ISyE 6762 (Winter 2017)

Stochastic Processes II

Prof. D. A. Goldberg
Georgia Institute of Technology

LaTeXer: W. Kong
http://wwkong.github.io
Last Revision: April 28, 2017

Table of Contents

These notes are currently a work in progress, and as such may be incomplete or contain errors.

ACKNOWLEDGMENTS:

Special thanks to *Michael Baker* and his LaTeX formatted notes. They were the inspiration for the structure of these notes.

Abstract

The purpose of these notes is to provide the reader with a secondary reference to the material covered in ISyE 6762.

Recommended Readings

* Sid Resnick's *A Probability Path*

* Grimmett and Stirzaker

* Ross' Blue Book on Stochastic Processes

Network Flows' Books

Schrijver's Book on Integer Programming

Useful Books

* Counter-examples in Real Analysis

* Counter-examples in Probability

Course Description

* General techniques in probability theory

* Probability as a modeling tool

* Modes of convergence

* Continuous Time Markov Chains (CTMC)

* Martingales

* Brownian Motion

Grading

25% Homework

45% Midterms (15% each for 3 sittings)

30% Final Exam

Office Hours

Friday, 3-4:30pm

Tuesday, 3-4pm

1 Probability Theory

1.1 Modes of Convergence

Definition 1.1. A sequence of r.v.s $\{X_{n,n\geq 1}\}$ converges **in distribution** to a r.v. X_∞ if for all x which are **continuity points** (points where the c.d.f. is continuous) of the c.d.f. of X_∞, it holds that

$$\lim_{n\to\infty} P(X_n \leq x) = P(X_\infty \leq x)$$

Example 1.1. Suppose for $n \geq 1$, $X_n = \{1 + \frac{1}{n}$ with probability (w.p.) 1. Through limiting arguments, one can show that the limit of the c.d.f.s for X_n converge to the function

$$F(x) = \begin{cases} 0 & x \leq 1 \\ 1 & x > 0 \end{cases}$$

which is not a c.d.f..

Definition 1.2. A set S is countable if there exists an injective map $f : S \mapsto \mathbb{N}$.

Fact 1.1. *The set of rational numbers $\mathbb{Q}$ is countable.*

Claim 1.1. A c.d.f. can only have "very few" (countable) discontinuity points.

Proof. Note that:

1) A c.d.f. is a **uniformly bounded** (u.b.) (in $[0,1]$) non-decreasing function.

2) For every non-decreasing u.b. function f with domain $\mathbb{R}$,

$$f^-(x) = \lim_{z \to x^-} f(z) \text{ and } f^+(x) = \lim_{z \to x^+} f(z)$$

exist for all $x \in \mathbb{R}$. In our setup, $f^-(x) = \sup_{z<x} f(z)$ and $f^+(x) = \inf_{x<z} f(z)$. □

Let $D_n = \{x \in \mathbb{R} : f^+(x) - f^-(x) \geq \frac{1}{n}\}$. Clearly, $|D_n| \leq n$. Now since

$$\begin{aligned} f \text{ is discontinuous at } x &\iff f^+(x) - f^-(x) > 0 \\ &\iff \exists n \geq 1 \text{ s.t. } f^+(x) - f^-(x) \geq 1/n \\ &\iff \exists n \geq 1 \text{ s.t. } x \in D_n \\ &\iff x \in \bigcup_{n=1}^{\infty} D_n \end{aligned}$$

then the set of discontinuities is $\bigcup_{n=1}^{\infty} D_n$ which is a countable collection of countable sets. Hence, the set of discontinuities of a c.d.f. is countable.

Example 1.2. Let $q(i)$ be i^{th} rational in an ordering of rationals under the ordering $\leq$ where $q(i) < q(i+1)$ for all $i \in \mathbb{N}$. Consider the random variable X which takes on the value $q(i)$ with probability 2^{-i} for $i = 1, 2, ...$.

Example 1.3. Consider the sequence of random variables $\{X_n\}$ where $X_n \sim \text{Ber}(\frac{1}{n})$. This limit will converge to a constant 0 random variable but there is some probability that for a given index N, we may see $X_n = 1$ for some $n > N$.

Example 1.4. Consider the sequence $\{U\}_{n\in\mathbb{N}}$ where $U \sim \text{Unif}(0,1)$. This sequence has a limit with probability 1, but the limit is random.

1.2 Probability Spaces

Example 1.5. Suppose that $P(X = Y = 0) = 8, P(X = Y = 2) = 2$. Generate a $U(0,1)$ variable w and note that we can simulate

$$X(w) = \begin{cases} 2 & \text{, if } w \in [0, 0.2) \\ 0 & \text{, otherwise} \end{cases}, Y(w) = \begin{cases} 2 & \text{, if } w \in [0, 0.2) \\ 0 & \text{, otherwise} \end{cases}$$

What if we want the same marginals but with $X \perp Y$? We can create:

$$X(w) = \begin{cases} 2 & \text{, if } w \in \bigcup_{k=0}^{9} \left[\frac{k}{10}, \frac{k}{10} + 0.02\right] \\ 0 & \text{, otherwise} \end{cases}, Y(w) = \begin{cases} 2 & \text{, if } w \in \bigcup_{k=0}^{9} \left[\frac{k}{100}, \frac{k}{100} + 0.002\right] \\ 0 & \text{, otherwise} \end{cases}$$

Example 1.6. To generate an infinite sequence of random uniform variables $\{U_i\}_{i\in\mathbb{N}}$ from a single uniform random variable U is to denote the i^{th} digit of U_j to be $(p_j)^i$ where p_j is the j^{th} prime number.

Definition 1.3. A **sample space** Ω is the set of possible outcomes for the "underlying randomness"; here, we will default to $[0,1]$.

A **filtration** $\mathcal{F}$ is a set of sets which subsets for which we are obliged to assign probabilities to. For $[0,1]$ this defaults to $\mathcal{B}_{[0,1]}$.

A **probability measure** $\mathcal{P} : \mathcal{F} \mapsto [0,1]$ is an assignment of probabilities to sets in $\mathcal{F}$ consistent with the laws of probability. It is sufficient to assign $\mathcal{P}$ to all intervals (Caratheodory Extension Theorem).

A **probability space** is a triple of the previous three $(\Omega, \mathcal{F}, \mathcal{P})$.

A **random variable** is any function $f : \Omega \mapsto \mathbb{R}$ satisfying the "measurability" property with respect to $\mathcal{F}$.

An **event** is an element of the filtration $\mathcal{F}$. Note that because they are elements of the filtration, we can assign probabilities to these events through $\mathcal{P}$.

Definition 1.4. A sequence of random variables $\{X_{n,n\geq 1}\}$ **converges in probability** to a non-defective limiting random variable X_∞ if
$$\forall \epsilon \geq 0, \lim_{n\to\infty} P(|X_n - X_\infty| > \epsilon) = 0$$

Example 1.7. Consider the sequence of iid random variables $\{B_n\}_{n\in\mathbb{N}}$ where $B_n \sim \mathrm{Ber}(1-\frac{1}{n})$ and the sequences of iid random variables $\{U\}_{n\in\mathbb{N}}, \{U_n\}_{n\in\mathbb{N}}$ where $U, U_n \sim \mathrm{Unif}(0,1)$. Define $X_n = B_n \cdot U$ and $Y_n = B_n \cdot U_n$ where

$$X_n = \begin{cases} U & \text{w.p. } 1-\frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

converges in probability to U while Y_n does not. To

see why X_n does, note that for all $\epsilon > 0, n \geq 1$,

$$P(|X_n - X_\infty| > \epsilon) \leq 1/n$$

where $X_\infty = U$ and because

$$\begin{aligned} &\{|X_n - X_\infty| > \epsilon\} \subseteq \{|X_n - X_\infty| \neq 0\} \\ \implies &P(|X_n - X_\infty| > \epsilon) \leq 1/n. \end{aligned}$$

To see that Y_n does not, we first make the following claim: $\{Z_n, n \geq 1\}$ converges in probability to Z_∞ implies that

$$\forall \epsilon > 0, \lim_{n\to\infty} P(|Z_n - Z_{n+1}| > \epsilon) = 0.$$

To see this, note that

$$\begin{aligned} &\{|Z_n - Z_{n+1}| > \epsilon\} \subseteq \{|Z_n - Z_\infty| > \epsilon/2\} \cup \{|Z_{n+1} - Z_\infty| > \epsilon/2\} \\ \implies &P(|Z_n - Z_{n+1}| > \epsilon) \leq P(|Z_n - Z_\infty| > \epsilon/2) + P(|Z_{n+1} - Z_\infty| > \epsilon/2) \\ \implies &\limsup_{n\to\infty} P(|Z_n - Z_{n+1}| > \epsilon) \leq \limsup_{n\to\infty} P(|Z_n - Z_\infty| > \epsilon/2) + \limsup_{n\to\infty} P(|Z_{n+1} - Z_\infty| > \epsilon/2) \end{aligned}$$

and the result on Y_n follows as the consecutive differences do not converge.

