# Lecture Notes on Martingales

by Russell Lyons

Extended from the book by Patrick Billingsley

Suppose we gamble on a sequence of fair games. Let $X_n$ be our fortune after $n$ plays and $\mathscr{F}_n := \sigma(X_1, \ldots, X_n)$ or more; e.g., we might use other randomness. Fairness is $X_n = E[X_{n+1} \mid \mathscr{F}_n]$. Unfavorable gambling has $\geq$ and favorable has $\leq$. Note $\mathscr{F}_n \subseteq \mathscr{F}_{n+1}$. One way to be favorable is to get handed some money before each fair game, where the amount has conditional expectation $\geq 0$ and the future games remain fair.

**Definition.** Let $X_n \in L^1(\Omega, \mathscr{F}, P)$ and $\mathscr{F}_n \subseteq \mathscr{F}$ be $\sigma$-fields. We call $\langle \mathscr{F}_n \rangle$ a ***filtration*** if $\mathscr{F}_n \subseteq \mathscr{F}_{n+1}$. We call $\langle X_n \rangle$ ***adapted*** to $\langle \mathscr{F}_n \rangle$ if $X_n \in \mathscr{F}_n$. We call $\langle (X_n, \mathscr{F}_n) \rangle$ a ***martingale*** (***submartingale***; ***supermartingale***) if $\langle X_n \rangle$ is adapted to the filtration $\langle \mathscr{F}_n \rangle$ and

$$\forall n \qquad X_n = E[X_{n+1} \mid \mathscr{F}_n] \text{ a.s. } (\leq\ ;\ \geq)$$

We'll later see the reason for the names of the prefixes; it is related to subharmonic functions (convex functions in the case of $\mathbb{R}$).

We call $\langle X_n \rangle$ a ***martingale*** if there is some filtration with respect to which it is a martingale. In such a case, $\mathscr{F}_n := \sigma(X_1, \ldots, X_n)$ always works. Note that $E[X_m \mid \mathscr{F}_n] = X_{m \wedge n}$ a.s. and $E[X_n]$ is the same for all $n$.

**Example 35.1.** Sums of independent random variables.

**Example 35.5.** If $Z \in L^1(\Omega, \mathscr{F}, P)$ and $\langle \mathscr{F}_n \rangle$ is a filtration, then $\langle (E[Z \mid \mathscr{F}_n], \mathscr{F}_n) \rangle$ is a martingale.

**Extra Credit Homework:** Show that $\langle n\mathbf{1}_{[0,1/n]}\,;\ n \geq 1 \rangle$ is a martingale on $[0,1]$ with respect to Lebesgue measure.

**Homework:** 35.2, 6 (correction: $X_0 = E[X_1]$. Also, show that $\langle Z_n \rangle$ is predictable, i.e., for all $n$, $Z_n$ is $\mathscr{F}_{n-1}$-measurable.)

**Example 35.2.** Let $\nu$ be a finite, signed measure on $(\Omega, \mathscr{F}, P)$ and $\langle \mathscr{F}_n \rangle$ be a filtration. Suppose that $\nu \restriction \mathscr{F}_n \ll P \restriction \mathscr{F}_n$ for all $n$. Let $X_n := d(\nu \restriction \mathscr{F}_n)/d(P \restriction \mathscr{F}_n)$. Then $\langle (X_n, \mathscr{F}_n) \rangle$ is a martingale:

$$\begin{aligned} E[X_{n+1} \mid \mathscr{F}_n](P \restriction \mathscr{F}_n) &= (X_{n+1}P) \restriction \mathscr{F}_n = (X_{n+1}P \restriction \mathscr{F}_{n+1}) \restriction \mathscr{F}_n \\ &= (\nu \restriction \mathscr{F}_{n+1}) \restriction \mathscr{F}_n = \nu \restriction \mathscr{F}_n = X_n(P \restriction \mathscr{F}_n). \end{aligned}$$

**Example 35.3.** Let $P$ be Lebesgue measure on $(0,1]$ and $\mathscr{F}_n := \sigma(I_k^{(n)}\,;\ 0 \leq k < 2^n)$, where $I_k^{(n)} := (k2^{-n}, (k+1)2^{-n}]$. Since $(A \in \mathscr{F}_n, P(A) = 0) \Rightarrow A = \varnothing$, we have for all $\nu$ that $\nu \restriction \mathscr{F}_n \ll P \restriction \mathscr{F}_n$. In this case,

$$X_n = \sum_{k=0}^{2^n - 1} \mathbf{1}_{I_k^{(n)}} \nu(I_k^{(n)})/P(I_k^{(n)}) = \sum_{k=0}^{2^n - 1} \mathbf{1}_{I_k^{(n)}} \nu(I_k^{(n)}) 2^n.$$

**Example 35.8.** If $\langle (X_n, \mathscr{F}_n) \rangle$ is a martingale, then $\langle (|X_n|, \mathscr{F}_n) \rangle$ is a submartingale.

This is a special case of:

**Theorem 35.1.** *Let $\langle X_n \rangle$ be integrable random variables adapted to a filtration $\langle \mathscr{F}_n \rangle$, $\varphi$ be convex on an interval containing all the ranges of $X_n$, and $\varphi \circ X_n \in L^1(P)$. Then $\langle (\varphi \circ X_n, \mathscr{F}_n) \rangle$ is a submartingale if either*
(i) *$\langle (X_n, \mathscr{F}_n) \rangle$ is a martingale; or*
(ii) *$\langle (X_n, \mathscr{F}_n) \rangle$ is a submartingale and $\varphi$ is increasing.*

*Proof.* In both cases, we have $\varphi(X_n) \le \varphi\big(E[X_{n+1} \mid \mathscr{F}_n]\big)$. By Jensen's inequality, this is $\le E\big[\varphi(X_{n+1}) \mid \mathscr{F}_n\big]$. ∎

We now depart from the book.

**Wald's Equation.** *Let $Z_n \in L^1(P)$ for $n \ge 1$, $\tau$ be an $\mathbb{N}$-valued random variable, and $\mu \in \mathbb{R}$. Suppose that*
(i) *for each $n \ge 1$, we have $P[\tau \ge n] > 0 \Rightarrow E[Z_n \mid \tau \ge n] = \mu$ ($\le$; $\ge$)*
*and*
(ii) *one of the following holds:*
(a) $\forall n \ \ Z_n \ge 0$*;*
(b) $\sup_{n\,;\,P[\tau \ge n] > 0} E\big[|Z_n| \bigm| \tau \ge n\big] < \infty$ *and* $E[\tau] < \infty$*;*
(c) $E\big[|\sum_{n=1}^{\tau} Z_n|\big] < \infty$ *and* $\lim_{n\to\infty} E\big[\sum_{k=1}^{n} \mathbf{1}_{[\tau > n]}\big] = 0$.

