PROBABILITY THEORY: LECTURE NOTES 6

CHANDRA NAIR

Disclaimer: These notes are created from a variety of great references in the field, most notably, Varadhan's lecture notes and Dembo's lecture notes.

9. Martingales

In this section we will study an important class of sequences of random variables that arise naturally in many settings, and unnaturally in other settings as a tool to prove certain results.

Definition 1. Given a measurable space $(\Omega, \mathcal{F})$ a *filtration* is an increasing sequence of σ-algebra's $\mathcal{F}_i \subset \mathcal{F}$, i.e. if $i \leq j$ then $\mathcal{F}_i \subseteq \mathcal{F}_j$.

Definition 2. A sequence of random variables $\{X_i\}$ is said to be *adapted* to the filtration $\{\mathcal{F}_i\}$ of sigma-algebra's if X_i is $\mathcal{F}_i$-measurable for every i.

Definition 3. A sequence $\{X_i\}, i \geq 1$ of integrable random variables, adapted to a filtration $\{\mathcal{F}_i\}, i \geq 1$, is said to be a *martingale* if, for all $i \geq 1$,

$$\mathrm{E}(X_{i+1}|\mathcal{F}_i) = X_i \ a.s.$$

A sequence $\{X_i\}, i \geq 1$ of integrable random variables, adapted to a filtration $\{\mathcal{F}_i\}, i \geq 1$, is said to be a *sub-martingale* if, for all $i \geq 1$,

$$\mathrm{E}(X_{i+1}|\mathcal{F}_i) \geq X_i \ a.s.$$

A sequence $\{X_i\}, i \geq 1$ of integrable random variables, adapted to a filtration $\{\mathcal{F}_i\}, i \geq 1$, is said to be a *super-martingale* if, for all $i \geq 1$,

$$\mathrm{E}(X_{i+1}|\mathcal{F}_i) \leq X_i \ a.s.$$

Remark 1. Usually, given a sequence of random variables $\{X_n\}$, we can define a natural filtration according to $\mathcal{F}_n = \sigma(X_1, .., X_n)$. Hence, unless otherwise specified, this will be the underlying filtration.

The following statements are easy to verify:

(1) If $\{X_i\}$ is a sequence of independent integrable random variables. Define $S_n = X_1 + \cdots + X_n$. Then $\{\alpha(S_n - \mathrm{E}(S_n)) + \beta\}$ is a martingale.
(2) If $\{X_i\}$ is a martingale and Φ is a concave function such that $\Phi(X_i)$ is integrable, then $\{\Phi(X_i)\}$ is a super-martingale.

    Proof. From Jensen's inequality $\mathrm{E}(\Phi(X_{i+1})|\mathcal{F}_i) \leq \Phi(\mathrm{E}(X_{i+1}|\mathcal{F}_i)) = \Phi(X_i)\ a.s.$. □

(3) Similarly, if $\{X_i\}$ is a martingale and Φ is a non-decreasing convex function such that $\Phi(X_i)$ is integrable, then $\{\Phi(X_i)\}$ is a sub-martingale.

Date: December 18, 2019.

(4) Given a filtration $\mathcal{F}_i$, if X is integrable then $X_i := \mathrm{E}(X|\mathcal{F}_i)$ is a martingale sequence.
(5) If $\{X_i\}$ is a martingale, then for $i \leq j$, $\mathrm{E}(X_j|\mathcal{F}_i) = X_i$ *a.s.*. The equality becomes the appropriate inequality for *sub*-martingales and *super*-martingales.

Definition 4. Given a filtration $\{\mathcal{F}_i\}$, a sequence $\{X_i\}, i \geq 1$ of integrable random variables is said to be a *predictable* if, for all $i \geq 1$, X_{i+1} is $\mathcal{F}_i$-measurable.

Theorem 1 (Doob's inequality)**.** *Suppose X_n is a martingale (or a non-negative sub-martingale) sequence of length n. Define*

$$A_l := \{\omega : \sup_{1 \leq j \leq n} |X_j(\omega)| \geq l\}.$$

Then

$$P(A_l) \leq \frac{1}{l} \int_{A_l} |X_n| dP \leq \frac{1}{l} \mathrm{E}(|X_n|).$$

Proof. First note that $\{|X_n|\}$ is a sub-martingale. Then we partition A_n as follows. Let

$$B_j = \{\omega : |X_1(\omega)| < l, \ldots, |X_{j-1}(\omega)| < l, |X_j(\omega)| \geq l\},\ 1 \leq j \leq n.$$

Note that $B_j \in \mathcal{F}_j$. Further since $\{|X_n|\}$ is a sub-martingale, we have that $\mathrm{E}(|X_n||\mathcal{F}_j) \geq |X_j|$ *a.s.*, implying that

$$\int_{B_j} |X_n| dP = \int_{B_j} \mathrm{E}(|X_n||\mathcal{F}_j) dP \geq \int_{B_j} |X_j| dP \geq l P(B_j).$$

