1 Conditional Expection

We start with some preliminaries, things known or less known.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space

Definition 1.1. *We define the Lebesgue spaces*
$L^p(\Omega, \mathcal{F}, \mathbb{P}) = L^p = \{X : \Omega \to \mathbb{R} \mid X \text{ is } \mathcal{F} \text{ measurable and } \mathbb{E}(|X|^p) < \infty\}$
where $1 \leq p < \infty$.
$L^\infty = \{X : \Omega \to \mathbb{R} \mid X \text{ is } \mathcal{F} \text{ measurable and } \exists C > 0 : \mathbb{P}(|X| \leq C) = 1\}$ *is the space of almost surely bounded random variables.*

Theorem 1.1 (Hölder's inequality). *Let $X \in L^P, Y \in L^q$ for conjugate $p, q \geq 1$ (i.e. $1/p + 1/q = 1$) then we have $XY \in L^1$ moreover Hölder's inequality holds true*

$$\mathbb{E}(|XY|) \leq \mathbb{E}(|X|^p)^{\frac{1}{p}} \mathbb{E}(|X|^q)^{\frac{1}{q}}$$

Remark 1.1.

1. Hölder's inequality is the reason why probability theory behaves "nicer" in consideration of Lebesgue spaces, i.e. let $0 < r < s$ and set $p = \frac{s}{r}$ then for the conjugate $q = \frac{p}{p-1}$ we can apply Hölder's inequality to $|X|^r$ and the constant 1 function.

$$\mathbb{E}(|X|^r) \leq \mathbb{E}(|X|^{rp})^{1/p} = \mathbb{E}(|X|^s)^{r/s} \implies \mathbb{E}(|X|^r)^{\frac{1}{r}} \leq \mathbb{E}(|X|^s)^{\frac{1}{s}}$$

   in particular if X is a random variable such that $X \in L^s$ then it always follows that $X \in L^r$ for any $r < s$. In general measure theory, where the spaces don't need to be of finite measure, this is not the case at all.

2. It is an important Theorem of Riesz that states that for all $1 \leq p \leq \infty$ the L^p spaces are complete, in particular they are Banach spaces. Especially for $p = 2 = q$ the space L^2 is even a Hilbert space.

Proposition 1.1 (Jensen's inequality). *Let $X \in L^1$ and $\varphi : \mathbb{R} \to \mathbb{R}$ a convex function, then we have*

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X))$$

Proposition 1.2 (Existence of the conditional expection). *Let $\mathcal{G} \subset \mathcal{F}$ be a subsigma-algebra. We define $Z = \mathbb{E}(Z \mid \mathcal{G})$ to be the unique $\mathcal{G}$-measurable variable such that*

$$\mathbb{E}(Z 1_A) = \mathbb{E}(X 1_A) \text{ for all } A \in \mathcal{G}$$

2   Discrete time Martingales

Definition 2.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, an increasing sequence of σ-algebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}$ is called a filtration on Ω. We call the space $(\Omega, \mathcal{F}, (\mathcal{F})_{n \in \mathbb{N}}, \mathbb{P})$ a filtered probability space.*

Definition 2.2. *A stochastic process $X = (X_n)_{n \in \mathbb{N}}$ is called adapted if X_n is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$. Moreoever, an adapted process is called a martingale if $X_n \in L^1$ for all $n \in \mathbb{N}$ and*

$$\mathbb{E}(X_n \mid \mathcal{F}_m) = X_{n \wedge m} \text{ for all } n, m \geq 0$$

Remark 2.1. Analogeously we define sub-martingales if instead of an equality we have $\geq$ above, or a super-martingale for $\leq$ respectively.

Definition 2.3. *A stopping time T is a random variable $T : \Omega \to \mathbb{N}_0^\infty$ such that $\{T \leq n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$. The sigma-algebra at the stopping time T is then defined as*

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for all } n \in \mathbb{N}\}$$

Remark 2.2. The sigma algebra at the stopping time T $\mathcal{F}_T$ encodes the information up to the random time T. In other words, if we interprete our filtered probability space as an random experiment, then the maximum information that can be found until the random time T is in $\mathcal{F}_T$.