*Then*

$$E\bigg[\sum_{n=1}^{\tau} Z_n\bigg] = \mu\, E[\tau], \quad (\le\ ;\ \ge)$$

*where $0 \cdot \infty := 0$.*

*Proof.* In case (ii)(a), we have by Tonelli's theorem (or the MCT) that

$$\begin{aligned} E\bigg[\sum_{n=1}^{\tau} Z_n\bigg] &= E\bigg[\sum_{n=1}^{\infty} Z_n \mathbf{1}_{[n \le \tau]}\bigg] = \sum_{n=1}^{\infty} E[Z_n \mathbf{1}_{[\tau \ge n]}] \\ &= \sum_{n=1}^{\infty} E[Z_n \mid \tau \ge n] P[\tau \ge n] = \sum_{n=1}^{\infty} \mu\, P[\tau \ge n] = \mu\, E[\tau]. \end{aligned} \tag{N1}$$

In case (ii)(b), we have

$$\begin{aligned} E\bigg[\sum_{n=1}^{\infty} |Z_n| \mathbf{1}_{[n \le \tau]}\bigg] &= \sum_{n=1}^{\infty} E\big[|Z_n| \bigm| \tau \ge n\big] P[\tau \ge n] \\ &\le E[\tau] \sup_n E\big[|Z_n| \bigm| \tau \ge n\big] < \infty, \end{aligned}$$

whence we may apply Fubini's theorem as in (N1).

Finally, in case (ii)(c), the LDCT gives

$$\begin{aligned} E\left[\sum_{k=1}^{\tau} Z_k\right] &= \lim_{n\to\infty} E\left[\sum_{k=1}^{\tau} Z_k \mathbf{1}_{[\tau\le n]}\right] = \lim_{n\to\infty} E\left[\sum_{k=1}^{n} Z_k \mathbf{1}_{[k\le\tau\le n]}\right] \\ &= \lim_{n\to\infty} E\left[\sum_{k=1}^{n} Z_k \big(\mathbf{1}_{[\tau\le k]} - \mathbf{1}_{[\tau>n]}\big)\right] = \lim_{n\to\infty} \sum_{k=1}^{n} \mu\, P[\tau\ge k] = \mu\, E[\tau]. \end{aligned}$$

(Note the case $\mu = 0$ and $E[\tau] = \infty$.) ∎

**Definition.** Given a filtration $\langle \mathscr{F}_n \rangle$, a random variable $\tau \in \mathbb{N} \cup \{\infty\}$ is called a ***stopping time*** with respect to $\langle \mathscr{F}_n \rangle$ if $\forall n \in \mathbb{N}\ [\tau = n] \in \mathscr{F}_n$. Equivalently, $\forall n \in \mathbb{N}\ [\tau \le n] \in \mathscr{F}_n$.

**Homework:** Show that if $\tau_1$ and $\tau_2$ are stopping times with respect to $\langle \mathscr{F}_n \rangle$, then so are $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$, and $\tau_1 + \tau_2$.

**Corollary.** *Let $\langle (X_n, \mathscr{F}_n) \rangle$ be a martingale (submartingale; supermartingale) and $\tau$ an a.s. finite stopping time with respect to $\langle \mathscr{F}_n \rangle$. If*
 (a) $\exists c \le 0\ \forall n\ X_{n+1} - X_n \ge c$ *a.s. and* $E[\tau] < \infty$
*or*
 (b) $E\big[|X_\tau|\big] < \infty$ *and* $\lim_n E[X_n \mathbf{1}_{[\tau>n]}] = 0$,
*then*

$$E[X_\tau] = E[X_0]. \quad (\ge\ ;\ \le)$$

*Proof.* In case (a), apply Wald's equation with $Z_n := X_n - X_{n-1} - c$ for $n \ge 1$ and $\mu := -c$. In case (b), apply Wald's equation with $Z_n := X_n - X_{n-1}$ for $n \ge 1$ and $\mu := 0$. Note that $[\tau \ge n] = [\tau \le n-1]^{\mathrm{c}} \in \mathscr{F}_{n-1}$. ∎

**Example.** If $\sup \tau < \infty$, then the hypotheses (b) hold. More generally, if $\{X_{\tau\wedge n}\,;\ n \ge 0\}$ is uniformly integrable, then the hypotheses (b) hold, because $E\big[|X_\tau|\big] = \lim_{n\to\infty} E\big[|X_{\tau\wedge n}|\big]$ and $E[X_n \mathbf{1}_{[\tau>n]}] = E[X_{\tau\wedge n} \mathbf{1}_{[\tau>n]}]$, where the latter random variables are uniformly integrable.

Here are some nice applications.

**Example.** Let $\langle X_n \rangle$ be simple random walk on $\mathbb{Z}$ starting from $X_0 = 0$. Let $\tau_a := \inf\{n\,;\ X_n = a\}$. Then $E[\tau_a] = \infty$ for $a \ne 0$ since the conclusion of the corollary does not hold. But if $\tau := \tau_a \wedge \tau_{-b}$ for $a, b > 0$, then $X_{\tau\wedge n}$ are bounded, so uniformly integrable, whence

$$0 = E[X_\tau] = a\, P[\tau_a < \tau_{-b}] + b\, P[\tau_{-b} < \tau_a],$$

giving $P[\tau_a < \tau_{-b}] = b/(a+b)$.

Here, we assumed that $\tau < \infty$ a.s. This is implied by a martingale convergence theorem, Theorem 35.5, but we can also prove it as follows, at the same time as we calculate $E[\tau]$.

Problem 35.2 shows that $\big\langle \big(X_n^2 - n, \sigma(X_1, \ldots, X_n)\big) \big\rangle$ is a martingale. Fix $N \in \mathbb{N}$. Then $\tau \wedge N$ is a bounded stopping time, so we may apply the corollary to get

$$0 = E[X_0^2 - 0] = E[X_{\tau\wedge N}^2 - \tau \wedge N] = E[X_{\tau\wedge N}^2] - E[\tau \wedge N],$$

i.e., $E[\tau \wedge N] = E[X^2_{\tau\wedge N}]$. Now let $N \to \infty$. The right-hand side is bounded, whence so is the left-hand side; this shows that $E[\tau] < \infty$, whence $\tau < \infty$ a.s. Furthermore, the MCT and the BCT give

$$E[\tau] = E[X^2_\tau] = a^2 \frac{b}{a+b} + b^2 \frac{a}{a+b} = ab.$$

**Definition.** Let $p(\cdot,\cdot)$ be a transition probability on a state space, $S$. A function $f\colon S \to \mathbb{R}$ is called ***harmonic*** (***subharmonic***; ***superharmonic***) if

$$\forall x \in S \;\; f(x) = \sum_{y\in S} p(x,y) f(y). \quad (\le\,;\;\ge)$$

If $\langle X_n \rangle$ is a Markov chain on $S$ with these transition probabilities and $f$ is harmonic (subharmonic; superharmonic), then $\langle f(X_n)\rangle$ is a martingale (submartingale; supermartingale).

**Example.** Consider the asymmetric random walk $\langle X_n \rangle$ on $\mathbb{Z}$ with $p(n, n+1) = p$ and $p(n, n-1) = q := 1-p$. Then $f(n) := (q/p)^n$ is harmonic, whence $\langle (q/p)^{X_n} \rangle$ is a martingale. If $\tau := \tau_a \wedge \tau_{-b}$, then this martingale is bounded up to time $\tau$, whence the corollary applies:

$$1 = E\big[(q/p)^{X_0}\big] = E\big[(q/p)^{X_\tau}\big] = (q/p)^a \, P[\tau_a < \tau_{-b}] + (q/p)^{-b} \, P[\tau_{-b} < \tau_a].$$

We can solve this to get another derivation of the probabilities in gambler's ruin when $p \ne q$. (Again, the martingale convergence theorem will apply that $\tau < \infty$ a.s., but it is also a consequence of the following homework.)

