36-752 Advanced Probability Overview	Spring 2018

7. Martingales

Instructor: Alessandro Rinaldo

Associated reading: Chapter 6 of Ash and Doléans-Dade; Sec 5.2, 5.3, 5.4 of Durrett.

Overview

Martingales are elegant and powerful tools to study sequences of dependent random variables. It is originated from gambling, where a gambler can adjust the bet according to the previous results. In a simple version, assume a gambler bets 1 dollar in the first game. If he wins, then he stops playing. Otherwise he doubles the bet until he wins. If each game is i.i.d coin toss with non-zero winning probability and the gambler has infinite amount of money, then he will win one dollar with probability one.

1 Martingales

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Definition 1 (Filtration and Martingales). *Let $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots$ be a sequence of sub-σ-field's of $\mathcal{F}$. We call $\{\mathcal{F}_n\}_{n=1}^\infty$ a* filtration *. If $X_n : \Omega \to I\!R$ is $\mathcal{F}_n$-measurable for every n, we say that $\{X_n\}_{n=1}^\infty$ is* adapted *to the filtration. If $\{X_n\}_{n=1}^\infty$ is adapted to a filtration $\{\mathcal{F}_n\}_{n=1}^\infty$, and if $\mathrm{E}|X_n| < \infty$ for all n and $\mathrm{E}(X_{n+1}|\mathcal{F}_n) = X_n$ for all n, then we say that $\{X_n\}_{n=1}^\infty$ is a* martingale *relative to the filtration. Alternatively we say that $\{(X_n, \mathcal{F}_n)\}_{n=1}^\infty$ is a martingale. If $X_n \leq \mathrm{E}(X_{n+1}|\mathcal{F}_n)$ for all n, we say that $\{(X_n, \mathcal{F}_n)\}_{n=1}^\infty$ is a* submartingale. *If the inequality goes the other way, it is a* supermartingale.

Proposition 2. *A martingale is both a submartingale and a supermartingale. $\{X_n\}_{n=1}^\infty$ is a submartingale if and only if $\{-X_n\}_{n=1}^\infty$ is a supermartingale.*

Example 3 (Sums of independent r.v.'s). *Let $\{Y_n\}_{n=1}^\infty$ be a sequence of independent random variables with finite mean. Let $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$ and $X_n = \sum_{j=1}^n Y_j$. If $\mathrm{E}(Y_n) = 0$ for every n, then $\{(X_n, \mathcal{F}_n)\}_{n=1}^\infty$ is a martingale. If $\mathrm{E}(Y_n) \geq 0$ for every n, then we have a submartingale, and if $\mathrm{E}(Y_n) \leq 0$ for every n, we have a supermartingale.*

Example 4 (R-N derivatives). *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{\mathcal{F}_n\}_{n=1}^\infty$ be a filtration. Let ν be a finite measure on $(\Omega, \mathcal{F})$ such that for every n, ν has a density X_n with* *respect to P when both are restricted to $(\Omega, \mathcal{F}_n)$. Then $\{X_n\}_{n=1}^\infty$ is adapted to the filtration. To see that we have a martingale, we need to show that for every n and $A \in \mathcal{F}_n$*

$$\int_A X_{n+1}(\omega) dP(\omega) = \int_A X_n(\omega) dP(\omega). \tag{1}$$

Since $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, each $A \in \mathcal{F}_n$ is also in $\mathcal{F}_{n+1}$. Hence both sides of Equation (1) equal $\nu(A)$.

Example 5 (Likelihood ratio – simple case). *As a more specific example of Example 4, let $\Omega = \mathbb{R}^\infty$ and let $\mathcal{F}_n = \{B \times \mathbb{R}^\infty : B \in \mathcal{B}^n\}$. That is, $\mathcal{F}_n$ is the collection of cylinder sets corresponding to the first n coordinates (the σ-field generated by the first n coordinate projection functions). Let P be the joint distribution of an infinite sequence of iid standard normal random variables. Let ν be the joint distribution of an infinite sequence of iid exponential random variables with parameter 1. For each n, when we restrict both P and ν to $\mathcal{F}_n$, ν has the density*

$$X_n(\omega) = \begin{cases} (2\pi)^{n/2} \exp\left(\sum_{j=1}^n [\omega_j^2/2 - \omega_j]\right) & \text{for } \omega_1, \ldots, \omega_n > 0, \\ 0 & \text{otherwise,} \end{cases}$$

with respect to P. It is easy to see that

$$\mathrm{E}(X_{n+1}|\mathcal{F}_n) = X_n \mathrm{E}(\sqrt{2\pi} \exp(\omega_{n+1}^2/2 - \omega_{n+1}) I_{(0,\infty)}(\omega_{n+1})) = X_n.$$

Example 6 (Likelihood ratio – general case). *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{Y_n\}_{n=1}^\infty$ be a sequence of random variables and $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$. Suppose that, for each n, $\mu_{Y_1,\ldots,Y_n}$ has a strictly positive density p_n with respect to Lebesgue measure λ^n. Let Q be another probability on $(\Omega, \mathcal{F})$ such that $Q((Y_1, \ldots, Y_n)^{-1}(\cdot))$ has a density q_n with respect to λ^n for each n. Define*

$$X_n = \frac{q_n(Y_1, \ldots, Y_n)}{p_n(Y_1, \ldots, Y_n)}.$$

It is easy to check that $\{(X_n, \mathcal{F}_n)\}_{n=1}^\infty$ is a martingale.

Example 7 (Convex transformation). *Let $\{(X_n, \mathcal{F}_n)\}_{n=1}^\infty$ be a martingale. Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function such that $\mathrm{E}[\phi(X_n)]$ is finite for all n. Define $Y_n = \phi(X_n)$. Then $\{(Y_n, \mathcal{F}_n)\}_{n=1}^\infty$ is a submartingale according to the conditional version of Jensen's inequality.*

Example 8 (Lévy martingale). *Let $\{\mathcal{F}_n\}_{n=1}^\infty$ be a filtration and let X be a random variable with finite mean. Define $X_n = \mathrm{E}(X|\mathcal{F}_n)$. By the law of total probability we have a martingale. Such a martingale is sometimes called a* Lévy martingale.

