STOCHASTIC MODELS FOR WEB 2.0

VIJAY G. SUBRAMANIAN

Email to v-subramanian@northwestern.edu or vijaygautamsubramanian@gmail.com, pointing out errors or hard to understand passages or providing comments, is welcome.

Disclaimer and other information

These are meant to be a rough set of notes and so could have some informal language. They give a quick reference to the background and material for the course but are not meant to be comprehensive. We will provide proofs for certain results but for most details we will refer the reader to the appropriate book or article.

Notation: $\mathbb{N} = \{1, 2, \dots\}$ is the set of natural numbers, $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$ the set of whole numbers, $\mathbb{R} = (-\infty, \infty)$ is the real line, $\mathbb{Z}_+ = \{0\} \cup \mathbb{N}$ is the set of non-negative whole numbers, and $\mathbb{R}_+ = [0, \infty)$ is the set of non-negative real numbers. Things of importance will be marked with $*$ with increasing number corresponding to increasing importance.

Sources: Further discussion and details can be found in the following books:

- D. Williams, "Probability with martingales," Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991.
- J. R. Norris, "Markov chains," Cambridge Series in Statistical and Probabilistic Mathematics, 2, Cambridge University Press, Cambridge, 2006.
- P. Brémaud, "Markov chains: Gibbs fields, Monte Carlo simulation, and queues," Texts in Applied Mathematics, 31, Springer-Verlag, New York, 1999.
- S. N. Ethier and T. G. Kurtz, "Markov processes: Characterization and convergence," Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986.
- P. Billingsley, "Probability and measure," Third edition, Wiley Series in Probability and Mathematical Statistics, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1995.
- O. Kallenberg, "Foundations of modern probability," Second edition, Probability and its Applications (New York), Springer-Verlag, New York, 2002.

1. Probability Fundamentals

Modern probability as set up by Kolmogorov is based on measure theory. Without going into details of measure theory, our objective is to collect together certain facts and to elucidate certain key principles and the intuition.

1.1. σ-algebra and probability measures.

We let the sample-space be Ω with the sample point being $\omega \in \Omega$. The allowed events are given by a collection of subsets $\mathcal{F}$ of Ω; if $\mathcal{F}$ is a

σ-algebra , then $(\Omega, \mathcal{F})$ is deemed a measurable space. *Of course, this begs the question of what a σ-algebra is?*

We start by describing an algebra of sets. A collection of subsets $\tilde{\mathcal{F}}$ of Ω is called an algebra if the following hold:

(1) $\Omega \in \tilde{\mathcal{F}}$;
(2) If $F \in \tilde{\mathcal{F}}$, then $F^C \in \tilde{\mathcal{F}}$ where $F^C := \Omega \setminus F = \{w \in \Omega : w \notin F\}$ is the complement of F (in Ω); and
(3) If $F, G \in \tilde{\mathcal{F}}$, then $F \cup G \in \tilde{\mathcal{F}}$. Inductively, this generalizes to closure (in terms of set containment) over finite unions.

Examples:

(1) $\tilde{\mathcal{F}} = \{\emptyset, \Omega\}$, $\tilde{\mathcal{F}} = 2^\Omega$ (the power set of Ω, i.e., the set of all subsets of Ω).
(2) $\Omega = [0,1]$ and $\tilde{\mathcal{F}} = \{$All finite unions of sets of the form $[a,b], [a,b), (b,a], (a,b)$ where $a, b \in [0,1]\}$.

A collection of subsets $\mathcal{F}$ of Ω is called a σ-algebra if the following hold:

(1) $\mathcal{F}$ is an algebra of sets.
(2) $\mathcal{F}$ is closed under countable unions, i.e., if $\{F_i\}_{i\in\mathbb{N}}$ is a countable collection of sets from $\mathcal{F}$, then $\cup_{i\in\mathbb{N}} F_i \in \tilde{\mathcal{F}}$.

Examples:

(1) $\tilde{\mathcal{F}} = \{\emptyset, \Omega\}$, $\tilde{\mathcal{F}} = 2^\Omega$ (the power set of Ω, i.e., the set of all subsets of Ω).
(2) $(***)$ $\Omega = \mathbb{R}$ and $\mathcal{F}$ the Borel σ-algebra (smallest σ-algebra containing all the open and closed sets of $\mathbb{R}$). We will discuss this in more detail soon.
(3) $(\star)$ $\Omega = \mathbb{R}$ and $\mathcal{F} := \{F \subseteq \Omega : \text{either } F \text{ or } F^C \text{ is countabe}\}$.

For a collection of sets $\tilde{\mathcal{F}}$ (not necessarily an algebra), $\sigma(\tilde{\mathcal{F}})$ denotes the smallest σ-algebra that contains $\tilde{\mathcal{F}}$. If $\tilde{\mathcal{F}} = \{\text{open subsets of } \mathbb{R}\}$, then $\sigma(\tilde{\mathcal{F}})$ is the Borel σ-algebra of $\mathbb{R}$, denoted by $\mathcal{B}(\mathbb{R})$. Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces. For a given function $f : \Omega_1 \mapsto \Omega_2$ where $f^{-1}(B) := \{\omega \in \Omega_1 : f(\omega) \in B\} \in \mathcal{F}_1$ for all $B \in \mathcal{F}_2$, i.e., the inverse images, we can now define $\sigma(f)$ which is the smallest σ-algebra that contains all the inverse images; note that $\sigma(f) = \sigma(\{f^{-1}(B) : B \in \mathcal{F}_2\})$.

Cartesian products: Let X and Y be two sets, then the Cartesian product is defined as $X \times Y := \{(x,y) : x \in X, y \in Y\}$. Similarly we define $\prod_{i=1}^n X_i := \{(x_1, \ldots, x_n) : x_i \in X_i \ \forall i = 1, \ldots, n\}$ and $\prod_{i\in\mathbb{N}} X_i := \{(x_1, x_2, \ldots) : x_i \in X_i \ \forall i \in \mathbb{N}\}$. Eg. $\{0,1\}^2$, $\{0,1\}^{\mathbb{N}}$, $\mathbb{N}^2$, $\mathbb{N}^{\mathbb{N}}$, $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^{\mathbb{N}}$ and $\mathbb{R}^{\mathbb{R}_+}$. Having defined Cartesian products we define a product σ-algebra using projections. Given a Cartesian product $\prod_{i\in\mathbb{N}} \Omega_i$ where $\{\mathcal{F}_i\}_{i\in\mathbb{N}}$ is the accompanying collection of σ-algebra , we denote the product σ-algebra $\prod_{i\in\mathbb{N}} \mathcal{F}_i$. The projection map π_j in coordinate $j \in \mathbb{N}$ is a function mapping $\prod_{i\in\mathbb{N}} \Omega_i$ to Ω_j which takes point $(\omega_1, \omega_2, \ldots)$ to ω_j. Then the product σ-algebra is given by $\sigma(\pi_1, \pi_2, \ldots)$. Using this it is easy to see that $\prod_{i\in\mathbb{N}} \mathcal{F}_i$ is also $\sigma(\{\prod_{i\in\mathbb{N}} B_i : B_i \in \mathcal{F}_i \ \forall i \in \mathbb{N}\})$. One can show that $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R})^n$. However, $\mathcal{B}(\mathbb{R}^{\mathbb{N}}) \neq \mathcal{B}(\mathbb{R})^{\mathbb{N}}$; since we deal with the product σ-algebra , we will always be using the latter.

π-systems: A set $\mathcal{S} \subseteq 2^\Omega$ is called a π-system if it is closed under finite intersections. Eg. $(**)$ $\Pi(\mathbb{R}) := \{(-\infty, x] : x \in \mathbb{R}\}$. An important fact is that $\sigma(\Pi(\mathbb{R})) = \mathcal{B}(\mathbb{R})$.

Probability measure: Given a measurable space $(\Omega, \mathcal{F})$, a non-negative set function $\mathbb{P} : \mathcal{F} \mapsto [0,1]$ is called a probability measure if

(1) (Normalized) If $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\emptyset) = 0$; and

(2) (Countably additive) If $\{F_i\}_{i\in\mathbb{N}}$ is a countable and disjoint ($F_i \cap F_j = \emptyset \; \forall i \neq j \in \mathbb{N}$) collection of sets from $\mathcal{F}$, then

$$\mathbb{P}(\cup_{i\in\mathbb{N}} F_i) = \sum_{i=1}^{\infty} \mathbb{P}(F_i).$$

We then call $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space.
Examples:

(1) $\Omega = \{0,1\}$, $\mathcal{F} = 2^\Omega = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$. Given $p \in [0,1]$ we can define $\mathbb{P}$ by the following values $\{0, 1-p, p, 1\}$ assigned to each element of $\mathcal{F}$.

(2) $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$. Given $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ we define the probability of set $B \in \mathcal{F}$ by

$$\mathbb{P}(B) = \int_{-\infty}^{\infty} 1_B(x) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

where we have

$$1_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise} \end{cases}$$

Note that for $B = [a,b]$ ($a < b \in \mathbb{R}$), i.e., an interval we get

$$\mathbb{P}([a,b]) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

This is a Gaussian probability distribution - $\mathcal{N}(\mu, \sigma^2)$.

(3) $\Omega = [0,1]$ and $\mathcal{F} = \mathcal{B}([0,1])$ (defined in the same manner as $\mathcal{B}(\mathbb{R})$). For $B \in \mathcal{F}$ we set $\mathbb{P}(B)$ to be the Lebesgue measure of B. For an interval $[a,b]$ with $0 \leq a \leq b \leq 1$, we get $\mathbb{P}([a,b]) = b - a$, i.e., the uniform probability distribution.

$(***)$ Fact: If $\mathcal{S} \subseteq 2^\Omega$ is a π-system that generates $\mathcal{F}$, i.e., $\sigma(\mathcal{S}) = \mathcal{F}$, and we have two probability measures $\mathbb{P}_1$ and $\mathbb{P}_2$ such that $\mathbb{P}_1(A) = \mathbb{P}_2(A)$ for all $A \in \mathcal{S}$, then $\mathbb{P}_1$ and $\mathbb{P}_2$ agree on $\mathcal{F}$. In other words, defining a probability measure on a π-system uniquely defines it over the σ-algebra generated by the π-system.

$(***)$ $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$. Let $\mathcal{S} = \Pi(\mathbb{R})$, then defining $\mathbb{P}$ on elements of $\Pi(\mathbb{R})$ uniquely specifies it for $\mathcal{B}(\mathbb{R})$! Note that each element of $\Pi(\mathbb{R})$ is given as $(-\infty, x]$ for some $x \in \mathbb{R}$ so we only need to know the following as a function of x,

$$\mathbb{P}((-\infty, x]) =: F(x) \quad \text{(Cumulative distribution function)}$$

1.2. **Random variables.** Let $(\Omega, \mathcal{F})$ be a measurable space; usually left abstract for convenience and richness [1]. **A random variable is a function $X : \Omega \mapsto \mathbb{R}$ such that X is (Borel) measurable, i.e., $X^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{B}(\mathbb{R})$.** In general, let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces, then $X : \Omega_1 \mapsto \Omega_2$ is a random variable if it is $\mathcal{F}_1/\mathcal{F}_2$ measurable, i.e., if $X^{-1}(B) \in \mathcal{F}_1$ for all $B \in \mathcal{F}_2$. If $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$, then given a random variable X we can define an induced probability measure $\mathbb{P}_X$ by assigning values based on the inverse map, i.e., $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$ for all $B \in \mathcal{B}(\mathbb{R})$.

Examples: We will assume $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$, $\mathbb{P}$ is Lebesgue measure on $[0,1]$ (same as uniform distribution denoted by $Leb([0,1])$). We will generate random variables by showing the mapping for each $\omega \in \Omega$.

[1] We will provide a rich enough example for all our purposes later on.

(1) Given ω, expand it in binary form as $0.\omega_1\omega_2\ldots$, then $X(\omega) = \omega_1$ is a random variable with respect to $(\Omega, \mathcal{F})$ or $(\{0,1\}, 2^{\{0,1\}})$. What is the induced measure? $\mathbb{P}_X(X=0) = \mathbb{P}_X(X=1) = 1/2$.
(2) Given ω, expand it in decimal form as $0.\omega_1\omega_2\ldots$, then $X(\omega) = \omega_1$ is a random variable with respect to $(\Omega, \mathcal{F})$ or $(\{0,1,\ldots,9\}, 2^{\{0,1,\ldots,9\}})$. What is the induced probability measure? $\mathbb{P}_X(X=0) = \mathbb{P}_X(X=1) = \cdots = \mathbb{P}_X(X=9) = 1/10$.
(3) Define $X(\omega) = -\log(\omega)$. Here we can take the measurable space to be $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. What is the induced probability measure? Verify that this yields that exponential distribution.

The probability space above, i.e., $([0,1], \mathcal{B}([0,1]), Leb([0,1]))$, is rich enough to be the canonical space that we will work with in most of this class. Thus, for convenience it'd be fine to imagine this space when we take an abstract $(\Omega, \mathcal{F}, \mathbb{P})$ to construct our random variables.

σ-algebra revisited - Earlier we defined $\sigma(f)$ for a special class of functions. Note that the f there was a measurable function. Thus, we actually defined $\sigma(X)$ for a random variable X. Thus, $\sigma(X) = \sigma(\{\omega \in \Omega : X(\omega) \in B\} : B \in \mathcal{B}(\mathbb{R})\})$. This can be easily generalized to a sequence of random variables $\{X_i\}_{i\in\mathcal{I}}$ (where $\mathcal{I}$ is at most a countable set) as follows $\sigma(\{X_i\}_{i\in\mathcal{I}}) = \sigma(\{\omega \in \Omega : X_i(\omega) \in B\} : i \in \mathcal{I}, B \in \mathcal{B}(\mathbb{R})\})$. A similar generalization holds if the random variables take values in another measurable space or different measurable spaces for each random variable. *Way to view or understand*: For $\omega \in \Omega$, $\{X_i(\omega)\}_{i\in\mathcal{I}}$ is the observed sequence, then $\sigma(\{X_i\}_{i\in\mathcal{I}})$ consists of all events/sets $F \in \mathcal{F}$ which one can use the said sequence to decide with certainty if $\omega \in F$ or not.
Examples:

(1) $\Omega = \{0,1\}^2$ and $\mathcal{F} = 2^\Omega$. Three random variables: X being the first bit and Y being the last bit and $Z = X \oplus Y$ (convince yourself that these are random variables). What are $\sigma(X)$, $\sigma(Y)$ and $\sigma(Z)$? $X^{-1}(\{0\}) = \{(0,1),(0,0)\}$ and $X^{-1}(\{1\}) = \{(1,0),(1,1)\}$. Therefore, $\sigma(X) = \{\emptyset, \{(0,1),(0,0)\}, \{(1,0),(1,1)\}, \Omega\}$ which is strictly smaller than $\mathcal{F}$. Work out the other examples on your own.
(2) $\Omega = [0,1]$ and $\mathcal{F} = \mathcal{B}([0,1])$. Three random variables given by

$$X(\omega) = \omega, \qquad Y(\omega) = \omega^2,$$
$$Z(\omega) = \begin{cases} 0 & \text{if } \omega \in [0,0.5) \\ 1 & \text{if } \omega \in [0.5,1] \end{cases}$$

It is easy to see that $\sigma(Z) = \{\emptyset, [0,0.5), [0.5,1], [0,1]\}$. What about $\sigma(X)$ and $\sigma(Y)$?

$(***)$ **Independence** - Let X and Y be two random variables. *When are they independent?* We will present the most general definition and later on show how it corresponds to what is usually presented.

Definition 1. *A set of σ-algebra $\{\mathcal{G}_i\}_{i\in\mathbb{N}}$ (sub σ-algebra of $\mathcal{F}$) are said to be independent if whenever $G_i \in \mathcal{G}_i$ ($i \in \mathbb{N}$) and for every distinct $i_1, \ldots, i_n$ ($n \in \mathbb{N}$), then*

$$\mathbb{P}(\cap_{j=1}^n G_{i_j}) = \prod_{j=1}^n \mathbb{P}(G_{i_j})$$

A sequence of random variables $\{X_i\}_{i\in\mathbb{N}}$ *are said to be independent if* $\{\sigma(X_i)\}_{i\in\mathbb{N}}$ *is independent. A sequence of events* $\{F_n\}_{n\in\mathbb{N}}$ *are said to be independent if the indicator functions (bounded random variables)* $\{1_{F_n}\}_{n\in\mathbb{N}}$ *are independent. Note that* $\sigma(F_n) = \sigma(1_{F_n}) = \{\emptyset, F_n, F_n^C, \Omega\}$.

1.3. **Expectation.** Let X be a random variable and $\mathbb{P}_X$ the induced probability measure. Then the expectation of X, $\mathbb{E}[X]$, is the integral of X with respect to $\mathbb{P}_X$ (if it exists), i.e.,

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x d\mathbb{P}_X(x)$$

If $Y = f(X)$ for some Borel function f, then the expectation of Y (if it exists) is given by

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} f(x) d\mathbb{P}_X(x)$$

Examples:

(1) Let $\mathbb{P}_X$ be a discrete probability measure, i.e., $\mathbb{P}_X = \sum_{i\in\mathbb{N}} p_i \delta_{x_i}$ for some (at most) countable $\{x_i\} \subset \mathbb{R}$ with $\sum_{i\in\mathbb{N}} p_i = 1$, where δ_x is the Dirac measure at x, i.e., for every $B \in \mathcal{B}(\mathbb{R})$ we have

$$\delta_x(B) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise} \end{cases}$$

Then $\mathbb{E}[Y] = \sum_{i\in\mathbb{N}} p_i f(x_i)$ if either all terms are non-negative or if the series is absolutely summable.

(2) Let $\mathbb{P}_X$ have a density (with respect to Lebesgue measure), i.e., there exists p_x a non-negative, Borel measurable and integrable function such that

$$\mathbb{P}_X(B) = \int_{-\infty}^{\infty} 1_B(x) p_X(x) dx,$$

then $\mathbb{E}[Y] = \int_{-\infty}^{\infty} f(x) p_X(x) dx$ if the integral exists.

We also define independence via expectation. A sequence of random variables $\{X_i\}_{i\in\mathbb{N}}$ is said to be independent if and only if for every distinct $i_1, \ldots, i_n$ ($n \in \mathbb{N}$) and every collection of bounded measurable functions $\{f_{i_k}\}_{k=1,\ldots,n}$ (each f_{i_k} begin $\sigma(X_{i_k})$-measurable) we have

$$\mathbb{E}[\prod_{k=1}^{n} f_{i_k}(X_{i_k})] = \prod_{k=1}^{n} \mathbb{E}[f_{i_k}(X_{i_k})]$$

Contrast this with the previous definition. This product decomposition also holds for every integrable function in the case of independence.

Items for review: We will assume knowledge of these from now onwards.

- Mean, variance, covariance and relationship between these.
- Definition of moments, moment generating function, probability generating function.
- Cauchy-Bunyakovsky-Schwarz inequality, Jensen's inequality.
- Markov inequality, Chebyshev inequality, Chernoff bound.
- Fubini's theorem and Tonelli's theorem on exchange order of integration for multiple integrals.

Details of the above to be added to notes at a later date.

1.4. **Modes of convergence.** We assume that we have a sequence of random variables $\{X_i\}_{i\in\mathbb{N}}$ and we will investigate means by which the sequence converges to another random variable X. Lots of questions arise though. *Since a random variable is a function, are we talking about convergence for every $\omega \in \Omega$? Since we also have a probability $\mathbb{P}$ on $(\Omega, \mathcal{F})$, can we somehow use it to determine convergence? What about some convergence in terms of the induced probability measures, i.e., $\{\mathbb{P}_{X_i}\}_{i\in\mathbb{N}}$ and $\mathbb{P}_X$? What about convergence in terms of means (if they exist)? What about other moments?* There is a notion of convergence attached to each of these questions. The relationship between these will be explored in the exercises. For use later on we define for a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$, the following two quantities $\limsup x_n := \inf_m \sup_{n\geq m} x_n$ and $\liminf x_n := \sup_m \inf_{n\geq m} x_n$. Both of these always exist and if they are equal, then $\lim_{n\to\infty} x_n$ exists.

Sure convergence - Here one insists that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for all $\omega \in \Omega$. This is generally very strong a notion. Take our usual probability space $(\Omega, \mathcal{F}, \mathbb{P})$. It is easy to see that one can change the value of each X_n at a countable number of ω to get $\tilde{X}_n$ such that sure convergence does not hold another. However, X_n and $\tilde{X}_n$ are much the same: same means and other moments (if they exist, and even existence or not has same answer), same induced probability measure, and $\mathbb{P}(|X_n - \tilde{X}_n| < \epsilon) = 0$ for all $\epsilon > 0$. Note that we modified the random variables on a set of probability 0 (with respect to $\mathbb{P}$). This idea leads to next notion of convergence.

$(***)$**Almost sure/almost everywhere convergence (a.s./a.e.)** - Here one includes $\mathbb{P}$ in the definition to say that convergence occurs if

$$\mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$$

This is denoted by either of the following

$$\lim_{n\to\infty} X_n = X \ a.s. \text{ OR } \lim_{n\to\infty} X_n \stackrel{a.s.}{=} X$$

This notion of convergence extends to a general metric space but requires that X_n and X come from the same space and take values in the same space. Question:*Since* $\lim_{n\to\infty} X_n = X$ *a.s., does it hold that* $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$ *(when the expectations exist)? In other words, can we interchange the order of integration and limits?* In general, NO!! On our usual probability space define the following sequence

$$X_n(\omega) = \begin{cases} 4n^2\omega & \text{if } \omega \in [0, \frac{1}{2n}] \\ 4n^2(\frac{1}{n} - \omega) & \text{if } \omega \in [\frac{1}{2n}, \frac{1}{n}] \\ 0 & \text{if } \omega \in (\frac{1}{n}, 0] \end{cases}$$

Note that $\lim_{n\to\infty} X_n = 0$ *a.s.* but for all n we have $\mathbb{E}[X_n] = 1 \neq 0$! However, there are conditions under which we can interchange limits and expectation. These are the outcome of the following results (Proofs can be found in standard textbooks):

(1) (*Monotone convergence theorem): Let* $X_n \geq 0$ a.s. *and* $X_n \uparrow X$ *(increase), then* $\mathbb{E}[X_n] \uparrow \mathbb{E}[X] \leq \infty$.
(2) *(Fatou's lemmas): Let* $X_n \geq 0$ a.s., *then we have (a) (Commonly used)* $\mathbb{E}[\liminf X_n] \leq \liminf \mathbb{E}[X_n]$*; and (b) if, in addition,* $X_n \leq Y$ *for all* n *with* $\mathbb{E}[Y] \leq \infty$*, the* $\mathbb{E}[\limsup X_n] \geq \limsup \mathbb{E}[X_n]$.
(3) *(Bounded convergence theorem): If* $|X_n| \leq K$ a.s. *for all* n*, then* $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X] \leq K$.

(4) *(Dominated convergence theorem): If* $|X_n| \leq Y$ a.s. *for all* n *with* $\mathbb{E}[Y] < \infty$*, then* $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X] \leq \mathbb{E}[Y]$.

(5) *(Scheffe's Lemma): If* $\{X_n\}_{n\in\mathbb{N}}$ *and* X *integrable (finite expectations), then* $\lim_{n\to 0} \mathbb{E}[|X_n - X|] = 0$ *if and only if (iff.)* $\lim_{n\to\infty} \mathbb{E}[X_n] = E[X]$ *or* $\lim_{n\to\infty} \mathbb{E}[|X_n|] = E[|X|]$.

$(***)$ **Convergence in probability** - This holds if for all $\epsilon > 0$, we have $\mathbb{P}(|X_n - X| > \epsilon) = 0$. This denoted by $\lim_{n\to\infty} X_n = X$ p OR $\lim_{n\to\infty} X_n \stackrel{p}{=} X$. This notion of convergence extends to a general metric space but requires that X_n and X come from the same space and take values in the same space. Again, one cannot interchange limits and expectation. Here the above conditions for $(a.s)$ convergence apply.

$(***)$ **Convergence in distribution** - $\lim_{n\to\infty} X_n \stackrel{d}{=} X$ if $\lim_{n\to\infty} F_n(x) = F(x)$ at all points of continuity of F, where F_n is the cdf of X_n and F the cdf of F. *Why the restriction on points of continuity of F?* Let X_n be uniform in $[0, \frac{1}{n}]$, then $\lim_{n\to\infty} X_n \stackrel{d}{=} 0$. Note that $\lim_{n\to\infty} F_n(x) = 1 = F(x)$ if $x > 0$ and $\lim_{n\to\infty} F_n(x) = 0 = F(x)$ if $x < 0$, but $F_n(0) = 0$ for n while $F(0) = 1$. The same definition carries through to vector valued random variables. However, for random processes we need to consider what is known as weak convergence. Note that we only need X_n and X to take values on the same space but they could be coming from different probability spaces. Once again, one cannot interchange limits and expectation. However, the interchange can be safely used for bounded continuous functions.

$(***)$ *Skorokhod representation theorem* - In great generality (depending only on properties of the probability measure of X), if $\lim_{n\to\infty} X_n \stackrel{d}{=} X$, then one can find a common probability space and random variables $\{\tilde{X}_n\}_{n\in\mathbb{N}}$ and $\tilde{X}$ with same induced probability measures (only marginals and not joint distributions) as $\{X_n\}_{n\in\mathbb{N}}$ and X, respectively, such that $\lim_{n\to\infty} \tilde{X}_n \stackrel{a.s.}{=} \tilde{X}$.

$(***)$ *Convergence in mean* - This holds if $\lim_{n\to\infty} \mathbb{E}[|X_n - X|] = 0$.

$(***)$ *Convergence in mean square (m.s.)* - This holds if $\lim_{n\to\infty} \mathbb{E}[|X_n - X|^2] = 0$.

$(***)$ **Borel-Cantelli Lemmas** - The first lemma is the following.

Lemma 1. *Let* $\{F_n\}_{n\in\mathbb{N}}$ *be a sequence of events (from $\mathcal{F}$) such that* $\sum_{n\in\mathbb{N}} \mathbb{P}(F_n) < \infty$*, then* $\mathbb{P}(\limsup F_n) = 0$*, where* $\limsup F_n := \cap_m \cup_{n\geq m} F_n$*, i.e.,* $\{\omega : \omega \in F_n$ *for infinitely many* $n\}$.

In words, we say that $\limsup F_n$ is all the ω that are in F_n infinitely often $(i.o.)$. We also define $\liminf F_n = \cup_m \cap_{n\geq m} F_n$, i.e., the $\{\omega : \omega \in F_n$ for all large $n\}$ for the ω that are eventually in F_n. *Do Fatou's lemmas say something here?*
The second lemma is the following.

Lemma 2. *If* $\{F_n\}_{n\in\mathbb{N}}$ *be a sequence of independent events, then* $\sum_{n\in\mathbb{N}} \mathbb{P}(F_n) = \infty$ *implies that* $\mathbb{P}(\limsup F_n) = 1$.

Items to review -

- Read up proofs of both Borel-Cantelli Lemmas; also note the Kolmogorov $0-1$ law.
- Review Strong Law and Weak Law of Large Numbers.
- Review Central Limit Theorem - deviations from mean.
- Review Cramér's Theorem - large deviations from mean.

$(***)$ **Random process/Stochastic process** - Given a fully ordered index set $\mathcal{I}$ (usually $\mathbb{N}, \mathbb{Z}, \mathbb{Z}_+, \mathbb{R}, \mathbb{R}_+$) and a (measurable) state space $(E, \mathcal{E})$, **a stochastic process is function $X : \mathcal{I} \times \Omega \mapsto E$ such that for each $i \in \mathcal{I}$ we have $X(i, \cdot) : \Omega \mapsto E$ is an E-valued random variable.**
Examples:

(1) $\{X_i\}_{i\in\mathbb{N}}$ with each X_i being an $([0,1],\mathcal{B}([0,1]))$ random variable, say uniformly distributed. In addition, one can insist that the random variables be independent.
(2) (Poisson process) $\{N_t\}_{t\in\mathbb{R}_+}$ where we have for all $t \in \mathbb{R}_+$ that $N_t \in \mathbb{N}$ is a random variable that has the Poisson distribution (with parameter t). Note that a Poisson process has independent increments with jumps of height 1.
(3) (Brownian motion/Weiner process) $\{W_t\}_{t\in\mathbb{R}_+}$ where we have for all $t \in \mathbb{R}_+$ that $W_t \in \mathbb{N}$ is a random variable that has the Gaussian distribution (with mean 0 and variance t). A Weiner process also has independent increments and has continuous sample paths.

1.5. **Conditional expectation.** $(***)$ Let X be an integrable random variable and $\mathcal{G}$ a sub σ-algebra of $\mathcal{F}$, i.e., $\mathcal{G} \subseteq \mathcal{F}$ and also a σ-algebra . Then the conditional expectation of X with respect to $\mathcal{G}$ is a $\mathcal{G}$-measurable integrable random variable $\mathbb{E}[X|\mathcal{G}]$ such that

$$\mathbb{E}[X1_G] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]1_G] \quad \forall G \in \mathcal{G}$$

If $\tilde{X}$ is any other random variable with these properties, then $\tilde{X} = \mathbb{E}[X|\mathcal{G}]$ *a.s.*; hence, called a version of the conditional expectation. In the above definition, it suffices to satisfy the equations for a π-system that contains Ω and generates $\mathcal{G}$. Given two random variables X and Y, we have $\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$. Eg. the regular expectation $\mathbb{E}[X]$ is conditional expectation with respect to the simplest σ-algebra $\{\emptyset,\Omega\}$.
Property 1: (Tower property) If $\mathcal{H}$ is a sub σ-algebra of $\mathcal{G}$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.
Property 2: If Y is a $\mathcal{G}$-measurable and bounded random variable, then $\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$. Also, if $X, Y \geq 0$ and $\max(\mathbb{E}[X],\mathbb{E}[Y],\mathbb{E}[XY]) < \infty$ with Y being $\mathcal{G}$-measurable, then $\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$.
Property 3: (Independence) If $\mathcal{H}$ is independent of $\sigma(X,\mathcal{G})$, then $\mathbb{E}[X|\sigma(\mathcal{G},\mathcal{H})] = \mathbb{E}[X|\mathcal{G}]$. From this one can show the following useful facts: $\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[X]$ and $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.

$(***)$ *Filtrations* - A collection $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ of sub σ-algebra of $\mathcal{F}$ with a fully ordered index set $\mathcal{I}$ is called a filtration if for all $i, j \in \mathcal{I}$ with $i \leq j$, we have $\mathcal{F}_i \subseteq \mathcal{F}_j$. Define $\mathcal{F}_\infty = \sigma(\cup_{i\in\mathcal{I}})$. Eg. Let $\{X_i\}_{i\in\mathcal{I}}$ be a stochastic process, then $\mathcal{F}_i = \sigma(X_k : k \leq i)$ is a filtration; it is called the natural filtration of $\{X_i\}_{i\in\mathcal{I}}$. If $\mathcal{I}$ is $\mathbb{R}$ or $\mathbb{R}_+$, we may demand other properties for the filtration but this too advanced for this course.

$(***)$ *Adapted process* - A stochastic process $\{X_i\}_{i\in\mathcal{I}}$ is said to be adapted to a filtration $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ if for all $i \in \mathcal{I}$, X_i is $\mathcal{F}_i$-measurable. Eg. $\{X_i\}_{i\in\mathcal{I}}$ is always adapted to its natural filtration.

$(***)$ *Stopping time* - A function $T : \Omega \mapsto \mathcal{I}$ is a stopping time if we have either of the following two equivalent definitions hold,

(1) The event $\{T \leq i\}$ is $\mathcal{F}_i$ measurable for $i \in \mathcal{I} \cup \infty$, i.e, $\{\omega : T(\omega) \leq i\} \in \mathcal{F}_i$;
(2) The event $\{T = i\} = \{\omega : T(\omega) = n\} \in \mathcal{F}_i$ for all $i \in \mathcal{I} \cup \{\infty\}$.

Intuition: Let $\{X_i\}_{i\in\mathbb{N}}$ be a stochastic process and $\{\mathcal{F}_i\}_{i\in\mathbb{N}}$ be its natural filtration, then $\{T = n\} = f_n(X_1,\ldots,X_n)$ for some Borel measurable f_n, i.e., one can decide based only on $\{X_1,\ldots,X_n\}$ and no future is needed.
Example: Let $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$ be a filtration and $\{X_i\}_{i\in\mathbb{Z}_+}$ an adapted stochastic process, then $T = \inf\{n \geq 0 : X_n \in B\}$ (for some $B \in \mathcal{F}$), i.e., the time $\{X_i\}_{i\in\mathbb{Z}_+}$ first enters B (Hitting time), is a stopping time. By convention we have $\inf(\emptyset) = \infty$ so $T = \infty$ holds only if $\{X_n\}_{n\in\mathbb{Z}_+}$ never enters B. It is easy to see that $\{T \leq n\} = \cup_{0\leq k\leq n}\{X_n \in B\} \in \mathcal{F}_n$ for all $n \in \mathbb{Z}_+$. In contrast, the random variable $L = \sup\{n \leq 100 : X_n \in B\}$ (with $\sup(\emptyset) = 0$ by

convention for this case [2]). In general, L is not a stopping time. We will see many examples of stopping times in this course.

[2] If $\mathcal{I}$ were $\mathbb{Z}$, then we will take $\sup(\emptyset) = -\infty$.

1.6. **Exercises.** These are exercises for Section 1 with the highlighted ones being for extra-credit. The introductory text gives the context of the question.

(1) *Algebra* - Show that $\emptyset \in \tilde{\mathcal{F}}$, and if $F_1, \ldots, F_n \in \tilde{\mathcal{F}}$, then $\cup_{i=1}^n F_n \in \tilde{\mathcal{F}}$ and $\cap_{i=1}^n F_n \in \tilde{\mathcal{F}}$. *Hint*: Use DeMorgan's laws.

(2) *Algebra* - If $\{F_i\}_{i\in\mathbb{N}}$ is a countable collection of sets from $\tilde{\mathcal{F}}$, does it hold that $\cup_{i\in\mathbb{N}} F_i \in \tilde{\mathcal{F}}$? *Hint*: work with the following - $\Omega = \mathbb{N}$ and $\tilde{\mathcal{F}} := \{F \subseteq \Omega :$ either F or F^C is finite$\}$. First verify that said $\tilde{\mathcal{F}}$ is indeed an algebra and then test the hypothesis. *Addendum*: For a σ-algebra show that $\cap_{i\in\mathbb{N}} F_i \in \tilde{\mathcal{F}}$ also holds.

(3) *σ-algebra* - Prove that example $(\star)$ is a σ-algebra . Compare it to $2^{\mathbb{R}}$.

(4) *π-systems* - Prove that $\{\prod_{i\in\mathbb{N}} B_i : B_i \in \mathcal{F}_i \ \forall i \in \mathbb{N}\}$ is a π-system; note that this generates $\prod_{i\in\mathbb{N}} \mathcal{F}_i$. Prove that $\sigma(\Pi(\mathbb{R})) = \mathcal{B}(\mathbb{R})$.

(5) *π-systems* - For a given function $f : \Omega_1 \mapsto \Omega_2$ where $f^{-1}(B) := \{\omega \in \Omega_1 : f(\omega) \in B\} \in \mathcal{F}_1$ for all $B \in \mathcal{F}_2$, show that $\{f^{-1}(B) : B \in \mathcal{F}_2\}$ is a π-system.

(6) *Random variables* - Prove that $X(\omega) = -\log(\omega)$ is a random variable and prove that one gets the exponential distribution. Prove that $X(\omega) = \cos(\omega)$ is a random variable and derive the distribution. *Hint*: use a π-system or continuity of the functions.

(7) *σ-algebra revisited* - Work out the examples not covered in class. *Hint*: in some cases you could a π-system or continuity.

(8) *σ-algebra revisited* - Let X and Y be two random variables. Prove that Y is $\sigma(X)$ measurable if and only if $Y = f(X)$ for some Borel measurable function f. [**Extra credit**]

(9) *Independence* - Let us use our canonical probability space. For $\omega \in [0,1]$ we expand in binary form as $0.\omega_1\omega_2\ldots$. Let $X_i(\omega) = \omega_i$. Show the independence of X_1 and X_2. Show the independence of $\{X_i\}_{i\in\mathbb{N}}$. What is the distribution of each random variable?

(10) *Independence* - For the same example, we now make the following definitions

$$\begin{aligned}
X_1(\omega) &= 0.\omega_1\omega_3\omega_6\omega_{10}\ldots \\
X_2(\omega) &= 0.\omega_2\omega_5\omega_9\ldots \\
X_3(\omega) &= 0.\omega_4\omega_8\ldots \\
X_4(\omega) &= 0.\omega_7\ldots \\
&\vdots
\end{aligned}$$

where we have used a bijection mapping from $\mathbb{N}^2$ to $\mathbb{N}$. Show that X_i is a random variable for all $i \in \mathbb{N}$. Argue that $\{X_i\}_{i\in\mathbb{N}}$ is independent and identically distributed (*i.i.d.*) with the uniform distribution on $[0,1]$. Use this to generate an $\mathbb{R}$-valued *i.i.d.* sequence of any given distribution F. *Hint*: Since $F : \mathbb{R} \mapsto [0,1]$ (with additional properties), invert it and use it to define the required random variables. Note that this uses the simplest version of the Skorokhod representation. [**Extra credit**]

(11) *Convergence* - Show that $\limsup X_n$, $\liminf X_n$ are measurable functions (first show for sup and inf of random variables and then use π-systems), convince yourself that sums and difference of measurable functions are measurable, and finally show that the set where $\lim_{n\to\infty} X_n$ exists is measurable. This shows that we can assign a probability in the definition of *a.s.* convergence without any worry.

(12) *Convergence* - In example from class where $\lim_{n\to\infty} X_n \overset{a.s.}{=} X$ but $\mathbb{E}[X_n]$ does not converge to $\mathbb{E}[X]$, show that the dominated convergence theorem does not apply: plot $Y(\omega) = \sup_n X_n(\omega)$ (is it a random variable?) and show that $\mathbb{E}[Y] = \infty$.

(13) *Convergence* - Show the following: *a.s.* $\Rightarrow p \Rightarrow d$ and convergence in mean square $\Rightarrow$ convergence in mean $\Rightarrow p$; this is short-hand for one type of convergence implying the other, naturally, assuming that all are from the same probability space. For the moment-based convergence results use Jensen's inequality, Markov inequality and Chebyshev inequality. For parts of this, it the next exercise may also be used.

(14) *Convergence* - If $\lim_{n\to\infty} X_n \overset{p}{=} X$ and $\sum_{n\in\mathbb{N}} \mathbb{P}(|X_n - X| > \epsilon) < \infty$ for all $\epsilon > 0$, then $\lim_{n\to\infty} X_n \overset{a.s.}{=} X$. *Hint*: Use the first Borel-Cantelli Lemma. Show how this implies that if $\lim_{n\to\infty} X_n \overset{p}{=} X$, then every subsequence $\{X_{n_m}\}$ has a subsequence such that $\{X_{n_{m_k}}\}$ that converges *a.s.*. [**Extra credit**]

(15) *Convergence* - Let $\lim_{n\to\infty} X_n \overset{p}{=} X$ with $|X_n| \le Y$ *a.s.* with Y integrable, then show that $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$ and $\lim_{n\to\infty} \mathbb{E}[|X_n - X|] = 0$. *Hint*: use the result above.

(16) *Conditional expectation* - If X, Y have a joint density $f_{X,Y}(x, y)$ and X is integrable, define the conditional density $f_{X|Y}(x|y)$ as

$$f_{X|Y}(x|y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{if } f_Y(y) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Define $g(y) := \int_{x\in\mathbb{R}} x f_{X|Y}(x|y)dx$. Show that $g(Y)$ is a version of the conditional expectation $\mathbb{E}[X|Y]$. *Hint*: use Fubini's theorem. What is the equivalent expression in the discrete case? Note that this shows how the definition in class includes previous cases.

(17) *Conditional expectation* - Prove follow-up statements of *Property 3*.

(18) *Conditional expectation* - State and prove convergence theorems for conditional expectation. State and prove a conditional form of Jensen's theorem. [**Extra credit**]

2. Martingales

Sources: The sources for this chapter are

- D. Williams, "Probability with martingales," Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991.
- R. Durrett, "Probability: Theory and examples," Second edition. Duxbury Press, Belmont, CA, 1996.

Stopping times - Before discussing martingales, we will point out a few more properties of stopping times. Let $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$ be the filtration of interest, then these are:

(1) Any $n \in \mathbb{Z}_+$ is a stopping time;
(2) If S and T are two stopping times (on the same filtration), then so are $T \vee S := \max(S,T)$, $T \wedge S := \min(S,T)$ and $S+T$;
(3) If $\{T_i\}_{i\in\mathbb{N}}$ are a sequence of stopping times (with respect to the same filtration), then $\sup_{n\in\mathbb{N}} T_n$ and $\inf_{n\in\mathbb{N}} T_n$ also stopping times.

2.1. Martingale.

Let $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$ be a filtration and $\{X_i\}_{i\in\mathbb{Z}_+}$ a random process taking values in $\mathbb{R}$. Define $\mathcal{F}_\infty := \sigma(\cup_{i\in\mathbb{Z}_+}\mathcal{F}_i)$.