Observing that $A = \sqcup_j B_j$ completes the proof. ∎

Lemma 1. *If X, Y are two non-negative random variables on the same probability space such that*

$$\mathrm{P}(Y \geq l) \leq \frac{1}{l} \int_{Y \geq l} X dP$$

then for every $p > 1$

$$\int Y^p dP \leq \left(\frac{p}{p-1}\right)^p \int X^p dP.$$

Proof. Let $F(x) = \mathrm{P}(Y \leq x)$ be the distribution function. Let $T(x) = 1 - F(x) = P(Y > x)$. Then

$$\begin{aligned}
\int Y^p dP &= \int_0^\infty y^p dF(y) \\
&= -\int_0^\infty y^p dT(y) \\
&= p\int_0^\infty y^{p-1} T(y) dy && \text{(integration by parts)} \\
&\leq p\int_0^\infty y^{p-2}\left(\int_{Y>y} X dP\right) dy \\
&= p\int X\left(\int_0^Y y^{p-2} dy\right) dP \\
&= \frac{p}{p-1}\int XY^{p-1} dP \\
&\leq \frac{p}{p-1}\left(\int X^p dP\right)^{\frac{1}{p}}\left(\int Y^p dP\right)^{\frac{p-1}{p}} && \text{(Hölder).}
\end{aligned}$$

If $\mathrm{E}(Y^p) < \infty$, we are done. Otherwise, obtain the result for $Y_m = \min\{Y, m\}$ and pass to the limit to get the desired inequality. □

Corollary 1 (Doob)**.** *Let X_n is a sub-martingale sequence of length n. Define $S = \sup_{1\leq j\leq n} |X_j(\omega)|$. Then*

$$\int S^p dP \leq \left(\frac{p}{p-1}\right)^p \int |X_n|^p dP.$$

Definition 5 (Martingale Transform)**.** Given a *predictable* process C_n and a martingale sequence $\{X_n\}$, the martingale transform is given by

$$(C \circ X)_n = \sum_{m=1}^{n} C_m(X_m - X_{m-1}).$$

Remark 2. This is the discrete analogue of a stochastic integral. Note that if C_m is bounded, then $(C \circ X)_n$ is a Martingale(why?).

Definition 6 (Stopping Time)**.** A mapping $T : \Omega \mapsto \{0, 1, 2, 3, ..., \infty\}$ is called a *stopping time* with respect to a filtration $\{\mathcal{F}_n\}_{n\geq 0}$ if

$$\{\omega : T(\omega) \leq n\} \in \mathcal{F}_n, \quad \forall n.$$

Theorem 2. *If X_n is a super-martingale sequence and T is a stopping time, then the stopped process*

$$X_n^T(\omega) := X_{T(\omega)\wedge n}(\omega)$$

is a super-martingale. Consequently,

$$\mathrm{E}\left(X_{T\wedge n}\right) \leq \mathrm{E}(X_0).$$

If X_n is a martingale, then $\{X_n^T\}$ is a martingale and further equality holds above.

Proof. Observe that X_n^T is $\mathcal{F}_n$ measurable from the definitions. Now

$$\begin{aligned}
\mathrm{E}(X_{T\wedge(n+1)}|\mathcal{F}_n) &= \mathrm{E}(X_{T\wedge(n+1)}1_{T\le n} + X_{T\wedge(n+1)}1_{T>n}|\mathcal{F}_n) \\
&= \mathrm{E}(X_{T\wedge n}1_{T\le n} + X_{n+1}1_{T>n}|\mathcal{F}_n) \\
&= X_{T\wedge n}1_{T\le n} + 1_{T>n}\,\mathrm{E}(X_{n+1}|\mathcal{F}_n) \ a.s. \\
&\le X_{T\wedge n}1_{T\le n} + 1_{T>n}X_n \ a.s. \\
&= X_{T\wedge n}.
\end{aligned}$$

The proof for the martingale case follows with the inequality in the above step replaced with equality (from definition). □

Theorem 3 (Doob's Optional Stopping Time Theorem)**.** *The following results hold:*

1. *Let X_n be a supermartingale and T is a stopping time. If any of the following conditions is satisfied:*
   (a) T is bounded
   (b) X_n is uniformly bounded and T is finite almost surely
   (c) $E(T) < \infty$ and $|X_n - X_{n-1}|$ is uniformly bounded (say by K)
   then $E(X_T) \le E(X_0)$.
2. *If X_n is a martingale and any of the above three conditions hold, then* $\mathrm{E}(X_T) = \mathrm{E}(X_0)$.
3. *Let $\{C_n\}$ be a bounded predictable sequence and $\{X_n\}$ be a martingale such that $|X_n - X_{n-1}|$ is uniformly bounded, and T is a stopping time such that* $\mathrm{E}(T) < \infty$*. Then* $\mathrm{E}\left((C \circ X)_T\right) = 0$.
4. *If X_n is a non-negative super-martingale and T is almost surely finite, then* $\mathrm{E}(X_T) \le \mathrm{E}(X_0)$.

Proof. We will establish each part in sequence.

1. Here X_n is a supermartingale.
   (a) Let $T \le N$, then $X_{T\wedge N} = X_T$. The inequality follows from the previous theorem.
   (b) If T is finite almost surely, then $X_{T\wedge n} \to X_T$ almost surely. Since X_n is uniformly bounded (hence $X_{T\wedge n}, X_T$ are also bounded by the same constant), bounded convergence theorem implies that $E(X_{T\wedge n}) \to \mathrm{E}(X_T)$. Applying previous theorem completes the proof.
   (c) Note that
   $$\begin{aligned}
   \mathrm{E}(|X_{T\wedge n} - X_0|) &= \mathrm{E}\left(\left|\sum_{k=1}^{T\wedge n} X_k - X_{k-1}\right|\right) \\
   &\le \sum_{k=1}^{T\wedge n} \mathrm{E}(|X_k - X_{k-1}|) \le \mathrm{E}(K(T\wedge n)) \le \mathrm{E}(KT) < \infty.
   \end{aligned}$$
   Hence dominated convergence theorem says that
   $$\mathrm{E}(X_T - X_0) = \mathrm{E}(\lim_n (X_{T\wedge n} - X_0)) = \lim_n \mathrm{E}(X_{T\wedge n} - X_0) \le 0.$$
2. Apply the earlier part to X_n and $-X_n$ to get the desired equality.
3. We know that $(C \circ X)_n$ is a zero-mean martingale. Further the given conditions imply that $|(C\circ X)_n - (C\circ X)_{n-1}|$ in uniformly bounded. Hence, from previous parts we are done.