Example 9 (Gambler's ruin). *Consider Example 3 again. Think of Y_n as being the amount that a gambler wins per unit of currency bet on the nth play in a sequence of games. Let Y_0 denote the gambler's initial fortune which we can assume is a known value, and let*

$\mathcal{F}_0$ be the trivial σ-field. (We could let Y_0 be a random variable and let $\mathcal{F}_0 = \sigma(Y_0)$, but then we would also have to expand $\mathcal{F}_n$ to $\sigma(Y_0, \ldots, Y_n)$.) Suppose that the gambler devises a system for determining how much $W_n \geq 0$ to bet on the nth play. We assume that W_n is $\mathcal{F}_{n-1}$ measurable for each n. This forces the gambler to choose the amount to bet before knowing what will happen. Now, define $Z_n = Y_0 + \sum_{j=1}^{n} W_j Y_j$. Since

$$\mathrm{E}(W_{n+1}Y_{n+1}|\mathcal{F}_n) = W_{n+1}\mathrm{E}(Y_{n+1}|),$$

and $W_{n+1} \geq 0$, we have that $\mathrm{E}(W_{n+1}Y_{n+1}|\mathcal{F}_n)$ is $\geq$, $=$, or $\leq$ than $W_n Y_n$ depending on whether $\mathrm{E}(Y_{n+1})$ is $\geq$, $=$, or $\leq$ than Y_n, respectively. That is, $\{(Z_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ is a submartingale, a martingale, or a supermartingale according as $\{(Y_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ is a submartingale, a martingale, or a supermartingale. This result is often described by saying that gambling systems cannot change whether a game is favorable, fair, or unfavorable to a gambler.

Definition 10 (Previsibility). *A sequence $\{W_n\}_{n=1}^{\infty}$ of random variables such that W_n is $\mathcal{F}_{n-1}/\mathcal{B}^1$-measurable is called* previsible *. (If there is no $\mathcal{F}_0$, then require that W_1 be constant.)*

Example 11 (Martingale transform). *Let $\{(X_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ be a martingale, and let $\{W_n\}_{n=1}^{\infty}$ be previsible. Define $Z_1 = X_1$ and $Z_{n+1} = Z_n + W_{n+1}(X_{n+1} - X_n)$ for $n \geq 1$. Then Z_n is $\mathcal{F}_n/\mathcal{B}^1$-measurable and*

$$\mathrm{E}(Z_{n+1}|\mathcal{F}_n) = Z_n + W_{n+1}\mathrm{E}(X_{n+1} - X_n|\mathcal{F}_n) = Z_n,$$

for all $n \geq 1$. This makes $\{(Z_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ a martingale. This is called a martingale transform. *Example 9 is an example of this.*

Theorem 12 (Doob decomposition). *$\{(X_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ is a submartingale if and only if there is a martingale $\{(Z_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ and a nondecreasing previsible process $\{A_n\}_{n=1}^{\infty}$ with $A_1 = 0$ such that $X_n = Z_n + A_n$ for all n. The decomposition is unique (a.s.).*

Proof: For the "if" direction, notice that

$$\mathrm{E}(Z_{n+1} + A_{n+1}|\mathcal{F}_n) = Z_n + A_{n+1} \geq Z_n + A_n = X_n.$$

For the "only if" direction, Define $A_1 = 0$ and

$$A_n = \sum_{k=2}^{n} \left(\mathrm{E}(X_k|\mathcal{F}_{k-1}) - X_{k-1}\right),$$

for $n > 1$. Also, define $Z_n = X_n - A_n$. Because $\mathrm{E}(X_k|\mathcal{F}_{k-1}) \geq X_{k-1}$ for all $k > 1$, we have $A_n \geq A_{n-1}$ for all $k > 1$, so $\{A_n\}_{n=1}^{\infty}$ is nondecreasing. Also, $\mathrm{E}(X_k|\mathcal{F}_{k-1})$ is $\mathcal{F}_{n-1}/\mathcal{B}^1$-

measurable for all $1 < k \le n$, so $\{A_n\}_{n=1}^{\infty}$ is previsible. Finally, notice that

$$\begin{aligned}
\mathrm{E}(Z_{n+1}|\mathcal{F}_n) &= \mathrm{E}(X_{n+1}|\mathcal{F}_n) - A_{n+1} \\
&= \mathrm{E}(X_{n+1}|\mathcal{F}_n) - \sum_{k=2}^{n+1} [\mathrm{E}(X_k|\mathcal{F}_{k-1}) - X_{k-1}] \\
&= X_n - \sum_{k=2}^{n} [\mathrm{E}(X_k|\mathcal{F}_{k-1}) - X_{k-1}] = Z_n,
\end{aligned}$$

so Z_n is a martingale.

For uniqueness, suppose that $X_n = Y_n + W_n$ is another decomposition so that Y_n is a martingale and W_n is previsible. Then write

$$\begin{aligned}
\sum_{k=2}^{n} [\mathrm{E}(X_k|\mathcal{F}_{k-1}) - X_{k-1}] &= \sum_{k=2}^{n} [\mathrm{E}(Y_k + W_k|\mathcal{F}_{k-1}) - X_{k-1}] \\
&= \sum_{k=2}^{n} (Y_{k-1} + W_k - X_{k-1}) \\
&= \sum_{k=2}^{n} (W_k - W_{k-1}) = W_n.
\end{aligned}$$

It follows that $W_k = A_k$ a.s., hence $Y_k = Z_k$ a.s. ■

The previsible process in Theorem 12 is called the *compensator* for the submartingale.

2 Stopping Times

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

Definition 13 (Stopping times). *A positive*[1] *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. A special σ-field, $\mathcal{F}_\tau$ is defined by*

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \le k\} \in \mathcal{F}_k, \text{ for all finite } k\}.$$

If $\{X_n\}_{n=1}^{\infty}$ is adapted to the filtration and if $\tau < \infty$ a.s., then X_τ is defined as $X_{\tau(\omega)}(\omega)$. (Define X_τ equal to some arbitrary random variable X_∞ for $\tau = \infty$.)

