Applied Stochastic Processes Notes

Prof. Vincent Tassion
Transcription: Trevor Winstral

Spring Semester 2021

Contents

Chapter 0

Introduction

Mathematical Definition of Stochastic Processes We want to describe a process evolving in time. The most relevant for us will be: Discrete time ($I = \mathbb{N}$) and Continuous time ($I = \mathbb{R}$).

Definition 0.1. Let (E, ξ) be a measurable space. A discrete stochastic process with state space E is a collection $X = (X_n)_{n\in\mathbb{N}}$ of RVs with values in E.

Definition 0.2. A continuous stochastic process is a collection $(X_t)_{t\in\mathbb{R}_+}$ of RVs with values in E.

In this class we will work with jump processes, ie when E is finite or countable. We will work with:

(i) Discrete time Markov Chains $I = \mathbb{N}$ and E finite or countable

(ii) Poisson renewal processes $I = \mathbb{R}_+$ and $E = \mathbb{N}$

(iii) Continuous Markov Chains $I = \mathbb{R}_+$ and E finite or countable

We will not work with Brownian Motion.

Example 0.1 (Simple Random Walk). State Space $\mathbb{Z}^d$, x, y are neighbors $\iff \|x - y\|_1 = 1$. An electron is starting at 0, and each step it jumps uniformly to one of the neighbors. How should we define this?

Definition 0.3 (SRW)**.** Let $(Z_n)_{n\in\mathbb{N}}$ iid, $\mathbb{P}[Z_n = \pm e_i] = \frac{1}{2d}$ where e_i is 1 in the i'th slot. $X_n := \sum_{k=1}^{n} Z_n = X_n + Z_{n+1}, X_0 = 1$. $\forall m, n X_m$ and X_n are dependent. The X_n do satisfy the Markov property: Conditional on $X_n = x$ then $(X_{m+n})_{n\geq 0}$ is a SRW starting at x independent of $(X_1, ..., X_m)$.

Will the SRW return to 0?

Theorem 0.1 (Polya)**.**
If $d = 1, 2$ then $\mathbb{P}\left[(X_n)\text{ visits } x \text{ infinitely many times}\right] = 1$
If $d \geq 3$ then $\mathbb{P}\left[(X_n)\text{ visits } x \text{ only finitely many times}\right] = 1$

Example 0.2 (Poisson Process)*.* We want to define and study N_t the number of cars passing a point during $[0, t]$.

Definition 0.4. T_1 = passage of time of the first car, T_2 = time between car 1 and car 2, etc.

- (T_i) are iid
- (T_i) are memoryless: $\mathbb{P}\left[T_1 \geq t + s | T_1 \geq s\right] = \mathbb{P}\left[T_1 \geq t\right]$
- Regularity: $\mathbb{P}\left[T_1 \geq s\right]$ is 'nice'

This implies that $\mathbb{P}\left[T_1 \geq s\right] = e^{-\lambda s}, \quad \lambda > 0$

Let $(T_i)_{i\geq 1}$ iid $exp(\lambda)$ RV. $N_t = \sum_{i\geq 1} \chi_{T_1+\ldots+T_i\leq t}$
Dependencies:

- $N_{t+s} - N_t \sim Pois(\lambda s)$
- Markov Property

LLN: $\frac{N_t}{t} \rightarrow_{t\to\infty} \frac{1}{\lambda}$

Chapter 1

Markov Chains and Generalities

Framework: $(\Omega, \mathcal{F}, \mathbb{P})$ Probability Space, E finite or countable set with the σ-algebra 2^E
Outset We would like to define a class of processes such that the evolution of the process is memoryless, but still location dependent. This means that the way a process continues past this point in time, does not depend on how it got to where it is now, but only on where it is at this point in time.

Definition 1.1. Let $X = (X_n)_{n\in\mathbb{N}}$ be a sequence of random variables in E. We say that X is a time homogeneous Markov Chain (MC) if:

(i) For all $n \geq 0$ and $x_1, \ldots, x_{n+1} \in E$

$$\boxed{\mathbb{P}\left[X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n\right] = \mathbb{P}\left[X_{n+1} = x_{n+1} \mid X_n = x_n\right].}$$

(ii) For all $m, n \geq 0$ and $x, y \in E$

$$\boxed{\mathbb{P}\left[X_{n+1} = y \mid X_n = x\right] = \mathbb{P}\left[X_{n+1=y} \mid X_n = x\right].}$$

Note: By convention when we write $\mathbb{P}[A \mid B]$ we assume $\mathbb{P}[B] > 0$.

Remark 1.1. The first condition is equivalent to

$$\forall f : E \to \mathbb{R} \text{ bounded, } \mathbb{E}\left[f(X_{n+1}) \mid X_0, \ldots, X_n\right] = \mathbb{E}\left[f(X_{n+1}) \mid X_n\right].$$

Remark 1.2. The first condition is equivalent to for all $f : E \to \mathbb{R}$ bounded,

$$\boxed{\mathbb{E}\left[f(X_{n+1}) \mid X_0, \ldots, X_n\right] = \mathbb{E}\left[f(X_{n+1}) \mid X_n\right].}$$

Example 1.1. If X_n are i.i.d. in E then (X_n) is a Markov Chain.

Example 1.2. SRW on $\mathbb{Z}^d$. // You mentioned that this should be a definition, I think we define this in the introduction (which we haven't done yet in Latex, but we did it in the course in this way //

1.1 Transition Probabilities

Definition 1.2. // This is the only instance where I replaced p with P to denote the collection in this chapter, to see it in usage, look at the next chapter where I have replaced all instances //

A *transition probability* is a collection $P = (p_{x,y})_{x,y \in E}$ such that:

- For any $x, y \in E$: $p_{x,y} \in [0, 1]$, and
- $\sum_{y \in E} p_{x,y} = 1$.

There are a few different representations of transition probabilities.

Graph For E finite or countable, we could set the vertices of a weighted oriented graph to the elements of E, and the edges to $(x, y) \in E^2$ with the weights p_{xy}. Note here that the sum of the weights of the edges leaving a vertex is equal to 1.

Matrix Say $E = \{1, \ldots, N\}$ and $p = (p_{ij})_{1 \leq i,j \leq N}$ with $P_{ij} \geq 0$ and $\sum_j p_{ij} = 1$. We call this a stochastic matrix.

Operator If E is finite or infinite then for all $f \in L^\infty(E)$ define the functor $Pf \in L^\infty(E)$ by $Pf(x) = \sum_{y \in E} P_{x,y} f(y)$ with $P \geq 0$ (for all $f \geq 0 : Pf \geq 0$) and satisfies $P1 = 1$.

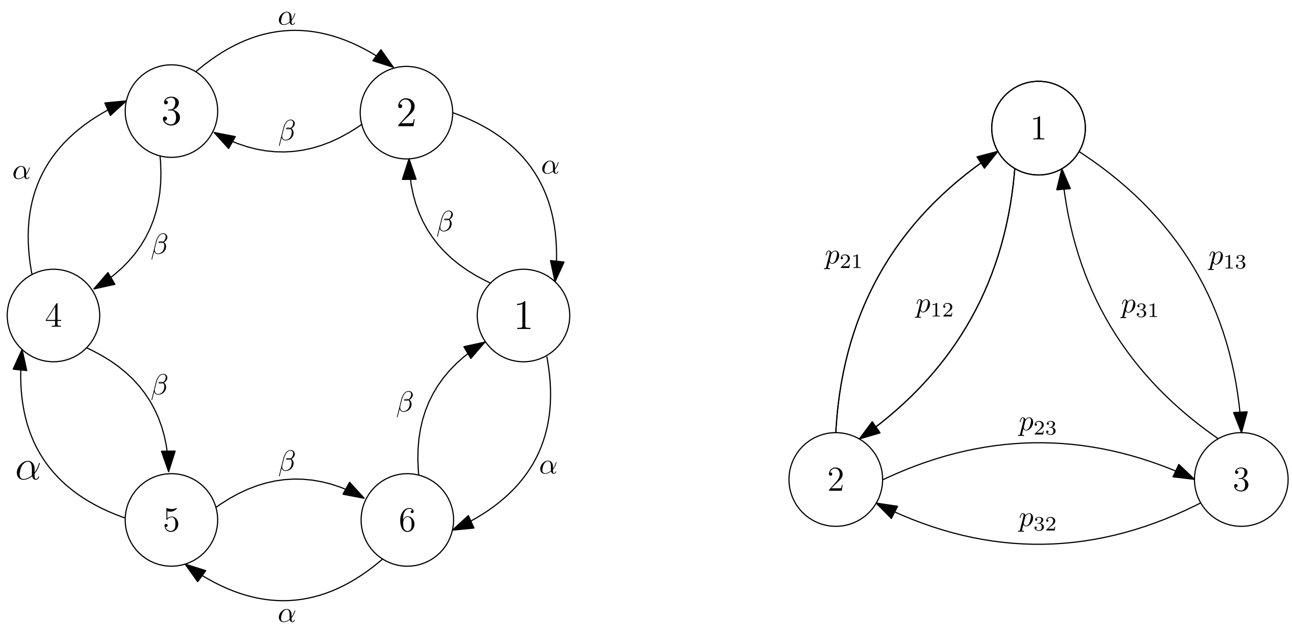

Figure 1.1: 6-state (asymmetric) Markov Chain and 3-state Markov Chain

Definition 1.3. Let p be a transition probability, μ a distribution on E, a sequence $(X_n)_{n\geq 0}$ of random variables with values in E is a Markov Chain with initial distribution μ and transition probability p (written $\mathrm{MC}(\mu, p)$) if for every $x_0, \ldots, x_n \in E$

$$\boxed{\mathbb{P}\left[X_0 = x_0, \ldots, X_n = x_n\right] = \mu(x_0) p_{x_0,x} \cdots p_{x_{n-1},x_n}.}$$

Proposition 1.3. *Let $X = (X_n)_{n\geq 0}$ sequence of random variables with values in E. We have*

$$(X \textit{ is a Markov Chain }) \iff (\exists \mu, p \textit{ such that } X \textit{ is a } MC(\mu, p)).$$

Proof. $\implies$: Let $\mu =$ be the law of X_0 and set

$$p_{xy} = \begin{cases} \mathbb{P}\left[X_{n+1} = y \mid X_n - x\right] & \text{if } \exists n : \mathbb{P}\left[X_n = x\right] > 0 \\ \mathbb{1}_{x=y} & \text{else.} \end{cases}$$

By homogeneity, p_{xy} is well-defined. Furthermore, for every $x_0, \ldots, x_n \in E$ we have

$$\begin{aligned} &\mathbb{P}\left[X_0 = x_0, \ldots, X_n = x_n\right] \\ &\quad = \mathbb{P}\left[X_n = x_n \mid X_0 = x_0, \ldots, X_{n-1} = x_{n-1}\right] \mathbb{P}\left[X_0 = x_0, \ldots, X_{n-1} = x_{n-1}\right] \\ &\quad = \mathbb{P}\left[X_0 = x_0\right] \prod_{i=1}^{n} \mathbb{P}\left[X_i = x_i \mid X_0 = x_0, \ldots, X_{i-1} = x_{i-1}\right] \\ &\quad = \mu(x_0) \prod_{i=1}^{n} \mathbb{P}\left[X_i = x_i \mid X_{i-1} = x_{i-1}\right] = \mu(x_0) \prod_{i=1}^{n} p_{x_{i-1}x_i}. \end{aligned}$$

It remains to check that P is a transition probability. Let $x \in E$. If there exists $n \geq 0$ such that $\mathbb{P}\left[X_n = x\right] > 0$, then

$$\sum_{y\in E} p_{xy} = \sum_{y\in E} \mathbb{P}\left[X_{n+1} = y \mid X_n = x\right] = 1.$$

Otherwise,

$$\sum_{y\in E} p_{xy} = \sum_{y\in E} \mathbb{1}_{x=y} = 1.$$

$\impliedby$: Assume $(X_n)_{n\geq 0}$ is a $\mathrm{MC}(\mu, p)$. Let $n \geq 0$ and $x_0, \ldots, x_{n+1}$ such that $\mu(x_0) p_{x_0x_1} \cdots p_{x_{n-1}x_n} > 0$. We have that

$$\mathbb{P}\left[X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n\right] = \frac{\mathbb{P}\left[X_0 = x_0, \ldots, X_{n+1} = x_{n+1}\right]}{\mathbb{P}\left[X_0 = x_0, \ldots, X_n = x_n\right]} = p_{x_nx_{n+1}}.$$

Now let $n \geq 0$ and $y \in E$ such that $\mathbb{P}[X_n = x] > 0$.

$$\begin{aligned}
&\mathbb{P}[X_{n+1} = y \mid X_n = x] \\
&\quad = \sum_{(u_0,\dots,u_{n-1})\in E^n} \mathbb{P}[X_{n+1} = y \mid X_0 = u_0, \dots, X_{n-1} = u_{n-1}, X_n = x] \cdot \\
&\qquad\qquad\qquad \mathbb{P}[X_0 = u_0, \dots, X_{n-1} = u_{n-1} \mid X_n = x] \\
&\quad = p_{xy} \sum_{(u_0,\dots,u_{n-1})\in E^n} \mathbb{P}[X_0 = u_0, \dots, X_{n-1} = u_{n-1} \mid X_n = x] = p_{xy}.
\end{aligned}$$

This concludes that X fulfills the two properties of a Markov Chain.

$\Longrightarrow$: If X_n is a Markov Chain, then set

$$p_{xy} = \begin{cases} \mathbb{P}[X_{n+1} = y \mid X_n - x] & \text{if } \exists n : \mathbb{P}[X_n = x] > 0 \\ \mathbb{1}_{x=y} & \text{else} \end{cases}.$$

We have that $\sum_{y\in E} p_{xy} = 1$, as the conditional probability is a probability measure itself, and $p_{xy} \geq 0$, for every $x, y \in E$ for the same reason. Thus we have that the collection of $(p_{xy})_{x,y\in E}$ forms a transition probability. Setting $\mu(x) = \mathbb{P}[X_0 = x]$, which is also clearly a probability measure on E. Now we only have to show that X_n is a MC(μ, p). For every $x_0, \dots, x_n \in E$, and every $n \geq 0$, we have

$$\begin{aligned}
&\mathbb{P}[X_0 = x_0, \dots, X_n = x_n] \\
&\quad = \mathbb{P}[X_n = x_n \mid X_0 = x_0, \dots, X_{n-1} = x_{n-1}]\, \mathbb{P}[X_0 = x_0, \dots, X_{n-1} = x_{n-1}] \\
&\quad = \mathbb{P}[X_0 = x_0] \prod_{i=1}^{n} \mathbb{P}[X_i = x_i \mid X_0 = x_0, \dots, X_{i-1} = x_{i-1}] \\
&\quad = \mu(x_0) \prod_{i=1}^{n} \mathbb{P}[X_i = x_i \mid X_{i-1} = x_{i-1}] = \mu(x_0) \prod_{i=1}^{n} p_{x_{i-1}x_i}.
\end{aligned}$$

Thus we have proven this implication by using the Markov property of Markov Chains.

$\Longleftarrow$: Here we have to demonstrate the two properties of a Markov Chain, the Markov property and homogeneity. For homogeneity we have

$$\begin{aligned}
&\mathbb{P}[X_{n+1} = y \mid X_n = x] \\
&\quad = \sum_{(u_0,\dots,u_{n-1})\in E^n} \mathbb{P}[X_{n+1} = y \mid X_0 = u_0, \dots, X_{n-1} = u_{n-1}, X_n = x] \cdot \\
&\qquad\qquad\qquad \mathbb{P}[X_0 = u_0, \dots, X_{n-1} = u_{n-1} \mid X_n = x] \\
&\quad = p_{xy} \sum_{(u_0,\dots,u_{n-1})\in E^n} \mathbb{P}[X_0 = u_0, \dots, X_{n-1} = u_{n-1} \mid X_n = x] = p_{xy}.
\end{aligned}$$

here we have implicitly assumed that $\mathbb{P}[X_n = x] > 0$, as without this the conditional probability we are taking is not well-defined. In the case where this is not true, the fulfillment of the homogeneity property is trivial as the transition probability is constantly 1.

For the Markov Property we have

$$\begin{aligned}\mathbb{P}[X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n] &= \frac{\mathbb{P}[X_0 = x_0, \dots, X_{n+1} = x_{n+1}]}{\mathbb{P}[X_0 = x_0, \dots, X_n = x_n]} \\ &= \frac{\mu(x_0) p_{x_0 x_1} \cdots p_{x_n x_{n+1}}}{\mu(x_0) p_{x_0 x_1} \cdots p_{x_{n-1} x_n}} \\ &= p_{x_n x_{n+1}} = \mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n].\end{aligned}$$

Where it is important to note that, again, we have implicitly assumed that $\mu(x_0) p_{x_0 x_1} \cdots p_{x_{n-1} x_n} > 0$. □

Question Given μ, p does a MC(μ, p) always exist?

1.2 Existence

Theorem 1.4. *Let p be a transition probability on E. Then there exist:*

(i) a measurable space (Ω, F),

(ii) a collection of probability measures $(\mathbb{P}_x)_x$ on (Ω, F), and

(iii) a sequence of random variables $(X_n)_{n \geq 0}$ on $(\Omega, \mathcal{F})$ such that for all $x \in E$, under $\mathbb{P}_x$, (X_n) is MC(δ_x, p).

Proof. We first fix a measure μ on E with $\mu(x) > 0$ for every x and construct a MC(μ, p) on some abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Consider X_0 a random variable with law μ, $U_1, U_2, \dots$ i.i.d uniform random variable on $[0, 1]$. One can construct a measurable function $\Phi : E \times [0, 1] \to E$ such that for any $x \in E$ $\mathbb{P}[\Phi(x, U) = y] = p_{xy}$. To achieve this, order $E = \{x_1, x_2, \dots\}$ and define for every $s_{i,j} = \sum_{k<j} p_{x_i x_k}$ for every i, j, then set $\Phi(x_i, u) = x_j$ if $s_{ij} \leq u < s_{ij} p_{x_i x_j}$. Define by induction, for every $n \geq 0$

$$X_{n+1} = \Phi(X_n, U_{n+1}).$$

Then we have for every $x_0, \dots, x_n \in E$

$$\begin{aligned}\mathbb{P}[X_0 = x_0, \dots, X_n = x_n] &= \mathbb{P}[X_0 = x_0, \Phi(x_0, U_1) = x_1, \dots, \Phi(x_{n-1}, U_n) = x_n] \\ &= \mu(x_0) p_{x_0 x_1} \cdots p_{x_{n-1} x_n}.\end{aligned}$$

Now define for every $x \in E$ $\mathbb{P}_x = \mathbb{P}[\,\cdot \mid X_0 = x]$, this is well defined as $\mu(x) > 0$. Then we have that for all $x \in E$

$$\mathbb{P}_x[X_0 = x_0, \ldots, X_n = x_n] = \delta_x(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n}.$$

We consider a measure μ on E such that for every $x \in E : \mu(x) > 0$, some abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let X_0 be a random variable with distribution μ. Let $U_1, U_2, \ldots$ be i.i.d. uniform random variables on $[0, 1]$. Our goal is to use these uniform random variables to produce the probabilities given by the transition probabilities, in a way similar to Sklar's Theorem (knowledge of Sklar's is not needed here). To do this we enumerate $E = \{x_i, i > 0\}$ and set $s_{ij} = \sum_{k<j} p_{x_ix_k}$. Note here that $s_{i,j+1} - s_{i,j} = p_{x_ix_j}$. Finally, set

$$\Phi : E \times [0, 1] \to E;\ (x_i, u) \mapsto x_j \text{ if } u \in (s_{ij}, s_{i,j+1}].$$

Now we have X_0 as needed and the tools to construct the sequence of random variables, along with the collection of probability measures we want.

We now have that $\mathbb{P}[\Phi(x, U_1) = y] = p_{xy}$. So if we set $X_{n+1} = \Phi(X_n, U_{n+1})$ for every $n > 0$ (by induction), we find that

$$\begin{aligned}\mathbb{P}[X_0 = x_0, \ldots, X_n = x_n] &= \mathbb{P}[X_0 = x_0, \Phi(x_0, U_1) = x_1, \ldots, \Phi(x_{n-1}, U_n) = x_n] \\ &= \mu(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n},\end{aligned}$$

by independence.

Now if we define $\mathbb{P}_x$ as $\mathbb{P}[\,\cdot \mid X_0 = x]$, then we have for every $x \in E$ that

$$\mathbb{P}_x[X_0 = x_0, \ldots, X_n = x_n] = \delta_x(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n}.$$

□

Framework for the rest of the chapter E is finite or countable, p transition probability, $(\Omega, \mathcal{F}, (P_x)_{x\in E})$ Probability Spaces, $(X_n)_{n\geq 0}$ random variables such that it is a MC(δ_x, p) under P_x.

For μ a probability measure on E we write $\mathbb{P}_\mu = \sum_x \mu(x)\mathbb{P}_x$.

1.3 Simple Markov Property

Remark 1.5. Under $\mathbb{P}_\mu$, $X = (X_n)_{n\geq 0}$ is MC(μ, p).

$$\mathbb{P}_\mu[X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, .X_n = x_n] = \mathbb{P}_\mu[X_{n+1} = x_{n+1} \mid X_n = x_n] = \mathbb{P}_{x_n}[X_1 = x_{n+1}]$$

i.e. conditional on $X_n = x$, x_{n+1} is sampled like the first step of a MC(δ_x, p) independent of the past.

