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 interpret our filtered probability space as an random experiment, then the maximum information that can be found until the random time T sets in is $\mathcal{F}_T$.

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*

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