[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 $\mathcal{F}_0 = \{\Omega, \varnothing\}$. Every martingale can also be extended by defining $X_0 = \mathrm{E}(X_1)$. For the rest of the course, we will assume that the lowest possible value for a stopping time is 1.

Example 14. *Let* $\{X_n\}_{n=1}^{\infty}$ *be adapted to the filtration and let* $\tau = k_0$*, a constant. Then* $\{\tau = n\}$ *is either* Ω *or* $\varnothing$ *and it is in every* $\mathcal{F}_n$*, so* τ *is a stopping time. Also,*

$$A \cap \{\tau \le k\} = \begin{cases} A & \text{if } k_0 \le k, \\ \varnothing & \text{if } k_0 > k. \end{cases}$$

So $A \cap \{\tau \le k\} \in \mathcal{F}_k$ *for all finite* k *if and only if* $A \in \mathcal{F}_{k_0}$*. So* $\mathcal{F}_\tau = \mathcal{F}_{k_0}$*.*

Example 15 (First passage). *Let* $\{X_n\}_{n=1}^{\infty}$ *be adapted to the filtration. Let* B *be a Borel set and let* $\tau = \inf\{n : X_n \in B\}$*. As usual,* $\inf \varnothing = \infty$*. For each finite* n*,*

$$\{\tau = n\} = \{X_n \in B\} \bigcap_{k<n} \{X_k \in B^C\} \in \mathcal{F}_n.$$

So, τ *is a stopping time.*

We can show that τ and X_τ are both $\mathcal{F}_\tau$ measurable. For example, to show that X_τ is $\mathcal{F}_\tau$ measurable, we must show that, for all $B \in \mathcal{B}^1$ $X_\tau^{-1}(B) \in \mathcal{F}$ and for all $1 \le k < \infty$, $\{\tau \le k\} \cap X_\tau^{-1}(B) \in \mathcal{F}_k$. Now,

$$X_\tau^{-1}(B) = \bigcup_{k=1}^{\infty} \left(\{\tau = k\} \cap [X_k^{-1}(B)]\right) \cup \left(\{\tau = \infty\} \cap X_\infty^{-1}(B)\right) \in \mathcal{F}.$$

This shows that X_τ is $\mathcal{F}$-measurable. Next, fix k and write

$$\{\tau \le k\} \cap X_\tau^{-1}(B) = \bigcup_{j=1}^{k} \left[X_j^{-1}(B) \cap \{\tau = j\}\right] \in \mathcal{F}_k.$$

This proves that X_τ is $\mathcal{F}_\tau$ measurable. Suppose that τ_1 and τ_2 are two stopping times such that $\tau_1 \le \tau_2$. Let $A \in \mathcal{F}_{\tau_1}$. Since $A \cap \{\tau_2 \le k\} = A \cap \{\tau_1 \le k\} \cap \{\tau_2 \le k\}$ for every event A, it follows that $A \cap \{\tau_2 \le k\} \in \mathcal{F}_k$ and $A \in \mathcal{F}_{\tau_2}$. Hence $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$. As an example, let τ be an arbitrary stopping time (not necessarily finite a.s.) and define $\tau_k = \min\{k, \tau\}$ for finite k. Then τ_k is a finite stopping time with $\tau_k \le \tau$. Hence X_{τ_k} is $\mathcal{F}_{\tau_k}$ measurable for each k and so X_{τ_k} is $\mathcal{F}_\tau$ measurable. Similarly, $\tau_k \le k$ so that $\mathcal{F}_{\tau_k} \subseteq \mathcal{F}_k$ and X_{τ_k} is $\mathcal{F}_k$ measurable.

Example 16 (Gambler's ruin). *The gambler in Example 9 can try to build a stopping time into a gambling system. For example, let* $\tau = \min\{n : Z_n \ge Y_0 + x\}$ *for some integer* $x > 0$*. This would seem to guarantee winning at least* x*. There are two possible drawbacks. One is that there may be positive probability that* $\tau = \infty$*. Even if* $\tau < \infty$ *a.s., it might require unlimited resources to guarantee that we can survive until* τ*. For example, let* $Y_0 = 0$ *and let* Y_n *have equal probability of being 1 or* -1 *all* n*. So, we stop as soon as we have won* x *more than we have lost. If we modify the problem so that we have only finite resources (say* k *units) then this becomes the classic* gambler's ruin problem. *The probability of achieving* $Z_n = x$ *before* $Z_n = -k$ *is* $k/(k + x)$*, which goes to 1 as* $k \to \infty$*. So, if we have unlimited resources, the probability is 1 that* $\tau < \infty$*, otherwise, we may never achieve the goal. If the probability of winning on each game is less than 1/2, then* $P(\tau = \infty) > 0$*.*

Suppose that we start with a martingale $\{(X_n, \mathcal{F}_n)\}_{n=1}^\infty$ and a stopping time τ. We can define

$$X_n^* = \begin{cases} X_n & \text{if } n \le \tau, \\ X_\tau & \text{if } n > \tau \end{cases} = X_{\min\{\tau,n\}}.$$

We can call this the *stopped martingale*. It turns out that $\{X_n^*\}_{n=1}^\infty$ is also a martingale relative to the filtration. First, note that $X_{\min\{\tau,n\}}$ is $\mathcal{F}_n$ measurable. Next, notice that

$$\begin{aligned} \mathrm{E}(|X_n^*|) &= \sum_{k=1}^{n-1} \int_{\{\tau=k\}} |X_k| dP + \int_{\{\tau \ge n\}} |X_n| dP \\ &\le \sum_{k=1}^{n} \mathrm{E}(|X_k|) < \infty. \end{aligned}$$

