Lecture 17&18: Sigma Fields and Martingales

Disclaimer

- This is a very informal treatment of the concept of Martingales
- In particular, the intuitions are specific to discrete-time martingales and discrete spaces
- Interested readers are referred to study σ-algebras for a more formal treatment of this material

In these Two Lectures

- We shall introduce the concept of Martingales
- We shall study Discrete-time Martingales over Discrete Spaces
- Specifically, we shall study Doob's martingale
- In the next lecture, we shall study Azuma's inequality

σ-Field

Let Ω be a (discrete) sample space with probability distribution p. That is, for any $x \in \Omega$, the value $p(x)$ represents the probability associated with the element x.

Definition

A *σ-field* $\mathcal{F}$ on Ω is a collection of subsets of Ω such that the following constraints are satisfied.

1. $\mathcal{F}$ contains $\emptyset$ and Ω, and
2. $\mathcal{F}$ is closed under union, intersection, and complementation.

Example σ-Fields

- $\mathcal{F}_0 = \{\emptyset, \Omega\}$ is a σ-field
- Suppose $\Omega = \{0,1\}^n$
- Let $\mathcal{F}_1 = \mathcal{F}_0 \cup \left\{ \{0\} \times \{0,1\}^{n-1}, \{1\} \times \{0,1\}^{n-1} \right\}$. Note that $\mathcal{F}_1$ is also a σ-field. In general, we can write $\mathcal{F}_1$ as the following set

$$\left\{ S \times \{0,1\}^{n-1} \colon S \subseteq \{0,1\} \right\}$$

  We are using the convention that is $S = \emptyset$, then $S \times \{0,1\}^{n-1} = \emptyset$.
- Let $\mathcal{F}_2 = \left\{ S \times \{0,1\}^{n-2} \colon S \subseteq \{0,1\}^2 \right\}$. Note that $\mathcal{F}_2$ has 16 elements, and $\mathcal{F}_1 \subset \mathcal{F}_2$. It is easy to verify that $\mathcal{F}_2$ is a σ-field.
- In general, consider the following σ-field, for $0 \leqslant k \leqslant n$.

$$\mathcal{F}_k = \left\{ S \times \{0,1\}^{n-k} \colon S \subseteq \{0,1\}^k \right\}$$

Smallest Set containing any Element

- Let $x \in \Omega$
- Consider a σ-field $\mathcal{F}$ on Ω
- The smallest set in $\mathcal{F}$ containing x, represented by $\mathcal{F}(x)$, is the intersection of all sets in $\mathcal{F}$ that contain x. Formally, it is the following set
$$\mathcal{F}(x) := \bigcap_{S \in \mathcal{F} \colon x \in S} S$$
- For example, let $n = 5$, $x = 01001$, and consider the σ-field $\mathcal{F}_2$ on Ω. In this case, the smallest set $\mathcal{F}_2(x)$ in $\mathcal{F}_2$ that contains x is $\{01\} \times \{0, 1\}^{n-2}$.

$\mathcal{F}$-Measurable

- Let $f : \Omega \to \mathbb{R}$ be an arbitrary function

Definition ($\mathcal{F}$-Measurable)

The function f is *$\mathcal{F}$-measurable* if, for all $x \in \Omega$ and $y \in \mathcal{F}(x)$, we have $f(x) = f(y)$, where $\mathcal{F}(x)$ represents the smallest subset in $\mathcal{F}$ containing x