(4) This is a direct consequence of Fatou's Lemma and that $\mathrm{E}(X_{T\wedge n}) \leq \mathrm{E}(X_0)$.

□

Lemma 2. *If T is a stopping time and N be a number such that $P(T \leq n+N|\mathcal{F}_n) > \epsilon$ a.s. then $E(T) < \infty$.*

Proof. Observe that

$$\begin{aligned} P(kN < T) &= \mathrm{E}(1_{T>(k-1)N} 1_{T>(k-1)N+N}) \\ &= \mathrm{E}(\mathrm{E}(1_{T>(k-1)N} 1_{T>(k-1)N+N}|\mathcal{F}_{(k-1)N})) \\ &= \mathrm{E}(1_{T>(k-1)N}\, \mathrm{E}(1_{T>(k-1)N+N}|\mathcal{F}_{(k-1)N}) \\ &\leq (1-\epsilon)\, \mathrm{E}(1_{T>(k-1)N}) = (1-\epsilon)P(T > (k-1)N). \end{aligned}$$

Hence by induction $P(T > kN) \leq (1-\epsilon)^k$.

Now note that

$$T \leq N 1_{T\leq N} + \sum_{k=2}^{n} kN 1_{(k-1)N \leq T \leq kN}.$$

Taking expectations we obtain

$$\mathrm{E}(T) \leq N \left(1 + \sum_{k=2}^{n} k(1-\epsilon)^{k-1}\right) = \frac{N}{\epsilon^2}.$$

□

Example 9.1. Consider a Monkey typing characters on a 26-key keyboard. Assume that the characters typed are i.i.d. and each characters is uniformly chosen. What is the expected time before the Monkey types "WOWOWOW"?

Solution. The expected time is $26^7 + 26^5 + 26^3 + 26$. A beautiful argument using Martingales for such problems was developed by Li(1980).

Let $X_n, n \geq 1$ denote the characters typed by the monkey. Define a stopping time according to

$$T = \inf\{n : (X_{n-6}, .., X_n) = (W, O, W, O, W, O, W)\}.$$

Clearly this stopping time satisfies the previous lemma with $N = 7$ and $\epsilon = (26)^{-7}$. Hence $\mathrm{E}(T) < \infty$. Define the following doubly-indexed process $Y_{n,m}, 1 \leq m, m-1 \leq n$ according to the following: $Y_{m-1,m} = 1$, $Y_{m+l-1,m} = 26^{l\wedge 7}$ for $l \geq 1$ if $(X_m, .., X_{m+(l\wedge 7)-1})$ matches the first $(l \wedge 7)$ characters of "WOWOWOW", else set $Y_{m+l-1,m} = 0$. Clearly $Y_{n,m}$ is $\mathcal{F}_n$ measurable. Further $\mathrm{E}(Y_{n+1,m}|\mathcal{F}_{n+1}) = Y_{n,m}$ for all $n \geq m-1$. Further if $n \geq m+6$ then $Y_{n+1,m} = Y_{n,m}$, and $Y_{n,m} \leq (26)^7$.

Observe that $M_n = \sum_{m=1}^{n+1} Y_{n,m} - n - 1$ is a zero-mean martingale. Further observe that

$$\begin{aligned} M_n - M_{n-1} &= \sum_{m=1}^{n} Y_{n,m} - Y_{n-1,m} \\ &= \sum_{m=1\vee(n-6)}^{n} Y_{n,m} - Y_{n-1,m}. \end{aligned}$$

Hence $|M_n - M_{n-1}| \leq 2 \times 7 \times (26)^7$ (can improve this easily). Therefore, we can apply part $c)$ of Doob's optional stopping time theorem and obtain

$$\mathrm{E}(M_T) = 0 \implies \mathrm{E}\left(\sum_{m=1}^{T+1} Y_{T,m} - T - 1\right) = 0.$$

From the definition of the stopping time and $Y_{n,m}$, we see that $Y_{T,T-6} = (26)^7, Y_{T,T-4} = (26)^5, Y_{T,T-2} = (26)^3, Y_{T,T} = 26, Y_{T,T+1} = 1$, and the rest are zeros. Therefore $E(T) = 26^7 + 26^5 + 26^3 + 26$. □

Here is a simple example that demonstrates the following: X_n is a martingale with bounded increments, i.e $|X_n - X_{n-1}| \leq K$. T is a stopping time that is almost surely finite. Yet $\mathrm{E}(X_T) \neq \mathrm{E}(X_0)$. Consider a symmetric random walk in 1-dimension with $X_0 = 0$. Let X_n denote the location of the random walk at time n. Let $T = \inf\{n : X_n = 1\}$. Then clearly, if T is finite almost surely, then $\mathrm{E}(X_T) = 1 \neq \mathrm{E}(X_0) = 0$. To show that T is finite almost surely, define a sequence of events

$$S_n = \{\omega : X_n(\omega) = 1\}.$$

Clearly T is finite almost surely if $\mathrm{P}(\cup_n S_n) = 1$. In fact, we can see that $\mathrm{P}(S_n \ i.o.) = 1$ since $\mathrm{P}(S_{2n+1}) = \binom{2n+1}{n}\frac{1}{2^{2n+1}} = \Theta(1/\sqrt{n})$ which sums to infinity (now apply Borell-Cantelli).

