MATH GU 4156 **Advanced Probability Theory** Columbia University (Fall 2023)

Lecture 5: Martingales: Definitions and First Results

Lecturer: ***Ioannis Karatzas*** *Scribes: Heyuan Yao*

5.1 Martingale and Examples

In our discussion of filtration and stopping times, there was no need to introduce a probability measure. This becomes necessary now.

We start with a **filtered probability space** $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{F} = \{\mathcal{F}_n\}_{n\in\mathbb{N}_0}$. A sequence $\mathcal{X} = \{X_n\}_{n\in\mathbb{N}_0}$ of random variables, adapted to $\mathbb{F}$ and integrable ($\mathbb{E}|X_n|< \infty$, $\forall n \in \mathbb{N}_0$) is called **martingale (resp., submartingale, supermartingale)** if

$$\mathbb{E}(X_m|\mathcal{F}_n) = X_n \quad \text{(resp, } \geq, \leq) \tag{5.1}$$

holds $\mathbb{P}$-a.e. for every $n \in \mathbb{N}_0,\ m \in \mathbb{N}_0$.

By the tower property of conditional expectations, it is enough to verify (5.1) for $m = n + 1$.

It is hard to believe, yet true, that you can build almost the entire edifice of Probability Theory based on such a simple property as (5.1).

It is a direct consequence of (5.1) that we have

$$\mathbb{E}(X_m) = \mathbb{E}(X_n) \quad \text{(resp, } \geq, \leq)$$

for a martingale (resp., submartingale, supermartingale): Conservation Law.

Here are some examples.

Lévy Martingale: $X_n = \mathbb{E}(\xi|\mathcal{F}_n)$, $n \in \mathbb{N}_0$ with $\xi in \mathbb{L}^1$.

Random Walk: Suppose $\xi_1, \xi_2, ...$ are independent, integrable with $\mathbb{E}(\xi_j) = \alpha$, $\forall j \in \mathbb{N}$ and define

$$S_0 = 0; \quad S_n = \sum_{j=1}^{n} \xi_j \quad (n \in \mathbb{N}).$$

Then $X_n = S_n - \alpha n,\ n \in \mathbb{N}_0$ is a martingale. And if $\alpha = 0$, then so is $M_0 = 1;\ \ M_n = \frac{1}{\alpha^n}\xi_1 ... \xi_n,\ n \in \mathbb{N}$.

Wald Martingale: In addition, suppose the $\xi_1, \xi_2, ...$ are also identically distributed, with moment generating function

$$\phi(\theta) := \mathbb{E}(e^{\theta \xi_1}), \ \ \theta \in \mathbb{R}$$

well-defined. Then

$$W_0 = 0; \ \ W_n = \frac{e^{\theta S_n}}{\phi^n(\theta)}; \ \ n \in \mathbb{N}$$

is a martingale.

Convexity: Suppose $\mathcal{X} = \{X_n\}_{n \in \mathbb{N}_0}$ is a martingale (resp., submartingale) with $\mathbb{E}|f(X_n)| < \infty$, $\forall n \in \mathbb{N}_0$ for some $f : \mathbb{R} \to \mathbb{R}$ convex (resp., convex increasing). Then $f(\mathcal{X})$ is a submartingale.

5.2 Fundamental Results

Here is an important result. Its continuous-time analogue is fundamental.

Doob Decomposition: Every submartingale $\mathcal{X} = (X_n)_{n \in \mathbb{N}_0}$ can be written as $X_n = M_n + A_n$, $n \in \mathbb{N}_0$ with $\mathcal{M} = \{M_n\}_{n \in \mathbb{N}_0}$ a martingale and $\mathcal{A} = A_n$ nondecreasing:

$$0 \leq A_0 \leq A_1 \leq A_2 \leq ... \leq A_n \leq A_{n+1} \leq ...$$

This $\mathcal{A}$ can actually be chosen predictable; and with this proviso, the decomposition is unique.

Proof: Define $A_0 := 0$, $A_{n+1} := \sum_{k=0}^{n} [\mathbb{E}(X_{k+1}|\mathcal{F}_k) - X_k]$ $(n \in \mathbb{N}_0)$ obviously increasing, predictable. Then $M_n := X_n - A_n,\ n \in \mathbb{N}_0$ is a martingale; indeed, we have $M_{n+1} - M_n = X_{n+1} - \mathbb{E}(X_{n+1}|\mathcal{F}_n)$.

With two such decompositions we have $X_n = M'_n + A'_n = M_n" + A_n"$, $n \in \mathbb{N}_0$, so $Z_n := M'_n - M_n" = A_n" - A'_n$, $n \in \mathbb{N}_0$ is both predictable and a martingale; therefore constant. Bust this constant is $Z_0 = A_0" - A'_0 = 0 - 0 = 0$; uniqueness. ■

Without predictability, uniqueness fails. We shall see this very vividly when we study square-integrable martingales.

Here is a very important notion, that will stay with us from now on. One of its incarnations is the stochastic integral of the Itô Calculus.