**Homework:** Let $\langle X_n \rangle$ be asymmetric random walk on $\mathbb{Z}$ with $p > q$, $s := \sqrt{q/p}$, and $r := 2\sqrt{pq}$. Show that $\langle s^{X_n} r^{-n} \rangle$ is a martingale. Deduce that if $X_0 = 0$, then $P[\tau_a \ge n] \le s^{-a} r^n$ for $a \ge 1$.

**Example.** What is $E[\tau_a]$ when $p > q$ and $a \ge 1$? We use another martingale: $\langle X_n - (p-q)n \rangle$. Fix $N \in \mathbb{N}$ and define $\tau := \tau_a \wedge N$. Then $\tau$ is bounded, so $0 = E[X_0 - (p-q)0] = E[X_\tau - (p-q)\tau]$, i.e., $(p-q)E[\tau_a \wedge N] = E[X_{\tau_a \wedge N}]$. By the MCT, $E[\tau_a \wedge N] \to E[\tau_a]$ as $N \to \infty$. Using the notation of the homework, we have

$$|E[X_{\tau_a\wedge N}] - a| \le E\big[|X_{\tau_a \wedge N} - a|; \tau_a > N\big] \le (N+a)\, P[\tau_a > N] \le (N+a) s^{-a} r^N \to 0$$

as $N \to \infty$, so $E[X_{\tau_a \wedge N}] \to a$ as $N \to \infty$. In conclusion,

$$E[\tau_a] = \frac{a}{p-q}.$$

**Homework:** For $\langle X_n \rangle$ be asymmetric random walk on $\mathbb{Z}$ with $p > q$, calculate $\mathrm{Var}(\tau_a)$ for $a > 0$. *Hint:* Use problem 35.2.

**Example.** Consider an ordinary deck of 26 red and 26 black cards, well shuffled. The cards are turned over one by one. At any time, and once only, we may bet that the next card is red. What's the best strategy to maximize our chance to be correct?

Let $A_k$ be the event that the $k$th card is red. If $\tau$ is the time we bet, then $[\tau = k] \in \sigma(A_1, \ldots, A_{k-1}) =: \mathscr{F}_k$, i.e., $\tau$ is a stopping time. At time $k$, the chance of winning would be $X_k := P[A_k \mid \mathscr{F}_k]$, so the chance of winning is $E[X_\tau]$. Now

$$E[X_{k+1} \mid \mathscr{F}_k] = E\big[P[A_{k+1} \mid \mathscr{F}_{k+1}] \bigm| \mathscr{F}_k\big] = P[A_{k+1} \mid \mathscr{F}_k] = X_k,$$

because all cards we have not seen are in random order. Thus, $\langle X_k \,;\ 1 \le k \le 52\rangle$ is a martingale, so $E[X_\tau] = E[X_1] = 1/2$. All strategies are equally effective.

Having learned about this martingale, we realize that $\forall k\ \ X_k = E[X_{52} \mid \mathscr{F}_k]$. But $X_{52} = P[A_{52} \mid \mathscr{F}_{52}] = \mathbf{1}_{A_{52}}$, so $X_k = P[A_{52} \mid \mathscr{F}_k]$. This shows more directly that any strategy is equivalent to betting on the last card. That's why it does not matter. In fact, we can see this without any use of martingales: betting on the next card has the same chance of winning as betting on the last card, no matter what we have seen, because all cards we have not seen are in random order.

Recall problem 16.9: $\mathcal{X}$ is uniformly integrable iff $\sup\big\{E\big[|X|\big] \,;\ X \in \mathcal{X}\big\} < \infty$ and $\forall \epsilon > 0\ \ \exists \delta > 0\ \ P(A) < \delta \Rightarrow \sup\big\{E\big[|X| \,;\ A\big] \,;\ X \in \mathcal{X}\big\} < \epsilon$.

**Proposition.** *If $\mathcal{X}$ is UI and $\mathcal{G}$ is a collection of sub-$\sigma$-fields, then $\big\{E[X \mid \mathscr{G}] \,;\ X \in \mathcal{X},\ \mathscr{G} \in \mathcal{G}\big\}$ is UI.*

**Example.** $\mathcal{X} = \{Z\}$ for some $Z \in L^1(P)$ and $\mathcal{G}$ is a filtration.

*Proof.* Let $M := \sup\big\{E\big[|X|\big] \,;\ X \in \mathcal{X}\big\}$. Given $\epsilon > 0$, choose $\alpha$ such that

$$\sup\big\{E\big[|X| \,;\ A\big] \,;\ X \in \mathcal{X},\ P(A) \le M/\alpha\big\} < \epsilon.$$

Now for $X \in \mathcal{X}$ and $\mathscr{G} \in \mathcal{G}$, we have

$$P\big[|E[X \mid \mathscr{G}]| \ge \alpha\big] \le \alpha^{-1}\, E\big[|E[X \mid \mathscr{G}]|\big] \le \alpha^{-1}\, E\big[E[|X| \mid \mathscr{G}]\big] = \alpha^{-1}\, E\big[|X|\big] \le M/\alpha.$$

Hence, with $A := \big\{|E[X \mid \mathscr{G}]| \ge \alpha\big\}$, we have

$$E\big[|E[X \mid \mathscr{G}]| \,;\ A\big] \le E\big[E[|X| \mid \mathscr{G}]\mathbf{1}_A\big] = E\big[E[|X|\mathbf{1}_A \mid \mathscr{G}]\big] = E\big[|X|\mathbf{1}_A\big] < \epsilon. \qquad \blacksquare$$

Let $\langle \mathscr{F}_n \rangle$ be a filtration and $\tau$ a stopping time. What is the information up to time $\tau$, $\mathscr{F}_\tau$? On the part of $\Omega$ where $\tau = n$, it is $\mathscr{F}_n$. Thus, we define

$$\mathscr{F}_\tau := \big\{A \in \mathscr{F} \,;\ \forall n\ \ A \cap [\tau = n] \in \mathscr{F}_n\big\}.$$

This is a $\sigma$-field and $\tau \in \mathscr{F}_\tau$.

**Homework:** Show that if $\langle X_n \rangle$ is adapted to $\langle \mathscr{F}_n \rangle$ and $\tau$ is a finite stopping time, then $X_\tau \in \mathscr{F}_\tau$.

**Homework:** Show that if $\tau_1$ and $\tau_2$ are finite stopping times, then $\mathscr{F}_{\tau_1 \wedge \tau_2} = \mathscr{F}_{\tau_1} \cap \mathscr{F}_{\tau_2}$.

