Recall:

Doob's Convergence Theorem

Let X_n be a supermartingale with $\sup_n \mathbb{E}(|X_n|) < \infty$. Then, $X_\infty = \lim_{n \to \infty} X_n$ exists a.s. and is finite.

We will now explore martingales with stronger assumptions:

L^2 - Martingale

In the following, we consider martingales X_n with finite second moment: $\mathbb{E}(X_n^2) < \infty$.

We define the inner product $\langle U, V \rangle = \mathbb{E}(UV)$ and have the following orthogonality property:

for $s \leq t \leq u \leq v$ and an L^2-martingale M_n

$$\langle M_t - M_s , M_v - M_u \rangle = 0$$

"Increments at different times are independent"

Proof: $\mathbb{E}(M_v - M_u \mid \mathcal{F}_k)$

$= \mathbb{E}(M_v \mid \mathcal{F}_k) - \mathbb{E}(M_u \mid \mathcal{F}_k) = M_k - M_k = 0$

for all $k \leq u \leq v$.

Likewise, $\mathbb{E}(M_t - M_s \mid \mathcal{F}_k) = 0$ for all $k \leq s \leq t$.

Consider

$$\mathbb{E}\big(\underbrace{(M_t - M_s)}_{\mathcal{F}_t\text{-measurable}}(M_v - M_u) \mid \mathcal{F}_t\big) = (M_t - M_s)\,\mathbb{E}(M_v - M_u \mid \mathcal{F}_t)$$

$$= (M_t - M_s) \cdot 0 = 0 \text{ (a.s.)}$$

$$\Rightarrow \mathbb{E}\big((M_t - M_s)(M_v - M_u)\big) = \mathbb{E}\big(\mathbb{E}((M_t - M_s)(M_v - M_u) \mid \mathcal{F}_t)\big)$$

$$= \mathbb{E}(0) = 0$$

So increments over disjoint intervals are orthogonal w.r.t. $\langle , \rangle$.

If we write $M_n = M_0 + (M_1 - M_0) + (M_2 - M_1) + \ldots + (M_n - M_{n-1})$

then all the summands are pairwise orthogonal and Pythagoras' theorem gives us

$$\mathbb{E}(M_n^2) = \mathbb{E}(M_0^2) + \mathbb{E}((M_1 - M_0)^2) + \ldots + \mathbb{E}((M_n - M_{n-1})^2)$$

So $\mathbb{E}(M_n^2) < \infty \iff \sum_{n=1}^{\infty} \mathbb{E}((M_n - M_{n-1})^2) < \infty$.

Here also $\mathbb{E}(|M_n|) \leq \sqrt{\mathbb{E}(M_n^2)} < \infty$

so the convergence theorem applies:

$M_n \to M_\infty$ a.s.

It also holds that $\mathbb{E}((X_\infty - X_n)^2) = \|X_\infty - X_n\|_2^2$

tends to 0. That is, $M_n \to M_\infty$ with respect

to the norm $\|\cdot\|_2$).

One can verify this as follows:

$$\mathbb{E}((M_{n+r} - M_r)^2) = \sum_{k=r+1}^{n+r} \mathbb{E}((M_k - M_{k-1})^2)$$

by orthogonality.

Now let $n \to \infty$: $\mathbb{E}((M_\infty - M_r)^2)$

$$= \mathbb{E}(\lim_{n\to\infty} (M_{n+r} - M_r)^2) \underset{\text{Fatou}}{\leq} \liminf_n \mathbb{E}((M_{n+r} - M_r)^2)$$

$$= \sum_{k=r+1}^{\infty} \mathbb{E}((M_k - M_{k-1})^2) < \infty.$$

Now as $r \to \infty$, it follows that

$\mathbb{E}((M_\infty - M_r)^2) \to 0.$

Now consider the special case where M_n is a sum of independent random variables $X_1, X_2, \dots, X_n$.
$M_0 = 0, \quad M_n = X_1 + X_2 + \dots + X_n$ with $\sigma_k^2 = Var(X_k) < \infty$. If $\mathbb{E}(X_k) = 0$ for all k, then M_n is a martingale.