Definition 2. *Relative to* $(\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}, \mathbb{P})$*, an adapted process* $\{X_i\}_{i\in\mathbb{Z}_+}$ *that is integrable, i.e.,* $\mathbb{E}[|X_n|] < \infty$ *for all* $n \in \mathbb{Z}_+$*, is called*

(1) *a martingale if* $\mathbb{E}[X_n|\mathcal{F}_{n-1}] = X_{n-1}$ a.s. *for all* $n \in \mathbb{N}$*;*
(2) *a submartingale if* $\mathbb{E}[X_n|\mathcal{F}_{n-1}] \geq X_{n-1}$ a.s. *for all* $n \in \mathbb{N}$*;*
(3) *a supermartingale if* $\mathbb{E}[X_n|\mathcal{F}_{n-1}] \leq X_{n-1}$ a.s. *for all* $n \in \mathbb{N}$*;*

The above can be written as $\mathbb{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}] \gtreqless 0$, so that in an average sense a submartingale increases, a martingale stays constant and a supermartingale decreases.

Examples:

(1) Let $\{X_i\}_{i\in\mathbb{N}}$ be an independent and mean 0 sequence. Then the process $\{S_i\}_{i\in\mathbb{Z}_+}$ with $S_0 = 0$ and $S_n = \sum_{i=1}^n X_i$ for $n \in \mathbb{N}$ is a martingale with respect to the natural filtration of $\{X_i\}_{i\in\mathbb{N}}$, i.e., $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$ for $n \in \mathbb{N}$. Clearly, $\{S_i\}_{i\in\mathbb{Z}_+}$ is adapted and integrable and by the inherent independence, we get
$$\mathbb{E}[S_n|\mathcal{F}_{n-1}] = S_{n-1} + \mathbb{E}[X_n|\mathcal{F}_{n-1}] = S_{n-1} + \mathbb{E}[X_n] = S_{n-1}$$
where each equality is to be interpreted in an *a.s.* sense. Thus, we get a martingale. Note that $\{S_i\}_{i\in\mathbb{Z}_+}$ is also a martingale with respect to its natural filtration; note here that $\sigma(S_1, S_2, \ldots, S_n) = \sigma(X_1, X_2, \ldots, X_n)$ (*why?*). If the means are all positive, then we get a submartingale, whereas we get a supermartingale if the means are all negative.
(2) Let $\{X_i\}_{i\in\mathbb{N}}$ be an independent, non-negative and mean 1 sequence. Then the process $\{S_i\}_{i\in\mathbb{Z}_+}$ with $S_0 = 1$ and $S_n = \prod_{i=1}^n X_i$ for $n \in \mathbb{N}$ is a martingale with respect to the natural filtration of $\{X_i\}_{i\in\mathbb{N}}$. Again, adaptedness and integrability follow easily, and from the inherent independence we get
$$\mathbb{E}[S_n|\mathcal{F}_{n-1}] = \mathbb{E}[S_{n-1}X_n|\mathcal{F}_{n-1}] = S_{n-1}\mathbb{E}[X_n|\mathcal{F}_{n-1}] = S_{n-1}$$
Note that we used the property that S_{n-1} is non-negative and $\mathcal{F}_{n-1}$-measurable along with independence of $\sigma(X_n)$ and $\mathcal{F}_{n-1}$ and $\mathbb{E}[X_n] = 1$. Therefore, we again get a martingale. If the means are all greater than 1, then we get a submartingale, whereas we get a supermartingale if the means are all less than 1. Here too we could restrict

attention to the natural filtration of $\{S_i\}_{i\in\mathbb{Z}_+}$, but we may not have $\sigma(S_1, S_2, \ldots, S_n)$ being equal to $\sigma(X_1, X_2, \ldots, X_n)$ (*why?*).

(3) Let Y be an integrable random variable. Given a filtration $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$ define the following random process $X_i = \mathbb{E}[Y|\mathcal{F}_i]$ for all $i \in \mathbb{Z}_+$; clearly $\{X_i\}_{i\in\mathbb{Z}_+}$ is adapted and integrable. By the Tower property of conditional expectation we get

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] = \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_n]|\mathcal{F}_{n-1}] = \mathbb{E}[Y|\mathcal{F}_{n-1}] = X_{n-1}$$

Again we have a martingale.

(4) If $\{X_i\}_{i\in\mathbb{Z}_+}$ is a submartingale, then $\{Y_i\}_{i\in\mathbb{Z}_+}$ with $Y_i = -X_i$ is a supermartingale, and *vice-versa*. Thus, if $\{X_i\}_{i\in\mathbb{Z}_+}$ is a martingale, then it is both a submartingale and a supermartingale.

Note that the Tower property yields the following relationship when $m < n$:

$$\mathbb{E}[X_n|\mathcal{F}_m] = \mathbb{E}[\mathbb{E}[X_n|\mathcal{F}_{n-1}]|\mathcal{F}_m] \gtreqless \mathbb{E}[X_{n-1}|\mathcal{F}_m] \gtreqless \cdots \gtreqless X_m,$$

so the relationship expressed by martingale definitions holds over multiple steps too. We also get $\mathbb{E}[X_n] \gtreqless \mathbb{E}[X_m]$; this is usually used with $m = 0$. Since X_0 is measurable with respect to every $\mathcal{F}_i$, we can consider $Y_i = X_i - X_0$ and see that it satisfies the equivalent martingale definition of X_i, which means that we can usually assume that our martingale starts at 0.

Let us look at the above processes in a different manner. Imagine we play a game at every time $n \in \mathbb{N}$ with our net income per unit stake at time n being $X_n - X_{n-1}$. Then the martingale represents a fair game (on average we gain as much as we lose) where a submartingale is favourable to us and a supermartingale is unfavourable to us; note that all casinos operate like a supermartingale! Now imagine staking more than a unit amount, i.e., at time $n \in \mathbb{N}$ the stake is C_n; we allow the stake to be negative as well. Note that C_n can only be $\mathcal{F}_{n-1}$-measurable as our stake cannot take into account anything that happens at time n. Such a sequence of random variables $\{C_n\}_{n\in\mathbb{N}}$ with C_n being $\mathcal{F}_{n-1}$-measurable is said to be *previsible*. Now our net winnings at time $n \in \mathbb{N}$ is given by $Y_n = \sum_{i=1}^{n} C_i(X_i - X_{i-1}) =: (C \bullet X)_n$; note that $Y_n - Y_{n-1} = C_n(X_n - X_{n-1})$. By definition set $Y_0 = 0$. This is called the martingale transform of X by C and is a discrete-integral. Note that Y_n is $\mathcal{F}_n$-measurable. This yields the following result.

Lemma 3. *Let $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$ be a filtration, $\{X_n\}_{n\in\mathbb{Z}_+}$ a random process that is adapted to the filtration, and $\{C_n\}_{n\in\mathbb{N}}$ be bounded and previsible. Define $Y_0 = 0$ and $Y_n = (C \bullet X)_n$ for $n \in \mathbb{N}$. If $\{X_n\}_{n\in\mathbb{Z}_+}$ is a martingale, then $\{Y_n\}_{n\in\mathbb{Z}_+}$ is also a martingale. In addition, if $\{C_n\}_{n\in\mathbb{N}}$ is non-negative, then $\{X_n\}_{n\in\mathbb{Z}_+}$ being a submartingale (supermartingale) implies the same for $\{Y_n\}_{n\in\mathbb{Z}_+}$.*

Proof. We know that $Y_n - Y_{n-1} = C_n(X_n - X_{n-1})$. Now by property of conditional expectation and previsibility and boundedness of C, we have

$$\mathbb{E}[Y_n - Y_{n-1}|\mathcal{F}_{n-1}] = \mathbb{E}[C_n(X_n - X_{n-1})|\mathcal{F}_{n-1}] = C_n\mathbb{E}[(X_n - X_{n-1})|\mathcal{F}_{n-1}]$$

Now both results are immediate. □

We will apply this result to prove one part of Doob's optional sampling theorem. Let T be a stopping time with respect to $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$. Set $C_n^T = 1_{T\geq n}$, i.e., bet 1 until T and then stop. We have the following result from Lemma 3.

Theorem 1. *[Doob's optional sampling Theorem A] If $\{X_n\}_{n\in\mathbb{Z}_+}$ is martingale (supermartingale) with respect to filtration $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$, then the stopped process $X^T = \{X_{T\wedge n} : n \in \mathbb{Z}_+\}$*

is also a martingale (supermartingale) with respect to the same filtration. This then implies the following:

$$\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0] \;\forall n \in \mathbb{Z}_+ \text{ if } X \text{ is a supermartingale; and}$$
$$\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0] \;\forall n \in \mathbb{Z}_+ \text{ if } X \text{ is a martingale}$$

Proof. Note that $C_n^T \in \{0,1\}$ so it is bounded and non-negative. The process $\{C_n^T\}_{n\in\mathbb{N}}$ is previsible; we have $\{C_n^T = 0\} = \{T \leq n-1\} \in \mathcal{F}_{n-1}$ for all $n \in \mathbb{N}$, since T is a stopping time. The result then follows from Lemma 3. □

Note that this doesn't say anything about $\mathbb{E}[X_T]$ and relating it to $\mathbb{E}[X_0]$. We have a cautionary example here. Let $\{X_i\}_{i\in\mathbb{N}}$ be *i.i.d.* random variables that assume values $\{-1,1\}$ with equal probability. Let $\{S_n\}_{n\in\mathbb{Z}_+}$ be the sequence of partial sums, i.e., $S_0 = 0$ and $S_n = \sum_{i=1}^n X_i$ when $n \in \mathbb{N}$. Then $\{S_n\}_{n\in\mathbb{Z}_+}$ is a martingale with respect to the natural filtration of $\{X_i\}_{i\in\mathbb{N}}$ where we set $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Define $T := \inf\{n \in \mathbb{Z}_+ : S_n = 1\}$, i.e., the hitting time of state 1. It can be shown that $\mathbb{P}(T < \infty) = 1$ while $\mathbb{E}[T] = \infty$; we will revisit this when we discuss Markov chains. Now, even though $\mathbb{E}[S_{T\wedge n}] = \mathbb{E}[S_0] = 0$ for all $n \in \mathbb{N}$, we have $\mathbb{E}[S_T] = 1 \neq 0 = \mathbb{E}[S_0]$. Thus, without any additional conditions we cannot say when $\mathbb{E}[S_T] = \mathbb{E}[S_0]$ when $\{S_i\}_{i\in\mathbb{Z}_+}$ is a martingale and T a stopping time (both with the same filtration). Some sufficient conditions that allow us to make such a claim are in the following result.

Theorem 2. *[Doob's optional sampling Theorem B] Given a filtration $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$ and a stopping time T, we have the following results:*

1. *Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a supermartingale. Then X_T is integrable and $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$ under each of the following conditions:*
   - (a) *T is bounded (for some $N \in \mathbb{N}$, $T(\omega) \leq N \forall\omega$);*
   - (b) *X is bounded (for some $K \in \mathbb{R}_+$, $|X_n(\omega)| \leq K \;\forall n, \omega$), and T is* a.s. *finite ($\mathbb{P}(T < \infty) = 1$); and*
   - (c) *$\mathbb{E}[T] < \infty$, and, for some $K \in \mathbb{R}_+$, $\{X_i\}_{i\in\mathbb{Z}_+}$ has bounded increments, i.e., $|X_n(\omega) - X_{n-1}(\omega)| \leq K \;\forall\omega \& n \in \mathbb{N}$.*
2. *Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a martingale. If either of (a), (b) or (c) above hold, then $\mathbb{E}[X_T] = \mathbb{E}[X_0]$;*
3. *Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a martingale with bounded increments ($|X_n(\omega) - X_{n-1}(\omega)| \leq K_1 \;\forall\omega, n \in \mathbb{N}$). Let $\{C_n\}_{n\in\mathbb{N}}$ be a previsible process bounded by K_2 and let $\mathbb{E}[T] < \infty$. Then $\mathbb{E}[(C \bullet X)_T] = \mathbb{E}[(C \bullet X)_0] = 0$; and*
4. *Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a non-negative, supermartingale and let T be* a.s. *finite, then $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$.*

Proof. We will prove each statement in order. Let us start with (1). We have $X_{T\wedge n}$ being integrable by Theorem 1 with $\mathbb{E}[X_{T\wedge n} - X_0] \leq 0$. With assumption (i), take $n = N$ and the result follows. For assumption (i), note that T being *a.s.* finite implies that $\lim_{n\to\infty} T \wedge n = T$. Thus, $\lim_{n\to\infty} X_{T\wedge n} = X_T$ with $X_{T\wedge n}$ being bounded. Therefore, by the bounded convergence theorem we have $\lim_{n\to\infty} \mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_T] \leq \mathbb{E}[X_0]$. With assumption (iii) note the following

$$X_{T\wedge n} - X_0 = \sum_{i=1}^{T\wedge n}(X_i - X_{i-1})$$
$$\Rightarrow |X_{T\wedge n} - X_0| \leq \sum_{i=1}^{T\wedge n}|X_i - X_{i-1}| \leq K(T\wedge n) \leq KT$$

Since T is integrable and since $\lim_{n\to\infty}(X_{T\wedge n} - X_0) = X_T - X_0$ (*a.s.*), by the dominated convergence theorem we have $\lim_{n\to\infty}\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_T] \leq \mathbb{E}[X_0]$.

The result for (2) follows by applying (1) to X and to $-X$; since both are martingales, they are also supermartingales.

From Lemma 3 we already know that $Y = (C \bullet X)$ is such that $\{Y_n\}_{n\in\mathbb{Z}_+}$ and $\{Y_{T\wedge n}\}_{n\in\mathbb{Z}_+}$ are martingales and $Y_n - Y_{n-1} = C_n(X_n - X_{n-1})$, so we have

$$|Y_n - Y_{n-1}| \leq |C_n||X_n - X_{n-1}| \leq K_1K_2$$

Thus, the result (3) follows by applying (2)(iii).

For the last result, we already know that $\lim_{n\to\infty} X_{T\wedge n} = X_T$ (*a.s.*), $X_{T\wedge n} \geq 0$ (as $\{X_i\}_{i\in\mathbb{Z}_+}$ is non-negative), and $\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0]$. Now we can apply Fatou's lemma to show the following

$$\mathbb{E}[X_T] = \mathbb{E}[\liminf X_{T\wedge n}] \leq \liminf \mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0]$$

□

Theorems 1 and 2 are important results that are frequently applied.

2.2. **Doob's Forward Convergence Theorem.** Here we will state and prove a convergence theorem for martingales. This will be an extremely useful result that we will apply many times over.

Doob's Upcrossing Lemma - Let us pick two numbers $a, b \in \mathbb{R}$ with $a < b$. Let $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$ be a filtration and $\{X_i\}_{i\in\mathbb{Z}_+}$ an adapted process. Define a previsible process $\{C_i\}_{i\in\mathbb{N}}$ as follows:

$$C_1 = 1_{X_0<a}$$
$$C_n = 1_{\{C_{n-1}=1\}}1_{\{X_{n-1}\leq b\}} + 1_{\{C_{n-1}=0\}}1_{\{X_{n-1}<a\}} \quad n \geq 2$$

Again we set $Y = (C \bullet X)$. For $N \in \mathbb{N}$ define $U_N[a, b](\omega)$ to be number of upcrossings of $[a, b]$ made by $\{X_n(\omega)\}_{n\in\mathbb{Z}_+}$, which is the $k \in \mathbb{Z}_+$ such that we can find $0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_k < t_k \leq N$ with $X_{s_i}(w) < a$ and $X_{t_i}(\omega) > b$. What does a sample-path of Y look like? *Need to add Figure 11.1 from D. Williams' book here.* For $a \in \mathbb{R}$ define $a_+ = a \vee 0$ and $a_- = (-a) \vee 0 = -(a \wedge 0)$; these are positive and the negative parts of a. From the figure it is easy to see the following inequality (*Prove this*)

$$Y_N(\omega) \geq (b - a)U_N[a, b](\omega) - (X_N(\omega) - a)_-$$

We now have the lemma.

Lemma 4. *[Doob's Upcrossing Lemma] Let $\{X_i\}_{i\in\mathbb{P}}$ be a super martingale, then for all $N \in \mathbb{Z}_+$ we have*

$$(b - a)\mathbb{E}[U_N[a, b]] \leq \mathbb{E}[(X_N - a)_-]$$

Proof. By construction C is previsible, bounded and non-negative, so by Lemma 3 $\{Y_i\}_{i\in\mathbb{Z}_+}$ is a supermartingale and $\mathbb{E}[Y_N] \leq \mathbb{E}[Y_0] = 0$. Now the result follows by using the pictorial inequality above. ☐

From this we have an important corollary.

Corollary 1. *Let* $\{X_i\}_{i\in\mathbb{Z}_+}$ *be a super martingale with* $\sup_n \mathbb{E}[|X_n|] < \infty$*. For* $a, b \in \mathbb{R}$ *with* $a < b$*, define* $U_\infty[a,b] = \lim_N \uparrow U_N[a,b]$ *(sequence on right is non-decreasing). Then we have*

$$(b-a)\mathbb{E}[U_\infty[a,b]] \leq |a| + \sup_n \mathbb{E}[|X_n|] < \infty,$$

which then implies that $\mathbb{P}(U_\infty[a,b] < \infty) = 1$*, i.e.,* $U_\infty[a,b]$ *is* a.s. *finite.*

Proof. By Lemma 4 we have

$$(b-a)\mathbb{E}[U_N[a,b]] \leq \mathbb{E}[(X_N - a)_-] \leq |a| + \mathbb{E}[|X_N] \leq |a| + \sup_n \mathbb{E}[|X_n|] < \infty$$

Now we let $N \to \infty$ and use the monotone convergence theorem to assert the convergence of the expectations, i.e., $\lim_{N\to\infty} \mathbb{E}[U_N[a,b]] = \mathbb{E}[U_\infty[a,b]]$. The result. then follows. ☐

Alternate statement of Doob's Upcrossing Lemma - Sometimes the Upcrossing Lemma is also stated as follows.

Lemma 5. *[Doob's Upcrossing Lemma] If* $\{X_i\}_{i\in\mathbb{Z}_+}$ *is a submartingale, then*

$$(b-a)\mathbb{E}[U_N[a,b]] \leq \mathbb{E}[(X_N - a)_+] - \mathbb{E}[(X_0 - a)_+]$$

Proof. Let $Y_n = a + (X_n - a)_+$ for all $n \in \mathbb{Z}_+$, then $\{Y_n\}_{n\in\mathbb{Z}_+}$ is a submartingale [3]. Note that this upcrosses $[a,b]$ the same number of times as $\{X_n\}_{n\in\mathbb{Z}_+}$, so we have $(b-a)U_N[a,b] \leq (C \bullet Y)_N$ (for the same previsible process $\{C_n\}n \in \mathbb{N}$ as in proof of Lemma 4). Let $D_n = 1 - C_n$ for all $n \in \mathbb{N}$. Now $Y_n - Y_0 = (C \bullet Y)_N + (D \bullet Y)_N$ and by Lemma 3 we also have $\mathbb{E}[(D \bullet Y)_N] \leq \mathbb{E}[(D \bullet Y)_0] = 0$ so that $\mathbb{E}[(C \bullet Y)_N] \leq \mathbb{E}[Y_N - Y_0]$ which yields the result. ☐

We are now ready to state and prove an important convergence theorem.

Theorem 3. *Let* $\{X_i\}_{i\in\mathbb{Z}_+}$ *be a super martingale with* $\sup_n \mathbb{E}[|X_n|] < \infty$*, then* $X_\infty = \lim_{n\to\infty} X_n$ a.s. *exists and is finite.*

Proof. For all ω define $X_\infty(\omega) = \limsup X_n(\omega)$ (X_∞ is $\mathcal{F}_\infty$-measurable). Note that X_∞ as defined takes values in $[-\infty, \infty]$ and the result that we'd like to prove is that $X_\infty = \lim_{n\to\infty} X_n$ *a.s.*

Since we allow convergence in $[-\infty, \infty]$, the only way there is no convergence (for an ω) is if $\liminf X_n(\omega) < \limsup X_n(\omega)$. So we have

$$\begin{aligned}\Lambda &:= \{\omega : \{X_n(\omega)\}_{n\in\mathbb{Z}_+} \text{ doesn't converge to a limit in } [-\infty,\infty]\} \\ &= \{\omega : \liminf X_n(\omega) < \limsup X_n(\omega)\} \\ &= \cup_{a,b\in\mathbb{Q}:a<b} \underbrace{\{\omega : \liminf X_n(\omega) < a < b < \limsup X_n(\omega)\}}_{=:\Lambda_{a,b}}\end{aligned}$$

where $\mathbb{Q}$ is the set of rational numbers (which is countable).

[3]This is a homework problem.

Since the sequence is in the vicinity of $\liminf X_n(\omega)$ and $\limsup X_n(\omega)$ (as close as necessary) infinitely often, we have

$$\Lambda_{a,b} \subseteq \{\omega : U_\infty[a,b](\omega) = \infty\},$$

but Corollary 1 shows that $\mathbb{P}(\{\omega : U_\infty[a,b](\omega) = \infty\}) = 0$. Therefore, using the countable decomposition we get

$$\mathbb{P}(\Lambda) \leq \sum_{a,b \in \mathbb{Q}: a<b} \mathbb{P}(\Lambda_{a,b}) = 0.$$

Therefore, $X_\infty = \lim_{n\to\infty} X_n$ exists *a.s.* in $[-\infty, \infty]$.

Now use Fatou's Lemma to yield

$$\mathbb{E}[|X_\infty|] = \mathbb{E}[\liminf |X_n|] \leq \liminf \mathbb{E}[|X_n|] \leq \sup_n \mathbb{E}[|X_n|] < \infty,$$

which then implies that $\mathbb{P}(X_\infty < \infty) = 1$. □

Here is Corollary of this result that reveals its generality.

Corollary 2. *The following hold:*

(1) *If* $\{X_i\}_{i\in\mathbb{Z}_+}$ *is a submartingale/martingale with* $\sup_n \mathbb{E}[|X_n|] < \infty$*, the result holds;*
(2) *If* $\{X_i\}_{i\in\mathbb{Z}_+}$ *is submartingale/supermartingale/martingale that is* a.s. *is bounded (by* K*), then the result holds; and*
(3) *If* $\{X_i\}_{i\in\mathbb{Z}_+}$ *is a non-negative supermartingale, then the result holds.*

Proof. For (1) we prove the result by using $-X$. In (2) note that boundedness implies $\sup_n \mathbb{E}[|X_n|] \leq K < \infty$. Finally, for (3) we have $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0]$ for all $n \in \mathbb{Z}_+$ by non-negativity and the supermartingale property, and then $\sup_n \mathbb{E}[|X_n|] \leq \mathbb{E}[X_0] < \infty$. □

As with earlier examples, we should keep in mind that $\mathbb{E}[X_n]$ may not converge to $\mathbb{E}[X_\infty]$. An alternate statement of the martingale convergence theorem (Theorem 3) is the following.

Theorem 4. *If* $\{X_n\}_{n\in\mathbb{Z}_+}$ *is a submartingale with* $\sup \mathbb{E}[(X_n)_+] < \infty$*, then* $\{X_n\}_{n\in\mathbb{Z}_+}$ *converges to a limit* X a.s. *with* $\mathbb{E}[|X|] < \infty$.

The proof is the same as Theorem 3 except for the last part where one has to use Fatou's lemma twice. One also has to use the fact that $\mathbb{E}[(X_n)_-] = \mathbb{E}[(X_n)_+] - \mathbb{E}[X_n] \leq \mathbb{E}[(X_n)_+] - \mathbb{E}[X_0]$ where the last inequality follows by the submartingale property.

2.3. **Azuma-Hoeffding Inequality.** Here we will prove the Azuma-Hoeffding inequality while mentioning some uses of it. The proof will be based on Doob's submartingale inequality that we'll introduce first. For a random variable X and a set $F \in \mathcal{F}$, we define $\mathbb{E}[X; F] := \mathbb{E}[X 1_F]$ (if the expectation exists).

Theorem 5. *[Doob's Submartingale Inequality] Let* $\{X_i\}_{i\in\mathbb{Z}_+}$ *be a non-negative submartingale (with respect to some filtration* $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$*). Then, for* $c > 0$*, we have*

$$c\mathbb{P}\left(\sup_{0\leq k\leq n} X_k \geq c\right) \leq \mathbb{E}\left[X_n; \left\{\sup_{0\leq k\leq n} X_k \geq c\right\}\right] \leq \mathbb{E}[X_n], \quad \forall n \in \mathbb{Z}_+$$

Proof. Fix an $n \in \mathbb{Z}_+$ and let $F = \{\sup_{0\leq k\leq n} X_k \geq c\}$. Then F is a disjoint union ($F = \sqcup_{0\leq k\leq n} F_i$) of the following

$$\begin{aligned} F_0 &= \{X_0 \geq c\} \\ F_k &= (\cap_{0\leq i\leq k-1}\{X_i < c\}) \cap \{X_k \geq c\}, \quad 1 \leq k \leq n \text{ with } n \in \mathbb{N} \end{aligned}$$

It is easy to see that each $F_k \in \mathcal{F}_k$, and $X_k \geq c$ on F_k. Therefore,

$$\mathbb{E}[Z_n; F_k] \geq \mathbb{E}[Z_k; F_k] \geq c\mathbb{P}(F_k),$$

where we used the submartingale property in the first inequality and non-negativity in the second. Summing over $k = 0, 1, \dots, n$ yields,

$$\mathbb{E}[Z_n; F] \geq c\mathbb{P}(F),$$

and since $\mathbb{E}[Z_n; F] \leq \mathbb{E}[Z_n]$, the result follows. □

The main application of this result follows from the next Lemma.

Lemma 6. *Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a martingale (with respect to some filtration $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$) and let $c : \mathbb{R} \mapsto \mathbb{R}$ be a convex function with $\mathbb{E}[|c(X_n)|] < \infty$ for all $n \in \mathbb{Z}_+$, then $\{c(X_i)\}_{i\in\mathbb{Z}_+}$ is a submartingale.*

Proof. Apply the conditional form of Jensen's inequality. □

We can now state and prove the Azuma-Hoeffding inequality.

Theorem 6. *[Azuma-Hoeffding Inequality] Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a martingale null at 0 (with respect to some filtration $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$) such that for some sequence $\{c_n\}_{n\in\mathbb{N}}$ of positive constants, we have*

$$|X_n - X_{n-1}| \leq c_n, \quad \forall n \in \mathbb{N},$$

then, for $x > 0$,

$$\begin{aligned} \mathbb{P}\left(\sup_{0\leq k\leq n} X_k \geq x\right) &\leq \exp\left(-\frac{1}{2}\frac{x^2}{\sum_{k=1}^n c_k^2}\right) \\ \mathbb{P}(|X_n| \geq x) &\leq 2\exp\left(-\frac{1}{2}\frac{x^2}{\sum_{k=1}^n c_k^2}\right) \end{aligned}$$

Before proving this theorem we prove a simple lemma.

Lemma 7. *Let X be a random variable with $a \leq X \leq b$, then for any $\theta \in \mathbb{R}$,*

$$\log \mathbb{E}[e^{\theta X}] \leq \theta\mathbb{E}[X] + \frac{\theta^2(b-a)^2}{8}$$

Proof. First note that $\log \mathbb{E}[e^{\theta X}] = s\mathbb{E}[X] + \log \mathbb{E}[e^{\theta(X-\mathbb{E}[X])}]$, so it suffices to show that with $\mathbb{E}[X] = 0$ and $a \leq X \leq b$ (where $a \leq 0$ and $b \geq 0$), we have

$$\log \mathbb{E}[e^{\theta X}] \leq \frac{\theta^2(b-a)^2}{8}$$

Since $e^{\theta x}$ is a convex function, by definition we have

$$e^{\theta x} \leq \frac{x-a}{b-a}e^{\theta b} + \frac{b-x}{b-a}e^{\theta a}, \quad \forall x \in [a, b]$$

Now taking expectations and using $\mathbb{E}[X] = 0$, we get

$$\mathbb{E}[e^{\theta x}] \leq \frac{b}{b-a}e^{\theta a} - \frac{a}{b-a}e^{\theta b}$$

Now $f(\theta) = \log\left(\frac{b}{b-a}e^{\theta a} - \frac{a}{b-a}e^{\theta b}\right)$ is a convex function of $\theta \in \mathbb{R}$ (see non-negativity of second derivative below). Note that $f(0) = 0$. We get the following

$$\frac{df}{d\theta}(\theta) = \frac{-ab(e^{\theta b} - e^{\theta a})}{be^{\theta a} - ae^{\theta b}}$$
$$\frac{d^2 f}{d\theta^2}(\theta) = \frac{-ae^{\theta b}}{be^{\theta a} - ae^{\theta b}} \frac{be^{\theta a}}{be^{\theta a} - ae^{\theta b}}(b-a)^2$$

Note that $\frac{df}{d\theta}(0) = 0$ and $\frac{d^2 f}{d\theta^2}(\theta) = p(1-p)(b-a)^2$ where $p = \frac{be^{\theta a}}{be^{\theta a} - ae^{\theta b}} \in [0,1]$. Thus, $\frac{d^2 f}{d\theta^2}(\theta) \leq \frac{(b-a)^2}{4}$. It then follows using Taylor's theorem that $f(\theta) \leq \frac{\theta^2(b-a)^2}{8}$, which using the monotonicity of $\exp(\cdot)$ gives our result. □

We should note that in the intermediate step we are bounding the moment generating function of X by a random variable that only takes values a and b but has the same expectation. We will revisit this in the exercises.

Now we prove the Azuma-Hoeffding inequality.

Proof of Theorem 6. Since $e^{\theta x}$ is a convex function of x for all $\theta \in \mathbb{R}$, using Lemma 6 we get that $\{e^{\theta X_i}\}_{i \in \mathbb{N}}$ is a submartingale; note that it is also non-negative. By the monotonicity of $\exp(\cdot)$ we also know the following for $\theta \in \mathbb{R}_+$,

$$\mathbb{P}\left(\sup_{0 \leq k \leq n} X_k \geq x\right) = \mathbb{P}\left(\sup_{0 \leq k \leq n} e^{\theta X_k} \geq e^{\theta x}\right)$$

Now by Doob's Submartingale inequality (Theorem 5), we get that

$$\mathbb{P}\left(\sup_{0 \leq k \leq n} X_k \geq x\right) \leq e^{-\theta x}\mathbb{E}[e^{\theta X_n}] \quad \forall \theta \in \mathbb{R}_+$$

Consider $\{e^{\theta X_n}\}_{i \in \mathbb{N}}$, we have by the null 0 and bounded increments property that

$$X_n = \sum_{i=1}^{n} X_i - X_{i-1}, \text{ and}$$
$$|X_n| \leq \sum_{i=1}^{n} |X_i - X_{i-1}| \leq \sum_{i=1}^{n} c_i < \infty \quad \forall n \in \mathbb{N},$$

which also implies that $e^{\theta X_n} \in [\exp(-\theta \sum_{i=1}^n c_i), \exp(\theta \sum_{i=1}^n c_i)]$ for all $n \in \mathbb{N}$. Now using adaptedness, boundedness and Lemma 7, we get

$$\begin{aligned}
\mathbb{E}[e^{\theta X_n}] &= \mathbb{E}\left[e^{\theta(X_n - X_{n-1} + X_{n-1})}\right] \\
&= \mathbb{E}\left[e^{\theta X_{n-1}}\mathbb{E}\left[e^{\theta(X_n - X_{n-1})}|\mathcal{F}_{n-1}\right]\right] \\
&\leq \mathbb{E}\left[e^{\theta X_{n-1}}e^{\frac{\theta^2 c_n^2}{2}}\right] = e^{\frac{\theta^2 c_n^2}{2}}\mathbb{E}[e^{\theta X_{n-1}}] \\
&\leq e^{\frac{\theta^2 \sum_{i=1}^n c_n^2}{2}}
\end{aligned}$$

Thus, we have

$$\mathbb{P}\left(\sup_{0\leq k\leq n} X_k \geq x\right) \leq e^{-\theta x} e^{\frac{\theta^2 \sum_{i=1}^n c_n^2}{2}}$$

Since the right side holds for all $\theta \in \mathbb{R}_+$, minimizing over θ yields our result; by elementary calculus, the minimizer is $\theta^* = \frac{x}{\sum_{i=1}^n c_n^2}$. Since $\{X_n \geq x\} \subseteq \{\sup_{0\leq k\leq n} X_k \geq x\}$, the second result also follows; the factor 2 comes from the same bound on $-X_n$ (which is also a martingale). □

Sometimes we may need to use an extension of Theorem 6, which we present without proof.

Theorem 7. *[Bernstein's inequality for martingales] Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a martingale null at 0 (with respect to some filtration $\{\mathcal{F}_i\}_{i\in\mathbb{Z}_+}$) such that for some $K > 0$, we have*

$$|X_n - X_{n-1}| \leq K, \quad \forall n \in \mathbb{N},$$

Denote the sum of conditional variances by

$$\Sigma_n^2 = \sum_{i=1}^n \mathbb{E}[(X_i - X_{i-1})^2 | \mathcal{F}_{i-1}]$$

Then for all constants $t, s > 0$, we have

$$\mathbb{P}\left(\sup_{0\leq k\leq n} X_k > t \text{ and } \Sigma_n^2 \leq s\right) \leq \exp\left(-\frac{t^2}{2(s + Kt/3)}\right)$$

2.4. **Exercises.** Please show all your work.

(1) *Stopping times*: Prove the three properties listed about stopping times.

(2) *Martingales*: Suppose $\{X_n\}_{n\in\mathbb{Z}_+}$ is a martingale with respect to filtration $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$, show that $\{X_n\}_{n\in\mathbb{Z}_+}$ is also a martingale with respect to its natural filtration. *Hint*: Use the definition of conditional expectation and the Tower property.

(3) *Martingales and Jensen*: Prove the following: If $\{X_n\}_{n\in\mathbb{Z}_+}$ is a submartingale and $\phi(\cdot)$ an increasing convex function with $\mathbb{E}[|\phi(X_n)|] < \infty$ for all $n \in \mathbb{N}$, then $\{\phi(X_n)\}_{n\in\mathbb{N}}$ is a submartingale. *Hint*: Use the conditional form of Jensen's inequality. Use this result to prove these two results. First, if $\{X_n\}_{n\in\mathbb{Z}_+}$ is a submartingale, then so is $\{(X_n - a)_+\}_{n\in\mathbb{Z}_+}$ for any $a \in \mathbb{R}$. Second, if $\{X_n\}_{n\in\mathbb{Z}_+}$ is a supermartingale, then so is $\{X_n \wedge a\}$ for any $a \in \mathbb{R}$. Generalize this to show the following: if $\{X_n\}_{n\in\mathbb{Z}_+}$ and $\{Y_n\}_{n\in\mathbb{Z}_+}$ are submartingales (on the same filtration), then so is $\{X_n \vee Y_n\}_{n\in\mathbb{Z}_+}$. (Last part is **Extra credit**.)

(4) *Martingales*: Let $\{X_n\}_{n\in\mathbb{N}}$ be an independent sequence with $\mathbb{E}[X_n] = 0$ and $\mathrm{var}(X_n) = \sigma_n^2$ for all $n \in \mathbb{N}$; define $S_n = \sum_{m=1}^n X_m$, $s_n^2 = \sum_{m=1}^n \sigma_m^2$ (with $S_0 = s_0 = 0$). Then $\{S_n^2 - s_n^2\}_{n\in\mathbb{Z}_+}$ is a martingale.

(5) *Martingales and stopping times*: Let S and T be stopping times with $S \leq T$. Define the process $\{1_{(S,T]}(n)\}_{n\in\mathbb{N}}$ by

$$1_{(S,T]}(n,\omega) = \begin{cases} 1 & \text{if } S(\omega) < n \leq T(\omega) \\ 0 & \text{otherwise} \end{cases}$$

Prove that $\{1_{(S,T]}(n)\}_{n\in\mathbb{N}}$ is previsible, and deduce that if $\{X_n\}_{n\in\mathbb{Z}_+}$ is a supermartingale, then $\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_{S\wedge n}]$ for all $n \in \mathbb{Z}_+$. *(Assume a common filtration.)*

(6) *Martingales and the switching principle*: Suppose $\{X_n^1\}_{n\in\mathbb{Z}_+}$ and $\{X_n^2\}_{n\in\mathbb{Z}_+}$ be two supermartingales (on the same filtration), and N is a stopping time (with same filtration) such that $X_N^1 \geq X_N^2$. Then show the following

$$Y_n = X_n^1 1_{\{N>n\}} + X_n^2 1_{\{N\leq n\}} \quad n \in \mathbb{Z}_+$$
$$Z_n = X_n^1 1_{\{N\geq n\}} + X_n^2 1_{\{N<n\}} \quad n \in \mathbb{Z}_+$$

are supermartingales.

(7) *Martingale convergence*: Consider the following process: $X_0 = 0$, when $X_{k-1} = 0$, let

$$X_k = \begin{cases} 1 & \text{w.p. } \frac{1}{2k} \\ -1 & \text{w.p. } \frac{1}{2k} \\ 0 & \text{w.p. } 1 - \frac{1}{k} \end{cases},$$

and when $X_{k=1} \neq 0$, let $X_k = kX_{k-1}$ with probability $\frac{1}{k}$ and 0 otherwise. Show that $\lim_{n\to\infty} X_n \stackrel{p}{=} 0$. Now use the second Borel-Cantelli lemma to show that $\{X_n\}_{n\in\mathbb{Z}_+}$ cannot converge to 0 in the *a.s.* sense. Simulate this process.

(8) *Martingale convergence*: Give an example of a martingale $\{S_n\}_{n\in\mathbb{Z}_+}$ with $\lim_{n\to\infty} S_n \stackrel{a.s.}{=} -\infty$. *Hint*: Let $S_n = X_1 + \cdots + X_n$ for independent but not identically distributed $\{X_n\}_{n\in\mathbb{N}}$ with $\mathbb{E}[X_n] = 0$ for all $n \in \mathbb{N}$. Simulate your example. [**Extra credit**]

(9) *Martingale convergence*: Let $\{X_n\}_{n\in\mathbb{Z}_+}$ and $\{Y_n\}_{n\in\mathbb{Z}_+}$ be positive integrable and adapted to filtration $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$. Suppose

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq (1 + Y_n)X_n$$

with $\sum_{n\in\mathbb{Z}_+} Y_n < \infty$ *a.s.*. Prove that $\{X_n\}_{n\in\mathbb{Z}_+}$ converges *a.s.* to a finite limit. *Hint*: Find a closely related supermartingale and use the convergence theorem.

(10) *Optional sampling*: A monkey breaks into an office and takes a liking to the computer and keyboard on the desk. It starts typing capital letters $\{A, B, \ldots, Z\}$ picking them at random. How long will it take on the average for the monkey to type $TATTATTAT$ (consecutive sequence)? *Hint*: Just before each time $n \in \mathbb{N}$, a new gambler arrives and bets \$1 that the n^{th} letter will be T. In case this doesn't happen (a loss), the gambler leaves. If, instead, the event happens (a win), the gambler receives 26 times what the bet was, i.e., \$26; all of the winnings are then bet on the event that the $(n+1)^{th}$ letter will be A. Again, if it is a loss, then the gambler leaves (empty-handed), but if it is a win, the gambler gets \$$26^2$ all of which is then bet on the event that the $(n+2)^{th}$ letter will be T. This procedure repeats through the $TATTATTAT$ sequence. Let $\mathcal{T}$ be the first time by which the monkey has produced the consecutive sequence $TATTATTAT$. Show that $\mathcal{T}$ is a stopping time that is *a.s.* finite. Then using Doob's optional sampling theorem show why $\mathbb{E}[\mathcal{T}] = 26^9 + 26^6 + 26^3 + 26$. What is the answer if the monkey had to produce $ABRACADABRA$ instead? [**Extra credit**]

(11) *Submartingale inequality*: Prove Kolmogorov's inequality, i.e., let $\{X_n\}_{n\in\mathbb{N}}$ be an independent sequence with $\mathbb{E}[X_n] = 0$ and $\mathbb{E}[X_n^2] = \sigma_n^2 < \infty$ and let $S_n = \sum_{n=1}^{m} X_n$ with $S_0 = 0$. Show that

$$\mathbb{P}\left(\sup_{0\le m\le n} |S_m| \ge x\right) \le \frac{\mathrm{var}(S_n)}{x^2}$$

Hint: Note that $\{S_n\}_{n\in\mathbb{Z}_+}$ is a martingale, so define an appropriate submartingale using Jensen's inequality and apply Doob's inequality.

(12) *Submartingale inequality*: Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be a martingale with $X_0 = 0$ and $\mathbb{E}[X_n^2] < \infty$ for all $n \in \mathbb{Z}_+$. Show that

$$\mathbb{P}\left(\sup_{0\le m\le n} X_m \ge x\right) \le \frac{\mathbb{E}[X_n^2]}{\mathbb{E}[X_n^2] + x^2}$$

Hint: Use the fact that $\{(X_n + c)^2\}_{n\in\mathbb{Z}_+}$ is a submartingale and optimize over c.

(13) *Example from class*: Let $\{X_i\}_{i\in\mathbb{N}}$ be *i.i.d.* random variables that assume values $\{-1, 1\}$ with equal probability. Let $\{S_n\}_{n\in\mathbb{Z}_+}$ be the sequence of partial sums, i.e., $S_0 = 0$ and $S_n = \sum_{i=1}^{n} X_i$ when $n \in \mathbb{N}$. Then $\{S_n\}_{n\in\mathbb{Z}_+}$ is a martingale with respect to the natural filtration of $\{X_i\}_{i\in\mathbb{N}}$ where we set $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Define $T := \inf\{n \in \mathbb{Z}_+ : S_n = 1\}$, i.e., the hitting time of state 1. For $\theta \in \mathbb{R}$, $\mathbb{E}[e^{\theta X}] = \cosh(\theta)$ so that $\mathbb{E}[e^{\theta X}/\cosh(\theta)] = 1$ for all $n \in \mathbb{N}$. Using this show that $M_n^\theta = (\cosh(\theta))^{-n} e^{\theta S_n}$ forms a martingale. Now consider $\theta > 0$, use the optional sampling theorem and the bounded convergence theorem to show that $\mathbb{E}[M_T^\theta] = 1 = \mathbb{E}[(\cosh(\theta))^{-T} e^{\theta}]$. Now argue using either the monotone convergence theorem or bounded convergence theorem that $\mathbb{P}(T < \infty) = 1$. Using the quantities derived previously find out the probability generating function of T. Using the probability generating function prove that $\mathbb{E}[T] = +\infty$. [**Extra credit**] We will revisit the same problem when studying Markov chains.