Notation $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.

Theorem 1.6 (Simple Markov Property (SiMP))**.** *Let μ be a distribution on E. Let $x \in E, k \in \mathbb{N}$. For every $f : E^{\mathbb{N}} \to \mathbb{R}_+$ measurable and bounded, for every Z bounded which is $\mathcal{F}_k$ measurable random variable, we have*

$$\boxed{\mathbb{E}_\mu \left[f((X_{k+n})_{n\geq 0}) Z \mid X_k = x_k \right] = \mathbb{E}_{x_k} \left[f((X_n)_{n\geq 0} \right] \mathbb{E}_\mu \left[Z \mid X_k = x \right].}$$

Proof. First note that using $Z = \mathbb{1}_{X_0 = x_0, \ldots, X_{k-1} = x_{k-1}}$ we only have to prove that

$$\mathbb{E}_\mu \left[f((X_{k+n})_{n\geq 0}) \mid X_0 = x_0, \ldots, X_k = x_k \right] = \mathbb{E}_{x_k} \left[f((X_n)_{n\geq 0}) \right].$$

We will proceed using induction. Approximate f by step functions f_k, using linearity, we only have to show our claim for the function $\mathbb{1}_A$ with $A \subset E^{\mathbb{N}}$ measurable, i.e.

$$\mathbb{P}_\mu \left[(X_{k+n})_{n\geq 0} \in A \mid X_0 = x_0, \ldots, X_k = x_k \right] = \mathbb{P}_{x_k} \left[(X_n)_{n\geq 0} \in A \right].$$

The collection of sets of the form $A = \{w \in E^{\mathbb{N}} : w_0 = y_0, \ldots, w_N = y_N\}$ for $N \geq 0$ and $y_0, \ldots, y_N \in E$ form a π-system generating the σ-algebra. Furthermore, on such sets

$$\begin{aligned}
&\mathbb{P}_\mu \left[(X_{k+n})_{n\geq 0} \in A \mid X_0 = x_0, \ldots, X_k = x_k \right] \\
&\quad = \mathbb{P}_\mu \left[X_k = y_0, \ldots, X_{k+N} = y_N \mid X_0 = x_0, \ldots, X_k = x_k \right] \\
&\quad = \frac{\mu(x_0) p_{x_0 x_1} \cdots p_{x_{k-1} x_k} \delta_{x_k}(y_0) p_{y_0 y_1} \cdots p_{y_{N-1} y_N}}{\mu(x_0) p_{x_0 x_1} \cdots p_{x_{k-1} x_k}} \\
&\quad = \delta_{x_k}(y_0) p_{y_0 y_1} \cdots p_{y_{N-1} y_N} \\
&\quad = \mathbb{P}_{x_k} \left[(X_n)_{n\geq 0} \in A \right].
\end{aligned}$$

Dynkin's Lemma then allows us to extend this property to the entire σ-algebra. $\square$

Corollary 1.7. *Let μ be a distribution on E, $x \in E$, $k \in \mathbb{N}$, for all $f : E^{\mathbb{N}} \to \mathbb{R}$ measurable and bounded:*

$$\boxed{\mathbb{E}_\mu \left[f((X_{k+n})_{n\geq 0} \mid X_k = x \right] = \mathbb{E}_x \left[f((X_n)_{n\geq 0} \right].}$$

1.4 n-Step Transition Probabilities

Definition 1.4. For every $n \geq 0$, $x, y \in E$, define $p_{xy}^{(n)} = P_x[X_n = y]$.

Proposition 1.8 (Chapman Kolmogorov (CK))**.**

$$\forall m, n \geq 0 \quad \forall x, y \in E \quad \boxed{p_{xy}^{(m+n)} = \sum_{z \in E} p_{xz}^{(m)} p_{zy}^{(n)}.}$$

TODO: Figure

Proof. Fix m, n and $x, y \in E$.

$$\begin{aligned} p_{xy}^{(m+n)} &= \mathbb{P}_x\left[X_{m+n} = y\right] = \sum_{z \in E} \mathbb{P}_x\left[X_{m+n} \mid X_m = z\right] \mathbb{P}_x\left[X_m = z\right] \\ &\overset{\text{(SiMP)}}{=} \sum_{z \in E} \mathbb{P}_z\left[X_n = y\right] \mathbb{P}_x\left[X_m = z\right] = \sum_{z \in E} p_{xz}^{(m)} p_{zy}^{(n)}. \end{aligned}$$

□

Proposition 1.9. *Assume E is finite: The matrix $(p_{ij}^{(n)})_{ij \leq 0}$ is equal to P^n. For every n, μ a distribution on E, and any $f : E \to \mathbb{R}$ we have*

$$\mathbb{E}_\mu\left[f(X_n)\right] = \mu P^n f,$$

for any $n \geq 0$, with $f = [f(1), \ldots, f(n)]^T$.

Proof. The first equation follows from $p_{ik}^{(n+1)} = \sum_j p_{ij}^{(n)} p_{jk}$ by induction. For the second equation, use the definition of the expectation

$$\mathbb{E}_\mu\left[f(X_n)\right] = \sum_{y \in E} f(y) \mathbb{P}_\mu\left[X_n = y\right] = \sum_{x, y \in E^2} \mu(x) \underbrace{\mathbb{P}_x\left[X_n = y\right]}_{= p_{xy}^{(n)}} f(y).$$

□

1.5 Stationary Distributions

Motivation: write μ_n as the law of X_n under P_μ, $\mu_0 = \mu$ and $\mu_{n+1} = \mu_n P$. For n large μ_n is a fixed point of the map $\lambda \to \lambda P = \left(\sum_{x \in E} \lambda(z) p_{xy}\right)_{y \in E}$

Definition 1.5. Let π be a distribution on E, we say that π is stationary (for p) if for $y \in E$

$$\boxed{\pi(y) = \sum_{x \in E} \pi(x) p_{xy}.}$$

Linear Algebra interpretation If E is finite and we write $\pi = [\pi(1), \ldots, \pi(n)]^T$, then

$$\boxed{\pi \text{ is stationary} \iff \pi P = \pi,}$$

i.e. π is a left eigenvector of P for the eigenvalue 1.

Probabilistic interpretation If π is a stationary distribution, then for all $n \geq 0$

$$P_\pi[X_n = x] = \pi(x).$$

No matter how far along you are in the chain, the probability that you land on a value x is equal to the probability that you start at x.

1.6 Reversibility

Definition 1.6. A distribution π on E is said to be reversible (for p) if for any $x, y \in E$

$$\boxed{\pi(x)p_{xy} = \pi(y)p_{yx}.}$$

The probability of starting at y and going to x is equal to the probability of starting at x and going to y. More generally, one can prove by induction that π is reversible if and only if for every n and any $x_0, \ldots, x_n$

$$\mathbb{P}_\pi\left[X_0 = x_0, \ldots, X_n = x_n\right] = \mathbb{P}_\pi\left[X_0 = x_n, \ldots, X_n = X_0\right].$$

Motivation We want an easy criterion for invariance, such reversible systems appear often in physics.

Proposition 1.10. *Let π be a distribution on E, if π reversible, then π is stationary.*

Proof.

$$\sum_{x \in E} \pi(x)p_{xy} = \sum_{x \in E} \pi(y)p_{yx} = \pi(y)\sum_{x \in E} p_{yx} = \pi(y).$$

□

Example 1.3 (Gas in Containers (Ehrenfest Model)). Imagine there are two containers A and B with gas particles, between them is a small hole through which the particles can pass through. At every step a single particle is selected uniformly at random and passes through this hole. To represent this mathematically, let X_n be the number of particles in A at time n, and let there be N total particles. We assume that the system in time homogeneous (time plays no role in its evolution, only its current state) and is memoryless (again only the current state of the system plays a role). This gives us the inspiration to model X_n as a Markov Chain. The transition probabilities are given by $p_{x,x+1} = 1 - \frac{x}{N}$, as in order for X_n to grow by 1, the randomly selected particle must be from container B; this occurs with probability $\frac{\#\text{ of particles in } B}{\#\text{ of total particles}} = \frac{N-x}{N}$. The only other option is for the amount of particles in A to decrease by 1, by the fact that the transition probabilities must sum to 1 we find: $p_{x,x-1} = \frac{x}{N}$. Now we wonder if it is possible to find a stationary distribution, this would represent the equilibrium distribution of particles (see the different interpretations above). To find this distribution, we instead simplify and see if we can find a reversible distribution, i.e. $\pi(x)p_{x,x+1} = \pi(x+1)p_{x+1,x}$. We then use this to calculate $\pi(x)$ explicitly and see if this defines a proper distribution.

$$\pi(x+1) = \frac{\pi(x)(1-\frac{x}{N})}{\frac{x+1}{N}} = \pi(x)\frac{N-x}{x+1} \overset{\text{(Induction)}}{=} \pi(0)\frac{N\cdots(N-x)}{(x+1)!}.$$

Thus we find that $\pi(x) = \binom{N}{x}\pi(0)$, π should define a distribution. This entails that the total mass of π be 1, i.e. $\sum_{x\in E}\pi(x) = 1$. Hence we find

$$\pi(0) = \left(\sum_{x\in E}\binom{N}{x}\right)^{-1} = \frac{1}{2^N}.$$

Hence, $\pi(x) = \binom{N}{x}\frac{1}{2^N}$, the binomial distribution; which is (as we have shown) reversible. When X_{n+1} is distributed like X_n (equilibrium) then the number of particles in A is distributed as $Bin(N, \frac{1}{2})$.

1.7 Communication Classes

Here we will see p as a weighted oriented graph.

Definition 1.7. Let $x, y \in E$. We say that "y can be reached from x" if there exists an $n \geq 0$ such that $p^{(n)}_{xy} > 0$ and we write $x \to y$. Furthermore, we say that "x and y communicate" if $y \to x$ and $x \to y$, and we write $x \leftrightarrow y$.

Proposition 1.11. $\leftrightarrow$ *is an equivalence class on E.*

Proof. Trivial $\square$

Definition 1.8. The equivalence classes of $\leftrightarrow$ are called communication classes, and if there is a single unique communication class for a chain p, we say that p is irreducible.

Motivation We will see that p irreducible implies that p has at most one stationary distribution.

Definition 1.9. A communication class C is closed if for any $x, y \in E$

$$x \in C, x \to y \implies y \in C.$$

"If one starts in C, one never leaves."

1.8 Strong Markov Property

Definition 1.10. Let $T : \Omega \to \mathbb{N} \cup \{+\infty\}$ random variable with values in $\mathbb{N} \cup \{+\infty\}$. We say that T is an $(\mathcal{F}_n)$-stopping time if

$$\forall n \in \mathbb{N}\{T = n\} \in \mathcal{F}_n.$$

Example 1.4 (Stopping Times). $H_A = \min\{n \geq 0 : X_n \in A\}$ (for $A \subset E$) and $H_x = \min\{n \geq 0 : X_n = x\}$ are stopping times.

Definition 1.11. Let T be a stopping time.

$$\mathcal{F}_T = \{A \in \mathcal{F} : \forall n \in \mathbb{N} : \{T = n\} \cap A \in \mathcal{F}_n\}.$$

Theorem 1.12 (Strong Markov Property (SiMP))**.** *Let μ be a distribution on E, T an $\mathcal{F}_n$-stopping time. Let $x \in E$, then for all $f : E^{\mathbb{N}} \to \mathbb{R}$ measurable and bounded, and Z which are $\mathcal{F}_T$ measurable and bounded, we have:*

$$\boxed{\mathbb{E}_\mu \left[f((X_{T+n})_{n\geq 0}) \cdot Z \mid T < \infty, X_T = x \right] = \mathbb{E}_x \left[f((X_n)_{n\geq 0}) \right] \mathbb{E}_\mu \left[Z \mid T < \infty, X_T = x \right].}$$

"Conditioned on $\{T < \infty, X_T = x\}$, $(X_{T+n})_{n\geq 0}$ is a MC(δ_x, p) independent of $\mathcal{F}_T$ "

Proof. We will multiply each side of the equation by $\mathbb{P}\left[T < \infty, X_T = x\right]$.

$$\begin{aligned}
\mathbb{E}_\mu \left[f((X_{T+n})_{n\geq 0}) Z \mathbb{1}_{T<\infty, X_T = x} \right] &= \sum_{k\geq 0} \mathbb{E}_\mu \left[f((X_{k+n})_{n\geq 0} Z \mathbb{1}_{T=k, X_T=k} \right] \\
&= \sum_{k\geq 0} \mathbb{E}_\mu \left[f((X_{k+n})_{n\geq 0}) Z \mathbb{1}_{T=k} \mid X_k = x \right] \mathbb{P}_\mu \left[X_k = x \right] \\
&\overset{\text{(SiMP)}}{=} \sum_{k\geq 0} \mathbb{E}_x \left[f((X_n)_{n\geq 0}) \right] \mathbb{E}_\mu \left[Z \mathbb{1}_{T=k, X_k=x} \right] \\
&= \mathbb{E}_x \left[f((X_n)_{n\geq 0} \right] \sum_{k\geq 0} \mathbb{E}_\mu \left[Z \mathbb{1}_{T=k, X_k=x} \right] = \mathbb{E}_x \left[f((X_n)_{n\geq 0}) \right] \mathbb{E}_\mu \left[Z \mathbb{1}_{T<\infty, X_T=x} \right].
\end{aligned}$$

∎

Application Reflection Principle for the SRW.

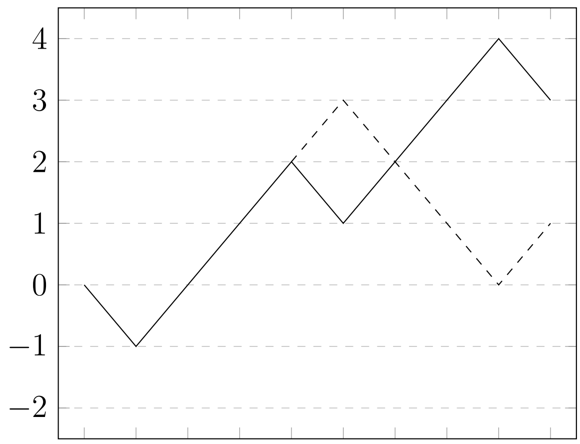

Figure 1.2: Example of a reflected simple random walk for $a = 2$.

Consider the SRW on $\mathbb{Z}$

Proposition 1.13. *Let $k \geq 0$ even, $a \geq 1$ odd. We have*

$$\mathbb{P}_0\left[\max_{0\leq m\leq k} X_m \geq a\right] = \mathbb{P}_0\left[|X_k| \geq a\right].$$

Proof. Define $H_a = \min\{n \geq 1 : X_n = a\}$, this is a stopping time.

$$\mathbb{P}_0\left[\max_{0\leq m\leq k} X_m \geq a\right] = \mathbb{P}_0\left[H_a \leq k\right] = \mathbb{P}_0\left[X_k > a\right] + \mathbb{P}_0\left[H_a \leq k, X_k < a\right].$$

Now our goal is to show that the term on the right is equal to $\mathbb{P}_0\left[X_k > a\right]$, as $2\mathbb{P}_0\left[X_k > a\right] = \mathbb{P}_0\left[|X_k| > a\right]$ by symmetry. We can go from $>$ to $\geq$ because a is even and k is odd. At this point we note that X_{H_a+n} is distributed as $a + (a - X_{H_a+n}) = 2a - X_{H_a+n}$. Geometrically, this means that if we only look at the walk after hitting a, the walk has the same distribution if we inverse the direction of each step: 'looking at the path after hitting a, we cannot tell if it is the normal or the inverted step walk'. We have

$$\begin{aligned}\mathbb{P}_0\left[H_a \leq k, X_k < a\right] &= \sum_{m=0}^{k} \mathbb{P}_0\left[X_k < a, H_a = m\right] = \sum_{m=0}^{k} \mathbb{P}_a\left[X_{k-m} < a\right]\mathbb{P}_0\left[H_a = m\right] \\ &= \sum_{m=0}^{k} \mathbb{P}_a\left[X_{k-m} > a\right]\mathbb{P}_0\left[H_a = m\right] = \sum_{m=0}^{k} \mathbb{P}_0\left[X_{k-m} > a, H_a = m\right] \\ &= \mathbb{P}_0\left[X_k > a, H_a \leq k\right] = \mathbb{P}_0\left[X_k > a\right].\end{aligned}$$

□

Conclusion Now we have properly defined a Markov Chain, shown its existence, and introduced some concepts to help classify different types of chains. Importantly, we have also introduced the transition probability framework.

Chapter 2

Markov Chains: Long Time Behavior

// I have set p to P to denote the collection of transition probabilities here, to see if you like it. //

Outset With the tools and classification concepts introduced previously, we would like to expand upon these to rigorously classify chains.

Framework: E finite or countable, $p = (p_{xy})x, y \in E$ transition probabilities, $(\Omega, F, (\mathbb{P}_x)_{x\in E})$, $X = (X_n)_{n\geq 0} \sim MC(\delta_x, P)$ under $\mathbb{P}_x$, $\mathbb{P}_\mu = \sum \mu(x)\mathbb{P}_x$.

Questions:

- When does there exist a stationary distribution?
- What is the behavior of X_n for n large?
- If we fix $x \in E$, will the chain visit x infinitely many times?

2.1 Recurrence/Transience

Notation $H_x = \min\{n \geq 1 : X_n = x\}$

Definition 2.1. Let $x \in E$, we say that:

- x is recurrent if $\boxed{\mathbb{P}_x [H_x < \infty] = 1}$.
- x is transient if $\boxed{\mathbb{P}_x [H_x < \infty] < 1}$.

Notation: For $x \in E$ write $V_x = \sum_{n\geq 0} \mathbb{1}_{X_n=x}$, i.e. the total number of visits.

Theorem 2.1 (Dichotomy Theorem). *$x \in E$:*

- *if x is recurrent, then* $\boxed{V_x = +\infty}$ $\mathbb{P}_x$*-a.s.*.
- *if x is transient, then* $\boxed{\mathbb{E}_x [V_x] < \infty}$.

Remark 2.2. It is impossible that $\mathbb{P}_x [V_x < \infty] > 0$ and $\mathbb{E}_x [V_x] = +\infty$.

Definition 2.2. $\rho_x = \mathbb{P}_x [H_x < \infty]$, if x is recurrent then $\rho_x = 1$, otherwise if x is transient $\rho_x < 1$. Thus the number of visits is a geometric RV with parameter $\rho_x < 1$.

// You didn't include this as a definition, and just as part of the statement of the lemma.//

Lemma 2.3. *For every $i \geq 0, x \in E$, we have $\mathbb{P}_x [V_x \geq i] = \rho_x^i$.*

Proof (Lemma). Define for every $i \geq 0$, $T_i = \min\{n > 0 : \sum_{k=1}^{n} \mathbb{1}_{X_k = x} = i\}$ 'the time of the i-th visit of x'. T_i is a stopping time because $\{T_i = n\} = \{\sum_{k=1}^{n-1} \mathbb{1}_{X_k = x} = i-1, X_n = x\} \in \mathcal{F}_n$.

$$\begin{aligned}
\mathbb{P}_x [V_x \geq i] &= \mathbb{P}_x [T_i < \infty, T_i < \infty] \\
&= \mathbb{P}_x [T_i < \infty | T_i < \infty, X_{T_i} = x] \, \mathbb{P}_x [T_i < \infty] \\
&\overset{\text{(SMP)}}{=} \mathbb{P}_x \left[T_{(1)} < \infty\right] \rho_x^{i-1} = \rho_x^i.
\end{aligned}$$

// In your notes you use T_i for $H_x^{(i)}$, I am using H as we have used it previously and use it again later on, furthermore we know that it is a stopping time already. //

We will proceed by induction over i. Define $H_x^{(i)}$ to be the i-th hit time of x. For $i = 0$ the claim is clear.

$$\begin{aligned}
\mathbb{P}_x [V_x \geq i+1] &= \mathbb{P}_x [V_x \geq i+1, V_x \geq i] = \mathbb{P}_x \left[H_x^{(i+1)} < \infty, H_x^{(i)} < \infty\right] \\
&= \mathbb{P}_x \left[H_x^{(i+1)} < \infty | H_x^{(i)} < \infty, X_{H_x^{(i)}} = x\right] \mathbb{P}_x \left[H_x^{(i)} < \infty\right] \\
&\overset{\text{(SMP)}}{=} \mathbb{P}_x \left[H_x^{(1)} < \infty\right] \rho_x^i = \rho_x^{i+1}.
\end{aligned}$$

□

Proof (Theorem). For x recurrent:

$$\mathbb{P}_x [V_x = \infty] = \mathbb{P}_x \left[\bigcap_{i=0}^{\infty} \{V_x \geq i\}\right] = \lim_{i \to \infty} \mathbb{P}_x [V_x \geq i] = \lim_{i \to \infty} \rho_x^i = 1.$$

For x transient:

$$\begin{aligned}\mathbb{E}_x[V_x] &= \sum_{k=0}^{\infty} k\mathbb{P}_x[V_x = k] = \sum_{k=1}^{\infty}\sum_{j=1}^{k} \mathbb{P}_x[V_x = k] = \sum_{j=1}^{\infty}\sum_{k=j}^{\infty} \mathbb{P}_x[V_x = k] \\ &= \sum_{j=1}^{\infty} \mathbb{P}_x[V_x \geq j] = \sum_{j=1}^{\infty} \rho_x^k = \frac{\rho_x}{1-\rho_x} < \infty.\end{aligned}$$

// I added this bit in the equation when I wrote this proof before, I think this should go in the appendix now// □

Proposition 2.4. *If E is finite, then there exists a recurrent state $x \in E$.*

Proof.

$$\sum_{x\in E} V_x = \sum_{x\in E}\sum_{n\geq 0} \mathbb{1}_{X_n = x} = \sum_{n\geq 0}\sum_{x\in E} \mathbb{1}_{X_n = x} = \sum_{n\geq 0} 1 = \infty$$

Fix some $y \in E$.

$$\sum_{x\in E} \mathbb{E}_y[V_x] = \mathbb{E}_y\left[\sum_{x\in E} V_x\right] = \infty.$$

Thus we know there exists $x \in E$ such that $\mathbb{E}_y[V_x] = \infty$. Using that $V_x = V_x \mathbb{1}_{H_x<\infty}$, we find

$$\infty = \mathbb{E}_y[V_x \mathbb{1}_{H_x<\infty}] \stackrel{\text{(SMP)}}{=} (1 + \mathbb{E}_x[V_x])\mathbb{P}_y[H_x < \infty] \leq \mathbb{E}_x[V_x].$$

Therefore, $\mathbb{E}_x[V_x] = \infty$, which concludes that x is recurrent.

