Applied Stochastic Processes Notes

Trevor Winstral

Spring Semester 2021

Contents

Chapter 1

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 1.1. 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 1.2. 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 1.1 (Simple Random Walk). State Space $\mathbb{Z}^d$, x, y are neighbors $\Leftrightarrow \|x - y\|_1 = 1$. An electron is starting at 0, and each step it jumps uniformly to one of the neighbors. How should we define this?

Definition 1.3 (SRW)**.** Let $(Z_n)_{n\in\mathbb{N}}$ iid, $\mathbb{P}\left[Z_n = \pm e_i\right] = \frac{1}{2d}$ 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 1.1 (Polya)**.**
If $d = 1, 2$ then $\mathbb{P}[(X_n)$ visits x infinitely many times$] = 1$
If $d \geq 3$ then $\mathbb{P}[(X_n)$ visits x only finitely many times$] = 1$

Example 1.2 (Poisson Process)*.* We want to define and study N_t the number of cars passing a point during $[0, t]$.

Definition 1.4. 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: $\mathbb{P}[T_1 \geq t + s | T_1 \geq s] = \mathbb{P}[T_1 \geq t]$
- Regularity: $\mathbb{P}[T_1 \geq s]$ is 'nice'

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

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

Chapter 2

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 2.1. A sequence $X_n, n \in \mathbb{N}$ of random variables with values in E is a (time homogeneous) Markov Chain (MC) if:

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

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

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

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

By convention when we write $\mathbb{P}\left[A|B\right]$ we assume $\mathbb{P}\left[B\right] > 0$.

Remark 2.1. The first condition is equivalent to for all $f : E \to \mathbb{R}$ bounded, $\mathbb{E}\left[f(X_{n+1})|X_0, \ldots, X_n\right] = \mathbb{E}\left[f(X_{n+1})|X_n\right]$

Example 2.1. If X_n are i.i.d. in E then X_n is a Markov Chain

Example 2.2. SRW on $\mathbb{Z}^d$

2.1 Transition Probabilities

Motivation In a finite state space $E = \{1, 2, 3\}$, then we write the probability to go from 1 to 2 as p_{12}. We would like to write these probabilities in a matrix.

Definition 2.2. A transition probability is a collection $p = (p_{x,y})_{x,y \in E}$ such that:

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

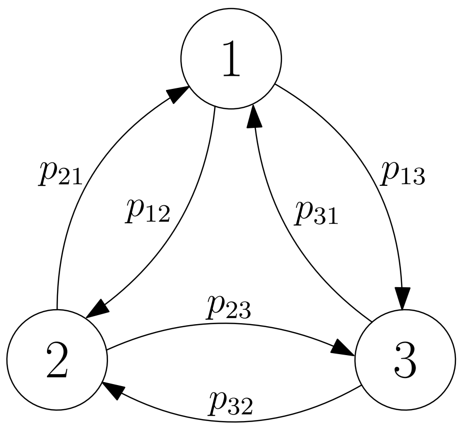
We could also represent this as a weighted directional graph with vertices E and weighted oriented edges: $\{(x, y) \in E : p_{x,y} > 0\}$. We know there is a 1 to 1 correspondence between directional graphs and matrices.

Matrix So 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 $\forall f \in L^\infty(E)$ define $Pf \in L^\infty(E)$ by $Pf(x) = \sum_{y \in E} P_{x,y} f(y)$ with $P \geq 0$ ($\forall f \geq 0 : Pf \geq 0$) and satisfies $P1 = 1$.

Definition 2.3. Let p be a transition probability, ψ 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 MC(ψ, p)) if for $x_0, \ldots, x_n \in E$

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



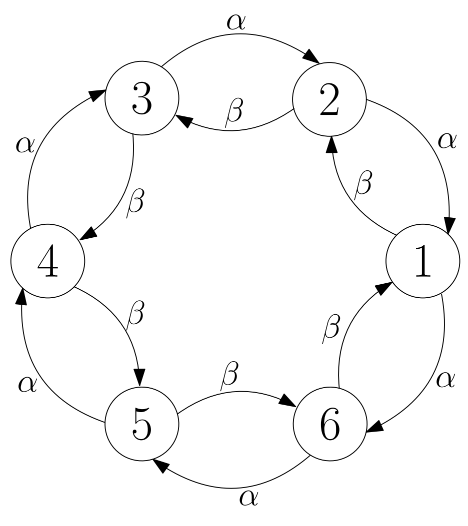


Figure 2.1: 3-State Markov Chain and 6-State (asymmetric) Markov Chain

Proposition 2.2. *Let $(X_n)_{n\geq 0}$ sequence of random variables with values in E:*

$$(X_n)_{n\geq 1} \text{ is a Markov Chain } \Leftrightarrow \exists \mu, p \text{ such that } (X_n)_n \text{ is a } MC(\mu, p).$$

Proof. $\implies$: If X_n is a Markov Chain, then set $p_{xy} = \mathbb{P}\left[X_{n+1} = y | X_n = x\right]$. $\sum_{y\in E} p_{xy} = 1$, as the conditional probability is a probability measure itself, and $p_{xy} \geq 0, \forall x, y \in E$ for the same reason. Thus we have that the collection of $(p_{xy})_{x,y\in E}$ form a transition probability. Setting $\mu(x) = \mathbb{P}\left[X_0 = x\right]$, 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 $x_0, \ldots, x_n \in E$, and every $n \geq 0$, we have

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

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

$\impliedby$: Here we have to show the two aspects of a Markov chain, the Markov Property and homogeneity. For homogeneity we have

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

here we have implicitly (sneakily) assumed that $\mathbb{P}\left[X_n = x\right] > 0$, as without this the conditional probability we are taking is not well-defined.

