STATS 310B: Theory of Probability II	**Winter 2016/17**
Lecture 11: February 14	
Lecturer: Sourav Chatterjee	*Scribes: Kenneth Tay*

11.1 Square Integrable Martingales

Definition 11.1 *Let $\{Z_n, \mathcal{F}_n\}_{n\geq 1}$ be a martingale. It is called* ***square-integrable*** *is $\mathbb{E}Z_n^2 < \infty$ for all n.*

Definition 11.2 *Let Z_n be a square-integrable martingale, and let $X_n = Z_n - Z_{n-1}$. (Put $Z_0 = \mathbb{E}Z_1$.) Define $\sigma_n^2 = \mathbb{E}[X_n^2 \mid \mathcal{F}_{n-1}]$. It can be checked that $Z_n^2 - \sum_{i=1}^n \sigma_i^2$ is a martingale.*

$\{A_n\} := \left\{\sum_{i=1}^n \sigma_i^2\right\}$ is called the ***predictable quadratic variation*** *of $\{Z_n\}$. Note that A_n is $\mathcal{F}_{n-1}$-measurable, non-negative and increasing.*

Definition 11.3 *$Z_n^2 = (Z_n^2 - A_n) + A_n$ is called the* ***Doob decomposition*** *of Z_n^2.*

More generally, if $\{Y_n\}$ is a submartingale, let $A_n = \sum_{j=1}^n \mathbb{E}[Y_j - Y_{j-1} \mid \mathcal{F}_{j-1}]$. A_n is a non-negative, increasing predictable process. Then $Y_n - A_n$ is a martingale. $Y_n = (Y_n - A_n) + A_n$ is the ***Doob decomposition*** *of Y_n.*

That is to say, every submartingale can be decomposed into the sum of a martingale and a non-negative increasing predictable process.

Note that $\mathbb{E}A_n = \mathbb{E}[Z_n^2]$ (as $Z_n^2 - A_n$ is a mean 0 martingale). But A_n is non-negative and increasing, so

$$\sup E[Z_n^2] = \lim_{n\to\infty} \mathbb{E}[Z_n^2] = \lim_{n\to\infty} \mathbb{E}A_n = \mathbb{E}\left[\sum_{j=1}^\infty \sigma_j^2\right].$$

Theorem 11.4 *For a square-integrable martingale $\{Z_n\}$, $\lim Z_n$ exists and is finite a.s. on the set $\left\{\sum_{n=1}^\infty \sigma_n^2 < \infty\right\}$.*

Proof: Take some $a > 0$, let $\tau = \inf\{n : A_{n+1} > a\}$. By Wald's Lemma, $\mathbb{E}[Z_{\tau\wedge n}^2] = \mathbb{E}[A_{\tau\wedge n}]$ for all n.

By definition of τ, $\mathbb{E}[A_{\tau\wedge n}] \leq a$. So

$$\mathbb{E}|Z_{\tau\wedge n}| \leq \sqrt{\mathbb{E}[Z_{\tau\wedge n}^2]} \leq \sqrt{a}$$

for all n. $Z_{\tau\wedge n}$ is a martingale which is bounded above by a constant, so $\lim Z_{\tau\wedge n}$ exists (by Martingale Convergence Theorem) and is finite a.s. (Fatou's Lemma).

So, on the set $\{\tau = \infty\}$, $\lim Z_n$ exists and is finite a.s. But

$$\{\tau = \infty\} = \left\{ \sum_{j=1}^{\infty} \sigma_j^2 \le a \right\}.$$

Take union over $a = 1, 2, \dots$ to get the required result. ■

11.2 Ergodic Theory

Definition 11.5 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A measurable map $\varphi : \Omega \to \Omega$ is called a* ***measure-preserving transform*** *if $P(\varphi^{-1}(A)) = P(A)$ for all $A \in \mathcal{F}$.*

We will denote $\varphi^n := \varphi \circ \dots \circ \varphi$. If φ is measure-preserving, so is φ^n for every n.

11.2.1 Examples of measure-preserving transforms

- Let $(\Omega, \mathcal{F}, P) = ([0,1), \text{Borel}, \text{Leb})$, $\theta \in \Omega$, and let $\varphi(x) = x + \theta \pmod 1$. Take any $A \in \mathcal{F}$. Let $B = \{x \in A : x - \theta \in [0,1)\}$, $C = \{x \in A : x - \theta + 1 \in [0,1)\}$. Then $B \cap C = \emptyset$, $B \cup C = A$, so $P(A) = P(B) + P(C)$. We then have

$$\begin{aligned}\varphi^{-1}(A) &= \{x \in [0,1) : x + \theta \in A\} \cup \{x \in [0,1) : x + \theta - 1 \in A\} \\ &= \text{translate of } B + \text{translate of } C,\end{aligned}$$

  and hence has the same measure as A. Thus φ is measure-preserving.

- Same probability space as before, but $\varphi(x) = 2x \pmod 1$. Then

$$\begin{aligned}\varphi^{-1}(A) &= \{x \in [0,1) : 2x \bmod 1 \in A\} \\ &= \left\{x : x \in \frac{1}{2}A \cap [0, 1/2)\right\} \text{ or } \left\{x : x \in \left(\frac{1}{2} + \frac{1}{2}A\right) \cap [1/2, 1)\right\}.\end{aligned}$$

  The 2 sets on the RHS are disjoint and each have measure $\frac{1}{2}P(A)$. Thus φ is measure-preserving.

