Lecture 13: March 20

Lecturer: Alessandro Rinaldo *Scribes: Maria Jahja*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Last Time: Martingales

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\{(\mathcal{F}_n, X_n)\}_{n=1}^{\infty}$ be a martingale where $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$ is a sequence of sub-σ-fields of $\mathcal{F}$ and $\{X_n\}_{n=1}^{\infty}$ is a collection of random variables such that X_n is an $\mathcal{F}_n$-meas. random variable. Explicitly, we say $\{X_n\}_{n=1}^{\infty}$ is a martingale relative to the filtration $\{\mathcal{F}_n\}_{n=1}^{\infty}$ if $X_n : \Omega \mapsto \mathbb{R}$ is $\mathcal{F}_n$-measurable and

1. $\mathbb{E}[|X_n|] < \infty$ for all n, and
2. $\mathbb{E}[|X_{n+1}|\mathcal{F}_n] = X_n$ for all n.

If $\mathbb{E}[|X_{n+1}|\mathcal{F}_n] \geq X_n$ for all n, we say $\{(\mathcal{F}_n, X_n)\}_{n=1}^{\infty}$ is a submartingale. If $\{(\mathcal{F}_n, X_n)\}_{n=1}^{\infty} \leq X_n$, we call it a supermartingale.

13.1 Stopping Times

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\{\mathcal{F}_n\}_{n=1}^{\infty}$ be a filtration.

Definition 13.1 (Stopping times). *A positive*[1] *(possibly extended) integer valued random variable τ is called a* stopping time *with respect to the filtration if $\{\tau = n\} \in \mathcal{F}_n$ for all finite n.*

If $\{X_n\}_{n=1}^{\infty}$ is adapted to the filtration and $\tau < \infty$ a.s.[2]*, then we define X_τ as $X_{\tau(\omega)}(\omega)$. A special σ-field $\mathcal{F}_\tau$ is defined by*

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq k\} \in \mathcal{F}_k \quad \forall \text{ finite } k\}$$

where τ is measurable with respect to $\mathcal{F}_\tau$ and X_τ is $\mathcal{F}_\tau$-measurable.

Example 16 (Gambler's ruin). Suppose there is a gambling system where $\{X_n\}$ are independent Rademacher random variables such that $\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 1/2$. We define how much money was won or lost by $Z_n = \sum_{i=1}^{n} X_i$. Suppose the gambler chooses the stopping time $\tau = \min\{n : Z_n \geq 0\}$, for some target integer $x > 0$. In other words, the gambler keeps playing until x money is won.

This seems like a desirable strategy which guarantees winning at least x. However, there are two drawbacks.

[1] If your filtration starts at $n = 0$, you can allow stopping times to be nonnegative valued. Indeed, if your filtration starts at an arbitrary integer k, then a stopping time can take any value from k on up. There is a trivial extension of every filtration to one lower subscript. For example, if we start at $n = 1$, we can extend to $n = 0$ by defining $\mathcal{F}_0 = \{\Omega, \emptyset\}$. Every martingale can also be extended by defining $X_0 = \mathbb{E}(X_1)$. We will assume that the lowest possible value for a stopping time is 1.

[2] If $\tau = \infty$, let X_τ be equal to some arbitrary random variable X_∞.

1. In order to play this strategy, the gambler must have an unlimited reserve of money. First, note that we will stop as soon as we have won x more than we have lost. If the gambler starts with k dollars, the probability of winning $Z_n = x$ before $Z_n = -k$ is $k/(k+x)$, which goes 1 as $k \to \infty$. Therefore, given unlimited resources, we have that $\tau < \infty$ with probability 1. Otherwise, if the gambler has finite capital, this goal may never be achieved and Z_n can be arbitrarily negative.

2. Further suppose the game is unfair, i.e. $\mathbb{P}(X_n = -1) > \mathbb{P}(X_n = 1)$. Then $\mathbb{P}(\tau = \infty) > 0$, meaning there is a positive probability the gambler will never stop.

13.2 Sequences of Stopping Times

Suppose τ_1 and τ_2 are two stopping times such that $\tau_1 \leq \tau_2$. Let $A \in \mathcal{F}_{\tau_1}$. Since $A \cap \{\tau_2 \leq k\} = A \cap \{\tau_1 \leq k\} \cap \{\tau_2 \leq k\}$, it follows that $A \cap \{\tau_2 \leq k\} \in \mathcal{F}_k$, and $A \in \mathcal{F}_{\tau_2}$. This implies that $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

Now imagine we have a martingale $\{(\mathcal{F}_n, X_n)\}_{n=1}^{\infty}$ and a sequence of a.s. finite stopping times $\{\tau_n\}_{n=1}^{\infty}$ such that $\tau_j \leq \tau_{j+1}$ for all j. Then we can construct a sub-martingale $\{(\mathcal{F}_{\tau_n}, X_{\tau_n})\}_{n=1}^{\infty}$. This is known as the *optional sampling theorem.*

Theorem 13.2 (Optional sampling theorem) *Let $\{(\mathcal{F}_n, X_n)\}_{n=1}^{\infty}$ be a (sub)martingale and let $\{\tau_n\}_{n=1}^{\infty}$ be a sequence of stopping times such that $\tau_n \leq M_n$ for all n a.s, where M_n is a finite constant. Then $\{(\mathcal{F}_{\tau_n}, X_{\tau_n})\}_{n=1}^{\infty}$ is a (sub)martingale.*

Proof: (Optimal sampling theorem).