Proposition 1.1. *Convergence in probability $\implies$ Convergence in distribution.*

Proposition 1.2. *Convergence in distribution to a constant $\implies$ Convergence in probability.*

Proof. We have

$$\begin{aligned} P(|X_n - c| > \epsilon) &\leq P(X_n \leq c - \epsilon) + P(X_n > c + \epsilon) \\ &= \underbrace{P(X_n \leq c - \epsilon)}_{\to 0} + 1 - \underbrace{P(X_n \leq c + \epsilon)}_{\to 1} \end{aligned}$$

□

Definition 1.5. Suppose $\{A_n, n \geq 1\}$ is a sequence of events. We define the event "A_n **happens infinitely often**" (written "A_n i.o.") as $\bigcap_{n=1}^\infty \bigcup_{m=n}^\infty A_m$ or interpreted as "no matter how far out in n you go, at least one more event after n will happen". We say $\{X_n, n \geq 1\}$ converges **almost surely** (a.s.) to a non-defective random variable $X_\infty \iff$ for all $\epsilon > 0$, letting $A_{n,\epsilon} = \{|X_n - X_\infty| > \epsilon\}$ we have
$$P(A_{n,\epsilon} \text{ i.o.}) = 0.$$

More generally,
$$\{X_n, n \geq 1\} \text{ converges a.s. } \iff P(\lim_{n\to\infty} X_n \text{ exists}) = 1$$
or for all $w \in \Omega'$ where $P(\Omega \backslash \Omega') = 0$ we have $X_\infty(w) = \lim_{n\to\infty} X_n(w)$.

Axiom 1. *(Countable additivity) For any $P-$measure and collection of mutually disjoint events $\{A_n, n \geq 1\}$ we have*

$$P\left(\bigcup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty P(A_n)$$

Proposition 1.3. *If $\{A_n, n \geq 1\}$ is a monotone increasing collection of events, i.e. $A_n \subseteq A_{n+1}$ for all n, in which case $A_n = \bigcup_{i=1}^n A_i$, it holds that*

$$P\left(\bigcup_{k=1}^{\infty} A_n\right) = \lim_{N\to\infty} P\left(\bigcup_{n=1}^{N} A_N\right) = \lim_{N\to\infty} P(A_N).$$

Proposition 1.4. *If $\{A_n, n \geq 1\}$ is a monotone decreasing collection of events, i.e. $A_n \supseteq A_{n+1}$ for all n, in which case $A_n = \bigcap_{i=1}^n A_i$, it holds that*

$$P\left(\bigcap_{k=1}^{\infty} A_n\right) = \lim_{N\to\infty} P\left(\bigcap_{n=1}^{N} A_N\right) = \lim_{N\to\infty} P(A_N).$$

1.3 Borel-Cantelli Lemma

Lemma 1.1. *(Borel-Cantelli Lemma) For any sequence of events $\{A_n, n \geq 1\}$, we have*

$$\sum_{n=1}^{\infty} P(A_n) < \infty \implies P(A_n \text{ i.o.}) = 0$$

Proof. Directly,

$$\begin{aligned} P(A_n \text{ i.o.}) &= P\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m\right) \\ &= \lim_{N\to\infty} P\left(\bigcup_{m=N}^{\infty} A_m\right) \\ &\leq \lim_{N\to\infty} P\left(\sum_{m=N}^{\infty} A_m\right) \to 0 \end{aligned}$$

□

Example 1.8. The converse is generally not true. Consider $\Omega = [0,1]$ and $A_n = I(U \in [0, 1/n])$. Then

$$\begin{aligned} P(A_n \text{ i.o.}) &= P\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m\right) \\ &= P\left(\bigcap_{n=1}^{\infty} A_n\right) \\ &= \lim_{N\to\infty} P(A_N) \\ &= \lim_{N\to\infty} \frac{1}{N} \end{aligned}$$

from the monotone decreasing property of A_n and the continuity of measures.

Proposition 1.5. *(Reverse Borel-Cantelli Lemma) If you have a mutual independent collection of events $\{A_n, n \geq 1\}$ then*

$$\sum_{n=1}^{\infty} P(A_n) = \infty \implies P(A_n \text{ i.o.}) = 1.$$

Proof. Directly,

$$\begin{aligned}
P\left([A_n \text{ i.o.}]^c\right) &= P\left(\left[\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m\right]^c\right) \\
&= P\left(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m^c\right) \\
&= \lim_{N\to\infty} P\left(\bigcap_{m=N}^{\infty} A_m^c\right) \\
&= \lim_{N\to\infty} \prod_{m=N}^{\infty} P(A_m^c) \\
&= \lim_{N\to\infty} \prod_{m=N}^{\infty} [1 - P(A_m)] \\
&\le \lim_{N\to\infty} \prod_{m=N}^{\infty} e^{-P(A_m)} \\
&= \lim_{N\to\infty} e^{-\sum_{m=N}^{\infty} P(A_m)} \\
&= \lim_{N\to\infty} 0 = 0
\end{aligned}$$

□

Example 1.9. Consider the set of i.i.d. random variables $\{X_n, x \ge 1\}$ where

$$X_n = \begin{cases} 1 & \text{w.p. } 1/n \\ 0 & \text{w.p. } 1 - 1/n \end{cases}$$

Let $A_n = I(X_n = 0)$ where $\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty}(1 - 1/n) = \infty$ and from the previous result, $P(A_n \text{ i.o.}) = 1$. In addition, if $B_n = I(X_n = 0)$ then $\sum_{n=1}^{\infty} P(B_n) = \sum_{n=1}^{\infty}(1/n) = \infty$ and $P(B_n \text{ i.o.})$. Hence, there is an infinite occurrence of 1 and 0 in the sequence of $\{X_n, n \ge 1\}$ and X_n does not converge almost surely.

Remark 1.1. There exist random variables with infinite mean, but with probability 1 they are finite. For example, we could use

$$P(X = n) = \frac{c}{n^2}, c = 6/\pi^2$$

where $P(X \ge N) = c \cdot \sum_{n=N}^{\infty} 1/n^2$ but $E[X] = c\sum_{n=1}^{\infty} 1/n = \infty$.

<u>More forms of a.s. convergence</u>

- (SLLN) $\{X_i, i \ge 1\}$ i.i.d., $E[X_1]$ exists and is finite $\implies \frac{\sum_{i=1}^{n} X_i}{n} \overset{a.s.}{\to} E[X_1]$.
- (WLLN) ?
- (CLT) $[(\sum_{i=1}^{n} X_i - nE[X])/(\sigma_x\sqrt{n})] \overset{d}{\to} N(0, 1)$

1.4 Interchanging Limit and Expectation

Problem 1.1. Suppose $\{X_n, n \ge 1\}$ converges a.s. to X_∞. Is it true that

$$\liminf_{n\to\infty} E[X_n] = \limsup_{n\to\infty} E[X_n] = E[X_\infty]?$$

NO. Consider $\{X_n, n \ge 1\}$ independent and defined by

$$X_n = \begin{cases} 2^n & \text{w.p. } 2^{-n} \\ 0 & \text{w.p. } 1 - 2^{-n} \end{cases}.$$

We have $\liminf_{n\to\infty} E[X_n] = \limsup_{n\to\infty} E[X_n] = 1$ but by the Borel-Cantelli Lemma, $X_n \overset{a.s.}{\to} 0$ and $E[X_\infty] = 0$. Worse, if we define

$$X_n = \begin{cases} 2^n & \text{w.p. } 2^{-n} \\ 0 & \text{w.p. } 1-2^{-n} \end{cases}, n \text{ is even}$$

$$X_n = \begin{cases} 2\cdot 2^n & \text{w.p. } 2^{-n} \\ 0 & \text{w.p. } 1-2^{-n} \end{cases}, n \text{ is odd}$$

then $1 = \liminf_{n\to\infty} E[X_n] \neq \limsup_{n\to\infty} E[X_n] = 2$.

Lemma 1.2. *(Fatou's Lemma [bounded from below]) If $\{X_n, n \geq 1\}$ is a sequence of r.v.s on some $P-$space, and on some $P-$space there exists a non-negative r.v. Z such that $P(X_n \geq Z) = 1$ for all n and $E[Z] < \infty$, then*

$$E\left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} E[X_n].$$

Lemma 1.3. *(Reverse Fatou's Lemma [bounded from above]) If $\{X_n, n \geq 1\}$ is a sequence of r.v.s on some $P-$space, and on some $P-$space there exists a non-negative r.v. Z such that $P(X_n \leq Z) = 1$ for all n and $E[Z] < \infty$, then*

$$E\left[\limsup_{n\to\infty} X_n\right] \geq \limsup_{n\to\infty} E[X_n].$$

Theorem 1.1. *(Dominated Convergence) Suppose that $\{X_n, n \geq 1\}$ is a sequence of r.v.s on a common probability space and $\{X_n, n \geq 1\}$ converges a.s. to X_∞. Suppose on some $P-$space there exists Z such that $E[Z] < \infty, P(|X_n| \leq Z) = 1$ for all n. Then, $\lim\limits_{n\to\infty} E[X_n]$ exists and*

$$\lim_{n\to\infty} E[X_n] = E\left[\lim_{n\to\infty} X_n\right] = E[X_\infty].$$

Theorem 1.2. *(Monotone Convergence) If $\{X_n, n \geq 1\}$ is a sequence of r.v.s such that $P(X_n \leq X_{n+1}) = 1$ for all n, then $\{X_n, n \geq 1\}$ converges a.s. [which comes from the fact that X_n is monotone] and $\lim\limits_{n\to\infty} E[X_n] = E[X_\infty]$.*

Example 1.10. Suppose that U is a fixed $U[0,1]$ r.v. and

$$X_n = \frac{n}{1+n^2\sqrt{U}}.$$

Then since $X_n \overset{a.s.}{\to} 0$ and

$$\frac{n}{1+n^2\sqrt{U}} \leq \frac{1}{\sqrt{U}}$$

with $E[1/\sqrt{U}] = 2 < \infty$ we can apply Dominated Convergence to get $\lim\limits_{n\to\infty} E[X_n] = E\left[\lim\limits_{n\to\infty} X_n\right]$. Note that if we had

$$X_n = \frac{n}{1+n^2U^2} \implies \limsup_{n\to\infty} E[X_n] = \liminf_{n\to\infty} E[X_n] = \frac{\pi}{2}.$$

