Class Notes on Martingales and Related Topics

(Based on Çınlar's *Probability and Stochastics*)

Simran Bilkhu

August 22, 2025

Contents

1 Martingales

1.1 Basic Definitions

Definition 1.1 (Sub-, Super-, and Martingales)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{F} = (\mathcal{F}_t)_{t\in T}$ a filtration, and $X = (X_t)_{t\in T}$ a real–valued, $\mathcal{F}$–adapted process with $\mathbb{E}|X_t| < \infty$ for all t. Then

- X is an $\mathcal{F}$*–submartingale* if $\mathbb{E}[X_t - X_s \mid \mathcal{F}_s] \geq 0$ for all $s < t$.
- X is an $\mathcal{F}$*–supermartingale* if $-X$ is an $\mathcal{F}$–submartingale.
- X is an $\mathcal{F}$*–martingale* if $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$ for all $s < t$.

Remark 1.2. For submartingales the defining inequality can be restated as $\mathbb{E}[X_t \mid \mathcal{F}_s] \geq X_s$. Analogous rewrites hold for supermartingales and martingales.

Example 1.3 (A Gambler's Fortune)**.** *Let X_t denote a gambler's fortune after t rounds of a game. The gambler would hope X is a sub-martingale so that the best guess of future wealth according to the most recent information at least as large as the current fortune. In practice the game is usually unfavorable and one merely has a super martingale.*

1.2 Constructions and Elementary Lemmas

Lemma 1.4 (Convex Transformation)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be convex and X an $\mathcal{F}$–martingale such that $f(X_t) \in L^1$ for each t. Then $f \circ X$ is an $\mathcal{F}$–submartingale.*

Proof. Fix $s < t$. By Jensen's inequality for conditional expectations,

$$\mathbb{E}\big[f(X_t) \mid \mathcal{F}_s\big] \geq f\big(\mathbb{E}[X_t \mid \mathcal{F}_s]\big) = f(X_s),$$

proving the submartingale property. □

Lemma 1.5 (Change of Filtration)**.** *Let X be a martingale with respect to a filtration $\mathcal{F}$ and let $\mathcal{G}$ be the (smaller) filtration generated by X. Then X is also a martingale with respect to $\mathcal{G}$.*

Proof. Because X is $\mathcal{G}$–adapted and integrable, it suffices to show the martingale equality. For $s < t$,

$$\mathbb{E}[X_t - X_s \mid \mathcal{G}_s] = \mathbb{E}\big[\mathbb{E}[X_t - X_s \mid \mathcal{F}_s] \mid \mathcal{G}_s\big] = \mathbb{E}[0 \mid \mathcal{G}_s] = 0.$$

□

Lemma 1.6 (Discrete–Time Characterisation)**.** *Assume $T = \mathbb{N}$. A process X is a martingale iff*

$$\mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n] = 0 \quad (\forall n \in \mathbb{N}).$$

Proof. ($\Rightarrow$) follows immediately from the martingale equality. Conversely, if the displayed conditional expectations vanish then for $s < t$

$$X_t - X_s = \sum_{k=s}^{t-1}(X_{k+1} - X_k), \qquad \mathbb{E}[X_t - X_s \mid \mathcal{F}_s] = \sum_{k=s}^{t-1}\mathbb{E}[X_{k+1} - X_k \mid \mathcal{F}_s] = 0,$$

so $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$. □

2 Uniform Integrability

2.1 Definition and Basic Facts

Definition 2.1 (Uniform Integrability)**.** A family $\mathcal{K} \subset L^1$ is *uniformly integrable (UI)* if

$$k(b) := \sup_{X \in \mathcal{K}} \mathbb{E}\big[|X| \mathbf{1}_{\{|X| > b\}}\big] \xrightarrow[b \to \infty]{} 0.$$

Remark 2.2. If $\mathcal{K}$ is dominated by an integrable r.v. Z (i.e. $|X| \leq Z$ a.s. for all $X \in \mathcal{K}$), then $\mathcal{K}$ is UI by the dominated convergence theorem.

2.2 Equivalent Conditions

Theorem 2.3 (UI Equivalences)**.** *For a collection $\mathcal{K}$ of random variables.*

(a) *$\mathcal{K}$ is uniformly integrable.*

(b) $h(b) := \sup_{X \in \mathcal{K}} \int_b^\infty \mathbb{P}(|X| > y)\, dy \xrightarrow[b \to \infty]{} 0.$

(c) *There exists an increasing convex* $f : \mathbb{R}_+ \to \mathbb{R}_+$ *with* $\lim_{x\to\infty} \dfrac{f(x)}{x} = \infty$ *such that* $\sup_{X\in\mathcal{K}} \mathbb{E}[f(|X|)] < \infty$.

