Unit 11: Backwards Martingales

Instructor: Quan Zhou

11.1 Convergence of backwards martingales

Definition 11.1. A sequence of random variables $(X_n)_{n\leq 0}$ adapted to an increasing sequence of σ-algebras $(\mathcal{F}_n)_{n\leq 0}$[1] is said to be a backwards martingale (or reversed martingale) w.r.t $(\mathcal{F}_n)_{n\leq 0}$, if for each $n \leq -1$, we have (i) $\mathsf{E}|X_n| < \infty$, and (ii) $\mathsf{E}[X_{n+1} \,|\, \mathcal{F}_n] = X_n$, a.s.

Remark 11.1. The first term in $(X_n)_{n\leq 0}$ is X_0, the second is X_{-1}, and so on. We are particularly interested in what happens as $n \to -\infty$. Compared to forward martingales, the key difference is that the σ-algebra $\mathcal{F}_n$ decreases as $n \to -\infty$. The following lemma explains why the theory for backwards martingales is easier.

Lemma 11.1. *Let $(X_n)_{n\leq 0}$ be a backwards martingale. Then, $(X_n)_{n\leq 0}$ is uniformly integrable.*

Proof. For any $n \leq 0$, we have $\mathsf{E}[X_0 \,|\, \mathcal{F}_n] = X_n$. Hence, the result follows from Theorem 8.4. □

Theorem 11.1. *For a backwards martingale $(X_n)_{n\leq 0}$, as $n \to -\infty$, X_n converges a.s. and in L^1 to some $X_{-\infty}$.*

Proof. For $n \leq 0$, let $U_n^{a,b}$ be the number of upcrossings of $[a, b]$ completed by $X_n, \dots, X_0$. Lemma 5.1 yields that

$$(b - a)\mathsf{E}U_n^{a,b} \leq \mathsf{E}(X_0 - a)^-.$$

Since $\mathsf{E}|X_0| < \infty$, letting $n \to -\infty$ yields that $\mathsf{E}U_{-\infty}^{a,b} < \infty$. Mimicking the proof of Theorem 5.1, we see that $X_{-\infty} = \lim_{n\to-\infty} X_n$ exists a.s. (but is possibly infinite). Fatou's lemma yields $\mathsf{E}|X_{-\infty}| \leq \liminf_{n\to-\infty} \mathsf{E}|X_n|$, which is finite, since $(X_n)_{n\leq 0}$ is uniformly integrable by Lemma 11.1. The uniform integrability also implies that he convergence also occurs in L^1. □

Lemma 11.2. *Let $(X_n)_{n\leq 0}$ be a backwards martingale, $X_{-\infty} = \lim_{n\to-\infty} X_n$, and $\mathcal{F}_{-\infty} = \cap_{n\leq 0}\mathcal{F}_n$. Then, $X_{-\infty} = \mathsf{E}[X_0 \,|\, \mathcal{F}_{-\infty}]$.*

[1] "Increasing" means that $\cdots \subset \mathcal{F}_{-2} \subset \mathcal{F}_{-1} \subset \mathcal{F}_0$.

Proof. Clearly, $X_{-\infty} \in \mathcal{F}_n$ for each $n \leq 0$. Hence, $X_{-\infty} \in \mathcal{F}_{-\infty}$. To prove the claim, it remains to show that for any $A \in \mathcal{F}_{-\infty}$, $\mathsf{E}[X_{-\infty}\mathbb{1}_A] = \mathsf{E}[X_0\mathbb{1}_A]$. Since $A \in \mathcal{F}_{-\infty}$ implies $A \in \mathcal{F}_n$ for each n, we have

$$X_n\mathbb{1}_A = \mathbb{1}_A\,\mathsf{E}[X_0 \,|\, \mathcal{F}_n] = \mathsf{E}[X_0\mathbb{1}_A \,|\, \mathcal{F}_n].$$

Taking expectations on both sides yields $\mathsf{E}[X_n\mathbb{1}_A] = \mathsf{E}[X_0\mathbb{1}_A]$. Theorem 11.1 implies $X_n\mathbb{1}_A$ converges in L^1 to $X_{-\infty}\mathbb{1}_A$. Hence, $\mathsf{E}[X_{-\infty}\mathbb{1}_A] = \mathsf{E}[X_0\mathbb{1}_A]$. □