Remark 1.2. If for all $\epsilon > 0$ and $n \in \mathbb{N}$ we have

$$P\left(\limsup_{n\to\infty} X_n < \epsilon\right) = P\left(\liminf_{n\to\infty} X_n > -\epsilon\right) = 1,$$

then we claim that $P\left(\lim\limits_{n\to\infty} X_n = 0\right) = 1$. Now in general if $P(-\epsilon < Z < \epsilon) = 1$, then

$$P(Z \neq 0) = P\left(\bigcup_{k=1}^{\infty}\left\{|Z| \geq \frac{1}{k}\right\}\right) \leq \sum_{k=1}^{\infty} P\left(|Z| \geq \frac{1}{k}\right) = \sum_{k=1}^{\infty}\left[1 - P\left(|Z| \leq \frac{1}{k}\right)\right] = 0$$

and $P(Z = 0) = 1$. So by symmetric arguments on

$$P\left(-\epsilon < \limsup_{n\to\infty} X_n < \epsilon\right) = 1$$
$$P\left(-\epsilon < \liminf_{n\to\infty} X_n < \epsilon\right) = 1$$

we get

$$P\left(\liminf_{n\to\infty} X_n = 0\right) = P\left(\limsup_{n\to\infty} X_n = 0\right) = 1 \implies \left(\lim_{n\to\infty} X_n = 0\right) = 1.$$

2 Continuous Time Markov Chains (CTMC)

Set-up

- Let $\{X_i : i \geq 0\}$ be the set of states visited by your discrete time Markov chain (DTMC) where X_0 is the initial state. In state i, you stay $Z_i \sim \exp(\gamma_i)$ where $P(Z_i > x) = e^{-\gamma_i x}, E[Z_i] = 1/\gamma_i$.
- Let $\{E_i, i \geq 0\}$ be a collection of i.i.d. $\exp(1)$ r.v.s.
- Let $X(t)$ be where the CTMC is at the time t.

Remark 2.1. We may construct our CTMC such that the amount of time W_i spent in the i^{th} state visited is E_i/γ_{X_i} since $E_i/\gamma_{X_i} \sim \exp(\gamma_{X_i})$.

Corollary 2.1. *We may construct $\{X(t), t \geq 0\}$ such that the time of the k^{th} transition equals*

$$T_k = \sum_{i=0}^{k-1} W_i = \sum_{i=0}^{k-1} \frac{E_i}{\gamma_{X_i}}, T_0 = 0.$$

Conjecture 2.1. *(?) We may think to construct $X(t)$ such that $X(t) = \sum_{k=0}^{\infty} X_k \times I\left(t \in [T_k, T_{k+1})\right)$ where $X(t)$ is an example of a càdlàg function. [This is WRONG]*

Example 2.1. (Chain Reaction) At time 0, there is a single particle (Gen. 0) with its own $\exp(1)$ clock. When its clock goes off, it disintegrates, and 2 new particles arise, each with its own $\exp(1)$ clock (Gen. 1). When the first Gen 1. clock goes off, all of Gen 1. disintegrates, each replaced by 2 new particles, each with its own $\exp(1)$ clock (Gen. 2). This process iterates for all subsequent generations, exploding by a factor of 2 each generation.

Let L_k be the lifetime of the k^{th} generation. Let $X(t)$ be the generation at time t.

(Q1) What is the distribution of L_k?

* (A1) L_k is the minimum of 2^k independent $\exp(1) \implies L_k \sim \exp(2^k)$.

(Q2) If $X(t)$ a CTMC? If so, give $\{\gamma_i\}$ and DTMC transition probabilities.

* (A2) For the DTMC, $P_{k,k+1} = 1$ for $k \geq 0$ and for the CTMC, $P(X_0 = 0) = 1, \gamma_i = 2^i, i \geq 0$.

* From our previous conjecture, we define

$$T_k = \sum_{i=0}^{K} W_i = \text{time at which generation } k \text{ ends},$$

Observe that

$$P\left(\exists k : T_k \geq 8\right) \leq \sum_{i=0}^{\infty} P(T_i \geq 8) \tag{1}$$

$$P\left(\exists k : T_k \geq 8\right) \leq P\left(\lim_{k\to\infty} T_k \geq 8\right) \tag{2}$$

where $\lim_{k\to\infty} T_k$ exists because for each $\omega \in \Omega$ we have that $\{T_k(\omega)\}_{k\in\mathbb{N}}$ is a monotone increasing sequence. Monotonicity is also what ensures the validity of (2). In general (1) is intractable as it will mostly be ∞. For example, consider

$$T_k = \begin{cases} 9 & \text{w.p. } 1/2 \\ 1 & \text{w.p. } 1/2 \end{cases}$$

where (1) will give a bound of ∞ while (2) will give a bound of $1/2$. Using (2), Markov's inequality and monotone convergence, we get

$$P\left(\lim_{k\to\infty} T_k \geq 8\right) \leq \frac{E\left[\lim_{k\to\infty} T_k\right]}{8} = \frac{\lim_{k\to\infty} E\left[T_k\right]}{8} = \frac{2}{8} = \frac{1}{4}$$

and $P\left(\exists k : T_k \geq 8\right) < 1$.

Definition 2.1. Letting $T_\infty = \lim_{k\to\infty} T_k$, we say that a chain for which $P(T_\infty = \infty) = 1$ for each initial starting state is **regular**. If $\{T_\infty < \infty\}$, we say that **explosion** has occurred, and occurs at T_∞.

Remark 2.2. Note that

$$\begin{aligned}
P(\text{no explosion}) &= P\left(\bigcap_{i\geq 1}\{\exists k : T_k \geq i\}\right) \\
&= \lim_{i\to\infty} P\left(\exists k : T_k \geq i\right) \\
&\leq \lim_{i\to\infty} \frac{2}{i} = 0
\end{aligned}$$

Remark 2.3. One sufficient condition for regularity is $\sup_i \gamma_i < \infty$. To see this, define $T_k = \sum_{i=0}^k \text{expo}_i(1)/\gamma_{X_i}$ where with probability 1, $T_k \geq \sum_{i=0}^k \text{expo}_i(1)/\sup_i \gamma_i$. WLOG, suppose $\sup_i \gamma_i = 1$. Then

$$P\left(\liminf_{k\to\infty} \frac{T_k}{k} \geq 1\right) = 1$$

from the SLLN where we know

$$\frac{\sum_{i=0}^k \text{expo}_i(1)}{k} \overset{a.s.}{\to} 1.$$

Now if $P(T_\infty < \infty) > 0$ then

$$P\left(\liminf_{k\to\infty} \frac{T_k}{k} = 0\right) = 1$$

which is impossible.

<u>Why is a CTMC Markovian?</u>

In other words, for all $t_1 < t_2 < ... < t_{n+1}$ and $k_1, ..., k_{n+1}$, why is

$$\begin{aligned}
&P(X(t_{n+1}) = k_{n+1} | X(t_1) = k_1, ..., X(t_n) = k_n) \\
=&P(X(t_{n+1}) = k_{n+1} | X(t_n) = k_n)
\end{aligned}$$

for a CTMC? The proof is very long, so we'll skip this (it's in Resnik's book somewhere).

2.1 Queuing Systems

Example 2.2. Consider a M/M/1 queue with arrival rate distributed Poisson(λ) and service time Expo(μ). Let $Q(t)$ be the number of customers at time t. A way to model this as a CTMC is

$$\begin{aligned}
&\gamma_0 = \lambda, \gamma_i = \lambda + \mu, i \geq 1 \\
&P_{0,1} = 1 \\
&P_{i,i-1} = \frac{\mu}{\mu + \lambda} \\
&P_{i,i+1} = \frac{\lambda}{\mu + \lambda}.
\end{aligned}$$

This tells us that the next event is an arrival is independent of which event occurs first. In particular, if $X_1 \sim \text{expo}(\lambda_1)$ and $X_2 \sim \text{expo}(\lambda_2)$, then

$$\begin{aligned}
&P(\min(X_1, X_2) \leq z, \min(X_1, X_2) = X_1) \\
=&\int_{x=0}^{x=z}\int_{y=x}^{\infty} \lambda_1\lambda_2 e^{-\lambda_1 x - \lambda_2 y}\, dy\, dx \\
=&\int_{x=0}^{x=z} \lambda_1 e^{-(\lambda_1+\lambda_2)x}\, dx \\
=&\frac{\lambda_1}{\lambda_1+\lambda_2}\left(1 - e^{-(\lambda_1+\lambda_2)z}\right)
\end{aligned}$$

and hence

$$\begin{aligned}
&P(\min(X_1, X_2) \leq z | \min(X_1, X_2) = X_1) \\
=&\frac{\frac{\lambda_1}{\lambda_1+\lambda_2}\left(1 - e^{-(\lambda_1+\lambda_2)z}\right)}{\frac{\lambda_1}{\lambda_1+\lambda_2}} \\
=&1 - e^{-(\lambda_1+\lambda_2)z} \\
=&P(\min(X_1, X_2) \leq z)
\end{aligned}$$

2.2 Chapman-Kolmogorov Equations

Define $P_{ij}(t) = P(X(t) = j | X(0) = i)$ and roughly

$$\begin{aligned}
P_{ij}(t+h) =& P\left(\text{starting in } i\text{; you are in } j \text{ at time } t\text{, and you stay there}\right) + \\
& P\left(\text{starting in } i\text{; you are somewhere else at time } t\text{, and jump to } j \text{ in } [t, t+h]\right).
\end{aligned}$$