// I am not sure the 1+ part of the above equation is correct as we have defined V_x to be the number of visits starting from $n = 0$, so the hit that happens at H_x is already included in $\mathbb{E}_x[V_x]$.//

Thus we know there exists $x \in E$ such that $\mathbb{E}_y[V_x] = \infty$, since the sum on the left is over a finite index set (E finite). We can write $V_x = V_x \mathbb{1}_{H_x<\infty}$, we find that (using the Strong Markov Property)

$$\infty = \mathbb{E}_y[V_x] = \mathbb{E}_y[V_x \mathbb{1}_{H_x<\infty}] = \mathbb{E}_x[V_x]\,\mathbb{P}_y[H_x < \infty],$$

because a chain started from y is the same (in the distribution sense) after hitting x as a chain started from x. $\mathbb{P}_y[H_x < \infty]$ must be ≤ 1, thus the term of the left must be equal to ∞, implying that $\mathbb{E}_x[V_x] = \infty$. □

2.2 Recurrence/Transience for the SRW on $\mathbb{Z}^d$

SRW on $\mathbb{Z}^d$: $E = \mathbb{Z}^d$, $p_{xy} = \frac{1}{2d}$ *if* $\|x - y\|_1 = 1, 0$ *else*

Theorem 2.5 (Polya)**.** *For the SRW, every state is recurrent if $d = 1, 2$, otherwise they are transient.*

Proof. Let $(Z_k)_{k>0}$ be i.i.d. random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}[Z_i = \pm e_i] = \frac{1}{2d}$ for $i = 1 \dots d$ and e_i being the unit vectors in $\mathbb{Z}^d$. Next, define $X_n = \sum_{k=1}^n Z_k$, which is a $MC(\delta_0, P)$. Then we find the following

$$\mathbb{E}[V_0] = \mathbb{E}\left[\sum_{n>0} \mathbb{1}_{X_n=0}\right] = \sum_{n>0} \mathbb{P}[X_n = 0].$$

Idea: $\mathbb{P}[X_n = x] = \mathbb{P}[Z_1 + \ldots + Z_n = x] = \sum_{\delta_1+\ldots+\delta_n=x} \mathbb{P}[Z_1 = \delta_1] \cdots \mathbb{P}[Z_n = \delta_n]$ is not easy to calculate. Instead, we use could use the Fourier Transform to link $\mathbb{E}\left[e^{iX_n}\right] = \mathbb{E}\left[e^{iZ_1}\right]^n$ to $(\mathbb{P}[X_n = x])_{x\in\mathbb{Z}^d}$. Define $\phi(\xi) = \mathbb{E}\left[e^{i\xi\cdot Z_1}\right]$ for ξ in $\Pi^d = [-\pi, \pi)^d$. Then we have

$$\phi(\xi) = \frac{1}{2d}\sum_{i=1}^d (e^{i\xi\cdot e_1} + e^{-i\xi\cdot e_1}) = \frac{1}{d}\sum_{i=1}^d \cos(\xi_i).$$

Fixing $n \geq 0$, we have (by independence) that the characteristic function of X_n is

$$\varphi_{X_n}(\xi) = \mathbb{E}\left[e^{i\xi\cdot X_n}\right] = \mathbb{E}\left[e^{i(\xi\cdot Z_1+\ldots+\xi\cdot Z_n)}\right] = \phi(\xi)^n.$$

We can now take advantage of the Fourier Transform by using the Fourier inversion, giving

$$\mathbb{P}[X_n = 0] = \frac{1}{(2\pi)^d}\int_{\Pi^d} \phi(\xi)^n d\xi.$$

Check this by directly computing

$$\begin{aligned}\int_{[0,2\pi)^d} \phi(\xi)^n d\xi = \int_{[-\pi,\pi)^d} \sum_{x\in\mathbb{Z}^d} \mathbb{P}[X_n = x] &= \sum_{x\in\mathbb{Z}^d} \mathbb{P}[X_n = x] \int_{[0,2\pi)^d} e^{i\xi\cdot x} d\xi \\ &= \begin{cases} (2\pi)^d & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases}.\end{aligned}$$

Therefore,

$$\begin{aligned}(2\pi)^d \sum_{n\geq 0} \mathbb{P}[X_n = 0] &= \sum_{n>0} \int_{\Pi^d} \phi(\xi)^n d\xi \overset{\text{(MCT)}}{=} \lim_{\alpha\uparrow 1} \sum_{n>0} \int_{\Pi^d} (\alpha\phi(\xi))^n d\xi \\ &\overset{\text{(Fub)}}{=} \lim_{\alpha\uparrow 1} \int_{\Pi^d} \frac{1}{1 - \alpha\phi(\xi)} d\xi \overset{\text{(MCT)}}{=} \int_{\Pi^d} \frac{1}{1 - \phi(\xi)} d\xi.\end{aligned}$$

We can see that for any $\xi_i \in [-\pi, \pi)$ we have the inequality $\frac{\xi_i^2}{6} \leq 1 - \cos(\xi_i) \leq \frac{\xi_i^2}{2}$.

With this, we find that $\frac{1}{6d}\|\xi\|_2^2 \leq 1 - \phi(\xi) \leq \frac{1}{2d}\|\xi\|_2^2$, finally giving us

$$\sum_{n\geq 0} \mathbb{P}\left[X_n = 0\right] < \infty \Longleftrightarrow \int_{B_1(0)} \frac{d\xi}{\|\xi\|_2^2} < \infty \Longleftrightarrow d > 2.$$

The final equivalence can be justified by using a change of variables and homogeneity. Define $A_i = B_0(2^{-i}) \setminus B_0(2^{-(i+1)})$ for every i. Next use the change of variable $\psi = 2^i\xi$, we find that

$$\int_{A_i} \frac{d\xi}{\|\xi\|^2} = \int_{A_0} \frac{2^{2i}}{\|\psi\|^2}(2^i)^{-d} d\psi = (2^i)^{2-d} \underbrace{\int_{A_0} \frac{d\psi}{\|\psi\|^2}}_{=:I_0}.$$

Therefore

$$\int_{B_0(1)} \frac{d\xi}{\|\xi\|^2} = \sum_{i=0}^{\infty} \int_{A_i} \frac{d\xi}{\|\xi\|^2} = I_0 \sum_{i=0}^{\infty} (2^i)^{2-d}.$$

Which is finite if and only if $d > 2$. $\square$

2.3 Classification of States

Theorem 2.6. *Let $x, y \in E$ such that $x \to y$. If x is recurrent then y is recurrent and $\mathbb{P}_x\left[H_y < \infty\right] = \mathbb{P}_y\left[H_x < \infty\right] = 1$. In particular $x \leftrightarrow y$.*

Proof. We want to use that every time the chain visits x, it has a non-zero probability to visit y after that, visiting x infinitely often should ensure that y is also visited infinitely often. Assume $y \neq x$ and x recurrent. Let $z_1, \ldots, z_{k-1}$ be distinct elements of E, not equal to x or y such that $p_{xz_1} \cdots p_{z_{k-1}y} > 0$. Then we have

$$\begin{aligned} 0 &= \mathbb{P}_x\left[H_x = \infty\right] \geq \mathbb{P}_x\left[X_1 = z_1, \ldots, X_k = 1, \forall n > 0\ X_{k+n} \neq x\right] \\ &\overset{\text{(SiMP)}}{=} \underbrace{\mathbb{P}_x\left[X_1 = z_1, \ldots, X_k = y\right]}_{>0} \underbrace{\mathbb{P}_y\left[\forall n > 0\ X_n \neq x\right]}_{\mathbb{P}_y[H_x = \infty]}. \end{aligned}$$

Thus $\mathbb{P}_y\left[H_x < \infty\right] = 1$. Next, we have to show that y is recurrent. Choose m, n such that $p_{xy}^{(n)}, p_{yx}^{(m)} > 0$, we have

$$\mathbb{E}_y\left[V_y\right] = \sum_{k>0} p_{yy}^{(k)} \geq \sum_{k>0} p_{yy}^{(m+k+n)} \overset{\text{(CK)}}{\geq} \underbrace{p_{yx}^{(m)}}_{>0} \underbrace{\left(\sum_{k>0} p_{xx}^{(k)}\right)}_{=\infty} \underbrace{p_{xy}^{(n)}}_{>0}.$$

Hence, y is recurrent. To show that $\mathbb{P}_x\left[H_y < \infty\right] = 1$, use the same argument as above, but with the roles of x and y swapped ($y \to x$, y recurrent), as before. $\square$

Remark 2.7. x recurrent and $x \neq y$ then $x \to y$ if and only if $\mathbb{P}_x[\exists n : X_n = y] > 0$ if and only if $\mathbb{P}_x[H_y < \infty] = 1$

Corollary 2.8. *Let C communication class for p. Either x is recurrent for every $x \in E$, or every $x \in E$ is transient.*

Proof. If $x \leftrightarrow y$, we have that x recurrent if and only if y is recurrent. □

Remark 2.9. We call a class which is comprised of recurrent states a recurrent class, and one comprised of transient states a transient class.

Corollary 2.10. *A recurrent class is always closed.*

Proof. C recurrent, $x \in C$, if $x \to y$ then we must have $y \to x$ (otherwise x wouldn't be recurrent), therefore $y \in C$. □

Corollary 2.11. *Thus we have an intuitive criterion for transience: if $x \to y$ but $y \nrightarrow x$, then x is transient. 'If we start at x and can get to a state, from which we cannot return to x, then x is transient'.*

Remark 2.12. In general it is possible to find disjoint subsets of E, T and $(R_k)_{k>0}$ such that T is the class of transient states, and R_k are recurrent classes, with $E = T \cup \bigcup_{k>0} R_k$. Then we can broadly classify the behavior of the chain by differentiating if X_n starts in some R_k and if X_n starts in T. In the former case the chain remains in R_k forever. If X_n starts in T, either it remains in T forever, or at some point it moves into an R_k and remains there forever.

2.4 Positive/Null Recurrence

Notation For $x \in E$ write $m_x = \mathbb{E}_x[H_x]$.

Definition 2.3. Let $x \in E$ be a recurrent state. We say that x is:

- positive recurrent if $\boxed{m_x < \infty}$
- null recurrent if $\boxed{m_x = +\infty}$.

Theorem 2.13. *Let $x, y \in E, x \leftrightarrow y$. Then*

$$\boxed{\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{xy}^{(k)} = \frac{1}{m_y}}.$$

Remark 2.14. Write $V_y^{(n)} = \sum_{k=1}^n \mathbb{1}_{X_k=y}$, "The number of visits to y up to time n". Thus the sum in the theorem is "Expected proportion of time spent at y".

//I think going from starting the sum at 0 to starting the sum at 1, causes confusion (at least it confused many people I know), so sticking to one definition would be better//

If y is transient, or null recurrent ($m_y = \infty$), the theorem tells us that $\lim_{n\to\infty} \mathbb{E}_x\left[\frac{V_y^{(n)}}{n}\right] = 0$: "null density of visits". Otherwise, y is positive recurrent and the expected density of visits is strictly positive.

Definition 2.4 (inter-visit times)**.** Let $y \in E$. Define $H_y^0 = H_y$ and for all $i \geq 1 : H_y^i = \min\{n \geq 1 : X_{H_y^0+\ldots+H_y^{i-1}+n} = y\}$ if $H_y^{i-1} < \infty$, else $+\infty$.

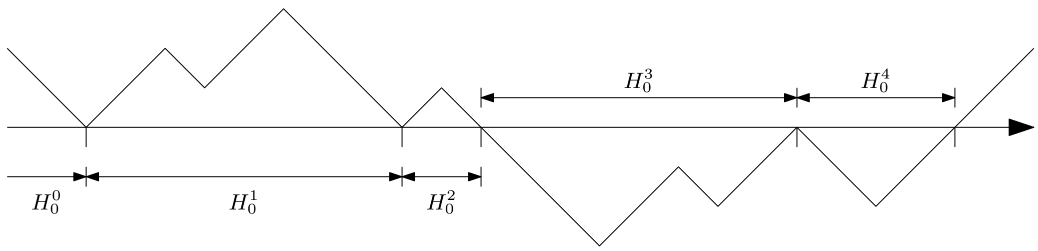

Figure 2.1: Inter-arrival times for the Simple Random Walk.

Lemma 2.15. *Let x, y with $x \leftrightarrow y$, assume y is recurrent. Then for every $j \geq 1$ and $t_0, \ldots, t_j \in \mathbb{N}$:*

$$\mathbb{P}_x\left[H_y^0 = t_0 \ldots H_y^j = t_j\right] = \mathbb{P}_x\left[H_y = t_0\right]\mathbb{P}_y\left[H_y = t_1\right]\ldots\mathbb{P}_y\left[H_y = t_j\right].$$

Under P_x, $H_y^1, H_y^2, \ldots$ are i.i.d. with law $\mathbb{P}_x\left[H_y^i = t\right] = \mathbb{P}_y\left[H_y = t\right]$.

Proof. Write $H^i = H_y^i$, we will prove via induction over j. We have $\mathbb{P}_x\left[H^0 = t\right] = \mathbb{P}_x\left[H_y = y\right]$ for all t, showing for $j = 0$. Let $j \geq 0$ and assume that the equation holds. First note that

$$\mathbb{P}_x\left[H^0 < \infty, \ldots, H^j < \infty\right] = \sum_{t_0,\ldots,t_j} \mathbb{P}_x\left[H^0 = t_0, \ldots, H^i = t_i\right] = \mathbb{P}_x\left[H_y < \infty\right]\mathbb{P}_y\left[H_y < \infty\right]^i = 1$$

$$\mathbb{P}_x\left[H^0 < \infty, \ldots, H^j < \infty\right] = \sum_{t_0,\ldots,t_j} \mathbb{P}_x\left[H^0 = t_0, \ldots, H^j = t_j\right] = \mathbb{P}_x\left[H_y < \infty\right]\mathbb{P}_y\left[H_y < \infty\right]^j = 1.$$

Therefore $T = H^0 + \ldots + H^j$ is finite $\mathbb{P}_x$-a.s. and $X_T = y$. Hence, we have that for every $t_0, \ldots, t_{j+1} \in \mathbb{N}$ that

$$\begin{aligned}
\mathbb{P}_x\left[H^0 = t_0, \ldots, H^{j+1} = t_{j+1}\right] &= \mathbb{P}_x\left[H^0 = t_0, \ldots, H^{j+1} = t_{j+1} | T < \infty, X_T = y\right] \\
&\overset{\text{(SMP)}}{=} \mathbb{P}_x\left[H^0 = t_0, \ldots, H^j = t_j\right] \mathbb{P}_y\left[\min\{n > 0 : X_n = y\} = t_{j+1}\right] \\
&= \mathbb{P}_x\left[X_y = t_0\right] \mathbb{P}_y\left[H_y = t_1\right] \cdots \mathbb{P}_y\left[H_y = t_{j+1}\right].
\end{aligned}$$

□

Proof (Theorem). **Case 1:** y transient: we know that $\mathbb{E}_y[V_y] < \infty$, thus (Strong Markov Property) $\mathbb{E}_x[V_y] < \infty$. Hence for every $n > 0$,

$$\frac{\mathbb{E}_x\left[V_y^{(n)}\right]}{n} \leq \frac{\mathbb{E}_x[V_x]}{n} \to 0.$$

Case 2: y recurrent: using the lemma, we know that the random variables H^j are i.i.d. under $\mathbb{P}_x$ and fulfill $\mathbb{E}_x[H^1] = \mathbb{E}_y[H_y] = m_y$. Then we can use the Law of Large Numbers and $\mathbb{P}_x[H^0 < \infty] = 1$ we find $\mathbb{P}_x$-a.s.,

$$\lim_{i \to \infty} \frac{H^0 + \ldots + H^i}{i} = m_y.$$

Note that this includes the case of $m_y = \infty$ by truncation. Now we write $N_n = V_y^{(n)}$ (the number of visits to y at time n).

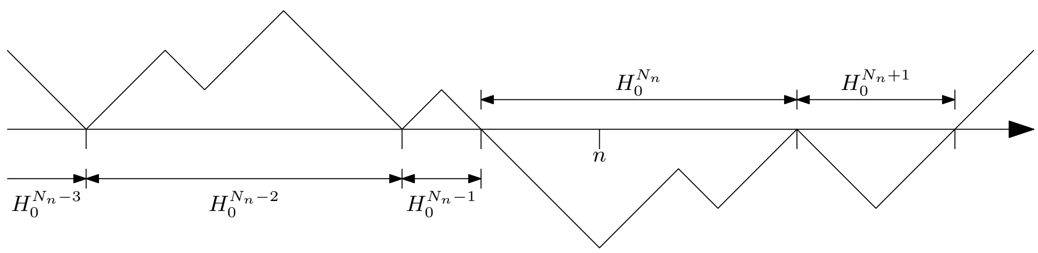

Figure 2.2: Inter-arrival times of 0 around time n of the Simple Random Walk.

Following directly from the definition of N_n we have that for any $n > 0$ that

$$H^0 + \ldots + H^{N_n - 1} \leq n < H^0 + \ldots + H^{N_n}.$$

Hence, for every $n > 0$

$$\frac{N_n}{H^0 + \dots H^{N_n}} < \frac{V_y^{(n)}}{n} \leq \frac{N_n}{H^0 + \dots + H^{N_n - 1}}.$$

The upper and lower bounds each converge to $\frac{1}{m_y}$ almost surely. Hence, we can conclude that $\mathbb{E}_x\left[\frac{V_y^{(n)}}{n}\right] \to \frac{1}{m_y}$ by the Dominated Convergence Theorem. $\square$

Proposition 2.16 (Classification of recurrent classes)**.** *Let R be a recurrent class. Then either:*

- *for all $x \in R$, x is positive recurrent, or*
- *for all $x \in R$, x is null recurrent.*

Proof. Fix $x, y \in E$ with $x \leftrightarrow y$ and x positive recurrent. Fix $k \geq 0$ with $p_{xy}^{(k)} > 0$. By Chapman-Kolmogorov, we have for all $j > 0$

$$p_{xy}^{(k+j)} \geq p_{xx}^{(j)} p_{xy}^{(k)}.$$

Thus

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} p_{xy}^{(i)}}_{\to \frac{1}{\mathbb{E}_y[H_y]}} \geq \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n-k} p_{xx}^{(j)}\right)}_{\to \mathbb{E}_x[H_x]} \underbrace{p_{xy}^{(k)}}_{>0}.$$

Therefore, $\frac{1}{\mathbb{E}_y[H_y]} > 0$ and y is positive recurrent. $\square$

Proposition 2.17. *Let R be a recurrent class, if R is finite, then R is positive recurrent. In particular, if E is finite, then every recurrent state is positive recurrent.*

Proof. Fix $x \in R$, since R is closed we have for every $n > 0$

$$1 = \mathbb{P}_x\left[X_n \in R\right] = \sum_{y \in R} p_{xy}^{(n)}.$$

Hence,

$$1 = \sum_{y \in R} \frac{1}{n} \sum_{k=1}^{n} p_{xy}^{(k)} \to \sum_{y \in R} \frac{1}{\mathbb{E}_y\left[H_y\right]}.$$

Thus, there must be a $y \in R$ such that $\mathbb{E}_y\left[H_y\right] < \infty$, implying that the entire class is positive recurrent. $\square$

2.5 Stationary Distributions for Irreducible Chains

Theorem 2.18. *Assume that p is irreducible.*

- *If the chain is transient or null recurrent, then there is no stationary distribution;*
- *if the chain is positive recurrent, then there exists a unique stationary distribution given by*

$$\boxed{\pi(x) = \frac{1}{\mathbb{E}_x[H_x]}}.$$

Proof. We will begin by assuming a stationary distribution π exists. Then for every $x \in E$ we have for all $n > 0$

$$\pi(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{P}_\pi[X_k = x] = \sum_{y\in E} \pi(y)\frac{1}{n}\sum_{k=1}^{n} \mathbb{P}_y[X_k = x] \to \sum_{y\in E} \pi(y)\frac{1}{\mathbb{E}_x[H_x]}.$$

Note that this also shows uniqueness of the stationary distribution.