Theorem 2.1 (Optional stopping theorem - weak version)**.** *Let X be a martingale. Let $X^T := (X_n^T)_{n \in \mathbb{N}}$ be given by $X_n^T := X_{n \wedge T}$ then X^T is also a martingale. Moreover if S, T are almost surely finite stopping times with $S \leq T$ almost surely, then*

$$\mathbb{E}(X_T \mid \mathcal{F}_S) = X_S$$

Theorem 2.2 (Martingale convergence Theorem)**.** *Let X be a super-martingale that is bounded in L^1 (i.e. $\sup_{n \in \mathbb{N}} \mathbb{E}(|X_n|) < \infty$). Then there exists a random variable $X_\infty \in L^1$ such that $X_n \to X_\infty$ almost surely, moreover*

$$\mathbb{E}(|X_\infty|) \leq \sup_{n \in \mathbb{N}} \mathbb{E}(|X_n|)$$

Corollary 2.1. *Let X be a super-martingale that is bounded from below (i.e. $X_n \geq c$ a.s. for some $c \in \mathbb{R}$ for all $n \in \mathbb{N}$). Then X_n converges almost surely to $X_\infty \in L^1$.*

Theorem 2.3 (Behaviour of Martingales with bounded increments)**.** *Let $X = (X_n)_{n \in \mathbb{N}}$ be a martingale with bounded increments (i.e. $|X_{n+1} - X_n| \leq C$). Let $\mathcal{C} := \{\limsup X_n = \liminf X_n \in \mathbb{R}\}$, $\mathcal{O} := \{\limsup X_n = +\infty\} \cap \{\liminf X_n = -\infty\}$, then $\mathbb{P}(\mathcal{O} \cup \mathcal{C}) = 1$.*

Lemma 2.1 (Doob's inequalities). *Let X be a non-negative submartingale and $X_n^* := \max_{k \leq n} X_k$ for $n \in \mathbb{N}$, then we have for any $\lambda > 0$*

$$\lambda \mathbb{P}(X_n^* > \lambda) \leq \mathbb{E}(X_n 1_{X_n^* > \lambda}) \leq \mathbb{E}(X_n)$$

In addition, for $p > 1$, we have

$$\|X_n^*\|_p \leq \frac{p}{p-1} \|X_n\|_p$$

Theorem 2.4 (Closed martingale convergence theorem). *Let X be a martingale and $p > 1$. Then the following statements are equivalent:*

1. *$\sup_{n \in \mathbb{N}} \|X_n\|_p < \infty$*

2. *X_n converges almost surely and in L^p to $X_\infty \in L^p$*

3. *There exists a random varaible $X_\infty \in L^p$ such that*

$$\mathbb{E}(X_\infty \mid \mathcal{F}_n) = X_n \text{ for all } n \in \mathbb{N}$$

If any (and consequently all) of these conditions hold true, we say that X is a closed martingale in L^p.

Definition 2.4. *A family $\mathcal{H}$ of random variables is called uniformly integrable (UI) if*

$$\lim_{\lambda \to \infty} \sup_{X \in \mathcal{H}} \mathbb{E}(|X| 1_{|X| > \lambda}) = 0$$

Exercise: Prove that $\mathcal{H}$ is UI if there exists $G : [0, \infty) \to [0, \infty)$ non-decreasing such that

1. $\lim_{t \to \infty} \frac{G(t)}{t} = \infty$

2. $\sup_{X \in \mathcal{H}} \mathbb{E}(G(X)) \infty$

Theorem 2.5 (UI convergence Theorem). *Let $(X_n)_{n \in \mathbb{N}}$ a stochastic process, then the following statements are equivalent:*

1. *$(X_n)_{n \in \mathbb{N}}$ is UI and $X_n \to X_\infty$ in probability.*

2. *$X_n \to X_\infty$ in L^1.*

Theorem 2.6 (Characterisation of UI martingales through closedness). *Let X be a martingale. The following assertions are equivalent:*