Suppose that $\sup_i \gamma_i < \infty$ and $\gamma^* := 1 + \sup_i \gamma_i$. Precisely, in short form notation,

$$P_{ij}(t+h) = \sum_{k=0}^{\infty} P(X(t+h) = j | X(t) = k) P_{ik}(t) = \sum_{k=0}^{\infty} P_{kj}(h) P_{ik}(t)$$

from the memoryless property. In long form notation, **for** $k \neq j$,

$$\begin{aligned}
P(X(t+h) = j | X(t) = k) =& P(X(t+h) = j, N[t, t+h] = 1 | X(t) = k) + && (15) \\
& P(X(t+h) = j, N[t, t+h] \geq 2 | X(t) = k) && (16)
\end{aligned}$$

where $N[t, t+h]$ is the number of transitions that are made in $[t, t+h]$. Now,

$$\begin{aligned}
P(X(t+h) = j, N[t, t+h] \geq 2 | X(t) = k) &\stackrel{(*)}{\leq} P\left(\bigcup_{k=2}^{\infty}\left\{\sum_{i=1}^{k} \text{expo}(\gamma^*) \leq h\right\}\right) \\
&\leq P(\text{expo}_1(\gamma^*) + \text{expo}_2(\gamma^*) \leq h) \\
&\leq P(\text{expo}_1(\gamma^*) \leq h) P(\text{expo}_2(\gamma^*) \leq h) \\
&= (1 - e^{-\gamma^* h})^2
\end{aligned}$$

where $(*)$ is because the chain with all rates being γ^* is more likely to have two or more transitions than our original chain. Next, since

$$|Y| \leq 1 \implies 1 + y \leq e^y \leq 1 + y + y^2$$

then

$$h < \frac{1}{\gamma^*} \implies \gamma^* h \geq 1 - e^{-\gamma^* h} \geq 1 + \gamma^* h - (\gamma^*)^2 h^2 \implies (1 - e^{-\gamma^* h})^2 \leq (\gamma^*)^2 h^2$$

and hence

$$0 \leq P(X(t+h) = j, N[t,t+h] \geq 2 | X(t) = k) \leq (\gamma^*)^2 h^2.$$

Next,

$$\begin{aligned} P(X(t+h) = j, N[t,t+h] = 1 | X(t) = k) &= P_{kj} \int_0^h \gamma_k e^{-\gamma_k x} \underbrace{e^{-\gamma_j (h-x)}}_{P(\text{expo}(\gamma_j) > h-x)} \, dx \\ &= P_{kj} \gamma_k e^{-\gamma_j h} \int_0^h e^{(\gamma_j - \gamma_k)x} \, dx. \end{aligned}$$

Now $1 - \gamma^* h \leq e^{-\gamma^* h} \leq 1$ and $|\gamma_j - \gamma_k| \leq |\gamma_* - 1| \leq \gamma_*$ then

$$1 - \gamma^* h \leq e^{-\gamma^* x} \leq e^{(\gamma_j - \gamma_k)} \leq e^{\gamma^* x} \leq 1 + \gamma^* h + (\gamma^*)^2 h^2 \leq 1 + 2(\gamma^*)^2 h$$

for small enough h. This gives us

$$h - \gamma^* h^2 \leq \int_0^h e^{(\gamma_j - \gamma_k)x} \, dx \leq h + 2(\gamma^*)^2 h^2$$

and the bound for the integral expression is

$$P_{kj}\gamma_k (1 - \gamma^* h)(h - \gamma^* h^2) \leq P_{kj} \gamma_k e^{-\gamma_j h} \int_0^h e^{(\gamma_j - \gamma_k)x} \, dx \leq P_{kj} \gamma_k \cdot 1 \cdot (h + 2(\gamma^*)^2 h^2)$$

or equivalently for all $h \in \left(0, \frac{1}{2\gamma^*}\right)$,

$$\begin{aligned} & P_{kj}\gamma_k h (1 - \gamma^* h)^2 \leq P_{kj} \gamma_k e^{-\gamma_j h} \int_0^h e^{(\gamma_j - \gamma_k)x} \, dx \leq P_{kj}\gamma_k h (1 + 2(\gamma^*)^2 h) \\ \Longrightarrow\ & P_{kj}\gamma_k h - 2\gamma^* P_{kj}\gamma_k h^2 \leq (15) \leq P_{kj}\gamma_k h + 2(\gamma^*)^2 P_{kj}\gamma_k h^2 \\ \Longrightarrow\ & P_{kj}\gamma_k h - 2(\gamma^*)^3 h^2 \leq (15) \leq P_{kj}\gamma_k h + 2(\gamma^*)^3 h^2 \\ \Longrightarrow\ & P_{kj}\gamma_k h - 3(\gamma^*)^3 h^2 \leq (15) + (16) \leq P_{kj}\gamma_k h + 3(\gamma^*)^3 h^2. \end{aligned}$$

Now **for** $k = j$, we have (15) equal to $P(\text{expo}(\gamma_k) > h) = e^{-\gamma_k h}$ while the bounds remain the same, namely

$$\begin{aligned} & 1 - \gamma_j h \leq (15) + (16) \leq 1 - \gamma_j h + 2\gamma_j^2 h^2 \\ \Longrightarrow\ & 1 - \gamma_j h - 3(\gamma^*)^3 h^2 \leq (15) + (16) \leq 1 - \gamma_j h + 3(\gamma^*)^3 h^2. \end{aligned}$$

Hence,

$$\begin{aligned} P_{ij}(t+h) \geq & \sum_{k \neq j} \left(P_{kj}\gamma_k h - 3(\gamma^*)^3 h^2 \right) P_{ik}(t) \\ & + \left(1 - \gamma_j h - 3(\gamma^*)^3 h^2 \right) P_{ij}(t) \end{aligned}$$

and

$$\begin{aligned} P_{ij}(t+h) \leq & \sum_{k \neq j} \left(P_{kj}\gamma_k h + 3(\gamma^*)^3 h^2 \right) P_{ik}(t) \\ & + \left(1 - \gamma_j h + 3(\gamma^*)^3 h^2 \right) P_{ij}(t) \end{aligned}$$

Now,

$$\begin{aligned} \frac{P_{ij}(t+h) - P_{ij}(t)}{h} \geq & \sum_{k \neq j} \left(P_{kj}\gamma_k P_{ik}(t) - 3(\gamma^*)^3 h P_{ik}(t) \right) \\ & - \gamma_j P_{ij}(t) - 3(\gamma^*)^3 h P_{ij}(t) \end{aligned}$$

and

$$\frac{P_{ij}(t+h)-P_{ij}(t)}{h} \leq \sum_{k\neq j}\left(P_{kj}\gamma_k P_{ik}(t)+3(\gamma^*)^3hP_{ik}(t)\right) -\gamma_j P_{ij}(t)+3(\gamma^*)^3hP_{ij}(t).$$

Next, since the γ_i are bounded, no explosion occurs and

$$\sum_{k\neq j}|P_{kj}\gamma_k P_{ik}(t)| \leq \gamma^*\sum_{k\neq j}|P_{ik}(t)| < \infty$$
$$\sum_{k\neq j}\left|3(\gamma^*)^3hP_{ik}(t)\right| \leq 3(\gamma^*)^3h\sum_{k\neq j}|P_{ik}(t)| < \infty.$$

Therefore, we conclude that

$$P'_{ij}(t) = \left(\sum_{k\neq j}P_{kj}\gamma_k P_{ik}(t)\right)-\gamma_j P_{ij}(t).$$

Example 2.3. (M/M/∞ Queue) Consider a scenario of infinite servers with arrivals distributed as Poisson(λ) and service as Expo(μ). Then, $\gamma_k=\lambda+k\mu$.

Exercise 2.1. Express $P_{ij}(t)$ as an integral involving $\sum$, $\{P_{ik}, k\neq i\}, \{P_{kj}(s), k\neq j\}$ and some $\gamma's$ and $e^{(\cdot)}$ (*Hint*: Condition on T_1, the time of the first transition).

The form should be:

$$P_{ij}(t)=\delta_{ij}e^{-\gamma_i t}+\int_0^t\sum_{k\neq i}P_{ik}\cdot P_{kj}(t-y)\gamma_i e^{-\gamma_i y}dy$$

Note that if we let $u=t-y$ then

$$\begin{aligned}P_{ij}(t)&=\delta_{ij}e^{-\gamma_i t}+\int_0^t\sum_{k\neq i}P_{ik}\cdot P_{kj}(u)\gamma_i e^{-\gamma_i (t-u)}du\\ &= e^{-\gamma_i t}\left(\delta_{ij}+\int_0^t\sum_{k\neq i}P_{ik}\cdot P_{kj}(u)\gamma_i e^{\gamma_i u}du\right)\end{aligned}$$

and using the product rule,

$$\begin{aligned}P'_{ij}(t)&=-\gamma_i e^{-\gamma_i t}\left(\delta_{ij}+\int_0^t\sum_{k\neq i}P_{ik}\cdot P_{kj}(u)\gamma_i e^{\gamma_i u}du\right)+e^{-\gamma_i t}\sum_{k\neq i}P_{ik}P_{kj}(t)\gamma_i e^{\gamma_i t}\\ &=-\gamma_i P_{ij}(t)+\sum_{k\neq i}P_{ik}P_{kj}(t)\gamma_i.\end{aligned}$$

Define a matrix A called the generator matrix

$$A_{ij}=\begin{cases}\gamma_i P_{ij}, & \text{if } i\neq j\\ -\gamma_i, & \text{otherwise}\end{cases}$$

where $P'(t)=A\cdot P(t)$. What if we condition on the “last transition ”? Then, roughly,

$$P_{ij}(t)=\delta_{ij}e^{-\gamma_i t}+\int_0^t\sum_{k\neq j}\underbrace{P_{ik}(t-y)}_{(1)}\underbrace{\gamma_k P_{kj}dy}_{(2)}\cdot\underbrace{e^{-\gamma_j y}}_{(3)}$$

where we interpret the terms as follows:

(1) At time $(t-y)$ we are in k

(2) We transition (instantaneously) “at” time $(t-y)$

(3) We stay in j for a final y time units

(last jump occurs from $k \to j$ at time $t-y$)

Using the same substitution scheme, we will get $P'(t) = P(t)\cdot A$.