Now if we assume that the chain is transient or null recurrent, then using Dominated Convergence Theorem, we have that $\pi(x) = \frac{1}{\mathbb{E}_x[H_x]} = 0$. This is a contradiction to $\sum_{x\in E} \pi(x) = 1$, therefore no stationary distribution can exist.

If, instead, we assume that the chain is positive recurrent, we have the same formula as before for $\pi(x)$ as the only possible candidate for the stationary distribution. So if we can show that π indeed defines a stationary distribution (unlike in the previous case), then we will be done. First, we fix $x > 0$ and find the inequality for all $y \in E$

$$\begin{aligned}
\frac{1}{\mathbb{E}_y[H_y]} &= \lim_{n\to\infty} \frac{1}{n}\sum_{j=k}^{n} p_{yy}^{(j)} \\
&= \lim_{n\to\infty} \sum_{x\in E}\left(\frac{1}{n}\sum_{j=k}^{n} p_{yx}^{(j-k)}\right) p_{xy}^{(k)} \\
&\overset{\text{(Fatou)}}{\geq} \sum_{x\in E} \liminf_{n\to\infty}\left(\frac{1}{n}\sum_{j=k}^{n} p_{yx}^{(j-k)}\right) p_{xy}^{(k)} \\
&= \sum_{x\in E} \frac{1}{\mathbb{E}_x[H_x]} p_{xy}^{(k)}.
\end{aligned}$$

Analogously for fixed x

$$1 = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}_x[X_j \in E] = \lim_{n\to\infty} \sum_{y\in E}\frac{1}{n}\sum_{j=1}^{n} \mathbb{P}_x[X_j = y] \overset{\text{(Fatou)}}{\geq} \sum_{y\in E} \frac{1}{\mathbb{E}_y[H_y]}.$$

So we would like to prove that these two inequalities are actually inequalities. First we sum the first inequality over y and get

$$\sum_{y\in E}\frac{1}{\mathbb{E}_y[H_y]} \geq \sum_{y\in E}\left(\sum_{x\in E}\frac{1}{\mathbb{E}_x[H_x]}p_{xy}^{(k)}\right) = \sum_{x\in E}\frac{1}{\mathbb{E}_x[H_x]}.$$

Thus the inequality must be an equality. Also note that if we can show that π is a distribution, this also shows that it is stationary, we have for every $k > 0$ and for all $y \in E$.

$$\frac{1}{\mathbb{E}_y[H_y]} = \sum_{x\in E}\frac{1}{\mathbb{E}_x[H_x]}p_{xy}^{(k)}.$$

We can use this to show that the second inequality is actually an equality. Fix $y \in E$ and note that $\frac{1}{\mathbb{E}_y[H_y]} > 0$ by positive recurrence. We have

$$\begin{aligned}\frac{1}{\mathbb{E}_y[H_y]} &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\left(\sum_{x\in E}\frac{1}{\mathbb{E}_x[H_x]}p_{xy}^{(k)}\right)\\ &= \lim_{n\to\infty}\sum_{x\in E}\frac{1}{\mathbb{E}_x[H_x]}\left(\frac{1}{n}\sum_{k=1}^{n}p_{xy}^{(k)}\right)\\ &\overset{\text{(DCT)}}{=} \sum_{x\in E}\frac{1}{\mathbb{E}_x[H_x]}\frac{1}{\mathbb{E}_y[H_y]}.\end{aligned}$$

Hence, $\pi(x) = \frac{1}{\mathbb{E}_x[H_x]}$ defines a distribution, which is stationary. □

TODO: Applications

2.6 Periodicity

Definition 2.5. Let $x \in E$. The period of x is defined by

$$\boxed{d_x = \gcd\{n > 0 : p_{xx}^{(n)} > 0\}}.$$

By convention $\gcd(\emptyset) = \infty$.

Proposition 2.19. *Let x, y be arbitrary elements of E then $x \leftrightarrow y$ implies that $d_x = d_y$.*

Proof. Let $x \neq y$. We prove that $d_y | d_x$. First let us fix $k, l \geq 0$ such that $p_{yx}^{(k)}, p_{xy}^{(l)} > 0$. Since $p_{yy}^{(k+l)} \geq p_{yx}^{(k)}p_{xy}^{(l)} > 0$ we have that $d_y | k + l$. Now we show that d_y is a common divisor of $\{n > 0 : p_{xx}^{(n)} > 0\}$, this will imply our claim. For every $n > 0$ satisfying $p_{xx}^{(n)} > 0$, we have

$$p_{yy}^{(k+l+n)} \geq p_{yx}^{(k)}p_{xx}^{(n)}p_{xy}^{(l)} > 0,$$

hence $d_y | k + l + n$. Since $d_y | k + l$, we also have $d_y | n$. □

Consequence If p is irreducible we have for arbitrary $x, y \in E$ that $d_x = d_y$.

Definition 2.6. We say that the chain p is aperiodic if for every $x \in E$

$$\boxed{d_x = 1}.$$

Proposition 2.20. *Let x be in E. We have $d_x = 1$ if and only if there is an $n_0 \geq 1$ such that for every $n \geq n_0$ we have that $p_{xx}^{(n)} > 0$.*

Proof. This follows from a lemma out of Number Theory, see the Appendix for more details. □

TODO: Proof, possibly include number theory lemma in appendix, in my opinion the idea can be explained in a hand wavy fashion intuitively and the algebra distracts from the concept here.

2.7 Product Chain

TODO: This needs some work with how we phrase it

Goal Define two Markov Chains: X_n a MC(μ, P) and $\tilde{X}_n$ a MC(ν, P) on the same probability space such that $X_n = \tilde{X}_n$ for n large.

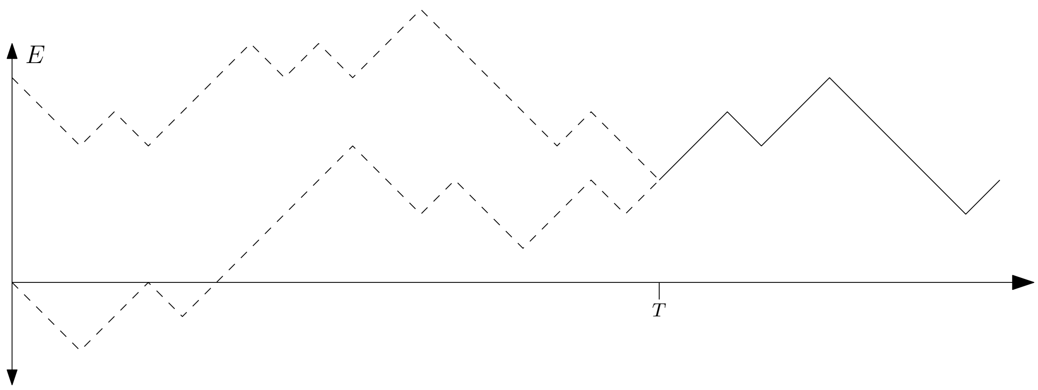

Figure 2.3: A coupling of two simple random walks started from 6 and 0

To achieve that we first consider two independent chains X and Y. We then show that the chains meet almost surely (using some assumptions on p) at some random time T. Then we ask that the chains follow the same trajectory for $t > T$. In order to introduce a suitable probability space, we consider the product chain.

Definition 2.7 (Product Chain)**.** Define for every $\omega = (x, y),\ \omega' = (x', y') \in E^2$

$$\boxed{\overline{p_{\omega,\omega'}} = p_{xx'}p_{yy'}}.$$

Remark 2.21. To see that $\overline{p}$ is a transition probability, calculate

$$\sum_{\omega' \in E} \overline{p}_{\omega\omega'} = \sum_{x',y' \in E} p_{xx'} p_{yy'} = 1.$$

Notation Consider:

- $(\Omega, F, (P_\omega)_{\omega \in E^2})$ Probability Spaces,
- $(W_n)_{n \geq 0} = ((X_n, Y_n))_{n \geq 0}$ a random variable on $(\Omega, \mathcal{F})$ such that for all $\omega \in E^2$, W_n is a $MC(\delta_\omega, \overline{P})$ under $\mathbb{P}_\omega$.

Remark 2.22. If μ, ν are distributions on E, then $\mu \otimes \nu$ is a distribution on E^2.

$$\boxed{P_{\mu \otimes \nu} = \sum_{(x,y) \in E^2} \mu(x) \nu(y) P_{(x,y)}}.$$

Proposition 2.23. *Let μ, ν be distributions on E. Under $P_{\mu \otimes \nu}$:*

- *$(X_n)_{n \geq 0}$ is a $MC(\mu, P)$;*
- *$(Y_n)_{n \geq 0}$ is a $MC(\nu, P)$.*

Proof. For every $k \geq 0$ and $x_0, \ldots, x_k, y_0, \ldots, y_k \in E$ we have

$$\begin{aligned}
&\mathbb{P}_{\mu \otimes \nu} \left[X_0 = x_0, \ldots, X_k = x_k, Y_0 = y_0, \ldots Y_k = y_k \right] \\
&\quad = \mathbb{P}_{\mu \otimes \nu} \left[W_0 = (x_0, y_0), \ldots, W_k = (x_k, y_k) \right] \\
&\quad = \mu(x_0) p_{x_0 x_1} \cdots p_{x_{k-1} x_k} \nu(y_0) p_{y_0 y_1} \cdots p_{y_{k-1} y_k}.
\end{aligned}$$

Summing over all possible $y_0, \ldots, y_k$ in E, implies that $(X_n)_n$ is a MC(μ, P), and equivalently that $(Y_n)_n$ is a MC(ν, P).

Now to show independence, we need to show that for all measurable sets $A, B \subset E^{\mathbb{N}}$

$$\mathbb{P}_{\mu \otimes \nu} \left[X \in A, Y \in B \right] = \mathbb{P}_{\mu \otimes \nu} \left[X \in A \right] \mathbb{P}_{\mu \otimes \nu} \left[Y \in B \right].$$

Our calculation from before shows that this equality holds for all sets of the form $A = \{(x_0, \ldots, x_n)\} \times E^{\mathbb{N}}$, $B = E^{\mathbb{N}} \times \{(y_0, \ldots, y_n)\}$. Therefore, it holds for all cylindrical sets, and thus, by Dynkin's Lemma, for all measurable sets. $\square$

Proposition 2.24. *If p is irreducible, aperiodic, and positive recurrent then $\overline{p}$ is irreducible, aperiodic, and positive recurrent.*

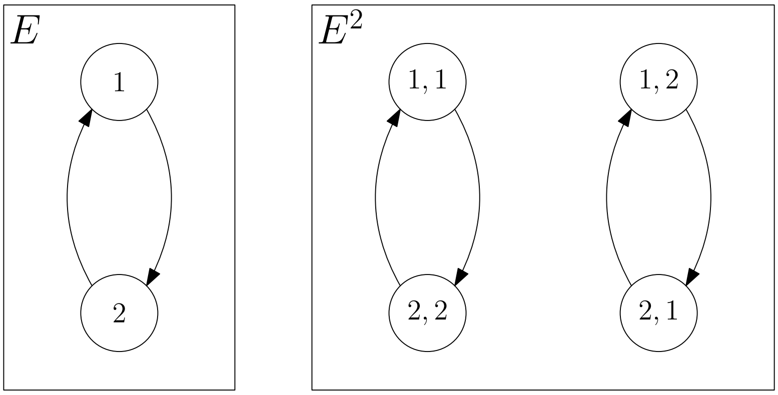

Figure 2.4: Example of an irreducible chain, with a reducible product chain.

Remark 2.25. Aperiodic is important! p irreducible does not imply that $\overline{p}$ irreducible! Take $E = \{1, 2\}$ and $p_{12} = p_{21} = 1$, the product chain here is no longer irreducible.

Proof. Let $w = (x, y)$ and $w' = (x', y') \in E^2$. By irreducibility we can choose $k, l \geq 0$ such that $p^{(k)}_{xx'}, p^{(l)}_{yy'} > 0$. Then for every $n \gg \max(k, l)$ we have

$$\overline{p}^{(n)}_{ww'} = p^{(n)}_{xx'} p^{(n)}_{yy'} \geq p^{(k)}_{xx'} p^{(n-k)}_{x'x'} p^{(l)}_{yy'} p^{(n-l)}_{y'y'} > 0.$$

This holds as the two terms $p^{(n-k)}_{x'x'}$ and $p^{(n-l)}_{y'y'}$ are strictly positive for n large enough.

Since p is irreducible and positive recurrent, it must admit a stationary distribution π. For every $(y, y') \in E^2$ we then have

$$\pi(y)\pi(y') = \sum_{x \in E} \pi(x) p_{xy} \sum_{x' \in E} p_{x'y'} = \sum_{(x,x')\ in E^2} p_{xy} p_{x'y'}.$$

Showing that $\pi \otimes \pi$ is stationary for $\overline{p}$, implying that $\overline{p}$ is positive recurrent. □

Proposition 2.26. *If p is irreducible and aperiodic, then $\overline{p}$ is irreducible and aperiodic.*

Proof. Let $w = (x, y)$ and $w' = (x', y') \in E^2$. By irreducibility we can choose $k, l \geq 0$ such that $p^{(k)}_{xx'}, p^{(l)}_{yy'} > 0$. Then for every $n \gg \max(k, l)$ we have

$$\overline{p}^{(n)}_{ww'} = p^{(n)}_{xx'} p^{(n)}_{yy'} \geq p^{(k)}_{xx'} p^{(n-k)}_{x'x'} p^{(l)}_{yy'} p^{(n-l)}_{y'y'} > 0.$$

This holds as the two terms $p^{(n-k)}_{x'x'}$ and $p^{(n-l)}_{y'y'}$ are strictly positive for n large enough. □

Proposition 2.27. *If p is irreducible, aperiodic, and positive recurrent, then $\overline{p}$ is irreducible, aperiodic, and positive recurrent.*

Proof. We only have to show that the product chain is positive recurrent, as the other properties follow from the previous proposition. Since p is irreducible and positive recurrent, it must admit a stationary distribution π. For every $(y, y') \in E^2$ we then have

$$\pi(y)\pi(y') = \sum_{x\in E} \pi(x)p_{xy} \sum_{x'\in E} p_{x'y'} = \sum_{(x,x')\ inE^2} p_{xy}p_{x'y'}.$$

Showing that $\pi \otimes \pi$ is stationary for $\overline{p}$, implying that $\overline{p}$ is positive recurrent. □

Definition 2.8. $T = \min\{n \geq 0 : X_n = Y_n\}$ is a stopping time.

Remark 2.28. In fact for $A = \{(x, y) \in E^2 : x = y\}$ (which is measurable) $T = H_A$, so T is a stopping time.

Proposition 2.29. *For μ, ν distributions on E, $n \geq 0$:*

$$\boxed{\sum_{x\in E} |\mathbb{P}_\mu [X_n = x] - \mathbb{P}_\nu [Y_n = x] | \leq 2\mathbb{P}_{\mu\otimes\nu} [T > n]}.$$

Proof. We consider the product Markov Chain $W_n = (X_n, Y_n)$ under $\mathbb{P}_{\mu\otimes\nu}$. We then define, for every n

$$\tilde{X}_n = \begin{cases} Y_n & \text{for } n < T \\ X_n & \text{for } n \geq T \end{cases}.$$

We now show that $(\tilde{X}_n)$ is a MC(ν, P) under $\mathbb{P} = \mathbb{P}_{\mu\otimes\nu}$. Let $n \geq 0$ and $x_0, \dots, x_n \in E$. Now we distinguish between possible values for T and find

$$\mathbb{P}\left[\tilde{X}_0 = x_0, \dots, \tilde{X}_n = x_n\right] = \sum_{k\in\mathbb{N}\cup\{\infty\}} \mathbb{P}\left[\tilde{X}_0 = x_0, \dots, \tilde{X}_n, T = k\right].$$

If $k > n$, the summand is equal to

$$\nu(x_0)p_{x_0x_1} \cdots p_{x_{n-1}x_n} \cdot \mathbb{P}[T = k|Y_0 = x_0, \dots, Y_n = x_n].$$

If $k \leq n$, the summand is equal to

$$\begin{aligned}
&\mathbb{P}\left[\underbrace{Y_0 = x_0, \dots, Y_k = x_k, T = k}_{\in\mathcal{F}_T}, X_{T+1} = x_{k+1}, \dots X_{T+n-k} = x_n\right] \\
&\overset{\text{(SMP)}}{=} \mathbb{P}[Y_0 = x_0, \dots Y_k = x_k, T = k]\, \mathbb{P}_{(x_k,x_k)}[X_1 = x_{k+1}, \dots, X_{n-k} = x_n] \\
&= \nu(x_0)p_{x_0x_1} \cdots p_{x_{k-1}x_k}\mathbb{P}[T = k|Y_0 = y_0, \dots, Y_k = x_k]\, p_{x_kx_{k+1}} \cdots p_{x_{n-1}x_n} \\
&= \nu(x_0)p_{x_0x_1} \cdots p_{x_{n-1}x_n}\mathbb{P}[T = k|Y_0 = x_0, \dots, Y_n = x_n].
\end{aligned}$$

To justify the last equality, we used the independence between (X_n) and (Y_n) to write

$$\begin{aligned}\mathbb{P}\left[T=k|Y_0=x_0,\dots,Y_k=x_k\right] &= \mathbb{P}\left[\forall i<k\ X_i\neq x_i, X_k=x_k\right]\\ &= \mathbb{P}\left[T=k|Y_0=x_0,\dots,Y_n=x_n\right].\end{aligned}$$

Finally using that $\sum_{k\in\mathbb{N}\cup\{\infty\}}\mathbb{P}\left[T=k|Y_0=x_0,\dots,Y_n=x_n\right]=1$ we obtain

$$\mathbb{P}\left[\tilde{X}_0=x_0,\dots,\tilde{X}_n=x_n\right]=\nu(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n}.$$

Using the coupling between X and $\tilde{X}$ to conclude that for every $n\geq 0$

$$\begin{aligned}\sum_{x\in E}\left|\mathbb{P}_\mu\left[X_n=x\right]-\mathbb{P}_\nu\left[X_n=x\right]\right| &= \sum_{x\in E}\left|\mathbb{P}\left[X_n=x\right]-\mathbb{P}\left[\tilde{X}_n=x\right]\right|\\ &= \sum_{x\in E}\Big|\mathbb{P}\left[X_n=x,T\leq n\right]+\mathbb{P}\left[X_n=x,T>n\right]\\ &\qquad -\mathbb{P}\left[\tilde{X}_n=x,T\leq n\right]+\mathbb{P}\left[\tilde{X}_n=x,T>n\right]\Big|\\ &\leq \sum_{x\in E}\mathbb{P}\left[X_n=x,T>n\right]+\mathbb{P}\left[\tilde{X}_n=x,T>n\right]\\ &= 2\mathbb{P}\left[T>n\right]\end{aligned}$$

□

// I would separate the proof into a lemma and then proof of the proposition//

Lemma 2.30. *$\tilde{X}_n=Y_n\mathbb{1}_{\{T<n\}}+X_n\mathbb{1}_{\{T\geq n\}}$ is a $MC(\nu,P)$.*

Proof. We consider the product Markov Chain $W_n=(X_n,Y_n)$ under $\mathbb{P}_{\mu\otimes\nu}$. We then define, for every n

$$\tilde{X}_n=\begin{cases}Y_n & \text{for } n<T\\ X_n & \text{for } n\geq T\end{cases}.$$

We now show that $(\tilde{X}_n)$ is a MC(ν,P) under $\mathbb{P}=\mathbb{P}_{\mu\otimes\nu}$. Let $n\geq 0$ and $x_0,\dots,x_n\in E$. Now we distinguish between possible values for T and find

$$\mathbb{P}\left[\tilde{X}_0=x_0,\dots,\tilde{X}_n=x_n\right]=\sum_{k\in\mathbb{N}\cup\{\infty\}}\mathbb{P}\left[\tilde{X}_0=x_0,\dots,\tilde{X}_n,T=k\right].$$

If $k > n$, the summand is equal to

$$\nu(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n}\cdot\mathbb{P}\left[T = k|Y_0 = x_0,\dots,Y_n = x_n\right].$$

If $k \le n$, the summand is equal to

$$\begin{aligned}
&\mathbb{P}\left[\underbrace{Y_0 = x_0,\dots,Y_k = x_k, T = k}_{\in\mathcal{F}_T}, X_{T+1} = x_{k+1},\dots X_{T+n-k} = x_n\right]\\
&\overset{\text{(SMP)}}{=} \mathbb{P}\left[Y_0 = x_0,\dots Y_k = x_k, T = k\right]\mathbb{P}_{(x_k,x_k)}\left[X_1 = x_{k+1},\dots,X_{n-k} = x_n\right]\\
&= \nu(x_0)p_{x_0x_1}\cdots p_{x_{k-1}x_k}\mathbb{P}\left[T = k|Y_0 = y_0,\dots,Y_k = x_k\right]p_{x_kx_{k+1}}\cdots p_{x_{n-1}x_n}\\
&= \nu(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n}\mathbb{P}\left[T = k|Y_0 = x_0,\dots,Y_n = x_n\right].
\end{aligned}$$

To justify the last equality, we used the independence between (X_n) and (Y_n) to write

$$\begin{aligned}
\mathbb{P}\left[T = k|Y_0 = x_0,\dots,Y_k = x_k\right] &= \mathbb{P}\left[\forall i < k\ X_i \neq x_i, X_k = x_k\right]\\
&= \mathbb{P}\left[T = k|Y_0 = x_0,\dots,Y_n = x_n\right].
\end{aligned}$$

Finally using that $\sum_{k\in\mathbb{N}\cup\{\infty\}}\mathbb{P}\left[T = k|Y_0 = x_0,\dots,Y_n = x_n\right] = 1$ we obtain

$$\mathbb{P}\left[\tilde{X}_0 = x_0,\dots,\tilde{X}_n = x_n\right] = \nu(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n}.$$

□

Proof (Proposition). We can now take advantage of the coupling between X and $\tilde{X}$ to conclude that for every $n \ge 0$

$$\begin{aligned}
\sum_{x\in E}\left|\mathbb{P}_\mu\left[X_n = x\right] - \mathbb{P}_\nu\left[X_n = x\right]\right| &= \sum_{x\in E}\left|\mathbb{P}\left[X_n = x\right] - \mathbb{P}\left[\tilde{X}_n = x\right]\right|\\
&= \sum_{x\in E}\Big|\mathbb{P}\left[X_n = x, T \le n\right] + \mathbb{P}\left[X_n = x, T > n\right]\\
&\qquad - \mathbb{P}\left[\tilde{X}_n = x, T \le n\right] + \mathbb{P}\left[\tilde{X}_n = x, T > n\right]\Big|\\
&\le \sum_{x\in E}\mathbb{P}\left[X_n = x, T > n\right] + \mathbb{P}\left[\tilde{X}_n = x, T > n\right]\\
&= 2\mathbb{P}\left[T > n\right]
\end{aligned}$$

□

2.8 Convergence for Irreducible Aperiodic Chains

Theorem 2.31. *Assume p is irreducible and aperiodic, and admits a stationary distribution π. Then for every distribution μ on E and $x \in E$*

$$\boxed{\lim_{n\to\infty} \mathbb{P}_\mu [X_n = x] = \pi(x)}.$$

Equivalently: Under $\mathbb{P}_\mu : X_n \overset{(law)}{\to} X_\infty$ where $X_\infty \sim \pi$.
Equivalently: For all $f : E \to \mathbb{R}$ bounded: $\lim_{n\to\infty} \mathbb{E}_\mu [f(X_n)] = \int_E f d\pi$.