1. *X is UI.*

2. *X converges almost surely and in L^1 to $X_\infty \in L^1$.*

3. *There exists $X_\infty \in L^1$ such taht $\mathbb{E}(X_\infty \mid \mathcal{F}_n) = X_n$, i.e. the martingale is closed.*

Theorem 2.7 (Optional stopping - strong version). *Let X be a closed (and thus UI) martingale. Then for any stopping times S and T, one has*

$$\mathbb{E}(X_T \mid \mathcal{F}_S) = X_{T \wedge S}$$

Definition 2.5. *Let $\mathcal{F}$ be a sigma-algebra and consider a sequence $\cdots \subset \mathcal{H}_2 \subset \mathcal{H}_1 \subset \mathcal{F}$ of sigma-algebras (i.e. we lose information over time). Let $Y_1 \in L^p(\Omega, \mathcal{H}_1, \mathbb{P})$ for some $p \in [1, \infty)$. Then $Y = (Y_n)_{n \in \mathbb{N}}$ is called a backward-martingale if $Y_n = \mathbb{E}(Y_1 \mid \mathcal{H}_n)$ for all $n \in \mathbb{N}$.*

Theorem 2.8 (Convergence theorem for backward-martingales)**.**
Let $Y = (Y_n)_{n \in \mathbb{N}}$ be a backward-martingale. Then Y converges almost surely and in L^p to $Y_\infty = \mathbb{E}(Y_1 \mid \mathcal{H}_\infty)$ where $\mathcal{H}_\infty = \cap_{n \in \mathbb{N}} \mathcal{H}_n$.

Theorem 2.9 (Kolmogorov's Law of Large Numbers)**.** *Let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of i.i.d. random variables with $\xi_1 \in L^1$. Let $S_n = \xi_1 + \cdots + \xi_n$ be the simple random walk, then we have*

$$\frac{S_n}{n} \to \mathbb{E}(X_1) \text{ almost surely and in } L^1$$

Proof. **Exercise.** Kolmogorov's LLN can be shown in it's full generality (i.e. for L^1 RV) by using backwards martingales and Kolomogorov's 0-1 Law. $\qquad\square$

Theorem 2.10 (Kolmogorov's 0-1 Law)**.** *Let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of i.i.d. random variables and let $\mathcal{H}_n = \sigma(\xi_j : j \geq n)$. Then, the sigma algebra $\mathcal{H}_\infty := \cup_{n \in \mathbb{N}} \mathcal{H}_n$ is trivial, i.e. for every event $A \in \mathcal{H}_\infty$ we have $\mathbb{P}(A) \in \{0, 1\}$.*

Next we will remind the very important Borel-Cantelli Lemmas. We recall the following definiton. Let $(A_n)_{n \in \mathbb{N}}$ be a sequence of events in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ then we define

$$A = \limsup_{n \to \infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{i \geq n}^{\infty} A_i = \{A_n \text{ infinitely often}\}$$

$$A^c := \liminf_{n \to \infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{i \geq n}^{\infty} A_i = \{A_n \text{ eventually}\}$$

We remark that the following is true

$$\omega \in \bigcap_{n=1}^{\infty} \bigcup_{i \geq n}^{\infty} A_i \iff \forall n \in \mathbb{N}, \exists i \geq n : \omega \in A_i$$

$$\omega \in \bigcup_{n=1}^{\infty} \bigcap_{i \geq n}^{\infty} A_i^c \iff \exists n \in \mathbb{N}, \forall i \geq n : \omega \in A_i$$

Theorem 2.11 (Borel-Cantelli Lemma 1)**.** *If $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) < \infty$, then we have $\mathbb{P}\{A_n \ i.o.\} = 0$.*

Theorem 2.12 (Borel-Cantelli Lemma 2)**.** *If the events A_n are independent and $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) = \infty$, then we have $\mathbb{P}\{A_n \ i.o.\} = 1$*

Theorem 2.13 (Extension of Borel Cantelli Lemma)**.** *Let $\mathcal{F} = (\mathcal{F}_n)_{n=1}^{\infty}$ be a Filtration and suppose $A_n \in \mathcal{F}_n$, then we have*

$$\{A_n \ i.o.\} = \left\{ \sum_{n \geq k} \mathbb{P}(A_n \mid \mathcal{F}_{n-1}) = \infty \right\}$$

for all $k \in \mathbb{N}$ up to a set of measure zero.