[*Start unfinished work*]

Remark 2.4. Let $\Sigma_{i,k,j}^{(t)} = \{\text{in } i \text{ at } 0,\ \text{in } j \text{ at } t,\ \text{first transition to } k\}$. Then

$$P_{ij}(t) = \sum_{k\neq i} P\left(\Sigma_{i,k,j}^{(t)}\right)$$

Claim 2.1. For $i \neq j$, we have

$$P\left(\Sigma_{i,k,j}^{(t)}\right) = \int_0^t P_{ik}P_{jk}(t-y)\gamma_i e^{-\gamma_i y}dy.$$

Proof. Let T_1 be the time of the first jump. We observe that

$$\begin{aligned}\Sigma_{i,k,j}^{(t)} &= I\begin{pmatrix}T_1\le t, \text{ first jump to } k\\ \text{AND}\\ \text{starting in } k \text{ at } T_1 \text{ you are}\\ \text{in } j \text{ at } t-T_1 \text{ time units later}\end{pmatrix}\\ &= I(T_1\le t)\cdot I(\text{first jump to } k)\cdot I\begin{pmatrix}\text{starting in } k \text{ at } T_1 \text{ you are}\\ \text{in } j \text{ at } t-T_1 \text{ time units later}\end{pmatrix}.\end{aligned}$$

Hence we may write

$$\begin{aligned}P\left(\Sigma_{i,k,j}^{(t)}\right) &= E\left[I(T_1\le t)\cdot I(\text{first jump to } k)\cdot I\begin{pmatrix}\text{starting in } k \text{ at } T_1 \text{ you are}\\ \text{in } j \text{ at } t-T_1 \text{ time units later}\end{pmatrix}\right]\\ &= P_{ik}E\left[I(T_1\le t)\cdot P_{kj}(t-T_1)\right]\end{aligned}$$

[*End unfinished work; left unfinished by the professor on purpose*] □

Notation 1. Let T_1 be the time of our first jump and T_2 be the time between the first and second jump.

Theorem 2.1. *For $i \neq j$ we have*

$$P_{ij}(t) = \int_0^t \sum_{k\neq i} P_{ik}P_{kj}(t-y)\gamma_i e^{-\gamma_i y}dy$$

Lemma 2.1. *We have*

$$P(X(t)=j, X(T_1)=k) = \int_0^t P_{ik}P_{kj}(t-y)\gamma_i e^{-\gamma_i t}dy$$

Proof. First remark that for all $n \geq 2/t$,

$$\begin{aligned}P(X(t)=j, X(T_1)=k) = &\sum_{m=0}^{\lfloor nt\rfloor-1} P\left(X(t)=j, X(T_1)=k, T_1\in\left[\frac{m}{n},\frac{m+1}{n}\right)\right) +\\ &\underbrace{P\left(X(t)=j, X(T_1)=k, T_1\in\left[\frac{\lfloor nt\rfloor}{n}, t\right)\right)}_{=C}.\end{aligned}$$

Let us bound $P\left(X(t)=j, X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right)\right)$. Directly we have

$$\begin{aligned}
&P\left(X(t)=j, X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right)\right) \\
=&\underbrace{P\left(X(t)=j, X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), X\left(\frac{m+1}{n}\right)=k\right)}_{=A}+ \\
&\underbrace{P\left(X(t)=j, X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), X\left(\frac{m+1}{n}\right)\neq k\right)}_{=B}
\end{aligned}$$

and we have

$$\begin{aligned}
A= \quad & P\left(X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), X\left(\frac{m+1}{n}\right)=k\right)\times \\
& P\left(X(t)=j \middle| X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), X\left(\frac{m+1}{n}\right)=k\right) \\
= \quad & P\left(X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), X\left(\frac{m+1}{n}\right)=k\right)\times P_{kj}\left(t-\frac{m+1}{n}\right).
\end{aligned}$$

Now,

$$\begin{aligned}
&P\left(X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), X\left(\frac{m+1}{n}\right)=k\right) \\
\geq &P\left(X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), T_2 \geq \frac{1}{n}\right) \\
= &\left(\int_{\frac{m}{n}}^{\frac{m+1}{n}} \gamma_i e^{-\gamma_i s} ds\right) P_{ik} e^{-\frac{\gamma_k}{n}} \\
\geq &\left(\int_{\frac{m}{n}}^{\frac{m+1}{n}} \gamma_i e^{-\gamma_i s} ds\right) P_{ik}\left(1-\frac{\gamma_k}{n}\right) \\
\geq &\frac{\gamma_i e^{-\gamma_i\left[\frac{m+1}{n}\right]}}{n} P_{ik}\left(1-\frac{\gamma_k}{n}\right) \\
= &\frac{\gamma_i e^{-\gamma_i\left[\frac{m+1}{n}\right]}}{n} P_{ik} - \frac{\gamma_i\gamma_k}{n^2}
\end{aligned}$$

and similarly

$$\begin{aligned}
&P\left(X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), X\left(\frac{m+1}{n}\right)=k\right) \\
\leq &P\left(X(T_1)=k, T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right)\right) \\
&\vdots \\
\leq &\frac{\gamma_i e^{-\gamma_i\left[\frac{m}{n}\right]}}{n} P_{ik} \\
= &\frac{\gamma_i e^{-\gamma_i\left[\frac{m+1}{n}\right]}}{n} P_{ik} e^{\frac{\gamma_i}{n}} \\
= &\frac{\gamma_i e^{-\gamma_i\left[\frac{m+1}{n}\right]}}{n} P_{ik} + \frac{2\gamma_i^2}{n^2}
\end{aligned}$$

Hence,

$$\frac{\gamma_i e^{-\gamma_i\left[\frac{m+1}{n}\right]}}{n} P_{ik} P_{kj}\left(t-\frac{m+1}{n}\right) - \frac{\gamma_i\gamma_k}{n^2} \leq A \leq \frac{\gamma_i e^{-\gamma_i\left[\frac{m+1}{n}\right]}}{n} P_{ik} P_{kj}\left(t-\frac{m+1}{n}\right) + \frac{2\gamma_i^2}{n^2}.$$

Next,

$$\begin{aligned}
&0 \le B \le P\left(T_1 \in \left[\frac{m}{n}, \frac{m+1}{n}\right), T_2 \le \frac{1}{n}\right) \le \frac{\gamma_i \gamma_k}{n^2} \\
&0 \le C \le \frac{\gamma_i}{n}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\frac{1}{n}\sum_{m=0}^{\lfloor nt \rfloor - 1} P_{ik} P_{kj}\left(t - \frac{m+1}{n}\right)\gamma_i e^{-\gamma_i \frac{m+1}{n}} - \frac{\gamma_i \gamma_k t}{n} \\
\le& P(X(t) = j, X(T_1) = k) \\
\le& \frac{1}{n}\sum_{m=0}^{\lfloor nt \rfloor - 1} P_{ik} P_{kj}\left(t - \frac{m+1}{n}\right)\gamma_i e^{-\gamma_i \frac{m+1}{n}} + \frac{2\gamma_i^2 t}{n} + \frac{\gamma_i}{n}
\end{aligned}$$

Equivalently, for all $n \ge \frac{2}{t} + \gamma_i$ we have

$$\begin{aligned}
&\frac{1}{n}\sum_{m=0}^{\lfloor nt \rfloor - 1} P_{ik} P_{kj}\left(t - \frac{m+1}{n}\right)\gamma_i e^{-\gamma_i \frac{m+1}{n}} - \frac{3t\gamma_i(\gamma_i + \gamma_k) + \gamma_i}{n} \\
\le& P(X(t) = j, X(T_1) = k) \\
\le& \frac{1}{n}\sum_{m=0}^{\lfloor nt \rfloor - 1} P_{ik} P_{kj}\left(t - \frac{m+1}{n}\right)\gamma_i e^{-\gamma_i \frac{m+1}{n}} + \frac{3t\gamma_i(\gamma_i + \gamma_k) + \gamma_i}{n}.
\end{aligned}$$

We will now show that $P_{kj}(z)$ is a Lipschitz continuous function. Now,

(1) $P_{kj}(t_1 + h) \ge P_{kj}(t_1) \pm \gamma_k h$

(2) $A_h = \{T_1 \le h\}, A_h^c = \{T_1 > h\}$

(3) We have

$$\begin{aligned}
P_{kj}(t_1 + h) &= P(X(t_1 + h) = j | A_h) P(A_h) + P(X(t_1 + h) = j | A_h^c) P(A_h^c) \\
&\ge P(X(t_1 + h) = j | A_h^c) P(A_h^c) \\
&= P_{kj}(t_1) e^{-\gamma_k h} \\
&\ge P_{kj}(t_1)(1 - \gamma_k h) \\
&\ge P_{kj}(t_1) - \gamma_k h
\end{aligned}$$

and

$$\begin{aligned}
P_{kj}(t_1 + h) &= P(X(t_1 + h) = j | A_h) P(A_h) + P(X(t_1 + h) = j | A_h^c) P(A_h^c) \\
&\le 1 - e^{-\gamma_k h} + P_{kj}(t_1) e^{-\gamma_k h} \\
&\le P_{kj}(t_1) + \gamma_k h
\end{aligned}$$

where (3) $\implies$ (1). Hence,

$$|P_{kj}(t_1 + h) - P_{kj}(t_1)| \le \gamma_k h$$

and $P_{kj}(\cdot)$ is Lipschitz continuous for all k and therefore, we can construct the Riemann integral representation

$$P_{ij}(t) = e^{-\gamma_i t}\left(\delta_{ij} + \sum_{k \ne i} \int_0^t P_{ik} P_{kj}(u) \gamma_i e^{\gamma_i u}\, du\right).$$

