# Chapter 4

# Conditioning and Martingale

Aithus C. Mao on 12th March 2024

The martingale replace the process of completely independence with similar repetition.

## 4.1 Conditioning

Let $X$ be a random variable on a probability space $(E, \mathscr{E}, \mathbb{P})$ with $\mathbb{E}|X| \leq \infty$.

**Definition 4.1.1** (Conditional Expectation w.r.t. a $\sigma$-algebra)**.** *Let $\mathscr{E}'$ be a sub-$\sigma$-algebra w.r.t. $\mathscr{E}$, the conditional expectation $\mathbb{E}[X|\mathscr{E}']$ is any random variable $Y$ on $\mathscr{E}'$ such that*

$$\text{for all } A \in \mathscr{E}',\ \int_A X d\mathbb{P} = \int_A Y d\mathbb{P}.$$

**Definition 4.1.2** (Conditional Expectation w.r.t. a Random Variable)**.** *Given two random variables $X$ and $Y$ on $(E, \mathscr{E}, \mathbb{P})$, the conditional expectation $\mathbb{E}[X|Y]$ is any random variable $Z$ on $(E, \sigma(Y), \mathbb{P})$ such that for all $A \in \sigma(Y)$, $\int_A X d\mathbb{P} = \int_A Z d\mathbb{P}$.*

**Lemma 4.1.1** (Uniqueness)**.** *All conditional expectations of one r.v. on a $\sigma$-algebra (or on another r.v.) is a.s. equal.*

*Proof.* We will only prove the case of conditioning on a $\sigma$-algebra, and the case of r.v. would be straightforward. Let $\mathscr{E}'$ be a sub-$\sigma$-algebra w.r.t. $\mathscr{E}$, $Y$ and $Z$ be two r.v.s that are $X$ conditioning on $\mathscr{E}'$.

Let $A \coloneqq \{x|Y(x) > Z(x)\}$. It can be easily verified that $A$ is $\mathscr{E}'$-measurable. Thus, $\int_A Y d\mathbb{P} = \int_A Z d\mathbb{P}$, i.e., $\int_A (Y-Z)d\mathbb{P} = 0$. As $Y(x) > Z(x)$ for all $x$ in $A$, we must have that $\mathbb{P}(A) = 0$. Similarly, we can also prove that $\mathbb{P}(\{x|Y(x) < Z(x)\}) = 0$. Consequently, $\mathbb{P}(\{x|Y(x) \neq Z(x)\}) = 0$, and this completes the proof. □

**Remark 4.1.1.** As stated in math stackexchange: The existence of conditional expectation is more difficult. The proofs I've seen either use the Radon-Nikodym theorem, or the Riesz representation theorem in Hilbert space. Any measure-theoretic probability book will have a proof.

**Lemma 4.1.2.** *Two random variables are independent iff the $\sigma$-algebras generated by them are independent.*

*Proof.* C. Mao-TODO □

## 4.2 Martingale

A martingale is a stochastic process where the previous r.v.s stacking on previous r.v.s by adding details.

**Definition 4.2.1** (Martingale)**.** *A real-valued stochastic process $X = (X_t)_{t\in\mathbb{T}}$ is called a martingale if $X$ is adapted to a filtration $\mathscr{F} = (\mathscr{F}_t)_{t\in\mathbb{T}}$*[*], *$X_t$ is finite integrable for all $t \in \mathbb{T}$, and*

$$\mathbb{E}[X_t - X_s|\mathscr{F}_s] = 0$$

*a.s. whenever $s < t$.*

**Definition 4.2.2** (Martingale Difference Sequence)**.** *A real-valued stochastic process $X = (X_t)_{t\in\mathbb{T}}$ is called a martingale difference sequence if $X$ is adapted to a filtration $\mathscr{F}$, each $X_t$ is finite integrable, and*

$$\mathbb{E}[X_t|\mathscr{F}_s] = 0$$

*a.s. whenever $s < t$.*

[*] A filtration is an index $\sigma$-algebras that the former is sub-$\sigma$-algebras of the latter. $X_t$ is $\mathscr{F}_t$ measurable for all $t \in \mathbb{T}$.

## 4.3 Lemma

Given two random variables $X, Y$ on $\mathscr{E}$ with $\sigma(X) \subseteq \sigma(Y)$, what would it be like if we apply conditioning on $\sigma(X)$ (i.e., $X$) to functions like $f(X,Y)$? $f(X,Y)$ shall be measurable w.r.t. $\sigma(Y)$.