**Proposition.** *Let $\langle (X_n, \mathscr{F}_n) \rangle$ be a martingale (submartingale; supermartingale) and $\tau$ be a finite stopping time. Then for all $n$,*

$$E[X_n \mid \mathscr{F}_\tau] = X_{\tau \wedge n} \ a.s. \ (\geq \ ; \ \leq)$$

*Proof.* It's sufficient to do the submartingale case. Note that $X_{\tau \wedge n} \in \mathscr{F}_\tau$. Since both sides of the desired inequality are in $\mathscr{F}_\tau$, it suffices to show that for all $A \in \mathscr{F}_\tau$,

$$E[X_n \, ; \ A] \geq E[X_{\tau \wedge n} \, ; \ A].$$

This holds by the following calculation:

$$\begin{aligned} E[X_n \mathbf{1}_A] &= \sum_k E[X_n \underbrace{\mathbf{1}_A \mathbf{1}_{[\tau = k]}}_{\in \mathscr{F}_k}] \geq \sum_k E[X_{n \wedge k} \mathbf{1}_A \mathbf{1}_{[\tau = k]}] \\ &= \sum_k E[X_{n \wedge \tau} \mathbf{1}_A \mathbf{1}_{[\tau = k]}] = E[X_{n \wedge \tau} \mathbf{1}_A]. \end{aligned}$$

∎

**Homework:** Show that if $\langle (X_n, \mathscr{F}_n) \rangle$ is a submartingale and $\tau$ is a stopping time, then $\langle (X_{\tau \wedge n}, \mathscr{F}_n) \rangle$ is a submartingale.

**Corollary.** *If $\langle X_n \rangle$ is a UI martingale and $\tau$ is a finite stopping time, then $\{X_{\tau \wedge n}\}$ is UI.*

*Proof.* Apply the preceding two propositions. ∎

**Corollary.** *If $\langle X_n \rangle$ is a UI submartingale and $\tau$ is a finite stopping time, then $\{X_{\tau \wedge n}\}$ is UI.*

*Proof.* Since $|X_{\tau \wedge n}| \leq |X_\tau| + |X_n|$, it suffices to show that $E\big[|X_\tau|\big] < \infty$. Now $\langle X_n^+ \rangle$ is a submartingale, so by the proposition, $E[X_{\tau \wedge n}^+] \leq E[X_n^+]$. Also, $E[X_0] \leq E[X_{\tau \wedge n}]$ by the corollary to Wald's equation. Hence,

$$E\big[|X_{\tau \wedge n}|\big] = 2\,E[X_{\tau \wedge n}^+] - E[X_{\tau \wedge n}] \leq 2\,E[X_n^+] - E[X_0] \leq 2\,E\big[|X_n|\big] - E[X_0].$$

Fatou's lemma gives the result. ∎

**Homework:** 35.7.

We'll now usually just write our results for the submartingale case.

**Proposition.** *If $\langle X_n \rangle$ is a submartingale and $\tau_1 \leq \tau_2 \leq j \in \mathbb{N}$ are bounded stopping times, then $E[X_{\tau_1}] \leq E[X_{\tau_2}]$.*

*Proof.* Write $Z_n := X_n - X_{n-1}$. Then $X_n = X_0 + \sum_{k=1}^n Z_k$, so

$$E[X_{\tau_2} - X_{\tau_1}] = E\Bigg[\sum_{k=\tau_1+1}^{\tau_2} Z_k\Bigg] = E\Bigg[\sum_{k=1}^{j} Z_k \mathbf{1}_{[\tau_1 < k \leq \tau_2]}\Bigg] \geq 0$$

since $[\tau_1 < k \leq \tau_2] \in \mathscr{F}_{k-1}$. ∎

**The Optional Sampling Theorem.** *If $\tau_1 \le \tau_2$ are finite stopping times and $\langle (X_{\tau_2 \wedge n}, \mathscr{F}_n) \rangle$ is a UI submartingale, then*

$$X_{\tau_1} \le E[X_{\tau_2} \mid \mathscr{F}_{\tau_1}] \text{ a.s.}$$

Note that by a preceding homework and corollary, $\langle (X_{\tau_2 \wedge n}, \mathscr{F}_n) \rangle$ is a UI submartingale if $\langle (X_n, \mathscr{F}_n) \rangle$ is a UI submartingale.

*Proof.* First, we show that $E[X_{\tau_1}] \le E[X_{\tau_2}]$. Fix $j \in \mathbb{N}$. Then $\tau_1 \wedge j \le \tau_2 \wedge j \le j$ are bounded stopping times, whence the preceding proposition gives

$$E[X_{\tau_1 \wedge j}] = E[X_{\tau_2 \wedge (\tau_1 \wedge j)}] \le E[X_{\tau_2 \wedge (\tau_2 \wedge j)}] = E[X_{\tau_2 \wedge j}].$$

Now $E[X_{\tau_2 \wedge j}] \to E[X_{\tau_2}]$ since $\{X_{\tau_2 \wedge j}\}$ is UI. Furthermore, by the preceding corollary, $\{X_{\tau_1 \wedge j}\} = \{X_{\tau_2 \wedge (\tau_1 \wedge j)}\}$ is UI, whence $E[X_{\tau_1 \wedge j}] \to E[X_{\tau_1}]$. Therefore $E[X_{\tau_1}] \le E[X_{\tau_2}]$.

To show the finer conclusion desired, it suffices to show that $\forall A \in \mathscr{F}_{\tau_1}$ $E[X_{\tau_1}\,;\, A] \le E[X_{\tau_2}\,;\, A]$. By the homework, $\mathscr{F}_{\tau_1} = \mathscr{F}_{\tau_1 \wedge \tau_2} \subseteq \mathscr{F}_{\tau_2}$. Given $A \in \mathscr{F}_{\tau_1}$, set $\tau := \tau_1 \mathbf{1}_A + \tau_2 \mathbf{1}_{A^c}$. This is a stopping time since for all $n$,

$$[\tau \le n] = \underbrace{\big([\tau_1 \le n] \cap A\big)}_{\in \mathscr{F}_n} \cup \underbrace{\big([\tau_2 \le n] \cap A^c\big)}_{\in \mathscr{F}_n}.$$

Also, $\tau \le \tau_2$, so by the result proved in the first paragraph, we have

$$0 \le E[X_{\tau_2} - X_\tau] = E[X_{\tau_2} - X_{\tau_1} \mathbf{1}_A - X_{\tau_2} \mathbf{1}_{A^c}] = E[X_{\tau_2} \mathbf{1}_A - X_{\tau_1} \mathbf{1}_A]. \qquad \blacksquare$$

When does a submartingale converge a.s.? Note that the heuristic that it tends to increase suggests that it always has an extended-real limit. However, this is not so, even if it is a.s. bounded:

**Example.** Let $\langle x_n \rangle$ be any sequence in $\mathbb{R}$. Then there is a martingale $\langle X_n \rangle$ that behaves like $\langle x_n \rangle$, i.e., $X_n - x_n$ is eventually constant a.s. (stabilizes a.s.). To see this, let $z_n := x_n - x_{n-1}$ and $\langle Z_n \rangle$ be independent, mean-0 random variables with $\sum_n P[Z_n \ne z_n] < \infty$. Set $X_n := \sum_{k=1}^n Z_k$. Then $\langle X_n \rangle$ is a martingale and by the Borel–Cantelli lemma, $X_n - x_n$ stabilizes a.s.