Note that $\exists K_\epsilon < \infty$ such that $\sum_{k \geq K_\epsilon} P_{ik} \leq \frac{\epsilon}{4}$ and so $\forall \epsilon > 0$, $\exists K_\epsilon < \infty$

$$\begin{aligned}\left|\sum_{k\neq i} P_{ik}P_{kj}(x) - \sum_{k\neq i} P_{ik}P_{kj}(u)\right| &= \left|\sum_{k=1}^{\infty} P_{ik}\left[P_{kj}(x) - P_{kj}(u)\right]\right| \\ &\leq \left|\sum_{k=1}^{K_\epsilon} P_{ik}\left[P_{kj}(x) - P_{kj}(u)\right]\right| + \frac{\epsilon}{2}\end{aligned}$$

for all $x, u > 0$. Now pick $\delta > 0$ such that $|x-u| < \delta/(\gamma_+ K_\epsilon 2)$ where $\gamma_+ = \max_{i=1,\ldots,K_\epsilon}$. Then,

$$\left|\sum_{k=1}^{K_\epsilon} P_{ik}\left[P_{kj}(x) - P_{kj}(u)\right]\right| \leq K_\epsilon \gamma_+ |x-u| \leq \frac{\epsilon}{2}$$

and hence

$$\left|\sum_{k\neq i} P_{ik}P_{kj}(x) - \sum_{k\neq i} P_{ik}P_{kj}(u)\right| \leq \epsilon.$$

Alternatively, since $\sum_{k=1}^{K_\epsilon} P_{ik}P_{kj}(\cdot)$ is continuous, we can always choose a $\delta > 0$ for $|x-u| > \delta$ such that

$$\left|\sum_{k=1}^{K_\epsilon} P_{ik}\left[P_{kj}(x) - P_{kj}(u)\right]\right| \leq \frac{\epsilon}{2}.$$

Previously, we explicitly chose the $\delta > 0$. Note that we have proven that **the infinite convex sum of bounded continuous functions is also continuous**. □

Remark 2.5. We can then conclude that if

$$A_{ij} = \begin{cases}\gamma_i P_{ij} & i \neq j \\ -\gamma_i & i = j\end{cases}$$

then

$$\begin{aligned}P'(t) &= A \cdot P(t) \\ P'(t) &= P(t) \cdot A\end{aligned}$$

where the latter needs some conditions imposed (the more thorough derivation) while the former is for general set ups. The above two, together, are called the **Chapman-Kolmogorov** (C-K) **differential equations**. The solution of the first equation can be verified to be

$$P(t) = e^{At} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n$$

using $P(0) = I$ as an initial condition. In fact,

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

Note that

$$\frac{dP}{dt} = \sum_{n=0}^{\infty} \frac{d}{dt}\left(\frac{t^n}{n!}A^n\right) = A\sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!}A^{n-1} = A \cdot P(t).$$

As A commutes with itself, $\frac{dP}{dt} = P(t) \cdot A$ as well.

2.3 Uniformization

Suppose we have a CTMC with rates $\{\gamma_i\}$, DTMC transition probabilities $\{P_{ij}\}$, no self-loops, and $\sup_i \gamma_i < \nu$. If we want to create a different CTMC <u>with</u> self loops so every state has rate ν and the the chain has the "same dynamics" as the original chain, then set

$$\begin{aligned}
&\tilde{\gamma}_i = \nu \\
&\begin{cases} \nu P_{ij} = \gamma_i \tilde{P}_{ij} & i \neq j \\ \nu(1-\tilde{P}_{ii}) = \gamma_i & i = j \end{cases} \implies \tilde{P}_{ij} = \begin{cases} \frac{\gamma_i P_{ij}}{\nu} & i \neq j \\ 1 - \frac{\gamma_i}{\nu} & i = j \end{cases}.
\end{aligned}$$

where $\tilde{P}$ is the probability transition matrix of the new DTMC and $\tilde{P}^0 = I$. Note that

$$P(t) = \sum_{k=0}^{\infty} \tilde{P}^k \left(\frac{e^{-\nu t}(vt)^k}{k!} \right)$$

as the number hits from $[0,t]$ is a Poisson process as the interarrival times are Expo(ν). Since $\tilde{P} = I + A/\nu$, then

$$\begin{aligned}
P(t) &= \sum_{k=0}^{\infty} \left(I + \frac{A}{\nu} \right)^k \left(\frac{e^{-\nu t}(vt)^k}{k!} \right) \\
&= e^{-\nu t} \sum_{k=0}^{\infty} \frac{((\nu t + A)t)^k}{k!} \\
&= e^{-\nu t} e^{(\nu I + A)t} \\
&= e^{-\nu t} e^{\nu I t + At} \\
&= e^{-\nu t} e^{\nu I t} e^{At} \\
&= e^{-\nu t} e^{\nu t} I e^{At} \\
&= e^{At}
\end{aligned}$$

Now suppose that you have a CTMC with self-loops. A new CTMC with no self-loops and the same dynamics will have

$$\begin{aligned}
&\tilde{\gamma}_i = \gamma_i(1 - P_{ii}) \\
&\tilde{\gamma}_i \tilde{P}_{ij} = \gamma_i P_{ij} \implies \tilde{P}_{ij} = \frac{\gamma_i P_{ij}}{\gamma_i(1-P_{ii})} = \frac{P_{ij}}{1-P_{ii}}
\end{aligned}$$

and hence

$$P'_{ij}(t) = \left(\sum_{k \neq j} P_{ik}(t) \gamma_k P_{kj}(t) \right) - P_{ij}(t) \gamma_j (1 - P_{jj}).$$

Proposition 2.1. *(Semi-group property) We have*

$$\begin{aligned}
P(s+t) &= P(s) \cdot P(t) = P(t) \cdot P(s) \\
P_{ij}(s+t) &= \sum_{k=1}^{\infty} P_{ik}(t) \cdot P_{kj}(s).
\end{aligned}$$

3 Steady-State Distributions and Pseudo-Markov Chains

Definition 3.1. A **pseudo-Markov-chain (PMC)** $Z(t)$ is a stochastic process with countable state space and

(1) There exists an underlying DTMC

(2) There is a distribution function F_i pegged to each state i

(3) Whenever you get to state i, you stay for a time i with distribution function F_i

Notation 2. Define

$$\begin{aligned}
T_{ij} &= \text{time having just entered state } i, \text{ to first enter state } j \\
T_{ii} &= \text{time having just entered state } i, \text{ to first enter state } i \\
T_i &= \text{time for one stay in state } i \\
\mu_{ij} &= E[T_{ij}], \mu_i = E[T_i] \in (0,\infty) \\
O_j^i &= E\left[\# \text{ of visits to } j \text{ on an } i \to i \text{ sojourn}\right] = E\left[\sum_{n=0}^{\tau^i-1} I(X_n = j)|X_0 = i\right] \\
\tau^i &= \inf\{n \geq 1 : X_n = i | X_0 = i\} \\
\hat{O}_j^i &= E\left[\text{time in } j \text{ on an } i \to i \text{ sojourn}\right] = E\left[\sum_{n=0}^{\tau^i-1} W_n I(X_n = j)|X_0 = i\right] = \mu_j O_j^i
\end{aligned}$$

Note that $O_i^i = 1$ and let $O^i = [O_1^i, O_2^i, ...]$, the occupation vector.

Remark 3.1. If we have a recurrent, irreducible DTMC, $O_j^i < \infty$. This is because if we hit j from i, there is a probability p_j^i that we hit i (while in state j) before we hit j again. Hence the number of hits is Geo(p_j^i) if we hit j before i plus Bernoulli(p_i^j).

Remark 3.2. $\mu_{ii} = \sum_j \hat{O}_j^i$, which is true from Fubini's Theorem, even in the case where both sides are ∞.

Theorem 3.1. *(Thm 10) For any non-lattice PMC, $Z(t)$, with irreducible recurrent underlying DTMC. For any initial state j, target state i, and reference state s,*

$$\lim_{t\to\infty} P(Z(t) = i | Z(0) = j) = \frac{\hat{O}_i^s}{\mu_{ss}}$$

and (Thm 11)

$$P\left(\lim_{t\to\infty} \frac{f(i,j,t)}{t} = \frac{\hat{O}_i^s}{\mu_{ss}}\right) = 1$$

where $f(i,j,t)$ is the time in i on $(0,t)$ if you start in j.

Remark 3.3. Note that

$$\frac{\hat{O}_i^s}{\mu_{ss}} \neq E\left[\text{fraction of time in state } i \text{ on an } s \to s \text{ sojourn}\right].$$

Think of the process which switches from s to i with probabilities 1 and spends time unit 1 in state s and time units X (pick a random variable) in state i.