For the Markov Property we have

$$\begin{aligned}
\mathbb{P}\left[X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_n = X_n\right] &= \frac{\mathbb{P}\left[X_0 = x_0, \ldots, X_{n+1} = x_{n+1}\right]}{\mathbb{P}\left[X_0 = x_0, \ldots, X_n = x_n\right]} \\
&= \frac{\mu(x_0) p_{x_0x_1} \cdots p_{x_nx_{n+1}}}{\mu(x_0) p_{x_0x_1} \cdots p_{x_{n-1}x_n}} \\
&= p_{x_nx_{n+1}} = \mathbb{P}\left[X_{n+1} = x_{n+1} | X_n = x_n\right].
\end{aligned}$$

Where it is important to note that, again, we have implicitly assumed that $\mu(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n} > 0$.

□

Question Given μ, p does $\mathrm{MC}(\mu, p)$ always exist (as a Markov Chain)?

2.2 Existence

Theorem 2.3. *Let p be a transition probability on E. Then there exists:*

(i) a measurable space (Ω, F)

(ii) a collection of probability measures $(P_x)_x$ on (Ω, F)

(iii) a sequence of random variables $(X_n)_{n\geq 0}$ on (Ω, F) such that for all $x \in E$, under P_x, (X_n) is $MC(\delta_x, p)$

There are 2 approaches to prove this.

One could set $\Omega = E^{\mathbb{N}}$, $\mathbb{P}\left[(x_0, \ldots, x_n) | x \in E^n\right] = \delta_x(x_0)p_{x_0,x_1}\cdots p_{x_{n-1},x_n}$. Instead we will work as follows:

Proof. We consider a measure μ on E such that $\forall x \in E : \mu(x) > 0$ on some abstract probability space (this part is of technical relevance) $(\Omega, \mathcal{F}, \mathbb{P})$. Now we look at a random variable X_0 with distribution μ, and $U_1, U_2, \ldots$ independent, uniformly distributed, random variables on $[0, 1]$ (we know these exist from previous probability lectures). 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 $E = \{x_i, i \in I\}$ where I is our index set (eg. $\{1, 2, \ldots, n\}$ or $\mathbb{N}$) 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 $\Phi : E \times [0, 1] \to E; (x_i, u) \mapsto x_j$ if $u \in (s_{i,j}, s_{i,j+1}]$, a measurable function. 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.

What these tools have given us is that $\mathbb{P}\left[\Phi(x, U_1) = y\right] = p_{xy}$. So if we set $X_{n+1} = \Phi(X_n, U_{n+1})$ for every $n > 0$ (by induction), we find that

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

by independence.

Now if we define $\mathbb{P}_x$ as $\mathbb{P}\left[\,\cdot\,|X_0 = x_0\right]$, then we have $\forall x \in E$ $\mathbb{P}_x\left[X_0 = x_0, \ldots, X_n = x_n\right] = \delta_x(x_0)p_{x_0x_1}\cdots p_{x_{n-1}x_n}$.

□

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

For μ Prob measure on E we write $P_\mu = \sum_x \mu(x) P_x$

2.3 Simple Markov Property

Under P_μ $(X_n)_{n\geq 0}$ is MC(μ, p). $P_\mu[X_{n+1} = x_{n+1}|X_0 = x_0, \ldots, .X_n = x_n] = P_\mu[X_{n+1} = x_{n+1}|X_n = x_n] = P_{x_n}[X_1 = x_{n+1}]$ i.e. Conditional on $X_n = x$, x_{n+1} is sampled like the first step of a MC(δ_x, p) indep of the past.

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

Theorem 2.4 (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:*

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

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 $\mathbb{E}_\mu\left[f((X_{k+n})_{n\geq 0})|X_0 = x_0, \ldots, X_k = x_k\right] = \mathbb{E}_{x_k}\left[f((X_n)_{n\geq 0})\right]$. We will proceed using measure theoretic induction (see any book on measure theory). Approximate f by step functions f_k, using linearity, we only have to show our claim for the function $\mathbb{1}_A$ with $A \subset E^{\mathbb{N}}$, i.e.

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

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 our σ-algebra. Furthermore, on such sets

$$\begin{aligned}
&\mathbb{P}_\mu\left[(X_{k+n})_{n\geq 0} \in A|X_0 = x_0, \ldots, X_k = x_k\right] \\
&\quad = \mathbb{P}_\mu\left[X_k = y_0, \ldots, X_{k+N} = y_N|X_0 = x_0, \ldots, X_k = x_k\right] \\
&\quad = \frac{\mu(x_0)p_{x_0x_1}\cdots p_{x_{k-1}x_k}\delta_{x_k}(y_0)p_{y_0y_1}\cdots p_{y_{N-1}y_N}}{\mu(x_0)p_{x_0x_1}\cdots p_{x_{k-1}x_k}} \\
&\quad = \delta_{x_k}(y_0)p_{y_0y_1}\cdots p_{y_{N-1}y_N} \\
&\quad = \mathbb{P}_{x_k}\left[(X_n)_{n\geq 0} \in A\right].
\end{aligned}$$

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

Corollary 2.5. *Let μ be a distribution on E, $x \in E$, $k \in \mathbb{N}$, $\forall f : E^{\mathbb{N}} \to \mathbb{R}$ measurable and bounded:*

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

2.4 n-Step Transition Probabilities

Definition 2.4. For every $n \geq 0$, $x, y \in E$, define $p_{xy}^{(n)} = P_x[X_n = y]$

Proposition 2.6 (Chapman Kolmogorov (CK))**.**

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

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

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

□

Remark 2.7. If E is finite:

- The matrix $(p_{ij}^{(n)})_{ij \leq 0} = P^n$
- For μ a distribution on E the following holds for any $f : E \to \mathbb{R}$
$$\mathbb{E}_\mu\left[f(X_n)\right] = \mu P^n f,$$
for any $n \geq 0$, with $f = [f(1), \ldots, f(n)]^T$.