We now give some results that are related to the Central Limit Theorem (CLT). We will first recall the classical result.

Theorem 2.14 (CLT)**.** *Let $X_1, X_2, \ldots$ be a sequence of i.i.d. L^2 random variables with $\mathbb{E}(X) = \mu$ and $\mathrm{Var}(X) = \sigma > 0$. Let $S_n = X_1 + \cdots + X_n$, then we have*

$$\frac{S_n - n\mu}{n\sigma} \implies \mathcal{N}(0,1)$$

It is our goal to generalize this result, i.e. weaken the conditions and then give a central limit theorem for martingales.

The outline is the following: Let $\{X_{n,k} : 1 \leq k \leq J(n), n \in \mathbb{N}_{\geq 1}\}$ be an array of centered random variables. Here K can be random, however it is almost surely finite with $J(n) \to \infty$ as $n \to \infty$.

Let $(\xi_j)_{j \in \mathbb{N}}$ be a sequence of (centered?) i.i.d. random variables, we then set

$$X_{n,k} = \frac{\xi_k}{\sqrt{n}}$$

Suppose that $X_{n,k} \in L^2(\mathcal{F})$ and let

$$S_n := \sum_{k=1}^{J(n)} X_{n,k}$$

Theorem 2.15 (Mc Leish)**.** *In the situation as above, let us define $T_n = \prod_{k=1}^{J(n)}(1 + itX_{n,k})$ and suppose that*

1. *T_n is UI and $\mathbb{E}(T_n) \to 1$ as $n \to \infty$.*

2. *$\sum_{k=1}^{J(n)} X_{n,k}^2 \to 1$ in Probability.*

3. *$\max_{k \leq n} |X_{n,k}| \to 0$ in Probability.*

Then $S_n \implies \mathcal{N}(0,1)$ as $n \to \infty$.

Theorem 2.16 (Martingale Central Limit Theorem)**.** *Let $X_{n,k}$ where $1 \leq k \leq m_n$ (m_n is deterministic with $m_n \to \infty$ as $n \to \infty$) be a $\mathcal{F}_{n,k}$-measurable martingale difference array. Suppose that*

1. $\mathbb{E}(X_{n,k} \mid \mathcal{F}_{n,k-1}) = 0$

2. $V_n := \sum_{j=1}^{m_n} X_{n,j}^2 \to 1$ *in Probability.*

3. *(Lindenberg Condition) For any $\epsilon > 0$ we have that*

$$\lim_{n \to \infty} \sum_{k=1}^{n} \mathbb{E}(X_{n,k}^2 1_{|X_{n,k}| > \epsilon}) = 0$$

Then we have as $n \to \infty$

$$S_n = \sum_{k=1}^{m_n} X_{n,k} \implies \mathcal{N}(0,1)$$

3 Stochastic processes

Let $(H_i)_{i \in I}$ be a sequence of complete, separable metric spaces, where I is an arbitrary index set.

Definition 3.1 (Product space)**.** *We define the product topology for $H = \prod_{i \in I} H_i$ to be the coarsest (that is the weakest/smallest) topology such that all the maps $\pi_i : H \to H_i$ are continuous.*

Remark 3.1.