Proposition 3.1. *(SLLN for Alternating Renewal Processes) Consider a delayed alternating renewal process $(T, \{(Z_i, Y_i)\}_{i=1}^{\infty})$ where $\{(Z_i, Y_i)\}_{i=1}^{\infty}$ are mutually independent and identically distributed, Z_i is ON, Y_i is OFF, and T is a.s. finite (waiting time T occurs first, then Z_1, then Y_1, then Z_2, then Y_2, etc.). Then if $Z_1 + Y_1$ is non-lattice, and $E[Z_1] < \infty$ we have*

$$\lim_{t\to\infty} P(ON(t)) = \frac{E(Z_1)}{E(Z_1) + E(Y_1)}.$$

Proposition 3.2. *(SLLN For Reward Renewal Processes) Consider a delayed reward renewal process $(T, \{(X_i, R_i)\}_{i=1}^{\infty})$ where $\{(X_i, R_i)\}_{i=1}^{\infty}$ are mutually independent and identically distributed, R_i is the reward for period X_i, and T is a.s. finite. Also we suppose that during a given interval, reward is earned over time in an "arbitrary" (as long as $\uparrow$ and continuous) way. Then we have*

$$P\left(\lim_{t\to\infty} \frac{R(t)}{t} = \frac{E[R]}{E[X]}\right) = 1.$$

Proof. (Thm 11) Set up a reward renewal process with X_i as an $s \to s$ sojourn with reward as the time spent in state i.

(Thm 10) Set up an alternating renewal process with T as the time from $j \to i$, Z_1 is the time spent in i until we hit a state that is not i, Y_1 is the time spent to get back to i. We have $Z_1 + Y_1 = T_{ii}$ and the result follows for $s = i$. By Thm 11, the result of Thm 10 holds for any s. □

Claim. Recall the C-K equation

$$P'(t) = P(t) \cdot A$$

where

$$A_{ij} = \begin{cases} \frac{1}{\mu_i} P_{ij}, & i \neq j \\ -\frac{1}{\mu_i}, & i = j \end{cases}.$$

(1) From Thm 10 we get

$$\bar{\eta} = P(\infty) = \lim_{t\to\infty} P(t) = \lim_{t\to\infty} \begin{pmatrix} P_{11}(t) & P_{12}(t) & \cdots & P_{1n}(t) \\ P_{21}(t) & \ddots & & \vdots \\ \vdots & & & \\ P_{m1}(t) & \cdots & & P_{mn}(t) \end{pmatrix} = \lim_{t\to\infty} \begin{pmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \\ \eta_1 & \eta_2 & \cdots & \eta_n \\ \eta_1 & \eta_2 & \cdots & \eta_n \\ \eta_1 & \eta_2 & \cdots & \eta_n \\ \eta_1 & \eta_2 & \cdots & \eta_n \end{pmatrix}$$

and

$$\lim_{t\to\infty} P'(t) = P(\infty) \cdot A.$$

Now suppose that $\lim_{t\to\infty} P'_{ij}(t) = \epsilon > 0$ which implies that $\exists T_\epsilon < \infty$ such that for all $t \geq T_\epsilon$, $P'_{ij}(t) \geq \frac{\epsilon}{2}$. Thus $P_{ij}(t + T_\epsilon) \geq t \cdot \frac{\epsilon}{2}$. So there exists $T'_\epsilon < \infty$ such that $P(T'_\epsilon) > 1$ which is impossible. So

$$\lim_{t\to\infty} P'_{ij}(t) = 0 \implies \bar{\eta} \cdot A = 0$$

and they are the unique solution to

$$\begin{cases} \eta \cdot A = 0 \\ \sum_i \eta_i = 1 \\ \eta_i \geq 0 \end{cases}.$$

Fact 3.1. *(F1) For any recurrent, irreducible DTMC, $O^i = O^i \cdot P$ for all i.*

Fact 3.2. *(F2) For any recurrent, irreducible DTMC, the system of equations $V = V \cdot P$ always has an uncountably infinite number of non-negtaive, non-zero solutions. For any fixed $s \in S$, one may recover all non-negative, non-zero solutions as $\{c \cdot O^s, c > 0\}$.*

Claim 3.1. (Claim 1) For a PMC with irreducible, recurrent DTMC, the set of non-negative, non-zero solutions to $V \cdot A = 0$ is $\{C \cdot \hat{O}^s, c > 0\}$ for any state s.

Proof. If $V \cdot A = 0$ then

$$\forall j, \frac{1}{\mu_j} \cdot v_j = \sum_i \frac{1}{\mu_i} v_i P_{ij}$$

and given a vector v, let z^v be the unique vector such that $z_i^v = \frac{1}{\mu_i} z_i$ and hence

$$V \cdot A = 0 \iff z^v = z^v \cdot P.$$

So,

$$\begin{aligned} & v \text{ is a non-negative, non-zero solution to } V \cdot A{=}0 \\ \iff & z^v \text{ is a non-negative, non-zero solution to } z^v = z^v \cdot P. \end{aligned}$$

By (F2),

$$\begin{aligned} & z^v \text{ is a non-negative, non-zero solution to } z^v = z^v \cdot P \\ \iff & \exists c > 0 \text{ s.t. } z^v = c \cdot O^s \\ \iff & \frac{1}{\mu_i} v_i = c \cdot O_s^i \\ \iff & v_i = c\mu_i O_s^i \\ \iff & v = c \cdot \hat{O}^s. \end{aligned}$$

□

Claim 3.2. (Claim 2) V is a (non-negative) solution to $V \cdot A = 0$ if and only if Z^v (defined as $Z_i^v = \frac{1}{\mu_i} v_i$) is a (non-negative, non-zero) solution to $Z^v = Z^v \cdot P$.

Theorem 3.2. *Suppose that $X(t)$ is a PMC with state space (s.s.) S and irreducible recurrent underlying DTMC satisfying our non-lattice assumptions. Suppose A is the underlying generating matrix. Then the following dichotomy holds:*

* ***[Ergodic Case]*** *If the system of equations $V \cdot A = 0$ has a non-negative solution V^1 such that $\sum_i V_i^1 = 1$ the for all $i \in S$, $\mu_{ii} < \infty$ and*

$$\lim_{t\to\infty} P(X(t) = i | X(0) = j) = V_i^1 \in (0,1), \forall i \in S$$

* ***[Non-Ergodic Case]*** *Else $\mu_{ii} = \infty$ for all $i \in S$ and*

$$\lim_{t\to\infty} P(X(t) = i | X(0) = j) = 0, \forall i \in S$$

Proof. [Ergodic Case] By Claim 2, for any $s \in S$, $\exists c_s$ such that $V^1 = c_s \cdot \hat{O}^s$. By the fact that $\mu_{ii} = \sum_j \hat{O}_j^i$, we have

$$\mu_{ss} = \frac{1}{c_s}\sum_i V_i^1 = \frac{1}{c_s} \implies c_s = \frac{1}{\mu_{ss}} \implies V^1 = \frac{\hat{O}^s}{\mu_{ss}} \implies V_i^1 = \frac{\hat{O}_i^s}{\mu_{ss}}.$$

[Non-Ergodic Case] Alternatively, suppose $\nexists V^1$ such that $V^1 \cdot A = 0$, V^1 non-negative, $\sum_i V_i^1 = 1$. Suppose for contradiction that $\exists i$ such that $\mu_{ii} < \infty$. Let $V^1 = \frac{\hat{O}^i}{\mu_{ii}}$ and note that by claim 1, V^1 is in our solution set which contradicts our initial assumption.

Thus, $\eta_i = \frac{\hat{O}_i^s}{\mu_{ss}}$ for any state s. $\square$

Corollary 3.1. *If in addition P is positive recurrent, then the PMC is ergodic $\iff$ $\sum_i \pi_i \mu_i < \infty$ in which case*

$$\lim_{t\to\infty} P\left(X(t) = i | X(0) = j\right) = \frac{\pi_i \mu_i}{\sum_j \pi_j \mu_j}.$$

Example 3.1. Suppose that $P_{01} = 1$, $P_{i,i+1} = P_{i,i+1} = \frac{1}{2}$ for all i and $\gamma_i = 2^i$ with $\mu_i = 2^{-i}$. Note that in the DTMC, if $E[T_{10}] < \infty$, then

$$E[T_{10}] = \frac{1}{2} + \frac{1}{2}\left(1 + E[T_{20}]\right)$$

and since $T_{20} \sim T_{21} + T_{10}$ then

$$E[T_{10}] = 1 + E[T_{10}]$$

so it must be that $E[T_{10}] = \infty$. If $P_{i,i-1} = \frac{3}{4}, P_{i,i+1} = \frac{1}{4}$, however, we get $E[T_{10}] = 2$. Back to the $P_{i,i-1} = P_{i,i+1} = \frac{1}{2}$ case, note that

$$A = \begin{bmatrix} -1 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 2 & -2^2 & 2 & & \\ & & 2^2 & -2^3 & 2^2 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix}$$

and so $\eta \cdot A = 0$ implies

$$\begin{cases} \sum_i \eta_i & = 1 \\ -\eta_0 + \eta_1 & = 0 \\ \eta_0 - 2\eta_1 + 2\eta_2 & = 0 \\ \eta_1 - 2^2\eta_2 + 2^2\eta_3 & = 0 \\ \vdots & \\ 2^{k-2}\eta_{k-1} - 2^k\eta_k + 2^k\eta_{k+1} & = 0 \\ \vdots & \end{cases} \implies \begin{cases} \sum_i \eta_i & = 1 \\ \eta_1 & = \eta_0 \\ \eta_2 & = \frac{1}{2}\eta_1 = \frac{1}{2}\eta_0 \\ \eta_3 & = \frac{2-1}{2^2}\eta_0 = 2^{-2}\eta_0 \\ \vdots & \\ \eta_k & = 2^{-(k-1)}\eta_0 \\ \vdots & \end{cases} \implies \eta_0 = \frac{1}{3}, \eta_k = \frac{1}{3 \cdot 2^{k-1}}.$$

Now in the CTMC case, $\frac{1}{2}T_{10} \sim T_{21}$ so

$$E[T_{10}] = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\left[\frac{1}{2} + E[T_{20}]\right] = \frac{1}{2} + \frac{3}{4}E[T_{10}] \implies E[T_{10}] = 2$$

and hence

$$\mu_0 = E[T_{00}] = E[T_{01}] + E[T_{10}] = 3.$$

Thus, $\eta_0 = \frac{\hat{O}_0^0}{\mu_{00}} = \frac{1}{3}$.