(14) *Hoeffding bound versus large deviations bound*: Let $\{X_i\}_{i\in\mathbb{N}}$ be independent such that $|X_i| \le a$ and $\mathbb{E}[X_i] = 0$ for all $i \in \mathbb{N}$. Let $\{S_n\}_{n\in\mathbb{Z}_+}$ be the partial sums process which is a martingale. Thus, we can get a bound for $\mathbb{P}(S_n \ge x)$ using the Azuma-Hoeffding

inequality [4]. Note that as a part of the proof we showed that $\mathbb{E}[e^{\theta X_i}] \leq \cosh(\theta)$ which is the moment generating function of a random variable Y that assumes values $-a$ and a with equal probability. Use this idea and Cramer's Theorem/Chernoff bound to find an alternate bound for $\mathbb{P}(S_n \geq x)$. Compare the two bounds: is one better than the other, and if so, prove it.

[4] This is called the Hoeffding bound.

3. Discrete-time Markov Chains

Sources: The sources for this chapter are

- J. R. Norris, "Markov chains," Cambridge Series in Statistical and Probabilistic Mathematics, 2, Cambridge University Press, Cambridge, 2006.
- P. Brémaud, "Markov chains: Gibbs fields, Monte Carlo simulation, and queues," Texts in Applied Mathematics, 31, Springer-Verlag, New York, 1999.
- R. Durrett, "Probability: Theory and examples," Second edition. Duxbury Press, Belmont, CA, 1996.

3.1. **Markov chains.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space [5] and assume that $(I, \mathcal{I})$ is a measurable space where I is a finite or countable set and $\mathcal{I} = 2^I$; often $I = \mathbb{N}/\mathbb{Z}_+/\mathbb{Z}$ (or subset) but we can have multi-dimensional and symbolic values as well. An element $i \in I$ is called a state and I is called the state-space. A matrix/operator $P = (p_{ij} : i, j \in I)$ is called stochastic if $P_{ij} \geq 0$ and $\sum_{j\in I} P_{ij} = 1$ for all $i \in I$, i.e., $(P_{ij} : j \in I)$ is a distribution on I. One can also represent P by a figure. For the most part we will only discuss time-invariant Markov chains (i.e., with P not a function of time), and only note that many of the definitions (not all though) can also be carried over to the time-varying case.

A random process $\{X_i\}_{i\in\mathbb{Z}_+}$ taking values in I is a discrete-time Markov chain with initial distribution μ and transition matrix P, (μ, P) Markov, if

(1) X_0 has distribution μ; and
(2) For $n \geq 0$, conditional on $X_n = i$, X_{n+1} has distribution $(P_{ij} : j \in I)$.

Equivalently, $\{X_i\}_{i\mathbb{Z}_+}$ is a Markov chain with initial distribution μ and transition matrix P if for all $\{i_0, i_1, \ldots, i_N\}$ $(N \in \mathbb{Z}_+)$ we have

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_N = i_N) = \lambda_{ij} P_{i_0 i_1} P_{i_1 i_2} \cdots P_{i_{N-1} i_N}$$

Define λP and P^n by matrix mulitplcation [6]. Denote by $P^{(n)}_{ij} = \mathbb{P}(X_n = j|X_0 = i)$; sometimes we will also use $\mathbb{P}_i(X_n = j)$ to denote the same [7]. The Markov property also yield the follow (semi-group) property that is sometimes called the Chapman-Kolmogorov equation,

$$P^{(n+m)}_{ij} = \sum_{k\in I} P^{(n)}_{ik} P^{(m)}_{kj} = \sum_{k\in I} P^{(m)}_{ik} P^{(n)}_{kj} \quad \forall n, m \in \mathbb{Z}_+$$

We now state a few properties:

(1) Conditional on $X_m = i$ (for some $m \in \mathbb{Z}_+$) we have that $\{X_{m+n}\}_{n\in\mathbb{Z}_+}$ is (δ_i, P) Markov and is independent of the random variables $(X_0, X_1, \ldots, X_m)$. Note that by the conditioning X_m is a constant, and hence, independent of everything.
(2) We have $\mathbb{P}(X_n = j) = (\mu P^n)_j$ and $\mathbb{P}_i(X_n = j) = \mathbb{P}(X_{n+m} = j|X_m = j) = (P^n)_{ij} = P^{(n)}_{ij}$; the latter is the n-step transition matrix.
(3) Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a (μ, P) Markov chain and let $f : I \mapsto I'$ be a measurable map, then $\{f(X_i)\}_{i\in\mathbb{Z}_+}$ need not be a Markov chain. As a consequence, if we lump states together and relabel, then we need not produce a Markov chain.

[5] In the homework, we will show how it is sufficient for our purposes to consider our typical example.
[6] Homework will be to show that these exist even in the countable case.
[7] Note that this is same as assuming that the initial distribution μ equals δ_i, the Dirac measure at i.

We say that state i leads to state j (denoted by $i \to j$) if $\mathbb{P}_i(X_n = j \text{ for some } j) > 0$. We say that state i communicates with state j (denoted by $i \leftrightarrow j$) if $i \to j$ and $j \to i$. The following properties hold:

(1) For distinct states i and j, either of these statements are equivalent: $i \to j$; $P_{ij}^{(n)} > 0$ for some $n \geq 0$; and $P_{i_o i_1} P_{i_1 i_2} \cdots P_{i_{n-1} i_n} > 0$ for some finite sequence of states $(i_0, i_1, \ldots, i_n)$ with $i_0 = i$ and $i_n = j$.

The property of communication ($\leftrightarrow$) is an equivalence relationship that partitions I into communicating classes where states communicate with each other. Such a class C is said to be closed if for all $i \in C$, if $i \to j$ for some $j \in I$, then $j \in C$, i.e., loosely put there is no way to escape C. A state i is said to be absorbing if $\{i\}$ is a closed class. A transition matrix P is said to be irreducible if I is the only communicating class, i.e., for all $i, j \in I$ there exists an $n \geq 0$ such that $P_{ij}^{(n)} > 0$.

3.2. **Hitting times.** Given $A \subset I$ we define $H^A = \inf\{n \in \mathbb{Z}_+ : X_n \in A\}$ (with the value being $+\infty$ if $X_n \notin A$ for all $n \in \mathbb{Z}_+$) to be the time that the Markov chain $\{X_n\}_{n \in \mathbb{Z}_+}$ hits the set A; note that this is a stopping time (what is the filtration?). The probability that starting in state i and hitting A in finite time is defined to be

$$h_i^A = \mathbb{P}(H^A < +\infty)$$

We denote the mean amount of time to hit set A when starting from i to be $k_i^A = \mathbb{E}[H_i^A]$; note that k_i^A is infinite if $h_i^A < 1$ and can be infinite even if $h_i^A = 1$. The main task we will consider now is to derive recursive relationships for h_i^A and k_i^A.

Theorem 8. *The sequence of hitting probabilities $\{h_i^A\}_{i \in I}$ is the minimal non-negative solution to the following system of linear equations*

$$h_i^A = \begin{cases} 1 & \text{if } i \in A; \\ \sum_{j \in I} P_{ij} h_j^A & \text{otherwise} \end{cases}$$

Proof. If $X_0 = i \in A$, then $H^A = 0$ and the first part of the result follows. If $X_0 = i \notin A$, then it follows that $H^A \geq 1$ so that we can consider reaching some state at time 1 and then repeat the assessment. By the Markov property we get

$$\mathbb{P}_i(H^A < \infty | X_1 = j) = \mathbb{P}_j(H^A < \infty) = h_j^A,$$

and therefore we get

$$\begin{aligned} \mathbb{P}_i(H^A < \infty) &= \sum_{j \in I} \mathbb{P}_i(H^A < \infty, X_1 = j) \\ &= \sum_{j \in I} \mathbb{P}_i(H^A < \infty | X_1 = j) \mathbb{P}_i(X_1 = j) = \sum_{j \in I} P_{ij} h_j^A \end{aligned}$$

For minimality, see the book by Norris. □

Theorem 9. *The sequence of mean hitting times $\{k_i^A\}_{i \in I}$ is the minimal non-negative solution to the following system of linear equations*

$$k_i^A = \begin{cases} 0 & \text{if } i \in A; \\ 1 + \sum_{j \notin A} P_{ij} k_j^A & \text{otherwise} \end{cases}$$

Proof. The proof proceeds just the same as for Theorem 8. If $X_0 = i \in A$, then $H^A = 0$ and $k_i^A = 0$. If $X_0 = i \notin A$, then $H^A \geq 1$ and by the Markov property we now have

$$\mathbb{E}_i[H^A|X_1 = j] = 1 + \mathbb{E}_j[H^A] = 1 + k_j^A$$

when the state at time 1 is j, and so we get

$$k_i^A = \sum_{j\in I} \mathbb{E}_i(H^A|X_1 = j)\mathbb{P}_i(X_1 = j) = 1 + \sum_{j\notin A} P_{ij}k_j^A$$

For minimality, see the book by Norris. □

This argument can be extended to the probability generating function, and hence the distribution as well; we will have to condition on the event $\{H^A < +\infty\}$.

Let us apply this to the Optional Sampling problem: $\{X_i\}_{i\in\mathbb{N}}$ is *i.i.d.* taking values $\{-1, +1\}$ with equal probability and $\{S_i\}_{i\in\mathbb{Z}_+}$ is the partial sums process which we analyzed using martingale theory; in fact, the partial sums process is also a Markov chain with $\mu = \delta_0$ and for $i \in \mathbb{Z}$, $P_{i,i+1} = P_{i,i-1} = 1/2$. Let A be $\{1\}$; note that the skip-free nature of the process allows us to take $A = \{i \geq 1\}$ as well, and we will do so. Let us find h_0^A and k_0^A. The equations are:

$$h_i^A = \begin{cases} 1 & i \geq 1 \\ \frac{h_{-1}^A+1}{2} & i = 0 \\ \frac{h_{i-1}^A+h_i^A}{2} & i < 0 \end{cases}, \qquad k_i^A = \begin{cases} 0 & i \geq 1 \\ 1 + \frac{k_{-1}^A}{2} & i = 0 \\ 1 + \frac{k_{i-1}^A+k_{i+1}^A}{2} & i < 0 \end{cases}$$

It is easy to argue that $h_i^A \equiv 1$ as one can reach A from any state $i \leq 0$ in finite time by a sequence of +1s; note that this is also the solution given by the equations. Similarly, we can iterate and find that $k_{-i}^A = (1 + i)(k_0^A - i)$ for all $i < 0$, and since $k_{-i}^A \geq k_0 \geq 1$, it follows that the only non-negative solution is $k_i^A = +\infty$ for $i \leq 0$. Note that we have established our required result in a much simpler manner.

3.2.1. *Strong Markov property.* By discussion earlier, we know that conditioning on $X_m = i$ we have $\{X_{m+n}\}_{n\in\mathbb{Z}_+}$ being a (δ_i, P) Markov chain along with independence with respect to the past. What happens when we condition on $X_T = i$ where T a random time? Do we still get a (δ_i, P) Markov chain? Obviously it doesn't hold for an arbitrary random time: let us say that the property hold for T, then it cannot hold for $T - 1$ as we will be jumping to i at T irrespective of P. We have the following theorem.

Theorem 10. *Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be a (μ, P) Markov chain and let T be a stopping time associated with $\{X_n\}_{n\in\mathbb{Z}_+}$, i.e., T is stopping time with respect to the natural filtration of $\{X_n\}_{n\in\mathbb{Z}_+}$. Then, conditioned on $T < \infty$ and $X_T = i$, $\{X_{T+n}\}_{n\in\mathbb{Z}_+}$ is a (δ_i, P) Markov chain that is independent of $(X_0, X_1, \ldots, X_T)$.*

Proof. If $B \in \sigma(X_0, X_1, \ldots, X_T)$, then $B \cap \{T = m\} \in \sigma(X_0, X_1, \ldots, X_m)$ for $m \in \mathbb{Z}_+$. Now by the Markov property we get

$$\begin{aligned}\mathbb{P}(\{X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_T = i\}) \\ = \mathbb{P}_i(\{X_0 = j_0, X_1 = j_1, \ldots, X_n = j_n\})\mathbb{P}(B \cap \{T = m\} \cap \{X_T = i\})\end{aligned}$$

Now summing over $m \in \mathbb{Z}_+$ and dividing by $\mathbb{P}(T < \infty, X_T = i)$ (to get the conditional probability), we get

$$\begin{aligned}\mathbb{P}(\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B | \{T < \infty, X_T = i\}) \\ = \mathbb{P}_i(\{X_0 = j_0, X_1 = j_1, \dots, X_n = j_n\})\mathbb{P}(B | \{T < \infty, X_T = i\})\end{aligned}$$

□

3.3. **Recurrence and Transience.** State i is said to be recurrent if

$$\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 1.$$

State i is said to be transient if

$$\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 0.$$

The Kolmogorov 0-1 law ensures that there is no other possibility.

For a fixed state i let define a sequence of stopping times $\{T_j^i\}_{j \in \mathbb{Z}_+}$ as follows:

$$T_j^i(\omega) = \begin{cases} 0 & \text{if } j = 0 \\ \inf\{n > T_{j-1}^i(\omega) : X_n = i\} & \text{otherwise} \end{cases}$$

Using these we can define the sequence of excursion times $\{S_j^i\}_{j \in \mathbb{N}}$ by

$$S_j^i(\omega) = \left(T_j^i(\omega) - T_{j-1}^i(\omega)\right) 1_{\{T_{j-1}^i(\omega) < \infty\}}$$

Using the Strong Markov property it can be argued that conditional on $T_{j-1}^i < \infty$, S_j^i has the same distribution as T_1^j, the first passage time to j which we also denote as T_j.

Let $1_{\{X_n = j\}}$ be the random variable that is 1 if $X_n = j$ and 0 otherwise. Then the total number of visits to i is given by $V_i = \sum_{n=0}^{\infty} 1_{\{X_n = i\}}$. Note that $\mathbb{E}_i[V_i] = \sum_{n=0}^{\infty} \mathbb{P}_i(X_n = i) = \sum_{n=0}^{\infty} P_{ii}^{(n)}$. Let $f_i := \mathbb{P}_i(T_i < \infty)$, then $\mathbb{P}_i(V_i > r) = f_i^r$. This then implies the following result that we present without proof.

Theorem 11. *One of the following two statements occurs:*

1. *if $\mathbb{P}_i(T_i < \infty) = 1$, then state i is recurrent and $\mathbb{E}_i[V_i] = \sum_{n=0}^{\infty} P_{ii}^{(n)} = \infty$;*
2. *if $\mathbb{P}_i(T_i < \infty) < 1$, then state i is transient and $\mathbb{E}_i[V_i] = \sum_{n=0}^{\infty} P_{ii}^{(n)} < \infty$*

In other words, a state is either recurrent or transient.

We then have an important property: transience or recurrence is a property of a communicating class, i.e., either all states are recurrent or they are transient. Additionally, we have the following: every recurrent class is closed and every finite closed class is recurrent. The proofs use Theorem 11. Finally, if P is irreducible and recurrent, then $\mathbb{P}(T_i < \infty) = 1$ for all $i \in I$.

3.4. **Invariant distributions, Positive recurrence and Ergodic theorem.** A measure [8] π is said to be invariant if $\pi P = \pi$. If π is the invariant distribution, i.e., invariant and probability distribution, of a Markov chain with transition matrix P, then given that $\{X_i\}_{i \in \mathbb{Z}_+}$ is (π, P) Markov, then so is $\{X_{i+m}\}_{i \in \mathbb{Z}_+}$ for every $m \in \mathbb{Z}_+$, i.e., starting a Markov chain with the invariant distribution yields a stationary random process. Often, the invariant distribution will also be deemed the equilibrium distribution. It is easy to see that for I finite, if $\lim_{n \to \infty} P_{ij}^{(n)} = \pi_j$, then π is an invariant distribution.

[8] In this case a sequence of non-negative numbers, the sum of which can be finite or infinite.

We will now find an invariant measure for an irreducible and recurrent transition matrix P. Fix a state k and for every state $i \in I$ define γ_i^k to be the expected time spent in state i between visits to state k, i.e.,

$$\gamma_i^k = \mathbb{E}_k \left[\sum_{n=0}^{T_k - 1} 1_{\{X_n = i\}} \right]$$

We have the following result.

Theorem 12. *Let P be irreducible and recurrent, then*

(1) $\gamma_k^k = 1$;
(2) $\gamma^k = (\gamma_i^k : i \in I)$ *satisfies* $\gamma^k = \gamma^k P$; *and*
(3) $0 < \gamma_i^k < \infty$ *for all* $i \in I$.

Proof. The first part is obvious. For the second part we first use the fact that T_k is a stopping time, and so the event $\{T_k \geq n\}$ is contained in $\sigma(X_0, X_1, \ldots, X_{n-1})$. This and the Markov property then imply that

$$\mathbb{P}_k(X_{n-1} = i, X_n = j \text{ and } T_k \geq n) = \mathbb{P}_k(X_{n-1} = i, T_k \geq n) P_{ij}.$$

Since P is recurrent, under $\mathbb{P}_k$ we have $T_k < \infty$ and $X_0 = X_{T_k} = k$ *a.s.* so that

$$\begin{aligned}
\gamma_i^k &= \mathbb{E}_k \left[\sum_{n=1}^{T_k} 1_{\{X_n = i\}} \right] = \mathbb{E}_k \left[\sum_{n=1}^{\infty} 1_{\{X_n = i, T_k \geq n\}} \right] \\
&= \sum_{n=1}^{\infty} \mathbb{P}_k(X_n = i, T_k \geq n) \\
&= \sum_{n=1}^{\infty} \sum_{j \in I} \mathbb{P}_k(X_{n-1} = j, X_n = i, T_k \geq n) \\
&= \sum_{j \in I} \sum_{n=1}^{\infty} \mathbb{P}_k(X_{n-1} = j, X_n = i, T_k \geq n) \\
&= \sum_{j \in I} \sum_{n=1}^{\infty} \mathbb{P}_k(X_{n-1} = j, T_k \geq n) P_{ji} \\
&= \sum_{j \in I} P_{ji} \sum_{n=0}^{\infty} \mathbb{P}_k(X_n = j, T_k - 1 \geq n) \\
&= \sum_{j \in I} P_{ji} \mathbb{E}_k \left[\sum_{n=0}^{T_k - 1} 1_{\{X_n = j\}} \right] = \sum_{j \in I} \gamma_j^k P_{ji}.
\end{aligned}$$

Since P is irreducible, there is an $n, m \geq 0$ such that $P_{ik}^{(n)} > 0$ and $P_{ki}^{(m)}$ for every $i \in I$. Using the recurrence above, we have $\gamma_i^k \geq \gamma_k^k P_{ki}^{(m)} > 0$ and $\gamma_i^k P_{ik}^{(n)} \leq \gamma_k^k \leq 1$ which proves our claim. $\square$

In fact, one can prove a stronger statement that every invariant measure will be a multiple of γ^k.

State i is said to be positive recurrent if $m_i = \mathbb{E}_i[T_i] < \infty$; this obviously subsumes recurrence. If $m_i = +\infty$, then state i is said to be null recurrent. One can then show that for an irreducible P, positive recurrence holds for every state, P has an invariant distribution π such that $\pi_i = \frac{1}{m_i}$, and these are equivalent statements. Note that $\sum_{j\in I}\gamma^i_j = m_i$ so that m_i being finite implies that γ^i is summable. If I is finite and P is irreducible, then we can provide an easy proof for positive recurrence through analytical techniques. Considerable effort goes into establishing transience, recurrence and positive recurrence of the states of a given Markov chain when the state-space is not finite.

Thus, far we have established when there exists an invariant distribution. However, in applications we'd also be interested in finding out when $P^{(n)}_{ij} \to \pi_j$ as $n \to \infty$. A state i is said to be aperiodic if $P^{(n)}_{ii} > 0$ for all sufficiently large n. For an irreducible P having one state being aperiodic implies that all states are aperiodic as follows from the lemma below.

Lemma 8. *If P is irreducible with an aperiodic state i, then for all states j, k we have $P^{(n)}_{jk} > 0$ for sufficiently large n. As a consequence, all states are aperiodic.*

Proof. Since P is irreducible, there exist $r, s \geq 0$ with $P^{(r)}_{ji} > 0$ and $P^{(s)}_{ik} > 0$. Since i is aperiodic, for sufficiently large n we have $P^{(n)}_{ii} > 0$. This then implies that $P^{(r+n+s)}_{jk} \geq P^{(r)}_{ji}P^{(n)}_{ii}P^{(s)}_{ik} > 0$. □

We now state and prove the main convergence theorem for Markov chains.

Theorem 13. *Let P be irreducible and aperiodic with invariant distribution π. Let μ be any distribution and let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a (μ, P) Markov chain. Then*

$$\lim_{n\to\infty} P(X_n = j) = \pi_j \quad \forall j \in I.$$

In particular, this implies that $\lim_{n\to\infty} P^{(n)}_{ij} = \pi_j$ for all $i, j \in I$.

Proof. The proof is using a general technique known as coupling. Let $\{Y_i\}_{i\in\mathbb{Z}_+}$ be a (π, P) Markov chain that is independent of $\{X_i\}_{i\in\mathbb{Z}_+}$. Fix a state b and set

$$T = \inf\{n \in \mathbb{N} : X_n = Y_n = b\}.$$

First we will prove that $\mathbb{P}(T < \infty) = 1$. Note that the process $W_n = (X_n, Y_n)$ is also a Markov chain with transition matrix $\tilde{P}_{(i,k)(j,l)} = P_{ij}P_{kl}$ and initial distribution $\tilde{\mu}_{(i,k)} = \mu_i\pi_k$. By the aperiodicity of P, we have for all (i, k) and (j, l) that $\tilde{P}^{(n)}_{(i,k)(j,l)} = P^{(n)}_{ij}P^{(n)}_{kl} > 0$ for n sufficiently large. Thus, $\tilde{P}$ is irreducible. Also, $\tilde{P}$ has an invariant distribution $\tilde{\pi}$ that is given by $\tilde{\pi}_{(i,k)} = \pi_i\pi_k$ so that $\tilde{P}$ is positive recurrent. Now T is the first passage time to (b, b), and hence, it follows that $\mathbb{P}(T < \infty) = 1$.

Now we will define a new process Z_n as follows

$$\forall n \in \mathbb{Z}_+ \text{ set } Z_n = \begin{cases} X_n & \text{if } n < T; \\ Y_n & \text{if } n \geq T. \end{cases}$$

Thus, Z_n follows X_n until the time both X_n and Y_n hit k after which it follows Y_n. (Need to add Figure on pg. 42 from J. Norris Book.) By the strong Markov property, we know that $\{(X_{T+n}, Y_{T+n})\}_{n\in\mathbb{Z}_+}$ and $\{(Y_{T+n}, X_{T+n})\}_{n\in\mathbb{Z}_+}$ are $(\delta_{(b,b)}, \tilde{P})$ Markov chains that are independent of the past up to T. Then it follows that $\{(Z_n, Z'_n)\}_{n\in\mathbb{Z}_+}$ is a $(\tilde{\mu}, \tilde{P})$ Markov chain where $Z'_n = X_n 1_{\{n\geq T\}} + Y_n 1_{\{n<T\}}$. Thus, it also follows that $\{Z_n\}_{n\in\mathbb{Z}_+}$ is a (μ, P) Markov chain.

By the definition of Z_n we know that

$$\mathbb{P}(Z_n = j) = \mathbb{P}(X_n = j, n < T) + \mathbb{P}(Y_n = j, n \geq T),$$

but since $\{Z_n\}_{n\in\mathbb{Z}_+}$ is (μ, P) Markov, we also have $\mathbb{P}(Z_n = j) = \mathbb{P}(X_n = j)$. Therefore,

$$\begin{aligned}|\mathbb{P}(X_n = j) - \pi_j| &= |\mathbb{P}(Z_n = j) - \mathbb{P}(Y_n = j)| \\ &= |\mathbb{P}(X_n = j, n < T) - \mathbb{P}(Y_n = j, n < T)| \leq \mathbb{P}(T > n)\end{aligned}$$

Since $\mathbb{P}(T < \infty) = 1$, the right goes to 0 as $n \to \infty$, which finishes the proof. □

Note that we got a strong result that also yields the rate of convergence. The last inequality can be turned into a total variational metric inequality.

Without the aperiodic assumption we get the following results.

Theorem 14. *Let P be irreducible. There is an integer $d \geq 1$ and a partition $I = \sqcup_{k=0}^{d-1} C_k$ such that*

(1) $P_{ij}^{(n)} > 0$ *only if $i \in C_r$ and $j \in C_{\bmod (r+n,d)}$ for some $r \in \{0, 1, \ldots, d-1\}$;*
(2) $P_{ij}^{(nd)} > 0$ *for all sufficiently large n, for all $i, j \in C_r$ and for all $r \in \{0, 1, \ldots, d-1\}$.*

The integer d is called the period of the Markov chain/transition matrix.

Theorem 15. *Let P be irreducible of period d where $C_0, C_1, \ldots, C_{d-1}$ is the partition of the state-space as in Theorem 14. Let λ be a distribution with $\sum_{i\in C_0} \lambda_i = 1$. Suppose that $\{X_n\}_{n\in\mathbb{Z}_+}$ is (λ, P) Markov. Then for $r = 0, 1, \ldots, d-1$ and $j \in C_r$ we have*

$$\lim_{n\to\infty} \mathbb{P}(X_{nd+r} = j) = \frac{d}{m_j},$$

where m_j is the expected return time to j (mean first passage time). In particular, for $i \in C_0$ and $j \in C_r$ we have $\lim_{n\to\infty} P_{ij}^{(nd+r)} = \frac{d}{m_j}$

Note that this theorem covers both the positive recurrent and null recurrent case; for the latter the limit is 0.

Reversibility: For some results it is useful to look at time in reverse. We will do so for Markov chains that are in equilibrium (argue why this is necessary) by running them backwards. We have a basic result.

Theorem 16. *Let P be irreducible and have an invariant distribution π. Suppose that $\{X_n\}_{0\leq n\leq N}$ is (π, P) Markov and set $Y_n = X_{N-n}$. Then $\{Y_n\}_{0\leq n\leq N}$ is $(\pi, \hat{P})$ Markov, where $\hat{P}$ is given by*

$$\pi_j \hat{P}_{ji} = \pi_i P_{ij} \text{ for all } i, j \in I.$$

The process $\{Y_n\}_{0\leq n\leq N}$ is called the time-reversal of $\{X_n\}_{0\leq n\leq N}$. A stochastic/transition matrix and a measure λ are said to be in detailed balance if

$$\lambda_j P_{ji} = \lambda_i P_{ij} \text{ for all } i, j \in I.$$

It is easy to verify that λ is invariant for P, i.e., $\lambda = \lambda P$. Let $\{X_n\}_{0\leq n\leq N}$ be (π, P) Markov with P irreducible, then $\{X_n\}_{0\leq n\leq N}$ is called reversible if, for all $N \geq 1$, $\{X_{N-n}\}_{0\leq n\leq N}$ is also (π, P) Markov. An important result then says that $\{X_n\}_{0\leq n\leq N}$ is reversible if and only if π (a distribution) and P are in detailed balance.

Ergodic theorem: Denote by $V_i(n) = \sum_{k=0}^{n-1} 1_{\{X_k=i\}}$ the numbers of visits to i before time n. The result we seek is the following.

Theorem 17. *Let P be irreducible and μ be any distribution. If $\{X_i\}_{i\in\mathbb{Z}_+}$ is (μ, P) Markov, then*

$$\lim_{n\to\infty} \frac{V_i(n)}{n} = \frac{1}{m_i} \quad a.s.$$

Moreover, if P is positive recurrent (with invariant distribution π), then for any bounded function $f : I \mapsto \mathbb{R}$, we have

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} f(X_k)}{n} = \bar{f} \left(:= \sum_{i\in I} \pi_i f_i = \mathbb{E}_\pi[f(X)] \right) \quad a.s..$$

Proof. If P is transient, then with probability 1, the total number V_i of visits to state i is finite, so $\frac{V_i(n)}{n} \le \frac{V_i}{n} \to_{n\to\infty} 0 = \frac{1}{m_i}$. Thus, we now assume that P is recurrent and start by fixing a reference state b. By the Strong Markov property, it is sufficient to have the initial distribution be δ_i.

Consider the excursion times $\{S_i\}_{i\in\mathbb{N}}$ which are *i.i.d.* finite random variables with $\mathbb{E}_i[S_i^{(r)}] = m_i$. By the definition note that

$$\sum_{k=1}^{V_i(n)-1} S_i^{(k)} \le n-1 < n \le \sum_{k=1}^{V_i(n)} S_i^{(k)}.$$

Hence,

$$\frac{\sum_{k=1}^{V_i(n)-1} S_i^{(k)}}{V_i(n)} \le \frac{n}{V_i(n)} \le \frac{\sum_{k=1}^{V_i(n)} S_i^{(k)}}{V_i(n)}$$

We now get the first part of the result by the recurrence of P (as $\mathbb{P}(V_i(n) \to_{n\to\infty} \infty) = 1$) and the Strong Law of Large Numbers.

For the second part, assume (without loss of generality) that $|f| \le 1$, then for any $J \subseteq I$ we have

$$\begin{aligned}
\left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \bar{f} \right| &= \left| \sum_{i\in I} \left(\frac{V_i(n)}{n} - \pi_i \right) f_i \right| \\
&\le \sum_{i\in J} \left| \frac{V_i(n)}{n} - \pi_i \right| + \sum_{i\in J^C} \left| \frac{V_i(n)}{n} - \pi_i \right| \\
&\le \sum_{i\in J} \left| \frac{V_i(n)}{n} - \pi_i \right| + \sum_{i\in J^C} \left(\frac{V_i(n)}{n} + \pi_i \right) \\
&\le \sum_{i\in J} \left| \frac{V_i(n)}{n} - \pi_i \right| + \frac{\sum_{i\in J^C} V_i(n)}{n} + \sum_{i\in J^C} \pi_i \\
&\le \sum_{i\in J} \left| \frac{V_i(n)}{n} - \pi_i \right| + 1 - \frac{\sum_{i\in J} V_i(n)}{n} + \sum_{i\in J^C} \pi_i \\
&\le \sum_{i\in J} \left| \frac{V_i(n)}{n} - \pi_i \right| + \sum_{i\in J} \left(\pi_i - \frac{V_i(n)}{n} \right) + 2 \sum_{i\in J^C} \pi_i \\
&\le 2 \sum_{i\in J} \left| \frac{V_i(n)}{n} - \pi_i \right| + 2 \sum_{i\in J^C} \pi_i
\end{aligned}$$

Given $\epsilon > 0$, choose J finite such that $\sum_{i \in J^C} \pi_i < \epsilon/4$, and then $N = N(\omega)$ so that, for $n \geq N(\omega)$ we have

$$\sum_{i \in J} \left| \frac{V_i(n)}{n} - \pi_i \right| < \epsilon/4;$$

note that J being finite ensures that $N(\omega)$ is finite too. Now for $n \geq N(\omega)$ we have

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \bar{f} \right| < \epsilon,$$

which finishes the proof. □

3.5. **Exercises.** Please show all your work.

(1) For the countable case make precise the definitions of P^n and λP where P is a stochastic matrix and λ is a distribution. *Hint:* You need to show that the associated series are finite.

(2) Let $\{X_i\}_{i\in\mathbb{Z}_+}$ be a (μ, P) Markov chain. Prove that the following are Markov chains and determine the parameters: (1) $Y_i = (X_i, X_{i+1})$ for $i \in \mathbb{Z}_+$; and (2) $Y_n = X_{kn}$ for some fixed $k > 0$.

(3) Let $\{Z_i\}_{i\in\mathbb{N}}$ be *i.i.d.* taking values in $\{0, 1\}$ with probability $(1-p, p)$. Let $\{S_i\}_{i\in\mathbb{Z}_+}$ be the partial sums process. Which of these is a DTMC: (a) $X_n = Z_{n+1}$; (b) $X_n = S_n$; (c) $X_n = S_0 + S_1 + \cdots + S_n$; and (d) $X_n = (S_n, S_0 + S_1 + \ldots + S_n)$. In all cases find the state-space and the parameters. For the cases that are not a DTMC, show explicitly.

(4) Let T_i be the first passage time to state i. Prove the following recursion using the Markov property and definition of stopping times:

$$\mathbb{P}_j(X_n = i) = \sum_{m=1}^{n} \mathbb{P}_j(T_i = m) P_{ii}^{(n-m)}$$

For any state $i \in I$, show the following $\sum_{m=0}^{n} \mathbb{P}_i(X_m = i) \geq \sum_{m=k}^{n+k} \mathbb{P}_i(X_m = i)$ for any $n, k \in \mathbb{Z}_+$. *Hint:* Use the following random time $T_i(k) = \inf\{n > k : X_n = i\}$ and a version of the recursion above to prove the result - is this a stopping time? The second part is **Extra credit**.

(5) Let I be finite. Let P be irreducible. Then prove that P has an invariant distribution. Use linear algebra (eigenvalues and eigenvectors) to prove the result. *Note that similar arguments can be made in the non-irreducible case.* Let A be an $I \times I$ matrix of all ones and $\mathbf{1}$ the row vector of all ones (with I columns). Then show that an invariant distribution π for P satisfies $\pi(\mathbf{I} - P + A) = \mathbf{1}$ where $\mathbf{I}$ is the $I \times I$ identity matrix. Under the irreducibility assumption further show that $\mathbf{I} - P + A$ is invertible so that one can solve for the invariant distribution quite easily using matrix inversion.

(6) *Markov chains and martingales*: Let $A, B \subset I$ with $A \cap B = \emptyset$. Let $\tau_A = \inf\{n \geq 0 : X_n \in A\}$ and similarly τ_B. Define the following function $h(x) = \mathbb{P}_x(\tau_A < \tau_B)$ for all $x \in I$. Suppose $\mathbb{P}_x(\tau_{A\cup B} < \infty) > 0$ (where $\tau_{A\cup B}$ is defined in the same manner as τ_A or τ_B) for all $x \notin A \cup B$, then show the following recursion

$$h(x) = \sum_{y\in I} P_{xy} h(y) \quad \forall x \notin A \cup B.$$

Show that any function $g(\cdot)$ that satisfies the above recursion is such that $\{g(X_{n\wedge\tau_{A\cup B}})\}$ is a martingale.

(7) *Markov chains and martingales*: Let $A \subset I$. Define function $g(x) = \mathbb{E}_x[\tau_A]$. Assume that $\mathbb{P}_x(\tau_A < \infty) > 0$. Now show the following recursion:

$$g(x) = 1 + \sum_{y\in I} P_{xy} g(y) \quad \forall x \notin A$$

Show that any function $h(\cdot)$ satisfying the recursion above is such that $\{h(X_{n\wedge\tau_A}) + n \wedge \tau_A\}$ is a martingale.

(8) *Convergence to equilibrium*: Let I be finite and P be irreducible and aperiodic. Show that there exists an $r < 1$ and $C < \infty$ such that $\mathbb{P}(T > n) \leq Cr^n$ where T is the

coupling time used in the proof in class. *Hint*: First consider the case of $P_{ij} > 0$ for all i, j and then consider the general case by taking a suitable power of P. Thus, we get exponentially fast convergence. **Extra credit**

(9) *Construction of Markov chains:* Assume we're given a stochastic matrix P and a distribution μ. Let $([0,1], \mathcal{B}([0,1]), Leb([0,1]))$ be the underlying probability space. Using the exercise from Homework 1, we know that we can generate a countable number of independent random variables with any desired distribution. Our random variables will take values in $(I, \mathcal{I})$. Let $(X_0, Y_{i,n} : i \in I, n \in \mathbb{N})$ be independent with values in I. Let X_0 be distributed as per μ and $\{Y_{i,n}\}_{m\in\mathbb{N}}$ be *i.i.d.* as per $P_{i\bullet}$ (the distribution corresponding to row i of the stochastic matrix P) for all $i \in I$. Construct sequence $\{X_n\}_{n\in\mathbb{Z}_+}$ as follows starting from X_1:

$$X_n(\omega) = Y_{X_{n-1}(\omega),n}(\omega) \quad \forall n \in \mathbb{N}$$

Show that $\{X_n\}_{n\in\mathbb{Z}_+}$ is a (μ, P) Markov chain. Note that we're not being economical with the random variables, so how can one modify the logic above so as to use as many of the generated random variables as possible? *Hint:* Think how you would simulate such a process. Along the same lines show that with a sequence of independent random variables $(X_0, Y_i : i \in \mathbb{N})$ with X_0 distributed as μ and every Y_i being uniformly distributed in $[0,1]$, show that the following recursive construction yields a Markov chain

$$X_n = G(X_{n-1}, Y_n) \quad \forall i \in \mathbb{N}$$

where $G(\cdot,\cdot) : I \times [0,1] \mapsto I$. What are the parameters? **Extra credit**

4. Continuous-time Markov Chains

Sources: The sources for this chapter are

- J. R. Norris, "Markov chains," Cambridge Series in Statistical and Probabilistic Mathematics, 2, Cambridge University Press, Cambridge, 2006.
- P. Brémaud, "Markov chains: Gibbs fields, Monte Carlo simulation, and queues," Texts in Applied Mathematics, 31, Springer-Verlag, New York, 1999.
- R. Durrett, "Probability: Theory and examples," Second edition. Duxbury Press, Belmont, CA, 1996.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space. Our stochastic processes, say $\{X_t\}_{t\in\mathbb{R}_+}$, will take values in a countable space I. We will discuss notions of classes, irreducibility, recurrence and convergence for continuous-time Markov chains. However, to avoid some technical measure theoretic issues, we will take a few results as given. We will start this part with the Poisson process. All processes that we will work with are assumed to be right-continuous with left-hand limits (*rcll* or *cadlag*, in short), i.e., under the countable state-space assumption we also have that for all $\omega \in \Omega$ and for all $t \geq 0$, we have an $\epsilon > 0$ such that $X_s(\omega) = X_t(\omega)$ for all $t \leq s \leq t + \epsilon$.

One issue that we will have to deal with is explosive versus non-explosive processes. From the above it is clear that our stochastic processes will stay in a state for non-zero amount of time and can be represented as a jump process. This implies one of three possibilities: only a finite number of jumps; an infinite number of jumps but only a finite number in each interval of time; and an infinite number of jumps in a finite amount of time. The last case is where a process is said to be explosive and one defines time ζ to be the explosion time. For our purpose we will assume that explosive processes reach a special (final/terminal) state ∞ at the explosion time and remain there. We will find conditions to exclude explosive processes; note that the first case is not of interest either. Similar to hitting times of states of a DTMC, we now define jump times $\{J_i\}_{i\in\mathbb{Z}_+}$, holding times $\{S_i\}_{i\in\mathbb{N}}$ and jump chain $\{Y_i\}_{i\in\mathbb{Z}_+}$ as follows:

$$\begin{aligned}
\forall i \in \mathbb{Z}_+ \quad & J_i = \begin{cases} 0 & \text{if } i = 0; \\ \inf\{t \geq J_{i-1} : X_t \neq X_{J_{i-1}}\} & \text{otherwise} \end{cases} \\
\forall i \in \mathbb{N} \quad & S_i = (J_i - J_{i-1}) 1_{\{J_{i-1} < \infty\}} \\
\forall i \in \mathbb{Z}_+ \quad & Y_i = X_{J_i}
\end{aligned}$$

At this point we don't ascribe any other structure to our stochastic process. Note that $S_i > 0$ for all $i \in \mathbb{N}$. If $J_{n+1} = \infty$ for some n, then set $X_\infty = X_{J_n}$, and otherwise it is not defined. The explosion time $\zeta = \sup_{n\in\mathbb{Z}_+} J_n = \sum_{n=1}^{\infty} S_n$.

4.1. **Poisson process.**

The first process that we will consider is one where $\{S_i\}_{i\in\mathbb{N}}$ is *i.i.d.* $\exp(\lambda)$ and $X_{J_i} = i$ for all $i \in \mathbb{Z}_+$. This is called a Poisson process of rate λ; we will denote a Poisson process by N. Note that the strong law of large numbers then implies that $\zeta = \infty$ (*a.s.*) and the process is non-explosive; also since $\mathbb{P}(S_i < \infty) = 1$ for all $i \in \mathbb{N}$, there have to be an infinite number of jumps.

We have a bunch of properties for $\{N_t\}_{t\in\mathbb{R}_+}$ with rate λ. The first is the Markov property.