2.5 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 2.5. 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

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

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

Probabilistic interpretation If π is a stationary distribution, then $\forall n \geq 0 \ P_\pi[X_n = x] = \pi(x)$

Basically, 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.

2.6 Reversibility

Definition 2.6. 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 $\Leftrightarrow \forall n; \forall x_0, \ldots, x_n$: $\mathbb{P}_\pi[X_0 = x_0, \ldots, X_n = x_n] = \mathbb{P}_\pi[X_0 = x_n, \ldots, X_n = X_0]$.

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

Proposition 2.8. *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 2.3 (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 $Bin(N, \frac{1}{2})$.

2.7 Communication Classes

Here we will will see p as a weighted oriented graph.

Definition 2.7. Let $x, y \in E$. Write:

- $x \to y$ if $\exists n \geq 0$ st $p_{xy}^{(n)} > 0$ "y can be reached from x"
- $x \leftrightarrow y$ if ($x \to y$ and $y \to x$) "x and y communicate"

Proposition 2.9. *$\leftrightarrow$ is an equivalence class on E*

Proof. Let $x, y, z \in E$ and $m, n \geq 0$ such that $p_{xy}^{(m)} > 0$ and $p_{yz}^{(n)} > 0$.

(i) Transitivity: $p_{xz}^{(m+n)} \geq p_{xy}^{(m)} p_{yz}^{(n)} > 0$ thus we know that $x \to z$, we can apply the same argument in the other direction as well.

(ii) Reflexivity: $p_{xx}^{(0)} = \mathbb{P}_x[X_0 = x] = 1 > 0$, hence $x \leftrightarrow x$.

(iii) Symmetry: Trivial.

$\square$

Definition 2.8.

- The equivalence classes of $\leftrightarrow$ are called communication classes.
- The chain p is called irreducible if there is a unique communication class.

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

Definition 2.9. A communication class C is closed if for any $x, y \in E$

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

i.e. if you start in C you never leave.

2.8 Strong Markov Property

Definition 2.10. 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 for all $n \in \mathbb{N}$:

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

Example 2.4 (Stopping Times). $H_A = \min\{n \geq 0 : X_n \in A\}$ (for A measurable) and $H_x = \min\{n \geq 0 : X_n = x\}$ are stopping times.

Definition 2.11. 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 2.10 (Strong Markov Property (SMP)). *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:*

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

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

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

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

□

Application Reflection Principle for the SRW.

Consider the SRW on $\mathbb{Z}$:

Proposition 2.11. *Let $k \geq 0$ even, $a \geq 1$ odd:* $\mathbb{P}_0 \left[\max_{0\leq m\leq k} X_m \geq a \right] = \mathbb{P}_0 \left[|X_k| \geq a \right]$

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

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

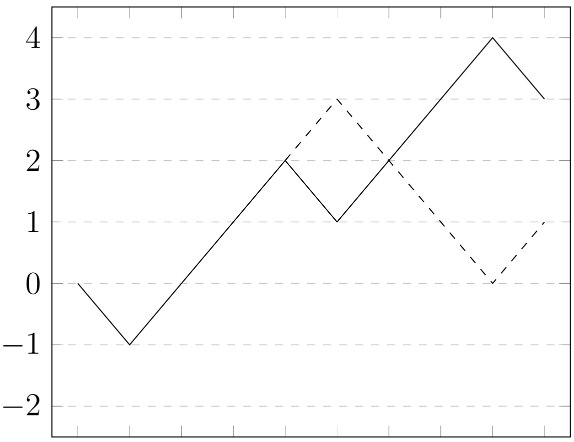


Figure 2.2: Example of a reflected simple random walk for $a = 2$

Now our goal is to show the term on the right is equal to $\mathbb{P}_0[X_k > a]$, as $2\mathbb{P}_0[X_k > a] = \mathbb{P}_0[|X_k| > a]$ 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'.

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

□

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 3

Markov Chains: Long Time Behavior

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?

3.1 Recurrence/Transience

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

Definition 3.1. Let $x \in E$, we say that:

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

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

Theorem 3.1 (Dichotomy Theorem)**.** *$x \in E$:*

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

Remark 3.2. It is impossible that $\mathbb{P}_x [V_x < \infty] > 0$ and $\mathbb{E}_x [V_x] = +\infty$.

Definition 3.2. $\rho_x = \mathbb{P}_x [H_x < \infty]$, 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 $\rho_x < 1$

Lemma 3.3. *For every* $i \geq 0, x \in E$*, we have* $\mathbb{P}_x\left[V_x \geq i\right] = \rho_x^i$.