Yet a mild hypothesis assures convergence:

**Theorem 35.5 (Doob's Martingale Convergence Theorem).** *Let $\langle X_n \rangle$ be a submartingale (or supermartingale) with $K := \sup_n E\big[|X_n|\big] < \infty$. Then $X := \lim_{n\to\infty} X_n$ exists a.s. and $E\big[|X|\big] \le K$.*

To prove this, we consider the number of upcrossings of intervals $[\alpha, \beta]$: Given $X_1, \ldots, X_n$, set $\sigma_1 := \min\{j \in [1, n]\,;\, X_j \le \alpha \text{ or } j = n\}$, $\tau_1 := \min\{j \in (\sigma_1, n]\,;\, X_j \ge \beta \text{ or } j = n\}$, $\sigma_2 := \min\{j \in (\tau_1, n]\,;\, X_j \le \alpha \text{ or } j = n\}$, $\tau_2 := \min\{j \in (\sigma_2, n]\,;\, X_j \ge \beta \text{ or } j = n\}$, etc. These are clearly stopping times. The upcrossings are the intervals $[\sigma_i, \tau_i]$ for which $X_{\sigma_i} \le \alpha$ and $X_{\tau_i} \ge \beta$. Note that each upcrossing is responsible for an increase of $\ge \beta - \alpha$.

**Theorem 35.4 (Doob's Upcrossing Inequality).** *If $\langle X_1, \ldots, X_n \rangle$ is a submartingale and $\alpha < \beta$, then the number $U$ of upcrossings of $[\alpha, \beta]$ satisfies*

$$E[U] \le \frac{E\big[(X_n - \alpha)^+\big]}{\beta - \alpha}.$$

*Proof.* Since $\langle (X_j - \alpha)^+ \rangle$ is a submartingale, we may take $\langle X_j \rangle$ to be nonnegative and $\alpha = 0$. (Note that $U$ doesn't change.) Because all $X_j \ge 0$, we have

$$X_n = X_{\tau_n} \ge X_{\tau_n} - X_{\sigma_1} = \sum_{k=1}^{n} \big(X_{\tau_k} - X_{\sigma_k}\big) + \sum_{k=2}^{n} \big(X_{\sigma_k} - X_{\tau_{k-1}}\big).$$

Since $n \ge \sigma_k \ge \tau_{k-1}$, the proposition (on optional sampling) gives $E[X_{\sigma_k} - X_{\tau_{k-1}}] \ge 0$. Also, $X_{\tau_k} - X_{\sigma_k} \ge \beta$ for each of the $U$ upcrossings and is $\ge 0$ otherwise. Hence $\sum_{k=1}^{n} \big(X_{\tau_k} - X_{\sigma_k}\big) \ge \beta\, U$, so $E[X_n] \ge E[\beta\, U]$, i.e., $E[U] \le E[X_n]/\beta$. ∎

*Proof of Theorem 35.5.* To show that $\lim X_n$ exists a.s. in $[-\infty, \infty]$, it suffices to show that the number of upcrossings of each $[\alpha, \beta]$ is finite a.s. for $\alpha, \beta \in \mathbb{Q}$ with $\alpha < \beta$. Let $U_n$ be the number of upcrossings of $[\alpha, \beta]$ by time $n$. Then

$$E[U_n] \le \frac{E\big[(X_n - \alpha)^+\big]}{\beta - \alpha} \le \frac{E\big[|X_n|\big] + |\alpha|}{\beta - \alpha} \le \frac{K + |\alpha|}{\beta - \alpha}.$$

Since $U_n \uparrow$, the MCT shows that $\lim U_n < \infty$ a.s., as desired.

Let $X := \lim X_n$. Then $|X| = \lim |X_n|$ and Fatou's lemma gives $E\big[|X|\big] \le K$, so $|X| < \infty$ a.s. ∎

**Corollary.** *If $\langle X_n \rangle$ is a nonnegative supermartingale, then $X := \lim_{n\to\infty} X_n$ exists a.s. and $E[X] \le E[X_1]$.* ∎

**Homework:** Let $\langle Y_n \rangle$ be IID with law $(\delta_1 + \delta_{-1})/2$.
(a) Find $B_n \in \sigma(Y_1, \ldots, Y_n)$ such that $P(B_n) \to 0$ and $P(B_n \text{ i.o.}) = 1$.
(b) Define $X_0 := 0$, $X_{n+1} := X_n(1 + Y_{n+1}) + \mathbf{1}_{B_n} Y_{n+1}$ for $n \ge 1$. Show that $\langle X_n \rangle$ is a martingale.
(c) Show that $X_n \to 0$ in probability.
(d) Show that $P[X_n \text{ converges}] = 0$.

It is very important to know when $E[\lim X_n] = \lim E[X_n]$. Recall from the corollary to Theorem 16.14 that if $X_n \to X$ a.s., then $\{X_n\}$ is UI iff $X_n \to X$ in $L^1$ (in which case $E[X_n] \to E[X]$). Thus:

**Proposition.** *If $\langle X_n \rangle$ is a submartingale, then the following are equivalent:*
(i) *$\{X_n\}$ is UI.*
(ii) *$\lim X_n$ exists a.s. and in $L^1$.*
(iii) *$\lim X_n$ exists in $L^1$.*

*Proof.* (ii) ⇒ (iii) is trivial. (ii) ⇒ (i) by the general corollary and (i) ⇒ (ii), (iii) ⇒ (ii) by Theorem 35.5. ∎

However, there is a martingale $\langle X_n \rangle$ that is bounded in $L^1$ with $E[X_n] = 0 = E[\lim X_n]$, yet $E\big[|X_n|\big] \not\to 0$. For example, let $X_1 = X_2 = X_3 = \cdots \sim (\delta_1 + \delta_{-1})/2$. It is also possible to have an example with $X_n \to 0$ a.s.; see the later extra credit problem.

We get an additional equivalence for martingales:

**Theorem.** *If $\langle (X_n, \mathscr{F}_n) \rangle$ is a martingale, then $X_n \to X$ in $L^1$ iff $X_n = E[X \mid \mathscr{F}_n]$ and $X$ has a version in $\sigma\big(\bigcup_n \mathscr{F}_n\big)$.*

Write $\mathscr{F}_\infty := \sigma\big(\bigcup_n \mathscr{F}_n\big)$; we also write this as $\mathscr{F}_n \uparrow \mathscr{F}_\infty$ and as $\mathscr{F}_\infty = \bigvee_n \mathscr{F}_n$. We say that $\langle (X_n, \mathscr{F}_n) \rangle$ is ***closed*** by $(X, \mathscr{F}_\infty)$.

*Proof.* $\Rightarrow$: Because $X_n \xrightarrow{L^1} X$, we have $X_n \mathbf{1}_A \xrightarrow{L^1} X \mathbf{1}_A$ for all $A$. Take $A \in \mathscr{F}_m$. Then

$$E[X\,;\ A] = \lim E[X_n\,;\ A] = E[X_m\,;\ A].$$

That is, $X_m$ is a version of $E[X \mid \mathscr{F}_m]$. Since all $X_n \in \mathscr{F}_\infty$, also $X \in \mathscr{F}_\infty$.