Proof. $(a\Rightarrow b)$ Using the identity $|X|\mathbf{1}_{\{|X|>b\}} = \int_b^{|X|} dy\, \mathbf{1}_{\{|X|>y\}}$ and Fubini's theorem,

$$\mathbb{E}[|X|\mathbf{1}_{\{|X|>b\}}] = \int_b^\infty \mathbb{P}(|X| > y)dy \geq \int_b^\infty \mathbb{P}(|X| > y)dy.$$

Taking suprema yields $h(b) \leq k(b) \to 0$.

$(b\Rightarrow c)$ Because $h(b) \to 0$ we can pick $0 < b_0 < b_1 < \cdots \to \infty$ with $h(b_n) \leq 2^{-n}h(0)$. Define the step function $g(x) = \sum_{n\geq 0} \mathbf{1}_{[b_n,\infty)}(x)$ and its antiderivative $f(x) = \int_0^x g(y)dy$. Then f is increasing, convex, and $f(x)/x \to \infty$. To emphasize the last two points visually:

Figure 1 illustrates this when $b_n = n$.

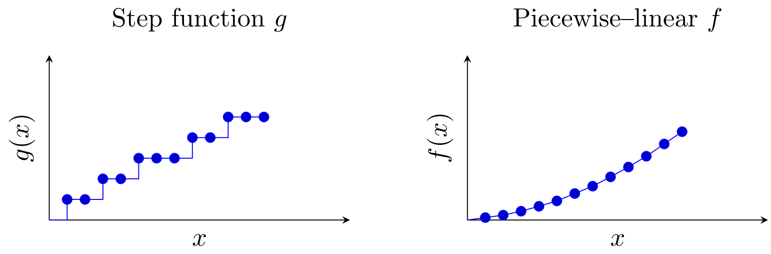

Figure 1: Left: g jumps by b_n at each integer. Right: f integrates g, forming increasingly steep linear segments, with slopes equal to $g(x)$. Convexity is evident as slopes grow without bound.

This polyline picture makes the inequality $f(x)/x \to \infty$ evident: since slopes diverge, any line through the origin eventually lies below f.

Moreover,

$$\mathbb{E}[f(|X|)] = \sum_{n\geq 0} \mathbb{E} \int_{b_n}^\infty \mathbb{I}_{(|X|>y)} dy \leq \sum_{n\geq 0} h(b_n) \leq 2h(0).$$

Supremum over X gives (c).

$(c\Rightarrow a)$ Let f be as described. Then f is a borel function (since f is increasing). Then,

$$|X|\mathbf{1}_{\{|X|>b\}} \leq f(|X|)\frac{1_{\{|X|>b\}}}{f(|X|)} \leq f \circ |X| \sup_{y>b} \frac{y}{f(y)}$$

By taking expectations, and suprema, we have:

$$k(b) \leq \sup_{X\in\mathcal{K}} \mathbb{E}[f(|X|)] \cdot \sup_{y>b} \frac{y}{f(y)} \xrightarrow[b\to\infty]{} 0$$

since $f(x)/x \to \infty$. ◻

Proposition 2.4. *Let* $Z \in L^1$ *and set* $X_t = \mathbb{E}[Z \mid \mathcal{F}_t]$. *Then* $\{X_t : t \in T\}$ *is a uniformly integrable martingale.*

Proof. Consider the collection $\mathcal{K}_Z = \{Z\}$. By theorem 2.3, there exists a convex function f such that $\mathbb{E}f(Z) < \infty$. We show that this means that

$$\sup_{X_t\in\mathcal{K}} \mathbb{E}\left(f|X_t|\right) < \infty$$

and apply theorem 2.3 to conclude that the collection is indeed uniformly integrable. Indeed,

$$\mathbb{E}\big[f\big(|\mathbb{E}(Z \mid \mathcal{F}_t)|\big)\big] \;\leq\; \mathbb{E}\big[f\big(\mathbb{E}(|Z| \mid \mathcal{F}_t)\big)\big] \qquad (|\cdot| \text{ inside expectation})$$

$$\leq \mathbb{E}\big[\mathbb{E}\big[f(|Z|) \mid \mathcal{F}_t\big]\big] \qquad \text{(Jensen's inequality)}$$

$$= \mathbb{E}\big[f(|Z|)\big] < \infty. \qquad \text{(By the tower property)}$$

□

3 Stochastic Kernels

3.1 Definition

Definition 3.1. Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. A *Stochastic kernel* (or transition kernel) from E to F is a map $K : E \times \mathcal{F} \to [0, 1]$ such that

(i) For each fixed $x \in E$, $A \mapsto K(x, A)$ is a probability measure on $(F, \mathcal{F})$.