Proof (Lemma). 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.

$$\begin{aligned}
\mathbb{P}_x\left[V_x \geq i+1\right] = \mathbb{P}_x\left[V_x \geq i+1 \wedge V_x \geq i\right] &= \mathbb{P}_x\left[H_x^{(i_1)} < \infty \wedge H_x^{(i)} < \infty\right] \\
&= \mathbb{P}_x\left[H_x^{(i+1)} < \infty | H_x^{(i)} < \infty, X_{H_x^{(i)}} = x\right] \mathbb{P}_x\left[H_x^{(i)} < \infty\right] \\
&\overset{\text{StMP}}{=} \mathbb{P}_x\left[H_x^{(1)} < \infty\right] \rho_x^i = \rho_x^{i+1}
\end{aligned}$$

□

Proof (Theorem). For x recurrent:

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

For x transient:

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

To justify (*) intuitively, we will write the values we are summing over in a table. In the sum on the left we sum over each row first (the inner sum), collect these values in a column, and then sum over that column (outer sum); meanwhile for the RHS we first sum over each column, collect these values in a row, and then sum over that row.

$$\begin{array}{lllll}
\mathbb{P}_x\left[V_x = 1\right] & & & & \sum = \sum_{j=1}^{1} \mathbb{P}_x\left[V_x = 1\right] \\
\mathbb{P}_x\left[V_x = 2\right] & \mathbb{P}_x\left[V_x = 2\right] & & & \sum = \sum_{j=1}^{2} \mathbb{P}_x\left[V_x = 2\right] \\
\mathbb{P}_x\left[V_x = 3\right] & \mathbb{P}_x\left[V_x = 3\right] & \mathbb{P}_x\left[V_x = 3\right] & & \sum = \sum_{j=1}^{3} \mathbb{P}_x\left[V_x = 3\right] \\
 & & & & \vdots \\
\vdots & \vdots & & & \sum_{k=1}^{\infty}\sum_{j=1}^{k} \mathbb{P}_x\left[V_x = k\right] \\
\sum_{k=1}^{\infty} \mathbb{P}_x\left[V_x = k\right] & \sum_{k=2}^{\infty} \mathbb{P}_x\left[V_x = k\right] & \dots & \sum_{j=1}^{\infty}\sum_{k=j}^{\infty} \mathbb{P}_x\left[V_x = k\right] &
\end{array}$$

Such tricks with sums will be used again. □

Proposition 3.4. *If* E *is finite, then there exists a recurrent state* $x \in E$.

Proof. Fix some $y \in E$.

$$\sum_{x\in E} V_x = \sum_{n=0}^{\infty}\sum_{x\in E} \mathbb{1}_{X_n=x} = \sum_{n\geq 0} 1 = \infty$$
$$\sum_{x\in E} \mathbb{E}_y\left[V_x\right] = \mathbb{E}_y\left[\sum_{x\in E} V_x\right] = \infty$$

Thus we know $\exists x \in E$ such that $\mathbb{E}_y\left[V_x\right] = \infty$ since the sum on the left is over a finite index set (E finite). Since we can write $V_x = V_x \mathbb{1}_{H_x<\infty}$, we find that (Strong Markov Property) $\infty = \mathbb{E}_y\left[V_x\right] = \mathbb{E}_y\left[V_x \mathbb{1}_{H_x<\infty}\right] = \mathbb{E}_x\left[V_x\right]\mathbb{P}_y\left[H_x < \infty\right]$, because a walk started from y is the same (in the distribution sense) after hitting x for the first time as a walk started from x. $\mathbb{P}_y\left[H_x < \infty\right]$ must be ≤ 1, thus the term of the left must be equal to $\infty \implies \mathbb{E}_x\left[V_x\right] = \infty$. □

3.2 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 3.5. *For the SRW, every state is recurrent if $d = 1, 2$, otherwise they are transient.*

3.3 Classification of States

Theorem 3.6. *Let $x, y \in E$ st $x \to y$. 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 $x \leftrightarrow y$.*

Remark 3.7. $x \neq y : x \to y \Leftrightarrow \mathbb{P}_x\left[\exists n : X_n = y\right] \Leftrightarrow \mathbb{P}_x\left[H_y < \infty\right]$

Corollary 3.8. *Let C communication class for p. Either $\forall x$: x is recurrent, or $\forall x$: x is transient.*

Corollary 3.9. *A recurrent class is always closed.*

3.4 Positive/Null Recurrence

Notation $x \in E : m_x = \mathbb{E}_x\left[H_x\right]$

Definition 3.3. Let $x \in E$ be a recurrent state. We say that:

- positive recurrent if $m_x < \infty$
- null recurrent if $m_x = +\infty$

Theorem 3.10. *Let* $x, y \in E, x \leftrightarrow y$. *Then* $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p_{xy}^{(k)} = \frac{1}{m_y}$

Remark 3.11. Write $V_y^{(n)} = \sum_{k=1}^{n} \chi_{X_k=y}$, "The number of visits to y up to time n". Thus the sum in the theorem is "Expected proportion of time spent at y".