9.1. **Convergence theorems.**

Definition 7 (Upcrossing)**.** Given a sequence of random variables $\{X_n\}$ and two numbers $a < b$, we define a non-negative integer-valued non-decreasing sequence of random variables $U_n[a,b]$ according to

$$U_n[a,b] = \max\{k : \exists\ 0 \leq s_1 < t_1 < \cdots < s_k < t_k \leq n,\ s/t\ X_{s_i} \leq a, X_{t_i} > b\}.$$

If $\{X_n\}$ is adapted to the filtration $\{\mathcal{F}_n\}$ then $\{U_n[a,b]\}$ is also adapted to the filtration $\{\mathcal{F}_n\}$.

Lemma 3 (Doob's Upcrossing inequality)**.** *(1) Let $\{X_n\}$ be a supermartingale. Then for any $n \geq 1$ and $a < b$*

$$(b-a)\,\mathrm{E}(U_n[a,b]) \leq \mathrm{E}((X_n - a)_-) \leq \mathrm{E}(|X_n|) + |a|.$$

(2) Let $\{X_n\}$ be a submartingale. Then for any $n \geq 1$ and $a < b$

$$(b-a)\,\mathrm{E}(U_n[a,b]) \leq \mathrm{E}((X_n - a)_+) - \mathrm{E}((X_0 - a)_+) \leq \mathrm{E}(|X_n|) + |a|.$$

Proof. (1) Define a predictable process inductively as follows

$$C_1 = 1_{X_0 \leq a}$$
$$C_n = 1_{C_{n-1}=1}1_{X_{n-1}\leq b} + 1_{C_{n-1}=0}1_{X_{n-1}\leq a}.$$

In words, C_n is a sequence that takes a value 1, starting from an instance the process goes below a, till the instance the process goes above b for the first time. Then C_n becomes zero, and turns to 1 only when the process goes below a again.

Define $Y_0 = 0$ and $Y_n = (C \circ X)_n, n \geq 1$ and note that this is a supermartingale. Clearly,

$$Y_n \geq (b-a)U_n[a,b] - \max\{0, a - X_n\}(=: (X_n - a)_-).$$

Taking expectations and noting that $\mathrm{E}(Y_n) \le \mathrm{E}(Y_0) = 0$ we obtain the result.

(2) Let $Z_n = (X_n - a)_+ + a$. Observe that Z_n is a sub-martingale and further it has the same number of up-crossings as X_n. Define $Y_n = (C \circ Z)_n, n \ge 1$ as before. Clearly

$$Y_n \ge (b-a)U_n[a,b].$$

Define similarly $\tilde{Y}_n = ((1-C)\circ Z)_n, n \ge 1$. Clearly $\tilde{Y}_n$ is also a sub-martingale and $\mathrm{E}(\tilde{Y}_n) \ge 0$ (verify). Now $Y_n + \tilde{Y}_n = Z_n - Z_0$, hence $\mathrm{E}(Y_n) \le \mathrm{E}(Z_n - Z_0) = \mathrm{E}((X_n - a)_+) - \mathrm{E}((X_0 - a)_+)$.

□

Theorem 4 (Martingale Convergence Theorem)**.** *Let X_n be a super(or sub)-martingale with $\sup_n \mathrm{E}(|X_n|) < \infty$, then X_n will converge almost surely to a limit.*

Proof. Clearly we have

$$\begin{aligned}\{\omega : \liminf_n X_n(\omega) < \limsup_n X_n(\omega)\} &= \cup_{a,b\in Q, a<b}\{\omega : \liminf_n X_n(\omega) < a < b < \limsup_n X_n(\omega)\} \\ &= \cup_{a,b\in Q,a<b}\{\omega : U_\infty[a,b] = \infty\}\end{aligned}$$

From Doob's upcrossing inequality and monotone convergence theorem we have

$$(b-a)\,\mathrm{E}(U_\infty[a,b]) \le \sup_n \mathrm{E}(|X_n|) + |a| < \infty.$$

Hence for every $a, b \in Q, a < b$ we have $P(\{\omega : U_\infty[a,b] = \infty\}) = 0$. □

Corollary 2. *If $\{X_n\}$ is a non-negative supermartingale, then $E(|X_n|) = \mathrm{E}(X_n) \le \mathrm{E}(X_0) < \infty$. Hence it always converges.*

Theorem 5. *Let $X_n, n \ge 0$ be a martingale with bounded increments, i.e. $|X_n - X_{n-1}| \le K$ for all n, ω. Define*

$$C = \{\omega : \lim_n X_n(\omega) \text{ exists and is finite}\}$$
$$D = \{\omega : \liminf_n X_n(\omega) = -\infty \text{ and } \limsup_n X_n(\omega) = \infty\}$$

Then $P(C \cup D) = 1$.