Finally, let $A \in \mathcal{F}_n$. Then $A \cap \{\tau > n\} \in \mathcal{F}_n$, so

$$\int_{A \cap \{\tau > n\}} X_{n+1} dP = \int_{A \cap \{\tau > n\}} X_n dP,$$

because $X_n = \mathrm{E}(X_{n+1}|\mathcal{F}_n)$. It now follows that

$$\begin{aligned} \int_A X_{n+1}^* dP &= \int_{A \cap \{\tau > n\}} X_{n+1} dP + \int_{A \cap \{\tau \le n\}} X_\tau dP \\ &= \int_{A \cap \{\tau > n\}} X_n dP + \int_{A \cap \{\tau \le n\}} X_\tau dP \\ &= \int_A X_n^* dP. \end{aligned}$$

It follows that $X_n^* = \mathrm{E}(X_{n+1}^*|\mathcal{F}_n)$ and the stopped martingale is also a martingale. Notice that $\lim_{n\to\infty} X_n^* = X_\tau$ a.s., if $\tau < \infty$ a.s.

Because we can use a constant stopping time to stop a martingale, it follows that martingale theorems will apply to finite sequences of random variables as well as infinite sequences.

Example 17. *Consider the stopping time in Example 16 with $x = 1$. That is τ is the first time that a gambler, betting on iid fair coin flips, wins 1 more than he/she has lost. This $\tau < \infty$ a.s. It follows that $\lim_{n\to\infty} X_n^* = X_\tau$ a.s. However, $\mathrm{E}(X_n^*) = 0$ for all n while $\mathrm{E}(X_\tau) = 1$ because $X_\tau = 1$ a.s.*

3 Optional Sampling

Let $\{(X_n, \mathcal{F}_n)\}_{n=1}$ be a martingale. Consider a sequence of a.s. finite stopping times $\{\tau_n\}_{n=1}^\infty$ such that $1 \le \tau_j \le \tau_{j+1}$ for all j. Then we can construct $\{(X_{\tau_n}, \mathcal{F}_{\tau_n})\}_{n=1}^\infty$ and ask whether or not it is a martingale. In general, an unpleasant integrability condition is needed to prove this. We shall do a simplified case.

Theorem 18 (Optional sampling theorem). *Let $\{(X_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ be a (sub)martingale. Suppose that for each n, there is a finite constant M_n such that $\tau_n \le M_n$ a.s. Then $\{(X_{\tau_n}, \mathcal{F}_{\tau_n})\}_{n=1}^{\infty}$ is a (sub)martingale.*

The unpleasant integrability condition that can replace $P(\tau_n \le M_n) = 1$ is the following: For every n,

- $P(\tau_n < \infty) = 1$,
- $\mathrm{E}(|X_{\tau_n}|) < \infty$, and
- $\liminf_{m\to\infty} \mathrm{E}(|X_m| I_{(m,\infty)}(\tau_n)) = 0$.

Proof: [Theorem 18] Without loss of generality, assume that $M_n \le M_{n+1}$ for every n. Since $\tau_n \le M_n$,

$$\mathrm{E}(|X_{\tau_n}|) = \sum_{k=1}^{M_n} \int_{\{\tau_n = k\}} |X_k| dP \le \sum_{k=1}^{M_n} \mathrm{E}(|X_k|) < \infty.$$

We already know that X_{τ_n} is $\mathcal{F}_{\tau_n}$ measurable. Let $A \in \mathcal{F}_{\tau_n}$. We need to show that $\int_A X_{\tau_{n+1}} dP (\ge) = \int_A X_{\tau_n} dP$. Write

$$\int_A [X_{\tau_{n+1}} - X_{\tau_n}] dP = \int_{A \cap \{\tau_{n+1} > \tau_n\}} [X_{\tau_{n+1}} - X_{\tau_n}] dP.$$

Next, for each $\omega \in \{\tau_{n+1} > \tau_n\}$, write

$$X_{\tau_{n+1}}(\omega) - X_{\tau_n}(\omega) = \sum_{\text{All } k \text{ such that } \tau_n(\omega) < k \le \tau_{n+1}(\omega)} [X_k(\omega) - X_{k-1}(\omega)].$$

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

$$\int_A [X_{\tau_{n+1}} - X_{\tau_n}] dP = \int_A \sum_{k=2}^{M_{n+1}} I_{\{\tau_n < k \le \tau_{n+1}\}} (X_k - X_{k-1}) dP.$$

Since $A \in \mathcal{F}_{\tau_n}$ and $\{\tau_n < k \le \tau_{n+1}\} = \{\tau_n \le k-1\} \cap \{\tau_{n+1} \le k-1\}^C$, it follows that

$$B_k = A \cap \{\tau_n < k \le \tau_{n+1}\} \in \mathcal{F}_{k-1},$$

for each k. So

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

because $X_{k-1} (\le) = \mathrm{E}(X_k | \mathcal{F}_{k-1})$ and $B_k \in \mathcal{F}_{k-1}$. ■

4 Martingale Convergence

The upcrossing lemma says that a submartingale cannot cross a fixed nondegenerate interval very often with high probability. If the submartingale were to cross an interval infinitely often, then its lim sup and lim inf would have to be different and it couldn't converge.