**Lemma 4.3.1.** *If $f(X,Y) = XY$, we shall have that, for all $A \in \sigma(X)$,*

$$\int_A \mathbb{E}[XY|X]\, d\mathbb{P} = \int_A X\mathbb{E}[Y|X] d\mathbb{P}.$$

*Proof.* I shall only prove the case that both $X$ and $Y$ are simple functions on $E$, and the extension to other integrable functions should be easy. Let $R(X)$ and $R(Y)$ be the possible values taken by $X$ and $Y$. For any $x \in R(X)$,

$$\begin{aligned}\int_{X^{-1}(x)} \mathbb{E}[XY|X] d\mathbb{P} = \int_{X^{-1}(x)} xY d\mathbb{P} = x \int_{X^{-1}(x)} Y d\mathbb{P} \\ =& x \int_{X^{-1}(x)} \mathbb{E}[Y|X] d\mathbb{P} \\ =& \int_{X^{-1}(x)} x\mathbb{E}[Y|X] d\mathbb{P}.\end{aligned}$$

Thus, for any $A \in \sigma(X)$, by the disjoint property of $X^{-1}(x)$ for $x \in R(X)$,

$$\begin{aligned}\int_A \mathbb{E}[XY|X] d\mathbb{P} =& \sum_{x \in X(A)} \int_{X^{-1}(x)} \mathbb{E}[XY|X] d\mathbb{P} \\ =& \sum_{x \in X(A)} \int_{X^{-1}(x)} x\mathbb{E}[Y|X] d\mathbb{P} \\ =& \int_A X\mathbb{E}[Y|X] d\mathbb{P}.\end{aligned}$$

C. Mao-TODO: The proof would be much simpler if we use the uniqueness of the conditional expectation. □

## 4.4 Azuma-Hoeffding (Azuma's) Inequality

**Theorem 4.4.1** (Azuma-Hoeffding Inequality). *Suppose $(X_t)_{t\in[N]}$ is a super-martingale adapting to $(\mathscr{F}_t)_{t\in[N]}$ and*[†]

$$|X_k - X_{k-1}| \leq c_k, \tag{4.1}$$

*almost surely for all $k \in [1, N]$. Then for all positive integers $N$ and all positive real $\epsilon$,*

$$P(X_N - X_0 \geq \epsilon) \leq \exp\Big(\frac{-\epsilon^2}{2\sum_{k=1}^N c_k^2}\Big).$$

*proof sketch.* We list the key components of the proof in the following.

- For any $A \in \mathscr{E}' \subseteq \mathscr{E}$, $X \in \mathscr{E}'$, $Y \in \mathscr{E}$, and a measure $\mu$ on $\mathscr{E}$,

$$\int_A f(X,Y)d\mu = \int_A \mathbb{E}\big[f(X,Y)|\mathscr{E}'\big]d\mu.$$

  We can take the whole set $E$ as $A$.[‡]
- We can thus use the induction,

$$\begin{aligned}\mathbb{E}\Big[\exp\big(\lambda\sum_{k=1}^N[X_k - X_{k-1}]\big)\Big] =& \mathbb{E}\Big[\mathbb{E}\Big[\exp\big(\lambda\sum_{k=1}^N[X_k - X_{k-1}]\big)\big|\mathscr{F}_{N-1}\Big]\Big] \\ =& \mathbb{E}\Big[\exp\Big(\lambda\sum_{k=1}^{N-1}[X_k - X_{k-1}]\Big)\cdot \mathbb{E}\big[\exp\big(\lambda[X_N - X_{N-1}]\big)\big|\mathscr{F}_{N-1}\big]\Big].\end{aligned}$$

(By Lemma 4.3.1)

  By Eq. (4.1) and the definition of super-martingales, it holds almost surely that

$$\mathbb{E}\big[\exp\big(\lambda[X_k - X_{k-1}]\big)\big|\mathscr{F}_{N-1}\big] \leq \exp\Big(\frac{c_N^2\lambda^2}{8}\Big).$$

  We can do inductions on $[2, N]$ and get $\mathbb{E}\Big[\exp\big(\lambda\sum_{k=1}^N[X_k - X_{k-1}]\big)\Big] \leq \exp\Big(\frac{\lambda^2\sum_t c_t^2}{8}\Big)$. We can apply the Chernoff bound and get the desired result.

□

† $[N] = \{0, 1, 2 \ldots N\}$, $[M, N] = \{M \ldots N\}$. The super-martingale requires that $\mathbb{E}[X_t - X_s|\mathscr{F}_s] \geq 0$ a.s. whenever $t \geq s$.

‡ This is due to the definition of the conditioning.

## 4.5 Fun Facts

**Fact 1.** For two $\sigma$-algebras $\mathscr{E}$ and $\mathscr{E}'$ with $\mathscr{E}' \subseteq \mathscr{E}$, we only have $\mathscr{E}'$-measurable $\to$ $\mathscr{E}$-measurable. Yet $\mathscr{E}$-measure implies $\mathscr{E}'$-measure.