If y is transient, 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 visit"

Definition 3.4 (inter-visit times). Let $y \in E$. Define $H_y^0 = H_y$ and $\forall i \geq 1 : H_y^i = min\{n \geq 1 : X_{H_y^0+\ldots+H_y^{i-1}+n} = y\}$ if $H_y^{i-1} < \infty$, else $+\infty$

Lemma 3.12. *Let* x, y *st.* $x \leftrightarrow y$, *assume* y *is recurrent. Then* $\forall j \geq 1, t_0 \ldots t_j \in \mathbb{N}$*:*

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

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

Proposition 3.13 (Classification of recurrent classes)**.** *Let* R *be a recurrent class. Then either:*

- $\forall x \in R : x$ *is positive recurrent*
- $\forall x \in R : x$ *is null recurrent*

Proposition 3.14. *Let* R *be a recurrent class, if* R *is finite, then* R *is positive recurrent.*

3.5 Stationary Distributions for Irreducible Chains

Theorem 3.15. *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* $\forall x \in E : \pi(x) = \frac{1}{\mathbb{E}_x[H_x]}$

3.6 Periodicity

Definition 3.5. Let $x \in E$. The period of x is defined by $d_x = gcd\{n \geq 0 : p_{xx}^{(n)} > 0\}$

Proposition 3.16. *Let* $x, y \in E : x \leftrightarrow y \implies d_x = d_y$

Consequence if p is irreducible we have $\forall x, y \in E : d_x = d_y$

Definition 3.6. We say that the chain p is aperiodic if $\forall x \in E : d_x = 1$

Proposition 3.17. *Let* $x \in E$. *We have* $d_x = 1 \Leftrightarrow \exists n_0 \geq 1$ *st* $\forall n \geq n_0 : p_{xx}^{(n)} > 0$

3.7 Coupling Method

What is coupling? Define probability measures μ_1, μ_2 on the same space F_1, F_2. A coupling between μ_1 and μ_2 is a probability measure $\overline{\mu}$ on $F_1 \times F_2$, $\overline{\mu}(A \times F_2) = \mu_1(A)$ and vice versa.

X_1, X_2 two random variables on $(\Omega, F, \mathbb{P})$, $X_1 \sim \mu_1, X_2 \sim \mu_2$, the law of (X_1, X_2) is a coupling!

Goal Define two MCs: $X_n \sim MC(\mu, p)$, $\quad \tilde{X_n} \sim MC(\nu, p)$ on the same probability space st $X_n = \tilde{X_n}$ for n large.

Definition 3.7 (Product Chain)**.** Define $\forall \omega = (x, y), \quad \omega' = (x', y') \in E^2$: $\overline{p_{\omega,\omega'}} = p_{xx'}p_{yy'}$

Notation Consider:

- $(\Omega, F, (P_\omega)_{\omega \in E^2})$ Probability Spaces
- $(W_n)_{n \geq 0} = ((X_n, Y_n))_{n \geq 0}$ RV on Ω, F st $\forall \omega \in E^2 : W_n$ is a $MC(\delta_\omega, \overline{p})$ under P_w

Remark 3.18. If μ, ν are distributions on E, then $\mu \otimes \nu$ is a distribution on E^2. $P_{\mu \otimes \nu} = \sum_{(x,y) \in E^2} \mu(x)\nu(y)P_{(x,y)}$

Proposition 3.19. *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* $MC(\nu, p)$

Proposition 3.20. *If p is irreducible and aperiodic, then $\overline{p}$ is irreducible and aperiodic.*

Remark 3.21. Aperiodic is important! p irreducible $\not\Rightarrow \overline{p}$ irreducible.

Proposition 3.22. *If p is irreducible, aperiodic, and positive recurrent, then $\overline{p}$ is irreducible, aperiodic, and positive recurrent.*

Definition 3.8. $T = min\{n \geq 0 : X_n = Y_n\}$ a stopping time.

Proposition 3.23. $\forall \mu, \nu$ *distributions on E:*

$$\forall n \geq 0 \sum_{x \in E} |\mathbb{P}_\mu [X_n = x] - \mathbb{P}_\nu [Y_n = x]| \leq 2\mathbb{P}_{\mu \otimes \nu} [T > n]$$

Lemma 3.24. $\tilde{X_n} = Y_n \chi_{\{T<n\}} + X_n \chi_{\{T \geq n\}}$ *is a* $MC(\nu, p)$

3.8 Convergence for Irreducible Aperiodic Chains

Theorem 3.25. *Assume p is irreducible and aperiodic, and admits a stationary distribution π. Then for every distribution μ on E :* $\lim_{n \to \infty} \mathbb{P}_\mu [X_n = x] = \pi(x), \forall x \in E$.

Equivalently: Under $P_\mu : X_n \overset{(law)}{\to} X_\infty$ *where* $X_\infty \sim \pi$

Equivalently: $\forall f : E \to \mathbb{R}$ *bdd:* $\lim_{n \to \infty} \mathbb{E}_\mu [f(X_n)] = \int_E f d\pi$

Note This theorem is important!!

Theorem 3.26. *Assume that p is irreducible, aperiodic, and null recurrent or transient. Then for every distribution μ and every $x \in E$:* $\lim_{n\to\infty} \mathbb{P}_\mu [X_n = x] = 0$

Lemma 3.27. *$\overline{p}$ irreducible and recurrent, then $\forall \mu$ distribution on E : $\forall i \geq 0, \forall x \in E$:* $\lim_{n\to\infty} |\mathbb{P}_\mu [X_n = x] - \mathbb{P}_\mu [X_{n+i} = x] | = 0$

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 4

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', ie 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 $(A_t,\ a(t) = \mathbb{E}[A_t])$.

Case 1: $T_1 ... \sim Exp(\lambda)$: $m(t) = t\lambda$, $E_t \sim Exp(\lambda), e(t) = \frac{1}{\lambda}$, $A_t \sim Exp(\lambda)$.

Case 2: More complicated.

4.1 Definition and First Properties

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

Definition 4.1. The continuous 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.