Theorem 11.2. *Let X be an integrable random variable and $\mathcal{F}_{-\infty} = \cap_{n\leq 0}\mathcal{F}_n$. Then, $\mathsf{E}[X \,|\, \mathcal{F}_n]$ converges to $\mathsf{E}[X \,|\, \mathcal{F}_{-\infty}]$ a.s. and in L^1.*

Proof. Define $X_n = \mathsf{E}[X \,|\, \mathcal{F}_n]$ for $n \leq 0$, which is a backwards martingale. Hence, $X_n \to X_{-\infty}$ a.s. and in L^1, where $X_{-\infty} = \mathsf{E}[X_0 \,|\, \mathcal{F}_{-\infty}]$ by Lemma 11.2. But $\mathsf{E}[X_0 \,|\, \mathcal{F}_{-\infty}] = \mathsf{E}[X \,|\, \mathcal{F}_{-\infty}]$ by the tower property. □

Exercise 11.1. Let $(X_n)_{n\leq 0}$ be a backwards martingale with $\mathsf{E}[|X_0|^p] < \infty$ for some $p > 1$. Show that X_n converges to $X_{-\infty}$ (as $n \to -\infty$) in L^p.

11.2 Application of backwards martingales

Theorem 11.3. *Let $Z_1, Z_2, \dots$ be i.i.d. integrable random variables with $\mathsf{E}Z_1 = \mu$. Let $S_n = Z_1 + \cdots + Z_n$. Then S_n/n converges to μ a.s. and in L^1 as $n \to \infty$.*

Proof. For each $n \geq 1$, define $X_{-n} = S_n/n$, and $\mathcal{F}_{-n} = \sigma(S_n, Z_{n+1}, Z_{n+2}, \dots)$. Then, $(X_n)_{n\leq -1}$ is adapted to $(\mathcal{F}_n)_{n\leq -1}$. We show that it is indeed a backwards martingale. For $n \geq 1$, we have

$$\begin{aligned}\mathsf{E}[X_{-n} \,|\, \mathcal{F}_{-(n+1)}] &= \mathsf{E}\left[\frac{S_n}{n} \,\middle|\, \mathcal{F}_{-(n+1)}\right] \\ &= \mathsf{E}\left[\frac{S_{n+1} - Z_{n+1}}{n} \,\middle|\, \mathcal{F}_{-(n+1)}\right] \\ &= \frac{S_{n+1}}{n} - \mathsf{E}\left[\frac{Z_{n+1}}{n} \,\middle|\, \mathcal{F}_{-(n+1)}\right].\end{aligned}$$

The i.i.d. assumption implies that

$$\mathsf{E}\left[\frac{Z_{n+1}}{n} \,\middle|\, \mathcal{F}_{-(n+1)}\right] = \mathsf{E}\left[\frac{Z_{n+1}}{n} \,\middle|\, S_{n+1}\right] = \frac{S_{n+1}}{n(n+1)}.$$

where the last equality follows from the fact that $\mathsf{E}[Z_1 \,|\, S_{n+1}] = \cdots = \mathsf{E}[Z_{n+1} \,|\, S_{n+1}]$. Hence,

$$\mathsf{E}[X_{-n} \,|\, \mathcal{F}_{-(n+1)}] = \frac{S_{n+1}}{n} - \frac{S_{n+1}}{n(n+1)} = \frac{S_{n+1}}{n+1} = X_{-(n+1)}.$$

Hence, by Theorem 11.1, as $n \to \infty$, $X_{-n} = S_n/n$ converges a.s. and in L^1 to some $X_{-\infty}$. Note that the convergence in L^1 implies $\mathsf{E}[X_{-\infty}] = \mu$. Since $X_{-\infty}$ is in the tail σ-algebra generated by $(Z_n)_{n\geq 1}$, Kolmogorov's zero-one law implies that $X_{-\infty}$ must be a constant; that is, $X_{-\infty} = \mu$, a.s. □

Definition 11.2. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be given by $\Omega = \{(\omega_1, \omega_2, \dots) \colon \omega_i \in \mathbb{R}\}$ and $\mathcal{F} = \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \times \cdots$. Define $X_n(\omega) = \omega_n$. Let $\mathbb{S}_n$ be the permutation group on $\{1, 2, \dots, n\}$. Given $\pi \in \mathbb{S}_n$ and $A \in \mathcal{F}$, define

$$\pi^{-1}A = \{\omega \in \Omega \colon (\omega_{\pi(1)}, \dots, \omega_{\pi(n)}, \omega_{n+1}, \dots) \in A\}.$$

Let $\mathcal{E}_n$ be the σ-algebra generated by all events A such that $A = \pi^{-1}A$ for every $\pi \in \mathbb{S}_n$. Let $\mathcal{E} = \cap_{n\geq 1}\mathcal{E}_n$ be the exchangeable σ-algebra.