3.1 Stationary Measure (CTMC)

Theorem 3.3. *For any ergodic CTMC, for all $j \in S$ and $t \geq 0$, we have*

$$P_j(\infty) = \sum_{i \in S} P_i(\infty) P_{ij}(t).$$

Proof. For any fixed $j, k, t \geq 0$ we have

$$\sum_{i \in S} P_i(\infty) P_{ij}(t) = \sum_{i \in S} P_{ij}(t) \lim_{s \to \infty} P_{ki}(s), \forall k$$

and

$$\sum_{i \in S} P_{ij}(t) \lim_{s \to \infty} P_{ki}(s) = \lim_{s \to \infty} \sum_{i \in S} P_{ij}(t) P_{ki}(s) = \lim_{s \to \infty} P_{kj}(s+t) = P_j(\infty).$$

□

3.2 Birth-Death Process

This is:

* A CTMC on $\mathbb{Z}^+$ (or any countable subset of $\mathbb{Z}^+$) such that $|i-j| > 1 \implies P_{ij} = 0$.

* $\lambda_i = \gamma_i P_{i,j+1}$ (birth rate)

* $\mu_i = \gamma_i P_{i,j-1}$ (death rate)

* $S = \{i_m, i_{m+1}, ..., i_M\}$ where i_M may be ∞

* $\mu_{i_m} = 0$ and if $i_M < \infty$ then $\lambda_{i_M} = 0$

* The generator matrix is

$$A = \begin{bmatrix} -(\lambda_1+\mu_1) & \lambda_1 & 0 & & & \\ \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & & \\ 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & 0 & \\ & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & & & \end{bmatrix}$$

and hence $\eta_{i+1} = \frac{\lambda_i}{\mu_{i+1}} \eta_i$ for all $i \in S$ (Proof: easily by induction).

* If

$$\sum_{k=i_m+1}^{i_M} \frac{\prod_{k=i_m}^{i-1} \lambda_k}{\prod_{k=i_{m+1}}^{i} \mu_k} < \infty$$

the CTMC is ergodic and for all $i \in S$,

$$P_i(\infty) = \frac{\prod_{k=i_m}^{i-1} \lambda_k}{\prod_{k=i_{m+1}}^{i} \mu_k} \times \left(\sum_{k=i_m+1}^{i_M} \frac{\prod_{k=i_m}^{i-1} \lambda_k}{\prod_{k=i_{m+1}}^{i} \mu_k} \right)^{-1}$$

* In a M/M/∞ queue, we have birth rate $\lambda_i = \lambda$ and $\mu_i = i\mu$, and we have the limiting distribution Poisson $\left(\frac{\lambda}{\mu}\right)$.

* In a M/M/1 queue, we have birth rate $\lambda_i = \lambda$ and $\mu_i = \mu$, and we have the limiting distribution Geo $\left(1 - \frac{\lambda}{\mu}\right)$ if $\lambda < \mu$.

4 Poisson Processes

* $N(t)$ is the number of events up until time t

A **Poisson process** (P.P.) with rate λ has the following equivalent formulations (I, II, III):

I. This is a CTMC on $\mathbb{Z}^+$ such that $\gamma_i = \lambda$ for all $i \geq 0$ and $P_{i,i+1} = 1$ for all $i \geq 0$.

II. The following are true:

(1) $N(t)$ is a counting process, i.e. $N(t) \geq 0$ for all t, $N(t) \in \mathbb{Z}^+$ for all t, $N(t) \uparrow$, $N(0) = 0$

(2) $\forall t_1 < t_2 < ... < t_n$, $[N(t_{i+1}) - N(t_i), i = 1, ..., n-1]$ are mutually independent [**Independent Increments**]

(3) $\forall t_1 < t_2$, we have $N(t_2) - N(t_1) \sim \text{Pois}(\lambda(t_2 - t_1))$.

III. The following are true:

(1) $N(t)$ is a counting process

(2) The increments are stationary

(3) We have independent increments

(4) For all $t \geq 0$ we have $\lim_{h \downarrow 0} h^{-1} P\left(N(t+h) - N(t) \geq 2\right) = 0$

(5) For all $t \geq 0$ we have $\lim_{h \downarrow 0} h^{-1} P\left(N(t+h) - N(t) = 1\right) = \lambda$

Proposition 4.1. *(I. Merged P.P. are P.P.) If $P_1(t)$ and $P_2(t)$ are independent P.P. with rates λ_1, λ_2, then*

$$P_3(t) = P_1(t) + P_2(t)$$

where P_3 is a P.P. with rate $(\lambda_1 + \lambda_2)$.

Proposition 4.2. *(II. Uniform order statistics) Suppose $P(t)$ is a P.P. with rate λ and suppose you condition on there having been exactly k events on $[0, T]$ for some k, T. Then the times of those events, view as an UNORDERED SET, has the same distribution as k independent $U[0, T]$ random variables.*

Example 4.1. Suppose you are engaged in battle with the Zonarkians in the arena of doom (a.o.d.) which is a length 1 unit interval $[0, 1]$.

* At time 0, the Zonarkian General Thantos sets an Expo(γ) clock.

* At time 0, a Zonarkian spaceship zooms by and distributes Zonarkian troopers in the a.o.d. as a rate λ P.P.

* A time 0, each Zonarkian trooper (if any) stars its own Expo clock where the rate of a trooper's clock is equal to its position.

* Let T_Z be the time the first Zonarkian clock goes off, whereby if a clock finishes the Earth is destroyed

* At time 0, seeing the Zonarkians are attacking, you call Jeff Goldblum who starts cracking Zonarkian code. It takes him exactly 10 units of time to crack the code and disarm the Zonarkian bombs

What is $P(T_Z > 10)$? Let $E_i \sim \text{Expo}\left(U[0, 1]\right)$. This turns out to be

$$\begin{aligned}
P\left(\min_{1 \leq i \leq N} (E_i) > 10\right) &= \sum_{k=0}^{\infty} \left(\min_{1 \leq i \leq N} (E_i) > 10 | N = k\right) \frac{e^{-\lambda}\lambda^k}{k!} \\
&= \sum_{k=0}^{\infty} \prod_{i=0}^{k} P\left(E_i > 10\right) \times \frac{e^{-\lambda}\lambda^k}{k!} \\
&= e^{-10\gamma - \lambda} \sum_{k=0}^{\infty} \prod_{i=1}^{k} P\left(E_i > 10\right) \times \frac{\lambda^k}{k!} \\
&= e^{-10\gamma - \lambda} \sum_{k=0}^{\infty} \prod_{i=1}^{k} \frac{\lambda^k}{k!} P^k\left(E_1 > 10\right) \\
&= e^{-10\gamma - \lambda + \lambda P(E_1 > 10)}
\end{aligned}$$

Now

$$P(E_1 > 10) = \int_0^1 e^{-10x}\, dx = \frac{1 - e^{-10}}{10}.$$

4.1 Non-Homogeneous Poisson Processes

A **non-homogeneous Poisson process** is a process where:

(1) $N(t)$ is a a counting process

(2) Independent increments

(3) $\forall t \geq 0$, $\lim_{h \downarrow 0} h^{-1} P(N(t+h) - N(t) \geq 2) = 0$

(4) $\forall t \geq 0$, $\lim_{h \downarrow 0} h^{-1} P(N(t+h) - N(t) \geq 1) = \lambda(t)$

We have

$$\# \text{ of events in } (t_1, t_2) \sim \text{Poisson}\left(\int_{t_1}^{t_2} \lambda(s)\, ds \right).$$

Theorem 4.1. *(Le Cam's Theorem) Roughly, the occurrence of a large number of rare events is roughly Poisson.*

Proposition 4.3. *Suppose $P(t)$ is a non-homogeneous Poisson Process. Suppose each event in $P(t)$ is classified as I or II where the probability an event at time t is type I is $p_1(t)$ (and these classifications are independent across events). Then $P_1(t)$, the # of type I events on $[0, t]$, and $P_2(t)$, the # of type II events on $[0, t]$, are independent non-homogeneous P.P. with rates $\lambda(t)p_1(t)$ and $\lambda(t)[1 - p_1(t)]$ respectively.*

Example 4.2. Suppose we have a M/M/∞ queue. We want to know the distribution of people in the server at time T. It turns out that

$$\begin{aligned} p_1(t) &= P(\text{arrival at } t \text{ is still in the server at } T) \\ &= e^{-\mu(T-t)} \end{aligned}$$

and $\lambda(t) = \lambda$. So our distribution is

$$\begin{aligned} \text{Poisson}\left(\int_0^T \lambda e^{-\mu(T-t)} dt \right) &= \text{Poisson}\left(\lambda \int_0^T e^{-\mu x} dx \right) \\ &= \text{Poisson}\left(\lambda \left[\frac{1 - e^{-\mu T}}{\mu} \right] \right). \end{aligned}$$

Example 4.3. Suppose we have a M/G/∞ queue with general service time c.d.f. F and mean $E[S]$. If we take $T \to \infty$ then the steady state number of jobs in the system is

$$\text{Poisson}\left(\lambda \int_0^\infty (1 - F(x))\, dx \right) = \text{Poisson}\left(\lambda E[S] \right)$$

Remark 4.1. The number of people who have left the server is INDEPENDENT of the number of people who are still in service (crazy!).