Example 4.1. (i) $pp(\lambda), \lambda > 0, T_i \sim Exp(\lambda)$

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

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

Proposition 4.1. *$N = (N_t)_{t\geq 0}$ is a counting process with jump times $S_1, S_2, ...$ and $\lim_{t\to\infty} N_t = +\infty$.*

Proposition 4.2. *There exists* $c > 0$ *st* $\forall t \geq 0 : \mathbb{E}\left[e^{cN_t}\right] \leq e^{\frac{1+t}{c}}$*, thus the expectation is finite* $\forall t$.

Theorem 4.3 (Law of Large Numbers)**.** *We have* $\lim_{t\to\infty} \frac{N_t}{t} = \frac{1}{\mu}$.

4.2 Renewal Function

Definition 4.2. The renewal function is defined by $\forall t \geq 0 : m(t) = \mathbb{E}[N_t]$.

Remark 4.4. $m(t) < \infty$ because N_t has exponential moment (you can use Jensen).

Proposition 4.5. $m(t)$ *is non-decreasing, non-negative, and right continuous.*

Theorem 4.6 (Elementary Renewal Theorem)**.** $\lim_{t\to\infty} \frac{m(t)}{t} = \frac{1}{\mu}$

4.3 Blackwell's Renewal Theorem

Definition 4.3. We say the law of T_1 is arithmetic if $\exists a > 0 : \mathbb{P}[T_1 \in a\mathbb{Z}] = 1$. It is non-arithmetic if this probability is < 1.

Theorem 4.7 (Blackwell)**.** *Assume that the law of* T_1 *is non-arithmetic, then* $\lim_{t\to\infty} m(t+h) - m(t) = \frac{h}{\mu}$.

Remark 4.8. $\frac{m(t)}{t} \approx \frac{m(\lfloor t\rfloor)}{\lfloor t\rfloor} = \frac{1}{\lfloor t\rfloor}\sum_{k=1}^{\lfloor t\rfloor} m(k) - m(k-1) \overset{Blackwell}{\rightarrow} \frac{1}{\mu}$. "Blackwell is stronger than elementary renewal."

4.4 Renewal Equation

Lebesgue-Stieltjes Integral

Notation $\mathcal{M} = \{f : \mathbb{R}_+ \to \mathbb{R}_+, \text{right continuous, non-decreasing}\}$ 'measures on $\mathbb{R}_+$ '. $\nu((a,b]) = f(b) - f(a)$

For all $h \in L^1(df)$ or $h \geq 0$ meas, we can define $\int h df$.

Example 4.2.

- $m \in \mathcal{M} \to \int hdm$ can be defined
- If T is a RV on $\mathbb{R}_+$ the $F_T(t) = \mathbb{P}[T \leq t]$

Definition 4.4. Let $G \in \mathcal{M}$. Let $h : \mathbb{R}_+ \to \mathbb{R}$ st either $\forall t : \int_0^t |h(t-s)|dG(s) < \infty$ or $h \geq 0$ a.e. we define:

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

Remark 4.9. Let X, Y be two indep RV on $\mathbb{R}_+$. Then with F_X, F_Y their respective cdf's:

$$\begin{aligned}\mathbb{P}[X+Y \le t] &= \int_{s=0}^{t} \mathbb{P}[X+s \le t]\, dF_y(s) \\ &= \int_0^t F_X(t-s) dF_Y(s)\end{aligned}$$

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

Why is this useful?

$$\begin{aligned}m(t) =& \mathbb{E}[N_t] = \mathbb{E}\left[\sum_n \mathbb{1}_{T_1+\ldots T_n \le t}\right] \\ =& \sum_n F_{T_1+\ldots+T_n}(t) = F^{*n}(t).\end{aligned}$$

Renewal Equation

Definition 4.5. Let $h : \mathbb{R}_+ \to \mathbb{R}$ meas. loc. bdd, $g : \mathbb{R}_+ \to \mathbb{R}$ st $\forall t \ge 0 : \int_0^t |g(t-s)| dF(s) < \infty$. We say that g is a solution of the (h, F) renewal equation if:

$$\forall t \ge 0 : g(t) = h(t) + \int_0^t g(t-s) dF(s)$$

Proposition 4.10 (First Example)**.** *m is a solution of the (F, F) renewal equation, ie $m = F + m * F$.*

Example 4.3 (Excess Time, 2nd Example). $E_t = S_{N_{t+1}} - t$, the time left to wait until next renewal. Define for $x \ge 0$, $e_x(t) = \mathbb{P}[E_t \le x]$. We can separate e_x into 2 parts, one for the probability if there has already been a renewal before time t, and one if that hasn't occured: $e_x(t) = \mathbb{P}[T_1 > t, E_t \le x] + \mathbb{P}[T_1 \le t, E_t \le x] = A + B$.

$A = \mathbb{P}[T_1 > t, T_1 \le t + x] = F(t+x) - F(t)$. Observe that E_t is meas wrt $T_1, T_2, \ldots$. $E_t = \phi_t(T_1, T_2, \ldots)$.

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

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.

Exercise Show that the age $a_x(t) = \mathbb{P}[A_t \le x]$ is the solution to some (h, F) renewal equation.

Well-Posedness of the Renewal Equation

Theorem 4.11. *Let $h : \mathbb{R}_+ \to \mathbb{R}$ meas, loc bdd. Then there exists a unique $g : \mathbb{R}_+ \to \mathbb{R}$ meas, loc bdd, solution of $g = h + g * F$ given by $g = h + h * m$.*

Intuitive Proof. Assume g is a solution.