- Intuitively, the function f is constant over all the elements of $\mathcal{F}_2(x)$, for any $x \in \Omega$
- For example, let $n = 5$ and consider the σ-field $\mathcal{F}_2$ on Ω
- As we have seen, we have $\mathcal{F}_2(x) = \{x_1x_2\} \times \{0,1\}^{n-2}$, where x_1 and x_2 are, respectively, the first and the second bits of x. That is, $\mathcal{F}_2(x)$ is the set of all n-bit strings that begin with x_1x_2.
- Let $f(x)$ be the total number of 1s in the first two coordinates of x. This function is $\mathcal{F}_2$-measurable
- Let $f(x)$ be the expected number of 1s over all strings whose first two bits are x_1x_2. This function is also $\mathcal{F}_2$-measurable
- Let $f(x)$ be the total number of 1s in the first three bits of x. This function is not $\mathcal{F}_2$-measurable, because $x = 00000$ and $y = 00100$ satisfy $y \in \mathcal{F}_2(x)$ but $f(x) \neq f(y)$

Conditional Expectation

- Let p be a probability distribution over the sample space Ω
- Let $\mathcal{F}$ be a σ-field on Ω
- Let $f\colon \Omega \to \mathbb{R}$ be an arbitrary function
- We define the conditional expectation as a function $\mathbb{E}\left[f|\mathcal{F}\right] : \Omega \to \mathbb{R}$ defined as follows

$$\mathbb{E}\left[f|\mathcal{F}\right](x) := \frac{1}{\sum_{y\in\mathcal{F}(x)} p(y)} \sum_{y\in\mathcal{F}(x)} f(y)\cdot p(y)$$

- We emphasize that the function f need not be $\mathcal{F}$-measurable to define the expectation in this manner!
- Note that $\mathbb{E}\left[f|\mathcal{F}\right](x) = \mathbb{E}\left[f|\mathcal{F}\right](y)$, for all $y \in \mathcal{F}(x)$. That is, the function $\mathbb{E}\left[f|\mathcal{F}\right]$ is $\mathcal{F}$-measurable!

Filtration

Let Ω be a sample space with probability distribution p

Definition (Filtration)

A sequence of σ-fields $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_n$ on Ω is a filtration if

$$\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$$

Note that when $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, then the σ-fields $\mathcal{F}_i$ on Ω defined below forms a filtration.

$$\mathcal{F}_i = \{S \times \Omega_{i+1} \times \cdots \times \Omega_n \colon S \subseteq \Omega_1 \times \cdots \times \Omega_i\}$$

Beginning of “Intuition Slides”

Sample Space

- As time progresses, new information about the sample is revealed to us
- At time 1, we learn the value of ω_1 of the random variable $\mathbb{X}_1$
- At time 2, we learn the value of ω_2 of the random variable $\mathbb{X}_2$
- As so on. At time t, we learn the value of ω_t of the random variable $\mathbb{X}_t$
- By the end of time n, we know the value ω_n of the last random variable $\mathbb{X}_n$
- At this point, $f(\mathbb{X}_1, \ldots, \mathbb{X}_n)$ can be calculated, where $f : \Omega \to \mathbb{R}$ is a function that we are interested in

Examples

- Balls and Bins. At time i we find out the bin ω_i where the ball i lands
- Coin tosses. At time i we find out the outcome ω_i of the i-th coin toss
- Hypergeometric Series. At time i we find out the color ω_i of the i-th ball drawn from the jar (where sampling is being carried out without replacement)
- Bounded Difference Function. At time i we find out the outcome ω_i of the i-th variable of the input of the function f.

Filtration

- In a filtration, the σ-field $\mathcal{F}_k$ represents the knowledge we have after knowing the outcomes $(\omega_1, \ldots, \omega_k)$
- For instance, the σ-field $\mathcal{F}_0$ on Ω represents "we know nothing about the sample"
- For instance, the σ-field $\mathcal{F}_n$ on Ω represents "we know everything about the sample"
- In general, the σ-field $\mathcal{F}_k$ on Ω represents "we know the first k coordinates of the sample"

Tree Representation

- Think of a rooted tree
- For every internal node, the outgoing edges represent the various possible outcomes in the next time step
- Leaves represent that the entire sample is already known
- The sequence of outcomes $(\omega_1, \ldots, \omega_n)$ represents a "root-to-leaf" path
- Consider a filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$. The set $\mathcal{F}_k(x)$ corresponding to this root-to-leaf path is the depth-k node on this path