Without loss of generality, assume that $M_n \leq M_{n+1}$ for each n. We first show that $\mathbb{E}[|X_{\tau_n}|]$ is finite. Since $\tau_n \leq M_n$ for all n,

$$\mathbb{E}[|X_{\tau_n}|] = \sum_{k=1}^{M_n} \int_{\{\tau_n = k\}} |X_k| dP \leq \sum_{k=1}^{M_n} \mathbb{E}[|X_k|] < \infty$$

using that $\mathbb{E}[|X_k|] < \infty$.

Next, to prove the second condition $\mathbb{E}[X_{\tau_{n+1}} | \mathcal{F}_{\tau_n}] \ (\geq) = X_{\tau_n}$ holds, we first note that X_{τ_n} is $\mathcal{F}_{\tau_n}$-measurable. Let $A \in \mathcal{F}_{\tau_n}$. We would like to show that

$$\int_A X_{\tau_{n+1}} dP \overset{(\geq)}{=} \int_A X_{\tau_n} dP \quad \text{for all } n.$$

Write

$$\int_A (X_{\tau_{n+1}} - X_{\tau_n}) dP = \int_{A \cap \{\tau_n \leq \tau_{n+1}\}} (X_{\tau_{n+1}} - X_{\tau_n}) dP.$$

For each $\omega \in \{\tau_n \leq \tau_{n+1}\}$, we have that

$$X_{\tau_{n+1}}(\omega) - X_{\tau_n}(\omega) = \sum_{k:\ \tau_n(\omega) < k \leq \tau_{n+1}(\omega)} (X_k(\omega) - X_{k-1}(\omega)).$$

The smallest k such that $\tau_n < k$ is $k = 2$. Therefore,

$$\int_{A \cap \{\tau_n > \tau_{n+1}\}} (X_{\tau_{n+1}} - X_{\tau_n}) dP = \int_A \sum_{k=2}^{M_{n+1}} \mathbf{1}\{\tau_n < k \leq \tau_{n+1}\}(X_k - X_{k-1}) dP.$$

Since $A \in \mathcal{F}_{\tau_n}$ and $\{\tau_n < k \le \tau_{n+1}\} = \{\tau_n \le k-1\} \cap \{\tau_{n+1} \le k-1\}^{\mathsf{c}}$, it follows that $B_k \in A \cup \{\tau_n < k \le \tau_{n+1}\} \in \mathcal{F}_{k-1}$ for any k. Hence,

$$\begin{aligned}\int_A (X_{\tau_{n+1}} - X_{\tau_n}) dP &= \sum_{k=2}^{M_{n+1}} \int_{B_k} (X_k - X_{k-1}) dP \\ &\overset{(\geq)}{=} \sum_{k=2}^{M_{n+1}} \int_{B_k} (X_k - \mathbb{E}[X_k|\mathcal{F}_{k-1}]) dP \\ &= 0\end{aligned}$$

since $B_k \in \mathcal{F}_{k-1}$ and $\mathbb{E}[X_k|\mathcal{F}_{k-1}] \overset{(\geq)}{=} X_{k-1}$. We conclude that

$$\int_B X_{k-1} dP = \int_B X_k dP$$

for all $B \in \mathcal{F}_{k-1}$. ■

Corollary 13.3 (Submartingale maximal inequality corollary) *If $X_1, X_2, \ldots, X_n$ is a submartingale with respect to the filtration $\{\mathcal{F}_n\}$, then for all $\alpha > 0$*

$$\mathbb{P}\left(\max_{i=1,\ldots,n} X_i \ge \alpha\right) \le \frac{1}{\alpha}\mathbb{E}[|X_n|].$$

We could also use Markov's inequality to say that if $X_i \ge 0\ \forall i$, then $\mathbb{P}(\max_i X_i \ge \alpha) \le \alpha^{-1}\mathbb{E}[\max_i X_i]$. This is a worse bound.

The corollary also extends Kolmogorov's maximal inequality, which says that if $X_1, X_2, \ldots, X_n$ are zero-mean independent random variables with finite variance, then

$$\mathbb{P}\left(\max_i |\sum_{i=1}^k X_i| \ge \alpha\right) \le \frac{\mathrm{Var}(S_n)}{\alpha^2}$$

where $S_n = \sum_{i=1}^n X_i$.

Proof: (Submartingale maximal inequality corollary).