Lemma 19 (Upcrossing lemma). *Let $\{(X_k, \mathcal{F}_k)\}_{k=1}^n$ be a submartingale. Let $r < q$, and define V to be the number of times that the sequence $X_1, \ldots, X_n$ crosses from below r to above q. Then*

$$\mathrm{E}(V) \leq \frac{1}{q-r}\left(\mathrm{E}|X_n| + |r|\right). \tag{2}$$

We will only give an outline of the proof of Lemma 19. Let $Y_k = \max\{0, X_k - r\}$. Then V is the number of times that Y_k moves from 0 to above $q - r$, and $\{(Y_k, \mathcal{F}_k)\}_{k=1}^\infty$ is a submartingale. It is easy to see that V is at most the sum of the upcrossing increments divided by $q - r$. That is,

$$V \leq \frac{1}{q-r}\sum_{k=2}^{n}(Y_k - Y_{k-1})I_{E_k},$$

where E_k is the event that the path is crossing up at time k. Notice that $E_k \in \mathcal{F}_{k-1}$ for all k. Hence, for each $k \geq 2$,

$$\mathrm{E}([Y_k - Y_{k-1}]I_{E_k}) = \int_{E_k}(Y_k - Y_{k-1})dP = \int_{E_k}[\mathrm{E}(Y_k|\mathcal{F}_{k-1}) - Y_{k-1}]dP.$$

Because $\mathrm{E}(Y_k|\mathcal{F}_{k-1}) - Y_{k-1} \geq 0$ a.s. by the submartingale property, we can expand the integral from E_k to all of Ω to get

$$\mathrm{E}([Y_k - Y_{k-1}]I_{E_k}) \leq \int [\mathrm{E}(Y_k|\mathcal{F}_{k-1}) - Y_{k-1}]dP = \mathrm{E}(Y_k - Y_{k-1}).$$

It follows that $(q - r)\mathrm{E}(V) \leq \mathrm{E}(Y_n) - \mathrm{E}(Y_1) \leq \mathrm{E}(Y_n)$ because $Y_1 \geq 0$. Because $\max\{0, x\}$ is a convex function of x, $\mathrm{E}(Y_n) \leq \mathrm{E}(|X_n|) + r$. The full proof is in the Appendix.

Theorem 20 (Martingale convergence theorem). *Let $\{(X_n, \mathcal{F}_n)\}_{n=1}^\infty$ be a submartingale such that $\sup_n \mathrm{E}|X_n| < \infty$. Then $X = \lim_{n\to\infty} X_n$ exists a.s. and $\mathrm{E}|X| < \infty$.*

Proof: Let $X^* = \limsup_{n\to\infty} X_n$ and $X_* = \liminf_{n\to\infty} X_n$. Let $B = \{\omega : X_*(\omega) < X^*(\omega)\}$. We will prove that $P(B) = 0$. We can write

$$B = \bigcup_{r<q,\ r,q \text{ rational}} \{\omega : X^*(\omega) > q > r > X_*(\omega)\}.$$

Now, $X^*(\omega) > q > r \geq X_*(\omega)$ if and only if the values of $X_n(\omega)$ cross from being below r to being above q infinitely often. For fixed r and q, we now prove that this has probability 0;

hence $P(B) = 0$. Let V_n equal the number of times that $X_1, \ldots, X_n$ cross from below r to above q. According to Lemma 19,

$$\sup_n \mathrm{E}(V_n) \leq \frac{1}{q-r}\left(\sup_n \mathrm{E}(|X_n|) + |r|\right) < \infty.$$

The number of times the values of $\{X_n(\omega)\}_{n=1}^{\infty}$ cross from below r to above q equals $\lim_{n\to\infty} V_n(\omega)$. By the monotone convergence theorem,

$$\infty > \lim_n \mathrm{E}(V_n) = \mathrm{E}(\lim_{n\to\infty} V_n).$$

It follows that $P(\{\omega : \lim_{n\to\infty} V_n(\omega) = \infty\}) = 0$.

Since $P(B) = 0$, we have that $X = \lim_{n\to\infty} X_n$ exists a.s. Fatou's lemma says $\mathrm{E}(|X|) \leq \liminf_{n\to\infty} \mathrm{E}(|X_n|) \leq \sup_n \mathrm{E}(|X_n|) < \infty$. ∎

Example 21 (Random walk). *For the random walk martingale of Example 3, if the Y_n's are iid with finite variance σ^2, then $X_n/\sqrt{n}$ converges in distribution so X_n can't converge a.s. To check how the condition of Theorem 20 is violated, the Markov inequality says that*

$$\frac{\mathrm{E}(|X_n|)}{\sqrt{n}c} \geq P(|X_n| > c\sqrt{n}) \to 2\left[1 - \Phi\left(\frac{c}{\sigma}\right)\right],$$

for each positive c. So, eventually $\mathrm{E}(|X_n|) \geq c\sqrt{n}[1 - \Phi(c/\sigma)]$ and $\lim_{n\to\infty} \mathrm{E}(|X_n|) = \infty$. However, if $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n) < \infty$, then the condition of Theorem 20 holds. Indeed, the Basic L^2 Convergence Theorem already told us that the sum converges a.s.

Example 22 (Lévy martingale). *For the Lévy martingale of Example 8,*

$$\mathrm{E}(|X_n|) = \mathrm{E}(|\mathrm{E}[X|\mathcal{F}_n]|) \leq \mathrm{E}\mathrm{E}(|X||\mathcal{F}_n) = \mathrm{E}(|X|) < \infty,$$

for all n, so the martingale converges. In Theorem 24, we can say even more about the limit.

We need the following result of uniform integrability of Lévy martingales before we can identify the limit of a Lévy martingale. Recall that uniform integrability allows the exchange of limit and integral under finite measure (Theorem 20 of Lecture Notes Set 3).

Lemma 23. *Let $\{\mathcal{F}_n\}_{n=1}^{\infty}$ be a sequence of σ-fields. Let $\mathrm{E}(|X|) < \infty$. Define $X_n = \mathrm{E}(X|\mathcal{F}_n)$. Then $\{X_n\}_{n=1}^{\infty}$ is a uniformly integrable sequence.*

Proof: Since $\mathrm{E}(X|\mathcal{F}_n) = \mathrm{E}(X^+|\mathcal{F}_n) - \mathrm{E}(X^-|\mathcal{F}_n)$, and the sum of uniformly integrable sequences is uniformly integrable, we will prove the result for nonnegative X. Let $A_{c,n} = \{X_n \geq c\} \in \mathcal{F}_n$. So $\int_{A_{c,n}} X_n(\omega)dP(\omega) = \int_{A_{c,n}} X(\omega)dP(\omega)$. If we can find, for every $\epsilon > 0$, a C such that $\int_{A_{c,n}} X(\omega)dP(\omega) < \epsilon$ for all n and all $c \geq C$, we are done. This is achieved using absolute continuity and the detail is a homework problem. ∎