- **(Bernoulli Shift.)** Let $\Omega = \{0,1\}^{\mathbb{N}}$, where $\mathbb{N} = \{0, 1, \dots\}$. Let $\mathcal{F}$ be the product σ-algebra, i.e. the smallest σ-algebra under which all coordinate maps are measurable. Let P be the law of an infinite i.i.d. Bernoulli(1/2) sequence. Define the Bernoulli shift operator

$$\varphi((\omega_0, \omega_1, \dots)) := (\omega_1, \omega_2, \dots).$$

  Then

$$\begin{aligned}P(\varphi^{-1}(A)) &= P((X_1, X_2, \dots) \in A) \\ &= P((X_0, X_1, \dots) \in A) \qquad \text{(as the 2 infinite sequences have the same law)} \\ &= P(A).\end{aligned}$$

- More generally, let $X_0, X_1, \dots$ be i.i.d. random variables taking values in a set S. Then we can take $S = \Omega^{\mathbb{N}}$, $\mathcal{F}$ to be the product σ-algebra, P to be the law of $(X_0, X_1, \dots)$, φ to be the shift operator. Then φ is measure-preserving.

Definition 11.6 *Let φ be measure-preserving on $(\Omega, \mathcal{F}, P)$. Let $\mathcal{I} = \{A \in \mathcal{F} : \varphi^{-1}(A) = A\}$. Then $\mathcal{I}$ is a σ-algebra, and is called the* ***invariant σ-algebra of*** *φ.*

Definition 11.7 *A measure-preserving transform φ is called* ***ergodic*** *if for all $A \in \mathcal{I}$, $P(A) = 0$ or 1.*

Proposition 11.8 *The Bernoulli shift operator is ergodic.*

Proof: Let $Y_n(\omega) = \omega_n$ be the coordinate maps. Then $\omega = (\omega_0, \omega_1, \dots) = (Y_0(\omega), \dots)$. The Y_n's are i.i.d. Bernoulli(1/2) variables.

Take any $A \in \mathcal{I}$. Then

$$\begin{aligned} A = \varphi^{-1}(A) &= \{\omega : \varphi(\omega) \in A\} \\ &= \{\omega : (\omega_1, \omega_2, \dots) \in A\} \\ &= \{(Y_1, Y_2, \dots) \in A\} \\ &\in \sigma(Y_1, Y_2, \dots). \end{aligned}$$

But similarly, $A = \varphi^{-1}(A)$ implies $A = \varphi^{-n}(A)$ for all n, which in turn implies that $A = \sigma(Y_n, Y_{n+1}, \dots)$ for all n. Thus A belongs to the tail σ-algbebra of $Y_0, Y_1, \dots$. By Kolmogorov's 0-1 Law, $P(A)$ must be 0 or 1, and so φ is ergodic.

■

Theorem 11.9 *Let $\Omega = [0, 1)$, $\varphi(x) = x + \theta \pmod 1$. Then φ is ergodic if and only if θ is irrational.*

Proof: First, suppose that θ is rational, i.e. $\theta = m/n$ with $n > m \geq 0$. Take a set $B \subseteq [0, 1/n)$ such that $0 < P(B) < \frac{1}{n}$. Let $A = \bigcup_{k=0}^{n-1} \left(B + \frac{k}{n}\right)$.

Then $0 < P(A) < 1$. We can also check that $\varphi^{-1}(A) = A$. Hence φ is not ergodic.

Next, suppose that θ is irrational. Take any $A \in \mathcal{B}([0, 1))$ such that $\varphi^{-1}(A) = A$. We want to show that $P(A) \in \{0, 1\}$, where P is the Lebesgue measure on $[0, 1)$.

Plan of the proof: Consider the function 1_A on $[0, 1)$ and show that 1_A is almost surely a constant.

Consider the Fourier coefficients of this function, $c_k = \int_0^1 1_A(x) e^{-2\pi i k x} dx$ for $k \in \mathbb{Z}$. Fact from Fourier analysis: $\lim_{n\to\infty} \sum_{k=-n}^{n} c_k e^{2\pi i k x} \to 1_A(x)$ for almost all $x \in [0, 1)$. With this fact, it suffices to show that $c_k = 0$ for all non-zero k.

For $k \neq 0$, we have

$$\begin{aligned}
c_k &= \int_0^1 1_A(x) e^{-2\pi i k x} dx \\
&= \int_0^1 1_A(\varphi(x)) e^{-2\pi i k x} dx && \text{(since } 1_A \circ \varphi = 1_A\text{)} \\
&= \int_0^{1-\theta} 1_A(x+\theta) e^{-2\pi i k x} dx + \int_{1-\theta}^1 1_A(x+\theta-1) e^{-2\pi i k x} dx \\
&= \int_\theta^1 1_A(y) e^{-2\pi i k (y-\theta)} dy + \int_0^\theta 1_A(y) e^{-2\pi i k (y+1-\theta)} dy \\
&= e^{2\pi i k \theta} \int_0^1 1_A(y) e^{-2\pi i k y} dy \\
&= e^{2\pi i k \theta} c_k.
\end{aligned}$$

Since θ is irrational, $\exp(2\pi i k \theta)$ can never be 1 if $k \neq 0$. Thus, $c_k = 0$ if $k \neq 0$.

∎