<u>Theorem</u> If $\sum \sigma_k^2 < \infty$, then $\sum_{k=1}^{\infty} X_k = \lim_{n\to\infty} M_n$ exists and is almost surely finite.

Proof: $$\sum_{k=1}^{\infty} \mathbb{E}((M_k - M_{k-1})^2) = \sum_{k=1}^{\infty} \underbrace{\mathbb{E}(X_k^2)}_{= Var(X_k)} = \sum_{k=1}^{\infty} \sigma_k^2.$$

So convergence follows. [Why? Work out details] $\square$

Remark: If the X_k are also uniformly bounded, the converse also holds: If the sum $\sum X_k$ converges a.s., then $\sum \sigma_k^2 < \infty$.

[Why? Exercise]

Example: Let $X_1, X_2, \ldots$ be random variables with $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$, and consider the random sum $\sum_{k=1}^{\infty} a_k X_k$, $\sup_k |a_k| < \infty$.

Note that $Var(a_k X_k) = \mathbb{E}((a_k X_k)^2) = a_k^2$.

So the theorem above shows that the random sum converges (a.s.) if and only if $\sum a_k^2 < \infty$.

Strong law of large numbers for $\mathcal{L}^2$ random var.

We will combine our $\mathcal{L}^2$ martingale results with results from real analysis:

Cesàro's lemma: If b_n is a seq of non-neg. reals with $b_n \uparrow \infty$ and v_n is a convergent sequence of reals with $v_n \to v_\infty$, then

$$\frac{1}{b_n} \sum_{k=1}^{n} (b_k - b_{k-1}) v_k \to v_\infty .$$

Note: WLOG, $b_0 = 0$ and then $\sum_{k=1}^{n} \frac{b_k - b_{k-1}}{b_n} = 1$, so LHS is a weighted average of v_k.

<u>Kronecker's lemma:</u> Let b_n be a non-neg seq of reals with $b_n \uparrow \infty$. Let x_n be an arbitrary seq. of reals and write $s_n = x_1 + \dots + x_n$. If $\sum_{n=1}^{\infty} \frac{x_n}{b_n}$ converges, then $s_n / b_n \to 0$.

Let Y_n be a sequence of independent rand. variables with $\mathbb{E}(Y_n) = 0$ and $Var(Y_n) < \infty$ for all $n \in \mathbb{N}$. If $\sum_{n=1}^{\infty} \frac{Var(Y_n)}{n^2} < \infty$ then $\sum_{n=1}^{\infty} \frac{Y_n}{n}$ converges a.s.

This is because $Var(Y_n / n) = \frac{Var(Y_n)}{n^2}$ and we can apply the prev. convergence theorem.

Kronecker's lemma with $b_n = n$ and $x_n = Y_n$ gives

$$\frac{s_n}{b_n} = \frac{\sum_{k=1}^{n} Y_k}{n}$$ converges for a.e. $\omega \in \Omega$.

<u>Remark:</u> The strong law of large numbers holds for all Y_n s.t. $\sum \frac{Var(Y_n)}{n^2} < \infty$ (rather than $\mathbb{E}(Y_n^4) \le K$)

Remark: If X_n is an i.i.d. seq. of random variables with mean μ and variance $\sigma^2 = Var(X_n)$, then $Y_n = X_n - \mu$ satisfies:

$$\left.\begin{array}{l} \bullet\ \mathbb{E}(Y_n) = 0 \\ \bullet\ Var(Y_n) = \sigma^2 \end{array}\right\} \Rightarrow \sum_{n\in\mathbb{N}} \frac{Var(Y_n)}{n^2} = \sigma^2 \sum_{n\in\mathbb{N}} \frac{1}{n^2} < \infty.$$

Hence,
$$\frac{X_1 + X_2 + \dots + X_n}{n} = \frac{Y_1 + Y_2 + \dots + Y_n}{n} + \mu \to \mu$$
almost surely.

We will slightly tweak this method with a truncation approach:

Kolmogorov's truncation lemma:

Let (X_n) be a seq. of iid random variables. Assume $X \sim X_n$ is integrable and $\mathbb{E}(X) = \mu$. Write
$$Y_n = \begin{cases} X_n & \text{if } |X_n| \leq n \\ 0 & \text{otherwise.} \end{cases}$$
Then the following hold:

1) $\mathbb{E}(Y_n) \to \mu$ as $n \to \infty$,
2) $\mathbb{P}(Y_n = X_n \text{ for all but finitely many } n) = 1$,
3) $\sum_{n\in\mathbb{N}} \frac{Var(Y_n)}{n^2} < \infty$.