Measurable with respect to a σ-Field

- Consider a filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$
- A function f being $\mathcal{F}_k$-measurable implies that f is constant over all leaves of the subtree rooted at $\mathcal{F}_k(x)$
- A random variable $\mathbb{F}_k = f(\mathbb{X}_1, \ldots, \mathbb{X}_n)$ will be $\mathcal{F}_k$-measurable if the value of $f(\mathbb{X}_1, \ldots, \mathbb{X}_n)$ depends only on $\mathbb{X}_1 = \omega_1, \ldots, \mathbb{X}_k = \omega_k$

End of "Intuition Slides"

Definition (Martingale Sequence)

Let $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$ be a filtration. The sequence $(\mathbb{F}_0, \mathbb{F}_1, \ldots, \mathbb{F}_n)$ forms a martingale with respect to the filtration if

1. $\mathbb{F}_t$ is $\mathcal{F}_t$-measurable, for $0 \leqslant t \leqslant n$, and
2. $\mathbb{E}\left[\mathbb{F}_{t+1} \mid |\mathcal{F}_t\right] = (\mathbb{F}_t|\mathcal{F}_t)$, for all $0 \leqslant t < n$.

- Note that given $\mathcal{F}_t = (\omega_1, \omega_2, \ldots, \omega_t)$, the value of $\mathbb{F}_t$ is fixed. So, we can write $\mathbb{E}\left[\mathbb{F}_t|\mathcal{F}_t\right](x)$ in short as $(\mathbb{F}_t|\mathcal{F}_t)(x)$
- Note that given $\mathcal{F}_t = (\omega_1, \omega_2, \ldots, \omega_t)$, the outcome of $\mathbb{F}_{t+1}$ is not yet fixed and is (possibly) random
- The second equation in the definition is an "equality of two functions." It means that $\mathbb{E}\left[\mathbb{F}_{t+1}|\mathcal{F}_t\right](x)$ is equal to $(\mathbb{F}_t|\mathcal{F}_t)$ for all $x \in \Omega$

Example

- Consider tossing a coin that gives heads with probability p, and tails with probability $(1-p)$, independently n times
- $\mathcal{F}_t$ is the outcome of the first t coin-tosses
- Let $\mathbb{S}_t$ represent the number of heads in the first t coin tosses
- Note that $\mathbb{S}_t(x)$ is fixed given $\mathcal{F}_t(x)$, where $x \in \Omega$
- Note that $(\mathbb{S}_{t+1}|\mathcal{F}_t)(y) = (\mathbb{S}_t|\mathcal{F}_t)(y) + 1$ with probability p (for a random y that is consistent with $\mathcal{F}_t(x)$), else $(\mathbb{S}_{t+1}|\mathcal{F}_t)(y) = (\mathbb{S}_t|\mathcal{F}_t)(y)$
- Therefore, $\mathbb{E}\left[\mathbb{S}_{t+1}|\mathcal{F}_t\right] = (\mathbb{S}_t|\mathcal{F}_t)(x) + p$
- So, the sequence $(\mathbb{S}_0, \mathbb{S}_1, \ldots, \mathbb{S}_n)$ is <u>not</u> a martingale sequence with respect to the filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$

Example

- Let f be a function and we consider a filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$
- Let $\mathbb{F}_t$ be the following random variable

$$\mathbb{F}_t = \mathbb{E}\left[f(\omega_1, \ldots, \omega_t, \mathbb{X}_{t+1}, \ldots, \mathbb{X}_n)\right],$$

  where $\omega_1, \ldots, \omega_t$ are the first t outcomes of $x \in \Omega$
- First, prove that $\mathbb{F}_t$ is $\mathcal{F}_t$ measurable
- Finally, prove that $(\mathbb{F}_0, \ldots, \mathbb{F}_n)$ is a martingale with respect to the filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$