1. This construction is equivalent to the construction of the weak topology.

2. This guarantees that the Borel σ-Algebra $\mathcal{B}(H)$ coincides with the product σ-algebra

$$\sigma \left(\left\{ \prod_{i \in I} A_i : A_i \in \mathcal{B}(H_i) \text{ for } i \in K, A_i = H_i \text{ for } i \notin K, K \subset F \text{ finite} \right\} \right)$$

that is, the product σ-algebra on $\mathbb{R}^{\mathbb{N}}$ is generated by

$$\bigcup_{n=1}^{\infty} \{ A_1 \times A_2 \times \cdots \times A_n \times \mathbb{R} \times \mathbb{R} \times \cdots \mid A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R}) \}$$

Definition 3.2 (Product measure)**.** *Let μ_i be a probability measure on H_i, we define the product measure $\mu = \bigotimes_{i \in I} \mu_i$ on H to be the unique measure such that*

$$\mu \left(\prod_{i \in I} A_i \right) = \int 1_{\prod_{i \in I} A_i} d\mu = \prod_{i \in I} \mu_i(A_i)$$

This implies that

$$\int \prod_{j \in K} f_j(\pi_j(\omega)) \mu(d\omega) = \prod_{j \in K} \int f_j d\mu_j, \text{ for all } f_j \in L^1(H_j, \mathcal{B}(H_j), \mu_j)$$

under the assumption that $A_i \in \mathcal{B}(H_i)$ for $i \in K$ and $A_i = H_i$ for $i \notin K$.

It is by no means clear that such a unique product measure exists at all, this is however part of the celebrated Kolmogorov Extension theorem which we will state next.

Theorem 3.1 (Kolmogorov's Extension Theorem). *Let $K \subset I$ be a finite subset and let μ_K be a probability measure on $\prod_{i \in K} H_i$. Suppose that for any $K \subset J$ where J is another finite subset of I the following consistency condition holds:*

$$\mu_J \left(\prod_{i \in J} A_i \right) = \mu_K \left(\prod_{i \in K} B_i \right), \ \ with \ A_i = \begin{cases} B_i \in \mathcal{B}(H_i), & if \ i \in K \\ H_i, & if \ i \notin K \end{cases}$$

Then there exists a unique measure on $H = \prod_{i \in I} H_i$ such that for all finite subsets $K \subset I$ one has

$$\mu \left(\prod_{i \in I} A_i \right) = \mu_K \left(\prod_{i \in K} B_i \right), \ \ with \ A_i = \begin{cases} B_i \in \mathcal{B}(H_i), & if \ i \in K \\ H_i, & if \ i \notin K \end{cases}$$

Remark 3.2.

1. Here $(\mu_K)_{K \subset I, K \text{ finite}}$ is called the collection of finite dimensional marginal distributions of μ.

2. Kolmogorov extension theorem says that, given the finite dimensional marginales (μ_K) as above, if they satisfy the consistency condition as requested in the theorem, then there exists a unique extension μ to the whole product space H with marginale (μ_K)

3. The Theorem is also known as Kolmogorov's existence theorem or Kolmogorov's consistency theorem. It guarantees taht a suitably "consistent" (here consistency is meant as in the theorem) collection of finite-dimensional distributions will define a stochastic process.

Definition 3.3. *A stochastic process $X = (X_i)_{i \in I}$, where I is an arbitraty index sex, is a collection of random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $X_i : \Omega \to H_i$ is $\mathcal{F}$-measurable.*
The law of X is a probability measure μ on $H = \prod_{i \in I} H_i$ such that for all finite $K \subset I$, the marginal law μ_K of μ is given by

$$\mu_K \left(\prod_{j \in K} A_j \right) = \mathbb{P}(X_j \in A_j, j \in K) \ for \ all \ A_j \in \mathcal{B}(H_j)$$

Remark 3.3. We can always choose $(\Omega, \mathcal{F}, \mathbb{P}) = (H = \prod_{i \in I} H_i, \mathcal{B}(H), \mu)$, then we have $X_i = \pi_i$ where π_i are the coordinate maps, i.e.

$$\pi_i : \begin{cases} H & \longrightarrow H_i \\ \omega & \longmapsto \omega_i \end{cases}$$

This is sometimes called the *canonical process.*

There is a section about the Galton-Watson process, it is an important exercise and will be left open on this summary.