Proof. Consider the product chain $(X_n, Y_n)_{n\geq 0}$ as before. We know that $\overline{P}$ is irreducible and positive recurrent, furthermore the stopping time $T = \min\{n \geq 0 : X_n = Y_n\}$ is $\mathbb{P}_{\mu\otimes\pi}$-a.s. finite. To check this last claim, simply note that $T \leq H_{(a,a)}$ for any a fixed. Then we have that for every $x \in E$

$$|\mathbb{P}_\mu [X_n = x] - \pi(x)| = |\mathbb{P}_\mu [X_n = x] - \mathbb{P}_\pi [X_n = x]| \leq 2\mathbb{P}_{\mu\otimes\pi} [T > n] \to 0.$$

□

Theorem 2.32. *Assume that P is irreducible, aperiodic, and null recurrent or transient. Then for every distribution μ and every $x \in E$*

$$\boxed{\lim_{n\to\infty} \mathbb{P}_\mu [X_n = x] = 0}.$$

Lemma 2.33. *$\overline{P}$ irreducible and recurrent, then for every μ distribution on E, any $i \geq 0$, and every $x \in E$*

$$\lim_{n\to\infty} |\mathbb{P}_\mu [X_n = x] - \mathbb{P}_\mu [X_{n+i} = x] | = 0$$

Proof. Define the distribution $\mu_i(y) = \mathbb{P}_\mu [X_i = y]$, 'the i-step initial distribution' '$\mu_i = \mu P^i$'. Next, observe that

$$\mathbb{P}_{\mu_i} [X_n = x] = \sum_{y\in E} \mu_i \mathbb{P}_y [X_n = x] \overset{\text{(SiMP)}}{=} \sum_{y\in E} \mathbb{P}_\mu [X_i = y] \mathbb{P}_\mu [X_{n+i} | X_i = y] = \mathbb{P}_\mu [X_{n+i} = x]$$

Now, if we consider the product chain $(X_n, Y_n)_{n\geq 0}$ under $\mathbb{P}_{\mu\otimes\mu_i}$ and define the stopping time $T = \min\{n \geq 0 : X_n = Y_n\}$. Here, we have that $T < \infty$ almost surely as $\overline{P}$ is recurrent. Hence we find that

$$|\mathbb{P}_\mu [X_n = x] - \mathbb{P}_\mu [X_{n+i} = x]| = |\mathbb{P}_\mu [X_n = x] - \mathbb{P}_\mu [X_n = x]| \leq 2\mathbb{P}_{\mu\otimes\mu_i} [T > n] \to 0.$$

□

Proof (Theorem). **Case 1:** Assume $\overline{P}$ transient. Consider the product chain (X_n, Y_n) under $\mathbb{P}_{\mu\otimes\mu}$, since (x,x) is a transient state the last visit time $L = \max\{n \geq 0 : (X_n, Y_n) = (x,x)\}$ is finite $\mathbb{P}_{\mu\otimes\mu}$-a.s. Hence,

Assume $\overline{P}$ transient. If we look at the product chain (X_n, Y_n) under $\mathbb{P}_{\mu\otimes\mu}$, we can see that (x,x) is a transient state. Thus the time of the last visit $L = \max\{n \geq 0 : (X_n, Y_n) = (x,x)\}$ is finite $\mathbb{P}_{\mu\otimes\mu}$-a.s. (if this was not almost sure, then we would have a non-zero probability that (x,x) is revisited infinitely often, thus by the Dichotomy Theorem (x,x) would be recurrent). Hence,

$$\mathbb{P}_\mu\left[X_n = x\right]^2 = \mathbb{P}_{\mu\otimes\mu}\left[X_n = x, Y_n = x\right] \leq \mathbb{P}_{\mu\otimes\mu}\left[L \geq n\right] \to 0.$$

Case 2: Assume $\overline{P}$ is null recurrent, fix $y \in E$. We would like to prove that $p_{yx}^{(n)} \to 0$. To do this fix $\epsilon > 0$ and choose k such that

$$\frac{1}{k+1}\sum_{i\leq k} p_{xx}^{(i)} < \epsilon.$$

We can choose such a k using the density of visits theorem and that (x,x) is null recurrent. Now define the stopping time $H = \min\{j \geq n : X_j = x\}$, 'the first hit time of x after time n'. So for every $n \geq 0$ we have

$$\frac{1}{k+1}\sum_{i=0}^{k}\mathbb{P}_\mu\left[X_{n+1} = x\right] \leq \frac{1}{k+1}\sum_{i=0}^{k}\mathbb{P}_\mu\left[X_{H+i} = x\right] \overset{\text{(SMP)}}{=} \frac{1}{k+1}\sum_{i=0}^{k}\mathbb{P}_x\left[X_i = x\right] \leq \epsilon.$$

The first inequality can be justified by noticing that the probability of the chain hitting x after n and before H is 0. Hence,

$$\begin{aligned}
\mathbb{P}_\mu\left[X_n = x\right] &= \frac{1}{k+1}\sum_{i=0}^{k}\mathbb{P}_\mu\left[X_n = x\right] \\
&\leq \underbrace{\frac{1}{k+1}\sum_{i=0}^{k}\left|\mathbb{P}_\mu\left[X_n = x\right] - \mathbb{P}_\mu\left[X_{n+i} = x\right]\right|}_{\to 0} + \underbrace{\frac{1}{k+1}\sum_{i=0}^{k}\mathbb{P}_\mu\left[X_{n+i} = x\right]}_{\leq\epsilon}.
\end{aligned}$$

Using the lemma ($\overline{P}$ is irreducible and recurrent) we find that

$$\limsup_{n\to\infty}\mathbb{P}_\mu\left[X_n = x\right] \leq \epsilon.$$

□

Conclusion We previously asked the following questions:

- If we fix $x \in E$, will the chain visit x infinitely many times?
- What is the behavior of X_n for n large?

Now we are equipped to answer them using our ideas of recurrence/transience and the theorem for existence (and uniqueness) of stationary distributions for an irreducible chain. We were also found that using coupling we find that if we let the chain evolve for a long time, then the distribution of X_n actually converges to the stationary distribution (where this distribution is 0 everywhere if a stationary distribution does not exist).

Chapter 3

Renewal Processes

Outset We want to model replacement times of a machine. First we wait T_1 until we replace it, then we wait T_2 until replacing the replacement, and so on.
Questions: After time t, how many replacements did we have to make (N_t)? What about the expected number $m(t) = \mathbb{E}[N_t]$? What about the 'excess time', i.e. if we are at time t, how long until the next replacement (E_t, $e(t) = \mathbb{E}[E_t]$)? Or the age of the machine (A_t, $a(t) = \mathbb{E}[A_t]$). //I think we should elaborate on case 1 more, and leave out case 2, these correspond to the what we said in the lecture and not what is specifically in your notes, where you have multiple examples.//

Case 1: $T_1, T_2, \dots$ $Exp(\lambda)$ random variables: $m(t) = t\lambda$, E_t also $Exp(\lambda)$, then $e(t) = \frac{1}{\lambda}$, A_t is also $Exp(\lambda)$.

Case 2: More complicated.

3.1 Definition and First Properties

Framework $(\Omega, \mathcal{F}, \mathbb{P})$ Probability space, $T_1, T_2, \dots$ i.i.d. random variables on $\mathbb{R}_+$ 'inter-arrival times', such that $\mathbb{P}[T_i = 0] < 1$, $\mu = \mathbb{E}[T_1] \in (0, \infty]$. $F(t) = \mathbb{P}[T_1 \leq t]$, $S_n = \sum_{i=1}^n T_i$, $S_0 = 0$ 'renewal times'.

Definition 3.1. The continuous time stochastic process $(N_t)_{t\geq 0}$ defined by:

$$\forall t \geq 0 : N_t = \sum_{k=1}^{\infty} \mathbb{1}_{S_k \leq t}$$

is called the *renewal process with arrival distribution F*.

Remark 3.1. It is important to differentiate between continuous time and continuous stochastic processes. Continuous time stochastic process are defined for every $t \in \mathbb{R}$ and may include jumps, meanwhile continuous stochastic processes follow continuous trajectories (eg. Brownian Motion) and are not the subject of this course.

Example 3.1. (i) $pp(\lambda), \lambda > 0,\ T_i$ a $Exp(\lambda)$ random variable,

(ii) $(T_i)_{i\geq 1}$ i.i.d. $Exp(\lambda)$, $(X_i)_{i\geq 1}$ i.i.d. $Ber(\frac{1}{2})$, $T_i' = X_i T_i$, where (T_i) and (X_i) are independent.

(iii) 'Fat Tailed' $\mathbb{P}\left[T_i \geq t\right] = \frac{1}{\sqrt{1+t}} \mathbb{1}_{t\geq 0}$

Definition 3.2. Let $N = (N_t)_{t\geq 0}$ be a continuous time stochastic process with values in $\mathbb{R}$. We say that N is a *counting process* if the following holds a.s.

(i) $N_0 = 0$ a.s.,

(ii) $t \mapsto N_t$ is non-decreasing, right continuous, with values in $\mathbb{N}$.

Proposition 3.2. *$N = (N_t)_{t\geq 0}$ is a counting process*[1] *with jump times $S_1, S_2, \ldots$ and* $\lim_{t\to\infty} N_t = +\infty$.

//Is there a reason to use $+\infty$ rather than just ∞?. Also we have not defined a counting process at this point in time, because we are doing renewal processes before PP.//

Proof. Since $\mathbb{P}\left[T_i > 0\right] > 0$, there exists $\alpha > 0$ such that $\mathbb{P}\left[T_i \geq \alpha\right] > 0$. Indeed $\mathbb{P}\left[T_1 > 0\right] = \mathbb{P}\left[\bigcup_{\alpha\in\mathbb{Q}_+} \{T_i \geq \alpha\}\right]$. We have

$$\sum_{i>0} \mathbb{P}\left[T_i \geq \alpha\right] = \infty.$$

Therefore, by the Borel-Cantelli lemma, $\mathbb{P}\left[A\right] = 1$. Where

$$A = \{\omega : T_i(\omega) \geq \alpha \text{ for infinitely many } i\}.$$

For every $\omega \in A$, $S_n(\omega) \overset{n\to\infty}{\to} \infty$, and therefore

$$t \mapsto N_t(\omega) = \sum_{n>0} \mathbb{1}_{S_n(\omega)\leq t}$$

is a non-decreasing function with values in $\mathbb{N}$. $\square$

[1] without the condition that $N_0 = 0$ a.s.

Proposition 3.3. *There exists $c > 0$ such that*

$$\forall t \geq 0 \ \mathbb{E}\left[e^{cN_t}\right] \leq e^{\frac{1+t}{c}}$$

Remark 3.4. In particular, for every $t \geq 1$, we have

$$\mathbb{E}\left[e^{c\frac{N_t}{t}}\right] \overset{(Jensen)}{\leq} \mathbb{E}\left[e^{cN_t}\right]^{\frac{1}{t}} \leq e^{\frac{2}{c}}$$

and

$$\mathbb{E}\left[\left(\frac{N_t}{t}\right)^d\right] \leq \frac{k!}{c^k} e^{\frac{2}{c}}.$$

Proof. As before, we can pick $\alpha > 0$ such that $\mathbb{P}\left[T_1 \geq \alpha\right] > 0$. For every $i > 0$, define

$$T_i' = \alpha \mathbb{1}_{T_i \geq \alpha}.$$

We have $T_i' \leq T_i$ a.s. and (T_i') are i.i.d. random variables with

$$T_i' = \begin{cases} \alpha & \text{with probability } \beta \\ 0 & \text{with probability } 1-\beta \end{cases}$$

where $\beta = \mathbb{P}\left[T_1 \geq \alpha\right] > 0$. Define $S_n' = \sum_{i=1}^n T_i'$ and the renewal process

$$N_t' = \sum_{n>0} \mathbb{1}_{S_n' \leq t}.$$

As in example this example isn't here yet, we have that

$$N_t' \overset{\text{(law)}}{=} X_0 + \sum_{i=1}^{\lfloor \frac{t}{\alpha} \rfloor} (1 + X_i),$$

where (X_i) are geometric random variables with success parameter β. Therefore, for $c > 0$ such that $(1-\beta)e^c < 1$, we have for all $t \geq 0$

$$\begin{aligned}
\mathbb{E}\left[e^{cN_t'}\right] &= e^{c\lfloor \frac{t}{\alpha} \rfloor} \prod_{i=0}^{\lfloor \frac{t}{\alpha} \rfloor} \mathbb{E}\left[e^{cX_i}\right] \\
&\leq e^{\frac{ct}{\alpha}} \left(\frac{\beta}{1-(1-\beta)e^c}\right)^{1+\frac{t}{\alpha}} \\
&\leq \left[\left(\frac{e^c}{1-(1-\beta)e^c}\right)^{\frac{1}{\alpha}}\right]^{1+t} \\
&\leq e^{\frac{1+t}{c}}
\end{aligned}$$

for c small enough (independent of t). To get the second inequality, choose $\alpha \leq 1$ and use that $1+\frac{t}{\alpha} \leq \frac{1+t}{\alpha}$ with this condition. This choice of α is justified as we are finding an upper bound, and α only appears in the denominators, thus we can choose it as small as needed. $\square$

Theorem 3.5 (Law of Large Numbers). *Write $\mu = \mathbb{E}[T_1]$, then we have*

$$\lim_{t\to\infty} \frac{N_t}{t} = \frac{1}{\mu} \ a.s.$$

Remark 3.6. If $\mu = \infty$, then $\lim_{t\to\infty} \frac{N_t}{t} = 0$ a.s.

Proof. // In your notes, you have $1/\mu$ and μ mixed up, I have made the correction here without putting it in blue, as I believe we spoke about this in the exercise class. //

Case 1: $\mu < \infty$. By the strong law of large numbers, we have

$$\lim_{n\to\infty} \frac{S_{n+1}}{n+1} = \lim_{n\to\infty} \frac{S_n}{n+1} = \mu \text{ a.s.}$$

Notice that for every t

$$S_{N_t} \leq t \leq S_{N_t+1}.$$

Therefore,

$$\underbrace{\frac{S_{N_t}}{N_t+1}}_{\to\mu} \leq \frac{t}{N_t+1} < \underbrace{\frac{S_{N_t+1}}{N_t+1}}_{\to\mu}.$$

Where the convergences are almost sure. Therefore $\lim_{t\to\infty} \frac{1+N_t}{t} = \frac{1}{\mu}$ a.s., which implies that $\lim_{t\to\infty} \frac{N_t}{t} = \frac{1}{\mu}$ a.s.

Case 2: $\mu = \infty$. Define $T_i^{(k)} = \min(k, T_i)$ for $k \geq 1$. This way we have $T_i^{(k)} \leq T_i$ a.s. and $T_i^{(k)} \uparrow T_i$ as $k \to \infty$ a.s. Consider the renewal process $N_t^{(k)}$ associated to the times $(T_i^{(k)})_{i\geq 1}$. Since $\mathbb{E}\left[T_i^{(k)}\right] \leq k < \infty$, we can apply case 1 to obtain that

$$\forall k \quad \lim_{t\to\infty} \frac{N_t^{(k)}}{t} = \frac{1}{\mathbb{E}\left[T_1^{(k)}\right]} \text{ a.s.}$$

Since $T_i^{(k)} \leq T_i$ a.s., we have $N_t^{(k)} \geq N_t$ a.s. Hence,

$$\forall k \quad \frac{1}{\mathbb{E}\left[T_1^{(k)}\right]} \geq \limsup_{t\to\infty} \frac{N_t}{t} \text{ a.s.}$$

Now, by monotone convergence, we have

$$\lim_{k\to\infty} \mathbb{E}\left[T_1^{(k)}\right] = \mathbb{E}\left[T_1\right] = \infty,$$

and the two equations above conclude the proof. □

Theorem 3.7 (Central Limit Theorem)**.** *Assume that* $\mathbb{E}\left[T_1^2\right] < \infty$*. Write* $\mu = \mathbb{E}\left[T_1\right]$, $\sigma^2 = Var(T_1)$*. Then, assuming* $\sigma > 0$*, we have*

$$\frac{N_t - \frac{t}{\mu}}{\sigma\sqrt{\frac{t}{\mu^3}}} \xrightarrow[t\to\infty]{\textit{(law)}} \mathcal{N}(0,1)$$

// Proof not in notes, I believe this was an exercise, I could include that.//

Example 3.2 (Renewal Reward Process)*.* Let $(D_i)_{i\geq 1}$ be i.i.d. random variables with $D_i \geq 0$ and $\mathbb{E}\left[D_i\right] < \infty$. Define for every t

$$R_t = \sum_{i\geq 1} D_i \mathbb{1}_{S_i \leq t}.$$

We call this the reward process, and D_i the reward at time S_i. Then for t large

$$\frac{R_t}{t} = \underbrace{\frac{1}{N_t}\sum_{i=1}^{N_t} D_i}_{\overset{\text{(LLN)}}{\to}\mathbb{E}[D_1]} \underbrace{\frac{N_t}{t}}_{\to\frac{1}{\mu}}.$$

Therefore, $\frac{R_t}{t} \to \frac{\mathbb{E}[D_1]}{\mu}$ a.s.

3.2 Renewal Function

Definition 3.3. The renewal function is defined by

$$\boxed{\forall t \geq 0 \quad m(t) = \mathbb{E}\left[N_t\right].}$$

Motivation $m(t) = \mathbb{E}\left[\text{Number of points in the interval } [0,t]\right]$ where a point is a renewal time.

Remark 3.8. $m(t) < \infty$ because N_t has exponential moments (Jensen).

Proposition 3.9. *$m(t)$ is non-decreasing, non-negative, and right continuous.*

Proof. We have $N_{t+s} - N_t \downarrow 0$ as $s \downarrow 0$. Therefore $m(t+s) - m(t) \to 0$ by monotone convergence. The other properties are obvious ~~obvious~~ trivial. □

Exercise Try to draw m(t) for the previous TODO examples.

Remark 3.10. The previous proposition implies that there exists a unique measure ν on $\mathbb{R}_+$ such that

$$\forall t \quad \nu([0,t]) = m(t).$$

$\nu(B) = \mathbb{E}\,[\text{Number of points on the set } B]$ for B measurable.

Notation Let G be a right continuous non-decreasing function on $\mathbb{R}_+$. Write dG for the corresponding Lesbesgue-Stieltjes measure. For $h \in L^1(dG)$ or h measurable and non-negative write

$$\int_{\mathbb{R}_+} hdG = \int_{\mathbb{R}_+} h(x)dG(x)$$

for the corresponding integral.

Definition 3.4 (Convolution Operator)**.** Let G be a right continuous non-decreasing function on $\mathbb{R}_+$. Let $h : \mathbb{R}_+ \to \mathbb{R}$ be such that for all $t \geq 0$ $\int_0^t |h(t-s)|dG(s) < \infty$ or h measurable non-negative. For every $t \geq 0$, define

$$(h * G)(t) = \int_0^t h(t-s)dG(s).$$

Remark 3.11. If X, Y are two independent random variables on $\mathbb{R}_+$ with distribution functions F_X, F_Y respectively, then

$$F_{X+Y} = F_X * F_Y.$$

The proof is left as an exercise.