Theorem 18. *Let $\{N_t\}_{t\in\mathbb{R}_+}$ be a Poisson process of rate λ, then, for any $s \geq 0$, the process $\{N_{t+s} - N_s\}_{t\in\mathbb{R}_+}$ is a Poisson process of rate λ that is, in addition, independent of $\{N_t\}_{t\in[0,s]}$.*

Proof. We will prove the result by conditioning on $N_s = i$ for all $i \in \mathbb{Z}_+$. Set $\tilde{N}_t = N_{t+s} - N_s$. Note that

$$\{N_s = j\} = \{J_j \le s < J_{j+1}\} = \{J_j \le s\} \cap \{S_{j+1} > s - J_j\}$$

Conditioned on this event, we have

$$X_r = \sum_{n=1}^{j} 1_{S_n \le r} \quad \forall r \in [0, s],$$

and for the process $\tilde{N}$ we can define the jump times $\{\tilde{S}_i\}_{i\in\mathbb{N}}$ as follows

$$\tilde{S}_i = \begin{cases} S_{j+1} - s + J_j & \text{if } i = 1 \\ S_{j+i} & \text{otherwise} \end{cases}$$

By the memoryless property of the exponential distribution and the independence of $\{S_i\}_{i\in\mathbb{N}}$, conditioned on $N_s = i$ we get that $\{\tilde{S}_i\}_{i\in\mathbb{N}}$ is also an *i.i.d.* $\exp(\lambda)$ sequence that is independent of $\{N_r\}_{r\in[0,s]}$. Hence, $\tilde{N}$ is a Poisson process of rate λ. Note that this holds for all $i \in \mathbb{Z}_+$, and so without the conditioning too. □

We will state without proof the Strong Markov property of a Poisson process. As before, for a filtration $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$, a random time T is a stopping time if $\{T \le t\} \subset \mathcal{F}_t$.

Theorem 19. *Let $\{N_t\}_{t\in\mathbb{R}_+}$ be a Poisson process of rate λ and T is a stopping time associated with N, then conditioned on the event $T < \infty$, we get that $\{N_{t+T} - N_T\}_{t\in\mathbb{R}_+}$ is a Poisson process of rate λ that is independent of $\{N_t\}_{t\in[0,T]}$.*

Effectively we have shown that the process $\{N_t - \lambda t\}_{t\in\mathbb{R}_+}$ is a martingale and the optional sampling theorem applies. This is an important fact to remember.

For a process X, we say that it has stationary increments if the distribution of $X_{t+s} - X_s$ is the same as that of X_t for all $t \ge 0$. Similarly we say that the increments are independent if given $n \in \mathbb{N}$ and $0 \le t_1 < t_2 < \cdots < t_n < \infty$, the random variables $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ are independent. Then we have the following characterizations of a Poisson process (without proof).

Theorem 20. *Let $\{N_t\}_{t\in\mathbb{R}_+}$ be a right-continuous and increasing process taking values in $\mathbb{Z}_+$ and starting from 0. Let $0 < \lambda < \infty$. Then the following three conditions are equivalent:*

(1) *(jump chain+holding time definition) the holding times $\{S_i\}_{i\in\mathbb{N}}$ of N are* i.i.d. $\exp(\lambda)$ *and the jump chain is given by $Y_n = n$ for all n;*
(2) *(infinitesimal definition) N has independent increments and, as $s \downarrow 0$, uniformly in t,*

$$\mathbb{P}(N_{t+s} - N_t = 0) = 1 - \lambda s + o(s), \quad \mathbb{P}(N_{t+s} - N_t = 1) = \lambda s + o(s),$$

where $o(s)$ stands for some function $f(s)$ such that $f(s)/s \to 0$ as $s \downarrow 0$; and
(3) *(transition probability description) N has stationary and independent increments, and, for each $t \ge 0$, N_t has the Poisson distribution of parameter λt.*

Any process N that satisfies any of these conditions is called a Poisson process (of rate λ).

Another property of a Poisson process is the following: conditioned on $N_t = n$, the location of the n jumps in $[0, t]$ has the same distribution as the ordered sample of size n of the uniform distribution on $[0, t]$.

4.2. **Continuous time Markov chains.** The Poisson process is one example of a CTMC, but as we will see later on, every other CTMC can be derived from a Poisson process. However, first we start with the definition of a rate matrix (also a Q-matrix or generator). A matrix/operator Q is called a rate matrix if the following hold:

(1) $0 \leq q_{ii} < \infty$ for all $i \in I$;
(2) for all $i \in I$, $q_{ij} \geq 0$ for all $j \in I \setminus \{i\}$;
(3) $\sum_{j \in I} q_{ij} = 0$ for all $i \in I$.

Define $q_i = \sum_{j \in I: j \neq i} q_{ij}$, then $q_{ii} = -q_i$. We will now obtain a stochastic matrix Π from Q which we will call the jump matrix as follows for all $i, j \in I$,

$$\Pi_{ij} = \begin{cases} q_{ij}/q_i & \text{if } j \neq i \text{ \& } q_i \neq 0; \\ 0 & \text{if } j \neq i \text{ \& } q_i 0 \end{cases},$$

$$\Pi_{ii} = \begin{cases} 0 & \text{if } q_i \neq 0; \\ 1 & \text{otherwise} \end{cases}.$$

As for the Poisson process, we now have three equivalent characterization of a CTMC $\{X_t\}_{t \in \mathbb{R}_+}$. The first is via the jump chain and holding times. A stochastic process X is called a (μ, Q) Markov chain (for some distribution μ) if the jump chain $\{Y_i\}_{i \in \mathbb{Z}_+}$ is a (μ, Π) DTMC and, for every $n \geq 1$, conditioned on $(Y_0, Y_1, \dots, Y_{n-1})$, the holding times $S_1, S_2, \dots, S_n$ are independent exponential random variables with parameters $q_{Y_0}, q_{Y_1}, \dots, q_{Y_{n-1}}$ respectively. The second construction is as follows. Choose $X_0 = Y_0$ with distribution μ and also choose an array $(T_n^i : n \geq 1, i \in I)$ of *i.i.d.* $\exp(1)$ random variables. Then, inductively for $n \in \mathbb{Z}_+$, if $Y_n = i$ we set

$$S_{n+1}^j = T_{n+1}^j / q_{ij}, \quad \text{for } j \neq i$$
$$S_{n+1} = \inf_{j \neq i} S_{n+1}^j$$
$$Y_{n+1} = \begin{cases} j & \text{if } S_{n+1}^j = S_{n+1} < \infty \\ i & \text{if } S_{n+1} = \infty. \end{cases}$$

The third characterization will be using Poisson processes. Here we start with the initial state $X_0 = Y_0$ chosen with distribution μ and a family of independent Poisson processes $\{(N_t^{ij})_{t \in \mathbb{R}_+} : i, j \in I, i \neq j\}$ with Poisson process N^{ij} having rate q_{ij}. Then set $J_0 = 0$ and define inductively for $i \in \mathbb{Z}_+$

$$J_{n+1} = \inf\{t > J_n : N_t^{Y_n j} \neq N_{J_n}^{Y_n j} \text{ for some } j \neq Y_n\}$$
$$Y_{n+1} = \begin{cases} j & \text{if } J_{n+1} < \infty \text{ \& } N_{J_{n+1}}^{Y_n j} \neq N_{J_n}^{Y_n j} \\ i & \text{if } J_{n+1} = \infty. \end{cases}$$

We will use this characterization later on to prove some asymptotic results (Kurtz's Theorem). Note that we can use one Poisson process to generate the required family of Poisson processes if $\gamma = \sup_{i \in I} q_i < \infty$; at the jumps of the Poisson process of rate γ when state is i, choose to stay in state i with probability $1 - q_i/\gamma$ or jump to state $j \neq i$ with probability q_{ij}/γ. This extremely useful procedure is called uniformization.

Explosion: We will now discuss when processes are explosive and when not. In general, we say that a rate matrix Q is explosive if the associated Markov chain is such that $\mathbb{P}_i(\zeta <$

$\infty) > 0$ for some $i \in I$; otherwise we call it non-explosive. We start with some simple conditions that are easy to check.

Theorem 21. *Let $\{X_t\}_{t\in\mathbb{R}_+}$ be a (μ, Q) Markov chain. Then X does not explode if any one of the following conditions holds:*

(1) *I is finite;*
(2) $\sup_{i\in I} q_i < \infty$*;*
(3) *$X_0 = i$, and i is recurrent for the jump chain.*

Proof. Define $T_n = q(Y_{n-1})S_n$, then $\{T_i\}_{i\in\mathbb{N}}$ are *i.i.d.* $\exp(1)$ and independent of the jump chain Y. In the first two cases, we have $\gamma = \sup_{i\in I} q_i < \infty$ and

$$q\zeta \geq \sum_{n=1}^{\infty} T_n = \infty$$

with probability 1 (SLLN). In the third case, we know that Y visits i infinitely often, at (integral) times $\{M_i\}_{i\in\mathbb{N}}$ (obtained similar to the passage times), then

$$q_i\zeta \geq \sum_{n=1}^{\infty} T_{M_n} = \infty$$

with probability 1. □

Necessary and sufficient conditions for Q to be explosive are now discussed as a consequence of the following theorem.

Theorem 22. *Let $\{X_t\}_{t\in\mathbb{R}_+}$ be a Markov process with rate matrix Q and let ζ be the explosion time of X. Fix a $\theta > 0$ and define $z_i = \mathbb{E}_i[\exp(-\theta\zeta)]$. Then z satisfies:*

(1) $|z_i| \leq 1$*;*
(2) *$Qz = \theta z$.*

In addition, z is the maximal solution that satisfies both properties.

Proof. The first is obvious so we discuss the second part. Fix a state i and condition on $X_0 = i$. Then consider the first jump, the time of it J_1 which is $\exp(q_i)$ and the next state j with probability $\mathbb{P}_i(X_{J_1} = j) = \Pi_{ij}$ are independent. Additionally, by the Markov property of the jump chain at time $n = 1$, conditional on $X_{J_1} = k$, the process $\{X_{J_1+t}\}_{t\in\mathbb{R}_+}$ is (δ_k, Q) Markov and independent of J_1. Using these facts we get

$$\begin{aligned}\mathbb{E}_i[e^{-\theta\zeta}|X_{J_1} = k] &= \mathbb{E}_i\left[e^{-\theta J_1}e^{-\theta\sum_{n=2}^{\infty} S_n}|X_{J_1} = k\right] \\ &= \int_0^{\infty} e^{-\theta t}q_i e^{-q_i t}dt\ \mathbb{E}_k[e^{-\theta\zeta}] = \frac{q_i z_k}{q_i + \theta}.\end{aligned}$$

Now we have

$$z_i = \sum_{k\in I:k\neq i} \mathbb{P}_i(X_{J_1} = k)\mathbb{E}_i[e^{-\theta\zeta}|X_{J_1} = k] = \sum_{k\in I:k\neq i} \frac{q_i\Pi_{ik}z_k}{q_i + \theta} = \sum_{k\in I:k\neq i} \frac{q_{ik}z_k}{q_i + \theta}$$

Recognizing that $q_i = -q_{ii}$ and with some minor manipulations, we get

$$\theta z_i = \sum_{k\in I} q_{ik}z_k.$$

The maximality part can be found in J. Norris' book. □

We then have a corollary that gives the necessary and sufficient conditions for non-explosiveness.

Corollary 3. *For each $\theta > 0$, the following are equivalent:*

(1) *Q is non-explosive;*
(2) *$Qz = \theta z$ and $|z_i| \leq 1$ for all i, imply that $z \equiv 0$.*

Proof. If the first condition holds, then $\mathbb{P}_i(\zeta = \infty) = 1$ so that $\mathbb{E}_i[e^{-\theta\zeta}] = 0$. Using Theorem 22 we see that any other solution of $Qz = \theta z$ such that $|z_i| \leq 1$ implies that $|z_i| \leq \mathbb{E}_i[e^{-\theta\zeta}] = 0$, proving the result in one direction. If the second condition holds, then by Theorem 22 we know that the particular choice of $\mathbb{E}_i[e^{-\theta\zeta}]$ also has to be 0 for all i. This then implies that $\mathbb{P}_i(\zeta = \infty) = 1$. $\square$

Strong Markov property: We state without proof (to steer clear of measure-theoretic issues) the Strong Markov property of CTMCs.

Theorem 23. *Let $\{X_t\}_{t\in\mathbb{R}_+}$ be a (μ, Q) Markov chain and let T be a stopping time associated with X. Then, conditioned on $T < \infty$ and $X_T = i \in I$, the process $\{X_{T+t}\}_{t\in\mathbb{R}_+}$ is a (δ_i, Q) Markov chain that is independent of $\{X_t\}_{t\in[0,T]}$.*

Forward and backward equations: Our description of CTMCs thus far is lacking in some regards. Given X being a (μ, Q) Markov chain, we still haven't specified how to obtain $\mathbb{P}(X_t = j)$ or even $\mathbb{P}_i(X_t = j)$. We start with the finite state-space case where there is an easy answer.

Theorem 24. *Let $\{X_t\}_{t\in\mathbb{R}_+}$ be a (μ, Q) Markov chain with a finite state-space I. Then we have following additional equivalent characterizations of the process:*

(1) *(infinitesmal definition) for all $t, h \geq 0$, conditional on $X_t = i$, X_{t+h} is independent of $\{X_s\}_{s\in[0,t]}$ and, as $h \downarrow 0$, uniformly in t for all j*
$$\mathbb{P}(X_{t+h} = j | X_t = i) = \delta_{ij} + q_{ij}h + o(h);$$
(2) *Let $P(t) = e^{Qt}$ (matrix exponential) with $P(0) = I$, then $P(t)$ is the unique solution to the forward equation*
$$\frac{d}{dt}P(t) = P(t)Q, \quad P(0) = I,$$
the backward equation
$$\frac{d}{dt}P(t) = QP(t), \quad P(0) = I,$$
and leads to the following characterization of the process X, for all $n \in \mathbb{Z}_+$ and all times $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n+1}$ and all states $i_0, i_1, \ldots, i_{n+1}$ we have
$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = P_{i_n i_{n+1}}(t_{n+1} - t_n).$$

For the countable case we cannot use the matrix exponential characterization. Thus, we have the following two results.

Theorem 25. *Let Q be a rate matrix. Then the backward equation*
$$\frac{d}{dt}P(t) = QP(t), \quad P(0) = I$$
has a minimal non-negative solution $\{P(t)\}_{t\in\mathbb{R}_+}$, and this solution forms a matrix/operator semi-group, i.e., $P(s)P(t) = P(s + t)$ for all $s, t \geq 0$.

Theorem 26. *A process $\{X_t\}_{t\in\mathbb{R}_+}$ that is a (μ, Q) Markov chain where $\{P(t)\}_{t\in\mathbb{R}_+}$ is the non-negative minimal semigroup solution to the backward solution is also characterized by the following, for all $n \in \mathbb{Z}_+$ and all times $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n+1}$ and all states $i_0, i_1, \ldots, i_{n+1}$ we have*

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = P_{i_n i_{n+1}}(t_{n+1} - t_n).$$

Theorems 25 and 26. We will only prove that the backward equation; for the rest of the details please see book by J. Norris. Let $P_{ij}(t) = \mathbb{P}_i(X_t = j)$. The semigroup property is easy to argue following the logic of the Chapman-Kolmogorov equations. Conditioning on $X_0 = i$, we have $J_1 \sim \exp(q_i)$ and X_{J_1} $(\pi_{ik} : k \in I)$. Then conditional on $J_1 = s$ and $X_{J_1} = k$ we have $\{X_{t+s}\}_{t\in\mathbb{R}_+}$ being a (δ_k, Q) Markov chain. Then we have the following

$$\mathbb{P}_i(X_t = j, t < J_1) = e^{-q_i t}\delta_i(j), \text{ and}$$
$$\mathbb{P}_i(J_1 \leq t, X_{J_1} = k, X_t = j) = \int_0^t q_i e^{-q_i s} \pi_{ik} P_{kj}(t-s) ds,$$

which then yields

$$\begin{aligned} P_{ij}(t) &= \mathbb{P}_i(X_t = j, t < J_1) + \sum_{k\neq i} \mathbb{P}_i(J_1 \leq t, X_{J_1} = k, X_t = j) \\ &= e^{-q_i t}\delta_i(j) + \sum_{k\neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} P_{kj}(t-s) ds \end{aligned}$$

Setting $u = t - s$ and rearranging we get

$$e^{q_i t} P_{ij}(t) = \delta_i(j) + \sum_{k\neq i} \int_0^t q_i e^{q_i u} \pi_{ik} P_{kj}(u) du$$

We can interchange the summation and integral by monotone convergence (or Tonelli's theorem) to get

$$e^{q_i t} P_{ij}(t) = \delta_i(j) + \int_0^t \sum_{k\neq i} q_i e^{q_i u} \pi_{ik} P_{kj}(u) du$$

We get a few results from this: (i) $P_{ij}(t)$ is continuous in t for all i, j; and (ii) the integrand is a uniformly converging sum of continuous functions so that $P_{ij}(t)$ is also differentiable. Differentiating we get

$$e^{q_i t}\left(q_i P_{ij}(t) + \frac{d}{dt} P_{ij}(t)\right) = \sum_{k\neq i} q_i e^{q_i t} \pi_{ik} P_{kj}(t),$$

which after rearranging yields

$$\frac{d}{dt} P_{ij}(t) = \sum_{k\in I} q_{ik} P_{kj}(t).$$

□

We will now prove the forward equation. First we prove a simple time-reversal result

Lemma 9. *The following holds*

$$q_{i_n}\mathbb{P}(J_n \le t < J_{n+1}|Y_0 = i_0, Y_1 = i_1, \dots, Y_n = i_n)$$
$$= q_{i_o}\mathbb{P}(J_n \le t < J_{n+1}|Y_0 = i_n, \dots, Y_{n-1} = i_1, Y_0 = i_0)$$

Proof. Conditional on $Y_0 = i_0, \dots, Y_n = i_n$, we know that holding times $S_1, \dots, S_{n+1}$ are independent with $S_k \sim \exp(q_{i_{k-1}})$. Using this we get the left side to be

$$\int_{\Delta(t)} q_{i_n} \exp\left(-q_{i_n}\left(t - \sum_{i=1}^{n} s_i\right)\right) \prod_{k=1}^{n} q_{i_{k-1}} \exp(-q_{i_{k-1}} s_k) ds_k$$

where $\Delta(t) = \{(s_1, \dots, s_n) : \sum_{i=1}^n s_i \le t \text{ and } s_1, \dots, s_n \ge 0\}$. Now we make the following substitutions: $u_1 = t - \sum_{i=1}^n s_i$ and $u_k = s_{n-k+2}$ for $k = 2, \dots, n$, and find that the above expression can be rewritten as

$$\int_{\Delta(t)} q_{i_0} \exp\left(-q_{i_0}\left(t - \sum_{i=1}^{n} u_i\right)\right) \prod_{k=1}^{n} q_{i_{n-k+1}} \exp(-q_{i_{n-k+1}} u_k) du_k$$
$$= q_{i_o}\mathbb{P}(J_n \le t < J_{n+1}|Y_0 = i_n, \dots, Y_{n-1} = i_1, Y_0 = i_0)$$

□

We will now prove the forward equation.

Theorem 27. *The minimal non-negative solution $\{P(t)\}_{t\in\mathbb{R}_+}$ of the backward equation is also the minimal non-negative solution of the forward equation*

$$\frac{d}{dt}P(t) = P(t)Q, \quad P(0) = I.$$

Proof. Let $\{X_t\}_{t\in\mathbb{R}_+}$ be a (μ, Q) Markov chain. From the holding times and jump chain characterization we get

$$P_{ij}(t) = \mathbb{P}_i(X_t = j)$$
$$= \sum_{n=0}^{\infty} \sum_{k \ne j} \mathbb{P}_i(J_n \le t < J_{n+1}, Y_{n-1} = k, Y_n = j)$$

From Lemma 9 we get

$$\mathbb{P}_i(J_n \le t < J_{n+1}|Y_{n-1} = k, Y_n = j) = \frac{q_i}{q_j}\mathbb{P}_j(J_n \le t < J_{n+1}|Y_1 = k, Y_n = i)$$
$$= \frac{q_i}{q_j}\int_0^t q_j e^{-q_j s}\mathbb{P}_k(J_{n-1} \le t - s < J_n|Y_{n-1} = i)ds$$
$$= q_i \int_0^t e^{-q_j s}\frac{q_k}{q_i}\mathbb{P}_i(J_{n-1} \le t - s < J_n|Y_{n-1} = k)ds$$

where we use the Markov property of the jump chain in the second equality and Lemma 9 once again in the third equality. Using this we get

$$\begin{aligned}
&P_{ij}(t)\\
&\quad= \delta_i(j)e^{-q_i t} + \sum_{n=1}^{\infty}\sum_{k\neq j}\int_0^t \mathbb{P}_i(J_{n-1} \leq t-s < J_n | Y_{n-1} = k)\mathbb{P}_i(Y_{n-1} = k, Y_n = j)q_k e^{-q_j s} ds\\
&\quad= \delta_i(j)e^{-q_i t} + \sum_{n=1}^{\infty}\sum_{k\neq j}\int_0^t \mathbb{P}_i(J_{n-1} \leq t-s < J_n | Y_{n-1} = k)\pi_{kj} q_k e^{-q_j s} ds\\
&\quad= \delta_i(j)e^{-q_i t} + \int_0^t \sum_{k\neq j}\sum_{n=1}^{\infty} \mathbb{P}_i(J_{n-1} \leq t-s < J_n | Y_{n-1} = k)\pi_{kj} q_k e^{-q_j s} ds\\
&\quad= \delta_i(j)e^{-q_i t} + \int_0^t \sum_{k\neq j} P_{ik}(t-s) q_{kj} e^{-q_j s} ds
\end{aligned}$$

where we have used monotone convergence theorem to exchange the integral and the summations. Now set $u = t - s$ and multiply on both sides by $e^{q_i t}$ to get the integral form of the forward equation,

$$P_{ij}(t)e^{q_j t} = \delta_i(j) + \int_0^t \sum_{k\neq j} P_{ik}(u) q_{kj} e^{q_j u} du$$

From the proof of the backward equation we know that

$$e^{q_i t}P_{ij}(t) = \delta_i(j) + \int_0^t \sum_{k\neq i} P_{kj}(u) q_{ik} e^{q_i u} du$$

so that $e^{q_i t}P_{ik}(t)$ is increasing in t for all i, k. Therefore, only one of the following two statements is true,

$$\sum_{k\neq j} P_{ik}(u) q_{kj} \text{ converges uniformly for all } u \in [0,t]; \textit{or}$$

$$\sum_{k\neq j} P_{ik}(u) q_{kj} = \infty \text{ for all } u \geq t$$

From the forward equation and using the finiteness of $P_{ij}(t)e^{q_j t}$ for all $t \in \mathbb{R}_+$, only the first condition can hold. We also know from the backward equation that $P_{ij}(t)$ is continuous for all i, j. Thus, by uniform continuity, we can differentiate the forward equation to get

$$\begin{aligned}
e^{q_j t}\left(q_j P_{ij}(t) + \frac{d}{dt}P_{ij}(t)\right) &= \sum_{k\neq j} P_{ik}(t) q_{kj} e^{q_j t}; \text{ or}\\
\frac{d}{dt}P_{ij}(t) &= \sum_{k\in I} P_{ik}(t) q_{kj},
\end{aligned}$$

which is the forward equation. For the minimality part we refer to the book by J. Norris. □

4.3. **Class structure, recurrence and transience.** The jump chain and holding times characterization of a CTMC allows us to fallback on definitions used for the jump DTMC. Now we say that i leads to j, i.e., $i \to j$ if $\mathbb{P}_i(X_t = j \text{ for some } t \geq 0) > 0$, and (communication) $i \leftrightarrow j$ if $i \to j$ and $j \to i$. We have a theorem relating properties of the CTMC and the jump DTMC.

Theorem 28. *For distinct states i and j the following are equivalent:*

(1) *$i \to j$ in the CTMC;*
(2) *$i \to j$ in the jump DTMC;*
(3) *$q_{i_0 i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} > 0$ for some sequence of pairwise distinct states and n such that $i_0 = i$ and $i_n = j$;*
(4) *$P_{ij}(t) > 0$ for all $t \geq 0$; and*
(5) *$P_{ij}(t) > 0$ for some $t \geq 0$.*

We will not prove the result but just remark that this shows that the situation is simpler for a CTMC as there is no periodicity issue. Also, communication classes, closed classes and irreducibility follow exactly from the jump chain DTMC.

Let $A \subset I$. Then define the hitting time of A to be $D^A = \inf\{t \geq 0 : X_t \in A\}$ with the usual convention that $\inf \emptyset = \infty$. By our construction using jump chains and holding times it is easy to argue that $\{D^A < \infty\} = \{H^A < \infty\}$ where H^A is the hitting of set A in the jump DTMC, and on this set we have $D^A = J_{H^A}$. From this it is clear that $\mathbb{P}_i(D^A < \infty) = \mathbb{P}_i(H^A < \infty) = h_i^A$. Thus, using the result for DTMCs we have the following result.

Theorem 29. *The vector of hitting probabilities h^A is the minimal non-negative solution to the following system of linear equations*

$$h_i^A = 1 \text{ for } i \in A$$
$$\sum_{j \in I} q_{ij} h_j^A = 0 \text{ for } i \notin A.$$

Let us define the expected hitting time to be $k_i^A = \mathbb{E}_i[D^A]$; note that this will be different from $\mathbb{E}_i[H^A]$ for the jump DTMC as we need to account for the time spent in each state. This yields the following result.

Theorem 30. *Assume that $q_i > 0$ for all $i \notin A$. Then the vector of mean hitting times k^A is the minimal non-negative solution to the following system of linear equations*

$$k_i^A = 0 \text{ for } i \in A$$
$$-\sum_{j \in I} q_{ij} k_j^A = 1 \text{ for } i \notin A.$$

Proof. If $X_0 = i \in A$, then $D^A = 0$ and $k_i^A = 0$. If $X_0 = i \notin A$, then one has to wait till the chain jumps from state i so that $D^A \geq J_1$. Thus, by the Markov property of the jump DTMC

$$\mathbb{E}_i[D^A - J_1 | Y_1 = j] = \mathbb{E}_j[D^A],$$

which then leads to

$$k_i^A = \mathbb{E}_i[D^A] = \mathbb{E}_i[J_1] + \sum_{j \neq i} \mathbb{E}_i[D^A - J_1 | Y_1 = j] \mathbb{P}_i(Y_1 = j)$$
$$= q_i^{-1} + \sum_{j \neq i} \pi_{ij} k_j^A$$

Rearranging we get the result. The proof of minimality can be found in the book of J. Norris. □

We define a state i to be recurrent if

$$\mathbb{P}_i(\{t \geq 0 : X_t = i\} \text{ is unbounded}) = 1,$$

and to be transient if

$$\mathbb{P}_i(\{t \geq 0 : X_t = i\} \text{ is unbounded}) = 0.$$

Note that if the the process explodes starting from i, then i cannot be recurrent. We can relate recurrence and transience of a state i in the CTMC to the corresponding attribute of the state in the jump DTMC.

Theorem 31. *We have:*

1. *if i is recurrent for the jump chain, then i is recurrent for the CTMC;*
2. *if i is transient for the jump chain, then i is transient for the CTMC;*
3. *every state is either recurrent or transient; and*
4. *recurrence and transience are class properties.*

Proof. (i) Assume that i is recurrent for the jump DTMC, then we know that the CTMC is non-explosive, $J_{n_m} \to \infty$ and $X_{J_{m_n}} = Y_{m_n} = i$ infinitely often so that $\{t \geq 0 : X_t = i\}$ is unbounded w.p. 1.
(ii) Assume that i is transient for the jump DTMC. If $X_0 = i$, then $N = \sup\{n \geq 0; Y_n = i\} < \infty$ which then implies that $\{t \geq 0 : X_t = i\}$ is bounded by J_{N+1} which is finite (w.p. 1) because $\{Y_n\}_{n=0,1,\cdots,n}$ cannot include a state j with $q_j = 0$ (i.e. an absorbing state).
The remaining properties now follow from results about DTMCs. □

As for DTMCs we now define the first passage time of $\{X_t\}_{t \in \mathbb{R}_+}$ to state i as $T_i = \inf\{t \geq J_1 : X_t = i\}$. As was the case of DTMCs we have the following dichotomy.

Theorem 32. *Exactly one of the two statements holds*

1. *if $q_i = 0$ or $\mathbb{P}_i(T_i < \infty) = 1$, then i is recurrent and $\int_0^\infty P_{ii}(t)dt = \infty$;*
2. *if $q_i > 0$ and $\mathbb{P}_i(T_i < \infty) < 1$, then i is transient and $\int_0^\infty P_{ii}(t)dt < \infty$;*

Proof. If $q_i = 0$, then starting in i, the process X never leaves state i. Thus, i is recurrent and $P_{ii}(t) = 1$, which implies $\int_0^\infty P_{ii}(t)dt = \infty$. Therefore, we assume that $q_i > 0$. Let N_i denote the first passage time of the jump DTMC to state i. Then it is clear that $\mathbb{P}_i(N_i < \infty) = \mathbb{P}_i(T_i < \infty)$, so that i is recurrent if and only if $\mathbb{P}_i(T_i < \infty) = 1$ (using Theorem 31 and results about DTMCs).

Let $\Pi_{ij}^{(n)}$ be the (i,j) entry of Π^n. We have the following using Tonelli's theorem.

$$\begin{aligned}\int_0^\infty P_{ii}(t)dt &= \int_0^\infty \mathbb{E}_i[1_{X_t=i}]dt = \mathbb{E}_i\left[\int_0^\infty 1_{X_t=i}dt\right] \\ &= \mathbb{E}_i\left[\sum_{n=0}^\infty S_{n+1}1_{Y_n=i}\right] = \sum_{n=0}^\infty \mathbb{E}_i[S_{n+1}1_{Y_n=i}] \\ &= \sum_{n=0}^\infty \mathbb{E}_i[S_{n+1}|Y_n=i]\mathbb{P}_i(Y_n=i) = \frac{1}{q_i}\sum_{n=0}^\infty \Pi_{ii}^{(n)}\end{aligned}$$

Now the conclusions about $\int_0^\infty P_{ii}(t)dt$ follow from the DTMC results. □

We have the following result whose proof is left as exercise.

Theorem 33. *Let $s > 0$ be given and set $Z_n = X_{ns}$, then it follows that:*

(1) *if i is recurrent for X, then i is recurrent for $\{Z_n\}_{n\in\mathbb{Z}_+}$; and*
(2) *if i is transient for X, then i is transient for $\{Z_n\}_{n\in\mathbb{Z}_+}$.*

4.4. **Invariant distribution, convergence and ergodic theorem.** A non-negative vector λ is deemed an invariant measure for rate matrix Q if $\lambda Q = 0$. We have an important result relating invariant measures of the CTMC and the jump DTMC.

Theorem 34. *Given a rate matrix Q with jump chain transition matrix Π and a measure λ, then the following are equivalent:*

(1) *λ is invariant with respect to Q; and*
(2) *μ is invariant with respect to Π where $\mu_i = \lambda_i q_i$ for all $i \in I$.*

The proof is straightforward, but the conclusions are quite profound. Using results about DTMCs, we can then prove that the invariant measure for an irreducible and recurrent Q matrix is unique up to scalar multiples, i.e., fixing a given state i, there is only one invariant measure with $\lambda_i = 1$.

Similar to the DTMC case, we say that a state i is positive recurrent if either $q_i = 0$ or the expected return time $m_i = \mathbb{E}_i[T_i]$ is finite. If state i is recurrent but $q_i > 0$ and $\mathbb{E}_i[T_i] = \infty$, then we say that state i is null recurrent. We have the following theorem.

Theorem 35. *Let Q be an irreducible rate matrix, then the following are equivalent:*

(1) *every state is positive recurrent;*
(2) *some state i is positive recurrent; and*
(3) *Q is non-explosive and has an invariant distribution π.*

Moreover, when (3) holds, then $m_i = \frac{1}{\lambda_i q_i}$ for all i.

Proof. If I is not a singleton, then irreducibility forces $q_i > 0$ for all i. Obviously (1) implies (2) so we will start with proving that (2) implies (3). Define the vector μ^i as follows

$$\mu_j^i = \mathbb{E}_i \int_0^{T_i\wedge\zeta} 1_{X_s=j}ds,$$

which is the amount time spent in state j before the explosion or returning to state i. By monotone convergence, we have $\sum_{j\in I}\mu_j^i = \mathbb{E}_i[T_i\wedge\zeta]$.

Denote by N_i the first passage time of the jump DTMC to state i. By Fubini's theorem we have

$$\begin{aligned}
\mu^i_j &= \mathbb{E}_i\left[\sum_{n=0}^{\infty} S_{n+1} 1_{\{Y_n=j;n<N_i\}}\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_i[S_{n+1} 1_{\{Y_n=j;n<N_i\}}] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_i[S_{n+1}|Y_n=j]\mathbb{E}_i[1_{\{Y_n=j;n<N_i\}}] \\
&= \frac{1}{q_j}\sum_{n=0}^{\infty} \mathbb{E}_i[1_{\{Y_n=j;n<N_i\}}] \\
&= \frac{1}{q_j}\mathbb{E}_i\left[\sum_{n=0}^{N_i-1} 1_{Y_n=j}\right] = \frac{\gamma^i_j}{q_j}
\end{aligned}$$

where γ^i_j is the expected (discrete) time steps in j between visits to i for the jump DTMC. Since (2) holds, i is recurrent and hence, X is non-explosive. We know that $\gamma^\Pi = \gamma^i$ so that μ^i is invariant for Q. Note, however, that μ^i has finite total mass as $\sum_{j\in I}\mu^i_j = \mathbb{E}_i[T_i] = m_i$, which yields an invariant distribution λ by setting $\lambda_j = \frac{\mu^i_j}{m_i}$.

Proving that (3) implies (1) is also straight forward. As Q is irreducible, so is Π. Fix an $i \in I$ with $\lambda_i > 0$ and set $\nu_j = \frac{\lambda_j q_j}{\lambda_i q_i}$; then notice that $\nu_i = 1$ and ν is an invariant measure for Π. Using a result for DTMCs, we then have that $\nu_j \geq \gamma^i_j$ for all $j \in I$. Thus,

$$\begin{aligned}
m_i &= \sum_{j\in I}\mu^i_j = \sum_{j\in I}\frac{\gamma^i_j}{q_j} \\
&\leq \sum_{j\in I}\frac{\nu_i}{q_j} = \frac{\sum_{j\in I}\lambda_j}{\lambda_i q_i} = \frac{1}{\lambda_i q_i} < \infty
\end{aligned}$$

Therefore, i is positive recurrent, and hence recurrent. Since Q is irreducible, Q is recurrent and so is Π. Thus, ν is the only invariant measure for Π with $\nu_i = 1$ and irreducibility of Π implies that $\nu_j > 0$ for all j, which then means that $\lambda_j > 0$ for all j and the same argument shows that every j is positive recurrent. □

Important: The non-explosiveness is critical. Consider the following CTMC on $\mathbb{Z}_+$ with $q_i > 0$ and we have

$$q_{ij} = \begin{cases} q_0 & \text{if } i = 0 \ \&\ j = 1 \\ pq_i & \text{if } i > 0 \ \&\ j = i+1 \\ (1-p)q_i & \text{if } i > 0 \ \&\ j = i-1 \end{cases}$$

where $0 < p < 1$. For this chain an invariant measure is given by the following formula: $\nu_0 = \frac{p}{q_0}$ and $\nu_i = \frac{1}{q_i}(\frac{p}{1-p})^i$. If q_i is constant, then when $p < 1/2$ (i.e., $p < 1-p$) suffices to produce an invariant distribution for the chain; note that the chain is non-explosive. On the other hand if $q_i = 2^i$ and $1 < \frac{p}{1-p} < 2$, then one can once again produce an invariant distribution. However, it is easy to show that X is transient. Therefore, X has to be explosive. We will revisit this example in the exercises.

Now we set about showing that a measure λ that solves $\lambda Q = 0$ is invariant.

Theorem 36. *Let Q be irreducible and recurrent, and let λ be a measure. Let $s > 0$ be given. Then the following are equivalent:*

(1) $\lambda Q = 0$*; and*
(2) $\lambda P(s) = \lambda$.

Proof. The finite state case is straightforward using the backward equation as

$$\frac{d}{ds}\lambda P(s) = \lambda \frac{d}{ds} P(s) = \lambda Q P(s),$$

so $\lambda Q = 0$ implies that $\lambda P(s) = \lambda P(0) = \lambda$ for all s. Since $P(s)$ is irreducible and recurrent, any μ that is invariant has to be a scalar multiple of λ. From the forward equation (or from the fact that $P(s) = e^{Qs}$) we know that Q and $P(s)$ commute, and therefore

$$0 = \frac{d}{ds}\lambda = \frac{d}{ds}\lambda P(s) = \lambda Q P(s) = \lambda P(s) Q = \lambda Q.$$

For the countable state-space case, we have to argue differently. Since Q is recurrent, it is non-explosive and $P(s)$ is also recurrent and irreducible. Therefore, any λ satisfying (1) or (2) is unique up to scalar multiples. From the proof of Theorem 35, fixing i we know that if we set

$$\mu^i_j = \mathbb{E}_i \int_0^{T_i} 1_{X_t = j} dt,$$

then $\mu^i Q = 0$. We will now show that $\mu^i P(s) = \mu^i$. By the strong Markov property at T_i (which is a consequence of the strong Markov property of the jump DTMC), we get

$$\mathbb{E}_i \left[\int_0^s 1_{X_t = j} dt \right] = \mathbb{E}_i \left[\int_{T_i}^{T_i + s} 1_{X_t = j} dt \right]$$

Therefore, using Fubini's theorem,

$$\begin{aligned}
\mu^i_j &= \mathbb{E}_i \left[\int_s^{T_i + s} 1_{X_t = j} dt \right] \\
&= \mathbb{E}_i \left[\int_0^{\infty} 1_{\{X_{s+t} = j, t < T_i\}} dt \right] \\
&= \int_0^{\infty} \mathbb{P}_i(X_{s+t} = j, t < T_i) dt \\
&= \int_0^{\infty} \sum_{k \in I} \mathbb{P}_i(X_t = k, t < T_i) P_{kj}(s) dt \\
&= \sum_{k \in I} P_{kj}(s) \mathbb{E}_i \left[\int_0^{T_i} 1_{X_t = k} dt \right] \\
&= \sum_{k \in I} \mu^i_k P_{kj}(s)
\end{aligned}$$

Note that this finishes the proof. □

We have an additional theorem whose proof is quite elementary.

Theorem 37. *Let Q be an irreducible non-explosive rate matrix with an invariant distribution λ. If $\{X_t\}_{t\in\mathbb{R}_+}$ is (λ, Q) Markov, then so is $\{X_{s+t}\}_{t\in\mathbb{R}_+}$ for any $s \in \mathbb{R}_+$.*

Convergence to equilibrium: Since we have no periodicity, we expect a simpler result to hold. First we prove a lemma about the uniform continuity of the transition probabilities $P_{ij}(t)$.

Lemma 10. *Let Q be a rate matrix with semigroup $P(t)$, then for all $t, h \in \mathbb{R}_+$ we have*

$$|P_{ij}(t+h) - P_{ij}(t)| \leq 1 - e^{-q_i h}.$$

Proof. Note that

$$\begin{aligned}
|P_{ij}(t+h) - P_{ij}(t)| &= \left| \sum_{k\in I} P_{ik}(h)P_{kj}(t) - P_{ij}(t) \right| \\
&= \left| \sum_{k\neq i} P_{ik}(h)P_{kj}(t) - (1-P_{ii}(h))P_{ij}(t) \right| \\
&\leq \max\left(\sum_{k\neq i} P_{ik}(h)P_{kj}(t), (1-P_{ii}(h))P_{ij}(t) \right) \\
&\leq 1 - P_{ii}(h) \leq \mathbb{P}_i(J_1 \leq h) = 1 - e^{-q_i h}
\end{aligned}$$

□

The convergence theorem is as follows.

Theorem 38. *Let Q be an irreducible non-explosive rate matrix with semigroup $P(t)$ and invariant distribution λ, then for all states i, j we have $\lim_{t\to\infty} P_{ij}(t) = \lambda_j$.*

Proof. Let $\{X_t\}_{t\in\mathbb{R}_+}$ be (δ_i, Q) Markov. Fix $h \geq 0$ and consider the h-skeleton DTMC $\{Z_n\}_{n\in\mathbb{Z}_+}$ where $Z_n = X_{nh}$. From the characterization of the Markov process we have

$$\mathbb{P}(Z_{n+1} = i_{n+1} | Z_0 = i_0, Z_1 = i_1, \ldots, Z_n = i_n) = P_{i_n i_{n+1}}(h),$$

so that Z is $(\delta_i, P(h))$ Markov. Now irreducibility implies that $P_{ij}(h) > 0$ for all i, j, i.e., $P(h)$ is aperiodic and λ is also the invariant distribution of $P(h)$. Thus, by the convergence theorem for DTMCs we get that $\lim_{n\to\infty} P_{ij}(h) = \lambda_j$ for all i, j.

Next we will use the uniform continuity from Lemma 10 to prove the general result. First, we fix a state i. Given $\epsilon > 0$ we can find an h such that

$$1 - e^{-q_i h} \leq \frac{\epsilon}{2} \quad \forall\ s \in [0, h],$$

then we can find an N such that

$$|P_{ij}(nh) - \lambda_j| \leq \frac{\epsilon}{2} \quad \forall\ n \geq N.$$

For $t \geq Nh$ we have $nh \leq t < (n+1)h$ for some $n \geq N$ and so

$$|P_{ij}(t) - \lambda_j| \leq |P_{ij}(t) - P_{ij}(nh)| + |P_{ij}(nh) - \lambda_j| \leq \epsilon,$$

which proves our result. □

Time reversal: We present results for this section without proof and gloss over some technicalities so that the reversed process is still an *rcll* process.