Proof: 1) $|Y_n| \leq |X_n|$ and hence

$\mathbb{E}(|Y_n|) \leq \mathbb{E}|X_n| = \mathbb{E}|X| < \infty$. Thus by dominated convergence, $\mathbb{E}(Y_n) \to \mathbb{E}(X) = \mu$.

2)

$\mathbb{P}(Y_n \neq X_n) = \mathbb{P}(|X_n| > n)$. Thus

$$\sum_{n \geq 1} \mathbb{P}(Y_n \neq X_n) = \sum_{n \in \mathbb{N}} \mathbb{P}(|X_n| > n)$$

$$= \sum_{n \in \mathbb{N}} \mathbb{P}(|X| > n)$$

$$= \sum_{n \in \mathbb{N}} \mathbb{E}(I_{\{|X| > n\}})$$

$$= \mathbb{E}\Big(\underbrace{\sum_{n \in \mathbb{N}} I_{\{|X| > n\}}}_{\#\text{ of integers smaller than } |X|}\Big) \leq \mathbb{E}|X| < \infty$$

The statement then follow with B.C. lemma.

3) We have $Var(Y_n) = \mathbb{E}(Y_n^2) - \mathbb{E}(Y_n)^2 \leq \mathbb{E}(Y_n^2)$

So $$\sum_{n \in \mathbb{N}} \frac{Var(Y_n)}{n^2} \leq \sum_{n \in \mathbb{N}} \frac{\mathbb{E}(Y_n^2)}{n^2} = \sum_{n \in \mathbb{N}} \frac{\mathbb{E}(X_n^2 \cdot I_{\{|X_n| \leq n\}})}{n^2}$$

$$= \sum_{n \in \mathbb{N}} \frac{\mathbb{E}(|X|^2 \cdot I_{\{|X| \leq n\}})}{n^2} = \mathbb{E}\Big(|X|^2 \sum_{n \in \mathbb{N}} \frac{1}{n^2} I_{\{|X| \leq n\}}\Big)$$

$$= \mathbb{E}\Big(|X|^2 \sum_{n \geq |X|} \frac{1}{n^2}\Big) \underset{(*)}{\leq} \mathbb{E}\Big(|X|^2 \frac{2}{\max\{1, |X|\}}\Big) = \mathbb{E}(2|X|) < \infty$$

where $(*)$ follows from:

$$\frac{1}{n^2} \leq \frac{2}{n(n+1)} = \frac{2}{n} - \frac{2}{n+1} \quad \text{and}$$

$$\sum_{n \geq k} \frac{1}{n^2} \leq \sum_{n \geq k} \left(\frac{2}{n} - \frac{2}{n+1}\right) = \left(\frac{2}{k} - \frac{2}{k+1}\right) + \left(\frac{2}{k+1} - \frac{2}{k+2}\right) + \dots$$

$$= 2/k. \qquad \square$$

Finally:

<u>Kolmogorov's</u> strong <u>law</u> of <u>large numbers</u> (LLN)

Let $X_1, X_2, \dots$ be independent, identically distributed random variables with $\mathbb{E}(X_i) = \mu$. Then,

$$\frac{1}{n}(X_1 + X_2 + \dots + X_n) \xrightarrow[n \to \infty]{} \mu. \quad \text{a.s.}$$

<u>Proof</u>: Define Y_n as above (truncation).

Note that $\frac{1}{n}(X_1 + \dots + X_n)$ and $\frac{1}{n}(Y_1 + \dots + Y_n)$ a.s. have the same limit as they only differ finitely many times (by 2)).