Remark 11.2. Let $\mathcal{T} = \cap_{n\geq 1}\sigma(X_n, X_{n+1}, \dots)$ be the tail σ-algebra generated by $(X_n)_{n\geq 1}$. Then, $\mathcal{T} \subset \mathcal{E}$, but not vice versa; that is, a tail event must be exchangeable, but an exchangeable event may not be a tail event.

Theorem 11.4 (Hewitt-Savage zero-one law)**.** *Consider the setting of Definition 11.2. If $X_1, X_2, \dots$ are i.i.d. and $A \in \mathcal{E}$, then $\mathsf{P}(A) = 0$ or 1.*

Sketch of proof. Suppose for any $n \geq 1$ and bounded function f,

$$\mathsf{E}[f(X_1, \dots, X_n) \,|\, \mathcal{E}] = \mathsf{E}[f(X_1, \dots, X_n)]. \tag{1}$$

This implies that $\mathcal{E}$ is independent of $\sigma(X_1, \dots, X_n)$ for every n. Then, one can use the argument in the proof of Kolmogorov's zero-one law to show that $\mathcal{E}$ is independent of $\mathcal{F}$, which proves the asserted the result.

To prove (1), define $Y_{-m} = \mathsf{E}[f(X_1, \dots, X_n) \,|\, \mathcal{E}_m]$ for each $m \geq n$. Since $\mathcal{E}_m$ is monotone decreasing, $(Y_{-m})_{m\geq n}$ is a backwards martingale with respect to $(\mathcal{E}_m)_{m\geq n}$. Hence, $Y_{-m} \to Y_{-\infty} = \mathsf{E}[f(X_1, \dots, X_n) \,|\, \mathcal{E}]$ a.s. It is not difficult to prove, using the i.i.d. assumption, that

$$\mathsf{E}[f(X_1, \dots, X_n) \,|\, \mathcal{E}_m] = \frac{(m-n)!}{m!} \sum_{a \in \mathcal{S}([m], n)} f(X_{a_1}, \dots, X_{a_n}), \tag{2}$$

where $[m] = \{1, 2, \ldots, m\}$, and

$$\mathcal{S}(I, n) = \{(a_1, \ldots, a_n) \colon a_1, \ldots, a_n \text{ are distinct, and } \forall\, i, a_i \in I\}.$$

A straightforward calculation shows that, a.s.,

$$\frac{(m-n)!}{m!} \left\{ \sum_{a \in \mathcal{S}([m], n)} f(X_{a_1}, \ldots, X_{a_n}) - \sum_{a \in \mathcal{S}([m] \setminus \{1\}, n)} f(X_{a_1}, \ldots, X_{a_n}) \right\}$$

converges to zero as $m \to \infty$. That is, the limit of $\mathsf{E}[f(X_1, \ldots, X_n) \mid \mathcal{E}_m]$ is independent of X_1. By repeating this argument, we find that $\mathsf{E}[f(X_1, \ldots, X_n) \mid \mathcal{E}]$ is independent of $\sigma(X_1, \ldots, X_n)$, which implies (1). □

Theorem 11.5 (de Finetti's Theorem)**.** *Consider the setting of Definition 11.2, and assume $X_1, X_2, \ldots$ are exchangeable; that is, for any n and $\pi \in \mathbb{S}_n$, $(X_1, \ldots, X_n)$ and $(X_{\pi(1)}, \ldots, X_{\pi(n)})$ have the same distribution. Then, conditional on $\mathcal{E}$, $X_1, X_2, \ldots$ are i.i.d.*

Proof. Proof is omitted. □

Theorem 11.6. *If $X_1, X_2, \ldots$ are exchangeable and take values in $\{0, 1\}$, then there exists a probability distribution μ on $[0, 1]$ such that*

$$\mathsf{P}(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0) = \int_0^1 y^k (1-y)^{n-k} \mu(\mathrm{d}y).$$

Proof. Proof is omitted. □

Exercise 11.2. Consider the setting of Definition 11.2. Find an event which is in $\mathcal{E}$ but not necessarily in $\mathcal{T}$.

Exercise 11.3. Let $X_1, X_2, \ldots$ be exchangeable with $\mathsf{E}[X_1^2] < \infty$. Prove that $\mathsf{E}[X_1 X_2] \geq 0$.

References

[1] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.