Notation $F^{*k} = \underbrace{F * \ldots * F}_{k \text{ times}}$. This is useful for the distribution function of $S_n = T_1 + \ldots + T_n$.

Proposition 3.12. *For every* $t \geq 0$

$$m(t) = \sum_{k\geq 1} F^{*k}(t).$$

Proof.

$$m(t) = \mathbb{E}\left[\sum_{k\geq 1} \mathbb{1}_{S_k \leq t}\right] = \sum_{k\geq 1} \mathbb{P}\left[S_k \leq t\right] = \sum_{k\geq 1} F^{*k}(t).$$

$\square$

Theorem 3.13 (Elementary Renewal Theorem)**.**

$$\boxed{\lim_{t\to\infty}\frac{m(t)}{t}=\frac{1}{\mu}.}$$

Proof. We already have $\lim_{t\to\infty}\frac{N_t}{t}=\frac{1}{\mu}$ a.s. Furthermore, we have seen that $\sup_{t\geq 1}\mathbb{E}\left[\left(\frac{N_t}{t}\right)^2\right]<\infty$. Hence $\frac{N_t}{t}$ is uniformly integrable and

$$\lim_{t\to\infty}\frac{m(t)}{t}=\lim_{t\to\infty}\mathbb{E}\left[\frac{N_t}{t}\right]=\mathbb{E}\left[\lim_{t\to\infty}\frac{N_t}{t}\right]=\frac{1}{\mu}.$$

□

3.3 Renewal with Delay

In general, the number of arrivals in $[a,a+t]$ depends on a, 'no stationary increments'. **Idea** Introduce a delay. The time $\overline{T}_1$ is chosen with a different distribution from $\overline{T}_2,\overline{T}_3,\ldots$

We consider $(\overline{T}_i)_{i>0}$ independent random variables on $\mathbb{R}_+$ with

(i) $\overline{T}_1\sim dG$,

(ii) $\overline{T}_i\sim dF$ for $i>1$.

TODO: Figure

Example 3.3 (Shifted Renewal with Delay). $T_1,T_2,\ldots$ i.i.d. as in the previous sections. Fix $t>0$, then define

$$\overline{T}_1=S_{n_{t+1}}-t,\ \overline{T}_i=S_{N_t+i}-S_{N_t+i-1}, i>1.$$

Definition 3.5.

$$\overline{S}_i=\overline{T}_1+\ldots+\overline{T}_i,\ \text{for } i>0$$
$$\overline{N}_t=\sum_{i>0}\mathbb{1}_{\overline{S}_i\leq t},\ \text{for } t\geq 0.$$

$(\overline{N}_t)_{t\geq 0}$ is called a renewal process with distribution function F and delay function G.

Goal Compute $\overline{m}(t)$, the renewal function associated to $\overline{N}_t$ and find G such that $\overline{m}$ is linear. **Notation** $\overline{m}(t)=\mathbb{E}\left[\overline{N}_t\right]$. // Should this be in display?//

As in the previous section, we can prove the following

Proposition 3.14.

$$\forall t \geq 0 \quad \overline{m}(t) = \sum_{i \geq 0} G * F^{*i}(t).$$

The Laplace transform behaves 'nicely' with the convolution operator, and the Laplace transform of $\overline{m}$ can be easily computed.

3.4 Intermezzo: Laplace Transform

// This is not a labeled section in your notes, but is included //
Notation $\mathcal{M} = \{f : \mathbb{R}_+ \to \mathbb{R}_+, \text{ right continuous, non-decreasing}\}$, 'non-negative measures on $\mathbb{R}_+$'. $\nu((a,b]) = f(b) - f(a)$ the corresponding Lebesgue-Stieltjes measure.

Definition 3.6. Let $f \in \mathcal{M}$. For every $s \geq 0$, define

$$\boxed{Lf(s) = \int_0^\infty e^{-sx} df(x).}$$

Remark 3.15. If $f = F_Y$ distribution function of a non-negative random variable Y, then $Lf(s) = \mathbb{E}\left[e^{-sY}\right]$.

Proposition 3.16. *For every $f, g \in \mathcal{M}$, we have*

$$L_{f*g} = L_f \cdot L_g.$$

Remark 3.17. If X, Y are two independent random variables on $\mathbb{R}_+$ then

$$\mathbb{E}\left[e^{-s(X+Y)}\right] = \mathbb{E}\left[e^{-sX}\right]\mathbb{E}\left[e^{-sY}\right].$$

Remark 3.18. If $f, g \in \mathcal{M}$, then $f * g$ is well defined and $f * g \in M$.

Proof. For every $h \geq 0$ measurable, we have

$$\int h d(f * g) = \int \int h(x+y) df(x) dg(y).$$

In particular, this is true for $h = \mathbb{1}_{(a,b]}$, because

$$\begin{aligned}
(f * g)(b) - (f * g)(a) &= \int f(b-y) - f(a-y) dg(y) \\
&= \int \int \mathbb{1}_{x \in (a-y, b-y]} df(x) dg(y) \\
&= \int \int \mathbb{1}_{(a,b]}(x+y) df(x) dg(y).
\end{aligned}$$

Thus, it is also true for general $h \geq 0$ by approximation. In particular for $h(x) = e^{-sx}$, we have

$$\int e^{-sz} d(f * g)(z) = \int \int e^{-s(x+y)} df(x) dg(y) \overset{\text{(Fubini)}}{=} \int e^{-sx} df(x) \int e^{-sy} dg(y).$$

□

Corollary 3.19.

$$L_{\overline{m}} = \frac{L_G}{1 - L_F}.$$

Proof. By monotone convergence, we have

$$L_{\overline{m}} = \sum_{i \geq 0} L_{G * F^{*i}}.$$

By induction $L_{G * F^{*i}} = L_G \cdot L_F^i$. Hence for all $t > 0$ (because $L_F(t) < 1$)

$$L_{\overline{m}} = \sum_{i \geq 0} L_G(t) L_F(t)^i = L_G(t) \sum_{i \geq 0} L_F(t)^i = \frac{L_G(t)}{1 - L_F(t)}.$$

This equality extends to $t = 0$ since both terms are infinite. □

Definition 3.7. Consider the delay function defined by

$$\forall t \geq 0 \quad G(t) = \frac{1}{\mu} \int_0^t (1 - F(x)) dx.$$

Remark 3.20.

$$\int_0^\infty (1 - F(x)) dx = \int_0^\infty \mathbb{P}\left[T_1 > x\right] dx \overset{\text{(Fubini)}}{=} \mathbb{E}\left[\int_0^\infty \mathbb{1}_{T_i > x} dx\right] = \mathbb{E}\left[T_1\right] = \mu.$$

Theorem 3.21. *Assume $\mu < \infty$. For the renewal process with delay function G, we have*

$$\boxed{\overline{m}(t) = \frac{t}{\mu}}$$

for every $t \geq 0$.

'The process is stationary.'

Lemma 3.22. *Let $m_1, m_2 \in \mathcal{M}$ as assume that*

$$\forall t > 0 \quad L_{m_1}(t) = L_{m_2}(t) < \infty,$$

Then $m_1 = m_2$.

Proof (Lemma). Admitted. □

Proof (Theorem). For $s > 0$, notice that for every h measurable and bounded

$$\int_0^\infty h(x)dG(x) = \int_0^\infty h(x)(1-F(x))\frac{dx}{\mu}.$$

This is done as usual, first showing for $h = \mathbb{1}_{[a,b]}$ and then concluding by approximation. In particular, for every $s > 0$

$$\begin{aligned} L_G(s) &= \int_0^\infty e^{-sx}dG(x) = \int_0^\infty e^{-sx}\underbrace{(1-F(x))}_{=\mathbb{P}\left[\overline{T}_2>x\right]}\frac{dx}{\mu} \\ &= \frac{1}{\mu s}\int_0^\infty se^{-sx}\mathbb{P}\left[\overline{T}_2 > x\right]dx \\ &\overset{\text{(Fubini)}}{=} \frac{1}{s\mu}\mathbb{E}\left[\int_0^\infty se^{-sx}\mathbb{1}_{\overline{T}_2>x}dx\right] \\ &= \frac{1}{s\mu}\mathbb{E}\left[\int_0^{\overline{T}_2} se^{-sx}dx\right] \\ &= \frac{1}{s\mu}\mathbb{E}\left[1-e^{-s\overline{T}_2}\right] = \frac{1-L_F(s)}{s\mu}. \end{aligned}$$

Therefore

$$\begin{aligned} L_{\overline{m}}(s) &= \frac{1-L_F(s)}{s\mu}\frac{1}{1-L_F(s)} = \frac{1}{s\mu} \\ &= \frac{1}{\mu}\int_0^\infty e^{-sx}dx = L_{\frac{1}{\mu}I}. \end{aligned}$$

Where I is the identity function. By the lemma, we conclude that for all $t \geq 0$ $\overline{m}(t) = \frac{1}{\mu}t$. □

3.5 Blackwell's Renewal Theorem

Definition 3.8. We say that the law of T_1 is *non-arithmetic* if

$$\forall a > 0 \quad \mathbb{P}\left[T_1 \in a\mathbb{Z}\right] < 1.$$

Definition 3.9. We say the law of T_1 is *arithmetic* if there exists $a > 0$ such that

$$\mathbb{P}\left[T_1 \in a\mathbb{Z}\right] = 1.$$

It is *non-arithmetic* if this probability is < 1.

Theorem 3.23 (Blackwell's Renewal Theorem)**.** *Assume that the law of T_1 is non-arithmetic, then for all $h \geq 0$*

$$\boxed{\lim_{t\to\infty} m(t+h) - m(t) = \frac{h}{\mu}.}$$

// This is from your lecture, but not in the notes.//

Remark 3.24.

$$\frac{m(t)}{t} \approx \frac{m(\lfloor t \rfloor)}{\lfloor t \rfloor} = \frac{1}{\lfloor t \rfloor}\sum_{k=1}^{\lfloor t \rfloor} m(k) - m(k-1) \overset{\text{(Blackwell)}}{\to} \frac{1}{\mu}.$$

"Blackwell is stronger than elementary renewal."

Proof (Sketch). Consider $\overline{T}_i,\ i > 0$ independent and independent of $(T_i)_{i>0}$ where $\overline{T}_1$ has law dG, and $\overline{T}_i$ has law dF for $i > 1$. Then the renewal function associated to these inter-arrival times is

$$\forall t \quad \overline{m}(t) = \frac{1}{\mu}t.$$

We call this 'stationary'. Write $S_k = \sum_{i\leq k} T_i$ and $\overline{S}_k = \sum_{i\leq k} \overline{T}_i$.

Now claim that for $\epsilon > 0$, a.s. there exists $K > 1$ (random) such that

$$|S_k - \overline{S}_k| \leq \epsilon.$$

We admit this.

//TODO: Figure //

Consider

$$\tilde{T}_i = \begin{cases} \overline{T}_i, & i \leq k \\ T_i, & i > K. \end{cases}$$

Then the renewal process associated to $(\tilde{(T)}_i)_{i>0}$ is a delayed process with renewal function

$$\tilde{m}(t) = \frac{t}{\mu}.$$

Furthermore for t large ($t > K$), we have

$$N_{t+h} - N_t \approx \tilde{N}_{t+h} - \tilde{N}_t.$$

Therefore

$$m(t+h) - m(t) = \mathbb{E}\left[N_{t+h} - N_t\right] \approx \mathbb{E}\left[\tilde{N}_{t+h} - \tilde{N}_t\right] = \frac{h}{\mu}.$$

□

3.6 Renewal Equation

Definition 3.10. Let $h : \mathbb{R}_+ \to \mathbb{R}$ be measurable locally bounded, $g : \mathbb{R}_+ \to \mathbb{R}$ such that for all $t \geq 0$ $\int_0^t |g(t-s)|dF(s) < \infty$. We say that g is a solution of the (h, F) renewal equation if

$$\boxed{\forall t \geq 0 \quad g(t) = h(t) + \int_0^t g(t-s)dF(s),}$$

ie. $g = h + g * F$.

Proposition 3.25. *m is a solution of the (F, F) renewal equation, ie. $m = F + m * F$.*

Proof 1.

$$m = \sum_{i>0} F^{*i} = F + \sum_{i>1} F^{*(i-1)} * F = F + \underbrace{\left(\sum_{i>1} F^{*(i-1)}\right)}_{m} *F.$$

□

Proof 2. For $t \geq 0$, we have

$$m(t) = \mathbb{E}\left[\sum_{k>0} \mathbb{1}_{T_1+\ldots+T_k \leq t}\right] = \mathbb{P}\left[T_1 \leq t\right] + \underbrace{\mathbb{E}\left[\sum_{k>1} \mathbb{1}_{T_1+\ldots+T_k \leq t}\right]}_{(\star)}$$

$$\begin{aligned}(\star) &\overset{\text{(Fubini)}}{=} \sum_{k>1} \mathbb{E}\left[\mathbb{1}_{T_1+\ldots+T_k \leq t}\right] \overset{\text{(Indep.)}}{=} \sum_{k>1} \int_0^t \mathbb{E}\left[\mathbb{1}_{s+T_2+\ldots+T_k \leq t}\right] dF(s) \\ &= \int_0^t m(t-s)dF(s).\end{aligned}$$

□

Example 3.4 (Generalization). Let $E = \{(s_i)_{i>0} : s_1 \leq s_2 \leq \ldots, \ s_i \to \infty\} \subset \mathbb{R}^{\mathbb{N}}$, where $\mathbb{R}^{\mathbb{N}}$ is equipped with the product σ-algebra. Let $\Phi : E \to \mathbb{R}$ be measurable such that

$$\forall t \geq 0, \ i = 1, 2 \quad \mathbb{E}\left[|\Phi(S_i - t, S_{i+1} - t, \ldots)|\right] < \infty.$$

Define $\phi(t) = \mathbb{E}\left[\Phi(S_1 - t, S_2 - t, \ldots)\right]$ for all $t \geq 0$.

// TODO: Possible figure? //

Proposition 3.26. *ϕ is a solution of the (h, F) renewal equation, where for all $t \geq 0$*

$$h(t) = \mathbb{E}\left[\Phi(S_1 - t, S_2 - t, \ldots)\right] - \mathbb{E}\left[\Phi(S_2 - t, S_3 - t, \ldots)\mathbb{1}_{T_1 \leq t}\right],$$

ie. for all $t \geq 0$ $\phi(t) = h(t) + \int_0^t h(t-s)dF(s)$.

Proof.

$$\begin{aligned}
\mathbb{E}\left[\Phi(S_1 - t, S_2 - t, \ldots)\right] &= h(t) + \mathbb{E}\left[\Phi(S_2 - t, S_3 - t, \ldots)\mathbb{1}_{T_1 \leq t}\right] \\
&= h(t) + \mathbb{E}\left[\Phi(T_1 + T_2 - t, T_1 + T)2 + T_3 - t, \ldots)\mathbb{1}_{T_1 \leq t}\right] \\
&= h(t) + \int_0^t \mathbb{E}\left[\Phi(s + T_2 - t, s + T_2 + T_3 - t, \ldots)\right] dF(s) \\
&= h(t) + \int_0^t \phi(t-s)dF(s).
\end{aligned}$$

□

Example 3.5 (Application 1). m is a solution of the (F, F) renewal equation. To use this, we have to bring it into the proper form. $N_t = \Phi(S_1 - t, S_2 - t, \ldots)$ where $\Phi(s_1, s_2, \ldots) = \sum_i \mathbb{1}_{s_i \leq 0}$. Hence, $m(t) = \mathbb{E}\left[N_t\right]$ is the solution of the (h, F) renewal equation with

$$h(t) = \mathbb{E}\underbrace{\left[\Phi(S_1 - t, S_2 - t, \ldots) - \mathbb{1}_{T_1 \leq t}\Phi(S_2 - t, S_3 - t, \ldots)\right]}_{=\begin{cases} 0, & T_1 > t \\ 1, & T_1 \leq t \end{cases}}$$

ie. $h(t) = \mathbb{P}\left[T_1 \leq t\right]$.

Example 3.6 (Application 2). For $t \geq 0$, define $E_t = S_{N_{t+1}} - t$, the time left to wait until next renewal. Define for $x \geq 0$, $e_x(t) = \mathbb{P}\left[E_t \leq x\right]$ for all $t \geq 0$.

First we will find a solution without the proposition. We can separate e_x into two parts, one for the probability if there has already been a renewal before time t, and one if that hasn't occurred $e_x(t) = \mathbb{P}\left[T_1 > t, E_t \leq x\right] + \mathbb{P}\left[T_1 \leq t, E_t \leq x\right] = A + B$.

$$A = \mathbb{P}\left[T_1 > t, T_1 \leq t + x\right] = F(t+x) - F(t).$$

Observe that E_t is measurable with respect to $(T_1, T_2, \ldots)$. Write $\phi_t(T_1, T_2, \ldots) = E_t$.

$$\begin{aligned}
\mathbb{P}\left[T_1 \leq t, E_t \leq x\right] =& \mathbb{P}\left[T_1 \leq t, \phi_t(T_1, T_2, \ldots) \leq x\right] \\
=& \int_0^t \mathbb{P}\left[\phi_t(s, T_2, \ldots) \leq x\right] dF(s) = \int_0^t \mathbb{P}\left[E_{t-s} \leq x\right] dF(s) \\
=& \int_0^t e_x(t-s)dF(s) = (e_x * F)(t)
\end{aligned}$$

Thus $e_x(t) = h_x(t) + (e_x * F)(t)$ with $h_x(t) = F(t+x) - F(t)$. So e_x is a solution of the (h_x, F) renewal equation.

To use the proposition, we have to bring our problem into the correct form. We have that $e_x(t) = \mathbb{E}\left[\Phi(S_1 - t, S_2 - t, \ldots)\right]$ where $\Phi(s_1, s_2, \ldots) = \mathbb{1}_{\min\{s_i :\, s_i \geq 0\} \leq x}$. Then e_x is a solution of the (h, F) renewal equation, where

$$\begin{aligned} h(t) &= \mathbb{E}\,\underbrace{\left[\Phi(S_1 - t, \ldots) - \mathbb{1}_{T_1 \leq t}\Phi(S_2 - t, \ldots)\right]}_{=\begin{cases} 0, & T_1 \leq t \text{ or } T_1 > t + x \\ 1, & t < T_1 \leq t + x \end{cases}} \\ &= \mathbb{P}\left[t < T_1 \leq t + x\right] = F(t+x) - F(t). \end{aligned}$$

3.7 Well-Posedness of the Renewal Equation

Theorem 3.27. *Let $h : \mathbb{R}_+ \to \mathbb{R}$ be measurable, locally bounded. Then there exists a unique $g : \mathbb{R}_+ \to \mathbb{R}$ measurable, locally bounded, and a solution of $g = h + g * F$ is given by $g = h + h * m$.*

Intuitive Proof. This is only to show how to get the idea that $g = h + h * m$ is a solution. Assume g is a solution, then we have

$$\begin{aligned} g =& h + g * F \\ =& h + (h + g * F) * F \\ &\vdots \\ \overset{(*)}{=}& h + h * F + h * F^{*2} + h * F^{*3} + \ldots \\ =& h + h * m \end{aligned}$$

□

Rigorous Proof. **Existence** $g = h + h * m$ is measurable and locally bounded, because h is. We have

$$\begin{aligned} h + g * F &= h + (h + h * m) * F \\ &= h + h * \underbrace{(F + m * F)}_{=m} = g. \end{aligned}$$

Uniqueness Let g_1, g_2 be two solutions of the (h, F) renewal equation. Then $g_1 - g_2 = (g_1 - g_2) * F = (g_1 - g_2) * F^{*n}$. We have for every $t \geq 0$

$$|g_1(t) - g_2(t)| = \left| \int_0^t (g_1 - g_2)(t - s) dF^{*n}(s) \right| \leq \sup_{[0,t]} |g_1 - g_2| \int_0^t dF^{*n}(s).$$

Where we can see the integral term is equal to $\mathbb{P}[T_1 + \ldots + T_n \leq t]$ which converges to 0. Hence $g_1 = g_2$. □

3.8 Asymptotic Behavior

In this section we assume that the law of T_1 is non-arithmetic.
Motivation Let g be the solution of the (h, F) renewal equation, what is the asymptotic behavior of $g(t)$ for $t \to \infty$?
Case 1 $h = \mathbb{1}_{[a,b]}$ for $0 \leq a \leq b$ and g the solution of the (h, F) renewal equation.

$$\begin{aligned} g(t) &= h(t) + \int_0^t h(t-s) dm(s) = h(t) + \int_{t-b}^{t-a} h(s) dm(s) \\ &= \underbrace{h(t)}_{\to 0} + \underbrace{m(t-a) - m(t-b)}_{\overset{\text{(Blackwell)}}{\to} \frac{b-a}{\mu}}. \end{aligned}$$

Where it was assumed that t was large in the last two equalities. Hence $\lim_{t\to\infty} g(t) \frac{1}{\mu} \int_0^\infty h(s) ds$.
Question How does this generalize?
Idea Extend to simple functions $\sum \lambda_i \mathbb{1}_{I_i}$ (this is easy), then try to extend to directly integrable Riemann functions.