Let n be a positive integer, and define stopping times $\tau_2 = n$ and $\tau_1 = \min\{k \ge 0 : X_k \ge \alpha\}$. If there is no such k, set $\tau_1 = n$.

Let $M_k = \max_{i=1,\ldots,k} X_i$. Then for all k, $\{M_n \ge \alpha\} \cap \{\tau_1 \le k\} = \{M_k \ge \alpha\} \in \mathcal{F}_k$. It follows that $\{M_k \ge \alpha\} \in \mathcal{F}_{\tau_1}$. So,

$$\alpha\mathbb{P}(M_n \ge \alpha) \le \int_{\{M_n \ge \alpha\}} X_{\tau_1}\ dP$$

since $X_{\tau_1} \ge \alpha$. Using the optional sampling theorem, we have that $X_{\tau_1} \le \mathbb{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}]$. Therefore, we have that

$$\int_{\{M_n \ge \alpha\}} X_{\tau_1}\ dP \le \int_{\{M_n \ge \alpha\}} \mathbb{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}]\ dP.$$

Since $\{M_n \geq \alpha\} \in \mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$,

$$\begin{aligned}
\int_{\{M_n \geq \alpha\}} \mathbb{E}[X_{\tau_2} | \mathcal{F}_{\tau_1}] \, dP &= \int_{\{M_n \geq \alpha\}} X_{\tau_2} \, dP \\
&= \int_{\{M_n \geq \alpha\}} X_n \, dP \\
&\leq \int_{\{M_n \geq \alpha\}} X_n^+ \, dP \qquad \text{where } X_n^+ \text{ is } \max\{X_n, 0\} \\
&\leq \mathbb{E}[X_n^+] \leq \mathbb{E}[|X_n|].
\end{aligned}$$

■

13.3 Convergence of Random Variables

We will discuss three types of convergence for random variables,

1. Convergence in probability,
2. Convergence in L_p,
3. Convergence almost surely.

Convergence in distribution is a weaker form of convergence that will be discussed later in the course.

Definition 13.4 (Convergence in probability). *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A sequence $\{X_n\}$ of random variables* converges in probability *to a random variable X when, for any $\varepsilon > 0$,*

$$P(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) \to 0 \text{ as } n \to \infty.$$

We denote this as $X_n \xrightarrow{p} X$.

In other words, it is when the measure of all ω such that $|X_n(\omega) - X(\omega)| > \varepsilon$ goes to zero. Note that this does not imply that $X_n(\omega) \to X(\omega)$ $\forall \omega$ in a set of zero probability.

Example 13.3. Let X_n take values in $\{0, 1\}$ such that $\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = \frac{1}{2}\frac{n+1}{n}$. Let $X \sim \text{Bernoulli}(1/2)$. Does $X_n \xrightarrow{p} X$?

Answer: In general, no. Consider these two scenarios.

1. Take $X_n \perp\!\!\!\perp X$ for all n. Then $\forall \varepsilon \in (0, 1)$ we have that

$$\begin{aligned}
\mathbb{P}(|X_n - X| \geq \varepsilon) = \mathbb{P}(X_n + X = 1) &= \mathbb{P}(X_n = 0 \cap X = 1) + \mathbb{P}(X_n = 1 \cap X = 0) \\
&= \frac{1}{4}\frac{n+1}{n} + \frac{1}{4}\frac{n-1}{n} = \frac{1}{2}.
\end{aligned}$$

   Hence, $X_n \stackrel{p}{\nrightarrow} X$.

2. However, if we take $\mathbb{P}(X_n = 1 | X = 1) = 1$ and $\mathbb{P}(X_n = 1 | X = 0) = 1/n$, then

$$\mathbb{P}(X_n = 1) = \frac{1}{2}\frac{n+1}{n}$$

and $\forall \varepsilon \in (0,1)$,

$$\begin{aligned}\mathbb{P}(|X_n - X| \geq \varepsilon) &= \mathbb{P}(X_n = 1 | X = 0)\mathbb{P}(X = 0) + \mathbb{P}(X_n = 0 | X = 1)\mathbb{P}(X = 1) \\ &= \frac{1}{2n} \to 0 \text{ as } n \to \infty.\end{aligned}$$

In this case, $X_n \xrightarrow{p} X$.

Most common statement. Usually, we are more interested in if X_n converges in probability to some constant c, rather than a random variable.

We can extend convergence in probability to general measures and distances.

Definition 13.5 (Convergence in measure). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $\{f_n\}$ be a sequence of measurable functions on (Ω, f) taking values on the metric space $(\mathcal{X}, d)$. Let f be a measurable function taking values in $(\mathcal{X}, d)$. Then we say f_n* converges in measure *to f if, for any $\varepsilon > 0$,*

$$\mu(\{\omega : d(f_n(\omega), f(\omega)) > \varepsilon\}) \to 0 \textit{ as } n \to \infty.$$

If we take μ to be a probability, convergence in measure is called convergence in probability.