Theorem 24 (Lévy's theorem). *Let $\{\mathcal{F}_n\}_{n=1}^{\infty}$ be an increasing sequence of σ-fields. Let $\mathcal{F}_\infty$ be the smallest σ-field containing all of the $\mathcal{F}_n$'s. Let $\mathrm{E}(|X|) < \infty$. Define $X_n = \mathrm{E}(X|\mathcal{F}_n)$ and $X_\infty = \mathrm{E}(X|\mathcal{F}_\infty)$. Then $\lim_{n\to\infty} X_n = X_\infty$, a.s.*

Proof: By Lemma 23, $\{X_n\}_{n=1}^{\infty}$ is a uniformly integrable sequence. Let Y be the limit of the martingale guaranteed by Theorem 20. Since Y is a limit of functions of the X_n, it is measurable with respect to $\mathcal{F}_\infty$. It follows from uniform integrability that for every event A, $\lim_{n\to\infty} \mathrm{E}(X_n I_A) = \mathrm{E}(Y I_A)$. Next, note that, for every m and $A \in \mathcal{F}_m$,

$$\begin{aligned}
\int_A Y dP &= \lim_{n\to\infty} \int_A \mathrm{E}(X|\mathcal{F}_n) dP \\
&= \lim_{n\to\infty} \int_A X_n dP \\
&= \int_A X dP,
\end{aligned}$$

where the last equality follows from the fact that $A \in \mathcal{F}_n$ for all $n \geq m$ so $\int_A X_n dP = \int_A X dP$ because $X_n = \mathrm{E}(X|\mathcal{F}_n)$. Since $\int_A Y dP = \int_A X dP$ for all $A \in \mathcal{F}_m$ for all m, it holds for all A in the field $\mathcal{F} = \bigcup_{n=1}^{\infty} \mathcal{F}_n$. Since $|X|$ is integrable and $\mathcal{F}$ is a field, we can conclude that the equality holds for all $A \in \mathcal{F}_\infty$, the smallest σ-field containing $\mathcal{F}$. The equality $\mathrm{E}(X I_A) = \mathrm{E}(Y I_A)$ for all $A \in \mathcal{F}_\infty$ together with the fact that Y is $\mathcal{F}_\infty$ measurable is precisely what it means to say that $Y = \mathrm{E}(X|\mathcal{F}_\infty) = X_\infty$. ■

Obviously there is an analogous result for super martingales.

Lemma 25 (Nonnegative supermartingale). *Let $\{(X_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ be a nonnegative supermartingale. Then X_n converges a.s. to a random variable with finite mean.*

Proof: Let $Y_n = -X_n$. Then $\{(Y_n, \mathcal{F}_n)\}_{n=1}^{\infty}$ is a submartingale.

$$\mathrm{E}(|Y_n|) = \mathrm{E}(X_n) = \mathrm{E}[\mathrm{E}(X_n|\mathcal{F}_{n-1})] \leq \mathrm{E}(X_{n-1}).$$

It follows that $\mathrm{E}(|Y_n|) \leq \mathrm{E}(X_1) < \infty$ for all n, so Theorem 20 applies and Y_n converges a.s. Trivially $-Y_n = X_n$ also converges. ■

In Section 6 we shall see an important example where $\mathcal{F}_\infty \neq \mathcal{F}$. Before this, we shall first introduce reversed martingales.

5 Reversed Martingales

Definition 26 (Reversed Martingales). *For $n = -1, -2, \ldots$, let sub-σ-field's $\mathcal{F}_{n-1} \subseteq \mathcal{F}_n$, suppose that X_n is $\mathcal{F}_n$ measurable, $\mathrm{E}(|X_n|) < \infty$, and $\mathrm{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1}$. Then $\{(X_n, \mathcal{F}_n)\}_{n=-1}^{-\infty}$ is a* reversed martingale.

An equivalent way to think about reversed martingales is through a decreasing sequence of σ-field's $\{\mathcal{F}_n\}_{n=1}^{\infty}$ such that $\mathcal{F}_{n+1} \subseteq \mathcal{F}_n$ for $n \geq 1$. The proofs of the next two theorems are similar to the corresponding theorems for forward martingales.

Theorem 27 (Reversed martingale convergence theorem). *If* $\{(X_n, \mathcal{F}_n)\}_{n=-1}^{-\infty}$ *is a reversed martingale, then* $X = \lim_{n\to-\infty} X_n$ *exists a.s. and* $\mathrm{E}(X) = \mathrm{E}(X_{-1})$.

Proof: Just as in the proof of Theorem 20, we let V_n be the number of times that the finite sequence $X_n, X_{n+1}, \ldots, X_{-1}$ crosses from below a rational r to above another rational q (for $n < 0$). Lemma 19 says that

$$\mathrm{E}(V_n) \leq \frac{1}{q-r}\left(\mathrm{E}(|X_{-1}|) + |r|\right) < \infty.$$

As in the proof of Theorem 20, it follows that $X = \lim_{n\to-\infty} X_n$ exists with probability 1. Since $X_n = \mathrm{E}(X_{-1}|\mathcal{F}_n)$ for each $n < -1$, Lemma 23 says that

$$\mathrm{E}(X) = \lim_{n\to-\infty} \mathrm{E}(X_n) = \mathrm{E}(X_{-1}). \quad \square$$

■

Notice that reversed martingales are all of the Lévy type. Not surprisingly, there is a version of Lévy's theorem 24 for reversed martingales. We state it in terms of decreasing σ-field's.