Proof. Since $X_n - X_0$ is also a martingale with bounded increments, we assume w.l.o.g that $X_0 = 0$. For any $M > 0$, define the stopping time $T_M = \inf\{n : X_n < -M\}$. Then $X_{n\wedge T_M} + K + M$ is a martingale, and further $X_{n\wedge T_M} + M + K \ge 0$. Applying Corollary 2 we see that $X_{n\wedge T_M} + K + M$ has a finite limit almost surely. This implies that on

$$\{\omega : T_M = \infty\},$$

the sequence X_n has a finite limit almost surely. Taking $M \to \infty$ along the integers we see that the sequence X_n has a finite limit almost surely on $\{\omega : \inf_n X_n(\omega) > -\infty\}$. Similarly by taking $-X_n$ we can argue that the sequence X_n has a finite limit almost surely on $\{\omega : \sup_n X_n(\omega) < \infty\}$. If $\inf_n X_n = -\infty$ and $\sup_n X_n = +\infty$ then clearly $\omega \in D$ completing the proof. □

Theorem 6. *Let $\{X_n\}$ be a Martingale adapted to the filtration $\{\mathcal{F}_n\}$. If $\sup_n \mathrm{E}(X_n^2) < \infty$ then $X_n \to X_\infty$ a.s. and $\mathrm{E}((X_n - X_\infty)^2) \to 0$. Further $\mathrm{E}(X_\infty^2) < \infty$ and $X_n = \mathrm{E}(X_\infty|\mathcal{F}_n)$ a.s..*

Proof. W.l.o.g. let us center the Martingale and assume $X_0 = 0$. Further define $Y_n = X_n - X_{n-1}, n \geq 1$. Observe that

$$\mathrm{E}(Y_n X_{n-1}) = \mathrm{E}(X_{n-1}(X_n - X_{n-1})) = \mathrm{E}(\mathrm{E}(X_{n-1}(X_n - X_{n-1})|\mathcal{F}_{n-1})) = 0$$

In the last step we used that $\mathrm{E}(X_{n-1}X_n|\mathcal{F}_n) = X_{n-1}\,\mathrm{E}(X_n|\mathcal{F}_n) = X_{n-1}^2$ *a.s.* with the justification that tower property holds as $X_{n-1}X_n$ is integrable (Cauchy-Schwatrz). Hence by induction

$$\mathrm{E}(X_n^2) = \sum_{m=1}^{n} \mathrm{E}(Y_m^2).$$

Since $\sup_n \mathrm{E}(X_n^2) < \infty$ we have $\sup_n \mathrm{E}(|X_n|) < \infty$; and Doob's convergence theorem yields the almost sure convergence. Note that (Fatou yields)

$$\mathrm{E}((X_\infty - X_n)^2) \leq \lim_{m\to\infty} \mathrm{E}((X_m - X_n)^2) = \sum_{m=n+1}^{\infty} \mathrm{E}(Y_m^2) \to 0$$

as $n \to \infty$ as $\sum_{n\geq 1} \mathrm{E}(Y_n^2)$ is finite. Fatou also yields $\mathrm{E}(X_\infty^2) < \infty$.

Note that for $m > n$

$$\mathrm{E}((X_\infty - X_m)|\mathcal{F}_n) = \mathrm{E}(X_\infty|\mathcal{F}_n) - X_n \ a.s.$$

Since $\mathrm{E}((X_\infty - X_m)^2) \to 0$ we have $X_n = \mathrm{E}(X_\infty|\mathcal{F}_n)$ *a.s.* □

Lemma 4. *Let X_n be integrable and X be integrable. Then $E(|X_n - X|) \to 0$ if (i) $X_n \to X$ in probability and (ii) $\{X_n\}$ is uniformly integrable.*

Proof. Let $Y_n^K = X_n 1_{|X_n|\leq K} + \mathrm{sgn}(X_n)K$ and $Y^K = X 1_{|X|\leq K} + \mathrm{sgn}(X)K$. Since $|Y_n^K - Y^K| \leq |X_n - X|$ we see that $Y_n^K \to Y^K$ in probability and hence from bounded convergence theorem $\mathrm{E}(|Y_n^K - Y^K|) \to 0$.
Note that UI implies $\sup_n E(|X_n - Y_n^K|) \leq \sup_n \mathrm{E}(|X_n| 1_{|X_n|>K}) \to 0$ and $K \to \infty$.
Now

$$\begin{aligned}
\limsup_n \mathrm{E}(|X_n - X|) &\leq \limsup_n \mathrm{E}(|X_n - Y_n^K|) + \limsup_n \mathrm{E}(|Y_n^K - Y^K|) + \mathrm{E}(|X - Y^K|) \\
&\leq \sup_n \mathrm{E}(|X_n| 1_{|X_n|>K}) + \mathrm{E}(|X| 1_{|X|>K}),
\end{aligned}$$

and the right hand side tends to 0 as $K \to \infty$. □

Theorem 7. *Let $\{X_n\}$ be an uniformly-integrable Martingale adapted to the filtration $\{\mathcal{F}_n\}$. Then $X_n \to X_\infty$ a.s. and $\mathrm{E}(|X_n - X_\infty|) \to 0$. Further $\mathrm{E}(|X_\infty|) < \infty$ and $X_n = \mathrm{E}(X_\infty|\mathcal{F}_n)$ a.s..*

Proof. $\{X_n\}$ is U-I implies $\sup_n \mathrm{E}(|X_n|) < \infty$. Hence, from convergence theorem, $X_n \to X_\infty$ a.s. and $\mathrm{E}(|X_\infty|) < \infty$. From previous lemma, we also have $\mathrm{E}(|X_n - X_\infty|) \to 0$. Note that for $m > n$

$$\mathrm{E}((X_\infty - X_m)|\mathcal{F}_n) = \mathrm{E}(X_\infty|\mathcal{F}_n) - X_n \ a.s.$$