$\Leftarrow$: We have seen that $\{X_n\}$ is UI, so let $X_\infty := \limsup X_n$. We must show that $X = X_\infty$. Since $\bigcup \mathscr{F}_n$ is a $\pi$-system, it suffices to check that $\forall A \in \bigcup \mathscr{F}_n \ \ E[X\,;\ A] = E[X_\infty\,;\ A]$. Indeed, if $A \in \mathscr{F}_n$, then $E[X\,;\ A] = E[X_n\,;\ A] = E[X_\infty\,;\ A]$, as we have just seen in the other part of the proof. ∎

**Theorem 35.6.** *If $\mathscr{F}_n \uparrow \mathscr{F}_\infty$ and $Z \in L^1(P)$, then $E[Z \mid \mathscr{F}_n] \to E[Z \mid \mathscr{F}_\infty]$ a.s.*

*Proof.* Apply the preceding to $X := E[Z \mid \mathscr{F}_\infty]$. ∎

**Corollary (Lévy's 0-1 Law).** *If $\mathscr{F}_n \uparrow \mathscr{F}_\infty$ and $A \in \mathscr{F}_\infty$, then $P[A \mid \mathscr{F}_n] \to \mathbf{1}_A$ a.s.* ∎

**Corollary (Kolmogorov's 0-1 Law).** *If $\{\mathscr{F}_n\}$ are independent, then the tail $\sigma$-field of $\{\mathscr{F}_n\}$ is trivial.*

*Proof.* Let $\mathscr{G}_n := \sigma(\mathscr{F}_1, \ldots, \mathscr{F}_n)$. If $A$ is in the tail $\sigma$-field, then $A \in \bigvee_n \mathscr{G}_n$, so $\mathbf{1}_A = \lim P(A \mid \mathscr{G}_n) = \lim P(A)$ since $A$ is independent of $\mathscr{G}_n$. Thus, $P(A) \in \{0, 1\}$. ∎

**Homework:** 35.4 (In fact, show that if $\exists \epsilon > 0 \ \ \sum_n P[Y_n > 1 + \epsilon] = \infty$, then $X = 0$ a.s.)

**Homework:** Let $Z_n$ be IID, mean-0 random variables. Let $\theta \in L^1$ and $Y_n := \theta + Z_n$. Show that $E[\theta \mid Y_1, \ldots, Y_n] \to \theta$ a.s.

**Homework:** (Pólya's urn) An urn contains 1 red ball and 1 green ball. Choose one ball at random and replace it with another of the same color. Repeat indefinitely. Show that the fraction of red balls in the urn has an a.s. limit. Extra credit: identify the limit distribution.

**Extra Credit Homework:** Give a martingale $\langle X_n \rangle$ such that $\sup E\big[|X_n|\big] < \infty$, for all $k$, $E[X_k] = 0 = \lim X_n$ a.s., but $E\big[|X_n|\big] \not\to 0$.

Recall Example 35.2: Let $\nu$ be a finite, signed measure on $(\Omega, P)$ and $\langle \mathscr{F}_n \rangle$ be a filtration. Suppose that $\nu{\restriction}\mathscr{F}_n \ll P{\restriction}\mathscr{F}_n$ for all $n$. Let $X_n := d(\nu{\restriction}\mathscr{F}_n)/d(P{\restriction}\mathscr{F}_n)$. We saw that $\langle (X_n, \mathscr{F}_n) \rangle$ is a martingale. Let $X := \limsup X_n$. Write $\mathscr{F}_\infty := \bigvee_{n \geq 1} \mathscr{F}_n$.

**Theorem 35.7.** (i) *If* $\nu\restriction\mathscr{F}_\infty \ll P\restriction\mathscr{F}_\infty$*, then* $X := d(\nu\restriction\mathscr{F}_\infty)/d(P\restriction\mathscr{F}_\infty)$*.*
(ii) *If* $\nu\restriction\mathscr{F}_\infty \perp P\restriction\mathscr{F}_\infty$*, then* $P[X = 0] = 1$*.*

This is not hard to deduce from Theorems 35.5 and 35.6—read p. 470.

Also recall Example 35.3, the specialization to $P$ being Lebesgue measure on $(0, 1]$, $\mathscr{F}_n := \sigma(I_k^{(n)}\,;\ 0 \le k < 2^n)$, where $I_k^{(n)} := (k2^{-n}, (k+1)2^{-n}]$. We always have that $\nu\restriction\mathscr{F}_n \ll P\restriction\mathscr{F}_n$ and

$$X_n = \sum_{k=0}^{2^n-1} \mathbf{1}_{I_k^{(n)}} \nu(I_k^{(n)}) 2^n.$$

We get that $X_n \to F'$ $P$-a.s., where $F(\omega) := \nu\big((0, \omega]\big)$ is function of bounded variation: see Example 35.10 on p. 471.

**Theorem 35.3 (Doob's Maximal Inequality).** *If* $\langle X_1, \ldots, X_n\rangle$ *is a submartingale and* $\alpha > 0$*, then*

$$P[\max_{i\le n} X_i \ge \alpha] \le \alpha^{-1} E[X_n\,;\ \max_{i\le n} X_i \ge \alpha] \le \alpha^{-1} E\big[|X_n|\big].$$

This extends Kolmogorov's inequality (Theorem 22.4), since if $S_i$ are the partial sums of independent, mean-0 random variables, then $\langle S_i^2\rangle$ is a submartingale.

*Proof.* Let $\tau_2 := n$ and $\tau_1 := \min\{i \le n\,;\ X_i \ge \alpha \text{ or } i = n\}$. Let $M_k := \max_{i\le k} X_i$. Then for all $k$,

$$[M_n \ge \alpha] \cap [\tau_1 \le k] = [M_k \ge \alpha] \in \mathscr{F}_k,$$

i.e., $[M_n \ge \alpha] \in \mathscr{F}_{\tau_1}$. Thus, the optional sampling theorem gives

$$\alpha\, P[M_n \ge \alpha] \le E[X_{\tau_1}\,;\ M_n \ge \alpha] \le E[X_{\tau_2}\,;\ M_n \ge \alpha] \le E\big[|X_n|\big]. \qquad \blacksquare$$

**Homework:** Show that if $\tau$ is a finite stopping time and $\langle (X_{\tau\wedge n}, \mathscr{F}_n)\rangle$ is a UI submartingale, then for all $\alpha > 0$,

$$P\big[\sup_n X_{\tau\wedge n} \ge \alpha\big] \le \alpha^{-1} E[X_\tau\,;\ \sup_n X_{\tau\wedge n} \ge \alpha] \le \alpha^{-1} E\big[|X_\tau|\big].$$