Theorem 39. *Let Q be an irreducible non-explosive rate matrix with invariant distribution λ. Given $T \in (0,\infty)$, let $\{X_t\}_{0\le t\le T}$ be (λ, Q) Markov. Set $\hat{X}_t = X_{T-t}$. Then the process $\{\hat{X}_t\}_{0\le t\le T}$ is $(\lambda, \hat{Q})$ Markov, where $\hat{Q}$ is given by $\hat{q}_{ji} = \frac{\lambda_i}{\lambda_j} q_{ij}$. In addition, $\hat{Q}$ is irreducible and non-explosive with invariant distribution λ.*

The $\hat{X}$ process is called the time-reversal of X. A rate matrix Q and a measure λ are said to be in detailed balance if $\lambda_i q_{ij} = \lambda_j q_{ji}$ for all i, j; it then follows that λ is invariant for Q. Finally, we have the following result.

Theorem 40. *Let Q be an irreducible non-explosive rate matrix with invariant distribution λ. Let $\{X_t\}_{t\in\mathbb{R}_+}$ be (λ, Q) Markov. Then the following are equivalent:*

(1) *X is reversible, i.e., for all $T > 0$, $\hat{X}$ is also (λ, Q) Markov;*
(2) *Q and λ are in detailed balance.*

Ergodic theorem: We will now prove the final result that parallels the ergodic theorem for DTMCs.

Theorem 41. *Let Q be irreducible and let μ be any distribution. If $\{X_t\}_{t\in\mathbb{R}_+}$ is (μ, Q) Markov, then*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t 1_{X_s=i} ds = \frac{1}{m_i q_i} \quad a.s.$$

where $m_i = \mathbb{E}_i[T_i]$ is the expected return time to state i. Moreover, if Q is positive recurrent, then for any bounded function $f : I \mapsto \mathbb{R}$ we have

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X_s) ds = \bar{f} \left(=: \sum_{i\in I} \lambda_i f_i \right) \quad a.s.$$

where λ is the unique invariant distribution.

Proof. As before if Q is transient, then the total time spent in state i is finite, so

$$\frac{1}{t}\int_0^t 1_{X_s=i} ds \le \frac{1}{t}\int_0^\infty 1_{X_s=i} ds \to_{t\to\infty} 0 = \frac{1}{m_i q_i}.$$

Henceforth, we will assume that Q is recurrent. Fix a state i. Since the first passage time to i is finite *a.s.*, the long-run proportion of time in state i is the same even if we start accounting for it after the first time one reaches state i. Thus, by the strong Markov property it is sufficient to consider $\mu = \delta_i$.

Denote by M_i^n the length of the n^{th} visit to i, by T_i^n the time of the n^{th} return to i and by L_i^n the length of the n^{th} excursion to i. Thus, for $n \in \mathbb{Z}_+$ we have

$$\begin{aligned} M_i^{n+1} &= \inf\{t > T_i^n : X_t \ne i\} - T_i^n \\ T_i^{n+1} &= \inf\{t > T_i^n + M_i^{n+1} : X_t = i\} \\ L_i^{n+1} &= T_i^{n+1} - T_i^n. \end{aligned}$$

By the strong Markov property (of the jump chain) at the stopping times T_i^n for $n \geq 0$, we can deduce that $\{L_i^n\}_{n\in\mathbb{N}}$ are *i.i.d.* with mean m_i, and $\{M_i^n\}_{n\in\mathbb{N}}$ are *i.i.d.* $\exp(q_i)$. Hence, by the strong law of large numbers we have

$$\lim_{n\to\infty} \frac{\sum_{m=1}^n L_i^m}{n} = m_i \quad a.s.$$
$$\lim_{n\to\infty} \frac{\sum_{m=1}^n M_i^m}{n} = \frac{1}{q_i} \quad a.s.$$

and hence

$$\lim_{n\to\infty} \frac{\sum_{m=1}^n M_i^m}{\sum_{m=1}^n L_i^m} = \frac{1}{m_i q_i} \quad a.s.$$

As a consequence, note that $\lim_{n\to\infty} T_i^n/T_i^{n+1} = 1$ *a.s.* Now for $T_i^n \leq t < T_i^{n+1}$ we have

$$\frac{T_i^n}{T_i^{n+1}} \frac{\sum_{m=1}^n M_i^m}{\sum_{m=1}^n L_i^m} \leq \frac{1}{t}\int_0^t 1_{X_s=i} ds \leq \frac{T_i^{n+1}}{T_i^n} \frac{\sum_{m=1}^{n+1} M_i^m}{\sum_{m=1}^{n+1} L_i^m},$$

so letting $t \to \infty$ (which makes $n \to \infty$) we have,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t 1_{X_s=i} ds = \frac{1}{m_i q_i} \quad a.s.$$

In the positive recurrent case we can write

$$\frac{1}{t}\int_0^t f(X_s)ds - \bar{f} = \sum_{i\in I} f(i) \left(\frac{1}{t}\int_0^t 1_{X_s=i} ds - \lambda_i\right)$$

where $\lambda_i = \frac{1}{m_i q_i}$. We then have the result using the same argument as in the DTMC case. □

Note once again that the null recurrence case is also covered, i.e., T_i being finite *a.s.* but $m_i = \infty$ so that $\frac{1}{m_i} = 0$ (by convention).

4.5. **Exercises.** Please show all your work.

(1) *Forward-Backward equations*: Solve the forward and backward equations for the Poisson process. Show that the solutions are the same and yield a semi-group.

(2) *Discrete-time sampling*: Prove Theorem 33.

(3) *Positive and null recurrence of CTMC and jump DTMC*: From Theorem 34 we know that an invariant measure for a CTMC yields an invariant measure for the jump DTMC and *vice-versa*. Show by example that we can have any of the four possibilities in terms of the CTMC and DTMC being positive/null recurrent. Use birth-death chains for the examples, i.e., CTMC on $\mathbb{Z}_+$ with $q_i > 0$ and where we have

$$q_{ij} = \begin{cases} q_0 & \text{if } i = 0 \ \& \ j = 1 \\ p_i q_i & \text{if } i > 0 \ \& \ j = i+1 \\ (1-p_i)q_i & \text{if } i > 0 \ \& \ j = i-1 \end{cases}$$

where $0 < p_i < 1$ for all $i \in \mathbb{N}$.

(4) *Explosiveness and stationary distributions*: Consider example after Theorem 35. Verify irreducibility of Q. Prove that ν is an invariant measure. When $q_i \equiv$constant and $p < 1/2$, show that chain is non-explosive and ν can be normalized to produce an invariant distribution; what does Theorem 35 tell us now? Now consider case of $q_i = 2^i$ and $1 < \frac{p}{1-p} < 2$. Prove that Q is explosive and also transient.

5. Criteria for Positive Recurrence, Transience and Kurtz's Theorem

We will start with the Levy martingales associated with a DTMC, and also mention the same for CTMC. Then we will present conditions for recurrence and transience using martingales, ending with the Foster-Lyapunov criterion. Finally, we will prove Kurtz's theorem that associates a differential equation with a class of CTMCs.

The sources for this chapter are:

- J. R. Norris, "Markov chains," Cambridge Series in Statistical and Probabilistic Mathematics, 2, Cambridge University Press, Cambridge, 2006.
- P. Brémaud, "Markov chains: Gibbs fields, Monte Carlo simulation, and queues," Texts in Applied Mathematics, 31, Springer-Verlag, New York, 1999.
- R. Durrett, "Probability: Theory and examples," Second edition. Duxbury Press, Belmont, CA, 1996.
- Lecture Notes on Stochastic Stability by Lyapunov Methods, S. Foss and T. Konstantopoulos, LMS/EPSRC Short Course Stability, Coupling Methods and Rare Events, Heriot-Watt University, Edinburgh, 4-9 September 2006. `http://www2.math.uu.se/~takis/L/StabLDC06/notes/SS_LYAPUNOV.pdf`
- Lecture Notes on Communication Networks Analysis, Chapter 2, B. Hajek, December 2006. `http://www.ifp.illinois.edu/~hajek/Papers/networkanalysisDec06.pdf`
- S. P. Meyn and R. L. Tweedie, "Markov chains and stochastic stability," Second edition, Cambridge University Press, Cambridge, 2009.
- S. Asmussen, "Applied probability and queues," Second edition, Applications of Mathematics (New York), 51, Stochastic Modelling and Applied Probability, Springer-Verlag, New York, 2003.

5.1. **Martingales and Markov chains.** All the proofs in this section will be for the discrete-time case, mainly to steer clear of measure-theoretic issues.

Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be a (μ, P) DTMC. For a function $f : I \mapsto \mathbb{R}$ we define the following

$$(P^n f)(i) = \mathbb{E}_i[f(X_n)] = \sum_{j\in I} P_{ij}^{(n)} f_j \quad \forall i \in I, n \geq 0$$

Then we have the following important characterization of a DTMC.

Theorem 42. *[Levy martingale] Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be an I-valued random process and let P be a stochastic matrix. Denote the natural filtration by $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$. Then the following are equivalent:*

1. *$\{X_n\}_{n\in\mathbb{Z}_+}$ is a (μ, P) DTMC;*
2. *for all bounded functions $f : I \mapsto \mathbb{R}$, the following process is a $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$ martingale that is null at 0:*

$$M_n^f := f(X_n) - f(X_0) - \sum_{m=0}^{n-1} (Pf)(X_m) - f(X_m) = f(X_n) - f(X_0) - \sum_{m=0}^{n-1} (P - \mathbf{I}) f(X_m)$$

where $\mathbf{I}$ is the identity mapping.

Proof. Assume that (1) holds. Since f is a bounded function, we have

$$|(Pf)(i)| = \left|\sum_{j\in I} P_{ij} f_j\right| \le \sup_j |f_j| \ \Rightarrow\ |M_n^f| \le 2(n+1)\sup_j |f_j| < \infty,$$

proving the integrability requirement.

Let $A_n = \{X_0 = i_0, \dots, X_n = i_n\}$; note that $A \in \mathcal{F}_n$, and in fact the collection of all such sets is a π-system that generates $\mathcal{F}_n$. The Markov property of $\{X_n\}_{n\in\mathbb{Z}_+}$ now implies that

$$\mathbb{E}[f(X_{n+1})|A_n] = \mathbb{E}_{i_n}[f(X_1)] = (Pf)(i_n),$$

which then implies that

$$\mathbb{E}[M_{n+1}^f - M_n^f|A_n] = \mathbb{E}[f(X_{n+1}) - (Pf)(X_n)|A_n] = 0,$$

and since sets like A_n generate $\mathcal{F}_n$, $\{M_n^f\}_{n\in\mathbb{Z}_+}$ is a martingale.

Now assume that (2) holds, i.e., for every bounded function f we have

$$\mathbb{E}[M_{n+1}^f - M_n^f|X_0 = i_0, \dots, X_n = i_n] = \mathbb{E}[M_{n+1}^f - M_n^f|A_n] = 0.$$

Take f to be $1_{i_{n+1}}$ and note that we get

$$\mathbb{P}(X_{n+1} = i_{n+1}|X_0 = i_0, \dots, X_n = i_n) = P_{i_n i_{n+1}},$$

which proves that $\{X_n\}_{n\in\mathbb{Z}_+}$ is Markov with transition matrix P. □

We now provide another set of martingales by dropping the boundedness requirement on f as follows.

Theorem 43. *Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be a (μ, P) DTMC. Suppose we have a function $f : \mathbb{Z}_+ \times I \mapsto \mathbb{R}$ that, for all $n \in \mathbb{Z}_+$, satisfies both*

$$\mathbb{E}[f(n, X_n)] < \infty \text{ and } (Pf)(n+1, i) := \sum_{j\in I} P_{ij} f(n+1, j) = f(n, i)\ \forall n \in \mathbb{Z}_+ \& i \in I,$$

then $\{M_n\}_{n\in\mathbb{Z}_+}$ such that $M_n = f(n, X_n)$ is a martingale (with the natural filtration of $\{X_n\}_{n\in\mathbb{Z}_+}$).

The proof is the same as that of the first part of Theorem 42 so we skip it. Theorem 43 is a very useful result. Consider the case where $\{S_n\}_{n\in\mathbb{Z}_+}$ with $S_0 = 0$ is the partial sums process of $\{X_n\}_{n\in\mathbb{N}}$ an *i.i.d.* sequence taking values $\{-1, 1\}$ with equal probability. We know that $\{S_n\}_{n\in\mathbb{Z}_+}$ is a martingale. Consider two functions $f(i) = i$ and $g(n, i) = i^2 - n$. Since $|S_n| \le n$ (at best n +1s or −1s) for all n, we get $\max(\mathbb{E}[|f(S_n)|], \mathbb{E}[|g(n, S_n)|] < \infty$. It is easy to verify that

$$(Pf)(i) = (i-1)/2 + (i+1)/2 = i = f(i),$$
$$(Pg)(n+1, i) = (i-1)^2/2 + (i+1)^2/2 - (n+1) = i^2 - n = g(n, i),$$

which then implies that both $\{f(S_n)\}_{n\in\mathbb{Z}_+}$ and $\{g(n, S_n)\}_{n\in\mathbb{Z}_+}$ are martingales. Consider the stopping time $T = \inf\{n \ge 0 : X_n = -a \text{ or } X_n = b\}$ for $a, b \in \mathbb{N}$. Now Doob's optional sampling theorem A says that

$$0 = \mathbb{E}[X_0] = \mathbb{E}[X_{T\wedge n}]$$
$$0 = \mathbb{E}[Y_0] = \mathbb{E}[Y_{T\wedge n}] = \mathbb{E}[X_{T\wedge n}^2] - \mathbb{E}[T \wedge n].$$

Therefore, setting p to be the probability of hitting $-a$ before b, letting $n \to \infty$ we get $\mathbb{E}[T \wedge n] \to \mathbb{E}[T]$ (monotone convergence) and $\mathbb{E}[X_{T\wedge n}] \to \mathbb{E}[X_T]$ and $\mathbb{E}[X_{T\wedge n}^2] \to \mathbb{E}[X_T^2]$

(bounded convergence); equivalently by Doob's optional sampling theorem B (*a.s.* finiteness of T is easy to show). Using this we get

$$0 = \mathbb{E}[X_T] = b - p(a+b), \text{ and}$$
$$\mathbb{E}[T] = \mathbb{E}[X_T^2] = a^2 p + b^2(1-p) = ab$$

For a CTMC we get a similar result as Theorem 42 that we present without proof.

Theorem 44. *Let $\{X_t\}_{t\in\mathbb{R}_+}$ be a random process with values in I and let Q be a rate matrix. Then the following are equivalent*

1. $\{X_t\}_{t\in\mathbb{R}_+}$ *is* (μ, Q) *Markov;*
2. *for every bounded function* $f : I \mapsto \mathbb{R}$ *the following random process* $\{M_t^f\}_{t\in\mathbb{R}_+}$ *given by*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \left[\sum_{j\in I} q_{X_s j}\Big(f(j) - f(X_s)\Big)\right] ds$$

is a martingale (null at 0*) with the natural filtration of* $\{X_t\}_{t\in\mathbb{R}_+}$.

The operator A that takes bounded functions to bounded functions defined by

$$(Af)(i) := \sum_{j\in I} q_{ij}\Big(f(j) - f(i)\Big) = \sum_{j\neq i} q_{ij}\Big(f(j) - f(i)\Big) = (Qf)(i)$$

is called the generator of the CTMC X; a similar definition holds for DTMCs where we will get $P - \mathbf{I}$, note that it has all the properties of a rate matrix.

5.1.1. *Harmonic functions and martingales.* Given a stochastic matrix P and a function $h : I \mapsto \mathbb{R}$, we call h harmonic if $Ph = h$, subharmonic if $Ph \geq h$ and superharmonic if $Ph \leq h$. Note that our previous results show that given $\{X_n\}_{n\in\mathbb{Z}_+}$ a (μ, P) DTMC, if $h(X_n)$ is integrable for all n or if h is non-negative, then $\{h(X_n)\}_{n\in\mathbb{Z}_+}$ is either a martingale, a submartingale or a supermartingale. Since finite state DTMCs with irreducible stochastic matrices are necessarily positive recurrent, from now onwards we will assume that the state-space is necessarily countably infinite.

We then have the following important result.

Theorem 45. *An irreducible and recurrent stochastic matrix P has no non-negative superharmonic or bounded subharmonic functions besides the constant functions.*

Proof. Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be a (μ, P) DTMC. If h is a non-negative superharmonic (resp., bounded subharmonic) function, then $\{h(X_n)\}_{n\in\mathbb{Z}_+}$ is a non-negative supermartingale (resp., bounded submartingale). Therefore, by the martingale convergence theorem, the process converges to a finite limit, say Y. Since P is irreducible and recurrent, $\{X_n\}_{n\in\mathbb{Z}_+}$ visits any state $i \in I$ infinitely often, which then implies that $Y = h(i)$ almost surely for all $i \in I$. Thus, the only way this holds is if h is a constant. □

5.2. **Transience and recurrence conditions.** We start with a theorem that provides conditions for transience. It is a consequence of Theorem 45.

Theorem 46. *A necessary and sufficient condition for an irreducible stochastic matrix P to be transient is the existence of some state (conventionally called 0) and of a bounded function $h : I \mapsto \mathbb{R}$, not identically null and satisfying*

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

Proof. Let T_0 be the return time to 0, i.e., the first passage time to 0. Consider the function $h_j = \mathbb{P}_j(T_0 = \infty)$. This is a bounded function and it can be easily verified that it satisfies the above recursion. If P is transient, then h_j is non-trivial which proves the necessity of the condition.

For the converse, assume that h is a not identically null and bounded function that satisfies the above recursion. Define another function $\tilde{h}$ as follows: $\tilde{h}(0) = 0$ and $\tilde{h}(j) = h(j)$ for all $j \neq 0$. Let $\alpha = \sum_{i \in I} P_{0i} \tilde{h}(i)$. Without loss of generality (eg. by changing signs if necessary), we can assume that $\alpha \geq 0$. Then it is easy to see that $\tilde{h}$ is a subharmonic function (bounded by construction). Now, if P were to be recurrent, then by Theorem 45 we would need $\tilde{h}$ to be a constant, i.e., $\tilde{h}(i) \equiv 0$ for all $i \in I$. However, this contradicts the non-triviality assumption of h. □

Next we present a necessary and sufficient condition for recurrence.

Theorem 47. *Let P be an irreducible stochastic matrix and suppose there exists a function $h : I \mapsto R$ such that $\{i : h(i) < K\}$ is finite for all finite K, and such that*

$$\sum_{k \in I} P_{ik} h(k) \leq h(i) \quad \forall i \notin F,$$

for some finite subset $F \subset I$. Then P is recurrent.

Proof. Since $\{i : h(i) < 0\}$ is finite, $\inf h(i) > -\infty$ and therefore, by adding a constant if necessary, we can assume without loss of generality that h is non-negative. Let τ_F be the return time to F, i.e., first passage time to F, and define $Y_n = h(X_n) 1_{\{n < \tau_F\}}$. First note that $1_{\{n<\tau_F\}} \geq 1_{\{n+1<\tau_F\}}$ so $Y_{n+1} \leq h(X_{n+1}) 1_{\{n<\tau_F\}}$. Then noting that τ_F is a stopping time and the Markov property we get for $i \notin F$

$$\begin{aligned} E_i[Y_{n+1}|X_0, \dots, X_n] &\leq \mathbb{E}_i[h(X_{n+1}) 1_{\{n<\tau_F\}} | X_0, \dots, X_n] \\ &= 1_{\{n<\tau_F\}} \mathbb{E}_i[h(X_{n+1})|X_0, \dots, X_n] = 1_{\{n<\tau_F\}} \mathbb{E}_i[h(X_{n+1})|X_n] \\ &= 1_{\{n<\tau_F\}} \sum_{k \in I} P_{X_n k} h(k) \leq 1_{\{n<\tau_F\}} h(X_n) = Y_n, \end{aligned}$$

where all relationships hold *a.s.* (strictly speaking $\mathbb{P}_i$-*a.s.*). Therefore, $\{Y_n\}_{n \in \mathbb{Z}_+}$ is a non-negative supermartingale. By the martingale convergence theorem, $\lim_{n\to\infty} Y_n = Y$ exists *a.s.* and is finite.

Now assume that P is transient. Then any finite subset of I must be visited only a finite number of times. Therefore, for any K, $\{n \in \mathbb{Z}_+ : h(X_n) < K\}$ is finite. This then implies that $\lim_{n\to\infty} h(X_n) = \infty$ *a.s.* (strictly speaking $\mathbb{P}_j$-*a.s.* for any $j \in I$). For this to be compatible with Y being $\mathbb{P}_i$-*a.s.* finite for all $i \notin F$, we must have $\mathbb{P}_i(\tau_F < \infty) = 1$ for all $i \notin F$. Since F is a finite set, some state in F must be recurrent, yielding a contradiction. □

While we proved sufficiency of the condition, this is also necessary for recurrence. We omit the proof by pointing the reader to Theorem 9.4.2 in Meyn and Tweedie book.

We now present sufficient conditions for transience and for absence of positive recurrence.

Theorem 48. *Let P be an irreducible stochastic matrix and let $h : I \mapsto \mathbb{R}$ be a function such that*

$$\mathbb{E}_i[h(X_1)] = (Ph)(i) = \sum_{k \in I} P_{ik} h(k) \leq h(i) \quad \forall i \notin F,$$

for some set F (not assumed to be finite). Moreover, suppose there exists $i \notin F$ such that

$$h(i) < h(j) \quad \forall j \in F.$$

Then: (i) if h is bounded, then P is transient; and (ii) if F is a finite set, h is bounded above and

$$\sum_{k \in I} P_{jk} |h(k) - h(j)| \leq A \quad \forall j \in I,$$

for some $A < \infty$, then the chain cannot be positive recurrent (i.e., it is either null recurrent or transient).

Proof. First we prove (i). Let τ_F be the return/first passage time to F and pick an $i \notin F$ that satisfies $h(i) < h(j)$ for all $j \in F$. Setting $Y_n = h(X_{n \wedge \tau_F})$ we have

$$\mathbb{E}_i[Y_{n+1}|X_0, \ldots, X_n] = \mathbb{E}_i[1_{\{n<\tau_F\}} h(X_{n+1})|X_0, \ldots, X_n] + \mathbb{E}_i[1_{\{n \geq \tau_F\}} h(X_{\tau_F})|X_0, \ldots, X_n]$$

The term inside the second (conditional) expectation is $1_{\{n \geq \tau_F\}} h(X_{\tau_F}) = 1_{\{n \geq \tau_F\}} Y_n$; now note that $1_{\{n \geq \tau_F\}} h(X_{\tau_F})$ is $\sigma(X_0, \ldots, X_n)$ measurable. Therefore, we have

$$\begin{aligned} \mathbb{E}_i[Y_{n+1}|X_0, \ldots, X_n] &= 1_{\{n<\tau_F\}} \mathbb{E}_i[h(X_{n+1})|X_0, \ldots, X_n] + 1_{\{n \geq \tau_F\}} h(X_{\tau_F}) \\ &\leq 1_{\{n<\tau_F\}} h(X_n) + 1_{\{n \geq \tau_F\}} h(X_{\tau_F}) = h(X_{n \wedge \tau_F}) = Y_n \end{aligned}$$

Thus, under $\mathbb{P}_i$ we get that $\{Y_n\}_{n \in \mathbb{Z}_+}$ is a bounded supermartingale. By the martingale convergence theorem, the limit Y of Y_n exists and is finite ($\mathbb{P}_i$-*a.s.*). Thereafter, an application of the bounded convergence theorem then says that $\mathbb{E}_i[Y] = \lim_{n \to \infty} \mathbb{E}_i[Y_n]$, and since $\mathbb{E}_i[Y_n] \leq \mathbb{E}_i[Y_0] = h(i)$ (by supermartingale property), we have $\mathbb{E}_i[Y] \leq h(i)$.

Now, if τ_F were $\mathbb{P}_i$-*a.s.* finite, then Y_n would eventually be stuck at $h(j)$ for some $j \in F$, and therefore by the definition of i, we would have $\mathbb{E}_i[Y] > h(i)$, which would contradict the previous inequality. Therefore, $\mathbb{P}_i(\tau_F < \infty) < 1$, which concludes the proof of (i).

For (ii), we chose $j \in F$ such that $\alpha := \mathbb{P}_j(\tau_i < \tau_F) > 0$, where i is the special state identified earlier. Note that such a j has to exist by irreducibility of P and finiteness of F. Then we have $\mathbb{E}_j[\tau_j] \geq \mathbb{E}_j[\tau_F] \geq \alpha(\mathbb{E}_j[\tau_i] + \mathbb{E}_i[\tau_F]) \geq \alpha \mathbb{E}_i[\tau_F]$ so that it suffices to show that $\mathbb{E}_i[\tau_F] = \infty$. Since h is bounded above, we can assume without loss of generality that $h \leq 0$ (by subtracting the upperbound, if necessary) so that the required integrability holds.

Assume that $\mathbb{E}_i[\tau_F] < \infty$. Then, it follows that $\mathbb{P}_i(\tau_F < \infty) = 1$ and by condition on weighted sum of differences of h, we get

$$\mathbb{E}_i\left[\sum_{n=1}^{\tau_F} |Y_n - Y_{n-1}|\right] = \mathbb{E}_i\left[\sum_{n=1}^{\infty} 1_{\{\tau_F \geq n\}} \mathbb{E}_i\Big[|Y_n - Y_{n-1}|\Big| X_0, \ldots, X_{n-1}\Big]\right] \leq A\mathbb{E}_i[\tau_F] < \infty,$$

where $Y_n = h(X_{n\wedge\tau_F})$, as in the first part. Thus, by Fubini we can interchange summation and expectation to get

$$\mathbb{E}_i[Y_{\tau_F}] = \mathbb{E}_i[Y_0] + \mathbb{E}_i\left[\sum_{n=1}^{\tau_F} Y_n - Y_{n-1}\right] = h(i) + \sum_{n=1}^{\infty} \mathbb{E}_i[(Y_n - Y_{n-1})1_{\{\tau_F \geq n\}}]$$
$$= h(i) + \sum_{n=1}^{\infty} \mathbb{E}_i\left[1_{\{\tau_F\geq n\}}\mathbb{E}_i\Big[Y_n - Y_{n-1}\Big|X_0,\ldots,X_{n-1}\Big]\right] \leq h(i),$$

using the property that $(Ph)(k) \leq h(k)$ for all $k \notin F$ (since the first time hit F is at time τ_F). This contradicts $Y_{\tau_F} > h(i)$. Hence, $\mathbb{E}_i[\tau_F] = \infty$. □

For many results to follow we will view a given chain only when it is in a given finite set F. Let τ_k^F be the time of the k^{th} return of the given DTMC $\{X_n\}_{n\in\mathbb{Z}_+}$ to F with $\tau_F = \tau_1^F$. We define an associated process $\{X_k^F\}_{k\in\mathbb{Z}_+}$ by setting $X_k^F = X_{\tau_k^F}$ (where we also assume that $X_0^F \in F$ so that $\tau_0^F = 0$). If the original chain is irreducible and recurrent, then $\{X_k^F\}_{k\in\mathbb{Z}_+}$ is a DTMC by the strong Markov property, with a somewhat complicated transition matrix. Nevertheless, it is easy to reason that $\{X_n\}_{n\in\mathbb{Z}_+}$ being irreducible and recurrent implies that $\{X_k^F\}_{k\in\mathbb{Z}_+}$ is also irreducible and positive recurrent. This then yields a criterion for positive recurrence of $\{X_n\}_{n\in\mathbb{Z}_+}$ that is very useful.

Lemma 11. *Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be irreducible and $F \subset I$ be finite. Then $\{X_n\}_{n\in\mathbb{Z}_+}$ is positive recurrent if $\mathbb{E}_i[\tau_F] < \infty$ for all $i \in F$.*

Proof. For each $i \in F$ define $N_i = \inf\{k \geq 1 : X_k^F = i\}$, i.e., the first passage time in the restricted DTMC. Also set $L_k = \tau_k^F - \tau_{k-1}^F$ for $k \in \mathbb{N}$. Then with $m = \max_{j\in F}\mathbb{E}_j[\tau_F] < \infty$ we have for $i \in F$ (by the strong Markov property) that

$$\mathbb{E}_i[\tau_i] = \mathbb{E}_i\left[\sum_{k=1}^{N_i} L_k\right] = \mathbb{E}_i\left[\sum_{k=1}^{\infty} L_k 1_{\{k\leq N_i\}}\right]$$
$$= \sum_{k=1}^{\infty}\mathbb{E}_i[1_{\{k\leq N_i\}}\mathbb{E}[L_k|X_0,\ldots,X_{\tau_{k-1}^F}]]$$
$$\leq m\sum_{k=1}^{\infty}\mathbb{E}_i[1_{k\leq N_i}] = m\mathbb{E}_i[N_i].$$

Since F is finite and $\{X_k^F\}_{k\in\mathbb{Z}_+}$ irreducible, we have $\{X_k^F\}_{k\in\mathbb{Z}_+}$ being positive recurrent. Thus, $\mathbb{E}_i[N_i] < \infty$, implying $\mathbb{E}_i[\tau_i] < \infty$ and the positive recurrence of $\{X_n\}_{n\in\mathbb{Z}_+}$. □

We now present an important result that proves positive recurrence. This is the celebrated Foster-Lyapunov criterion. In the sequel we will also present the most general statement of this type.

Theorem 49. *[Foster-Lyapunov] Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be a DTMC with an irreducible transition matrix P. If for some function $V : I \mapsto \mathbb{R}$ and some $\epsilon > 0$, we have $\inf_{i\in I} V(x) > -\infty$ and*

$$\sum_{k\in I} P_{jk}V(k) < \infty \quad \forall j \in F$$
$$\sum_{k\in I} P_{jk}V(k) \leq V(j) - \epsilon \quad \forall j \notin F$$

for a finite set F. Then the chain is positive recurrent.

Note that V being bounded below implies that without loss of generality we can assume that $V : I \mapsto \mathbb{R}_+$, and we do so in the remainder. The last condition is also written in the following more illuminating forms

$$(PV)(j) - V(j) = ((P - \mathbf{I})V)(j) \leq -\epsilon + b1_F(j)$$
$$\mathbb{E}[V(X_{n+1}) - V(X_n)|X_n = j] \leq -\epsilon + b1_F(j)$$

for some finite constant b. In other words, the one-step drift of the DTMC is such that outside F, it is strictly decreasing; loosely put there is pull towards the finite set F.

Proof. Let $X_0 = i \notin F$ and define $Y_n = V(X_n)1_{\{\tau_F > n\}}$. Then it is easy to see that we have on $\{\tau_F > n\}$

$$\mathbb{E}_i[Y_{n+1}|X_0, \ldots, X_n] \leq 1_{\{\tau_F > n\}}\mathbb{E}_i[V(X_{n+1})|X_0, \ldots, X_n] \leq 1_{\{\tau_F > n\}}V(X_n) - \epsilon 1_{\{\tau_F > n\}} \leq Y_n,$$

and on $\{\tau_F \leq n\}$ we have $Y_n = Y_{n+1} = 0$. Thus, $\{Y_n\}_{n \in \mathbb{Z}_+}$ is a non-negative supermartingale. Therefore, iterating we get

$$0 \leq E_i[Y_{n+1}] \leq \mathbb{E}_i[Y_n] - \epsilon \mathbb{P}_i(\tau_F > n) \leq \cdots \leq \mathbb{E}_i[Y_0] - \epsilon \sum_{k=0}^{n} \mathbb{P}_i(\tau_F > k)$$

In other words,

$$\sum_{k=0}^{n} \mathbb{P}_i(\tau_F > k) \leq \frac{V(i)}{\epsilon}$$

Letting $n \to \infty$, this then implies that $\mathbb{E}_i[\tau_F] \leq \frac{h(i)}{\epsilon}$. Now for $j \in F$, we have

$$\mathbb{E}_j[\tau_F] = 1 + \sum_{i \notin F} P_{ji}\mathbb{E}_i[\tau_F] \leq 1 + \frac{1}{\epsilon}\sum_{i \notin F} P_{ji}V(i)$$

which is again finite. The positive recurrence result now follows from Lemma 11. □

We have a simple corollary that is known as Pakes' Lemma.

Corollary 4. *[Pakes' Lemma] Let $\{X_n\}_{n \in \mathbb{Z}_+}$ be an irreducible DTMC on $I = \mathbb{N}$ such that for all $n \in \mathbb{Z}_+$ and all $i \in I$,*

$$\mathbb{E}[X_{n+1}|X_n = i] < \infty, \text{ and}$$
$$\limsup_{i \to 0} \mathbb{E}[X_{n+1} - X_n|X_n = i] < 0.$$

Then $\{X_n\}_{n \in \mathbb{Z}_+}$ is positive recurrent.

Proof. Let the left side of the last inequality be -2ϵ where $\epsilon > 0$. The last inequality also says that for i sufficiently large, greater than some i_0, $\mathbb{E}[X_{n+1} - X_n|X_n = i] < -\epsilon$. We can now apply Theorem 49 with $V(i) = i$ and $F = \{i : i \leq i_0\}$ to prove the result. □

From drift conditions we can obtain bounds on performance. The simplest result here is the following.

Proposition 1. *Suppose P is an irreducible and positive recurrent stochastic matrix with stationary distribution π, and let f, g and V be non-negative functions on I such that*

$$(PV)(i) \leq V(i) - f(i) + g(i) \quad \forall i \in E.$$

If $\mathbb{E}_\pi[g(X)] < \infty$ and $\mathbb{E}_\pi[V(X)] < \infty$, then we also have $\mathbb{E}_\pi[f(X)] \leq \mathbb{E}_\pi[g(X)] < \infty$, where X is an I-valued random variable with distribution π.

Proof. In vector form, the relationship above can be written as

$$f \leq V - PV + g \Rightarrow P^k f \leq P^k V - P^{k+1}V + P^k g \quad \forall k \in \mathbb{Z}_+$$

Therefore, we also get

$$\sum_{k=1}^{n} P^k f \leq PV - P^{n+1}V + \sum_{k=1}^{n} P^k g \leq PV + \sum_{k=1}^{n} P^k g$$

Multiplying by π on the left, we get

$$n\mathbb{E}_\pi[f(X)] \leq \mathbb{E}_\pi[V(X)] + n\mathbb{E}_\pi[g(X)]$$

Now dividing by n on both sides and letting $n \to \infty$, we get the result. Note that the basic relationship already gave us the bound $\mathbb{E}_\pi[f(X)] \leq \mathbb{E}_\pi[V(X)] + \mathbb{E}_\pi[g(X)]$. The extra work was to tighten this bound. □

We will now generalize most of these results. First we present a useful lemma.

Lemma 12. *Suppose P is a stochastic matrix, and V, f and g are non-negative functions such that $PV - V \leq -f + g$. Then for any initial state i_0 and any stopping time τ,*

$$\mathbb{E}_{i_0}\left[\sum_{k=0}^{\tau-1} f(X_k)\right] \leq V(i_0) + \mathbb{E}_{i_0}\left[\sum_{k=0}^{\tau-1} g(X_k)\right]$$

where $\{X_k\}_{k\in\mathbb{Z}_+}$ is (δ_{i_0}, P) Markov.

Proof. The relationship $PV - V \leq -f + g$ implies that

$$\mathbb{E}[V(X_{k+1})|X_0, \ldots, X_k] + f(X_k) \leq V(X_k) + g(X_k).$$

Let $\tau^n = \min(\tau, n, \inf\{k \geq 0 : V(X_k) \geq n\})$ which is again a stopping time. Using the fact that $1_{\{\tau^n > k\}} \geq 1_{\{\tau^n > k+1\}}$ and the fact that

$$\mathbb{E}_{i_0}[V(X_{k+1})1_{\{\tau^n > k+1\}}] \leq \mathbb{E}_{i_0}[V(X_{k+1})1_{\{\tau^n > k\}}] = \mathbb{E}_{i_0}[\mathbb{E}[V(X_{k+1})|X_0, \ldots, X_k]1_{\{\tau^n > k\}}]$$

we get

$$\mathbb{E}_{i_0}[V(X_{k+1})1_{\{\tau^n > k+1\}}] + \mathbb{E}_{i_0}[f(X_k)1_{\{\tau^n > k\}}] \leq \mathbb{E}_{i_0}[V(X_k)1_{\{\tau^n > k\}}] + \mathbb{E}_{i_0}[g(X_k)1_{\{\tau^n > k\}}]$$

The definition of τ^n implies that all the terms in the relationship above are 0 for $k \geq n$ and $\mathbb{E}_{i_0}[V(X_k)1_{\{\tau^n > k\}}] \leq n < \infty$ for $k < n$. Thus, adding the above over all k we get

$$\mathbb{E}_{i_0}\left[\sum_{k=0}^{\tau^n-1} f(X_k)\right] \leq V(i_0) + \mathbb{E}_{i_0}\left[\sum_{k=0}^{\tau^n-1} g(X_k)\right]$$

Letting $n \to \infty$ and appealing to the monotone convergence theorem yields the required result. □

Using this we get a different version of the Foster-Lyapunov criterion.

Theorem 50. *[Alternate Foster-Lyapunov stability criterion] Suppose P is an irreducible stochastic matrix, $V : I \mapsto \mathbb{R}_+$ and F is a finite subset of I. If there exist $\epsilon > 0$ and a constant b (non-negative) such that $PV - V \leq -\epsilon + b1_F$, then P is positive recurrent.*

Proof. Let $f \equiv \epsilon$, $g = b1_F$ and $\tau_F = \inf\{t \geq 1 : X(t) \in F\}$ (the first passage time to F). Then Lemma 12 implies that $\epsilon\mathbb{E}_i[\tau_F] \leq V(i) + b$ for any $i \in I$, in particular for any $i \in F$. Now using Lemma 11 we get the desired result. ☐

This now yields the following generalization of Proposition 1 where we don't assume that $\mathbb{E}_\pi[V(X)] < \infty$.

Proposition 2. *[Moment Bound] Suppose P is an irreducible and positive recurrent stochastic matrix with stationary distribution π, and let f, g and V be non-negative functions on I such that*

$$(PV)(i) \leq V(i) - f(i) + g(i) \quad \forall i \in E.$$

Then $\mathbb{E}_\pi[f(X)] \leq \mathbb{E}_\pi[g(X)]$ where X is an I-valued random variable with distribution π.

Proof. Fix a state i_0 and let $\{X_k\}_{k\in\mathbb{Z}_+}$ be (δ_{i_0}, P) Markov. Let $\tau_m^{i_0}$ be the time of the m^{th} return to state i_0, then we have

$$\mathbb{E}_{i_0}\left[\sum_{k=0}^{\tau_m^{i_0}-1} f(X_k)\right] = mE[\tau_{i_0}]\mathbb{E}_\pi[f(X)]$$

$$\mathbb{E}_{i_0}\left[\sum_{k=0}^{\tau_m^{i_0}-1} g(X_k)\right] = mE[\tau_{i_0}]\mathbb{E}_\pi[g(X)]$$

Lemma 12 applied with stopping time $\tau_m^{i_0}$ yields $mE[\tau_{i_0}]\mathbb{E}_\pi[f(X)] \leq V(i_0) + mE[\tau_{i_0}]\mathbb{E}_\pi[g(X)]$. Dividing by $mE[\tau_{i_0}]$ on both sides and letting $m \to \infty$ yields the required result. ☐

This and Theorem 49 then yield the following useful corollary.

Corollary 5. *[Combined Foster-Lyapunov criterion with moment bound] Let P be an irreducible stochastic matrix and suppose V, f and g are non-negative functions on I such that $PV - V \leq -f + g$. Suppose, in addition, that for some $\epsilon > 0$ the set F defined by $F := \{i \in I : f(i) < g(i) + \epsilon\}$ is finite. Then X is positive recurrent and $\mathbb{E}_\pi[f(X)] \leq \mathbb{E}_\pi[g(X)]$.*

Proof. Let $b = \max_{i\in C}(g(i) + \epsilon - f(i))$. Then V, F, b and ϵ satisfy the hypotheses of Theorem 49 so that P is positive recurrent. Now we can apply Proposition 2 to prove $\mathbb{E}_\pi[f(X)] \leq \mathbb{E}_\pi[g(X)]$. ☐

We will now present and prove a significant generalization of the Foster-Lyapunov criterion for positive recurrence. We start with a set of assumptions on functions (non-negative) V, (non-negative) g and h:

(1) (L0) V is unbounded from above: $\sup_{i\in I} V(i) = \infty$.
(2) (L1) h is bounded from below: $\inf_{i\in I} h(i) > -\infty$.
(3) (L2) h is eventually positive: $\liminf_{V(i)\to\infty} h(i) > 0$.
(4) (L3) g is locally bounded from above: $G(N) = \sup_{\{i:V(i)\leq N\}} g(i) < \infty$ for all $N > 0$.
(5) (L4) g is eventually bounded by h: $\limsup_{V(i)\to\infty} g(i)/h(i) < \infty$.

Then the theorem is as follows.

Theorem 51. *Suppose that the drift of V in $g(i)$ steps satisfies*

$$\mathbb{E}_i[V(X_{g(i)}) - V(X_0)] \leq -h(i) \quad \forall i \in I$$

where $\{X_n\}_{n\in\mathbb{Z}_+}$ is a (δ_i, P) DTMC and the functions V, g and h satisfy (L0)-(L4). Let

$$\tau_N = \inf\{n \geq 1 : V(X_n) \leq N\}.$$

Then there exists $N_0 > 0$ such that for all $N \geq N_0$ and any $i \in I$, we have

$$\mathbb{E}_i[\tau_N] < \infty \text{ and } \sup_{V(i)\leq N} \mathbb{E}_i[\tau_N] < \infty.$$

The result proves that the set $B_N = \{i \in I : V(i) \leq N\}$ is positive recurrent. Thus, if P is irreducible and B_N is a finite subset of I, then the chain is positive recurrent using Lemma 11. That this is a generalization of other criteria can be seen as follows:

(1) (Pakes' Lemma) - here $I = \mathbb{N}$, $V(i) = i$, $g(i) \equiv 1$ and $h(i) = \epsilon - b1_{\{j\leq N_0\}}(i)$.
(2) (Foster-Lyapunov criterion) - here I is general, $g(i) \equiv 1$ and $h(i) = -\epsilon - b1_{\{V(j)\leq N_0\}}(j)$.
(3) (Dai's criterion or Fluid limits criterion) - here I is general, $g(i) = \lceil V(i) \rceil$ (where $\lceil t \rceil = \inf\{n \in \mathbb{Z} : t \leq n\}$) and $h(i) = \epsilon V(i) - C_1 1_{\{j:V(j)\leq C_2\}}(i)$
(4) (Meyn-Tweedie criterion) - here $h(i) = g(i) - C_1 1_{\{j:V(j)\leq C_2\}}(i)$.
(5) Such state-dependent drift conditions were first proposed by Fayolle, Malyshev and Menshikov.