Since $\mathrm{E}(|X_\infty - X_m|) \to 0$ we have

$$\mathrm{E}(|\,\mathrm{E}(X_\infty|\mathcal{F}_n) - X_n|) \leq \mathrm{E}(\mathrm{E}(|X_\infty - X_m||\mathcal{F}_n)) = \mathrm{E}(|X_\infty - X_m|) \to 0.$$

Hence $X_n = \mathrm{E}(X_\infty|\mathcal{F}_n)$ *a.s.* □

Theorem 8. *Fix $p > 1$. Let $\{X_n\}$ be a Martingale adapted to the filtration $\{\mathcal{F}_n\}$. If $\sup_n \mathrm{E}(|X_n|^p) < \infty$ then $X_n \to X_\infty$ a.s. and $\mathrm{E}(|X_n - X_\infty|^p) \to 0$. Further $\mathrm{E}(|X_\infty|^p) < \infty$ and $X_n = \mathrm{E}(X_\infty|\mathcal{F}_n)$ a.s..*

Proof. $\sup_n \mathrm{E}(|X_n|^p) < \infty$ implies uniform integrability. Further Martingale convergence theorem implies almost sure convergence to a random variable X_∞ and Fatou implies X_∞ satisfies $\mathrm{E}(|X_\infty|^p) < \infty$. Previous theorem implies convergence in L_1.

Further a similar argument as before yields that

$$\mathrm{E}((X_\infty - X_m)|\mathcal{F}_n) = \mathrm{E}(X_\infty|\mathcal{F}_n) - X_n \ a.s.$$

Since $\mathrm{E}(|X_\infty - X_m|) \to 0$ we have $X_n = \mathrm{E}(X_\infty|\mathcal{F}_n)\ a.s.$

Define Y^K as before and let $Y_n^K = \mathrm{E}(Y^K|\mathcal{F}_n)$. Note that $\|X_n - Y_n^K\|_p \leq \|X - Y^K\|_p$. Hence

$$\|X_n - X\|_p \leq \|X_n - Y_n^K\|_p + \|X - Y^K\|_p + \|Y^K - Y_n^K\|_p \leq 2\|X - Y^K\|_p + \|Y^K - Y_n^K\|_p.$$

Now Y_n^K is a bounded martingale sequence. Hence Y_n^K converges to Y_∞^K a.s. and in L_2. Further $\mathrm{E}(Y_\infty^K|\mathcal{F}_n) = Y_n$ a.s.

What is left is to show that $Y_\infty^K = Y^K$ almost surely. By construction, $X_\infty, Y^K, Y_\infty^K$ are $\sigma(\cup_n \mathcal{F}_n)$-measurable. Further since, for every $\mathcal{F}_n$

$$\mathrm{E}(Y_\infty^K|\mathcal{F}_n) = \mathrm{E}(Y^K|\mathcal{F}_n) a.s$$

we have that, for all n and for all $A \in \mathcal{F}_n$

$$\int_A Y_\infty^K dP = \int_A Y^K dP.$$

The collection of all A for which the above equality holds is a monotone class (here use integrability of Y_∞^K and Y^K) and we are done. □

Theorem 9 (Doob's decomposition)**.** *Let $\{X_n\}, n \geq 0$ be a process adapted to $\{\mathcal{F}_n\}$. Then we can express*

$$X_n = X_0 + M_n + A_n,$$

where M_n is a martingale null at zero and A_n is a predictable. Further this decomposition is unique almost surely. Finally $\{X_n\}$ is a sub-martingale if and only if A_n is a non-decreasing sequence, almost surely.

Proof. Note that if such a decomposition exists, then

$$\mathrm{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) = A_n - A_{n-1}.$$

Therefore, by telescopic sum,

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

since $A_0 = 0$. What remains to be verified is that $\{X_n - X_0 - A_n\}$ is a Martingale, but this is immediate. The uniqueness comes from the construction and essential uniqueness of the conditional expectation. The sub-martingale consequence is also immediate from the construction. □

Definition 8 (Angle-bracket process)**.** Let $\{X_n\}$ be a square-integrable martingale which is null at zero. Then $\{X_n^2\}$ is a sub-martingale and it has a Doob's decomposition leading to a predictable process $\{A_n\}$. Then we define $A_n = \langle X_n \rangle$.

Note that if $\mathrm{E}(A_\infty) < \infty$ then $\sup_n \mathrm{E}(X_n^2) \leq \mathrm{E}(A_\infty) < \infty$ and X_n converges almost-surely and in L_2.

Theorem 10. *Let $\{X_n\}$ be a square-integrable martingale which is null at zero. Let $A_n = \langle X_n \rangle$. Then*

(i) If $A_\infty(\omega) < \infty$ then $\lim_n X_n(\omega)$ exists (except possibly on a null set)
(ii) If X_n has uniformly bounded increments, the converse is true: i.e. when $\lim_n X_n(\omega)$ exists and is finite, then $A_\infty(\omega) < \infty$ (except possibly on a null set)

Proof. (*i*) Define a stopping time

$$S(k) = \inf\{n : A_{n+1} > k\}.$$

Then note that

$$A_{n\wedge S(k)} = A_{S(k)} 1_{S(k)\leq n-1} + A_n 1_{S(k)>n-1},$$

hence $A_{n\wedge S(k)}$ is predictable.