(ii) For each fixed $A \in \mathcal{F}$, the function $x \mapsto K(x, A)$ is $\mathcal{E}$–measurable.

Intuitively $K(x, \cdot)$ is the 'law' of the next state given the current state x. Kernels generalise stochastic matrices to arbitrary measurable spaces.

Remark 3.2. In the above definition, if $(E, \mathcal{E}) = (F, \mathcal{F})$, K is called a Markov kernel.

3.2 A Physical Analogy

Imagine you are standing at position x on a straight road looking at a passing car of fixed length L (Figure 2). The set A represents angles (or retinal positions) under which portions of the car are visible. The value $K(x, A)$ can be interpreted as the *fraction of the car's length that you see under angles in* A, i.e. the 'apparent size' of A from viewpoint x. As you move, the measure $K(x, \cdot)$ changes continuously: nearby observers assign similar 'size' to any fixed angular window A. Requirement (ii) above—measurability in x— makes sure that this dependence is somewhat smooth.

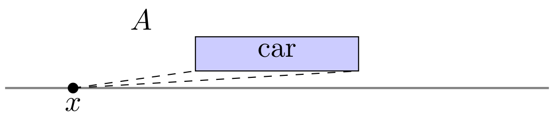

Figure 2: Visualising $K(x, A)$: the portion of the car seen from position x within viewing angles A. Moving the observer changes the measure smoothly.

Remark 3.3. We use the notation $Kf(x)$ as a sort of integration oparator with respect to the measure $K(x, \cdot)$:

$$Kf(x) = \int_E K(x, dy) f(y)$$

Remark 3.4. In finite, discrete cases where $E = F = \{1, \ldots, d\}$, a kernel is just a stochastic matrix $P = (p_{ij})$ and the above notation reduces to matrix vector multiplication.

4 Markov Chains

4.1 Definitions

Definition 4.1 (Finite State Space)**.** Given a transition matrix $P = (p_{ij})_{i,j \in E}$ on a finite set E, a process $X = (X_n)_{n \in \mathbb{N}}$ is a *Markov chain with transition matrix* P if

$$\mathbb{P}\big(X_{n+1} = j \mid X_0 = i_0, \ldots, X_n = i\big) = p_{ij} \quad (\forall n, i, j).$$

Definition 4.2 (General State Space)**.** Let $(E, \mathcal{E})$ be measurable and P a Markov kernel. An E–valued process X is a *Markov chain with kernel* P if X is adapted to some filtration $\mathcal{F}$ and, for every bounded measurable $f \geq 0$,

$$\mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n] = (Pf)(X_n), \qquad n \in \mathbb{N}.$$

4.2 Martingale Characterisation

4.3 Definitions

Let $(E, \mathcal{E})$ be a measurable space and P a Markov kernel. A process $X = (X_n)_{n \in \mathbb{N}}$ adapted to a filtration $\mathcal{F} = (\mathcal{F}_n)$ is a *Markov chain with transition kernel P (with respect to $\mathcal{F}$)* if for every bounded measurable $f \geq 0$,

$$\mathbb{E}\big[f(X_{n+1}) \mid \mathcal{F}_n\big] = (Pf)(X_n), \qquad n \in \mathbb{N}.$$

In the finite–state case this reduces to the familiar transition–matrix formulation $\mathbb{P}(X_{n+1} = j \mid X_n = i) = p_{ij}$.

4.4 Martingale Characterisation

The following result gives a concise martingale test for the Markov property.

Theorem 4.3 (Martingale Characterisation)**.** *Let X be adapted to $\mathcal{F}$ and let P be a Markov kernel on $(E, \mathcal{E})$. Then X is a Markov chain with kernel P (with respect to $\mathcal{F}$) if and only if for every bounded $f \in \mathcal{E}_+$,*

$$M_n \ := \ \sum_{m=0}^{n} f(X_m) \ - \sum_{m=0}^{n-1} (Pf)(X_m), \qquad n \in \mathbb{N},$$

is an $\mathcal{F}$–martingale.

Proof. (⇒) Assume X has the Markov property. For $n \geq 0$,

$$M_{n+1} - M_n \ = \ f(X_{n+1}) - (Pf)(X_n).$$

Taking conditional expectation with respect to $\mathcal{F}_n$ and using the Markov property yields

$$\mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] = \mathbb{E}\big[f(X_{n+1}) \mid \mathcal{F}_n\big] - (Pf)(X_n) = (Pf)(X_n) - (Pf)(X_n) = 0.$$

Which means that M martingale.