Proof of Theorem 51. From the drift condition, it is clear that $V(i) - h(i) \geq 0$ for all $i \in I$. We choose N_0 such that $\inf_{V(i)>N_0} h(i) > 0$. Then, for $N \geq N_0$, we set

$$d = \sup_{V(i)>N} g(i)/h(i), \quad H = -\inf_{i\in I} h(i), \quad c = \inf_{V(i)>N} h(i).$$

Define an increasing sequence of stopping time $\{T_n\}_{n\in\mathbb{Z}_+}$ recursively by

$$T_0 = 0, \quad T_n = T_{n-1} + g(X_{T_{n-1}}) \; \forall n \in \mathbb{N}.$$

By the strong Markov property, the sequence $\{Y_n\}_{n\in\mathbb{Z}_+}$ given by $Y_n = X_{T_n}$ forms a DTMC (possibly time-inhomogeneous) with (can be proved via induction) $\mathbb{E}_i[V(Y_{n+1})] \leq \mathbb{E}_i[V(Y_n)] + H$, which then implies that $\mathbb{E}_i[V(Y_n)] < \infty$ for all $n \in \mathbb{Z}_+$ and $i \in I$. Define the stopping time

$$\gamma = \inf\{n \geq 1 : V(Y_n) \leq N\}$$

Note that $\tau_N \leq T_\gamma$ *a.s.* and thus, proving that $\mathbb{E}_i[T_\gamma] < \infty$ will prove the required result. Set $\mathcal{E}_n$ to be the net accumulation of V between 0 and $\gamma \wedge n$, i.e.,

$$\mathcal{E}_n = \sum_{k=0}^{\gamma\wedge n} V(Y_k) = \sum_{k=0}^{n} V(Y_k) 1_{\{\gamma\geq k\}}$$

Now we are in the traditional set-up and can determine the expected drift, i.e.,

$$\begin{aligned}
\mathbb{E}_i[\mathcal{E}_n - \mathcal{E}_0] &= \mathbb{E}_i\left[\sum_{k=1}^{n} \mathbb{E}[V(Y_k)1_{\{\gamma\geq k\}}|Y_0,\dots,Y_{k-1}]\right] \\
&= \mathbb{E}_i\left[\sum_{k=1}^{n} 1_{\{\gamma\geq k\}}\mathbb{E}[V(Y_k)|Y_0,\dots,Y_{k-1}]\right] \\
&\leq \mathbb{E}_i\left[\sum_{k=1}^{n} 1_{\{\gamma\geq k\}}(V(Y_{k-1}) - h(Y_{k-1}))\right] \\
&\leq \mathbb{E}_i\left[\sum_{k=1}^{n+1} 1_{\{\gamma\geq k-1\}}V(Y_{k-1})\right] - \mathbb{E}_i\left[\sum_{k=1}^{n} 1_{\{\gamma\geq k\}}h(Y_{k-1})\right] \\
&= \mathbb{E}_i[\mathcal{E}_n] - \mathbb{E}_i\left[\sum_{k=1}^{n} 1_{\{\gamma\geq k\}}h(Y_{k-1})\right]
\end{aligned}$$

where we used $V(i) - h(i) \geq 0$. From the above we get

$$\mathbb{E}_i\left[\sum_{k=1}^{n} 1_{\{\gamma\geq k\}}h(Y_{k-1})\right] \leq \mathbb{E}_i[V(Y_0)] = V(i)$$

First we assume $V(i) > N$. Then it follows that $V(Y_k) > N$ for $k < \gamma$ by the definition of γ. Therefore, $h(Y_k) \geq c > 0$ for $k < \gamma$. Using this to obtain

$$c\mathbb{E}_i\left[\sum_{k=0}^{n} 1_{\{\gamma>k\}}\right] \leq V(i) + c,$$

where we added $c\mathbb{E}_i[1_{\{\gamma>n\}}] \leq c$ on both sides to get the bound. So by the monotone convergence theorem we have

$$\mathbb{E}_i[\gamma] \leq \frac{V(i)+c}{c} < \infty$$

Using $h(i) > dg(i)$, since $V(i) > N$, we also have

$$\sum_{k=0}^{\gamma-1} h(Y_i) \geq d\sum_{k=0}^{\gamma-1} g(Y_i) = dT_\gamma$$

From the above we get $T_\gamma < \infty$ *a.s.* and

$$\mathbb{E}_i[\tau_N] \leq \mathbb{E}_i[T_\gamma] \leq \frac{V(i)+c}{cd} < \infty.$$

We now need to consider the case of $V(i) \leq N$. By conditioning on Y_1, using the Markov property of $\{Y_n\}_{n\in\mathbb{Z}_+}$ and the calculations above, we have

$$\begin{aligned}\mathbb{E}_i[\tau_N] \leq \mathbb{E}_i[T_\gamma] &\leq g(i) + \sum_{j:V(j)>N} P^1_{ij}\mathbb{E}_j[T_\gamma] \\ &\leq g(i) + \sum_{j:V(j)>N} P^1_{ij}\frac{V(j)+c}{cd} \\ &\leq g(i) + \sum_{j\in I} P^1_{ij}\frac{V(j)+c}{cd} \\ &\leq g(i) + \frac{V(i)+H+c}{cd} \leq G(N) + \frac{N+H+c}{cd} < \infty\end{aligned}$$

where $P^1_{ij} := \mathbb{P}_i(X_{g(i)} = j)$ for all $j \in I$, is the transition probability for the first transition of the Markov chain $\{Y_n\}_{n\in\mathbb{Z}_+}$; note that we used the drift inequality above as

$$\mathbb{E}_i[V(X_{g(i)})] = \sum_{j\in I} P^1_{ij}V(j) \leq V(i) - h(i) \leq V(i) + H$$

□

The condition (L4) is not just a technical condition as the time horizon $g(i)$ over which the drift is computed shouldn't be too large compared to the estimate $h(i)$ for the size of the drift itself. The following simple example demonstrates this. Let $I = \mathbb{N}$ and let the transition probabilities be

$$P_{11} = 1,\ P_{k,k+1} = p_k,\ P_{k,1} = 1 - p_k \quad \forall k \geq 2$$

where $p_k \in (0, 1)$ for all $k \geq 2$ and $\lim_{k\to\infty} p_k = 1$; for concreteness we take $p_k = 1 - \frac{1}{k}$ for $k \geq 2$. Thus, we have jumps either of size 1 or $-(k-1)$ till the first time state 1 is hit. Also note that state 1 is absorbing while the rest are not, i.e., once the Markov chain hits this state, it remains there, and so 1 can be the only recurrent state.

Let $\{X_k\}_{k\in\mathbb{Z}_+}$ be (δ_k, P) Markov where $k \geq 2$. Using $1 - x \leq e^{-x}$ for $x \in [0, 1]$ we get that

$$\mathbb{P}_k(X_{n+1} = X_n + 1 \text{ for all } n \geq 0) = \prod_{l=k}^{\infty} p_k \leq e^{-\sum_{l=k}^{\infty}\frac{1}{l}} = 0$$

implying recurrence of the state 1. However, if we set $\tau = \inf\{n : X_n = 1\}$ to be the hitting time of state 1, then we find that

$$\mathbb{P}_k(\tau = n) = \begin{cases} \frac{1}{k} & \text{if } n = 1; \\ \frac{1}{k+n-1}\prod_{l=0}^{n-2}(1 - \frac{1}{k+l}) = \frac{k-1}{(k+n-1)(k+n-2)} & \text{otherwise.}\end{cases}$$

Now it is easy to see that $\mathbb{E}_k[\tau] = \infty$ for all $k \geq 2$; note that $\mathbb{P}_k(\tau > n) = \frac{k-1}{k-1+n}$. Therefore, state 1 cannot be positive recurrent. Let us choose $V(k) = \log(1 \vee \log(k))$ and $g(k) = k^2$. Then we have

$$\mathbb{E}_k[V(X_{g(k)})] = \mathbb{P}_k(\tau > g(k))V(k + k^2) = \frac{k-1}{k-1+k^2}\log(1 \vee \log(k + k^2))$$

Now it is easy to verify that $\mathbb{E}_k[V(X_{g(k)}) - V(k)] \leq -h(k)$ where $h(k) = c_1V(k) - c_2$ where c_1 and c_2 are positive constants. In this case L(0)-L(3) hold but L(4) fails.

5.3. **Stability criteria for continuous time processes.** We now present the equivalent of the Foster-Lyapunov criterion for CTMCs. Now the drift is defined using the rate matrix Q (remember equivalence of Q and $P - \mathbf{I}$) as follows for $X(t) = i$

$$(QV)(i) = \sum_{j \in I} q_{ij}(V(j) - V(i)) = \sum_{j \in I : j \neq i} q_{ij}(V(j) - V(i)).$$

The result we seek is the following.

Theorem 52. *[Foster-Lyapunov criterion for CTMCs] Suppose $V : I \mapsto \mathbb{R}_+$ and F is a finite subset of S. Let Q be an irreducible rate matric. Then the following hold:*

1. *if $QV(i) \leq 0$ on $I \setminus F$, and $\{i : V(i) \leq K\}$ is finite for all K, then $\{X_t\}_{t \in \mathbb{R}_+}$ is recurrent;*
2. *suppose for some $b > 0$ and $\epsilon > 0$ that*
$$(QV)(i) \leq -\epsilon + b 1_F(i) \quad \forall i \in I$$
Assume further that $\{i : V(i) \leq K\}$ is finite for all K or that $\{X_t\}_{t \in \mathbb{R}_+}$ is non-explosive. Then $\{X_t\}_{t \in \mathbb{R}_+}$ is positive recurrent.

The result is proved using the discrete-time results (Theorems 47 and 49) by analyzing the jump DTMC. First we have a version of Lemma 12.

Lemma 13. *Suppose $QV \leq -f + g$ on I where f and g are non-negative functions. Fix an initial state i_0, let N be a stopping time for the jump DTMC $\{Y_n\}_{n \in \mathbb{Z}_+}$, and let J_N be the time of the N^{th} jump of $\{X_t\}_{t \in \mathbb{R}_+}$. Then*

$$\mathbb{E}_{i_0}\left[\int_0^{J_N} f(X(t))dt\right] \leq V(i_0) + \mathbb{E}_{i_0}\left[\int_0^{J_N} g(X(t))dt\right]$$

Proof. Let D denote the diagonal matrix with entries q_i. Since the jump DTMC transition matrix Π is given by $\Pi = D^{-1}Q + \mathbf{I}$, the condition $QV \leq -f + g$ implies that $\Pi V - V \leq -\tilde{f} + \tilde{g}$, where $\tilde{f} := D^{-1}f$ and $\tilde{g} := D^{-1}g$. Now applying Lemma 12 to the jump DTMC yields

$$\mathbb{E}_{i_0}\left[\sum_{k=0}^{N-1} \tilde{f}(Y_k)\right] \leq V(i_0) + \mathbb{E}_{i_0}\left[\sum_{k=0}^{N-1} \tilde{g}(Y_k)\right]$$

From the holding-times and jump-chain description of $\{X_t\}_{t \in \mathbb{R}_+}$ it is easy to see that

$$\mathbb{E}_{i_0}\left[\sum_{k=0}^{N-1} \tilde{f}(Y_k)\right] = \mathbb{E}_{i_0}\left[\int_0^{J_N} f(X(t))dt\right] \text{ and } \mathbb{E}_{i_0}\left[\sum_{k=0}^{N-1} \tilde{g}(Y_k)\right] = \mathbb{E}_{i_0}\left[\int_0^{J_N} g(X(t))dt\right]$$

The result then follows. ∎

Proof of Theorem 52. Part (1) is quite straightforward. The condition $QV \leq 0$ is equivalent to the condition $\Pi V - V \leq 0$, which using Theorem 47 implies that the jump DTMC $\{Y_n\}_{n \in \mathbb{Z}_+}$ is recurrent. This then implies recurrence for $\{X_t\}_{t \in \mathbb{R}_+}$ too.

Now consider part (2). First note that on $I \setminus F$ we have $QV \leq -\epsilon < 0$. Thus, from part (1) we get recurrence of $\{X_t\}_{t \in \mathbb{R}_+}$; therefore, it is non-explosive in either case. Let $f(i) \equiv -\epsilon$ and $g(i) = 1_F(i)$. Fix an $i_0 \in F$ and set $N = \inf\{k \geq 1 : Y_k \in F\}$ (i.e., first passage time to F in the jump DTMC starting out in $i_0 \in F$) and J_N to be N^{th} jump. Then Lemma 13 implies that $\epsilon \mathbb{E}_{i_0}[J_N] \leq V(i_0) + b/q_{i_0}$. Since J_N is finite *a.s.* and since $\{X_t\}_{t \in \mathbb{R}_+}$ is non-explosive, we must have that J_N is the first time to hit set F after leaving state i_0. Since F is finite, this then implies that the first passage time to F beginning from any state in F is finite. Now the continuous time equivalent of Lemma 11 proves the positive recurrence of $\{X_t\}_{t\in\mathbb{R}_+}$. □

We can now get other equivalent results as follows.

Proposition 3. *Suppose V, f and g are non-negative functions on I and suppose that $QV \le -f + g$ on I. In addition, assume that Q is an irreducible and positive-recurrent rate matrix with stationary distribution π. Then $\bar{f} := \pi f \le \bar{g} := \pi g$.*

Corollary 6. *Let Q be an irreducible rate matrix and V, f and g are non-negative functions on I such that $QV \le -f + g$ on I and the set $F = \{i : f(i) < g(i) + \epsilon\}$ is finite for some $\epsilon > 0$. Suppose also that $\{i : V(i) \le K\}$ is finite for all K. Then Q is positive recurrent and $\pi f \le \pi g$ where π is the (unique) invariant distribution of Q.*

5.4. **Kurtz's theorem.** This result presents a differential equation based approximation to a class of CTMCs with specific assumptions. The special class is general enough to model logistic growth, epidemics and chemical reactions; in fact, the formal proof of mass action kinetics (also knows as the Chemical Master Equation) is a consequence of Kurtz's theorem (Theorem 2.1, Chapter 11 in reference that follows). The material for this section comes from Chapter 11, S. N. Ethier and T. G. Kurtz. 1986. Markov processes: Characterization and convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, New York.

The main result is based on characterization of CTMCs via Poisson processes that we already discussed; the exact representation is obtained by random time changes. For this section we constrain the state-space to be $\mathbb{Z}^d$ for some $d \in \mathbb{N}$. If Q is the rate matrix and $\{N_l\}_{l\in\mathbb{Z}^d}$ be a collection of independent rate 1 Poisson processes. Let X_0 be some fixed point in $\mathbb{Z}^d$, ζ be the explosion time and ∞ the exploded state, then we have the (δ_{X_0}, Q) CTMC $\{X_t\}_{t\in\mathbb{R}_+}$ given by

$$X_t = \begin{cases} X_0 + \sum_{l\in\mathbb{Z}^d} l N_l \left(\int_0^t q_{X_s,X_s+l}\, ds\right) & \text{if } t < \zeta \\ \infty & \text{if } t \ge \zeta \end{cases}$$

See Section 4, Chapter 6 in Ethier & Kurtz book for more details. Note that we have changed notation for the Poisson process from N_t^l to $N_l(t)$ where $l \in \mathbb{Z}^d$ and $t \in \mathbb{R}_+$ for the sake of readability.

For the result we consider a sequence of CTMCs indexed by $n \in \mathbb{N}$. We will be interested in how the sequence of processes behave as $n \to \infty$. We will start with a simple result that follows from the strong law of large numbers. Let $\{Z_n\}_{n\in\mathbb{N}}$ be a random process taking values in $\mathbb{R}$ that satisfies the strong law of large numbers, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Z_i = \mu \quad a.s.$$

for some $\mu \in \mathbb{R}$; an example is an *i.i.d.* sequence that is integrable with mean μ. Fix a $T > 0$ and for all $t \in [0, T]$ define $S_t^n = \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} Z_k$, then we have the following result.

Proposition 4. *The sequence of processes $\{S_t^n\}_{t\in[0,T]}$ and $n \in \mathbb{N}$ converges to $\{\mu t\}_{t\in[0,T]}$ such that uniform norm converges to* 0 a.s.*, i.e.,*

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |S_t^n - \mu t| = 0 \quad \text{a.s.}$$

Proof. It is possible to justify that $\sup_{t\in[0,T]} |S_t^n - \mu t|$ is a random variable for every $n \in \mathbb{N}$ but we take this as given here. Note that $S_0^n \equiv 0$ for all $n \in \mathbb{N}$. Thus, we fix a $t \in (0,T]$ and consider $n \geq \lceil 1/t \rceil$. Now notice that

$$\begin{aligned}
|S_t^n - \mu t| &= \left| \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} Z_k - \mu t \right| \\
&= \left| \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} Z_k - \mu \frac{\lfloor nt \rfloor}{n} + \mu \left(t - \frac{\lfloor nt \rfloor}{n} \right) \right| \\
&\leq \frac{\lfloor nt \rfloor}{n} \left| \frac{1}{\lfloor nt \rfloor} \sum_{k=1}^{\lfloor nt \rfloor} (Z_k - \mu) \right| + \mu \left| t - \frac{\lfloor nt \rfloor}{n} \right| \\
&\leq T \left| \frac{1}{\lfloor nt \rfloor} \sum_{k=1}^{\lfloor nt \rfloor} (Z_k - \mu) \right| + \frac{\mu}{n}
\end{aligned}$$

Now for a fixed $\epsilon > 0$ we can choose (fixing a sample-path that satisfies the strong law of large numbers) a K^* such that for all $K \geq K^*$

$$\left| \frac{1}{K} \sum_{k=1}^{K} (Z_k - \mu) \right| \leq \frac{\epsilon}{4T}$$

Now for $n \geq \frac{4\mu}{\epsilon}$ the second term is less than equal to $\frac{\epsilon}{4}$. Thus, if $t > \frac{K^*}{n}$, then first term is less than $\frac{\epsilon}{4}$ for all $n \geq \frac{4\mu}{\epsilon}$. Now for $0 \leq t \leq \frac{K^*}{n}$ we have

$$|S_t^n - \mu t| \leq \frac{1}{n} \left| \sum_{k=1}^{K^*} Z_k \right| + \frac{\mu K^*}{n}$$

where we note that the right side goes to 0 as $n \to \infty$. Thus, taking n large enough (and greater than $\frac{4\mu}{\epsilon}$) we can make the right side less than $\frac{\epsilon}{2}$. Therefore we have for n large enough that $\sup_{t\in[0,T]} |S_t^n - \mu t| \leq \epsilon$, which proves the result. □

This proof is the same as that of Theorem 4, "Ordinary CLT and WLLN Versions of $L = W$," Peter W. Glynn and Ward Whitt, Mathematics of Operations Research, 13(4), 1988, pp. 674–692. This yields the following corollary that we will use later on.

Corollary 7. *Let* $\{N(t)\}_{t\in\mathbb{R}_+}$ *be a rate* 1 *Poisson process. Then for all* $T \in \mathbb{R}_+$ *we have* $\lim_{n\to\infty} \sup_{0\leq t\leq T} \left| \frac{1}{n} N(nt) - t \right|$.

Proof. First note that $N(nt) = \sum_{i=1}^{\lfloor nt \rfloor} (N(i) - N(i-1)) + N(nt) - N(\lfloor nt \rfloor)$. By the independent increments property of the Poisson process each of these segments is independent and $N(i) - N(i-1)$ is distributed as $Poisson(1)$ for all $i \in \mathbb{N}$ while $N(nt) - N(\lfloor nt \rfloor)$ is distributed as $Poisson(nt - \lfloor nt \rfloor)$ which is smaller than $N(\lceil nt \rceil) - N(\lfloor nt \rfloor)$ which is distributed as $Poisson(1)$. Therefore, we have

$$\sup_{0\leq t\leq T} \left| \tfrac{1}{n} N(nt) - t \right| \leq \sup_{0\leq t\leq T} \left| \tfrac{1}{n} N(\lfloor nt \rfloor) - t \right| + \frac{N(\lceil nt \rceil) - N(\lfloor nt \rfloor)}{n}$$

The first term goes to 0 *a.s.* owing to Proposition 4 (since an *i.i.d.* sequence of $Poisson(1)$ random variables satisfies the strong law of large numbers) and it is easy to see that the second term also goes to 0 *a.s.*, and the independence allows us to combine these results without any trouble. □

Kurtz's theorem applies to a class of CTMCs with a specific form for the rate matrix. We start with a collection of non-negative functions β_l, $l \in \mathbb{Z}^d$, defined on a subset E of $\mathbb{R}^d$ such that $\sum_{l\in\mathbb{Z}^d} \beta_l(x) < \infty$ for all $x \in E$. Setting $E_n = E \cap \{\frac{k}{n} : k \in \mathbb{Z}^d\}$, we require that $x \in E_n$ and $\beta_l(x) > 0$ imply $x + \frac{l}{n} \in E_n$. We define a set of CTMCs $\{\hat{X}^n\}_{n\in\mathbb{N}}$ to be a density dependent family if process $\hat{X}^n$ has state-space E_n and rate matrix given by

$$q^n_{x,y} = n\beta_{n(y-x)}(x) \quad x, y \in E_n;$$

in other words, a transition of size l/n (with $l \in \mathbb{Z}^d$) from state $\frac{k}{n}$ occurs at rate $n\beta_l(\frac{k}{n})$. From the random time change representation we then have that

$$\hat{X}^n_t = \begin{cases} \hat{X}^n_0 + \sum_{l\in\mathbb{Z}^d} lN_l\left(n\int_0^t \beta_l\left(\frac{\hat{X}^n_s}{n}\right)\ ds\right) & \text{if } t < \zeta^n \\ \infty & \text{if } t \geq \zeta^n \end{cases}$$

where $\hat{X}^n_0$ is the initial condition and τ^n_∞ is the explosion time. Now set $F(x) = \sum_{l\in\mathbb{Z}^d} l\beta_l(x)$ and $X^n_t = \frac{\hat{X}^n_t}{n}$ for all $t \in \mathbb{R}_+$. Then we obtain

$$X^n_t = \begin{cases} X^n_0 + \sum_{l\in\mathbb{Z}^d} l\frac{N_l\left(n\int_0^t \beta_l(X^n_s)\ ds\right)}{n} & \text{if } t < \zeta^n \\ \infty & \text{if } t \geq \zeta^n \end{cases}$$

Consider the case of $t < \zeta^n$ where we can further write the following

$$X^n_t = X^n_0 + \sum_{l\in\mathbb{Z}^d} l\left(\frac{N_l\left(n\int_0^t \beta_l\left(X^n_s\right)\ ds\right)}{n} - \int_0^t \beta_l\left(X^n_s\right)\ ds\right) + \int_0^t F(X^n_s)ds,$$

where we used non-negativity of β_l to exchange the integral and summation. Note the term in the parenthesis is (conditionally) a difference between a Poisson random variable and its mean. Using this we have the following theorem.

Theorem 53. *Suppose that for each compact subset K of E, we have*

$$\sum_{l\in\mathbb{Z}^d} \|l\|_1 \sup_{x\in K} \beta_l(x) < \infty,$$

and there exists an $M_K > 0$ such that

$$\|F(x) - F(y)\|_1 \leq M_K\|x - y\|_1 \quad x, y \in K.$$

Suppose X^n is such that $\lim_{n\to\infty} X^n_0 = x_0$, and $\{x_t\}_{t\in\mathbb{R}_+}$ satisfies

$$x_t = x_0 + \int_0^t F(x_s)ds \quad t \in \mathbb{R}_+.$$

Then for every $t \geq 0$

$$\lim_{n\to\infty} \sup_{0\leq s\leq t} \|X^n_s - x_s\|_1 = 0 \quad a.s.$$

Implicitly the result assume global existence for solution to differential equation $\frac{d}{dt}x_t = F(x_t)$ with x_0 given. The Lipschitz assumption on F guarantees this. The notation $\|\cdot\|_1$ stands for the l_1 norm, i.e., the sum of the absolute values of the entries.

Proof of Theorem 53. Since $\{x_s\}_{0\le s\le t}$ is a continuous function, for every fixed $t \ge 0$ the values taken by $\{x_s\}_{0\le s\le t}$ lie in a compact set. Now, for the desired convergence it is enough if we consider a neighbourhood of $\{x_s\}_{0\le s\le t}$ (again compact) so that without loss of generality we can assume that $\bar{\beta}_l = \sup_{x\in E}\beta_l(x)$ satisfies $\sum_{l\in\mathbb{Z}^d}\|l\|_1\bar{\beta}_l < \infty$, and that there exists a fixed $M > 0$ such that

$$\|F(x) - F(y)\|_1 \le M\|x-y\|_1 \quad x, y \in E.$$

Using this we get

$$\begin{aligned}
\epsilon_t^n &:= \sup_{0\le u\le t}\left|X_u^n - X_0^n - \int_0^u F(X_s^n)ds\right| \\
&\le \sum_{l\in\mathbb{Z}^d}\|l\|_1\frac{1}{n}\sup_{0\le u\le t}\left|N_l\left(n\int_0^u \beta_l\left(X_s^n\right)\ ds\right) - n\int_0^u \beta_l\left(X_s^n\right)\ ds\right| \\
&\le \sum_{l\in\mathbb{Z}^d}\|l\|_1\frac{1}{n}\sup_{0\le u\le t}\left|N_l\left(nu\bar{\beta}_l\right) - nu\bar{\beta}_l\right| \\
&\le \sum_{l\in\mathbb{Z}^d}\|l\|_1\frac{1}{n}\sup_{0\le u\le t} N_l\left(n\bar{\beta}_l u\right) + n\bar{\beta}_l u \\
&\le \sum_{l\in\mathbb{Z}^d}\|l\|_1\left(\frac{1}{n}N_l\left(n\bar{\beta}_l t\right) + \bar{\beta}_l t\right)
\end{aligned}$$

where the inequalities hold term by term from the second inequality onwards; note that the second inequality follows via monotonicity as

$$\begin{aligned}
&\sup_{0\le u\le t}\left|N_l\left(n\int_0^u \beta_l\left(X_s^n\right)\ ds\right) - n\int_0^u \beta_l\left(X_s^n\right)\ ds\right| \\
&\quad= \sup_{0\le u\le n\int_0^t \beta_l(X_s^n)ds}|N_l(u) - u| \le \sup_{0\le u\le n\bar{\beta}_l t}|N_l(u) - u| = \sup_{0\le u\le t}|N_l(n\bar{\beta}_l u) - n\bar{\beta}_l u|
\end{aligned}$$

Now noting that the process on the right in the bound on ϵ_t^n has independent increments, and using the strong law of large numbers we get

$$\lim_{n\to\infty}\sum_{l\in\mathbb{Z}^d}\|l\|_1\left(\frac{1}{n}N_l\left(n\bar{\beta}_l t\right) + \bar{\beta}_l t\right) = \sum_{l\in\mathbb{Z}^d}\|l\|_1\lim_{n\to\infty}\left(\frac{1}{n}N_l\left(n\bar{\beta}_l t\right) + \bar{\beta}_l t\right) = 2t\sum_l\|l\|_1\bar{\beta}_l < \infty$$

where the limits hold in an *a.s.* sense. Of course, this implies that we can interchange the limit and summation. The term by term inequalities then imply that we can also interchange limit and summation in the third expression and using Corollary 7 we get that $\lim_{n\to\infty}\epsilon_t^n = 0$ *a.s.*

Using the Lipschitz nature of F now yields

$$\|X_t^n - x_t\|_1 \le \|X_0^n - x_0\|_1 + \epsilon_t^n + \int_0^t M\|X_s^n - x_s\|_1 ds,$$

and then by using the Gronwall-Bellman inequality we finally get

$$\|X_t^n - x_t\|_1 \leq (\|X_0^n - x_0\|_1 + \epsilon_t^n)\, e^{Mt}$$

Therefore, the result follows as $n \to \infty$. □

5.5. **Exercises.** Please show all your work. For all exercises simulate an example.

(1) Consider network shown in Figure 1. Assume discrete-time, let arrivals $\{A_i\}_{i\in\mathbb{N}}$ be *i.i.d.* taking values in $\mathbb{Z}_+$ such that $\mathbb{E}[A_i^2] < \infty$. Server 1 can serve up to R_1 arrivals and server 2 can serve up to R_2 arrivals with both R_1 and R_1 in $\mathbb{N}$. Queues hold arrivals if they cannot be served and queues cannot go negative. After service from server 1 arrivals go to server 2. Assume that arrivals in a slot can only be served in the next slot. Show that $\{(Q1_n, Q2_n)\}_{n\in\mathbb{Z}_+}$ is a DTMC; assume $Q1_0 = Q2_0 = 0$. What are its parameters? Can you write down a state recursion? Show that DTMC is transient if $\mathbb{E}[A_i] > \min(R_1, R_2)$ (*Hint*: use the strong law of large numbers). It is possible to prove that DTMC is positive recurrent if $\mathbb{E}[A_i] < \min(R_1, R_2)$. However, show that using a quadratic Lyapunov function, i.e., $V(Q1_n, Q2_n) = Q1_n^2 + aQ1_nQ2_n + bQ2_n^2$, one cannot prove this result. In other words, find the best conditions for positive recurrence using a quadratic Lyapunov function by varying a and b. **Important**: a and b cannot be chosen arbitrarily.

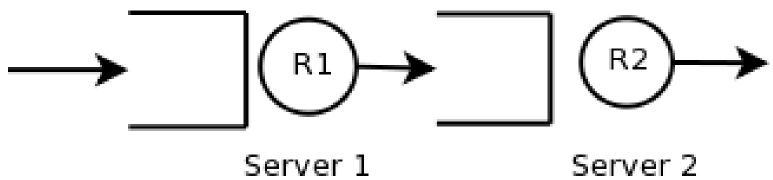

FIGURE 1. Tandem queues

(2) *Lamperti's criterion*: Consider a DTMC in $\mathbb{Z}_+$ with $\lim_{x\to\infty} x\mathbb{E}[X_{n+1} - X_n | X_n = x] = -c$ and $\lim_{x\to\infty} \mathbb{E}[(X_{n+1} - X_n)^2 | X_n = x] = b$ where b and c are positive constants. Use an appropriate Lyapunov function to show that the DTMC is positive recurrent if $2c > b$. Can you prove it is transient if $2c < b$? **Extra credit**

(3) *Lindley's recursion*: Let $\{A_n\}_{n\in\mathbb{N}}$ taking values in $\mathbb{Z}$ be *i.i.d.* with $\mathbb{E}[A_i] < 0$ and $\mathbb{E}[A_i^2] < \infty$. Show that the DTMC given by $X_{n+1} = \max(X_n + A_{n+1}, 0)$ for $n \in \mathbb{Z}_+$ is positive recurrent; assume that $X_0 = 0$. Can you show that it cannot be positive recurrent if $\mathbb{E}[A_i] \geq 0$ (without assuming $\mathbb{E}[A_i^2] < \infty$)? Can you show positive recurrence without assuming $\mathbb{E}[A_i^2] < \infty$? The last part is **Extra credit**.

(4) Consider queueing system in Figure 2. Arrivals are assumed to be Bernoulli with parameter p and can be routed to either queue. Service (when possible, i.e., when the queue is non-empty) is also Bernoulli with parameter R_1 for server 1 and R_2 for server 2. Let $Q1_n$ and $Q2_n$ be the queue-length of each queue; assume $Q1_0 = Q2_0 = 0$. The routing policy that we follow is to route to the shorter queue with ties broken in favour of queue 1. Assume $p = 0.7$ and $R_1 = R_2 = 0.4$. Show that $V(Q1_n, Q2_n) = Q1_n + Q2_n$ does not satisfy the Foster-Lyapunov criterion for any choice of b and finite set F. Is the DTMC positive recurrent for given p, R_1 and R_2? Can you guess the conditions for positive recurrence? Show that $V(Q1_n, Q2_n) = Q1_n^2 + Q2_n^2$ allows one to use the Foster-Lyapunov criterion to prove positive recurrence. The last part is **Extra credit**.

(5) Take the birth-death CTMC, i.e., Q matrix on $\mathbb{Z}_+$ such that $q_{01} = \lambda$ and for $i \in \mathbb{N}$ we have $q_{i,i+1} = \lambda$ and $q_{i,i-1} = \mu$ with every other entry (except the diagonal ones) being 0. Show using the Foster-Lyapunov condition that the CTMC is positive recurrent

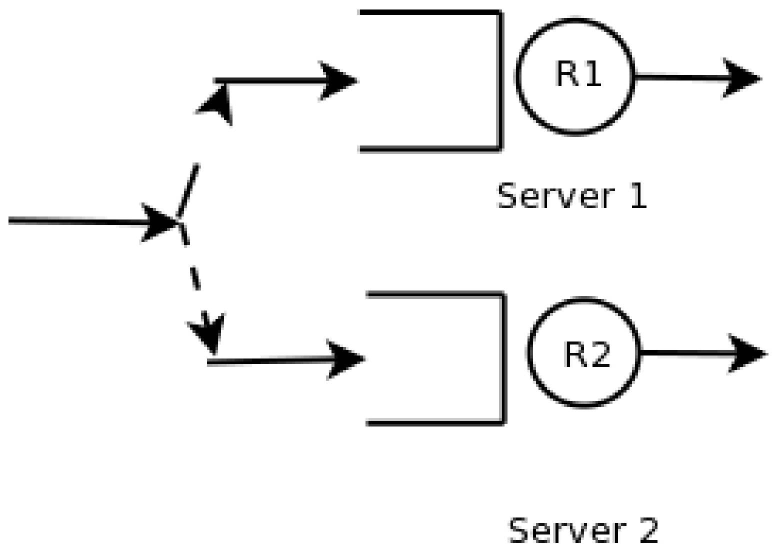

FIGURE 2. Shortest queue routing

if $\lambda < \mu$. Can you also show that it cannot be positive recurrent for $\lambda \geq \mu$? *Hint*: Use uniformization to convert to a DTMC. The last part is **Extra credit**.

6. Controlled Markov Chains

The sources for this section are:

- J. R. Norris, "Markov chains," Cambridge Series in Statistical and Probabilistic Mathematics, 2, Cambridge University Press, Cambridge, 2006.
- P. R. Kumar and P. Varaiya, "Stochastic systems: Estimation, identification and adaptive control," Prentice Hall, Englewood Cliffs, N. J., 1986.
- S. M. Ross, "Applied Probability Models with Optimization Applications," Dover Publications, 1970.

Thus far our Markov chains were governed by fixed rules. However, in real systems there is considerable flexibility in designing the rules/taking actions so as to meet some objective (eg., reach destination while not having an accident in the shortest time without speeding). From real systems it is also clear that actions taken in the past influence the future evolution of the state, and therefore the future actions.

A general means of modelling this control aspect in a stochastic framework is via controlled Markov chains. We will stick to the discrete-time setting. Let us suppose we are given some distribution λ on state-space I and an action space A such that for every action $a \in A$ we have a transition matrix $P(a)$ and a cost function $c(a) : I \mapsto \mathbb{R}$; note that the cost of an action also depends on the state. This set-up is the input for a Markov decision process, although we may not always get a Markov process! To actually get a process, we have to decide on a methodology to pick actions. This is called a policy which is sequence of functions $u_n : I^{n+1} \mapsto A$ for $n \in \mathbb{Z}_+$; we have assumed that a policy can depend on the current state and the entire past state but in general, it could also depend on the past actions directly (instead of through the state). Each policy then determines how the process evolves, i.e., the probability law $\mathbb{P}^u$ for the random process $\{X_n\}_{n\in\mathbb{Z}_+}$ with values in I by

(1) $\mathbb{P}^u(X_0 = i_0) = \lambda_{i_0}$;
(2) $\mathbb{P}^u(X_{n+1} = i_{n+1}|X_0 = i_0, \ldots, X_n = i_n) = P_{i_n,i_{n+1}}(u_n(i_0, \ldots, i_n))$.

We also distinguish something called a stationary policy $u : I \mapsto A$ where $u_n(i_0, \ldots, i_n) = u(i_n)$ for every $n \in \mathbb{Z}_+$. Note that under a stationary policy, the probability law $\mathbb{P}^u$ makes $\{X_n\}_{n\in\mathbb{Z}_+}$ a DTMC with transition probabilities $P^u_{ij} = P_{ij}(u(i_n))$.

We suppose that a cost $c(i, a) = c_i(a)$ is incurred when action a is chosen in state i. We will assume that $c(i, a) \geq 0$. Then with a policy u we can associate an expected total cost starting in state i which is

$$V^u(i) = \mathbb{E}^u_i\left[\sum_{n=0}^{\infty} c\big(X_n, u_n(X_0, \ldots, X_n)\big)\right];$$

we use the index u to remind us that the probability law is determined by policy u! We also define something call the value function, namely, $V^*(i) = \inf_u V^u(i)$ which is the minimum expected cost starting from i. The basic problem in Markov decision processes is to find out the u that will yield the minimum expected cost.

We will make some technical assumptions to prove our results. These are

(1) *(Assumption 1)*: for all $i, j \in I$, the functions $c_i : A \mapsto [0,\infty)$ and $P_{ij} : A \mapsto [0, 1]$ are continuous;
(2) *(Assumption 2)*: for all i and all $B < \infty$, the set $\{a : c_i(a) \leq B\}$ is compact;
(3) *(Assumption 3)*: for each i, for all but finitely many j, for all $a \in A$ we have $P_{ij}(a) = 0$.

A simple case where (1) and (2) hold is when A is finite set; we will often make this assumption.

We first develop the function $V_n(i)$ for $n \in \mathbb{N}$ and $i \in I$; assume $V_0(i) \equiv 0$ for all i. This is iteratively defined as follows for every $i \in I$:

$$V_1(i) = \inf_{a \in A} c(i,a)$$
$$V_{n+1}(i) = \inf_{a \in A} \{c(i,a) + \sum_{j \in I} P_{ij}(a) V_n(j)\} \quad n \geq 1$$

By induction we can prove that $V_n(i) \leq V_{n+1}(i)$ for all i, i.e., $V_n(i)$ increases to a limit $V_\infty(i)$ (maybe be infinite). Note also that we have

$$V_{n+1}(i) \leq c(i,a) + \sum_{j \in I} P_{ij}(a) V_n(j) \quad \forall a,$$

and by letting $n \to \infty$ we have

$$V_\infty(i) \leq \inf_{a \in A} \{c(i,a) + \sum_{j \in I} P_{ij}(a) V_\infty(j)\}$$

We hazard the guess that $V_\infty(i)$ is $V^*(i)$ as at every time we are choosing the best action assuming there is finite amount of time to go. This is not true in general but does follow under our assumptions.

Lemma 14. *There is a stationary policy u such that*

$$V_\infty(i) = c(i, u(i)) + \sum_{j \in I} P_{ij}(u(i)) V_\infty(j).$$

Proof. If $V_\infty(i) = \infty$, then this true for any policy, so let us assume that $V_\infty(i) \leq B < \infty$. Clearly there is no reason to take actions that will cost more than B. Then we have

$$V_{n+1}(i) = \inf_{a \in K} \left\{ c(i,a) + \sum_{j \in J} P_{ij}(a) V_n(j) \right\}$$

where K is the compact set $\{a : c(i,a) \leq B\}$ and J is the finite set $\{j : P_{ij}(a) \not\equiv 0\}$. Now, by continuity, the infimum is attained and

$$V_{n+1}(i) = c(i, u_n(i)) + \sum_{j \in I} P_{ij}(u_n(i)) V_n(j)$$

for some $u_n(i) \in K$. By compactness of K, there is a convergent subsequence $u_{n_k}(i) \to u(i)$ (for some $u(i) \in K$). Passing to the limit as $k \to \infty$ in the equation above we get the desired result. □

We state without proof a simpler statement of Theorem 4.2.3 from J. Norris' book on Markov chains.

Theorem 54. *Let $\{X_n\}_{n \in \mathbb{Z}_+}$ be a DTMC with transition matrix P and let $(c_i : i \in I)$ be a set of non-negative values/costs associated with each state. Set $\phi_i = \mathbb{E}_i \left[\sum_{n=0}^{\infty} c(X_n)\right]$ (which exists but can be infinite). Then we have the following results:*

(1) *the function* $\phi = (\phi_i : i \in I)$ *satisfies*

$$\phi_i = c(i) + \sum_{j \in I} P_{ij}\phi_j \quad \forall i \in I;$$

(2) *if* $\psi = (\psi_i; i \in I)$ *is another non-negative function that satisfies*

$$\psi_i \geq c(i) + \sum_{j \in I} P_{ij}\psi_j \quad \forall i \in I,$$

then $\psi_i \geq \phi_i$ *for all* $i \in I$.

A main result of this section is the following.

Theorem 55. *We have*

(1) $V_n(i) \uparrow V^*(i)$ *as* $n \to \infty$ *for all* i;

(2) *if* u^* *is any stationary policy such that* $a = u^*(i)$ *minimizes*

$$c(i,a) + \sum_{j \in I} P_{ij}(a)V^*(j)$$

for all $i \in I$, *then* u^* *is optimal, i.e.,* $V^{u^*}(i) = V^*(i)$ *for all* i.