Observe that

$$\mathrm{E}(X_{n\wedge S(k)}^2 - A_{n\wedge S(k)}|\mathcal{F}_{n-1}) = \left(X_{(n-1)\wedge S(k)}^2 - A_{(n-1)\wedge S(k)}\right).$$

Thus $A_{n\wedge S(k)} = \langle X_{n\wedge S(k)} \rangle$. Since $A_{n\wedge S(k)}$ is bounded, hence $\mathrm{E}(A_{\infty\wedge S(k)}) \leq k$ implying $X_{n\wedge S(k)}$ converges almost surely. This implies X_n converges almost surely on $\{\omega : S(k) = \infty\}$. Taking $k \to \infty$ we have that If $A_\infty(\omega) < \infty$ then $\lim_n X_n(\omega)$ exists.

(ii) Define a new stopping time

$$T(k) = \inf\{n : |X_n| > k\}.$$

Hence, similar to before,

$$\mathrm{E}(X^2_{n\wedge T(k)}) = \mathrm{E}(A_{n\wedge T(k)}) \le (k+K)^2.$$

There on $\{\omega : T(k) = \infty\}$ we know that A_∞ is finite. Take $k \to \infty$. □

Theorem 11 (Levy's extension of Borel-Cantelli Lemmas)**.** *Let $\{\mathcal{F}_n\}$ be a filtration and let $G_n \in \mathcal{F}_n$. Define*

$$Z_n = \sum_{k=1}^{n} 1_{G_k}.$$

Let $B_n = \mathrm{E}(1_{G_n}|\mathcal{F}_{n-1})$ and let

$$Y_n = \sum_{k=1}^{n} B_k.$$

Then, almost surely,

- $Y_\infty < \infty$ *implies* $Z_\infty < \infty$
- $Y_\infty = \infty$ *implies* $\frac{Z_n}{Y_n} \to 1$.

Proof. Observe that

$$\mathrm{E}(Z_n - Y_n|\mathcal{F}_{n-1}) = Z_{n-1} - Y_{n-1},$$

hence $M_n := Z_n - Y_n$ is a Martingale. Let $A_n = \langle M_n \rangle$. Note that

$$\begin{aligned} A_n &= \sum_{k=1}^{n} \mathrm{E}(M_k^2 - M_{k-1}^2|\mathcal{F}_{k-1}) = \sum_{k=1}^{n} \mathrm{E}((M_k - M_{k-1})^2|\mathcal{F}_{k-1}) \\ &= \sum_{k=1}^{n} (1_{G_k} - B_k)^2|\mathcal{F}_{k-1}) = \sum_{k=1}^{n} B_k(1-B_k) \le Y_n. \end{aligned}$$

If $Y_\infty < \infty$ then $A_\infty < \infty$ and from previous theorem, $\lim_n M_n$ exists and hence $Z_\infty < \infty$.

To get the second part, note that if $A_\infty < \infty$ then $\lim_n M_n$. Hence if $Y_\infty = \infty$ then $\frac{Z_n}{Y_n} \to 1$ (trivially). The more interesting case is when $A_\infty = \infty$.

Show that the Martingale transform

$$U_n = \sum_{k=1}^{n} \frac{M_k - M_{k-1}}{1 + A_k}$$

is a square integrable martingale, where $U_0 = 0$.

Now observe that (justify each step since I am leaving out some calculations)

$$\begin{aligned}
\langle U_n \rangle) &= \sum_{k=1}^{n} \mathrm{E}((U_k - U_{k-1})^2 | \mathcal{F}_{k-1}) \\
&= \sum_{k=1}^{n} \mathrm{E}\left(\frac{M_k^2 - M_{k-1}^2}{1 + A_k} | \mathcal{F}_{k-1} \right) \\
&= \sum_{k=1}^{n} \frac{A_k - A_{k-1}}{(1 + A_k)^2} \\
&\le \sum_{k=1}^{n} \left(\frac{A_k}{1 + A_k} - \frac{A_{k-1}}{1 + A_{k-1}} \right) = \mathrm{E}\left(\frac{A_n}{1 + A_n} \right) \le 1.
\end{aligned}$$

Since $\langle U_n \rangle \le 1$, hence U_n converges to a finite limit. Hence, Kronecker's lemma says that $\frac{M_n}{1+A_n} \to 0$. □

Lemma 5 (Kronecker)**.** *Let* $\sum_{m=1}^{n} x_m$ *be a convergent. It a sequence of positive numbers* $b_n \uparrow \infty$ *then*

$$\frac{\sum_{k=1}^{n} b_k x_k}{b_n} \to 0.$$

Proof. Let $s_n = \sum_{m=1}^{n} x_m$ The key is to express

$$\sum_{k=1}^{n} \frac{b_k x_k}{b_n} = \sum_{k=1}^{n} \frac{b_k (s_k - s_{k-1})}{b_n} = s_n - \sum_{k=1}^{n-1} \frac{(b_{k+1} - b_k) s_k}{b_n}.$$

Given $\epsilon > 0$ we know that $\exists n_0$ such that $|s_n - s| \le \epsilon, \forall n \ge n_0$. Hence for $n > n_0$ we have

$$\begin{aligned}
\left| \sum_{k=1}^{n} \frac{b_k x_k}{b_n} \right| &\le |s_n - s| + \sum_{k=1}^{n-1} \frac{(b_{k+1} - b_k)|s_k - s|}{b_n} \\
&\le \epsilon + \frac{b_n - b_{n_0}}{b_n} \epsilon + \frac{1}{b_n} \sum_{k=1}^{n_0} (b_{k+1} - b_k)|s_k - s|.
\end{aligned}$$

Taking $\limsup_n$ we obtain an upper bound of 2ϵ and since $\epsilon > 0$ is arbitrary we are done. □

Theorem 12 (Kakutani's theorem for Product Martingales)**.** *Let* $X_1, X_2, \cdots$ *be independent non-negative random variables, each of mean 1. Define* $M_0 = 1$ *and let*

$$M_n := X_1 \cdots X_n.$$

Then $\{M_n\}$ *is a non-negative martingale; hence* $M_\infty := \lim_n M_n$ *exists almost surely.*

Secondly, the following five statements are equivalent:

(*a*) $\mathrm{E}(M_\infty) = 1$
(*b*) $M_n \to M_\infty$ *in* L^1
(*c*) $\{M_n\}$ *is UI*
(*d*) $\prod_n a_n > 0$, *where* $a_n = \mathrm{E}(\sqrt{X_n})$
(*e*) $\sum_n (1 - a_n) < \infty$.