<u>Transform:</u> For random sequences $\mathcal{M} = \{M_n\}_{n \in \mathbb{N}_0}$, $\Theta = \{\theta_n\}_{n \in \mathbb{N}_0}$ adapted and predictable, respectively, we call the sequence $\mathcal{I} = \Theta \cdot \mathcal{M}$, defined by

$$I_0 = 0; \quad I_n = \sum_{k=1}^{n} \theta_k (M_k - M_{k-1}), \;\; n \in \mathbb{N};$$

the transform of $\mathcal{M}$ by Θ.

Proposition 5.1 (Stability of Martingales under Predictable Transform) *With $\mathcal{M}$, Θ above, suppose $\mathbb{E}(|\theta_k(M_k - M_{k-1})|) < \infty, \;\; \forall k \in \mathbb{N}$. Then $\mathcal{I} = \Theta \cdot \mathcal{M}$ is a*

- *martingale, if $\mathcal{M}$ is a martingale;*
- *supermartingale (resp, submartingale), if $\Theta \geq 0$ and $\mathcal{M}$ is a supermartingale (resp, submartingale).*

Proof: Follows directly from

$$\mathbb{E}(I_{n+1}|\mathcal{F}_n) - I_n = \mathbb{E}[\theta_{n+1}(M_{n+1} - M_n)|\mathcal{F}_n] = \theta_{n+1}\mathbb{E}[(M_{n+1} - M_n)|\mathcal{F}_n].$$

■

Proposition 5.2 (Stability of Martingales under Stopping) *If $(X_n)_{n \in \mathbb{N}_0}$ is a (super)(sub)martingale, then so is $(X_{T \wedge n})_{n \in \mathbb{N}_0}$ for any stopping time T.*

Proof: One way to stop a sequence, is to "freeze" its future increments:

$$X_{T \wedge n} = X_0 + \sum_{k=1}^{n} (X_k - X_{k-1}) \underbrace{\mathcal{I}_{T \geq k}}_{\theta_k}, \;\; n \in \mathbb{N}_0.$$

But this is the transform of $\mathcal{X}$ via the sequence $\Theta = \{\theta_k\}_{k \in \mathbb{N}}$ with $0 \leq \theta_k := \mathcal{I}_{T \geq k} = 1 - \mathcal{I}_{T < k} = \mathcal{I}_{T \geq k-2}$ predictable! The claim follows from the precious proposition. ■

5.3 Optimal Sampling

We have seen already that a martingale $(X_n)_{n\in\mathbb{N}_0}$ has a constant expectations:

$$\mathbb{E}(X_n) = \mathbb{E}(X_0), \quad \forall n \in \mathbb{N}.$$

This is, in a very real sense, a conservation law.

The question then arises: does this property extend to stopping times? That is, if T is a stopping time if the underlying filtration for with $\mathbb{E}(X_T)$ can be defined well, do we have

$$\mathbb{E}(X_T) = \mathbb{E}(X_0)? \tag{5.2}$$

It does not take long, to realize that this does not always hold. Take, for instance, the simple symmetric random walk on the integer lattice, started at $X_0 = 0$, and wait until the first time T it hits the level 1. We have seen already (and we shall prove again presently, by different techniques) theta $\mathbb{P}(T < \infty) = 1$, thus $\mathbb{P}(X_T = 1) = 1$. But then $1 = \mathbb{E}(X_T \neq \mathbb{E}(X_0) = 0)$, defeating the conjecture.

It becomes clear now that, if we want (5.2) to work, we need to impose conditions. Either on the stopping time, or on the martingale, or on both.

Theorem 5.3 (Doob's Optimal Sampling (Baby OST)) *On a given filtered probability space, consider a supermatingale $\mathcal{X} = \{X_n\}_{n\in\mathbb{N}_0}$ and a stopping time T. We have then*

$$\mathbb{E}(X_T) = \mathbb{E}(X_0),$$

provided that either

(i) T is bounded (i.e., $\mathbb{P}(T \leq m) = 1$, for some $m \in \mathbb{N}$); or

(ii) X is bounded (i.e., $\mathbb{P}(|X_N(\omega) \leq K|, \forall n \in \mathbb{N}_0) = 1$ for some $k \in (0,\infty)$); or

(iii) $\mathbb{E}(T) < \infty$ and X has bounded increments (i.e., $\mathbb{P}(|X_n(\omega) - X_{n-1}(\omega)| \leq K, \forall n \in \mathbb{N}) = 1$).

And if $\mathcal{X}$ is a martingale, the display becomes $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

Proof: From the Proposition (5.2), we have for every $n \in \mathbb{N}$:

$$\mathbb{E}(X_{T\wedge n} - X_0) \leq 0. \tag{5.3}$$

For (i), we can take $n = m$ and be done.

For (ii), we can let $n \to \infty$ in (5.3) and appeal to the DCT.

And for (iii), we write

$$|X_{T\wedge n} - X_0| \leq \sum_{k=1}^{T\wedge n} |X_k - X_{k-1}| \leq KT, \quad \mathbb{P}\text{-a.e.};$$

and because $\mathbb{E}(T) \leq \infty$, the DCT applies again, and leads to the result upon letting $n \to \infty$, and leads to the result in (5.3). ∎

There is no telling how for one can go using just this very humble result; there are fancier versions, of course, but this is already a gem.