Definition 3.11. $h : \mathbb{R}_+ \to \mathbb{R}_+$ measurable, h is called *directly Riemann Integrable* (dRi) if

$$\forall \Delta > 0 \quad \sum_{k=0}^{\infty} \Delta \sup_{[k\Delta,(k+1)\Delta]} h < \infty$$

and

$$\lim_{\Delta \to 0} \Delta \sum_{k=0}^{\infty} \sup_{[k\Delta,(k+1)\Delta]} h = \lim_{\Delta \to \infty} \Delta \sum_{k=0}^{\infty} \inf_{[k\Delta,(k+1)\Delta]} h.$$

$h : \mathbb{R}_+ \to \mathbb{R}_+$ is dRi if and only if $h_+ = \max(h, 0)$ and $h_- = \max(-h, 0)$ are dRi.

Remark 3.28 (Counter Example). $h = \sum_{k>0} \mathbb{1}_{[k,k+2^{-k}]}$ is integrable, but is not dRi.
TODO: Figure

Proposition 3.29. *Let $h : \mathbb{R}_+ \to \mathbb{R}$ be measurable. Assume that h is continuous at a.e. $t \in \mathbb{R}$ and there exists H non-decreasing such that $0 \leq |h| \leq H$ and $\int_0^\infty H < \infty$. Then h is dRi.*

Remark 3.30. In particular if h is bounded, continuous at a.e. $t \in \mathbb{R}$, and vanishes outside a compact set, the h is dRi.

// The proof is omitted and given in Sznitman's notes, I think we should include it here.//

Theorem 3.31 (Smith Key Renewall Theorem)**.** *Let h be dRi, F non-arithmetic. Then $g = h + h * m$ satisfies*

$$\boxed{\lim_{t\to\infty} g(t) = \frac{1}{\mu}\int_0^\infty h(u)du.}$$

Remark 3.32. The case $h = \mathbb{1}_{[0,b]}$ corresponds to the Blackwell Theorem.

The idea of the proof is to use an approximation of h by functions of the form $h_{c,\Delta} = \sum_{k\geq 0} c_k \mathbb{1}_{[k\Delta,(k+1)\Delta)}$.

Proof. since h is dRi we have

$$\sum_k \sup_{[k,k+1]} h < \infty.$$

Hence $h(t) \to 0$. Therefore it suffices to prove

$$\lim_{t\to\infty} \int_0^t h(t-s)dm(s) = \frac{1}{\mu}\int_0^t h(u)du.$$

Let $\Delta > 0$ such that $F(\Delta) < 1$.

Assume $h = \sum_{k\geq 0} c_k \mathbb{1}_{[k\Delta,(k+1)\Delta)}$ with $c_k \geq 0$ and $\sum_{k\geq 0} c_k < \infty$. By monotone convergece

$$h(t-s)dm(s) = \sum_{k\geq 0} \underbrace{c_k[m(t-k\Delta) - m(t-k\Delta-\Delta)]}_{h_k(t)}.$$

Observe that for every $u \geq \Delta$

$$\begin{aligned}1 \geq F(u) = m(u) - \int_0^u F(u-s)dm(s) &= \int_0^u (1-F(u-s))dm(s) \\ &\geq \int_{u-\Delta}^u \underbrace{(1-F(u-s))}_{\geq 1-F(\Delta)} dm(s) \geq (1-F(\Delta))\left(m(u)-m(u-\Delta)\right).\end{aligned}$$

In the first equality, it was used that m is the solution of the (F,F) renewal equation. Hence for every t and every k

$$h_k(t) \leq \frac{c_k}{1-F(\Delta)},$$

by distinguishing between $t-k\Delta \geq \Delta$ and $t-k\Delta < \Delta$, and using that m is non-decreasing, vanishing on $(-\infty,0)$. By dominated convergence

$$\lim_{t\to\infty} \sum_{k\geq 0} h_k(t) = \sum_{k\geq 0} \underbrace{\lim_{t\to\infty} h_k(t)}_{\overset{\text{(Blackwell)}}{=} c_k \frac{\Delta}{\mu}} .$$

Hence $\lim_{t\to\infty} \int_0^t h(t-s)dm(s) = \sum_{k=0}^{\infty} c_k \frac{\Delta}{\mu} = \frac{1}{\mu}\int_0^\infty h(u)du$.

Now assume $h \geq 0$ dRi. Let $\Delta > 0$ such that $F(\Delta) < 1$. Write

$$\underline{h}_\Delta = \sum_{k\geq 0} (\inf_{[k\Delta,(k+1)\Delta]} h)\mathbb{1}_{[k\Delta,(k+1)\Delta)}$$

$$\overline{h}_\Delta = \sum_{k\geq 0} (\sup_{[k\Delta,(k+1)\Delta]} h)\mathbb{1}_{[k\Delta,(k+1)\Delta)}.$$

We have for every t

$$\int_0^t h(t-s)dm(s) \leq \int_0^t \overline{h}_\Delta(t-s)dm(s) \to \frac{1}{\mu}\int_0^t \overline{h}_\Delta(u)du.$$

Hence

$$\limsup_{t\to\infty} \int_0^t h(t-s)dm(s) \leq \frac{1}{\mu}\int_{\mathbb{R}} \overline{h}_\Delta(u)du.$$

Since

$$\left|\int_{\mathbb{R}} \overline{h}_\Delta(u)du - \int_{\mathbb{R}} h(u)du\right| \leq \sum_{k\geq 0} \Delta \left(\overline{h}_\Delta(k\Delta) - \underline{h}_\Delta(k\Delta)\right) \overset{\Delta\to 0}{\longrightarrow} 0,$$

where the limit is due to h being dRi. We can let Δ tend to 0 in the equation above (with $\limsup$) to obtain

$$\limsup_{t\to\infty} \int_0^t h(t-s)dm(s) \leq \frac{1}{\mu}\int_{\mathbb{R}} h(u)du,$$

and equivalently

$$\liminf_{t\to\infty}\int_0^t h(t-s)dm(s) \geq \frac{1}{\mu}\int_{\mathbb{R}} h(u)du.$$

$$\frac{1}{\mu}\int_{\mathbb{R}} h(u)du \leq \liminf_{t\to\infty}\int_0^t h(t-s)dm(s) \leq \limsup_{t\to\infty}\int_0^t h(t-s)dm(s) \leq \frac{1}{\mu}\int_{\mathbb{R}} h(u)du.$$

□

// This was done in class //

Application Let $\mu < \infty$. Let E_t be the excess time (time until next renewal) and $e_x(t) = \mathbb{P}[E_t \leq x]$. What is $\lim_{t\to\infty} e_x(t)$? We know that $e_x = h_x + e_x * F$, where $h_x(t) = F(t+x) - F(t)$.

Remark 3.33. $\mu = \mathbb{E}[T_1] = \int_0^\infty \mathbb{P}[T_1 > t]\, dt$

With this we have that $h_x(t) \leq 1 - F(t) = \mathbb{P}[T_1 > t]$, and $1 - F(t)$ is non-increasing in t and continuous a.e. (because it is the difference of two monotone functions).

$$\int_0^\infty \mathbb{P}[T_1 > t]\, dt = \mathbb{E}[T_1] = \mu < \infty.$$

Thus (by the proposition) h_x is dRi. Now we can apply the theorem and get that

$$\lim_{t\to\infty}\mathbb{P}[E_t \leq x] = \frac{1}{\mu}\int_0^\infty h_x(t)dt = \frac{1}{\mu}\int_0^\infty F(t+x) - F(t)dt,$$

with $F(t+x) - F(t) = \mathbb{E}\left[\mathbb{1}_{T_1\in(t,t+x]}\right]$, we find that the limit is equal to

$$\frac{1}{\mu}\int_0^\infty \mathbb{E}\left[\mathbb{1}_{T_1\in(t,t+x]}\right] dt = \frac{1}{\mu}\mathbb{E}\left[\int_0^\infty \mathbb{1}_{t\in[T_1-x,T_1)}\right] dt = \frac{1}{\mu}\mathbb{E}\left[\int_{\max\{T_1-x,0\}}^{T_1} dt\right] = \begin{cases} T_1, & T_1 \leq x \\ x, & T_1 > x. \end{cases}$$

Thus for t large: $\mathbb{P}[E_t \leq x] \approx \frac{1}{\mu}\mathbb{E}[\min\{T_1, x\}]$.

Remark 3.34. $G(x) = \frac{1}{\mu}\mathbb{E}[\min\{T_1, x\}]$ is the delay distribution in the proof of Blackwell's Theorem.

Conclusion We have now used renewal processes to define a general structure to model a real life process mathematically. Using this object enabled us to implement the LLN and make statements about the asymptotic behavior of such processes over large periods of time.

Chapter 4

General Poisson Point Processes

Reference Lectures on the Poisson Process (Penrose), Poisson Processes (Kingman)

4.1 Introduction

Question How can we represent points on $\mathbb{R}_+$ mathematically?

(i) A set of points $\mathcal{S} = \{S_1, S_2, ...\}$

(ii) 'Time point of view', ie $T_1, T_2, ...$ where T_i = time between the $(i-1)$'th and i'th point.

(iii) Cadlag formulation with values in $\mathbb{N}$. N_t = number of points in $[0, t]$.

(iv) Measure $N : \mathcal{B}(\mathbb{R}_+) \to \mathbb{N}$ with $N(A)$ = number of points in A.

Goal Define $\Omega \to$'set of points'. For a general state space $\mathbb{R}^2, [0,1]^2$, a manifold, etc. (ii) and (iii) are specific to $\mathbb{R}_+$, so they do not generalize. (i) is not very easy to describe. (iv) is actually nice, so we will use this point of view.

Framework (E, d) a Polish space (separable, complete, metric space). $\mathcal{E}$ Borel σ-algebra. $\mu : \sigma$ finite measure on $(E, \mathcal{E})$, ie $\exists B_i \uparrow E : \mu(B_i) < \infty$ where $B_i \uparrow E \iff B_1 \subset B_2 \subset ... : \bigcup_{i \geq 1} B_i = E$.

Example 4.1. Of such spaces:

(i) $E = \{0\}, \mu = \delta_0$

(ii) $E = \mathbb{R}_+, \mu = \lambda \mathcal{L}$

(iii) $E = \mathbb{R}^2, \mu(dx) = \frac{1}{\pi} e^{-|x|^2} dx$ 'Gaussian'

Goal We wish to define a point process on $(E, \mathcal{E})$ where the 'number of points around x ' $\approx \mu(dx)$ on $\mathbb{R}_+$.

4.2 Point Processes

Notation $\mathcal{N} = \{\nu : \nu = \sigma-$finite measure st $\forall B \in \mathcal{E} : \nu(B) \in \mathbb{N} \cup \{+\infty\}\}$. **Measure Structure** Let $\mathcal{B}(\mathcal{N})$ be the σ-algebra generated by the sets $\{\nu \in \mathcal{N} : \nu(B) = k\} = \mathcal{N}_k$ for $B \subset E$ meas and $k \in \mathbb{N}$. $\rightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ measured space.

Proposition 4.1. *Let $\mathcal{N}_{<\infty} = \{\nu \in \mathcal{N} : \nu(E) < \infty\}$, there exists meas maps $\tau : \mathcal{N}_{<\infty} \rightarrow \mathbb{N}, X_i : \mathcal{N}_{<\infty} \rightarrow E$ st $\forall \nu \in \mathcal{N}_{<\infty} : \nu = \sum_{i=0}^{\tau(\nu)} \delta_{X_i(\nu)}$.*

Definition 4.1. A point process on $(E, \mathcal{E})$ is a RV N with values in $\mathcal{N}$. 'N is a random σ-finite measure', $N \leftrightarrow$ 'random set of points'.

This means $N : \Omega \rightarrow \mathcal{N}$ meas, for any fixed $B \subset E : N(B) : \Omega \rightarrow \mathbb{N} \cup \{+\infty\}$ is measurable. Thus a stochastic process corresponds to $(N(B))_{B \in \mathcal{E}}$. '$N(B)$ = number of points in B'.

Example 4.2. Point Processes:

- $N = 0$ a.s. $\rightarrow$ empty set
- $E = [0, 1], X$ RV on $[0, 1]$. $N = \delta_X$ is a point process.
- $X_1, ...X_n$ iid RV on $[0, 1]$, $N = \delta_{X_1} + ... + \delta_{X_n}$ is a point process.

4.3 Poisson Point Processes

Setup $(E, \mathcal{E})$ Polish, μ fixed σ-finite measure (think of $\lambda\mathcal{L}$), $\mathcal{N} = \{\sigma$ finite counting measure$\}$, $(\Omega, F, \mathbb{P})$ abstract prob space.

Definition 4.2. A Poisson process with intensity μ on $(E, \mathcal{E})$ ($ppp(\mu)$) is a point process st:

(i) $\forall B_1...B_k \subset E$ meas and disjoint: $N(B_1)...N(B_k)$ are indep.

(ii) $\forall B \subset E$ meas $N(B) \sim Pois(\mu(B))$.

4.4 Existence and Uniqueness

Question Does there always exist a $ppp(\mu)$ on E?

Spaces with finite measure

Proposition 4.2. *Let $Z \sim Pois(\mu(E))$, $(X_i)_{i \geq 1}$ iid where $X_i \sim \frac{\mu(.)}{\mu(E)}$. Then $N = \sum_{i=1}^{Z} \delta_{X_i}$ is a $ppp(\mu)$ on E.*

Superposition

Lemma 4.3. *Let* $\lambda = \sum_{i=1}^{\infty} \lambda_i, \lambda_i \geq 0$. $X_i \sim Pois(\lambda_i), i \geq 1$ *indep, then* $X = \sum_{i=1}^{\infty} X_i \sim Pois(\lambda)$.

Theorem 4.4. *Let* $N_i, i \geq 1$ *be a sequence of indep ppp*(μ_i) *where* μ_i *and* $\mu = \sum_{i=1}^{\infty} \mu_i$ *are* σ*-finite measures. Then* $N = \sum_{i=1}^{\infty} N_i$ *is a ppp*(μ).

Corollary 4.5. μ σ*-finite measure on* $(E, \mathcal{E})$*, then* $\exists$ *ppp*(μ) *on* E.

Uniqueness

Let N be a *ppp*(μ) on E, define $P_N =$ law of N ($\rightarrow$ a probability meas on $\mathcal{N}$).

Proposition 4.6. *Let* N, N' *be two ppp*(μ) *on* $(E, \mathcal{E})$ *then* $P_N = P_{N'}$.

Theorem 4.7 (Representation of ppp as Proper Processes)**.** *Let* N *be a ppp*(μ) *on* $(E, \mathcal{E})$*, there exists some RV* $\tau \in \mathbb{N} \cup \{+\infty\}$ *st:* $X_n \in E, n \geq 1 : N = \sum_{i=1}^{\tau(}$

4.5 Laplace Functional

N a random meas on $(E, \mathcal{E})$ for $u : E \rightarrow \mathbb{R}$ what should we interpret $\int_E u dN$ as?

Lemma 4.8. $X \sim Pois(\lambda), \lambda > 0$*, then* $\forall u \geq 0 : \mathbb{E}\left[e^{-uX}\right] = exp(-\lambda(1 - e^{-u}))$.

Definition 4.3. Let N be a point process on $(E, \mathcal{E})$, for every $u : E \rightarrow \mathbb{R}_+$ define $L_N(u) = \mathbb{E}\left[exp(-\int u(x)N(dx)\right]$

Remark 4.9. $\int_E u(x)N(dx) = \int_E u dN$ is a RV.

Theorem 4.10 (Characterization via Laplace Functional)**.** *Let* μ σ*-finite meas on* $(E, \mathcal{E})$*. Let* N *be a point process on* E*. TFAE:*

(i) N *is a ppp*(μ)

(ii) $\forall u : E \rightarrow \mathbb{R}_+$ *meas:* $L_N(u) = exp(-\int_E 1 - e^{-u(x)}\mu(dx))$

4.6 Simple Processes

Remark 4.11. For $x \in E$, $\{x\}$ is meas. because E is Polish.

Definition 4.4. A measure $\eta \in \mathcal{N}$ is said to be simple if $\forall x \in E : \eta(\{x\}) \leq 1$.

Proposition 4.12. *$\{\eta : \eta$ is simple$\}$ is measurable in $\mathcal{N}$.*

Theorem 4.13. *Assume that μ is a diffuse ($\forall x : \mu(\{x\} = 0$) σ finite measure. Then every ppp(μ) is simple a.s.*

Consequence $\exists \tau,\ X_i$ RV, $X_i \neq X_j$ if $i \neq j$ a.s.: $N = \sum_{i=1}^{\tau} \delta_{x_i}$ a.s.

4.7 Mapping and Restriction

$(E, \mathcal{E}), (F, \mathcal{F})$ Polish spaces, μ σ-finite measure on E, $T : E \to F$ meas, $T\#\mu$ push forward measure of μ under T $[T\#\mu(B) = \mu(T^{-1}(B))]$.

Theorem 4.14. *Assume that $T\#\mu$ is σ-finite. Let N be a ppp(μ) on E, then $T\#N$ is a ppp$(T\#\mu)$ on F.*

Example 4.3. $E = \mathbb{R},\ F = \mathbb{Z},\ T : E \to F; x \to \lfloor x \rfloor,\ \mu = \mathcal{L},\ T\#\mu = |.|$.

Notation If ν is a measure on E, $C \subset E$ meas. $\nu_C : \nu(. \cap C)$

Theorem 4.15 (Restriction)**.** *Let $C_1, C_2, ... \subset E$ meas. and disjoint. If N is a ppp(μ) on E, then $N_{C_1}, N_{C_2}...$ are indep ppp with resp. intensities $\mu_{C_1}, \mu_{C_2}, ...$*

4.8 Marking

Motivation Cars on a highway, at time 0 the position of the cars is a *ppp*(1) on $\mathbb{R}$ (that means on average 1 car per kilometer of highway). We put an observer (Olga) at 0 on $\mathbb{R}$.

Case 1: All of the cars have speed 50km/h, we want to study X = number of cars seen by Olga in 1 hour. What is the law of X? $X \sim Pois(50)$.

Case 2: The cars have a random speed $\sim \mathcal{U}([50, 100])$. What is the law of X? It may at first seem complicated, but it is not!

Framework $(E, \mathcal{E})$ Polish, $\mu = \sigma$-finite. $(F, \mathcal{F}, \nu)$ Polish, Probability space.

Definition 4.5. Let $N = \sum_{i=1}^{\tau} \delta_{X_i}$ a *ppp*(μ) on E. Y_i iid RV with law ν indep of N. The marked point process is the PP on $E \times F$ defined by $M = \sum_{i=1}^{\tau} \delta_{(X_i, Y_i)}$.

Remark 4.16. X_i corresponds to the position of the cars in Case 2, and Y_i to their speeds.

Theorem 4.17. *The marked process is a ppp*$(\mu \otimes \nu)$.

Conclusion The General PPP we have defined gives us a very general way to talk about a random processes on a large class of spaces (Polish), which fulfill a Markov-like property. This tool will allow us to make much stronger statements in more specific cases.

4.9 Standard Poisson Process

In discrete time processes $(X_n)_{n\in N}$, the law is characterised by the law of $(X_{n_1}, ..X_{n_k}; n_1...n_k \in \mathbb{N})$. In continuous time processes we have $(X_t)_{t\geq 0}$, we need to define $X_t : \forall t \in \mathbb{R}$ which is not countable.
Outset We would like to define a renewal process which also fulfills the Markov property, enabling us to not have. Furthermore we would like a simple continuous time process which is in some way a 'universal' stationary process on $\mathbb{R}_+ \to \mathbb{N}$ with independent increments and jumps of size 1. We would also like to see if any of the ideas from the previous chapter can be specified to this context.

Applications Queuing processes, insurance claims, compound Poisson process.

Framework $(\Omega, F, \mathbb{P})$ probability space, time space: $\mathbb{R}_+ = [0, \infty)$

There are 2 points of view: random points on $\mathbb{R}_+$ (reminiscent of PPP) or continuous time stochastic process (renewal process).

4.10 Exponential Random Variables

Note We will use the 2nd point of view here.

Definition 4.6. Let $\lambda > 0$, a real RV T is exponential with parameter λ (we write $T \sim Exp(\lambda)$) if it has density $f(t) = \lambda e^{-\lambda t}\chi_{\{t\geq 0\}}$. $\iff \forall t \geq 0 \mathbb{P}\left[T > t\right] = e^{-\lambda t}$

Proposition 4.18 (Memoryless Property)**.** *Let* $\lambda > 0$ *and* $T \sim Exp(\lambda)$*. Then* $\forall s, t \geq 0$: $\mathbb{P}\left[T > s + t | T > t\right] = \mathbb{P}\left[T > s\right]$

Proposition 4.19 (Minimum of indep Exponentials)**.** *Let* $n \geq 0, T_1...T_n$ *indep with* $T_i \sim Exp(\lambda_i), \lambda_i > 0$:

- $min\{T_1...T_n\} \sim Exp(\lambda_1 + ... + \lambda_n)$
- $\mathbb{P}\left[T_1 = min\{T_1...T_n\}\right] = \frac{\lambda_1}{\lambda_1 + ... + \lambda_n}$

Reminder X a real RV with density f, Y a RV with values in some measurable space E indep of X. Then $\forall \phi : \mathbb{R} \times E \to \mathbb{R}$ meas + bdd we have: $\mathbb{E}\left[\phi(X,Y)\right] = \int_0^\infty \mathbb{E}\left[\phi(x,Y)\right] f(x)dx$

Proposition 4.20 (Sum of Exponentials)**.** *Let $\lambda > 0, n \geq 1$. Let $T_1 ... T_n$ be iid $Exp(\lambda)$ RVs. Then $S_n = T_1 + ... + T_n$ is $\Gamma(n, \lambda)$ distributed. ie S_n is continuous with density $f_{S_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$*

We can check that $\Gamma(1,t) = Exp(\lambda)$

4.11 Definition of Poisson Processes

Setup $\lambda > 0, (T_i)_{i \geq 0}$ iid $Exp(\lambda), S_n = T_1 + ... + T_n$

Definition 4.7. The stochastic process $N = (N_t)_{t \geq 0}, N_t = \sum_{i=1}^\infty \chi_{S_i \leq t}$ is called the Poisson process with intensity λ ($pp(\lambda)$). The RVs $T_1, T_2, ...$ are the inter-arrival times and $S_1, S_2, ...$ the arrival times/jump times.