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

We must only show that $(*)$ can be made rigorous. Otherwise this is just an intuitive proof, we can use this as a way to find a candidate for g, and then prov that it is actually a legitimate solution as follows. □

Rigorous Proof. $g = h + h * m$ is meas. loc. bdd., because h is. We have $h + g * F = h + (h + h * m) * F = h + h * F + h * m * F = h + h * (F + m * F) = h + h * m = g$

Uniqueness g_1, g_2 are 2 solutions, 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 $\mathbb{P}\left[T_1 + ... + T_n \leq t\right]$ which converges to 0. □

4.5 Asymptotic Behavior

Motivation We want to study the behavior of $g(t)$ when t is large and when g is a solution to the (h, F) renewal equation.

Case 1 $h = \mathbb{1}_{[a,b]}$, and g a solution. $g(t) = h(t) + \int_0^t h(t - s) dm(s)$. $h(t - s) = \mathbb{1}_{[a,b]}(t - s) = \mathbb{1}_{s \in [t-b, t-a]}$. So $g(t) = h(t) + m(t - a) - m(t - b)$ and with Blackwell's Theorem we find that this tends towards $0 + \frac{b-a}{\mu}$. Now we need to figure out how this generalizes.

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 4.6. $h : \mathbb{R}_+ \to \mathbb{R}_+$ meas., h is directly Riemann Integrable (dRi) if $\forall \Delta > 0 : \sum_{k=0}^{\infty} \Delta sup_{[k\Delta, (k+1)\Delta]} h < \infty$ 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 iff h_+ and h_- are dRi. See notes for example for integrable but not dRi function.

Proposition 4.12. *Let $h : \mathbb{R}_+ \to \mathbb{R}$. Assume that h is continuous at a.e. $t \in \mathbb{R}$, $\exists H$ non-decreasing st $0 \leq |h| \leq H$ and $\int_0^\infty H < \infty$, then h is dRi.*

Theorem 4.13 (Smith Key Renewall Theorem)**.** *Let h be dRi, F non-arithmetic. Then $g = h + h * m$ satisfies $lim_{t\to\infty} g(t) = \frac{1}{\mu}\int_0^\infty h(u)du$.*

Remark 4.14. 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 $h_{c,\Delta} = \sum_{k\geq 0} c_k \mathbb{1}_{[k\Delta,(k+1)\Delta)}$.

Application Let $\mu < \infty$. Let E_t be the excess time (time until next renewal) and $e_x(t) = \mathbb{P}\left[E_t \leq x\right]$. 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 4.15. $\mu = \mathbb{E}\left[T_1\right] = \int_0^\infty \mathbb{P}\left[T_1 > t\right] dt$

With this we have that $h_x(t) \leq 1 - F(t) = \mathbb{P}\left[T_1 > t\right]$, and $1 - F(t)$ is non-increasing in t and continuous ae (because it is the difference of two monotone functions). $\int_0^\infty \mathbb{P}\left[T_1 > t\right] dt = \mathbb{E}\left[T_1\right] = \mu < \infty$. So (by the proposition) h_x is dRi. Now we can apply the theorem and get that $lim_{t\to\infty}\mathbb{P}\left[E_t \leq x\right] = \frac{1}{\mu}\int_0^\infty h_x(t)dt = \frac{1}{\mu}\int_0^\infty F(t+x) - F(t)dt$, 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 $\frac{1}{\mu}\int_0^\infty \mathbb{E}\left[\mathbb{1}_{T_1\in(t,t+x]}\right] dt = \frac{1}{\mu}\mathbb{E}\left[\int_0^\infty \mathbb{1}_{t\in[T_1-x,T_1)}\right] dt = \frac{1}{\mu}\mathbb{E}\left[\int_{max\{T_1-x,0\}}^{T_1} dt\right] = T_1$ if $T_1 \leq x$ and x if $T_1 > x$. Thus we get for t large: $\mathbb{P}\left[E_t \leq x\right] \approx \frac{1}{\mu}\mathbb{E}\left[min\{T_1, x\}\right]$.

Remark 4.16. $G(x) = \frac{1}{\mu}\mathbb{E}\left[min\{T_1, x\}\right]$ 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 5

General Poisson Point Processes

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

5.1 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 $B_i \uparrow E \Leftrightarrow B_1 \subset B_2 \subset ... : \bigcup_{i \geq 1} B_i = E$.

Example 5.1. 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 $\mathbb{R}_+$.

5.2 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 5.1. *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 5.1. 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 5.2. 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.

5.3 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 5.2. 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 $N(B) \sim Pois(\mu(B))$.

5.4 Existence and Uniqueness

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

Spaces with finite measure

Proposition 5.2. *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 5.3. *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 5.4. *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 5.5. 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 5.6. *Let* N, N' *be two ppp*(μ) *on* $(E, \mathcal{E})$ *then* $P_N = P_{N'}$.

Theorem 5.7 (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:* $X_n \in E, n \geq 1 : N = \sum_{i=1}^{\tau(}$

5.5 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 5.8. $X \sim Pois(\lambda), \lambda > 0$*, then* $\forall u \geq 0 : \mathbb{E}\left[e^{-uX}\right] = exp(-\lambda(1 - e^{-u}))$.

Definition 5.3. 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 5.9. $\int_E u(x)N(dx) = \int_E u dN$ is a RV.

