classes 3 and 4 as follows: after service in queue 4 for the first time, a class 1 customer changes to class 3 and the second time, the same customer changes to class 3 and the second time, the same customer changes to class 4 (from class 3). With this, $p_{41}^{13} = 1$, $p_{41}^{34} = 1$ and $p_{40}^{44} = 1$ satisfies the constraints of class 1 customers. We can now use the same theory to obtain the stationary distribution by finding total arrival rate at each node and computing $\rho_i$'s and multiplying distributions for individual queues.
7.5 Problems

Problem 1: (Random Walks) \( \{X_n\} \) iid, \( E[X_1] = \mu \), \( 0 < \mu < \infty \). \( S_0 = 0 \), \( S_n = \sum_{k=1}^{n} X_k \).

1. Show \( P[v(t) < \infty] = 1 \) for all \( 0 < t < \infty \).
2. Show \( \{v(t) > n\} = \{M_n \leq t\} \) and \( v(t) \to \infty \) a.s as \( t \to \infty \).
3. Using strongly ascending ladder heights, show \( \frac{v(t)}{t} \to \frac{1}{\mu} \) a.s.

Problem 2: (GI/M/1 queue) Let \( \{A_n\} \) i.i.d. interarrival time to a queue with a general distribution. The service times are i.i.d. \( \text{exp}(\mu) \).

1. Consider \( q_n \) = the queue length just before \( n \)th arrival. Study the stability conditions for it.
2. Using (1) obtain the stability conditions for the waiting time process. Also for \( q_t \) and \( V_t \).

Problem 3: (Queue with priority) Consider a queue with two classes of traffic. Class-1 gets Poisson arrivals with rate \( \lambda_1 \) and class-2 with rate \( \lambda_2 \). All the service times are i.i.d. with a general distribution and mean \( \frac{1}{\mu} \). Class-1 has preemptive resume priority over class-2.

1. Study the stability (Existence of stationary distribution etc.).
2. Find probability under stationarity that customer of class-\( i \), \( i = 1, 2 \) experiences zero delay in the queue.
3. Find the mean delay of each class under stationarity.

Problem 4: Consider two queues in tandem with Poisson arrivals with rate \( \lambda \) to \( Q_1 \). The service times are exponential i.i.d. with rate \( \mu_i \) in \( Q_i \). Let \( \lambda < \mu_i, i = 1, 2 \).

1. Compute the stationary distribution of queue length seen by arrivals to \( Q_2 \). Buffer length of each queue is infinite.
2. Assume \( Q_1 \) has infinite buffer but \( Q_2 \) has finite buffer of size \( N \). Find the stationary probability of buffer overflow at \( Q_2 \).
3. Let \( Q_i \) has finite buffer \( N_i < \infty, i = 1, 2 \). Compute the stationary probability of buffer overflow in \( Q_i, i = 1, 2 \). Also compute the mean sojourn time \( E_x S_i \) in \( Q_i, i = 1, 2 \). Also compute stationary distribution of \( S_i \).

Problem 5: Consider a single queue with Poisson arrivals with rate \( \lambda \), Processor sharing, mean service time \( \frac{1}{\mu} \). After completion of service a customer is fed back with probability \( p \) and leaves the system with probability \( 1-p \). Assume the system is under stationarity.

1. Find the distribution of queue length seen by an external arrival. Also find distribution of its sojourn time.
2. Find the distribution of queue length seen by a customer fed back.
3. Which of the following flows are Poisson:
   - Fed back customers.
• Customers completing service at the server.
• Customers leaving the network.

4. Solve (1)-(3) if the queue length buffer is of length $N$. In part-3 also check the flow of external customers entering the queue. Also compute the probability of external arrivals getting lost at the queue and the probability of a fed back customer getting lost.

**Problem 6:** Consider an open Jackson network with three nodes with exogenous arrivals to each queue as Poisson with rate $\lambda_i, i = 1, 2, 3$ and exponential i.i.d. service rates $\mu_i, i = 1, 2, 3$. The Markovian routing probabilities are $p_{12} + p_{13} = 1, p_{21} + p_{23} = 1, p_{30} = 1$.

1. Find conditions for $q(t) = (q_1(t), q_2(t), q_3(t))$ to be a stable Markov chain.
2. On which arcs in the network the flows are Poisson under stationarity.
3. Find the mean sojourn time in the network under stationarity.
4. Find distribution of sojourn time in $Q_3$ under stationarity.
5. Find the distribution of sojourn time on a visit of a customer from $Q_1$ to $Q_2$.
6. Answer all above if $Q_3$ has a general service time distribution with Process sharing.

**Problem 7:** (Window flow control) A source transmits its packets through a queue (router) to the destination via a window flow control mechanism. The window size is $N$. Packets from the source enter $Q_1$ and are served in i.i.d. $\text{exp}(\mu_1)$ in FCFS fashion. Whenever the destination receives a packet, it immediately releases an acknowledgement in $Q_2$. Service times in $Q_2$ are i.i.d. $\text{exp}(\mu_2)$. At any time the sum of packets and acks in the system is $N$. Whenever an ack reaches source, it releases the next packet to $Q_1$. (This models a simplistic version of TCP and is a closed queueing network.)

1. Find conditions for stationary distribution of queue lengths in network.
2. Find rate at which packets are released by source.
3. Find the mean sojourn time of packets in $Q_1$.

**Problem 8:** Consider a $GI/GI/1$ queue with last come first serve preemptive resume discipline. Whenever a customer arrives, server leaves other customers and starts serving a new one. Whenever a server completes a service it goes back to previous customer to complete its service. Find conditions for its stability. Show its mean sojourn time equals busy period of $GI/GI/1$ with FCFS.