Now

$$\frac{1}{n}(Y_1 + \dots + Y_n) = \frac{(Y_1 - \mathbb{E}(Y_1)) + \dots + (Y_n - \mathbb{E}(Y_n))}{n} + \frac{1}{n}\left(\mathbb{E}(Y_1) + \dots + \mathbb{E}(Y_n)\right)$$

The first summand satisfies previous criteria:

$$\mathbb{E}(Y_j - \mathbb{E}(Y_j)) = 0, \quad \sum_j \frac{Var(Y_j - \mathbb{E}(Y_j))}{j^2} = \sum \frac{Var\, Y_j}{j^2}$$

which is finite by (3). Hence

$$\lim_{n\to\infty} \frac{1}{n}\left(Y_1 + \dots + Y_n\right) = \lim_{n\to\infty} \frac{1}{n}\left(\mathbb{E}(Y_1) + \dots + \mathbb{E}(Y_n)\right)$$

which equals μ by Cesàro's lemma and (1). $\square$

Doob decomposition

Recall:

$$\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) \begin{cases} \geq X_{n-1} & \text{sub-} \\ = X_{n-1} & \text{martingale.} \\ \leq X_{n-1} & \text{super-} \end{cases}$$

Let X_n be an adapted process w.r.t. $(\mathcal{F}_n)$. Then we can always find a previsible process A_n and a martingale M_n s.t.

$X_n = X_0 + M_n + A_n$ and $M_0 = A_0 = 0$.

This decomposition is unique (up to a null set).

From this we also get

X_n is a super-/submartingale

$\Updownarrow$

A_n is decreasing / increasing a.s.

<u>Proof:</u> Suppose we are given the decomposition:

$X_n - X_{n-1} = M_n - M_{n-1} + A_n - A_{n-1}$.

This gives $\mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1})$

$= \mathbb{E}(M_n - M_{n-1} \mid \mathcal{F}_{n-1}) + \mathbb{E}(A_n - A_{n-1} \mid \mathcal{F}_{n-1})$

$= \quad 0 \quad + \quad A_n - A_{n-1}$.

martingale previsible

Thus $A_n = \sum_{k=1}^{n} A_k - A_{k-1} = \sum_{k=1}^{n} \mathbb{E}(X_k - X_{k-1} \mid \mathcal{F}_{k-1})$

is uniquely determined and so is $M_n = X_n - X_0 - A_n$ (a.s.).

Conversely, one can check that this choice of M_n, A_n works. $\square$

Uniform Integrability

Problem: Given $X_n \to X_\infty$, when can we say that $\mathbb{E}(X_n) \to \mathbb{E}(X_\infty)$?

Example: $X_n = \begin{cases} n^2 & \text{with prob } 1/n^2 \\ 0 & \text{otherwise.} \end{cases}$

Then $\mathbb{E}(X_n) = 1$. Since $\sum_n P(X_n \neq 0) = \sum_n 1/n^2 < \infty$ we have $X_n \to X_\infty = 0$ a.s.

But $\mathbb{E}(X_\infty) = 0 \neq 1 = \lim_n \mathbb{E}(X_n)$!

Uniform integrability is a key condition that allows exchange of $\mathbb{E}$ and lim.

Lemma: Let X be an integrable random variable. For every $\varepsilon > 0$, there exists $\delta > 0$ such that for all events E with $P(E) < \delta$, we have $\mathbb{E}(|X|; E) = \mathbb{E}(|X| \cdot I_E) < \varepsilon$.

(This is a special case of Egorov's theorem.)

Proof: Suppose this was not the case:

For some $\varepsilon_0 > 0$ there exists a sequence of events E_n s.t. $\mathbb{P}(E_n) < 2^{-n}$ but $\mathbb{E}(|X| \cdot I_{E_n}) \geq \varepsilon_0$. Since $\sum_n \mathbb{P}(E_n) < \infty$, the B.C. lemma implies that only finitely many E_n occur. Let $F = \limsup_{n \to \infty} E_n$.

Then $\mathbb{P}(F) = 0$.

Hence $\mathbb{E}(|X| \cdot I_F) = 0$.

But by the reverse Fatou lemma:

$$\limsup_{n \to \infty} \mathbb{E}(|X| I_{E_n}) \leq \mathbb{E}(|X| \limsup_n I_{E_n})$$

$$= \mathbb{E}(|X| \cdot I_F) = 0$$

But the LHS is bounded below by $\varepsilon_0 > 0$, a contradiction $\lightning$

$\square$