If one of the statements fail to hold then $\mathrm{P}(M_\infty = 0) = 1$.

Proof. Note that $0 < a_n \leq 1$ (the second inequality follows by Jensen). The equivalence between (d) and (e) is straightforward. Define

$$Y_n = \sqrt{\frac{M_1 M_2 \cdots M_n}{a_1 a_2 \cdots a_n}}.$$

Since Y_n is a non-negative Martingale we know that Y_n converges almost-surely to a finite limit. Hence if $\prod_n a_n = 0$, then we must have $M_\infty = 0$ almost surely. Argue (c) does not hold.

On the other hand if $\prod_n a_n > 0$, then $\lim_n \mathrm{E}(Y_n^2) < \infty$. Now

$$\mathrm{E}(\sup_{1 \leq k \leq n} M_k) \leq \mathrm{E}(\sup_{1 \leq k \leq n} Y_k^2) \leq 4\,\mathrm{E}(Y_n^2).$$

The last step is Doob's inequality for submartingales, specialized to L^2. Hence monotone convergence theorem yields

$$\mathrm{E}(\sup_n M_n) < \infty.$$

Defining $M_* = \sum_n M_n$, since $\mathrm{E}(M_*) < \infty$, we see that $\{M_n\}$ is UI and implies the first two parts. □

Theorem 13 (Law of iterated Logarithm for Gaussians)**.** *Let $\{X_n\}$ be i.i.d. random variables, each distributed as $\mathcal{N}(0,1)$. Let $S_n = \sum_{i=1}^n X_i$. Then*

$$\limsup_n \frac{S_n}{\sqrt{2n \log\log n}} = 1 \ a.s.$$

Proof. Let $h_n = \sqrt{2n \log\log n}, n \geq 3$. Further it is easy to see that

$$\mathrm{E}(e^{\theta S_n}) = e^{\frac{1}{2}\theta^2 n}.$$

Now $\{e^{\theta S_n}\}$ is a sub-martingale, hence

$$P(\sup_{1 \leq k \leq n} S_k \geq l) \leq e^{-\theta l} e^{\frac{1}{2}\theta^2 n}.$$

Optimizing over θ yields

$$P(\sup_{1 \leq k \leq n} S_k \geq l) \leq e^{-\frac{l^2}{2n}}.$$

Take $K > 1$ and $l_n = Kh(K^{n-1})$. Now observe that

$$P\left(\sup_{1 \leq k \leq K^n} S_k \geq l_n\right) \leq e^{-\frac{l_n^2}{2K^n}} = e^{-K \log\log K^{n-1}} = \frac{1}{((n-1)\log K)^K}.$$

From B-C 1, we have, almost surely, for all $K^{n-1} \leq k \leq K^n$, and n large-enough $S_k \leq \sup_{1 \leq k \leq K^n} S_k \leq l_n = Kh(K^{n-1}) \leq Kh(k)$. Hence $\limsup_k \frac{S(k)}{k} \leq K$ almost surely. Since $K > 1$ is arbitrary, we are done with the upper bound.

Let N be an integer larger than 1. Note that $\frac{S(N^{n+1}) - S(N^n)}{\sqrt{N^{n+1} - N^n}}$ is a Gaussian. Now from Lemma below we have

$$P(F_n) := P\left(\frac{S(N^{n+1}) - S(N^n)}{\sqrt{N^{n+1} - N^n}} \geq (1-\epsilon)\frac{h(N^{n+1} - N^n)}{\sqrt{N^{n+1} - N^n}}\right) \geq c(n \log N)^{-(1-\epsilon)^2}.$$

From B-C 2, infinitely often, we have

$$S(N^{n+1}) - S(N^n) \geq (1-\epsilon)h(N^{n+1} - N^n)$$

infinitely often. However from earlier part $S(N^n) \geq -Kh(N^n)$ for $K > 1$ eventually (in n), so infinitely often we have

$$S(N^{n+1}) \geq (1-\epsilon)h(N^{n+1} - N^n) - Kh(N^n).$$

Therefore

$$\limsup_k \frac{S(k)}{h(k)} \geq \limsup_n \frac{S(N^{n+1})}{h(N^{n+1})} \geq (1-\epsilon) + \theta(1/\sqrt{N}).$$

Letting $N \to \infty$ completes the proof. □

Lemma 6. *Let X be standard normal. Then*

$$P(X > x) \geq (x + x^{-1}) \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Proof. Let $\phi(x)$ be the density. Then $(x^{-1}\phi(x))' = -(1 + x^{-2})\phi(x)$. Hence

$$x^{-1}\phi(x) = -\int_x^\infty (y^{-1}\phi(y))' dy = \int_x^\infty (1 + y^{-2})\phi(y) dy \leq (1 + x^{-2}) P(X > x).$$

□