Theorem 28 (Lévy Theorem for reversed martingales). *Let* $\{\mathcal{F}_n\}_{n=1}^{\infty}$ *be a decreasing sequence of* σ*-fields. Let* $\mathcal{F}_\infty = \bigcap_{n=1}^{\infty} \mathcal{F}_n$*. Let* $\mathrm{E}(|X|) < \infty$*. Define* $X_n = \mathrm{E}(X|\mathcal{F}_n)$ *and* $X_\infty = \mathrm{E}(X|\mathcal{F}_\infty)$*. Then* $\lim_{n\to\infty} X_n = X_\infty$ *a.s.*

Proof: It is easy to see that $\{(X_{-n}, \mathcal{F}_{-n})\}_{n=-1}^{-\infty}$ is a reversed martingale and that $\mathrm{E}(|X_1|) < \infty$. By Theorem 27, it follows that $\lim_{n\to-\infty} X_{-n} = Y$ exists and is finite a.s. To prove that $Y = X_\infty$ a.s., note that $X_\infty = \mathrm{E}(X_1|\mathcal{F}_\infty)$ since $\mathcal{F}_\infty \subseteq \mathcal{F}_1$. So, we must show that $Y = \mathrm{E}(X_1|\mathcal{F}_\infty)$. Let $A \in \mathcal{F}_\infty$. Then

$$\int_A X_n(\omega)dP(\omega) = \int_A X_1(\omega)dP(\omega),$$

since $A \in \mathcal{F}_n$ and $X_n = \mathrm{E}(X_1|\mathcal{F}_n)$. Once again, using Lemma 23, it follows that

$$\lim_{n\to\infty} \int_A X_n(\omega)dP(\omega) = \int_A Y(\omega)dP(\omega) = \int_A X_1(\omega)dP(\omega),$$

hence $Y = \mathrm{E}(X_1|\mathcal{F}_\infty)$. ■

Theorem 28 allows us to prove a strong law of large numbers that is even more general than the usual version. The greater generality comes from the fact that it applies to sequences that are not necessarily independent.

6 Exchangeability and de Finetti Theorem

A sequence of random quantities $\{X_n\}_{n=1}^{\infty}$ is *exchangeable* if, for every n and all distinct $j_1, \ldots, j_n$, the joint distribution of $(X_{j_1}, \ldots, X_{j_n})$ is the same as the joint distribution of $(X_1, \ldots, X_n)$.

Example 29 (Conditionally iid random quantities). *Let $\{X_n\}_{n=1}^{\infty}$ be conditionally iid given a σ-field $\mathcal{C}$. Then $\{X_n\}_{n=1}^{\infty}$ is an exchangeable sequence. The result follows easily from the fact that*

$$\mu_{X_{j_1},\ldots,X_{j_n}|\mathcal{C}} = \mu_{X_1,\ldots,X_n|\mathcal{C}}, \quad a.s.$$

Example 30. *Let $\{X_n\}_{n=1}^{\infty}$ be Bernoulli random variables such that*

$$P(X_1 = x_1, \ldots, X_n = x_n) = \frac{1}{(n+1)\binom{n}{y}},$$

where $y = \sum_{j=1}^{n} x_j$. One can show that this specifies consistent joint distributions. One can also check that the X_n's are not independent.

$$\begin{aligned} P(X_1 = 1) &= \frac{1}{2}, \\ P(X_1 = 1, X_2 = 1) &= \frac{1}{3} \neq \left(\frac{1}{2}\right)^2. \end{aligned}$$

Theorem 31 (Strong law of large numbers). *Let $\{X_n\}_{n=1}^{\infty}$ be an exchangeable sequence of random variables with finite mean. Then $\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} X_j$ exists a.s. and has mean equal to $\mathrm{E}(X_1)$. If, the X_j's are independent, then the limit equals $\mathrm{E}(X_1)$ a.s.*

Proof: Define $Y_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ and let $\mathcal{F}_n$ be the σ-field generated by all function of $(X_1, X_2, \ldots)$ that are invariant under permutations of the first n coordinates. (For example, Y_n is such a function.) Let $Z_n = \mathrm{E}(X_1|\mathcal{F}_n)$. Theorem 28 says that Z_n converges a.s. to $\mathrm{E}(X_1|\mathcal{F}_\infty)$, where $\mathcal{F}_\infty = \bigcap_{n=1}^{\infty} \mathcal{F}_n$. We prove next that $Z_n = Y_n$, a.s. Since Y_n is $\mathcal{F}_n$ measurable, we need only prove that, for all $A \in \mathcal{F}_n$, $\mathrm{E}(I_A Y_n) = \mathrm{E}(I_A X_1)$. Notice that I_A can be written as a function $h(X_1, X_2, ...)$, a function of $X_1, X_2, \ldots$ that is invariant under permutations of $X_1, \ldots, X_n$. By exchangeability, for all $j = 1, \ldots, n$, $X_1 h(X_1, X_2, \ldots)$ has the same distribution as $X_j h(X_j, X_2, \ldots, X_{j-1}, X_1, X_{j+1}, \ldots)$. But

$$h(X_j, X_2, \ldots, X_{j-1}, X_1, X_{j+1}, \ldots) = h(X_1, X_2, \ldots) = I_A,$$

by permutation invariance of h. Hence, for all $j = 1, \ldots, n$, $I_A X_j$ has the same distribution as $I_A X_1$. It follows that

$$\mathrm{E}(I_A X_1) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}(I_A X_j) = \mathrm{E}(I_A Y_n).$$

Clearly $\mathrm{E}(X_1|\mathcal{F}_\infty)$ has mean $\mathrm{E}(X_1)$. If the X_n's are independent, then the limit, being measurable with respect to the tail σ-field, must be constant a.s., by Kolmogorov 0-1 law. The constant must equal the mean of the random variable, which is $\mathrm{E}(X_1)$. ■

Example 32. *In Example 30, we know that Y_n converges a.s., hence it converges in distribution. We can compute the distribution of Y_n exactly: $P(Y_n = k/n) = 1/(n+1)$ for $k = 0, \ldots, n$. Hence, Y_n converges in distribution to uniform on the interval $[0,1]$, which must be the distribution of the limit. The limit is not a.s. constant.*

There is a very useful theorem due to deFinetti about exchangeable random quantities that relies upon the strong law of large numbers. To state the theorem, we need to recall the concept of "random probability measure" that was introduced in Example 40 of Lecture Notes Set 4. Let $(\mathcal{X}, \mathcal{B})$ be a Borel space, and let $\mathcal{P}$ be the set of all probability measures on $(\mathcal{X}, \mathcal{B})$. We can think of $\mathcal{P}$ as a subset of the function space $[0,1]^{\mathcal{B}}$, hence it has a product σ-field. Recall that the product σ-field is the smallest σ-field such that for all $B \in \mathcal{B}$, the function $f_B : \mathcal{P} \to [0,1]$ defined by $f_B(Q) = Q(B)$ is measurable. These are the coordinate projection functions.