Theorem 5.10 (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:* $L_N(u) = exp(-\int_E 1 - e^{-u(x)}\mu(dx))$

5.6 Simple Processes

Remark 5.11. For $x \in E$, $\{x\}$ is meas. because E is Polish.

Definition 5.4. A measure $\eta \in \mathcal{N}$ is said to be simple if $\forall x \in E : \eta(\{x\}) \leq 1$.

Proposition 5.12. $\{\eta : \eta$ *is simple*$\}$ *is measurable in* $\mathcal{N}$.

Theorem 5.13. *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.

5.7 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 5.14. *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 5.3. $E = \mathbb{R},\ F = \mathbb{Z},\ T : E \to F; x \to \lfloor x \rfloor,\ \mu = \mathcal{L},\ T\#\mu = |.|$.

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

Theorem 5.15 (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 $\mu_{C_1}, \mu_{C_2}, ...$*

5.8 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 5.5. 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 5.16. X_i corresponds to the position of the cars in Case 2, and Y_i to their speeds.

Theorem 5.17. *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.

5.9 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).

5.10 Exponential Random Variables

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

Definition 5.6. Let $\lambda > 0$, a real RV T is exponential with parameter λ (we write $T \sim Exp(\lambda)$) if it has density $f(t) = \lambda e^{-\lambda t}\chi_{\{t\geq 0\}}$. $\Leftrightarrow \forall t \geq 0 \mathbb{P}[T > t] = e^{-\lambda t}$

Proposition 5.18 (Memoryless Property)**.** *Let $\lambda > 0$ and $T \sim Exp(\lambda)$. Then $\forall s, t \geq 0$: $\mathbb{P}[T > s + t | T > t] = \mathbb{P}[T > s]$*

Proposition 5.19 (Minimum of indep Exponentials)**.** *Let $n \geq 0, T_1...T_n$ indep with $T_i \sim Exp(\lambda_i), \lambda_i > 0$:*

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

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: $\mathbb{E}[\phi(X, Y)] = \int_0^\infty \mathbb{E}[\phi(x, Y)] f(x)dx$

Proposition 5.20 (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 $\Gamma(1, t) = Exp(\lambda)$

5.11 Definition of Poisson Processes

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

Definition 5.7. 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 $\mathbb{P}[N_t = n] = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$

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

5.12 Markov Property

Theorem 5.21 (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 $(N_u)_{0\leq u\leq t}$.*

5.13 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 5.8. 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 5.22 (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 $t_1 ... t_k$.

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

5.14 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 5.9. 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 5.4. $pp(\lambda)$ is a counting process.

Remark 5.23. 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 5.24. *Let $\lambda > 0 :$ Let N be a counting process, the following are equivalent:*

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

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

Remark 5.25. By def N is a $pp(\lambda) \Leftrightarrow N$ is a counting process with jumps of size 1 a.s. and $S_1, S_2 - S_1, ...$ are iid $exp(\lambda)$.

5.15 Microscopic Characterization

Theorem 5.26. *Let N be a counting process, let $\lambda > 0$. TFAE:*

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

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

5.16 Properties of Poisson Process

Theorem 5.27 (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 5.28 (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 5.29. 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 6

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 $\mathbb{P}\left[X_{t+h} = y \mid X_t = \right.$ for h small the pr equal to 1.

6.1 Definition via Generator

Definition 6.1. 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 6.2. 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 6.3. 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 6.4. 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 6.5. 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 6.1. 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 6.1 (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?

6.2 Non-Rigorous Section: The Constructive Approach

Example 6.2 (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 6.3 (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 $T_1 = min\{T_{12}, T_{13}\} \sim Exp(\alpha+\beta)$. $\mathbb{P}\left[\text{jump from } 1 \to 2\right] = \mathbb{P}\left[T_1 = T_{12}\right] = \frac{\alpha}{\alpha+\beta} = \frac{q_{12}}{q(1)}$. The skeleton $(\overline{X_n})$ is a Discrete time MC with transition probabilities $\kappa_{xy} = \frac{q_{xy}}{q(x)}$.

6.3 Definition by Skeleton and Holding Time

Note q is a fixed generator.

Discrete Chain Associated to 2

Definition 6.6. 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 6.2. κ is transition probability (check for the cases where $q(x) = 0$ and $q(x) \neq 0$).

Example 6.4. (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 6.7. We say that q has no explosion if $\forall$ choice of $\mu : S_\infty = +\infty$ a.s.

Remark 6.3. 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 6.4. $pp(\lambda)$ is a birth chain with $\lambda_i = \lambda$.

Theorem 6.5. *The birth chain q has no explosion $\Leftrightarrow \sum_{i\geq 1} \frac{1}{\lambda_i} = \infty$.*

Non-Explosion Characterization

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

Theorem 6.6. *For $x \in E$, let $Y = (Y_n^{(x)})_{n\geq 0}$ be a $MC(\mu, \kappa)$. Then q has no explosion $\Leftrightarrow \forall x \sum_{n\geq 0} \frac{1}{q(Y_n^{(x)})} < \infty$ a.s.*

Remark 6.7. $\sum_{n\geq 0} \frac{1}{q(Y_n)}$ 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 6.8 (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)

6.4 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.

6.5 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 $x, y \in E$: $p_{xy}(t) = \mathbb{P}_x[X_t = y]$. In the discrete case this corresponds to $p_{xy}^{(n)} = p_{xy}(t)$.

Remark 6.9. We have

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

Proposition 6.10 (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 6.11 (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 6.12. *If E is finite, we have $\forall t \geq 0 : P(t) = exp(tQ)$.*