4 Markov Processes

Definition 4.1. *A stochastic process $(X_t)_{t \in I}$ is a Markov if for all $f : \mathcal{H} \to \mathbb{R}_+$ measurable ($\mathcal{H}$ is a metric space) and for all $s < t$ we have*

$$\mathbb{E}(f(X_t) \mid \mathcal{F}_s^X) = \mathbb{E}(f(X_t) \mid \sigma(X_s)), \ \text{where } \mathcal{F}_s^X := \sigma(X_t : t \leq s)$$

Definition 4.2. *A process $(X_n)_{n \in \mathbb{N}}$ is a time-homogenous Markov chain if there exists a measurable map $Q : \mathcal{H} \to \mathcal{M}(\mathcal{H}) := \{probability\ measures\ on\ \mathcal{H}\}$ such that $\mathbb{P}(X_n \in A \mid X_{n-1} = x) = Q(x, A)$ for all Borel sets $A \subset \mathcal{H}$. We call Q the transition probability of the Markov chain. $\mathcal{H}$ is called the state space of the Chain, moreover $x \in \mathcal{H}$ is called a state of the chain.*

Proposition 4.1. *$\mathcal{M}(\mathcal{H})$ equipped with the topology of weak convergence is a completely metrizable, separable space.*

Recall that a sequence $\mu_n \in \mathcal{M}(\mathcal{H})$ converges weakly to $\mu \in \mathcal{M}(\mathcal{H})$ if and only if

$$\int f d\mu_n \to \int f d\mu, \ \text{for all } f \in C_b(\mathcal{H} \to \mathbb{R})$$

Given Q as above, then there exists a Markov chain with $\mathbb{P}_\nu(X_0 \in A) = \nu(A)$ and transition probability Q. We define for all $n \in \mathbb{N}$

$$\mathbb{P}_\nu(x_0 \in A_0, x_1 \in A_1, \ldots, x_n \in A_n) = \int_{A_0 \times \cdots \times A_n} \nu(dx_0) Q(x_0, dx_1) \ldots Q(dx_{n-1}, dx_n)$$

By Kolmogorov's extension theorem, there exists a unique probability measure $\mathbb{P}_\nu$ on $(\Omega, \mathcal{F}_\infty)$, where $\mathcal{F}_\infty = \bigvee_{n=1}^\infty \mathcal{F}_n$ is the product σ-algebra on Ω, with marginal distribution as above.

Definition 4.3. *We define the shift operator acting on Ω by $\Theta : \Omega \to \Omega$, $\Theta(\omega) = (\omega_1, \omega_2, \omega_3, \ldots)$ where $\omega = (\omega_0, \omega_1, \ldots)$.*
We then have $\Theta^n(\omega) = (\omega_{n+j})_{j=0}^\infty$ for all $n \in \mathbb{N}$.

Theorem 4.1 (Markov Property)**.** *Under $\mathbb{P}_\nu$, conditionally on the canonical filtration $\mathcal{F}_n = \sigma(\omega_0, \omega_1, \ldots, \omega_n)$, $\Theta^n(\omega)$ has law given by $\mathbb{P}_{\omega_n}$ and is independent of $\mathcal{F}_{n-1}$ for all $n \in \mathbb{N}_{\geq 1}$.*

Theorem 4.2 (Strong Markov Property)**.** *Let T be a stopping time such that $\mathbb{P}_\nu(T < \infty) > 0$. Conditionally on the event $\{T < \infty\}$ and $\mathcal{F}_T$ we have that Θ^n is a Markov Chain with law given by $\mathbb{P}_{X_T}$ and independent of $\mathcal{F}_{T-1}$*