(⇐) Conversely, suppose (M_n) is a martingale for every bounded f. From the displayed decomposition we again have

$$\mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] = \mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n] - (Pf)(X_n) = 0,$$

which forces $\mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n] = (Pf)(X_n)$. Since this holds for all bounded f, X satisfies the Markov property. □

4.5 Definition and Properties

Let $\mathcal{F}$ be a filtration over $\mathbb{R}_+$. Let $W = (W_t)_{t \in \mathbb{R}_+}$ be a continuous process with state space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and starting point $W_0 = 0$.

Definition 4.4. The continuous process W is called a Wiener process with respect to $\mathcal{F}$ if it is adapted to $\mathcal{F}$ and

$$\mathbb{E}_s f\left(W_{s+t} - W_s\right) = \int_{\mathbb{R}} dx \frac{1}{\sqrt{2\pi t}} e^{-x^2 \frac{1}{2t}} f\left(x\right).$$

for all $s, t \in \mathbb{R}_+$ and for all positive Borel functions f on $\mathbb{R}$

The above definition intales 2 things. First, the increments of the process are independent of $\mathcal{F}_f$. Second, the increments are normally distribution with mean zero and variance equal to the time difference.

4.6 Martingale Characterization

Proposition 4.5. *Let $\mathcal{F}$ be a filtration. A continuous process W is a Wiener process with respect to $\mathcal{F}$ if and only if, for every $r \in \mathbb{R}$,*

$$M_t \;:=\; \exp\bigl(rW_t - \tfrac{1}{2}r^2 t\bigr), \qquad t \geq 0,$$

is an $\mathcal{F}$–martingale.

Proof. Necessity. Suppose W is Wiener. Fix $r \in \mathbb{R}$ and $0 \leq s < t$. By independent increments,

$$W_t - W_s \;\sim\; \mathcal{N}\bigl(0, t-s\bigr) \quad \text{and is independent of } \mathcal{F}_s.$$

We note

$$\begin{aligned}
\mathbb{E}_s\left(M_t/M_s\right) &= \mathbb{E}_s\bigl[\exp\bigl(r(W_t - W_s) - \tfrac{1}{2}r^2(t-s)\bigr)\bigr] \\
&= \mathbb{E}_s\bigl[\exp\bigl(r(W_t - W_s)\bigr]\exp\bigl(-\tfrac{1}{2}r^2(t-s)\bigr) \\
&= 1 \quad \text{substituting in the MGF of a normal distribution with mean 0 and variance } (t-s)
\end{aligned}$$

Therefore

$$\mathbb{E}_s[M_t] = M_s\,\mathbb{E}_s\Bigl[\exp\bigl(r(W_t - W_s) - \tfrac{1}{2}r^2(t-s)\bigr)\Bigr] = M_s,$$

so (M_t) is a martingale.

Sufficiency. Conversely, assume (M_t) is a martingale for every r. Then the above holds for all $0 \leq s < t$ and $r \in \mathbb{R}$. Necessarily, this implies:

$$\mathbb{E}_s \exp\bigl(r(W_{s+t} - W_s)\bigr) = \exp\bigl(\tfrac{1}{2}r^2 t\bigr). \tag{*}$$

The right side is exactly the MGF of $\mathcal{N}(0,t)$, so $(*)$ implies $W_{s+t} - W_s \sim \mathcal{N}(0,t)$ and is independent of $\mathcal{F}_s$. $\square$

5 Poisson Processes

Definition 5.1 (Counting processes)**.** $N = (N_t)_{t\in\mathbb{R}_+}$ is a counting process if it has state space $\left(\mathbb{N}, 2^{\mathbb{N}}\right)$ whose path $t \to N_t\left(\omega\right)$ starts from $N_0\left(\omega\right) = 0$, is increasing and right continuous, and increases by jumps of size one only.

Remark 5.2. $N_t\left(\omega\right)$ is equal to the number of jumps of $s \to N_s\left(\omega\right)$ in the interval $(0, t]$

Definition 5.3 (Poisson Processes)**.** The counting process N is said to be a Poisson process with rate c with respect to $\mathcal{F}$ if it is adapted to $\mathcal{F}$ and

$$\mathbb{E}_s f\left(N_{s+t} - N_s\right) = \sum_{k=0}^{\infty} \frac{e^{-ct}\left(ct\right)^k}{k!}.$$

Theorem 5.4. *Let N be a counting process. It is a Poisson process with rate c, with respect to $\mathcal{F}$, iff*

$$M_t = N_t - ct, \qquad t \in \mathbb{R}_+.$$

if an $\mathcal{F}$ martingale

Portions of these notes were prepared with assistance from ChatGPT (OpenAI, 2025). For further referense, refer to E. Cinlar, Probability and Stochastics.