**Doob's $L^p$ Inequality.** *If* $\langle X_1, \ldots, X_n\rangle$ *is a nonnegative submartingale and* $p > 1$*, then*

$$\|\max_{i\le n} X_i\|_p \le \frac{p}{p-1}\|X_n\|_p.$$

*Proof.* Because of Theorem 35.3, it suffices to show that for any random variables $X, Y \ge 0$, if

$$\forall \alpha > 0 \quad P[X \ge \alpha] \le \alpha^{-1} E[Y\,;\ X \ge \alpha],$$

then

$$\forall p > 1 \quad \|X\|_p \le \frac{p}{p-1}\|Y\|_p.$$

Now

$$\begin{aligned}
\|X\|_p^p = E[X^p] &= \int_0^\infty P[X^p > s]\,ds = \int_0^\infty P[X > s^{1/p}]\,ds \\
&= \int_0^\infty pt^{p-1}\,P[X > t]\,dt \le \int_0^\infty pt^{p-2}\int_\Omega Y\mathbf{1}_{[X\ge t]}\,dP\,dt \\
&= \int_\Omega pY\int_0^\infty t^{p-2}\mathbf{1}_{[X\ge t]}\,dt\,dP = \int_\Omega \frac{p}{p-1}YX^{p-1}\,dP \\
&\le \frac{p}{p-1}\|Y\|_p\|X^{p-1}\|_q,
\end{aligned}$$

where $q^{-1} = 1 - p^{-1} = (p-1)/p$. That is,

$$E[X^p] \le \frac{p}{p-1}\,E[X^p]^{1/q}\|Y\|_p.$$

If $\|X\|_p < \infty$, then this gives the desired inequality. If not, then note that we may replace $X$ by $X \wedge K$ in the hypothesis: $[X \wedge K \ge \alpha] = [X \ge \alpha]$ for $\alpha \le K$ and $[X \wedge K \ge \alpha] = \varnothing$ for $\alpha > K$. This gives

$$\|X \wedge K\|_p \le \frac{p}{p-1}\|Y\|_p,$$

and now we may let $K \to \infty$. ∎

**Doob's $L^p$ Convergence Theorem.** *If $\langle X_n\rangle$ is a martingale or nonnegative submartingale, $p > 1$, and $\sup_n E\big[|X_n|^p\big] < \infty$, then $\exists X$ such that $X_n \to X$ a.s. and in $L^p$.*

Problem 35.4 gave a counter-example for $p = 1$. Also, doubling bets on fair games gives a martingale counter-example.

*Proof.* Since $\|X_n\|_1 \le \|X_n\|_p$, Theorem 35.5 gives $X_n \to X$ a.s. Since $|X_n|$ is a submartingale, Doob's $L^p$ inequality and the MCT (or Fatou's lemma) give $\sup|X_n| \in L^p$. Since $|X_n - X|^p \le 2^p \sup|X_n|^p \in L^1$, the LDCT gives $X_n \to X$ in $L^p$. ∎

**Homework:** Let $p > 1$, $A_n$ be independent with $P(A_n) = \left(\frac{n}{n+1}\right)^p$ for $n \ge 1$, and $X_n := n\mathbf{1}_{\bigcap_{k=1}^{n-1} A_k}$. Show that $\langle X_n\rangle$ is a supermartingale and $E[X_n^p] = 1$, but $X_n \to 0$ a.s. Thus, Doob's $L^p$ convergence theorem does not extend to nonnegative supermartingales.

A sub/super/martingale indexed by the negative integers is called a ***backwards*** or ***reversed*** sub/super/martingale. This adjective refers only to the index set! A filtration can be defined with respect to any poset.

**Example.** Let $\langle Y_n\rangle$ be IID in $L^1(P)$, $S_n := \sum_{k=1}^n Y_k$, $X_{-n} := S_n/n$, and $\mathscr{F}_{-n} := \sigma(S_n, S_{n+1}, \ldots)$. Then for $1 \le k \le n$, $E[Y_1 \mid \mathscr{F}_{-n}] = E[Y_k \mid \mathscr{F}_{-n}]$ by symmetry.

To prove this rigorously, we show that $E[Y_1\,;\ A] = E[Y_k\,;\ A]$ for all $A \in \mathscr{F}_{-n}$. Both sides here depend only on the joint distribution of $Y_1$ (or $Y_k$) and $S_n, S_{n+1}, \ldots$: see the proposition below. So assume that $P$ is a product measure, the law of $\langle Y_n\rangle$ on $\mathbb{R}^\infty$. Then there is a bijection of the probability space interchanging $Y_1$ and $Y_k$ but preserving $P$ and $S_m$ for all $m \ge n$. This proves the equality.

Since $\sum_{k=1}^n E[Y_k \mid \mathscr{F}_{-n}] = E[S_n \mid \mathscr{F}_{-n}] = S_n$, it follows that $E[Y_1 \mid \mathscr{F}_{-n}] = S_n/n = X_{-n}$. Thus, $\langle(X_{-n}, \mathscr{F}_{-n})\rangle$ is a UI reversed martingale.

**Proposition.** *Let $X\colon(\Omega,\mathscr{F},P)\to(\Omega',\mathscr{F}')$ and $Y\colon(\Omega,\mathscr{F},P)\to(\Omega'',\mathscr{F}'')$ be random variables and $f\colon(\Omega',\mathscr{F}')\to\mathbb{R}$ be Borel with $f(X)\in L^1(P)$. Then $E[f(X)\mid Y]$ depends only on $f$ and the law of $(X,Y)$. In other words, for every probability measure $\mu$ on $\mathscr{F}'\times\mathscr{F}''$ such that $\int|f(x)|\,d\mu(x,y)<\infty$, there is a Borel $g_\mu\colon(\Omega'',\mathscr{F}'')\to\mathbb{R}$ such that $g_{(X,Y)_*P}(Y)$ is a version of $E[f(X)\mid Y]$.*

*Proof.* Write $\bar f(x,y):=f(x)$ and $\bar{\mathscr{F}}'':=\{\Omega'\times A\,;\ A\in\mathscr{F}''\}$. Set $g_\mu$ to be the Radon–Nikodym derivative of $(\bar f\mu)\restriction\bar{\mathscr{F}}''$ with respect to $\mu\restriction\bar{\mathscr{F}}''$. A check of the definition of conditional expectation shows that this works: if $\nu:=(X,Y)_*P$, then for all Borel $B$,

$$\int_{[Y\in B]} g_\nu(Y)\,dP=\int_{[y\in B]} g_\nu(y)\,d\nu(x,y)=\int_{[y\in B]} f(x)\,d\nu(x,y)=\int_{[Y\in B]} f(X)\,dP. \qquad\blacksquare$$

Note: every reversed martingale is UI.

Write $\mathscr{F}_{-\infty}:=\bigcap_n\mathscr{F}_{-n}$.

**Theorem 35.8.** *Let $\langle(X_{-n},\mathscr{F}_{-n})\rangle$ be a reversed submartingale. Then its limit, $X_{-\infty}:=\lim_{n\to\infty}X_{-n}$, exists a.s. in $[-\infty,\infty)$ and $X_{-\infty}\le E[X_{-1}\mid\mathscr{F}_{-\infty}]$ a.s. If $\langle X_{-n}\rangle$ is a reversed martingale, then $X_{-\infty}\in(-\infty,\infty)$ a.s. and $X_{-\infty}=E[X_{-1}\mid\mathscr{F}_{-\infty}]$ a.s.*

*Proof.* Given $[\alpha,\beta]$, let $U_n$ be the number of upcrossings of $[\alpha,\beta]$ by $X_{-n},X_{-n+1},\ldots,X_{-1}$. Then the upcrossing lemma gives

$$E[U_n]\le\frac{E\big[|X_{-1}|\big]+|\alpha|}{\beta-\alpha}.$$

Hence $\sup U_n<\infty$ a.s., so $X_{-\infty}:=\lim X_{-n}$ exists a.s. in $[-\infty,\infty]$.