Proof. For any policy u we can write the following

$$\begin{aligned} V^u(i) &= \mathbb{E}_i^u \left[\sum_{n=0}^{\infty} c(X_n, u_n(X_0, \ldots, X_n)) \right] \\ &= c(i, u_0(i)) + \sum_{j \in I} P_{ij}(u_0(i)) V^{u[i]}(j) \end{aligned}$$

where $u[i]$ is the policy constructed using u and given by

$$u[i](i_0, \ldots, i_n) = u_{n+1}(i, i_0, \ldots, i_n) \quad \forall n \in \mathbb{Z}_+$$

Since $V^*(j)$ is a lower bound over all policies we get

$$V^u(i) \geq c(i, u_0(i)) + \sum_{j \in I} P_{ij}(u_0(i))V^*(j) \geq \inf_{a \in A} \left\{ c(i,a) + \sum_{j \in I} P_{ij}(a)V^*(j) \right\},$$

and taking an infimum over all policies we then obtain

$$V^*(i) \geq \inf_{a \in A} \left\{ c(i,a) + \sum_{j \in I} P_{ij}(a)V^*(j) \right\}$$

Now note that $V_0(i) = 0 \leq V^*(i)$ and assuming $V_n(i) \leq V^*(i)$ for all $i \in I$, it is easy to see that $V_{n+1}(i) \leq V^*(i)$. Therefore by induction and taking limits, we find that $V_\infty(i) \leq V^*(i)$ for all $i \in I$.

Let u^* be a stationary policy such that

$$V_\infty(i) \geq c(i, u^*(i)) + \sum_{j \in I} P_{ij}(u^*(i))V_\infty(j).$$

By Lemma 14 we know that such a policy exists. Now

$$V^{u^*}(i) = \mathbb{E}_i^{u^*} \left[\sum_{n=0}^{\infty} c(X_n, u^*(X_n)) \right],$$

so it can be shown that

$$V^{u^*}(i) = c(i, u^*(i)) + \sum_{j \in I} P_{ij}(u^*(i)) V^{u^*}(j)$$

and $V^{u^*}(i) \le V_\infty(i)$ (from Theorem 54). However, we know that $V^*(i) \le V^{u^*}(i)$ which then implies that $V_\infty(i) = V^*(i) = V^{u^*}(i)$ for all $i \in I$. □

Theorem 55 says a few remarkable things: (i) there exists a policy that meets the lower bound and, in addition, the optimal policy is stationary; and (ii) once the value function $V^*(i)$ is obtained (by computing $V_n(i)$ and taking $n \to \infty$), then solving the minimization problem in the second part of Theorem 55 yields the optimal policy. The latter procedure is called value iteration as the iteration is over the value function and the policy is determined at the end of the value iteration procedure; note that we followed a backward induction procedure for value iteration. Note that the value function satisfies a fixed point equation, which is the called the Bellman equation. The optimality rule for control actions is called the Bellman optimality principle. A similar iteration can be performed in the space of stationary policies. Given a stationary policy u and its value function V^u, we can obtain another stationary policy θu by choosing

$$\theta u(i) \in \arg\min_{a \in A} c(i, a) + \sum_{j \in I} P_{ij}(a) V^u(j)$$

Some natural questions are whether this procedure will converge and whether we obtain the optimal policy in the limit. Under some assumptions that does follow.

Theorem 56. *We have*

(1) $V^{\theta u}(i) \le V^u(i)$ *for all* $i \in I$, *i.e., the value function is lowered in each iteration; and*
(2) $V^{\theta^n u} \downarrow V^*(i)$ *as* $n \to \infty$ *for all* i, *provided that*

$$\mathbb{E}_i^{u^*}[V^u(X_n)] \to 0 \text{ as } n \to \infty \quad \forall i \in I,$$

where u^* *is the optimal (stationary) policy.*

Proof. Part (1): Using Theorem 54 we get for all $i \in I$

$$\begin{aligned} V^u(i) &= c(i, u(i)) + \sum_{j \in I} P_{ij}(u(i)) V^u(j) \\ &\ge c(i, \theta u(i)) + \sum_{j \in I} P_{ij}(\theta u(i)) V^u(j), \end{aligned}$$

and since for all $i \in I$ by Theorem 54 we also get for all $i \in I$

$$V^{\theta u}(i) = c(i, \theta u(i)) + \sum_{j \in I} P_{ij}(\theta u(i)) V^{\theta u}(j),$$

it follows from the second part of the same theorem that $V^{\theta u}(i) \le V^u(i)$ for all $i \in I$. Note that iterating this we get $V^{\theta^n u}(i) \le V^{\theta^{n-1} u}(i)$ for all $n \in \mathbb{Z}_+$ so that a limit exists as $n \to 0$ for all $i \in I$. It is also clear that we have

$$V^{\theta^n u}(i) \le c(i, a) + \sum_{j \in I} P_{ij}(a) V^{\theta^{n-1} u}(j) \quad \forall\ i \in I \& a \in A.$$

Part (2): Let $\{X_n\}_{n\in\mathbb{Z}_+}$ be a DTMC obtained by following the optimal stationary policy u^* and let $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$ be the natural filtration of this process. Fix an $N \geq 0$ and for $n = 0, 1, \ldots, N$ define the process

$$M_n^N = V^{\theta^{N-n}u}(X_n) + \sum_{k=0}^{n-1} c(X_k, u^*(X_k)).$$

Then we have

$$\begin{aligned}\mathbb{E}^{u^*}[M_{n+1}^N|\mathcal{F}_n] &= \sum_{j\in I} P_{X_n,j}(u^*(X_n))V^{\theta^{N-n-1}u}(j) + c(X_n, u^*(X_n)) + \sum_{k=0}^{n-1} c(X_k, u^*(X_k))\\ &\geq V^{\theta^{N-n}u}(X_n) + \sum_{k=0}^{n-1} c(X_k, u^*(X_k)) = M_n^N\end{aligned}$$

where we used the decreasing property of $V^{\theta^n u}$ with $a = u^*(X_n)$. Thus, it follows that $\mathbb{E}^{u^*}[M_{n+1}^N] \geq \mathbb{E}^{u^*}[M_n^N]$ so that we have

$$\begin{aligned}V^{\theta^N u}(i) = \mathbb{E}_i^{u^*}[M_0^N] &\leq \mathbb{E}_i^{u^*}[M_N^N]\\ &= \mathbb{E}_i^{u^*}[V^u(X_N)] + \mathbb{E}_i^{u^*}\left[\sum_{n=0}^{N-1} c(X_n, u^*(X_n))\right].\end{aligned}$$

Therefore, if $\mathbb{E}_i^{u^*}[V^u(X_n)] \to 0$ as $n \to \infty$ for all $i \in I$, the right side converges to $V^*(i)$ and so does $V^{\theta^N(u)}(i)$ (we also know $V^{\theta^N u}(i) \geq V^*(i)$ for all $i \in I$ and $N \in \mathbb{Z}_+$). ☐

Thus far we discussed minimization of expected total cost. Thus, if the optimal policy u^* were such that any of the recurrent states have non-zero value, then for any such state i we get $V^*(i) = \infty$. Note that the transient states contribute a finite cost. Since this is a major restriction, we have to consider modified versions of the cost to cover the recurrent case. A first step is to discount the value/cost of the future actions; the idea is that the value dissipates over time. Assume we are given a discount factor $\alpha \in (0, 1)$, then for a policy u we define the expected total discounted cost to be

$$V_\alpha^u(i) = \mathbb{E}_i^u\left[\sum_{n=0}^{\infty} \alpha^n c(X_n, u_n(X_0, \ldots, X_n))\right],$$

and the discounted value function $V_\alpha^*(i) = \inf_u V_\alpha^u(i)$. Note that the discount cost version reduces to the total cost problem if we enhance the state by adding an absorbing state ∂ and define a new Markov decision process by

$$\begin{aligned}\tilde{P}_{ij} &= \alpha P_{ij}(a), \quad \tilde{P}_{i\partial} = 1 - \alpha, \tilde{P}_{\partial\partial} = 1\\ \tilde{c}_i(a) &= c_i(a), \quad\ \ \tilde{c}_\partial(a) = 0\end{aligned}$$

Thus, the new process follows the old one until a geometrically distributed time (with parameter α and duration at least 1) when it jumps to ∂ and stays there without accumulating any cost/value. For later use we will deemed the value function for policy u in the new process $\tilde{V}(u)$ and the optimal value function $\tilde{V}^*$.

Similar to the total cost case we can define $V_{0,\alpha}(i) = 0$ and, inductively,

$$V_{n+1,\alpha}(i) = \inf_{a\in A}\left\{c(i,a) + \alpha\sum_{j\in I} P_{ij}(a)V_{n,\alpha}(i)\right\},$$

and given a stationary policy u, devise a new stationary policy $\theta_\alpha u$ by choosing

$$\theta_\alpha u(i) \in \arg\min_{a\in A} c(i,a) + \alpha\sum_{j\in I} P_{ij}(a)V_\alpha^u(i).$$

Then we have the following result.

Theorem 57. *Assume that the cost function $c(i,a)$ is uniformly bounded. Then we have*

(1) *$V_{n,\alpha}(i) \uparrow V_\alpha^*(i)$ as $n\to\infty$ for all $i\in I$;*
(2) *the value function V_α^* is the unique bounded solution to*

$$V_\alpha^*(i) = \inf_{a\in A}\left\{c(i,a) + \alpha\sum_{j\in I} P_{ij}(a)V_\alpha^*(i)\right\};$$

(3) *let u^* be the stationary policy such that $a = u^*(i)$ minimizes*

$$c(i,a) + \alpha\sum_{j\in I} P_{ij}(a)V_\alpha^*(j) \quad \forall i\in I,$$

then u^* is optimal, i.e., $V_\alpha^{u^*}(i) = V_\alpha^*(i)$ for all $i\in I$; and

(4) *for all stationary policies u we have $V_\alpha^{\theta_\alpha^n u}(i) \downarrow V_\alpha^*(i)$ as $n\to\infty$ for all $i\in I$.*

Proof. Since $V_\alpha^u = \tilde V^u$, $V_\alpha^* = \tilde V^*$ and $\theta_\alpha u = \tilde\theta_u$ (where $\tilde\theta$ is the θ operation in the new total cost Markov decision process), parts (1), (2) and (3) follow from Theorems 55 and 56, except for the uniqueness claim in (2). However, note that for any bounded solution V to the said equation in part (2), there is a stationary policy $\tilde u$ such that

$$V(i) = c(i,\tilde u(i)) + \alpha\sum_{j\in I} P_{ij}(\tilde u(i))V(j),$$

then by Theorem 4.2.5 in J. Norris' book, we have $V = V_\alpha^u$. Therefore $\theta_\alpha\tilde u = \tilde u$ and so if part (4) holds, then $\tilde u$ is optimal and $V = V_\alpha^*$.

Since $c(i,a) \le B < \infty$ (for some B), we have for any stationary policy u that

$$V_\alpha^u(i) = \mathbb{E}_i^u\left[\sum_{n=0}^\infty \alpha^n c(X_n, u(X_n))\right] \le \frac{B}{1-\alpha};$$

in fact this holds for any policy. Therefore, we also have

$$\tilde{\mathbb{E}}_i^{u^*}[\tilde V^u(X_n)] = \alpha^n\mathbb{E}_i^{u^*}[V_\alpha^u(X_n)] \le \frac{B}{1-\alpha}\alpha^n \to_{n\to\infty} 0,$$

where $\tilde{\mathbb{E}}$ refers to expectation in the new Markov decision process. Thus, Theorem 56 proves the required result. □

We could, instead, consider minimizing the long-run average cost as our criterion. For a policy u and for $n\in\mathbb{N}$, define $\bar V_n^u(i)$ as

$$\bar V_n^u(i) = \mathbb{E}_i^u\left[\frac{1}{n}\sum_{k=0}^{n-1} c(X_k, u_k(X_0,\dots,X_k))\right].$$

As in the discounted cost case we will assume $|c(i,a)| \leq B < \infty$ for all i and a, which makes $|\bar{V}_n^u(i)| \leq B$ but a limit may not exist. Therefore, the exact quantity that we'll attempt to find is $\bar{V}^*(i) = \inf_u \limsup \bar{V}_n^u(i)$. If we had a stationary policy u that resulted in a DTMC that was positive recurrent with equilibrium distribution π_u, then we know that $\lim_{n\to\infty} \bar{V}_n^u = \mathbb{E}_{\pi_u}[c(X, u(X))]$ and we could then try to find the best such policy. However, as before it is not clear that there is an optimal policy and a stationary one at that. In general, the long-run average cost criterion is tricky and the standard approach is to find conditions for a stationary policy to be optimal. The most general result here is the following.

Theorem 58. *Suppose we can find a constant $\bar{V}^*$ and a bounded function $W(i)$ such that*

$$\bar{V}^* + W(i) = \inf_{a\in A}\left\{c(i,a) + \sum_{j\in I} P_{ij}(a)W(j)\right\} \quad \forall i \in I.$$

Let u^* be a strategy such that $a = u^*(i)$ achieves the infimum above for each i. Then

1. *$\bar{V}_n^{u^*}(i) \to \bar{V}^*$ as $n \to \infty$ for all $i \in I$; and*
2. *$\liminf_{n\to\infty} \bar{V}_n^u(i) \geq \bar{V}^*$ for all i, for all u.*

Note that the optimal cost does not depend on i, which is to be expected from our earlier discussion on the ergodic theorem for Markov chains.

Proof. Fix a policy u and let $\{X_n\}_{n\in\mathbb{Z}_+}$ be the resulting process. Define $U_n = u_n(X_0, \ldots, X_n)$ (note that U_n is $\sigma(X_0,\ldots,X_n)$ measurable) and

$$M_n = W(X_n) - n\bar{V}^* + \sum_{k=0}^{n-1} c(X_k, U_k).$$

Setting $\mathcal{F}_n = \sigma(X_0,\ldots,X_n)$, we then obtain

$$\begin{aligned}\mathbb{E}^u[M_{n+1}|\mathcal{F}_n] &= M_n + \left\{c(X_n,U_n) + \sum_{j\in I} P_{X_n,j}(U_n)W(j)\right\} - (\bar{V}^* + W(X_n)) \\ &\geq M_n\end{aligned}$$

with equality if $u = u^*$; thus, $\{M_n\}_{n\in\mathbb{Z}_+}$ is a submartingale in general but a martingale when $u = u^*$. Therefore

$$W(i) = \mathbb{E}_i^u[M_0] \leq \mathbb{E}_i^u[M_n] = \mathbb{E}_i^u[W(X_n)] - n\bar{V}^* + n\bar{V}_n^u(i).$$

This then yields

$$\bar{V}^* \leq \bar{V}_n^u(i) + 2\frac{\sup_i |W(i)|}{n}.$$

This proves part (2) by letting $n \to \infty$. When $u = u^*$ we also have

$$\bar{V}_n^u(i) \leq \bar{V}^* + 2\frac{\sup_i |W(i)|}{n},$$

which proves part (1). □

Note that $|M_{n+1} - M_n| = |W(X_{n+1}) - W(X_n) - \bar{V}^* + c(X_n,U_n)| \leq 2(\sup_{i\in I}|W(i)| + B)$. Now using a theorem on martingales known as the martingale stability theorem, one can show a much stronger result, namely, $\liminf \frac{1}{n}\sum_{k=0}^{n-1} c(X_k,U_k) \geq \bar{V}^*$ *a.s.*, and equality for the optimal poilcy u^*, if u^* results in a positive recurrent DTMC.

If I and A are finite and one assumes that $P(a)$ is irreducible for every $a \in A$, then one can prove that the long-run average cost problem can be viewed as the discounted cost problem where the discount factor $\alpha \uparrow 1$. For this one fixes a specific state, called 0, and defines the following $W_\alpha^*(i) = V_\alpha^*(i) - V_\alpha^*(0)$ and $\tilde{V}_\alpha^* = (1-\alpha)V_\alpha^*(0)$. Then $W_\alpha(i) \to W(i) - W(0)$ and $\tilde{V}_\alpha^* \to \bar{V}^*$ as $\alpha \uparrow 1$.

The results for the discounted cost problem generalize to the scenario where Assumption 3 can be replaced by a condition that P_{ij} is uniformly continuous in a uniformly over j. If the action/control set A is finite, then we can remove Assumption 3 for both the discounted cost and long-run average cost problem results. The association between the discounted cost and long-run average cost problems with this new assumption carries through if $W_\alpha^*(i)$ is uniformly bounded for all $i \in I$ and $a \in A$; a sufficient condition for the latter is uniform (over i and α) boundedness of the mean first passage time to state 0 from any state $i \in I$ (including 0) for the optimal policy for each α.

7. Multi-armed Bandits and Applications

The sources for this section are:

- P. R. Kumar and P. Varaiya, "Stochastic systems: Estimation, identification and adaptive control," Prentice Hall, Englewood Cliffs, N. J., 1986.
- T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," Advances in Applied Mathematics, 6, 4–22, 1985.
- P. Auer, N. Cesa-Bianchi and P. Fischer, "Finite time analysis of the multi-armed bandit problem," Machine learning, 47 2/3:235–256, 2002.
- N. Cesa-Bianchi and G. Lugosi, "Prediction, learning, and games," Cambridge University Press, Cambridge, 2006.

We'll apply the results of the previous section to a specific problem called the multi-armed bandit problem. We will also be interested in non-classical formulations. Applications of this then follow. The general stochastic formulation is the following. We are given $N \in \mathbb{N}$ stochastic processes $\{X_k^n\}_{k\in\mathbb{Z}_+}$ and $n = 1, 2, \ldots, N$ taking values in $\mathbb{N}$ (after remapping if necessary). At each time k, we choose a control/action a from the action set $A = \{1, 2, \ldots, N\}$. If $U_k = n$, then the state $(X_{k+1}^1, \ldots, X_{k+1}^N)$ at time $k+1$ is given by

$$\begin{aligned} X_{k+1}^m &= X_k^m \quad \forall m \in A \setminus \{n\} \\ X_{k+1}^n &= j \quad \text{ with probability } P_{X_k^n, j} \end{aligned}$$

where $P = \{P_{ij}, i, j \in \mathbb{N}\}$ is a pre-specified transition matrix. The state is observed and the goal is to choose policy u that determines $\{U_k\}_{k\in\mathbb{Z}_+}$ so that

$$\mathbb{E}^u\left[\sum_{k=0}^{\infty} \alpha^k R(X_k^{U_k})\right]$$

is maximized, where $R : I \mapsto \mathbb{R}$ is a bounded reward function and $\alpha \in (0, 1)$.

In words we can describe the problem as follows: at each time we can pick one process to evolve in a Markovian fashion while the remaining processes remain frozen, and this must be done so as to maximize the expect total discounted reward where the reward at each time is a function of the state of the process that's picked. Specific examples of this will be mentioned after the main result is presented. We have for this problem what is called the exploration-exploitation tradeoff. We would like to play the arm whose state gives us the highest reward at any given time but its probabilistic evolution may lead to a bad system state in the future that may take a long time to recover from. Therefore, one needs to explore the states and find out good states (so that rewards are high for some period of time) and balance this with exploiting current good states.

We start by noting that Theorem 57 applies to this problem in the form generalized to countable state-space and finite action space. Therefore, the value function $V_\alpha^* : \mathbb{N}^N \mapsto \mathbb{R}$ solves the following dynamic programming equation

$$V_\alpha^*(x^1, \ldots, x^N) = \max_{a\in\{1,\ldots,N\}} \left\{ R(x^a) + \alpha \sum_{j=1}^{\infty} P_{x^a,j} V_\alpha^*(x^1, \ldots, x^{a-1}, j, x^{a+1}, \ldots, x^N) \right\}$$

The main result is the following theorem.

Theorem 59. *Let $\{Y_k\}_{k\in\mathbb{Z}_+}$ be a DTMC with transition matrix P taking values in $\mathbb{N}$. For each $i \in \mathbb{N}$ define*

$$\nu(i,\tau) := \frac{\mathbb{E}_i\left[\sum_{k=0}^{\tau-1}\alpha^k R(Y_k)\right]}{\mathbb{E}_i\left[\sum_{k=0}^{\tau-1}\alpha^k\right]} \text{ and}$$
$$\gamma(i) := \sup_{1\le\tau\le\infty} \nu(i,\tau)$$

Then the particular halting time

$$\tau_i : \inf\{k \in \mathbb{N} : \gamma(Y_k) < \gamma(i)\}$$

attains the supremum above for every $i \in I$.

For the bandit problem the index rule policy that chooses the process with the largest current index, i.e.,

$$U_k \in \arg\max_{l=1,\dots,N} \gamma(X_k^l)$$

is optimal.

The index function γ is called Gittins' index and the policy is the dynamic index assignment policy. For the finite state case there is means to calculate the Gittins index - see Chapter 11, Section 7 of the Kumar-Varaiya book.

Proof. We start by proving the first statement. Assume a simpler problem where we only have two arms with one of them fixed/static, at a position that yields reward R^*. Here the only state we need to be concerned with is the state of the stochastic arm, say 1. Thus, the value function is now given by

$$V_\alpha^*(i) = \max\left(R(i) + \alpha\sum_{j=1}^{\infty} P_{ij}V_\alpha^*(j), R^* + \alpha V_\alpha^*(i)\right),$$

where one plays the static arm if the second term equals the maximum. Now one can observe that if one chooses the static arm at any state i, since the state does not change, one continues with the static arm forever. States $i \in I$ for which it is best to use the static arm are such that $V_\alpha^*(i) = \frac{R^*}{1-\alpha}$, which we deem as M. Here we have

$$V(i,M) = \max\left(R(i) + \alpha\sum_{j\in I} P_{ij}V(j,M), M\right)$$

Thus, it is sufficient in this case to consider halting times (when we stop playing the stochastic arm) and for non-trivial solutions we assume that $R^* \in [\inf_{i\in\mathbb{N}} R(i), \sup_{i\in\mathbb{N}} R(i)]$. For a halting time τ, the cost is (abusing notation)

$$V(i,\tau,M) := V(i,\tau,R^*) = \mathbb{E}_i\left[\sum_{i=0}^{\tau-1}\alpha^t R(Y_k) + \alpha^\tau M\right],$$

where we can also write $\alpha^\tau M$ as $\sum_{t=\tau}^{\infty}\alpha^t R^*$. The optimal stationary policy for this two arm case (or the 1.5 arm case as it is sometimes called) for each M is easy to determine. We can partition the state space into three sets, namely,

(1) *strict continuation set*, which is given by $C_M = \{i : V(i,\tau,M) > M\}$, and where one plays the stochastic arm;
(2) *strict stopping set.* which is given by $S_M = \{i : M > R(i) + \alpha \sum_{j\in I} V(j,\tau,M)\}$, and where one plays the static arm; and
(3) *indifferent set*, which is given by $\partial_M = \{i : M = R(i) + \alpha \sum_{j\in I} V(j,\tau,M)\}$, and where one can play either arm.

Then the optimal policy is given by the stopping time $\tau(i,M)$ which is the first passage time to S_M. Note that as M increases, C_M becomes smaller, S_M becomes larger, and $\tau(i,M)$ decreases.

For halting time $\tau = 0$, we know that $V(i,\tau,M) \equiv M$ for all i. Therefore, the only way it is optimal to play the static arm at time 0 (for state i) is when $\sup_{\tau\geq 1} V(i,\tau,M) \leq M$. Now we can show the following

$$\begin{aligned} V(i,\tau,M) &= \mathbb{E}_i\left[\sum_{i=0}^{\tau-1}\alpha^t R(Y_k)\right] + \mathbb{E}_i[\alpha^\tau]M \\ &= \frac{\mathbb{E}_i\left[\sum_{i=0}^{\tau-1}\alpha^t R(Y_k)\right]}{\mathbb{E}_i[\sum_{i=0}^{\tau-1}\alpha^i]}\frac{\mathbb{E}_i[1-\alpha^\tau]}{1-\alpha} + \mathbb{E}_i[\alpha^\tau]M \\ &= \frac{\nu(i,\tau)}{1-\alpha}\mathbb{E}_i[1-\alpha^\tau] + \mathbb{E}_i[\alpha^\tau]M \\ &= \frac{\nu(i,\tau)}{1-\alpha}(1-\mathbb{E}_i[\alpha^\tau]) + \mathbb{E}_i[\alpha^\tau]M, \end{aligned}$$

which then implies that we need $\frac{\nu(i,\tau)}{1-\alpha} \leq M$ for all halting times $\tau \geq 1$. This then implies we need $\gamma(i) \leq (1-\alpha)M$. Thus, if we set $(1-\alpha)M(i) = \gamma(i)$ for each $i \in I$, then $V(i,M(i)) = M(i)$ solves the dynamic programming optimality equation from Theorem 57; note that it is a bounded solution. The characterization in terms of τ_i is now clear from the discussion above and the first part is proved.

We will now prove the second part, which will be done using an interchange argument and forward induction. Let u be the specific index rule policy for which

$$U_k = n \text{ if } n = \min\left\{m \in \arg\max_{l=1,\dots,N}\gamma(X_k^l)\right\},$$

where we break ties by choosing the smallest label. We now show that π, which is a stationary policy, is optimal.

Suppose that u chooses process i at time $k=0$, i.e., $U_0 = i$. Consider $j \neq i$ and let $\hat{u}$ be a non-stationary policy that chooses process j at time 0 and thereafter proceeds according to policy u. Under policy $\hat{u}$, process j will be chosen at times $0,1,\dots,\tau_j - 1$ where

$$\tau_j := \begin{cases} \inf\{k \geq 1 : \gamma(X_k^j) \leq \gamma(X_0^i)\} & \text{if } j > i; \\ \inf\{k \geq 1 : \gamma(X_k^j) < \gamma(X_0^i)\} & \text{if } j < i. \end{cases}$$

Thereafter, under $\hat{u}$, process i will be chosen at least for times $k = \tau_j, \tau_j+1, \dots, \tau_j+\tau_i-1$ where

$$\tau_j + \tau_i := \inf\{k \geq \tau_j + 1 : \gamma(X_k^i) \leq \gamma(X_{\tau_j}^i) = \gamma(X_0^i)\}.$$

Consider now another policy $\tilde{u}$, which chooses process i at times $k = 0,1,\dots,\tau_i - 1$, then process j at times $k = \tau_i, \tau_i+1,\dots,\tau_i+\tau_j-1$, and thereafter does exactly what $\hat{u}$ does.

Comparing the cost of policies $\hat{u}$ and $\tilde{u}$ we have

$$\begin{aligned}
&\mathbb{E}^{\tilde{u}}\left[\sum_{k=0}^{\infty}\alpha^k R(X_k^{U(k)})\right] - \mathbb{E}^{\hat{u}}\left[\sum_{k=0}^{\infty}\alpha^k R(X_k^{U(k)})\right] \\
&= \mathbb{E}^{\tilde{u}}\left[\sum_{k=0}^{\tau_i}\alpha^k R(X_k^i) + \sum_{k=\tau_i}^{\tau_i+\tau_j-1}\alpha^k R(X_k^j)\right] - \mathbb{E}^{\hat{u}}\left[\sum_{k=0}^{\tau_j}\alpha^k R(X_k^j) + \sum_{k=\tau_j}^{\tau_i+\tau_j-1}\alpha^k R(X_k^i)\right] \\
&= \mathbb{E}^{\tilde{u}}\left[\sum_{k=0}^{\tau_i}\alpha^k R(X_k^i)\right] + \mathbb{E}^{\tilde{u}}\left[\alpha^{\tau_i}\mathbb{E}^{\hat{u}}\left[\sum_{k=0}^{\tau_j-1}\alpha^k R(X_k^j)\right]\right] \\
&\quad - \mathbb{E}^{\hat{u}}\left[\sum_{k=0}^{\tau_j}\alpha^k R(X_k^j)\right] - \mathbb{E}^{\hat{u}}\left[\alpha^{\tau_j}\mathbb{E}^{\tilde{u}}\left[\sum_{k=0}^{\tau_i-1}\alpha^k R(X_k^i)\right]\right] \\
&= \mathbb{E}^{\tilde{u}}\left[\sum_{k=0}^{\tau_i-1}\alpha^k R(X_k^i)\right]\left(1-\mathbb{E}^{\hat{u}}\left[\alpha^{\tau_j}\right]\right) - \mathbb{E}^{\hat{u}}\left[\sum_{k=0}^{\tau_j-1}\alpha^k R(X_k^j)\right]\left(1-\mathbb{E}^{\tilde{u}}\left[\alpha^{\tau_i}\right]\right) \\
&\geq \gamma(X_0^i)\mathbb{E}^{\tilde{u}}\left[\sum_{k=0}^{\tau_i-1}\alpha^k\right]\left(1-\mathbb{E}^{\hat{u}}\left[\alpha^{\tau_j}\right]\right) - \gamma(X_0^j)\mathbb{E}^{\hat{u}}\left[\sum_{k=0}^{\tau_j-1}\alpha^k\right]\left(1-\mathbb{E}^{\tilde{u}}\left[\alpha^{\tau_i}\right]\right) \\
&= (\gamma(X_0^i)-\gamma(X_0^j))(1-\alpha)\left(1-\mathbb{E}^{\hat{u}}\left[\alpha^{\tau_j}\right]\right)\left(1-\mathbb{E}^{\tilde{u}}\left[\alpha^{\tau_i}\right]\right) \geq 0,
\end{aligned}$$

where we used the definition of the Gittins' index in the penultimate inequality; the first term is equal as $\tilde{u}$ follows u until $0 \leq k \leq \tau_i - 1$, while the second term is smaller than the term with the Gittins' index. Therefore, $\tilde{u}$, which coincides with u for $0 \leq k \leq \tau_i - 1$, is an improvement over $\hat{u}$. If at time τ_i, $\tilde{u}$ does not follow u, then by shifting the time origin to τ_i, we can repeat the argument above to obtain yet another policy that improves the reward while coinciding with u even at time τ_i. In this way we can obtain policies which coincide with u over arbitrary large initial segments of time, and which are all improvements over $\hat{u}$. We can, thus, conclude that u itself is better than $\hat{u}$. This shows that following u at time $k = 0$ is optimal. We can repeat this argument from any point that we diverge from u by shifting the origin to that time. Hence, this argument clinches the result. □

There are multiple proofs of the optimality of the Gittins' index policy, each proof reveals something new and allows for generalizations in the different directions. Some good sources for this are:

- E. Frostig, G. Weiss, "Four proofs of Gittins' multiarmed bandit theorem," Appl. Probab. Trust (1999).
- J. N. Tsitsiklis, "A short proof of the Gittins' index theorem," Ann. Appl. Probab. 4 (1994), no. 1, 194–199.

The multi-armed bandit problem models many real problems. First is the eponymous problem of playing slot machines. Say there are ten slot machines and *a priori* we have no idea of the actual chances of winning on any of them but you know the statistics, i.e., each machine has a fixed probability of winning that is chosen independently from some know distribution. Assuming we have a large pot of money to back us up, we have to decide which machine to play at any given time. The objective is to maximize the expected winnings. Only by playing a machine can we estimate the probability of winning. Over multiple plays of each machine, we get a better idea of the unknown probability of winning. A better example is clinical trials where we assume that there are three drugs/treatment options whose probability of successful treatment is unknown (except for a gross statistical knowledge). The goal is to design a sequence of trials to maximize the expected probability of successful treatment. Another application is in processing of jobs with different rewards where the processor only gets a finite amount of time to complete a task and the amount of time to complete any given task is random with a known distribution. Once a task is finished another one of the same kind appears. Routing problems are also examples if one assumes that the routes never overlap, then we need to decide which route to use for a given packet where the chance of it arriving at the destination by a given time is stochastic. The multi-armed bandit problem has many interesting modern-day applications as well. The first is related to online advertisements. An online advertiser stands to make revenue if and only if a user clicks on the ad and then buys the product being advertised. The advertisers negotiate with a search engine to come with the price that the advertisers have to pay to display the ad, and this revenue is paid only if the user buys a product. Thus, each ad is like a slot machine for the search engine, with some probability (which is user dependent) the ad gives a reward and otherwise nothing. Assuming that the search engine can only display one ad, this then becomes the standard multi-armed bandit problem. The second modern application is to dynamic spectrum access, perhaps for cognitive radio applications. Assume that we have many frequency bands available for transmission. These channels are for so-called primary users, e.g., TV broadcast, who take precedence. A secondary user has to sample the channel space to find an empty channel to use or one where the likelihood of a primary user operating is low at the current time. In addition, the rate that the secondary user receives (the reward) could be dependent on the channel. This can then be modeled by the multi-armed bandit problem. There are lots of problems in microeconomics that fall within the multi-armed bandit umbrella. An example is the work on optimal dynamic auctions by M. Pai and R. Vohra in 2008. Some other papers that consider applications in economics are: D. P. Foster and R. V. Vohra, "Asymptotic calibration," Biometrika 85 (1998), no. 2, 379–390; S. Hart and A. Mas-Colell, "A general class of adaptive strategies," J. Econom. Theory 98 (2001), no. 1, 26–54; S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," Econometrica 68 (2000), no. 5, 1127–1150; and D. Fudenberg and D. K. Levine, "The theory of learning in games," MIT Press Series on Economic Learning and Social Evolution, 2, MIT Press, Cambridge, MA, 1998.

We will discuss one application in some detail. First is the slot machine problem. We have N slot machines, $M_1, \dots, M_N$. For machine M_i, the probability of success (reward 1) on a play is θ_i with the probability of failure (reward 0) being $1 - \theta_i$. The parameters $(\theta_1, \dots, \theta_N)$ are assumed to be independent random variables with prior distributions given $P_1, \dots, P_N$. At each time only one machine can be played. As we play the different machines, the conditional probability distribution of the success probability parameter changes. Let $P_1^k, \dots, P_N^k$ be these conditional probability distributions given the observed past history up to time k. Assume that we choose machine M_n to play at time k. Then at time $k+1$, the conditional probabilities will be:

$$P_m^{k+1} = P_m^k \quad \forall m \in \{1, \dots, N\} \setminus \{n\}$$

$$P_n^{k+1}(d\theta_n) = \begin{cases} \frac{\theta_n P_n^k(d\theta_n)}{\int_0^1 \theta P_n^k(d\theta)} & \text{with probability } \int_0^1 \theta P_n^k(d\theta) \\ \frac{(1-\theta_n)P_n^k(d\theta_n)}{\int_0^1 (1-\theta) P_n^k(d\theta)} & \text{with probability } 1 - \int_0^1 \theta P_n^k(d\theta) \end{cases}$$

The expected reward for playing machine M_n at time k is $\int_0^1 \theta P_n^k(d\theta)$. The book by Kumar and Varaiya discusses many more applications in Chapter 11.

7.1. **No Regret Formulation.** We now discuss a variation of the multi-armed bandit problem initiated by Lai and Robbins. Here the aim to achieve asymptotically efficient adaptive allocation rules. The (simplest) set-up is the following. There are N arms and each play of arm i produces an $i.i.d$ reward with unknown (finite) mean μ_i. The rewards are assumed to be independent across arms as well, though not identically distributed. As always a policy, or allocation strategy, u is a rule that chooses the arm to play based on past rewards and actions/plays. Let $T_i(n)$ be the number of times machine i is played by policy u until time n. Then the regret of u after n plays is defined by

$$\mu^* n - \sum_{j=1}^{N} \mu_j \mathbb{E}^u[T_j(n)] \text{ where } \mu^* := \max_{1 \le i \le N} \mu_i,$$

i.e., it is the expected loss when compared to a genie-aided strategy that only plays the best arm; note that $\sum_{j=1}^N T_j(n) = n$ for all n. If the requirement for u is such that $\lim_{n \to 0} \frac{1}{n} \sum_{j=1}^N \mu_j \mathbb{E}^u[T_j(n)] = \mu^*$, i.e., the policy be asymptotically optimal (also known as Hannan consistency), then the class of policies one should explore will include ones that use the following approach: (exploration phase) experiment with the different arms sufficient number of times (sublinear in n) to get a good estimate of the mean and then (exploitation phase) play the best arm. Then the design problem changes to finding out a class of policies that will minimize the regret because one can easily construct a policy that yields the optimal long-run average reward. The minimum regret was first characterized by Lai and Robbins, and they found, for specific families of reward distributions (indexed by a single real parameter), policies satisfying

$$\mathbb{E}^u[T_j(n)] \le \left(\frac{1}{D(p_j || p^*)} + o(1) \right) \log(n)$$

where $o(1) \to 0$ as $n \to \infty$ and

$$D(p_j || p^*) := \int p_j(x) \log \left(\frac{p_j(x)}{p^*(x)} \right) dx$$

is the relative entropy (Kullback-Leibler divergence) between the reward density $p_j(\cdot)$ of an suboptimal machine j and the reward density p^* of the machine with the highest reward (in expectation). Thus, under this class of policies, the optimal machine is played exponentially more often than any other machine, asymptotically. Lai and Robbins also proved that this regret is the best possible when the reward distributions satisfy some mild assumptions, i.e., $\mathbb{E}^u[T_j(n)] \ge \log(n)/D(p_j||p^*)$ asymptotically for all (Hannan) consistent policies.

We will know present a simple policy that achieves this logarithmic regret (source is Auer-CesaBianchi-Fischer paper). The policy (called UCB) is the following:

- *Initialization*: Play each arm once;
- *Iterative step*: Play machine j that maximizes
$$\bar{x}_n^j + \sqrt{\frac{2\log(n)}{n^j}}$$
where $\bar{x}_n^j$ is the average reward obtained from machine j, n^j the number of times machine j has been played so far, and n is the total number of plays so far. After every play, update the variables and repeat.

Define $\Delta_i := \mu^* - \mu_i$ to be (expected) loss of machine i when compared to the optimum reward. Then the result is the following.

Theorem 60. *For all $N > 1$, if policy UCB is run on N arms having arbitrary reward distributions $P_1, \dots, P_N$ with support in $[0,1]$, then its expected regret after any number of n plays is at most*
$$\left(8 \sum_{i:\mu_i<\mu^*} \frac{\log(n)}{\Delta_i}\right) + \left(1 + \frac{\pi^2}{3}\right)\left(\sum_{j=1}^N \Delta_j\right)$$
where $\mu_1, \dots, \mu_N$ are the means of the distributions $P_1, \dots, P_N$.

Proof. Let $c_{t,s} := \sqrt{2\log(t)/s}$. Also define by the following
$$\bar{X}_n^i = \frac{1}{n}\sum_{t=1}^n X_t^i$$
where $\{X_t^i\}_{t\in\mathbb{N}}$ is the random process of rewards for arm i if it is played successively. For any machine i, we upper bound $T_i(n)$ on any sequence of plays. Let I_t denote the machine played at time t, then we have with l being an arbitrary positive integer that
$$\begin{aligned}
T_i(n) &= 1 + \sum_{t=N+1}^n 1_{\{I_t=i\}} \\
&\le l + \sum_{t=N+1}^n 1_{\{I_t=i, T_i(t-1)\ge l\}} \\
&\le l + \sum_{t=N+1}^n 1_{\{\bar{X}^*_{T^*(t-1)} + c_{t-1,T^*(t-1)} \le \bar{X}^i_{T_i(t-1)} + c_{t-1,T_i(t-1)},\, T_i(t-1)\ge l\}} \\
&\le l + \sum_{t=N+1}^n 1_{\{\min_{0<s<t} \bar{X}^*_s + c_{t-1,s} \le \max_{l\le s_i<t} \bar{X}^i_{s_i} + c_{t-1,s_i}\}} \\
&\le l + \sum_{t=1}^\infty \sum_{s=1}^{t-1} \sum_{s_i=l}^{t-1} 1_{\{\bar{X}^*_s + c_{t,s} \le \bar{X}^i_{s_i} + c_{t,s_i}\}}
\end{aligned}$$
Now we observe that $\bar{X}^*_s + c_{t,s} \le \bar{X}^i_{s_i} + c_{t,s_i}$ implies that at least one of the following must hold
$$\begin{aligned}
\bar{X}^*_s &\le \mu^* - c_{t,s} \\
\bar{X}^i_{s_i} &\ge \mu_i + c_{t,s_i} \\
\mu^* &< \mu_i + 2c_{t,s_i}.
\end{aligned}$$

Note that the converse statement is the following

$$\bar{X}^i_{s_i} + c_{t,s_i} < \mu_i + 2c_{t,s_i} \leq \mu^* < \bar{X}^*_s + c_{t,s}.$$

We can bound the probability of the first two events by using a version of the Azuma-Hoeffding inequality since both $s\bar{X}^*_s - s\mu^*$ and $s_i\bar{X}^i_{s_i} - s_i\mu_i$ are martingales null at 0 with increments in $[-\mu^*, 1-\mu^*]$ and $[-\mu_i, 1-\mu_i]$, respectively. This then yields the following

$$\mathbb{P}(\bar{X}^*_s \leq \mu^* - c_{t,s}) \leq \exp\left(-2\frac{s^2 c^2_{t,s}}{s}\right) = e^{-4\log(t)} = t^{-4}$$

$$\mathbb{P}(\bar{X}^i_{s_i} \leq \mu_i + c_{t,s_i}) \leq \exp\left(-2\frac{s_i^2 c^2_{t,s_i}}{s_i}\right) = e^{-4\log(t)} = t^{-4}$$

Also note that we have

$$\mu^* - \mu_i - 2c_{t,s_i} = \mu^* - \mu_i - 2\sqrt{\frac{2\log(t)}{s_i}} \geq \mu^* - \mu_i - \sqrt{\frac{\log(t)}{\log(n)}}\Delta_i \geq \mu^* - \mu_i - \Delta_i = 0$$

for $s_i \geq 8\log(n)/\Delta_i^2$ (since $n \geq t$). Thus, if we take $l = \lceil 8\log(n)/\Delta_i^2 \rceil$, then we cannot have $\mu^* < \mu_i + 2c_{t,s_i}$ and only the first two inequalities can hold. Thus, we get

$$\begin{aligned}
\mathbb{E}[T_i(n)] &\leq \lceil 8\log(n)/\Delta_i^2 \rceil + \sum_{t=1}^{\infty}\sum_{s=1}^{t-1}\sum_{s_i=l}^{t-1}\left(\mathbb{P}(\bar{X}^*_s \leq \mu^* - c_{t,s}) + \mathbb{P}(\bar{X}^i_{s_i} \leq \mu_i - c_{t,s_i})\right) \\
&\leq \lceil 8\log(n)/\Delta_i^2 \rceil + \sum_{t=1}^{\infty}\sum_{s=1}^{t-1}\sum_{s_i=l}^{t-1} 2t^{-4} \\
&\leq \lceil 8\log(n)/\Delta_i^2 \rceil + \sum_{t=1}^{\infty}\sum_{s=1}^{t-1}\sum_{s_i=1}^{t-1} 2t^{-4} \\
&\leq 8\log(n)/\Delta_i^2 + 1 + \frac{\pi^2}{3}
\end{aligned}$$

which completes the proof. □

This does not yield the best regret but is presented for ease of analysis. In the homework, we will look at other strategies.