Example 33 (Empirical probability measure). *Let $X_1, \ldots, X_n$ be random quantities taking values in $\mathcal{X}$. For each $B \in \mathcal{B}$, define $\mathbf{P}_n(\omega)(B) = \frac{1}{n}\sum_{j=1}^n I_B(X_j(\omega))$. For each ω, $\mathbf{P}_n(\omega)(\cdot)$ is clearly a probability measure, so $\mathbf{P}_n : \Omega \to \mathcal{P}$ is a function that we could prove was measurable, but that proof will not be given here. Theorem 31 says that $\mathbf{P}_n(\omega)(B)$ converges to $\mathrm{E}(I_B(X_1)|\mathcal{F}_\infty)(\omega)$ for all B and almost all ω. If we assume that the X_n's take values in a Borel space, then $\mathrm{E}(I_B(X_1)|\mathcal{F}_\infty) = \Pr(X_1 \in B|\mathcal{F}_\infty)$ is part of an rcd. This rcd plays an important roll in deFinetti's theorem.*

DeFinetti's theorem says that a sequence of random quantities is exchangeable if and only if it is conditionally iid given a random probability measure, and that random probability measure is the limit of the empirical probability measures of $X_1, \ldots, X_n$. That is, Example 29 is essentially the only example of exchangeable sequences. A simple proof can be found in Kingman (1978, Annals of Probability, Vol. 6, 183–197). A photocopy of the pages are attached at the end of this note.

Theorem 34 (DeFinetti's theorem). *A sequence $\{X_n\}_{n=1}^\infty$ of random quantities is exchangeable if and only if $\mathbf{P}_n$ (the empirical probability measure of $X_1, \ldots, X_n$) converges a.s. to a random probability measure $\mathbf{P}$ and the X_n's are conditionally iid with distribution Q given $\mathbf{P} = Q$.*

Example 35. *In Example 30, the empirical probability measure is equivalent to $Y_n = \sum_{k=1}^n X_k/n$, since Y_n is one minus the proportion of the observations less than or equal to 0. So $\mathbf{P}$ is equivalent to the limit of Y_n, the limit of relative frequency of 1's in the sequence. Conditional on the limit of relative frequency of 1's being x, the X_k's are iid with Bernoulli distribution with parameter x.*

A Upcrossing lemma

Proof: [Upcrossing lemma] Let $Y_k = \max\{0, X_k - r\}$ for every k so that $\{(Y_k, \mathcal{F}_k)\}_{k=1}^n$ is a submartingale. Note that a consecutive set of $X_k(\omega)$ cross from below r to above q if and only if the corresponding consecutive set of $Y_k(\omega)$ cross from 0 to above $q-r$. Let $T_0(\omega) = 0$ and define T_m for $m = 1, 2, \ldots$ as

$$\begin{aligned} T_m(\omega) &= \inf\{k \le n : k > T_{m-1}(\omega), Y_k(\omega) = 0\}, \text{ if } m \text{ is odd}, \\ T_m(\omega) &= \inf\{k \le n : k > T_{m-1}(\omega), Y_k(\omega) \ge q - r\}, \text{ if } m \text{ is even}, \\ T_m(\omega) &= n + 1, \text{ if the corresponding set above is empty.} \end{aligned}$$

Now $V(\omega)$ is one-half of the largest even m such that $T_m(\omega) \le n$. Define, for $k = 1, \ldots, n$,

$$R_k(\omega) = \begin{cases} 1 & \text{if } T_m(\omega) < k \le T_{m+1}(\omega) \text{ for } m \text{ odd}, \\ 0 & \text{otherwise.} \end{cases}$$

Then $(q-r)V(\omega) \le \sum_{k=1}^n R_k(\omega)[Y_k(\omega) - Y_{k-1}(\omega)] = \hat{X}$, where $Y_0 \equiv 0$ for convenience. First, note that for all m and k, $\{T_m(\omega) \le k\} \in \mathcal{F}_k$. Next, note that for every k,

$$\{\omega : R_k(\omega) = 1\} = \bigcup_{m \text{ odd}} \left(\{T_m \le k-1\} \cap \{T_{m+1} \le k-1\}^C\right) \in \mathcal{F}_{k-1}. \tag{3}$$

$$\begin{aligned} \mathrm{E}(\hat{X}) &= \sum_{k=1}^n \int_{\{\omega : R_k(\omega)=1\}} [Y_k(\omega) - Y_{k-1}(\omega)] dP(\omega) \\ &= \sum_{k=1}^n \int_{\{\omega : R_k(\omega)=1\}} [\mathrm{E}(Y_k|\mathcal{F}_{k-1})(\omega) - Y_{k-1}(\omega)] dP(\omega) \\ &\le \sum_{k=1}^n \int [\mathrm{E}(Y_k|\mathcal{F}_{k-1})(\omega) - Y_{k-1}(\omega)] dP(\omega) \\ &= \sum_{k=1}^n [\mathrm{E}(Y_k) - \mathrm{E}(Y_{k-1})] = \mathrm{E}(Y_n), \end{aligned}$$

where the second equality follows from Equation (3) and the inequality follows from the fact that $\{(Y_k, \mathcal{F}_k)\}_{k=1}^n$ is a submartingale. It follows that $(q-r)\mathrm{E}(V) \le \mathrm{E}(Y_n)$. Since $\mathrm{E}(Y_n) \le |r| + \mathrm{E}(|X_n|)$, it follows that Equation (2) holds. ■