Given $a<0$, $\langle(X_{-n}\vee a,\mathscr{F}_{-n})\rangle$ is a reversed submartingale. We claim it is UI. The rest then follows since $\forall A\in\mathscr{F}_{-\infty}$ $\forall n\ge1$ $E[X_{-n}\vee a\,;\ A]\le E[X_{-1}\vee a\,;\ A]$; taking $n\to\infty$ gives $E[X_{-\infty}\vee a\,;\ A]\le E[X_{-1}\vee a\,;\ A]$. Now take $a\to-\infty$.

To prove UI, let $c>0$. Then

$$c\,P[X_{-n}\vee a\ge c]\le E[X_{-n}\vee a\,;\ X_{-n}\vee a\ge c]\le E[X_{-1}\vee a\,;\ X_{-n}\vee a\ge c]\le E[X_{-1}^+].$$

Therefore $P[X_{-n}\vee a\ge c]\to0$ uniformly in $n$ as $c\to\infty$. Comparing the middle terms in these inequalities gives UI. $\blacksquare$

**Theorem 35.9.** *If $Z\in L^1(P)$ and $\langle\mathscr{F}_{-n}\rangle$ is a filtration, then $E[Z\mid\mathscr{F}_{-n}]\to E[Z\mid\mathscr{F}_{-\infty}]$ a.s.*

*Proof.* $\langle(E[Z\mid\mathscr{F}_{-n}],\mathscr{F}_{-n})\rangle$ is a reversed martingale, so its limit is $Z_{-\infty}=E\big[E[Z\mid\mathscr{F}_{-1}]\bigm|\mathscr{F}_{-\infty}\big]=E[Z\mid\mathscr{F}_{-\infty}]$. $\blacksquare$

**Example.** With the previous example, we have $S_n/n\to X_{-\infty}=E[X_{-1}\mid\mathscr{F}_{-\infty}]=E[Y_1\mid\mathscr{F}_{-\infty}]$ a.s. and in $L^1$. On the other hand, by Kolmogorov's 0-1 law, $\lim S_n/n$ is a constant a.s. Hence $X_{-\infty}=E[X_{-\infty}]=E\big[E[Y_1\mid\mathscr{F}_{-\infty}]\big]=E[Y_1]$ a.s. This gives a new derivation of the SLLN. While it doesn't give Etemadi's version that assumed only pairwise independence, it doesn't need full independence either—exchangeability is enough, though we don't then get that the limit is constant a.s.

Here, we call $\langle X_1, \ldots, X_n \rangle$ ***exchangeable*** if for all permutations $\pi$ of $\langle 1, 2, \ldots, n \rangle$, $\langle X_{\pi(1)}, \ldots, X_{\pi(n)} \rangle$ has the same distribution. We call $\langle X_i\,;\ i \ge 1 \rangle$ ***exchangeable*** if for all $n$, $\langle X_1, \ldots, X_n \rangle$ is exchangeable.

Suppose that all $X_i$ are indicators. It is not hard to give examples of exchangeable $\langle X_1, \ldots, X_n \rangle$; for example, pick $k$ uniformly out of $n$ to be 1. In fact, every exchangeable sequence $\langle X_1, \ldots, X_n \rangle$ of indicators is a mixture of these examples, because if $S_n := \sum_{k=1}^n X_k$, then the law of $\langle X_1, \ldots, X_n \rangle$ conditional on $S_n$ is the uniform measure on $S_n$ out of $n$. The analogue for infinite exchangeable sequences is that every example is a mixture of Bernoulli trials:

**Theorem 35.10 (de Finetti's Theorem).** *If $\langle X_i\,;\ i \ge 1 \rangle$ is an exchangeable sequence of indicators, then there is a random variable $\theta$ with values in $[0, 1]$ such that given $\theta$, $\langle X_i \rangle$ is Bernoulli($\theta$), i.e., for all $m$ and $u_i \in \{0, 1\}$, $P[X_1 = u_1, \ldots, X_m = u_m \mid \theta] = \theta^k (1-\theta)^{m-k}$ a.s., where $k := \sum_{i=1}^m u_i$.*

*Proof.* This is just a limit of the finite case: Define $\mathscr{F}_{-n} := \sigma(S_n, S_{n+1}, \ldots)$. Fix $m \ge 1$, and consider $n \ge m$. Because $\langle X_1, \ldots, X_n \rangle$ is exchangeable, so is its law conditional on $\mathscr{F}_{-n}$ by the preceding proposition. Given $u_i \in \{0, 1\}$, write $A := [X_1 = u_1, \ldots, X_m = u_m]$. Then with $k := \sum_{i=1}^m u_i$, we have a.s.

$$P(A \mid \mathscr{F}_{-n}) = \frac{S_n}{n} \cdot \frac{S_n - 1}{n-1} \cdots \frac{S_n - k + 1}{n - k + 1} \cdot \frac{n - S_n}{n - k} \cdot \frac{n - S_n - 1}{n - k - 1} \cdots \frac{n - S_n - m + k + 1}{n - m + 1}.$$

Taking $n \to \infty$ and using Theorem 35.9 gives $P(A \mid \mathscr{F}_{-\infty}) = \theta^k (1-\theta)^{m-k}$ a.s. Because this is $\theta$-measurable, the result follows by the tower property, Theorem 34.4. ∎

**Homework:** (Pólya's urn) An urn contains $r$ red balls and $g$ green balls. Choose one ball at random and replace it with another of the same color. Repeat indefinitely. Let $X_n$ be the indicator that the $n$th ball is red. Show that $\langle X_n \rangle$ is exchangeable. Extra credit: identify the distribution of the random variable $\theta$ that gives the mixture, in this case, of de Finetti's theorem.

**Homework:** With the notation of the proof of de Finetti's theorem, show that $\mathscr{F}_{-\infty} = \sigma(\theta)$.

**Homework:** Let $\langle X_i\,;\ i \ge 1 \rangle$ be simple random variables, each taking values in a finite set, $U$.

(a) Suppose that $\langle X_1, \ldots, X_n \rangle$ is exchangeable. Define $Z_n(u) := |\{i \le n\,;\ X_i = u\}|$. Describe the law of $\langle X_1, \ldots, X_n \rangle$ conditional on $Z_n := \langle Z_n(u)\,;\ u \in U \rangle$.

(b) Suppose that $\langle X_i\,;\ i \ge 1 \rangle$ is exchangeable. Show that $Z := \lim_{n\to\infty} Z_n/n$ exists a.s. *Hint:* Let $\mathscr{F}_{-n} := \sigma(Z_n, X_{n+1}, X_{n+2}, \ldots)$.

(c) With the assumptions and notation of part (b), show that given $Z$, $\langle X_i \rangle$ is IID, namely, for all $m$ and $u_i \in U$, $P[X_1 = u_1, \ldots, X_m = u_m \mid W] = \prod_{i=1}^m Z(u_i)$.

You may omit pp. 474–478.