Elementary Properties

- The mapping $t \to N_t$ is a.s. right continuous, with values in $\mathbb{N}$
- For fixed $t \geq 0$ $N_t \sim Pois(\lambda t)$ ie $\mathbb{P}\left[N_t = n\right] = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$

Comment "A property hold a.s." $\iff \exists$ meas set $A : \mathbb{P}\left[A\right] = 1$ and $\forall \omega \in A$ the property holds.

4.12 Markov Property

Theorem 4.21 (Markov Property of N)**.** *Fix $t \geq 0$, the stochastic process $N^{(t)} = (N_s^{(t)})_{s \geq 0}$ defined by $N_s^{(t)} = N_{t+s} - N_t$ is a Poisson process, independent of $(N_u)_{0 \leq u \leq t}$.*

4.13 Stationary and Independent Increments

Motivation We want to describe the law of $(N_{t_0}, ..., N_{t_k})$, the key here is that they are not totally independent. If we have 5 points at time t_0 then we know at time t_1 there will be at least 5 points. So we look at the law of $(N_{t_1} - N_{t_0}, ..., N_{t_k} - N_{t_{k-1}})$ ie the increments.

Definition 4.8. A stochastic process $(X_t)_{t \geq 0}$ is said to have indep and stationary increments if

- $\forall k \geq 1, \forall 0 = t_0 < ... < t_k : X_{t_1} - X_{t_0}, ..., X_{t_k} - X_{t_{k-1}}$ are indep
- $\forall s < t, \forall n \geq 0 : X_t - X_s \overset{law}{=} X_{t+h} - X_{s+h}$

Theorem 4.22 (Marginals of Poisson Process)**.** *We have the following:*

(i) $\forall k \geq 1, \forall 0 = t_0 < ... < t_k : N_{t_1} - N_{t_0}, ..., N_{t_k} - N_{t_{k-1}}$ *are indep*

(ii) $\forall s \leq t : N_t - N_s \sim Pois(\lambda(t-s))$

In particular $N = (N_t)_{t\geq 0}$ *has indep and stationary increments.*

We know the law of $(N_{t_1}, ..., N_{t_k}$ for every fixed $t_1 ... t_k$.

$$\begin{aligned}\mathbb{P}\left[N_{t_1} = m_1 ... N_{t_k} = m_k\right] =& \mathbb{P}\left[N_{t_1} = m_1, N_{t_2} - N_{t_1} = m_2 - m_1, ..., N_{t_k} - N_{t_{k-1}} = m_k - m_{k-1}\right] \\ =& \prod_{i=1}^{k} \frac{(\lambda(t_0 - t_{i-1}))^{m_i - m_{i-1}}}{m_i - m_{i-1}} e^{-\lambda(t_i - t_{i-1})}\end{aligned}$$

4.14 Finite Marginals Characterization

Motivation Let $(N_t)_{t\geq 0}$ a stochastic process. Does the last formula from above ensure that the process is $pp(\lambda)$? No, we can define $\tilde{N}_t = \sum_{i\geq 1} \chi_{S_i < t}$, we could also just change the value of the process as some random points, thus when we fix $t_1, ..., t_k$ we have 0 probability to see these.

In order to get a characterization we need to add some regularity assumptions.

Definition 4.9. Let $N = (N_t)_{t\geq 0}$ be a continuous stoch process with values in $\mathbb{R}$. We say that N is a counting process if the following holds a.s.:

(i) $N_0 = 0$ a.s.

(ii) $t \to N_t$ is non decreasing, right continuous, with values in $\mathbb{N}$

In this case, we can define the jump times by setting $S_1 = min\{t : N_t > 0\}$ and by induction $S_{i+1} = min\{t \geq S_i : N_t > N_{S_i}\}$.

Example 4.4. $pp(\lambda)$ is a counting process.

Remark 4.23. The condition (ii) is almost sure in the following manner: $\exists A$ meas. with $\mathbb{P}[A] = 1$ st $\forall \omega \in A : t \to N_t(w)$ is non decreasing, right continuous, with values in $\mathbb{N}$.

Theorem 4.24. *Let* $\lambda > 0$: *Let* N *be a counting process, the following are equivalent:*

(i) N *is* $pp(\lambda)$

(ii) $\forall k \geq 1, \forall t_0 = 0 < t_1 < ... < t_k, \forall n_1, ..., n_k \in \mathbb{N}$:
$\mathbb{P}\left[N_{t_1} - N_{t_0} = n_1, ..., N_{t_k} - N_{t_k-1} = n_k\right] = \prod_{i=1}^{k} \frac{(\lambda(t_i - t_{i-1}))^{n_i}}{n_i!} e^{-\lambda(t_i - t_{i-1})}$

Remark 4.25. By def N is a $pp(\lambda) \iff N$ is a counting process with jumps of size 1 a.s. and $S_1, S_2 - S_1, ...$ are iid $exp(\lambda)$.

4.15 Microscopic Characterization

Theorem 4.26. *Let* N *be a counting process, let* $\lambda > 0$. *TFAE:*

(i) N *is* $pp(\lambda)$

(ii) N *has indep and stationary increments and* $\mathbb{P}\left[N_t = 1\right] = \lambda t + o(t)$ *and* $\mathbb{P}\left[N_t \geq 2\right] = o(t)$

4.16 Properties of Poisson Process

Theorem 4.27 (Law of Large Numbers)**.** *Let* N *be a* $pp(\lambda), \lambda > 0$*, then:* $lim_{t\to\infty} \frac{N_t}{t} = \lambda$.

Motivation If we want to specify (and remove) certain points, for instance if the PP is describing arrival times at a bakery then say we want to differentiate between customers who are younger than 45 and those who are older. If we just look at one of these groups, what type of process are they?

Theorem 4.28 (Thinning)**.** *Let* $(N_t)_{t\geq 0} \sim pp(\lambda)$ *with jump times* $(S_i)_{i\geq 0}$. *Let* $(X_i)_{i\geq 0}$ *iid* $Ber(p)$ *indep of* N *(this is the differentiation, called the marking of* N*). Define* $N_t^1 = \sum_{i\geq 1} \chi_{S_i \leq t, X_i = 1}$ *and* $N_t^0 = \sum_{i\geq 1} \chi_{S_i \leq t, X_i = 0}$.
(N_t^0) *and* (N_t^1) *are indep Poisson processes with respective rates* $\lambda_0 = (1-p)\lambda, \lambda_1 = p\lambda$.

Let (N_t^0) and (N_t^1) be indep Poisson processes with respective rates $\lambda_0 > 0, \lambda_1 > 0$. Let $N_t = N_t^0 + N_t^1$.

Theorem 4.29. N_t *is a counting process and we define for every* i*:* $X_i = \mathbb{1}_{\{i\text{'th jump of } N_t \text{ is a jumping time of } N_t^1\}}$. *Then* N_t *is a* $pp(\lambda_0 + \lambda_1)$ *and* (X_i) *is a marking of* N *with* $\forall i : \mathbb{P}\left[X_i = 1\right] = \frac{\lambda_1}{\lambda_0 + \lambda_1}$.

Conclusion We successfully defined a renewal process with the Markov property, we also found that this object is also a PPP, thus giving us a process which has the asymptotic behavior (LLN, etc) from the renewal process perspective and getting the Strong and Weak Markov Property from the Poisson Point Process perspective.

Chapter 5

Continuous Time Markov Chains

Framework $(\Omega, \mathcal{F}, \mathbb{P})$ Probability space, E finite or countable.

Outset We will now be extending the theory of Discrete Markov Chains developed in Chapters 1 and 2 and generalizing the theory of Poisson Processes in Chapter 5. Instead of jumping at every step (studying $(X_n)_{n\in\mathbb{N}}$), we will now make jumps at random times on $\mathbb{R}_+$ with the continuous time MC $(X_t)_{t\geq 0}$ using times on $\mathbb{R}_+$.

Discrete Time MC	**Continuous Tim**
Time $\mathbb{N}$	$\mathbb{R}_+$
Initial Distribution $X_0 \sim \mu$	$X_0 \sim \mu$
Memoryless Property $\mathbb{P}\left[X_{n+1} = x_{n+1} \mid X_0 = x_0, ..., X_n = x_n\right] =$ $\mathbb{P}\left[X_{n+1} = x_{n+1} \mid X_n = x_n\right]$	$\forall t_0$ $\mathbb{P}\left[X_{t_{n+1}} = x_{n+1}\right.$ $\mathbb{P}\left[X\right.$
Transition Probabilities $\mathbb{P}\left[X_{n+1} = y \mid X_n = x\right] = p_{x,y}$	μ-scopic genera $\mathbb{P}\left[X_{t+h} = y \mid X_t = x\right.$ for h small the pr equal to 1.

5.1 Definition via Generator

Definition 5.1. Let $X = (X_t)_{t\geq 0}$ be a cont. time stochastic process with values in E. We say that X is a jump process without explosion if a.s.

(i) $t \mapsto X_t$ is right continuous

(ii) $\forall t > 0$ the number of discontinuity points of $s \mapsto X_s$ on $[0, t]$ is finite.

Definition 5.2. Jump times: $S_0 = 0, S_{i+1} = inf\{t > S_i, X_t = X_{S_i}\}$, with condition (ii) implying that $S_n \to \infty$ as $n \to \infty$ a.s.

Definition 5.3. Skeleton: $\forall n \in \mathbb{N} : \bar{X}_n := X_{S_n}$ if $S_n < \infty$, if $\exists n_0 : S_n = \infty \ \forall n \geq n_0$ then $\forall n \geq n_0 : X_n = X_{n_0 - 1}$.

Definition 5.4. A generator (Q-matrix) is a family $q = (q_{xy})_{x,y\in E}$ where:

(i) $q_{xy} \geq 0 \forall x \neq y$

(ii) $\forall x : \sum_{y\neq x} q_{xy} < \infty$

(iii) $q_{xx} = -q(x) = -\sum_{y\neq x} q_{xy}$

Definition 5.5. Let μ be a distribution on E, q a generator, let X be a jump process without explosion. We say that X is a $CTMC(\mu, q)$ (Continuous Time Markov Chain without explosion with initial distribution μ and generator q) if:

(i) $X_0 \sim \mu$

(ii) $\forall t_1 < ...t_{n+1} : \forall x_1, ..., x_{n=1} \in E : \mathbb{P}\left[X_{t_{n+1}} = x_{n+1} | X_{t_1} = x_1, ..., X_{t_n} = x_n\right] = \mathbb{P}\left[X_{t_{n+1}} = x_{n+1} | X_n = x_n\right]$

(iii) $\forall x, y \in E : \forall t > 0$: as $h \to 0^+$: $\mathbb{P}[X_{t+h} = y | X_t = x] = \delta_{xy} + q_{xy}h + o(h)$ uniformly in $t \geq 0, y \in E$.

Remark 5.1. In (iii): $\forall x, \exists \varphi_x : \mathbb{R}_+ \to \mathbb{R}_+$ st $\varphi_x(h) \overset{h\to 0^+}{\to} 0$ and $\forall h > 0, \forall y \in E : \mathbb{P}[X_{t+h} = y | X_t = x] = \begin{cases} 1 - q(x)h + h\varphi_{x,x,t}(h) \\ q_{xy}h + h\varphi_{x,y,t}(h) \end{cases}$ where $0 \leq \varphi_{x,z,t}(h) \leq \varphi_x(h)$.

Example 5.1 (Poisson Process). Let $(N_t)_{t\geq 0}$ be a $pp(\lambda)$. Then N is a $CTMC(\mu, q)$ with $\mu = \delta_0$ and $q = (q_{xy})_{x,y\in\mathbb{N}} = \lambda$ if $y = x + 1$, $-\lambda$ if $y = x$, and 0 otherwise.

Question Does $CTMC(\mu, q)$ exist for arbitrary μ and q?

5.2 Non-Rigorous Section: The Constructive Approach

Example 5.2 (2 State Markov Chain). $E = \{1, 2\}$, $q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$, $\alpha, \beta > 0$. $(X_t)_{t\geq 0}, X_t \sim CTMC(\delta_1, q)$? $X_0 = 1$, $T_1 \sim Exp(\alpha)$, $T_2 \sim Exp(\beta)$ (see notes for reasoning). This gives us the candidate $X_t = \begin{cases} 1, t \in [S_i, S_{i+1}) \\ 2, t \in [S_{i+1}, S_{i+2}) \end{cases}$.

Idea q_{xy} should represent the parameter for the time taken to jump from x to y. Since we want our process to have the Markov property, it is natural to see q_{xy} as the parameter in the exponential RV representing the waiting time to jump from x to y.

Example 5.3 (3 State Markov Chain). We start at $X_0 = 1$, we have probability α to jump to 2, and probability β to jump to 3. Thus we have $T_{12} \sim Exp(\alpha)$, $T_{13} \sim Exp(\beta)$, then we shall actually jump at $T_1 = min\{T_{12}, T_{13}\} \sim Exp(\alpha + \beta)$. $\mathbb{P}[\text{jump from } 1 \to 2] = \mathbb{P}[T_1 = T_{12}] = \frac{\alpha}{\alpha+\beta} = \frac{q_{12}}{q(1)}$. The skeleton $(\overline{X_n})$ is a Discrete time MC with transition probabilities $\kappa_{xy} = \frac{q_{xy}}{q(x)}$.

5.3 Definition by Skeleton and Holding Time

Note q is a fixed generator.

Discrete Chain Associated to 2

Definition 5.6. Let $x, y \in E$, if $q(x) > 0$ we define $\kappa_{xy} = \frac{q_{xy}}{q(x)}$ and $\kappa_{xx} = 0$, if $q(x) = 0$ then $\kappa_{xy} = \begin{cases} 0, x \neq y \\ 1, x = y \end{cases}$.

Remark 5.2. κ is transition probability (check for the cases where $q(x) = 0$ and $q(x) \neq 0$).

Example 5.4.
(i) The $pp(\lambda)$, with $\kappa_{i,i+1} = 1$.
(ii) The 2-State MC, with $\kappa_{1,2} = \kappa_{2,1} = 1$
(iii) The 3-State MC, more complicated (see notes).

Something can go wrong

Let μ probability measure on E, q generator. Our goal is to define (X_t) a $CTMC(\mu, q)$. Let $Y = (Y_n)$ be a discrete $MC(\mu, \kappa)$, $H_1, H_2, ...$ iid $Exp(1)$ RVs, set $T_i = \frac{1}{q(Y_i)} H_i$, conditional on Y $T_i \sim Exp(q(Y_i))$ and they are independent.

We define $S_i = T_1 + T_2 + ... + T_i$ for $i > 1$, and $X_t = Y_n$ if $t \in [S_n, S_{n+1})$. Now have we defined X_t for all $t \geq 0$? No, as $lim_{n\to\infty} S_n$ could be finite.

Definition 5.7. We say that q has no explosion if $\forall$ choice of $\mu : S_\infty = +\infty$ a.s.

Remark 5.3. This is only a condition on q.

Question Does there exist q with explosion? (Answer later)
Question If q has no explosion, is (X_t) a $CTMC(\mu, q)$? (Also later)

Birth Chain

$E = \mathbb{N}$, fix $(\lambda_i)_{i\geq 1}$, and $q_{i,i+1} = \lambda_i$, $q_{i,i} = -\lambda_i$, and otherwise $q_{i,j} = 0$. We get that $\kappa_{i,j} = \delta_{i,i-1}$, $Y_n = n$, and $T_i \sim Exp(\lambda_i)$. Now we set $S_\infty = \sum_{i=1}^{\infty} T_i$ and we ask, is $S_\infty < \infty$ or $S_\infty = \infty$ a.s.

Remark 5.4. $pp(\lambda)$ is a birth chain with $\lambda_i = \lambda$.

Theorem 5.5. *The birth chain q has no explosion* $\iff \sum_{i\geq 1} \frac{1}{\lambda_i} = \infty$.

Non-Explosion Characterization

Fix q a generator on E ($\kappa_{xy} = \frac{q_{xy}}{q(x)}$).

Theorem 5.6. *For $x \in E$, let $Y = (Y_n^{(x)})_{n\geq 0}$ be a $MC(\mu, \kappa)$. Then q has no explosion $\iff \forall x \sum_{n\geq 0} \frac{1}{q(Y_n^{(x)})} < \infty$ a.s.*

Remark 5.7. $\sum_{n\geq 0} \frac{1}{q(Y_n)}$ is a RV.

Application Sufficient Condition: q is non-explosive if

- E is finite (2 and 3 State MC)
- $inf_{x\in E:\ q(x)\neq 0} q(x) > 0$ (Poisson, 2 and 3 State MC)
- The chain κ is irreducible and recurrent.

Key Theorem

Theorem 5.8 (Characterization of CTMC)**.** *Let $X = (X_t)_{t\geq 0}$ be a jump process without explosion. Let q be a non-explosive generator. Then TFAE:*

(i) X is a $CTMC(\mu, q)$

(ii) The skeleton of X ($Y = \overline{X_n}$)is a discrete time $MC(\mu, \kappa)$ and conditioned on Y, the holding times satisfy $S_i - S_{i-1} \sim Exp(q(Y_i))$ are indep.

Consequences

- Existence of CTMC for non-explosive q
- Uniqueness of the law of a $CTMC(\mu, q)$ (if X, Y are $CTMC(\mu, q)$ then $\forall t_1 < ... < t_n$: $(X_{t_1}, ..., X_{t_n} \sim (Y_{t_1}, ..., Y_{t_n}))$
- There exist constructive algorithms (see Morris)

5.4 Markov Properties

Framework $(\Omega, \mathcal{F}, (\mathbb{P}_x)_{x\in E})$, $(X_t)_{t\geq 0}$ st under $\mathbb{P}_x$, X is $CTMC(\mu, q)$ with q non-explosive. (Such probability measures exist, take μ with $\mu(x) > 0 \forall x \in E$, consider $(X_t)_{t\geq 0} = CTMC(\mu, q)$ then let $\mathbb{P}_x = \mathbb{P}\left[.|X_0 = x\right]$.)

Simple Markov Property Fix $t \geq 0, x \in E$; Conditionally on $X_t = x$ we have that $(X_{t_s})_{s\geq 0}$ is a $CTMC(\delta_x, q)$ indep of $(X_n)_{n\leq t}$

Strong Markov Property The same applies if we replace t by a random stopping time T.

5.5 Transition Probabilities

$X = (X_t)_{t\geq 0}$ is a $CTMC(\delta_x, q)$ under $\mathbb{P}_x$, then we define for $t \geq 0$ and $x, y \in E$: $p_{xy}(t) = \mathbb{P}_x\left[X_t = y\right]$. In the discrete case this corresponds to $p_{xy}^{(n)} = p_{xy}(t)$.

Remark 5.9. We have

- $\forall t \geq 0 : (p_{xy}(t))_{x,y\in E}$ is a transition probability $\sum_y p_{xy}(t) = \sum_y \mathbb{P}_x\left[X_t = y\right] = 1$.
- $\forall x : p_{xx}(t) \geq e^{-q(x)t} \forall t$
- $\forall x, y \in E : p_{xx}(h) = 1 - q(x)h + o(h)$ and $p_{xy}(h) = q_{xy}h + o(h)$ for $x \neq y$.

Proposition 5.10 (Chapman Kolmogorov (CK) Equations)**.** $\forall t, s \geq 0 : p_{xy}(t+s) = \sum_z p_{xz}(t)p_{zy}(s)$

Question Knowing q, what is $p_{xy}(t)$?

Theorem 5.11 (Backward/Forward equations)**.** *$\forall x, y \in E : p_{xy}$ is C^1 on $\mathbb{R}_+$ and $\forall t \geq 0$ we have the backward equation:*

$$p'_{xy}(t) = \left(\sum_{z\neq x} q_{xy}p_{zy}(t)\right) - q(x)p_{xy}(t)$$

And the forward equation:

$$p'_{xy}(t) = \left(\sum_{z \neq y} p_{xz}(t) q_{zy}\right) - p_{xy}(t) q(y)$$

Application Let us look at what happens when E is finite ($E = \{1...k\}$). Then $P(t) = \begin{pmatrix} p_{11}(t) & ... & p_{1k}(t) \\ \vdots & & \vdots \\ p_{k1}(t) & ... & p_{kk}(t) \end{pmatrix}$ and $Q = \begin{pmatrix} q_{11} & ... & q_{1k} \\ \vdots & & \vdots \\ q_{k1} & ... & q_{kk} \end{pmatrix}$ So we get that $p'_{xy}(t) = \sum_{z \in E} q_{xz} p_{zy}(t) \implies P'(t) = QP(t)$ (from backward equation) we also get $P'(t) = P(t)Q$ (from forwards equation).

Theorem 5.12. *If E is finite, we have $\forall t \geq 0 : P(t) = exp(tQ)$.*