We will now present a different formulation of the multi-armed bandit problem. Until now the rewards from the arms changed in via a fixed random process. However, these could be chosen in an adversarial fashion, i.e., possibly a worst-case sequence of rewards. Even in this scenario we'd like to construct a policy that has the minimum regret. Here the rewards are assumed to be generated using a sequence $\{Y_t\}_{t\in\mathbb{N}}$ where Y_t can depend on all the actions until time $t-1$. The reward at time t for action i is $g(i, Y_t)$ for some function g taking values in $[0, 1]$; note that we don't get to see Y_t, else we could evaluate the gain for every arm. The objective is now to minimize weak regret, i.e., minimize loss of reward when compared to $\max_{i=1,\dots,N}\sum_{t=1}^{n} g(i, Y_t)$; note that the comparison is with a strategy that plays a fixed arm. The strategy that we will present is a randomized strategy that chooses action I_t at time t using a probability distribution λ_t where λ_t^i is the probability of choosing arm i. We will also have time-varying weights for each actions, which we denote by w_t^i for arm i at time $t \in \mathbb{Z}_+$. We choose three non-negative numbers β, η and γ, all less than 1. Then the algorithm (called MAB) is the following.

- *Initialization phase*: Set $w_0^i \equiv 1$ and $\lambda_1^i \equiv 1/N$;
- *Iteration phase*: For each $t \in \mathbb{N}$ repeat the following:
  1. Pick arm $I_t \in \{1, \ldots, N\}$ according to distribution λ_t;
  2. Calculate the estimated gains as follows

$$\hat{g}(i, Y_t) = \tilde{g}(i, Y(t)) + \frac{\beta}{\lambda_t^i} = \begin{cases} \frac{g(i,Y_t)+\beta}{\lambda_t^i} & \text{if } I_t = i \\ \frac{\beta}{\lambda_t^i} & \text{otherwise;} \end{cases}$$

  3. Update the weights $w_t^i = w_{t-1}^i e^{\eta \hat{g}(i,Y_t)}$ and $W_t = \sum_{i=1}^n w_t^i$;
  4. Calculate the updated probability distribution

$$\lambda_{t+1}^i = (1-\gamma)\frac{w_t^i}{W_t} + \frac{\gamma}{N} \quad \forall i \in \{1, \ldots, N\}.$$

Note that λ_t^i is never 0 for any i as we always add the term γ/N. Also note that

$$\tilde{g}(i, Y(t)) = \begin{cases} \frac{g(i,Y_t)}{\lambda_t^i} & \text{if } I_t = i \\ 0 & \text{otherwise} \end{cases}$$

is such that $\mathbb{E}[\tilde{g}(i, Y(t)) | I_1, \ldots, I_{t-1}] = g(i, Y_t)$, which then implies that $\tilde{g}(i, Y(t))$ is an unbiased estimator of the gain. However, by using $\hat{g}(i, Y_t)$ we are biasing the estimates. Define the true gain of strategy i till time n to be $G_n^i = \sum_{t=1}^n g(i, Y_t)$ and the (biased) estimated gain of strategy i till time n to be $\hat{G}_n^i = \sum_{t=1}^n \hat{g}(i, Y_t)$. Then we have the following Azuma-Hoeffding type result.

Lemma 15. *Let $\delta \in (0,1)$. For any $\beta \in [\sqrt{\log(N/\delta)/(nN)}, 1]$ and $i \in \{1, \ldots, N\}$, we have $\mathbb{P}(G_n^i > \hat{G}_n^i + \beta nN) \le \delta/N$.*

Proof. By Markov's inequality, we have

$$\mathbb{P}(G_n^i > \hat{G}_n^i + \beta nN) \le \mathbb{E}[\exp(\beta(G_n^i - \hat{G}_n^i))]e^{-\beta^2 nN}.$$

Since $\beta \ge \sqrt{\log(N/\delta)/(nN)}$, we have $e^{-\beta^2 nN} \le \delta/N$. Thus, we only need to prove that $\mathbb{E}[\exp(\beta(G_n^i - \hat{G}_n^i))] \le 1$. We denote the random variable of interest by M_t^i, i.e.,

$$M_t^i := \exp(\beta(G_n^i - \hat{G}_n^i)) = \exp\left(\beta\left(g(i, Y_t) - \tilde{g}(i, Y(t)) - \frac{\beta}{\lambda_t^i}\right)\right) M_{t-1}^i$$

For $t = 2, \dots, n$, we get the following

$$
\begin{aligned}
&\mathbb{E}[M_t^i | I_1, \dots, I_{t-1}] \\
&= M_{t-1}^i \mathbb{E}[\exp\left(\beta\left(g(i,Y_t) - \tilde{g}(i,Y(t)) - \frac{\beta}{\lambda_t^i}\right)\right) | I_1, \dots, I_{t-1}] \\
&\le M_{t-1}^i e^{-\frac{\beta^2}{\lambda_t^i}} \mathbb{E}[1 + \beta\left(g(i,Y_t) - \tilde{g}(i,Y(t))\right) + \beta^2\left(g(i,Y_t) - \tilde{g}(i,Y(t))\right)^2 | I_1, \dots, I_t] \\
&\qquad (\text{since } \beta \le 1, g(i,Y_t) - \tilde{g}(i,Y(t)) \le 1 \text{ and } e^x \le 1 + x + x^2 \text{ for } x \le 1) \\
&= M_{t-1}^i e^{-\frac{\beta^2}{\lambda_t^i}} \mathbb{E}[1 + \beta^2\left(g(i,Y_t) - \tilde{g}(i,Y(t))\right)^2 | I_1, \dots, I_t] \\
&\le M_{t-1}^i e^{-\frac{\beta^2}{\lambda_t^i}} \left(1 + \frac{\beta^2}{\lambda_t^i}\right) \\
&\qquad (\text{since } \mathbb{E}[\left(g(i,Y_t) - \tilde{g}(i,Y(t))\right)^2 | I_1, \dots, I_t] = \mathbb{E}[\tilde{g}(i,Y_t)^2 | I_1, \dots, I_t] - g(i,Y_t)^2 \le \frac{1}{\lambda_t^i}) \\
&\le M_{t-1}^i \ (\text{since } 1 + x \le e^x).
\end{aligned}
$$

In other words, $\{M_t^i\}_{t\in\mathbb{N}}$ is a supermartingale. Now we have

$$\mathbb{E}[M_1^t] = e^{-N\beta^2}\mathbb{E}[\exp(\beta(g(i,Y_1) - \tilde{g}(i,Y_1)))];$$

since $\beta \ge 0$ and $g(i,Y_1) \in [0,1]$, we can show that $\mathbb{E}[M_1^i] \le 1$ (use the same bounds as above). Using the supermartingale property we now get $\mathbb{E}[M_n^i] \le 1$, which proves the result. □

We now have theorem that shows good performance of the MAB policy.

Theorem 61. *For any $\delta \in (0,1)$ and for any $n \ge 8N\log(N/\delta)$, if the MAB policy is run with parameters*

$$\gamma \le \frac{1}{2}, \quad 0 < \eta < \frac{\gamma}{2N}, \text{ and } \beta \in \left[\sqrt{\frac{1}{nN}\log\left(\frac{N}{\delta}\right)}, 1\right],$$

then, with probability at least $1-\delta$, the (weak) regret is at most

$$n(\gamma + \eta(1+\beta)N) + \frac{\log(N)}{\eta} + 2nN\beta.$$

In particular, choosing

$$\beta = \sqrt{\frac{1}{nN}\log\left(\frac{N}{\delta}\right)}, \quad \gamma = \frac{4N\beta}{3+\beta}, \text{ and } \eta = \frac{\gamma}{2N},$$

one has the regret being at most

$$\frac{11}{2}\sqrt{nN\log\left(\frac{N}{\delta}\right)} + \frac{\log(N)}{2}.$$

Proof. We begin by noting that

$$\begin{aligned}\log\left(\frac{W_n}{W_0}\right) &= \log\left(\sum_{i=1}^{N} e^{\eta \hat{G}_n^i}\right) - \log(N)\\ &\geq \log\left(\max_{i=1,\dots,N} e^{\eta \hat{G}_n^i}\right) - \log(N)\\ &= \eta \max_{i=1,\dots,N} \hat{G}_n^i - \log(N).\end{aligned}$$

Since $\beta \leq 1$ and $\eta \leq \gamma/(2N)$, it is easy to verify that $\eta\hat{g}(i,Y_t) \leq 1$ for all $t = 1,\dots,N$. Therefore, we have

$$\begin{aligned}\log\left(\frac{W_t}{W_{t-1}}\right) &= \log\left(\sum_{i=1}^{N} \frac{w_{t-1}^i}{W_{t-1}} e^{\eta\hat{g}(i,Y_t)}\right)\\ &= \log\left(\sum_{i=1}^{N} \frac{\lambda_t^i - \frac{\gamma}{N}}{1-\gamma} e^{\eta\hat{g}(i,Y_t)}\right)\\ &\leq \log\left(\sum_{i=1}^{N} \frac{\lambda_t^i - \frac{\gamma}{N}}{1-\gamma}\big(1 + \eta\hat{g}(i,Y_t) + \eta^2\hat{g}^2(i,Y_t)\big)\right)\\ &\leq \log\left(1 + \frac{\eta}{1-\gamma}\sum_{i=1}^{N}\lambda_t^i\hat{g}(i,Y_t) + \frac{\eta^2}{1-\gamma}\sum_{i=1}^{N}\lambda_t^i\hat{g}^2(i,Y_t)\right)\\ &\qquad \left(\text{since } \sum_{i=1}^{N}\frac{\lambda_t^i - \frac{\gamma}{N}}{1-\gamma} = 1\right)\\ &\leq \frac{\eta}{1-\gamma}\sum_{i=1}^{N}\lambda_t^i\hat{g}(i,Y_t) + \frac{\eta^2}{1-\gamma}\sum_{i=1}^{N}\lambda_t^i\hat{g}^2(i,Y_t) \text{ (since } \log(1+x) \leq x \;\; \forall x > -1).\end{aligned}$$

By the definition of $\hat{g}(i,Y_t)$ we note that

$$\begin{aligned}\sum_{i=1}^{N}\lambda_t^i\hat{g}(i,Y_t) &= g(I_t,Y_t) + N\beta, \text{ and that}\\ \sum_{i=1}^{N}\lambda_t^i\hat{g}^2(i,Y_t) &= \sum_{i=1}^{N}\lambda_t^i\hat{g}(i,Y_t)\left(1_{\{I_t=i\}}\frac{g(i,Y_t)}{\lambda_t^i} + \frac{\beta}{\lambda_t^i}\right)\\ &= \hat{g}(I_t,Y_t)g(I_t,Y_t) + \beta\sum_{i=1}^{N}\hat{g}(i,Y_t)\\ &\leq (1+\beta)\sum_{i=1}^{N}\hat{g}(i,Y_t).\end{aligned}$$

Substituting this into the upper bound above, summing over $t = 1, \dots, n$ and denoting $\bar{G}_n = \sum_{t=1}^n g(I_t, Y_t)$, we obtain

$$\log\left(\frac{W_n}{W_0}\right) \le \frac{\eta}{1-\gamma}\bar{G}_n + \frac{\eta}{1-\gamma} nN\beta + \frac{\eta^2(1+\beta)}{1-\gamma}\sum_{i=1}^N \hat{G}_n^i.$$

Comparing the upper and lower bounds on $\log(W_n/W_0)$ and rearranging, we get

$$\begin{aligned}\bar{G}_n - (1-\gamma)\max_{i=1,\dots,N}\hat{G}_n^i &\ge -\frac{1-\gamma}{\eta}\log(N) - nN\beta - \eta(1+\beta)\sum_{i=1}^N \hat{G}_n^i \\ &\ge -\frac{\log(N)}{\eta} - nN\beta - \eta(1+\beta)N\max_{i=1,\dots,N}\hat{G}_n^i,\end{aligned}$$

which is another way of writing

$$\bar{G}_n \ge -\frac{\log(N)}{\eta} - nN\beta + \big(1-\gamma-\eta(1+\beta)N\big)\max_{i=1,\dots,N}\hat{G}_n^i.$$

By Lemma 15 and the union bound, with probability at least $1-\delta$,

$$\max_{i=1,\dots,N}\hat{G}_n^i \ge \max_{i=1,\dots,N} G_n^i - \beta nN$$

whenever $\beta \in [\sqrt{\log(N/\delta)/(nN)}, 1]$. Thus, with probability at least $1-\delta$, we have

$$\bar{G}_n \ge -\frac{\log(N)}{\eta} - nN\beta\big(2-\gamma-\eta(1+\beta)N\big) + \big(1-\gamma-\eta(1+\beta)N\big)\max_{i=1,\dots,N} G_n^i.$$

By the choice of parameters $1-\gamma-\eta(1+\beta)N \ge 0$. Therefore the regret is at most

$$\begin{aligned}\max_{i=1,\dots,N} G_n^i - \bar{G}_n &\le \frac{\log(N)}{\eta} + nN\beta\big(2-\gamma-\eta(1+\beta)N\big) + \big(\gamma+\eta(1+\beta)N\big)\max_{i=1,\dots,N} G_n^i \\ &\le \frac{\log(N)}{\eta} + 2nN\beta + \big(\gamma+\eta(1+\beta)N\big)n,\end{aligned}$$

which concludes the proof. □

7.2. **Other generalizations.** Other generalizations of the multi-armed bandit problem are:

- *Multiple plays and switching costs*: In the multi-armed bandit problem, only one arm was picked at a time. In applications, it is often necessary to model multiple plays. In addition, switching arms may also have a cost associated with it. In either case, the simple index policy is no longer optimal. In this space, most of the work has been on producing good no-regret policies. Some references are:
  (1) V. Anantharam, P. Varaiya, and J. Walrand, "Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays. I. I.I.D. rewards," IEEE Trans. Automat. Control 32 (1987), no. 11, 968–976.
  (2) V. Anantharam, P. Varaiya, and J. Walrand, "Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays. II. Markovian rewards," IEEE Trans. Automat. Control 32 (1987), no. 11, 977–982.
  (3) R. Agrawal, M. Hegde and D. Teneketzis, "Multi-armed bandit problems with multiple plays and switching cost," Stochastics Stochastics Rep. 29 (1990), no. 4, 437–459.

(4) R. Agrawal, M. Hegde and D. Teneketzis, "Asymptotically efficient adaptive allocation rules for the multiarmed bandit problem with switching cost,"" IEEE Trans. Automat. Control 33 (1988), no. 10, 899–906.
(5) S. Guha and K. Munagala, "Multi-armed bandits with metric switching costs," Proceedings of 36th International Colloquium on Automata, Languages and Programming. Rhodes, Greece, 2009.

- *Restless bandits*: In the multi-armed bandit problem, the arms that are not picked remain static. If the arms that are not picked can also change, then the problem is called the restless bandits problem. This has been proved to be a hard problem to solve. It is conjectured that a variation of the Gittins' index (called Whittle's index) will lead to close to optimal performance. This is an important area of research. Some references are:
  (1) P. Whittle, "Restless bandits: Activity allocation in a changing world," A celebration of applied probability. J. Appl. Probab. 1988, Special Vol. 25A, 287–298.
  (2) J. Niño-Mora, "Dynamic priority allocation via restless bandit marginal productivity indices," TOP 15 (2007), no. 2, 161–198.
  (3) S. Guha, K. Munagala and P. Shi, "Approximation algorithms for restless bandit problems," J.ACM, 58(1), 2010.
- *Others*: These are generalizations that yield faster computational solutions, generalizes to semi-Markov processes, etc. Some references are:
  (1) P. Varaiya, J. Walrand, and C. Buyukkoc, "Extensions of the Multi-armed Bandit Problem," IEEE Trans. Autom. Control, AC-30, 426–439, 1985.
  (2) M. Katehakis and A. Veinott, "The multi-armed bandit problem: decomposition and computation," Math. Oper. Res., 12(2), 262–268, 1987.
  (3) I. Sonin, "A generalized Gittins index for a Markov chain and its recursive calculation," Statistics and Probability Letters, 78, 1526–1533, 2008.
  (4) J. Niño-Mora, "A $(2/3)^n$ Fast-Pivoting Algorithm for the Gittins Index and Optimal Stopping of a Markov Chain," INFORMS Journal of Computing, 19(4), 596–606, 2007.

8. Peer-to-Peer Networking

The sources for this section are:

- L. Massoulié and M. Vojnoviè, "Coupon replication systems," IEEE/ACM Trans. Networking, 16(3):603–616, 2008.
- B. Hajek and J. Zhu, "The missing piece syndrome in peer-to-peer communication," Proc. IEEE ISIT 2010, June 2010, pp. 1748-1752.
- J. Zhu and B. Hajek, "Stability of a peer-to-peer communication system," To appear in Proceedings of ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing, June 2011.

Fix $N \in \mathbb{N}$, then let $\mathcal{N} = \{1, \ldots, N\}$ and $2^{\mathcal{N}}$ be the power set of $\mathcal{N}$. Each $n \in \mathcal{N}$ is a piece/part of a file in a file-sharing network such as BitTorrent, eDonkey or Kazaa. These networks have a distributed architecture via an overlay network. The main goal is for participants of the network to download the file that is uploaded to the network without a centralized server to access the file from. Often some users who have downloaded the whole file stay on the network and gives out chunks to neighbours who contact them, although a majority leave as soon they've downloaded the file. Since a large portion of today's Internet traffic is generated by these applications, it is useful to analyze their performance. A key part of the protocol of these applications is choosing the piece of the file that one user requests from another. BitTorrent uses what is called the rarest piece first, i.e., user A downloads the latest piece of the file from user B that user A does not possess. The goal of this section is to analyze a mathematical model of a file-sharing protocol using techniques that we developed earlier. We will model the system in continuous time where each user is called a peer to highlight the distributed nature of the system. The papers above provide a good overview of this application area.

Consider following random process. A type C peer is one that has pieces corresponding to $C \subseteq \mathcal{N}$; a type $\mathcal{N}$ peer is called a peer seed. If a type C peer receives a piece $i \notin C$, then its type changes to $C \cup \{i\}$; note that types never reduce. Arrivals from the outside of type C form a rate λ_C Poisson process such that $\lambda_{total} := \sum_{C \in 2^{\mathcal{N}}} \lambda_C > 0$; typically one will assume that $\lambda_\emptyset$ or λ_i (for $\lambda_{\{i\}}$, i.e., users with just one piece) for all $i \in \mathcal{N}$ are the only non-zero arrival rates. The idea is that users do a search when they first think of download a file, pick up some pieces and then join the file-sharing experience with users possessing the whole file doing so only to help others get the file. We assume that there exists exactly one fixed seed in the network that has type $\mathcal{N}$ that never ever leaves the network. The fixed seed contacts a uniformly chosen peer at instances of a U_s rate Poisson process and gives it a piece that it doesn't already possess, choosing uniformly amongst such pieces; none of the peers can contact the fixed seed. Each peer in the system (not the fixed seed) contacts a randomly chosen peer (again not the fixed seed) at instances of a rate μ Poisson process and uploads a piece that the contacted peer does not possess, again choosing uniformly amongst such pieces. Note that we are not modeling the rarest-piece-first policy in order to keep the state-space simple. If a peer becomes a peer seed and if $\gamma \in (0, \infty)$, then it can leave the network once an exponential clock with parameter γ ticks, where the clock starts as soon as the peer seed status is achieved. If $\gamma = \infty$, then peer seeds leave instantly and in this case we also assume that $\lambda_{\mathcal{N}} = 0$.

Now it is easy to see from the description that we have a CTMC with state being $\mathbf{n} = (n_C : C \subseteq \mathcal{N})$, where n_C denotes the number of type C peers. Denote by $n_{total} := \sum_{C \subseteq \mathcal{N}} n_C$

to be the total of number peers in the system when the state is $\mathbf{n}$ where we never count the fixed seed. We let $\mathbf{e}_C$ denote a vector with dimensions the same as $\mathbf{n}$, with a one in position C and zero in every other coordinate. The positive entries of the rate/generator matrix $Q = (q(\mathbf{n}, \mathbf{n}') : \mathbf{n}, \mathbf{n} \in \mathbb{Z}_+^{2^N})$ are given by

$$\begin{aligned} q(\mathbf{n}, \mathbf{n} + \mathbf{e}_C) &= \lambda_C \\ q(\mathbf{n}, \mathbf{n} - \mathbf{e}_{\mathcal{N}}) &= \gamma n_{\mathcal{N}} \\ q(\mathbf{n}, \mathbf{n} - \mathbf{e}_C + \mathbf{e}_{C\cup\{i\}}) &= \frac{n_C}{n_{total}} \left(\frac{U_s}{N - |C|} + \mu \sum_{B \subset \mathcal{N}: i \in B} \frac{n_B}{|B \setminus C|} \right) 1_{\{n_C \geq 1, n_{total} \geq 1\}} \text{ if } C \subsetneq \mathcal{N} \text{ and } i \notin C. \end{aligned}$$

This resulting Markov process is $\{\mathbf{N}(t)\}_{t\in\mathbb{R}_+}$.

First we will analyze this system taking a large system limit and applying Kurtz's theorem (Theorem 53) look at the result ordinary differential equation. Let K be the parameter by which we will scale the systems. In the K^{th} system, let the arrival rate for type C peers be $\lambda_C^K = K\lambda_c$ and the sampling rate of the fixed seed be $U_s^K = KU_s$. We will also assume that the initial state in the K^{th} system is such that $K^{-1}\mathbf{N}^K(0) = x(0)$. Now we have the following

$$\begin{aligned} q^K(\mathbf{n}, \mathbf{n} + \mathbf{e}_C) &= K\lambda_C \\ q^K(\mathbf{n}, \mathbf{n} - \mathbf{e}_{\mathcal{N}}) &= \gamma K \frac{n_{\mathcal{N}}}{K} \\ q^K(\mathbf{n}, \mathbf{n} - \mathbf{e}_C + \mathbf{e}_{C\cup\{i\}}) &= K \frac{\frac{n_C}{K}}{\frac{n_{total}}{K}} \left(\frac{U_s}{N - |C|} + \mu \sum_{B \subset \mathcal{N}: i \in B} \frac{\frac{n_B}{K}}{|B \setminus C|} \right) 1_{\{\frac{n_C}{K} \geq \frac{1}{K}, \frac{n_{total}}{K} \geq \frac{1}{K}\}} 1_{\{C \subsetneq \mathcal{N}, i \notin C\}}, \end{aligned}$$

Let us denote the Markov process of scale K as $\{\mathbf{N}^K(t)\}_{t\in\mathbb{R}_+}$. Note that we are in the density-dependent setting where it is possible that we can apply Theorem 53. First we write down the form for $F(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}_+$ with $x_{total} := \sum_{C \subseteq \mathcal{N}} x_C$, as follows:

$$\begin{aligned} &F(\mathbf{x}) \\ &= \sum_{C \subsetneq \mathcal{N}} \mathbf{e}_C \Bigg[\lambda_C - U_s \frac{x_C}{x_{total}} 1_{\{x_{total}>0\}} - \mu x_C \frac{\sum_{B \subseteq \mathcal{N}: |B\setminus C|>0} x_B}{x_{total}} 1_{\{x_{total}>0\}} \\ &\qquad + \frac{U_s}{N - |C| + 1} \frac{\sum_{i \in C} x_{C\setminus\{i\}}}{x_{total}} 1_{\{x_{total}>0\}} + \mu \sum_{i \in C} x_{C\setminus\{i\}} \frac{\sum_{B \subseteq \mathcal{N}: i \in B} \frac{x_B}{|B\setminus C|+1}}{x_{total}} 1_{\{x_{total}>0\}} \Bigg] \\ &\quad + \mathbf{e}_{\mathcal{N}} \left[\lambda_{\mathcal{N}} - \gamma x_{\mathcal{N}} + (U_s + \mu x_{\mathcal{N}}) \frac{\sum_{i \in \mathcal{N}} x_{\mathcal{N}\setminus\{i\}}}{x_{total}} 1_{\{x_{total}>0\}} \right] \end{aligned}$$

If $U_s = 0$, then we can show that $|\frac{\partial F_C(\mathbf{x})}{\partial x_B}| < \infty$ uniformly for all $\mathbf{x}$ (even for $\mathbf{x} = \mathbf{0}$) and for all $B, C \subseteq \mathcal{N}$, and this implies that $F(\cdot)$ is Lipschitz. Then noting that we satisfy the conditions of Theorem 53, we get for every $T \in \mathbb{R}_+$ that

$$\lim_{K\to\infty} \sup_{0 \leq t \leq T} \left\| \frac{1}{K} \mathbf{N}^K(t) - \mathbf{x}(t) \right\|_1 = 0 \quad a.s.$$

where component $x_C(\cdot)$ of $\{\mathbf{x}(t)\}_{t\in[0,T]}$ where $C \subsetneq \mathcal{N}$ satisfies

$$\frac{d}{dt}x_C(t) = \lambda_C - \mu x_C \frac{\sum_{B\subseteq\mathcal{N}:|B\setminus C|>0} x_B}{x_{total}} 1_{\{x_{total}>0\}} + \mu \sum_{i\in C} x_{C\setminus\{i\}} \frac{\sum_{B\subseteq\mathcal{N}:i\in B} \frac{x_B}{|B\setminus C|+1}}{x_{total}} 1_{\{x_{total}>0\}},$$

and when $C = \mathcal{N}$ the corresponding equation is

$$\frac{d}{dt}x_{\mathcal{N}}(t) = \lambda_{\mathcal{N}} - \gamma x_{\mathcal{N}} + \mu x_{\mathcal{N}} \frac{\sum_{i\in\mathcal{N}} x_{\mathcal{N}\setminus\{i\}}}{x_{total}} 1_{\{x_{total}>0\}}.$$

If $\gamma = \infty$, then we assume $\lambda_{\mathcal{N}} = 0$ and $x_{\mathcal{N}} \equiv 0$; this is the case for the flat system from the Massoulié-Vojnovič paper. Solution of these equations is non-trivial but can reveal what the system performance will be like. One can also make comments on positive recurrence and transience if the resulting ODE has a globally asymptotically stable point; trying a few examples it is easy to see that here this is not necessarily the case.

We will analyze the CTMC using more regular tools (Foster-Lyapunov) to prove positive recurrence and transience. The main result is the following.

Theorem 62. *Given $(\lambda_C : C \in \mathcal{N})$, U_s, μ, γ and the rate matrix Q, the following hold:*

(1) *the Markov process with generator matrix Q is transient for $0 < \gamma \le \mu$ when, for some piece $k \in \mathcal{N}$, we have $U_s + \sum_{C\subseteq\mathcal{N}:k\in C} \lambda_C = 0$, and also when $0 < \mu < \gamma \le \infty$, if for some $k \in \mathcal{N}$ we have*

$$\lambda_{total} > \left[U_s + \sum_{C\subseteq\mathcal{N}:k\in C} \lambda_C(N+1-|C|)\right] \frac{\gamma}{\gamma-\mu},$$

where $|C|$ is the number of elements in set C;

(2) *the Markov process with generator matrix Q is positive recurrent for $0 < \gamma \le \mu$ when for all $k \in \mathcal{N}$, we have $U_s + \sum_{C\subseteq\mathcal{N}:k\in C} \lambda_C > 0$, and also when $0 < \mu < \gamma \le \infty$, if for all $k \in \mathcal{N}$, we have*

$$\lambda_{total} < \left[U_s + \sum_{C\subseteq\mathcal{N}:k\in C} \lambda_C(N+1-|C|)\right] \frac{\gamma}{\gamma-\mu}.$$

In both cases, for $\gamma = \infty$ the inequalities are interpreted by taking limits as γ increases to infinity.

We will only outline/sketch the proofs of this theorem.

Theorem 62, Part (1). We consider the case of $0 < \mu < \gamma < \infty$ and without loss of generality assume that the inequality is true for $k = 1$, i.e.,

$$\begin{aligned}\Delta &= \sum_{C\subseteq\mathcal{N}} \lambda_C - \left[U_s + \sum_{C\subseteq\mathcal{N}:1\in C} \lambda_C(N+1-|C|)\right] \frac{\gamma}{\gamma-\mu} \\ &= \sum_{C:1\notin C} \lambda_C - U_s \frac{\gamma}{\gamma-\mu} - \sum_{C\subseteq\mathcal{N}:1\in C} \lambda_C \frac{(N-|C|)\gamma+\mu}{\gamma-\mu} > 0\end{aligned}$$

Now partition the peers into five groups as follows:

(1) *Normal young peer*: this is a peer that does not have piece one and does not have at least one other piece;

(2) *Infected peer*: this is a peer that obtained piece one after arriving, but before obtaining all the other pieces. Once a peer is infected, it remains infected until it leaves the system; it is considered to be infected even when it is a peer seed;
(3) *Gifted peer*: this is a peer that arrives with piece one. A gifted peer is gifted for its entire time in the system, and even when it is a peer seed;
(4) *One-club peer*: this is a peer that has all pieces except piece one, i.e., of type $\{2, \ldots, N\}$; and
(5) *Former one-club peer*: this is a peer in the system that was at some earlier time a one-club peer. Note that a former one-club peer is a peer seed but the converse is not true, infected peers and gifted peers can be peer seeds too.

Consider an initial state where there are many peers in the system, and all of them are one-club peers. Piece one can arrive into the system from the outside in two ways: uploads by the fixed seed or arrivals of gifted peers. Initially we ignore the effect of normal young peers getting piece one (and becoming infected). Most of the uploads by the fixed seed are uploads of piece one to one-club peers. Each such upload creates a new peer seed, which on the average will upload piece one to μ/γ more one-club peers (geometric number of such uploads occur), and each of the resulting peers will upload piece one to another μ/γ one-club peers, and so forth, as in a branching process. Thus, each upload of a piece by the fixed seed finally results in about $\gamma/(\gamma-\mu)$ departures from the one-club. Each gifted peer, with type C on arrival, where $1 \in C$, will directly upload to, on average, $K - |C| + \mu/\gamma$ one-club peers, and these will become peers seeds that then initiate the branching process type reduction in one-club peers, so that the total expected number of one-club departures caused by the type C gifted peer is $(K - |C| + \mu/\gamma)\gamma/(\gamma - \mu)$. The set of one-club peers increases with the arrival of peers without piece one. Taking the difference of the arrival rate and departure rate of one-club peers, one gets Δ which is strictly positive so that over time the one-club peer set increases without bound.

The discussion above neglected the possibility that normal young peers can also receive piece one, and thereby, creating infected peers. An infected peer can upload to one-club peers, creating former one-club peers, and to normal young peers, creating more infected peers. This results in another branching process evolution of infected peers and former one-club peers. However, the expected number of infected offspring of a former one-club peer or an infected peer converges to zero, as the fraction of one-club peers converges to one. Hence, when the one-club peer group is large enough, the existence of infected peers is negligible; it will not affect the growth of the one-club peer group. □

Note from the logic described above that when the departure rate of one-club peers is large enough, then the system can recover from high loads of one-club peers.

Theorem 62, Part (2). Again we will only consider the case of $0 < \mu < \gamma < \infty$. This relies on using a Lyapunov function and the Foster-Lyapunov criterion. Let r, d, β and α be positive constants (left unspecified), with r and β small, d large, and α less than but close to one. Let $\mathcal{E}_C = \{B : B \in \mathcal{N}, B \subseteq C\}$ is the collection of all subsets of C, i.e., collection of types of peers which may become type C peers in the future. Let $\mathcal{H}_C = \{B : B \in \mathcal{N}, |B \setminus C| > 0\}$

be the set of types of peers who can help type C peers. Using these define

$$E_C = \sum_{B \in \mathcal{E}_C} n_B$$
$$H_C = \frac{\gamma}{\gamma - \mu} \sum_{B \in \mathcal{H}_C} (K - |B| + \mu/\gamma) n_B.$$

Let $\phi : \mathbb{R}_+ \mapsto \mathbb{R}_+$ a function with parameters d and $beta$ be given by

$$\phi(x) = \begin{cases} 2d + \frac{1}{\beta} - x & \text{if } x \in [0, 2d] \\ \frac{\beta}{2}\left(x - 2d - \frac{1}{\beta}\right)^2 & \text{if } x \in \left(2d, 2d + \frac{1}{\beta}\right] \\ 0 & \text{if } x > 2d + \frac{1}{\beta} \end{cases}$$

Then the candidate Lyapunov function is given by

$$V(\mathbf{n}) = \sum_{C \subseteq \mathcal{N}} r^{|C|} T_C, \text{ with } T_C = \begin{cases} \frac{1}{2}E_C^2 + \alpha E_C \phi(H_C) & \text{if } C \neq \mathcal{N} \\ \frac{1}{2} n_{\mathcal{N}}^2 & \text{if } C = \mathcal{N} \end{cases}$$

It can be shown that this has negative drift (QV) as required by the Foster-Lyapunov criterion for CTMCs. When H_C is small, then the drift of $\frac{1}{2}E_C^2$ is positive and the drift of the term $\alpha E_C \phi(H_C)$ is sufficiently negative to compensate. The main idea is to consider sets like the one-club peer and show negative drift of those. □

8.1. **Exercises.** These are exercises for the last three sections. Please show all your work.

(1) *Total cost criterion*: Prove that $V_n(i) \leq V_{n+1}(i)$.

(2) *Total cost criterion*: In Theorem 55, prove statement that $V_n(i) \leq V^*(i)$ implies that $V_{n+1}(i) \leq V^*(i)$.

(3) *Multi-armed bandits*: The goal is to simulate four no-regret policies for 10,000 steps and see which is best. Take $N = 10$ and assume that the rewards are *i.i.d.* random variables $\{\theta_t^i\}_{t\in\mathbb{N}}$ for arm i with θ_t^i for arm i chosen in $[x, x + dx]$ with probability proportional to $x^{i-1}dx$ for $x \in [0, 1]$ with the rewards being independent across arms. Let $\mu_i = \mathbb{E}[\theta_t^i]$ be the mean reward for machine i, $\mu^* = \max_{i=1,\ldots,N} \mu_i$ and $\Delta_i = \mu^* - \mu_i$. The algorithms to consider are:

(a) Algorithm 1 is UCB that we analyzed in class. It is the following:
- *Initialization*: Play each arm once;
- *Iterative step*: Play arm j that maximizes
$$\bar{x}_n^j + \sqrt{\frac{2\log(n)}{n^j}}$$
where $\bar{x}_n^j$ is the average reward obtained from machine j, n^j the number of times machine j has been played so far, and n is the total number of plays so far. After every play, update the variables and repeat.

(b) The second algorithm is a refinement of UCB called UCB2. Pick $\alpha \in (0, 1)$ (parameter) and set $\tau(r) := \lceil (1 + \alpha)^r \rceil$ for $r \in \mathbb{Z}_+$ and
$$a_{n,r} := \sqrt{\frac{(1+\alpha)\log\left(\frac{en}{\tau(r)}\right)}{2\tau(r)}}.$$
Then do the following:
- *Initialization phase*: Set $r_j = 0$ for all arms and play each machine once;
- *Iteration phase*: Select arm j that maximizes $\bar{x}_n^j + a_{n,r_j}$, where $\bar{x}_n^j$ is the average reward obtained from arm j, a_{n,r_j} the quantity described above and n the total number of plays thus far. Play chosen arm j exactly $\tau(r_j + 1) - \tau(r_j)$ times. Increment r_j by 1 and repeat.

(c) The third algorithm is called ϵ_n-GREEDY which chooses the maximum of $\bar{x}_n^j$ with probability $(1-\epsilon_n)$ and a random arm otherwise. The parameters are $c > 0$ and $d \in (0, 1)$ with $d \leq \min_{i:\mu_i<\mu^*} \Delta_i$. Then choose $\epsilon_n = \min(1, \frac{cN}{d^2 n})$.

(d) The fourth algorithm is MAB that we analyzed in class. The reward at time t for arm i is also denoted by $g(i, \theta_t) = \theta_t^i$ where θ_t is the vector process of rewards across all arms. The parameters of the algorithm are $\delta \in (0, 1)$, $\beta = \sqrt{\frac{1}{nN}\log\left(\frac{N}{\delta}\right)}$, $\gamma = \frac{4N\beta}{3+\beta}$ and $\eta = \frac{\gamma}{2N}$ where n is the duration that the algorithm is run for. The description of the algorithm is as follows:
- *Initialization phase*: Set $w_0^i \equiv 1$ and $\lambda_1^i \equiv 1/N$;
- *Iteration phase*: For each $t \in \mathbb{N}$ repeat the following:
  (i) Pick arm $I_t \in \{1, \ldots, N\}$ according to distribution λ_t;
  (ii) Calculate the estimated gains as follows
$$\hat{g}(i, \theta_t) = \tilde{g}(i, \theta_t) + \frac{\beta}{\lambda_t^i} = \begin{cases} \frac{g(i,\theta_t)+\beta}{\lambda_t^i} & \text{if } I_t = i \\ \frac{\beta}{\lambda_t^i} & \text{otherwise;} \end{cases}$$

(iii) Update the weights $w_t^i = w_{t-1}^i e^{\eta \hat{g}(i,\theta_t)}$ and $W_t = \sum_{i=1}^n w_t^i$;
(iv) Calculate the updated probability distribution

$$\lambda_{t+1}^i = (1-\gamma)\frac{w_t^i}{W_t} + \frac{\gamma}{N} \quad \forall i \in \{1,\dots,N\}.$$

What is the best arm? Simulate the optimum policy too. Use the same sample-paths for all algorithms and in each case have enough runs to get good estimates of the average regret of each policy. Take note of statistics of how many times each arm is played as a function of time over the 10,000 steps.

(4) *Peer-to-peer network*: Consider following CTMC. Fix $N \in \mathbb{N}$ (say 4), then let $\mathcal{N} = \{1,\dots,N\}$ and $2^{\mathcal{N}}$ be the power set of $\mathcal{N}$. A type C peer is one that has pieces corresponding to $C \subseteq \mathcal{N}$; a type $\mathcal{N}$ peer is called a peer seed. If a type C peer receives a piece $i \notin C$, then its type changes to $C \cup \{i\}$; types never reduce. Arrivals from the outside of type C form a rate λ_C Poisson process such that $\lambda_{total} := \sum_{C \in 2^{\mathcal{N}}} \lambda_C > 0$. There exists exactly one fixed seed in the network that has type $\mathcal{N}$ that never ever leaves the network. The fixed seed contacts a uniformly chosen peer at instances of a U_s rate Poisson process and gives it a piece that it doesn't already possess, choosing uniformly amongst such pieces. Each peer in the system (not the fixed seed) contacts a randomly chosen peer (again not the fixed seed) at instances of a rate μ Poisson process and uploads a piece that the contacted peer does not possess, again choosing uniformly amongst such pieces. If a peer becomes a peer seed and if $\gamma \in (0,\infty)$, then it can leave the network once an exponential clock with parameter γ ticks, where the clock starts as soon as the peer seed status is achieved. If $\gamma = \infty$, then peer seeds leave instantly and we also assume that $\lambda_{\mathcal{N}} = 0$. Simulate this CTMC under the following different settings (choose your own parameters with U_s, λ_C for all $C \subseteq \mathcal{N}$, $\mu < \infty$, mention them and show how they fit the categories below):
(a) The case where $0 < \gamma \le \mu$ and where, for some piece $k \in \mathcal{N}$, we have $U_s + \sum_{C \subseteq \mathcal{N}: k \in C} \lambda_C = 0$;
(b) The case where $0 < \mu < \gamma \le \infty$ and for some $k \in \mathcal{N}$ we have

$$\lambda_{total} > \left[U_s + \sum_{C \subseteq \mathcal{N}: k \in C} \lambda_C (N + 1 - |C|) \right] \frac{\gamma}{\gamma - \mu},$$

where $|C|$ is the number of elements in set C;
(c) The case where $0 < \gamma \le \mu$ and where for all $k \in \mathcal{N}$, we have $U_s + \sum_{C \subseteq \mathcal{N}: k \in C} \lambda_C > 0$; and
(d) The case where $0 < \mu < \gamma \le \infty$ and where for all $k \in \mathcal{N}$, we have

$$\lambda_{total} < \left[U_s + \sum_{C \subseteq \mathcal{N}: k \in C} \lambda_C (N + 1 - |C|) \right] \frac{\gamma}{\gamma - \mu}.$$

Comment on the different scenarios - make guesses regarding positive recurrence or transience from simulations.