We now work (hopefully) with $\mathcal{H}$ a countable (or finite) state space, the transition probability Q is then represented as a matrix, i.e. we have

$$Q(x, A) = \sum_{y \in A} Q_{xy}, \text{ for all } x \in \mathcal{H}$$

A probability measure on $\mathcal{H}$ is a vector such that we have

$$\nu(A) = \sum_{y \in A} \nu_y$$

We have

$$\mathcal{M}(\mathcal{H}) = \{\mu \in \mathbb{R}^{\mathcal{H}} : \mu_x \geq 0 \text{ for all } x \in \mathcal{H} \text{ and } \sum_{x \in \mathcal{H}} \mu_x = 1\}$$

Definition 4.4. *We define for all $x \in \mathcal{H}$*

$R_x := \inf\{n \geq 1 : \omega_n = x\}$, *i.e. the first return to the state x.*

$N_x := \#\{n \geq 0 : \omega_n = x\}$, *i.e. the number of visits of the state x.*

A state $x \in \mathcal{H}$ is called recurrent if $\mathbb{P}_x(R_x < \infty) = 1$, else x is called transient.

Proposition 4.2. *If x is recurrent, then $\mathbb{P}(N_x = \infty) = 1$.*
If x is transient, then we have

$$\mathbb{E}_x(N_x) = \frac{1}{\mathbb{P}_x(R_x = \infty)} < \infty$$

Definition 4.5. *The Green Kernel of the Markov Chain is defined as*

$$G_{xy} := \mathbb{E}_x(N_y) = \sum_{n \in \mathbb{N}} Q_{xy}^n = \sum_{n \in \mathbb{N}} \mathbb{P}_x(\omega_n = y)$$

i.e. the expected number of visits of y starting from the state x.
(Recall that $Q_{xy}^n = \mathbb{P}_x(\omega_n = y)$.)

Proposition 4.3. *Let $x, y \in \mathcal{H}$, $x \neq y$, then the following holds*

1. *$G_{xx} = \infty$ if and only if x is recurrent.*

2. *$G(x, y) = 0$ if and only if $\mathbb{P}_x(R_y < \infty) = 0$.*
   (i.e. impossible to go from x to y.)

3. *If $\mathbb{P}_x(R_y < \infty) > 0$, then $G(x, y) = G(y, y)\mathbb{P}_x(R_y < \infty)$.*

Proof. See Exercises! $\qquad\square$

Corollary 4.1. *Let $R := \{x \in \mathcal{H} : \mathbb{P}_x(R_x < \infty) = 1\}$ the set of recurrent states. We have that if $x \in R$ and $G(x, y) > 0$, then $y \in R$ and also $G(y, x) > 0$, $\mathbb{P}_y(R_x < \infty) = 1$.*

Remark 4.1. The above corollary yields that the being recurrent gives equivalent classes on R. We have $R = \cup_{j \in I} R_j$ and the sets R_j are disjoint, the R_j's are called the recurrence classes. We define the equivalence relation $x \sim y$ if $G(x, y) > 0$. It follows that $x \in R_j$, then $G(x, y) > 0$ if and only if $y \in R_j$.

Theorem 4.3 (Classification of states). *The following holds true*

1. *If $x \in R_j$ for some $j \in I$, then*

$$\begin{cases} \mathbb{P}_x(N_y = \infty) = 1, & \text{for all } y \in R_j \\ \mathbb{P}_x(N_y = 0) = 1, & \text{for all } y \notin R_j \end{cases}$$

2. *Let $T = \inf\{n \geq 0 : \omega_n \in R\}$ be the first time, then some weird statement that should be reformulated.*

Definition 4.6. *The chain is called irreducible if $G_{xy} > 0$ for all $x, y \in \mathcal{H}$. Irreducibility means that either we have $R = \mathcal{H}$ or $R = \emptyset$.*

Corollary 4.2. *If the chain is irreducible, one of the following applies*

1. *All states are recurrent, in particular $\mathbb{P}_x(N_y = \infty, \text{ for all } y \in \mathcal{H}\} = 1$ for all $x \in \mathcal{H}$.*

2. *All states are transient, in particular $\mathbb{P}_x(N_y < \infty, \text{ for all } y \in \mathcal{H}\} = 1$ for all $x \in \mathcal{H}$.*